Review

Cassava mosaic geminiviruses: actual knowledge and perspectives

BASAVAPRABHU L. PATIL AND CLAUDE M. FAUQUET*

International Laboratory for Tropical Agricultural Biotechnology (ILTAB), Danforth Plant Science Center, 975 N. Warson Rd., St. Louis, MO 63132, USA

SUMMARY

Cassava mosaic disease (CMD) caused by cassava mosaic geminiviruses (CMGs) is one of the most devastating crop diseases and a major constraint for cassava cultivation. CMD has been reported only from the African continent and Indian subcontinent despite the large-scale cultivation of cassava in Latin America and several South-East Asian countries. Seven CMG species have been reported from Africa and two from the Indian subcontinent and, in addition, several strains have been recognized. Recombination and pseudo-recombination between CMGs give rise not only to different strains, but also to members of novel virus species with increased virulence and a new source of biodiversity, causing severe disease epidemics. CMGs are known to trigger gene silencing in plants and, in order to counteract this natural host defence, geminiviruses have evolved suppressor proteins. Temperature and other environmental factors can affect silencing and suppression, and thus modulate the symptoms. In the case of mixed infections of two or more CMGs, there is a possibility for a synergistic interaction as a result of the presence of differential and combinatorial suppressor proteins. In this article, we provide the status of recent research findings with regard to the CMD complex, present the molecular biology knowledge of CMGs with reference to other geminiviruses, and highlight the mechanisms by which CMGs have exploited nature to their advantage.

CASSAVA AND CASSAVA MOSAIC DISEASE

Cassava (*Manihot esculenta* Crantz, Family Euphorbiaceae), which originates from Latin America, is a major source of food for more than 700 million people in tropical developing countries and is cultivated in a total global area of 18.6 million hectares with a total production of 238 million tonnes, half of which is contributed

by the African continent (FAOSTAT, 2008). The status of cassava cultivation today is changing from subsistence farming to an industrialized system designed to process cassava into a diverse spectrum of products, including starch, sago grains, flour, chips, animal feed and, potentially, biofuel, all derived from a crop that has the ability to grow in poor soils (Thresh, 2006). Unfortunately, cassava is affected by several pests and diseases and is vulnerable to at least 20 different viral diseases as a result of its vegetative propagation. Cassava mosaic disease (CMD), caused by cassava mosaic geminiviruses (CMGs) (Family *Geminiviridae*: Genus *Begomovirus*), transmitted by the whitefly *Bemisia tabaci*, is one of the most important viral diseases and a major constraint for cassava production in Africa and the Indian subcontinent (Legg and Fauquet, 2004; Legg *et al.*, 2006).

The earliest report of CMD was from Tanzania by Warburg (1894) and it was first proposed to be a viral disease by Zimmermann (1906). This virus was initially called cassava latent virus (Bock *et al.*, 1981) until Stanley and Gay (1983) first published its sequence and Koch's postulate was satisfied (Bock and Woods, 1983), when it was renamed the African cassava mosaic virus (ACMV). Recent reports suggest that the CMD pandemic has affected at least nine countries in East and Central Africa, covering an area of 2.6 million square kilometres and causing an estimated annual economic loss of US\$1.9–2.7 billion, and it has been termed the most damaging plant virus disease in the world, causing famine and the death of thousands of people (Legg *et al.*, 2006). Although most of the molecular studies concerning CMGs have focused mainly on ACMV, it would be difficult and illogical to restrict this review only to this virus species, because it is now known that members of several different geminivirus species cause CMD and are replacing ACMV in nature. This article summarizes the current understanding, highlights recent research findings and discusses the new concepts concerning CMGs from both Africa and the Indian subcontinent.

TAXONOMY AND NOMENCLATURE

Recently, taxonomic guidelines have been developed to provide a framework on which to base the definitions of species and

*Correspondence: E-mail: iltab@danforthcenter.org
Website: <http://www.danforthcenter.org/iltab>

Fig. 1 Geographical distribution of cassava-infecting geminiviruses. Representatives of all species and strains of cassava mosaic geminiviruses (CMGs) in Africa and the Indian subcontinent. The green colour indicates the area under cassava cultivation and each different coloured dot represents a unique species. The virus name abbreviations are given in the text.

strains. Based on this new approach, in which the sequence identity demarcation between members of different species has been set at 89% for the DNA-A component of begomoviruses, seven African and two Indian CMG species have been recognized (Fauquet *et al.*, 2008; Table S1, see Supporting Information). The corresponding virus names are African cassava mosaic virus (ACMV), East African cassava mosaic virus (EACMV), East African cassava mosaic Cameroon virus (EACMCV), East African cassava mosaic Kenya virus (EACMKV), East African cassava mosaic Malawi virus (EACMMV), East African cassava mosaic Zanzibar virus (EACMZV) and South African cassava mosaic virus (SACMV), reported from Africa, and Indian cassava mosaic virus (ICMV) and Sri Lankan cassava mosaic virus (SLCMV), identified from the Indian subcontinent. In addition to these species, several strains have been recognized: East African cassava mosaic virus-Uganda (EACMV-UG), -Kenya (EACMV-KE), -Tanzania (EACMV-TZ); South African cassava mosaic virus-South Africa (SACMV-ZA), -Madagascar (SACMV-MG); East African cassava mosaic Cameroon virus-Cameroon (EACMCV-CM), -Tanzania (EACMCV-TZ); Indian cassava mosaic virus-India (ICMV-IN), -Kerala (ICMV-Ker); Sri Lankan cassava mosaic virus-India (SLCMV-IN), -Sri Lanka (SLCMV-LK) (Fig. 1; Table S1).

GEOGRAPHICAL DISTRIBUTION AND EXPANSION

Although the cassava geminivirus pandemic continues to expand, CMD has been reported only from the African continent and Indian subcontinent despite the large-scale cultivation of cassava in Latin America and many South-East Asian countries (Fig. 1; Fargette *et al.*, 2006). The absence of CMD in South America and several countries in South-East Asia, despite the occurrence of other geminivirus diseases, has been mainly attributed to the inability of the polyphagous *B. tabaci* B biotype to colonize cassava effectively in this part of the world (Carabali *et al.*, 2005).

Cassava mosaic geminiviruses in Africa

The earliest reports on the increasing prevalence and geographical expansion of the disease in both East and West African countries appeared in the 1920s and 1930s (Legg and Thresh, 2000). In the early 1990s, a new epidemic of severe CMD occurred in Uganda, which progressed steadily from the north-east towards the south-west, and the progress of the epidemic

'front' was mapped during regular monitoring surveys in 1992 (Legg and Thresh, 2000). The subsequent discovery of the association of a novel recombinant virus, the EACMV Uganda strain (EACMV-UG), in severely affected plants from the epidemic area boosted research interest in CMD (Deng *et al.*, 1997; Zhou *et al.*, 1997). This was followed by an extensive series of surveys in different parts of the African continent during which the co-occurrence of ACMV and EACMV was discovered in western Kenya and north-western Tanzania (Ariyo *et al.*, 2005). From these surveys, many virus isolates were sequenced and new species related to EACMV (EACMV-like viruses) were identified from East and West African countries. The sequence analyses of both genomic components of these viruses revealed the presence of recombined fragments, a feature commonly found in begomoviruses (Padidam *et al.*, 1999b).

The initial surveys conducted during the early period of the epidemic in central-southern Uganda showed that ACMV occurred more frequently than EACMV-UG, although there were some incidences of mixed infections (Harrison *et al.*, 1997). However, subsequent surveys revealed more frequent occurrences of EACMV-UG than ACMV and a significantly decreased proportion of mixed infections (Legg *et al.*, 2006). Such a drastic change has been attributed to the selection of planting material by the farmers, in which they avoided severely diseased plants that were usually dually infected. It has also been hypothesized, but not proven, that EACMV-UG DNA-A, which contains a fragment of the ACMV coat protein (CP), might be more efficiently transmitted by whitefly vectors, thus leading to its widespread distribution in Africa (Pita *et al.*, 2001b).

