Available online at http://bjas.bajas.edu.iq
https://doi.org/10.37077/25200860.2024.37.2.16
College of Agriculture, University of Basrah

Basrah Journal of Agricultural Sciences

ISSN 1814 – 5868 Basrah J. Agric. Sci., 37(2), 206-220, 2024 E-ISSN: 2520-0860

Complete Genome Sequence of *Turnip yellows polerovirus* from Iraq: Insight into the Factors Affecting the Genetic Variability of TuYV Populations

Muhannad Al-Waeli[1,*], Shirin Farzadfar[2] & Reza Pourrrahim[2]

[1] Department of Plant Protection, College of Agriculture, University of Basrah, Iraq.
[2] Plant Virus Research Department, Iranian Research Institute of Plant Protection (IRIPP), Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran.

***Corresponding author email**: **M.A.**: muhannad.khalf@uobasrah.edu.iq;
F.S.:farzadfar2002@yahoo.com**; P.R.**:pourrrahim@yahoo.com.

Received 26th July 2024; Accepted 12th November 2024; Available online 31st December 2024

Abstract: *Turnip yellows virus* (TuYV), is one of the most important viruses belonging to the *Polerovirus* genus. Incidence, genetic variability, and population structure of TuYV have not yet been studied in mid-Eurasia Iraq. Different brassica fields in Babylon, Basrah, Najaf, and Qadisiyeh provinces of Iraq were surveyed for TuYV infection. A total of 149 symptomatic leaf samples were checked for TuYV infection using reverse transcription polymerase chain reaction (RT-PCR). Furthermore, total RNA from three symptomatic rapeseed (*Brassica napus*) leaf samples from Basrah-Iraq was used for RNA-Seq. The diversity and population structure of TuYV and the evolutionary forces that shape these populations have been considered. High prevalence infections (52.34%) of TuYV were detected in surveyed fields. An Iraqi isolate was fully sequenced and showed the highest identity 92.68% to a Swedish isolate. Full genome sequences indicated two main phylogroups, however, using ORF0, TuYV isolates clustered in four clades. Clade I was further divided into two subclades (IA and IB), where the Iraqi isolate fell into IB. The recent distribution of TuYV populations in GI and GIII can be inferred from high haplotype diversity and low genetic differentiation. The *Fst* and *Nm* values indicate that the gene flow was low, thus the opportunity for divergence in populations via genetic drift is enhanced. By selective pressure, it was found that amino acid substitutions also contribute to TuYV evolution. This is the first evidence of the TuYV full genome sequence in mid-Eurasian Iraq and highlights the importance of recombination and selection pressure in the evolution of TuYV. Analysis of these variations is necessary for making advances in control strategies of viral diseases to prevent their spread.

Keywords: full genome, Iraq, molecular analysis, *Turnip yellows virus*.

Introduction

Turnip yellows virus (TuYV), is one of the most important viruses belonging to the *Polerovirus* genus (*Solemoviridae* family) (Filardo *et al.,* 2021; Sõmera *et al.,* 2021; LaTourrette *et al.,* 2021). At least 60 plant species belonging to *Brassicaceae*, *Fabaceae*, *Amaranthaceae,* and *Asteraceae* have been reported as hosts of TuYV, and about 17 aphid species can transmit TuVY in a persistent, non-circulative manner (Stevens *et al.,* 2008). The green peach aphid (*Myzus persicae*) is the major and efficient vector of TuYV, and a high

Copyright © 2024 Author(s). This is an open access article distributed under the terms of the Creative Commons CC BY 4.0 License (https://creativecommons.org/licenses/by/4.0/).

(up to 90%) transmission rate in some cases has been reported (Schliephake *et al.,* 2000; Newbert, 2016). Similar to other poleroviruses TuYV infection is limited to the phloem, as such; symptoms generally include stunting, yellowing, leaf malformations, and discoloration of the main leaf vein (Fiallo-Olive *et al.,* 2018). TuYV infection has been shown to reduce seed yield in single rape seed plants by 40 to 50% (Schroeder, 1994) and cause yield losses in rape seed crops of 11 to 46% (Jones *et al.,* 2007).

The TuYV genome consists of a single-stranded positive-sense RNA, protected by a virus protein genome-linked (VPg) cap at the 5′ untranslated region (UTR) and has organized six overlapping open reading frames (ORFs- P0 to P5). The P0 protein encoded by ORF0 is a non-conserved RNA silencing suppressor protein (Bortolamiol *et al.,* 2007). RNA-dependent RNA polymerase (RdRp) is encoded by a translational merging of ORF1 (P1) along ORF2 (P2). The major coat protein (P3) is necessary for virus assembly and efficient transmission of TuYV by aphid vectors and is encoded by ORF3 (Smirnova *et al.,* 2015). This protein is conserved within poleroviruses (Asare-Bediako, 2011; Newbert, 2016; Filardo *et al.,* 2021). The CP-read through (CP-RT) protein (ORF5), formed by the translational fusion of ORF3 and it has been involved in vector transmission, virus movement, and accumulation (Smirnova *et al.,* 2015). Two proteins P3a and P4 encoded by ORF3a and ORF4, are necessary for long-distance and cell-to-cell movement, respectively.

Phylogenetic analysis of TuYV has been studied using different isolates mostly from Europe and Australia (Newbert, 2016; Pimenta *et al.,* 2024). Using whole genome sequence analysis TuYV isolates fell into three clear groups, which were distinct from the TuYV-Fl sequence as an original isolate (Newbert, 2016). In addition, three clades have been also indicated for TuYV populations using ORF0 (the most various ORFs of the TuYV genome) (Pimenta *et al.,* 2024). Based on the ORF0 sequence analysis, all Chinese isolates are separated from the other European and Australian isolates belonging to the other clades (1 and 2), as well as TuYV-FL and WA-1 (Newbert, 2016; Pimenta *et al.,* 2024). Umar *et al.* (2022) reported 11 distinct TuYV groups based on their country of origin. However, specific grouping based on host species origin or geographic isolation was not obtained using phylogenetic population analysis based on ORF0, except for Chinese isolates (Pimenta *et al.,* 2024). In addition, TuYV population analysis using ORF3 divided worldwide isolates into three major clades, A, B, and C (Umar *et al.,* 2022). Recombination is a major factor in the appearance of new species, host range expansion, and emergence of resistance-breaking strains among RNA viruses (Garcia-Ruiz, 2018; LaTourrette *et al.,* 2021). RNA recombination is frequent in poleroviruses and plays an important role in the emergence of new species or strains by both inter and intra-species combinations (Pagan & Holmes, 2010; Umar *et al.,* 2022). At least 6 recombination events have been reported including those within ORF0 (events 1 and 2) and between ORF0 and ORF3 (events 3 to 6) (Pimenta *et al.,* 2024).

Major crops in Iraq are cereal, vegetables, millet, industrial plants, and fruit. The cultivated area of different brassica crops in Iraq is 23,818 hectares, including cauliflower, cabbage, turnip, and radish with 6,049, 4,401, 5,748, and 7,620 hectares, respectively, which constitutes about 4% of all vegetables cultivated in Iraq (CSO Iraq, 2021). Although the cultivated area of other brassica crops has not been determined separately, rapeseed and mustard are also among the crops that their cultivation are increasing in Iraq. Numerous diseases and pests also emerged and spread along with this growth (Al-Zehebawi *et al.,* 2019; Fadhil & Thamer, 2023). Despite the economically important role of brassica infecting viruses, molecular analysis of these pathogens is less understood in this area, possibly due to a lack of genetic sequence data. This study was carried out to assess the presence of TuYV infecting brassica crops in Iraq. The knowledge on identification of TuYV in the major brassica growing areas is noticeable for expanding effective control plans in these crops. Furthermore, the diversity and population structure of TuYV and the evolutionary forces that shape these populations have been considered.

