مهندسی ژنتیک و ایمنی زیستی
ISSN 2588-5073 چاپی
ISSN 2588-5081 الکترونیکی
دوره ۷، شماره ۱، بهار و تابستان ۹۷
صفحه ۱۱-۱

Genetic Engineering and Biosafety Journal
Volume 7, Number 1

تجزیه و تحلیل تبارزائی و تنوع ژنتیکی ویروس پیچیدگی برگ زرد گوجه فرنگی در عراق براساس ژن V1

Phylogenetic analysis and genetic variation of Tomato yellow leaf curl virus based on the V1 gene in Iraq

مهند الوائلی[۱،۲]، اکبر دیزجی *[۱]، غلامحسین مصاحبی [۱]، اکبر آهنگران [۳]

Mahnad Al-Waeli[1,2], Akbar Dizadji*[1], Gholam Hossein Mossahebi[1], Akbar Ahangaran[3]

[۱] دانشجوی دکتری، دانشیار و استاد گروه گیاهپزشکی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران
[۲] گروه گیاهپزشکی، دانشکده کشاورزی، دانشگاه بصره، بصره، عراق
[۳] استادیار، سازمان حفظ نباتات کشور، تهران، ایران

[1] Ph.D. student, Associate Professor and Professor, Department of Plant Protection, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran
[2] Department of Plant Protection, Faculty of Agriculture, Basrah University, Iraq
[3] Assistant Professor, Plant Protection Organization, Tehran, Iran

* نویسنده مسئول، پست الکترونیکی: adizaji@ut.ac.ir

(تاریخ دریافت: ۹۷/۲/۲۸ - تاریخ پذیرش: ۹۷/٤/۲٦)

چکیده

ویروس پیچیدگی برگ زرد گوجه فرنگی (Tomato yellow leaf curl virus, TYLCV) از مهمترین بیمارگرهای مناطق گرمسیری و نیمه گرمسیری است. طی سال های ۱۳۹٦- ۱۳۹۵، تعداد ۳۹۳ نمونه گوجه فرنگی دارای علائم بیماری پیچیدگی برگ زرد گوجه فرنگی از شش استان مختلف عراق جمع آوری شد. در آزمون ساندویچ دوطرفه الایزا ((DAS-ELSA، ۵۵ نمونه (١٤ درصد) با آنتی بادی های اختصاصی TYLCV واکنش مثبت نشان داد. پس از استخراج دی. ان. ا کل، آلودگی ۲۱ (از ۵۵) نمونه با واکنش مثبت در الایزا به TYLCV با واکنش زنجیره ای پلیمراز (PCR) تایید و ژن پروتئین پوششی V1 ١٦جدایه از ویروس با استفاده از جفت آغازگر اختصاصی تکثیر، همسانه سازی و تعیین ترادف شد. براساس همردیف سازی ترادف نوکلئوتیدی، بالاترین میزان یکسانی نوکلئوتیدی ژن پروتئین پوششی این جدایه ها (۹۹/٦- ۹٤/۱ درصد) با جدایه های این ویروس از کویت (KJ830841) و عراق (JQ354991) تعیین شد. در تحلیل تبارزایی براساس ترادف نوکلئوتیدی V1، جدایه های عراق، بدون ارتباط با منشا جغرافیایی، در دو گروه و چهار زیرگروه قرار گرفتند. شاخص های تنوع ژنتیکی براساس این ناحیه ژنومی ویروس حاکی از وجود تنوع ژنتیکی زیاد در زیرجمعیت های عراقی بود. نسبت جانشینی های نامترادف به جانشینی های مترادف Pi(a)/Pi(s) کمتر از یک (۱>) این ژن، نشان دهنده تاثیر فشار انتخابی منفی روی آن می باشد. همچنین جریان ژنی کمی بین دو زیرجمعیت مرکزی و جنوبی عراق تعیین شد که نشان دهنده تمایز ژنتیکی جدایه ها در این نواحی از کشور می باشد. این مطالعه اولین پژوهش در خصوص بررسی پراکنش جغرافیایی و تنوع ژنتیکی TYLCV در عراق می باشد. تعیین ترادف نوکلئوتیدی ژنوم کامل جدایه ها به منظور تعیین سویه ویروس ضروری می باشد.

واژه‌های کلیدی

بگوموویروس،
تنوع ژنتیکی،
عراق،
گوجه فرنگی

[DOR: 20.1001.1.25885073.1397.7.1.6.5]

Introduction

Tomato yellow leaf curl disease (TYLCD) is a most popular destructive diseases of tomato (*Solanum lycopersicum*) in tropical and subtropical regions. TYLCD caused by at least 10 virus species belong to *Begomovirus* genus in the *Geminiviridae* Family (Brown *et al.,* 2015). *Tomato yellow leaf curl virus* (TYLCV) is the most widespreaded of these begomoviruses (Lefeuvre *et al.,* 2010). TYLCV genome consists of one or two single-stranded circular DNA molecules (DNA-A and DNA-B) encoding for two open-reading frames (ORF) on the virion-sense (V1 and V2) and four on the complementary- sense strand (C1-C4) (Pradhan *et al.,* 2017). This virus is critically transmitted by *Bemisia tabaci* in circulative persistent manner, seeds, and infected dicotyledonous plants such as tomato, pepper, cucumber, watermelon, and several weeds (Mnari-Hattab *et al.,* 2014; Shirazi *et al.,* 2014; Kil *et al.,* 2016; Quiñones *et al.,* 2002; Reina *et al.,* 1999; Papayiannis *et al.,* 2011). TYLCV infected tomato plants showing symptoms like, severe leaf curling, yellowing, and stunting (Shirazi *et al.,* 2014). The incidence of TYLCV may reach up to 100% and can cause economic losses in the range of 50-90% especially when the virus infects tomato plants in early growing stage (Al-Ani *et al.,* 2011).