From the work of Ndunguru *et al.* (2005) it is evident that EACMVs may be indigenous to Tanzania, thus making East Africa the cradle of origin for many cassava viruses, and, in particular, the Eastern Arc and Coastal Forests of Tanzania/Kenya as the main hot spots for CMD begomovirus biodiversity in Africa. Thus, it is possible that the geminiviruses infecting the local hosts were spread throughout Africa for millions of years and later colonized cassava after its introduction in the 16th century and onwards. In addition to the transmission of CMD by whiteflies, human intervention can be another important cause for rapid geographical movement and exchange of geminiviruses. Recombination between members of different virus species has been a driving force for biodiversity, as exemplified by the EACMV/SACMV gradient between East Africa and South Africa, favoured by a natural corridor along the Eastern Rift Valley (Fig. S1, see Supporting Information; Ndunguru *et al.*, 2005). Very recently, Bull *et al.* (2006) conducted an exhaustive survey of the genetic diversity of CMGs across the major cassava-growing areas of Kenya, mainly representing EACMV and EACMVZ, as well as a novel begomovirus named East African cassava mosaic Kenya virus (EACMKV). The DNA-B components of these EACMV-like

viruses were much less diverse than their corresponding DNA-A components, but clustered as western and eastern (coastal) groups.

Comparison of the CMG sequences from Africa indicates that all isolates of ACMV, irrespective of their geographical origin, are clustered together with little variation in their genomic sequence. In contrast, the genomes of EACMV-like viruses are more genetically diverse as a result of the frequent recombination between members of different species (Fig. 3). In addition, there is a large range of phenotypic symptom variation for each of these virus isolates, irrespective of their location and their belonging to a particular species (Bull *et al.*, 2007; Pita *et al.*, 2001a). Thus, the explosion of the cassava mosaic pandemic in Africa is currently explained by a conjunction of several mechanisms: (i) the key synergistic interaction between two viruses (discussed in detail later in this article), leading to a huge increase in viral titre in the top of the plants, thus enhancing the whitefly transmission capacity; (ii) the CP recombinant fragment that may provide an advantage to EACMV-UG, as it has been shown that the whitefly transmission epitopes are coded by this fragment and that there has been co-adaptation, as far as the rate of transmission is concerned, between the local whiteflies and viruses (Maruthi *et al.*, 2002); and (iii) the tremendous increase in whitefly populations adapted to cassava in Uganda, where the epidemic started.

Cassava mosaic geminiviruses on the Indian subcontinent

In India, cassava is mainly grown in three southern states: Kerala, Tamil Nadu and Andhra Pradesh. In addition, it is sparsely cultivated in several other parts of India with a total production of 7.7 million tonnes (FAOSTAT, 2008). The first report of CMD in the Indian subcontinent was documented by Abraham (1956) and later elaborated by Alagianagalingam and Ramakrishnan (1966). A similar disease in Sri Lanka caused by a different begomovirus, Sri Lankan cassava mosaic virus (SLCMV), was reported by Austin (1986). The first clone of Indian cassava mosaic virus (ICMV-Ker) was obtained by Hong *et al.* (1993) and, subsequently, Saunders *et al.* (2002) cloned SLCMV. The DNA-A and DNA-B components of ICMV share 65% and 30% nucleotide sequence identity with ACMV DNA-A and DNA-B, respectively. Although SLCMV has similar iteron sequences to those of ACMV, it shares a greater identity with ICMV.

Patil and Dasgupta (2005) and Patil *et al.* (2005) reported the presence of SLCMV in southern India through differential polymerase chain reaction (PCR) studies, and developed a distribution map for ICMV and SLCMV. Subsequently, several ICMV and SLCMV isolates from southern India were cloned and their biodiversity and infectivity were studied (Dutt *et al.*, 2005; Patil *et al.*, 2007; Rothenstein *et al.*, 2006). Most of the CMGs cloned from

India were isolates of SLCMV with lower diversity compared with the variability of CMGs from Africa (Rothenstein *et al.*, 2006). Based on DNA-A sequence data, all of these isolates were classified as isolates of ICMV or SLCMV, and their phylogenetic analysis revealed recombination events between ancestors of ICMV and SLCMV (Rothenstein *et al.*, 2006).

THE WHITEFLY VECTOR AND CASSAVA MOSAIC DISEASE TRANSMISSION

The abundance of the whitefly vector (*B. tabaci*; Hemiptera: Aleyrodidae) is the second major distinguishing biological feature of the CMD pandemic in Africa, particularly at the epidemic 'front' between severely affected and relatively unaffected areas (Legg and Fauquet, 2004). CMD was first shown to be transmitted by *Bemisia* sp. in Africa and it is now considered to be a whitefly complex, with different populations occurring in different regions (Brown *et al.*, 1995). The CP of a geminivirus is specifically adapted for transmission by the local whitefly population, which explains the antigenic similarity of the CPs of begomoviruses from the same area (Harrison and Robinson, 1999) and co-adaptation between CMGs and their local whitefly populations (Maruthi *et al.*, 2002). Mutation studies made in ACMV to disrupt the CP showed that ACMV was infectious in plants, either by mechanical or agro-inoculation, but was not transmitted by whiteflies (Klinkenberg *et al.*, 1989). In addition, it has been shown that efficient whitefly transmission of ACMV requires products of both genomic components; the results of Liu *et al.* (1997) indicated that virus accumulation and transmission to new plants were determined by DNA-B (mostly *BC1* coding for the movement protein) and virus acquisition was determined by CP; it was clear that the loss of whitefly transmissibility could be caused by the prolonged vegetative propagation of a virus isolate.

The whitefly B-biotype is more fecund and has an extremely broad host range (Colvin *et al.*, 2004), which might have contributed to the transmission of new viruses from weed hosts to cultivated crop plants, thereby leading to the emergence of a number of geminivirus diseases throughout the world. Since the onset of the CMD epidemic in Uganda, an increased whitefly population has been reported, mainly comprising two distinct clusters of *B. tabaci* (Ug1 and Ug2) (Legg *et al.*, 2002; Sseruwagi *et al.*, 2005). In India, two distinct biotypes of whitefly, cassava biotype and the sweet potato biotype, have been reported through sequencing of the mitochondrial cytochrome oxidase I gene (*mtCOI*) (Lisha *et al.*, 2003). The inability of the B-biotype to colonize cassava in the Americas has been postulated as a major reason for the absence of CMD in that region (Carabali *et al.*, 2005). Although the problem of CMD has been partially resolved using resistant cassava cultivars, the abundant whitefly populations still continue to fuel the virus pandemic across Africa. It

is not known why these whiteflies have adapted so well to cassava, but it is important that we improve our understanding of whitefly behaviour and adaptation, as they are essential elements in the emergence of new diseases.

PLANT HOST RANGE OF CASSAVA MOSAIC GEMINIVIRUSES

The host range of viruses is also considered to be an important criterion to differentiate between geminiviruses and to better understand their epidemiology. Each CMG has different preferences for different hosts in addition to their natural host, cassava. Previous studies have shown that the experimental host range of ACMV is largely restricted to the Solanaceae family, within which it is more readily transmitted to members of the genera *Nicotiana* and *Datura* (Bock and Woods, 1983). Both ICMV and SLCMV also infect other species of *Nicotiana* (for example, *N. tabacum* and *N. glutinosa*), and SLCMV, in particular, is highly virulent with a very broad host range extending to *Arabidopsis*, as reported recently (Mittal *et al.*, 2008), and even to *Ageratum conyzoides* in association with ageratum yellow vein betasatellite (Saunders *et al.*, 2002). The host range of EACMV has not been studied extensively, but it is not known to infect *Nicotiana* species, except for the commonly used experimental host *N. benthamiana*, and this is also the case with SACMV (Berrie *et al.*, 2001). Thus, CMGs can be classified into two major groups based on their host range: ACMV, ICMV and SLCMV forming one class infecting all species of *Nicotiana* in addition to cassava, although with different levels of virulence, and EACMV-like viruses and SACMV in another class, which infects only *N. benthamiana* and cassava. Recently, there has been a report of an association of ACMV and EACMV-Ug with *Manihot glaziovii*, a wild species of cassava, which has been used as a source of resistance in CMD breeding programmes (Sserubombwe *et al.*, 2008).

