Research Article

Nucleotide Sequence and Bioinformatics Analysis of the Complete Genome of an Egyptian Isolate of Zucchini Yellow Mosaic Virus

Fatma S. Abdel Razek[1], Ahmed Mahdy[2*], Samar S.A. El-Masry[1], Shafik Ibrahim[3], Shrouk E.E. Farg[1] and Atef Sadik[1]

[1]Department of Agricultural Microbiology, Laboratory of Virology, Faculty of Agriculture, Ain Shams University, P.O. Box 68, Hadayek Shobra, Cairo, Egypt; [2]Department of Agricultural Microbiology, Faculty of Agriculture, Zagazig University, 44511 Zagazig, Egypt; [3]Agricultural Genetic Engineering Research Institute, Agricultural Research Center, 9 Gamaa St., P.O. Box, 12619, Giza, Egypt.

Abstract | This study aimed to characterize the complete genome of an Egyptian strain of Zucchini Yellow Mosaic Virus (ZYMV), identify key genetic elements linked to aphid transmission, and analyze its phylogenetic relationship with global ZYMV strains. A single necrotic local lesion technique was used for the biological purification of a Zucchini yellow mosaic virus (ZYMV) strain from the Eskandarani squash cultivar, and the purification was confirmed by reverse-transcriptase polymerase chain reaction (RT-PCR). Flexuous filament virions with a size of 11 × 730 nm were purified. The complete nucleotide sequence of the ZYMV genome, consisting of 9591 nucleotides, was obtained and deposited in GenBank (accession numbers: LC795783.1 for nucleotides and BFD45315.1 for protein). The genome encodes a polyprotein of 3080 amino acids (~350 kDa). The genome of the Egy-1920 strain contains ten genes, mainly *P1, HC, P3, 6K1, C1, 6K2, NIa-VPg, NIa-Pro, NIb*, and *cp*, with sizes ranging from 156 to 1902 nucleotides. The strain exhibited high nucleotide (92.58–99.93 %) and protein (95.68–99.81 %) sequence identity compared to 45 global ZYMV strains, showing significant homology to Taiwanese and Chinese strains. Phylogenetic analysis placed the Egy-1920 strain in close relation to Taiwanese ZYMV strains. Bioinformatics analysis of the helper component (HC) protein revealed three aphid-transmission motifs. The genome length variation (9243-9947 nts) among ZYMV strains was also discussed. The results pave the way for the development of reliable diagnostic techniques and the detection of possible therapy targets for ZYMV control. The findings of this study will contribute to the development of enhanced diagnostic methods and potential therapeutic strategies for effective ZYMV control.

Received | January 29, 2025; **Revised** | March 10, 2025; **Accepted** | March 29, 2025; **Published** | April 13, 2025
***Correspondence** | Ahmed Mahdy, Department of Agricultural Microbiology, Faculty of Agriculture, Zagazig University, 44511 Zagazig, Egypt; **Email:** micromicro2000@gmail.com
Citation | Razek, F.S.A., A. Mahdy, S.S.A. El-Masry, S. Ibrahim, S.E.E. Farg and A. Sadik. 2025. Nucleotide sequence and bioinformatics analysis of the complete genome of an Egyptian isolate of zucchini yellow mosaic virus. *Novel Research in Microbiology Journal,* 9(2): 92-115.
DOI | https://dx.doi.org/10.17582/journal.NRMJ/2025/9.2.92.115
Keywords | Aphid transmission, Biological purification, Bioinformatics analysis, RT-PCR, Squash, ZYMV

Copyright: 2025 by the authors. Licensee ResearchersLinks Ltd, England, UK.
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Introduction

Viruses are among the most significant plant pathogens, causing substantial productivity losses in agricultural crops (Velasco *et al*., 2020; Ahsan *et al*., 2023; Tatineni and Hein, 2023). Within the *Cucurbitaceae* family, several crops like squash (*Cucurbita* spp.) hold immense economic value

worldwide. However, the production of these crops faces a growing threat from viral diseases, especially Zucchini Yellow Mosaic Virus (ZYMV). ZYMV is a major cucurbits pathogen in cucurbits, which can cause output losses of up to 100 %, posing a serious challenge for farmers in several regions across the globe (Lecoq *et al*., 2009; Sharma, 2023). Within the *Potyviridae* family, ZYMV is classified as a member of the *Potyvirus* genus, having been first recorded in Italy in 1973 (Bubici *et al*., 2020). The virus causes rapid epidemics in cucurbits such as zucchini, pumpkin, and cucumber, and is primarily disseminated through aphids in a non-persistent manner (Katis *et al*., 2006; Perotto *et al*., 2018).

The genome of ZYMV is a single-stranded, positive-sense RNA molecule that is approximately 9.6 kilo base (kb) in length. It encodes a single large open reading frame (ORF), which is processed into a polyprotein of around 3080 amino acids. Three viral proteases cleave this polyprotein into 10 functional proteins essential for replication, movement, and transmission. These include P1(protease), HC-Pro (helper component/protease), P3, 6K1, CI (cylindrical inclusion protein), 6K2, NIa (nuclear inclusion protein A), VPg (genome-linked viral protein), NIb (nuclear inclusion protein B), and CP (coat protein) (Dong *et al*., 2022). Additionally, a small overlapping ORF, known as PIPO, is translated in the +2 reading frame within the P3 region, which plays a role in the virus's replication cycle (Chung *et al*., 2008). The 5′ and 3′ un-translated regions (UTRs) of the genome also contain regulatory elements that facilitate cap-independent translation that is essential for efficient production of viral proteins (Kneller *et al*., 2006; Bernet and Elena, 2015).

To better understand of ZYMV and its genetic diversity, researchers have sequenced several isolates of the virus from different parts of the world. Lee and Wong (1998) sequenced a Singapore isolate (ZYMV-S), which was 9603 nucleotides long, revealing a polyprotein of 3082 amino acids. Their analysis identified a mutation in the HC-Pro protein which may affect aphid transmission by changing the conserved motif from K-I-T-C to K-L-S-C. Further sequence comparisons showed that ZYMV-S shared 65-98 % identity with other isolates, with higher similarity to strains from California and Reunion Island. Notably, the 5′ UTR exhibited 67-72 % identity with these isolates, while the 3′ UTR expressed 82-94 % identity. This sequence variability, particularly in the P1 protein, offered insights into strain differentiation and potyvirus taxonomy.

The Taiwan isolate (ZYMV-TW-TN3) sequenced by Dong *et al*. (2022) also contributed to the understanding of the virus's genetic diversity. The genome length was 9591 nucleotides (nt), encoding a polyprotein of 3080 amino acids, in addition, it exhibited significant variation in the P1 protein (59.0-93.2 % identity). The 5′ and 3′ UTRs showed distinct nucleotide identities (61.6-83.3 % and 90.4-95.7 %, respectively). Phylogenetic analysis classified ZYMV-TW-TN3 as a genotype I isolate, showing closer relationships with U.S. isolates (genotype II) and further emphasizing the genetic variation across different geographical regions.

In Europe, Glasa and Pittnerová (2006) sequenced the Slovak ZYMV-Kuchyna isolate, which was 9593 nucleotides long, similar in size to other isolates. The analysis revealed 90.4-98.8 % nucleotide identity and 78-98.8 % amino acid identity compared to 12 other ZYMV isolates. This study underscores the importance of understanding the conserved and variable regions of the genome, particularly when developing strategies for disease control and vector management.

Another study reported by Choi *et al*. (2007) sequenced the genome of a ZYMV isolate from hollyhock (ZYMV-A). The genome length was again 9593 nt and the polyprotein was 3080 amino acids long. This isolate exhibited conserved proteolytic cleavage sites and the key KITC and DAG motifs in the HC-Pro and CP genes, respectively, which were essential for aphid transmission. Phylogenetic analysis of ZYMV-A virus revealed it to be a distinct group within ZYMV, establishing it as a unique strain at the molecular level. The presence of these conserved motifs confirmed its potential for vector transmission.

In the Kurdistan region, Maghamnia *et al*. (2018) sequenced a ZYMV strain infecting squash, which also had a genome length of 9593 nt. This strain showed 79.6-98.8 % nucleotide identity with other ZYMV isolates and exhibited high homology with Central European strains. Their analyses highlighted the importance of genome sequencing in understanding ZYMV's evolution, genetic diversity, and potential for strain-specific diagnostics.

Links Researchers

Finally, Chinnadurai *et al.* (2021) characterized four ZYMV isolates from Trinidad and Tobago by complete genome sequencing. Phylogenetic analysis revealed 5.9-6.0 % nucleotide and 7.7-7.9 % amino acid sequence divergence from the most closely related isolates (from Israel and Slovakia). Based on these genetic differences, they proposed a new genotype, ZYMV-Trini, which displayed the greatest divergence in the HC-Pro gene. This study confirmed that ZYMV can be transmitted through seeds, though aphid transmission *via* Aphis gossypii remained the more common route.

These studies highlight the extensive genetic diversity among ZYMV isolates, with genomic variations impacting transmission, pathogenicity, and host range (Lecoq *et al*., 2009; Ali and Kobayashi, 2010). Sequencing the entire genomes of ZYMV strains from diverse geographical regions is essential for devising targeted control strategies and deepening our understanding of virus evolution and epidemiology (Lecoq and Desbiez, 2012). The objective of this study was to focus on the complete nucleotide sequence and bioinformatics analysis of an Egyptian ZYMV isolate, by comparing this isolate with others from various regions and hosts.

Materials and Methods

Source of viral isolate

The Egyptian ZYMV isolate was obtained from naturally infected squash (*Cucurbita pepo* cv. *Eskandarani*) showing symptoms of yellow mosaic, vein banding, and leaf deformation. Identification as ZYMV was confirmed at the Laboratory of Virology, Ain Shams University, Cairo, Egypt.

Biological purification and propagation

The isolate was purified using the single local lesion (SLL) technique (Antoniw and White, 1986). Sap from crushed lesions was used to inoculate squash plants, which were then incubated in a greenhouse for 15 d.

Purification and confirmation

ZYMV was further purified following the procedure conducted by Brakke (1961). Infected leaves were homogenized in extraction buffer (0.1 M phosphate, pH 7.0, with 1 % PVP and 0.02 M sodium sulfate), filtered, and subjected to differential centrifugation. The supernatant was centrifuged at 8,000 × g to remove debris, followed by high-speed centrifugation at 40,000 × g to pellet virus particles. Sucrose gradient (30 %–60 %) centrifugation at 100,000 × g separated the virus, which was dialyzed, concentrated, and stored at −80 °C. Negative staining was performed following Brenner and Horne (1959). A drop of purified virus was placed on a carbon-coated copper grid, left to adhere, blotted, stained with 2 % uranyl acetate for 30 sec, and then air-dried. The sample was observed under a transmission electron microscopy (TEM) to confirm virus morphology.

Reverse transcription-polymerase chain reaction (RT-PCR) confirmation

Total RNA was extracted from the infected leaves, and RT-PCR was performed using ZYMV-specific primers (Kuan *et al*., 2014), namely *ZYU-F* (5'-AT-GTCTCAAGGTCGAGTA-3') and *ZYD1186-R* (5'-TTATTAGGTCCAGCACGGC-3'). The PCR conditions were as follows: initial denaturation: 95 °C for 3 min followed by denaturation: 95 °C for 30 sec, annealing: 55 °C for 30 sec (primer-specific), and extension: 72 °C for 30 sec to 1 min. The final cycle of the 35 cycles was extended for 5 min at 72 °C. Amplified DNA fragments were resolved on a 1 % Agarose gel, stained, and visualized under UV light.

Full genome sequencing

cDNAs were synthesized using the total RNA extracted from the purified virus preparation in the presence of primers specific to the ZYMV genome, as detailed in Table 1 (Abdel Aleem *et al*., 2021). Overlapping genomic segments were then amplified and sequenced using Sanger sequencing (ABI Prism 3730). The complete genome was assembled using Geneious software, and sequence alignment along with phylogenetic analysis was performed using MEGA software (Kumar *et al*., 2018).

Results

Biological purification and molecular confirmation of the viral isolate

The ZYMV isolate was biologically purified using the single local lesion technique on *Chenopodium amaranticolor*, followed by propagation on squash plants (*Cucurbita pepo* cv. *Eskandarani*), as shown in Figure 1. After 14 d post-propagation on squash plants, an electron micrograph of the partially purified ZYMV particles stained with 2 % uranyl acetate revealed the virus morphology as flexuous filamentous particles.