The spreading of TYLCV in the eastern Mediterranean region had been limited till 1980 (Lefeuvre *et al.,* 2010). Later, the virus was reported worldwide in many countries such as America, Africa, Middle East, Iran, Japan, Australia, and China (Czosnek *et al.,* 1990; Cohen & Antignus, 1994; Hajimorad *et al.,* 1996; Czosnek & Laterrot, 1997; Polston & Anderson, 1997; Shirazi *et al.,* 2014). According to the Lefeuvre *et al.,* (2010), the first appearance of TYLCVs most probably occured in some place in the Middle East between 1930s and 1950s and the worldwide spread was initiated in the 1980s. (Lefeuvre *et al.,* 2010).

For the first time, TYLCV was identified in Iraq by Makkouk (1978), and in 2013 the complete genome sequence was determined from tomato plants of Iraq (Al-Kuwaiti *et al.,* 2013). It has been suggested that Iran and its neighboring regions were both the center of TYLCV variation and the site of the most intensive current TYLCV progression (Lefeuvre *et al.,* 2010). Although Iraq and Iran as two neighbors with mutual borders, there is not enough data on TYLCV isolates and strains distribution in Iraq. Genetic variability of plant viruses is crucial for comprehension of virus development and the genetics of virus-host plant interactions and can be helpful in the planning of management strategies (García-Andrés *et al.,* 2007). The aims of the present study were to determine the occurrence and genetic diversity of TYLCV isolates in tomato fields of Iraq, study genetic structure of TYLCV among subpopulations of Iraq and other countries and evaluate phylogenetic and genetic relationships of TYLCV isolates based on theV1 gene sequences.

Materials and Methods

Sampling

During 2014-2015, a total of 393 symptomatic tomato samples showing yellowing, leaf curl and stunting appearance (Fig. 1) were collected from 47 tomato fields located in the main tomato producing regions in Iraq, include Karbala, Babil, Najaf, Qadisiyah, Dhi-Qar, and Basrah (Fig. 2). In addition, four different regions including Zubair, Garmat-Ali, Abulkhaseeb, and Safuan were selected for sampling in Basrah.

DAS-ELISA

All samples were investigated for TYLCV infection using commercial complete kit (Bioreba AG, Switzerland) in double antibody sandwich enzyme–linked immuno sorbent assay (DAS-ELISA) according to the manufacturer's instruction (Bioreba, Basel, Switzerland). Samples were considered positive if the mean absorbance values were equal or greater than three times of those for the absorbance of the negative control. All buffers as well as positive and negative controls used in these assays were provided by the company.

PCR amplification, cloning and sequencing

Based on ELISA results, total DNA was extracted from 200 mg leaf tissue of all ELISA positive samples using CTAB method according to Lodhi *et al.,* (1994), then used to amplify an 789 bp fragment covering full length of V1 (CP) gene by using a pair of degenerate primers, V1 (CP) Forward 5'GAATTCATGTCGAAGCGWCCMGGCGA3' and V1 (CP) Reverse 5'GAATTCTTAATTTKRTAYTGAATCATAGAA3' (with minor modifications, Kim *et al.,* 2011). The possible occurrence of second genomic component DNA-B and DNA-β satellites in leaf samples were tested by using specific primer pairs PBL1v2040/PCRc1 and Beta01/Beta02 respectively (Rojas *et al.,* 1993; Briddon *et al.,* 2002). The extracted PCR products were ligated into a pTG19-T cloning vector (Vivantis, Malaysia) or pJET 1.2/blunt cloning vector (CloneJET PCR cloning kit; Thermo Scientific, Germany). Following extraction of plasmid by GeneAll Spin Miniprep kit (GeneAll kit, South Korea) according to the manufacturer's protocol, the recombinant plasmids containing target DNA were investigated by digestion reaction or PCR colony amplification using (M13 forward/T7 reverse or pJET1.2 F/pJET1.2 R) vector universal primer pairs. The verified plasmids were subjected to nucleotide sequencing in both directions using vector universal primer pairs which was done by Bioneer Inc. (South Korea).

