Association of two novel viruses with chlorotic fleck disease of ginger

A.I. Bhat[1] | K.P. Naveen[1] | N.S Pamitha[1] | R.P.Pant[2]

[1]ICAR-Indian Institute of Spices Research, Kozhikode 673012, Kerala, India
[2]ICAR-Indian Agricultural Research Institute, New Delhi 110012, India

Correspondence
A.I. Bhat, ICAR-Indian Institute of Spices Research, Kozhikode 673012, Kerala, India. Email: aib65@yahoo.co.in; Tel: +91-495-2731410

Received: 25 October 2019; revised manuscript received: 18 May 2020; accepted: 6 June 2020.

Summary

Chlorotic fleck disease of ginger, the causal virus of which was unknown so far is an important production constraint of ginger in India and other parts of the world. In the present study, two new RNA viruses were discovered in chlorotic fleck affected plant by the virome analysis using highthroughput sequencing of small RNA (sRNA) and transcriptome. The highthroughput sequencing results were verified through reverse transcriptase polymerase chain reaction (RT–PCR) using total RNA from infected plants and primers designed from the contigs. Cloning and sequencing of the RT-PCR products of one of the virus resulted in a sequence of 4143 bases that showed similarities to panicoviruses and machlomoviruses. The complete genome of this virus contained six open reading frames (ORFs) that potentially encode proteins of 43, 104, 8, 7, 15, and 27 kDa. Based on the genomic and phylogenetic analysis, this virus is predicted to be a new member of the family *Tombusviridae* for which the name ginger chlorotic fleck-associated virus 1 (GCFaV-1) is proposed. Similarly, cloning and sequencing of the RT-PCR products of other virus resulted in a sequence of 5514 bases that showed similarities to

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/aab.12615

ampeloviruses. Based on the genomic and phylogenetic analysis, this virus is predicted to be a new member of the genus *Ampelovirus* for which the name ginger chlorotic fleck-associated virus 2 (GCFaV-2) is proposed. A reliable RT–PCR and SYBR Green–based quantitative RT–PCR assays were developed for the detection of both viruses in plants that would aid in the identification and propagation of virus-free ginger plants. Additional investigations are required to elucidate the relationship between the symptoms and viral infections.

Keywords

Ampelovirus, detection, ginger, identification, RNA sequencing, RT-PCR, small RNA sequencing, Tombusviridae

1 | Introduction

Ginger (*Zingiber officinale* Rosc.) (Family: *Zingiberaceae*) is an herbaceous crop, perennial but cultivated as an annual. The aerial portion is referred to as pseudostem and rhizome, the ginger of commerce is the modified stem for the storage of food. The rhizome is the material used for vegetative propagation (Ravindran & Babu, 2005). Ginger production is mainly distributed over India, China, Nigeria, Nepal, Indonesia, Thailand, Bangladesh, Cameroon, Japan, and the Philippines (FAO, 2020). India ranks first concerning area, spanning about 46.1% of the world's total area, followed by Nigeria (19.3%), China (12.6%), Nepal (6.2%), Thailand (2.7%), Indonesia (2.7%) and Bangladesh (2.6%). A virus causing symptoms of yellowish and dark green mosaic and stunting of ginger in Korea has been identified as a strain of cucumber mosaic virus (CMV) based on host range, electron microscopy, and serology (So, 1980; Kitajima

& Pozzer, 1994). Thomas (1986) detected a new virus namely ginger chlorotic fleck virus (GCFV) with possible affinities to *Sobemovirus* in the ginger imported into Australia from many countries. Fan et al. (1999) reported isolation and characterization of an isolate of tobacco mosaic virus (TMV) infecting ginger in the Shandong province of China. In India, the first occurrence of a viral disease in ginger was published by Nambiar & Sarma (1974). The disease is widespread in almost all ginger growing areas of the country with an incidence ranging from 10–90%. The disease is characterized by light green to bright yellow intraveinal chlorotic flecks which later coalesce together turning entire vein chlorotic. Mottling on the pseudostem and stunting of plants are the other associated symptoms of the disease. Our efforts to detect viruses through RT-PCR using primers specific to viruses reported on ginger elsewhere such as CMV, TMV, and *Sobamovirus* did not give results. Scanning of the literature revealed that highthroughput sequencing (HTS) could provide an unbiased approach to plant disease diagnosis without prior knowledge of the host or virus (Quan et al., 2008; Adams et al., 2009; Kreuze et al., 2009; Barba et al., 2014; Cao et al., 2019).

The *Closteroviridae,* which includes flexuous viruses with a single-stranded and long positive-sense RNA, is divided into four genera: *Closterovirus, Crinivirus, Ampelovirus*, and *Velarivirus*. Ampeloviruses have long filamentous particles of about 1400 to 2200 nm in length that contain positive-sense ssRNA of 16.9 to 17.9 kb coding for 10-13 proteins (Martelli et al., 2012; Fuchs et al., 2020). The first open reading frame (ORF) 1a encodes a putative RNA helicase while the ORF 1b encodes an RNA-dependent RNA polymerase, overlaps ORF1a, and lacks an initiation codon (Ling et al., 2004). Other ORFs encoded in the 5' to 3' direction include a small hydrophobic protein, a heat shock protein 70-like protein, 61.0 kDa protein and coat protein (CP), two diverged copies of the CP (CPd1 and CPd2) and two other proteins of unknown

function (Komínek et al., 2005). Open reading frame (ORF) 1a and ORF 1b are expressed *via* a +1 ribosomal frameshift to produce a fusion protein containing RNA dependent RNA polymerase (RdRp) while other ORFs are translated from a set of nested 3´ co-terminal subgenomic RNAs. The International Committee on Taxonomy of Viruses (ICTV, 2019) listed 12 species in the genus and some other species are awaiting recognition. Ampeloviruses are further separated into two phylogenetically distinct clades- subgroup I and II based on the genome size and number of ORFs (Martelli et al., 2012).

