First genome analysis and molecular characterization of Chickpea chlorotic dwarf virus Egyptian isolate infecting squash

Inas Farouk Fahmy · Omnia Taha ·
Abdel Nasser El-Ashry

Received: 8 October 2014 / Accepted: 29 December 2014 / Published online: 14 February 2015
© Indian Virological Society 2015

Abstract This study aims to identifying and characterizing some molecular properties of geminiviruses co-infection in squash field crop cultivated in Egypt. Squash crops observed to be heavily infected with several insect vectors, also severe chlorosis and stunting was observed. Electron microscopic analysis has revealed geminate capsid particles which indicate the infection of Geminiviruses, especially SqLCV which represent an economic problem to squash field crop in Egypt. We have investigated possible mixed infections with different plant viruses associated with chlorotic stunt diseases and or other genus groups of geminiviruses. The main objective of this study is to investigate the recombination events, possible recombinants and variants among these genera in the same family differing in vector transmission. This is the first report of the molecular characterization, phylogenetic analysis and putative recombination events of the full length genome of the Chickpea Chlorotic Dwarf Mastrevirus in Egypt. And the first report of co-infection with another begomovirus infecting squash plants. A full length clone of both viruses were isolated and characterized at the molecular level. The

complete nucleotide sequence of DNA-A was determined (2,572 bp) and submitted to the genbank under accession no. KF692356. The isolate from Egypt has about 97.8 % homology with the Chickpea chlorotic dwarf virus (CpCDV) isolate from Syria DNA-A isolate FR687959, a 83.2 % homology with the Sudan isolate AM933134 and a 82.7 % homology with Pakistan isolate FR687960. To best of our knowledge this is the first report of complete genome of CpCDV that infect squash plants in Egypt and worldwide.

Keywords Rolling circle amplification (RCA) · CpCDV · Masterviruses · Recombination analysis

Introduction

The family Geminiviridae contains four genera that are separated from each other by their genome organization and their transmitting vector. Genus *Begomovirus*, *Curtovirus*, *Topocovirus* and *Mastrevirus*. Genus *Mastrevirus* (type species maize streak virus, MSV) [36] this genus contains the economically important MSV that is endemic in sub-Saharan Africa. Masterviruses are transmitted by leafhoppers in a circulative, non-propagative manner.

All members have narrow host ranges and, with exception of *Tobacco yellow dwarf virus* (TYDV) and *Bean yellow dwarf virus* (BeYDV) which infect dicotyledenous species, their host ranges are limited to species in the poaceae.

Chickpea chlorotic dwarf virus (CpCDV) was first reported in *Cicer arietinum*; from India; by [17]. Typical of all mastreviruses, CpCDV is a circular; ~2.7-Kb monopartite single-stranded DNA genome encodes only four proteins. As described in earlier studies the

Electronic supplementary material The online version of this article (doi:10.1007/s13337-014-0246-4) contains supplementary material, which is available to authorized users.

I. F. Fahmy (✉) · O. Taha
Phytopathogen Vector Interaction Lab, Department of
Microbiology, Agricultural Research Center (ARC), Agricultural
Genetic Engineering Research Institute, Giza, Egypt
e-mail: inasfaroukfahmy2000@gmail.com;
inasfaroukfahmy@gmail.com

A. N. El-Ashry
Department of Plant Pathology, Faculty of Agriculture, Institute
für Nutzpflanzenwissenschaften und Ressourcen Schutz,
University of Bonn, Bonn, Germany

transcription of masterviruses is bi-directional through a long intergenic region. The virion-sense expression comes out with the movement protein (MP) and a coat protein (*Cp*), and the complementary-sense expression of the replication-associated proteins, Rep and RepA. Whereas the MP and *Cp* are involved in virus movement and encapsidation [25], Rep is an crucial initiator of virus replication, and RepA plays regulatory role of host and viral gene transcription [15, 16]. Masterviruses family adopts host DNA replication and double-stranded DNA break repair proteins in their genome replication via rolling-circle mechanism [2] and recombination-dependent mechanism [8].