Links Researchers

Table 1: *A set of primer pairs designed for amplification of ZYMV genomic fragments for molecular analysis, based on complete sequence of an Egyptian ZYMV strain deposited in GenBank (MT383108.1).*

Primer pairs	Primer codes	Sequences (5'→3')	Regions	Amplified fragments (pb)
PP-01	ZYMV-F01	AAATTAAAACAAATCACAAAGA	2-23	311
	ZYMV-R01	AGAATGTCAATGCTTTTGTTTG	334-312	
PP-02	ZYMV-F02	GTGTTTACAAACAAGCAATCTA	62-84	637
	ZYMV-R02	AACAGGGATATTTTTATTGCGT	721-699	
PP-03	ZYMV-F03	GGTTAATATAGTGGCACCTGGC	229-241	768
	ZYMV-R03	GATCACCAGAAGCTCCTATAAC	1009-987	
PP-04	ZYMV-F04	GCGAGACATACACTCACCTTCA	743-765	1266
	ZYMV-R04	GATCACCAGAAGCTCCTATAAC	2031-2009	
PP-05	ZYMV-F05	ACAAGCACTCACATGAAGCAAA	1442-1464	589
	ZYMV-R04	GATCACCAGAAGCTCCTATAAC	2031-2009	
PP-06	ZYMV-F06	ATATTATTGCTTGGCATGATTT	2532-2554	769
	ZYMV-R05	CATCGCTCTCTCTTTTTCTTCC	3301-3279	
PP-07	ZYMV-F07	TAGAGGAGCAGTAGGTTCTGGA	3879-3901	645
	ZYMV-R06	TTTGGCTAGTTCCATGCGTTTC	4524-2502	
PP-08	ZYMV-F08	GCCACAGAAGCAGCAGCCTTAT	4756-4778	565
	ZYMV-R07	TGTCGAAGTGATTCCGCTTCAT	5321-5299	
PP-09	ZYMV-F08	GCCACAGAAGCAGCAGCCTTAT	4756-4778	1126
	ZYMV-R08	AATTCTCAGGCTCCACACCATA	5904-5882	
PP-10	ZYMV-F09	CGGACTTCGCGGTAAATGGGAT	5559-5581	1250
	ZYMV-R09	TGCGTGGCTATAAAATCATCAG	6831-6809	
PP-11	ZYMV-F10	CATGTCGAGTTGGAGAGCAAAT	6247-6269	584
	ZYMV-R09	TGCGTGGCTATAAAATCATCAG	6831-6809	
PP-12	ZYMV-F11	TCATCTAAGAATTATTTTGTCC	6782-6803	1042
	ZYMV-R10	CAACCCACCAATCCTCCATATA	7824-7802	
PP-13	ZYMV-F12	GAGAATAGACTTATCTTCTTTG	8014-8036	469
	ZYMV-R11	CTTGATGGAGGGCTTGTAGGTA	8505-8483	
PP-14	ZYMV-F13	GAAAGACAAAGAAGATGACAAA	8584-8606	539
	ZYMV-R12	TAGGTTTCGAAGCAAACCATAC	9145-9123	
PP-15	ZYMV-F14	GATGTTAATAGAAACATGCACA	9331-9353	240
	ZYMV-R13	AGGCTTGCAAACGGAGTCTAAT	9593-5971	

RT-PCR analysis confirmed the presence of ZYMV in the infected plants as illustrated in Figure 2. An 837 bp PCR product was successfully amplified from the ZYMV-infected sample (Lane S), corresponding to the expected ZYMV sequence, compared to the positive control (Lane +ve), which was absent in the negative control (Lane -ve). A 50 bp DNA ladder (Lane M) was used to estimate the product size (Figure 2).

Primers designed to amplify the complete genome of the viral isolate

A total of fifteen primer pairs were designed to amplify overlapping segments of the ZYMV genome, as detailed in the materials and methods (Table 1). These primers targeted specific regions of the ZYMV genome to enable comprehensive molecular analysis. Amplification results demonstrated the successful amplification of all 15 overlapping segments, measuring 311, 637, 768, 1266, 589, 769, 645, 565, 1126, 1250, 584, 1042, 469, 539, and 240 bp, respectively (Figure 3). The cDNA used for amplification was synthesized from RNA extracted from a purified virus preparation.

Sequencing of the amplified products and their analysis

The amplified fragments were sequenced and analyzed collectively, resulting in the reconstruction of the full-length genome of the Egyptian ZYMV strain.

Links Researchers

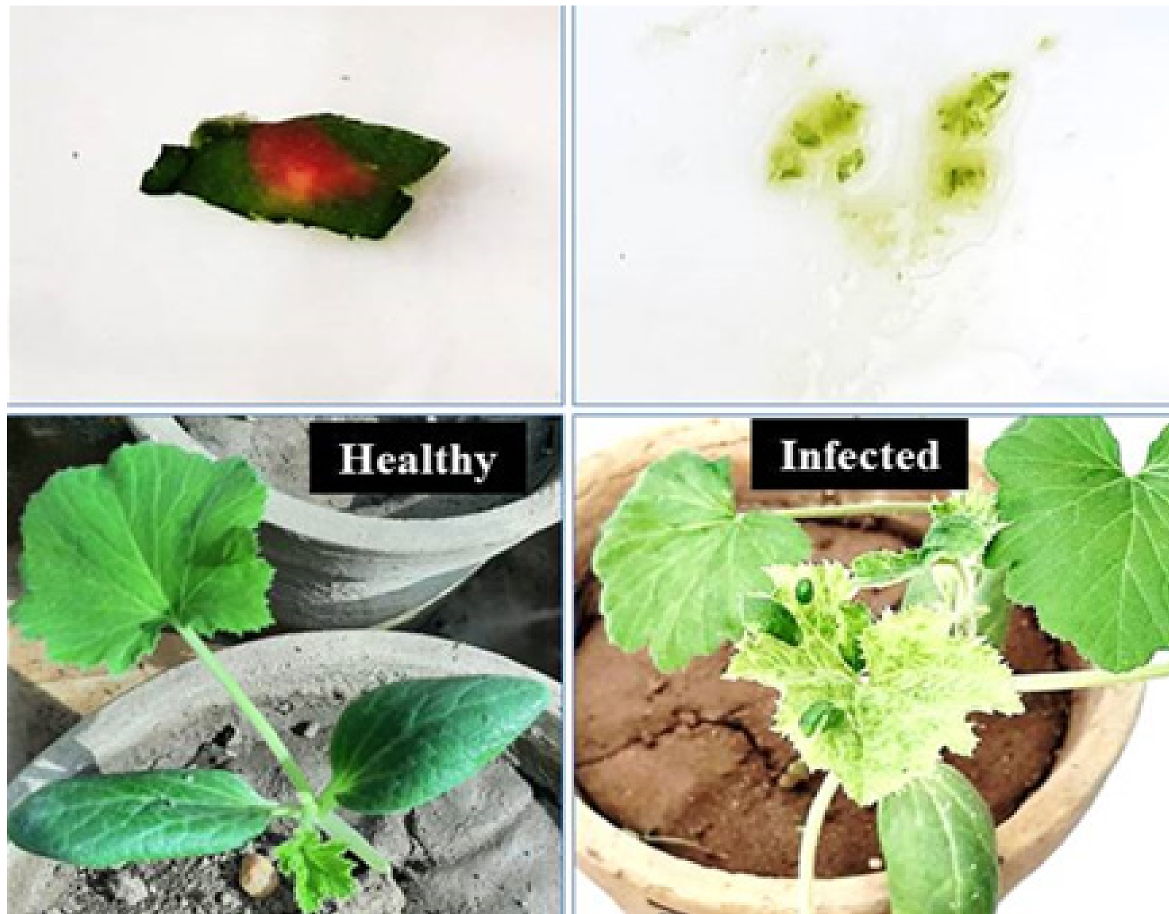


Figure 1: *Biological purification of the Zucchini yellow mosaic virus isolate was performed using the single local lesion technique on the diagnostic host Chenopodium amaranticolor and then propagated on squash plants (Cucurbita pepo cv. Eskandarani).*

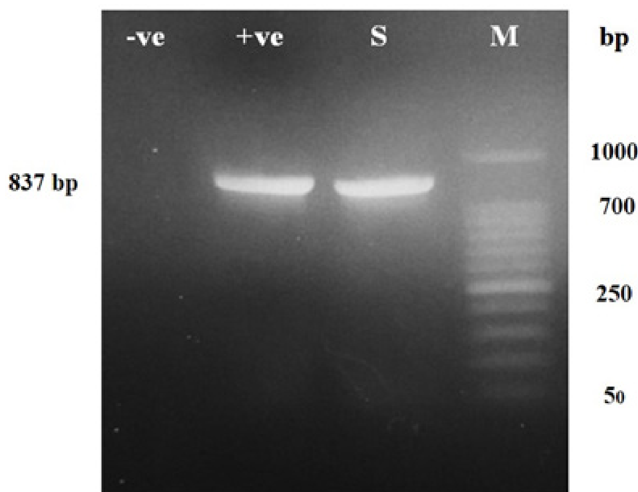
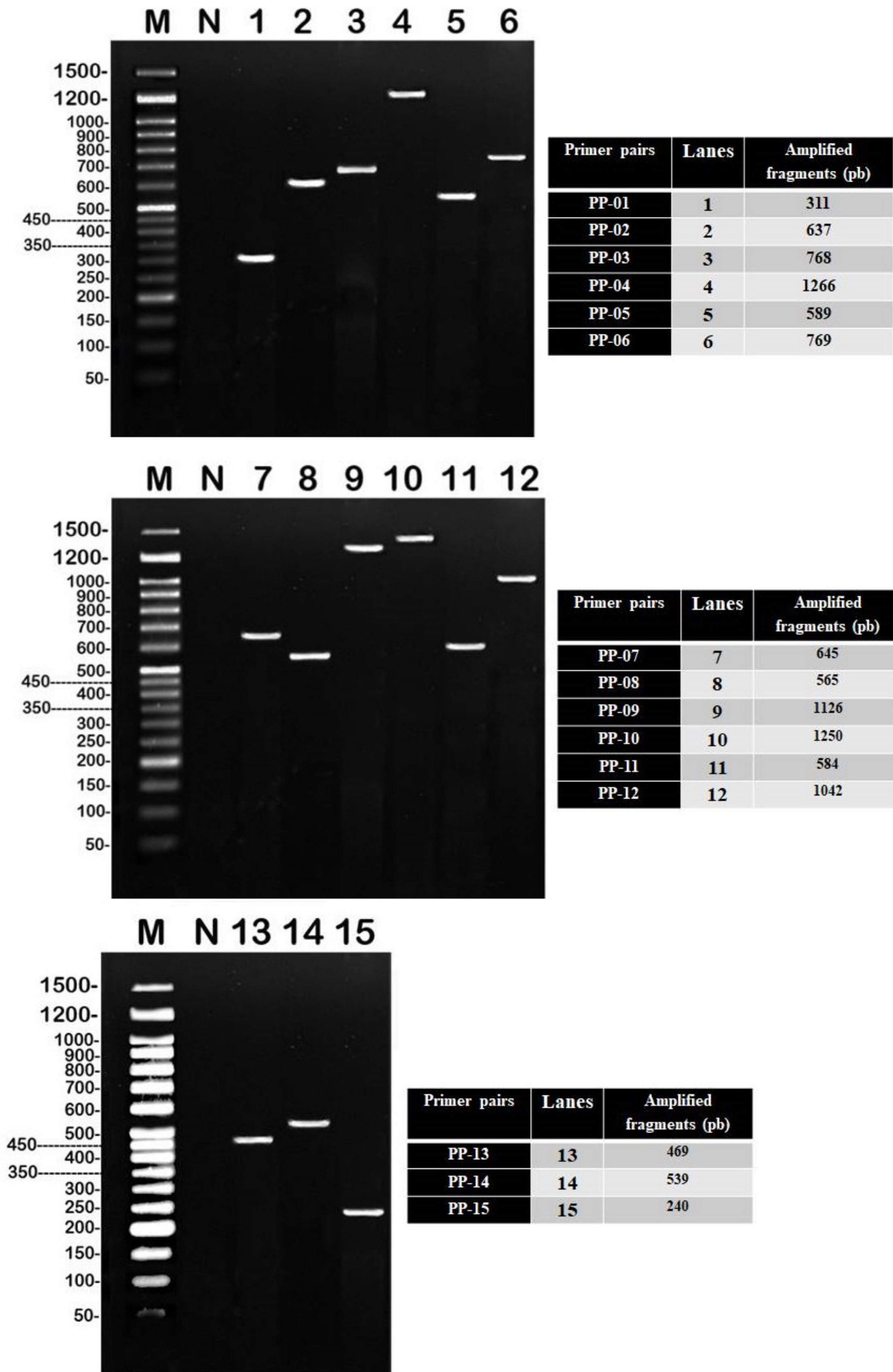


Figure 2: *An agarose gel (1.2 %) stained with ethidium bromide confirms the reverse transcription-polymerase chain reaction amplification of the Zucchini yellow mosaic virus isolate. A distinct 837 bp PCR product was observed in the ZYMV-infected plant sample (Lane S), consistent with the expected size. The positive control (Lane +ve) also showed an 837 bp band, verifying the assay's accuracy, while no amplification was detected in the negative control (Lane -ve), confirming the absence of contamination. M, 50 bp DNA ladder.*

Dividing the genome into manageable overlapping fragments allowed for efficient sequencing and accurate assembly. Overlapping regions further ensured high coverage and verification of sequence integrity during the assembly process. This primer set was a valuable tool for studying genetic variations across ZYMV isolates. Additionally, the complete genome sequence obtained through this approach provided critical insights into the molecular biology, genetic diversity, and epidemiology of ZYMV.

Data in Table 1 provides a detailed breakdown of the genomic structure of the ZYMV-Egy-1920 strain, recorded in GenBank under accession ID LC795783.1. The genome spans 9,591 nucleotides, consistent with typical ZYMV genome sizes, which generally ranged from 9.5 to 10 kb. Each genomic segment and corresponding gene product is listed with its respective accession, nucleotide length, protein length, and protein identifier.

