Volume 3, Issue 5, 2025

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

GENETIC DIVERSITY AND PHYLOGENETIC ANALYSIS OF CUCUMBER MOSAIC VIRUS ISOLATES FROM CUCURBITS IN THE PUNJAB REGION OF PAKISTAN

Shehreen Sohail[*1], Taufiq Ahmad[2], Mehvish Sarwar[3], Kaleem Ullah[4], Amna Saleem[5], Fareeha Sohail[6], Batool Yasin[7], Husnain Karim Riaz[8]

[1, 7, 8]Department of Microbiology, University of Central Punjab, Lahore, Pakistan
[2]Department of Eastern Medicine, Faculty of Pharmacy and Health Sciences University of Baluchistan, Quetta
[3] National Institute of Health, Pakistan
[4, 5]Department of Biotechnology, University of Central Punjab, Lahore, Pakistan
[6]Department of Microbiology, Government College University, Lahore, Pakistan

[*1]shehreen.baig19@gmail.com

DOI: https://doi.org/10.5281/zenodo.15878422

Keywords
Cucumber mosaic virus, Tobacco mosaic virus, Cucumovirus, Crops, GMOs

Article History
Received on 08 Aril 2025
Accepted on 29 June 2025
Published on 14 July 2025

Copyright @Author
Corresponding Author: * Shehreen Sohail

Abstract

Cucumber mosaic virus (CMV), which is a highly destructive plant virus that affects nearly 1000 types of plants worldwide, causing significant economic losses in agricultural crops. The genus Cucumovirus is responsible for a variety of plant diseases, including aphids, fungi, viruses, and parasites. In this study, a total of 825 species and 118 genera are included. The objective of detecting and identifying the Pakistani isolates of Cucumber mosaic virus involves the development and optimization of a reliable diagnostic assay or protocol that can effectively detect the virus from infected plant samples. The ultimate goal is to establish the phylogenetic relationship of Pakistani isolates and sensitive and specific detection method that can be used to accurately identify the presence of Cucumber mosaic virus in Pakistani isolates. Cucumber mosaic virus (CMV) infection in cucurbit crops in the Lahore district of Pakistan. RNA isolated from infected plants was used in reverse transcriptase PCR (RT-PCR) to amplify about 600 bp DNA fragments specific to the virus. The amplified PCR products were purified using a commercial kit and then sequenced. Multiple sequence alignment and phylogenetic analysis were performed to compare the nucleotide and amino acid sequences of the CP gene of the Pakistani CMV isolates with those of other CMV isolates in NCBI GenBank. The results showed that the Pakistani isolates were closely related to isolates from other countries and formed a separate group based on amino acid sequence analysis. The study demonstrates that CMV is a common problem in cucurbit crops in Pakistan, causing characteristic symptoms such as leaf deformation, reduced size, severe mosaic, and blistering. This research also concludes by stating the need for further research to understand economic impact and training farmer communities of Cucumber mosaic virus to manage disease through adaptation of resistant cultivars and vector population control in Pakistan.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

INTRODUCTION

Viruses, the atomically concealed infectious creatures, are frequently the cause of global devastation. The simplest description of viruses is a protein encapsulating nucleic acid (DNA/RNA) fragment/fragments. This nucleic acid actually conveys genetic information in the form of codon sequences for three or more protein kinds. Viruses are not active outside of their hosts due to their obligatory nature. Instead, they rely on their cells to penetrate and multiply. This is one of the reasons why many regard viruses as non-living. They infect a wide range of organisms, including fungi, bacteria, mammals, and plants, but they are quite particular to their hosts. Tobacco mosaic virus (TMV) was discovered in the nineteenth century by Martinus Beijerinck and his colleague Dmitrii Iwanowski as the first plant virus (Scholthof et al., 2008). A wide variety of fungal, bacterial, animal, and plant pathogenic viruses have been found since then.

According to recent estimates, over 4,000 viral species are capable of infecting over 1,000 different plant species. The primary rationale for studying plant pathogenic viruses is their negative influence on agricultural productivity. Historically, viruses were assumed to only impact people, animals, and plants.

According to Lapidot (2014), viruses are the primary cause of about half of the developing plant diseases that lower the yield of almost all crops and vegetables, including cucurbits. Vegetable plant diseases present a significant hindrance to the efficient production, use and export of vegetables. These diseases not only decrease the quantity of the crops, but they also deteriorate the quality of the products (Atanda et al., 2011).Vegetables are the highest-paying crops on the earth, including Pakistan. Different types of vegetables, such as edible roots, seeds, stems, or organic items, have different effects on an individual's eating pattern. Because of good soil and vegetable growing conditions, Pakistan produces massive volumes of veggies throughout the year. However, members of this family are consumed as vegetables in our daily lives and are referred to as revitalizing due to their easy access and require hardly any effort (Shrivastava et al., 2014).

