RESEARCH ARTICLE

How to Cite: Pozharskiy, A., Kostyukova, V., Gritsenko, D. (2026). First report and complete genome sequence of zucchini yellow mosaic virus from cucumber in Kazakhstan. Contig, 1(3), 202617. DOI: 10.66273/3134-6359.2026.1.3.005

UDC: 578.864

Received: 05/09/2026 | Accepted: 05/10/2026 | Published online: 06/10/2026

First report and complete genome sequence of zucchini yellow mosaic virus from cucumber in Kazakhstan

Alexandr Pozharskiy[1], Valeriya Kostyukova[1], Dilyara Gritsenko[1,2,3,*]

[1] Laboratory of Molecular Biology, Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan

[2] Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan

[3] Research Center AgriBioTech, Almaty 050040, Kazakhstan

*** Corresponding author: d.gritsenko@ipbb.kz | ORCID: 0000-0001-6377-3711**

Abstract

Zucchini yellow mosaic virus (ZYMV) was detected for the first time in Kazakhstan in cucumber (*Cucumis sativus* L.). RNA sequencing using the Oxford Nanopore platform was performed for virus identification and genomic characterization. A complete genome sequence of isolate Pcu-2026KZ, 9580 nt in length, including a complete 9242-nt polyprotein ORF, was obtained. BLAST analysis showed 98.5% nucleotide identity with the French isolate C05-205. Phylogenetic analysis of 124 complete ZYMV genomes revealed five major clusters. Isolate Pcu-2026KZ was placed within a cluster predominantly comprising Asian isolates and formed a strongly supported branch with the French isolate C05-205. Its phylogenetic position may indicate a relationship with European and Asian ZYMV lineages; however, the origin and introduction pathway of the virus into Kazakhstan require further investigation. The potential spread of ZYMV among Cucurbitaceae crops is also of interest, as the same cluster predominantly included isolates detected in *Cucurbita* spp. These findings provide the first complete genome characterization of ZYMV from Kazakhstan and highlight the need for further surveillance of the virus in cultivated crops.

Keywords: ZYMV; *Potyvirus*; *Cucumis sativus*; RNA-Seq

Introduction

Cucumber (*Cucumis sativus* L.) is one of the widely cultivated vegetable crops; however, its productivity is substantially affected by viral diseases. Among the viruses infecting cucurbit crops, Zucchini yellow mosaic virus (ZYMV; genus *Potyvirus*, family *Potyviridae*) is a widespread viral pathogen [1]. ZYMV infection is characterized by the development of mosaic symptoms, chlorosis, and leaf deformation, as well as impaired plant growth and reduced fruit quality [2]. The virus is efficiently transmitted mainly by aphids in a non-persistent manner, facilitating its rapid spread in agricultural ecosystems [3,4]; the virus also can be transmitted with seeds [5] and via mechanical contact [6], making the spread of infection more difficult to control.

ZYMV exhibits substantial genetic diversity among isolates from different geographical regions and host plants. Comparative studies have shown that the degree of genetic similarity among isolates can vary

considerably, while phylogenetic analyses can reveal geographically distinct viral groups. For example, comparison of the Iranian ZYMV-Kurdistan isolate with previously reported complete genome sequences revealed 79.6–98.8% nucleotide sequence similarity, with the isolate occupying a distinct position among genetically differentiated ZYMV groups [7]. Studies of isolates from Australia, Asia, and Oceania have also identified recombination events, indicating a complex pattern of genetic variation within the virus [8,9,10]. These findings indicate that characterization of ZYMV based solely on individual genomic regions may not fully capture the genetic diversity of the virus. Analysis of complete genome sequences enables the simultaneous assessment of genetic divergence, phylogenetic relationships, and potential recombination events, which is particularly important for studying isolates from regions with limited genomic data.

