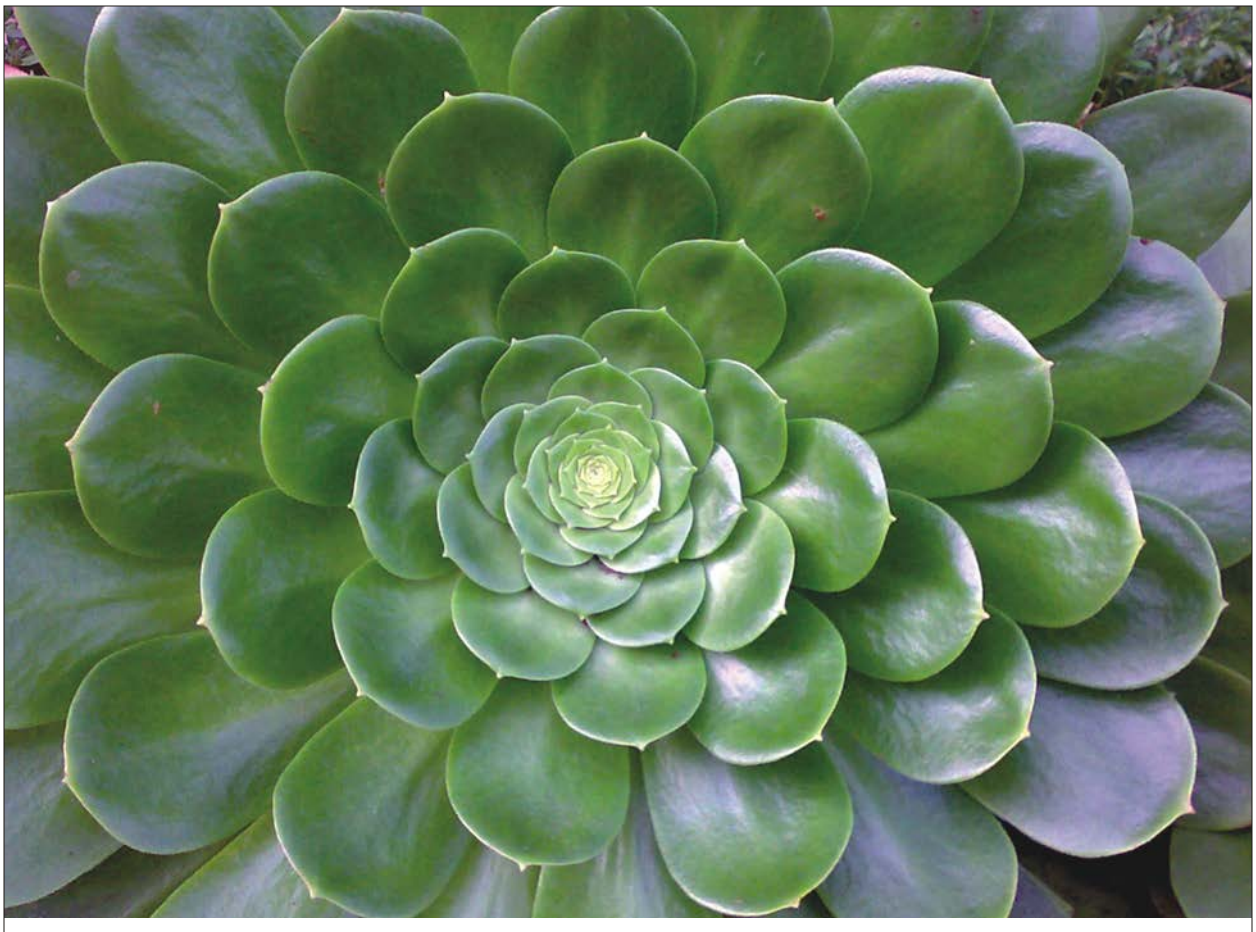

Journal of
Plant Sciences

ISSN 1816-4951

Academic
Journals Inc.

www.academicjournals.com

OPEN ACCESS Journal of Plant Sciences

ISSN 1816-4951
DOI: 10.3923/jps.2017.5.16

Research Article
Genetic Diversity of Tomato Yellow Leaf Curl-like Viruses in Saudi Arabia

[1,2]Adel Abdelsabour Rezk and [3]Khalid Abdullah Alhudaib

[1]Department of Biotechnology, College of Agricultural and Food Science, King Faisal University, Saudi Arabia
[2]Department of Virology, Plant Pathology Institute, Agriculture Research Center, Egypt
[3]Department of Plant Protection, College of Agricultural and Food Science, King Faisal University, 31982 Al-Ahsa, Saudi Arabia

Abstract

Background and Objective: Tomato Yellow Leaf Curl Disease (TYLCD) is one of the most devastating plant diseases in the different regions of the world and caused by a complex of several virus species in genus *Begomovirus,* family Geminiviridae. This study aimed to detect distribution genetic diversity of TYLCV in different location in Saudi Arabia. **Materials and Methods:** Leaf samples of 476 of tomato plants showing symptoms of tomato yellow leaf curl disease were collected during 2012-2013 from six different locations in Saudi Arabia (Alqasim, Alhasa, Aljawf, Hail, Jizan and Tabuk) to study the incidence and genetic biodiversity of Tomato Yellow Leaf Curl Virus (TYLCV). The PCR analysis was used to detect and amplify TYLCV sequences using degenerate and virus-specific primers. **Results:** The TYLCV was detected in samples collected from Alqasim, Alhasa, Hail and Jizan with infection percentages of 93.5, 92, 88 and 78%, respectively; TYLCV was not detected in samples collected from Aljawf and Tabuk. The full length of the viral genome was amplified and sequenced for four isolates, one each from Alqasim, Alhasa, Hail and Jizan. Comprehensive analysis of the virus sequences were performed including sequence alignment, GC content, homology and phylogeny analysis. These analysis showed that all sequences of TYLCV were between 2771-2780 nucleotides in length encoding either 6 open reading frames in TYLCV-Has and TYLCV-Jiz, or 7 open reading frame TYLCV-Hail and TYLCV-Qas. The GC content ranged between 0.41-0.42. Nucleotide identity was up to 95% between TYLCV-Jiz, TYLCV-Hail and TYLCV-Qas but only 83% with TYLCV-Has. **Conclusion:** The TYLCV-Jiz, TYLCV-Qas and TYLCV-Hail have higher sequence identity to Tomato Leaf Curl Sudan Virus (TLCSV) and TYLCV-Has has more identity to the tomato yellow leaf curl virus of Iran.

Key words: TYLCV, tomato, *Begomovirus*, whitefly, genetic diversity, geminiviruses

Received: October 25, 2016 Accepted: November 28, 2016 Published: December 15, 2016

Citation: Adel Abdelsabour Rezk and Khalid Abdullah Alhudaib, 2017. Genetic diversity of tomato yellow leaf curl-like viruses in Saudi Arabia. J. Plant Sci., 12: 5-16.

