Next-generation sequencing-based detection reveals Erysiphe necator-associated virus 1 in okra plants

Mariam hashem Al-helu[1]*, Zhongtian Xu[2], Jun-Min Li[2], Adnan A. Lahuf[1]

[1]Plant Protection Department, Agriculture College, University of Kerbala, Karbala, Iraq
[2] Institute of Plant Virology, Ningbo University, Ningbo 315211, China
***Corresponding author e-mail: mariam.h@s.uokerbala.edu.iq**
https://doi.org/10.59658/jkas.v11i1.1450

Received:
Feb. 17, 2024

Accepted:
Mar. 12, 2024

Published:
Mar. 18, 2024

Abstract

Erysiphe necator is a devastating fungal phytopathogen that has caused powdery mildew on several genera within the Vitaceae family worldwide. In the current investigation, next-generation sequencing of RNA extracted from the okra crop, besides the bioinformatics analysis, was operated to examine the virome associated with the okra crop. Although *E. necator* fungus was not identified in the okra tissues, the mycovirus-related sequence of Erysiphe necator-associated virus 1 was detected. This mycovirus sequence showed a high similarity ranging between 80.38-93.48% and 100% coverage with the four distinct global isolates of the Erysiphe necator-associated virus 1. This genomic data postulates comprehension of the existence of the mycoviruses in plant hosts without their primary host, the fungus. To our knowledge, this is the first identification of Erysiphe necator-associated virus 1 in Iraq.

Keywords: Mycoviruses, *Erysiphe necator*, Next-generation sequencing.

Introduction

Fungal viruses, also known as mycoviruses, are viruses that may infect fungi or replicate inside fungal cells. Since the first recorded identification of a mycovirus in 1962 [1], various mycoviruses have been detected across various fungal species, including some important in plant pathology [2]. Mycoviruses typically consist of double-stranded RNA (dsRNA). However, some may have positive single-stranded RNA (+ssRNA), negative single-stranded RNA (-ssRNA), or circular single-stranded DNA genomes [3]. Mycoviruses are classified into 13 families, with double-stranded RNA-containing viruses falling into five families and single-stranded RNA-containing viruses falling into six families. Viruses with RNA reverse-transcribing genomes may be categorized into two distinct families, while viruses having single-stranded circular genomes are classified under the Genomoviridae family [4]. Several mycovirus-associated double-stranded RNAs remain unclassified within any known family. The examination of dsRNA obtained from fungal tissue is a prevalent technique employed to identify and describe mycoviruses. This is because single-stranded RNA viruses generate dsRNA replicative intermediates as part of their life cycle, but fungal tissue

usually lacks large dsRNA molecules [5].

Erysiphe necator Schw. (formerly known as Unicinula necator (Schw.) Burr.) is a notable plant pathogen that causes powdery mildew disease in plants of the Vitaceae family, such as grapevines (genus Vitis) [6]. The existence of this may lead to significant reductions in grape production and quality. Specifically, the *V. vinifera* species is highly susceptible to this fungus. Mycoviruses that cause hypovirulence in *E. necator* may provide viable options for reducing dependence on chemical fungicides [7]. Viruses have been detected by the utilization of enzyme-linked immunosorbent assays (ELISA) and/or reverse-transcription polymerase chain reaction (RT-PCR). RT-PCR has greater sensitivity compared to ELISA, which increases the probability of underestimating the viral load in wild plants when primarily depending on ELISA [8]. Nevertheless, the efficacy of RT-PCR is limited by the prerequisite of prior acquaintance with viral sequences [9]. The absence of established techniques further complicates the detection of viruses in plants using RT-PCR and ELISA. In conjunction with bioinformatics, next-generation sequencing (NGS) methods enable concurrently identification of both recognized and unrecognized viruses and mycoviruses in crops. This methodology has been effectively utilized in diverse research endeavors and has the potential for identifying mycoviral sequences in plant specimens[10, 11, 12].