Despite the widespread distribution of ZYMV in different regions of the world, information on its circulation and genetic characteristics in Kazakhstan remains limited. In particular, no published data are currently available on the complete genome sequence of a ZYMV isolate detected in cucumber plants in Kazakhstan. Obtaining and analyzing such a sequence will allow the genetic relationship of the Kazakh isolate to previously characterized viral populations to be determined and will contribute to a better understanding of ZYMV genetic diversity and geographical distribution.

The present study reports ZYMV in cucumber plants for the first time in Kazakhstan and provides genetic characteristics of the obtained isolate based on its complete viral genome sequence, including assessment of its phylogenetic position and genetic similarity to previously characterized ZYMV isolates.

Materials and Methods

A cucumber plant (*Cucumis sativus* L.) showing pronounced symptoms of viral infection was sampled from a private garden in spring 2025 and transported to the laboratory for further molecular genetic analysis. Total RNA was extracted using the Monarch Total RNA Miniprep Kit (New England Biolabs, USA) according to the manufacturer's instructions. The quality of the extracted RNA was assessed by electrophoresis on a 1.5% agarose gel, while RNA concentration was determined using a Qubit 4 fluorometer (Thermo Fisher Scientific, USA).

Direct RNA sequencing libraries were prepared using the Oxford Nanopore Technologies (ONT) Direct RNA Sequencing Kit (SQK-RNA004) according to the manufacturer's protocol. Sequencing was performed on a MinION platform (Oxford Nanopore Technologies) using MinKNOW software v24.11.10. Basecalling was performed using Dorado v7.6.8 integrated into MinKNOW.

The obtained RNA reads were converted to DNA sequence and filtered to include only the sequences of length ≥1000 bp using the SeqKit tool [11] and further mapped against the reference dataset of all plant viruses available in the NCBI Virus database [12] using BWA v0.7.17-r1188 [13]. The mapping statistics were collected and the consensus sequence of the ZYMV genome was constructed by SAMtools v1.19.2 [14]. The identity of the obtained sequence was confirmed by NCBI BLAST search (https://blast.ncbi.nlm.nih.gov/). The obtained genome sequence was annotated manually based on the reference sequence for ZYMV (NC_003224.1).

For the phylogenetic analysis, 124 complete genomic sequences of ZYMV were retrieved from NCBI Virus, as well as ten reference sequences of *Potyvirus* species isolated from plants of the family Cucurbitaceae (an outgroup). The maximum likelihood tree based on the GTR G+I model was calculated using MEGA v12.1.2 [15] with adaptive bootstrap (error threshold 0.05).

Results and Discussion

During a survey of cucumber plants (*Cucumis sativus* L.) in a private garden in spring 2025, one sample exhibiting pronounced symptoms characteristic of viral infection on fruits was collected. The symptoms consisted of pale yellow spots (chlorosis) on the fruit skin and the moderate fruit wilting (Figure 1, B, D). The leaves displayed less pronounced chlorotic spots (Figure 1, A, C). To confirm the viral etiology of the disease and identify the causative agent, total RNA was extracted from the sample, followed by high-throughput sequencing using the Oxford Nanopore platform. The resulting sequences were used to identify viral components and subsequently assemble the viral genome.