GENOME ORGANIZATION AND GENE REGULATION

Geminiviruses are a large and diverse group of plant viruses characterized by geminate shaped particles of 30 × 20 nm that replicate their circular single-stranded DNA (ssDNA) genome (with a size of 2.7–2.8 kb) via double-stranded DNA (dsDNA) intermediates, either by rolling circle replication or a recombination-dependent replication mechanism in the nuclei of infected cells (Hanley-Bowdoin *et al.*, 1999; Jeske *et al.*, 2001). The Family *Geminiviridae* is divided into four genera on the basis of genome organization, host range and insect transmission (Hull, 2002). CMGs belong to the genus *Begomovirus* which contains whitefly-transmitted, dicotyledonous plant-infecting

Fig. 2 Genome organization of cassava mosaic geminiviruses (genus *Begomovirus*). The genome is split into two components, termed DNA-A (left) and DNA-B (right). DNA-A comprises six open reading frames (ORFs), and each ORF encodes a specific protein. *AC1*, replication-associated protein (Rep); *AC2*, transcriptional activator protein (TrAP); *AC3*, replication enhancer protein (REn); *AC4*, RNA-silencing suppressor; *AV1*, coat protein (CP); *AV2*, precoat protein. DNA-B has two ORFs: *BV1* encodes nuclear-shuttle protein (NSP), and *BC1* encodes movement protein (MP). C, Complementary-sense ORFs; V, virion-sense ORFs. The non-coding intergenic region, also referred to as the common region (CR), with the stem-loop structure is also depicted. The fragmented circles outside of the virus genome represent the DI-DNAs of Indian cassava mosaic virus (ICMV) and Sri Lankan cassava mosaic virus (SLCMV) with their respective positions; dark and light grey lines represent DNA-A and DNA-B sequences, respectively. (Adapted from Patil *et al.*, 2007.)

geminiviruses (Fauquet *et al.*, 2008). The DNA-A encodes four genes in the complementary sense (*AC1–AC4*) and two in the virion sense (*AV1* and *AV2*), responsible for replication, transcription enhancement and encapsidation functions, whereas DNA-B encodes one gene in the virion and complementary senses (*BV1* and *BC1*), respectively, required for intra- and inter-cellular movement (Fig. 2; Hull, 2002; Jeske, 2009).

The geminate particles which encapsidate the ACMV genomic components are known to release their genomic DNA through breaks occurring at the top and shoulders of the virus particles (Bottcher *et al.*, 2004; Kittelmann and Jeske, 2008). The begomovirus genomic components DNA-A and DNA-B share a common region of approximately 200 nucleotides having a high nucleotide sequence identity of more than 80% (Harrison and Robinson, 1999). The common region contains several regulatory elements, including two TATA motifs and also multiple copies of *cis*-elements known as iterons, which are the binding sites for the replication-associated protein (Rep) (Hanley-Bowdoin *et al.*, 1999). Different types of iteron sequences have been identified in CMGs and they can be classified into three groups: the ACMV type with isolates of ACMV, EACMV and SLCMV; the EACMV type encompassing all the other EACMV-like viruses and SACMV; and the ICMV type with ICMV isolates alone (Fig. 4).

Transcriptional regulation in begomoviruses has been studied extensively in both transgenic plants and protoplast systems (Frey *et al.*, 2001; Shivaprasad *et al.*, 2005). The dsDNAs formed during rolling circle replication serve as transcriptional templates and the geminivirus genomes are transcribed in a bidirectional manner by the bidirectional promoters located in the intergenic regions (Hanley-Bowdoin *et al.*, 1999). The begomovirus open reading

frame (ORF) *AC2* is a viral transcription factor that trans-activates the late viral genes *AV1* and *BV1* (Sunter and Bisaro, 1992), and recently it has been shown that *AC2* of ACMV trans-activates several plant genes (Lacatus and Sunter, 2008; Trinks *et al.*, 2005). In ACMV and other Old World begomoviruses, the *AV1* gene (CP) is overlapped at the 5' end by another ORF (*AV2*). Characterization of *AV2* of the CMGs indicated it to possess the redundant function of *AV2* in monopartite begomoviruses functioning as a movement protein; presumably, it is an evolutionary relic of a monopartite geminivirus that may still increase virus fitness, but is no longer indispensable in a bipartite genome (Bull *et al.*, 2007; Rothenstein *et al.*, 2007). In the case of the ACMV genome, five putative promoter regions have been identified and tested for their transcriptional activity in both protoplasts and transgenic plants, and, in contrast with DNA-A, the complementary-sense DNA-B promoter shows weak activity and the viral-sense promoter appears to be regulated by host factors (Frey *et al.*, 2001; Haley *et al.*, 1992; Zhan *et al.*, 1991). However, despite all of these studies, there are still major gaps in our understanding of the structural organization of the geminivirus promoter, gene regulation and transcription.

RECOMBINATION, PSEUDO-RECOMBINATION AND TRANS-REPLICATION: THE SOURCE OF MOLECULAR DIVERSITY

Recombination between geminiviruses was first revealed with the molecular analysis of EACMV-UG, the virus responsible for the pandemic on cassava in Africa (Deng *et al.*, 1997; Zhou *et al.*,

1997). Soon after, it was established that geminiviruses and, in particular, begomoviruses are subject to very intense recombination rates, leading to a rapidly changing and large molecular diversity (Lefeuvre *et al.*, 2009; Padidam *et al.*, 1999b) to respond to changes in the environment and to invade new ecological niches. CMGs are no exception, and Fig. 3 depicts the recombination maps of DNA-A and DNA-B of CMGs. However, it is striking that ACMV seems to be an exception to this rule, as all the available ACMV sequences originating from several places in Africa do not show any evidence of recombination. This could result from the lack of ancestral parents of this virus, or could indicate that not all begomoviruses are subject to recombination, although ACMV donated 550 nucleotides to an EACMV isolate to create EACMV-UG, associated with the CMD pandemic in Africa. By contrast, EACMV-like viruses seem to be highly prone to recombination, as shown by the diverse number of strains encompassing stretches of DNA from various known origins (Fig. 4).

The complementary-sense gene transcription initiation and termination sites and virion strand origins of replication on the circular DNA molecules at which the recombination repairs of double-stranded genome breakages are resolved have been identified as the recombination hot spots in geminiviruses (Cromie *et al.*, 2001; Lefeuvre *et al.*, 2007, 2009; Varsani *et al.*, 2008). The intergenic regions showed a greater tendency to be moved between the genomes than did the actual genes and, among the coding regions, the complementary-sense genes showed relatively higher frequency of recombination relative to the virion-sense genes. It has also been shown that there are very few recombination events occurring within the stretches of genome coding for the structural proteins, and most of the recombinations occur on the periphery of the ORFs (Lefeuvre *et al.*, 2009).

Viruses can acquire additional DNA components through a mechanism called regulon grafting, whereby the transfer of intergenic regions harbouring replicational and transcriptional control elements to further DNAs occurs, and this has been reported in the case of ACMV (Jovel *et al.*, 2004; Roberts and Stanley, 1994). It has been hypothesized that the DNA-B component of SLCMV is actually a captured ICMV-B, following component acquisition by a hypothetical monopartite strain of SLCMV (Saunders *et al.*, 2002). This hypothesis illustrates how the compatibility of DNA-A and DNA-B can be obtained, enhanced and maintained, and indicates a way in which a begomovirus might acquire a new DNA-B during mixed infection of two begomoviruses, even if they are too distantly related to produce a viable pseudo-recombinant. Thus, this mechanism provides an explanation for the evolution of bipartite begomoviruses from monopartite ancestors (Rojas *et al.*, 2005).

Pseudo-recombination is a term used to describe the situation in which the DNA-A and DNA-B of the two components of a geminivirus originate from two different geminiviruses.

In general, pseudo-recombination occurs when there is re-assortment of the genomic DNA-A and DNA-B of isolates of the same begomovirus species, but not those of distinct begomovirus species (Chakraborty *et al.*, 2008). Pseudo-recombination in which the DNA-A of one virus coexists with and trans-replicates the DNA-B of another virus has been reported for CMGs occurring in Uganda. The most frequent naturally occurring CMG infections in Uganda involve the co-occurrence of EACMV-UG2 DNA-A and EACMV-UG3 DNA-B (Pita *et al.*, 2001a). EACMZV seems to be a recombination product of EACMV with an unknown begomovirus from which the 5' terminal Rep coding sequence and part of the intergenic region have been captured, whereas EACMV and SACMV recombined to form members of a new species, EACMKV (Bull *et al.*, 2006) (Fig. 4). EACMZV has 'GGAGA' as the iteron sequence and is incompatible with other EACMVs, but surprisingly ACMV and EACMZV did not form a viable pseudo-recombinant despite having identical iteron sequences, indicating that there could be additional factors required for trans-replication (Fig. 4). In contrast, both EACMV and EACMKV, sharing the same iteron sequences (GGGGG), can form viable pseudo-recombinants (Bull *et al.*, 2007).

The diversification of geminiviruses mainly occurs because of recombination and the exchange of genomic components in mixed infections (Padidam *et al.*, 1999b). CMGs are known to undergo recombination in mixed infections (Fig. 4) and the survival of such recombined molecules depends on several factors, including their ability to replicate (Pita *et al.*, 2001b). Trans-replication is an important mechanism, wherein the DNA-A component of a given geminivirus is capable of replicating a DNA-B component of a heterologous virus, thereby offering an additional possibility for new interaction and providing a new source of CMG biodiversity. Trans-replication and trans-complementation have been experimentally demonstrated among CMGs that exhibit distinguishing characteristics; moreover, ACMV DNA-B genes can functionally interact with TGMV genomic components (Saunders and Stanley, 1995; Stanley *et al.*, 1985).