The *Tombusviridae,* which includes isometric viruses with a smooth or granular appearance of about 28–30 nm diameter with a single-stranded positive-sense RNA of about 4 kb in size, is divided into three subfamilies, 16 genera and 76 species (ICTV, 2019). The subfamily, *Calvusvirinae* consist of a single genus *Umbravirus* whose members lack coat protein ORF while the subfamily, *Regressovirinae* also has a single genus *Dianthovirus* whose members have a bisegmented genome. The rest of the 14 genera belong to the subfamily, *Procedovirinae* with 64 species. The 14 genera of this subfamily are *Alphacarmovirus, Alphanecrovirus, Aureusvirus, Avenavirus, Betacarmovirus, Betanecrovirus, Gallantivirus, Gammacarmovirus, Macanavirus, Machlomovirus, Panicovirus, Pelarspovirus, Tombusvirus,* and *Zeavirus.* The genomes of viruses in this subfamily are nonsegmented and encode four to seven ORFs that include RNA-dependent RNA polymerase (RdRp), accessory protein to assist replication, one to three movement proteins to assist virus transport and coat protein (Scheets, 2013; Scheets et al., 2015)

2 | Materials and methods

2.1 | Virus isolates

For initial identification and characterization of the viruses, infected ginger plants belonging to the variety, IISR–Varada collected and maintained through vegetative propagation under insect-proof glasshouse conditions at the ICAR-Indian Institute of Spices Research, Kozhikode, Kerala, India was used. The same virus-infected plant (positive control) was used for the optimization of RT-PCR and SYBR Green-based quantitative RT-PCR (qRT–PCR) assays along with virus-free (confirmed through RT-PCR test) (negative control) ginger plant. Both assays were validated using field samples of different popular varieties of ginger collected from ginger growing regions of Karnataka and Kerala states of India. Besides, the ginger germplasm accessions collected from different agro-climatic regions and maintained at the ICAR-Indian Institute of Spices Research, India were also used for validating the assays.

2.2 | Electron microscopy

Leaf dip transmission electron microscopy of infected ginger leaves was carried out at the Virology Unit, ICAR-Indian Agricultural Research Institute, New Delhi using the negative staining method with 2% uranyl acetate (pH 4.5) (Milne, 1984).

2.3 | Small RNA sequencing (sRNA–seq) and RNA sequencing (RNA–seq)

Two approaches namely, small RNA sequencing (sRNA–seq) and RNA sequencing (RNA–seq) were used to identify viruses associated with the disease. Total RNA isolated using Xcelgen Plant RNA Mini Kit from three pooled leaf samples of virus-infected ginger plants was used for sRNA–seq and RNA–seq. The sRNA libraries were prepared using Illumina TruSeq Small RNA Prep Kit from 1 µg total RNA. The cDNA was PCR amplified using a common

primer and index primer to create cDNA construct. The cDNA construct was purified through 6% PAGE gel, the library was sequenced using 1 × 50 bp chemistry by Illumina Hiseq 2500 platform at the Xcelris Labs Ltd, Ahmedabad, India. Total reads obtained were filtered to remove adapter sequences using Trimmomatic-v3.6 followed by size selection (16–28 bases) using cutadapt v1.9. The size selected reads were used for the identification of viruses through Velvet 1.1.07 and CLC genomics workbench v6. The *de novo* assembled contigs obtained from the above were subjected to BLASTN and BLASTX program using non–redundant plant virus database. For RNA–seq, library preparations were constructed using total RNA depleted of rRNA using NEB Next UltraTM RNA library preparation kit for Illumina. The cDNA was amplified and the library was sequenced using 2 × 150 bp chemistry by Illumina Hiseq 2500 platform at the Nucleome Informatics Private Limited, Hyderabad, India. The *de novo* assembled contigs using BinPacker and rnaSPAdes were subjected to BLASTN and BLASTX program against the non-redundant plant virus database.

2.4 | RT–PCR, cloning and Sanger sequencing

To confirm the results of sRNA–seq and RNA–seq, primers were designed based on the sequence of the specific contigs that mapped to the viral genomes (Supplementary Table 1) and subjected to RT-PCR using total RNA isolated from a virus-infected and healthy ginger plant as a template. Total RNA was isolated as per the protocol of Deepa et al. (2013). The RT–PCR reaction was carried out in 1x *Taq* assay buffer that contained 1.5 mM $MgCl_2$, 10 mM dithiothreitol, 400 µM dNTP mix, 10 pM each of forward and reverse primers, 1 U of RNase inhibitor, 1.25 U of MuMLV reverse transcriptase, 0.75 U of *Taq* DNA polymerase and 1 µL total RNA as the template with a final volume of 50 µL. Before the addition of the RNA

template to the reaction mixture, it was heated to 80°C for 10 min and rapidly cooled down in ice for 10 min to make RNA linear. Single-step RT–PCR was carried out in Eppendorff's Master Cycler Gradient by initially holding the sample at 42°C for 45 min (cDNA synthesis) followed by 35 cycles at 94°C for 30 s (denature), 50–58°C (depending Tm of the primer pair) for 30 s (primer annealing) and 72°C for 1–2 min (depending on the expected product size) (DNA synthesis) and a final extension at 72°C for 10 min. To obtain the 3' end of the sequence, the total RNA was polyadenylated using *E. coli* Poly(A) Polymerase and used as the template in a 3' RACE experiment using 5′/3′ RACE Kit (Roche Applied Science). 5' RACE was employed to obtain the 5' end of viral RNA. Experimental controls were set up using total RNA from virus-free plant and nuclease-free water instead of template RNA. The reaction products were subjected to electrophoresis and visualized under UV light. The resulting RT–PCR products were cloned into pTZ57R/T (Fermentas, USA) as per manufacturer's protocol and two clones each were sequenced from both directions by Sanger dideoxy chain termination method at the automated sequencing facility available at Agri Genome, Kochi, India.

2.5 | Sequence analyses

The sequences were assembled with Seqaid Version 3.6 (Peltola et al., 1984). ORFs in the sequence data were predicted by ORF finder [www.ncbi.nlm.nih.gov/projects/gorf] and translated into amino acid residues using the Expasy tool. Sequences of the complete genome of available ampeloviruses and viruses under the family *Tombusviridae* were retrieved from the NCBI database and used for analysis (Supplementary Table 2 and 3). The analyses were carried out in the form of nucleotide and translated amino acid sequences. Sequences were aligned using clustal X (Thompson et al., 1997) and percent identity was calculated using Clustal omega

(www.ebi.ac.uk/Tools/msa/clustalo).Conserved domain/motif of putative protein was detected by the Conserved Domain Database (CDD) of NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd /wrpsb.cgi) and Pfam (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan). Phylogenetic relationships between novel viruses and a range of related viruses were determined by clustal X with a bootstrap analysis of 1000 replicates.

