



Genome characterization of citrus yellow vein-clearing virus: limited heterogeneity of viral genomes in *Mandarivirus*-infecting different citrus species

Ram Prasanna Meena^{1,3} · K. Prabha² · V. K. Baranwal¹

Received: 15 December 2018 / Accepted: 18 August 2019 / Published online: 30 August 2019
© King Abdulaziz City for Science and Technology 2019

Abstract

Citrus yellow vein-clearing virus (CYVCV) is a mandarivirus infecting citrus producing yellow vein-clearing symptoms. The leaf samples collected during surveys of different citrus-growing areas in India exhibited diverse symptoms and 40% of the plants were positive for CYVCV in RT-PCR, indicating the wide distribution of the virus in India. It was reported for first time that CYVCV infects kinnow mandarin and sweet oranges and produces chlorotic ringspots symptoms identical to Indian citrus ringspot virus (ICRSV). The complete genome sequences of CYVCV infecting four citrus cultivars have been deciphered through overlapping primers. All the four genomes comprise of 7531 nucleotides excluding the 3' poly (A) tail. The sequence identity of genomes of four CYVCV isolates in the present study ranged from 95.2 to 99.8% with genome sequences of 31 CYVCV isolates available in public domain and the mean genomic diversity was 0.017, indicating low level of heterogeneity. The phylogenetic analysis revealed that CYVCV isolates from India, Pakistan, and Turkey were clustered in the same clad apart from China isolates. The least normalized dN/dS mean value (0.092) indicated that RdRP region evolved under relatively stronger selection constraints than the other five coding regions of CYVCV. The four intragenic putative recombination events detected in RDP4 program occurred naturally in CYVCV genome, indicating the evolutionary progress of the virus. Tajima's and Fu and Li's *D* parameters were performed using genomic sequences in DnaSP v5 program and the retrieved negative values indicated the presence of limited genetic variability in CYVCV genomes. To the best of our knowledge, this is the first comprehensive report on molecular characterization of CYVCV from India. It will be helpful in understanding the evolutionary relationship of CYVCV and ICRSV.

Keywords Citrus yellow vein-clearing virus · Mandarivirus · Detection · Genetic heterogeneity · Recombination

Introduction

Citrus is one of the most important fruit crop grown worldwide. The crop is invaded by large number of viruses, bacteria, viroids, phytoplasmas, and fungal pathogens, but huanglongbing (HLB) (caused by *Candidatus Liberibacter*

asiaticus) and viruses are the major constraints in citrus production. Among the viral diseases, yellow vein clearing is an emerging and rapidly spreading disease in major citrus-growing areas of South Asian countries. The yellow vein-clearing disease was first reported on lemon and sour orange from Pakistan in 1988, later in 1997 on Etrog citron from Abohar, India (Catara et al. 1993; Ahlawat and Pant 2003) and recently from Turkey, China, and Iran (Chen et al. 2014; Bani and Aghajanzadeh 2017). In India, CYVCV came into light when a mixture of decorated and undecorated flexuous virus particles were observed under transmission electron microscopy (TEM) in ringspot symptomatic samples, processed with the heterologous antibodies of ICRSV (Byadgi and Ahlawat 1995). The TEM examined samples of Etrog citron showed the presence of flexuous filamentous virions of CYVCV with modal length of 685 nm and diameter of 13–14 nm (Alshami et al. 2003). In 2012, the first complete

✉ V. K. Baranwal
vbaranwal2001@yahoo.com

¹ Advanced Centre for Plant Virology Division of Plant Pathology, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India
² ICAR-Directorate of Floricultural Research, Pune 411005, India
³ Present Address: ICAR-Directorate of Medicinal and Aromatic Plants Research, Anand, Gujarat 387310, India

genome sequence of CYVCV was characterized from Turkey and the genome nucleotides showed ~74% sequence identity with ICRSV-K1 (Loconsole et al. 2012). The CYVCV was identified as a distinct species, closely related to ICRSV under the genus *Mandarivirus* in family *Alphaflexiviridae* of the order *Tymovirales* (Adams et al. 2014). In an earlier study, no amplification was observed in some of the ringspot exhibiting leaf samples in RT-PCR assay using the coat protein specific primers of ICRSV (Prabha and Baranwal 2012). This indicated that either there was high variability in CP of ICRSV or it could be a new species of *Mandarivirus*.

CYVCV produced typical vein-clearing symptoms on Etrog citron (Alshami et al. 2003). The two virus species, CYVCV and ICRSV, share common natural and experimental host, mechanical transmission, genome organisation, and viral morphology, and in nature, they can be present either singly or together in single host. Thus, it was not possible to distinguish them based on the TEM observation, symptoms and bioassay approaches. Therefore, for molecular detection and diagnosis, and to understand the complete characteristic of the virus and demarcation of exact species, genome sequencing is imperative. In this order, complete genome sequence of CYVCV from four citrus cultivars has been deciphered following the overlapping primer strategy. The molecular diversity, population demography, selection