Corresponding Author: Adel Abdelsabour Rezk, Department of Biotechnology, College of Agricultural and Food Science, King Faisal University, Saudi Arabia
Tel: 00966546387706

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

The genus *Begomovirus* belongs to family Geminiviridae and is subdivided into old and new world members. The old world members contain a monopartite genome that is about 2.8 kb, have homology with component DNA-A of the bipartite members and most of these members interact with a class of ssDNA satellite molecules known as alpha-satellites and beta-satellites. In contrast, most new world members have bipartite genome composed of two components (i.e., DNA-A and DNA-B components), each about 2.6 kb and both components are needed for infectivity[1,2]. Only a small number of bipartite *Begomoviruses* occur in the old world[3]. The primary species differentiation measure for *Begomovirus* classification is based on nucleotide sequences similarity; to be classified as a new species a newly described DNA-A component must share less than 89% nucleotide identity with another DNA-A component of a previously recognized *Begomovirus* species[4].

Tomato Yellow Leaf Curl Disease (TYLCD) is one of the most devastating plant diseases in the warm and temperate regions of the world and caused by a complex of several virus species in genus *Begomovirus*, family Geminiviridae. Collectively, these viruses are referred to as "Tomato Yellow Leaf Curl Viruses" (TYLCV) and tomato leaf curl viruses (ToLCV)[5,6]. Both are members of the old world group that have true monopartite genomes containing only a DNA-A like molecule that is sufficient to cause characteristic disease symptoms[7]. The first report of TYLCV[8] occurred in Israel in the early 1960s. Establishing clear guidelines for the analysis of full-length genomic sequences and standardized terminology for naming of new established species and strains, the International Committee on Taxonomy of Viruses (ICTV) Geminiviridae working group made proposals to improve taxonomic communications among users while leaving options for further improvements in the future that will serve the Gemini virus research community[9]. Taxonomically the viruses associated with TYLCD have been assigned to at least 6 species and 15 strains of viruses, at least 9 different virus species more or less related phylogenetically and strains of them and include tomato yellow leaf curl virus[10]. All of these viruses belong to the genus *Begomovirus* that includes geminiviruses transmitted in a circulative persistent manner by the whitefly *Bemisia tabaci* Gennadius: (Homoptera: Aleyrodidae). They typically encode six open reading frames (ORFs), four on the complementary (–) strand (C1, C2, C3 and C4) and two on the viral (+) strand (V1 and V2)[3,7].

Phylogenetic analysis of TYLCV genome sequences suggested that this virus probably arose somewhere in the Middle East between the 1930s and 1950s and that it spread globally at the beginning of the 1980s after the appearance of two strains (TYLCV-Mld and TYLCV-IL)[11,12]. A high degree of TYLCV diversity and recombination between strains and *Begomovirus* species has been reported in the Middle East[13]. The region around Iran appears to be the current center of TYLCV diversity and the site of the most intensive ongoing TYLCV evolution. However, further analysis also indicates that this region is epidemiologically isolated suggesting that novel TYLCV variants found there are probably not direct global threats[11]. The TYLCV was observed for first time in Saudi Arabia in 1978 and subsequently verified by serological and PCR methods[14]. Recently two different strains of TYLCV has been detected in pepper and common bean plants from different locations in Saudi Arabia[15]. No additional records identified the specific strain(s) of TYLCV in Saudi Arabia. Genetic diversity and regional distribution of TYLCV did not study in this area so far. Thus; studies were conducted to detect and identify TYLCV in different locations of Saudi Arabia and compare their genomic sequences with other strains available in GenBank. In the present study, four isolates of TYLCV were identified in different locations within Saudi Arabia ("Alqasim", "Alhasa", "Hail" and "Jizan") and their genomic sequences were analyzed.

MATERIALS AND METHODS

Plant samples collection: Leaf samples from 476 symptomatic tomato (*Solanum lycopersicum*) plants showing TYLCD symptoms (leaf curling, yellowing and stunting), as well as asymptomatic tomato plants were collected from six locations with distinct climates in Saudi Arabia: "Alqasim" (75 samples), "Alhasa" (138 samples), "Aljawf" (48 samples), Hail (93 samples), "Jizan" (69 samples) and "Tabuk" (53 samples) (Fig. 1). Samples were collected during autumn and spring seasons of 2012 and 2013. The collected samples were either processed immediately or stored at -20°C prior to analysis.

Nucleic acid extraction: Total genomic DNA was extracted from collected tomato plants as previously described by Anfoka *et al.*[16] and Alhudiab *et al.*[17]. In brief, 50 mg leaf tissue was ground in 500 µL extraction buffer (50 mM EDTA, 100 mM tris-HCl, 500 mM NaCl, 10 mM β-mercaptoethanol) and incubated at 65°C for 10 min. After adding 0.2 volumes of potassium acetate (5 M, pH 8.0), an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added. The mixture was vortexed and clarified by centrifugation at 6708×g for 20 min. An equal volume of isopropanol was added to the aqueous layer and the mixture was incubated for

Fig. 1: Map of Saudi Arabia

Stars refer to locations of collected samples

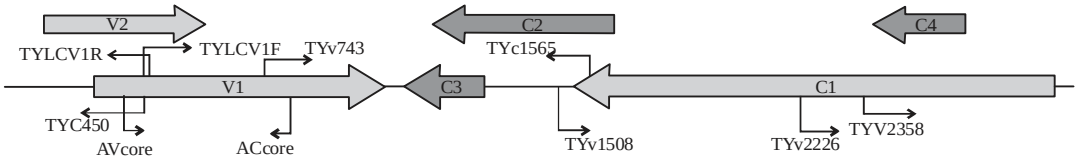

Fig. 2: A representation of the *Begomovirus* genome illustrating the anticipated PCR products obtained with the primers used in this study

The arrows show the locations of used primers to amplify the full length of Tomato Yellow Leaf Curl Virus (TYLCV)

Table 1: Primers for PCR amplification of tomato yellow leaf curl virus isolates

Primer name	Location	Sequences	Expected amplification	References
AVcore	V1-V2	GCCHATRTAYAGRAAGCCMAGRAT	~580 bp	Brown *et al.*[18]
ACcore		GGRTTDGARGCATGHGTACANGCC		
TYv2358	C1-V2-V1	TGGAAATGTGCTGACCTGGC	~870 bp	Present study
TYc450		TGGGCTTTCGGTACATGGGCCG		Present study
TYLCV1F	Full length	ACAGCCCATACAGCAGCCGTGC	~2780 bp	Present study
TYLCV1R		GCTGTCGAAGTTCAGCCTTCGGC		Present study
TYv743	V1	GTCCAGTTACCACTCCCTATGG	For sequencing	Present study
TYc1565	C2-C1	CCAAGAAGAAACCAGTCAGACG	For sequencing	Present study
TYv1508	C2-C1	TAGTATGAGCAGCCGCAGTC	For sequencing	Present study
TYv2226	C1	TCATTGGCTGATTGTTGTCCC	For sequencing	Present study

10 min at -20°C. After 15 min of centrifugation at 13148×g, the pellet was washed with 70% ethanol and resuspended in sterile deionized water.

