FEBS Letters 583 (2009) 1825–1832

FEBS Letters

journal homepage: www.FEBSLetters.org

Minireview

Evolution of geminiviruses and their satellites

Muhammad Shah Nawaz-ul-Rehman, Claude M. Fauquet *

ILTAB/Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132, USA

ARTICLE INFO

Article history:
Received 2 March 2009
Revised 26 May 2009
Accepted 27 May 2009
Available online 3 June 2009

Edited by Shou-Wei Ding

Keywords:
Virus
Satellite
Evolution
Origin
Diversity
Geminivirus

ABSTRACT

Geminiviruses and their satellites have circular single stranded DNA genomes, infecting many crops and weeds across the globe. To successfully invade new hosts, break host resistance, move virus particles within and between plants, geminiviruses and their satellites have evolved a coordinated network of protein interactions, showing a possible evolutionary path. Humans have played an important role in the last century to promote the emergence of many geminivirus diseases, thereby impacting their evolution. The greatest molecular diversity of geminiviruses and their satellites resides in Southeast Asia revealing a possible center of origin. This minireview leads us to a possible general grand scheme of their evolution.

1. Introduction

Geminivirus genomes are either monopartite (one circular single stranded (ss)DNA component also called as DNA-A) or bipartite (two circular ssDNA components of equal sizes-DNA-A and DNA-B), but in many instances they are accompanied by circular ssDNA alphasatellites and betasatellites. Therefore, we will consider in this minireview the whole "geminivirus" complex extended to these satellites, as the evolution of geminiviruses is intimately linked and dependent upon the existence and evolution of their satellites. It is probable that domestication of crops promoted the spread of geminivirus diseases from reservoir plants, but it is a reality that extensive cultivation of crop plants, promoted the explosion of many "geminivirus" diseases in the last century. On the functional and biological levels, it is increasingly clear that the level of inter-dependence between the viruses and their satellites is more complex than initially thought and it demonstrates the close relationships built over time of these different molecules and thereby their impact on their co-evolution [1]. We will firstly review the genetic elements of "geminiviruses" that are playing together to respond to natural changes over time. We will then review the environmental changes impacting "geminiviruses" and finally we will consider their diversity in the world. Adding all this information together, it is possible to offer a credible scenario for "geminivirus" evolution, from plasmids to monopartite geminiviruses infecting monocots and dicots, with satellites in the Old World (OW), to strict bipartite geminiviruses infecting dicots without satellites in the New World (NW).

2. The "geminivirus" actors

2.1. Geminiviruses

Viruses belonging to the family *Geminiviridae* are circular ssDNA molecules encapsidated in a twinned icosahedral capsid. Geminiviruses have been classified into four different genera: *Mastrevirus*, *Topocuvirus*, *Curtovirus* and *Begomovirus* based on their genome organization, hosts, and insect vectors [2].

Mastreviruses (type species *Maize streak* virus) are transmitted by leafhoppers in a non-propagative, persistent and circular manner and have a single genome component (monopartite) of 2.7 kb. Members of the genus *Mastrevirus* have been found only in the OW (Eastern Hemisphere, Europe, Africa, Asia and Australia) where they largely infect monocots such as maize, sugarcane and panicum, though there are some members that also infect dicots [2]. The genome of mastreviruses encodes four different genes. The open reading frames (ORFs) in the virion strand, V1 and V2, encode the coat protein (CP) and the movement protein (MP), respectively. The complementary sense strand encodes the replication associated C1 (Rep-A) and C2 (Rep-B) genes. The genes on sense and complementary sense are separated by a long intergenic region (LIR) and a short intergenic region (SIR) (Fig. 1).

* Corresponding author. Fax: +1 314 587 1956.
E-mail address: iltab@danforthcenter.org (C.M. Fauquet).

0014-5793/$36.00
doi:10.1016/j.febslet.2009.05.045

The genus *Topocuvirus* (type species *Tomato pseudo curly top virus*) contains a single monopartite virus nearly 3 kb in size, transmitted by treehoppers. Topocuviruses have been found only in the NW and appear to result from a recombination between mastreviruses and begomoviruses [3] (Fig. 1).

Curtoviruses (type species *Beat curly top virus*) are monopartite leafhopper transmitted geminiviruses. Curtoviruses have a genome size of nearly 3 kb and mainly infect dicot plants [2].

Begomoviruses (type species *Bean golden mosaic virus*) are whitefly transmitted (*Bemisia tabaci*) and only infect dicots. The genus *Begomovirus* includes monopartite (DNA-A genome of 2.8 kb, with six ORFs), bipartite (a DNA-B, equal in size to DNA-A, is also present) [2] (Fig. 1).

The genomes of topocuviruses, curtoviruses, and begomoviruses have similar features and genome organization. Based on their functions, their proteins have been named: pre-coat protein (V2/AV2), coat protein (CP, V1/AV1), replication enhancer protein (REn, C3/AC3), transcription activation protein (TrAP, C2/AC2), replication-associated protein (Rep, C1/AC1) and uncharacterized C4 or AC4 protein [2]. DNA-B components (only found in bipartite begomoviruses) encode only two proteins: movement protein (MP, BC1) and nuclear shuttle protein (NSP, BV1) (Fig. 1).

Based on the genome organization, the genetic diversity and the geographical distribution, begomoviruses have been divided into two groups: OW (Europe, Africa, Asia and Australia) and NW (America) begomoviruses (Figs. 1 and 2). All geminiviruses share a conserved nona-nucleotide (TAATATT/AC) where initiation of replication takes place [4]. Both DNA-A and DNA-B components of bipartite begomoviruses are essential for successful systemic infection where replication and transcription functions are provided by the DNA-A and movement functions are provided by the DNA-B. In some few cases among OW bipartite begomoviruses, DNA-A alone is sufficient for systemic infection and movement while the DNA-A component of NW begomoviruses are strictly dependent on the DNA-B component [2,3].

2.2. Alphasatellites

Begomovirus-associated alphasatellites (previously called DNA-1) are self-replicating satellite-like molecules, dependent on the helper virus for movement, encapsidation and vector transmission. There is no known specific function attributed to alphasatellites. They have been shown to be associated with monopartite OW begomovirus diseases such as cotton leaf curl disease in Pakistan and India, tomato yellow leaf curl disease in China and ageratum yellow vein disease from South East Asia [5]. Alphasatellites have a highly conserved genome organization, encompassing a replication-associated protein of nearly 36 kDa, an adenine-rich region of nearly 200 nts and an origin of replication (Ori) (including a conserved nona-nucleotide TAGTATT/AC), similar to nanoviruses (*Nanoviridae*; another family of circular ssDNA viruses) [5].

