MOLECULAR CHARACTERIZATION OF *SQUASH LEAF CURL VIRUS* EGYPTIAN ISOLATE AND GENERATION OF INFECTIOUS CLONE VIA ROLLING CIRCLE AMPLIFICATION

[12]

Inas. F. Fahmy

ABSTRACT

***Squash** leaf curl virus* (SqLCV) is one of the most important viruses affecting squash and cucurbit plants in Egypt. In order to produce resistant line against SqLCV, it has to be characterized on the molecular and biological levels. This study was done to produce an infectious clone of the respected virus. The complete molecular characterization, phylogenetic analysis, and putative recombination events in *Squash leaf curl virus* are reported. A typical leaf curl symptoms were observed upon inoculation, using the resulted clone, which proves that the inserted sequences are representing all functional domains that are needed to cause the infection.

The complete nucleotide sequence of SqLCV was determined as DNA- A (2653bp) and DNA- B (2619bp). The SqLCV isolate from Egypt has about 97.8% homology with the SqLCV isolates from *Squash leaf curl virus* isolate LB2 segment DNA-A isolate ENA|HM368373 registered in the genebank. While, the least level of homology with *Chenopodium leaf curl virus* isolate Gezira ENA |HM62651 with 74.4 %. For component B the highest homology was recorded with *Squash leaf curl virus* LB2 isolate DNA-B ENA|HM368374 with 97.2% and the least homology was Tomato severe rugose virus isolate Pi-1 DNA- B, ENA|HQ606468 with 68% homology. Bioinformatic analysis revealed revealed that the genome of SqLCV is a circular DNA bipartite genome DNA-A and DNA-B, which form six open reading frames. In DNA-A two proteins (V1 and V2) are located on virion-sense strand whereas four other proteins (C1, C2, C3,

C4) are located in the complementary sense. It has been proved that the DNA-B is more prone to increasing recombination events which could be interpreted as a result of pseudo-recombination. This recombination could be affected by the number of genes and the interaction of the virus component solely with the host plant or with the host plant in addition to the insect vector which provides a restriction to rapid evolution of a component compared to the other one.

Key words: *Begomoviruses*, templification, RCA, viral infectivity.

Agricultural Genetic Engineering Research Institute (AGERI), ARC, GIZA, Egypt

*** Corresponding author:**
Inas. F. Fahmy

INTRODUCTION

Whitefly-transmitted *Geminiviruses* have emerged as serious pathogens of agronomic and horticultural crops in Egypt and worldwide. *Geminiviruses* belong to an economically important family of plant viruses that are characterized by having small circular single-stranded DNA (ssDNA) genomes, encapsidated in twinned icosahedral particles **(Lazarowitz, 1992).** The *Geminiviridae* family consists of four genera that differ with respect to insect vector, host range and genome structure **(Fauquet *et al.*, 2000).** The genus *Begomovirus* represents the whitey-transmitted, dicot-infecting *Geminiviruses* that most commonly have bipartite genomes consisting of two ssDNA components, designated DNA A and DNA B. Some *Begomoviruses*, however, have a single genomic component **(Noris *et al.*, 1994),** which contains all functions necessary for replication and movement in the plant. *Squash leaf curl virus* reported as a severe viral disease of squash mainly (*Cucurbita pepo* L.) in these regions. It was isolated for the first time in Egypt from squash (*Cucurbita pepo* L. cv. Eskansarani) plants growing in Qaluobiya Governorate **(Farag *et al.*, 2005).** SqLCV virus is transmitted efficiently by whitefly,

Bemesia tabaci (Genn) **(Cohen *et al.*, 1983 and Mc Creight, 1984).**

The amplification of the viral genome was obtained by the use of DNA polymerase of the *Bacillus subtilis* bacteriophage Φ 29 **(Brukner *et al.*, 2006)** upon DNA isolation, and because *Begomviruses* replicate by the Rolling Circle Replication mechanism it is the ideal organism to be amplified employing the same technique which is considered to be the most effective method facilitates the screening and cloning of large number of *Begomoviruses*.

In order to study the biology of a virus at the molecular level, it is necessary to produce an infectious clone of the viral genome subject it to sequence to determine its coding potential. This will allow mutagenesis of specific regions of the genome and the subsequent introduction of the mutated or wild type genomes into host plants for further studying host–virus or insect-virus interactions. Also it can be used to study the effect of serial silencing experiments targeting essential virus genes to interfere with virus life cycle.

MATERIALS &METHODS

SqLCV isolates:

Squash plants (*Cucurbita pepo* L.) showing leaf curling, yellow mottling, stunting and reduced fruit set in different regions from crop fields in Giza, Elmansoria, Elboustan, Cairo were collected and stored in -80°C until analyzed for viral infection via PCR.

DNA extraction, templiPhi and restriction digestion:

Total DNA was extracted from infected plants using modified CTAB method; DNA was precipitated with isopropanol and used for preliminary PCR detection and further viral amplification using templiPhi technique amplification with done by the use of Φ 29 DNA polymerase the method used is reported by **(Fahmy *et al.*, 2011)**. A primer set was used to amplify 570 bp fragment to confirm virus infectivity with *Squash leaf curl virus*, the oligonucleotide degenerate primers were selected according to Brown *et al.* (2001) namely, V324 (+) primer corresponding to 5' GCCYAT RTA YAG RAA GCC MAG 3' and C889 (-) primer corresponding to 5' GGR TTD GAR GCA TGH GTA CAT G 3', presented in fig (2). A

strain collected from Elboustan area was selected for further amplification and cloning procedures, about 100 ng of total DNA isolated from infected plant and the templification was done as described by (**Fahmy *et al.*, 2011 and Packialakshmi *et al.*, 2010**). 2µl of the templiPhi product was run on agarose gel 1.2% to confirm amplification. 2µl of templiphied product were digested using different common restriction digestion enzymes (Fermentas, Fast digest), the enzymes used in the experiment were, EcoRI, HindIII, BamHI, SalI, SacI, Bgl II,PstI, NotI, XbaI, KpnI. Digestion was carried out for 20 min at 37°C with 2 units of enzyme. Digestion product was analyzed on agarose gel 1.2% to confirm virus size.

Cloning and sequencing

The restriction maps of both components, component -A and -B, were tested for having single cutters that can cut once in only one of the components. These enzymes were selected to linearize each component separately in order to facilitate the cloning of these components onto the pBlue script II (KS -) plasmid. The different cloned putative virus clones were sequenced to confirm their identity by analysis on the EMBL database then submitted to the genebank.

Table 1: primers used for sequencing of DNA A and B of SqLCV-eg.