Table 1. Characteristics of TYLCV isolates used in this study

Isolates	Accession no.	Collection date	Country
IRQ	JQ354991	2012	Iraq
Iran	AJ132711	1998	Iran
Ir2	EU085423	2008	Iran
Kahnooj	EU635776	2009	Iran
TOB	KT990213	2015	Iran
BOU	GU076454	2009	Iran
OM	GU076453	2009	Iran
Iran-IL	GU076446	2009	Iran
Mild	EU143745	2007	Jordan
cucumber	EF433426	2007	Jordan
Mild-Jordan	EF158044	2006	Jordan
Jordan	EF054893	2006	Jordan
Homra	JX444575	2012	Jordan
IL-JO	GQ861426	2009	Jordan
LBa4	EF185318	2006	Lebanon
Ra3	EF051116	2006	Lebanon
Egypt	AY594174	2004	Egypt
KISR-2	KJ830841	2014	Kuwait
KISR-4	KR108214	2015	Kuwait
Tm-8	HG969282	2014	Oman
Tm-3	HG941651	2014	Oman
Tom-92	LN680630	2014	Oman
Tom-96	LN680631	2014	Oman
GJ	AB636409	2011	South Korea
Goseong	JN680149	2011	South Korea
Is:Sz	AB110218	2003	Japan
SS11	JQ867092	2012	China
LNYK	KJ754194	2014	China
LN-17	KU517848	2016	China
SHIJIAZHUANG	KF612971	2013	China
KSCEg1	JX456643	2012	China
ISRAEL CN:SH	GU434144	2012	China
Mild	AF105975	1998	Portugal
Mild	AF071228	1998	Spain
Florida	AY530931	2004	USA
IQ:Dq-1	MF429928 (This study)	2015	Iraq-Dhi-Qar
IQ:Dq-2	MF429929 (This study)	2015	Iraq-Dhi-Qar
IQ:Dq-6	MF429930 (This study)	2015	Iraq-Dhi-Qar
IQ:Ka-4	MF429934 (This study)	2015	Iraq- Karbala
IQ:Ka-5	MF429935 (This study)	2015	Iraq- Karbala
IQ:Ka-42	MF429936 (This study)	2014	Iraq-Karbala
IQ:Ba-Kh2	MF429937 (This study)	2015	Iraq- Basrah
IQ:Ba-Kh6	MF429938 (This study)	2014	Iraq- Basrah
IQ:Ba-Sf2	MF429939 (This study)	2015	Iraq-Basrah
IQ:Ba-Sf11	MF429940 (This study)	2015	Iraq- Basrah
IQ:Ba-91	MF429942 (This study)	2014	Iraq- Basrah
IQ:Ba-100	MF429943 (This study)	2014	Iraq- Basrah
IQ:Ba-Zu53	MF429944 (This study)	2014	Iraq- Basrah
IQ:Ba-Zu82	MF429945 (This study)	2014	Iraq- Basrah
IQ:Na-4	MF429946 (This study)	2015	Iraq-Najaf
IQ:Na-32	MF429948 (This study)	2014	Iraq-Najaf
TYLCMLV	FM212663	2008	Cameroon

[DOR: 20.1001.1.25885073.1397.7.1.6.5]

Phylogenetic and genetic diversity analysis

DNA STAR Lasergene (SEQMAN and EDITSEQ, version 10) was used to assemble the contigs based on forward and reverse nucleotide sequence of amplicons The attained nucleotide sequences were compared with corresponding sequences by the BLASTn analysis software (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple alignment of the CP nucleotide sequences of 16 Iraqi TYLCV isolates with other TYLCV isolates which have retrieved from the GenBank (Table 1) was performed by MEGA 7 sequence analysis package and with the default parameters (Tamura *et al.*, 2013; Edgar, 2004). Model evaluation was carried out for the nucleotide sequence dataset within MEGA 7 and genetic distances of all the isolates within and between subpopluations were estimated by Kimura2-parameter (K2P) methods in MEGA 7. Phylogenetic tree based on the CP nucleotide sequences was constructed by Maximum-likelihood (ML) method implemented in MEGA 7 The robustness of the phylogenetic trees was assessed by 1000 bootstrap replicates.

Genetic diversity of TYLCV coat protein gene was appraised within and between variant geographical subpopulations (table 3). Genetic diversity analysis were evaluated by DnaSP version 5.0 program (Rozas *et al.*, 2003).

In order to study the natural selection, the proportion of synonymous substitutions per synonymous site (Pis) and the proportion of non-synonymous substitutions per non-synonymous site (Pia) were calculated using DnaSP 5.0 program by the method described by Nei & Gojobori (1986). The Pi(a)/Pi(s) ratio showed the selection pressure in progression for coat protein of the subpopulations of variant collection areas. Selection at individual codons was checked by the fixed effects likelihood (FEL) and internal fixed effects likelihood (IFEL) methods which were available in the DATAMONKEY server (http://www.datamonkey.org) (Kosakovsky-Pond & Frost, 2005). In order to classify the site as positively or negatively selections, the cut-off *P*-value was selected to be 0.1 and only selections were determined to be significant by both methods were considered as positive selections. The statistic F_{ST} (Weir & Cockerham, 1984) was used to estimate inter-subpopulation genetic differntation and the gene flow by using DnaSP 5.0 program.

Fig. 1. Symptoms on TYLCV infected tomato plants. A: leaf curling, B and C: yellowing, D: stunting.

[DOR: 20.1001.1.25885073.1397.7.1.6.5]

Fig. 2. Map of Iraq showing six main tomato producing provinces regions, surveyed for TYLCV presence.

Table 2. Location and the number of positive samples based on the DAS-ELISA

Geographical location		No. of fields	No. of samples	TYLCV infected No.	TYLCV infection rate (%)
Karbala (central Iraq)		9	73	11	15.0
Najaf (central Iraq)		8	60	10	16.6
Babil (central Iraq)		3	49	1	2.0
Qadisiyah (central Iraq)		3	15	0	0
Dhi-Qar (southern Iraq)		5	58	9	15.5
Basrah (southern Iraq)	Zubair	4	41	8	19.5
	Garmat-Ali	2	12	3	25.0
	Abulkhaseeb	3	29	8	27.5
	Al-jazeera	4	18	1	5.5
	Safuan	6	38	4	10.5
Total		47	393	55	13.9

Results and Discussion

Detection of TYLCV

Based on the DAS-ELISA results, 55 out of 393 symptomatic collected tomato samples (13.9%) reacted positively with TYLCV specific antibodies. The DAS-ELISA positive rates ranged from 0- 17.4% in symptomatic tomato samples collected from Qadisiyah and Basrah counties, respectively (Table 2).

