Asian Journal of Agriculture and Biology
https://doi.org/10.35495/ajab.2024.112

Original Article

Next-generation sequencing reveals genetic variation in ToCV infecting Pakistani tomato plants

Ahmed Raza[1,2,3*], Muhammad Taimoor Shakeel[4], Sidra Iqbal[5], Azhar Abbas Khan[6], Tahira Abbas[7], Ammarah Hasnain[8], Mahmoud Ahmed Amer[9], Abdul Majid[2], Adnan Ahmad[10], Ummad Ud Din Umar[11], Muhammad Umer Iqbal[12], Muhammad Naveed Aslam[4], Xifeng Wang[3], Mingzheng Duan[1*]

[1]Key Laboratory of Edible Fungi Resources Innovation Utilization and Cultivation, College of Agronomy and Life Sciences, Zhaotong,657000, China

[2]Plant Virology Lab, Crop Diseases Research Institute, National Agriculture Research Centre, Pakistan Agricultural Research Council, Islamabad, Pakistan

[3]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

[4]Department of Plant Pathology, Faculty of Agriculture & Environment, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

[5]Department of Plant Breeding and Genetics, University of Agriculture-Faisalabad, Constituent College Depalpur, Okara, 56300, Pakistan

[6]Department of Entomology, Faculty of Agricultural Sciences and Technology, University of Layyah, Pakistan

[7]Department of Horticulture, Faculty of Agricultural Sciences and Technology, University of Layyah, Pakistan

[8]Department of Biotechnology, Faculty of Biological Sciences, Lahore University of Biological and Applied Sciences, Pakistan

[9]Plant Protection Department, College of Food & Agriculture Sciences, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia

[10]GenEd, Pakistan Council of Scientific and Industrial Research, Lahore, 54600, Pakistan

[11]Department of Plant Pathology, Faculty of Agriculture Science and Technology, Bahauddin Zakariya University, Multan 60800, Pakistan

[12]Department of Plant Pathology, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan

**Corresponding author's email:* ahmedraza7128@parc.gov.pk; duanmingzheng@ztu.edu.cn

Received: 13June 2024 / Accepted: 20 April 2025 / Published Online: 28 May 2025

Abstract

Tomato chlorosis virus (ToCV), causing yellowing disease in tomato, is a phloem-limited, whitefly-transmitted crinivirus that mainly affects plants by reducing chlorophyll content. During a survey conducted in 2019, cirinivirus-specific yellowing symptoms were observed in the tomato fields in Multan, Khanewal, and Lodhran districts of Pakistan. To confirm the etiology of the yellowing disease of tomato, a total of 76 symptomatic and 22 asymptomatic samples were collected out of which 72% of the tested plants were found to be positive for ToCV using RT-PCR. The full genomic nucleotide sequences of two selected ToCV-isolates were obtained by next-generation sequencing (NGS). RNA1 and RNA2 of each isolate comprised 8594 and 8242 nucleotides, respectively. The NGS results were verified by sequencing the amplified overlapping fragments of RNA1 and RNA2 using specific primers designed in this study. In BLASTn analysis for ToCV-Pak1 and ToCV-Pak2, RNA1 from both isolates had the highest similarities (99.41-99.46%) with a South Korean isolate; RNA2 had the highest similarities of 99.67-99.64% with a Greek isolate for Pak1 and Turkish isolate for Pak2, respectively. Phylogenetic analysis indicated that RNA1 of the Pakistani isolates clustered together, forming a subclade, and RNA2 from one of the Pakistani isolates (MN869006) clustered with a Brazilian isolate. At the same time, the other one (MN869007) fell close to a Turkish isolate. Statistical analysis indicated a low polymorphic frequency in RNA1 but a higher polymorphic frequency in RNA 2 quantifying the degree of variation in the analyzed isolates which may

result in emergence of new strains with altered levels of virulence. One probable recombination event was detected in RNA1. Analysis revealed that RNA1 of Pak1 and Pak2 is recombinant between South Korean isolate (major parent) and Spanish isolates (minor parent) with recombinant breakpoints at 6172 and 6668 nucleotide positions from the 5′ end, respectively. This study provides the first full genomic analysis and the genetic diversity of ToCV-RNA1 infecting the tomato plant in Pakistan. Understanding the trends and rate of variation in the isolates may provide an insight into the development of resistance in tomato plants against tomato yellowing disease.

Keywords: ToCV, Phylogenetic analysis, Recombination, Genetic characterization, Tomato

How to cite this article:
Raza A, Shakeel MT, Iqbal S, Khan AA, Abbas T, Hasnain A, Amer MA, Abdul Majid, Ahmad A, Umar UD, Iqbal MU, Aslam MN, Wang X and Duan M. Next-generation sequencing reveals genetic variation in ToCV infecting Pakistani tomato plants. Asian J. Agric. Biol. 2025: 2024112. DOI: https://doi.org/10.35495/ajab.2024.112

Introduction

The Tomato yellowing disease (TYD) is one of the major threats to tomato production worldwide (Hanssen et al., 2010). TYD is attributed to two viruses, *Tomato chlorosis virus,* and *Tomato infectious chlorosis virus,* which are classified in the genus Criniviruses in the family *Closteroviridae*. Both viruses are phloem-limited, single-stranded, positive-sense, bipartite Criniviruses (Wisler et al., 1998; Wintermantel et al., 2005; Zhao et al., 2014). TYD exhibits symptoms of interveinal chlorosis, yellowing, and brittleness of the mature leaves, accompanied by bronzing and necrotic flecks at a later stage. Fruit does not show any symptoms, but the yield is significantly reduced because of the damage to the photosynthetic area Particularly, its prevalence in tomato crops has been recorded to be almost 100% (Fortes et al., 2012; Arruabarrena et al., 2014) that, lowers the yield of the crop due to depletion of chlorophyll in the leaves; however, it does not produce any symptoms on its fruits (Wisler et al., 1998; Wintermantel et al., 2005). It is difficult to differentiate between the symptoms expressed by tomato plants when infected with ToCV or TICV. However, they can be distinguished by differential symptoms expressed by each of these viruses on *Nicotiana clevelandii* as it produces necrotic flecks in addition to interveinal chlorosis when infected by TICV.

ToCV is a flexuous, rod-shaped particle, 650-850nm long with a bipartite genome: RNA1 (8595 nt) has four open reading frames (ORFs) and encodes the proteins mainly involved in replication: RNA2 (8242 nt) comprises nine ORFs that encode proteins to encapsidates the virus, and aid movement and vector transmission (Wintermantel et al., 2005; Zhao et al., 2025). Since ToCV was first reported in Florida in 1989 (Wisler et al., 1998), it is now distributed worldwide and infect multiple host plant species belonging to several plant families including, *Asteraceae, Amaranthaceae, Apocenaceae, Aizoaceae, Chenopodiaceae, Araliaceae, Malvaceae, Portulacaceae* (Shakeel et al., 2017) *Solanaceae* (Wisler et al., 1998; Fortes and Navas-Castillo, 2008; García-Cano et al., 2010; Fonseca et al., 2013; Fiallo-Olivé et al., 2014; Arruabarrena et al., 2015; Zhou et al., 2015; Shakeel et al., 2017), *Zygophyllaceae* (Lozano et al., 2004; Shakeel et al., 2017), *Plantaginaceae* (Solórzano-Morales et al., 2011; Orfanidou et al., 2014), *Acanthaceae* (Mamoun Abdel-Salam et al., 2019), *Boraginaceae* and *Mazaceae* (Kil et al., 2015), *Compositae* (Tsai et al., 2004), *Convolvulaceae* (Orfanidou et al., 2014; Kil et al., 2015), *Euphorbiaceae* and *Moraceae* (Mamoun Abdel-Salam et al., 2019), *Fumariaceae, Portulacaceae, Primulaceae, Rubiaceae, Primulaceae,* and *Oxalidaceae* (Orfanidou et al., 2014), *Phytolacaceae* (Solórzano-Morales et al., 2011; Kil et al., 2015), *Fabaceae* (Kil et al., 2015; Wang et al., 2018), *Cucurbitaceae,* and *Rutaceae* (Solórzano-Morales et al., 2011). ToCV holds a unique position among the plant viruses due to its ability to be transmitted by three whitefly species including *B. tabaci, Trialeurodes vaporariorum* and *T. abutiloneus (*Wisler et al., 1998; Karasev, 2000; Navas-Castillo et al., 2000; Wintermantel and Wisler, 2006; Navas-Castillo et al., 2011; Orfanidou et al., 2016**).** Apart from its transmissibility, its broader host range also facilitates its distribution and establishment that underscores its significance as serious plant pathogen.

Different studies focused on heat shock protein 70 homolog (HSP70h) and coat protein (CP) coding regions of ToCV showed both genes are highly conserved with as high as 98 to 100% amino acid sequence identity (Wintermantel and Wisler, 2006). In year 2022-23, tomatoes are grown over almost 68862 hectares in greenhouses and open fields in Pakistan producing 762736 tons (Ministry of National Food Security and Research, 2023). In Pakistan, studies on TYD related to criniviruses are limited, whereas whitefly-transmitted begomoviruses are well-studied (Tayyib et al., 2013). Recently, we reported the first case of ToCV infecting tomatoes in Pakistan (Raza et al., 2020). The objective of the present study was to fully characterize the complete genome of ToCV and investigate the genetic diversity of tomato isolates from Pakistan.