This study aimed to investigate okra viral infection in Iraq. Advanced genetic sequencing technology, namely high-throughput sequencing (HTS) (sometimes referred to as next-generation sequencing (NGS)), in conjunction with bioinformatic analysis, was applied to discover virome associated with okra crop.

Materials and Methods

Plant samples

A survey was conducted during the 2022-2023 season at different okra fields in Baghdad province, Iraq. The goal was to collect samples of okra plants displaying signs of viral infection. Thirty afflicted okra plants were chosen at random for harvest. Randomly collected leaf samples were obtained from many okra plants in each field. The samples were conserved and sent in RNAprotect Tissue Reagent (Qiagen, Germany) for further analysis using next-generation sequencing (Total RNA-seq) [13,14].

Virus Sequencing, Assembly, and Annotation

The RNeasy Plant Mini Kit (Qiagen, Hilden, North Rhine-Westphalia, Germany) isolated total RNA from 100 mg of symptomatic leaf tissue. After extracting RNA, ribosomal RNA was removed using the Ribo-Zero Plant kit. Following that, cDNA libraries were produced with Illumina's TruSeq Stranded whole RNA kit, and sequencing was carried out on an Illumina HiSeq 4000 platform, producing pair-ended reads (2 x 150 bp) at Macrogen Inc. (Teheran-ro, Gangnam-gu, Seoul, Republic of Korea). The raw readings were subjected to preprocessing using BBDuk adapter/quality trimming tool version 38.84 to remove adapters and low-quality reads. Those sequences aligned with the okra genome (*Abelmoschus esculentus*; GCA_0350

48815.1). The unmapped reads were built using the Velvet assembler version 1.2.10 in Geneious Prime® 2023.2.1 [15]. The contigs underwent BLAST searches against an adapted viral sequences database (obtained from NCBI in January 2024) and the whole NCBI non-redundant protein sequences (nr) database. All viral genomes, including their associated terminal sequences, were assembled using Geneious Prime 2023.0.1 [16]. Additionally, the unmapped reads were aligned in the reference-based assembly method with the 11 chromosome sequences of *Erysiphe necator* (CM045619.1 to CM045629.1) operating Geneious Prime 2023.0.1.

Phylogenetic Analysis

Phylogenetic studies employed RNA-dependent RNA polymerase (RdRp) sequences from recently discovered and previously recorded mycoviruses from different genera obtained from the NCBI GenBank. The MUSCLE method was used to perform alignments of nucleotide and amino acid sequences. The phylogenetic connections were determined using the neighbor-joining (NJ) technique in MEGA X, using 1,000 bootstrap replications [17].

Results and Discussion

High throughput sequencing and detection of mycovirus-related sequences

The sequence datasets obtained included 25,461,526 and 28,781,862 high-quality short-read sequences after trimming. The unmapped sequences were subjected to constructing contigs using de novo assembly, producing 55,704 contigs. Although there were no contigs similar to the chromosomes of *E. necator,* the BLASTn and BLASTx investigations identified contigs that closely resembled sequences of Erysiphe necator virus species (EnVSs). The BLASTn analysis of the Iraqi sequence displayed high identity ranging between 80.38 to 93.48% with sequences of EnVSs species 1,3,4 and 5 (Table 1). However, it was most similar to Erysiphe necator-associated virus 1 species, with percentage identity and query cover reaching 93.48% and 100%, respectively.

Table (1): Sequence and similarity analysis of the Iraqi Erysiphe necator associated virus 1 isolate Iraq-1 sequence with the other related EnVSs sequences

Description	Query Cover	Percentage identification	Accession Number
Erysiphe necator associated virus 1 isolate PMS13_DN222, complete genome	100%	93.48%	MN611681.1
Erysiphe necator associated virus 4 isolate PMS4_DN1011, complete genome	100%	84.13%	MN611684.1
Erysiphe necator associated virus 3 isolate PMS2_97, complete genome	100%	81.53%	MN611683.1