Recently, natural trans-replications within different EACMV strains (EACMV-UG2 DNA-A and EACMV-UG3 DNA-B) or between ACMV DNA-A and EACMV-UG3 DNA-B have been reported from Uganda with epidemiological implications (Pita *et al.*, 2001a,b). Evidence for trans-replication between EACMCV-CM DNA-A and ACMV-CM DNA-B and between EACMCV-CM DNA-A and EACMV-UG3 DNA-B has also been obtained in tobacco protoplasts (Pita *et al.*, 2001a). It appears that a DNA-A molecule can recruit a DNA-B molecule provided that there is compatibility between their Rep and iterons, although such an incompatibility can be overcome by the exchange of intergenic regions (Roberts and Stanley, 1994; Saunders *et al.*, 2001, 2002). Both infectivity studies and leaf disc assays in *N. benthamiana* have demonstrated the following: ACMV DNA-A trans-replicated DNA-B of SLCMV, but not that of

DNA-A Component

DNA-B Component

Fig. 3 Recombination linearized map of representatives of the nine species of cassava-infecting geminivirus and their strains. Each horizontal line represents the genotype of one virus isolate, and the different patterns and shades represent the tentative origins of the putative recombinant fragments. The length of the genome is indicated at the top of each diagram and the genome organization is depicted at the bottom; the abbreviated names of the virus isolates are listed on the left. The shades and patterns for the different original parental genome regions are provided below the genome organization. The numbers at the top of the diagram correspond to the length of the genome, starting at the nicking site of the conserved nona-nucleotide TAATATT/AC. The virus name abbreviations are given in the text.

Fig. 4 Sequence comparison of intergenic regions of cassava-infecting begomoviruses depicting the iteron sequences and their relative positions. Comparison of DNA-A components of representatives of all nine species and their strains, showing the location of iterons (highlighted with different shades for the different types of iteron) with reference to the TATA box and the stem-loop with the nona-nucleotide sequence embedded in it (boxed), and their orientation indicated by arrows, all located upstream of the stem-loop structure. Spaces (-) have been introduced to align the motifs, and each shade represents a unique iteron sequence. The virus name abbreviations are given in the text.

ICMV; SLCMV DNA-A trans-replicated ACMV DNA-B but not that of ICMV; and ICMV DNA-A was unable to trans-replicate either ACMV or SLCMV DNA-B (Saunders *et al.*, 2002). This highlights the importance of iteron sequences (Fig. 4) in forming a successful pseudo-recombination complex, and thus in playing an important role in the evolution of geminiviruses through recombination and pseudo-recombination events.

DEFECTIVE AND SATELLITE DNAs ASSOCIATED WITH CASSAVA MOSAIC DISEASE

In addition to genomic components, smaller sized DNAs, referred to as defective DNA (def-DNA), often occur naturally in several geminivirus-infected plants. Some def-DNAs have been shown to interfere with virus proliferation, as they are associated with a delay in and an attenuation of symptoms; consequently, they are named 'defective interfering DNAs (DI-DNAs)' (reviewed by Patil and Dasgupta, 2006). def-DNAs are usually half the size of the full-length genomic component and may be formed by sequence deletion, duplication, inversion or rearrangement of viral DNA and, in some cases, insertion of foreign sequences (Fig. 2; Stanley *et al.*, 1997). def-DNAs always contain the intergenic region, harbouring all *cis*-acting sequences required for replication, and, for bipartite viruses, def-DNAs are predominantly derived from the DNA-B component (Fig. 2). In most cases, the *BV1* ORF and the C-terminus of *BC1* are deleted (Stanley and Townsend, 1986). def-DNAs are mostly formed at higher frequency in laboratory host plants, such as *N. benthamiana*, and rarely in their natural hosts, indicating that they are not well adapted to the experimental hosts.

Stanley and Townsend (1986) first characterized the def-DNA forms associated with ACMV DNA-B as being usually encapsidated in isometric particles of 18×20 nm in size, whereas genomic DNA was encapsidated in geminate particles (Frischmuth *et al.*, 2001). Recently, DI-DNA derived from the DNA-A component of EACMV and SLCMV has been reported, albeit at a lower frequency (Ndunguru *et al.*, 2006; Patil *et al.*, 2007). Patil *et al.* (2007) also reported defective molecules generated from intercomponent recombination between DNA-A and DNA-B components of CMGs from India, in which the sites of deletions or recombinations were mapped to the potential recognition sites of replication-associated enzymes topoisomerase I and II. Symptom amelioration by DI-DNA is one of its most fascinating aspects, which could be mainly a result of the slowing down of the replication of the cognate helper virus caused by the sequestration of limited amounts of Rep proteins by the DI-DNA (Patil and Dasgupta, 2006).

In addition to the presence of def-DNA molecules as a sub-genomic component, there are satellite DNA molecules, termed alphasatellites and betasatellites, associated with the begomo-

virus disease complex, which mostly depend on the proteins encoded by the helper virus for their replication, movement and encapsidation (Briddon *et al.*, 2008). Hitherto, most of the ssDNA satellites reported have been associated with monopartite begomoviruses, but the recent discovery of resistance-breaking satellite molecules, termed 'satDNA', associated with CMD in Tanzania has posed serious threats to cassava cultivation (J. Ndunguru and C. M. Fauquet, personal communication). Experiments on the trans-replication of begomovirus betasatellites and alphasatellites by cassava-infecting geminiviruses and their further effect on symptom modulation have demonstrated that each species of CMG shows a differential interaction with the DNA satellites (B. L. Patil and C. M. Fauquet, personal observation).

VIRULENCE OF CASSAVA MOSAIC GEMINIVIRUSES

The aetiology of CMD has increased in complexity during the past decade, particularly on the African continent, reflected by the increasing number of novel species of CMGs from different locations. The presence of EACMV-based recombinants, such as EACMV-UG, has been correlated with the severe CMD epidemic in Uganda (Deng *et al.*, 1997; Zhou *et al.*, 1997), although, as of today, there is no proof that this recombination was responsible for the CMD pandemic.

The response of the host to virus infection is characterized by an initial onset of systemic symptoms from which the plant may or may not recover. The detailed analysis of representatives of four distinct CMG species by Chellappan *et al.* (2004b) has allowed us to distinguish two distinct infection types in cassava: a rapid onset from which plants recover (ACMV and SLCMV), and a slow onset from which plants do not recover (EACMV and EACMCV). In addition to gene silencing, the DNA-B component can have a profound effect on symptom modulation (Bull *et al.*, 2007). With the recent identification of two novel CMGs, i.e. EACMKV and EACMZV, and their isolates, it has been shown that symptom severity or symptom recovery could be a manifestation of the type of DNA-B component involved in the infection, which has been revealed to modulate the symptom phenotypes of these viruses (von Arnim and Stanley, 1992; Bull *et al.*, 2007). Thus, such symptom variability in infected cassava plants, which ranges from mild to extremely severe, has been attributed to variation in virus strains or their genes, virulence, host susceptibility or modification of vector activity (Fig. 5). Host susceptibility is variable, with some germplasm, such as the cassava landrace TME3, considered to be extremely resistant to CMGs, controlled by a major dominant resistance gene locus CMD-2 (Akano *et al.*, 2002). Although this landrace is highly resistant, it is not immune to infection from several CMGs, including some isolates of EACMV and EACMZV, which initially give significant symptoms

Fig. 5 Differential virulence of cassava mosaic geminivirus (CMG) isolates. Symptoms of isolates of African cassava mosaic virus (ACMV) (top panels) and East African cassava mosaic virus—Uganda (EACMV-UG) (bottom panels) reported from Uganda in 1998. Mild isolates and severe isolates are shown on the left and right, respectively.

in TME3; later, the cassava plants recover completely from the infection, possibly indicating that the CMD resistance in TME3 could be through the gene silencing mechanism (B. L. Patil and C. M. Fauquet, personal observation).