2.3. Betasatellites

Betasatellites are pathogenicity determinant molecules associated with various plant diseases exclusively caused by monopartite begomoviruses in the OW [1]. Betasatellites are completely dependent on the helper component (DNA-A) for their replication, encapsidation and transmission by whitefly vectors. Unlike alphasatellites, betasatellites are clearly associated with their specific helper component irrespective of hosts and geographical distribution [6]. In their composition, betasatellites, encode a single gene, BetaC1 (13 kDa protein), in the complementary strand of their genome, and contain an adenine-rich region of nearly 240 nts as well as a satellite-conserved region (SCR) of nearly 220 nts, which is highly conserved among all betasatellites known today [6] (Fig. 1).

Betasatellites share no sequence homology with the helper components except a similar nona-nucleotide (TAATATT/AC). Although some level of specificity exists for trans-replication of betasatellites, they generally have a loose specificity of trans-replication by different helper components. For example, Cotton leaf curl Multan virus (CLCuMuV) cannot trans-replicate Ageratum yellow vein betasatellite (AYVB), whereas Ageratum yellow vein virus (AYVV) can trans-replicate Cotton leaf curl Multan betasatellite (CLCuMuB) [7]. On the other hand, Tomato leaf curl betasatellite (ToLCB, a defective form of betasatellite, which lacks any ORF in its genome) has been reported to be trans-replicated by the African cassava mosaic virus (ACMV) and Beet curly top virus (BCTV), which are highly diverse in terms of their geographical distribution

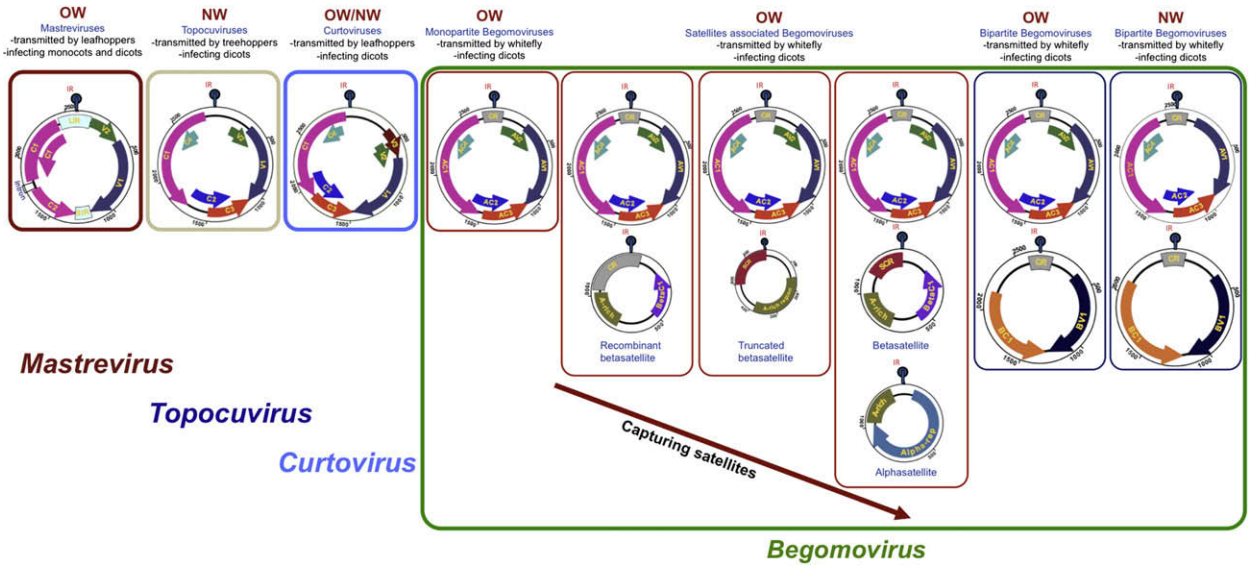

Fig. 1. Genome organization of geminiviruses. The family is composed of four genera: *Mastrevirus*, *Topocuvirus*, *Curtovirus* and *Begomovirus*. Mastreviruses represent the simplest genome with only four ORFs (V1, V2, C1, C2) coded by a single molecule. The viruses of the genera, *Topocuvirus*, *Curtovirus* have 6–7 ORFs (V1, V2, V3, C1, C2, C3, C4) coded by a single molecule. While begomoviruses have six or eight ORFs (AV1, AV2, AC1, AC2, AC3, AC4, BV1 and BC1) coded by one or two molecules. Begomoviruses are associated with satellite molecules such as alphasatellites, betasatellites and their recombinants or defectives molecules, coding respectively for one ORF (Alpha-Rep, BetaC1). LIR; Large intercistronic region, SIR; Small intercistronic region, IR; intercistronic region; CR; Common region, SCR; satellite common region.