The first record of CpCDV was reported to infect faba bean in Egypt in 2003 [26]. CpCDV is the only member of the *Geminivirus* that is reported to naturally infect faba bean, and to our knowledge no molecular characterization for the virus has been reported to the Egyptian isolate till to date. Also this is the first report on infection of *chlorotic stunt disease* (CDS) of squash plants worldwide which puts a huge pressure on possible recombination events between two genera in the same family which are transmitted by different vectors.

Materials and methods

CpCDV isolates

Different squash plants (*Cucurbita pepo* L.) with severe infections showing leaf curling, yellow mottling, stunting and reduced fruit set in different regions from crop fields in Giza governorate, Elmansoria region, 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 (Fermentas). 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. [3] 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 Mansoria area was selected for further amplification and cloning procedures as described by [10]. Finally, Φ 29 DNA polymerase was denatured at 65°C for 10 min. Two microliters of the templiPhi product

was run on a 1.2 % agarose gel to confirm amplification (Fig. 1). Two microliters of templified product were digested using different common restriction enzymes *EcoRI*, *HindIII*, *BamHI*, *SalI*, *SacI*, *PstI*, *XbaI*, *KpnI*, *NcoI* (Fermentas, fast digest). Digestion reaction was carried out for 20 min at 37°C with 2 U of enzyme. Digestion product was analyzed on 1.2 % agarose gel to confirm the virus size.

Cloning and sequencing

The enzymes that cut once in putative virus components A were chosen for further cloning steps in pBlue script II (KS−). Several putative virus clones were sent for sequencing to confirm their identity. After getting sequencing results, a set of primers were designed for walking to obtain the full sequence of viral genome as described in Table S1, the sequenced data were analyzed using BLASTN and submitted to the EMBL database.

Sequence analysis and construction of phylogenetic tree

The genomic sequences of DNA-A, -B molecules were assembled from contiguous sequences using the Vector NTI program (Qiagen), the open reading frames (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 (Qiagen) 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

Fig. 1 Restriction digestion of templified products from squash isolates sample no. 9, lane 1, 1 kb marker (Biolabs), restriction digestion with *SalI* (lane 2), *KpnI* (lane 3), *PstI* (lane 4), *BamHI* (lane 5), *XbaI* (lane 6), *EcoRI* (lane 7), *HindIII* (lane 8), *SacI* (lane 9)

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 [19–21] available at the website of EBI. And phylogenetic trees were constructed using the neighbor-joining method using the Recombination Detection Program version 4.0 (RDP4).

Possible gene-recombination events were analyzed using the RDP4 [29], available online (<http://darwin.uvigo.es/rdp/rdp.html>). The recombination analyses were applied using six automated methods: RDP [27, 29], GENECONV [35], BootScan [40], SiScan [12, 40], MaxChi and CHI-MAERA [31, 38], 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 CpCDV

In this study an infectious head-to-tail 2.7-mer of CpCDV A clone was prepared by the digestion of approximately 10 µg recombinant plasmid DNA using the restriction enzyme (*Bam*HI, for component A) 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 backbone. The linearized genome fragment was recovered from agarose as described using Qiax gel clean (Qiagen). 5 µg 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 using Qiax gel cleans (Qiagen). The insert and vector were ligated at a molar ratio of at least 10:1, respectively, using 5 U ligase (Promega, USA) and 100 ng of *pCambia* 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*. The clone should be stored in glycerol at −70 °C. A glycerol stock of the *Agrobacterium* strain LBA 4404 carrying dimers of the respective DNA-A genomic components of CpCDV were prepared according to [22]. Bacterial cultures harbouring plasmids with DNA-A was 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 CpCDV A construct was grown for 36–48 h and inoculated

into the axillary buds of tobacco (*Nicotiana tabacum*) plants as described previously [6]. The agroinoculated plants were evaluated for symptom appearance 21–45 days post-inoculation. 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 begomovirus namely V324 (+) and C889, the primer pair was selected according to [26].