PCR amplification of viral DNA: Extracted total DNA was used as a template for amplification by PCR with primers (Fig. 2, Table 1). The degenerate primer pair AVcore/ACcore[18] was used to determine positive samples for genus *Begomoviruses* members. The primer pair TYLCV1F/TYLCV1R (Fig. 2, Table 1) were designed according to nucleotide sequences in GenBank and based on previous data[11] to amplify full-length TYLCV sequences. Primers TYv2358 and TYc450 were used in nested

Table 2: GenBank accession numbers for tomato yellow leaf curl virus isolates sequences used for sequence analysis

Isolates	Accession No.	Country	Reference
TYLCV-Jiz	KC845301	Saudi Arabia	Present study
TYLCV-Hail	KF040453	Saudi Arabia	Present study
TYLCV-Has	KF435137	Saudi Arabia	Present study
TYLCV-Qas	KF561125	Saudi Arabia	Present study
Tomato yellow leaf curl Sudan	HE819244	Sudan	Khan *et al.*[21]
TYLCV-Mir-Sud	JN591385	Sudan	Khan *et al.* [21]
TYLCV-Jazira	AY044137	Sudan	Idris and Brown[22]
TYLCV-Shambat	AY044139	Sudan	Idris and Brown[22]
TYLCV-Had	JF919733	Sudan	Idris *et al.*[23]
TYLCV-Yemen	EF110891	Yemen	Unpublished data
TYLCV- Genaveh	GU076454	Iran	Lefeuvre *et al.*[11]
TYLCV- Minab	GU076442	Iran	Lefeuvre *et al.*[11]
TYLCV-Kahnooj1	GU076448	Iran	Lefeuvre *et al.*[11]
TYLCV-Kahnooj2	GU076449	Iran	Lefeuvre *et al.*[11]
TYLCV-Ir2	EU085423	Iran	Azizi *et al.*[24]
TYLCV-KSA46	HG530539	Saudi Arabia	Idris *et al.*[25]
TYLCV-Hwas	GU126513	Korea	Kim *et al.*[26]
TYLCV-Bos	GU325634	Korea	Kim *et al.*[26]
TYLCTHV-MM	AF206674	Myanmar	Unpublished data
TYLCV-DO	AF024715	Dominican Republic	Nakhla *et al.*[27]
TYLCV-mild Reunion	AJ865337	Reunion	Delatte *et al.*[28]
TYLCV-PR	AY134494	Puerto Rico	Bird *et al.*[29]
TYLCV-Egy	AY594174	Egypt	Abhary *et al.*[30]
TYLCV-Mild	X76319	Israel	Antignus and Cohen[31]
TYLCV-Mal	JN859135	Portugal	Monci *et al.*[32]
TYLCV-Sh10	EU031444	China	Liao *et al.*[33]
TYLCV- ZJHZ12	FN256256	China	Zhang *et al.*[34]
TYLCV-IL2	AB116631	Japan	Ueda *et al.*[35]
TYLCV- Guasave	FJ012358	Mexico	Unpublished data
TYLCV-Cuba	AJ223505	Cuba	Bejarano *et al.*[36]

*TYLCV: Tomato yellow leaf curl virus

PCR to amplify a fragment of 870 bp of TYLCV. The PCR mixture included 5 µL of 10× PCR buffer (100 mM tris-HCl, pH 8.3 15 mM $MgCl_2$, 500 mM KCl), 5 µL of 2.5 mM deoxynucleotide triphosphates each, 2 µL of 10 pM concentration of each forward and reverse primer, 2.5 U of GoTaq DNA polymerase (Promega, USA), 5 µL of template extracted genomic DNA and sterile distilled water to give a final volume of 50 µL. The PCR reactions were carried out in a thermal cycler with initial denaturation at 94°C for 2 min, followed by 35 cycles at 94°C for 1 min, 55°C for 2 min and 72°C for 2 min and final extension at 72°C for 10 min.

Cloning, sequencing and alignment analysis: The PCR products were purified from agarose gel using the Qiaquick gel extraction kit (Qiagen, Germany). The purified PCR products were cloned into pGEM-T Easy vector (Promega, USA) according to manufacturer's instructions. Plasmid DNA was prepared by the Mini preparation kit (Promega, USA) and sequenced (Macrogen Inc., Soul, Korea). Alignment analysis were performed using the online BLAST service of the National Center for Biotechnology Information (URL: http//www.ncbi.nih.gov/BLAST/). Primers TYv743, TYc1565, TYv1508 and TYv2226 were used to sequence the inner parts of the amplified DNA of TYLCV. Nucleotide sequences reported in this study were deposited in GenBank under the accession numbers KF561125 (isolate from "Alqasim"), KF435137 (isolate from "Alhasa"), KF040453 (isolate from "Hail") and KC845301 (isolate from "Jizan"). The GC content for the four isolates in this study was calculated using the Geecee software.

Phylogenetic and homology analysis: Phylogenetic tree construction was conducted from the multiple alignments using the DNAMAN software version 8 (Lynnon, Canada) using the neighbor-joining method[19] and the Jukes-Cantor distance correction method[20]. Sequences from clones of TYLCV were aligned and compared among themselves and with sequences of TYLCV available in GenBank and listed in (Table 2). The homology comparison of the candidate TYLCV sequences was carried out using the DNAMAN (version 8) software and genetic diversity was analyzed.

RESULTS AND DISCUSSION

Symptoms and geminivirus detection: In 2012 and 2013, leaf samples were collected from 476 tomato plants that showed typical TYLCD symptoms from six different locations in Saudi Arabia. Leaf symptoms included yellowing, curling,

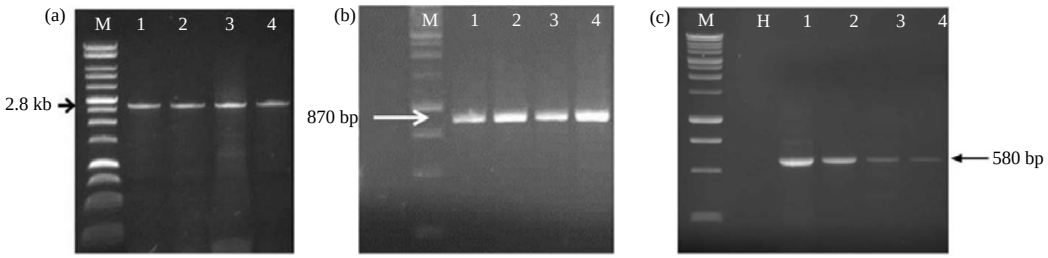

Fig. 3(a-c): Agarose gel electrophoresis of the PCR amplified of Tomato Yellow Leaf Curl Virus (TYLCV) genomic DNA. (a) PCR amplifications of full-length of TYLCV, (b) Nested PCR amplification and (c) PCR amplifications of CP using primers degenerate primers AVcore and ACcore
1, 2, 3 and 4: Tomato sample from Alhasa, Hail, Jizan and Alqasim respectively

stunting and malformations. To detect geminiviruses in symptomatic tomatoes, PCR analysis was performed using degenerate primers AVCore and ACCore. The expected product size (580 bp) was obtained from 330 out of 476 tomato samples as follows: 122 out of 138 (88%) of the collected samples from "Alhasa"; 69 out of 75 (92%) of collected samples from "Alqasim"; 87 out of 93 (94%) of tomato samples from "Hail"; 54 out of 69 (78%) of the collected tomato samples from "Jizan" and no bands were obtained with tomato samples collected from either "Aljowf" and "Tabuk". All samples that gave positive reaction with degenerate primers were retested with set of primer TYv2358/TYc450 as specific primer for TYLCV. The expected product size (870 bp) was obtained from all tested samples. This result indicated that all samples gave positive reaction with degenerate AVCore and ACCore were infected with TYLCV. Lower percentages of TYLCV infection were noted in all location compared to previous reports[14].