Material and Methods

Sample collection

In a survey of 16 open fields in Multan, Khanewal, and Lodhran districts from February to March 2019, 76 samples were collected from tomato plants showing typical TYD symptoms such as interveinal chlorosis, leaf brittleness, yellowing (Fig. 1A-1C) and 22 from asymptomatic plants. By district, thirty symptomatic and nine asymptomatic samples were collected from Multan, 25 symptomatic and seven asymptomatic from Khanewal, and 21 symptomatic and six

https://doi.org/10.35495/ajab.2024.112

asymptomatic from Lodhran (Fig 1D). All samples were preserved in an icebox and processed and analyzed at the Molecular Plant Virology Lab, Department of Plant Pathology,, Bahauddiin Zakariya University, Multan.

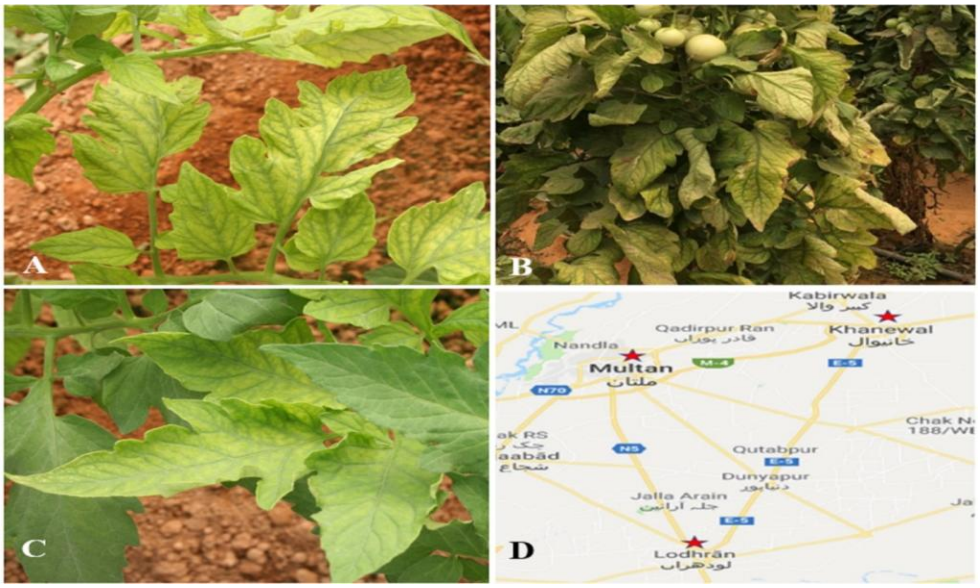

Figure-1. Symptoms on tomato plants in the field and collection areas in Pakistan. (A) Typical interveinal chlorosis on lower leaves of the plant. (B) Severe interveinal chlorosis, bronzing, and necrotic flecks on older leaves. (C) Early chlorosis on leaves of the infected plant (right) and leaves on the healthy plant (left). (D) District locations where samples were collected.

RNA extraction

For the detection of ToCV in the collected samples, RNA isolation was done from freshly collected leaf samples using Trizol reagent (Invitrogen, USA) following the manufacturer's protocol. The isolated RNA was used to synthesize cDNA with random primers in the RevertAid Premium Reverse Transcriptase kit (Thermo Fisher Scientific, USA).

Nested RT-PCR

The synthesized cDNA was subjected to Nested PCR using HS11/12 primers in the first step, and its product was used as a template in the second step using the TOC-5/TOC6 primers targeting HSP70 gene of ToCV (Table 1). Clathrin adaptor complexes medium subunit (CAC) is a reference gene of tomato plant that was targeted as an internal control for PCR and reverse complementation using the primer 5′-CCTCCGTTGTGATGTAACTGG-3′; 5′-ATTGGTGGAAAGTAACATCATCG-3′ that amplifies a product of 173bp. The amplified products were visualized on 1.5% agarose gel. The samples found infected with ToCV were selected for Next Generation Sequencing (NGS) (Dovas et al., 2002).

Table-1. Sequence of specific primers to amplify the ToCV isolates

Sr. No	Name of Primer	Sequence	Amplification Size	References
1	**HS11/HS12**	5 -GG(G/T)TT(A/G)GA(G/T)TT(C/T)GGTACTAC-3 5 -CC(G/T)CCACCAAA(A/G)TCGTA-3	**587-bp**	Dovas et al., 2002
2	**TOC5/TOC6**	5 -GGTTTGGATTTTGGTACTACATTCAGT-3 5 -AAACTGCCTGCATGAAAAGTCTC-3	**463-bp**	Dovas et al., 2002

Next generation sequencing

Two of the ToCV isolates, ToCV-Pak1 and ToCV-Pak2, were selected for NGS to obtain complete genome sequences. Trizol Reagent (Invitrogen, USA) was used to isolate a total RNA from the infected leaves following the manufacturer's protocol. The libraries of Small RNA-seq (sRNA) s were constructed using NEBNext Small RNA Library Prep Set for Illumina and NEBNext Multiplex Oligos for Illumina (New England Biolabs, USA). The constructed library was quantified by KAPA Library Quantification Kit (Roche, Switzerland), and an Agilent Bioanalyzer 2100 (Agilent Technologies, USA) was used for quality control. NovaSeq6000 system (Illumina, USA) was used for deep sequencing of the full genome of ToCV. High-quality raw reads of ToCV were stored in a FASTQ format and sequencing was done by Biomarker Technologies, Inc.

Bioinformatics analysis

To eliminate any potential contamination and adopter sequences, the software Cutadapt v1.9.1 was employed to remove fragments shorter than 17 and longer than 30 nucleotides. The raw sequences underwent quality testing using FastQC to assess their quality and to determine the presence of any potential issues. The program was used to carefully examine the raw sequences, and only those of high quality and error-free were selected. These selected reads were matched with sequences from the plant rRNA database (available at http://www.plantrdnadatabase.com/) using Bowtie version 1.2.2 (accessible at http://bowtie-bio.sourceforge.net/index.shtml).

During the alignment process, a maximum of three mismatches were allowed. Complete genome assembly was performed using VirusDetect version 1.7 (Zheng et al., 2017), utilizing unaligned reads. The reference for this study was the non-redundant (95%) plant virus database, which was obtained from GenBank. In this study, the Burrows-Wheeler Aligner (BWA) version 0.7.12-r1039 was employed to align small RNA (sRNA) reads with the reference sequence. The sequences used in this study include ToCV sequences, host sequences from *Solanum lycopersicum*, and constructed viral contigs (Li and Durbin, 2009). SAMtools version 1.9 (http://samtools.sourceforge.net/) was employed for reference-guided assembly. This software was utilized to process BWA alignments and produce alignment information at each place. This information was subsequently employed to facilitate the production of viral contigs. The assembly of viral small RNAs was conducted using Velvet (v1.2.09) software, following the methodology described in a previous publication (Singh et al., 2020). The software was accessed through the European Bioinformatics Institute website (https://www.ebi.ac.uk/).

Validation by Sanger sequencing

Newly assembled ToCV isolates identified from Pakistan were subjected to the Primer-Blast tool available at NCBI and was assessed to design the specific primers for the detection of local ToCV isolates [Fig. 2A]. For the validation of sequences obtained from NGS, complementary DNA was constructed from the total RNA extracted from the samples using Random Primer using Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) from Promega, USA. PCR reactions to detect TOCV in the collected samples were done using the conditions mentioned in the RT-PCR heading. Primer sequences, their product length and annealing temperature are given in Table S1. PCR amplicons were visualized using 1% agarose visualization. The obtained bands were executed and purified using the TIANquick Oligo Purification Kit (Tiangen Biotech, China). Later, they were ligated into pEASY-T5 Zero Cloning Vector (TransGen, China) following the manufacturer's instructions. The Ligated constructs were transformed into 50 µL Trans1-T1 Phage Resistant Chemically Competent Cells from Transgene Biotech (Beijing, China). Colony PCR using M13 forward and reverse primers was performed for the grown colonies, and the PCR-positive samples were sent to Sangon Biotech (China) for sequencing. The whole genome was validated by Sanger sequencing (Fig. 2A), and the validated RNA1 and RNA2 of both isolates were submitted to NCBI and were assigned to accession numbers MN869004 and MN869005 for RNA1 and MN869006 and MN869007 for RNA2 of ToCV-Pak1 and Pak-2, respectively.

Phylogenetic analysis

The validated complete sequence of RNA1 (8594 nt) and RNA2 (8242 nt) of the two ToCV isolates from Pakistan were subjected to BLAST and phylogenetic analysis. Furthermore, a BLAST search of the GenBank database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) provided a complete genome sequence of ToCV isolates from different countries including USA, China, Greece,

Turkey, Spain, Brazil, and South Korea. The complete genome sequence of TICV was used as an outgroup for the phylogenetic tree for RNA1 and RNA2. All the isolates (23 isolates of RNA1 and 26 isolates for RNA2), including the obtained Pakistani sequence and the sequences retrieved from GenBank, were aligned using Clustal W, and the phylogenetic tree was constructed using the maximum likelihood method and MEGA-X software (Kumar et al., 2016; Tamura and Nei, 1993). Branch support was maintained by setting the bootstrap value to 1000. Accession numbers and country of origin of the isolates from NCBI are mentioned in Table S2.