Results

It was discovered that the CpCDV-Eg was infecting squash plants in a mixed infection with SqLCV which lead to severe infection of squash plants in Mansoria, and Giza region. These severe symptoms on squash plants were thought to be due to association of *Beta* satellites which found to be a major contribution to symptom severity in most geminiviruses.

All collected samples were checked for geminivirus infection via PCR (data not shown). It has been proved that the templified product cut with *Bam*HI was representing the A component of *Squash Leaf Curl Virus*. This may lead us to the conclusion of either a point mutation has been occurred naturally in SqLCV-Mansoria strain abolishing the *Bam*HI cutting site, as the presence of component A was verified using *Cp* primers amplifying a 735 bp representing the *Cp* of the respected virus. Upon sequencing of three samples of A component, CpCDV was found as one of the A component samples.

The above shown Phi polymerase amplified products subjected to sequencing and it has been revealed that the product cut with *Bam*HI was representing CpCDV-Eg isolate submitted to the genbank under accession no. KF692356, and the product that cuts with *Eco*RI, *Sal*I, *Kpn*I, *Not*I (data not shown) which gave 2 bands one is about 1.7 kb, and the other is 1 kb, was representing *Squash Leaf Curl Virus* B component. While the enzymes *Pst*II, *Hind*III (lane 8), *Sac*I (lane 9) do not cut any of the templified products for both viruses. The enzyme *Xba*I gave two bands about 1.3 kb. In a former study done on another strain called El Boustan of SqLCV [9]. To the best of our knowledge this is the first time to report the infection of CpCDV to any variety of squash crop world-wide, which is not recorded as a host of this virus in all virus databases.

Sequence analysis and phylogenetic comparisons within Mastervirus genus

The analyzed squash samples revealed two viruses in mixed infection CpCDV and SqLCV. CpCDV sequence analysis has revealed that the virus consists of 2,572 nt and

was deposited in the genbank under accession number KF692356 (Fig. 2), the bioinformatics analysis using Vector NTI 11 Suite revealed four ORFs two in the virion sense which are the *Cp* from starts from nucleotide number 422–1,157 and the *MP* starts in the nucleotide position (131–400 nt) in the, while in the negative sense encodes the *RepA* from the position 1,533–2,558 and contains the *RepB* from 1,988 to 2,368.

All geminiviruses share a GC rich stem loop located in the intergenic region and the nona-nucleotide TAA-TATTAC composing the loop is the only conserved sequenced among all geminiviruses.

The expressed viral proteins similarity chart compared with representing isolates from several countries, such as Syria (FR687959), Pakistan (FR687960), India (JF831147), and Sudan (AM933134). The highest similarity with *RepA* is 99.3 % with the Syrian isolate while the lowest 85.4 % with the isolate from Pakistan.

An infectious clone of the CpCDV virus was produced in order to further study the virus and its infectivity competence to squash plants and to investigate the virus at the molecular level. Meanwhile, studying these details can help at last in producing resistant line against this damaging virus in a later stage.

Phylogenetic analysis (Fig. 4), indicated that the CpCDV-Eg isolate grouped most closely with a clade containing begomoviruses from Syria, Sudan and Pakistan, including the most closely related CpCDV-Syr-2. The most closely related viruses to these two isolates from Syria and from Egypt is several BeYDV (Fig. 3). We will find that through recombination analysis that the BeYDV ENAIDQ458791 representing the viral genome which both the CpCDV from Syria and Egypt has been evolved. Also located in the same clade the *Chickpea Chlorosis Virus* with