Materials & Methods

Plant materials and viral source

In the early autumn and winter of 2022 and 2023, different brassica fields in Babylon, Basrah, Najaf, and Qadisiyeh provinces of Iraq were surveyed for TuYV infection (Table 1). A total of 149 symptomatic leaf samples with mottle, mosaic, and yellows (Fig. 1) were collected and checked for TuYV infection using reverse transcription polymerase chain reaction (RT-PCR). In addition, five TuYV isolates from different hosts and geographic locations were selected and their major CP sequences (ORF3) were amplified.

Table (1): Number of samples were collected from four provinces in Iraq during 2022-2023.

Province	Samples					
	Turnip/infected	Cabbage	Radish	Rapeseed	Mustard	Total
Babylon	23/12	8/0	5/2	-	-	36/14
Basrah	8/3	-	20/17	35/24	11/11	74/44
Najaf	10/2	-	4/0	-	-	14/2
Qadisiyah	21/6	-	4/1	-	-	25/7
Total	62/23	8/0	33/20	35/24	11/11	149/78

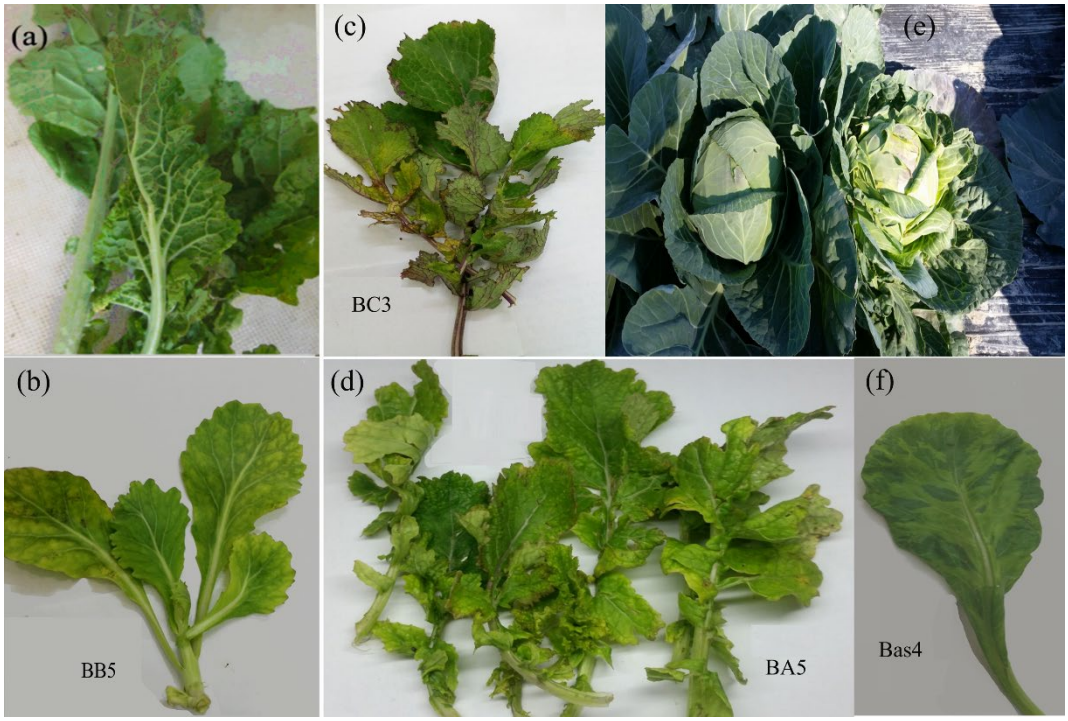

Fig. (1): Symptomatic leaf samples were collected from different brassica hosts in Iraq: (a) mustard; (b) rapeseed; (c and d) turnip; (e) cabbage; and (f) radish.

Total nucleic acid extraction and RT-PCR

Based on the previous studies, sequences of ORF0 and ORF3 were considered for the detection and molecular analysis of Iraqi TuYV isolates using different approaches (Hauser *et al.,* 2000; Farzadfar & Pourrahim, 2017; Umar *et al.,* 2022). Total nucleic acids (TNAs) extraction, was conducted using TRIzol (ThermoFisher, Invitrogen, USA). First-strand cDNA synthesis was done using the Easy cDNA Synthesis Kit (Cat. No. A101161, Parstous Co., Iran) at a final volume of 20 µl reaction, according to the manufacturer's instructions. The cDNA (3 µl) was amplified using 20 µM (20 pmol/µl) of TuYV ORF0 and ORF3

specific primers (Table 2) via 2XTaq PreMix (Cat. No. C101081, Parstous Co., Iran) in a 20µl reaction volume. After the first step of denaturation (5 min at 95°C), 40 cycles were applied, each of 30 s at 95°C, 30 s at 50°C, and 45 s at 72°C, followed by a final step of 10 min at 72°C. Amplified fragments were observed by agarose gel electrophoresis and their nucleotide sequence was determined.

Table (2): Sequence, position in the genome, target for amplification of primers used for RT-PCR amplification of TuYV.

Primer	Sequence	Position (nt)	amplification	Reference
ORF0F	CAAGTAACGTTTCTTTTAGCAGGT	91-114	681 bp	This study[a]
AB3	TCATACAAACATTTCGGTGTAGAC	748-771		Asare-Bediako, 2011
ORF3F	GAACAATCAATGGAAGAAGACGACCA	3491-3521	817 bp	This study[a]
ORF3R	TCTGGGAATACCATCGAGAATCGAGGTT	4281-4308		

[a]: The primers designed in this study

High throughput sequencing (HTS)

Total RNA from three symptomatic rapeseed leaf samples from Basrah-Iraq were used for RNA-Seq. Ribo-zero rRNA Removal Kit (Cat. No.: RZH1046, Epicentre, WI, the USA) was used for the deletion of rRNA from the total RNA. RNA-Seq libraries were constructed using TruSeq Stranded Total RNA for Illumina according to the manufacturer's instructions. The paired-end reads of 150 bp were generated by sequencing on the Illumina HiSeq 2000 platform (Novogene, China). CLC Genomics Workbench (version 20, QIAGEN, Venlo, The Netherlands) was used for data analysis of FASTQ files. Low-quality sequences (less than 15 nucleotide or ambiguous nucleotides more than 2) and adaptor's sequences were trimmed using Cutadapt (ver. 2.2). The assembled transcripts were mapped to nonchromosomal and chromosomal reference genomes of *B. rapa* (accession number GCA_002114115.1), to extract viral sequences from the entire transcriptome. After the elimination of host sequences, unmapped reads were kept and were used for viral genome *de novo* assembly using CLC Genomics Workbench (version 9, CLC Bio, Qiagen). CLC basic local alignment search tool (BLAST) was used to compare the unmapped contigs with the viral reference sequences available at NCBI (https://www.ncbi.nlm.nih.gov/genome/viruses). To assemble the complete viral genomes, the virus-associated contigs were aligned by pairwise alignment with the reference whole TuYV (NC_003743) genome sequence, using Geneious Prime 2019 with a 95% identity threshold for consensus sequence. The 5' and 3' terminal sequences of the TuYV was determined by SMARTer race 5'/3' kit (Cat. No. 634859, TaKaRa Biotechnology, Japan) according to the manufacturer's instructions. The complete genomes at viral ORFs and untranslated regions (UTRs) in the 5' and 3' ends of TuYV were obtained. The viral-related sequences were aligned using the ClustalX program for further analysis (Larkin et al., 2007).