Selection pressure and recombination analysis

The identification of insertion and deletion events, also known as InDels, was performed using DnaSP software (version 5.10) (Balasubramanian and Selvarajan, 2014). Recombination events in aligned ToCV sequences were identified using the RDP software (Balasubramanian and Selvarajan, 2014). The following methods were employed: RDP, Bootscan, Lard, GENECONV, MaxChi, Chimaera, Siscan, and 3SEQ. The default settings were used, along with a Bonferroni corrected P-value cut-off of $\alpha = 0.05$. Recombination breakpoints were deemed statistically significant when corroborated by four or more methodologies. The assessment of molecular diversity at segregating sites was conducted using statistical measures, including Tajima's D, Fu and Li's D*, Fu and Li's F*, as well as haplotype diversity and nucleotide diversity at all sites. This analysis was performed using DnaSP version 5.0 (Librado and Rozas, 2009).

Results

RT-PCR

ToCV was detected in 28 of 30 symptomatic samples from Multan, 24 of 25 symptomatic samples from Khanewal, and 19 of 21 symptomatic samples from Lodhran (Fig. S1). ToCV was not detected in any of the asymptomatic samples (Fig. S1). Details on the collected samples and ToCV presence are given in Table 2.

Table-2: Location, symptoms, and ToCV presence in symptomatic and asymptomatic samples of tomato collected from fields in three districts in Pakistan.

District	Location	Number of Samples		ToCV Infected
		Symptomatic	Asymptomatic	
Multan	Nawab Pur	5 Y, IC	2	4
	Nag Shah Chowk	7 IC	1	7
	Multan – Bahawalpur Road	4 IC	2	4
	Multan – Vehari Road	4 IC	1	4
	Band Bosan	4 IC	1	4
	Suraj Miani	6 Y, IC	2	5
	Total	**30**	**9**	**28**
Khanewal	Jamesabad	5 Y	1	5
	Jahanian Road	4 Y	-	3
	Makhdoom Pur	6 Y	-	6
	Mahni Sial	5 Y, Br	4	5
	Mian Channu	5 Y	2	5
	Total	**25**	**7**	**24**
Lodhran	Shahpur	4 Y	1	3
	Seekar	4 Y	2	4
	Khudai	5 Y	1	4
	Burhanpur	5 IC, Y	-	5
	Muhammad Sai	3 Y	2	3
	Total	**21**	**6**	**19**

Y: yellowing, IC: interveinal chlorosis, Br: brittleness, -: no collection.

Composition of sRNA populations

Seventeen million one-end 50 bp reads with high confidence were obtained from the total sRNA profile of two tomato samples using high-throughput sequencing. The reads for ToCV-Pak1 and ToCV-Pak2 were 10.4 million and 6.6 million, respectively. Approximately 10.5 million (~62%) accurate reads were used further for analysis (Table S3). Accurately read mapping with the reference genome of *Solanum lycopersicum* resulted in 85.35 % (8.9 million reads) host-specific reads and 14.65 % (1.5 million reads) virus-specific reads (Fig. S2A). The base pair length of clean reads reached from 17-30 nt with 21 nt reads as most abundant (~5 × 106), followed by 24-nt (~4 × 106) and 22-nt (~3 × 106) reads (Fig. S2B).

ToCV contigs in the samples

The objective is to do a contig analysis of the samples. The Velvet program was utilized to perform de novo assembly of reads that did not align with the tomato genome, resulting in the generation of 3501 contigs (Fig. 2B). When performing a comparison and annotation of the assembled contigs with the ToCV reference using BLAST, it was observed that over 86% of the contigs exhibited significant similarity to ToCV, a member of the *Closteroviridae* family.

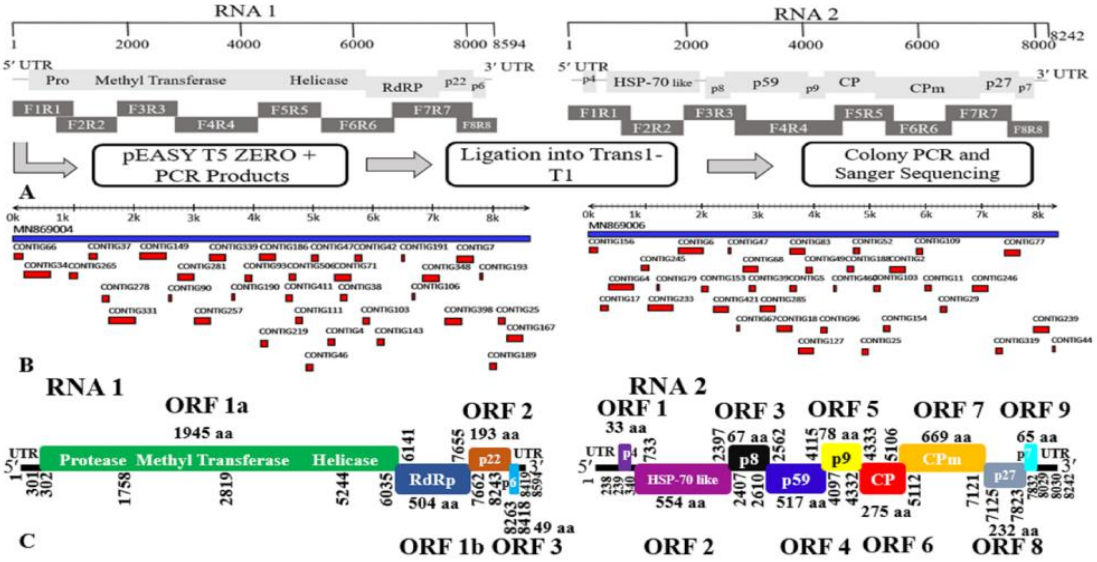

Figure-2. Flow chart from sequencing to full genome analysis of ToCV-Pak1 and ToCV-Pak2. (A) Schematic diagram of full genome amplification using specific primers and Sanger sequencing. (B) Different contigs of RNA1 and RNA2 of ToCV-Pak1 resulted from amplification of the ToCV's full-length genome. MN869004 is RNA1 while MN869006 is RNA2. (C) Genomic structure of Pakistani ToCV RNA1 and RNA2 indicating open reading frames, encoded proteins, and untranslated regions.

BLAST and phylogenetic analysis

The BLAST analysis of the correctly annotated RNA 1 of ToCV-Pak1 and ToCV-Pak2 revealed that RNA1 had the highest similarity (99.41-99.46%) with a South Korean isolate from tomato. In comparison, RNA2 of ToCV-Pak1 had the highest similarity (99.67%) with a Greek isolate. Interestingly, RNA2 of ToCV-Pak2 had 99.64% similarity with a Turkish isolate instead of a Greek isolate. The phylogenetic analysis of 23 complete sequences of RNA1 showed that ToCV is genetically conserved, and two major clades (Clade 1 and 2) were formed (Fig. 3A). Clade 2 contained a single isolate from the Taiwan region of China. In contrast, all other isolates, including the Pakistani isolate, were gathered in Clade 1. Clade 1 consisted of subclade 1A, which comprised all isolates from South Korea, Brazil, Greece, Turkey, China, and the United States; subclade 1B, which comprised only Pakistani isolates; and subclade A3, which comprised two isolates from Spain.

The phylogenetic tree based on RNA2 from 24 isolates from different countries and the two Pakistani isolates also showed two significant clades (Fig. 3B). Clade 1 consisted of 24 isolates within three subclades.

Pakistani isolates are grouped in Subclade 1A with isolates from Greece, Turkey, South Korea, Brazil, the United States, and China. Subclade 1B consisted mainly of European isolates (UK, Spain) but had one South Korean isolate. Subclade 1C had a single isolate (South Africa). Clade 2 consisted only of Chinese isolates (Fig. 3B).

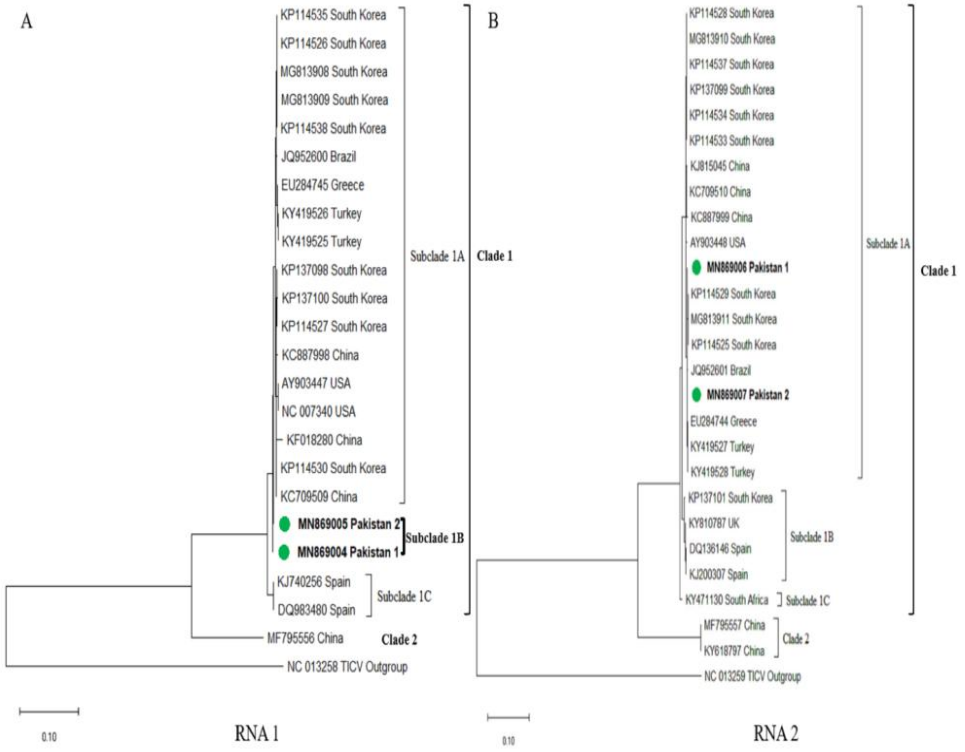

Figure-3. Phylogenetic analysis of RNA1 and RNA2 of Pakistani and worldwide ToCV isolates using the maximum likelihood method based on the Tamura-Nei model using Mega X. All positions with gaps and missing data were removed. Codon positions included were 1st+2nd+3rd+Noncoding. Tree based on nucleotide sequences for (A) RNA1 from 23 ToCV isolates and (B) RNA2 from 26 ToCV isolates, including two from Pakistan.