Sequencing, phylogenetic and genetic diversity analyses

The expected PCR product size around 789 bp was amplified in 21 samples, while no amplicons were produced from extracts of healthy tomato plants. These 21 samples were originated from four Iraq provinces, including five samples from Karbala, two samples from Najaf, three samples from Dhi-Qar, and 11 samples from Basrah. The second genomic DNA-B and DNA-β satellite were not detected in 21 TYLCV infected tomato samples by using PBL1v2040/PCRc1 and Beta01/Beta02 primers respectively. The amplicon of 16 (out of 21) TYLCV isolates were selected based on the geographical location (Fig. 2, Table 1). Their nucleotide sequences were determined and compared with other sequences in the GenBank. Multiple alignments of TYLCV coat protein genes showed high nucleotide sequence identity as 95.3- 99.8% among 16 new TYLCV isolates obtained in this study.

Additionally, new TYLCV isolates in present study shared the highest nucleotide sequence identity of 94.1-99.6% in CP gene with Kuwait and Iraq isolates. Model evaluation showed that Tamura three-parameter nucleotide substitution model (Tamura, 1992) with gamma distributed rates (T92+G) among sites was the best model. In phylogenetic tree based on the CP nucleotide sequences, TYLCV isolates were divided into two major groups that group I with four subgroups and group II (Fig. 3). Interestingly, isolates originated from Iraq were scattered in both group I (subgroups I and II) and group II suggesting potentially high genetic variation among the isolates. Five out of 16 isolates, originated from different districts of Karbala, Basrah, and Najaf, placed into group I, subgroup I, among them Ka-4 and Ba-91 were categorized far from other three Ba-100, Ba-Zu53 and Na-32 which constituted a separate clade. Ba-91 isolate was closely related to an Iraqi isolate (JQ354991) which was reported in 2012. (Al-Kuwaiti *et al.*, 2013). Isolates Ka-42, Ka-5, and Na-4 were categorized in the subgroup II, which were closely beside Jordan isolate. All the Chinese isolates were grouped in the second subgroup in group I beside the other isolates from New Zealand, Mexico, Israel, Japan, Spain, and Reunion which do not follow regular spearing pattern. The subgroups III and IV contains isolates from different locations such as Jordan, Egypt, Iran, and Oman, while none of 16 Iraqi isolates were categorized in these subgroups. TYLCV isolates with Oman and Iran origin were scattered in group I (subgroups III and IV) and group II. In group II, eight Iraqi isolates originated from Basrah and Dhi-Qar were categorized as an unique clade closely related with an isolate from Kuwait (KJ830841) with a strong bootstrap support (Fig 3).

Table 3. Genetic variability determinants of TYLCV isolates using DnaSP.

Subpopulation	No. of Seq.	S	Eta	Pi	K	h	hd	θ-W	Pi(a)	Pi(s)	Pi(a)/Pi(s)
Iraq(Central & Southern)	17	66	72	0.026	20.61	16	0.99±0.023	0.025	0.0090	0.087	0.103
South of Iraq	11	49	52	0.020	16.21	10	0.98 ± 0.046	0.021	0.0069	0.069	0.100
Central Iraq	6	38	38	0.020	15.66	6	1.00±0.126	0.022	0.0079	0.062	0.127
Kuwait-Oman	6	50	57	0.029	22.93	6	1.00±0.096	0.028	0.0096	0.097	0.098
Iran	7	68	74	0.035	27.28	7	1.00±0.076	0.035	0.0093	0.123	0.075
Jordan - Lebanon –Egypt	9	47	48	0.017	11.88	9	1.00±0.052	0.022	0.0041	0.053	0.077
China- Japan- Korea	9	31	33	0.013	10.22	9	1.00±0.052	0.014	0.0049	0.041	0.119

S, Number of polymorphic (segregating) sites Eta, Total number of mutations, Pi, Nucleotide diversity, K, Average number of nucleotide difference, h, Number of haplotypes, hd, Haplotype diversity, θ-W, Theta-W, Pis, Synonymous site, Pia, Nonsynonymous site

Table 4. F_{ST} values for pairs of geographical subpopulations of TYLCV

	A	B	C	D	E
B	0.37220				
C	0.23028	0.26427			
D	0.44658	0.010577	0.24710		
E	0.22028	0.25531	0.01439	0.23436	
F	0.53605	0.26959	0.28524	0.24174	0.27923

A- Southern Iraq subpopulation, B- Central Iraq subpopulation, C- Iran subpopulation
D- Jordan- Lebanon-Egypt subpopulation, E- Kuwait-Oman subpopulation, F- China-Korea-Japan subpopulation

The effective number of codons (ENC) was calculated as 51.239, indicated that codons usage bias was not well remarkable in these subpopulations and was seemingly maintained at a stable level. As shown in Table 3, the calculated Pi and hd values for all Iraqi isolates (0.026 and 0.99, respectively) likely for both Iraqi subpopulations (0.02 and 0.98- 1, respectively) were higher than 0.005 and 0.5, indicated a high genetic diversity in Iraqi subpopulations of TYLCV. The proportion between nucleotide variation values in synonymous and nonsynonymous positions (Pi(a)/Pi(s)=ω) represented selection pressure in evolution for coat protein gene of the geographical sub-populations of TYLCV. The maximum and minimum values of Pi(a)/Pi(s) ratios (0.127 and 0.075, respectively) were <1.0 (Table 3), indicated that this gene was under negative selection (Yang *et al.,* 2014). The computation of selection at individual codons by FEL and IFEL methods and the calculated ratio of Pi(a)/Pi(s) showed a little proportion of the sites were found to had evolved under negative selection. The Pi(a)/Pi(s) ratio was demonstrated low value for the TYLCV coat protein gene (P-value<0.1).