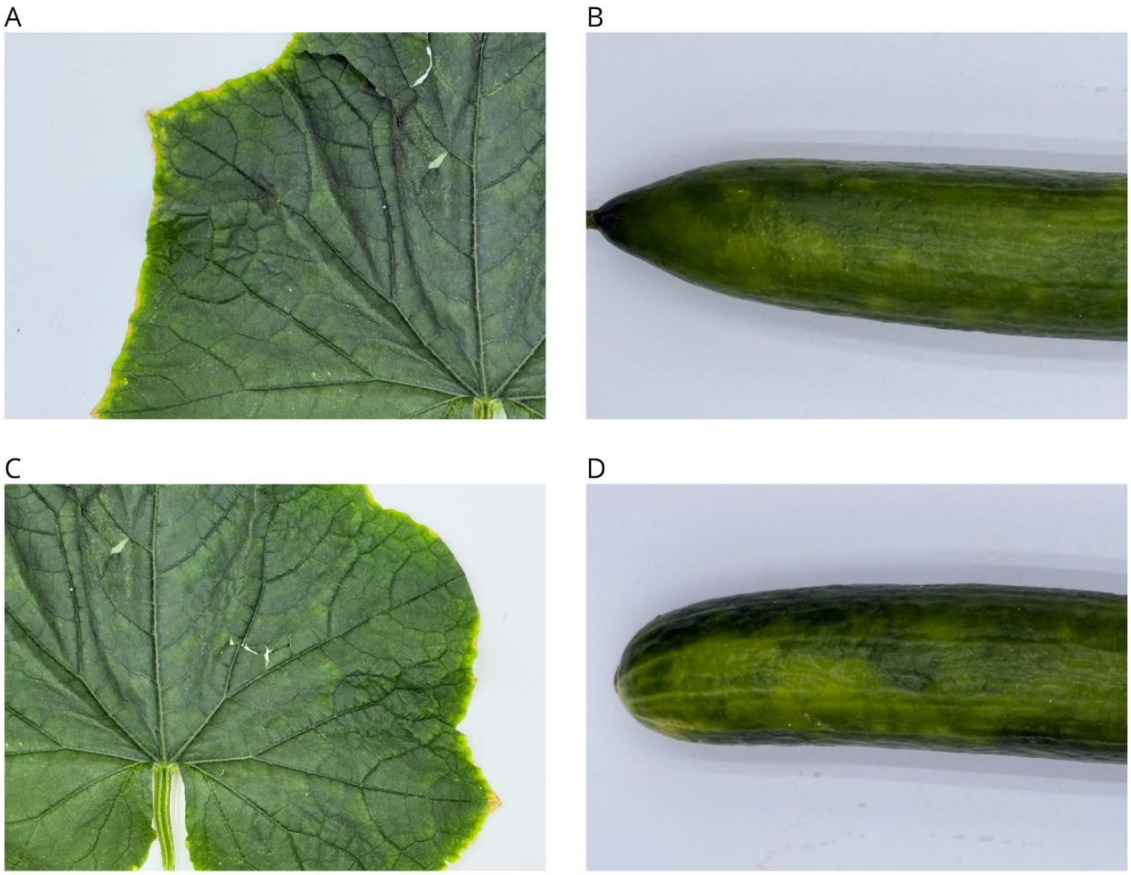


Figure 1. Symptoms of viral infection in cucumber (*Cucumis sativus* L.). Mosaic symptoms and chlorosis of the leaf blades (A, C) and uneven mosaic coloration of the fruits (B, D) observed in the examined sample.

As a result of sequencing, 9971 RNA reads of length above 1000 bp covering a total of 13,218,556 bases were obtained, of which 815 were matched against viral genomes. Of all sequences, 638 corresponded to ZYMV and covered 99.854% of the corresponding reference genome with the mean coverage depth 117.12 and average mapping quality 59.9; the average base quality was 17.2, which corresponds to 98.09% sequencing accuracy. The resulting genome sequence obtained as a consensus of all mapped reads had a total length of 9580 bases and contained a complete polyprotein ORF of length 9242 bases without errors, as well as all predicted protein domains of the ZYMV polyprotein (Figure 2). Taking into account completeness of the coding part of the genome and only a small length difference with the reference genome (9591 bases), the sequence was considered the complete genome of the found isolate (Pcu-2026KZ).

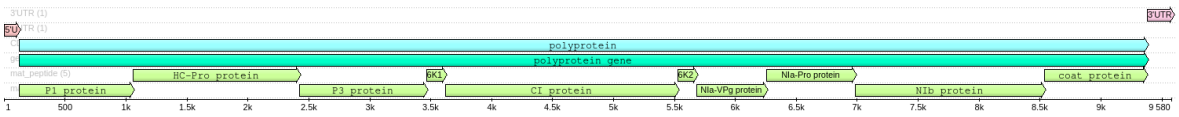


Figure 2. Annotated genome structure of ZYMV isolate Pcu-2026KZ: 5′- and 3′-UTR (pink), ORF (blue-green), polyprotein (cyan), individual protein domains (salad green).