Sequence analysis

The computer-assisted research determined that the genome of ToCV Pakistani isolates consists of two parts: RNA 1, which has a length of 8594 nucleotides, and RNA 2, which has a length of 8242 nucleotides. RNA 1 had four open reading frames: 302 ORF 1a 6139 (1758 methyltransferase-domain 2819, 5244 helicase-domain 6035), 6141 ORF 1b 7155, 7662 ORF2 8243 (encodes p22) and 8263 ORF3 8418 (encodes P5) (Fig. 2C). The sequence also contains untranslated regions of 301 nt at the 5′ end and 175 nt at the 3′ end. The nucleotide sequence of RNA 1 of Pakistani isolates shared 78-99.4% identities with isolates from other countries. The maximum identities (99.4%) were recorded with Brazilian (JQ952600), Greece (EU284745), and Korean (KP137098) isolates. RNA 2 (flanked by 5′NTR 238 and 8030 3′NTR 8242) contained nine open reading frames: 239 ORF1 340 (small hydrophobic protein P4), 733 ORF2 2397 (heat shock protein 70-like protein), 2407 ORF3 2610 (P8), 2562 ORF4 4115 (putative movement protein P59), 4097 ORF5 4333 (P9), 4333 ORF6 5106 (CP), 5112 ORF7 7121 (minor-coat-protein CPm, encapsidates virion tails), 7125 ORF8 7823 (P27), and 7832 ORF9 8029 (P7) (Fig. 2C). The nucleotide sequence of RNA 2 shared 97-99.6% identities with the other isolates. The maximum identities were recorded with isolates from Greece (EU284744), Korea (KP114529), and Turkey (KY419527).

https://doi.org/10.35495/ajab.2024.112

Selection pressure and recombination analysis

When the ToCV isolates were examined for InDels, seven were found for RNA1 and 14 for RNA2 sequences of studied ToCV populations. RNA 1 of the Pakistani ToCV isolates had 0.007-0.219 evolutionary distance with isolates from other countries; the highest evolutionary distance was 0.219 with a Chinese isolate (Taiwan region) (MF89556), and the lowest (0.007) with two Korean isolates (KP114538, MG813908). Similarly, RNA 2 of the Pakistani ToCV isolates had 0.003-0.217 evolutionary distance with isolates from other countries; the highest was 0.217 with two Chinese isolates (Taiwan region) (MF795557, KY618797), and the lowest was 0.003 with a Greece isolate (EU284744). No recombination event was detected in RNA 2 of either Pakistani isolate. However, RNA 1 of ToCV-Pak1 was found to be a recombinant of a Korean and a Spanish isolate as major and minor parents, respectively, with recombination breakpoints starting from 6172 to 8274 from 5′ ends. RNA1 of ToCV-Pak2 isolate was also found to be recombinant with the Korean and the Spanish isolates as major and minor parents, respectively, from 5′ end (Fig. 4A-4C).

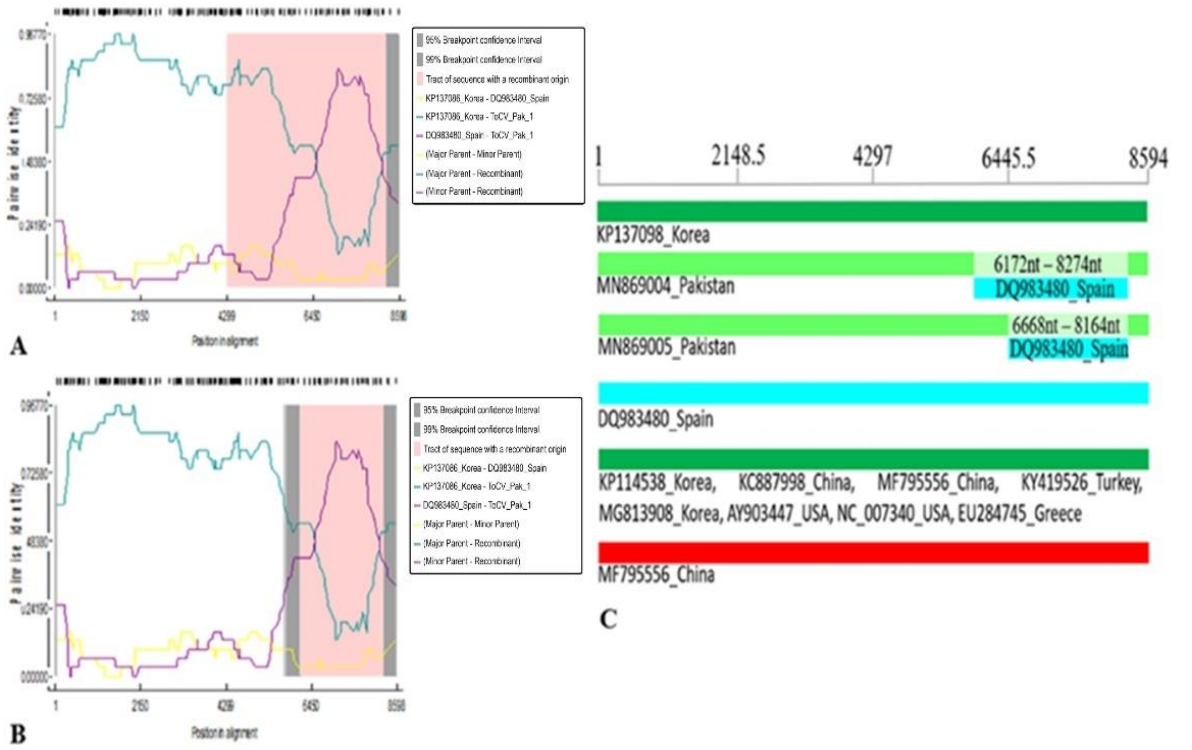

Figure-4. Recombinational analysis of RNA1 of Pakistani ToCV isolates. (A) Pak1-RNA1 recombination between a Korean and a Spanish isolates as major and minor parent, starting from nt 6172 and ending at nt 8274. (B) Pak2-RNA1 recombination between a Korean and a Spanish isolates as major and minor parent, starting from nt 6668 and ending at nt 8164. (C) ToCV-RNA1 recombination analysis of worldwide isolates indicating ToCV-Pak-RNA1 (MN869004, MN869005) as having recombination with a Korean (KP137098) and a Spanish (DQ983480) isolate.

The average P-values based on the methods RDP, GENECONV, BootScan, MaxiChi, Chemeara, and 3Seq were 1.230 x 10-12, 9.828 x 10-06, 5.596 x 10-12, 1.980 x 10-11 and 1.312 x 10-10, 9.325 x 10-15, respectively (Table 3).

Table-3: Average P-values by different detection methods used in recombination analysis

Methods	ToCV Pak 1	ToCV Pak 2
RDP	1.349X10-11	1.349 $X10^{-11}$
GENECONV	4.422 $X10^{-10}$	4.422 $X10^{-10}$
BootScan	8.818 $X10^{-06}$	8.818 $X10^{-06}$
MaxChi	5.624 $X10^{-03}$	5.624 $X10^{-05}$
Chimsera	9.115 $X10^{-06}$	9.115 $X10^{-06}$
SiScan	2.482 $X10^{-07}$	2.482 $X10^{-07}$
3Seq	4.121 $X10^{-10}$	4.121 $X10^{-10}$
LARD	0	0
PhylPro	0	0

Gene flow and genetic differentiation analysis

When the level of gene flow between the Pakistani and the other ToCV isolates from the world in GenBank was assessed, the fixation index (Fst) was 0.00564 (less than 0.33 standard value) in RNA1 and 0.11828 (less than 0.33 standard value) in RNA 2, suggesting frequent gene flow. In the genetic differentiation analysis of RNA 1, the value of the three permutation-based statistics Ks*, Z, and Snn were found to be significant, 4.86231, 34.27778, and 0.55556, respectively. In RNA 2, the Ks*, Z, and Snn values were significant, 5.42299, 49.33333, and 0.71429, respectively.