Erysiphe necator associated virus 5 isolate PMS5_DN3225, complete genome	100%	80.38%	MN611685.1
Riboviria sp. isolate H2_Bulk_36_scaffold_6157 sequence	100%	79.88%	MN033318.1
Riboviria sp. isolate H1_Rhizo_26_FD_scaffold_24795 sequence	93%	79.37%	MN032694.1

Additionally, the Iraqi Erysiphe necator-associated virus was further examined for open reading frames (ORFs), revealing a single extensive ORF that may encode proteins consisting of 209 amino acids (aa). Comparative searches of the NCBI database using the BLASTx program revealed that the proteins produced by this large open reading frame (ORFs) encode the RNA-dependent RNA polymerase (RdRp), showed sequence similarities ranging from 66.51% to 99.04% with RdRp of RNA viruses belonging to the five species of EnVSs (EnVS 1, EnVS 2, EnVS 3 EnVS 4, and EnVS 5; Table 2). Therefore, the obtained sequence was deposited into the GenBank database under accession number PP352232.1.

Table (2): Similarity of the RdRp protein sequence of the Iraqi Erysiphe necator-associated virus compared with other identical protein sequences

Description	Query Cover	Percentage identification	Accession Number
RNA-dependent RNA polymerase [Erysiphe necator associated virus 1]	100%	100.00%	QKI79983.1
RNA-dependent RNA polymerase [Erysiphe necator associated virus 5]	100%	84.91%	QKI79987.1
RNA-dependent RNA polymerase [Erysiphe necator associated virus 4]	100%	84.91%	QKI79986.1
RNA-dependent RNA polymerase [Erysiphe necator associated virus 3]	100%	84.77%	QKI79985.1
RNA-dependent RNA polymerase [Erysiphe necator associated virus 2]	99%	60.04%	QKI79984.1
RNA-dependent RNA polymerase [Riboviria sp.]	90%	55.85%	QDH86582.1

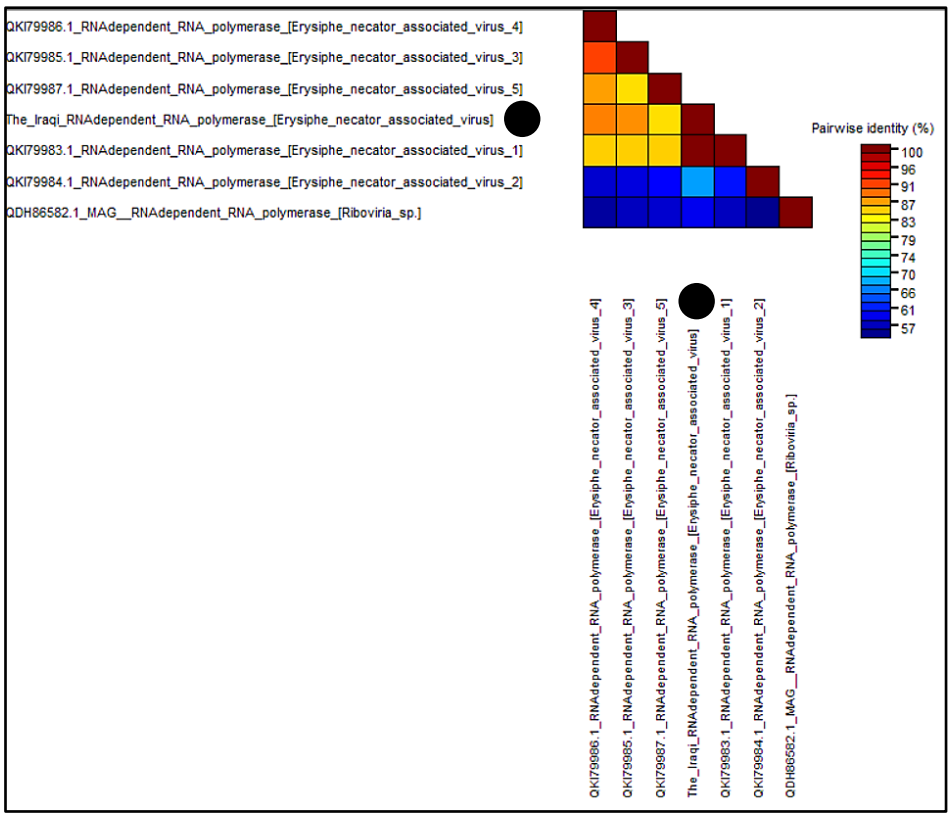

Figure (1): The similarity percentage among the RdRp protein sequence of the Iraqi EnVS (indicated with a black dot) compared with other identical protein sequences of related viruses. This figure was built using the Sequence Demarcation Tool version 1.2.