	Primer name	Nucleotide sequence
DNA A	sqT3a-2	5'ACGCGTCACGTCCGAAATAC3'
	sqT7a-2	5'AATTAACGATAAACGTCTTT3'
DNA B	sqT3b-2	5'TAATAAATTGCAGAAGTCCCGCA3'
	sqT7b-2	5'TGCCCCTTGGAGTCCCCTAAAGG3'

Table 2: Complete genome sequence of Squash leaf Curl Virus- Eg DNA A and B isolate used for comparison and phylogenetic analysis

ponent A	Euphorbia mosaic virus clone EF2E ;ENA|FJ807782	Pepper golden mosaic virus isolate PepGMV-Ser ;AY928516	Component B	Tomato mild mosaic virus isolate BR:Pda58:05 ; ENA|EU710753	Sida golden mosaic vir ENA|AJ250751
leaf curl virus isolate LB2 ENA|HM368373 on)	Cucurbit leaf curl virus ;ENA|AF224760	Sequence 265 from Patent WO2010123904;HI647378	Squash leaf curl virus isolate LB2 :ENA|HM368374	Macroptilium yellow mosaic virus isolate Hope Pastures ;ENA|EF585289	Tomato rugose mosaic viru ; ENA|AF291706
leaf curl virus isolate IL :ENA |HQ184436	Cucurbit leaf crumple virus - [Arizona] ;ENA|AF256200	Sequence 264 from Patent WO2010123904. ;HI647377	Squash leaf curl virus isolate IL: ENA| HQ184437	Sida golden mosaic Honduras virus ; ENA|Y11098	Sida micrantha mosaic v [Brazil:Mato grosso do sul: - [Bolivia] ; ENA|HM58
leaf curl virus isolate Cairo DQ285019	Tomato mild yellow leaf curl Aragua virus ; ENA|AY927277	Pepper golden mosaic virus strain Tamaulipas ;U57457	Squash leaf curl virus ; ENA|EF532621	Euphorbia yellow mosaic virus ; ENA|FJ619508	Macroptilium mosaic Puer virus-[Bean] ; ENA|AF4-
leaf curl virus :ENA|EF532620 i)	Pepper golden mosaic virus isolate Guanajuato1 ; ENA|GU128148	Pepper golden mosaic virus-[CR ;AF149227	Squash leaf curl virus isolate 10 s ; ENA|DQ285017	Okra yellow mosaic Mexico virus isolate Mezatepec 3 ; ENA|GU972604	Macroptilium yellow mosa isolate St. Thomas ENA|EF585291
leaf curl virus :ENA|M38183	Pepper golden mosaic virus isolate SLP1 ; ENA|GU564594	Sequence 266 from Patent WO2010123904. ;HI647379	Squash leaf curl virus isolate WAZ ;DQ285018	Macroptilium mosaic Puerto Rico virus ; ENA|AY044134	Tomato golden mosaic v Common ; ENA|M737
leaf curl virus isolate 10 ;ENA|DQ285016	Squash leaf curl Israel virus coat protein ; ENA|AY206998	Cabbage leaf curl virus ;U65529	Squash leaf curl virus ; ENA|M38182	Okra yellow mosaic Mexico virus isolate OkYMMV-TX ; ENA|HM035060	Abutilon mosaic Bolivia ENA|HM585446
leaf curl virus strain SLCV-WAZ :ENA 203	Pepper golden mosaic virus isolate Guanajuato2 ;GU128149	Euphorbia yellow mosaic virus - Goias [Brazil:GO:Luziania 8675:2009] ;JF756676	Squash mild leaf curl virus-[Imperial Valley] SLCV-R ; ENA|AF421553	Sinaloa tomato leaf curl virus ; ENA|AJ508783	Sida yellow mosaic Yucata ENA|DQ875873
leaf curl virus isolate SYR-C2: ENA 922	Rhynchosia golden mosaic Yucatan virus ;EU021216	Euphorbia yellow mosaic virus - Goias [Brazil:DF:Taquara:ABA3:2002] ;JF756670	Squash leaf curl virus isolate Cairo ;DQ285020	Macroptilium yellow mosaic virus isolate Mona ; ENA|EF582841	Chino del tomate virus-[ENA|AF101478
leaf curl virus isolate JOR-C1: ENA 3921	Tobacco yellow crinkle virus-[Capsicum:Cuba:2007] ; ENA|FJ222587	Euphorbia mosaic virus - A [Mexico:Yucatan:2004] clone pEuMV-YP-A ;DQ318937	Squash yellow mild mottle virus clone p98631-;AF440790	Rhynchosia golden mosaic Havana virus-[Cuba:Havana:28:2007] ; ENA|HM236369	Sida golden mosaic vir ENA|AF039841
chlorotic leaf curl virus-[Guatemala] AF325497	Rhynchosia golden mosaic Yucatan virus clone Desm-C6 ; ENA|GQ352453	Bean calico mosaic virus ;ENA|AF110190	Melon chlorotic leaf curl virus-[Guatemala] ; ENA|AF325498	Euphorbia mosaic virus Brazil Mato Grosso do sul ; ENA|FN435996	Euphorbia mosaic Venezue [Tomato:Aragua:2003 ENA|EF547938
yellow mild mottle virus-[CR] isolate 98-NA|AY064391	Pepper golden mosaic virus strain pepper mild tigre ; ENA|EF210556		Squash mild leaf curl virus isolate 4 ; ENA|DQ285015	Tomato severe rugose virus isolate Pi-1 ; ENA|HQ606468	Tomato mild leaf curl vi [Uberlandia] ; ENA|DQ3
mild leaf curl virus isolate ;DQ285014	Capsicum begomovirus Cuba/2007 clone 23E1 ;FJ174699		Cucurbit leaf curl virus ; ENA|AF224761	Euphorbia mosaic virus ; ENA|FN435998	
mild leaf curl virus-[Imperial Valley] -R ;ENA|AF421552	Rhynchosia mosaic Sinaloa virus ; ENA|DQ406672		Cucurbit leaf curl virus-[Arizona] ; ENA|AF327559	Cleome leaf crumple virus ; ENA|FN436000	
leaf curl virus isolate Giza coat protein DQ364057	Euphorbia mosaic virus Peru ; ENA|AM886131		Sida mosaic Bolivia virus 1 ; ENA|HM585442	Euphorbia yellow mosaic virus - Goias [Brazil:GO:Itaberai 5082B:2007] ; ENA|JF756677	
rbia mosaic virus - B o:Jalasco:Pepper:2005] :ENA|DQ520942	Pepper golden mosaic virus isolate PepGMV-Mo ; ENA|AY928512		Tomato yellow vein streak virus isolate Ba-3 ;EF417916	Tomato golden vein virus isolate DF[BR:Pip1799:03] ; ENA|JF803265	
rbia mosaic virus - [Jamaica:Wissadula] DQ395342	Tobacco yellow crinkle virus ;ENA|FJ213931		Wissadula golden mosaic St Thomas Virus ; ENA|EU158095	Euphorbia yellow mosaic virus - Goias [Brazil:GO:Acreuna AB3540:2005] ; ENA|JF756678	
rbia mosaic virus - A [Puerto urabo:1991] :ENA|AF068642	Euphorbia mosaic virus [CU: Havana 27: 2007] ;HQ896201		Bean calico mosaic virus ; ENA|AF110190	Tomato mottle Taino virus ;AF012301	
rbia mosaic virus - ca:Wissadula:Euphorbia:2004] FJ407052	Chenopodium leaf curl virus [VEM], complete genome:HM626516		Sinaloa tomato leaf curl virus ; ENA|AJ508782	Wissadula golden mosaic St Thomas Virus clone W132BFL5 ; ENA|GQ355487	

Sequence analysis and construction of phylogenetic tree:

The genomic sequences of DNA-A, DNA-B molecules were assembled from contiguous sequences using the Vector NTI program (Quiagen), the ORFs from the full length genomes A and B were predicted using the ORF finder program available at http://www. Ncbi.nlm.nih.gov/projects/gorf/ and confirmed by similarity search using BLAST and FASTA from EBI webpage http:// www. ebi.ac.uk/. The genome maps of DNA-A and B were constructed using Vector NTI program (Quiagen) ORFs were identified by using the Vector NTI Suite 11 program. To determine whether the molecules were similar to known viral/satellite sequences, they were initially analyzed by using the BLAST program available at the National Center for Biotechnology Information (http:// www.ncbi.nlm.nih.gov/BLAST/) and FASTA program available at the website (http://www.ebi.ac.uk/sss/fasta/). Sequences were aligned using the MAFFT program **(Katoh *et al.*, 2009) and (Katoh and Toh, 2010)** available at the website of EBI. Phylogeneitic trees were constructed using the neighbor-joining method using the Recombination Detection Program version 4.0 (RDP4).