61–64 % homology and the TYDV with a homology of 66.4 %. The cluster of *Panicum Streak Virus* is distantly related to our Egyptian isolate with 56 % homology. Recombination analysis as demonstrated in Fig. 4 has revealed that the overall genomes analysis of DNA-A of CpCDV-Eg and all the analyzed genomes listed in Table S2 showed a total of 47 unique recombination events, 22 unique events using the RDP [27, 29], 12 unique events using GENECONV [35] automated method and 30 using MaxChi [38]. For the CpCDV one potential recombination event detect by the RDP automated method [27, 29] at (breakpoint at base no. 1078 in alignment and ending at base no. 1474 in alignment) with a probability of $P = 1.404 \times 10^{-08}$, the recombinant CpCDV has a major parent CpCDV clone 8 ENAIDM900416 with a similarity of 84 %, and a minor parent BYDV ENAIDQ458791 with a similarity of 64.4 %. This recombination event was confirmed several times using MaxChi with a probability ($P = 1.5 \times 10^{-06}$), GENECONV with a probability ($P = 1.3 \times 10^{-02}$), BootScan probability ($P = 3.8 \times 10^{-05}$), CHIMAERA probability ($P = 4.7 \times 10^{-04}$), and SiScan probability ($P = 7.1 \times 10^{-12}$) see Fig. 4b.

In this study it has been observed upon sequencing of the full genome of CpCDV-Eg that the start codon of the *RepA* has been changed from ATG to GTG. It has been observed through several studies has proved for example in yeast in 2010, that thought investigating different genes like *ALA1* encodes alanyl-tRNA synthase and *GRS1* encodes glycyl-tRNA synthase use ACG and TTG as their alternative translation start codon. It has been proved by [5] that except for AAG and AGG, all triplets that differ from ATG by a single nucleotide were able to serve as initiator codons in *ALA1*. It has been concluded that from these initiator codons, TTG, CTG, ACG, and ATT that had ~50 % initiating activities compared to that of ATG, they indicated that the GTG, ATA, and ATC had ~20 % initiating activities compared to that of ATG. Unexpectedly, these non-AUG initiator codons showed different favoring actions toward different sequence backgrounds. Specifically, GTG was one of the most efficient non-ATG initiator codons [5, 7].

Discussion

In the present study, the CpCDV isolate infecting squash field crops in a mixed infection with SqLCV in Mansoria and Giza governorates in Egypt causing chlorosis, severe stunting and 100 % yield loss to the farmers in that area, CSD belongs to the mastervirus family particularly infecting chickpea present major threat to the crop production in North Africa and Asia [14]. Despite the huge effect of CSD and their wide distribution, very few species has been identified like in Sudan and Pakistan [24], and another

Fig. 2 ORFs distribution of A component *Chickpea chlorotic dwarf virus* using Vector NTI 11 program. *Cp* coat protein, *MP* movement protein, *RepA*, *B* replication protein A, B

Fig. 3 Phylogenetic tree based on the alignment of full DNA virus sequence showing the relationship of CpCDV 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)

three identified in possibly new species Australia [41]. Also the CpCDV from Syria has been directly identified after these new species [33].

CpCDV, a leafhopper-transmitted virus, was formerly reported in Egypt in New Disease Report in March 2002, in two samples of *Vicia faba* collected from the El-Nobarria governorate showing leaf rolling, yellowing and stunting with no further molecular characterisation. However severe chlorosis and stunting symptoms were observed on squash crop in Mansoria and Giza fields. By means of rolling circle replication (RCA) it has been proved that CpCDV infect Squash in a mixed infection with *Squash Leaf Curl Virus-Eg* isolate in Mansoria and Giza which in that case cause a lot of economical loss of squash and cucurbit yields in Egypt, may be extended to different regions. It has been reported formerly that the family cucurbitaceae is not one of the favored hosts for CpCDV [4] which point to a higher degree of evolution in the genetic make-up of the virus enhancing the putative adaptation response to infect a host such squash crop.