Links Researchers



Primer pairs	Lanes	Amplified fragments (pb)
PP-01	1	311
PP-02	2	637
PP-03	3	768
PP-04	4	1266
PP-05	5	589
PP-06	6	769

Primer pairs	Lanes	Amplified fragments (pb)
PP-07	7	645
PP-08	8	565
PP-09	9	1126
PP-10	10	1250
PP-11	11	584
PP-12	12	1042

Primer pairs	Lanes	Amplified fragments (pb)
PP-13	13	469
PP-14	14	539
PP-15	15	240

Figure 3: *An agarose gel (1.2 %) stained with ethidium bromide shows the reverse transcription-polymerase chain reaction amplification of overlapping genome segments of the Zucchini yellow mosaic virus isolate. Lanes 1–15 correspond to the PCR products amplified using the 15 primer pairs designed to amplify the full genome of the viral isolate. 'N' denotes the negative control, which is PCR without any template, and 'M' represents the DNA marker. The amplified fragments, generated using 15 primer pairs covering the full genome, are clearly visible. No amplification was observed in the negative control (Lane N), confirming the absence of contamination.*

Dividing the genome into manageable overlapping fragments allowed for efficient sequencing and accurate assembly. Overlapping regions further ensured high coverage and verification of sequence integrity during the assembly process. This primer set was a valuable tool for studying genetic variations across ZYMV isolates. Additionally, the complete genome sequence obtained through this approach provided critical insights into the molecular biology, genetic diversity, and epidemiology of ZYMV.

The ZYMV genome included two un-translated regions: A 5' UTR (138 nt) and a 3' UTR (213 nt), which served regulatory roles in viral RNA replication and protein synthesis but did not encode proteins. Between these UTRs, there were ten coding genes essential for the virus's lifecycle: *P1* protease: 930 nt, 310 aa; *HC-Pro*: 1368 nt, 456 aa; *P3* protein: 1038 nt, 346 aa; *6K1*: 156 nt, 52 aa; *C1* protein: 1902 nt, 634 aa; *6K2*: 159 nt, 53 aa; *NIa-VPg*: 570 nt, 190 aa; *NIa-Pro*: 729 nt, 243 aa; *NIb* protein: 1551 nt, 517 aa, and *cp* (CP): 837 nt, 279 aa (Figure 4 and Table 2).

The total translated proteins consisted of 3080 amino acids across all coding regions, aligning with known potyvirus characteristics. This structural overview was essential for understanding ZYMV-Egy-1920's genome organization and protein functions.

The nucleotide sequences of the ZYMV-Egy-1920 strain (GenBank: LC795783.1) and its polyprotein sequences (BFD45315.1) isolated from *Cucurbita pepo* cv. *Eskandarani* (Egypt) were compared to 45 similar ZYMV strains from various hosts and countries (Table 3 and Figure 5). The hosts included cucurbits (*Cucurbita pepo*, *Cucumis sativus*,

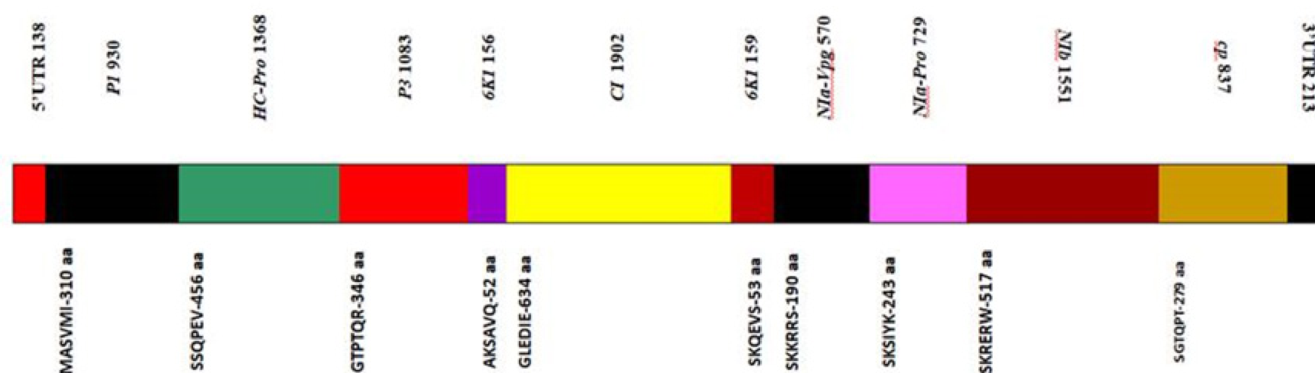


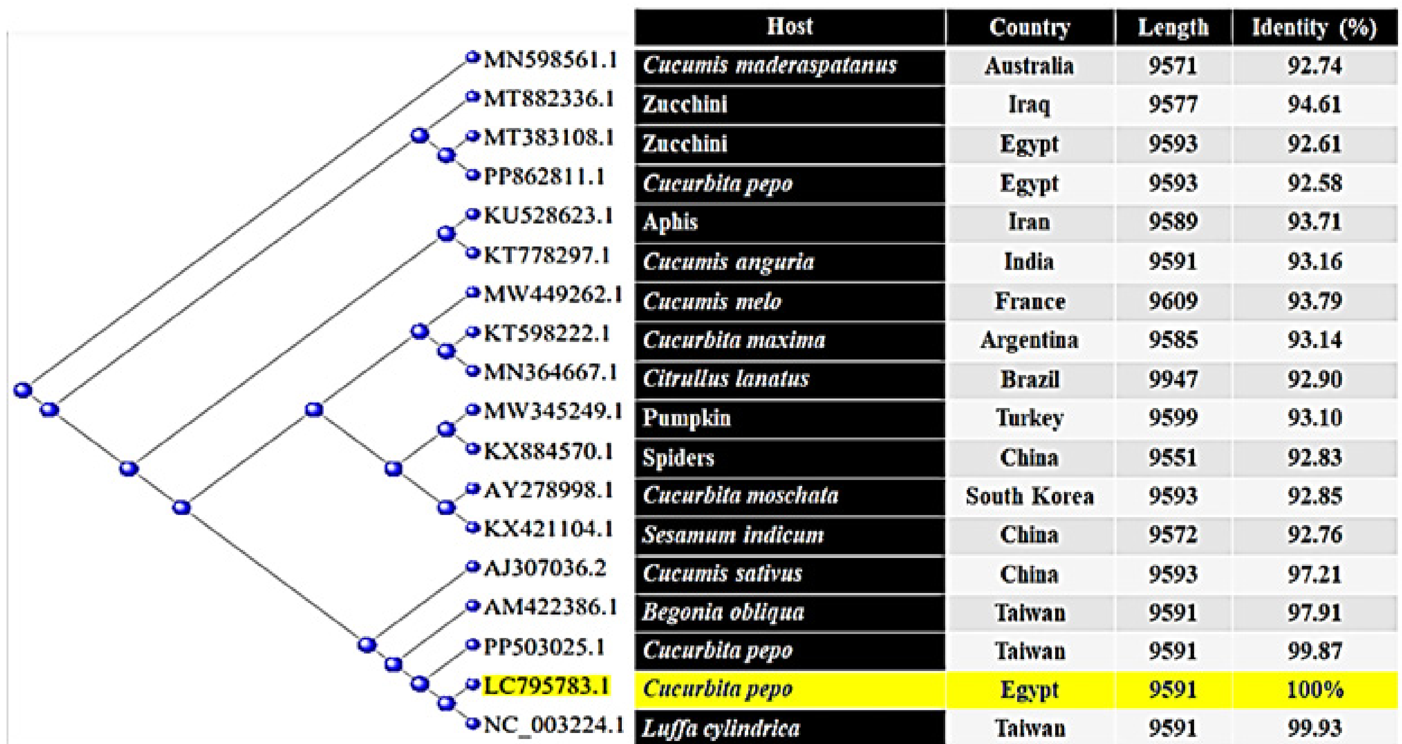
Figure 4: *Genome organization of Zucchini yellow mosaic virus-Egy-1920 strain (LC795783.1). The genome started with 5'UTR (138 nts), followed by P1 gene (930 nts), HC-Pro gene (1368 nts), P3 gene (1083 nts), 6K1 gene (156 nts), CI gene (1902 nts), 6K2 gene (159 nts), Nia-Pro gene (729 nts), Nib (1551 nts), cp gene (837 nts), and 3"UTR (213 nts).*

Table 2: *Structure of the complete genome of ZYMV-Egy-1920 strain (LC795783.1)*.

Accessions	Length	Region	Gene	Product	Protein_id	Length
LC799405.1	138	5'UTR	....	....	....	....
LC799406.1	930	Gene 01	P1	P1 protease	BFF82028.1	310
LC799407.1	1368	Gene 02	HC	HC-Pro protein	BFF82029.1	456
LC799408.1	1038	Gene 03	P3	P3 protein	BFF82030.1	346
LC799409.1	156	Gene 04	6K1	6K1 protein	BFF82031.1	52
LC799410.1	1902	Gene 05	C1	C1 protein	BFF82032.1	634
LC799411.1	159	Gene 06	6K2	6K2 protein	BFF82033.1	53
LC799412.1	570	Gene 07	NIa-VPg	NIa-VPg protein	BFF82034.1	190
LC799413.1	729	Gene 08	NIa-Pro	NIa-Pro protein	BFF82035.1	243
LC799414.1	1551	Gene 09	NIb	NIb protein	BFF82036.1	517
LC799415.1	837	Gene 10	cp	Coat protein	BFF82037.1	279
LC799416.2	213	3'UTR	....	....	....	....
Total	9591	12	10	10	10	3080

Table 3: *Geographic origin, source of isolation, and sequence identity percentages at the nucleotide and protein levels of 45 ZYMV isolates/strains compared to the ZYMV-Egy-1920 strain (LC795783.1 and BFD45315.1).*

Country	Source of isolation (Host)	Collection date	Nucleotide	Length (nts)	Identities (%)	Protein-id	Identities (%)
Taiwan	*Luffa cylindrica*	1999	NC_003224.1	9591	99.93	NP_477522.1	99.81
Taiwan	Cucurbita pepo	2024	PP503025.1	9591	99.87	WYC12591.1	99.71
China	Luffa cylindrica	2001	AJ316228.2	9592	98.00	CAC87635.2	98.70
Taiwan	Begonia obliqua	2006	AM422386.1	9591	97.91	CAM12729.1	98.51
China	Cucumis sativus	2001	AJ307036.2	9593	97.21	CAC85170.2	98.34
Iraq	Zucchini	2018	MT882336.1	9577	94.61	QWL14830.1	98.21
Spain	Cucurbita pepo	2016	KX499498.1	9592	94.55	AOW31897.1	97.34
Germany	Cucurbita pepo	2024	ON604832.1	9592	94.00	WBG54264.1	97.79
Brazil	Cucurbita pepo	2022	PP256252.1	9593	94.00	WVM33519.1	97.89
UK	Cucurbita pepo	2022	OM471983.1	9592	93.84	UOF93247.1	97.92
USA	Cucurbita pepo	2023	KC665627.1	9515	93.82	AGT95932.1	97.53
Japan	Cucumis sativus	2004	AB188115.1	9592	93.81	BAE75934.1	97.31
Israel	Cucumis sativus	1979	OR233209.1	9593	93.81	WMX25403.1	97.76
France	Cucumis melon	2010	MW449262.1	9609	93.79	QSM07176.1	97.14
France	Cucurbita pepo	2013	OQ847411.1	9243	93.77	WIW79931.1	96.92
Iran	Aphis	2016	KU528623.1	9589	93.71	ANW46658.1	95.68
Iran	Cucurbita pepo	2022	MF684760.1	9592	93.61	AXK59837.1	97.47
Hungary	Cucurbita pepo	2005	ON604841.1	9592	93.60	WBG54297.1	97.44
Slovakia	Cucurbita pepo	2021	DQ124239.1	9593	93.58	AAZ78317.1	97.24
Israel	Cucurbita pepo	2023	OL311706.1	9592	93.55	UOF93031.1	97.47
Spain	Cucumis melo	2013	OR879104.1	9579	93.52	WPR15590.1	97.60
Turkey	Cucumis sativus	2023	MW345248.1	9599	93.35	QTZ21713.1	97.08
Italy	Cucumis sativus	2003	OQ335839.1	9592	93.34	WIW79771.1	97.18
China	Cucumis melon	2012	AJ316229.2	9593	93.27	CAC87636.2	96.53
India	Cucumis anguria	2011	KT778297.1	9591	93.16	ALM55107.1	97.24
Argentina	Cucurbita maxima	2019	KT598222.1	9585	93.14	AMH40820.1	96.82
Italy	Cucurbita pepo	2020	MK956829.1	9611	93.14	QDF46332.1	97.18
Turkey	Pumpkin	1998	MW345249.1	9599	93.10	QTZ21714.1	97.18
Australia	Cucumis melo	2015	MN598580.1	9572	93.07	QID92256.1	96.95
Trinidad	Pumpkin	2018	MF072712.1	9594	93.03	AWX33671.1	97.50
Kenya	Pumpkin	2013	MT497463.1	9572	92.91	QXU64380.1	97.37
Brazi	Citrullus lanatus	2022	MN364667.1	9947	92.90	QLC27861.1	96.88
Sudan	Cucumis sativus	2018	OP357945.1	9592	92.88	UZN89765.1	97.01
South Korea	Cucurbita pepo	2003	MH042026.1	9575	92.87	AXY92165.1	97.18
South Korea	Cucurbita moschata	2013	AY278998.1	9593	92.85	AAQ17214.1	97.24
China	Spiders	2003	KX884570.1	9551	92.83	APG79042.1	96.98
Australia	Cucumis sativus	2010	MN598565.1	9572	92.82	QID92241.1	97.01
Australia	Cucurbita pepo	2023	MN598563.1	9572	92.79	QID92239.1	96.98
France	Zucchini	1992	MW449260.1	9604	92.78	QSM07174.1	97.01
China	Sesamum indicum	2016	KX421104.1	9572	92.76	ARN61640.1	97.24
China	Citrullus lanatus	2010	OQ136666.1	9572	92.75	WII96531.1	96.92
Australia	Cucumis maderaspatanus	2008	MN598561.1	9571	92.74	QID92237.1	96.82
Australia	Citrullus lanatus	2022	MN598564.1	9595	92.69	QID92240.1	96.92
Egypt	Zucchini	2018	MT383108.1	9593	92.61	QNS28122.1	95.78
Egypt	Cucurbita pepo	2021	PP862811.1	9593	92.58	XBR10375.1	95.68