Possible gene-recombination events were analyzed by using the Recombination Detection Program version 4.0(RDP4) **(Martin *et al.*, 2005)**, available online (http://darwin.uvigo.es/rdp/rdp.html). The recombination analyses were applied using six automated methods: RDP **(Martin *et al.*, 2005)**, GENECONV **(Padidam *et al.*, 1999)**, BootScan **(Salminen *et al.*, 1995)**, SiScan **(Gibbs *et al.*, 2000; Salminen *et al.*, 1995)**, MaxChi and CHIMAERA **(Posada and Crandall, 2001)**, default parameters were used in our application except that the option 'Reference sequence selection' was set at 'internal references only'.

Construction of an infectious clone of SqLCV A&B:

In this study an infectious head-to-tail 2.7-mer of SqLCV A&B clone was prepared by the digestion of approximately 10 µg recombinant plasmid DNA using the restriction enzyme (*Bam*HI, for component A) and (*Eco*RI, for component B) selected for the insert preparation in an additional enzyme that does not cut the viral genome but has a recognition site(s) inside the vector multiple cloning site. The linearized genome fragment was recovered from agarose as described using Qiax gel clean (Qiagen). 5µg of the Binary vector (pCambia) were digested with the enzyme used to prepare the viral genome fragment. The Binary vector was dephosphorylated and recovered after linearization from 0.7% agarose as described by the manufacturer instructions (Qiax gel clean (Qiagen). The insert and vector were Ligated at a molar ratio of at least 10:1, respectively using 5 U ligase and 100 ng of p Cambia binary vector. Ligated mix for both virus clones were then

transformed using high-efficiency DH5α competent cells. Clones containing a multimeric insert by restriction enzyme digestion using enzymes flanking the insert were selected, or by colony PCR confirmation. A recombinant clone containing a multimeric copy of the viral genome were selected and used for DNA purification and transformation of *Agrobacterium* **(Stenger *et al.*, 1991)**.The clones were stored in glycerol at −80°C. A glycerol stock of the *Agrobacterium* strain LBA 4404 carrying dimers of the corresponding insert, DNA-A and DNA-B genomic components of SqLCV were prepared according to **(Kheyr-Pour *et al.*, 1991)**. Bacterial cultures harbouring plasmids with DNA-A and respective DNA-B genomic components were mixed to equal amounts and left for 2-3 h at room temperature to sediment debris prior to agroinoculation.

Whole plant infectivity assays:

Cultures of *A. tumefaciens* harboring the pCambia- 2.7 SqLCV A&B construct were grown for 36-48 h and inoculated into the axillary buds of squash (*Cucurbita pepo* cv.) and tobacco (*Nicotiana tabacum*) plants as described previously **(Dry *et al*, 1993)**. The agroinoculated plants were evaluated for symptom appearance 21-45 days post-inoculation (dpi), (fig.1). Meanwhile, developing leaves were sampled from these plants and DNA was extracted and analyzed for the presence of the viral genome by PCR using the degenerate primers tested for the detection of SqLCV namely V324 (+) and C889, the primer pair was selected according to **Brown *et al.*, (2001)** (Table 1).

RESULTS

In this study a field isolate revealed severe symptoms have been sequenced, these symptoms compared with other previously identified *Squash leaf curl virus* collected from Cairo which shows mild symptoms on squash plants. The isolate was tested for the possibility of the association of ß satellites which are mainly responsible for the development of symptom severity of infected squash plants. It has been revealed that no association of ß satellites with SqLCV component A and B using universal ß satellites primers or using restriction digestion analysis (fig.2), which in that case would reveal 1.3-1.4 fragment kb for ß satellite DNA presence in our field isolate, a positive control was used to confirm ß absence.

The full length DNA A 2653 and the full length DNA B 2619bp of SqLCV collected from Elboustan region has a typical genome organization of the new world *Begomoviruses* with five ORFs one in the virion sense and four in the complementary sense for the DNA-A while for DNA-B the ORF finder of the program Vector NTI (invitrogen) has revealed two ORFs in the virion sense BV1 and BV2 and an ORF in the complementary sense.

DNA A of SqLCV-Eg severe isolate has a coding capacity of 2653bp of the SqLCV-Eg genome is shown in Fig. 3. In the virion sense orientation (plus-strand), two potential proteins (V1) are encoded. In the complementary sense orientation (minus-strand), four open reading frames (ORFs) are found (ORFs C1 to C4), fig. 3. By analogy with other WTGs, ORF AV1 encodes the coat protein from 212-978) (CP), ORF AV2 is the movement-associated protein, ORF AC1 (1460-2508) encodes a replication-associated protein (Rep) (Laufs et al., 1995), AC2 (1123-1519) protein (TrAP) transactivates virion-sense gene expression (Hung and Petty, 2001), AC3 protein (979-1383),(REn) is the replication enhancer protein (Pedersen and Hanley-Bowdoin, 1994) and ORF AC4 probably determines symptom severity **(Rigden *et al.*, 1994)**. The intergenic region (IR) includes the conserved nonanucleotide sequence TAATATTAC present in all geminiviruses sequenced so far **(Ikegami *et al.*, 1988)**.

The SqLCV isolate from Egypt has about 97.8% homology with the SqLCV isolates from *Squash leaf curl virus* isolate LB2 segment DNA-A isolate ENA|HM368373 registered in the genebank. While, the least level of homology with *Chenopodium leaf curl virus* isolate Gezira ENA |HM62651 with 74.4 %. While for component B *Squash leaf curl virus* LB2 isolate DNA-B ENA|HM368374 with 97.2% and the least homology was *Tomato severe rugose virus* isolate Pi-1 DNA- B, ENA|HQ606468 with 68% homology.

In total nucleotide sequence comparisons with other geminivurses, SqLCV-Eg was closely related to a previously recorded isolate from Cairo DQ285019 with 97.6%.

Recombination detection for DNA A:

DNA A of SqLCV shows one potential recombination event at (breakpoint at base No. 1321 and ending at base No. 1510) Maxichi with a probability ($P=9.99 \times 10^{-1}$), major parent is *Rhynchosia mosaic sinaloa virus* DNA-A ENA|DQ 4056203 and minor parent is *Chenopodium leaf curl virus* DNA-A [VEM] ENA|HM626516, complete genome. Only trace evidence of recombination was identified and *Rhynchosia mosaic sinaloa virus* DNA-A ENA|DQ 4056203 or *Chenopodium leaf curl virus* DNA-A [VEM] ENA|HM626516 may be actual recombinant.

Figure 1. A,Developed symptoms on *Nicotiana bentamina* inoculated with Agrobacterium harboring constructed *P Cambia* with SqLCV as an infectious clone. B Characteristic symptoms of collected samples from different regions in Egypt, Giza, Cairo, Elboustan, ELmansoria.