The two isolates found in the mixed infection have been characterized at the molecular level and possible recombination events within the same genus and other *Geminivirus* genus has been studied. Although the infection of CpCDV disease the causal agent of mild chlorosis and stunting of sugar beet in Iran has been reported as CpCDV [11, 34], no evidence on the infection of this *Mastrevirus* to any vegetable crop like squash worldwide has been reported. It is clearly that due to the wide geographical spreading of *Mastrevirus* in Africa and Asia that a potential biological diversity in the mastrevirus family is arisen, thus to be further studied and analyzed. The identified virus infecting squash is a typical virus to the genus *Mastrevirus*, it contains four predicted ORFs that possibly encode four gene products with greater degrees of similarity to corresponding proteins found in other dicot-infecting mastreviruses. Typically, for mastreviruses, the Rep is translated from a spliced transcript that functions through a fusion between the RepA and RepB ORFs. Generally the amino acid sequences of the predicted proteins and the complete nucleotide sequence show the highest levels of identity to other mastreviruses identified in southern Asia and North Africa [33]. Through our recombination studies and in most cases, except for the amino acid sequences of the virion-sense genes, the highest levels of sequence identity are to BeYDV, in our study through RDP the BeYDV ENAIDQ458791 several recombination events has been detected between the CpCDV Egyptian isolate and this above described isolate in Fig. 4b.

According to species demarcation threshold which known to be >78 % where the virus considered to belong to the same species, and it was further analyzed by [32], where the results suggested that mastrevirus genomes that

calculated to be >94 % should be considered members or variants of the same strain. The Egyptian isolate identified in this study found to have 97.8 % similarity to the Syrian isolate. A great emphasis has been emerged here stressed on the biological difference between the viral isolates would be classified into different strains. For example, whereas the “A” strain of MSV is clearly the only group of MSV variants that cause severe disease in maize [28, 42], the “A” strain of WDV favourably infects barley, whereas the “C” strain preferentially infects wheat [39], our Egyptian isolate of CpCDV could be assigned as a specific strain preferentially infects squash which is essentially a non-host species of this virus. Through recombination analysis performed using RDP4, it has been observed that the Egyptian isolate although more closely related to the Syrian isolate but a considerable part of the genome found to be originated from a different isolate of BeYDV assigned in the genbank under accession no ENAIDQ458791 (Fig. 4) and in the phylogenetic tree (Fig. 3); it is found to cluster with CpCDSYV with different strains of BeYDV in the same clade which supports the evidence that the virus has been arisen through several recombination events resulted from co-infection of the virus to the faba bean when both viruses are belonging to Mastrevirus. This could be supported as an evidence of correct recombination events when we know that the first record of CpCDV was reported to infect faba bean in Egypt in 2003 [26]. Through sequencing results put in the study and the through the RDP datasets, it has been noticed that the recombination patterns detected in the dicot-infecting mastreviruses compared to the monocot infecting mastrevirus are interspecies more than intra-species recombination patterns [23]. CpCDV is the only member of the *Geminivirus* that is reported to naturally infect faba bean. We are reporting here the first time of the co-infection of CpCDV of squash plants worldwide which puts a huge pressure on possible recombination events between two genera in the same family and transmitted by different vectors. It has been reported on the plant virus online description and lists from VIDE database, URL <http://biology.anu.edu.au/Groups/MES/videl/> [4], that the only susceptible host for CpCDV is the cucumber *Cucumis sativus*, while CpCDV has not been reported or observed infecting squash, pumpkin plants *Cucurbita maxima* or *C. pepo*. It has been also reported that the CpCDV has been detected infecting pepper in Oman and the first to be detected and analysed in the Arabian Peninsula [1] which could point the huge host adaptability through recombination events due to the intensive cropping as the case in Egypt.