Haplotype and nucleotide diversity

For evaluating the pattern of molecular diversity at segregating sites, in RNA 1, Tajima's D, Fu, and Li's D* values were negative (-2.10697, -2.55465, respectively). Still, Fu and Li's F* value was positive, indicating a low frequency of polymorphisms in RNA1 of the ToCV populations (Tajima, 1989; Hey and Harris, 1999; Tsompana et al., 2005). However, values of Fu and Li's D* and Fu and Li's F* for RNA 2 were positive (1.44122, 1.00479, respectively), and Tajima's D was negative (-0.61336), indicating a high polymorphic frequency in RNA 2 of understudied ToCV population.

Discussion

Criniviruses infect essential crop species in the family *Cucurbitaceae* and *Solanaceae*. Several criniviruses are considered potential agents causing epidemics worldwide (O'Sullivan et al., 2013). ToCV has been associated with TYD in open fields in the Multan region of Pakistan. Until now, ToCV-resistant varieties of any cultivated crops are not commercially available, and managing its vector poses an effective control measure to reduce the incidence and severity of this virus. However, it is evident that wild tomatoes produce relatively milder symptoms and have lower viral titer when tested in experimental and lab conditions (García-Cano et al., 2010; Shakeel et al., 2017; Mansilla-Córdova et al., 2018; Zhao et al., 2025). The virus can be inoculated naturally to the plants through whitefly while an effective approach of infectious clones have been adopted to inoculate the plants under experimental conditions that has boosted to findings of the host-virus interactions and the underlying mechanisms (Navas-Hermosilla et al., 2021)

The bipartite genome of Pakistani ToCV isolates was analyzed using computer-assisted methods. The study indicated that the genome consists of two RNA molecules, RNA 1 and RNA 2, which are 8594 and 8242 long nucleotides. These RNA molecules contain four and nine open reading frames (ORFs) and are bordered by untranslated sections at the 5′ and 3′ ends. The findings presented in this study are consistent with prominent genetic features that are unique to ToCV as reported in previous studies (Wintermantel and Wisler, 2006; Zhao et al., 2014). The Open Reading Frame 1a (ORF 1a) gene encodes a widely occurring multifunctional protein with a molecular weight of 221-kDa. This protein is translated into many domains, including protease, methyltransferase, and helicase, collectively called Pro-MT-HEL (Karasev, 2000; Lozano et al., 2007; Fiallo-Olivé and Navas-Castillo, 2019). The Open Reading Frame 1b (ORF 1b) encodes the RNA-dependent RNA polymerase, a protein with a molecular weight of 59 kDa. This area is highly conserved and may be found in several species of the genus Crinivirus. It is previously reported that hallmark gene array consists of ORFs for a putative small hydrophobic protein (SHP), coat protein (CP), minor coat protein (CPm), a heat shock protein homolog (Hsp70h), that is translated from

subgenomic RNAs which is a common phenomenon in the family *Closteroviridae* (Karasev et al., 1995; Klaassen et al., 1995; Karasev et al., 1996; Jelkmann et al., 1997; Fazeli and Rezaian, 2000).

Recombination events usually bring about genetic diversity among the viruses (Suzuki et al., 1998; Varsani et al., 2006). More specifically, an increased genetic variability and virus evolution occurring in the plant viruses is attributed to viral RNA-RNA recombination and/or reassortment of genomic segments (pseudo-recombination) that results in the emergence of new viral strains or species (Varsani et al., 2006; Chattopadhyay and Mandal, 2020). Nevertheless, in the study done on ToCV isolates detected from Pakistan, no recombination event was detected in RNA 2, but RNA 1 was found to be a recombinant of Brazilian isolate (JQ952600, as a major parent) and Spanish isolate (DQ983480, as a minor parent) suggesting no geographical association of the Pakistani ToCV isolates with neighboring countries. One reason for these findings may be the unavailability of complete sequence data for the genome of many ToCV isolates worldwide. Currently, the complete genomic sequence is available for only 23 ToCV isolates from just 11 countries in NCBI database, which seems insufficient to reveal a satisfactory insight into recombination. In the genetic differentiation analyses, the positive values of the neutrality test suggested a frequent gene flow between Pakistani ToCV isolates and the isolates reported from other parts of the world (fixation index: Fst), suggesting ToCV populations are under positive selection. RNA recombination can also be a process to purge deleterious mutations from viral populations that may pile up because of the high mutation rates of RNA viruses and thus manage somehow to escape loss fitness loss in small viral populations (Moya et al., 2004; Wu et al., 2025). These mutations have not been found to be very common, which might result from selection pressure that rules out most of the recombinants (Sztuba-Solińska et al., 2011). The findings of this research project the same trend that has been observed in the evolutionary studies of other criniviruses where distinct selection pressures act on different genomic components (Butković and González, 2022; Fiallo-Olivé and Navas-Castillo, 2019; LaTourrette and Garcia-Ruiz, 2022).

The value of Tajima's D was positive (2.10697) for RNA 1 and negative (-0.61336) for RNA 2. Fu and Li's D* and Fu and Li's F* tests were negative for RNA 1, suggesting a strong negative or purifying selection in the ToCV population. However, for RNA 2, these values were positive, indicating a weak positive or lack of purifying selection. To interpret the comparative genomics and role of genome-based evolution, it is necessary to highlight the number of substitutions per site, which depicts the extent of separating a pair of homologous sequences since they evolved and diverged from their typical ancestral sequence (Rosenberg, 2005). The highest evolutionary distance for RNA1 and RNA2 was 0.219 and 0.217, respectively, which indicates that the evolution is low in the case of ToCV. It can be concluded that the positive values of the neutrality test such as Tajima's D, Fu and Li's D*, and Fu and Li's F* for RNA2 predicts a weak positive selection and on-going evolutionary events which might indicate the virus potential to change its pathogenic behavior however in RNA1, the negative values refer its tendency to preserve the functional integrity of the viral genes that are related to replication process and key enzymatic proteins. The findings highlight that RNA 1 of ToCV is genetically conserved as compared to RNA 2 that might impact the adoptability and host. Other members of *Closteroviridae* such as *Citrus tristeza virus (CTV)* and *Grapevine leafroll-associated virus 3 (GLRaV-3)* exhibit recombination events with as similar trend, where they contribute to viral adaptability and fitness (Jeger et al., 2006; Bar-Joseph and Mawassi, 2013; Boulila, 2010; O'Sullivan et al., 2013; Rubio et al., 2020).

According to the recent studies conducted on the transcriptome ToCV, it has been established that the early infection of ToCV in the tomato plants instigates the plant immune response which results in alteration in the expression of flavonoid and steroids, photosynthesis-related genes, hormone signaling pathways that may affect the yield potential of the plants (García-Cano et al., 2010; Mansilla-Córdova et al., 2018). On the other hand, it is also evident that Hsp70 and sgt1 protein are crucially important in resistance development against ToCV as their suppression through viral induced gene silencing (VIGS) is strongly related to the enhanced accumulation of viral particles. These findings provide a solid basis of utilizing host factors to device novel strategies for enhanced resistance against ToCV infection.

To summarize, ToCV is emerging as a serious threat to tomato production in three districts in South Punjab of Pakistan. The complete genome of Pakistani isolates has been characterized, giving broad insight

into the ORFs and their encoded proteins. Multiple sequence alignment revealed the conserved nature of Pak-ToCV isolates. The recombination of RNA2 is unique and might play a role in the widespread spread of the virus in the Country. Further work should be done on large-scale data to look for the role of recombination in viral spread.

Conclusion

Tomato Chlorosis Virus is a semi-persistent whitefly-transmitted virus, continuously threatening tomato crops cultivated globally. Early detection using sensitive molecular tools such as RT-PCR and LAMP is essential for effective disease surveillance and management. The recombination events and ongoing genomic evolution may influence the virulence, host range, and adoption, virus-vector interactions, and diagnosis. Although conventional control measures offer partial control, studying the viral genome may strengthen disease control through crop resistance and vector suppression, which presents sustainable disease management. Future research should focus on understanding recombination-driven genetic diversity, enhancing field-ready diagnostics, and breeding for durable resistance to ensure effective and long-term ToCV management.

Acknowledgments

We thank Dr. Liu Wen Wen and Dr. Wang Hui for their help during the experimental work.

Disclaimer: None

Conflict of Interest: None.

Source of Funding: This work was jointly funded by the project of Scientific research start-up funds for doctoral talents of Zhaotong University - Mingzheng Duan, Grant number: 202406; Young Talent Project of Talent Support Program for the Development of Yunnan. Grant number: 210604199008271015.

Data Availability Statement

The relevant data is provided in the paper. Supplementary files also contain information to better understand the studies. However, any other relevant data/information regarding the research would be provided upon inquiry by the readers.

Supplementary Material

Figure S1: Agarose gel showing amplification of 463 bp of heat shock protein gene from 47 tomato field samples from Multan (Lanes 1-39) and Lodhran district (Lanes 40-66), Figure S2: sRNA profile of tomato samples of Pakistan, Table S1: Primers designed for full length amplification of RNA1 and RNA2 of ToCV, Table S2: GenBank Accession numbers for ToCV and TICV isolates used for genomic diversity and other analyses, Table S3: sRNA profile from tomato samples from Pakistan.

Contribution of Authors

Raza A, Shakeel MT, Iqbal S, Khan AA, Abbas T, Hasnain A, Amer MA, Abdul Majid, Ahmad A, Umar UD, Iqbal MU, Aslam MN, Wang X & Duan M: Conceptualized and designed the study.

Raza A, Shakeel MT, Ahmad A & Abdul Majid: Material preparation, data collection, and analysis.