Material and Method:

Squash leaf samples showing signs of leaf chlorosis, especially in intervascular areas, were randomly collected from farmers' fields in Lahore district by stratified sampling, stored on ice in plastic bags, and sent to the PCSIR Plant Molecular Pathology Laboratory. Sterilization along with distilled water was used to remove contaminants and thus effectively clean the samples. Samples were stored at 4°C for further processing. In addition, some parts of these samples were stored in desiccators containing calcium chloride and silica gel. In addition, some samples were stored at -80 °C. Each sample was appropriately labeled with the date of collection, place of collection, name of the collector, and details of symptoms. Cucumber mosaic virus (CMV) isolates were bio-purified by mechanical inoculation of CMV-infected leaf sap (homogenized in 0.1 M phosphate buffer) onto Cucurbitaceae leaves (5-6 leaf stage) pre-dusted with carbon dioxide for three successive transfer periods. Leaves kept in a greenhouse at 25°C and ~70% relative humidity developed local lesions and symptoms after 1-2 weeks. The given figure 1 showing symptoms from the collected samples.

Volume 3, Issue 5, 2025

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

Figure 1: Cucumber mosaic virus symptoms in collected samples

Molecular Characterization

Copies of the capsid protein gene from CMV virus isolates were synthesized using the previously described CMV-specific primers. Direct sequencing of the amplified andpurified CMV isolates by PCR made it possible to obtain the complete sequence of the capsid protein gene. The Korean company Macrogen sequenced the upper and lower strands using the popular Sanger dideoxy sequencing method. Sequences provided by the sequencing company were converted to amino acid sequences and both forms of sequences were matched to CMV isolates using BLAST, MEGA software and bioinformatics tools. In addition, after careful analysis, all sequences were submitted to GenBank.

RNA Extraction:

RNA content was isolated from 0.1 g collected plant leaf samples using the Sigma Tri- Reagent kit according to the protocol provided by the manufacturer. Leaf powder was prepared by grinding the leaves with a sterilized solution in liquid nitrogen and then poured into 2 ml DNase- and RNase-free microcentrifuge tubes. To extract RNA from the samples, 1 ml of Trizol reagent was added and mixed by shaking, followed by incubation at 25°C. Chloroform was added and the tubes were shaken and incubated, then centrifuged to separate the RNA in the upper aqueous phase. The RNA was precipitated with isopropanol, washed with ethanol, and air-dried. The RNA was suspended in water and heated to 60°C for 15 min. The isolated RNA was stored at -80°C and had a concentration of 500 ng/µL for both DNase and RNase. The quality of the RNA was verified by running it on a 1% agarose gel in TAE buffer (1X).

cDNA Synthesis and RCR:

CMV genomic RNA was synthesized from complementary DNA (isolated using the popular Trizole method) using a thermos-science countermeasure kit. Then, a DNA copy of the viral RNA was synthesized with a reverse transcriptase enzyme manufactured by Invitrogen using random hexamer reverse primers. The mixture containing 5 µl messenger RNA, 6 µl water, and 1 µl 20 pM primer was heated at 65°C for 5 min, and then quickly cooled by placing on ice. Then 1 µL, 4 µL, 1 µL and 1 µLof RNase inhibitor, 5X reaction buffer, 10 µM dNTP and RT enzyme were added,respectively. The mixture was heated to 42°C in 60 minutes and the reaction was complete at 70°C in 10 minutes. To use the cDNA in PCR at a later time, it was stored at -20°C. For CMV capsid gene amplification, a 50 µl PCR reaction mixture was prepared using 10x PCR buffer

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

(5 µl), cDNA template (4 µl), 25 mM MgCl2 (3 µl), and 10 µM dNTP (1 µl), nuclease-free water (35 µl), 1 µl each of 20 pM up and down primers and Taq polymerase; 500 units. PCR was conducted using the following temperature parameters: 5 min at 95°C, followed by 35 cycles of 94°C, 58°C, and 72°C for 60 seconds each, and finally 10 min at 72°C. Amplified PCR products were processed on a 1% agarose gel to check the results under UV light, and positive samples were stored at 4°C for extraction of the PCR product.