The BLAST search confirmed the species of the isolate as ZYMV; the closest match corresponded to the ZYMV isolate C05-205 from France (accession MW449263.1) with the identity rate 98.5%.

Phylogenetic analysis involving ZYMV sequences from around the world revealed a structure with five well-defined clusters shown in different colours in Figure 3, A. The cluster containing isolate Pcu-2026KZ (shown blue, Figure 3, B) consisted generally of isolates from Asia (China, Korea) which were obtained from *Cucurbita* sp., except for several isolates from *Luffa* sp., *Sesamum indicum*, and *Citrullus lanatus*. The proximity of isolate Pcu-2026KZ to the French isolate C05-205 was confirmed with 100% bootstrap support. The other sequences from the same sub-cluster originated from *Cucurbita* sp. from South Korea and Turkey.

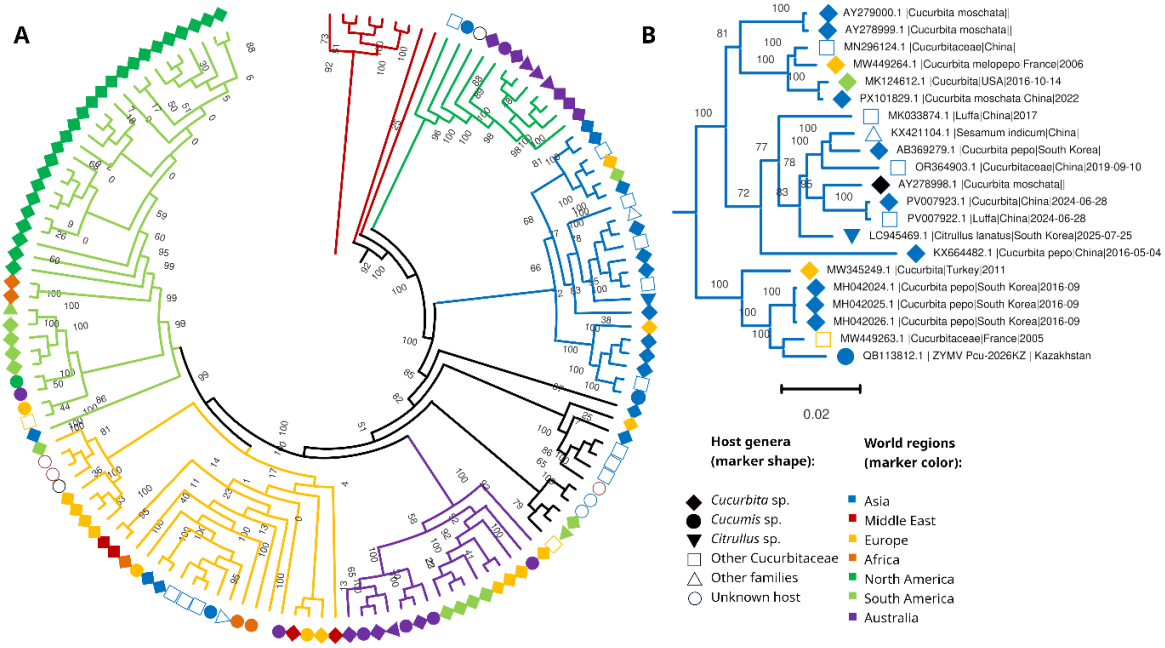


Figure 3. Maximum likelihood tree of ZYMV isolates: (A) complete tree of 135 sequences; outgroup of potyviruses from Cucurbitaceae hosts shown red, other colours indicate individual clusters; (B) cluster containing isolate Pcu-2026KZ.