Phylogenetic, genetic diversity, and recombination

Phylogenetic tree reconstruction based on the full genome (n=91), and ORF0 (n=303) was performed using the Neighbor-Joining (NJ) method implemented in MEGAX (Kumar *et al.,* 2018) and Neighbor-Net method in SplitsTree4 v.4.12.6 (Huson & Bryant, 2006) technique with 1,000 replicates to evaluate the statistical significance of branches. Dnadist implemented in Phylip 3.6 was used for the construction of a nucleotide diversity plot (Felsenstein, 2005). The six recombination detection programs implemented in RDP4 version 4.39 were used for recombination analysis using full genome sequences of TuYV, with default layout and Bonferroni corrected *P*-value cut-off of 0.05 (Martin *et al.,* 2015). Those recombination event(s) that were detected by at least three methods and *P*-value $\leq 1.0\times10^{-6}$ were considered.

Estimation of population genetic parameters

DnaSP v.6 was used for the estimation of population genetic parameters based on ORF0 (Rozas *et al.,* 2017). Nucleotide and haplotype diversities were determined for all populations as well as for each population. Furthermore, other parameters including nucleotide diversity (per site) (Pi), the average number of nt differences between the sequences (k), the number of haplotypes (h), and haplotype diversity (Hd) were also estimated for each population. Fu and Li's D (Fu and Li 1993) and Tajima's D (Tajima, 1989) values were calculated to find the demographic expansion. The genetic differentiation between populations was estimated using Kst*, Ks*, Z*, *Snn,* and *Fst* parameters according to geographical regions where TuYV was isolated (Rozas *et al.,* 2017; Hudson, 2000). There is no genetic differentiation if Kst* is near zero (null hypothesis) (Tsompana *et al.,* 2005). A smaller mean of Z* and *Snn* values indicates low genetic differentiation and vice versa (Hudson, 2000). The null hypothesis in Ks *, Kst*, and Z*, is rejected by a significant *P* value < 0.05 (Hudson 2000; Tsompana *et al.,* 2005). In addition, genetic differentiation between populations was estimated using *Fst* values, wherein it considered as very high for *Fst* > 0.25, high for 0.15 < *Fst* < 0.25, moderate for 0.05 < *Fst* < 0.15, and low for *Fst* < 0.05 (Hudson *et al.,* 1992). Concerning the criterion by Hudson (2000), gene flow is considered low for *Nm* < 1, high for 1 < *Nm* < 4, and very high for *Nm* > 4.

Selection pressure

The ratio of divergence at nonsynonymous and synonymous sites (shown as ω=dN/dS) was used to evaluate the degree to which selection affects ORF0 and ORF3. Models M0, M1a, M2, and M3 implemented in the CODEML program of the PAML4 package were used to determine the value of ω (Yang, 2007). Those sites with $\omega > 1$ and probability above 90% were considered as positively selected sites. Other sites with ($\omega < 1$), and ($\omega = 1$) were proposed as negative selection, and neutral evolution, respectively.

Results

Detection of TuYV

TuYV infections were detected in 78 out of 149 brassica symptomatic samples (52.34%) using RT-PCR and specific primers (Fig. 2a), indicating a high prevalence infection of this virus in brassica fields that were surveyed in Iraq. The highest infection percent among symptomatic samples was determined for mustard (100%), followed by rapeseed (68.6%), radish (67.0%), and Turnip (37.1%). However, no infection with this virus was detected in cabbage samples. Also, seven asymptomatic cauliflower samples were tested for TuYV infection, and the results were negative. Based on symptomatic samples, TuYV incidence were 59.5, 38.9, 28.0, and 14.3 percent in Basrah, Babylon, Qadisiyah, and Najaf provinces, respectively (Table 1). Except for IRQ-BaC, four TuYV isolates from Babylon (IRQBT3), Basrah (IRQBaR12), Gadisiyah (IRQGT21), and Najaf (IRQNR1) were selected and their *CP* genes were amplified (Table 2, Fig. 2b), and sequenced. The CP gene consists of 609 nucleotides in four Iraqi TuYV isolates and encodes CP with 202 amino acid residues.

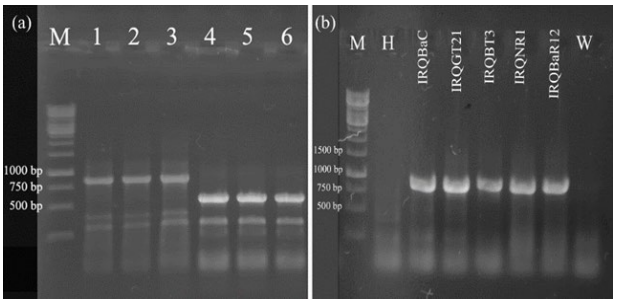

Fig. (2): Analysis of PCR products by agarose gel electrophoresis. M = GeneRuler 1 kb DNA ladder (Termofisher). The target-specific band size is around 681 bp for ORF0 (a) lanes 1, 2, 3, and 817 bp for ORF3 lanes 4, 5, 6. (b) The CP amplification reactions of five TuYV isolates: IRQBaC (from Basrah), IRQGT21 (from Gadisiyah), IRQBT3 (from Babylon), IRQNR1 (from Najaf), and IRQBaR12 (from Basrah), were shown. Healthy plant (H) and no-template (W) served as negative controls.

Genome assembly of TuYV

The whole transcriptome sequencing data were obtained using the Illumina platform. Entire reads before and after trimming were

58,971,338 and 47,707,525, rspectively. Assembly of the viral-related clean reads gave one contig 5606 nucleotide length which showed the highest identity 92.68% to a Swedish TuYV isolate (accession no.OP719311). This contig covered the nearly complete Iraqi TuYV isolate and represented a full-length TuYV genome, except for the 5' and 3' terminal sequences. Subsequently, ORFs, along with the 5' and 3' non-translated regions were achieved (Table 3). The Iraqi isolate (TuYV-IRQ-BaC) was most closely related to other TuYV isolates, and shared 87.90 to 92.68% nucleotide identity. Typically, of the genus *Polerovirus*, six ORFs were identified in this isolate. The complete genome of the TuYV-IRQ-BaC isolate was deposited in GenBank (accession no. PP586214). To confirm the presence of genomic sequence for which reads of TuYV were identified, RT-PCR assays were conducted using specific primers (Table 1). Expected PCR products were obtained and their sequences confirmed the *in silico* data.

Table (3): Genome features and ORFs of TuYV-IRQ-BaC isolate.

Type	length	Product
5'UTR	1-22	
ORF0	23-771	hypothetical protein ORF0
ORF1	165-1,987	RNA-dependent RNA polymerase
ORF3 + ORF5; CP + RTD	3,474-5,489	Aphid transmission
ORF3	3,475-4,082	Coat protein
ORF4	3,505-4,032	Movement protein
3'UTR	5,487-5,606	

Phylogenetic and Recombination analysis

Phylogenetic analyses using 91 full-length sequences, clustered TuYV isolates into two main phylogroups (Fig. 3a). Most of the TuYV isolates including IRQ-BaC fell into GI, but 10 TuYV isolates fell into small group GII. The phylogroup II consisted of seven Australian (MK103, MK106, MK109 isolated from pea, 5,514b, MK102, MK104 isolated from rapeseed, and MK113 isolated from lentil) along with one UK (isolated from pea), and two unknown TuYV isolates. Two dimensional nucleotide diversity plots also confimed two phylogroups where 0.00 to 18.4% and 18.4 to 27.6 % nucleotide diversities were estimated for GI and GII, respectively (Fig. 3b). Specific grouping based on host species origin or geographic isolation was not obtained using phylogenetic tree analysis. No recombination site was detected for the TuYV Iraqi IRQ-BaC isolate

Phylogenetic analysis via 303 ORF0 sequences separated TuYV isolates into four main clades according to NJ and Neighbor-Net trees (Fig. 4a, 4c). The largest clade I (CI) is divided into two subclades IA, and IB. Subclade IA consisted of just European isolates, whereas, isolates with different host origins from Australia and Europe fell into subclade IB and clade CIII (Fig. 4a, 4c). Two Iranian and Iraqi isolates from mid-Eurasia fell into subclade IB (Fig. 4c). All Chinese and one Polish isolates fell into clade II (CII). Our phylogenetic analysis indicates a new diverse clade (IV). This clade consists of 24 Australian TuYV isolates (isolated from pea and one orchid), one isolate from New Zealand, three isolates from South America, one isolate from South Africa, three European isolates G5, Geo15, and F98 from Germany, Greece, and France, respectively, which clustered with the reference isolate FL1 of TuYV (acc. no. X13063).