Gel Electrophoresis:

Gel electrophoresis was performed to separate the amplified PCR products. Buffer TAE (1X tris-acetic acid EDTA) was added to prepare a 1% agarose gel and then dyed with 100µg of intercalating dye with ethidium bromide per ml. In addition, ethidium bromide serves as a fluorescent dye and is therefore called an intercalation dye. The electrophoresis process was carried out in a buffer tank containing buffer TAE (1X tris-acetic acid EDTA) and/or TBE (1X tris-borate EDTA). Agarose gel was immersed in it. The PCR product (5 µl) was loaded into each well and mixed with 3 µl of dye for a 6-foldload. 4 µl of a 1 kb DNA ladder was added to the leftmost well manufactured by Thermo-science. A current of 120 V was applied to the gel reservoir to allow the DNA to move towards the negative electrode, giving the DNA enough time to migrate. A 1 kb DNA ladder was used to identify amplified fragments. Therefore, the separated PCR products were examined under UV light and then visualized (as shown in Figure 2). Three of the acquired sequences were deemed to be of high quality and were thus chosen for sequence analysis. There were 551 nucleotide readings in each isolate. BLASTn analysis identified these sequences as CMV CP genes found in the CMV genome

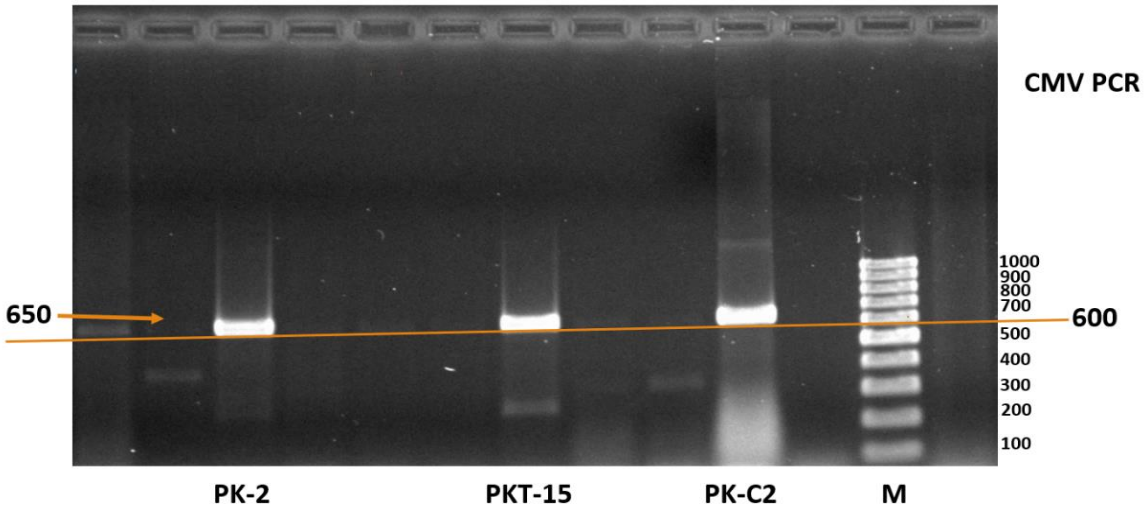

Figure 2: Gel bands visualization of isolates

Purification of PCR Products:

The amplified PCR products were purified using the Invitrogen PureLink® Purification Kit.

A 4X volume of binding buffer (B2) was used to dilute 50-100 µl of the PCR product and the mixture was poured onto a PureLink® spin column in a collection tube, then centrifuged at 5000 rpm for 60 seconds. The column was washed twice with 650 µl wash buffer and centrifuged at maximum speed. The columns were loaded with new DNase- and RNase-free tubes, and 50 µl of water was used to elute the isolated PCR products, which were then centrifuged at maximum speed for 2 minutes resulting in purified PCR product.

Sequence Analysis and Protein Translation:

The sequence was obtained from Macrogen Company, Korea, and CP gene-specificprimers were used to sequence each respective PCR product. The top and bottom thread sequences were obtained to monitor the accuracy of the sequence. Necessary corrections were made to the misread sequences by carefully observing the peaks of the ascending and descending strand sequence. In addition, weak