Computation of F_{ST}, the coefficient used to assess the level of genetic differentiation between pairs of subpopulations, showed that the values of F_{ST} among central Iraq with, Iran, Kuwait-Oman, Jordan-Lebanon-Egypt and China-Japan-Korea subpopulations were less than 0.33. In contrast, F_{ST} values between southern Iraq with central Iraq, Jordan-Lebanon-Egypt and China-Korea-Japan subpopulations were 0.37, 0.44 and 0.53, respectively (Table 4) as a proof of genetic differentiation.

Tomatoes are planting in most parts of Iraq, as a greenhouse and open field crop. The average production of the tomatoes has been steadily increased over the time, Basrah governorate was also known as an "efficient tomato cluster in Iraq" (Anonymous, 2007). Despite the occurrence and importance of viral infections in tomato plants, there is a little knowledge regarding the genetic information of the Iraqi TYLCV isolates. Now Iraq is one of most popular tomato production region in the world which explains the fact that further studies need about TYLCV in this country (Anonymous, 2007). In the present investigation, despite our expectations based on serious TYLCD symptom observations (leaf curling, yellowing, and stunting) in tomato fields, the overall DAS-ELISA positive rate was determined as 14% in central and southern regions of Iraq. Based on the DAS-ELISA results, unlike maximal ELISA positive rate (17.4%) among samples collected from Basrah, no samples with positive reaction were found with TYLCV specific antibody in Qadisiyah. The overall low ELISA positive rate of TYLCD symptomatic tomato samples in Iraq is in agreement with Yazdani-Khameneh *et al.,* (2016) results, which showed only 6.8% (10 out of 148) of symptomatic tomato samples in Iran reacted positively in ELISA with TYLCV antibodies. On the other hand, a lower incidence of TYLCV confirmed by PCR among ELISA positive samples (5.3%, 21 out of 393) in this study, is in accordance with detection of TYLCV in 12% of TYLCD symptomatic plants in Baghdad of Iraq (Al-Kuwaiti *et al.,* 2013). The overall low presence of TYLCV in TYLCD symptomatic tomato samples might be explained by the presence of other virus species, including other begomoviruses, which was described previously (Malekzadeh *et al.,* 2011). Inconsistent results of serological and molecular detection of TYLCV could be partly due to serological relationships among different *Begomovirus* species, which was reported before (Rishi, 2004; Yazdani-Khameneh *et al.,* 2016). In agreement with this idea, we detected *Tomato leaf curl Palampur virus* (ToLCPMV, MF038135) among TYLCV-ELISA positive samples (data not shown).

Our results showed that TYLCV was not limited to the central regions of Iraq such as Baghdad (Al-Kuwaiti *et al.,* 2013), Karbala and Najaf (present study) and its geographical distribution had been extended to southern regions of the country such as Basrah province, with highest positive rate in ELISA (17.4%) and PCR confirmation (7.9%). TYLCV outbreaks are dependent upon optimum weather conditions, vector and plant species (Fang *et al.,* 2013). Probably, *B. tabaci* plays the crucial role in TYLCV outbreaks in Iraq. So far, *B. tabaci* has been reported from Iraq but the insect distribution has not been studied clearly so far. Although TYLCV seed transmission of TYLCV-IL has been recently reported with the rate of 20- 100% (Kil *et al.,* 2016).

Following the serological and molecular screening of tomato samples, sixteen TYLCV-CP sequences were obtained from central and south of Iraq. In phylogenetic analysis, all TYLCV isolates originated divided into two major groups based on CP sequences as described before (Lefeuvre *et al.,* 2010).

[DOR: 20.1001.1.25885073.1397.7.1.6.5]