Two-dimensional nucleotide diversity plot also confirmed four main clades. Clade I divided into two subclades IA and IB, where the low nucleotide variability (0.0 to 3.0%) was estimated for subclade IA, followed by 3.0 to 6.0%, 9.0 to 12.1%, 6.0 to 12.1%, and 12.1 to 17.5% for IB, GII, GIII, and GIV, respectively (Fig. 4b).

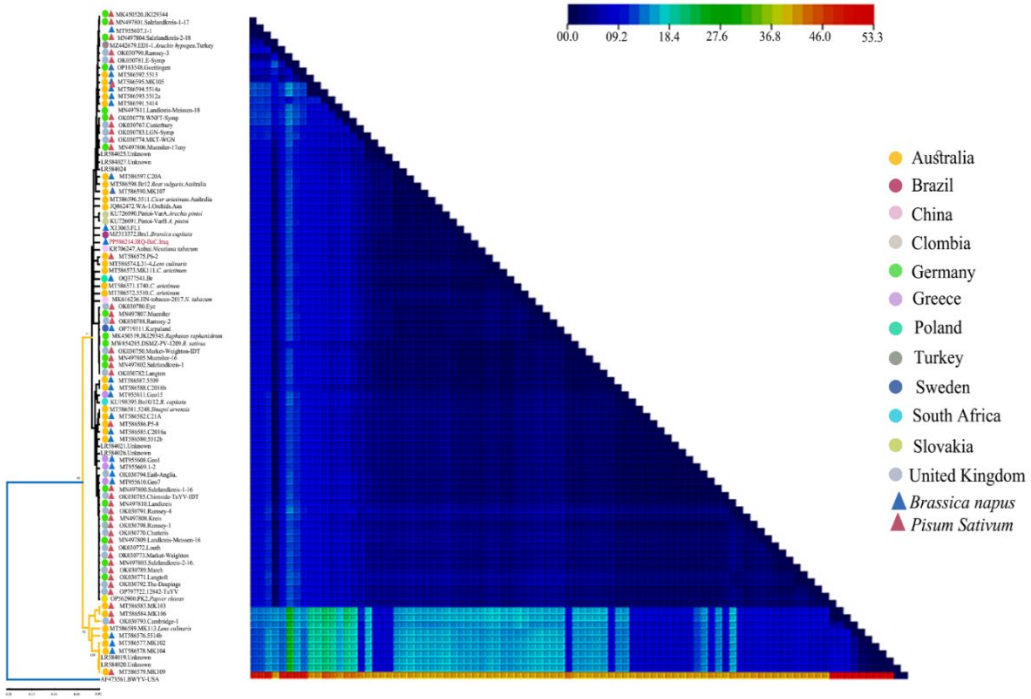

Fig. (3): The Phylogenetic tree was obtained from the whole genome sequences of 91 isolates of TuYV using the Neighbor-Joining (NJ) method. A genetic distance of 0.01 nt substitutions per site is shown by the scale bar. Bootstrap values indicate the percentage of 1,000 replications supporting the branching patterns shown. Phylogenetic analysis includes full/near-full genome *Beet western yellows virus* (BWYV) was used as an outgroup. Different colors were used for each contrary and also for the two main hosts (*Pisum sativum*, and *Brassica napus*). The Iraqi isolate showed by red color.

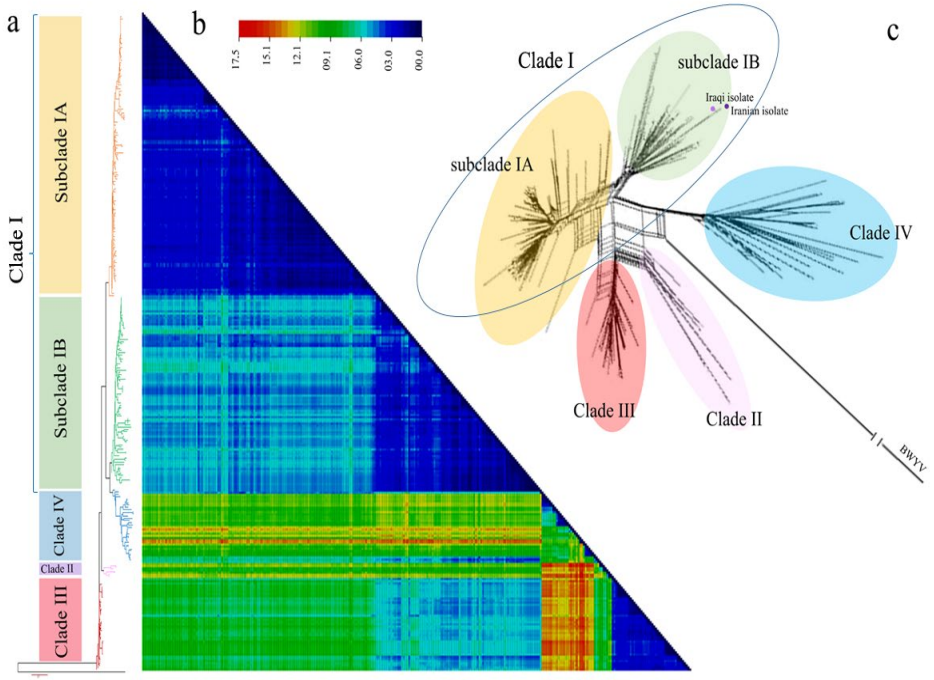

Fig. (4): Phylogenetic trees were obtained from the 303 ORF0 sequences using Neighbor-Joining (NJ) (a), and Neighbor-Net (c) methods. A genetic distance of 0.01 nt substitutions per site is shown by the scale bar. *Beet western yellows virus* (BWYV) was used as an outgroup. Two subclades IA and IB, CladeII, Clade III, and Clade IV indicated by orange, green, blue, pink, and red colors, respectively. Tow dimensional plot indicated based on different clades of ORF0 (b).

Estimation of population genetic parameters

Among the 303 isolates tested for ORF0, 259 distinct haplotypes were detected with 0.9982 haplotype diversity (Hd) for all sequences. The average number of nucleotide differences (k), and nucleotide diversity (Pi) for all isolates were 30.518, and 0.05929, respectively (Table 4). The highest haplotype diversity (Hd) was found for CII (all Chinese and one Polish isolate), followed by CIV. In addition, the average nucleotide diversity from high to low was in order of CIV, followed by CII, CIII, CIB, and CIA. Nucleotide differences within haplotypes (k) were highest in CII followed by CIV, CIA, and CIB populations (Table 4).

Table (4): Demography test statistics between the ORF0 of TuYV populations.