Based on the phylogenetic analysis, we suppose that isolate Pcu-2026KZ was introduced in Kazakhstan from Europe, as the closest isolate was from France; in turn, this isolate probably originated from Asia, as the common cluster suggests. The cucumber variety tested here, Rava/Meva, was imported from the Netherlands; however, the absence of ZYMV isolates from this country in the datasets does not allow us to verify whether the virus was imported along with the seed material. On the other hand, the cluster containing isolate Pcu-2026KZ consisted mainly of isolates infecting *Cucurbita* sp.; this may suggest that the virus could be initially introduced to Kazakhstan with the pumpkin seeds and then transmitted cross-specifically to cucumber plants via aphids. As the present data are limited by the only identified isolate, further screening is required to elucidate the distribution of ZYMV in Cucurbitaceae crops in Kazakhstan and thus create a basis for its monitoring.

The study illustrates the power of nanopore RNA-Seq in detection of viral pathogens in plants. The traditional PCR based detection is cheap and reliable but requires a preliminary assumption about the candidate virus to be tested in order to select correct primers and probes. The typical symptoms of viral infections in plants, like chlorosis and wilting, may ambiguously suggest different viruses as probable infecting agents and thus require conducting multiple PCR tests for all candidates, which may increase

total costs and required labour and thus reduce the advantages of the method. Generally, the specificity of PCR limits universality of the tests and at the same time does not guarantee absence of cross-specific results [16]. Moreover, PCR tests are based on particular regions of viral genomes and thus do not reflect the complete diversity of the virus, which is crucial for tracking and understanding its distribution. With a typical genome size of 5–10 kbp in plant viruses, PCR based genome amplification becomes more tedious and demanding for optimization of the experimental procedure. RNA-Seq, while more expensive, solves these problems and may provide complete sequences of viral genomes directly from total RNA of the host tissue. Nanopore sequencing becomes one of the standard tools for a range of applications in plant pathology and crop protection [17]. As the technology keeps getting developed, the increasing efficiency and decreasing costs per sample will further expand its applicability to wide scale virological surveys in crops and increase its impact on crop protection in Kazakhstan as well as worldwide.

Conclusions

In the present study, Zucchini yellow mosaic virus (ZYMV) was detected for the first time in Kazakhstan in a cucumber plant (*Cucumis sativus* L.), and the complete genome sequence of the Pcu-2026KZ isolate was obtained. Phylogenetic analysis placed the isolate within a cluster comprising predominantly Asian isolates, with the C05-205 isolate from France being the most closely related. These findings contribute to the understanding of the genetic diversity and geographic distribution of ZYMV and provide a basis for further monitoring of the virus in Cucurbitaceae crops in Kazakhstan.

Data Availability Statement

The genome sequence of ZYMV isolate Pcu-2026KZ is deposited in the NCBI Nucleotide database (https://www.ncbi.nlm.nih.gov/nucleotide/) under accession number QB113812.1.

Author Contributions

Conceptualization: A.P., D.G.; Methodology: A.P., D.G.; Formal analysis: A.P., V.K.; Investigation: A.P., V.K.; Writing – original draft: A.P., V.K. and D.G.; Writing – review & editing: A.P., V.K. and D.G.; Supervision: D.G.

Funding

This research was funded by the Ministry of Agriculture of the Republic of Kazakhstan within the framework of a targeted funding program BR22887230.