Many symptomatic samples gave negative result for *Begomovirus* detection by PCR; a similar result was previously noted[37,38]. The presence of TYLCD-like symptoms might be induced by another pathogen or by physiological stress as suggested by Tahiri *et al*.[37]. In addition, low concentration of the virus or uneven distribution of TYLCV in the plant may prevent its detection by PCR[38]. In addition, may those plants were infected with one or several other viruses which were not detected with the PCR test used. These factors may have contributed to the absence of detectable *Begomoviruses* from all 101 samples from "Aljowf and Tabuk". The virus may be absent because the insect vector, that is whitefly, is not present in that area; most *Begomoviruses* including TYLCV are transmitted only by the whitefly *B. tabaci*[89].

The molecular characterization of TYLCV isolates was conducted by using degenerate primers AVCore and ACCore to amplify a fragment of about 580 bp of the coat protein gene (CP) of TYLCV (Fig. 3c). Alignment with sequences from GenBank indicated that the TYLCV sequences detected in this study are most closely related to TYLCV from Iran, Yemen and Sudan.

PCR amplification of viral DNA: PCR analysis was performed using primers in Table 1 to amplify fragments and full length TYLCV from tomato samples of "Alqasim", "Alhasa", "Hail" and "Jizan". The expected sizes of TYLCV fragments as in Table 1 and full-length genome were amplified from the tested tomato tissues indicating the symptomatic tomato leaf tissue were infected by TYLCV (Fig. 3). The expected size (2.8 kb) of full-length genome of TYLCV was amplified using primers TYLCVF1 and TYLCVR2 from the infected plant tissue collected from "Alhasa", "Alqasim", "Hail" and "Jizan" (Fig. 3a). Nested PCR amplification of an 870 bp fragment with primers TYv2358 and TYc450 (Fig. 3b) supported the results obtained in Fig. 3a.

Multiple sequence alignment and homology analysis: The four amplified full-length amplicons from TYLCV isolates from "Alhasa", "Jizan", "Hail" and "Alqasim" were sequenced. The size of the obtained sequences of Saudi Arabia TYLCV isolates were 2775 nt for TYLCV-Qas and TYLCV-Hail, 2780 nt and 2771 nt for TYLCV-Jiz and TYLCV-Has, respectively. Full-length genome sequences of isolated TYLCV were analyzed and compared with other previous known sequences in DNA database obtained from GenBank as in Table 2. The previously reported length of TYLCV genome ranges from 2752-2794 nt[40]. The obtained sequences (Table 3) indicated that TYLCV-Hail and TYLCV-Qas share 97% identity and 95% identity to TYLCV-Jiz. Identity of TYLCV-Jiz to TYLCV-Has was only 95%. Alignment with other TYLCV strain sequences indicate that TYLCV-Qas, TYLCV-Hail and TYLCV-Jiz segregate in the same clade with TYLCV isolates from Sudan (JF919733, HE819244, JN591385, AY044139 and AY044137), Saudi Arabia

Table 3: Homology analysis of the genome sequences of isolated Tomato Yellow Leaf Curl Virus (TYLCV)

Isolate	Accession No.	Isolate location	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Has (%)	TYLCV-Qas (%)
TYLCV-Jiz	KC845301	Jizan	100.0			
TYLCV-Hail	KF040453	Hail	96.2	100.0		
TYLCV-Has	KF435137	Alhasa	82.8	82.5	100.0	
TYLCV-Qas	KF561125	Alqasim	94.8	97.0	82.8	100

*TYLCV: Tomato yellow leaf curl virus

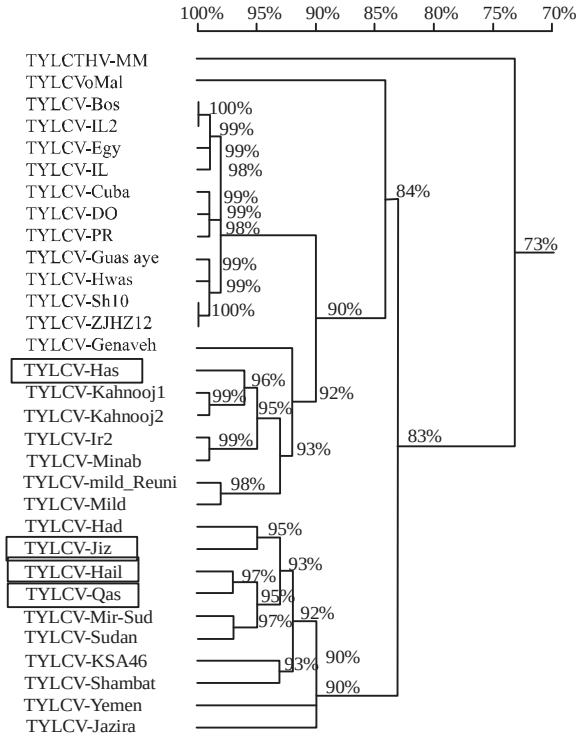

Fig. 4: Homology tree showing the identity percentage among the isolated tomato yellow leaf curl virus in this study and other strains and isolates that selected from GenBank

(HG530539) and Yemen (EF110891) with identity from 90-95%. Previous results indicated that the identity of partial coat protein of TYLCV from "Alqasim" and "Hail" in Saudi Arabia was more than 95% with both tomato leaf curl Sudan and Yemen viruses[17]. This variance may have occurred because prior studies compared partial sequence of the coat protein gene whereas this study examined the full sequence genome. Isolate TYLCV-Has segregates with distant clades of TYLCV isolates from Iran (GU076442, GU076448, GU076449, EU085423 and GU076454) with identity from 95-97% (Fig. 4). These results agreed with previous sequence analysis of 48 sequences of TYLCV from different various countries where homology ranged from 72.8-98.6% and the analyzed sequences contained typical structure of TYLCV[40].

Expression of sense-oriented genes is determined by the Intergenic Region (IR)[41]. The length of the IR (Table 4) may explain the variation in genome length of the Saudi TYLCV isolates. The identity of the IR is 99% between TYLCV-Jiz and TYLCV-Qas and both share 89% identity with TYLCV-Hail, but the IR of TYLCV-Has is only 81% identical to the other three Saudi isolates. This may reflect the geographical close proximity of Jizan to Sudan and Yemen and seedling transported to Riyadh, Alqasim and Hail. TYLCV detected in Alhasa is more related to those from Iran which is close to the Alhasa location.

The IR sequences of all four TYLCV have the potential to form a stem loop structure. In addition, they contain identical nucleotide sequences of 41 bases (AAAGCGGCCATCCGTA TAATATTACCGGATGGCCGCGCCCC) and including the conserved nona-nucleotide motif (TAATATTAC) present in all *Begomovirus* and involved in binding to replication associated protein gene (Rep). The arrangement of the conserved sequences that formed the stem-loop structure was common to all *Begomoviruses* and the inverted repeat sequences were important for the intermolecular recombination and circularization[26].

GC content: The GC content is 0.42 in TYLCV isolates of Alhasa (KF435137), Hail (KF040453) and Alqasim (KF561125) and 0.41 in isolate of Jizan (KC845301). The GC content of all the isolates and strains of TYLCV from different countries in this study ranged from 0.41-0.42 (Table 4). This result was in agreement with previous study[40] that found that the GC content of 48 sequences of TYLCV from 29 countries was 0.41 except four sequences were 0.42 from Jezira-Sudan, Ethiopia, Iran and Oman.