Accession	Host	Country	Length	Identity (%)
MN598561.1	*Cucumis maderaspatanus*	Australia	9571	92.74
MT882336.1	Zucchini	Iraq	9577	94.61
MT383108.1	Zucchini	Egypt	9593	92.61
PP862811.1	*Cucurbita pepo*	Egypt	9593	92.58
KU528623.1	Aphis	Iran	9589	93.71
KT778297.1	*Cucumis anguria*	India	9591	93.16
MW449262.1	*Cucumis melo*	France	9609	93.79
KT598222.1	*Cucurbita maxima*	Argentina	9585	93.14
MN364667.1	*Citrullus lanatus*	Brazil	9947	92.90
MW345249.1	Pumpkin	Turkey	9599	93.10
KX884570.1	Spiders	China	9551	92.83
AY278998.1	*Cucurbita moschata*	South Korea	9593	92.85
KX421104.1	*Sesamum indicum*	China	9572	92.76
AJ307036.2	*Cucumis sativus*	China	9593	97.21
AM422386.1	*Begonia obliqua*	Taiwan	9591	97.91
PP503025.1	*Cucurbita pepo*	Taiwan	9591	99.87
LC795783.1	*Cucurbita pepo*	Egypt	9591	100%
NC_003224.1	*Luffa cylindrica*	Taiwan	9591	99.93

Figure 5: *Distances tree of results shows the genetic relationship between ZYMVgp1 RNA, complete genome strain ZYMV-Egy-1920 (LC795783.1) isolated from Cucurbita pepo cv. Eskandarani, and 17 ZYMV isolates from 15 hosts belonging to 13 countries.*

Cucumis melo, *Luffa cylindrica*, *Cucurbita maxima*, and *Cucurbita moschata*), and non-cucurbits such as *Begonia obliqua*, Aphis, and Spiders. The sequence identities at the levels of nucleotides and protein ranged from 95.68 % to 99.81 % and from 95.68 % to 99.81 %, respectively, showing varying degrees of genetic similarity.

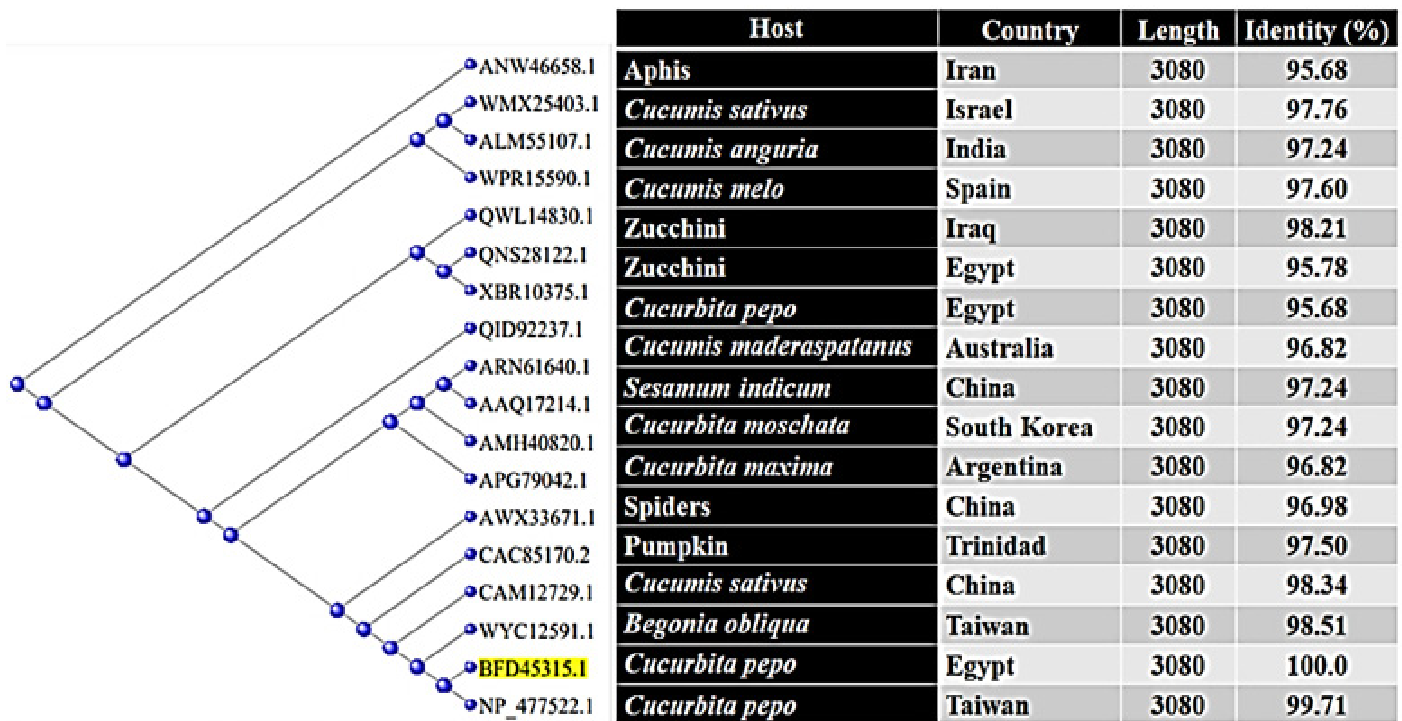
The highest identities of 99.93 % and 99.81 % based on each of nucleotides and protein, respectively, were found in a *Luffa cylindrica* isolate from Taiwan (collected at 1999), followed by a *Cucurbita pepo* isolate from Taiwan (2024, 99.87 % and 99.81 %). While, non-cucurbit hosts, such as Aphis (Iran, 2013, 93.71 % and 95.68 %) and Spiders (China, 2013, 92.83 % and 96.98 %), showed lower identities, suggesting more divergence in these isolates.

ZYMV isolates exhibited global distribution with higher identity values from Taiwan (two strains) and China (one strain) (99.93, 98, 99.87 % for nucleotides and 99.81, 99.71, 98.00% for protein, respectively).

The other two Egyptian ZYMV isolates (MT383108.1 and PP62811.1) had an identity of 92.61 % and 95.78 % and 92.58 % and 95.68 % for nucleotides and protein, respectively, indicating moderate divergence from the other strains. Isolates from several countries like Australia (92.79 % and 96.98 %) and France (92.79 % and 96.92 %) also showed lower identities, suggesting regional genetic variation.

At the level of polyprotein of the ZYMV isolate under investigation, as shown in Figure 6, isolates from Taiwan, China, and the UK showed identity percentages greater than 97 %, indicating close genetic similarity that is likely attributed to regional or host-based factors. The identities of the isolates from multiple countries such as Iran, Spain, and Turkey were generally in the range of 97 % to 96 %, while those isolates from Egypt showed a somewhat lower identity (95.68 %-95.78 %).

Genetic divergence was apparent in isolates Citrullus lanatus strains (from Brazil, China, and Australia), which showed relatively lower identity (96.88 %-96.92 %), possibly due to geographic and ecological variation. The isolates with higher identities were often more recent (e.g., Cucurbita pepo strains from Taiwan, 2024, 99.71%) or collected from regions with long-established cucurbit cultivation. In contrast, older isolates, such as those from Cucumis sativus (Japan, 2004) and Cucurbita pepo (Australia, 2010) tended to exhibit somewhat lower identity values



Host	Country	Length	Identity (%)
Aphis	Iran	3080	95.68
Cucumis sativus	Israel	3080	97.76
Cucumis anguria	India	3080	97.24
Cucumis melo	Spain	3080	97.60
Zucchini	Iraq	3080	98.21
Zucchini	Egypt	3080	95.78
Cucurbita pepo	Egypt	3080	95.68
Cucumis maderaspatanus	Australia	3080	96.82
Sesamum indicum	China	3080	97.24
Cucurbita moschata	South Korea	3080	97.24
Cucurbita maxima	Argentina	3080	96.82
Spiders	China	3080	96.98
Pumpkin	Trinidad	3080	97.50
Cucumis sativus	China	3080	98.34
Begonia obliqua	Taiwan	3080	98.51
Cucurbita pepo	Egypt	3080	100.0
Cucurbita pepo	Taiwan	3080	99.71

Figure 6: *Distances tree of results shows the genetic relationship between polyprotein of ZYMVgp1 strain ZYMV-Egy-1920 (BFD45315.1) isolated from Cucurbita pepo cv. Eskandarani compared to 17 strains of ZYMV isolated from 15 hosts belonging to 13 countries.*

(96.92 %-97.18 %). The ZYMV-Egy-1920 isolate from Egypt showed a lower identity compared to most other strains, with 95.78 % identity to a zucchini isolate from Egypt (2018) and 95.68 % identity to Cucurbita pepo from Egypt (2021). This may reflect regional adaptation or a more distant evolutionary path for the strains circulating in Egypt.

The experimental results showed that cucurbits, particularly *Cucurbita pepo* and *Cucumis sativus*, were the primary hosts, with high identity values indicating minimal genetic variation in those areas with intensive cucurbit cultivation (e.g., Taiwan and China). Non-cucurbit hosts showed more divergence, possibly due to cross-species transmission or adaptation to different ecological niches. Countries with isolated agricultural environments (e.g., Taiwan, China) exhibited more stable viral evolution, while regions with diverse agricultural practices (e.g., USA, France) showed greater genetic variability.

Countries with intensive cucurbit cultivation, such as Taiwan, China, and Israel, yield isolates with high identity values, suggesting that ZYMV strains in these regions have undergone limited genetic divergence. Conversely, regions with more diverse agricultural practices (e.g., Egypt and Brazil) have isolates with greater genetic variation.

More recent isolates from Taiwan (2024) and other regions (e.g., the UK and Brazil) have shown high identities with older strains from the same host species (e.g., Cucurbita pepo). However, isolates from non-cucurbit hosts (e.g., Aphis and Spiders) tend to show lower identities, suggesting that virus adaptation might occur more significantly in these non-cucurbit species.

In conclusion, ZYMV strains from cucurbits, especially Cucurbita pepo, show higher identity percentages, particularly in regions with a long history of cucurbit cultivation. In contrast, strains from countries like Egypt and Brazil display greater genetic diversity, highlighting the influence of host species and geographical location on virus evolution. The higher identity values observed in isolates from cucurbits, especially Cucurbita pepo and Cucumis sativus, suggest that the virus may exhibit greater genetic stability within these hosts.

In this study, the 5'UTR of ZYMV was highly conserved across different isolates (Figure 7), particularly in the first 60 nucleotides, reflecting its critical role in viral replication and translation initiation. Minor variations, mostly point mutations (e.g., A to T or C changes in isolates such as AJ307036.2, MT383108.1, and AM422386.1) were observed, but these did not significantly impact the region's function. These

variations were likely region-specific, as the isolates came from different geographic areas. Several strains, such as MF072712.1 and NC_003224.1, aligned perfectly with the LC799405.1 sequence, showing 100 % identity across 138 nucleotides. The overall high conservation of the 5' UTR suggested that it undergone evolutionary pressure to maintain its functionality which was essential for the viral life cycle. While the minor variations may provide insights into the viral evolution, they do not drastically alter the region's structure or function. This conservation makes the 5' UTR an ideal target for diagnostic tools like RT-PCR, though the small differences in some isolates, especially near positions 121-138, therefore, the diagnostic assays may require adjustments to accommodate new variations.