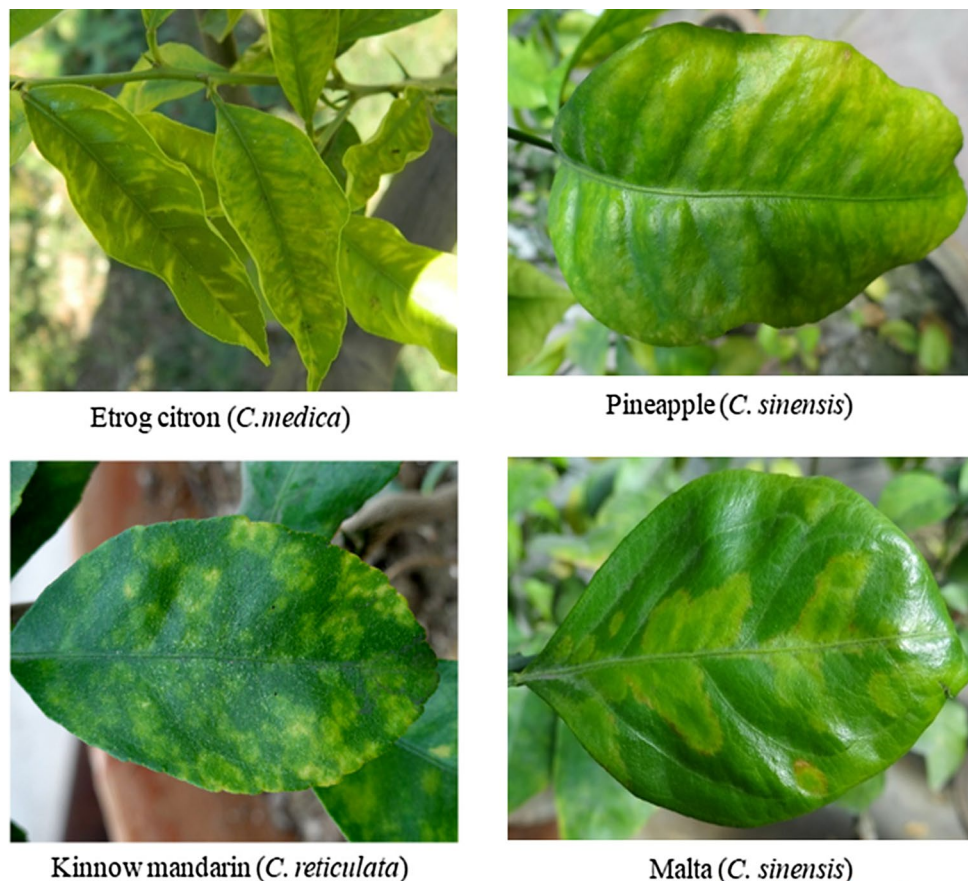
pressure, and recombination analysis were performed for better understanding of evolution and distinctiveness of citrus yellow vein-clearing virus.

Materials and methods

Survey, virus sources, and transmission electron microscopy

Surveys were conducted during 2013–2014 in 18 citrus orchards of major citrus-growing areas of Punjab, Maharashtra, Manipur, West Bengal, and Rajasthan in India and representative of 60 apparently infected symptomatic samples of different citrus cultivars was collected. Samples of four citrus cultivars viz, Etrog citron (*Citrus medica* var. Etrog) from Regional Fruit Research Station (PAU), Abohar; Pineapple sweet orange (*Citrus sinensis* L. Osbeck), PAU, Ludhiana, Punjab; Malta (*C. sinensis* L.) from Royal farm, Sri Ganganagar, Rajasthan and kinnow mandarin (*Citrus reticulata* Blanco) from Pune (Fig. 1) were selected to determine the full genome sequence of CYVCV. The fresh leaf samples of these plants were examined under TEM following the standard leaf dip method (Milne and Luisoni 1977). For all the selected four samples, bio indexing was repeated

Fig. 1 Symptoms on citrus host cultivars selected for determining the complete genome sequence of citrus yellow vein-clearing virus



twice. The scions of aforementioned were grafted, each in three replications, on 1-year-old sweet orange and kinnow mandarin seedlings raised from the seeds collected from disease-free certified mother plants. The observations on symptom development and disease progress were recorded at regular interval after grafting. A multiplex RT-PCR, developed by Meena and Baranwal (2016), was performed to screen these four samples for major citrus viruses and greening bacteria and these samples were confirmed free of virus infections other than CYVVCV.

RNA extraction and genome-sequencing strategy

Total RNA was extracted from 100 mg each of symptomatic leaf tissues from the four citrus cultivars using the commercial RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. The quality and quantity of RNA were measured using spectrophotometer (Nanodrop, UK) and gel electrophoresis.

First-strand cDNA was synthesized using 5 µL of RNA (1 µg) and 1 µL (10 mM) of the anti-sense primers, and the mixtures was incubated at 70 °C for 4 min and immediately chilled on ice for 5 min. Followed by remaining components were added and mixed gently by pipetting. The mixed contents of the tube incubated at 42 °C for 60 min and reaction terminated as per the instruction of SMART Scribe RT kit (TaKaRa/Clontech, Japan). To determine the complete

genome sequence of CYVVCV, the overlapping genomic regions were amplified using nine pair of primers designed from the reference sequence of CYVVCV Y1 (JX040635.1) (Table 1). The PCR assays were performed with a volume of 50 µL to amplify all the fragments, containing 5 µL 10 × Ex *Taq* buffer, 4 µL (2.5 mM) dNTP mix, 1.0 µL (25 mM) MgCl₂, 0.5 µL (5 u/µL) *Taq* DNA polymerase (TaKaRa, Japan), 1.0 µL (10 µM) of each forward and reverse primers, 4 µL of template Cdna, and sterile-distilled water in thin-walled 200 µL tubes. The 5' and 3' termini of the virus isolates were determined using the 5'/3' RACE kit (Roche, Switzerland) with specific CYVR-3'F/CYVR-5'R and oligo (dT) primers.

The amplified PCR products were electrophoresed on 1.2% agarose gel and purified using QIAquick gel elution kit (Qiagen, Hilden, Germany). The purified product was cloned with T&A-cloning system (Real Biotech Corporation, Taiwan) and minimum two positive clones for each amplicon sequenced in both directions (Eurofins Labs, Bangalore, Karnataka and Xcelris Genomics, Ahmedabad, India) for the full-genome sequencing.

In silico sequence analysis

BLASTN analysis (<https://blast.ncbi.nlm.nih.gov>) was performed to determine the retrieved sequences from the overlapping amplicons. This was followed by assembling of

Table 1 Name, polarity, sequence, and other properties of the sense and anti-sense primers and expected size of amplicon in RT-PCR generated by the species specific primers used in this study to determine complete nucleotide sequence of citrus yellow vein-clearing virus