GENE SILENCING, SUPPRESSION AND SYNERGISM IN CASSAVA MOSAIC GEMINIVIRUSES

One of the most important biological functions for RNA silencing, first established in plants, is an adaptive defence system

against viruses (Ding and Voinnet, 2007; Waterhouse *et al.*, 2001). Recently, it has been shown that all the CMGs investigated trigger post-transcriptional gene silencing (PTGS) in infected plants with the production of virus-specific short interfering RNAs (siRNAs) (Chellappan *et al.*, 2004b). This was attributed to the formation of dsRNA caused by the overlap at the 3' ends of transcripts produced from opposite polarity through bidirectional transcription, and also possibly by dsRNA formation in the secondary structure of the virus transcripts. The symptom recovery phenotype exhibited by ACMV and SLCMV in cassava has been correlated with the production of virus-derived siRNAs

through PTGS (Chellappan *et al.*, 2004b). Detailed analysis of the composition of these siRNAs has revealed that most of the siRNAs in ACMV-infected plants are 21 nucleotides in size; the majority are derived from the DNA-A component, particularly from a region that corresponds to the C-terminus of *AC1* overlapping with the N-terminus of the *AC2* gene, and those derived from the DNA-B component are mostly from the C-terminus of *BC1* (Chellappan *et al.*, 2004a). Virus-derived siRNAs implicating virus-induced PTGS have also been demonstrated for EACMCV, EACMV and ICMV, although the infected plants do not recover. These observations indicate that, although the plant responds to virus infection through the induction of gene silencing, the plant might not always recover from virus infection. The interaction of host gene silencing machinery with other viral factors or suppressors might counteract this effect, and thus might not lead to the recovery of plants from the infection.

In addition, virus–host interactions are strongly modified by environmental factors, particularly temperature and light conditions (Chellappan *et al.*, 2005b). In CMGs, irrespective of the recovery- and non-recovery-type viruses, there was a general trend for symptom recovery with an increase in temperature, suggesting increased PTGS or transcriptional gene silencing (TGS) at high temperatures (Chellappan *et al.*, 2004b). Recently, Akbergenov *et al.* (2006) have characterized siRNAs of three different size classes (21, 22 and 24 nucleotides), which were derived from both the coding and intergenic regions of ACMV.

To circumvent the constraints imposed by gene silencing triggered in the host plant, geminiviruses have evolved suppressor proteins that counteract this host defence response. Voinnet *et al.* (1999) first showed that *AC2* of ACMV, when transiently expressed, could suppress green fluorescent protein (GFP) silencing. Transient expression of ACMV *AC2* also induced the expression of about 30 genes in *Arabidopsis*, such as Werner exonuclease-like 1, which was capable of suppressing RNA silencing, thus suggesting that *AC2* suppresses silencing indirectly by activating the expression of cellular proteins (Trinks *et al.*, 2005). The *AC2* protein could suppress silencing by two independent mechanisms, one of which targets cytoplasmic RNA silencing and the other siRNA-directed methylation (Bisaro, 2006). In this regard, it is interesting to note that geminivirus *AC2* also interacts with adenosine kinase (ADK), an enzyme involved in sustaining methylation through methyltransferase activity, although it has not yet been investigated whether ADK is related to PTGS (Wang *et al.*, 2003). In addition to its interaction with ADK, *AC2* also suppresses sucrose non-fermenting 1 (SNF1) kinase, which plays a central role in the regulation of metabolism, indicating that SNF1-mediated responses constitute a novel defence pathway in plants (Hao *et al.*, 2003; Wang *et al.*, 2005). Thus, further research might unravel the intricate network between these diverse defence pathways.

The *AC4* gene is entirely embedded within *AC1* and, despite the high level of conservation of *AC1*, *AC4* is one of the least conserved genes of all the geminiviruses. Functional analysis of *AC4* has proved to be elusive, as the mutagenesis and/or transgenic expression of some *AC4* genes results in no phenotype, whereas others produce phenotypes consistent with a movement protein or a symptom determinant (Krake *et al.*, 1998; Vanitharani *et al.*, 2004). *AC4* has been found to be capable of suppressing gene silencing for some CMGs, enhancing disease symptoms. Transient studies on the suppression activity of *AC4* from four different CMGs in *N. benthamiana* 16c plants showed that two *AC4* proteins, from viruses giving recovery phenotypes in cassava (ACMV and SLCMV), showed suppressor activity with inhibition of GFP-specific siRNAs and increased accumulation of GFP mRNA, whereas two other *AC4* proteins from non-recovery-type viruses (EACMCV and ICMV) did not show similar silencing suppression activity. This may reflect considerable functional diversity within *AC4* sequences of distinct begomoviruses. Transgenic expression of *AC4* of ACMV, but not EACMCV, leads to developmental defects in *Arabidopsis* that resemble virus disease symptoms, which has been explained by its interference in the microRNA pathway (Chellappan *et al.*, 2005a). Conversely, in the case of non-recovery-type viruses, the *AC2* proteins were effective silencing suppressors, whereas those from recovery-type viruses were less effective (Vanitharani *et al.*, 2004).

Mixed infections of ACMV and EACMCV have been shown to be an important feature of the severe CMD first reported from Uganda and subsequently in neighbouring countries (Harrison *et al.*, 1997; Pita *et al.*, 2001a). A synergistic effect in a mixed infection occurs when there is a significant difference in *AC2* and *AC4* functions of these individual viruses. The difference in the phenotypes exhibited by these viruses further suggests that *AC2* and *AC4* may act at different steps in the silencing pathway, and that the effect of *AC4* could be more transient, allowing some hosts to overcome it. These experiments revealed that the synergism between ACMV and EACMCV was mediated by the complementation of two different PTGS suppressors, *AC4* and *AC2* present in ACMV and EACMCV, respectively, acting in a temporal and spatial manner (Fig. 6; Fondong *et al.*, 2000; Vanitharani *et al.*, 2004). Therefore, having more than one type of PTGS suppressor provides an advantage to viruses, synergistically interacting in mixed infections, and thus helps in the establishment of a severe/successful disease.

STRATEGIES FOR CONTROLLING CASSAVA MOSAIC DISEASE

Several approaches have been used to control CMD, including improved cultural practices, breeding of natural resistant varieties and improved biotechnological strategies, which have been reviewed extensively by Thresh and Cooter (2005), Thresh

Fig. 6 Symptom severity of cassava plants. Dual infection leads to synergistic severe disease and corresponding post-transcriptional gene silencing (PTGS) capacity of each virus. Second panel (from left to right): non-infected healthy control plant; plant infected with African cassava mosaic virus (ACMV); plant infected with East African cassava mosaic Cameroon virus (EACMCV); dual infection with both viruses, which leads to synergistic severe disease (candle-stick symptoms). The third panel indicates the function of AC2- and the bottom panel indicates the function of AC4-type PTGS suppressor proteins. EACMCV-AC2 (EAC2; third panel, second red image from right) and ACMV-AC4 (AAC4; bottom panel, third red image from right) are positive for PTGS suppression. Synergism is explained by dual suppression of PTGS (red panels on the right). Green control panels (left) represent frame-shift mutants of ACMV-AC4 and EACMCV-AC2 that have lost the capacity to suppress PTGS. ACMV-AC2 is considered to be negative, as PTGS suppression is weak, and EACMCV-AC4 is not a PTGS suppressor. (Adapted from Vanitharani *et al.*, 2005.)

(2006), Sudarshana *et al.* (2007) and Vanderschuren *et al.* (2007b), and thus they are described only briefly here. Of all the methods used for the control of CMD resistance, genetic engineering is a rapid method of transferring resistance genes to traditional cultivars, bypassing the possibility of the appearance

of undesired traits. AC1 is an excellent target for interference by the expression of mutant proteins because of its crucial functions in the replication cycle and its multiple interactions with both the viral and host proteins, and targeting it may also provide broader resistance against different geminiviruses (Brunetti *et al.*, 2001).

To engineer a virus-inducible hypersensitive reaction, attempts were made to express apoptotic non-viral proteins controlled by the ACMV promoter trans-activated by the AC2 protein on infection (Zhang *et al.*, 2003); however, the use of such proteins for practical purposes has biosafety concerns. Symptom amelioration by DI-DNA was tested as a possible strategy to control ACMV in both the model host *N. benthamiana* and in cassava, but did not lead to significant levels of resistance in cassava (Frischmuth and Stanley, 1993; Taylor *et al.*, 2003).

An alternative approach is antisense RNA technology, also known as RNAi/gene silencing technology, in which resistance to ACMV infection of cassava has been achieved with high efficacy by expressing anti-sense RNAs, or hairpin RNAs, against viral transcripts encoding Rep, TrAP and REN (Chellappan *et al.*, 2004a; Vanderschuren *et al.*, 2007a; Zhang *et al.*, 2005). However, the major concern in using gene silencing strategies to control viruses is the fact that many viruses encode one or more PTGS suppressors that could suppress the enhanced resistance mechanism (Vanitharani *et al.*, 2005) and, above all, that point mutations or recombinations in the target viruses could 'break' the engineered resistance; therefore, different strategies, such as the G5 strategy, are now being developed to counter this issue.