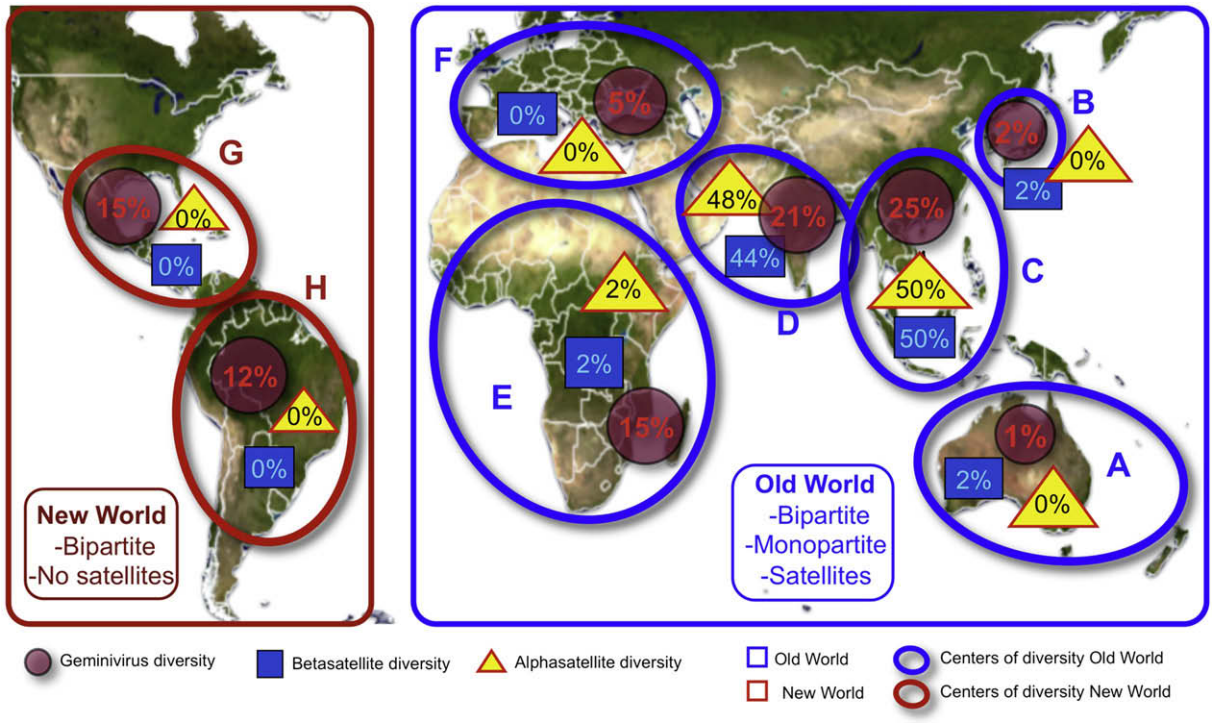

Fig. 2. Geographical distribution of geminiviruses. Based on the percentage of species for each type of molecule, eight different centers of diversification have been identified (A–H). The percentages in circles represent the percentage of geminivirus species, percentages in boxes represent the percentage of betasatellite species, and percentages in triangles represent the percentage of alphasatellite species. To differentiate the types of viruses and satellites, boxes and circles have been highlighted with different colors. A – Australia, B – Japan, C – China, D – Indian subcontinent, E – Africa, F – Mediterranean-European region, G – Central America, H – South America.

and molecular diversity, with BCTV belonging to *Curtovirus*, a different genus [8].

3. Common features between geminivirus components and satellites

The unique common features for all geminiviruses and their satellites is the rolling circle replication (RCR) mechanism and the presence of a stem loop with the nona-nucleotide origin of replication (TAATATT/AC or TAGTATT/AC for alphasatellites). RCR is a basic replication mechanism, for the replication of bacterial and archaeal plasmids as well as bacterial, human, animal and plant viruses, which depends on a replication-associated protein, possessing nicking and ligation functions of double stranded (ds)DNA replicative forms [9]. The geminivirus encoded Rep protein has been shown to perform the same functions as prokaryotic plasmid replication-associated proteins during the RCR [10].

Although geminiviruses and their satellites are highly diverse in nucleotide sequence of their genomes and ORFs [11,12], the genetic architecture, localization, length of individual genes at specific locus and the function of individual proteins are highly conserved among the members of each genus. For example, members of the genus *Mastrevirus* have four ORFs in their genomes, LIR and SIR and a Rep protein interrupted by an intron, irrespective of the species considered, monocot or dicot hosts and geographical distribution [2].

The TrAP proteins of the OW Tomato yellow leaf curl China virus (TYLCCNV-C2), African cassava mosaic virus (ACMV-AC2), Mungbean yellow mosaic virus (MYMV-AC2), and the NW Tomato golden mosaic virus, (TGMV-AC2) have all been associated with a transcription activator function of late viral genes [13–17]. Although small differences are present in their mode of action and regulation of genes, all TrAPs of these diverse viruses share the trans-activation of viral genes and nuclear localization. Similar conclusions can be drawn for all geminivirus and satellite genes.

4. Variable features of geminivirus components

All the NW bipartite viruses lack the AV2 ORF in their DNA-A component. Moreover, NW begomoviruses have a N-terminal nuclear localizing signal (NLS) PWRsMaGT in their CP, which is absent in OW begomoviruses [18]. The strong link between the absence of the AV2 ORF and NW begomoviruses suggest that all the NW begomoviruses evolved from a single parent after the Gondwana continental separation. The fact that OW and NW begomoviruses have a very similar genome organization, with the exception of AV2 ORF, shows that all begomoviruses have a common origin irrespective of their geographical distribution.

5. Genetic effects on “geminivirus” evolution

As for all viruses, the evolution of geminiviruses is subject to point mutations and recombination. The impact of point mutations has been studied for three different geminiviruses, Tomato yellow leaf curl China virus, East African cassava mosaic virus, and Maize streak virus and showed to be similar as for RNA viruses ($\sim 10^{-4}$ substitutions per site per year) [19–21]. In addition it was concluded that the most commonly observed substitution mutation is the transition mutation type rather than deletion or insertion mutations. The study demonstrates that the mutation rate of geminiviruses is dependent on the type of virus, host plant, age of the inoculated plant and inoculum homogeneity. More importantly, it is believed that recombination among different DNA-A components is the main source of molecular variation among geminiviruses [22] and can result in gain of virulence for the helpers and acquisition of new satellite molecules. A family recombination

analysis for a wide variety of geminivirus DNA-A components suggests that the CR is a hotspot of recombination, irrespective of their geographical origin, while the coding regions of geminiviruses are less vulnerable to recombination [22]. Furthermore, the CR of DNA-A component is not only exchanged between members of different geminivirus species, but is also exchanged with heterologous molecules such as betasatellites or DNA-B component [23,24]. In nature, exchange of CRs has been recorded for the Potato yellow mosaic virus (PYMV) and Potato yellow mosaic Panama virus (PYMPV) [24], where a simultaneous acquisition of a new CR occurred for both components. In another example, the DNA-A CR of Sri Lankan cassava mosaic virus (SLCMV) recombined with the DNA-B CR of Indian cassava mosaic virus (ICMV) [25]. Experimental evidence demonstrated that the recombination between DNA-A and DNA-B results in gain of virulence due to the increase of replication and systemic movement of DNA-A molecules [26].

In recent years, several begomoviruses showed a re-assortment of DNA-B components with DNA-A molecules of different diseases. This re-assortment has not only resulted in the gain of virulence, but also shifted the host range for the newly assorted combination. Melon chlorotic leaf curl virus (MCLCuV) can infect species within the families Cucurbitaceae, Fabaceae, and Solanaceae in the NW. Bean calico mosaic virus (BCaMV) is a bean-restricted virus and does not infect cucurbits in laboratory conditions. Interestingly, interaction of either DNA-A or DNA-B components of BCaMV with MCLCuV resulted in infection to many cucurbit species [27].

Recombination of DNA-A CR is not only limited to DNA-B component or other DNA-A molecules, but is also subject to recombination with satellites. CR recombinant betasatellites have been characterized for AYVV and TYLCCNV [23,28]. Similarly DNA-A recombinant CLCuMuB have also been found in non-cultivated cotton species in Pakistan (Nawaz-ul-Rehman et al., Unpublished). Such recombinant betasatellites lack the normal SCR but encode a functional BetaC1 gene responsible for pathogenicity. ToLCV-associated satellite from Australia lacks BetaC1 gene but comprise an A-rich region and SCR. Systemic infection in this case does not require the BetaC1 gene [1,8]. The presence of molecules with a BetaC1 deleted, or an SCR betasatellite replaced by a geminivirus CR, may represent evolutionary links between geminiviruses and betasatellites. Such recombinations might have resulted in the evolution of the most virulent betasatellites comprising an SCR, an A-rich region and a BetaC1 gene (Fig. 1).