Primer name	Polarity	Primer sequence (5'-3')	Length (nt.)	Annealing temp. (°C)	GC (%)	Amplified region	Expected amplicon size
CYVR-01F	+	ATGGCCACTATTAGGGGTGC	20	60	52.4	R1	1042 bp
CYVR-01R	-	CGAAGAAATAGTTGGCCATATG	22		60		
CYVR-02F	+	GGAGCTTGAACATACACCGG	21	62	55	R2	807 bp
CYVR-02R	-	CAGTCTTCGTCGTCACGCGTG	20		55		
CYVR-03F	+	GTTCTTCGGCCACCCACC	19	61	50	R3	906 bp
CYVR-03R	-	GAGACAGTTGTCTGGGCCG	19		45.2		
CYVR-04F	+	GCTACCTCGCTGAAGGTCTCG	21	62	57.5	2-4	1780 bp
CYVR-04R	-	GGCGCATGTACCGAGCCATAG	21		47.8		
CYVR-05F	+	CAATGGGTCAAGAAAGTCGAG	21	60	47.6	R4	417 bp
CYVR-05R	-	GTCGAAGGTCCGCCCTTCA	19		60		
CYVR-06F	+	GGAGATTATCGCCGGGTATCG	21	62	57.1	4-6	1763 bp
CYVR-06R	-	GTAGGTGCCTGTGCAACTCG	20		60		
CYVR-07F	+	ATGAGCTTCGACTACACTCAC	21	60	47.6	Cp	978 bp
CYVR-07R	-	TTAGATGTTGAAAGGGGTCGG	21		47.6		
CYVR-08F	+	GACCTCACGATGGACCACG	19	58	63	NaBP	407 bp
CYVR-08R	-	GGCACTAACAGGATTTTCGGTTAG	24		45.8		
CYVR-3'F	+	CCCTAGAGAGACACCCTAC	19	56	57.9	3'	577 bp + poly A
CYVR-5'R	-	GGTTCGATGTGTTGGTTYAG	20	58	49	5'	440 bp

Positions of degenerate nucleotide indicated by the following codes: M = A or C; Y = C or T residues

complete genome sequences of four CYVCV isolates using BioEdit 7.1.3.0 (Hall 1999). The conserved protein domains were identified by submitting the sequences to conserved domain database (CDD) of NCBI (<http://www.ncbi.nlm.nih.gov>), while the encoded proteins and cleavage sites were annotated using conserved domain architecture retrieval tool (CDART). For the comparative analysis of complete genome sequences four CYVCV isolates from this study, thirty-one CYVCV, two ICRSV (HQ324250; AF406744), one potato virus X (KJ434601), and one allium virus X (FJ670570) isolates were downloaded from NCBI database and used in the study. Multiple nucleotide and amino acid sequence alignment was performed using CLUSTAL W (Thompson et al. 1994) with default setting. Mean diversity of the population and the best fitted nucleotide model was determined with the MEGA 6.06. The phylogenetic evolutionary relationship among the sequences was inferred based on the complete nucleotide sequences of CYVCV across the world including four from this study and genomes of ICRSV. Genome sequence of PVX was an out-group member of the same family *Alphaflexiviridae*. The phylogenetic trees were constructed using the maximum likelihood, neighbor joining, and maximum evolution methods with 1000 bootstrap replications in MEGA 6.06 (Tamura et al. 2013).

Recombination and estimation of selection pressure

The recombination between citrus infecting mandariviruses was analyzed using the Recombination Detection Program v4.56 (RDP4) software (Martin et al. 2015). The putative recombination events, likely major and minor parents and breakpoints, were identified using the different algorithms viz., RDP, GENECONV, CHIMAERA, MAXCHI, BOOTSCAN, SISCAN, LARD, 3SEQ, and PhylPro implemented in the RDP4 software package with the default settings. The recombination events detected by at least three different methods were accepted.

The aligned sequences of coding regions were used to determine the natural selection pressure at molecular level by Datamonkey software (<http://www.datamonkey.org/>). The rate of non-synonymous substitutions per non-synonymous site (dN) to synonymous substitutions per synonymous site (dS) was analyzed separately for positive, negative, and neutral selection. The selection pressure was considered negative or purifying ($dN/dS < 1$), neutral ($dN/dS = 1$) and positive or adaptive selection ($dN/dS > 1$) for each data set. The site-specific selection pressures acting on different genes were determined by codon-based likelihood algorithms SLAC, FEL, and IFEL with the significant level of P value at 0.01 and a Fast, Unconstrained Bayesian Approximation (FUBAR) program using the GTR-nucleotide model with 0.9 threshold value implemented by Datamonkey server (<http://datamonkey.org>)

also used for inferring the site-specific selection (Murrell et al. 2013).

Population demographic analysis

Genetic differentiation and population demographic analysis of CYVCV isolates, using different permutation statistical tests (Hudson 2000), was performed with the default setting of DnaSPv5 (Librado and Rozas 2009). The demography of the population considered number of haplotypes, haplotype diversity, Tajima's D test of neutrality, and Fu and Li's D and F statistical tests. The number of haplotypes depends on the nucleotide variation. Hundred percent nucleotide sequence identity was considered as single haplotype. Nucleotide diversity estimated the pair wise differences, while haplotype diversity referred to the number of haplotypes in population.

Results

Field survey, TEM, and RT-PCR

The leaf samples of citrus cultivars collected during the surveys from different citrus orchards exhibited vein clearing, ringspots, chlorosis, water soaking, and leaf distortion symptoms. The samples collected from different parts of the country were tested by RT-PCR using CYVCV-specific primers, and in 40% of the leaf samples, CYVCV was detected. The four citrus cultivars exhibiting diverse symptoms viz., Etrog citron; conspicuous vein clearing, Pineapple; ringspot symptoms, Malta; chlorotic symptoms with prominent rings; and Kinnow mandarin large number of yellow rings spots (Fig. 1) were selected for complete genome sequencing of CYVCV.

All the four selected citrus samples when examined under transmission electron microscopy revealed the presence of flexuous virus particles in leaf dip preparation. The virion diameter ranged from 13 to 14 nm with a length of 580–785 nm confirmed the association of *mandarivirus* with the disease. Furthermore, amplification in RT-PCR assay using the CYVCV-specific primers (CP region) and sequencing of amplified product (978 bp) confirmed the presence of CYVCV in the four samples. These four samples were positive for CYVCV and were free from ICRSV, CTV, CYMV, and HLB bacterium in multiplex RT-PCR assay (Meena and Baranwal 2016). In graft inoculation study, symptoms of ringspots were observed on the kinnow mandarin and sweet orange samples.

Whole genome sequencing and sequence properties of CYVCV

The complete genomic information of CYVCV was determined from the four different citrus hosts showing variable symptoms. Nine overlapping primers pairs were used in RT-PCR and RACE profiling covered entire genome of the CYVCV, as listed in Table 1. The assembled sequence contigs of nine clones showed that the four CYVCV isolates genome consisted of 7531 nucleotides (nt) excluding 3'-poly (A) tail. These assembled sequences were deposited in the NCBI GenBank and accession numbers given as CYVCV_ECAI—KT696510; CYVCV_RMGI—KT696511; CYVCV_PALI—KT696512; and CYVCV_KPMI—KT696513. The BLASTN analysis of four CYVCV isolates showed 96–98% sequence identity with 99–100% query coverage (qc) to the reference sequence of CYVCV Y1 (JX040635.1) and 74% sequence identity with ICRSV isolates (AF406744) with 93% query coverage.