Population	N	h	Hd	K	Pi	Neutrality test and significance	
						Tajima's D	Fu and Li's D
All	303	259	0.9982	30.518	0.05589	-1.43044 ns	-2.10408 ns
CIA	129	110	0.9965	10.93883	0.02003	-1.68602 ns	-3.04318 *
CIB	92	75	0.9947	11.88127	0.02186	-1.87213 *	-3.45025 *
CII	7	7	1.0000	43.04762	0.07390	0.71564 ns	1.02931 ns
CIII	42	37	0.9919	11.35075	0.02192	-1.30963 ns	-2.29241 ns
CIV	33	30	0.9962	42.61364	0.07529	-0.82431 ns	-0.73484 ns

N – number of isolates; h – number of haplotypes; Hd – haplotype diversity; k – average number of nucleotide differences between sequences; Pi– nucleotide diversity (per site); *P < 0.05; ns – not significant.

The Fu's Fs test was not significant for different TuYV clades (P>0.05), except for two subclades IA and IB populations, as also indicated for Tajima's D test. Population genetic parameters were calculated according to the geographical isolation which indicates that TuYV populations are distinct from each other (Table 5). Pairwise F_{ST} values showed a strong genetic differentiation between TuYV populations (Table 5). The *Fst* varies from 0.15679 (between Oceania and West Asia populations) to 0.50982 (between West Asia and South America) populations. In addition, gene flow (*Nm*) between the populations varied from 1.34 (between Oceania and West Asia populations) to 0.24 (between West Asia and South America) populations. The highest gene flow was determined between Oceania and West Asia followed by Europe and Oceania (Table 5). However, low gene flow was found between other populations (*Nm* < 1). Genetic differentiation and frequent gene flow were also confirmed with Ks*, Kst* Z*, and *Snn* statistics values (Table 5).

Table (5): Genetic differentiation measurement between populations from different continents using ORF0 sequences.

Continent	Ks*	Kst*	*Z**	*Snn*	*P* value	F_{ST}	*Nm*
Europe & East Asia	2.90674	0.02034	8.66709	0.99454	2: 0.0845 ns	0.4236	0.34
Europe & Oceania	2.97904	0.06108	9.27984	0.96393	2: 0.0294 *	0.1934	1.04
Europe & South America	2.89630	0.01369	8.65613	0.98889	2: 0.0827 ns	0.4678	0.28
Europe & West Asia	2.89288	0.00603	8.66660	0.98889	2: 0.5876 ns	0.2865	0.62
East Asia & Oceania	3.40525	0.02451	7.66090	1.00000	2: 0.1416 ns	0.3730	0.42
East Asia & South America	3.67269	0.12070	2.02676	1.00000	0.3423 ns	0.4438	0.31
East Asia & West Asia	3.61474	0.08368	2.27720	1.00000	2: 0.3423 ns	0.3993	0.38
Oceania & South America	3.39570	0.01651	7.63192	0.99083	0.1381 ns	0.4032	0.37
Oceania & West Asia	3.39294	0.00424	7.66236	0.99083	2: 0.1381 ns	0.1567	1.34
West Asia & South America	3.42753	0.16730	1.20131	0.83333	2: 0.3062 ns	0.5098	0.24

Probability (*p*-value) obtained by the permutation test (PM test) with 1000 replicates. *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, P < 0.001; ns, not significant. The PM test was performed using DnaSP v. 6.10.04. *Fst* > 0.25 indicates infrequent gene flow; *Fst* < 0.25 suggests frequent gene flow. *Nm* is the migration fraction per generation.

Selection analyses

Model0 (M0) was used to assess selection pressures on two proteins P0 (ω = 0.432), and P3 (ω = 0.273), which indicates the mean ω ratio differed between them (Table 6). In addition, model3 (M3) suggested that a large set of sites (55.84%) were evolving under strong purifying selection (ω = 0.07096), 31.1% sites were under weak purifying selection (ω= 0.535) and 24 sites (13.1%) were under positive selection (ω = 1.92864). Although 12 sites under positive selection including (19Y, 117R, 124L, 127N, 145H, 159E, 164G, 176V, 181L, 118L, 230C, and 249 E) were detected in all phylogroups, howbeit, 37T and 69G were just found in clades CIV, and CII, respectively (Table 7).

For P3 the mean ω value was 0.273 and different sites were found under positive selection by M2 and M3 with $\omega \geq 1.0$ and posterior probability > 99.0 (Table 6). The M3 suggested that a large set of sites (57.10%) evolved under strong purifying selection (ω=0.07096), 31.4% of sites were under weak purifying selection (ω=0.53541) and 11.5% of sites were under positive selection (ω = 2.529). As indicated by M2 and M3, 20 sites were detected under positive selection, most of them were found in the 5' end of the TuYV CP gene (Table 6). Except 37P, 56I, and 16M which were just found in two Chinese isolates (HN, and Anhui, isolated from *Nicotiana tabacum*) (Table 7), other four sites 23I, 26N, 155S, and 160L were found among a few isolates including two Australian 1740, and 5510 (isolated from *Cicer arietinum*), a Polish (*B. napus*) and two Iranian (*Medicago sativa*) TuYV isolates (Table 7).

Table (6): The dn⁄ds (ω) values, log-likelihood (ln*L*) values, likelihood ratio test (LRT) statistics and positively selected amino acid sites undergoing different models of codon substitution used to investigate selection pressures on P0 and P3 proteins encoded by the *Turnip yellows virus* analyzed in this study.

	Models[1]	Parameter estimates	ω ratio	ln*L*	LRT[2]	Positively selected (amino acids) sites [3]
P0	M0	ω=0.43202	0.432	-11559.013		None
	M3	*p0*=0.558, *p1*=0.311, *p2*=0.131, *ω0*=0.071, *ω1*=0.535, *ω2*=1.928	0.458	-11149.185	P<0.01 (0.0)	(4I, 19Y, 23R, 37T, 54V, 69G, 79L, 108H, 117R, 127N, 145H, 149L, 170V, 173H, 176V, 181L, 228L, 230C, 246 E) ** (77K,124L,159E, 164 G,227Y)*
	M1	*p0*=0.71288, *p1*=0.287, *ω0*=0.115,*ω1*=1.000	0.368	-11208.281		Not allowed
	M2	*p0*=0.695, *p1*=0.218, *p2*=0.086, *ω0*=0.125, *ω1*=1.000, *ω2*=2.336	0.507	-11163.916		(79L, 108H, 127N,176V, 228L)** (23R, 37T, 124L, 159E, 164G,230C, 246E)*
P3	M0	ω=0.273	0.273	-5404.617		None
	M3	*p0*= 0.571, *p1*=0.314, *p2*=0.115, *ω0*=0.096, *ω1*=0.551, *ω2*=2.529	0.287	-5238.457	P<0.01 (0.0)	(7R, 20T, 23I, 26N, 34A, 37P, 49N, 53G, 54R,55T, 56I, 62G, 98M, 153R, 154R, 155S, 164M)** (57P,115V,160L)*
	M1	*p0*=0.865, *p1*=0.134, *ω0*=0.078, *ω1*=1.000	0.205	-5254.969		Not allowed
	M2	*p0*= 0.873, *p1*= 0.069, *p2*= 0.057, *ω0*= 0.091, *ω1*=1.000, *ω2*=2.338	0.283	-5239.631	P>0.05,(0.093)	(20T, 55T,154R)** 57P*

[1] Model descriptions M0 (one ration); M3 (discrete); M1a (nearly neutral); M2 (positive selection).
[2] LRTs of M3 vs. M0 are tests of heterogeneity of selection pressures among codon sites. M2 vs. M1 are tests of positive selection; all assess LRT statistics (2dlnL) against a chi-square distribution with the degrees of freedom equal to the difference in the number of parameters between the nested models being compared.
[3] Amino acid (codon) sites with higher posterior probabilities of P>95.0 (*) and P>99.0 (**) undergoing positive selection are shown.