In the present study, the alignment of the 3' UTR region of the ZYMVgp1 (ZYMV-Egy-1920 and LC799416.2) with the 10 other ZYMV strains showed high conservation (Figure 8), especially in the central and the terminal regions (positions 180-213). Most strains aligned perfectly with LC799416.2, indicating strong sequence conservation typical for those regions involved in viral replication and translation. Minor nucleotide variations, such as G to A and G to T substitutions in isolates like LC314648.1, AF127931.1, and AF435425.1, did not significantly impact the sequence structure. These differences may reflect the regional or the strain-specific variations, but the core 3' UTR sequence remained highly conserved across the diverse geographic origins. This makes the 3' UTR a reliable target for molecular diagnostics, although continued surveillance of emerging ZYMV strains is needed to account for any new genetic variations in diagnostic assays.

```
LC799405.1   1    AAATTAAAACAAATCACAAAGACTACAGAAATCAACGAACAAGCAGACGATTTCTAAACC   60
MF072712.1   2    ............................................................   61
MF072713.1   2    ............................................................   61
MF072714.1   2    ............................................................   61
NC_003224.1  1    ............................................................   60
PP503025.1   1    ............................................................   60
AJ307036.2   2    ..................................................A..T......   61
MT383108.1   2    ..................................................A..T......   61
PP862811.1   2    ..................................................A..T......   61
AM422386.1   2    ..................................................A..T......   61
AJ316228.2   2    ................................................A.A..T......   61
LC799405.1   61   GTGTTTACAAACAAGCAATCTATAATTCGCACAGCATCAAGAATTTCTGCAATCACTTTG   120
MF072712.1   62   ............................................................   121
MF072713.1   62   ............................................................   121
MF072714.1   62   ............................................................   121
NC_003224.1  61   ............................................................   120
PP503025.1   61   ............................................................   120
AJ307036.2   62   ............................................................   121
MT383108.1   62   ............................................................   121
PP862811.1   62   ............................................................   121
AM422386.1   62   ............................A...............................   121
AJ316228.2   62   .........................C..................................   121
LC799405.1   121  TTTATTTTAGACACAACA   138
MF072712.1   122  ..................   139
MF072713.1   122  ..................   139
MF072714.1   122  ..................   139
NC_003224.1  121  ..................   138
PP503025.1   121  ..................   138
AJ307036.2   122  ..................   139
MT383108.1   122  ..................   139
PP862811.1   122  ..................   139
AM422386.1   122  ..................   139
AJ316228.2   122  ..................   139
```

Figure 7: *Sequences producing alignment shows the differences between 5'UTR region of ZYMVgp1 strain ZYMV-Egy-1920 (LC799405.1) isolated from Cucurbita pepo cv. Eskandarani and the most similar 10 ZYMV strains documented in GenBank. This figure highlights genetic differences between the viral isolates, revealing variations in amplified regions. These differences provide insights into the unique characteristics of each isolate, with implications for viral evolution, transmission, and host interactions.*

```
LC799416.2  1     TAAAGGGTAGGTCACCTACCTAGGTTATCGATTCGCTGCCGACGTAATTCTAATATTTAC  60
PP503025.1  9379  ............................................................  9438
NC_003224.1 9379  ............................................................  9438
LC314648.1  838   .............G..............................................  897
AF127931.1  894   ............................................................  953
AF435425.1  838   .............G..............................................  897
AJ316227.1  1471  .............G..............................................  1530
AY611026.1  973   .............G..............................................  1032
AJ316228.2  9380  .............G...................................T..........  9439
AY995216.1  1465  .............G..............................................  1524
AM422386.1  9380  .............G..............................................  9439
LC799416.2  61    CGCTTTTTATGATATCTTTAGATTTCAGTGTGGGCCTCCCACCTTTAAAGCGTAAAGTTT  120
PP503025.1  9439  ............................................................  9498
NC_003224.1 9439  ............................................................  9498
LC314648.1  898   ......A......G..............................................  957
AF127931.1  954   ......A......G..............................................  1013
AF435425.1  898   ......A......G.....................T........................  957
AJ316227.1  1531  ......A......G..........................C...................  1590
AY611026.1  1033  ......A......G..............................................  1093
AJ316228.2  9440  ......A......G..............................................  9499
AY995216.1  1525  ......A......G..............................................  1584
AM422386.1  9440  ......A......G.....................................A........  9499
LC799416.2  121   ATGTTAGTTGTCCAGGAATGCCGTAGTCCTGTCGGAAGCTTTAGTGTGAGCCTCTCACGG  180
PP503025.1  9499  ............................................................  9558
NC_003224.1 9499  ............................................................  9558
LC314648.1  958   .................G..........................................  1017
AF127931.1  1014  .................G.......................C.................A  1073
AF435425.1  958   .................G..........................................  1017
AJ316227.1  1591  .................G..........................................  1650
AY611026.1  1094  .................G..........................................  1153
AJ316228.2  9500  .................G..........................................  9559
AY995216.1  1585  .................G.........................................A  1644
AM422386.1  9500  .................G..........................................  9559
LC799416.2  181   ATAAGCTCGAGATTAGACTCCGTTTGCAAGCCT  213
PP503025.1  9559  .................................  9591
NC_003224.1 9559  .................................  9591
LC314648.1  1018  ................................   1049
AF127931.1  1074  .................................  1106
AF435425.1  1018  .................................  1050
AJ316227.1  1651  .................................  1683
AY611026.1  1154  .................................  1186
AJ316228.2  9560  .................................  9592
AY995216.1  1645  .................................  1677
AM422386.1  9560  ................................   9591
```

Figure 8: *Sequences producing alignment shows the differences between 3'UTR region of ZYMVgp1 strain ZYMV-Egy-1920 (LC799416.2) isolated from Cucurbita pepo cv. Eskandarani and the most similar 10 ZYMV strains documented in GenBank. This figure highlights genetic differences between the viral isolates, revealing variations in amplified regions. These differences provide insights into the unique characteristics of each isolate, with implications for viral evolution, transmission, and host interactions.*

The motifs identified in the ZYMVgp1 protein, as shown in Table 4, provided valuable insights into potential functional domains that may play critical roles in the virus's lifecycle. These motifs are categorized by their confidence levels: low, moderate, and high. Each category suggests varying degrees of biological significance, with high-confidence motifs being the most likely to play important roles in enzymatic activity, protein interactions, or other functional processes.

Table 4: *List of Alpha Fold-USERSEQ1-F1-v4 motif hits for the listed predicted structural sequences of ZYMVgp1 protein Egy-1920 strain (BFD45315.1).*

Low confidence motifs		
PS51216	Motif	YKRTYKKERK
	Score	4.759
PS51181	Motif	RCYSDIIYLVNVCLVFSLVLQMSNTVRNMiaatreekERAMANKADENERTLMHMYHIF
	Score	4.574
Moderate confidence motifs		
PS51871	Motif	QTVLAPLNSLCTRVLKIARNKNIPVEMIGNKKARHTLTFKRFRGYFVGKVSVAHEEGRMR (from 170 to 310)
	Score	21.232
PS51192	Motif	EIASSSEGEFLVRGAVGSGKSTSLPAHLAKKGKVLLLEPTRPLAENVSRQLAGDPFFQNVT... (from 1236-1388)
	Score	18.940
PS51194	Motif	DMVQHGNNI--LVYVASYNEVDMLSKLLTERQFSVTKVDGRTMQLGKTTIETHGTSQKPH...
	Score	13.980 (Moderate confidence)
PS50507	Motif	WLYCHADGSQFDSSLTPALLNAVLIIRSFYMEDWWVGQEMLENLyAEIVYTPILAPDGTI
	Score	(Moderate confidence)
High and very high confidence motifs		
PS51744	Motif	MYIAKEGYCYLNIFLAMLVNVNENEAKDFTKMIRdVLIPMLGQWPSLMDVATAAYILGVF
	Score	48.310
PS51436	Motif	SKSIYKGVRDYNGISTIVCQLTNDSDGLKETMYGIGYGPIIITNGHLFRKNNGTLLVRSW
	Score	70.067

Low-Confidence Motifs (e.g., PS51216 and PS51181) had been identified with low confidence, indicating that their predicted structure or biological relevance was uncertain. Although these motifs contained basic amino acids such as Lysines (K) and Arginines (R), which are often involved in protein-protein or protein-DNA interactions, however, their low confidence suggests that the match may not be reliable enough to confirm functional importance. While these regions could represent areas of interest, further experimental validation is necessary to assess their true biological roles.

Moderate-Confidence Motifs (e.g., PS51871, PS51192, and PS51194) suggest potentially important functional regions, but their significance requires further structural or experimental validation. They may represent binding sites or other functional domains, but additional evidence is needed to confirm their involvement in the virus's biological processes.

High-Confidence Motifs (e.g., PS51744 and PS51436) were highly confident and likely represented biologically significant regions of the protein. They were expected to be involved in several key functions such as enzymatic activity, protein interactions, or binding to host cell factors. Given their high confidence, these motifs should be prioritized for further experimental validation to confirm their essential roles in the viral lifecycle.

Finally, the current identification of these motifs within the ZYMVgp1 protein provides important insights into potential functional regions involved in replication, host cell interaction, and enzymatic activity. The high-confidence motifs, particularly PS51744 and PS51436, are of primary interest and should be prioritized for further validation due to their high likelihood of biological significance. While moderate- and low-confidence motifs are also important and they may require additional experimental investigation to fully elucidate their roles. Understanding these motifs will be crucial for developing diagnostic tools and therapeutic strategies aimed at targeting the virus.

The phylogenetic analysis based on the complete genome sequence of ZYMV-Egy-1920 (LC795783.1) and 17 other isolates illustrates the genetic diversity, regional adaptation, and host-specific variations of ZYMV. The clustering of isolates based on host species and geographic region reflects the evolutionary

Links Researchers

dynamics of the virus, with cucurbit strains exhibiting more genetic stability and non-cucurbit strains showing greater variability. These findings underscore the importance of monitoring the genetic diversity of ZYMV strains to enhance the diagnostic accuracy and the design targeted control strategies for ZYMV management, particularly in regions with intensive cucurbit cultivation.

The distance tree of the ZYMV polyprotein sequences (BFD45315.1) compared to the most similar ZYMV strains deposited in GenBank provided clear evidence of host-specific genetic variation, with isolates from cucurbit hosts were clustering closely together and exhibiting minimal genetic divergence (Figure 6). In contrast, isolates from non-cucurbit hosts showed greater genetic variability, suggesting that cross-species transmission or adaptation to alternative hosts may drive the observed divergence. The regional distribution and genetic divergence of ZYMV strains highlight the importance of local agricultural practices, host availability, and environmental factors in shaping viral evolution. These findings contribute to the understanding of ZYMV's genetic diversity, which is essential for the development of effective diagnostic tools and the control strategies for cucurbit crops.

The sequence of the HC-Pro protein of the ZYMV (BFF82029.1) as a multifunctional protein showed several conserved motifs associated with protease function, such as the Serine (S), Cysteine (C), and Histidine (H) residues that are part of the catalytic triad commonly found in viral proteases. The Serine (S) residue in the sequence likely played a key role in this catalytic activity, helping to cleave the polyprotein into functional units during virus replication. KLSC motif was found within the HC-Pro sequence and was typically involved in protein-protein interactions and RNA binding. The PTK motif (Proline-Threonine-Lysine) was another important region of HC-Pro, which played a role in viral function and host interaction.

The coat protein (CP) of the Egyptian strain of ZYMV, as identified through the BFF82037.1 sequence, was of 279 amino acids long. Analysis of this sequence revealed several features that are consistent with known functions of CPs, particularly in RNA binding and capsid assembly. The segment observed at positions 121-129 ("VVMNGFMVW") was a characteristic hydrophobic stretch commonly found in viral CPs involved in RNA encapsidation. The charged residues found in the N-terminal region (e.g., "AGATKKDKEDDK") were likely involved in protein-RNA interactions, playing a crucial role in stabilizing the viral RNA within the capsid. An important feature in this CP sequence was the DAG motif (Asp-Ala-Gly) near the N-terminal region. This motif is well-conserved in potyviruses and has been linked to viral replication and movement.

Discussion

Biological purification and molecular confirmation

In this study, ZYMV was confirmed in infected plant samples using various diagnostic methods. The virus was biologically purified *via* the single local lesion technique on *Chenopodium amaranticolor*, a common method for isolating and concentrating virus particles (Mahmoud *et al*., 2022). The virus was then propagated in squash plants (*Cucurbita pepo* cv. *Eskandarani*), which facilitated virus multiplication, aligning with previous reports on ZYMV accumulation in squash (Simmons *et al*., 2013). Electron microscopy confirmed the expected flexuous filamentous morphology of ZYMV particles, with clear presence observed 14 d post-inoculation (Desbiez and Lecoq, 1997). The 2 % uranyl acetate staining effectively highlighted the virus at the ultrastructural level (Milne, 1984).

Reverse transcription-polymerase chain reaction (RT-PCR) analysis confirmed ZYMV presence with an 837 bp product, consistent with the expected size for ZYMV RNA (Coutts *et al*., 2011). The absence of PCR amplification in the negative control ruled out contamination (Gao *et al*., 2012). The combination of biological purification, electron microscopy, and RT-PCR provided comprehensive confirmation of ZYMV, validating the methods used in this study. These results pave the way for further investigation into the virus's genetic diversity and management strategies for affected crops (Lim and Brown, 2018; Alinizi *et al*., 2021).