Encoded protein: In the virion sense, there are two ORFs: AV1 and AV2. ORF AV1 encodes between 243-259 aa in all studied isolates and strains of TYLCV, including those of this study (Table 4). The multiple alignment of amino acids of the ORFs and percentages of identity were illustrated through the homology tree (Fig. 5, Table 5). Protein V1 of TYLCV-Jiz, TYLCV-Qas and TYLCV-Hail shared the highest aa sequence identity of 96% or more. In contrast, the identity AV1 of TYLCV-Has was only 76% with coat proteins of TYLCV-Jiz, TYLCV-Hail and TYLCV-Qas.

The amino acid analysis by PLAST illustrated that encoded protein V1 of TYLCV-Jiz shared with the highest aa sequence

Table 4: Genomic characteristics of tomato yellow leaf curl virus isolates

					No. of amino acids in each ORFs						
					Virion strand		Complimentary strand				
Isolates	Accession No.	Genome length	GC content	IR length	V1	V2	C1	C2	C3	C4	C5
TYLCV-Jiz	KC845301	2780	0.41	309	256	116	358	134	133	95	-
TYLCV-Hail	KF040453	2775	0.42	211	243	115	389	134	113	97	128
TYLCV-Has	KF435137	2771	0.42	297	257	115	358	134	133	99	-
TYLCV-Qas	KF561125	2775	0.42	208	243	115	389	134	113	97	128
TYLCV-Sudan	HE819244	2772	0.42	244	243	115	388	134	113	95	128
TYLCV-Mir-Sud	JN591385	2773	0.42	246	243	115	377	134	113	95	128
TYLCV-Jazira	AY044137	2779	0.42	310	256	115	358	134	133	99	128
TYLCV-Shambat	AY044139	2768	0.41	215	256	115	386	134	133	99	128
TYLCV-Had	JF919733	2780	0.42	311	256	115	358	134	133	95	128
TYLCV-Yemen	EF110891	2781	0.42	312	256	115	358	134	133	99	128
TYLCV- Genaveh	GU076454	2764	0.42	292	257	115	358	134	133	99	-
TYLCV- Minab	GU076442	2764	0.42	292	257	115	353	134	133	96	-
TYLCV-Kahnooj1	GU076448	2770	0.42	214	257	115	386	134	133	96	-
TYLCV-Kahnooj2	GU076449	2770	0.42	214	257	115	386	134	133	96	-
TYLCV-Ir2	EU085423	2776	0.42	292	257	115	357	134	133	99	-
TYLCV-KSA46	HG530539	2788	0.42	262	256	115	377	134	133	99	128
TYLCV-Hwas	GU126513	2781	0.41	315	257	115	356	134	133	96	-
TYLCV-Bos	GU325634	2774	0.41	459	255	115	130	134	133	96	-
TYLCTHV-MM	AF206674	2746	0.42	283	255	111	360	133	133	98	-
TYLCV-DO	AF024715	2781	0.41	315	257	115	356	134	133	96	-
TYLCV-mild Reunion	AJ865337	2791	0.41	319	257	115	358	134	133	99	-
TYLCV-PR	AY134494	2781	0.41	315	257	115	356	134	133	96	-
TYLCV-Egy	AY594174	2781	0.41	315	257	115	356	134	133	96	-
TYLCV-IL	X15656	2787	0.41	466	259	115	130	134	133	96	-
TYLCV-Mild	X76319	2790	0.41	475	257	115	154	134	133	99	-
TYLCV-Mal	JN859135	2782	0.41	313	256	114	358	134	133	99	-
TYLCV-Sh10	EU031444	2781	0.41	315	257	115	356	134	133	96	-
TYLCV- ZJHZ12	FN256256	2781	0.41	315	257	115	356	134	133	96	-
TYLCV-IL2	AB116631	2774	0.41	459	257	115	130	134	133	96	-
TYLCV- Guasave	FJ012358	2781	0.41	466	257	115	130	134	133	96	-
TYLCV-Cuba	AJ223505	2781	0.41	315	257	115	356	134	133	96	-

TYLCV: Tomato yellow leaf curl virus

identity 98% with coat protein of tomato leaf curl Sudan virus isolates (CD144955, AEX92963 and AGN52712). The V1 of TYLCV-Has shared 100% identity to the coat protein of TYLCV isolated from pepper in Saudi Arabia (AGZ91414) and 95% identity with the coat protein of TYLCV from Iran (ADG37310, ADG37280, ADG37292 and CAA10748). The V1 of TYLCV-Qas and TYLCV-Hail shared the highest identity (99%) with coat proteins of tomato leaf curl Sudan viruses (TLCSV) (CCH75843 and AEX92961). AV2 is partially overlapping with AV1 ended with CCC that encoding to amino acid proline and terminated with TGA stop code in most of Saudi isolates. The second ORF (AV2), which encodes to protein (pre-coat protein), is important for infection of the plant and playing a role in TYLCV as a suppressor of RNA silencing[42]. Association of V2 with viral genomic DNA molecules is suggesting that V2 functions as a DNA shuttling protein[43]. The AV2 encodes a protein V2 of 115 aa in TYLCV-Qas, TYLCV-Hail and TYLCV-Has but 116 aa in TYLCV-Jiz (Table 4). The size ORF AV2 was consistent with all studied isolates and strains of TYLCV encoding 115 aa except TYLCV-Mal (JN859135) (114 aa) and TYLCV-Jiz (116 aa). Protein V2 of TYLCV-Jiz shared 93% identity with V2 of TYLCV-Qas and 97% with V2 protein of TYLCV-Hail. Identities of V2 protein of TYLCV-Has with V2 protein with other Saudi isolates in this study were low at 88%. PLAST through GenBank demonstrate that protein V2 of TYLCV-Jiz and TYLCV-Hail shared the highest aa sequence identity with the V2 protein of TLCSV from Sudan (ABQ42709, AGZ95034, AEX92960 and AGN52711), whereas the V2 protein of TYLCV-Has shared the highest identity of 97% with V2 protein of isolates from Iran (ADG37297) and Oman (ABG90923). The V2 of TYLCV-Qas also shared 96% identity to V2 protein of TLCSV.

There were other four and in some cases, five ORFs on the complementary strand of the genome. The first ORF is AC1 that encodes protein C1 that functions in replication of the virus[44]. AC1encodes between 355-389 aa in most reported cases (Table 4). The C1 protein of TYLCV-Jiz shared 92% with