Table (7): Different sites were detected under positive selection according to each Clades of ORF0.

ORF0/Clades	Sites under positive selection
IA	4I, 54V, 77K, 108H, 149L,
IB	23R, 77K, 108H, 149L, 227Y
II	23R, 54V, **69G**, 77K, 79L, 108H, 149L, 170V, 173H, 227Y
III	54V, 77K, 79L, 108H, 149L, 170V, 173H, 227Y
IV	23R**, 37T**, 54V, 77K, 79L, 108H, 149L, 170V, 173H, 227Y
All Clades	19Y, 117R, 124L, 127N, 145H, 159E, 164G, 176V, 181L, 228L, 230C, 249E
P3 (*CP* gene)	Sites under positive selection
Isolate (geographical isolation)	
HN (China); Br (Poland); MK111 (Australia)	23I, 26N
HN and Anhui(China)	37P, 56I, 164M
HN (China), Br (Poland), 1740, and 5510 (Australia)	155S
HN (China), Br (Poland), 1740, and 5510 (Australia), IRTuYK132 and IRTuYL41 (Iran)	160L
Common*	7R, 20T, 34A, 49N, 53G, 54R, 55T, 56I, 57P, 62G, 98M, 115V, 154R

In ORF0, two sites 69G and 37T just were detected in CII and GIV, respectively, and highlighted in bold.
* These sites were found among many isolates of TuYV.

Discussion

In this study, some brassica fields were surveyed during the period from 2022 to 2023 and the TuYV was found to be widespread in four provinces in Iraq (West Asia). Nevertheless, the average incidence of this virus among symptomatic plants surveyed differed from 37.1% to 100% (Table 1). A high population of cabbage aphid (*Brevicoryne brassicae*) was observed in the fields surveyed. Although, the green peach aphid *M. persicae* is considered the most efficient vector for TuYV (Schliephake *et al.,* 2000), however, other aphid species, such as *B. brassicae*, *Aphis gossypii*, and *Macrosiphum euphorbiae*, transmits TuYV with lower rates (less than 10%) (Schliephake *et al.,* 2000; Stevens *et al.,* 2008; Pimenta *et al.,* 2024). The climate change and mild autumns provide prolonged active periods to insects including aphis, which contribute to the spread of viruses to winter crops (Puthanveed *et al.,* 2023).

Studying the genetic diversity of viruses increases our information about geographical origin, virulence variations, emerging new viruses or epidemics, and finally viral evolution. We used HTS to determine the full genome sequence of an isolate of TuYV from rapeseed in Basrah province, to understand the genetic variation of TuYV in Iraq (West Asia). An assembled genome, generated using HTS was 5,045 bp long and had the typical organization to TuYV. The complete genome of IRQ-BaC isolate (5,606 nucleotides) with six overlapping ORFs, shared 92.68 % identity to a Swedish TuYV isolated from rapeseed. Phylogenetic analysis using complete genome sequences indicates two distinct clades, with no specific grouping based geographical isolation or host species origin, consistent with earlier studies (Asare-Bediako, 2011; Newbert, 2016; LaTourrette *et al.,* 2021; Filardo *et al.,* 2021; Slavíková *et al.,* 2022; Umar *et al.,* 2022; Pimenta *et al.,* 2024). Most of the isolates including the Iraqi isolate fell into clade I with low nucleotide diversity. The wider distribution and prevalence of clade I isolates may be associated with higher fitness or transmission efficiency by germplasm or vectors (Gouveia *et al.,* 2011). Furthermore, strong selective pressures (e.g. due to fitness) or evolutionary founder effects or bottleneck events could be inferred from the lowest nucleotide diversity in this clade (Farooq *et al.,* 2013).

ORF0 has been used in different studies to investigate phylogenetic and molecular diversity in poleroviruses including TuYV (Hauser *et al.,* 2000; Asare-Bediako, 2011; Newbert, 2016; LaTourrette *et al.,* 2021; Filardo *et al.,* 2021; Slavíková *et al.,* 2022; Umar *et al.,* 2022; Pimenta *et al.,* 2024). In this study, TuYV isolates fell into four main clades, indicating a similar but not identical topology to that of ORF0 trees which were previously reported (Asare-Bediako, 2011; Newbert, 2016; Pimenta *et al.,* 2024). All the Chinese isolates and the Polish Br isolate (accession no. OQ377541) clustered in clade II with 9.0 to 12.1% nucleotide diversity. The highest nucleotide diversities were found in two clades III and GIV than in the subclades of IA, IB, and CII. The clade IV consisted of more different TuYV isolates from various geographic and host origins. However, some correlations with geographical isolation e. g. subclade IA and clade II may indicate the important role of geographic isolation which affects the genetic structure of the viral population (Yang *et al.,* 2021; Sun *et al.,* 2021). The environmental adaption can also be considered an important evolutionary force that shaped the genetic structure of TuYV.

Recombination hotspots were detected in P1, P1-P2, P3, and P3+P5 of poleroviruses, and indicates that the 5′ and 3′ halves of the genome could have different evolutionary histories (LaTourrette *et al.,* 2021; Umar *et al.,* 2022). In addition, Sun *et al.* (2021), indicated that critical biological functions controlled by the P3 gene could be affected by recombination and mutation. Our analysis also showed recombination breakpoints between the P1 (event 1), P1-P2 fusion (event 2), P3 (event 3), and P3+P5 (event 4), based on full sequences, however, no recombination site was detected for Iraqi IRQ-BaC isolate (Data not shown).

The recent distribution of the population can be inferred from high haplotype diversity and low genetic differentiation as indicated for TuYV populations in GI and GIII in comparison to GII and CIV (Table 4). Except

for two subclades IA and IB, the Fu and Li's D and Tajima's D tests were not significant for different TuYV populations ($P > 0.05$). In addition, positive values of Tajima's D and Fu and Li's D tests were found for clade II (The Chinese croup), but no statistically significant departure from neutrality ($P>0.05$) was shown (Table 4). The negative values of Tajima's D indicate a possible recent population expansion or selection variation events such as the founder effect or bottleneck that reduce genetic diversity. However, positive Tajima D values are associated with an excess of common variation in a region due to balancing selection, population sub-division, or sudden population contraction. *Fst* and *Nm* values indicate that the gene flow was low in TuYV populations (Sun *et al.,* 2021; this study), thus the opportunity for divergence in populations via genetic drift is enhanced (Jones & Wang, 2012). Low genetic diversity and low gene flow correlated with a weak genetic structure of populations. As poleroviruses are not seed-borne, therefore, high gene flow determined between Oceania and West Asia, and Europe and Oceania populations may illustrate the role of germplasm transmission between these far geographical regions.

Selection pressure on different poleroviruses' proteins including MP, CP-RT, CP, and P0 was evaluated and indicated that the genome of poleroviruses is under purifying selection and genetically stable, to preserve the long-term stability of biological structures and keeping the encoded protein functional (Asare-Bediako, 2011; Umar *et al.,* 2022). In this study, different sites were found under positive selection in two proteins P0 and P3 with posterior probability >99, which may show the key role of these sites in adaptability to environmental changes and new hosts as previously indicated for MP, and CP-RT of poleroviruses (LaTourrette *et al.,* 2021; Sun *et al.,* 2021; Umar *et al.,* 2022). Information about the biological functions of these proteins has rarely been demonstrated. Nevertheless, it is assumed that those sites under positive selection pressure contribute to increased fitness and host adaptability (Parto & Lartillot, 2018; Sun *et al.,* 2021).