Primer design, genome amplification, and genetic diversity

The design and successful application of the fifteen primer pairs for amplifying the overlapping segments of the ZYMV genome underscore the efficacy of this approach for viral genome assembly and analysis. Overlapping primer design has proven to

be a robust method for obtaining full-length viral genomes, offering high accuracy and reliability in sequence assembly through redundancy in coverage. The amplified regions (ranging from 240 to 1266 bp) provide manageable fragment sizes suitable for sequencing, ensuring efficient handling in downstream applications. This strategy is particularly advantageous for RNA viruses like ZYMV, which exhibit high mutation rates and genetic variability (Simmons *et al*., 2013). Overlapping amplification minimizes sequencing errors and enables the detection of recombination events and mutations (Hughes and Hughes, 2007; Beerenwinkel *et al*., 2012). This capability is essential for studying genetic diversity and evolutionary dynamics (Domingo *et al*., 2012).

The complete genome sequence of the Egyptian ZYMV strain added valuable data to the growing ZYMV database, aiding future studies on variability influenced by geography and hosts, with implications for virus transmissibility, host range, and control strategies (Simmons *et al*., 2013). The primers developed here can identify mutation hotspots and conserved regions, supporting diagnostic and resistance-breeding efforts (Revers and García, 2015).

These data also advance the broader study of potyviruses, which share conserved genome features and replication strategies. Insights from ZYMV improve our understanding of related viruses affecting key crops (Adams *et al*., 2005; Lucas, 2006). Additionally, high-coverage genome sequences aid in viral evolution studies, resistance prediction, and antiviral strategies like RNA interference (Akbar *et al*., 2022).

The complete genome sequence of the Egyptian isolate of Zucchini Yellow Mosaic Virus (ZYMV) provided important insights into the genetic diversity and evolutionary dynamics of the virus. The high sequence identity with Mediterranean and Middle Eastern isolates indicated the potential spread of ZYMV across these regions, facilitated by aphid vectors (Desbiez and Lecoq, 1997; Lecoq *et al*., 2009). The conserved regions of the genome, particularly in the HC-Pro and CP proteins, confirmed the role of these proteins in aphid transmission and virus-host interactions, as previously described in other ZYMV isolates (Revers and García, 2015; Dong *et al*., 2022). Sequence variability in regions such as P1 and NIa may contribute to the adaptation of the virus to the different environmental conditions and host plants, as observed in various previous studies on potyviruses (Lee and Wong, 1998; Krause-Sakate *et al*., 2005).

This study highlights the importance of continuous surveillance and molecular characterization of ZYMV isolates to better understand the virus's epidemiology and to develop more effective control strategies. The genomic data provided will also aid in the design of diagnostic tools for rapid detection of ZYMV in cucurbit crops (Chung *et al*., 2013; Maghamnia *et al*., 2018).

In this study, the single local lesion technique was successfully used to purify identify and characterize plant viruses. It involved inoculating plant tissues with a virus and observing the formation of localized lesions, typically necrotic spots, which resulted from a hypersensitive response to the viral infection. By analyzing the number, size, and appearance of these lesions, researchers can infer the virus's infectivity, virulence, and genetic diversity (Mahmoud *et al*., 2022).

The observed genetic divergence among ZYMV isolates in this study, particularly between cucurbit and non-cucurbit hosts, is consistent with the recent findings on the role of host specificity in shaping the genetic diversity of Potyviruses. Cucurbits, such as Cucurbita pepo and Cucumis sativus, exhibit more stable virus populations with lower genetic variation in regions with intensive cucurbit cultivation, such as Taiwan and China (Rabadán *et al*., 2023). This stability is likely due to the high prevalence of ZYMV in these regions, leading to limited genetic divergence of the strains.

The high identity percentages found in this study (up to 99.81 %) in isolates from cucurbit species like Cucurbita pepo and Luffa cylindrica in Taiwan further supported the idea that ZYMV strains were genetically stable in regions with long histories of cucurbit cultivation (Dunham *et al*., 2014).

In contrast, non-cucurbit hosts such as Aphis (Iran, 2013) and Spiders (China, 2013) exhibited lower identities (93.71 % and 92.83 %, respectively). This suggests more significant evolutionary divergence in strains infecting non-cucurbit hosts, which is consistent with findings from recent studies showing that Potyviruses exhibit higher genetic variability

when infecting alternative hosts. Such divergence is often attributed to cross-species transmission or adaptation to novel ecological niches (McLeish *et al*., 2022). The divergence observed in ZYMV strains from non-cucurbit hosts may be attributed to the virus's ability to evolve in response to the unique pressures of different hosts or ecological environments (Nishimura *et al*., 2019).

The geographic distribution of ZYMV strains, particularly the divergence observed between Egyptian isolates and those from other regions is in line with recent work on Potyviruses evolution. Isolates from Egypt, including ZYMV-Egy-1920, showed moderate divergence from Taiwanese and Chinese strains, with identity values ranging from 92.61 % to 95.78 % at the nucleotide level. This finding suggests some regional adaptation or evolutionary divergence of strains circulating in Egypt (Abdel Aleem *et al*., 2021). As reported by Spadotti *et al*. (2015), such regional variations in ZYMV strains are likely influenced by local agricultural practices, host species availability, and environmental factors that shape viral evolution.

The lower identity values of Egyptian isolates (95.78 % identity to a zucchini isolate from Egypt in 2018 and 95.68 % identity to Cucurbita pepo from Egypt in 2021) could reflect the unique evolutionary pressures in the Egyptian's agricultural environment, which may differ from those in other regions such as Taiwan and China, where cucurbit cultivation is more widespread and intensively managed (Ghanem *et al*., 2023). Moreover, these regional differences in genetic diversity may indicate that viral strains in Egypt and similar regions have been subject to more rapid evolutionary changes due to multiple factors, including varying host species, vector populations, or ecological niches (Alhajjar *et al*., 2020).

In terms of host range, this study confirmed that cucurbits, particularly Cucurbita pepo and Cucumis sativus, are the primary hosts of ZYMV, with high identity values observed in isolates from these species, particularly in regions with long histories of cucurbit cultivation. This observation supports the findings of recent studies indicating that Potyviruses exhibit limited genetic divergence within their primary hosts due to stable virus-host interactions over time (Dunham *et al*., 2014). Non-cucurbit hosts, however, show more genetic variation which may be attributed to cross-species transmission, suggesting that ZYMV may evolve more rapidly when transmitted by alternative hosts or vectors (Rabadán *et al*., 2023).

Countries with intensive cucurbit cultivation, such as Taiwan, China, and Israel, tend to have isolates with high identity percentages, suggesting that ZYMV strains in these regions have undergone limited genetic divergence (Desbiez, 2019). This is in contrast to regions with diverse agricultural practices and broader host species distributions, such as Egypt and Brazil, where isolates exhibit greater genetic variability (Spadotti *et al*., 2015). More recent isolates from Taiwan (2024) and other regions, such as the UK and Brazil, demonstrate high identity with older strains from the same host species (Cucurbita pepo), reinforcing the idea that viral evolution in stable agricultural environments leads to less genetic divergence over time (Ghanem *et al*., 2023).

The 5' un-translated region (UTR) of ZYMV was highly conserved across different isolates, reflecting its crucial role in viral replication and translation (Thurner *et al*., 2004; Chujo *et al*., 2015; Zhang *et al*., 2024). Our comparison of the 5' UTR of the ZYMVgp1 strain (ZYMV-Egy-1920, LC799405.1) with 17 other ZYMV isolates showed strong sequence identity, with 100 % match in several strains. This high conservation supports the use of the 5' UTR as a reliable target for molecular diagnostics, including RT-PCR assays, as it is stable across different strains (Thurner *et al*., 2004; Luigi *et al*., 2023).

While most strains exhibited minimal genetic variation, slight differences particularly in strains OQ847411.1and KX499498.1 may reflect viral adaptation or mutation, which could affect host specificity or pathogenicity (Syller, 2005). These variations highlight the importance of ongoing surveillance and the need for updated diagnostic tools to account for minor strain differences (Kuan *et al*., 2014).

Analysis of 3' UTR region

The alignment of the 3' UTR region of ZYMVgp1 (ZYMV-Egy-1920 and LC799416.2) with 10 other ZYMV strains showed high conservation, particularly in the central and terminal segments (positions 180-213). Minor nucleotide variations, such as G to A and G to T substitutions in strains like LC314648.1, AF127931.1, and AF435425.1, reflect geographic

or host-specific adaptations but do not disrupt the essential functions of the 3' UTR (Luigi *et al*., 2023). These minor variations are consistent with previous studies on plant viruses, which showed that such changes can affect viral fitness and host interactions (Chujo *et al*., 2015). The high conservation of this region supports its use in molecular diagnostics, as it remains stable across isolates. However, the observed genetic differences underscore the importance of ongoing surveillance to ensure diagnostic tools remain accurate (Thurner *et al*., 2004). Regular updates to diagnostic protocols may be needed to account for emerging variations in the virus genome.

The motifs identified in the ZYMVgp1 protein were categorized by their confidence levels-low, moderate, and high which suggest varying degrees of reliability regarding their biological significance. Understanding the functional relevance of these motifs can aid in identifying new therapeutic targets, as viral proteins often contain regions crucial for replication, host interaction, and enzymatic activity (Navarro *et al*., 2020).

The low-confidence motifs, such as PS51216 and PS51181, suggested that the corresponding regions may have functional relevance, but their biological importance is uncertain. Recent studies have highlighted that low-confidence predictions may sometimes overlook important motifs due to issues such as incomplete structural data or ambiguous sequence alignment (Sobhy, 2016). In some cases, motifs enriched in basic amino acids (e.g., lysines and arginines) can mediate protein-protein or protein-DNA interactions, yet low-confidence predictions may fail to identify these motifs accurately (Zheng *et al*., 2023). Therefore, these regions should be considered as potential sites for further investigation, particularly through experimental validation, such as mutagenesis or binding assays, to establish their role in viral function.

Moderate-confidence motifs, such as PS51871, PS51192, and PS51194, indicated areas with potential biological importance, but additional structural or experimental validation is needed to fully confirm their functional roles. This is consistent with the findings of recent studies, where moderate-confidence motifs often correlate with regions involved in binding or enzymatic activity, though their precise function may remain unclear without empirical support (Zhang *et al*., 2017). For example, motifs involved in host cell recognition or interaction with replication machinery may display moderate confidence scores due to inherent sequence variability across different virus strains.

High-confidence motifs, such as PS51744 and PS51436 were strongly predicted to be biologically significant regions, likely involved in key functions including enzymatic activity, protein interactions, or binding to host cell factors. High-confidence motifs have been well-documented in recent studies to play crucial roles in the viral lifecycle. For instance, the presence of conserved motifs within replication-associated proteins often correlates with essential enzymatic functions such as polymerase activity or RNA binding (Venkataraman *et al*., 2018). These motifs should be prioritized for further experimental validation, particularly through techniques such as co-immunoprecipitation (Co-IP) and mass spectrometry to identify binding partners, or functional assays to confirm their roles in viral replication and host interaction.

Recent work has highlighted the importance of such motifs in the development of antiviral strategies, where targeting key regions of viral proteins can inhibit replication and disrupt the virus-host interaction (Romero-López and Berzal-Herranz, 2013). However, moderate- and low-confidence motifs also warrant further study, as they could represent previously overlooked functional regions that contribute to viral pathogenicity.

In conclusion, while the high-confidence motifs identified in the ZYMVgp1 protein provided compelling targets for therapeutic development, but the full understanding of this protein's functionality requires further experimental validation across the spectrum of confidence levels. Future studies should aim to verify the biological roles of these motifs, providing deeper insights into viral mechanisms and contributing to the design of effective diagnostic tools and antiviral therapies (Mehetre *et al*., 2023).

Phylogenetic analysis of the complete genome

The phylogenetic analysis of the complete genome sequence of ZYMV strain ZYMV-Egy-1920 (LC795783.1) and 17 other isolates revealed significant genetic diversity, regional adaptation, and host-specific variations within the virus. The clustering

Links Researchers

of isolates based on host species and geographic region highlighted the stable viral populations in cucurbit crops, where strains exhibited minimal genetic divergence, in consistence with findings in regions with intensive cucurbit cultivation (Shrestha *et al*., 2021). This is particularly evident in isolates from Cucurbita pepo and Cucumis sativus, which formed a closely related clade with high sequence identity (up to 99.81 %).

In contrast, isolates from non-cucurbit hosts, such as aphids and spiders, showed greater genetic variability (93.71-92.83%), suggesting that cross-species transmission or adaptation to novel hosts may drive the observed divergence (McLeish *et al*., 2022). This finding aligns with previous studies indicating that viral evolution was often influenced by host shifts, which can lead to greater genetic diversity (Desbiez *et al*., 2002). The divergence observed in these non-cucurbit isolates underscores the potential role of alternative hosts in shaping viral genetic evolution, as also observed in other potyviruses (Nishimura *et al*., 2019).

Moreover, regional differences in ZYMV strains highlighted the influence of local agricultural practices, host species availability, and environmental factors on viral evolution (Abdel Aleem *et al*., 2021). These insights are crucial for improving the design of diagnostic tools and developing region-specific control strategies to mitigate the impact of ZYMV on cucurbit crops (Alinizi *et al*., 2021; Ghanem *et al*., 2023). Monitoring the genetic diversity of ZYMV strains is essential for effective management and prevention strategies, especially in areas with high agricultural intensification.