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

sequences were truncated at the 3' and 5' ends and clear sequences were extracted from each amplicon. The protein conformation of this sequence was verified using an online translation tool. The translated nucleotide and amino acid sequences were used for BLAST analysis, and isolates showing maximum and minimum sequence identity were downloaded from NCBI GenBank. The nucleotide sequences of the Pakistani isolates and sequences downloaded from GenBank were aligned using the ClustalX sequence alignment tool in MEGA 10 (Janssen et al., 2002). The Bio Edit software was used to calculate the nucleotide and amino acid sequence identities of poorly understood sequences, using the sequence identity matrix tool. The MEGA 10 program was used to perform phylogenetic analysis of aligned sequences. Phylogenetic trees were constructed using the neighbor-joining method, and 1000 bootstrap iterations in several thin-rooted trees were used to estimate group validity (Larkin et al., 2007). The RDP package was used to detect putative recombination in the nucleotide sequence of the CMV (Kumar et al., 2016). Chimaera, Bootscan, RDP, 3SEQ, Lard, MaxChi, Siscan, and GENECONV were used after Bonferroni's adjusted P-value threshold ($\alpha = 0.05$) and default parameters. The detection of recombination events was considered correct if it was confirmed by more than four of the above methods.

Phylogenetic Analysis

The 3 Pakistani CMV isolates were classified in cluster A and closely related to isolates from other countries, including KJ778898, KJ778899, MN340988, and MN330994, based on nucleotide sequence analysis of the CP gene. In contrast, based on amino acid sequence analysis, the Pakistani isolates formed a separate group from all other isolates, while the isolates from other countries formed a separate group. This information is based on a phylogenetic analysis of the CP gene sequences of 27 CMV isolates, including the 3 Pakistani isolates shown in Figure 3 and 4.

Table: 1: Nucleotide sequence properties of cp genes of Pakistani CMV isolates

Isolate	Characteristics					
	(Nt)	(AA)	A%	C%	G%	U%
CMVPK-C2	551	184	22.4%	25.1%	24.6%	27.9%
CMVPK-2	551	184	22.4%	24.6%	24.3%	28.7%
CMVPKT-15	551	184	22.7%	24.3%	24.1%	28.9%

The ExPasy translate tool was used to translate each isolate's sequence into 180 amino acids. In sequence analysis, nucleotide sequences from three Cucumber mosaic virus (CMV) isolates from Pakistan were employed. These sequences contained 22.4-22.7% Adenine, 24.3-25.1% Cytosine, 24.1-24.6% Guanine and 27.9-28.9% Uracil (Table 1). Nucleotide and amino acid sequences from Pakistani isolates were 96.5% to 98.3% comparable to each other, according to sequence comparison. While they shared 96.5% to 97.6% nucleotide and 88.5% to 98.3% amino acid sequence similarities with other CMV isolates in NCBI GenBank shown in table 2. MN340988 and MN330994 isolates showed higher percentages of nucleotide and amino acid identities, while the KJ778898 and KJ778899 isolates had lower nucleotide and amino acid identity percentages. Multiple sequence alignment of the amino acid sequences of all Pakistani isolates did not reveal any unique changes in any single location of the CP gene.

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

Table 2: List of CMV isolates downloaded from NCBI-GenBank

Acc. No.	Isolate	Country	Year	Host
MN340997	Ho	India	2019	*Capsicum annuum*
KJ778899	TN TNV SG1	India	2014	Snake gourd
KJ778898	TN TNAU SG1	India	2014	Snake gourd
MN340994	GU2	India	2019	*Capsicum annuum*
MN340988	GU1	India	2019	*Capsicum annuum*
MN340991	BA	India	2019	*Capsicum annuum*
HE962479	Vir	Italy	2012	*Capsicum sp.*
KU947030	ICAR-IISR PN25	India	2016	Black pepper cultivar Panniyur 1
KM272278	KO	India	2014	*Capsicum annuum*
MT395347	TNCR2-4	India	2020	Chilli
MT422733	TNCR2-3	India	2020	Chilli
MT422732	TNCR2-2	India	2020	Chilli
MT422731	TNCR2-1	India	2020	Chilli
AF150731	MB-CMV	Germany	1999	Cucumber
GU111228	New Delhi	India	2009	Tomato
MN495618	ATCMV2	India	2019	Chilli
KJ778900	TN NGK BoG 1	India	2014	Bottle gourd
KC019300	HB24	China	2012	Cucumber
KT004543	WN1	China	2015	Piper nigrum L (black pepper)
KC527698	RP14	South Korea	2013	Capsicum annuum L. (pepper)
MT661451	RB1	South Korea	2020	Capsicum annuum L.
MN422337	GTN	South Korea	2019	Capsicum annuum
KP033525	GTN	South Korea	2014	Pepper
MN422334	P1	South Korea	2019	Capsicum annuum