All the four genome sequences of CYVCV followed typical genomic organisation of mandariviruses, consisting of six ORFs with 5' untranslated region (UTR) of 80 nt and 3' UTR of 36 nt followed by a poly (A) tail. The genome size of CYVCV isolates was 7531 nt; however, the CYVCV Y1 isolate of Turkey (JX040635.1) was of 7529 nt length. The two extra bases (CC) and located at position 28 and 29 of genome and the 5' UTR of genome started with GAAAA sequence signature. ORF1-encoding RNA-dependent RNA polymerase (RdRP) exhibited all the replication-associated domains including methyltransferase, helicase, AlkB, RdRP_2, and GDD motif located at aa 1520–1522 sequence of CYVCV isolates. The multiple alignments of the RdRP-deduced amino acid sequences of CYVCV and ICRSV isolates revealed the significant variability in the core region. ORF2, ORF3, and ORF4 had partial overlapping and were identified to encode the triple gene block (TGB) proteins. ORF2 encode a protein known as TGB1 in CYVCV containing the domains of RNA viral helicase1 (pfam01443), NTPase-binding motifs, GAGKT (aa 30–34) and DEY (aa 76–78) which served as ATPase catalytic sites. ORF3 designated, as TGB2 encoded plant virus movement super family (pfam01307) in different viruses belonging to *Alphaflexiviridae*. The two highly conserved amino acid motifs, one at N-terminus (MPLQPPPDHTWA) and another in core region (RDTSRHVGDPHSLPFGGxYRDGSKVx-HYNSPR), were identified throughout the genus *Mandarinivirus*. ORF4 encoded the smallest protein homologous to '7 kD viral coat protein' of potexviruses and carlaviruses. ORF5 encoded coat protein in mandariviruses and is rich in G + C content ranging from 54.16 to 55.11%. The amino acid sequences homologous to conserved domain of '*Flexi_CP superfamily*' (pfem00286) with an *E* value of

1.40e–60 were observed. The characteristic motifs VWN (aa 226–228), PPANW (aa 236–240) and AAFDF (aa 252–256) of *Alphaflexiviridae* were also identified on CP. The ORF6 partially overlaps ORF5, encoded Viral_NaBP protein and characterized by a zinc finger-like motif (aa 175–190). The details on molecular features of all the ORFs of CYVCV isolates characterized in this study are presented in Table 2.

Nucleotide diversity and phylogenetic relationship

The sequence identity matrix showed that 36 CYVCV isolates shared 95.2–99.8% and the two ICRSV isolates shared 71.5–71.9% nucleotides sequence identities with the reference sequence CYVCV Y1 (JX040635.1). Among the characterized four Indian isolates, CYVCV-ECAI (KT696510.1) shared least, 95.1–95.5% sequence identity with other global isolates of CYVCV and 41% similarity with PVX, a representative member of *Alphaflexiviridae*. The mean diversity of nucleotide sequences was estimated with best fitted Tamura–Nei model using MEGA 6.6 software and a mean diversity of 0.017 was measured within 35 CYVCV isolates population.

In phylogenetic trees using complete genome sequences, CYVCV were grouped together in two group viz, groups I and II on a sister branch of ICRSV, and formed a distinct well supported clad. Furthermore, group II bifurcated into cluster IIa and IIb and CYVCV isolates from India were phylogenetically closer to the CYVCV isolates from Turkey and Pakistan in cluster IIa. CYVCV isolates from China grouped in cluster IIb distinctly in the same clade (Fig. 2). The CYVCV isolates did not show any host specificity or symptomatic relationship in phylogenetic clustering. CYVCV and ICRSV, both are closely associated member of genus *Mandarinivirus* infecting the common host citrus, but are distinct and segregated into two separate clusters.

Estimation of selection pressure

The selection pressure on all the six coding regions was computed by normalized dN/dS ratio, using the Datamonkey software. The normalized dN/dS value described in Table 3 suggested that all the six coding regions of CYVCV were evolved under varying degree of purifying or negative selection pressure. Results on selection pressure indicated that RdRP region of CYVCV evolved under the highest degree of purifying selection pressure with 0.09 mean value, and for TGBs and CP, the mean values were 0.11–0.25, while the NaBP coding region showed high degree of variability with 0.35 mean value, evolving under the least degree of purifying selection pressure (Table 3).

Table 2 BLAST analysis of the predicted genes nucleotide sequences and molecular features of the four citrus yellow vein-clearing virus isolates sequenced from India

Genome region	Indian CYVCV isolates		CYVCV-ECAI			CYVCV-RMGI			CYVCV-PALI			CYVCV-KMPI		
	Start	End	G+C content	MW (kDa)	Seq. id/QC (%)	G+C content	MW (kDa)	Seq. id/QC (%)	G+C content	MW (kDa)	Seq. id/QC (%)	G+C content	MW (kDa)	Seq. id/QC (%)
5' UTR	1	80	42.50		97/100	41.25		98/100	41.25		98/100	41.25		98/100
ORF1 (RdRp)	81	5030	51.52	186.43	96/100	51.41	186.43	97/100	51.27	186.48	97/100	51.47	186.37	97/100
ORF2 (TGB1)	5037	5714	51.62	24.96	94/100	51.18	25.14	97/100	50.19	25.04	96/100	51.03	25.15	96/100
ORF3 (TGB2)	5692	6018	50.46	12.09	94/100	51.07	12.02	97/100	50.76	12.01	97/100	50.76	12.03	97/100
ORF4 (TGB3)	5945	6127	49.73	6.18	98/100	48.63	6.18	97/100	48.63	6.208	98/100	48.63	6.208	98/100
ORF5 (CP)	6150	7127	55.11	35.66	96/100	54.19	35.69	98/100	54.16	35.62	98/100	54.16	35.72	98/100
ORF6 (NaBP)	6827	7495	54.11	25.49	96/100	53.66	25.43	99/100	53.57	25.29	98/100	53.96	25.44	98/100
3' UTR	7496	7531	29.85		100/97	30.56		100/99	30.48		100/99	30.56		99/100
Whole genome	1	7531	51.52		96/99	51.47		97/100	51.29		98/100	51.48		98/100