Conclusion

This study was conducted to investigate the incidence of TuYV infecting brassica crops in some provinces in Iraq. The knowledge on identification of TuYV in the major brassica growing areas is noticeable for expanding effective control plans in both these crops and other hosts. Furthermore, the diversity and population structure of TuYV and the evolutionary forces that shape these populations have been considered. Increasing the number of sequences in the future could help to investigate the consequences associated with these pathogens (e.g. synergistic interactions with other known long-term viruses). Untargeted descriptions of all viruses by HTS have made a big evolution in the discovery and characterization of viruses, which is necessary to develop early diagnosis, and precise identification of viral agents. This is the first evidence of the TuYV full genome sequence in mid-Eurasian Iraq and highlights the importance of recombination and selection pressure in the evolution of TuYV. The accurate pictures of phylogenetic relationships and comparisons among distantly related virus genomes present an important insight into basic evolutionary mechanisms. Additionally, analysis of these variations is necessary for making advances in control strategies of viral diseases to hinder their spread.

Acknowledgments

This work has been done under the joint project number 2-16-16-073-000817 between the Iranian Research Institute of Plant Protection (IRIPP) and the University of Basrah (Iraq). The support of the Plant Virus Research Department of IRIPP, and the Department of Plant Protection, College of Agriculture, University of Basrah, to carry out this project is kindly acknowledged.

Contributions of authors

Conceptualization, M. A., S. F., and R. P.; methodology, S. F; software, M.A., and S. F.; validation, R.P, and S.F.; formal analysis, M.A., and S.F.; investigation, M.A., R.P, and S.F.; writing-original draft preparation, R.P, review and editing, M.A., S.F., and R.P.;

visualization, All authors have read and agreed to the published version of the manuscript.

ORCID
M. Al-Waeli: 0000-0003-4944-7252
S. Farzadfar: 0000-0002-9024-9381
R. Pourrahim: 0000-0001-5175-6148

Conflicts of interest

The authors declare that they have no conflict of interest.

References

Al-Zehebawi, N. A., Al-Waily, D. S., & Al-Saad, L. A. (2019). The Application of manure (Poultry Wastes) and bio-formulations of *Trichoderma harzianum* and *T. viride* and their interaction to control root-knot disease on radish and chard. *Basrah Journal of Agricultural Sciences*, *32*(1), 88-98. https://doi.org/10.37077/25200860.2019.195

Asare-Bediako, E. (2011). Brassicaceae: *Turnip yellows virus* interactions. Doctoral dissertation, University of Warwick, Coventry. https://wrap.warwick.ac.uk/44041/

Bortolamiol, D., Pazhouhandeh, M., Marrocco, K., Genschik, P., & Ziegler-Graf, V. (2007). The *Polerovirus* F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. *Current Biology 17*(18), 1615-1621. https://doi.org/10.1016/j.cub.2007.07.061

CSO. (2021). Central Statistical Organization IRAQ. 15 pp. https://cosit.gov.iq/ar/agri-stat/veg-prod

Fadhil, A. A., & Thamer, N. K. (2023). Chemical Control of the Leaf Miner Scaptomyza flava Fallen (Diptera: Drosophilidae) and Determination of Acetamipride and Abamectin Residues on Radish Plant. *Basrah Journal of Agricultural Sciences*, *36*(2), 226-234. https://doi.org/10.37077/25200860.2023.36.2.17

Farooq, A. B. U., Ma, Y. Xi., Wang, Z., Zhuo, N., Wenxing, X., Wang, G. P., & Hong, N. (2013). Genetic diversity analyses reveal novel recombination events in Grapevine leafroll-associated virus 3 in China. *Virus Research 171*(1), 15-21. https://doi.org/10.1016/j.virusres.2012.10.014

Farzadfar, Sh., & Pourrahim, R. (2017). Molecular detection of *Turnip yellows virus* (TuYV) infecting alfalfa in Iran. *Australian Plant Disease Note 12*, 12. https://doi.org/10.1007/s13314-017-0236-7·

Felsenstein, J. (2005). PHYLIP (phylogeny inference package) version 3.6. Distributed by Author. Department of Genome Sciences, University of Washington, Seattle.

Fiallo-Olivé, E., Navas-Hermosilla, E., Ferro, C. G., Zerbini, F. M., & Navas-Castillo, J. (2018). Evidence for a complex of emergent poleroviruses affecting pepper worldwide. *Archives of Virology 163*(5), 1171-1178. https://doi.org/10.1007/s00705-018-3733-x

Filardo, F., Nancarrow, N., Kehoe, M., McTaggart, A., Congdon, B., Kumari, S. G., Aftab, M., Trębicki, P., Rodoni, B., Thomas, J., & Sharman, M. (2021). Genetic diversity and recombination between turnip yellows virus strains in Australia. *Archives of Virology 166*(), 813-829. https://doi.org/10.1007/s00705-020-04931-w

Garcia-Ruiz. P. (2018). Susceptibility genes to plant viruses. *Viruses 10*(9), 484. https://doi.org/10.3390/v10090484

Gouveia, P., Santos, M. T., Eiras-Dias, J. E., & Nolasco, G. (2011). Five phylogenetic groups identified in the coat protein gene of grapevine leafroll-associated virus 3 obtained from Portuguese grapevine varieties. *Archives of Virology 156*(3), 413-420. https://doi.org/10.1007/s00705-010-0878-7

Hauser, S., Weber, C., Vetter, G., Stevens, M., Beuve, M. & Lemaire, O. (2000). Improved detection and differentiation of poleroviruses infecting beet or rape by multiplex RT-PCR. *Journal of Virological Methods 89*(1-2), 11-21. https://doi.org/10.1016/s0166-0934(00)00203-2

Hudson, R. R. (2000). A new statistic for detecting genetic differentiation. *Genetics 155*(4), 2011-2014. https://doi.org/10.1093/genetics/155.4.2011

Hudson, R. R., Boos, D. D., & Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution 9*(1), 138-151. https://doi.org/10.1093/oxfordjournals.molbev.a040703

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution 23*(2), 254-267. https://doi.org/10.1093/molbev/msj030

Jones, O. R., & Wang, J. (2012). A comparison of four methods for detecting weak genetic structure from marker data. *Ecology and Evolution 2*(5), 1048-1055. https://doi.org/10.1002/ece3.237

Jones, R. A. C., Coutts, B. A. & Hawkes, J. (2007). Yield-limiting potential of beet western yellows virus in Brassica napus. *Australian Journal of Agricultural Research* 58 (8)788-801. https://doi.org/10.1071/AR06391

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGAX: Molecular evolutionary genetics analysis across Computing Platforms. *Molecular Biology and Evolution 35*(6), 1547-1549. https://doi.org/10.1093/molbev/msy096

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R. McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins. D. G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics 23*(21), 2947-2948. https://doi.org/10.1093/bioinformatics/btm404

LaTourrette, K., Holste, K. M., & Garcia-Ruiz, H. (2021). *Polerovirus* genomic variation. *Virus Evolution 7*(2), 1-18. https://doi.org/10.1093/ve/veab102

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution 1(1),* vev003. https://doi.org/10.1093/ve/vev003

Newbert, M., (2016). The genetic diversity of Turnip yellows virus in oilseed rape (Brassica napus) in Europe, pathogenic determinants, new sources of resistance and host range. Doctoral dissertation, University of Warwick, Coventry. https://wrap.warwick.ac.uk/id/eprint/79104/

Pagan, I., & Holmes, E. C. (2010). Long-term evolution of the Luteoviridae: time scale and mode of virus speciation. *Journal of Virology 84*(2), 6177-6187. https://doi.org/10.1128/JVI.02160-09