The non-specific DNA-binding activities of some proteins, such as the G5 protein from isolates of *Enterobacteriophage M13* which binds to ssDNA, competing with CP and affecting geminivirus movement, are being used to engineer broad-spectrum resistance to cassava-infecting geminiviruses and their satellites (Fig. S2, see Supporting Information; Padidam *et al.*, 1999a; J.S. Yadav, N. Taylor and C.M. Fauquet, personal communication). More recently, peptide aptamers, a class of recombinant proteins that bind to and inactivate the target protein, and artificial zinc finger proteins, which block the Rep protein, are being tested to develop geminivirus-resistant transgenics (Takenaka *et al.*, 2007; L.A. Lopez-Ochoa and L. Hanley-Bowdoin, personal communication).

Although further improvements in the group of transgene-derived CMD resistance strategies are required before field deployment, these approaches offer great promise. Most importantly, they can provide a valuable complement to conventional breeding approaches to control CMD, and could effectively allow farmers to continue to grow locally preferred cultivars. In view of the difficulties that have been experienced in some pandemic-affected areas in addressing farmer quality preferences, this represents a very important development. Currently, two major research institutes, namely the Donald Danforth Plant Science Center (St. Louis, MO, USA) and the Swiss Federal Institute of Technology (Zurich, Switzerland), in collaboration with several African countries, India and other laboratories across the globe, are using different biotechnological approaches to find a remedy for CMD, together with improving

other quality traits of cassava. The two Consultative Group on International Agricultural Research (CGIAR) institutes, CIAT (International Centre for Tropical Agriculture, Cali, Colombia) and IITA (International Institute of Tropical Agriculture, Ibadan, Nigeria), are using molecular markers to tag genes for CMD resistance from multiple sources of resistance as a means of pyramiding different genes for durable resistance (M. Fregene, personal communication).

CONCLUSION

In the last two decades, significant progress has been made in our knowledge of CMGs in terms of their structure, diversity, movement and interaction with their host. This has been supplemented by a wealth of information generated by genomic studies and the description of hundreds of virus sequences. Despite this progress, it has not been possible to understand the presence of this vast diversity in the case of EACMV-like viruses and ICMV in comparison with ACMV and SLCMV. The diversification and rapid spread of EACMV-like viruses, their high rate of recombination combined with synergistic interaction with other CMGs and the population explosion of their whitefly vector are major threats to cassava cultivation in Africa and the Indian subcontinent. In addition, there is a huge threat of CMD introduction in South America and South-East Asia. Thus, strict quarantine regulations need to be exercised to prevent the import of CMGs and cassava-adapted whitefly biotypes.

More recently, dual pandemics have been reported of two major viral diseases infecting cassava, 'cassava mosaic disease' and 'cassava brown streak disease', caused by an RNA virus (putative member of the genus *Ipomovirus* of the Family *Potyviriidae*), in the coastal regions of eastern Africa. The rapid spread of these diseases into eastern and central regions of Africa is alarming the cassava community (Alicai *et al.*, 2007).

Undoubtedly, the discovery of the gene silencing mechanism will have a profound impact on the understanding of plant-virus interactions and also on the control of plant viruses. Further exploitation of this wealth of knowledge is essential to improve the efficiency and durability of all existing strategies, each approach with its own limitations, as well as for developing potential new strategies to cope with the emergence of resistance-breaking, cassava-infecting geminiviruses and satellites. These novel approaches for the management of CMD appear to have great potential in the near future, a prospect that certainly justifies further investment in this technology to herald another green revolution.

ACKNOWLEDGEMENTS

The authors acknowledge Patricia Cosgrove, Dr Vasundhara Patil and Dr Martin Fregene for critical reading of the manuscript.

REFERENCES

- Abraham, A. (1956) *Tapioca Cultivation in India*. Pusa, New Delhi: Indian Council of Agricultural Research.
- Akano, O., Dixon, O., Mba, C., Barrera, E. and Fregene, M. (2002) Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor. Appl. Genet.* **105**, 521–525.
- Akbergrenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Jr, Hohn, T. and Pooggin, M.M. (2006) Molecular characterization of geminivirus-derived small RNAs in different plant species. *Nucleic Acids Res.* **34**, 462–471.
- Alagianagalingam, M.N. and Ramakrishnan, K. (1966) Cassava mosaic in India. *South Ind. Hort.* **14**, 71–72.
- Alicai, T., Omongo, C.A., Maruthi, M.N., Hillocks, R.J., Baguma, Y., Kawuki, R., Bua, A., Otim-Nape, G.W. and Colvin, J. (2007) Re-emergence of cassava brown streak disease in Uganda. *Plant Dis.* **91**, 24–29.
- Ariyo, O.A., Koebler, M., Dixon, A.G.O. and Winter, S. (2005) Molecular variability and distribution of cassava mosaic begomoviruses in Nigeria. *J. Phytopathol.* **153**, 226–231.
- von Arnim, A. and Stanley, J. (1992) Determinants of tomato golden mosaic virus symptom development located on DNA B. *Virology*, **186**, 286–293.
- Austin, M.D.N. (1986) Scientists identify cassava virus. *Asian Agribus.* **3**, 10.
- Berrie, L.C., Rybicki, E.P. and Rey, M.E. (2001) Complete nucleotide sequence and host range of South African cassava mosaic virus: further evidence for recombination amongst begomoviruses. *J. Gen. Virol.* **82**, 53–58.
- Bisaro, D.M. (2006) Silencing suppression by geminivirus proteins. *Virology*, **344**, 158–168.
- Bock, K.R. and Woods, R.D. (1983) The etiology of African cassava mosaic disease. *Plant Dis.* **67**, 994–995.
- Bock, K.R., Guthrie, E.J. and Figueiredo, G. (1981) A strain of cassava latent virus occurring in coastal districts of Kenya. *Ann. Appl. Biol.* **99**, 151–159.
- Bottcher, B., Unseld, S., Ceulemans, H., Russell, R.B. and Jeske, H. (2004) Geminite structures of African cassava mosaic virus. *J. Virol.* **78**, 6758–6765.
- Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X. and Fauquet, C.M. (2008) Recommendations for the classification and nomenclature of the DNA-beta satellites of begomoviruses. *Arch. Virol.* **153**, 763–781.
- Brown, J.K., Frohlich, D.R. and Rosell, R.C. (1995) The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Annu. Rev. Entomol.* **40**, 511–534.
- Brunetti, A., Tavazza, R., Noris, E., Lucioli, A., Accotto, G.P. and Tavazza, M. (2001) Transgenically expressed T-Rep of tomato yellow leaf curl Sardinia virus acts as a trans-dominant-negative mutant, inhibiting viral transcription and replication. *J. Virol.* **75**, 10 573–10 581.
- Bull, S.E., Briddon, R.W., Sserubombwe, W.S., Ngugi, K., Markham, P.G. and Stanley, J. (2006) Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. *J. Gen. Virol.* **87**, 3053–3065.
- Bull, S.E., Briddon, R.W., Sserubombwe, W.S., Ngugi, K., Markham, P.G. and Stanley, J. (2007) Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. *J. Gen. Virol.* **88**, 1624–1633.
- Carabali, A., Bellotti, A.C., Montoya-Lerma, J. and Cuellar, M.E. (2005) Adaptation of *Bemisia tabaci* biotype B (Gennadius) to cassava, *Manihot esculenta* (Crantz). *Crop Prot.* **24**, 643–649.
- Chakraborty, S., Vanitharani, R., Chattopadhyay, B. and Fauquet, C.M. (2008) Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India. *J. Gen. Virol.* **89**, 818–828.
- Chellappan, P., Masona, M.V., Vanitharani, R., Taylor, N.J. and Fauquet, C.M. (2004a) Broad spectrum resistance to ssDNA viruses associated with transgene-induced gene silencing in cassava. *Plant Mol. Biol.* **56**, 601–611.
- Chellappan, P., Vanitharani, R. and Fauquet, C.M. (2004b) 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.
- Chellappan, P., Vanitharani, R. and Fauquet, C.M. (2005a) MicroRNA-binding viral protein interferes with Arabidopsis development. *Proc. Natl. Acad. Sci. USA*, **102**, 10381–10386.
- Chellappan, P., Vanitharani, R., Ogbe, F. and Fauquet, C.M. (2005b) Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant Physiol.* **138**, 1828–1841.
- Colvin, J., Omongo, C.A., Maruthi, M.N., Otim-Nape, G.W. and Thresh, J.M. (2004) Dual begomovirus infections and high *Bemisia tabaci* populations: two factors driving the spread of a cassava mosaic disease pandemic. *Plant Pathol.* **53**, 577.
- Cromie, G.A., Connelly, J.C. and Leach, D.R. (2001) Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol. Cell*, **8**, 1163–1174.
- Deng, D., Otim-Nape, W.G., Sangare, A., Ogwai, S., Beachy, R.N. and Fauquet, C.M. (1997) Presence of a new virus closely related to East African cassava mosaic geminivirus, associated with cassava mosaic outbreak in Uganda. *Afr. J. Root Tuber Crops*, **2**, 23–28.
- Ding, S.W. and Voinnet, O. (2007) Antiviral immunity directed by small RNAs. *Cell*, **130**, 413–426.
- Dutt, N., Briddon, R.W. and Dasgupta, I. (2005) Identification of a second begomovirus, Sri Lankan cassava mosaic virus causing cassava mosaic disease in India. *Arch. Virol.* **150**, 2101–2108.
- FAOSTAT (2008) FAO database. Food and Agriculture Organization of the United Nations, Rome, Italy. <http://faostat.fao.org/site/567/default.aspx>.
- Fargette, D., Konate, G., Fauquet, C., Muller, E., Peterschmitt, M. and Thresh, J.M. (2006) Molecular ecology and emergence of tropical plant viruses. *Annu. Rev. Phytopathol.* **44**, 235–260.
- 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.
- Fondong, V.N., Pita, J.S., Rey, M.E.C., De Kochko, A., Beachy, R.N. and Fauquet, C.M. (2000) Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol.* **81**, 287.
- Frey, P.M., Schärer-Hernandez, N.G., Futterer, J., Potrykus, I. and Puonti-Kaerlas, J. (2001) Simultaneous analysis of the bidirectional African cassava mosaic virus promoter activity using two different luciferase genes. *Virus Genes*, **22**, 231–242.