As it has been mentioned previously with different genomic data collected through studies of CSD recombination has played a key role in evolution of dicot infecting Mastreviruses [13, 30]. The infection of CpCDV to squash

plants could be explained by the enormous pressure that can be faced in Mansoria region in Egypt where the squash field crop is planted in several areas as a unique cash crop very near to faba bean in several planting rounds puts a huge force to the vectors to transfer to another host plant and adapt to these hosts. Due to this pressure it may lead to host expansion and host tradeoffs. There are several examples have described an image signifying that in a homogenous host landscape, selection promotes the evolution of specialist viruses and the existence of a fitness load associated with host-range expansion. Nonetheless, it is fair mentioning that exceptions to this tradeoff rule exist [7]. Or this could be explained as a silencing suppression activity to regain the replication fitness of non-host virus as it has been proved for *Tobacco Mild Green Mosaic Virus* [18].

It has been observed also through different sequencing trails that the RepA gene start codon was GTG rather than ATG this has been also described by Perelygina et al. [37], that it was inferred that the B virus UL1 and UL2 genes use the GTG codon as a start codon at genomic sequence positions 9,072 and 9,814, respectively. Conservation of the UL1 GTG start codon was confirmed by sequence analysis of the 5'-region of the respected gene from two B virus clinical isolates.

An infectious clone of CpCDV has been prepared together with an infectious clone of *Squash Leaf Curl Virus* to be introduced in a future study to newly planted squash variety called Eskandarani, an old local Egyptian variety, to test for change of any genetic makeup through the continuous pressure due to its infectivity of the virus to squash plant, and we are intending to test the possible recombination events between the viruses arising through co-infection and to study the effect of recombination and or co-infection on symptoms severity and squash field crops all over Egypt agricultural areas. It has been observed that there are no sufficient sampling strategies in the Middle East area and in Africa enabling the detection of the specific and correct origin of the dicot-infecting master-virus, which hinder the evolutionary studies detecting possibility of host transfer pathways of these viruses. Success in reaching such information can lead to make a great emphasis on management strategies of these viruses for a specific area in the world.

Through this study and a complementary study in the future our objectives will concentrate on further analyzing the synergism of the two viruses in different areas. This strategy will enable us to investigate gene functionality thus provides a start towards studying the effect of various gene silencing experiments on various virus replication and movement and all functional proteins affecting virus life cycle, and whether it will affect both viruses through targeting silencing experiments of the possible recombination events of both viruses enabling the interference with its

lifecycle and therefore eliminating its spread and its deleterious effect on squash crop in Egypt. Also, elucidating the reasons behind CpCSV host range expansion to infect squash crop facilitating the application of new and effective management strategies based on wise scientific decisions.

Acknowledgments This work has been done by the full funding of the Agricultural Genetic Engineering Research Institute. A partial funding from the Science and Development Fund, STDF during a collection mission funded by Project #892 for whitefly and virus collection has contributed in enhancing our virus collection. The thanks also are extended to Eng. Nasser El-Torky, PICO Company Lab Executive Manager for his help in virus collection in Mansoria and El-Behira region.