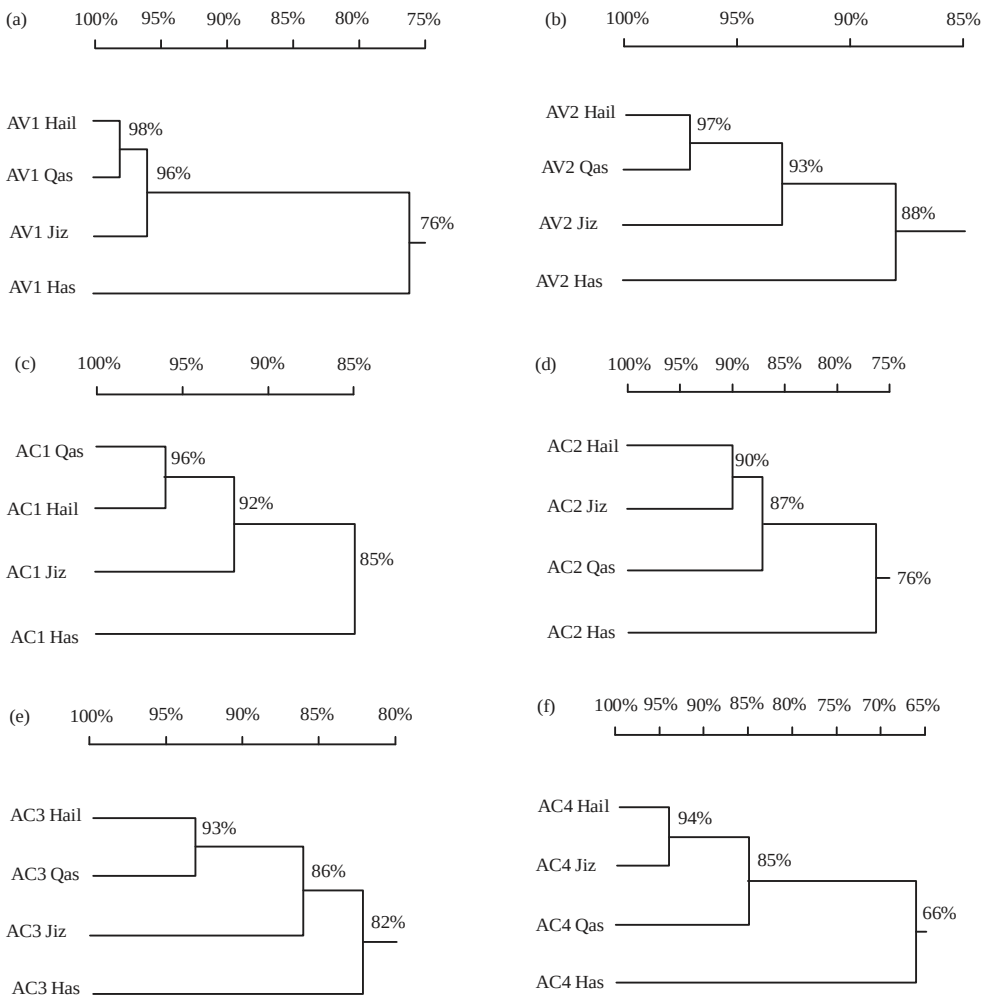

Fig. 5: Cluster multiple amino acids alignments and homology tree of the putative ORFs of isolated tomato yellow leaf curl virus from Hail, Alqasim, Jizan and Alhasa. (a-f) AV1, AV2, AC1, AC2, AC3 and AC4 proteins, respectively

C1 protein to both TYLCV-Hail and TYLCV- Qas. The result of PLAST through GenBank showed C1 protein of TYLCV-Jiz, TYLCV-Qas and TYLCV-Hail shared the highest identities (93-95%) with the C1 protein of TLCSV isolates from Sudan (AF156618, CCH75846 and AEX92966). As was the case for ORFs AV1 and AV2, TYLCV-Has is shared less identity to the other isolates in this study (Fig. 5). The highest identity of C1 protein of TYLCV-Has was 94% with C1 protein of TYLCV from Iran (ACF04176) and Jezira–Sudan (AAL05293). The second ORF on the complementary strand is AC2 that encodes protein C2; the C terminal of AC2 overlaps with ORF AC1. The AC2 encodes as mall protein (134 aa) in all TYLCV genomes included in this study (Table 4). The C2 protein of TYLCV-Jiz, TYLCV-Hail and TYLCV-Qas shared greater identity than with TYLCV-Has. PLAST through GenBank revealed that the C2 protein of TYLCV-Jiz and TYLCV-Hail shared identity greater than 98% with TLCSV (CCH75846, AFI56620 and AGZ95036) and the C2 protein of TYLCV-Has shared 99% with C2 protein of TYLCV from Iran (ACF04175). The third ORF on the complimentary strand (AC3) encodes a replication enhancer (REn) protein[45]. This protein consists of 113 aa in both isolates TYLCV-Hail and TYLCV-Qas but 133 aa in isolates TYLCV-Jiz and TYLCV-Has and all other studies isolates (Table 4). The C3 of TYLCV-Jiz shared highest identity 99% with TLCSV (AGZ95037) but the identities were 86% with C3 protein in both TYLCV-Hail and TYLCV-Qas and 82% with C3 protein of TYLCV-Has (Fig. 5). In addition, PLAST of GenBank sequences showed 94-96% identity of protein C3 from TYLCV-Hail and TYLCV-Qas with TLCSV (AEX92956) and 98% for C3 protein of TYLCV-Has with TYLCV from Iran (ADG37257).

The fourth ORF on the complementary strand is AC4 that encodes the C4 protein whose function was induction the

Table 5: Homology analysis of the amino acid sequences of the putative ORFs of tomato yellow leaf curl virus

	AC1				AC2				AC3				AC4			
ORFs	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)
AC1																
TYLCV-Jiz	100															
TYLCV-Hail	92	100														
TYLCV-Hasa	85	85	100													
TYLCV-Qas	92	96	85	100												
AC2																
TYLCV-Jiz					100											
TYLCV-Hail					90	100										
TYLCV-Hasa					76	76	100									
TYLCV-Qas					87	87	76	100								
AC3																
TYLCV-Jiz									100							
TYLCV-Hail									86	100						
TYLCV-Hasa									82	82	100					
TYLCV-Qas									86	93	82	100				
AC4																
TYLCV-Jiz													100			
TYLCV-Hail													94	100		
TYLCV-Hasa													66	66	100	
TYLCV-Qas													85	85	66	100

	AV1				AV2			
ORFs	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)	TYLCV-Jiz (%)	TYLCV-Hail (%)	TYLCV-Hasa (%)	TYLCV-Qas (%)
AV1								
TYLCV-Jiz	100							
TYLCV-Hail	96	100						
TYLCV-Hasa	76	76	100					
TYLCV-Qas	96	98	76	100				
AV2								
TYLCV-Jiz					100			
TYLCV-Hail					93	100		
TYLCV-Hasa					88	88	100	
TYLCV-Qas					93	97	88	100

*TYLCV: Tomato yellow leaf curl virus

disease symptoms[46]. As V2 protein and C2 protein, C4 protein was playing a role in suppression of gene silencing in plant[47]. The C4 proteins of TYLCV-Hail, TYLCV-Jiz and TYLCV-Qas shared greater than 85% identity but only 66% identity with TYLCV-Has (Fig. 5). Overall, of the five proteins encoded in the complementary sense, the C4 protein of all four Saudi isolates exhibited lower identities to other sequences available in GenBank.

The fifth ORF in the complementary sense was AC5; this ORF was found in only two isolates examined in this study: TYLCV-Hail and TYLCV-Qas (Table 4). This ORF was also found in other TYLCV strains. In addition, ORFAC5 was found in some Gemini viruses such as Watermelon chlorotic stunt virus (WmCSV) and in TLCSV isolated from beans in Saudi Arabia (unpublished data). AC5 protein of TYLCV-Hail shared identity 98% with C5 protein of TYLCV-Qas.