2.6 | Development of RT–PCR and SYBR Green-based quantitative RT–PCR (qRT–PCR) assays for the detection of viruses

Total RNA isolated from the positive control and negative control was subjected to single tube RT–PCR and qRT–PCR assays using primers designed to RdRp or coat protein gene of the viruses (Supplementary Table 1). RT–PCR was carried out as described above while, qRT–PCR reaction was carried out in a final volume of 25 µL each containing: 12.5 µL of 2 × QuantiFast^{TM} SYBR Green PCR Master mix (Qiagen, Hilden, Germany), 1.0 µL (1 µM µL^{-1}) of each forward and reverse primers (Supplementary Table 1) and 1 µL template (about 120 ng) and 50 U of Revert Aid reverse transcriptase (Fermentas, Maryland, USA). Thermocycling conditions consisted of an initial cDNA synthesis at 42°C for 45 min and initial denaturation step at 95°C for 5 min, followed by 35 cycles of 95°C for 20 s and 60°C for 20 s. Total RNA from negative control and water control (without template) were used to monitor the specificity of the primers and potential contamination within the PCR reagents. Following qRT-PCR, the amplicons were subjected to melt analysis from 60°C to 95°C. The specificity of products of the quantitative RT–PCR was also periodically verified by electrophoresis on a 1.5% agarose gel and by directly sequencing the gel-purified PCR product. To determine the sensitivity of detection, total RNA

was serially diluted from 10^0 to 10^{-10} and, 1 µL of each of the dilutions was subjected to both RT-PCR and qRT–PCR assays. Both RT–PCR and qRT–PCR assays were validated using 1 µL of total RNA isolated from ginger samples collected from different varieties representing the different regions, and ginger germplasm accessions along with positive and negative controls.

3 | Results

3.1 | Electron microscopy

Transmission electron microscopy of leaf dip preparation of infected ginger plant (Fig. 1a) showed the presence of two kinds of particles- striated flexuous particles of about 650 to 850 nm in length (Fig. 1b) and isometric particles with the smooth appearance of about 28–30 nm in diameter indicating the association of two kinds of viruses with the disease.

3.2 | Identification of the viruses through sRNA–seq and RNA–seq

In sRNA–seq, a total of 61,324,157 clean, size selected 16–26 bp reads were obtained. Velvet assembly of these sequences generated 14,408 and 3,275 contigs with length ranging in nucleotide (nt) from 33 to 625 (average 59 nt) and from 37 to 1527 (average 78 nt) in K-mer value 17 and 19 respectively. The CLC *de novo* assembly generated 12,089 contigs with length ranging in nt from 47 to 1221 (average 69 nt). A BLASTX search of the contigs revealed 636 contigs (with maximum contig length up to 462 bp) with maximum amino acid (aa) sequence identity with coat protein (CP), heat shock 70 (HSP70) like protein, p61, and RNA-dependent RNA polymerase (RdRp) genes of different ampeloviruses including grapevine leafroll–associated virus 4 (GLRaV-4) and pineapple mealybug wilt–associated virus 1 (PMWaV-1) and

PMWaV-3. Similarly, 94 contigs (with maximum contig length up to 2223 bp) showed maximum amino acid (aa) sequence identity with RdRp, movement protein and coat protein genes of panico-, machlomo-, pelarspo- and other genera of *Tombusviridae*.

In RNA–seq, 86,695,144 bp clean reads were obtained and *de novo* assembly of RNA reads generated 81,029 contigs with length ranging from 201 nt to 14698 nt (average contig length was 1092 nt). A BLAST search of the contigs against plant viral genomes available in the NCBI database revealed that 737 and 2087 contigs could be annotated through BLASTN and BLASTX respectively. As observed in sRNA–seq, these contigs showed sequence identities with both ampeloviruses (contig length ranging from 325 to 2850 nt) and members of the family, *Tombusviridae* (contig length ranging from 119 to 4270 nt). Thus, results of sRNA–seq and RNA–seq indicate the association of two viruses with the infected ginger plant.

3.3 | Confirmation of the virus belonging to the family *Tombusviridae* through RT–PCR, cloning and Sanger sequencing

Cloning and sequencing of RT–PCR products by the Sanger method produced a sequence of 4143 bases including the 5' and 3' RACE data for the virus (from the infected ginger plant) which was deposited in the GenBank with accession number MN581046. The complete nucleotide sequence of the viral genomic RNA consists of 4143 nt and contains six open reading frames (ORFs) (Fig. 2). The complete genome sequence of the present virus isolate when subjected to BLAST analysis brought several members of the family, *Tombusviridae* as closely related viruses. Further pairwise comparison of the complete nt sequence of virus isolate showed identities ranging from 34 to 47% with different genera of *Tombusviridae*, the highest (46–47%) being seen with members of the genus, *Panicovirus* followed by *Machlomovirus* (46%) (Table

1). The 5´ untranslated region (UTR) of the virus is 27 nt while 3´ UTR is 289 nt. The first ORF of the virus isolate potentially encodes a protein of 43 kDa (p43) that begins at AUG (nt 28–30) and terminates at an UAG codon (nt 1165–1167) (Fig. 2). If this amber codon is readthrough, the ORF extends to a downstream UAG (nt 2761–2763) giving rise to ORF 2 with a protein of 104 kDa (p104). The virus sequence 5' AAA UAG GGG 3' (1162–1170) surrounding the amber stop codon is consistent with the proposed relative efficiency of readthrough sequences (A/C/U) (A/U) A UAG (G/C) (G/A) (Skuzeski et al., 1990), and an exact match with the corresponding sequence present in the members of *Tombusviridae* (Guilley et al., 1985; Nutter et al., 1989; Mollow et al., 2013). The p104 showed homology with the protein of members of the family *Tombusviridae* (Table 1) and p61 contains the motifs characteristic of the viral RdRp including the highly conserved GDD box (Castano & Hernandez, 2005; Stenger & French, 2008; Mollow et al., 2013; Scheets, 2013; Tahir et al., 2017; Monger & Jeffries, 2019). The readthrough protein (p104) showed identities with members of *Tombusviridae*, the highest (50–52% at the nt and 47–49% at the aa level) with members of the genus, *Panicovirus* followed by *Machlomovirus* (52% and 44% at nt and aa level) (Table 1). The RdRp overlap with ORF 3, movement protein (MP) 1, a small 8 kDa protein (p8) (2721–2933 nt). The p8 also showed the highest homology of 26–37% at the aa level with MP1 of panicoviruses (Table 1). The Pfam and NCBI-CDD search showed that p8 contains the motifs characteristic for the *Tombusviridae* MP. The ORF 4 (MP2) initiates with non-canonical start codon ACG (2933–2935) to make 7 kDa protein (p7). The last adenine base of MP1 stop codon is the first base of MP2 start codon (nt 2933). This protein lacks significant similarity to the protein of known viruses. ORF 4 is expressed by a frameshift that would result in a 15 kDa fused protein, p15. The putative p15 protein has aa sequence homology of 21–32% to panicovirus movement protein. The fifth ORF (3117 to 3854) encodes the 27 kDa

CP (p27) that shows significant identities with CP of members of *Tombusviridae*, highest 40% at the nt level with machlomo and panicoviruses. The Pfam and NCBI-CDD search showed that p27 contains the motifs characteristic for the viral CP. The ORF 6 (3223–3621 nt) that overlaps the CP with a protein of 14 kDa (p14) has no significant similarity to the protein of known viruses. The CP is followed by the 3' untranslated region (3' UTR).