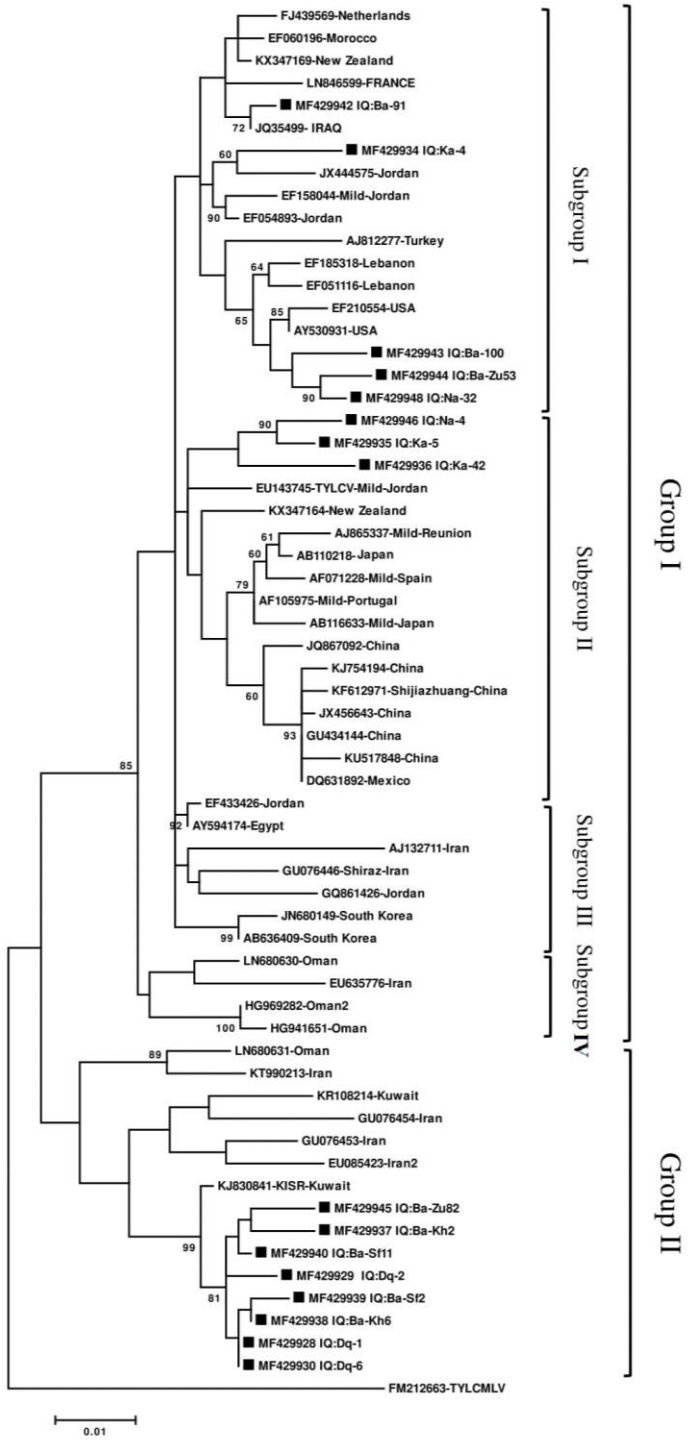

Figure 3. Phylogenetic tree based on nucleotide sequences of TYLCV V1 (CP) constructed using MEGA 7 with 1000 bootstrap replications. Only bootstrap values greater than 60% are shown. The tree was rooted with the CP nucleotide sequences of *Tomato yellow leaf curl Mali virus* (TYLCMV) as an out-group species.. The characterized isolates in this study were labeled in a black square.

The virus isolates obtained from central Iraq (Karbala, Najaf) and three isolates from southern Iraq (Basrah counties) grouped into two clades in subgroup I of group I. Contrariwise, the rest of southern Iraq originated TYLCV isolates (except three ones fell into group I) were organized as a separate clade in group II.

Our phylogenetic analysis showed that there was no correlation between isolates regarding the geographical

origin, in contrast with the results of Shirazi *et al.*, (2014). It is noteworthy to say that southern Iraqi isolates are more related with isolates from Iran, Oman, and Kuwait in comparison with the others isolates (Fig. 3), suggesting that these isolates were introduced to Iraq from southern borderlines.

Until now, nine TYLCV strains/variants , including TYLCV-IL from Israel and Iran, TYLCV-Mld from Israel, TYLCV-IR, TYLCV-Bou and TYLCV-Ker from Iran, TYLCV-OM from Oman and Iran, TYLCV-Gez from Sudan, TYLCV- KISR from Kuwait and TYLCV- PK from Pakistan, have been characterized worldwide (Lefeuvre *et al.*, 2010; Al-Ali *et al.*, 2015; Zaidi *et al.*, 2017). TYLCV-IL and TYLCV-Mld have the broadest geographical ranges of any TYLCV strains worldwide. Iran has the greatest number of strains (five out of the nine), this fact makes Iran possibly the center of TYLCV variant in the world (Lefeuvre *et al.*, 2010). Based on TYLCV coat protein gene phylogenetic tree, it seems that Iraqi isolates located in group I were close to the TYLCV-IL and TYLCV-Mld strains, the rest of isolates were in group II located around TYLCV-KISR from Kuwait. Ultimately, the assessment of the full length viral genome sequences of Iraqi isolates is needed in order to confirm their strains.

The Pi and hd values of both Iraqi subpopulations (Central and Southern) indicated a high genetic diversity in Iraqi subpopulations of TYLCV. The ratio between nonsynonymous and synonymous substitutions was less than 1, indicated that the CP gene was subjected to negative selection and was strongly conserved. This result was compatible with the findings obtained by Yang *et al.*, (2014). Estimated Pi(a)/Pi(s) values for central and southern Iraqi subpopulations were 0.127 and 0.10, respectively, indicated that central subpopulation was subjected to more intensive purifying selection than southern Iraqi subpopulation. Genetic variation within southern Iraq was the same as that of central Iraq (Pi= 0.02), while West Asia Iran and East Asia (China- Korea- Japan) subpopulations showed the highest and lowest variability with Pi value of 0.035 and 0.013, respectively. As shown in Table 3, Iran, (Kuwait-Oman) and Iraq subpopulations respectively were the most divergent TYLCV-subpopulations in this study. Our results were in accordance with previously reported studies for TYLCV isolates with Iran and Oman origins (Hosseinzadeh *et al.*, 2014; Shirazi *et al.*, 2014). Overall, these results were in agreement with Lefeuvre *et al.*, (2010), described emergence and universal spread of TYLCV were extremely rapid and Middle East in general, and the area surrounding Iran in specific, are probably the onging and past centers of current TYLCV diversification.

Low gene flow was observed between southern and centeral Iraq subpopulations (F_{ST} value of 0.37), demonstrating genetic differentiation among Iraqi subpopulations. Alternatively, the high gene flow existed among southern Iraq and Iran, Kuwait-Oman subpopulations, with F_{ST} values less than 0.33 (Yang *et al.*, 2014) in contrast with Jordan-Lebanon-Egypt and China-Korea-Japan subpopulations (Table 4).

According to our knowledge this is the first study that encountered the genetic variation of TYLCV in Iraq.