Khan AA, Abbas T & Iqbal MU: Performed data validation of the first draft.

Raza A, Shakeel MT, Ahmad A, Iqbal S & Abdul Majid: Data analysis and preparation of initial draft.

Amer MA, Umar UD, Aslam MN, Hasnain A, Wang X & Duan M: Contributed to final draft.

Wang X: Supervised the research conduction.

Duan M: Provided the financial liabilities.

All authors read and approved the final draft of manuscript.

References

Arruabarrena A, Rubio L, González-Arcos M, Maeso D, Fonseca M and Boiteux L, 2014. First report of *Tomato chlorosis virus* infecting tomato crops in Uruguay. Plant Dis., 98(10): 1445-1445. https://doi.org/10.1094/PDIS-11-13-1153-PDN.

Arruabarrena A, Rubio L, González-Arcos M, Maeso D, Sánchez-Campos S, Fonseca M and Boiteux L, 2015. First report of *Solanum sisymbriifolium* and *S. americanum* as natural weed hosts of *Tomato chlorosis virus* (genus Crinivirus) in South America. Plant Dis. 99(6): 895-895. https://doi.org/10.1094/PDIS-12-14-1315-PDN.

Balasubramanian V and Selvarajan R, 2014. Genetic diversity and recombination analysis in the

coat protein gene of *Banana bract mosaic virus*. Virus Genes. 48: 509-517. DOI 10.1007/s11262-014-1056-x.

Bar-Joseph M and Mawassi M, 2013. The defective RNAs of *Closteroviridae*. Front. microbiol. 4: 132. https://doi.org/10.3389/fmicb.2013.00132.

Boulila M, 2010. Selective pressure, putative recombination events and evolutionary relationships among members of the family *Closteroviridae:* A proposal for a new classification. Biochem Syst Ecol. 38(6): 1185-1192. https://doi.org/10.1016/j.bse.2010.12.002.

Butković A and González R, 2022. A brief view of factors that affect plant virus evolution. Front. Virol. 2: 994057. DOI: 10.3389/fviro.2022.994057

Chattopadhyay A and Mandal B, 2020. Hypotheses of virus origin and evolutionary patterns of plant viruses. In Applied plant virology (pp. 779-796). AP. https://doi.org/10.1016/B978-0-12-818654-1.00054-2

Dovas C, Katis N and Avgelis A, 2002. Multiplex detection of criniviruses associated with epidemics of a yellowing disease of tomato in Greece. Plant Dis. 86(12): 1345-1349. DOI: 10.1094/PDIS.2002.86.12.1345.

Fazeli C and Rezaian M, 2000. Nucleotide sequence and organization of ten open reading frames in the genome of *Grapevine leafroll-associated virus 1* and identification of three subgenomic RNAs. J Gen Virol. 81(3): 605-615. https://doi.org/10.1099/0022-1317-81-3-605.

Fiallo-Olivé E, Espino A, Botella-Guillén M, Gómez-González E, Reyes-Carlos J and Navas-Castillo J, 2014. Tobacco: a new natural host of *Tomato chlorosis virus* in Spain. Plant Dis. 98(8): 1162-1162. https://doi.org/10.1094/PDIS-01-14-0083-PDN

Fiallo-Olivé E and Navas-Castillo J, 2019. *Tomato chlorosis virus*, an emergent plant virus still expanding its geographical and host ranges. Mol. Plant Pathol. 20(9): 1307-1320.

Fonseca M, Boiteux L, Abreu H, Nogueira I and Pereira-Carvalho R, 2013. *Physalis angulata*: a new natural host of *Tomato chlorosis virus* in Brazil. Plant Dis. 97(5): 692-692. https://doi.org/10.1094/PDIS-06-12-0593-PDN.

Fortes I, Moriones E and Navas-Castillo J, 2012. *Tomato chlorosis virus* in pepper: prevalence in commercial crops in southeastern Spain and symptomatology under experimental conditions. Plant Pathol. 61(5): 994-1001. https://doi.org/10.1111/j.1365-3059.2011.02584.x.

Fortes I and Navas-Castillo J, 2008. Potato (*Solanum tuberosum*) is an experimental and natural host of *Tomato chlorosis virus*. 3rd Conference of the International Working Group on Legume and Vegetable Viruses. Ljubljana, Slovenia,

García-Cano E, Navas-Castillo J, Moriones E and Fernández-Muñoz R, 2010. Resistance to *Tomato chlorosis virus* in wild tomato species that impair virus accumulation and disease symptom expression. Phytopathol. 100(6): 582-592. https://doi.org/10.1094/PHYTO-100-6-0582.

Hanssen I, Lapidot M and Thomma B, 2010. Emerging viral diseases of tomato crops. MPMI. 23(5): 539-548. https://doi.org/10.1094/MPMI-23-5-0539.

Hey J and Harris E, 1999. Population bottlenecks and patterns of human polymorphism. Mol. Biol. Evol. 16(10): 1423-1426.

Jeger M, Seal SE and Van den Bosch F, 2006. Evolutionary epidemiology of plant virus disease. Adv. Virus Res. 67: 163-203. https://doi.org/10.1016/S0065-3527(06)67005-X

Jelkmann W, Fechtner B and Agranovsky A, 1997. Complete genome structure and phylogenetic analysis of *little cherry virus*, a mealybug-transmissible closterovirus. J Gen Virol. 78(8): 2067-2071. https://doi.org/10.1099/0022-1317-78-8-2067.

Karasev A, Boyko V, Gowda S, Nikolaeva O, Hilf M, Koonin E, Niblett C, Cline K, Gumpf D and Lee R, 1995. Complete sequence of the *citrus tristeza virus* RNA genome. Virol. J. 208(2): 511-520. https://doi.org/10.1006/viro.1995.1182.

Karasev A, 2000. Genetic diversity and evolution of closteroviruses. Annu. Rev. Phytopathol. 38(1): 293-324. https://doi.org/10.1146/annurev.phyto.38.1.293.

Karasev A, Nikolaeva O, Mushegian A, Lee R and Dawson W, 1996. Organization of the 3′-terminal half of *beet yellow stunt virus* genome and implications for the evolution of closteroviruses. Virol. J. 221(1): 199-207. https://doi.org/10.1006/viro.1996.0366.

Kil E, Lee Y, Cho S, Auh C, Kim D, Lee K, Kim M, Choi H, Kim C and Lee S, 2015. Identification of natural weed hosts of *Tomato chlorosis virus* in Korea by RT-PCR with root tissues. Eur. J. Plant Pathol. 142: 419-426. DOI:10.1007/s10658-015-0622-y.

Klaassen V, Boeshore M, Koonin E, Tian T and Falk B, 1995. Genome structure and phylogenetic analysis of *lettuce infectious yellows virus*, a whitefly-transmitted, bipartite closterovirus. Virol. J. 208(1): 99-110. https://doi.org/10.1006/viro.1995.1133.

Kumar S, Stecher G and Tamura K, 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33(7): 1870-1874. https://doi.org/10.1093/molbev/msw054.

LaTourrette K and Garcia-Ruiz H, 2022. Determinants of virus variation, evolution, and host adaptation. Pathogens. 11(9): 1039. https://doi.org/ 10.3390/pathogens11091039

Li H and Durbin R, 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinform. 25(14): 1754-1760. https://doi.org/10.1093/bioinformatics/btp324.

Librado P and Rozas J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinform. 25(11): 1451-1452. https://doi.org/10.1093/bioinformatics/btp187.

Lozano G, Moriones E and Navas-Castillo J, 2004. First report of sweet pepper (*Capsicum annuum*) as a natural host plant for *Tomato chlorosis virus*. Plant Dis. 88(2): 224-224. https://doi.org/10.1094/PDIS.2004.88.2.224A.

Lozano G, Moriones E and Navas-Castillo J, 2007. Complete sequence of the RNA1 of a European isolate of *tomato chlorosis virus*. Arch. Virol. 152(4): 839-842.

Mamoun Abdel-Salam AM, Rezk AA and Rehab A, 2019. Biochemical, Serological, Molecular and Natural Host Studies on *Tomato Chlorosis Virus* in Egypt. PJBS 22(2): 83-94. https://doi.org/10.3923/pjbs.2019.83.94.

Mansilla-Córdova P, Bampi D, Rondinel-Mendoza N, Melo P, Lourenção A and Rezende J, 2018. Screening tomato genotypes for resistance and tolerance to *Tomato chlorosis virus*. Plant Pathol. 67(5): 1231-1237. https://doi.org/10.1111/ppa.12826.

Ministry of National Food Security and Research, 2023. Crops Area & Production (District Wise), 2022-23 (p. 93). Government of Pakistan.

Moya A, Holmes E and González-Candelas F, 2004. The population genetics and evolutionary epidemiology of RNA viruses. Nat. Rev. Microbiol. 2(4): 279-288.

Navas-Castillo J, Camero R, Bueno M and Moriones E, 2000. Severe yellowing outbreaks in tomato in Spain associated with infections of *Tomato chlorosis virus*. Plant Dis. 84(8): 835-837. https://doi.org/10.1094/PDIS.2000.84.8.835.