Among different genera of *Tombusviridae*, the present virus isolate shared maximum identity of 47% in the complete genome, 52%, 40%, 47% and 46% in the nt sequence of RdRp, CP, MP1 and MP2 genes of panicoviruses (Table 1) while the identity of 46% in the complete genome, 52%, 40%, 51% and 42% in the nt sequence of RdRP, CP, MP1 and MP2 gene of machlomovirus. Similarly, the present virus isolate shared maximum aa identity of 49%, 27%, 37%, and 32% in the RdRp, CP, MP1, and MP2 proteins of panicoviruses (Table 1). The aa sequence identity of the present virus isolate to its closest relatives in far below the threshold (75% in the RdRp and CP) for discriminating distinct species in the family *Tombusviridae* (Scheets et al., 2015) indicating that the virus belongs to a new species in the family *Tombusviridae* for which the name ginger chlorotic fleck-associated virus 1 (GCFaV-1) is proposed.

Phylogenetic analysis of the complete genome of GCFaV–1, RdRp, MP1, and CP with the corresponding region from other viruses under the family *Tombusviridae* revealed the uniqueness of the GCFaV-1 (Fig. 3). Among different genera of the *Tombusviridae*, the present virus isolate showed a close relationship with panicoviruses and machlomoviruses in the subfamily, *Procedovirinae.* All known virus species of a genus showed closer relationships that are well separated from other genera.

3.4 | Confirmation of the virus belonging to the genus *Ampelovirus* through RT–PCR, cloning and Sanger sequencing

Cloning and sequencing of RT–PCR products by the Sanger method produced a sequence of 5514 bases of the virus (from the infected ginger plant) which was deposited in the GenBank with accession number MN581045. The BLASTN and BLASTX search of the virus isolate sequence showed identities with ampeloviruses. The sequenced region (5514 nt) contained coding regions for partial ORF 1b, RdRp (from 1 to1307 nt; 434 aa); ORF 2, small hydrophobic protein (from 1304 to 1462 nt; 52 aa); ORF 3, HSP70 (from 1446 to 3044 nt; 532 aa); ORF4, 61.0 kDa protein (from 3028 to 4668 nt; 546 aa) and partial ORF 5, CP (from 4731 to 5514 nt; 261aa) (Fig. 4). The sequenced region (5514 nt) of the present virus isolate showed identity ranging from 38 to 57% with different ampeloviruses, highest (57%) being observed with GLRaV-4 followed by PMWaV-1 and PMWaV-3 (56%) (Table 2). The nt and aa sequences of different ORFs of the present virus isolate shared identities in RdRp (43–61% and 32–63%), small hydrophobic protein (21–48% and 10–33%), HSP70 (40–60% and 22–48%), p61 (29–54% and 12–50%) and coat protein (35–61% and 21–65%) with 11 distinct species of ampeloviruses (Table 2). Sequence comparison clearly showed that the present virus isolate shared the highest identity in the RdRp and p61 genes with GLRaV-4 followed by PMWaV-3 and PMWaV-1 while the highest identity in the small hydrophobic protein, and HSP70 genes with PMWaV-1 followed by PMWaV-3 and GLRaV-4. The highest identity in the coat protein gene was found with PMWaV-3 followed GLRaV-4 and PMWaV-1 (Table 2). The Pfam and NCBI-CDD search showed the presence of motifs characteristic of viral RdRp, HSP70, viral heat shock protein HSP90 homologue and closterovirus CP in the corresponding proteins of the present virus isolate. The amino acid sequence of gene products such as RdRp, CP, HSP70h differing by more

than 25% is used as the species demarcation criteria in the genus *Ampelovirus* (Fuchs et al., 2020). Thus, maximum aa identity of 48–65% of the present virus isolate in RdRp, CP, HSP70h with different species of ampeloviruses indicate that the virus belongs to a new species in the genus *Ampelovirus* for which the name ginger chlorotic fleck-associated virus 2 (GCFaV-2) is proposed.

Phylogenetic analysis of the sequenced nucleotide region (5514 nt) of GCFaV–2, putative aa sequence of RdRp, HSP70, p61 and CP with the corresponding region from other ampeloviruses revealed closeness of GCFaV-2 with PMWaV-1, PMWaV-3, and GLRaV-4, all members of ampeloviruses in the subgroup II (Fig. 5). The members of the subgroup I of ampeloviruses formed a cluster that is well separated from members of subgroup II. Strains of known species such as GLRaV-4, GLRaV-3, GLRaV-1, and little cherry virus 2 (LChV-2) formed close individual clusters. Phylogenetic tree for all five proteins showed more or less similar topology grouping GCFaV–1, PMWaV-1, PMWaV-3, and GLRaV-4 in a single cluster.

3.5 | Development of diagnostic assays for the detection of GCFaV-1 and GCFaV-2

RT–PCR and qRT–PCR were able to detect both viruses, GCFaV-1, and GCFaV-2 in positive control alone as no product was seen with total RNA from the negative control. The identity of the amplicons was also confirmed by sequencing. The sensitivity of detection of GCFaV-1was up to 10^{-2} dilution of RNA for RT–PCR and 10^{-7} for qRT–PCR while it was 1 µL of undiluted RNA and 10^{-5} for the detection of GCFaV-2 in RT–PCR and qRT–PCR respectively (Supplementary Fig. 1). qRT–PCR was 10,000 times more sensitive than RT–PCR for the detection of both viruses. Validation of both RT–PCR and qRT–PCR assays using total RNA isolated from plants showing characteristic symptoms of the disease in different popular varieties

collected from ginger growing regions of Karnataka and Kerala states of India showed detection of both viruses in all samples except in the negative control (Supplementary Fig. 2). The specificity of the real-time PCR assay was confirmed through melt curve analysis which showed two peaks corresponding to the two viruses. In another study 80 asymptomatic plants of ginger germplasm accessions collected from diverse agro-climatic regions of India and maintained in the Field Gene Bank at the ICAR-Indian Institute of Spices Research, Kozhikode was screened through qRT-PCR. Of these, 38 accessions showed a positive reaction to both viruses indicating that non-expression of symptoms cannot be considered as a fool-proof criterion to categorize a plant as virus-free. The Ct values of virus-positive accessions for GCFaV-1 ranged from 5.23 to 24.08 and GCFaV-2 from 8.07 to 32.87. Forty-two accessions were found free from both viruses.