Although DNA-B components associated with NW begomoviruses are crucial for systemic infection, the DNA-B component among some OW begomoviruses is not essential and is exchangeable between geminiviruses, and with betasatellites, resulting in severe infection. For instance, the AYVB can be trans-replicated by SLCMV DNA-A alone and shows systemic and successful infection on ageratum plants [25]. Presence of an alternative movement function provided by betasatellites and exchange of CR with different helper components (DNA-A) therefore re-enforce the concept to consider DNA-B as a captured satellite-like molecule.

6. "Geminivirus" movement, a conserved and coordinated network of viral DNA transport system

In order to be replicated inside the nucleus, to spread from one cell to another cell and to disseminate from one plant to another plant, geminiviruses and their satellites have evolved a coordinated network of different movement associated proteins. Depending upon the type of virus, V1, V2 or AV1, C4, BV1 (NSP), BC1 (MP) and BetaC1 proteins have been reported for geminivirus movement functions [29,30]. A movement model for bipartite begomoviruses has been proposed for MP and NSP of Abutilon mosaic virus (AbMV), Cabbage leaf curl virus, Bean dwarf mosaic virus (BDMV) and Squash leaf curl virus (SLCV), which relies on the coordinated functions of NSP and MP. This model indicates that NSP shuttles the ssDNA from the nucleus to the plasma membrane at the periphery of the cell from where MP takes over and transfers the viral DNA to adjacent cells through plasmodesmata [29–34]. The complex of viral DNA movement in begomoviruses has the ability to bind ss-DNA or dsDNA and can transport the viral DNA to neighboring cells [32,33]. Additionally, BC1 and BV1 have limitations to bind with specific size and form of viral DNA [32]. Although, BV1-BC1 interaction is vital for virus movement, for some bipartite viruses such as SLCV, in the presence of defective or mutated BV1 gene, the CP can provide an alternative function to BV1 gene of DNA-B components [35]. The CP of monopartite begomoviruses such as TYLCV, are nuclear-targeted proteins and perform viral DNA movement functions in combination with C4 protein [33]. Thus for monopartite begomoviruses CP and C4 may act similar to BV1–BC1 interaction for virus movement.

The MSV encoded MP(AV2) does not directly bind to ssDNA but rather interacts with CP for viral DNA movement function [36]. The CP in this case localizes in the nucleus just like the NSP of begomoviruses. These examples illustrates that geminiviruses movement is mainly performed in coordinated fashion, which relies on the interaction of a nuclear targeting and cell membrane trafficking proteins.

Recently it has also been shown that the BetaC1 protein encoded by betasatellites can be an alternative to DNA-B component [29] although more experimental evidence is needed to determine whether a coordinated action of CP and BetaC1 results in the movement of viral DNA, or whether BetaC1 alone can perform virus movement functions. Nonetheless the BetaC1, like any other geminivirus DNA-binding MP, can bind both to dsDNA and ssDNA in a sequence-independent manner and localizes in the nucleus [14].

7. RNAi and "geminiviruses": evidence of differential evolution of "geminiviruses encoded gene silencing suppressors

The universal mechanism of RNA degradation by a specific mechanism called RNA interference (RNAi), or post-transcriptional gene silencing (PTGS) is also present in plants. RNAi is used as a defense mechanism against plant viruses. The key mechanism involves the conversion of viral transcripts into dsRNAs followed by the production of short interfering RNAs of varying lengths (21–24 nts) by different RNA cutting enzymes, called dicers [37–39]. Despite the fact that there is a priori no dsRNA molecules in the geminivirus cycle of infection, it has been clearly established that geminiviruses are target of the PTGS [40]. Different models have been presented to explain the possible phenomenon of short RNA production by begomovirus genomes [39]. Evidence for geminivirus gene silencing suppressor activity suggests that at least 4 proteins have this capacity, such as V2 of TYLCV; AC2 of East African cassava mosaic Cameron virus (EACMCV), TGMV, and TYLCCNV; AC4 of ACMV and SLCMV; and satellite encoded BetaC1 from AYVB and Tomato yellow leaf curl China betasatellite (TYLCCNB) [37,40–42]. Surprisingly, the PTGS suppressor function of these proteins is not conserved among all geminiviruses. For example, C2 of TYLCCNV is a suppressor of gene silencing, while C2 of TYLCV, another tomato infecting monopartite, is not [16]. Interestingly, none of these suppressors has been characterized as being as strong as those encoded by RNA viruses, such as P19 [37], which may in part support the choice of a multiprotein strategy adopted by geminiviruses. The existence of a range of molecularly unrelated suppressors of PTGS for geminiviruses illustrates the differential evolution of geminivirus encoded suppressors of gene silencing.

8. Environmental effects on "geminivirus" evolution

Insect vectors responsible for the spread of geminiviruses play an important role in geminivirus evolution. Persistent and circulative mode of virus spread within adult insects and physical interaction between insects and geminiviruses show the co-evolutionary adaptation between geminiviruses and their insect vectors.

Beet leafhoppers (*Circulifer tenellus*) are responsible for curly top disease spread in western areas of the United States during their annual migration period. A close relationship has been established between the seasons, the weeds, the crops and the insects [43–45]. Feeding the viruliferous beet leafhoppers on non-hosts such as corn still maintains the virus but the transmission ability is lost. Similarly, MSV can be detected and maintained in head, hemolymph and gut of the viruliferous leafhoppers up to 28 days after exposure to an infected plant [46].

Studies on the whitefly (*B. tabaci*) transmitted begomovirus, TYLCV, also showed the persistent and circulative mode of transmission [47]. Interestingly, in the unique case of TYLCV some viral propagative activities were found in whitefly, demonstrating a possible closer interaction between insects and geminiviruses [47]. A study done on cassava and cassava mosaic disease (CMD) in India and Africa indicated that there was co-adaptation of the virus and the vector, demonstrated by a higher fitness for transmission of the local viruses [48]. The CP of TYLCV and AbMV have been shown to directly bind with an homolog of GroEL protein, produced by an endosymbiotic bacteria in the whitefly hemolymph. Geminivirus evolution through their natural vector occurred only with the CP gene, which interacts with the insects for virus transmission [49].

Different surveys conducted for whitefly population on cassava plants in Africa revealed that the CMD pandemic is correlated with high population densities and high fecundity of B-biotype whitefly adapted to cassava [50]. CMD dissemination occurs not only through whitefly, but also through the spread of infected plant material by human migrations due to instability and wars, resulting in the CMD pandemic in Africa [50].

Studies conducted on the interaction of an infected host plant and invasive B-biotype whitefly indicated that there is a strong link between whitefly proliferation and virus infection on the host plants. Comparison of introduced and invasive B-biotype whitefly in China with indigenous non-invasive ZHJ1 whitefly revealed that B-biotype whitefly population density increases many fold compared to indigenous whitefly population when fed on TYLCCNV or Tobacco curly shoot virus (TbCSV) infected plants [51].