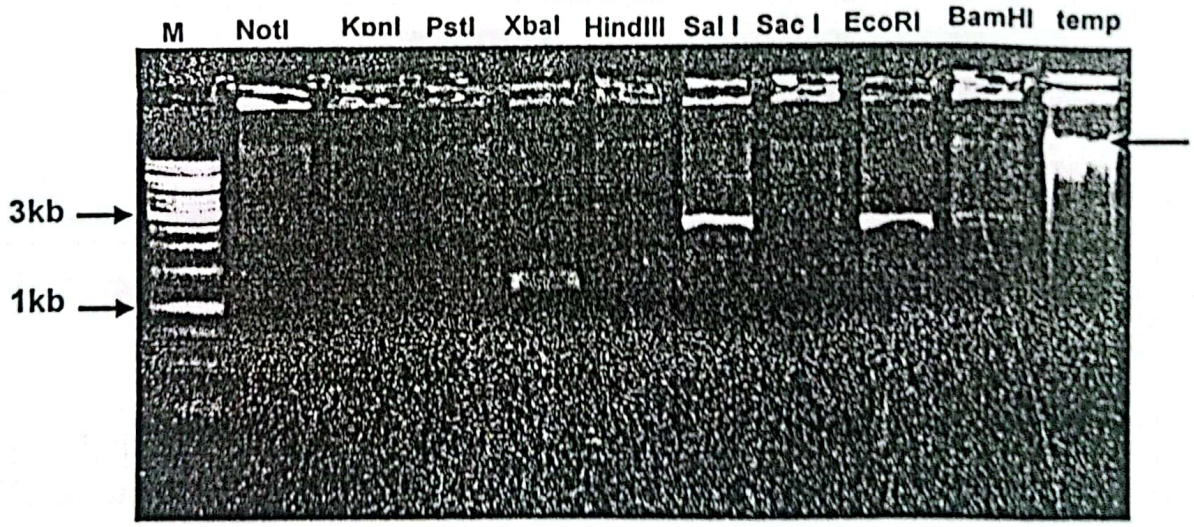

Figure 2. restriction digestion of amplified squash leaf curl virus, lane 1, 1 kb marker (Fermentas), restriction digestion with NotI (lane 2), KpnI (lane 3), PstI (Lane 4), XbaI (lane 5), HindIII (lane 6), SalI (lane 7), Sac I (lane 8), EcoRI (lane 9), BamHI (lane 10), templification product using Phi 29 (fermentas) lane 11. Note (templified products see arrow).

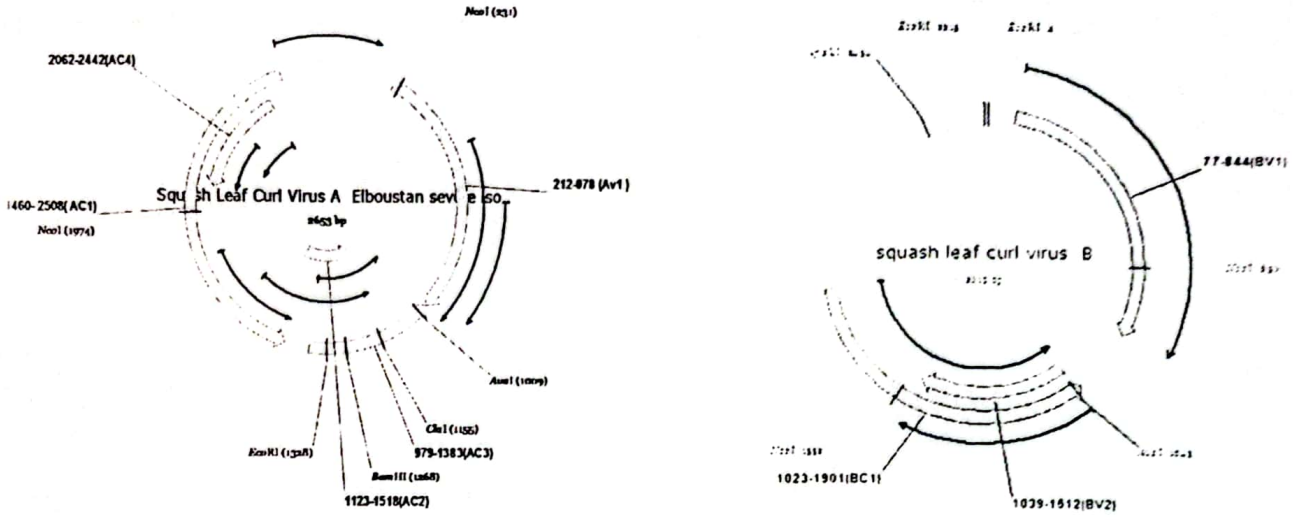

Figure 3. ORFs distribution of A and B component Squash Leaf Curl Virus using Vector NTI program.

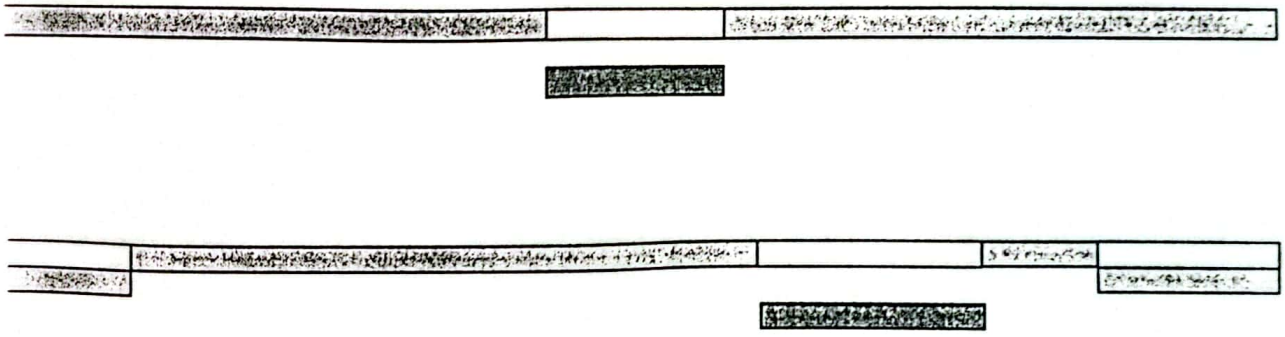

Figure 4. Schematic representation of recombination analysis using RDP3 (version 2.0) program. A) SqLCV component A, B) SqLCV component B.

There are several recombinants identified for SqLCV DNA-B as represented in fig. 4, one potential recombination event at (breakpoint at 1960 and ending at 0 and from 0 to 242) SiScan with a probability ($P=1.88X\ 10^{-7}$), major parent is

Squash leaf curl virus isolate LB2 HM|368374 and minor parent is *Squash leaf curl virus* ENA|EF532620, and *Squash leaf curl virus* ENA|EF532620 may be actual recombinant. Recombination event number 2 using SiScan at (breakpoint starting at 1449 and ending at 1932) with a probability ($P=9.8X\ 10^{-3}$), major parent is *Euphorbia mosaic virus* Brazil Mato Grosso do sul ENA|FN435996 with a percentage of 63.1% and minor parent *Tomato mottle taino virus* ENA|AF012301 and both of them may be actual recombinant.