4 | Discussion

Members of *Tombusviridae* and *Closteroviridae* are reported to infect both dicots and monocots. In the present study, we identified and characterized the complete genome of a novel tombusvirid (tentatively named as ginger chlorotic fleck-associated virus 1; GCFaV-1) and partial genome of an ampelovirus (tentatively name as ginger chlorotic fleck-associated virus 2; GCFaV-2) associated with chlorotic fleck disease of ginger, an important production constraint of ginger in India and other parts of the world (Nambiar & Sarma, 1974; Thomas, 1986). This result correlates with leaf dip transmission electron microscopy of infected ginger leaves which showed the presence of isometric particles of 28–30 nm diameter with a smooth appearance resembling a tombusvirid (Fig. 1c) (Scheets et al., 2015) and striated flexuous rod-shaped particles resembling closterovirid (Fig. 1b). The length of virus particles in the family

Closteroviridae known to range from 1400 to 2200 nm (Martelli et al., 2012; Fuchs et al., 2020). The smaller length of particles observed in the present study may represent broken particles.

The identity of the viruses was established by sRNA–seq and RNA–seq followed by verification through RT-PCR and Sanger sequencing. The complete genome of GCFaV-1 shared 34–47% with members of *Tombusviridae* and highest aa identities of 49% and 27% in the RdRp and CP of this virus fall in the range of <75% among known species of *Tombusviridae* (Scheets et al., 2015; Tahir et al., 2017; Zhuo et al., 2018). Therefore GCFaV-1 should be considered a member of a new species in the family *Tombusviridae*. The sequence analyses provide evidence of the close relationship of GCFaV-1 with members of panicoviruses and machlomoviruses (Table 1; Fig. 3). The GCFaV-1 genome has highly conserved sequence motifs in the RdRp, MP1 and CP characteristic of the members of *Tombusviridae* (Castano & Hernandez, 2005; Stenger & French, 2008; Mollow et al., 2013; Scheets, 2013; Tahir et al., 2017; Monger & Jeffries, 2019). Among all, GCFaV-1 protein sequences are most similar to their counterparts in panicoviruses (Table 1) although complete genome sequence identity is only 46–47% with members of the genus *Panicovirus*. Phylogenetic studies using complete genome and different ORFs are in agreement with percent na/aa identity that showed distinct nature of GCFaV-1 and its close clustering with members of panico- and machlomoviruses (Fig. 3)

The partial genome of GCFaV-2 shared 38–57% in the sequenced region with the members of ampeloviruses in the family *Closteroviridae*. The aa sequence identities of 32–63% in the RdRp, 22–47% in the HSP70 and 21–65% in the CP between GCFaV-2 and other ampeloviruses fall in the range of <75% among known species of the genus *Ampelovirus* (Martelli et al., 2102). Therefore, GCFaV-2 should be considered a member of a new species in the genus *Ampelovirus*. The GCFaV-2 genome also has highly conserved motifs in the RdRp,

HSP70, p61, and CP that are characteristic to the members of *Closteroviridae* (Ling et al., 2004; Martelli et al., 2012; Ito & Nakaune, 2016; Fuchs et al., 2020). Among all, GCFaV-2 sequences are most similar to their counterparts in the subgroup II of ampelovirues (Table 2; Fig. 5) (Martelli et al., 2012; Ito & Nakaune, 2016; McLaughlin et al., 2017). Phylogenetic studies are also in agreement with percent nt/aa identity that showed distinct nature of GCFaV-2 and its close clustering with members of the subgroup II of ampeloviruses (Fig. 5).

This is the first report of the occurrence of viruses belonging to *Tombusviridae* and *Ampelovirus* infecting ginger. Ginger is a monocot belonging to the family *Zingiberaceae*. Members of panicoviruses and machlomoviruses appear to be closer to GCFaV-1 are also known to infect monocots (Scheets, 2013; Tahir et al., 2017) while members of the subgroup II of ampeloviruses that infect both monocots and dicots appears to be closely related viruses to GCFaV-2 (Martelli et al., 2102). No vector is known for GCFaV-1 and GCFaV-2 and transmission mainly occur through vegetative propagating material. A few of the ampeloviruses are reported to be transmitted through mealybugs. Mealybug infestation of ginger is rarely seen and its role in the transmission of GCFaV-2 is yet to be studied. Similarly, many viruses in the family *Tombusvridae* are known to be transmitted mechanically. However, information on the mechanical transmission of GCFaV-1 is not reported so far. Hence efforts should be done on isolate GCFaV-1 through mechanical inoculation on to virus-free ginger. This would help in the comparative evaluation of ginger plants with GCFaV-1 alone and mixed infection by GCFaV-1 and GCFaV-2 to know the possible synergism between both viruses.

In the absence of a resistant variety, the use of virus-free plants is the best option to manage viral diseases. Ginger is mainly propagated by vegetative means that favours virus concentration over time and lead to virus spread to newer areas when used for planting. Ginger is

also propagated through micropropagation and microrhizomes (Inden et al., 1988; Balachandran et al., 1990; Sharma & Singh, 1995) for which selection of virus-free mother plants is important. As observed in the present study symptoms alone cannot serve as criteria to select virus-free plants for propagation. A rapid, high throughput, sensitive assay targeting virus coat protein, and nucleic acid is necessary for accurate detection of viruses (Jeong et al., 2014). The RT–PCR, and qRT–PCR assays developed in the present study can be used for sensitive detection of both viruses. Among the two, quantitative RT–PCR was 10,000 times more sensitive than conventional RT–PCR. However, so far no information on the natural variability of GCFaV-1 and GCFaV-2 are available. Hence to have a robust, fool-proof diagnostics, a large number of virus isolates representing different varieties and agro-climatic regions need to be sequenced and primers designed to the conserved region should be used in the detection assays.

Acknowledgements

Authors are thankful to Science and Engineering Research Board (SERB), Government of India for the funding (EMR/2016/001135); Head (Division of Crop Protection), Distributed Information sub-center and Director, ICAR-Indian Institute of Spices Research, Kozhikode, Kerala, India for facilities and Virology Unit, ICAR-Indian Agricultural Research Institute, New Delhi, India for electron microscopy work.

ORCID

ORCID.org/0000-0002-2637-1411 (A I Bhat)

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest

The authors declare that they have no conflict of interest.

Involvement of human participants and /or animals

The present research did not involve any experimentation on humans or animals.

References

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., & Boonham, N. (2009). Next-generation sequencing and metagenomic analysis: A universal diagnostic tool in plant virology. *Molecular Plant Pathology, 10*, 537–545.