These examples highlight the biological fitness of virus-vector interactions, which by mutual interaction can benefit the invasive virus vectors over the non-invasive vectors when they share the same biological niche [51].

More evidences of geminivirus and insect vector co-evolution have been demonstrated by the artificial exchange of CPs between the members of two different genera. Exchange of the whitefly transmitted ACMV–CP with leafhopper transmitted BCTV–CP resulted in the successful transmission of the chimeric ACMV by leafhoppers [49], demonstrating a close adaptation and co-evolution between vectors and viruses.

9. Human impact on "geminivirus" diseases

Humans play a major role in geminivirus disease spread either by transferring crops from their centers of origin to other ecological areas, by introducing new germplasm in different environments, by spreading viruses in the form of infected plants or by promoting the explosion of viruliferous insect vectors. The global spread of begomoviruses and invasion of local viruses on introduced crops such as tomato, cotton or cassava suggests that the current epidemics were a result of the sudden encounter of existing begomoviruses with newly introduced hosts or vectors rather than a rapid evolution of geminiviruses.

Geminiviruses have been emerging as a worldwide problem since the 1960s, and frequently, epidemics can be linked to human activities. Begomovirus spread is directly correlated with the wide spread of the B-biotype of *B. tabaci* Gennadius. The whitefly B-biotype, by colonizing more than 1000 host plants, has an enhanced capacity to maintain its populations and reproduce in new environmental conditions compared to the Q-biotype [52]. During the late 1980s, with the spread of silverleaf whitefly or sweet potato whitefly (*B. tabaci*) in the United States, a large number of crops such as tomato, cotton, cucurbits and beans have been devastated by geminivirus infections [53,54].

The worldwide dissemination of tomatoes exemplifies one of the best examples of human impact on geminivirus emergence and spread. Export of tomato seedlings from Israel to the Caribbean during the early 1990s resulted in the introduction of TYLCV in the NW [55,56]. Within 10 years, TYLCV spread throughout all the tomato growing areas of Central and North America and later in the rest of the world [55–57].

Similarly, introduction of cassava plants from South America (center of origin for cassava) to Africa in the 16th century resulted in several epidemics of CMD. These viruses have not been found in South America, therefore it is unlikely that they were moved along with cassava. The accepted concept is that cassava geminiviruses evolved from unidentified indigenous hosts in Africa and exploded on cassava [58].

Introduction of cotton crops from Mexico to the Indian sub-continent occurred in 1818. From 1818 to 1967, there was no noticeable incidence of CLCuD in the Indian sub-continent [59,60]. In the meantime, due to high yield of NW cotton and an increase in textile industry, areas of cotton also dramatically increased. The first incidence of CLCuD in Pakistan was recorded in 1967, subsequently the disease spread to central Pakistan as observed in 1985 [59,60]. Within 5 years, the disease appeared in epidemic proportions in Pakistan and in 2000 it was found in neighboring countries such as India and China. Today there are seven known species of geminiviruses associated with CLCuD, prevailing mainly in Pakistan and India [11].

As a result of natural selection by human interference, a dramatic adaptation of geminiviruses to new hosts and new ecological zones was observed in the 20th century.

10. World distribution of "geminiviruses" and center of diversification

According to the modern concepts of crop diversification, there are at least eight major centers of plant origin [61]. There is an enormous diversity of particular plant families in the centers of origin, and less diversity on the periphery. Geminivirus centers of diversification can be identified from the phylogenetic tree built from all sequences available [11] and can be described as the areas where a maximum number of geminivirus species have been identified so far, and a similar evaluation can be done for their satellites. We identified eight different geographical locations as centers of diversification for "geminiviruses" (Fig. 2). These centers are: A- Australia, B- Japan, C- South China, D- Indian subcontinent, E- Sub-saharan Africa, F- Mediterranean-European region, G- South America and H- Central America. Together, the Chinese and Indian centers host more than 46% of geminivirus species, 94% of betasatellite species and 98% of alphasatellite species (Fig. 2). The presence of this enormous diversity for geminiviruses and associated satellites points towards designating Indo-China as a center of origin for these viral and subviral entities.

It is also noteworthy that Corchorus yellow vein virus (CoYVV) and Corchorus golden mosaic virus (CoGMV), recently identified in Vietnam and India, have all the features of NW begomoviruses in spite of being the most remote viruses in this cluster [18]. These viruses also lack the AV2 ORF and have the PWRsMaGT motif in their CP gene. This strongly suggests that ancestors of NW-like begomoviruses were present in the OW even before the Gondwana separation.

Surprisingly, the diversity centers for geminiviruses do not match perfectly with those of host crops [61]. There is an overlap between the crop and virus centers for five out of the eight geminivirus centers, with only the Australian, African and South American centers differing. However, considering the tomato example, which originated from South America, only 15 different begomoviruses infect tomato there compared to the African continent, where 20 different representatives of begomovirus species infect tomato. Similarly, cultivated cotton in its native place of origin (Central America) is infected by a single bipartite begomovirus called Cotton leaf crumple virus (CLCrV) [62]. On the other hand, varieties of NW cultivated cotton in the Indian subcontinent are infected by members of at least seven species of begomoviruses [11]. Furthermore, neither cotton nor tomato are infected by the same species of viruses in their native place and their introduced geographical areas, pointing towards the invasion of imported virus-free crops by local geminiviruses. All of this data strongly support the theory that local viruses invaded endemic or introduced crops, suggesting rapid host changes.

11. "Geminivirus" evolution proposal

Because there are no fossils for viruses, it is very difficult to ascertain how old geminiviruses or their ancestors could be? The oldest record for symptoms of geminivirus infection is 1257 years old. A Japanese poem written in 752 AD describes the beauty of eupatorium plants in Japan. The yellow color of this plant is now attributed to a satellite associated begomovirus infection [63].

Because of the RCR mechanism that is common to geminiviruses and their satellites, to circoviruses and anelloviruses infecting animals, to microviruses infecting bacteria and to plasmids replicating in bacteria, archaea and algae, it has been hypothesized that the evolution of geminiviruses originated from these plasmid molecules [3]. Plasmids act as an extra chromosomal molecule in prokaryotes (such as bacteria) and ancient eukaryotes (such as red algae). It is generally admitted that the Rep protein has been the sole common protein throughout this long evolutionary process. This is seconded by the discovery of a very significant homology (>31% over 55 amino acids) between the Rep protein of the mastrevirus Wheat dwarf virus (WDV) and a Rep-like sequence encoded by a red algae (*Porphyra pulchera*) plasmid (Fauquet. CM; personal communication). As reviewed by others, if algae plasmids are considered to be the ancestors of geminiviruses, then geminivirus evolution can be dated back at least to 450 million years. It is unclear how a replicated plasmid in an algae evolved into an independent virus capable of movement between cells, tissues and plant hosts and it is a mystery how this pro-virus acquired de novo genes such as AC2, AC3, AC4, AV1 and AV2 to constitute the genome of today's geminiviruses. Once the efficient movement and insect transmission ability via the CP was obtained, geminivirus spread occurred at an alarming rate. Furthermore, human interference resulted in an unbalanced ecosystem, which resulted in the invasion of geminiviruses onto introduced domesticated crops.