References

1. Akhtar S, Khan AJ, Briddon RW. A distinct strain of *Chickpea chlorotic dwarf virus* infecting pepper in Oman. *Plant Dis*. 2014;98(2):286–286.
2. Boulton MI. Functions and interactions of mastrevirus gene products. *Physiol Mol Plant Pathol*. 2002;60:243–55.
3. Brown JK, Idris AM, Torres-Jerez I, Banks GK, Wyatt SD. The core region of the coat protein gene is highly useful for establishing the provisional identification of begomoviruses. *Arch Virol*. 2001;146:1581–98.
4. Brunt AA, Crabtree K, Dallwitz MJ, Gibbs AJ, Watson L, Zurcher EJ, editors. Plant viruses online: descriptions and lists from the VIDE database. Version: 20th August 1996. 1996 onwards. <http://biology.anu.edu.au/Groups/MES/vide/>.
5. Chang CP, et al. A single sequence context cannot satisfy all non-AUG initiator codons in yeast. *BMC Microbiol*. 2010;10:188.
6. Dry IB, Rigden JE, Krake LR, Mullineaux PM, Rezaian MA. Nucleotide sequence and genome organization of *tomato leaf curl geminivirus*. *J Gen Virol*. 1993;74:147–51.
7. Elena SF, Agudelo-Romero P, Lalic J. The evolution of viruses in multi-host fitness landscapes. *Open Virol J*. 2009;3:1–6.
8. Erdmann JB, Shepherd DN, Martin DP, Varsani A, Rybicki EP, Jeske H. Replicative intermediates of *maize streak virus* found during leaf development. *J Gen Virol*. 2010;91:1077–81.
9. Fahmy IF. Molecular characterization of *Squash Leaf Curl Virus* Egyptian isolate and generation of infectious clone via rolling circle amplification. *Egypt J Virol*. 2012;19:131–51.
10. Fahmy IF, Taha O, Eldogdog K. Generation of infectious clone of *Tomato Yellow leaf Curl Virus* Egyptian isolate using the Φ 29 DNA polymerase. *Egypt J Virol*. 2011;18:325–44.
11. Farzadfar Sh, Pourrahim R, Golnaraghi AR, Ahoonmanesh A. PCR detection and partial molecular characterization of Chickpea Chlorotic Dwarf Virus in naturally infected sugar beet plants in Iran. *J Plant Pathol*. 2008;90(2):247–51.
12. Gibbs MJ, Armstrong JS, Gibbs AJ. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*. 2000;16:573–82.
13. Hadfield J, Thomas JE, Schwinghamer MW, Krabberger S, Stainton D, Dayaram A, Parry JN, Pande D, Martin DP, Varsani A. Molecular characterisation of dicot-infecting mastreviruses from Australia. *Virus Res*. 2012;166:13–22.
14. Hamed AA, Makkouk KM. Occurrence and management of Chickpea chlorotic dwarf virus in chickpea fields in northern Sudan. *Phytopathol Mediterr*. 2002;41:193–8.