Parto, S., & Lartillot, N. (2018). Molecular adaptation in Rubisco: Discriminating between convergent evolution and positive selection using mechanistic and classical codon models. *PLos One 13*(2), e0192697. https://doi.org/10.1371/journal.pone.0192697

Pimenta, R. J., Macleod, K., Babb, R., Coleman, K., MacDonald, J., Asare-Bediako, Jenner, C. E., & Walsh, J. A. (2024). Genetic variation of *Turnip yellows virus* in Arable and vegetable brassica crops, perennial wild Brassicas, and aphid vectors collected from the Plants. *Plant Disease*, *108*(3), 616-623. https://doi.org/10.1094/PDIS-05-23-0906-RE

Puthanveed, V., Singh, K., Poimenopoulou, E., Pettersson, J. Siddique, A. B., & Kvarnheden, A. 2023. Milder autumns may increase risk for infection of crops with *Turnip yellows virus*. *Phytopathology 113*(9), 1788-1798. https://doi.org/10.1094/PHYTO-11-22-0446-V

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., GuiraoRico, S., Librado, P., Ramos-Onsins, S. E., & SanchezGracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution 34*(12), 3299-3302. https://doi.org/10.1093/molbev/msx248

Schliephake, E., Graichen, K., & Rabenstein, F. (2000). Investigations on the vector transmission of the *Beet mild yellowing virus* (BMYV) and the *Turnip yellows virus* (TuYV). *Journal Plant Disease and Protection 107*(1), 81-87. https://eurekamag.com/research/003/483/003483778.php

Schroeder, M. (1994). Investigations on the susceptibility of oilseed rape (Brassica-napus L, ssp napus) to diferent virus diseases. *Journal Plant Disease and Protection* 101, 567-589.

Slavíková, L., Ibrahim, E., Alquicer, G., Tomašechová, J., Šoltys, K., Glasa, M., & Kundu, J. K. (2022). Weed hosts represent an important reservoir of *Turnip yellows virus* and a possible source of virus introduction into oilseed rape crop. *Viruses 14*(11), 2511. https://doi.org/10.3390/v14112511

Smirnova, E., Firth, A. E., Miller, W. A., Scheidecker, D., Brault, V., Reinbold, C., Rakotondrafara, A. M., Chung, B, Betty Y-W., & Ziegler-Graff, V. (2015). Discovery of a small non-AUG-initiated ORF in poleroviruses and luteoviruses that is required for long-distance movement. *PLoS Pathology 11*(5), e1004868. https://doi.org/10.1371/journal.ppat.1004868

Sõmera, M., Fargette, D., Hébrard, E. & Sarmiento, C. (2021). ICTV Report Consortium 2021. ICTV Virus Taxonomy Profile: Solemoviridae. *Journal of General Virology 102*(12), 001707. https://doi.org/10.1099/jgv.0.001707

Stevens, M., McGrann, G., Clark, B., & Authority, H. (2008). *Turnip yellows virus* (syn *Beet western yellows virus*): An emerging threat to European oilseed rape production. Research Review 69 HGCA. Available via: https://ahdb.org.uk/turnip-yellows-virus-syn-beet-western-yellows-virus-an-emerging-threat-to-european-oilseed-rape-production

Sun, S. R., Chen, J. S., He, E. Q., Huang, M. T., Fu, H. Y., Lu, J. J., & Gao, S. J. (2021). Genetic variability and molecular evolution of *Maize yellow mosaic virus* populations from different geographic origins. *Plant Disease*, *105*(4), 896-903. https://doi.org/10.1094/PDIS-05-20-1013-RE

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

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

Umar, M.; Tegg, R.S.; Farooq, T.; Thangavel, T., & Wilson, C.R. (2022). Abundance of poleroviruses within Tasmanian pea crops and surrounding weeds, and the genetic diversity of TuYV isolates found. *Viruses 14*(8), 1690. https://doi.org/10.3390/v14081690

Yang, C., Zhao, Q., Wang, Y., Zhao, J., Qiao, L., Wu, B, Yan, S., Zheng, J., & Zheng Xi. 2021. Comparative analysis of genomic and transcriptome sequences reveals divergent patterns of codon bias in wheat and its ancestor species. *Frontier in Genetic 12*, 732432. https://doi.org/10.3389/fgene.2021.732432

Yang, Z. (2007). PAML4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution 24*(8), 1586-1591. https://doi.org/10.1093/molbev/msm088

التسلسل الكامل لجينوم *Turnip yellows polerovirus* من العراق : نظرة للعوامل التي تؤثر على التنوع الجيني لمجاميع فايروس TuYV

مهند الوائلي*[1]، شيرين فرزادفر*[2]، و رضا بوررحيم[2]

1. قسم وقاية النبات، كلية الزراعة، جامعة البصرة، العراق.
2. قسم بحوث فيروسات النبات، المعهد الإيراني لبحوث وقاية النبات(IRIPP) ، منظمة البحوث والتعليم والإرشاد الزراعي (AREEO)، طهران، إيران.

المستخلص : يعد فايروس اصفرار اللفت (*Turnip yellows virus*) أحد أهم الفايروسات التي تعود إلى جنس *Polerovirus*, لاتوجد اي دراسات تتناول الإصابة والتنوع الجيني والتركيب لفايروس اصفرار اللفت (TuYV) في أوراسيا الوسطى- العراق حتى الآن. تم اجراء مسح لحقول مختلفة لنباتات العائلة الصليبية في العراق وذلك للتحري عن فايروس اصفرار اللفت TuYV وباستخدام تقنية تفاعل البوليميراز المتسلسل العكسي(RT-PCR) reverse transcription polymerase chain reaction , تم فحص 149 عينة من الأوراق المصابة للكشف عن الفايروس TuYV, استخرج الحامض النووي الرايبي RNA لثلاث عينات من أوراق نبات السلجم من محافظة البصرة-العراق واستخدم لـــ RNA-Seq. تم دراسة التنوع التركيبي والبنائي لـــفايروس TuYV والقوى التطورية لهذه المجموعة من الفايروسات, اشارت النتائج الى انتشار الفايروس TuYV في الحقول التي تم اجراء المسح لها وبنسبة (52.34%), و أظهرت نتائج تحليل التتابعات النيوكليوتيدية الكامل للعزلة العراقية تشابه مع العزلة السويدية وبنسبة (92.68%), وتفرعت الشجرة الوراثية للتتابعات الجينوم الكامل إلى مجموعتين رئيسيتين، في حين تفرعت الشجرة الوراثية لعزلات الفايروس TuYV إلى أربع مجموعات باستخدام ORF0, وتفرعت المجموعة I إلى مجموعتين فرعيتين (IA وIB)، حيث وقعت العزلة العراقية في المجموعة IB. ومن خلال التوزيع لمجموعات فايروس TuYV في GI وGIII امكن الاستنتاج من وجود تنوع كبير في النمط الوراثي و انخفاض التمايز الجيني و تشير قيم *Fst* و*Nm* إلى أن (gene flow) كان منخفضًا ,وبالتالي فأن هنالك فرصة لتعزيز التباين والانحراف الوراثي ,وبالضغط الانتقائي وجد أن استبدال الأحماض الأمينية له دور في تطور فايروس TuYV, ويعد هذا البحث أول دراسة تتناول تسلسل الجينوم الكامل لـ TuYV في منطقة أوراسيا الوسطى العراق ويسلط الضوء على أهمية إعادة التركيب والضغط الانتقائي في تطور TuYV, ويعد تحليل هذه التباينات ضروريا لتحقيق تقدم في استراتيجيات السيطرة على الأمراض الفيروسية لمنع انتشارها.

كلمات مفتاحية: الجينوم الكامل; العراق; تحليل جزيئي; فايروس اصفرار اللفت.