Navas-Castillo J, Fiallo-Olivé E and Sánchez-Campos S, 2011. Emerging virus diseases transmitted by whiteflies. Annu. Rev. Phytopathol. 49: 219–248. https://doi.org/10.1146/annurev-phyto-072910-095235

Navas-Hermosilla E, Fiallo-Olivé E and Navas-Castillo J, 2021. Infectious clones of *Tomato chlorosis virus*: toward increasing efficiency by introducing the Hepatitis Delta virus ribozyme. Front. microbiol. 12: 693457. https://doi.org/10.3389/fmicb.2021.693457

O'Sullivan D, Giblin L, McSweeney P, Sheehan J and Cotter P, 2013. Nucleic acid-based approaches to investigate microbial-related cheese quality defects. Front. microbiol. 4: 1. https://doi.org/10.3389/fmicb.2013.00001.

Orfanidou C, Dimitriou C, Papayiannis L, Maliogka V and Katis N, 2014. Epidemiology and genetic diversity of criniviruses associated with tomato yellows disease in Greece. Virus Res. 186: 120-129. https://doi.org/10.1016/j.virusres.2013.12.013.

Orfanidou C, Pappi P, Efthimiou K, Katis N and Maliogka V, 2016. Transmission of *Tomato chlorosis virus* (ToCV) by Bemisia tabaci biotype Q and evaluation of four weed species as viral sources. Plant Dis. 100(10): 2043-

2049. https://doi.org/10.1094/PDIS-01-16-0054-RE.

Raza A, Shakeel MT, Umar UU, Tahir MN, Hassan AA, Katis NI and Wang X, 2020. First report of *tomato chlorosis virus* infecting tomato in Pakistan. Plant Dis. 104(2036): 10.1094.

Rubio L, Galipienso L and Ferriol I, 2020. Detection of plant viruses and disease management: Relevance of genetic diversity and evolution. *Front. Plant Sci*. 11, 1092. https://doi.org/10.3389/fpls.2020.01092

Rosenberg M, 2005. Evolutionary distance estimation and fidelity of pair wise sequence alignment. BMC Bioinformatics. 6: 1-9. doi:10.1186/1471-2105-6-102.

Shakeel MT, Al-Saleh MA, Amer MA, Al-Shahwan IM, Umar M, Dimou N, Orfanidou CG, Zakri AM and Katis NI, 2017. Molecular characterization and natural host range of *Tomato chlorosis virus* in Saudi Arabia. Plant Pathol. J. 99(2): 415-421. https://www.jstor.org/stable/44686786.

Singh K, Jarošová J, Fousek J, Huan C and Kundu JK, 2020. Virome identification in wheat in the Czech Republic using small RNA deep sequencing. J. Integr. Agric. 19(7): 1825-1833. https://doi.org/10.1016/S2095-3119(19)62805-4.

Solórzano-Morales A, Barboza N, Hernández E, Mora-Umaña F, Ramírez P and Hammond RW, 2011. Newly discovered natural hosts of *Tomato chlorosis virus* in Costa Rica. Plant Dis. 95(4): 497-497. https://doi.org/10.1094/PDIS-11-10-0836.

Suzuki Y, Gojobori T and Nakagomi O, 1998. Intragenic recombinations in rotaviruses. FEBS Lett. 427(2): 183-187. https://doi.org/10.1016/S0014-5793(98)00415-3.

Sztuba-Solińska J, Urbanowicz A, Figlerowicz M and 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.

Tajima F, 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genet. 123(3): 585-595. https://doi.org/10.1093/genetics/123.3.585.

Tamura K and Nei M, 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10(3): 512-526. https://doi.org/10.1093/oxfordjournals.molbev.a040023.

Tayyib M, Suhail A, Arif M and Khan MA, 2013. Diversity of whiteflies (*Aleyrodidae:* Homoptera) in Punjab (Pakistan) PhD thesis unpublished. 121 pp. DOI:10.13140/RG.2.2.15366.19528.

Tsai W, Shih S, Green S, Hanson P and Liu HY, 2004. First report of the occurrence of *Tomato chlorosis virus* and *Tomato infectious chlorosis virus* in Taiwan. Plant Dis. 88(3): 311-311. https://doi.org/10.1094/PDIS.2004.88.3.311B

Tsompana M, Abad J, Purugganan M and Moyer JW, 2005. The molecular population genetics of the *Tomato spotted wilt virus* (TSWV) genome. Mol. Ecol. 14(1): 53-66. https://doi.org/10.1111/j.1365-294X.2004.02392.x.

Varsani A, van der Walt E, Heath L, Rybicki E, Williamson A and Martin D, 2006. Evidence of ancient papillomavirus recombination. JGV. 87(9): 2527-2531. https://doi.org/10.1099/vir.0.81917-0.

Wang X, Feng J, Zang L, Yan Y, Yang Y and Zhu X, 2018. Natural occurrence of *Tomato chlorosis virus* in cowpea (*Vigna unguiculata*) in China. Plant Dis. 102(1): 254. https://doi.org/10.1094/PDIS-02-17-0288-PDN.

Wintermantel W, Wisler G, Anchieta A, Liu H, Karasev A and Tzanetakis I, 2005. The complete nucleotide sequence and genome organization of *Tomato chlorosis virus*. Arch. Virol. 150: 2287-2298. DOI 10.1007/s00705-005-0571-4.

Wintermantel WM and Wisler GC, 2006. Vector specificity, host range, and genetic diversity of *Tomato chlorosis virus*. Plant Dis. 90(6): 814-819 https://doi.org/10.1094/PD-90-0814.

Wisler G, Duffus J, Liu HY and Li R, 1998. Ecology and epidemiology of whitefly-transmitted closteroviruses. Plant Dis. 82(3): 270-280. https://doi.org/10.1094/PDIS.1998.82.3.270.

Wisler G, Li R, Liu HY, Lowry D and Duffus J, 1998. *Tomato chlorosis virus*: a new whitefly-transmitted, phloem-limited, bipartite closterovirus of tomato. Phytopathol. 88(5):

402-409. https://doi.org/10.1094/PHYTO.1998.88.5.402.

Wu K, Zhang S, Huang W, Du Z, Gao F and Guan X, 2025. Phylogeographic History of *Tomato Chlorosis Virus*. Viruses. 17(4): 457. https://doi.org/10.3390/v17040457

Zhao D, Xia T, Zhou T, Zhao L, Zhu X and Gong B, 2025. Inoculation method and disease evaluation of *Tomato chlorotic virus* (ToCV) in *Solanum lycopersicum*. Veg. Res. 5(1). https://doi.org/10.48130/vegres-0025-0002

Zhao LM, Li G, Gao Y, Liu YJ, Sun GZ and Zhu XP, 2014. Molecular detection and complete genome sequences of *Tomato chlorosis virus* isolates from infectious outbreaks in China. Phytopathol. 162(10): 627-634. https://doi.org/10.1111/jph.12236.

Zheng Y, Gao S, Padmanabhan C, Li R, Galvez M, Gutierrez D, Fuentes S, Ling KS, Kreuze J and Fei Z, 2017. VirusDetect: An automated pipeline for efficient virus discovery using deep sequencing of small RNAs. Virol. J. 500: 130-138. https://doi.org/10.1016/j.virol.2016.10.017.

Zhou Y, Yan J, Qiao G, Liu M, Zhang W and Li X, 2015. First report of *Tomato chlorosis virus* infecting eggplant (*Solanum melongena*) in China. Plant Dis. 99(11): 1657-1657 https://doi.org/10.1094/PDIS-01-15-0014-PDN.

Supplementary Material

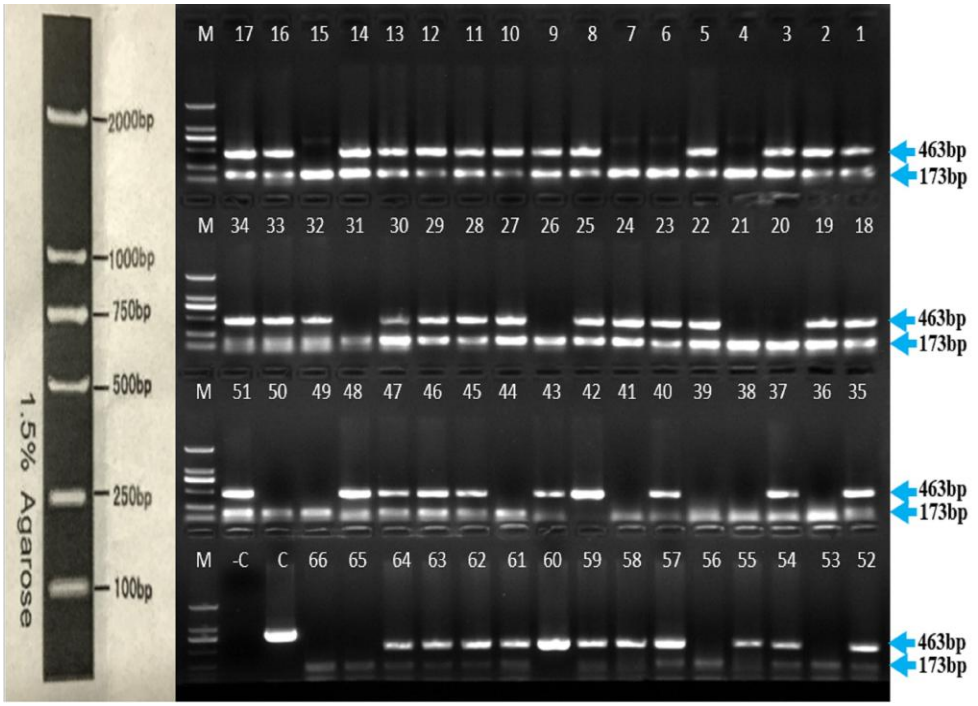

Figure-S1. Agarose gel showing amplification of 463 bp of heat shock protein gene from 47 tomato field samples from Multan (Lanes 1-39) and Lodhran district (Lanes 40-66); lanes 4, 6, 7, 15, 20, 21, 26, 31, 36, 38, 39, 41, 44, 49, 50, 53, 56, 65, and 66 were negative. Lanes 1-66 also show the amplification of 173 bp of CAC reference gene of tomato; M: marker; C: internal positive control. -C: internal negative control.