Phylogenetic analysis

The amino acid sequence of the putative RdRp from the virus linked to Iraqi EnVS was compared to analogous sequences from five additional Erysiphe necator-associated viruses discovered in GenBank. A phylogenetic tree was created using the NJ technique based on this multiple-sequence alignment. Figure 2 demonstrates the strong relationship between the Iraqi EnVS and Erysiphe necator-associated virus 1 (QKI79986.1). Within that group, they also clustered with other species of Erysiphe necator-associated virus 5, 3, and 4 (QKI79987.1, QKI79985.1, and QKI79986.1 respectively). This outcome strongly suggests that the Iraqi mycovirus belongs to the species Erysiphe necator-associated virus 1.

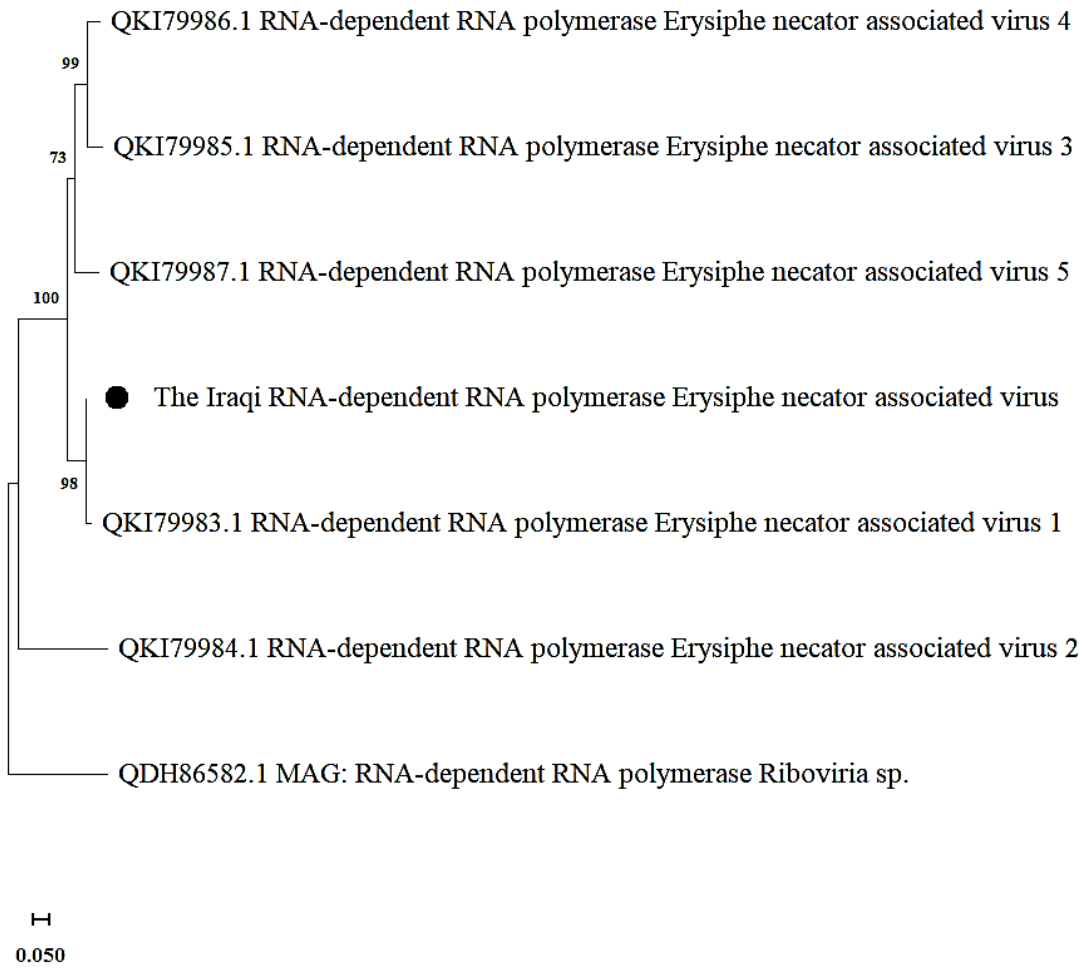

Figure (2): Phylogenetic analysis of the Iraqi EnVS (indicated with a black dot). A phylogenetic tree was built based on the RdRp amino acid sequences of Erysiphe necator viruses.

In the last 15 years, high-throughput sequencing (HTS) has been applied successfully to detect viruses in many hosts, including filamentous fungi [18, 19, 20]. In the current study, HTS was combined with bioinformatics analysis to investigate the occurrence of virome sequences associated with okra crops. This method enabled a comprehensive analysis of mycoviral sequences linked to the okra plant. Although previous reports discovered the existence of dsRNA and virus-like particles within the *E. necator* fungus [21, 22], comprehensive analyses regarding the mycoviruses within fungi are very limited, particularly in Iraq. However, the current investigation showed the presence of uncharacterized mycovirus-related sequences, assigned as Erysiphe necator virus sequences 1. The specific terminal nucleotides of these sequences were not determined. However, based on the size and characteristics of the open reading frame (ORF) that encodes RdRp sequences, it may be deduced that

these sequences are almost entirely intact. Further investigation will be required to confirm the accurate end nucleotides of these sequences.