Ethics approval and consent

Not applicable. The study did not involve humans or animals.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Desbiez C, Lecoq H (1997) Zucchini yellow mosaic virus. *Plant Pathology* 46 (6):809-829. DOI: 10.1046/j.1365-3059.1997.d01-87.x
2. Moya-Ruiz CD, Gómez P, Juárez M (2023) Occurrence, distribution, and management of aphid-transmitted viruses in cucurbits in Spain. *Pathogens* 12 (3):422. DOI: 10.3390/pathogens12030422
3. Simmons HE, Dunham JP, Zinn KE, et al. (2013) Zucchini yellow mosaic virus (ZYMV, Potyvirus): vertical transmission, seed infection and cryptic infections. *Virus Research* 176 (1):259-264. DOI: 10.1016/j.virusres.2013.06.016
4. Gal-On A (2007) Zucchini yellow mosaic virus: insect transmission and pathogenicity — the tails of two proteins. *Molecular Plant Pathology* 8 (2):139-150. DOI: 10.1111/j.1364-3703.2007.00381.x
5. Simmons HE, Holmes EC, Gildow FE, et al. (2011) Experimental verification of seed transmission of Zucchini yellow mosaic virus. *Plant Disease* 95 (6):751-754. DOI: 10.1094/PDIS-11-10-0843
6. Coutts BA, Kehoe MA, Jones RAC (2013) Zucchini yellow mosaic virus: contact transmission, stability on surfaces, and inactivation with disinfectants. *Plant Disease* 97 (6):765-771. DOI: 10.1094/PDIS-08-12-0769-RE
7. Maghamnia HR, Hajizadeh M, Azizi A (2018) Complete genome sequence of Zucchini yellow mosaic virus strain Kurdistan, Iran. *3 Biotech* 8 (3):147. DOI: 10.1007/s13205-018-1177-3
8. Maina S, Coutts BA, Edwards OR, et al. (2017) Zucchini yellow mosaic virus populations from East Timorese and northern Australian cucurbit crops: molecular properties, genetic connectivity, and biosecurity implications. *Plant Disease* 101 (7):1236-1245. DOI: 10.1094/PDIS-11-16-1672-RE
9. Alinizi HR, Mehrvar M, Zakiaghl M (2021) Analysis of the molecular and biological variability of Zucchini yellow mosaic virus isolates from Iran and Iraq. *Gene* 788:145674. DOI: 10.1016/j.gene.2021.145674
10. Maina S, Barbetti MJ, Edwards OR, et al. (2019) Zucchini yellow mosaic virus genomic sequences from Papua New Guinea: lack of genetic connectivity with northern Australian or East Timorese genomes, and new recombination findings. *Plant Disease* 103 (6):1326-1336. DOI: 10.1094/PDIS-09-18-1666-RE
11. Shen W, Le S, Li Y, Hu F (2016) SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE* 11 (10):e0163962. DOI: 10.1371/journal.pone.0163962
12. Brister JR, Ako-adjei D, Bao Y, Blinkova O (2015) NCBI viral genomes resource. *Nucleic Acids Research* 43 (D1):D571-D577. DOI: 10.1093/nar/gku1207
13. Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN]. DOI: 10.48550/arXiv.1303.3997
14. Li H, Handsaker B, Wysoker A, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25 (16):2078-2079. DOI: 10.1093/bioinformatics/btp352
15. Kumar S, Stecher G, Suleski M, et al. (2024) MEGA12: Molecular Evolutionary Genetics Analysis version 12 for adaptive and green computing. *Molecular Biology and Evolution* 41 (12):msae263. DOI: 10.1093/molbev/msae263
16. Malinowski T (2005) Potential problems with the reliability of PCR based diagnostic methods related to plant viruses sequence variation. *Phytopathologia Polonica* 35:125-139.
17. Sun K, Liu Y, Zhou X, et al. (2022) Nanopore sequencing technology and its application in plant virus diagnostics. *Frontiers in Microbiology* 13:939666. DOI: 10.3389/fmicb.2022.939666

Author Information

Full Name	Alexandr Pozharskiy
Department	
Institution	Laboratory of Molecular Biology, Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan
Email	aspozharsky@gmail.com
ORCID	0000-0002-2581-2860
Scopus ID	

Full Name	Valeriya Kostyukova
Department	
Institution	Laboratory of Molecular Biology, Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan
Email	valera.kostykova.15@gmail.com
ORCID	0000-0002-8922-964X
Scopus ID	

Full Name	Dilyara Gritsenko
Department	
Institution	Laboratory of Molecular Biology, Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan; Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan; Research Center AgriBioTech, Almaty 050040, Kazakhstan
Email	d.gritsenko@ipbb.kz
ORCID	0000-0001-6377-3711
Scopus ID	