- Frischmuth, T. and Stanley, J. (1993) Strategies for the control of geminivirus diseases. *Semin. Virol.* **4**, 329–337.
- Frischmuth, T., Ringel, M. and Kocher, C. (2001) The size of encapsidated single-stranded DNA determines the multiplicity of African cassava mosaic virus particles. *J. Gen. Virol.* **82**, 673–676.
- Haley, A., Zhan, S., Richardson, K., Head, K. and Morris, B. (1992) Regulation of the activities of African cassava mosaic virus promoters by the AC1, AC2, and AC3 gene products. *Virology*, **188**, 905–909.
- Hanley-Bowdoin, L., Settlege, S.B., Orozco, B.M., Nagar, S. and Robertson, D. (1999) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *CRC Crit. Rev. Plant Sci.* **18**, 71–106.
- Hao, L., Wang, H., Sunter, G. and Bisaro, D.M. (2003) Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. *Plant Cell*, **15**, 1034–1048.
- Harrison, B.D. and Robinson, D.J. (1999) Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). *Annu. Rev. Phytopathol.* **37**, 369–398.
- Harrison, B.D., Zhou, X., Otim-Nape, G.W., Liu, Y. and Robinson, D.J. (1997) Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Ann. Appl. Biol.* **131**, 437–448.
- Hong, Y.G., Robinson, D.J. and Harrison, B.D. (1993) Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted geminiviruses in cassava. *J. Gen. Virol.* **74**, 2437–2443.
- Hull, R. (2002) Genome organization. In: *Matthew's Plant Virology*, 4th edn. (Hull, R. ed.), pp. 171–224. London: Academic Press.
- Jeske, H. (2009) Geminiviruses. *Curr. Top. Microbiol. Immunol.* **331**, 185–226.
- 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.
- Jovel, J., Reski, G., Rothenstein, D., Ringel, M., Frischmuth, T. and Jeske, H. (2004) Sida micrantha mosaic is associated with a complex infection of begomoviruses different from Abutilon mosaic virus. *Arch. Virol.* **149**, 829–841.
- Kittelman, K. and Jeske, H. (2008) Disassembly of African cassava mosaic virus. *J. Gen. Virol.* **89**, 2029–2036.
- Klinkenberg, F.A., Ellwood, S. and Stanley, J. (1989) Fate of African cassava mosaic virus coat protein deletion mutants after agroinoculation. *J. Gen. Virol.* **70**, 1837–1844.
- Krake, L.R., Rezaian, M.A. and Dry, I.B. (1998) Expression of the tomato leaf curl geminivirus C4 gene produces viruslike symptoms in transgenic plants. *Mol. Plant-Microbe Interact.* **11**, 413–417.
- Lacatus, G. and Sunter, G. (2008) Functional analysis of bipartite begomovirus coat protein promoter sequences. *Virology*, **376**, 79–89.
- Lefevre, P., Martin, D.P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., Becker, N., Reynaud, B. and Lett, J.M. (2007) Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. *J. Gen. Virol.* **88**, 3458–3468.
- Lefevre, P., Lett, J.M., Varsani, A. and Martin, D.P. (2009) Widely conserved recombination patterns amongst single stranded DNA viruses. *J. Virol.* **83**, 2697–2707.
- Legg, J.P. and Fauquet, C.M. (2004) Cassava mosaic geminiviruses in Africa. *Plant Mol. Biol.* **56**, 585–599.
- Legg, J.P. and Thresh, J.M. (2000) Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus Res.* **71**, 135–149.
- Legg, J.P., French, R., Rogan, D., Okao-Okuja, G. and Brown, J.K. (2002) A distinct *Bemisia tabaci* (Gennadius) (Hemiptera: Sternorrhyncha: Aleyrodidae) genotype cluster is associated with the epidemic of severe cassava mosaic virus disease in Uganda. *Mol. Ecol.* **11**, 1219–1229.
- Legg, J.P., Owor, B., Sseruwagi, P. and Ndunguru, J. (2006) Cassava mosaic virus disease in East and Central Africa: epidemiology and management of a regional pandemic. *Adv. Virus Res.* **67**, 355–418.
- Lisha, V.S., Antony, B., Palaniswami, M.S. and Henneberry, T.J. (2003) *Bemisia tabaci* (Homoptera: Aleyrodidae) biotypes in India. *J. Econ. Entomol.* **96**, 322–327.
- Liu, S., Bedford, I.D., Briddon, R.W. and Markham, P.G. (1997) Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *J. Gen. Virol.* **78**, 1791–1794.
- 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.
- Mittal, D., Borah, B.K. and Dasgupta, I. (2008) Agroinfection of cloned Sri Lankan cassava mosaic virus DNA to *Arabidopsis thaliana*, *Nicotiana tabacum* and cassava. *Arch. Virol.* **153**, 2149–2155.
- Ndunguru, J., Legg, J.P., Aveling, T.A., Thompson, G. and Fauquet, C.M. (2005) Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. *Virol. J.* **2**, 21.
- Ndunguru, J., Legg, J.P., Fofana, I.B.F., Aveling, T.A.S., Thompson, G. and Fauquet, C.M. (2006) Identification of a defective molecule derived from DNA-A of the bipartite begomovirus of East African cassava mosaic virus. *Plant Pathol.* **55**, 2–10.
- Padidam, M., Beachy, R.N. and Fauquet, C.M. (1999a) A phage single-stranded DNA (ssDNA) binding protein complements ssDNA accumulation of a geminivirus and interferes with viral movement. *J. Virol.* **73**, 1609–1616.
- Padidam, M., Sawyer, S. and Fauquet, C.M. (1999b) Possible emergence of new geminiviruses by frequent recombination. *Virology*, **265**, 218–225.
- Patil, B.L. and Dasgupta, I. (2005) Diversity of viruses infecting some agriculturally important crop plants in India. In: *International Conference on Microbial Diversity: Current Perspectives and Potential Applications* (Satyanarayana, T. and Johri, B.N., eds), pp. 471–480. New Delhi, India: I.K. International Publishing House.
- Patil, B.L. and Dasgupta, I. (2006) Defective-interfering DNAs of plant viruses. *CRC Crit. Rev. Plant Sci.* **25**, 47–64.
- Patil, B.L., Rajasubramaniam, S., Bagchi, C. and Dasgupta, I. (2005) Both Indian cassava mosaic virus and Sri Lankan cassava mosaic virus are found in India and exhibit high variability as assessed by PCR-RFLP. *Arch. Virol.* **150**, 389–397.
- Patil, B.L., Dutt, N., Briddon, R.W., Bull, S.E., Rothenstein, D., Borah, B.K., Dasgupta, I., Stanley, J. and Jeske, H. (2007) Deletion and recombination events between the DNA-A and DNA-B components of Indian cassava-infecting geminiviruses generate defective molecules in *Nicotiana benthamiana*. *Virus Res.* **124**, 59–67.