.Table 2: Representing Squash Leaf Curl Isolates from different areas and their percentage of homology with the SqLCV Elboustan isolate.

Code	Replication-associated protein	aa	% homology
TR:F1C387_SLCV	AC1 Squash leaf curl virus (Israel) HQ184436	349	96.2
TR:D9J2V2_SLCV	AC1 Squash leaf curl virus (Lebanon) HM368373	348	96.5
TR:Q2XNA4_SLCV	AC1 Squash leaf curl virus (Cairo) DQ285019	348	96.2
TR:Q2XNB3_SLCV	AC1 Squash leaf curl virus isolate 10 USA:Imperial Valley, California DQ285016	348	95.8
TR:Q336J7_SLCV	AC1 Squash leaf curl virus WAZ GN=AC1 USA Arizona AF256203	348	95.1
SP:REP_SLCV	Squash leaf curl virus GN=AC1 USA M38183	347	95.5
TR:A5HJ01_SLCV	AC1 Squash leaf curl virus Jordan EF532620	348	91.6
TR:Q8QYY1_9GEMI	Squash mild leaf curl virus-[Imperial Valley] USA:California AF421552	349	84.0
TR:Q2XNC0_9GEMI	Squash mild leaf curl virus -[Imperial Valley] USA:California DQ285014	349	84.0
TR:F5CPN5_9GEMI	AC1 Euphorbia yellow mosaic virus - Goias [Brazil] JF756675	384	81.0
TR:F5CPN0_9GEMI	AC1 Euphorbia yellow mosaic virus - Goias [Brazil JF756674	384	81.0

```
trSquash        ------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
tr|F1C387|F1C38------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
tr|D9J2V2|D9J2V------------------------MPRNPNSFRLTARNLFLTYPRCDVPKEEVLEMLLH
tr|Q2XNA4|Q2XNA------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
tr|Q2XNB3|Q2XNB------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
tr|Q336J7|Q336J------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
sp|P29048|REP_S------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLEMLLH
tr|A5HJ01|A5HJ0------------------------MPRNPNSFRLTARNIFLTYPRCDVPKEEVLGELRI
tr|Q8QYY1|Q8QYY------------------------MPRNPNSFRFTARNIFLTYPQCDIPKDEAIQMLEH
tr|Q2XNC0|Q2XNC------------------------MPRNPNSFRFTARNIFLTYPQCDIPKDEAIQMLEH
tr|F5CPN5|F5CPNMGHQDTTKNRFLRGLPLVSWLIFSIMPRNSNSFRLTAKNIFLTYPRCDIPKDEVLQLLRD
tr|F5CPN0|F5CPNMGHQDTTKNRFLRGLPLVSWLIFTIMPRNPNSFRLTAKNIFLTYPRCDIPKDEVLQLLRD
                  ****.****:**:*:*****:**:**:*.:  *

trSquash        LSWSVVKPNYIRMAREEHSDGSPHLH-------------RG-FNYLENPTSRMLDFRPHS
tr|F1C387|F1C38LSWSVVKPNYIRVAREEHSDGSPHLHCLIQLSGKSNIKDAGLFFDLTHP-RRSARFHPNI
tr|D9J2V2|D9J2VLSWSVVKPNYIRVAREEHSDGSPHLHCLIQLSGKSNIKDAG-FFDLTHP-RRSARFHPNI
```

```
tr|Q2XNA4|Q2XNALSWSVVKPNYIRVAREEHSDGSPHLHCLIQLSGKSNIKDAG-FFDLTHP-RRSARFHPNI
tr|Q2XNB3|Q2XNBLSWSVVKPNYVRVAREEHSDGSPHLHCLIQLSGKSNIKDAG-FFDLTHP-RRSARFHPNI
tr|Q336J7|Q336JLSWSVVKPNYVRVAREEHSDGSPHLHCLIQLSGKSNIKDAG-FFDLTHP-RRSARFHPNI
sp|P29048|REP_SLSWSVVKPNYVRVAREEHSDGSPHLHCLIQLSGKSNIKDAG-FFDLTHP-RRSARFHPNI
tr|A5HJ01|A5HJ0LSVVSGKPNYVRVAREEHSDGAPHLTRLKQLSGKSNNKELG-IFPTNTQ-RRSARFHPKI
tr|Q8QYY1|Q8QYYLQWSVVKPTYIRVAREEHSDGFPHLHCLIQLSGKTNIKNAR-FFDLTHP-RRAACFHPNI
tr|Q2XNC0|Q2XNCLQWSVVKPTYIRVAREEHSDGFPHLHCLIQLSGKTNIKNAR-FFDLTHP-RRAACFHPNI
tr|F5CPN5|F5CPNLPWAVVKPTYIRVARELHADGFPHLHCLIQLSGKSNIKDAK-FFNLTHP-RRSAEFHPNA
tr|F5CPN0|F5CPNLPWTVVKPTYIRVARELHADGFPHLHCLIQLSGKSNIKDAR-FFNLTHP-RRSAEFHPNA
                *     **.*:*:*** *:** * *  :        *    *:*:
```

Figure 5. Amino acid sequence of the N terminus of the Rep protein colored parts representing Rep motifs I (FLTY) and II (HLH) conserved among different *Geminiviruses*. Tr squash representing our new studied isolate.

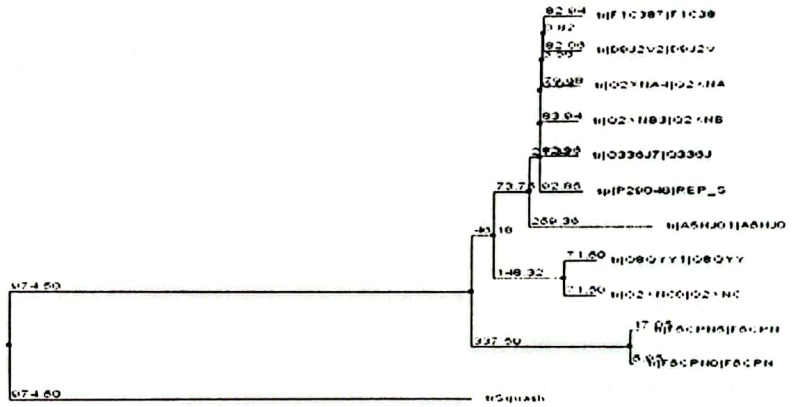

Figure 6. Phylogenetic tree based on the alignment of full Rep amino acid sequence showing the relationship of SqLCV AC1 on component A with other 11 *Begomoviruses*.

Phylogenetic analysis:

For Squash Leaf Curl DNA-A the phylogenetic analysis of severe Egyptian isolate with other *Begomoviruses* revealed that they are divided into three groups with an out group *Bean Calico Mosaic Virus* : ENA|AF110190, group no 3 is divided into two subgroups A and B group A comprised Euphporbia viruses and group B comprises all squash leaf curl viruses the nearest isolates is the *Squash leaf curl virus* isolate LB2 ENA|HM368373 (Lebanon) and Squash leaf curl virus isolate Cairo :ENA|DQ285019 isolate fig. (7). While for *Squash Leaf Curl* DNA-B the phylogenetic analysis of

severe Egyptian isolate with other *Begomoviruses* revealed that they are divided into three groups group one comprises all *Squash Leaf Curl Virus* the nearest isolate in the first group Squash leaf curl virus strain SLCV-WAZ USA:ENA |AF256203 fig. (8).