E. necator was not identified in the current investigation as a fungus associated with the okra crop. However, EnVS 1 was detected independently of its fungal host, presenting a potential discrepancy in the results. Nevertheless, a study [23] has demonstrated the capacity of mycoviruses to reproduce within plant cells, hence carrying significant evolutionary consequences. The study provides experimental data demonstrating that some mycoviruses, particularly those classified under the Partitiviridae and Totiviridae families, can reproduce within plant cells without altering their nucleotide sequence or indications of adapting to the host. The study also noted that the Partitivirus PaPV1, which could only replicate in specific plant cells, significantly increased RNA accumulation in BY2 cells, indicating host specificity.

This study demonstrates the presence of mycoviruses in okra plants, even in the absence of its typical fungus host, *E. necator*. This finding improves our understanding of the range of mycoviruses present in this economically significant host. The acquired insights might play a crucial role in predicting viral disease outbreaks in different crops and the transmission of viruses and their possible alternate hosts, including fungi. Additional research is required to clarify the connection between the mycovirus discovered in this study, the E. necator fungus, and the okra crop.

References

1) Hollings, M. (1962). Viruses associated with a dieback disease of cultivated mushrooms. *Nature*, 196, 962–965.

2) Ghabrial, S. A., & Suzuki, N. (2009). Viruses of plant pathogenic fungi. *Annual Review of Phytopathology*, 47, 353–384.

3) Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., & Suzuki, N. (2015). 50-plus years of fungal viruses. *Virology*, 479–480, 356–368.

4) King, A. M. Q., Adams, M. J., Carstens, E. B., & Lefkowitz, E. J. (2012). Virus taxonomy: Classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses (Vol. 9). Elsevier Academic Press, London, pp. 23–1208.

5) Zabalgogeazcoa, I., Benito, E. P., Ciudad, A. G., Criado, B. G., & Eslava, A. P. (1998). Double-stranded RNA and virus-like particles in the grass endophyte *Epichloe festucae*. *Mycology Research*, 102, 914–918.

6) Braun, U., & Cook, R. T. A. (2012). Taxonomic Manual of the Erysiphales (Powdery Mildews). CBS Biodiversity Series, 11, 1–707.