- Pita, J.S., Fondong, V.N., Sangaré, A., Kokora, R.N.N. and Fauquet, C.M. (2001a) Genomic and biological diversity of the African cassava geminiviruses. *Euphytica*, **120**, 115–125.
- Pita, J.S., Fondong, V.N., Sangaré, A., Otim-Nape, G.W., Ogwal, S. and Fauquet, C.M. (2001b) Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. Gen. Virol.* **82**, 655–665.
- Roberts, S. and Stanley, J. (1994) Lethal mutations within the conserved stem-loop of African cassava mosaic virus DNA are rapidly corrected by genomic recombination. *J. Gen. Virol.* **75**, 3203–3209.
- 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.
- Rothenstein, D., Haible, D., Dasgupta, I., Dutt, N., Patil, B.L. and Jeske, H. (2006) Biodiversity and recombination of cassava-infecting begomoviruses from southern India. *Arch. Virol.* **151**, 55–69.
- Rothenstein, D., Krenz, B., Selchow, O. and Jeske, H. (2007) Tissue and cell tropism of Indian cassava mosaic virus (ICMV) and its AV2 (precoat) gene product. *Virology*, **359**, 137–145.
- Saunders, K. and Stanley, J. (1995) Complementation of African cassava mosaic virus AC2 gene function in a mixed bipartite geminivirus infection. *J. Gen. Virol.* **76**, 2287–2292.
- 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.
- 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.
- Shivaprasad, P.V., Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T. and Pooggin, M.M. (2005) Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. *J. Virol.* **79**, 8149–8163.
- Sserubombwe, W.S., Briddon, R.W., Baguma, Y.K., Ssemakula, G.N., Bull, S.E., Bua, A., Alicai, T., Omongo, C., Otim-Nape, G.W. and Stanley, J. (2008) Diversity of begomoviruses associated with mosaic disease of cultivated cassava (*Manihot esculenta* Crantz) and its wild relative (*Manihot glaziovii* Mull. Arg.) in Uganda. *J. Gen. Virol.* **89**, 1759–1769.
- Sseruwagi, P., Legg, J.P., Maruthi, M.N., Colvin, J., Rey, M.E.C. and Brown, J.K. (2005) Genetic diversity of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) populations and presence of the B biotype and a non-B biotype that can induce silverleaf symptoms in squash, in Uganda. *Ann. Appl. Biol.* **147**, 253–265.
- Stanley, J. and Gay, M.R. (1983) Nucleotide sequence of cassava latent virus DNA. *Nature*, **301**, 260–262.
- Stanley, J. and Townsend, R. (1986) Infectious mutants of cassava latent virus generated in vivo from intact recombinant DNA clones containing single copies of the genome. *Nucleic Acids Res.* **14**, 5981.
- Stanley, J., Townsend, R. and Curson, S.J. (1985) Pseudorecombinants between cloned DNAs of two isolates of cassava latent virus. *J. Gen. Virol.* **66**, 1055–1061.
- Stanley, J., Saunders, K., Pinner, M.S. and Wong, S.M. (1997) Novel defective interfering DNAs associated with *ageratum* yellow vein geminivirus infection of *Ageratum conyzoides*. *Virology*, **239**, 87–96.
- Sudarshana, M.R., Roy, G. and Falk, B.W. (2007) Methods for engineering resistance to plant viruses. *Methods Mol. Biol.* **354**, 183–195.
- Sunter, G. and Bisaro, D.M. (1992) Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. *Plant Cell*, **4**, 1321–1331.
- Takenaka, K., Koshino-Kimura, Y., Aoyama, Y. and Sera, T. (2007) Inhibition of tomato yellow leaf curl virus replication by artificial zinc-finger proteins. *Nucleic Acids Symp. Ser. (Oxf)* **51**, 429–430.
- Taylor, N.J., Mason, M.V., Ogbe, F. and Fauquet, C.M. (2003) Cassava: an ancient crop for modern times [Compact Disc 3]. In: *Proceedings of the 5th International Meeting of the Cassava Biotechnology Network, 4–9 November, 2001, St. Louis, MO* (Fauquet, C.M. and Taylor, N.J., eds). St. Louis, Missouri, USA: Donald Danforth Plant Science Center.
- Thresh, J.M. (2006) Control of tropical plant virus diseases. *Adv. Virus Res.* **67**, 245–295.
- Thresh, J.M. and Cooter, R.J. (2005) Strategies for controlling cassava mosaic virus disease in Africa. *Plant Pathol.* **54**, 587–614.
- 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.
- Vanderschuren, H., Akbergenov, R., Pooggin, M.M., Hohn, T., Grisse, W. and Zhang, P. (2007a) Transgenic cassava resistance to African cassava mosaic virus is enhanced by viral DNA-A bidirectional promoter-derived siRNAs. *Plant Mol. Biol.* **64**, 549–557.
- Vanderschuren, H., Stupak, M., Futterer, J., Grisse, W. and Zhang, P. (2007b) Engineering resistance to geminiviruses-review and perspectives. *Plant Biotechnol. J.* **5**, 207–220.
- 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.
- Vanitharani, R., Chellappan, P. and Fauquet, C.M. (2005) Geminiviruses and RNA silencing. *Trends Plant Sci.* **10**, 144–151.
- Varsani, A., Shepherd, D.N., Monjane, A.L., Owor, B.E., Erdmann, J.B., Rybicki, E.P., Peterschmitt, M., Briddon, R.W., Markham, P.G., Oluwafemi, S., Windram, O.P., Lefevre, P., Lett, J.M. and Martin, D.P. (2008) Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *J. Gen. Virol.* **89**, 2063–2074.
- Voinnet, O., Pinto, Y.M. and Baulcombe, D.C. (1999) Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. USA*, **96**, 14 147–14 152.
- Wang, H., Hao, L., Shung, C.Y., Sunter, G. and Bisaro, D.M. (2003) Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. *Plant Cell*, **15**, 3020–3032.
- Wang, H., Buckley, K.J., Yang, X., Buchmann, R.C. and Bisaro, D.M. (2005) Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *J. Virol.* **79**, 7410–7418.
- Warburg, O. (1894) Die kulturpflanzen usambaras. *Mitt. Dtsch. Schutzgeb.* **7**, 131.
- Waterhouse, P.M., Wang, M. and Lough, T. (2001) Gene silencing as an adaptive defence against viruses. *Nature (London)*, **411**, 834–842.

- Zhan, X.C., Haley, A., Richardson, K. and Morris, B. (1991) Analysis of the potential promoter sequences of African cassava mosaic virus by transient expression of the beta-glucuronidase gene. *J. Gen. Virol.* **72**, 2849–2852.
- Zhang, P., Futterer, J., Frey, P., Potrykus, I., Puonti-Kaerlas, J. and Grissem, W. (2003) Engineering virus-induced African cassava mosaic virus resistance by mimicking a hypersensitive reaction in transgenic cassava. In: *Plant Biotechnology 2002 and Beyond* (Vasil, I.K. ed.), pp. 143–145. Dordrecht: Kluwer Academic Publishers.
- Zhang, P., Vanderschuren, H., Futterer, J. and Grissem, W. (2005) Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *Plant Biotechnol. J.* **3**, 385–397.
- Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G.W., Robinson, D.J. and Harrison, B.D. (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* **78**, 2101–2111.
- Zimmermann, A. (1906) Die Krauselerkrankheit des Maniok. *Pflanzer*, **2**, 145.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Linear recombination pattern along the Rift Valley in East Africa, from Uganda (left) to South Africa (right), for the cassava

mosaic geminivirus (CMG) DNA-A components, showing a gradient between East African cassava mosaic virus (EACMV)-type (purple colour) and South African cassava mosaic virus (SACMV)-type (green colour) sequences. The virus name abbreviations are given in the text.

Fig. S2 Example of cassava mosaic geminivirus (CMG)-resistant plants produced by genetic engineering. (A) Control 60444 cassava, not infected. (B) Control 60444 cassava, infected by East African cassava mosaic virus—Uganda (EACMV-UG). (C) Control 60444 cassava, infected by African cassava mosaic virus (ACMV). (D) Transgenic cassava plant resistant to ACMV via the gene silencing strategy. (E) Transgenic cassava plant resistant to EACMV-UG via the gene silencing strategy. (F) Transgenic cassava plant resistant to ACMV via the G5 strategy.

Table S1 List of full-length clones of cassava mosaic geminiviruses (CMGs) isolated from Africa and the Indian subcontinent. The species name, isolate name and their abbreviations, and the accession numbers for DNA-A and DNA-B, are provided.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.