Balachandran, S.M., Bhat, S.R., & Chandel, K.P.S. (1990). *In vitro* clonal multiplication of turmeric (*Curcuma* spp.) and ginger (*Zingiber officinale* Rosc.). *Plant Cell Reports, 8*, 521–524.

Barba, M., Czosnek, H., & Hadidi, A. (2014). Historical Perspective, Development and Applications of Next-Generation Sequencing in Plant Virology. *Viruses, 6*, 106–136.

Cao, M., Zhang, S., Li, M., Liu, Y., Dong, P., Li, S., Kuang, M., Li, R., & Zhou, Y. (2019). Discovery of four novel viruses associated with flower yellowing disease of green Sichuan pepper (*Zanthoxylum armatum*) by virome analysis. *Viruses, 11*, 696. Doi:10.3390/v11080696.

Castano, A., & Hernandez, C. (2005). Complete nucleotide sequence and genome organization of Pelargonium line pattern virus and its relationship with the family *Tombusviridae. Archives of Virology, 150*, 949–965

Deepa, K., Sheeja, T.E., Santhi, R., Sasikumar, B., Cyriac, A., Deepesh, P.V., & Prasath, D. (2013). A simple and efficient protocol for isolation of high quality functional RNA from different tissues of turmeric (*Curcuma longa* L.). *Physiological and Molecular Biology of Plants*, *20*, 263–271.

FAO. (2020). http://www.fao.org/faostat/en/#data/QC/visualize.

Fan, G.Q., Wang, J.Y., Guo, X.Q., & Zhu, H.C. (1999). Identification of *Tobacco mosaic virus* infecting ginger. *Journal of Shandong Agricultural University, 30*, 249–255.

Fuchs, M., Bar-Joseph, M., Candresse, T., Maree, H.J., Martelli, G.P., Melzer, M.J., Menzel, W., Minafra, A., Sabanadzovic, S. & ICTV Report Consortium. (2020). ICTV Virus Taxonomy Profile: *Closteroviridae*, *Journal of General Virology*, *101*, 364–365.

Guilley, H., Carrington, J.C., Balazs, E., Jonard, G., Richards, K., & Morris, T.J. (1985). Nucleotide sequence and genome organization of carnation mottle virus RNA. *Nucleic Acids Research*. *13*, 6663–6667.

Inden, H., Asahira, T., & Hirano, A. (1988). Micropropagation of ginger. *Acta Horticulturae*, *230*, 177–184.

International Committee on Taxonomy of Viruses (ICTV). (2019). *Virus Taxonomy: 2019 Release*. Available at: http://www.ictvonline.org/virusTaxonomy.asp.

Ito, T., & Nakaune, R., (2016). Molecular characterization of a novel putative ampelovirus tentatively named grapevine leafroll-associated virus 13. *Archives of Virology*, *161*, 2555–2559.

Jeong, J.J., Ju, H.J., & Noh, J. (2014). A review of detection methods for the plant viruses. *Research in Plant Disease, 20,* 173–181.

Kitajima, E.W., & Pozzer, L. (1994). Disease caused by viruses on sweet potato, beetroot, Dioscorea, ginger and yam, *Informe Agropecuario-Belo-Horizonte, 17,* 39–41.

Komínek, P., Glasa, M., & Bryxiova, M. (2005). Analysis of the molecular variability of grapevine leafroll-associated virus 1 reveals the presence of two distinct virus groups and their mixed occurrence in grapevines. *Virus Genes, 31,* 247–255.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., & Simon, R. (2009). Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology, 388,* 1–7.

Ling, K-S., Zhu, H-Y., & Gonsalves, D. (2004). Complete nucleotide sequence and genome organization of Grapevine leafroll-associated virus 3, type member of the genus *Ampelovirus*. *Journal of General Virology, 85,* 2099–2102.

Martelli, G.P., AbouGhanem-Sabanadzovic, N., Agranovsky, A.A., Al Rwahnih, M., Dolja, V.V., Dovas, C.I., Fuchs, M., Gugerli, P., Hu, J.S., Jelkmann, W., Katis, N.I., Maliogka, V.I., Melzer, M.J., Menzel, W., Minafra, A., Rott, M.E., Rowhani, A., Sabanadzovic, S., & Saldarelli, P. (2012). Taxonomic revision of the family Closteroviridae with special reference to the grapevine leafroll-associated members of the genus Ampelovirus and the putative species unassigned to the family. *Journal of Plant Pathology, 94,* 7–19.

McLaughlin, M., Lockhart, B., Jordan, R., Denton, G., & Mollov, D. (2017). Complete nucleotide sequence of clematis chlorotic mottle virus, a new member of the family *Tombusviridae*. *Archives of Virology, 162,* 1373–1379.

Milne, R.G. (1984). Electron microscopy for the identification of plant viruses in *in-vitro* preparation. *Methods in Virology, 7*, 7–120.

Mollov, D., Lockhart, B., & Zlesak, D.C. (2013). Complete nucleotide sequence of *Rosa rugosa* leaf distortion virus, a new member of the family *Tombusviridae*. *Archives of Virology*, *158*, 2617–2620.

Monger, W., & Jeffries, C. (2019). A new virus, classifiable in the family *Tombusviridae*, found infecting *Solanum tuberosum* in the UK. *Archives of Virology*, *163*, 1585–1594.

Nambiar, K.K.N., & Sarma, Y.R. (1974). Mosaic disease of ginger. *Arecanut and Spices Bulletin*, *6*, 3–4.

Nutter, R.C., Scheets, K., Panganiban, L.C., & Lommel, S.A. (1989). The complete nucleotide sequence of the maize chlorotic mottle virus genome. *Nucleic Acids Research, 17*, 3163–3177.

Peltola, H., Soderlund, H., & Ukkonen, E. (1984). SEQAID: a DNA sequence assembling program based on a mathematical model. *Nucleic Acids Research, 12*, 307–321.

Quan, P., Briese, T., Palacios, G., & Lipkin, W.I. (2008). Rapid sequence-based diagnosis of viral infection. *Antiviral Research*, *79*, 1–5.

Ravindran, P.N., & Babu, K.N. (2005). Ginger: The Genus *Zingiber*. Boca Raton, Florida: CRC Press. 545 pp.

Scheets, K. (2013). Infectious transcripts of an asymptomatic panicovirus identified from a metagenomic survey. *Virus Research*, *176*, 161–168

Scheets, K., Jordan, R., White, K.A., & Hernandez, C. (2015). Pelarspovirus, a proposed new genus in the family *Tombusviridae*. *Archives of Virology, 160*, 2385–2393.

Sharma, T.R., & Singh, B.M. (1995). *In vitro* microrhizome production in *Zingiber officinale* Rosc. *Plant Cell Reporter, 15*, 274–277.