The current Egyptian strain CP sequence of the ZYMV (BFF82037.1) showed several structural motifs consistent with the functional roles of the CP in RNA encapsidation, vector transmission, and plant cell-to-cell movement. The DAG motif in particular may contribute significantly to the virus's ability to spread within the host, reinforcing the potential importance of this sequence in the pathogenesis of ZYMV. Hydrophobic regions and RNA encapsidation in the coat protein (CP) facilitated the packaging of the viral RNA into the capsid (Desbiez *et al*., 2002; Gal-On, 2007). In addition, charged residues and protein-RNA interactions found in the N-terminal region were rich in basic residues (*i.e.*, lysines and arginines) that are essential for binding the negatively charged RNA genome, stabilizing it within the capsid during viral assembly (Zhang *et al*., 2017).

Molecular insights into zucchini yellow mosaic virus

The conserved DAG motif located near the N-terminus played a key role in viral replication and movement. This motif is essential for facilitating virus trafficking through plasmodesmata, enabling the cell-to-cell movement and overcoming the host resistance mechanisms (Velasco *et al*., 2020). Finally, this CP was reported to interact with aphid vectors during transmission, facilitating the acquisition and inoculation of the virus into new plants (Gadhave *et al*., 2020), and may also contribute to these interactions, supporting the virus's spread *via* aphids. Overall, The Egyptian strain CP sequence displayed key features that contribute to viral stability, cell-to-cell movement, and vector transmission, with the DAG motif playing a significant role in the virus's ability to spread within the host.

The HC-Pro protein of ZYMV played a critical role in the virus's replication, movement, and interaction with the host. HC-Pro is involved in viral polyprotein processing, RNA binding, and suppression of host defense mechanisms (Velasco *et al*., 2020; Hýsková *et al*., 2024). Several conserved motifs within the HC-Pro sequence, including the catalytic triad (Serine, Cysteine, and Histidine) and functional motifs such as KLSC and PTK, contributed to its complex functions.

One of the most significant roles of HC-Pro is its protease activity. The catalytic triad is essential for viral polyprotein cleavage, which is necessary for viral maturation and replication. These residues are involved in nucleophilic attack and stabilization of the transition state during proteolysis, a mechanism that is widely conserved in viral proteases (Syller, 2005; Adams *et al*., 2007). The Serine (S) residue in particular was critical in the enzymatic mechanism, acting as a nucleophile that facilitated peptide bond cleavage (Wang *et al*., 2000).

Recent studies have demonstrated that the KLSC motif in HC-Pro is critical for the interaction between HC-Pro and other viral proteins, as well as host proteins that assist in viral movement and replication (Do *et al*., 2023). The KLSC motif may enhance the stability of the viral RNA and participate in suppressing the host RNA-silencing mechanisms,

Links Researchers

which are critical for the antiviral defense in plants (Adams *et al*., 2007). Moreover, several interactions mediated by the KLSC motif likely play a role in the spread of the virus through the plant vasculature, facilitating the systemic infection (Wang *et al*., 2000; Dombrovsky *et al*., 2014).

The PTK motif is another important region of HC-Pro, contributing to the viral-host interaction, particularly in the manipulation of host immune responses. This motif helps HC-Pro to suppress RNA silencing by interfering with the host's antiviral defense mechanisms (Gal-On, 2007). The lysine residue in this motif may engage in electrostatic interactions with host cell proteins, possibly modifying host cellular pathways to favor viral replication (Gao *et al*., 2012). The PTK motif has also been implicated in the virus's ability to modulate host signaling pathways, facilitating viral spread and immune evasion (Gal-On, 2007).

Conclusions and Recommendations

Zucchini yellow mosaic virus (ZYMV) was confirmed through biological purification, electron microscopy, and RT-PCR, validating its presence in infected plant samples. The complete genome sequence of the Egyptian isolate revealed high sequence identity with the Mediterranean and the Middle Eastern strains; highlighting the viral potential spread *via* the aphid vectors. Conserved genomic regions, including HC-Pro and CP proteins, are crucial for viral replication, host interaction, and aphid transmission. Genetic variability among ZYMV isolates, particularly in cucurbit and non-cucurbit hosts, underscores the virus's adaptive evolution in response to host-specific and environmental pressures. Motif analysis in viral proteins, such as the HC-Pro catalytic triad and the CP DAG motif, revealed their essential roles in viral pathogenesis, including RNA binding, polyprotein processing, and vector transmission. This study emphasizes the importance of continuous surveillance and molecular characterization for effective control strategies. The findings of this study contribute to the design of reliable diagnostic tools and the identification of potential therapeutic targets for ZYMV management.

Overall, this study provides a comprehensive analysis of an Egyptian isolate of ZYMV, focusing on its complete genomic characterization, genetic features associated with aphid transmission, and phylogenetic relationships with global ZYMV strains. The research highlights the molecular diversity within the virus, offering insights into key genetic elements that facilitate transmission and host adaptation. By comparing the Egy-1920 strain with other ZYMV strains worldwide, this study identifies significant genetic homology, particularly with Taiwanese and Chinese strains, contributing to the understanding of ZYMV's global spread. The findings from this study lay the groundwork for the development of more accurate diagnostic techniques and provide valuable information for potential therapeutic strategies aiming at controlling ZYMV, a significant pathogen in crop production.

Acknowledgment

We would like to express our sincere gratitude to Dr. Hassan Abdel Fattah (Emeritus Professor of Agricultural Microbiology, Department of Agricultural Microbiology, Faculty of Agriculture, Zagazig University) and Dr. Medhat Kamel Ali (Emeritus Professor of Plant Pathology, Department of Plant Pathology, Faculty of Agriculture, Ain Shams University) for their invaluable support in saving chemicals and equipment for virus purification in this study. Their expertise, insightful feedback, and continuous assistance have significantly contributed to the success and quality of this work.

Novelty Statement

This study presents the first complete genome sequence of the Zucchini yellow mosaic virus (ZYMV) strain from the *Eskandarani* squash cultivar, providing valuable insights into its genetic composition and the identification of key aphid-transmission motifs in the helper component (HC) protein. The high nucleotide and protein sequence identities observed compared to other global ZYMV strains, particularly Taiwanese and Chinese strains, highlight the distinctiveness of the Egy-1920 strain. This study also expands the understanding of the genome length variation among the ZYMV strains, which has important implications for the development of novel diagnostic tools and therapeutic strategies for the ZYMV control.

Author's Contribution

Fatma S. Abdel Razek: Conceptualization, designed

the research, Investigation, Methodology, Data analysis, writing of the original draft and review and editing.
Ahmed Mahdy: Research design, investigation, and methodology, Data analysis and review and editing.
Samar S.A. El-Masry: Research design, Methodology. Data analysis, validation of results, writing the original draft and editing.
Shafik D. Ibrahim: Data analysis, writing the original draft and editing.
Shrouk E.E. Farg: Methodology and validation of results.
Atef S. Sadik: Supervision, validation of results and review and editing.

Ethical approval
This study was conducted in accordance with ethical guidelines, and no human or animal subjects were involved. Ethical approval was not required for this research.

Funding source
This research was conducted as part of a Ph.D. project and did not receive any external funding.

Conflict of interests
The authors have declared no conflicts of interest.

References

Abdel Aleem, E.E.A., Rabie, M. and Fattouh, F.A., 2021. Molecular characterisation of Zucchini yellow mosaic virus infecting *Cucurbita pepo* in Egypt. Plant Prot. Sci., 57(4): 263-270. https://doi.org/10.17221/191/2020-PPS

Adams, M.J., Antoniw, J.F. and Beaudoin, F., 2007. Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. Mol. Plant Pathol., 6(4): 471-487. https://doi.org/10.1111/j.1364-3703.2005.00296.x

Adams, M.J., Antoniw, J.F. and Fauquet, C.M., 2005. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch. Virol., 150(3): 459-479. https://doi.org/10.1007/s00705-004-0440-6

Ahsan, M., Ashfaq, M., Amer, M.A., Shakeel, M.T., Mehmood, M.A., Umar, M. and Al-Saleh, M.A., 2023. Zucchini yellow mosaic virus (ZYMV) as a serious biotic stress to cucurbits: Prevalence, diversity, and its implications for crop sustainability. Plants, 12(19): 3503. https://doi.org/10.3390/plants12193503

Akbar, S., Wei, Y. and Zhang, M.Q., 2022. RNA interference: Promising approach to combat plant viruses. Int. J. Mol. Sci., 23(10): 5312. https://doi.org/10.3390/ijms23105312

Alhajjar, K., Ibrahim, M., Khattab, S., El-Sayed, M. and Ahmed, F., 2020. Host range and genetic diversity of Zucchini yellow mosaic virus strains. Plant Dis., 104(10): 2407-2415.

Ali, A. and Kobayashi, M., 2010. Seed transmission of cucumber mosaic virus in pepper. J. Virol. Methods, 163(2): 234-237. https://doi.org/10.1016/j.jviromet.2009.09.026

Alinizi, H.R., Mehrvar, M. and Zakiaghl, M., 2021. Analysis of the molecular and biological variability of Zucchini yellow mosaic virus isolates from Iran and Iraq. Gene, 788: 145674. https://doi.org/10.1016/j.gene.2021.145674

Antoniw, J.F. and White, R.F., 1986. Changes with time in the distribution of virus and PR protein around single local lesions of TMV infected tobacco. Plant Mol. Biol., 6: 145-149. https://doi.org/10.1007/BF00021483

Beerenwinkel, N., Günthard, H.F., Roth, V. and Metzner, K.J., 2012. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Front. Microbiol., 3: 329. https://doi.org/10.3389/fmicb.2012.00329

Bernet, G.P. and Elena, S.F., 2015. Distribution of mutational fitness effects and of epistasis in the 5'untranslated region of a plant RNA virus. BMC Evol. Biol., 15: 1-13. https://doi.org/10.1186/s12862-015-0555-2

Brakke, M.K., 1961. Density gradient centrifugation and its application to plant viruses. Adv. Virus Res. Acad. Press, 7: 193-224. https://doi.org/10.1016/S0065-3527(08)60011-1

Brenner, S. and Horne, R.W., 1959. A negative staining method for high resolution electron microscopy of viruses. Biochim. Biophys. Acta, 34(1959): 103–110. https://doi.org/10.1016/0006-3002(59)90237-9

Bubici, G., Navarro, B., Carluccio, A.V., Ciuffo, M., Di Serio, F. and Cillo, F., 2020. Genomic sequence variability of an Italian Zucchini yellow mosaic virus isolate. Eur. J. Plant Pathol., 156(1): 325-332. https://doi.org/10.1007/s10658-019-01886-w

Chinnadurai, C., Kollam, M., Ramsubhag, A. and Jayaraman, J., 2021. Genome characterization

Links Researchers

of zucchini yellow mosaic virus infecting cucurbits reveals the presence of a new genotype in Trinidad and Tobago in the Caribbean region. Arch. Virol., 166(6): 1661-1669. https://doi.org/10.1007/s00705-021-05048-4

Choi, S.K., Yoon, J.Y. and Sohn, S.H., 2007. Analysis of the complete genome sequence of Zucchini yellow mosaic virus strain A isolated from Hollyhock. Plant Pathol. J., 23(4): 245-250. https://doi.org/10.5423/PPJ.2007.23.4.245

Chujo, T., Ishibashi, K., Miyashita, S. and Ishikawa, M., 2015. Functions of the 5′- and 3′-untranslated regions of Tobamovirus RNA. Virus Res., 206: 82-89. https://doi.org/10.1016/j.virusres.2015.01.028

Chung, B.N., Yoon, J.Y. and Palukaitis, P., 2013. Engineered resistance in potato against potato leafroll virus, potato virus A and potato virus Y. Virus Genes, 47: 86-92. https://doi.org/10.1007/s11262-013-0904-4

Chung, B.Y.W., Miller, W.A., Atkins, J.F. and Firth, A.E., 2008. An overlapping essential gene in the Potyviridae. Proc. Natl. Acad. Sci., 105(15): 5897-5902. https://doi.org/10.1073/pnas.0800468105

Coutts, B.A., Kehoe, M.A., Webster, C.G., Wylie, S.J. and Jones, R.A.C., 2011. Zucchini yellow mosaic virus: biological properties, detection procedures and comparison of coat protein gene sequences. Arch. Virol., 156: 2119-2131. https://doi.org/10.1007/s00705-011-1102-0

Desbiez, C. and Lecoq, H., 1997. Zucchini yellow mosaic virus. Plant Pathol., 46(6): 809-829. https://doi.org/10.1046/j.1365-3059.1997.d01-87.x

Desbiez, C., 2019. The never-ending story of cucurbits and viruses. VI Int. Symp. Cucurbits, 1294: 173-192. https://doi.org/10.17660/ActaHortic.2020.1294.23

Desbiez, C., Wipf-Scheibel, C. and Lecoq, H., 2002. Biological and serological variability, evolution and molecular epidemiology of Zucchini yellow mosaic virus (ZYMV, Potyvirus) with special reference to Caribbean islands. Virus Res., 85(1): 5-16. https://doi.org/10.1016/S0168-1702(02)00013-8

Do, D.H., Nguyen, T.B.N., Ha, V.C., Raja, J.A. and Yeh, S.D., 2023. Generation of attenuated Passiflora mottle virus through modification of the helper component protease for cross protection. Phytopathology, 113(8): 1605-1614. https://doi.org/10.1094/PHYTO-01-23-0007-R

Dombrovsky, A., Reingold, V. and Antignus, Y., 2014. Ipomovirus–an atypical genus in the family Potyviridae transmitted by whiteflies. Pest Manage. Sci., 70(10): 1553-1567. https://doi.org/10.1002/ps.3735

Domingo, E., Sheldon, J. and Perales, C., 2012. Viral quasispecies evolution. Microbiol. Mol. Biol. Rev., 76(2): 159–216. https://doi.org/10.1128/MMBR.05023-11

Dong, Z.X., Lin, C.C., Chen, Y.K., Chou, C.C. and Chen, T.C., 2022. Identification of an emerging cucumber virus in Taiwan using Oxford nanopore sequencing technology. Plant Methods, 18(1): 143. https://doi.org/10.1186/s13007-022-00976-x

Dunham, J.P., Simmons, H.E., Holmes, E.C. and Stephenson, A.G., 2014. Analysis of viral (zucchini yellow mosaic virus) genetic diversity during systemic movement through a *Cucurbita pepo* vine. Virus Res., 191: 172-179. https://doi.org/10.1016/j.virusres.2014.07.030

Gadhave, K.R., Gautam, S., Rasmussen, D.A. and Srinivasan, R., 2020. Aphid transmission of Potyvirus: The largest plant-infecting RNA virus genus. Viruses, 12(7): 1-22. https://doi.org/10.3390/v12070773

Gal-On, A., 2007. Zucchini yellow mosaic virus: Insect transmission and pathogenicity-the tails of two proteins. Mol. Plant Pathol., 8(2): 139-150. https://doi.org/10.1111/j.1364-3703.2007.00381.x

Gao, W., Wang, X. and Zhang, L., 2012. Molecular diagnostic methods for plant viruses. J. Plant Pathol., 94(2): 281–290.