Comparisons between Rep amino acid sequence of *Squash leaf curl virus* Elboustan isolate and other isolates reveals the highest sequence amino acid homology represent 96.5% with SqLCV from Lebanon HM368373 and then from Israel with a 96.2% the lowest homology was with *Euphorbia yellow mosaic virus* from Brazil with 81 % homology see table (2).

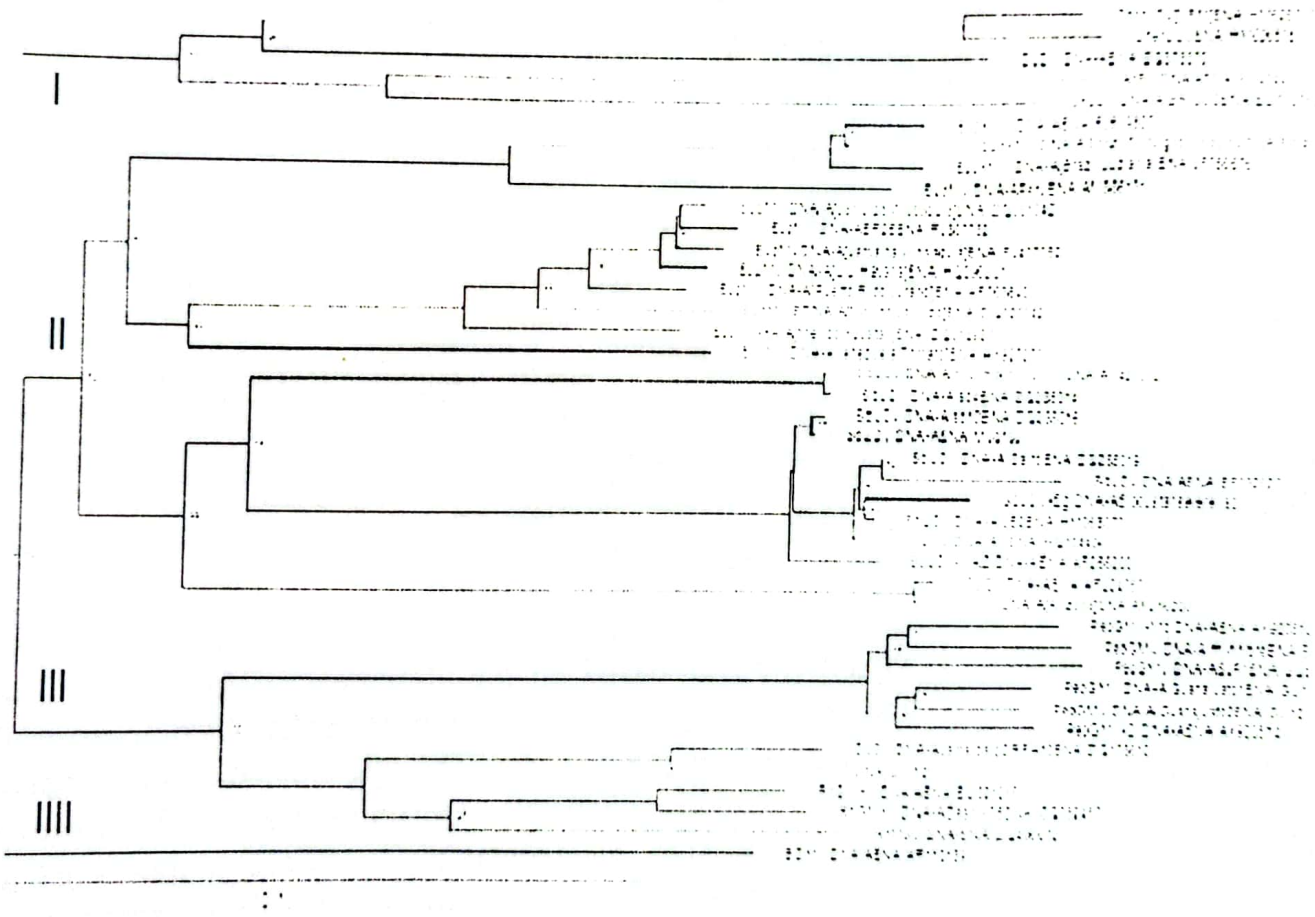

Figure 7. Phylogenetic tree based on the alignment of full DNA virus sequence showing the relationship of SqLCV component A with other 40 *Begomoviruses*. The tree was constructed using

the neighbor-joining method algorithms with 100 bootstrap replications in RDP. Horizontal distances are proportional to sequence distances (shown in scale bar); vertical distances are arbitrary. The number at each branch indicates the number of times in which the given branch is supported (level of confidence).