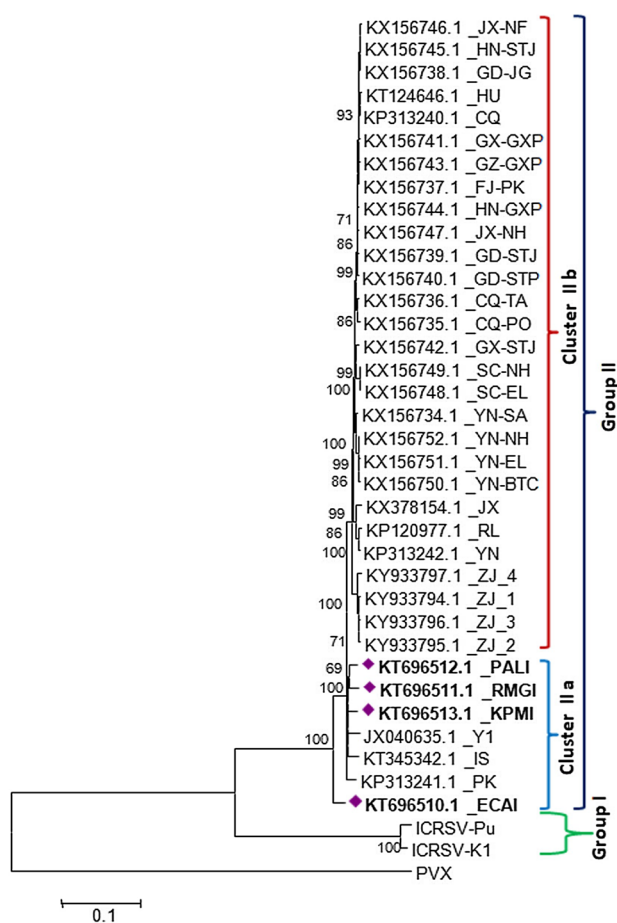


Fig. 2 Phylogenetic tree generated by the maximum likelihood (ML) method from the aligned nucleotide sequences of the global citrus yellow vein-clearing virus isolates, the members of the same genus *Mandarivirus* (Indian citrus ringspot virus), and a member of *Potexvirus* (PVX) as an out-group member of family *Alphaflexiviridae*. The tree was constructed with the lowest BIC score from the list using MEGA6 software. The bootstrap support values (1000 replicates) are shown at the nodes of the branches. The branches showing a bootstrap values less than 60% were collapsed. Sequences obtained from this study of the four CYVCV isolates from India demarcated bold and by pink color squares

Split network and recombination

The Neighbor-Net phylogenetic analysis performed for the sequenced data suggested the evolution and diversification between the CYVCV and ICRSV isolates. The recombination events and breakpoints were considered only after identification by at least three of the nine methods implemented in RDP4, as described in Table 4. The four putative recombination events were detected in analysis of complete genome sequences and KX156746.1 and KT696513.1 were identified as major parental sequences (Fig. 3). The CYVCV isolates KX156735.1, KX156742.1, and KT696512.1 (Indian isolate) were identified as recombinants. Moreover, we did not detect any inter

CYVCV–ICRSV recombinant event, indicating the stability of genome of both the virus species.

Neutrality test and population demography

The nucleotide sites' distribution on CYVCV genome was evaluated using the coding regions (ORFs) sequences of 35 CYVCV isolates. The mismatch nucleotide sites' distribution on sequences presented in Table 5 showed that the nucleotide variability was low and not uniform in coding regions. The Tajima's D , F_u and Li's D and F_u and Li's F values for all the coding regions were negative, suggesting that the CYVCV population is in the state of increasing the diversity. However, the values for all ORFs using different demographic algorithms viz., Tajima D , F_u and Li D , and F_u and Li F were recorded negative, while P value was < 1 , indicating that the population of the CYVCV was not significantly variable. However, the population size was small and P value was not significant, so the results were not able to provide conclusive evidences.

Discussion

The results of present study revealed that the CYVCV was not confined to Etrog citron and other lemon group. The virus is also present in other commercially important citrus cultivars such as kinnow mandarin and other sweet orange cultivars in India. CYVCV is generally associated with vein-clearing symptoms (Alshami et al. 2003), but in our study, CYVCV produced ringspots symptoms on kinnow mandarin and sweet orange in the absence of ICRSV. The failure of detection of CYVCV in different commercial cultivars of citrus with ring spot symptoms was due to lack of information that CYVCV can also cause ring spot symptoms. The virus has been reported from several South Asian countries and is an emerging problem for the citrus growers (Chen et al. 2014; Zhou et al. 2017).

Although yellow vein-clearing disease was reported in 1997 from India, molecular evidence and sequence information were unknown so far. In the present study, complete genome sequences of four CYVCV have been deciphered from the four citrus cultivars producing diverse symptoms. The sequenced genome has shown a typical genome organization of genus *Mandarivirus* and previously described for CYVCV Y1 isolate (Loconsole et al. 2012). However, the four CYVCV isolates from this study consisted of 7.531 nt, while the first genome sequence of CYVCV-Y1 consisted of 7529 nt excluding the 3'-poly (A) tail. The two extra nucleotides in 5' UTR of all the four CYVCV genome in this study in comparison with CYVCV Y1 isolate from turkey have also been reported in CYVCV sequences from China (Zhen et al. 2015).

Table 3 Estimates of selection pressure on six coding regions of the global citrus yellow vein-clearing virus isolates

Coding region	Number of codons	Normalized dN/dS ^a		Selection pressure					
		Log (L)	Mean	Positive		Negative		Neutral	
				Number	Percent	Number	Percent	Number	Percent
RdRp	1649	-12152.7	0.09267	1	0.06	458	27.77	1190	72.17
TGB1	225	-1717.98	0.11823	1	0.44	21	9.33	203	90.22
TGB2	108	-802.87	0.25498	0	0	9	8.33	99	91.67
TGB3	60	-442.22	0.23627	0	0	4	6.66	56	93.34
CP	325	-2416.36	0.21841	4	1.23	19	5.84	302	92.93
NaBP	222	-1461.49	0.35836	3	1.35	8	3.60	211	95.05