Acknowledgements

The authors would like to acknowledge the financial supports of University of Tehran and Basrah University for this research.

References

Al-Ali, E.H.M., Al-Hashash, H.K., Ben-Hejji, A.H., Al-Shayjji N. and Al-Aqeel, H.A. 2015. First complete genomic characterization and phylogeny of a new recombinant of *Tomato yellow leaf curl virus* (genus *Begomovirus*, family *Geminiviridae*) from Kuwait. *Archives of Virology*, 160:1823-1826.

Al-Ani, R.A., Adhab, M.A., Hamad, S.A. and Diwan, S.N. 2011. *Tomato yellow leaf curl virus* (TYLCV), identification, virus vector relationship, strains characterization and a suggestion for its control with plant extracts in Iraq. *African Journal of Agricultural Research*, 6(22):5149-5155.

Al-Kuwaiti, N., Otto, B., Collins, C., Seal, S. and Maruthi, M. 2013. Molecular characterisation and first complete genome sequence of *Tomato yellow leaf curl virus* (TYLCV) infecting tomato in Iraq. *New Disease Reports*, 27:17-17.

Anonymous. 2007. Iraq private sector growth and employment generation tomato paste in Iraq. USAID-IZDIHAR February 8:pp25.

Briddon, R.W., Bull, S.E., Mansoor, S., Amin I. and Markham P.G. 2002. Universal primers for the PCR-mediated amplification of DNA β, a molecule associated with some monopartite begomoviruses. *Molecular Biotechnology*, 20(3):315–318.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-Olivé, E., Briddon, R.W., Hernández-Zepeda, C., Idris, A. and Malathi, V.G. 2015. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Archives of Virology*, 160(6):1593-1619.

Cohen, S. and Antignus,Y. 1994. *Tomato yellow leaf curl virus*, a whitefly-borne *Geminivirus* of tomatoes. In Advances in disease vector research. Springer, New York, pp. 259-288.

Czosnek, H., Navot, N and Laterrot, H. 1990. Geographical distribution of *Tomato yellow leaf curl*

virus. A first survey using a specific DNA probe. *Phytopathologia Mediterranea*, 29: 1-6.

Czosnek, H. and Laterrot, H. 1997. A worldwide survey of tomato yellow leaf curl viruses. *Archives of Virology*, 142(7):1391-1406.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research***,** 32:1792–1797.

Fang, Y., Jiao, X., Xie, W., Wang, S., Wu, Q., Shi, X., Chen, G. , Su, **Q., Yang, X., Pan, H. and Zhang, Y**. 2013. *Tomato yellow leaf curl virus* alters the host preferences of its vector *Bemisia tabaci*. *Scientific Reports -UK* , 3:2876-80.

García-Andrés, S., Accotto, G.P., Navas-Castillo, J. and Moriones, E. 2007. Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. *Virology*, 359(2):302-312.

Hajimorad, M.R., Kheyr-Pour, A., Alavi, V., Ahoonmanesh, A., Bahar. M., Rezaian, M.A. and Gronenborn, B. 1996. Identification of whitefly transmitted *Tomato yellow leaf curl geminivirus* from Iran and a survey of its distribution with molecular probes. *Plant Pathology*, 45(3):418-425.

Hosseinzadeh, M.R., Shams-Bakhsh, M., Osaloo, S.K. and Brown, J.K. 2014. Phylogenetic relationships, recombination analysis, and genetic variability among diverse variants of *Tomato yellow leaf curl virus* in Iran and the Arabian Peninsula: further support for a TYLCV center of diversity. *Archives of Virology*., 159(3):485-497.

Kil, E.J., Kim, S., Lee, Y.J., Byun, H.S., Park, J., Seo, H., Kim, C.S., Shim, J.K., Lee, J.H., Kim, J.K. and Lee, K.Y. 2016. *Tomato yellow leaf curl virus* (TYLCV-IL): a seed-transmissible *Geminivirus* in tomatoes. *Scientific Reports -UK*, 6: 19013.

Kim, S.H., Oh, S., Oh, T.K., Park, J.S., Kim, S.C., Kim, S.H., Kim, Y.S., Hong, J.K., Sim, S.Y., Park, K.S., Lee, H.G., Kim, K.J. and Choi,C.W. 2011. Genetic diversity of tomato-infecting *Tomato yellow leaf curl virus* (TYLCV) isolates in Korea. *Virus.Genes*, 42:117–127.

Kosakovsky-Pond, S.L. and Frost,S.D.W. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*, 21:2531–2533.

Lefeuvre, P., Martin, D.P., Harkins, G., Lemey, P., Gray, A.J., Meredith, S., Lakay F., Monjane, A., Lett, J.M., Varsani, A. and Heydarnejad, J. 2010. The spread of *Tomato yellow leaf curl virus* from the Middle East to the world. *PLoSPathog*, 6(10): e1001164.

Lodhi, M.A., Ye, G.N., Weeden, N.F. and Reisch, B.I. 1994. A simple and efficient method for DNA extraction from grapevine cultivars and *Vitis* species. *Plant Molecular Biology Reporter*, 12(1):6-13.

Makkouk, K.M 1978. A study on tomato viruses in the Jordan Valley with special emphasis on tomato yellow leaf curl. *Plant Disease Reporter*, 62(3):259-262.

Malekzadeh, S., Bananej, K., Vahdat, A. 2011. Serological and molecular identification of *Tomato yellow leaf curl virus* in Khuzestan province of Iran. *Phytopathologia Mediterranea*.50:303−309.