Phylogenetic tree based on Amino acid and nucleic acid:

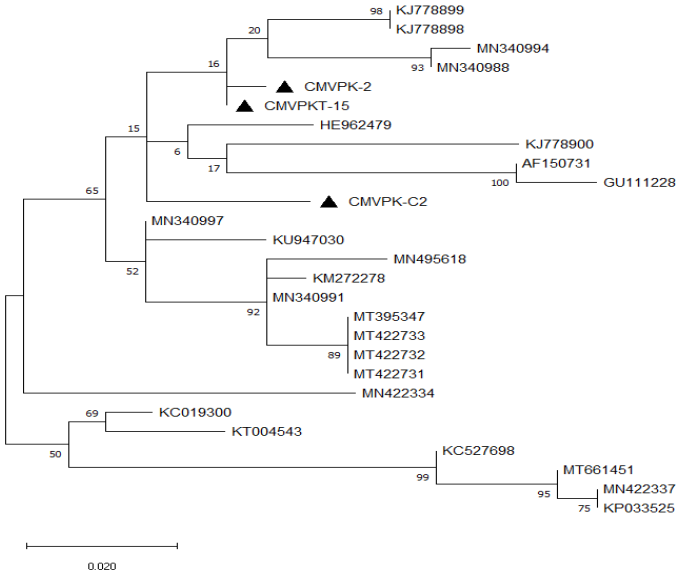

Figure 3: Phylogenetic tree of Cucumber mosaic virus (CMV) isolates based on Amino Acid

Volume 3, Issue 5, 2025

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

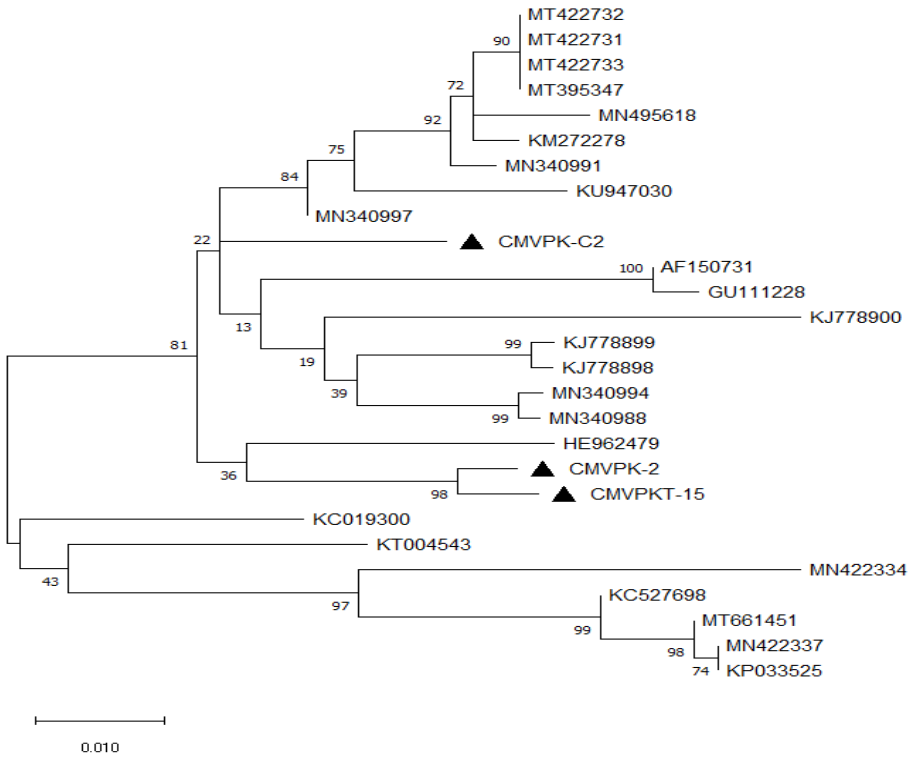

Figure 4: Phylogenetic tree of Cucumber mosaic virus isolates based on Nucleic Acid

Discussion:

Vegetables are one of the high-income-generating crops in Pakistan, providing income prospects for small and progressive farmers. Bitter melons, zucchini, cucumbers, squash, watermelons, pumpkins, cantaloupes, hairy melons, and gourds are the most widely farmed vegetables in Pakistan during the summer season. Aside from insect attacks and fungal or bacterial infections, viral diseases are tricky to handle because no physical or chemical means are sufficient. Proper and precise identification and molecular sequence data of these viral infections are critical for this aim. The molecular characterization of CMV infecting cucurbits was undertaken in this work. In Pakistan record of CMV infecting cucurbits is not too old (Ashfaq et al., 2014) but it has not have been vastly studied. The molecular research of plant viruses is not new; much has been done in the case of many plant viruses around the world. This virus was initially detected in the current investigation as causing severe interveinal chlorosis and leaf curling. CMV has also been associated with symptoms such as leaf deformation, reduced size, severe mosaic, and blistering. Based on CP gene nucleotide sequences, the molecular diversity of 29 CMV isolates, comprising 25 previously reported isolates from various areas of the world and four newly assimilated strains from Pakistan, was determined in this study. Detection, identification, and characterization of plant viruses at the molecular level using the CP gene is a standard approach used to investigate various plant viruses around the world (Tsompana et al., 2005) (Ahmad et al., 2022). Maximum identities of Pakistani isolates were discovered in this investigation with Chinese isolates. CMV isolates clustered with Chines isolates in phylogenetic analyses. These findings indicate that Pakistani isolates were most likely introduced to Pakistan by diseased seedlings of cucurbits or other Cucurbitaceae crops from adjacent countries. Pakistan imports a significant number of seeds, seed products, and plant materials

Frontier in Medical & Health Research

ISSN: (e) 3007-1607 (p) 3007-1593

from neighboring countries. Because local seed and planting material production in Pakistan is insufficient to meet our needs. Furthermore, the study done can serve as a starting point for future efforts to sort out genetic alterations and comprehend the genetic architecture of Pakistani isolates.

Conclusion:

The outcomes of this study demonstrated that Cucumber mosaic virus infects cucurbit crops in the Lahore district of Pakistan, as it does in other nations throughout the world. CMV causes characteristic symptoms such as leaf deformation, reduced size, severe mosaic, and blistering, as detected during sample collection field surveys. The acquired nucleotide sequences from collected samples were recognized as the CP gene of Cucumber mosaic virus (CMV) using NCBI BLASTn analysis. A geographical association was identified in phylogenetic tree reconstruction and nucleotide identity percentages, with Pakistani isolates clustering and showing higher identities with Indian isolates. The findings of this study clearly suggest that a complete assessment of diverse vegetable and other field crops in Pakistan for viruses is required to understand the pattern of genetic diversity and the occurrence of new strains or species of CMV. This will aid in the creation of long-term strategies for managing CMV and other infections. Unfortunately, information regarding plant viruses that infect crops and vegetables is scarce, and management strategies such as vegetable breeding and the development of GMOs to combat these unknown adversaries are inadequate. One example is the lack of or inaccessibility to pathogen- related molecular data.

REFERENCES:

Atanda, S. A., Pessu, P. O., Agoda, S., Isong, I. U., & Ikotun, I. (2011). The concepts and problems of post–harvest food losses in perishable crops. African Journal of Food Science, 5(11), 603-613.

Ashfaq, M., Iqbal, S., Mukhtar, T., & Shah, H. (2014). Screening for resistance to cucumber mosaic cucumovirus in chilli pepper. Journal of Animal and Plant Sciences, 24(3), 791-795.

Ahmad, A., Banat, F., Alsafar, H., & Hasan, S. W. (2022). Algae biotechnology for industrial wastewater treatment, bioenergy production, and high-value bioproducts. Science of The Total Environment, 806, 150585.

Janssen, D., Ruiz, L., Velasco, L., Segundo, E., & Cuadrado, I. M. (2002). Non-cucurbitaceous weed species shown to be natural hosts of Cucumber vein yellowing virus in south-eastern Spain. Plant Pathology, 51(6), 797-797.

Kumar, S., Chauhan, P. S., Agrawal, L., Raj, R., Srivastava, A., Gupta, S., ... &Nautiyal, C. S. (2016). Paenibacilluslentimorbus inoculation enhances tobacco growth and extenuates the virulence of Cucumber mosaic virus . PLoS One, 11(3), e0149980.

Scholthof, K. B. G. (2008). Tobacco mosaic virus: the beginning of plant virology. APSnet Features.

Shrivastava, D., Singhai, A. K., & Yadav, R. K. (2014). Effect of lime and rice husk ash on engineering properties of black cotton soil. Int. J. Eng. Res. & Sci. & Technology, 3(2), 292-296.

Tsompana, M., Abad, J., Purugganan, M., & Moyer, J. W. (2005). The molecular population genetics of the Tomato spotted wilt virus (TSWV) genome. Molecular Ecology, 14(1), 53-66.