^aThe mean dN/dS values < 1 indicate negative or purifying selection, dN/dS values = 1 indicate neutral selection, dN/dS values > 1 indicate that positive selection for each coding region-specific data set was determined by SLAC (single-likelihood ancestor counting). The positive and negative nucleotide selection sites were determined by FUBAR (A Fast, Unconstrained Bayesian AppRoximation) using the nucleotide GTR model with 0.9 threshold value

Table 4 Intragenic recombination events detected using RDP4 and each event is represented by a number

Sr. no	Recombinant	Break Point		Major parent	Minor parent	P value detected by RDP4 programs ^a
		5'	3'			
1	KX156735.1	5765	6373	KX156746.1	JX040635.1	R (2.238 × 10 ⁻⁰⁵); G (2.351 × 10 ⁻⁰³); B (8.074 × 10 ⁻⁰⁶); Se (3.927 × 10 ⁻⁰³)
2	KX156742.1	3356	6375	KX156746.1	KX156749.1	R (1.880 × 10 ⁻⁰⁴); G (1.462 × 10 ⁻⁰⁴); B (2.698 × 10 ⁻⁰⁷); M (1.214 × 10 ⁻⁰⁹); Si (2.137 × 10 ⁻⁰⁹); Se (1.338 × 10 ⁻⁰⁹)
3	KT696512.1	1936	2848	KT696513.1	KX156746.1	G (2.291 × 10 ⁻⁰³); B (2.796 × 10 ⁻⁰⁵); M (1.113 × 10 ⁻⁰²); C (7.640 × 10 ⁻⁰³); Si (2.980 × 10 ⁻⁰⁵)
4	KT696512.1	163	1084	Unknown	KT696511.1	R (9.474 × 10 ⁻⁰³); B (3.346 × 10 ⁻⁰³); C (4.154 × 10 ⁻⁰²)

Major and minor parental sequences indicate the donor sequences that contributed larger and smaller fractions of recombinant in citrus yellow vein-clearing virus isolates genomes

^aRecombination detecting program: R (RDP), G (GenConv), B (BootScan), M (MaxChi), C (Chimaera), Si (SiScan), Se (3Seq), L (LARD), P (PhylPro) in RDP4 software. The corresponding events supported by at least three different RDP4 methods were reported in the present study

The RdRp region (ORF1) of RNA viruses possesses all functional domains of replicase proteins, which are highly prone to nucleotides variability. The lack of proof-reading activity of RNA polymerases in RdRp region of RNA viruses provides the great potential for evolution, genetic variability, and adaptation into new environmental conditions due to high mutation rate and generation of variable population (Holland et al. 1982; García-Arenal and Fraile 2011). However, RdRp was relatively stable in CYVCV. The ORF2, 3, and 4 were found partially overlapped and encoded triple gene block proteins family. ORF3 designated, as TGB2 is the most conserved genomic region in CYVCV. The TGBs may have great potential of silencing suppressor, as recent studies revealed that the virus movement depends on the multiple functions of the proteins including the silencing suppression (Bayne et al. 2005; Solovyev et al. 2012). In another study (Duan et al. 2012), it has been demonstrated that the interaction of the silencing suppressor, 2b, encoded in cucumoviruses with the AGO proteins in vivo was required

for nucleolar targeting and redistribution of both 2b and AGO proteins in the nucleus. TGB1 protein and helicase proteins explored in endornavirus (Morozov and Solovyev 2003; Koonin and Dolja 2012) have the ability to suppress RNA silencing (Bayne et al. 2005; Senshu et al. 2011). As enough genomic information is available, the functional roles of these ORFs can be explored more easily in mandarivirus.

The coat protein (CP) of the RNA viruses is one of the most important genomic region considered for species demarcation, genetic diversity, and immunodiagnosics development (Boonham et al. 2002; Adams et al. 2014). The CP analysis in this study revealed that CYVCV shared the structural core and evolutionary origin with the ICRSV and other filamentous plant viruses in allexi, carla, and foveaviruses. The C-terminus of genomic region, nucleic acid-binding protein, characterized by a zinc finger-like motif found (CX2CX9HX4CXH) in CYVCV was identical to ICRSV. This region was also present in carla and allexiviruses, but shorter than the mandariviruses and the sequencing result

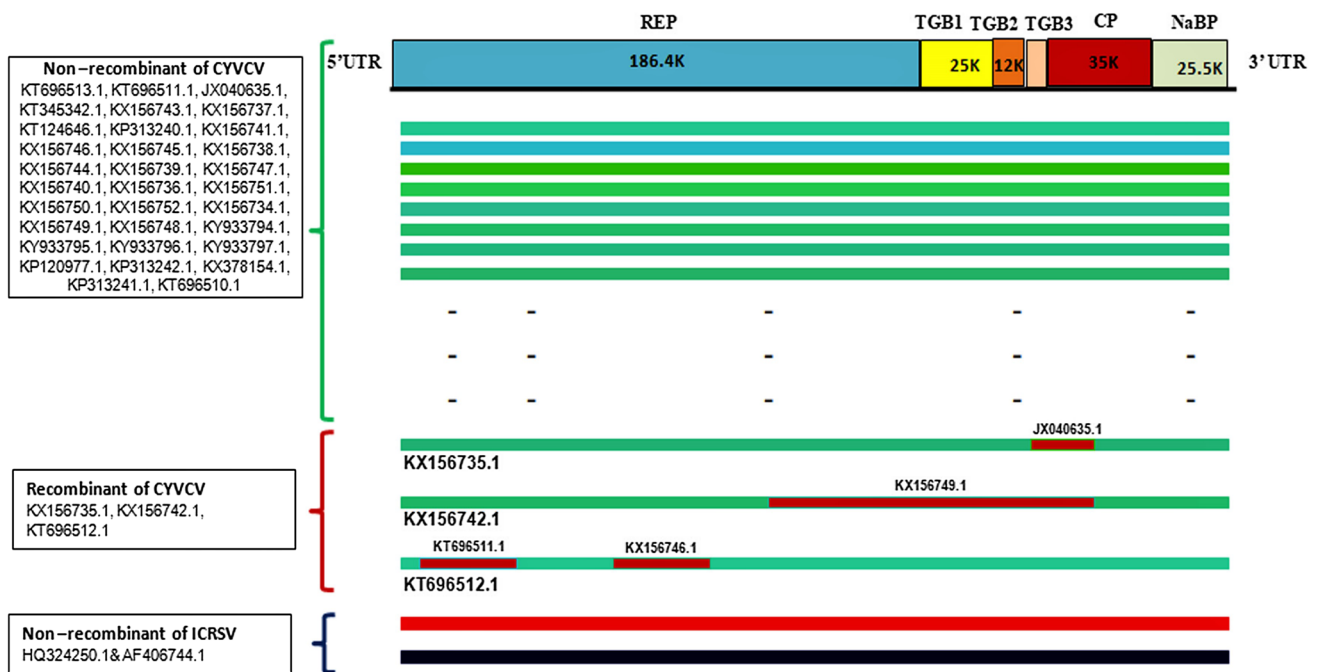


Fig. 3 Recombination maps of the citrus yellow vein-clearing virus genomes of the global isolates, including four from this study. Two isolates of Indian citrus ringspot virus, member of the same genus, *Mandarinivirus*, were also included to elucidate inter (CYVCV and ICRSV)-specific recombination patterns, but no events were reported.