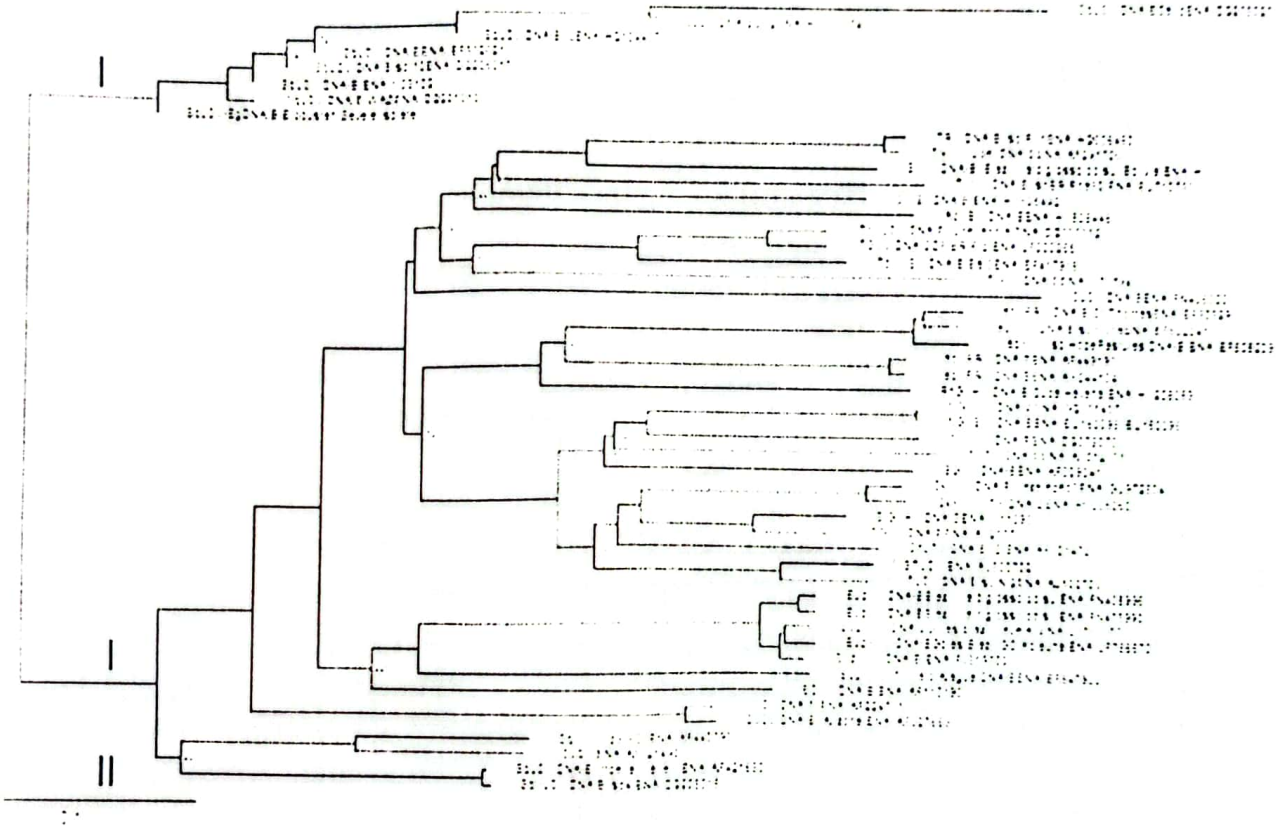

Figure 8. Phylogenetic tree based on the alignment of full DNA virus sequence showing the relationship of SqLCV component B with other 50 *Begomoviruses*. The tree was constructed using the neighbor-joining method algorithms with 100 bootstrap replications in RDP. Horizontal distances are proportional to sequence distances (shown in scale bar); vertical distances are arbitrary. The number at each branch indicates the number of times in which the given branch is supported (level of confidence).

DISCUSSION

Sequence and recombination analysis:

Sequence analysis of a full-length clone of SqLCV from Egypt:

In this study, the RCA using Φ29 DNA polymerase was used to characterize the full length genome and verifying the presence of the DNA ß which may associate with SqLCV causing severe symptoms. The main objective of this study is to generate an infectious clone of full virus sequence capable of infecting healthy plants at any time for preserving the Egyptian isolate. Another reason for preserving our isolate is its further usage in our studies in investigating plant resistance genes and exploring the effect of employing serial silencing experiments of different virus genes on suppressing virus genes thus interfering with virus life cycle to eliminate its fatal effect on squash and melon crops.

Through our experiments we have proved that developing these severe symptoms is not due to the association of DNA ß which has been always reported as a major cause of symptom severity in several research papers such as *Cotton leaf curl virus* from Pakistan **(Briddon *et al.*, 2001)**, or *Tomato leaf curl Java virus* from Indonesia **(Kon *et al.*, 2006)** which contribute to symptom severity.

A 33-base potential stem-loop forming region SqLCV-Eg severe isolate has a 33bp (AGCGGCCATCCGTATAATAT TACCGGATGGCCG) was found in the intergenic region (IR). It includes the conserved nona-nucleotide sequence TAATATTAC present in all *Geminiviruses* sequenced so far (Ikegami et al., 1988). The Nona-nucleotide sequence is marked in DNA B 2151 to 2159) and in DNA A from (1, 2- 2653) represented in fig 3. The first nucleotide sequence of SqLCV of Egypt was reported in 2006 **(Idris *et al.*, 2006)** characterized by the traditional abutting primers, while this is the first time to produce an infectious clone upon getting the virus sequence via RCA using Φ29 DNA polymerase in Egypt and the Arab world.

Rep is the only *Geminivirus* protein that plays the only key role for viral replication **(Rogers *et al.*, 1989)**. It is a multifunctional protein that mediates virus-specific recognition of its cognate origin **(Fontes *et al.*, 1994)** and

transcriptional repression. Rep initiates and terminates viral DNA synthesis (**Laufs *et al.*, 1995**) and plays significant role in inducing the accumulation of host replication factors in infected cells (**Nagar *et al.*, 1995**). Rep binds specifically to double-stranded DNA (dsDNA) at a repeated sequence in the 5` intergenic region of the viral genome, cleaves and ligates DNA within an invariant sequence in a hairpin loop of the plus-strand origin (**Laufs *et al.*, 1995**), and acts as a DNA helicase to unwind viral DNA during plus-strand replication (**Singh *et al.*, 2008**). Rep is involved in a variety of protein interactions. The formation of Rep homo-oligomers is required for its dsDNA binding and DNA helicase activities (**Orozco *et al.*, 1996**). Rep activity may also be modulated by interaction with AC3 (**Settlage *et al.*, 2001**), which is required for high levels of viral DNA accumulation with coat protein, which down regulates DNA cleavage and ligation activity *in vitro* (**Malik *et al.*, 2005**). Comparing Rep amino acid sequence of *Squash leaf curl virus* Elboustan isolate and other isolates revealed the highest sequence amino acid homology represent 96.5% with SqLCV from Lebanon HM368373 and then from Israel with a 96.2% the lowest homology was with *Euphorbia yellow mosaic virus* from Brazil with 81 % homology as clarified in the results, this limited evolution of Rep protein may be due to the complete dependence of the virus on this gene for its life maintenance and the conserved domains that affect the virus life cycle which should not be prone to dramatic changes.

The N terminus of Rep contains three conserved sequences called motifs I, II, and III that are characteristic of many rolling-circle initiators (**Koonin *et al.*, 1992**). Motif I (FLTY) is required for specific dsDNA binding while motif II (HLH) is a metal-binding site that may be involved in protein conformation and DNA cleavage. Motif III (YxxKD/E) is the catalytic site for DNA cleavage, with the hydroxyl group of the Y residue forming a covalent bond with the 5' phosphoryl group of the cleaved DNA strand (**Laufs *et al.*, 1995**).

The first recombinant identified for SqLCV DNA- B as mentioned in recombination analysis, which was a potential recombination event at (breakpoint at 1960 and ending at 0 and from 0

to 242) SiScan with a probability ($P=1.88X\ 10^{-7}$), major parent is *Squash leaf curl virus* isolate LB2 HM|368374 and minor parent is *Squash leaf curl virus* ENA|EF532620, and *Squash leaf curl virus* ENA|EF532620 may be actual recombinant, this data can be interpreted as that the Egyptian isolate respectively when we talk about DNA-B is mix of recombinants between two isolates the first one which has the major identity to our isolate is the SqLCV LB2 HM|368374 found in Lebanon and the SqLCV ENA|EF532620 from Jordan and this strain has been translocated to Egypt and Lebanon as well as to Jordan through our borders with Israel where it has got the virus from the new world. From this study it has been observed that the most identity of the whole virus with its two component is nearly close to LB2 isolate reported in Lebanon with a 97.8% identity and the virus has been first reported in Egypt in 2006, because *Squash leaf curl* Gemini-virus (SqLCV) was first observed in squash in California during 1977 to 1978 **(Flock and Mayhew, 1981)** and in cultivated buffalo ground in Arizona at about the same time **(Rosemey *et al.*, 1986)**. As it is a new world virus it has been transferred to Israel at the same time from the new world to the old world and it has been translocated to different countries in the Middle East and reported subsequently at first in Egypt by **(Farag *et al*, 2005)** and in Jordan by **(Al-Musa *et al.*, 2008)** and by **(Samsatly *et al.*, 2010)** in Lebanon and as the virus was first purified and characterized by **Cohen *et al.*, (1983)** it means that it has been translocated to the three countries and reported consequently.

Through recombination analysis we have noticed that more pressure has been found to affect DNA-B than DNA-A and it has more recombination events observed which could be interpreted as a result of pseudo-recombination. This might has happened due to the fewer genes that DNA-B encodes thus being more susceptible for variation as DNA-B evolves exclusively in response to the host plant pressure while DNA-A evolves in response to the host in addition to its insect vector which could provide more restrictions to its rapid diversification and thus co-evolution.