Taking into consideration all the information available about geminiviruses and their satellites, it would be possible to propose that the evolution of geminiviruses followed the path shown in Fig. 3. Replicated plasmids in red algae and other primitive life forms, managed to acquire new genes, allowing this molecule to become more independent from its host and eventually capable of infecting plants, probably monocots first, as a pre-mastrevirus.

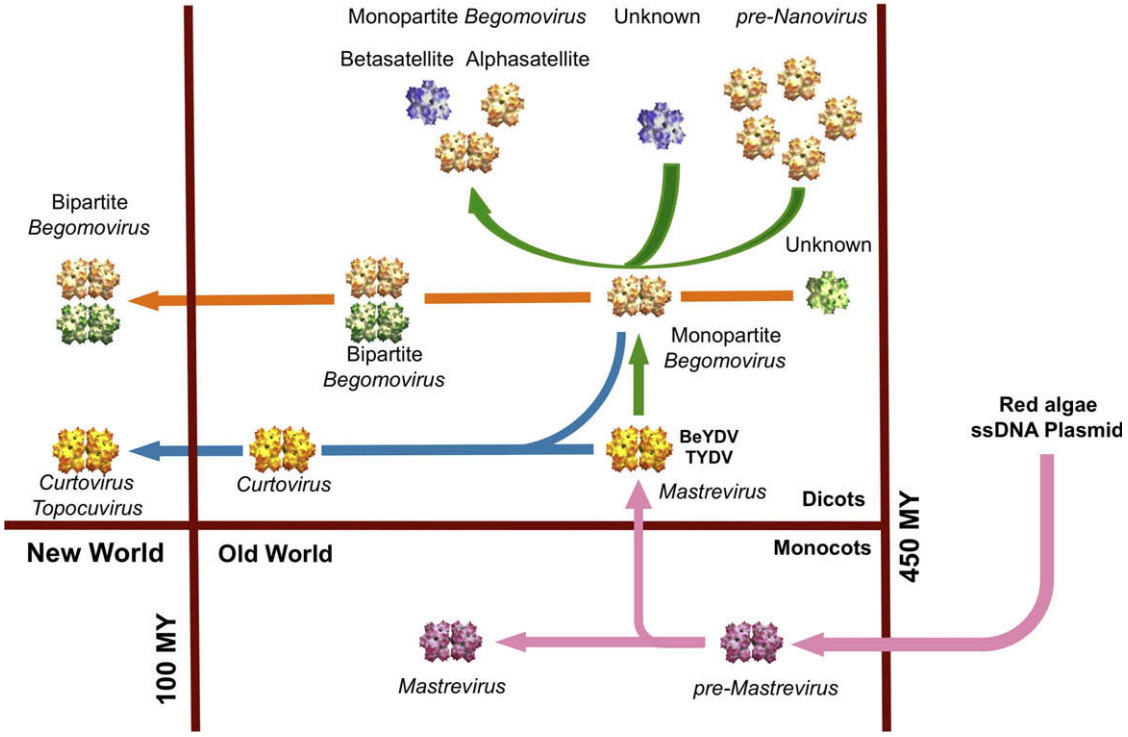

Fig. 3. Possible scenario of geminivirus (DNA-A and DNA-B) and satellite evolution, in space and time. The horizontal and vertical bars represent time, geographical or host domains. One hundred MY represents the estimated time of continental drift between the Old World and the New World.

This evolution must have coincided with acquisition of insect transmission. At some point in time, they managed to infect dicots but still had the same type of leafhopper vector, then acquired new genes to become a pre-monopartite and a monopartite, transmitted by whiteflies. This monopartite begomovirus, having the capacity to capture other molecules, acquired an alphasatellite from a pre-nanovirus, or a betasatellite from an unknown source. In the meantime, hybrids were formed through recombination between a mastrevirus infecting dicots and a monopartite begomovirus to form the ancestors of curtoviruses and topocuviruses. In a subsequent period, a monopartite managed to capture an ancestor of what is today a B-component, and this combination of two components was extremely successful to the point where such bipartite begomoviruses are only present on the American continent, following the drift of the continents around 125 million years ago.

Although such a scenario for the evolution of geminiviruses is generally satisfactory and supported by others [3], there are many questions left unanswered, such as why we have not been able to identify any of the possible precursors of these different phases, why there is only one category of begomoviruses and none of the satellites present on the American continent, and why there are no mastreviruses present on the American continent neither? It is probable that in the near future some of these questions will be answered thanks to the extended capacity in virus sequencing.

12. Conclusion

Evolution of viruses is difficult to explain because there are no ancient records that can be used to trace the biological activity of these molecules, therefore we must rely on existing viruses to rebuild their possible history. Plant geminiviruses are so small that most probably their genome, through mutation and recombination, has been rebuilt numerous times in the past, however their structure and size are remarkably stable, probably due to the accumulation of numerous biochemical, biophysical and biological constraints on very small proteins. Although there are molecular records for only a mere 25 years, there are indications that their remote ancestors existed 450 million years ago.

As time passes, we discover individual cases with intermediate characteristics, such as the NW-like bipartite begomoviruses isolated in Vietnam and India. These viruses possess all the features of NW begomoviruses while being the most remote of them on the NW branch of the phylogenetic trees, clearly indicating that they are ancestors of NW begomoviruses, and not a recent exportation of NW viruses into the OW. We can predict that new viruses isolated from remote places or remote hosts in the Tree of Life will unravel new intermediates in the geminivirus tree.

A fundamental question in virus evolution is: what is the origin of all the viral genes that have no homology to any of the genes known so for in the Tree of Life? What is the mechanism by which viruses acquire new genes? Furthermore, many of these genes are coding for multifunctional proteins, adding another level of complexity in their interaction with host proteins, and their de novo "creation". Geminiviruses and their satellites have at least four gene silencing suppressor proteins that are not related to each other and have complementary functions. It is now obvious that viral proteins are multifunctional with variable level of activities. The PTGS suppressors AC2 and AC4 of begomoviruses can vary from no suppressor activity to very high suppression activity. At least these features provide indications of different evolutionary potentials and re-enforce the concept that viruses are in a constant exploratory mode for survival.