The red colored boxes over the genome maps indicated the recombinant fragment and the other details, as described in Table 5. The accession numbers of non-recombinant and recombinant genomes of CYVCV and ICRSV mentioned in the square boxes at left side of the recombinant maps

Table 5 Neutrality tests, nucleotide, and haplotype diversity based on the genomic-coding region of citrus yellow vein-clearing virus using DnaSP v5 program

Genomic regions ^a	Nucleotide sites	Polymorphic sites (S)	G + Cn	Tajima's D	Fu and Li's D	Fu and Li's F	Fu F's	Haplotype diversity	Nucleotide diversity (Pi)
ORF1	4947	638	51.50	- 1.88782	- 2.46033	- 2.68751	- 6.372	1.000	0.01669
ORF2	678	110	51.20	- 1.98100	- 3.06741	- 3.19542	- 16.013	0.992	0.01938
ORF3	327	51	51.00	- 2.04126	- 2.71103	- 2.94325	- 16.034	0.963	0.01711
ORF4	183	30	48.90	- 1.63298	- 2.50672	- 2.61871	- 11.804	0.889	0.02109
ORF5	978	123	54.58	- 1.8701	- 3.09110	- 3.16882	- 12.783	0.992	0.01589
ORF6	669	63	53.6	- 2.08764	- 3.36914	- 3.47465	- 17.087	0.981	0.01024

The data sets represent coding regions of the global 35 CYVCV isolates genomic sequences. The Tajima D, Fu and Li D, Fu and Li F, Fu F's, etc. were performed using the DnaSP v5 software. Details of sequences used for analysis are: four isolates of CYVCV from India, 30 isolates from China, and one isolate from Turkey

from this study were also in agreement with the previous work (Rustici et al. 2002; Loconsole et al. 2012).

The study on genetic and molecular diversity of the viral pathogens provides a better understanding of the ecology, evolution, and biology of viruses (Wang et al. 2011); in this connection, an attempt was made to study the recombination, population dynamics using different statistical algorithms. The data obtained in the present study revealed the existence of low level of genetic diversity and support the evolutionary relationship in the virus population. The

phylogenetic analysis considering full genome sequences of CYVCV isolates from different countries showed that the virus isolates from India, Pakistan, and Turkey were more closer than China isolates, but, due to low number of differences at the nucleotides and amino acid residues, failed to form distinct clusters. The analysis showed that the mandarinviruses formed a separate cluster in the family *Alphaflexiviridae* (Martelli et al. 2007). CYVCV isolates formed a distinct well supported clad on the sister branch with the ICRSV species.

Recombination is an important factor of genetic variation and the frequent recombination leads to the process in shaping the fitness of population in plant viruses (Sztuba-Solinska et al. 2011). In the present study, only four intra-genic recombination events were observed that may increase the ability of CYVCV to sustain on diverse host and climatic regimes. Both, CYVCV and ICRSV, are closely associated member of the genus *Mandarivirus* and can be present together on a single host, but no intergenic recombination event was observed, indicating the stability of the genome of these species. Negative *D* values in estimating the selection pressure for all the six coding regions of CYVCV indicated the existence of low-frequency polymorphism. The genetic heterogeneity and population demography analysis indicated that CYVCV population is under the increasing state towards the heterogeneity. The mean *dN/dS* value in RdRp coding region was least (0.09267) with the highest negative selection pressure (27.77%), indicating that the functional constraints were much higher for RdRp coding region and it has evolved under relatively stronger selection constraints. Tajima's *D* and Fu and Li's *D* parameters principally indicate molecular evolution under different demographic model and the negative values indicate the existence of very low genetic variability. The codons under the positive selection or mutation affecting the functional activities of the virus genes have been identified by FUBER in potato virus Y and sugarcane mosaic virus (Li et al. 2013). The codons detected under the positive selection pressure in CYVCV and their functional activities need to be investigated further.

Conclusions

It is evident from the present study that CYVCV has a wide host range in citrus species. The virus is rapidly spreading and produces variable symptoms often identical with the ICRSV. The complete genome sequence information obtained in the present study will help to develop molecular detection and diagnosis and to understand the complete characteristics of the virus. The results obtained using all the statistical algorithms, indicated the existence of limited genetic variability in CYVCV genomes, and this will help in developing effective management strategies.

Acknowledgements The authors gratefully acknowledge the ICAR-OUTREACH project for funding. Authors are thankful to Head, Division of Plant Pathology and Director, Indian Agricultural Research Institute for providing necessary lab facilities. The first author is grateful to ICAR, New Delhi and ICAR-DMAPR, Anand, Gujarat for allowing to pursue PhD program. Special thanks are addressed to the anonymous reviewers for their constructive comments and suggestions to improve the manuscript.