Mnari-Hattab, M., Zammouri, S., Pellegrin, F. and Gauthier, N. 2014. Natural occurrence of *Begomovirus* recombinants associated with tomato yellow leaf curl disease co-existing with parental viruses in tomato crops and weeds in Tunisia. *Journal of Plant Pathology*, 96(1):195-200.

Nei, M. and Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3:418-426.

Papayiannis, L.C., Katis, N.I., Idris, A.M. and Brown, J.K. 2011. Identification of weed hosts of *Tomato yellow leaf curl virus* in Cyprus. *Plant Disease*, 95(2):120-125.

Polston, J.E. and Anderson, P.K. 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. *Plant Disease*, 81(12):1358-1369.

Pradhan, B., Van Tien, B., Dey, V. N. and Mukherjee, S.K. 2017. Molecular Biology of *Geminivirus* DNA Replication. Viral Replication, Avid Science, Chapter 1:pp.2-31

Quiñones, M., Fonseca, D., Martinez, Y. and Accotto, G.P. 2002. First report of *Tomato yellow leaf curl virus* infecting pepper plants in Cuba. *Plant Disease*, 86(1):73-73.

Reina, J., Morilla, G., Bejarano, E.R., Rodríguez, M.D. and Janssen, D. 1999. First report of *Capsicum annuum* plants infected by *Tomato yellow leaf curl virus*. *Plant Disease*, 83(12):1176-1176.

Rishi N. 2004. Current status of begomoviruses in the Indian subcontinent. *Indian Phytopathology*, 57:396–407.

Rojas, M. R., Gilbertson, R. L., Russell, D. R. and Maxwell, D. P. 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease*, 77:340-347.

Rozas, J., Sa´nchez-DelBarrio, J.C., Messeguer, X. and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 19:2496–2497.

Shirazi, M., Mozafari, J., Rakhshandehroo, F. and Shams-Bakhsh, M. 2014. Genetic diversity, host range, and distribution of *Tomato yellow leaf curl virus* in Iran. *Acta Virologica*, 58:128-136.

[DOR: 20.1001.1.25885073.1397.7.1.6.5]

Tamura, K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Molecular Biology and Evolution*, 9:678–87.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30 (12): 2725– 2729.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38:1358–1370.

Yang, X., Zhou, M., Qian, Y., Xie, Y. and Zhou, X. 2014. Molecular variability and evolution of a natural population of *Tomato yellow leaf curl virus* in Shanghai, China. *Journal of Zhejiang University-Science B*, 15(2):133-142.

Yazdani-Khameneh, S., Aboutorabi, S., Shoori, M., Aghazadeh, A., Jahanshahi, P., Golnaraghi, A. and Maleki, M. 2016. Natural Occurrence of *Tomato leaf curl New Delhi virus* in Iranian Cucurbit Crops. *The Plant Pathology Journal*, 32(3):201-8.

Zaidi, S.S., Shakir, S., Farooq, M., Amin, I. and Mansoor, S. 2017. First Report of a Novel Strain of *Tomato yellow leaf curl virus* Causing Yellow Leaf Curl Disease on Cluster Bean in Pakistan. *Plant Disease*, 101(6):1071-1071.

Genetic Engineering and Biosafety Journal
Volume 7, Number 1

Phylogenetic analysis and genetic variation of *Tomato yellow leaf curl virus* based on the V1 gene in Iraq

Mahnad Al-Waeli[1,2], Akbar Dizadji*[1], Gholam Hossein Mossahebi[1], Akbar Ahangaran[3]

[1] Ph.D. student, Associate Professor and Professor, Department of Plant Protection, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran
[2] Department of Plant Protection, Faculty of Agriculture, Basrah University, Iraq
[3] Assistant Professor, Plant Protection Organization, Tehran, Iran

* Corresponding author: Akbar. Dizadji, E- mail: adizaji@ut.ac.ir

Abstract

Tomato yellow leaf curl virus (TYLCV) is a supreme pathogen in tropical and subtropical areas. During 2014-2015, a total of 393 tomato leaf tissue samples showing tomato yellow leaf curl disease (TYLCD) symptoms were collected from six different provinces of Iraq. In double antibody sandwich enzyme-linked imunoserbent assay (DAS-ELSA) 55 out of 393 samples (14%) reacted positively with TYLCV-specific antibodies. Following total DNA extraction, the presence of TYLCV was verified in 21 out of 55 ELISA positive samples by PCR and coat protein gene (V1) of 16 TYLCV isolates was amplified using specific primers and cloned. Nucleotide sequences of Iraqi TYLCV isolates of V1 gene showed the highest nucleotide sequences identity of 94.1-99.6% in CP gene with Kuwait (KJ830841) and Iraq (JQ354991) isolates. In phylogenetic analysis based on V1 nucleotide sequences, Iraqi isolates were separated into two main groups and four subgroups, without correlation with geographical origin. Genetic diversity parameters based on V1 nucleotide sequences indicated a high genetic variability in Iraqi populations of TYLCV. The proportion of non-synonymous to synonymous nucleotide diversity was <1.0, indicating that this gene is under negative selection. A low gene flow was observed between southern and centric Iraq subpopulations, demonstrating genetic differentiation among Iraqi subpopulations. This is the first study dealing with the distribution and genetic variation of TYLCV in Iraq. Fulllength viral genome sequencing of TYLCV isolates of Iraq is necessary to differentiate virus strain(s) in Iraq.

Key words: Begomovirus, Genetic diversity, Iraq, Tomato.

[DOR: 20.1001.1.25885073.1397.7.1.6.5]