15. Hefferon KL, Moon Y-S, Fan Y. Multi-tasking of nonstructural gene products is required for bean yellow dwarf geminivirus transcriptional regulation. *FEBS J*. 2006;273:4482–94.
16. Hofer JMI, Dekker EL, Reynolds HV, Woolston CJ, Cox BS, Mullineaux PM. Coordinate regulation of replication and virion sense gene expression in wheat dwarf virus. *Plant Cell*. 1992;4:213–23.
17. Horn NM, Reddy SV, Roberts IM, Reddy DVR. ICTV virus description. *Ann Appl Biol*. 1993;122:467–79.
18. Ishibashi K, Meshi T, Ishikawa M. Gaining replicability in a non-host compromises the silencing suppression activity of *Tobacco mild green mosaic virus* in a host. *J Virol*. 2011;85(4):1893–5.
19. Katoh K, Standley DM. Multiple alignment of DNA sequences with MAFFT. *Mol Biol Evol*. 2013;30(4):772–80.
20. Katoh K, Toh H. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics (Oxf, Engl)*. 2010;26(15):1899–900.
21. Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol (Clifton, NJ)*. 2009;537:39–64.
22. Kheyr-Pour A, Bendahmane M, Matzeit V, Accotto GP, Crespi S, Navot N, Pichersky E, Zeidan M, Zamir D, Czosneck H. *Tomato yellow leaf curl virus*: a whitefly transmitted geminivirus with a single genomic component. *Virology*. 1991;185:151–61.
23. Krabberger S, Harkins GW, Kumari SG, Thomas JE, Schwinghamer MW, Sharman M, Collings DA, Briddon RW, Martin DP, Varsani A. Evidence that dicot-infecting mastreviruses are particularly prone to inter-species recombination and have likely been circulating in Australia for longer than in Africa and the Middle East. *Virology*. 2013;444:282–91.
24. Krabberger S, Mumtaz H, Claverie S, Martin DP, Briddon RW, Varsani A. Identification of an Australian-like dicot-infecting mastrevirus in Pakistan. *Arch Virol*. 2014. doi:[10.1007/s00705-014-2299-5](https://doi.org/10.1007/s00705-014-2299-5).
25. Lazarowitz SG, Pinder AJ, Damsteegt VD, Rogers SG. Maize streak virus genes essential for systemic spread and symptom development. *EMBO J*. 1989;8:1023–32.
26. Makkouk KM, Rizkallah L, Kumari SG, Zakiand M, Abul Enein R. First record of *Chickpea chlorotic dwarf virus* (CpCDV) affecting faba bean (*Vicia faba*) crops in Egypt New Disease Report. *Plant Pathol*. 2003;52:413.
27. Martin DP, Rybicki E. RDP: detection of recombination amongst aligned sequences. *Bioinformatics*. 2000;16:562–3.
28. Martin DP, Willment JA, Billharz R, Velders R, Odhiambo B, Njuguna J, James D, Rybicki EP. Sequence diversity and virulence in *Zea mays* of *maize streak virus* isolates. *Virology*. 2001;288:247–55.
29. Martin DP, Williamson C, Posada D. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*. 2005;21:260–2.
30. Martin DP, Briddon RW, Varsani A. Recombination patterns in dicot-infecting mastreviruses mirror those found in monocot-infecting mastreviruses. *Arch Virol*. 2011;156:1463–9.
31. Maynard Smith J. Analyzing the mosaic structure of genes. *J Mol Evol*. 1992;4:126–9.
32. Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A. A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family Geminiviridae). *Arch Virol*. 2012;1601–7.
33. Mumtaz H, Kumari SG, Mansoor S, Martin DP, Briddon RW. Analysis of the sequence of a dicot-infecting mastrevirus (family Geminiviridae) originating from Syria. *Virus Genes*. 2011;42(3):422–8.
34. Nahid N, Amin I, Mansoor S, Rybicki EP, van der Walt E, Briddon RW. Two dicot-infecting mastreviruses (family Geminiviridae) occur in Pakistan. *Arch Virol*. 2008;153:1441–51.
35. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 1999;265:218–25.
36. Palmer KE, Rybicki EP. The molecular biology of mastreviruses. In: Maramorosch K, Murphy FA, Shatkin AJ, editors. *Advances in virus research*, vol. 50. San Diego: Academic Press; 1997.
37. Perelegina L, Zhu L, Zurkühlen H, Mills R, Borodovsky M, Hilliard JK. Complete sequence and comparative analysis of the genome of herpes B virus (*Cercopithecine Herpesvirus 1*) from a rhesus monkey. *J Virol*. 2003;77(11):6167–77.
38. Posada D, Crandall KA. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci USA*. 2001;98:13757–62.
39. Ramsell JNE, Boulton MI, Martin DP, Valkonen JPT, Kvamheden A. Studies on the host range of the barley strain of Wheat dwarf virus using an agroinfectious viral clone. *Plant Pathol*. 2009;58:1161–9.
40. Salminen MO, Carr JK, Burke DS, McCutchan FE. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by BOOTSCANning. *AIDS Res Hum Retrovir*. 1995;11:1423–5.
41. Thomas J, Parry J, Schwinghamer M, Dann E. Two novel mastreviruses from chickpea (*Cicer arietinum*) in Australia. *Arch Virol*. 2010;155:1777–88.
42. Varsani A, Shepherd DN, Monjane AL, Owor BE, Erdmann JB, Rybicki EP, Peterschmitt M, Briddon RW, Markham PG, Oluwafemi S, Windram OP, Lefeuvre P, Lett JM, Martin DP. Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *J Gen Virol*. 2008;89:2063–74.