Skuzeski, J.M., Nichols, L.M., & Gesteland, R.F. (1990). Analysis of leaky viral translation termination codons *in vivo* by transient expression of improved/I-glucuronidase vectors. *Plant Molecular Biology*, 15, 65–79.

So, I.Y. (1980). Studies on ginger mosaic. *Korean Journal of Plant Protection, 19*, 67–72.

Stenger, D. C., & French E-R. (2008). Complete nucleotide sequence of a maize chlorotic mottle virus isolate from Nebraska. *Archives of Virology*, *153*, 995–997.

Tahir, M.N., Lockhart, B., Grinstead, S., & Mollov, D. (2017). Characterization and complete genome sequence of a panicovirus from Bermuda grass by high-throughput sequencing. *Archives of Virology, 162*, 1099–1102.

Thomas, J.E. (1986). Purification and properties of *Ginger chlorotic fleck virus*. *Annals of Applied Biology, 108*, 43–50.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., & Higgins, D.G. (1997). The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research, 24*, 4876–4882.

Zhuo, T., Zhu, Li-J., Lu, C-C., Jiang, C-Y., Chen, Zi-Y., Zhang, G., Wang, Z-H., Jovel, J. & Han, Y-H. 2018. Complete nucleotide sequence of jasmine virus H, a new member of the family *Tombusviridae*. *Archives of Virology*, *163*, 731–735.

Figure legends

Fig. 1(a) Symptoms of chlorotic fleck disease of ginger (b) Transmission electron micrograph of leaf dip preparations of infected ginger leaf showing flexuous particles, and (c) isometric particles.

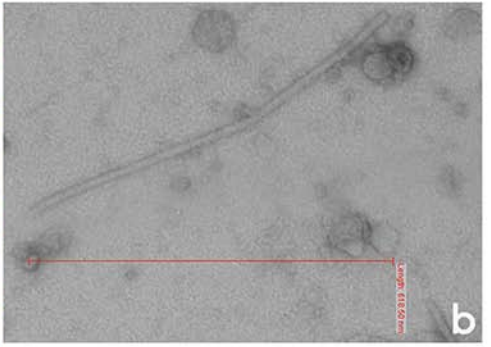
Fig. 2 (a) Putative genome map of ginger chlorotic fleck-associated virus 1 (GCFaV-1) infecting ginger. The different open reading frames (p48, p61, p8, p7, p27, p15) are indicated. The primers designed based on the highthroughput sequence data and their orientation is indicated through arrows. The sequence, orientation, and expected product size in RT–PCR is provided in Supplementary Table 1. (b) Agarose gel electrophoresis of RT–PCR products obtained with different primer pairs in the order listed in Supplementary Table 1.

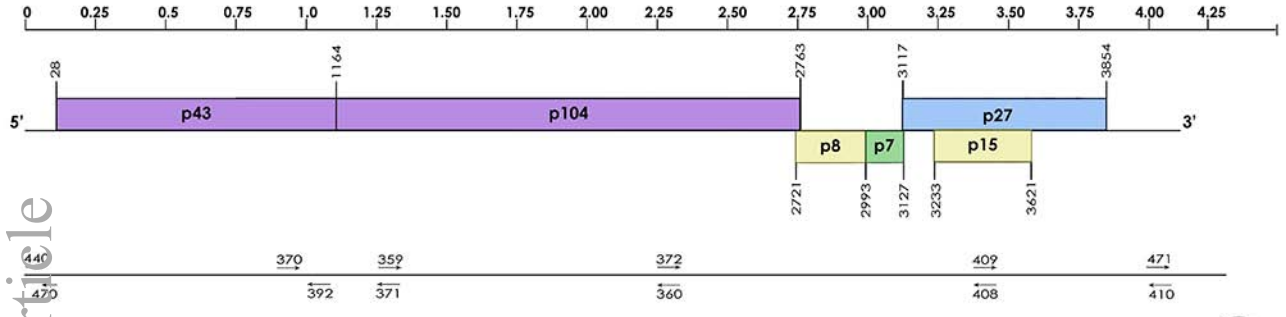
Fig. 3 Phylogenetic tree based on the complete genome (4143 nt) of ginger chlorotic fleck–associated virus 1(GCFaV-1) with viruses belonging to different genera and species of *Tombusviridae*. The designation given to each of the isolates and their GenBank accession numbers are provided in Supplementary Table 2.

Fig. 4 (a) Putative partial genome map of ginger chlorotic fleck-associated virus 2 (GCFaV-2) infecting ginger. The different open reading frames (RdRp, p8, HSP70, p61, and coat protein) are indicated. The primers designed based on the highthroughput sequence data and their orientation is indicated through arrows. The sequence, orientation, and expected product size in

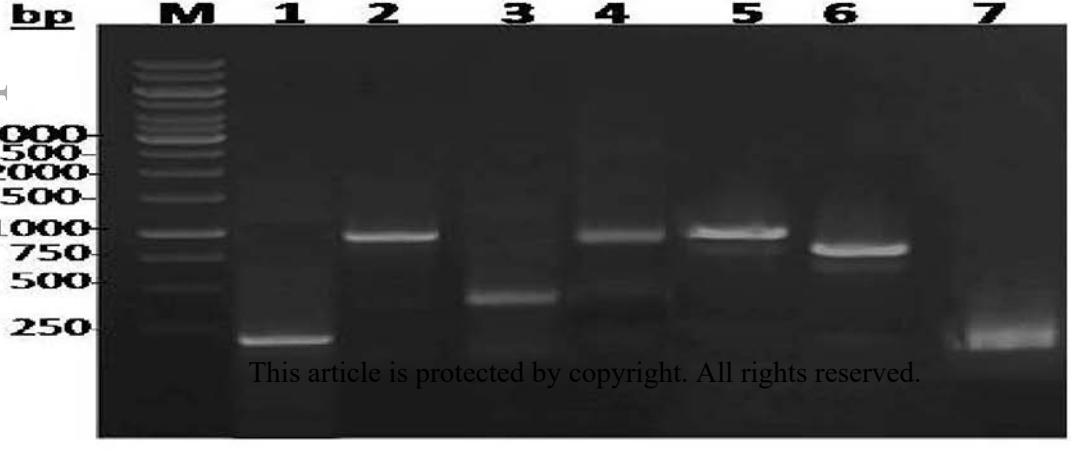
RT–PCR is provided in Supplementary Table 1. (b) Agarose gel electrophoresis of RT–PCR products obtained with different primer pairs in the order listed in Supplementary Table 1.