Geminiviruses are instrumental in demonstrating that a virus molecule can capture another related molecule to the point where we consider them a segmented genome. The discovery of alphasatellites gave us the evidence that a satellite can "jump" from a nanovirus to a geminivirus, and the betasatellites gave us the concept that a satellite can turn a non–infectious virus to a super virulent virus. The co-evolution evidence between betasatellites and their helpers clearly demonstrates that two physically independent molecules can conserve, for a very long period of time, a close relationship based on synergistic molecular interactions [64]. Finally, this leads to the possible idea that the DNA-B of bipartite begomoviruses might have been an unrelated satellite until it became, through recombination, a recognized segment of a geminivirus divided genome.

Acknowledgments

This work was supported by the National Academy of Sciences' Pakistan-US Science and Technology Cooperative Program, Contract PGA-7251-05-007 and by the Donald Danforth Plant Science Center.

References

[1] Briddon, R.W. and Stanley, J. (2006) Subviral agents associated with plant single-stranded DNA viruses. Virology 344, 198–210.
[2] Stanley, J., Bisaro, D.M., Briddon, R.W., Brown, J.K., Fauquet, C.M., Harrison, B.D., Rybicki, E.P. and Stenger, D.C. (2005) Family *Geminiviridae*. In Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses, pp. 301–326, Elsevier, London.
[3] Rojas, M.R., Hagen, C., Lucas, W.J. and Gilbertson, R.L. (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu. Rev. Phytopathol. 43, 361–394.
[4] Fontes, E.P.B., Gladfelter, H.J., Schaffer, R.L., Petty, I.T.D. and Hanley-Bowdoin, L. (1994) Geminivirus replication origins have a modular organization. Plant Cell 6, 405–416.
[5] Briddon, R.W. et al. (2004) Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus-DNA [beta] complexes. Virology 324, 462–474.
[6] Briddon, R.W. et al. (2003) Diversity of DNA β, a satellite molecule associated with some monopartite begomoviruses. Virology 312, 106–121.
[7] Saunders, K., Briddon, R.W. and Stanley, J. (2008) Replication promiscuity of DNA β, satellites associated with monopartite begomoviruses; deletion mutagenesis of the Ageratum yellow vein virus DNA-[beta] satellite localizes sequences involved in replication. J. Gen. Virol. 89, 3165–3172.
[8] Dry, I.B., Krake, L.R., Rigden, J.E. and Rezaian, M.A. (1997) A novel subviral agent associated with a geminivirus: The first report of a DNA satellite. Proc. Natl. Acad. Sci. USA 94, 7088–7093.
[9] Ilyina, T.V. and Koonin, E.V. (1992) Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res. 20, 3279–3285.
[10] Jeske, H., Lutgemeier, M. and Preiss, W. (2001) DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. EMBO J. 20, 6158–6167.
[11] Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M. and Zhou, X. (2008) Geminivirus strain demarcation and nomenclature. Arch. Virol. 153, 783–821.
[12] Briddon, R., Brown, J., Moriones, E., Stanley, J., Zerbini, M., Zhou, X. and Fauquet, C. (2008) Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. Arch. Virol. 153, 763–781.
[13] Trinks, D., Rajeswaran, R., Shivaprasad, P.V., Akbergenov, R., Oakeley, E.J., Veluthambi, K., Hohn, T. and Pooggin, M.M. (2005) Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J. Virol. 79, 2517–2527.
[14] Cui, X., Li, G., Wang, D., Hu, D. and Zhou, X. (2005) A begomovirus DNAβ-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. J. Virol. 79, 10764–10775.
[15] Sunter, G. and Bisaro, D.M. (2003) Identification of a minimal sequence required for activation of the tomato golden mosaic virus coat protein promoter in protoplasts. Virology 305, 452–462.
[16] Dong, X., van Wezel, R., Stanley, J. and Hong, Y. (2003) Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. J. Virol. 77, 7026–7033.
[17] Schnippenkoetter, W.H., Martin, D.P., Hughes, F.L., Fyvie, M., Willment, J.A., James, D., von Wechmar, M.B. and Rybicki, E.P. (2001) The relative infectivities and genomic characterisation of three distinct mastreviruses from South Africa. Arch. Virol. 146, 1075–1088.
[18] Ha, C., Coombs, S., Revill, P., Harding, R., Vu, M. and Dale, J. (2008) Molecular characterization of begomoviruses and DNA satellites from Vietnam: additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J. Gen. Virol. 89, 312–326.