Funding The author(s) received the financial support under the ICAR-OUTREACH program for the research work.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest about this manuscript and research.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Adams MJ, Lefkowitz EJ, King AMQ, Carstens EB (2014) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2014). *Arch Virol* 159:2831–2841. <https://doi.org/10.1007/s0070>
- Ahlawat YS, Pant RP (2003) Major viruses and virus-like diseases India, their diagnosis and management. *Annu Rev Plant Pathol* 2:447–474
- Alshami AAA, Ahlawat YS, Pant RP (2003) A hitherto unreported yellow vein clearing disease of citrus in India and its viral etiology. *Indian Phytopathol* 56(4):422–427
- Bani SMH, Aghajanzadeh S (2017) Occurrence of citrus yellow vein clearing virus in citrus species in Iran. *J Plant Pathol* 99(1):290
- Bayne EH, Rakitina DV, Morozov SY, Baulcombe DC (2005) Cell-to-cell movement of *Potato potexvirus X* is dependent on suppression of RNA silencing. *Plant J* 44:471–482. <https://doi.org/10.1111/j.1365-313X.2005.02539.x>
- Boonham N, Walsh K, Peterson S, North J, Smith P, Barker I (2002) The detection of tuber necrotic isolates of Potato virus Y and the accurate discrimination of PVY^O, PVY^N and PVY^C strains using RT-PCR. *J Virol Methods* 102:103–112
- Byadgi AS, Ahlawat YS (1995) A new viral ringspot disease in citrus (Citrus species in India). *IJAS* 65(10):61–68
- Catara A, Azzaro A, Davino M, Polizzi G (1993) Yellow vein clearing of lemon in Pakistan. In: Proceedings of 12th conference IOCV, Riverside, CA, pp 364–367
- Chen H, Li Z, Wang X, Zhou Y, Tang K, Zhou C, Zhao X, Yue J (2014) First report of *Citrus yellow vein clearing virus* on Lemon in Yunnan, China. *China Plant Dis* 98:1747. <https://doi.org/10.1094/PDIS-04-14-0343-PDN>
- Duan CG, Fang YY, Zhou BJ, Zhao JH, Hou WN, Zhu H (2012) Suppression of Arabidopsis ARGONAUTE1-mediated slicing, transgene-induced RNA silencing, and DNA methylation by distinct domains of the cucumber mosaic virus 2b protein. *Plant Cell* 24:259–274. <https://doi.org/10.1105/tpc.111.092718>
- García-Arenal F, Fraile A (2011) Population dynamics and genetics of plant infection by viruses. In: Caranta C, Aranda MA, Tepfer M, Lopez-Moya JJ (eds) Recent advances in plant virology. Caister Academic Press, Norfolk, pp 263–281
- Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, Vande Pol S (1982) Rapid evolution of RNA genomes. *Science* 215:1577–1588
- Hudson RR (2000) A new statistic for detecting genetic differentiation. *Genetics* 155:2011–2014
- Koonin EV, Dolja VV (2012) Expanding networks of RNA virus evolution. *BMC Biol*. <https://doi.org/10.1186/1741-7007-10-54>
- Li Y, Liu R, Zhou T, Fan Z (2013) Genetic diversity and population structure of *Sugarcane mosaic virus*. *Virus Res* 17:242–246. <https://doi.org/10.1016/j.virusres.2012.10.024>
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>

- Loconsole G, Onelge N, Potere O, Giampetruzzi A, Bozan O, Satar S, De Stradis A, Savino VN, Yokomi RK, Saponari M (2012) Identification and characterization of *Citrus yellow vein clearing virus*, a putative new member of the genus Mandarivirus. *Phytopathology* 102:1168–1175. <https://doi.org/10.1094/PHYTO-06-12-0140-R>
- Martelli GP, Adams MJ, Kreuze JF, Dolja VV (2007) Family Flexiviridae: a case study in virion and genome plasticity. *Annu Rev Phytopathol* 45:73–100. <https://doi.org/10.1146/annurev.phyto.45.062806.094401>
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* <https://doi.org/10.1093/ve/vev003>
- Meena RP, Baranwal VK (2016) Development of multiplex polymerase chain reaction assay for simultaneous detection of clostero-, badna- and mandari-viruses along with huanglongbing bacterium in citrus trees. *J Virol Methods* 235:58–64. <https://doi.org/10.1016/j.jviromet.2016.05.012>
- Milne RG, Luisoni E (1977) Rapid immuno electron microscopy of virus preparations. In: Maramorosch K, Koprowski H (eds) *Methods in virology*, vol 6. Academic Press, New York, pp 265–281
- Morozov SY, Solovyev AG (2003) Triple gene block modular design of a multifunctional machine for plant virus movement. *J Gen Virol* 84:1351–1366. <https://doi.org/10.1099/vir.0.18922-0>
- Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196–1205. <https://doi.org/10.1093/molbev/mst030>
- Prabha K, Baranwal VK (2012) The genome sequence of an isolate of *Indian citrus ringspot virus* infecting the sweet orange in India. *J Virol* 86:12446–12447. <https://doi.org/10.1128/JVI.02084-12>
- Rustici G, Milne RG, Accotto GP (2002) Nucleotide sequences, genome organization and phylogenetic analysis of *Indian citrus ringspot virus*. *Arch Virol* 147:2215–2224. <https://doi.org/10.1007/s00705-002-0875-6>
- Senshu H, Yamaji Y, Minato N, Shiraishi T, Maejima K, Hashimoto M, Miura C, Neriya Y, Namba S (2011) A dual strategy for the suppression of host antiviral silencing: two distinct suppressors for viral replication and viral movement encoded by potato virus M. *J Virol* 85(19):10269–10278. <https://doi.org/10.1128/JVI.05273-11>
- Solovyev AG, Kalinina NO, Morozov SY (2012) Recent advances in research of plant virus movement mediated by triple gene block. *Front Plant Sci* 3:276. <https://doi.org/10.3389/fpls.2012.00276>
- Szuba-Solińska J, Urbanowicz A, Figlerowicz M, Bujarski JJ (2011) RNA–RNA recombination in plant virus replication and evolution. *Annu Rev Phytopathol* 49:415–443. <https://doi.org/10.1146/annurev-phyto-072910-095351>
- Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22(22):4673–4680
- Wang J, Sharma AM, Duffy S, Almeida RPP (2011) Genetic diversity in the 3' terminal 4.7 kb region of *Grapevine leafroll-associated virus*. *Phytopathology* 101:445–450. <https://doi.org/10.1094/PHYTO-07-10-0173>
- Zhen S, Kurth EG, Peremyslov VV, Changyong Z, Dolja VV (2015) Molecular characterization of a *Citrus yellow vein clearing virus* strain from China. *Arch Virol* 160(7):1811–1813. <https://doi.org/10.1007/s00705-015-2423-1>
- Zhou Y, Chen HM, Cao MJ, Wang XF, Jin X, Liu KH, Zhou CY (2017) Occurrence, distribution and molecular characterization of *Citrus yellow vein clearing virus* in China. *Plant Dis* 101:137–143. <https://doi.org/10.1094/PDIS-05-16-0679-RE>