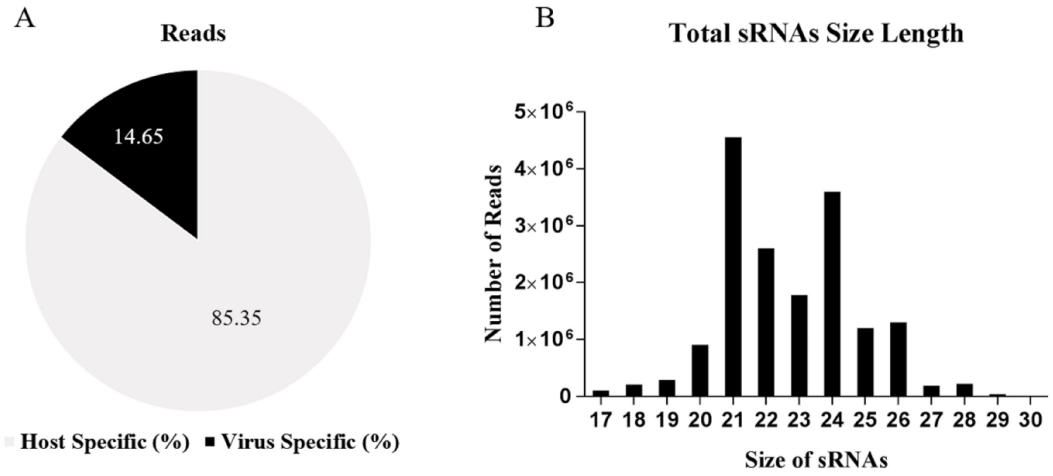

Figure-S2: sRNA profile of tomato samples of Pakistan. A) Percentage of host-specific and virus-specific reads in the field sample from Pakistan. B) Size length distribution of total sRNAs indicating 21-nt sRNA as most abundant.

https://doi.org/10.35495/ajab.2024.112

Table-S1: Primers designed for full length amplification of RNA1 and RNA2 of ToCV

Primer Name	Sequence (5′ - 3′)	Amplification Size (bp)	Annealing Temperature
ToCV-RNA1-F1	TAGCTTTAAGGCTTTCTGGTGTG	1276	54 °C
ToCV-RNA1-R1	TTACCGTTGGTGAGAGTAGC		
ToCV-RNA1-F2	GGCCAGAAGATGGTTATGGA	1211	55 °C
ToCV-RNA1-R2	TCATAAACCTCCACCGCTAC		
ToCV-RNA1-F3	AACTTCATCATGGCCAAACG	1151	55 °C
ToCV-RNA1-R3	ACAGTGACCTCTTGCTTGAA		
ToCV-RNA1-F4	AGGCTAAGTGTTACGATGGG	1393	53.5 °C
ToCV-RNA1-R4	GCCACCAACAGAATACACAC		
ToCV-RNA1-F5	GTGTGTATTCTGTTGGTGGC	1263	55.5 °C
ToCV-RNA1-R5	TTTGTCACCGTAACACTCCA		
ToCV-RNA1-F6	TGGAGTGTTACGGTGACAAA	1196	55.5 °C
ToCV-RNA1-R6	ATCTTCGGTTTCATCTCCCC		
ToCV-RNA1-F7	TGACGATTGGTTGCAGGATA	1616	54 °C
ToCV-RNA1-R7	ACCTAAAATAACCCGACCCC		
ToCV-RNA1-F8	AAGTTGGGGTCGGGTTATTT	411	52.5 °C
ToCV-RNA1-R8	AGATCTACCAAGGTCTACGTGT		
ToCV-RNA2-F1	CTAGTCCAGGTGTTTCCTGTGG	1254	55 °C
ToCV-RNA2-R1	GGTTCATTGACTATGCGACG		
ToCV-RNA2-F2	CAATGTGTCTGTTCCGGCTG	775	55 °C
ToCV-RNA2-R2	CGGAACAGACATGGGTTTTC		
ToCV-RNA2-F3	GAAAACCCATGTCTGTTCCG	1056	55 °C
ToCV-RNA2-R3	TCTTCAGGATGGTCGGCAAG		
ToCV-RNA2-F4	AAGGAACCTTGGCAGGTTGT	1680	53.5 °C
ToCV-RNA2-R4	GATGTTGAATCGTCCCTCCT		
ToCV-RNA2-F5	GGAAATAATGGGGGACAGTGCT	1404	55 °C
ToCV-RNA2-R5	GCAATTTGTTTTCGCGCACC		
ToCV-RNA2-F6	CAGTACCCGTTCGTGAAGTA	951	57.5 °C
ToCV-RNA2-R6	TGTCATCGGACTCAGAGGGT		
ToCV-RNA-F7	AGTGATCGTTCCAAGTCTGG	1216	55 °C
ToCV-RNA-R7	AATACAATGCTGCGTTCACGTC		
ToCV-RNA-F8	ACAAGATACGGGGTTGACGA	754	50 °C
ToCV-RNA-R8	CGACCTATTTATTTATATAC		

Table-S2: GenBank Accession numbers for ToCV and TICV isolates used for genomic diversity and other analyses

Accession Numbers	Virus	Country of origin	Host	Genome segment
MN869004	ToCV	**Pakistan**	Tomato	RNA1
MN869005	ToCV	**Pakistan**	Tomato	RNA1
KP114538	ToCV	South Korea	Tomato	RNA1
MG813908	ToCV	South Korea	Tomato	RNA1
AY903447	ToCV	USA	Tomato	RNA1
NC_007340	ToCV	USA	Tomato	RNA1
JQ952600	ToCV	Brazil	Tomato	RNA1
EU284745	ToCV	Greece	Tomato	RNA1
KP137098	ToCV	South Korea	Tomato	RNA1
KY419526	ToCV	Turkey	Tomato	RNA1
KC887998	ToCV	China	Tomato	RNA1
DQ983480	ToCV	Spain	Tomato	RNA1
MF795556	ToCV	China	Tomato	RNA1
KP114535	ToCV	South Korea	Tomato	RNA1
KP114526	ToCV	South Korea	Tomato	RNA1

KP137100	ToCV	South Korea	Tomato	RNA1
MG813909	ToCV	South Korea	Tomato	RNA1
KP114527	ToCV	South Korea	Tomato	RNA1
KY419525	ToCV	Turkey	Tomato	RNA1
KF018280	ToCV	China	Tomato	RNA1
KP114530	ToCV	South Korea	Tomato	RNA1
KJ740256	ToCV	Spain	Tomato	RNA1
KC709509	ToCV	China	Tomato	RNA1
NC_013258	TICV	USA	Tomato	RNA1
MN869006	ToCV	**Pakistan**	Tomato	RNA2
MN869007	ToCV	**Pakistan**	Tomato	RNA2
EU284744	ToCV	Greece	Tomato	RNA2
KP114529	ToCV	South Korea	Tomato	RNA2
KY419527	ToCV	Turkey	Tomato	RNA2
JQ952601	ToCV	Brazil	Tomato	RNA2
MG813911	ToCV	South Korea	Tomato	RNA2
KC709510	ToCV	China	Tomato	RNA2
DQ136146	ToCV	Spain	Tomato	RNA2
MF795557	ToCV	China	Tomato	RNA2
KY618797	ToCV	China	Tomato	RNA2
KJ815045	ToCV	China	Tomato	RNA2
AY903448	ToCV	USA	Tomato	RNA2
KY810787	ToCV	United Kingdom	Tomato	RNA2
KJ200307	ToCV	Spain	Tomato	RNA2
KP114528	ToCV	South Korea	Tomato	RNA2
KP114537	ToCV	South Korea	Tomato	RNA2
KP114534	ToCV	South Korea	Tomato	RNA2
KP114533	ToCV	South Korea	Tomato	RNA2
MG813910	ToCV	South Korea	Tomato	RNA2
KP137099	ToCV	South Korea	Tomato	RNA2
KP114525	ToCV	South Korea	Tomato	RNA2
KP137101	ToCV	South Korea	Tomato	RNA2
KC887999	ToCV	China	Tomato	RNA2
KY419527	ToCV	Turkey	Tomato	RNA2
KY419528	ToCV	Turkey	Tomato	RNA2
KY471130	ToCV	South Africa	Tomato	RNA2
NC_013259	TICV	USA	Tomato	RNA2

Table-S3: sRNA profile from tomato samples from Pakistan

Isolates	Overall	Trimmed	% Age	Host Specific (%)	Virus Specific (%)
ToCV-Pak1	10,401,985	6,256,489	60	84.6	15.4
ToCV-Pak2	6,594,967	4,223,236	64	86.1	13.9
Total	16,996,952	10,479,725	62	85.35	14.65