[19] Duffy, S. and Holmes, E. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J. Gen. Virol. doi:10.1099/vir.0.009266-0.
[20] Ge, L., Zhang, J., Zhou, X. and Li, H. (2007) Genetic structure and population variability of Tomato yellow leaf curl China virus. J. Virol. 81, 5902–5907.
[21] Isnard, M., Granier, M., Frutos, R., Reynaud, B. and Peterschmitt, M. (1998) Quasispecies nature of three Maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars. J. Gen. Virol. 79, 3091–3099.
[22] Padidam, M., Sawyer, S. and Fauquet, C.M. (1999) Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218–225.
[23] Tao, X. and Zhou, X. (2008) Pathogenicity of a naturally occurring recombinant DNA satellite associated with tomato yellow leaf curl China virus. J. Gen. Virol. 89, 306–311.
[24] Urbino, C., Polston, J.E., Patte, C.P. and Caruana, M.L. (2004) Characterization and genetic diversity of Potato yellow mosaic virus from the Caribbean. Arch. Virol. 149, 417–424.
[25] Saunders, K., Salim, N., Mali, V.R., Malathi, V.G., Briddon, R., Markham, P.G. and Stanley, J. (2002) Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. Virology 293, 63–74.
[26] Hou, Y.M. and Gilbertson, R.L. (1996) Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. J. Virol. 70, 5430–5436.
[27] Idris, A.M., Mills-Lujan, K., Martin, K. and Brown, J.K. (2008) Melon chlorotic leaf curl virus: characterization and differential reassortment with closest relatives reveal adaptive virulence in the squash leaf curl virus clade and host shifting by the host-restricted bean calico mosaic virus. J. Virol. 82, 1959–1967.
[28] Saunders, K., Bedford, I.D. and Stanley, J. (2001) Pathogenicity of a natural recombinant associated with ageratum yellow vein disease: implications for geminivirus evolution and disease aetiology. Virology 282, 38–47.
[29] Saeed, M., Zafar, Y., Randles, J.W. and Rezaian, M.A. (2007) A monopartite begomovirus-associated DNA beta satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. J. Gen. Virol. 88, 2881–2889.
[30] Hehnle, S., Wege, C. and Jeske, H. (2004) Interaction of DNA with the movement proteins of geminiviruses revisited. J. Virol. 78, 7698–7706.
[31] Wege, C. and Pohl, D. (2007) Abutilon mosaic virus DNA B component supports mechanical virus transmission, but does not counteract begomoviral phloem limitation in transgenic plants. Virology 365, 173–186.
[32] Gilbertson, R.L., Sudarshana, M., Jiang, H., Rojas, M.R. and Lucas, W.J. (2003) Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: Insights into DNA trafficking. Plant Cell 15, 2578–2591.
[33] Rojas, M.R., Jiang, H., Salati, R., Xoconostle-Czares, B., Sudarshana, M.R., Lucas, W.J. and Gilbertson, R.L. (2001) Functional analysis of proteins involved in movement of the monopartite begomovirus, tomato yellow leaf curl virus. Virology 291, 110–125.
[34] Guevara-González, R.G., Ramos, P.L. and Rivera-Bustamante, R.F. (1999) Complementation of coat protein mutants of pepper huasteco geminivirus in transgenic tobacco plants. Phytopathology 89, 540–545.
[35] Qin, S., Ward, B.M. and Lazarowitz, S.G. (1998) The bipartite geminivirus coat protein aids BR1 function in viral movement by affecting the accumulation of viral single-stranded DNA. J. Virol. 72, 9247–9256.
[36] Liu, H., Boulton, M.I., Thomas, C.L., Prior, D.A., Oparka, K.J. and Davies, J.W. (1999) Maize streak virus coat protein is karyophyllic and facilitates nuclear transport of viral DNA. Mol. Plant-Microbe Interact. 12, 894–900.
[37] Voinnet, O. (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. Trends Plant Sci. 13, 317–328.
[38] Ding, S.-W. and Voinnet, O. (2007) Antiviral immunity directed by small RNAs. Cell 130, 413–426.
[39] Vanitharani, R., Chellappan, P. and Fauquet, C.M. (2004) RNAi and gene silencing phenomenon mediated by viral suppressors in plants in: RNA Interference: From Basic Science to Drug Development (Appasani, K., Ed.), Cambridge University Press, Cambridge. p. 544.
[40] Chellappan, P., Vanitharani, R. and Fauquet, C.M. (2004) Short Interfering RNA accumulation correlates with host recovery in DNA virus-infected hosts, and gene silencing targets specific viral sequences. J. Virol. 78, 7465–7477.
[41] Zrachya, A., Glick, E., Levy, Y., Arazi, T., Citovsky, V. and Gafni, Y. (2007) Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology 358, 159–165.
[42] Vanitharani, R., Chellappan, P., Pita, J.S. and Fauquet, C.M. (2004) Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. J. Virol. 78, 9487–9498.
[43] Soto, M.J. and Gilbertson, R.L. (2003) Distribution and rate of movement of the curtovirus Beet mild curly top virus (Family *Geminiviridae*) in the beet leafhopper. Phytopathology 93, 478–484.
[44] Creamer, R., Luque-Williams, M. and Howo, M. (1996) Epidemiology and incidence of beet curly top geminivirus in naturally infected weed hosts. Plant Dis. 80, 533–535.
[45] Bennett, C.W. (1971) The Curly Top Disease of Sugar Beet and Other Plants, The American Phytopathological Society, St. Paul, MN.
[46] Lett, J.M., Granier, M., Hippolyte, I., Grondin, M., Royer, M., Blanc, S., Reynaud, B. and Peterschmitt, M. (2002) Spatial and temporal distribution of geminiviruses in leafhoppers of the genus *Cicadulina* monitored by conventional and quantitative polymerase chain reaction. Phytopathology 92, 65–74.
[47] Czosnek, H. and Ghanim, M. (2002) The circulative pathway of begomoviruses in the whitefly vector *Bemisia tabaci* – insights from studies with Tomato yellow leaf curl virus. Ann. Appl. Biol. 140, 215–231.
[48] Maruthi, M.N., Colvin, J., Seal, S., Gibson, G. and Cooper, J. (2002) Co-adaptation between cassava mosaic geminiviruses and their local vector populations. Virus Res. 86, 71–85.
[49] Briddon, R.W., Pinner, M.S., Stanley, J. and Markham, P.G. (1990) Geminivirus coat protein gene replacement alters insect specificity. Virology 177, 85–94.
[50] Legg, J.P. and Fauquet, C.M. (2004) Cassava mosaic geminiviruses in Africa. Plant Mol. Biol. 56, 585–599.
[51] Jiu, M., Zhou, X.-P., Tong, L., Xu, J., Yang, X., Wan, F.-H. and Liu, S.-S. (2007) Vector-Virus mutualism accelerates population increase of an invasive whitefly. PLoS ONE 2, e182.
[52] Dalton, R. (2006) Whitefly infestations: the Christmas invasion. Nature 443, 898–900.
[53] Henneberry, T.J., Jones, W.A., Perring, T.M. and Faust, R.M. (2000) Silverleaf whitefly: National research, action, and technology transfer plan (Formerly Sweetpotato whitefly, Strain B). Third Annual Review of the Second 5-Year Plan, US Dept. Agric., Agric. Res., p. 209.
[54] Perring, T.M., Cooper, A.D., Rodriguez, R.J., Farrar, C.A. and Bellows Jr., T.S. (1993) Identification of a whitefly species by genomic and behavioral studies. Science 259, 74–77.
[55] Duffy, S. and Holmes, E.C. (2007) Multiple introductions of the Old World begomovirus tomato yellow leaf curl virus into the New World. Appl. Environ. Microbiol. 73, 7114–7117.
[56] Polston, J., Bois, D., Serra, C. and Concepcion, S. (1994) First report of a tomato yellow leaf curl-like geminivirus in the Western Hemisphere. Plant Dis. 78, 831.
[57] Ueda, S., Takeuchi, S., Okabayashi, M., Hanada, K., Tomimura, K. and Iwanami, T. (2005) Evidence of a new Tomato yellow leaf curl virus in Japan and its detection using PCR. J. Gen. Plant Pathol. 71, 319–325.
[58] Ndunguru, J., Legg, J.P., Aveling, T.A.S., Thompson, G. and Fauquet, C.M. (2005) Molecular biodiversity and evolution of cassava geminiviruses in Africa Tanzania a center of origin and diversification. Virology J. 2, 21.
[59] Hafeezur, R., Khan, W.S., Khan, M.U.D. and Shah, M.K.N. (2001) Stability of cotton cultivars under leaf curl virus epidemic in Pakistan. Field Crops Res. 69, 251–257.
[60] Briddon, R.W. and Markham, P.G. (2000) Cotton leaf curl virus disease. Virus Res. 71, 151–159.
[61] Harlan, J.R. (1971) Agricultural origins: centres and non centres. Science 174, 468–474.
[62] Idris, A.M. and Brown, J.K. (2004) *Cotton leaf crumple virus* is a distinct western hemisphere begomovirus species with complex evolutionary relationships indicative of recombination and reassortment. Phytopathology 94, 1068–1074.
[63] Saunders, K., Bedford, I.D., Yahara, T. and Stanley, J. (2003) Aetiology: the earliest recorded plant virus disease. Nature 422, 831.
[64] Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z. and Fauquet, C.M. (2003) Characterization of DNAβ, associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. J. Gen. Virol. 84, 237–247.