On the bases of the obtained data in this study, TYLCV-Jiz, TYLCV-Qas and TYLCV-Hail were more related to Tomato leaf curl Sudan virus (TLCSV) and TYLCV-Has was more related to Tomato yellow leaf curl virus of Iran. Both viruses infect tomato and induce tomato yellow leaf curl-like disease. The close similarity of isolates TYLCV (Jiz, Hail and Qas) may be a consequence of the movement by tomato plants from "Jizan" to "Hail" and "Alqasim". The close proximity of "Jizan" to Yemen and Sudan may account for its similarity to TYLCV-Sudan; this virus was recorded in both Sudan and Yemen[22,23]. TYLCV-Has was more related to those strains of TYLCV in Iran, perhaps because the "Alhasa" oasis is situated close to Iran and there is agriculture exchange between the farmers in Alhasa and Iran.

Samples collected from the other location "Tabuk" and "Aljawf" (the north area of Saudi Arabia) were virus negative. This may be a consequence of the season during which samples were collected. Other strains related to viruses of the Mediterranean region (TYLCV-IL, TYLCV-Mild and TYLCV-Sardinia) are recorded there[16]. Future surveys in the North of Saudi Arabia will be necessary to evaluate the virus populations there. As known, the genome sequence of genus *Begomovirus* is very unstable and well known to be a group with high diversity. This is the first study of the molecular diversity for TYLCV in the Saudi Arabia. This study discovers the high diversity of TYLCV genome in this area that can beneficial for the viral control. This study will help the researcher to uncover the distribution of this virus in other area that many researchers were not able to explore.

CONCLUSION

Full-length genome sequences of different four isolates of TYLCV collected from different locations in Saudi Arabia were analyzed to study the genetic diversity of the isolated virus. Based on the obtained data in this study two different strains of TYLCV were found in Saudi Arabia. First that isolated from Jizan, Alqasim and Hail and this strain was more related to the Sudan strain. Second, that isolated from Alhasa and was more related to strain of Iran. Additional more, study is necessary to evaluate the population and distribution of virus in other locations especially in the North of Saudi Arabia. There are some other strains expected to be their especially those are recorded in Jordon such as Israeli and Sardinian isolates.

SIGNIFICANCE STATEMENTS

Begomovirus is very unstable in the genome sequences and well known to be a group with high diversity. This study has a level of importance because regional distribution of TYLCV in Saudi Arabia was not surveyed so far.

ACKNOWLEDGEMENTS

The author would like to thank the National Plan for Science, Technology and Innovation (MAARIFAH), King Abdullaziz City for Science and Technology, Saudi Arabia, (Grant No. 10-Bio1096-06) for funding this study. We would like to thank Mr. Wael Alarby and Mr. Mostufa Almgasla at Plant Protection Department and Ministry of Agriculture, for their assistance during the sampling of the research. Also we thank Dr. Ken Eastwell Washington State University for supporting this study. The authors are grateful to Pest and Plant Diseases Unit at King Faisal University for conducting this study.

REFERENCES

1. Rojas, M.R. and R.L. Gilbertson, 2008. Emerging Plant Viruses: A Diversity of Mechanisms and Opportunities. In: Plant Virus Evolution, Roossinck, M.J. (Ed.). Springer, Berlin, Heidelberg, ISBN: 978-3-540-75762-7, pp: 27-51.
2. Briddon, R.W., B.L. Patil, B. Bagewadi, M.S. Nawaz-ul-Rehman and C.M. Fauquet, 2010. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite *Begomoviruses*. BMC Evol. Biol., Vol. 10. 10.1186/1471-2148-10-97.
3. Khan, J.A., S. Akhtar, R. Briddon, U. Ammara, A. Al-Matrooshi and S. Mansoor, 2012. Complete nucleotide sequence of watermelon chlorotic stunt virus originating from Oman. Viruses, 4: 1169-1181.
4. Fauquet, C.M., R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini and X. Zhou, 2008. *Geminivirus* strain demarcation and nomenclature. Arch. Virol., 153: 783-821.

5. Abhary, M.K., B.L. Patil and C.M. Fauqet, 2007. Molecular Biodiversity, Taxonomy and Nomenclature of Tomato Yellow Leaf Curl-Like Viruses. In: Tomato Yellow Leaf Curl Disease, Management, Molecular Biology, Breeding for Resistance, Czosnek, H. (Ed.). Springer, Dordrecht, The Netherlands, pp: 85-118.
6. Diaz-Pendon, J.A., M.C. Nizares, E. Moriones, E.R. Bejarano, H. Czosnek and J. Navas-Castillo, 2010. Tomato yellow leaf curl viruses: Menage a trois between the virus complex, the plant and the whitefly vector. Mol. Plant Pathol., 11: 441-450.
7. Rey, M., J. Ndunguru, L. Berrie, M. Paximadis and S. Berry *et al*., 2012. Diversity of dicotyledonous-infecting geminiviruses and their associated DNA molecules in Southern Africa, including the South-West Indian Ocean islands. Viruses, 4: 1753-1791.
8. Cohen, S. and I. Harpaz, 1964. Periodic, rather than continual acquisition of a new tomato virus by its vector, the tobacco whitefly (*Bemisia tabaci* Gennadius). Entomol. Exp. Appl., 7: 155-166.
9. Brown, J.K., F.M. Zerbini, J. Navas-Castillo, E. Moriones and R. Ramos-Sobrinho *et al*., 2015. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Arch. Virol., 160: 1593-1619.
10. Moriones, E. and J. Navas-Castillo, 2008. Rapid evolution of the population of begomoviruses associated with the tomato yellow leaf curl disease after invasion of a new ecological niche: A review. Spanish J. Agric. Res., 6: 147-159.
11. Lefeuvre, P., D. Martin, G. Harkins, P. Lemey and A.J.A. Gray *et al*., 2010. The spread of tomato yellow leaf curl virus from the Middle East to the world. PLoS Pathogens, Vol. 6. 10.1371/journal.ppat.1001164.
12. Perefarres, F., M. Thierry, N. Becker, P. Lefeuvre, B. Reynaud, H. Delatte and J.M. Lett, 2012. Biological invasions of geminiviruses: Case study of TYLCV and *Bemisia tabaci* in Reunion Island. Viruses, 4: 3665-3688.
13. Khan, A.J., A.M. Idris, N.A. Al-Saady, M.S. Al-Mahruki, A.M. Al-Subhi and J.K. Brown, 2008. A divergent isolate of tomato yellow leaf curl virus from Oman with an associated DNA β satellite: An evolutionary link between Asian and the Middle Eastern virus-satellite complexes. Virus Genes, 36: 169-176.
14. Ajlan, A.M., G.A.M. Ghanem and K.S. Abdulsalam, 2007. Tomato yellow leaf curl virus (TYLCV) in Saudi Arabia: Identification, partial characterization and virus-vector relationship. Arab J. Biotechnol., 10: 179-192.
15. Rezk, A.A., 2016. Molecular characterization of tomato yellow leaf curl virus (TYLCV) infecting pepper and common bean. Int. J. Virol., 12: 1-9.
16. Anfoka, G., M. Abhary, F.H. Ahmad, A.F. Hussein and A. Rezk *et al*., 2008. Survey of tomato yellow leaf curl disease-associated viruses in the Eastern Mediterranean Basin. J. Plant Pathol., 90: 313-322.
17. Alhudiab, K., W. Alaraby and A. Rezk, 2014. Molecular characterization of tomato yellow leaf curl disease associated viruses in Saudi Arabia. Int. J. Virol., 10: 192-203.
18. Brown, J.K., A.M. Idris, I. Torres-Jerez, G.K. Banks and S.D. Wyatt, 2001. The core region of the coat protein gene is highly useful for establishing the provisional identification and classification of *Begomoviruses*. Arch. Virol., 146: 1581-1598.
19. Saitou, N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406-425.
20. Jukes, T.H. and C.R. Cantor, 1969. Evolution of Protein Molecules. In: Mammalian Protein Metabolism, Munro, H.N. (Ed.). Academic Press, New York, pp: 21-132.
21. Khan, A.J., S. Akhtar, A.K. Singh and R.W. Briddon, 2013. A distinct strain of Tomato leaf curl Sudan virus causes tomato leaf curl disease in Oman. Plant Dis., 97: 1396-1402.
22. Idris, A.M. and J.K. Brown, 2005. Evidence for interspecific-recombination for three monopartite begomoviral genomes associated with the tomato leaf curl disease from Central Sudan. Arch. Virol., 150: 1003-1012.
23. Idris, A.M., N.M. Abdullah and J.K. Brown, 2012. Leaf curl diseases of two solanaceous species in Southwest Arabia are caused by a monopartite begomovirus evolutionarily most closely related to a species from the Nile Basin and unique suite of beta satellites. Virus Res., 169: 296-300.
24. Azizi, A., J. Mozafari and M. Shams-Bakhsh, 2008. A new strain of tomato yellow leaf curl virus from Iran. GenBank: EU085423.2, National Center for Biotechnology Information, Bethesda, MD.
25. Idris, A., M. Al-Saleh, M.J. Piatek, I. Al-Shahwan, S. Ali and J.K. Brown, 2014. Viral metagenomics: Analysis of begomoviruses by illumina high-throughput sequencing. Viruses, 6: 1219-1236.
26. Kim, S.H., S. Oh, T.K. Oh, J.S. Park and S.C. Kim *et al*., 2011. Genetic diversity of tomato-infecting Tomato Yellow Leaf Curl Virus (TYLCV) isolates in Korea. Virus Genes, 42: 117-127.
27. Nakhla, M.K., D.P. Maxwell, R.T. Martinez, M.G. Carvalho and R.L. Gilbertson, 1994. Widespread occurrence of the eastern Mediterranean strain of tomato yellow leaf curl geminivirus in tomatoes in the Dominican Republic. Plant Dis., 78: 926-926.
28. Delatte, H., D.P. Martin, F. Naze, R. Goldbach, B. Reynaud, M. Peterschmitt and J.M. Lett, 2005. South West Indian ocean islands tomato begomovirus populations represent a new major monopartite begomovirus group. J. Gen. Virol., 86: 1533-1542.
29. Bird, J., A.M. Idris, D. Rogan and J.K. Brown, 2001. Introduction of the exotic *Tomato yellow leaf curl virus*-Israel in tomato to Puerto Rico. Plant Dis., 85: 1028-1028.