7) Qiu, W., Feechan, A., & Dry, I. (2015). Current understanding of grapevine defense mechanisms against the biotrophic fungus (*Erysiphe necator*), the causal agent of powdery mildew disease. *Horticulture Research*, 2, 15020.

8) Prendeville, H. R., Ye, X., Morris, T. J., & Pilson, D. (2012). Virus infections in wild plant populations are both frequent and often unapparent. *American Journal of Botany*, 99, 1033-1042.

9) Lacroix, C., Renner, K., Cole, E., Seabloom, E. W., Borer, E. T., & Malmstrom, C. M. (2016). Methodological guidelines for accurate detection of viruses in wild plant species. *Applied and Environmental Microbiology*, 82, 1966-1975.

10) Susi, H., Filloux, D., Frilander, M. J., Roumagnac, P., & Laine, A.-L. (2019). Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ*, 7, e6140.

11) Abass, M. O., & Lahuf, A. A. (2022). First report of the satellite DNA Beta associated with Tomato Yellow Leaf Curl Virus–Mild on Tomato in Iraq. *Plant Health Progress*, 23(4), 480-481.

12) Salman, M. D., & Lahuf, A. A. (2022). A new procedure to identify plant RNA viruses associated with the whitefly (*Bemisia tabaci*) using next-generation sequencing. *Arab Journal of Plant Protection*, 40(2), 169-174.

13) Mohammed, M. S., & Lahuf, A. A. (2023). First report of Cucumber vein yellowing virus in Iraq. *New Disease Report*, 47(2), 1-2.

14) Ye, Z.-X, Lahuf, A. A., Salman, M. D., Zhang, Y., & Li, J.-M. (2022). Draft genome sequence of a novel Iflavirus from leafhoppers (*Exitianus capicola*) in Iraq. *Microbiology Resource Announcement*, 11(5), 1-2.

15) Abass, M., & Lahuf, A. (2023). First report of tomato leaf curl Palampur virus infecting tomato in Iraq. *Journal of Plant Pathology*, 105, 385.

16) Abass, M. O., & Lahuf, A. (2023). High-throughput sequencing and bioinformatic analysis reveal presence of the endogenous pararetrovirus Tobacco vein clearing virus genome in the tomato (*Solanum lycopersicum*) host genome. *Arab Journal of Plant Protection*, 41(1), 77-84.

17) Al Rwahnih, M., Daubert, S., Urbez-Torres, J. R., Cordero, F., & Rowhani, A. (2011). Deep sequencing evidence from single grapevine plants reveals a virome dominated by mycoviruses. *Archives of Virology*, 156, 397–403.

18) Hadidi, A., Flores, R., Candresse, T., & Barba, M. (2016). Next-generation sequencing and genome editing in plant virology. *Frontiers in Microbiology*, 7, 1325.

19) Hadidi, A., Flores, R., Candresse, T., & Barba, M. (2016). Next-generation sequencing and genome editing in plant virology. *Frontiers in Microbiology*, 7, 1325.

20) Kondo, H., Hisano, S., Chiba, S., Maruyama, K., Andika, I. B., Toyoda, K., Fujimori, F., & Suzuki, N. (2016). Sequence and phylogenetic analyses of novel totivirus-like double-stranded RNAs from field-collected powdery mildew fungi. *Virus Research*, 213, 353–364.

21) Azzam, O. I., & Gonsalves, D. (1991). Detection of dsRNA from cleistothecia and conidia of the grape powdery mildew pathogen *Uncinula necator*. *Plant Disease*, 75, 964–967.

22) Pandey, B., Naidu, R. A., & Grove, G. G. (2018). Detection and analysis of mycovirus-related RNA viruses from grape powdery mildew fungus *Erysiphe necator*. *Archives of Virology*, 163(4), 1019–1030.

23) Nerva, L., Varese, G. C., Falk, B. W., & Turina, M. (2017). Mycoviruses of an endophytic fungus can replicate in plant cells: evolutionary implications. *Scientific Reports*, 7(1), 1908.