Ghanem, G.A., Mahmoud, A.M., Kheder, A.A., Gebily, D.A. and Osamy, A., 2023. Antiviral activities of three *Streptomyces* spp. against Zucchini yellow mosaic virus (ZYMV) infecting squash (*Cucurbita pepo* L.) plants. Egypt. J. Biol. Pest Contr., 33(1): 113. https://doi.org/10.1186/s41938-023-00750-8

Glasa, M. and Pittnerová, S., 2006. Complete genome sequence of a Slovak isolate of Zucchini yellow mosaic virus (ZYMV) provides further evidence of a close molecular relationship among central European ZYMV isolates. J. Phytopathol., 154(7-8): 436-440. https://doi.org/10.1111/j.1439-0434.2006.01124.x

Hughes, A.L. and Hughes, M.A., 2007. More

effective purifying selection on RNA viruses than in DNA viruses. Gene, 404(1-2): 117–125. https://doi.org/10.1016/j.gene.2007.09.013

Hýsková, V., Bělonožníková, K., Chmelík, J., Hoffmeisterová, H., Čeřovská, N., Moravec, T. and Ryšlavá, H., 2024. Potyviral helper-component protease: multifaced functions and interactions with host proteins. Plants, 13(9): 1-21. https://doi.org/10.3390/plants13091236

Katis, N.I., Tsitsipis, J.A., Lykouressis, D.P., Papapanayotou, A., Margaritopoulos, J.T., Kokinis, G.M., Perdikis, D.C. and Manoussopoulos, I.N., 2006. Transmission of Zucchini yellow mosaic virus by colonizing and non-colonizing aphids in Greece and new aphid species vectors of the virus. J. Phytopathol., 154(5): 293-302. https://doi.org/10.1111/j.1439-0434.2006.01096.x

Kneller, E.L.P., Rakotondrafara, A.M. and Miller, W.A., 2006. Cap-independent translation of plant viral RNAs. Virus Res., 119(1): 63-75. https://doi.org/10.1016/j.virusres.2005.10.010

Krause-Sakate, R., Zerbini, F.M., Brandon, R., Moreira, A. and Gonsalves, D., 2005. Analysis of the genetic diversity of Zucchini Yellow Mosaic Virus in Brazil. J. Gen. Virol., 86(12): 3557-3566.

Kuan, C.P., Deng, T.C., Huang, H.C., Chi, H.H. and Lu, Y.L., 2014. Use of reverse transcription loop-mediated isothermal amplification for the detection of Zucchini yellow mosaic virus. J. Phytopathol., 162(4): 238-244. https://doi.org/10.1111/jph.12175

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. MEGA X; Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol., 35(1): 1547-1549. https://doi.org/10.1093/molbev/msy096

Lecoq, H. and Desbiez, C., 2012. Viruses of cucurbit crops in the mediterranean region: An ever-changing picture. Adv. Virus Res., 84: 67-126. https://doi.org/10.1016/B978-0-12-394314-9.00003-8

Lecoq, H., Wipf-Scheibel, C., Chandeysson, C., Le Van, A., Fabre, F. and Desbiez, C., 2009. Molecular epidemiology of Zucchini yellow mosaic virus in France: An historical overview. Virus Res., 141(2): 190-200. https://doi.org/10.1016/j.virusres.2008.11.020

Lee, J. and Wong, F., 1998. Complete genome sequence of Zucchini yellow mosaic virus isolate from Singapore (ZYMV-S). Phytopathology, 88(10): 1001-1007. https://link.springer.com/article/10.1007/s13205-018-1177-3.

Lim, C.S. and Brown, C.M., 2018. Know your enemy: Successful bioinformatic approaches to predict functional RNA structures in viral RNAs. Front. Microbiol., 8: 2582. https://doi.org/10.3389/fmicb.2017.02582

Lucas, W.J., 2006. Plant viral movement proteins: Agents for cell-to-cell trafficking of viral genomes. Virology, 344(1): 169-184. https://doi.org/10.1016/j.virol.2005.09.026

Luigi, M., Manglli, A., Corrado, C.L., Tiberini, A., Costantini, E., Ferretti, L., Tomassoli, L. and Bertin, S., 2023. Development, validation, and application of reverse transcription Real-Time and droplet digital PCR assays for the detection of the Potyviruses watermelon mosaic virus and zucchini yellow mosaic virus in cucurbits. Plants, 12(12): 2364. https://doi.org/10.3390/plants12122364

Maghamnia, M., Shafiei, R. and Ahmad, J., 2018. Genome characterization of a Zucchini yellow mosaic virus strain infecting squash in Kurdistan. Mol. Plant Pathol., 19(4): 938-948. https://link.springer.com/article/10.1007/s13205-018-1177-3.

Mahmoud, S.Y., Farrag, E.S., Amer, M.A., Wardany, A. and Abdel-Haliem, M.E., 2022. Biological and molecular characterization of a distinct two Zucchini yellow mosaic potyvirus isolates from Upper Egypt. https://doi.org/10.21203/rs.3.rs-1987595/v1

McLeish, M.J., Zamfir, A.D., Babalola, B.M., Peláez, A., Fraile, A. and García-Arenal, F., 2022. Metagenomics show high spatiotemporal virus diversity and ecological compartmentalisation: Virus infections of melon, Cucumis melo, crops, and adjacent wild communities. Virus Evol., 8(2): veac095. https://doi.org/10.1093/ve/veac095

Mehetre, G.T., Leo, V.V., Singh, G., Sorokan, A., Maksimov, I., Yadav, M.K., Kalidas Upadhyaya, K., Hashem, A., Asma N. Alsaleh, A.N., Dawoud, T.M., Almaary, K.S. and Singh, B.P., 2021. Current developments and challenges in plant viral diagnostics: A systematic review. Viruses, 13(3): 412. https://doi.org/10.3390/v13030412

Milne, R.G., 1984. Electron microscopy for the identification of plant viruses in *in vitro*

preparations. Methods Virol. Elsevier, 7: 87-120. https://doi.org/10.1016/B978-0-12-470207-3.50010-3

Navarro, J.A., Serra-Soriano, M., Corachán-Valencia, L. and Pallás, V., 2020. A conserved motif in three viral movement proteins from different genera is required for host factor recruitment and cell-to-cell movement. Sci. Rep., 10(1): 4758. https://doi.org/10.1038/s41598-020-61741-5

Nishimura, S., Yamada, K., Ito, S., Fujita, M. and Taniguchi, T., 2019. Genetic diversity of Zucchini yellow mosaic virus strains from non-cucurbit hosts. J. Gen. Virol., 100(2): 142-152.

Perotto, M.C., Pozzi, E.A., Celli, M.G., Luciani, C.E., Mitidieri, M.S. and Conci, V.C., 2018. Identification and characterization of a new potyvirus infecting cucurbits. Arch. Virol., 163(3): 719-724. https://doi.org/10.1007/s00705-017-3660-2

Rabadán, M.P., Juárez, M. and Gómez, P., 2023. Long-term monitoring of aphid-transmitted viruses in melon and zucchini crops: Genetic diversity and population structure of cucurbit aphid-borne yellows virus and watermelon mosaic virus. Phytopathology, 113(9): 1761-1772. https://doi.org/10.1094/PHYTO-10-22-0394-V

Revers, F. and García, J.A., 2015. Molecular biology of potyviruses. Adv. Virus Res., 92: 101-199. https://doi.org/10.1016/bs.aivir.2014.11.006

Romero-López, C. and Berzal-Herranz, A., 2013. Unmasking the information encoded as structural motifs of viral RNA genomes: A potential antiviral target. Rev. Med. Virol., 23(6): 340-354. https://doi.org/10.1002/rmv.1756

Sharma, P., 2023. Epidemiology of potyviruses infecting crops of Cucurbitaceae. In Plant RNA Viruses. Academic Press. pp. 213-227. https://doi.org/10.1016/B978-0-323-95339-9.00016-8

Shrestha, S., Michael, V.N.E., Fu, Y. and Meru, G., 2021. Genetic loci associated with resistance to zucchini yellow mosaic virus in squash. Plants, 10(9): 1935. https://doi.org/10.3390/plants10091935

Simmons, H.E., Dunham, J.P., Zinn, K.E., Munkvold, G.P., Holmes, E.C. and Stephenson, A.G., 2013. Zucchini yellow mosaic virus (ZYMV, Potyvirus): Vertical transmission, seed infection and cryptic infections. Virus Res., 176(1-2): 259-264. https://doi.org/10.1016/j.virusres.2013.06.016

Sobhy, H., 2016. A review of functional motifs utilized by viruses. Proteomes, 4(1): 3. https://doi.org/10.3390/proteomes4010003

Spadotti, D.M.D.A., Wassano, D.T., Rezende, J.A.M., Camargo, L.E.A. and Inoue-Nagata, A.K., 2015. Biological and molecular characterization of Brazilian isolates of Zucchini yellow mosaic virus. Sci. Agricola, 72(2): 187-191. https://doi.org/10.1590/0103-9016-2014-0197

Syller, J., 2005. The roles and mechanisms of helper component proteins encoded by potyviruses and caulimoviruses. Physiol. Mol. Plant Pathol., 67(3-5): 119-130. https://doi.org/10.1016/j.pmpp.2005.12.005

Tatineni, S. and Hein, G.L., 2023. Plant viruses of agricultural importance: Current and future perspectives of virus disease management strategies. Phytopathology, 113(2): 117-141. https://doi.org/10.1094/PHYTO-05-22-0167-RVW

Thurner, C., Witwer, C., Hofacker, I.L. and Stadler, P.F. 2004. Conserved RNA secondary structures in Flaviviridae genomes. J. Gen. Virol., 85(5): 1113-1124. https://doi.org/10.1099/vir.0.19462-0

Velasco, L., Ruiz, L., Galipienso, L., Rubio, L. and Janssen, D., 2020. A historical account of viruses in intensive horticultural crops in the Spanish Mediterranean Arc: New challenges for a sustainable agriculture. Agronomy, 10(6): 860. https://doi.org/10.3390/agronomy10060860

Venkataraman, S., Prasad, B.V. and Selvarajan, R., 2018. RNA dependent RNA polymerases: Insights from structure, function and evolution. Viruses, 10(2): 76. https://doi.org/10.3390/v10020076

Wang, X., Ullah, Z. and Grumet, R., 2000. Interaction between zucchini yellow mosaic potyvirus RNA-dependent RNA polymerase and host poly-(A) binding protein. Virology, 275(2): 433-443. https://doi.org/10.1006/viro.2000.0509

Zhang, A., He, L. and Wang, Y., 2017. Prediction of GCRV virus-host protein interactome based on structural motif-domain interactions. BMC Bioinf., 18: 1-13. https://doi.org/10.1186/s12859-017-1500-8

Zhang, S., Yang, C., Qiu, Y., Liao, R., Xuan, Z.,

Ren, F., Dong, Y., Xie, X., Han, Y., Wu, D., Ramos-González, P. L., Freitas-Astúa, J., Yang, H., Zhou, C. and Cao, M., 2024. Conserved untranslated regions of multipartite viruses: Natural markers of novel viral genomic components and tags of viral evolution. Virus Evol., 10(1): veae004. https://doi.org/10.1093/ve/veae004

Zheng, K., Chen, S., Ren, Z. and Wang, Y., 2023. Protein arginine methylation in viral infection and antiviral immunity. Int. J. Biol. Sci., 19(16): 5292. https://doi.org/10.7150/ijbs.89498