30. Abhary, M.K., G.H. Anfoka, M.K. Nakhla and D.P. Maxwell, 2006. Post-transcriptional gene silencing in controlling viruses of the tomato yellow leaf curl virus complex. Arch. Virol., 151: 2349-2363.
31. Antignus, Y. and S. Cohen, 1994. Complete nucleotide sequence of an infectious clone of a mild isolate of Tomato Yellow Leaf Curl Virus (TYLCV). Phytopathology, 84: 707-712.
32. Monci, F., S. Sanchez-Campos, J. Navas-Castillo and E. Moriones, 2002. A natural recombinant between the geminiviruses tomato yellow leaf curl Sardinia virus and tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology, 303: 317-326.
33. Liao, B.L., Y. Liu, Y. Xie and X.P. Zhou, 2007. Genomic characterization of DNA-A and associated satellite DNA molecule of an isolate of Tomato yellow leaf curl China virus in *Solanum aculeatissimum*. Acta Phytopathol. Sin., 37: 138-143.
34. Zhang, Y., W. Zhu, H. Cui, Y. Qiu and K. Sha *et al*., 2008. Molecular identification and the complete nucleotide sequence of TYLCV isolate from Shanghai of China. Virus Genes, 36: 547-551.
35. Ueda, S., T. Kimura, M. Onuki, K. Hanada and T. Iwanami, 2004. Three distinct groups of isolates of *Tomato yellow leaf curl virus* in Japan and construction of an infectious clone. J. Gen. Plant Pathol., 70: 232-238.
36. Bejarano, E. R., 1998. Tomato yellow leaf curl virus-[Cuba] complete genome. GenBank: AJ223505.1, National Center for Biotechnology Information. http://www.ncbi.nlm.nih.gov/nuccore/AJ223505.
37. Tahiri, A., A. Sekkat, A. Bennani, M. Granier, G. Delvare and M. Peterschmitt, 2006. Distribution of tomato infecting begomoviruses and *Bemisia tabaci* biotypes in Morocco. Ann. Applied Biol., 149: 175-186.
38. Rotbi, M., A.P. de Castro, M.J. Diez and N. Elmtili, 2014. Identification and distribution of *Tomato yellow leaf curl virus* TYLCV and *Tomato yellow leaf curl Sardinia virus* TYLCSV infecting vegetable crops in Morocco. Afr. J. Biotechnol., 13: 1476-1483.
39. Weng, S.H. and C.W. Tsai, 2013. Insect transmission of tomato yellow leaf curl viruses. Proceedings of the International Symposium of Insect Vectors and Insect-Borne Disease, August 6-7, 2013, Tai Chung, Taiwan, pp: 255-274.
40. Wan, H.J., W. Yuan, R.Q. Wang, Q.J. Ye, M.Y. Ruan and Z.M. Li, 2015. Assessment of the genetic diversity of tomato yellow leaf curl virus. Genet. Mol. Res., 14: 529-537.
41. Gover, O., Y. Peretz, R. Mozes-Kosh, E. Maori, H. Rabinowitch and I. Sela, 2014. Only minimal regions of tomato yellow leaf curl virus (TYLCV) are required for replication, expression and movement. Arch. Virol., 159: 2263-2274.
42. Zrachya, A., E. Glick, Y. Levy, T. Arazi, V. Citovsky and Y. Gafni, 2007. Suppressor of RNA silencing encoded by *Tomato yellow leaf curl virus*-Israel. Virology, 358: 159-165.
43. Moshe, A., E. Belausov, A. Niehl, M. Heinlein, H. Czosnek and R. Gorovits, 2015. The *Tomato yellow leaf curl virus* V2 protein forms aggregates depending on the cytoskeleton integrity and binds viral genomic DNA. Scient. Rep., Vol. 5. 10.1038/srep09967.
44. Laufs, J., W. Traut, F. Heyraud, V. Matzeit, S.G. Rogers, J. Schell and B. Gronenborn, 1995. *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. Proc. Natl. Acad. Sci. USA., 92: 3879-3883.
45. Settlage, S.B., R.G. See and L. Hanley-Bowdoin, 2005. *Geminivirus* C3 protein: Replication enhancement and protein interactions. J. Virol., 79: 9885-9895.
46. Jupin, I., F. De Kouchkovsky, F. Jouanneau and B. Gronenborn, 1994. Movement of tomato yellow leaf curl geminivirus (TYLCV): Involvement of the protein encoded by ORF C4. Virology, 204: 82-90.
47. Luna, A.P., G. Morilla, O. Voinnet and E.R. Bejarano, 2012. Functional analysis of gene-silencing suppressors from tomato yellow leaf curl disease viruses. Mol. Plant Microb. Interact., 25: 1294-1306.