Fig. 5 Phylogenetic tree based on the sequenced region (5514 nt) of ginger chlorotic fleck–associated virus 2 (GCFaV-2) with the corresponding region of viruses belonging to different species of ampeloviruses. The designation given to each of the isolates and their GenBank accession numbers are provided in Supplementary Table 3.

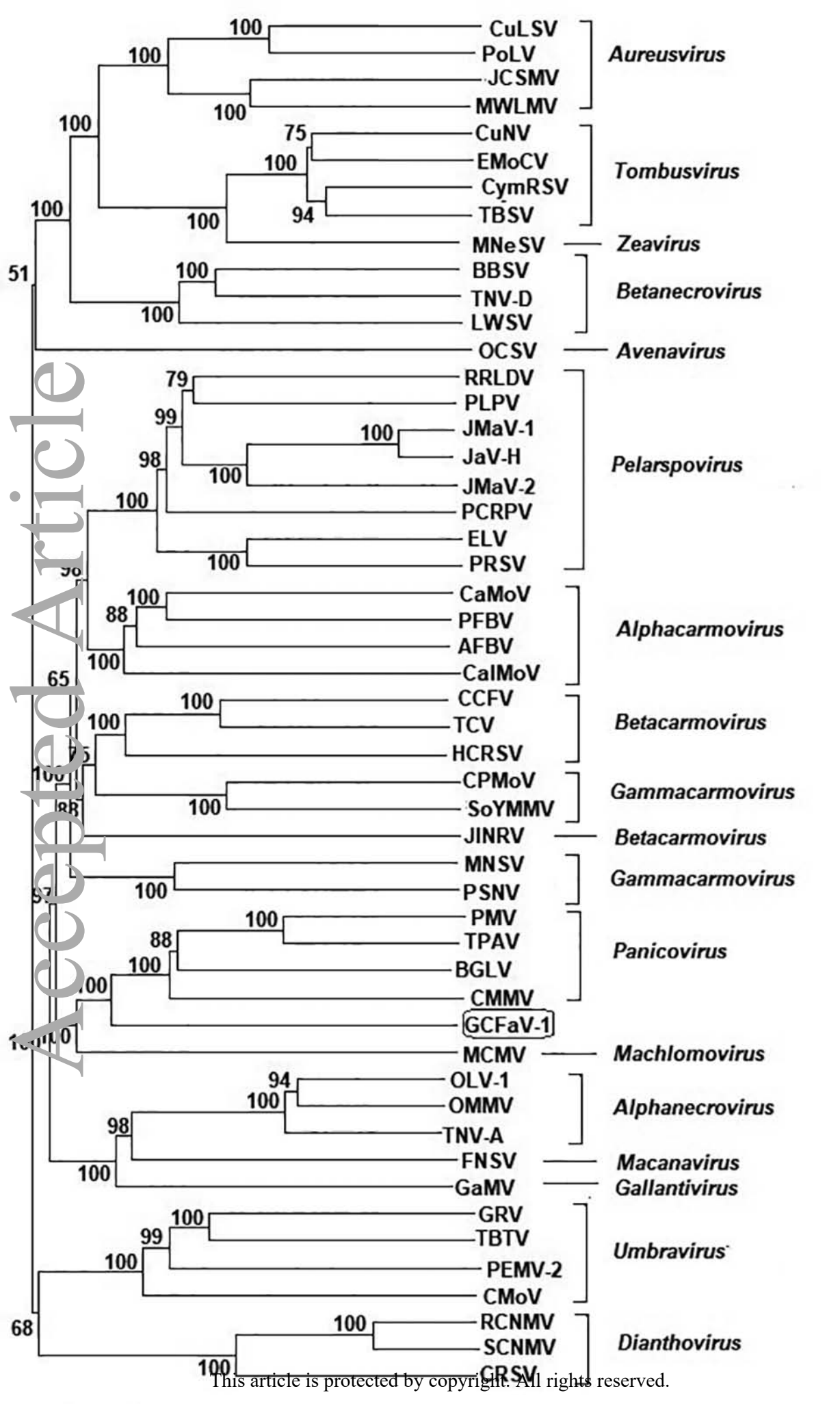


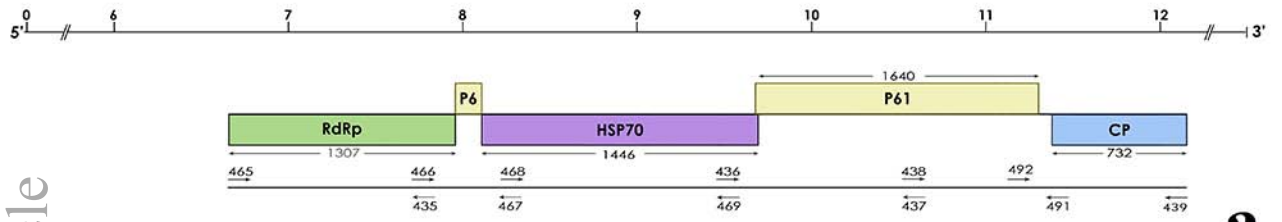


a

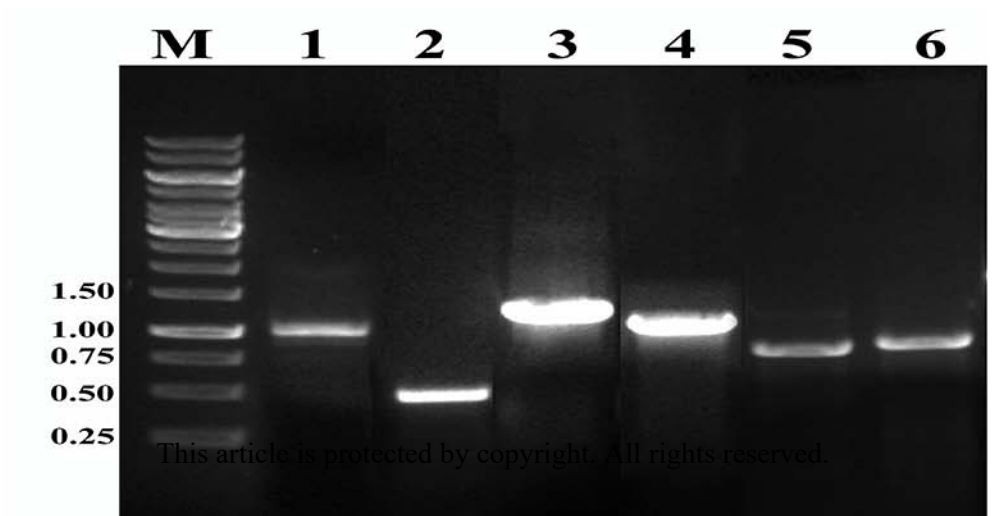


b





a



b