REFERENCES

Al-Musa, A.; G. Anfoka; S. Misbeh; M. Abhary and F.H. Ahmed (2008). Detection and molecular characterization of *Squash leaf curl virus* (SLCV) in Jordan. J. of Phytopathology, 156(5): 311.

Briddon R.W. (2001). Identification of DNA components required for induction of cotton leaf curl disease. *Virology* 285:234-243.

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 of *Begomoviruses*. Arch. Virol, 146: 1581-1598.

Brukner I.; D. Labuda and M. Krajinovic (2006). Phi29-based amplification of small genomes. Analytical Biochemistry, 354: 154-156.

Cohen S.; J.E. Duffus; R.C. Larson; Y.H. Lis and R.A. Flock (1983). Purification, serology and vectorrelationships of squash leaf curl virus a whitefly–transmitted Geminiviruses. Phytopathology, 73: 1669-1673.

Dry, I. B., Rigden, J. E., Krake, L. R., Mullineaux, P. M. & Rezaian, M. A. (1993). Nucleotide sequence and genome organization of tomato leaf curl geminivirus. *Journal of General Virology* 74, 147-151.

Fahmy I. F., Taha O., El-Dougdoug K., (2011). Generation of infectious clone of *Tomato Yellow leaf Curl Virus* Egyptian isolate using the Φ29 DNA polymerase,. *Egy.J. of Virol.Vol 8:18-37.*

Farag A.G., Amer M.A., Amin H.A., and Mayzad H.M., (2005). Detection of Bipartite Gemininviruses Causing *Squash Leaf Curl* Disease in Egypt using Polymerase Chain Reaction and Nucleotide Sequence. Egyptian J. Virol., 2: 239-354.

Fauquet C. M., Maxwell, D. P. Gronenborn, B. & Stanley J. (2000). Revised proposal for naming geminiviruses. *Archives of Virology* 145: 1743-1761.

Fontes E. P., Eagle P. A., Sipe P. S., Luckow V. A., and L. Hanley-Bowdoin.(1994). Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J. Biol. Chem. 269:8459–8465.

Gibbs M.J.; J.S. Armstrong and A.J. Gibbs (2000). Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics, 16: 573-582.

Hung, H.C. and I.T.D. Petty (2001). Functional equivalence of late gene promoters in *bean golden mosaic virus* with those in *tomato golden mosaic virus*. J. Gen. Virol., 82: 667-672.

Idris, A.M.; A. Abdel-Salam and J.K. Brown (2006). Introduction of the New World *Squash leaf curl virus*to squash (*Cucurbita pepo*) in Egypt: a potential threat to important food crops. Plant Disease, 90(9): 1262.

Ikegami, M.; T. Morinaga and K. Miura (1988). Potential gene products of *bean golden mosaic virus* have higher sequence homologies to those of *tomato golden mosaic virus* than those of *cassava latent virus*. Virus Genes, 1:191-203.

Katoh, K.; G. Asimenos and H. Toh (2009). Multiple alignment of DNA sequences with MAFFT. Methods in molecular biology, 537:39-64

Katoh K. and H. Toh (2010). Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics (Oxford, England), 26 (15):1899-900.

Kheyr-Pour, A.; M. Bendahmane; V. Matzeit; G.P. Accotto; S. Crespi and N. Navot; E. Pichersky; M. Zeidan; D. Zamir and H. Czosneck (1991). Tomato yellow leaf curl virus: a whitefly transmitted geminivirus with a single genomic component. Virology, 185: 151-161.

Kon, T.; S.H. Hidayat; S. Hase; H. Takahashi and M. Ikegami (2006). The natural occurrence of two distinct *Begomovirus* associated with DNAβ and a recombinant DNA in a tomato plant. Phytopathology, 96: 517-525.

Koonin, E.V. and T.V. Ilyina. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. J. Gen. Virol., 73: 2763-2766.

Laufs, J.; W. Traut; F. Heyrand; 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.

Lazarowitz, S.G. (1992). Geminiviruses: genome structure and gene function. Critical Reviews in Plant Sciences, 11(4): 327-349.

Malik, P.S.; V. Kumar; B. Bagewadi and S.K. Mukherjee (2005). Interaction between coat protein and replication initiation protein of Mung bean yellow mosaic India virus might lead to control of viral DNA replication. Virology, 337: 273-283.

Martin, D.P.; C. Williamson and D. Posada (2005). a RDP2: recombination detection and analysis from sequence alignments. Bioinformatics, 21: 260-262.

McCreight, J.D. (1984). Tolerance of *Cucurbita* spp. to *Squash Leaf Curl*. Cucurbit Genetics Cooperative Report, 7: 71-72.

Nagar, S.; T.J. Pedersen; K.M. Carrick; L. Hanley-Bowdoin and D. Robertson (1995). A geminivirus induces expression of a host DNA synthesis protein in terminally differentiated plant cells. Plant Cell, 7: 705-719.

Noris, E.; E. Hidalgo; G. P. Accotto and E. Moriones (1994). High similarity among the tomato yellow leaf curl virus isolates from the West Mediterranean Basin: the nucleotide sequence of an infectious clone from Spain. Archives of Virology, 135: 165-170.

Orozco, B.M. and L. Hanley-Bowdoin (1996). A DNA structure is required for geminivirus replication origin function. J. Virol., 70: 148-158.

Packialakshmi, R.M.; N. Srivastava; K. Girish and R.R. Usha (2010). Molecular characterization of a distinct begomovirus species from Vernonia cinerea and its associated DNA-b using the bacteriophage Φ29 DNA polymerase. Virus Genes, 41: 135-143.

Padidam, M.; S. Sawyer and C.M. Fauquet (1999). Possible emergence of new geminiviruses by frequent recombination. Virology, 265: 218-225.

Pedersen, T.J. and L. Hanley-Bowdoin (1994). Molecular characterization of the AL3 protein encoded by a bipartite

geminivirus. Virology, 202: 1070-1075.

Posada, D. and K.A. Crandall (2001). Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. Proc Natl. Acad. Sci., 98: 13757-13762.

Rigden; J.E.; L.B. Krake; M.A. Rezaian and I.B. Dry (1994). ORF C4 *of tomato leaf curl virus* is a determinant of symptom severity. Virology, 204: 847-850.

Salminen, M.O.; Carr, J.K.; D.S. Burke and F.E. McCutchan (1995). Identification of breakpoints in intergenotypic recombinants of HIV type 1 by BOOTSCANning. AIDS Res Hum Retroviruses, 11: 1423-1425.

Samsatly, J.; Y. Abou-Jawdah and H. Sobh (2010). The INSDC. Plant Protection, AUB, American University of Beirut-Bliss street–Hamra, Beirut, 7546, Lebanon.

Settlage, S.B.; A.B. Miller; W. Gruissem and L. Hanley-Bowdoin (2001). Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. Virology, 279: 570-576.

Singh, D.K.; P.S. Malik; N.R. Choudhury and S.K. Mukherjee (2008). MYMIV replication initiator protein (Rep.): roles at the initiation and elongation steps of MYMIV DNA replication. Virology, 380: 75-83.

Stenger, D.C.; G.N. Revington; M.C. Stevenson and D.M. Bisaro (1991). Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling circle replication of a plant viral DNA. Proc. Natl. Acad. Sci. USA, 88: 8029-8033.