SARAH JACQUELINE CAVALCANTI DA SILVA

DIVERSIDADE E ESTRUTURA GENÉTICA DE BEGOMOVÍRUS QUE INFECTAM PLANTAS DANINHAS NO NORDESTE BRASILEIRO

RECIFE, PE, BRASIL
FEVEREIRO, 2011

SARAH JACQUELINE CAVALCANTI DA SILVA

DIVERSIDADE E ESTRUTURA GENÉTICA DE BEGOMOVÍRUS QUE INFECTAM PLANTAS DANINHAS NO NORDESTE BRASILEIRO

Tese apresentada ao Programa de Pós-Graduação em Fitopatologia da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do grau de Doutor em Fitopatologia.

RECIFE-PE, BRASIL
FEVEREIRO, 2011

Ficha Catalográfica

S586d Silva, Sarah Jacqueline Cavalcanti da
Diversidade e estrutura genética de Begomovírus que infectam plantas daninhas no nordeste brasileiro / Sarah Jacqueline Cavalcanti da Silva. – 2011.
125 f. : il.

Orientador: Gilvan Pio-Ribeiro.
Tese (Doutorado em Fitopatologia) – Universidade Federal Rural de Pernambuco, Departamento de Agronomia, Recife, 2011.
Inclui apêndice e referências.

1. Begomovírus 2. Virologia vegetal 3. Plantas daninhas 4. Diversidade genética I. Pio-Ribeiro, Gilvan, Orientador II. Título

CDD 632.7

DIVERSIDADE E ESTRUTURA GENÉTICA DE BEGOMOVÍRUS QUE INFECTAM PLANTAS DANINHAS NO NORDESTE BRASILEIRO

SARAH JACQUELINE CAVALCANTI DA SILVA

COMITÊ DE ORIENTAÇÃO

Prof°. Dr. Gilvan Pio Ribeiro - UFRPE - Orientador
Prof°. Dr. Gaus Silvestre de Andrade Lima – UFAL – Co-orientador
Prof°. Dr. Francisco Murilo Zerbini (UFV) – Co-orientador

RECIFE – PE, BRASIL
FEVEREIRO, 2011

DIVERSIDADE E ESTRUTURA GENÉTICA DE BEGOMOVÍRUS QUE INFECTAM PLANTAS DANINHAS NO NORDESTE BRASILEIRO

SARAH JACQUELINE CAVALCANTI DA SILVA

Tese defendida e aprovada pela Banca Examinadora em 25 de fevereiro de 2011.

ORIENTADOR:

____

Profº. Dr. Gilvan Pio Ribeiro
(UFRPE)

EXAMINADORES:

____

Profº. Dr. Gaus Silvestre de Andrade Lima
(UFAL)

____

Profº. Dr. Francisco Murilo Zerbini
(UFV)

____

Dra. Genira Pereira de Andrade
(UFRPE)

____

Profª. Dra. Elvira Maria Régis Pedrosa
(UFRPE)

RECIFE - PE, BRASIL
FEVEREIRO – 2011

Aos meus pais, Egildo e Ivanete, e meus irmãos, Soraia e Eric, pelo amor, carinho, confiança e incentivo durante todos os momentos de minha vida. As minhas madrinhas Terezinha (Lêu) e Maria de Lourdes (Lurdinha),

Com carinho

DEDICO

AGRADECIMENTOS

A Deus por me dar força e sustentação durante as adversidades encontradas durante a realização deste trabalho;

À Universidade Federal Rural de Pernambuco e a Universidade Federal de Viçosa, pelo apoio institucional;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudo.

Aos Professores Iraíldes Pereira Assunção e Gaus Silvestre Andrade de Lima pela orientação, confiança, amizade e ensinamentos passados durante todo esse período. Sem eles nada disso seria possível;

Ao Profº. Francisco Murilo Zerbini por todos os ensinamentos, apoio e por acreditar em minha capacidade de realizar este trabalho (serei sempre grata);

Ao Profº. Gilvan Pio Ribeiro pelo conhecimento transmitido, amizade, apoio e dedicação;

À Dra. Poliane Alfenas-Zerbini pela amizade, apoio e conhecimentos a mim transmitidos;

Às minhas amigas irmãs, Erica Graciano, Kamila Câmara, Kátia Félix, Kirley Marques, Isadora França, Joyce Lima pelo amor incondicional que temos umas as outras;

A Braz Tavares da Hora Júnior pelos conhecimentos transmitidos, assistência nas análises dos dados, e pela amizade e apoio em todos os momentos da minha passagem em Viçosa (serei sempre grata);

A Robson José do Nascimento pela amizade e por ter me recebido tão bem em sua casa;

Aos meus amigos do Laboratório de Fitovirologia da Universidade Federal Rural de Pernambuco, Ana Paula, André, Diego, Genira, Jayne, Otagibson, Paula, Roger e Zilderlânia;

Aos meus amigos do Laboratório de Fitopatologia Molecular da Universidade Federal de Alagoas, Bárbara, Edilene, Édipo, Edlaine, Érica, Humberto, Igor, Isadora, João, Liliane, Luciana, Mariote e Wellington;

Aos meus amigos do Laboratório de Virologia Vegetal Molecular da Universidade Federal de Viçosa, Ana, André, Adriana, Alison, Álvaro, Amanda, Antônio, César, Daniela, Fernanda, Francisco, Larissa, Lenin, Joaquim, Jorge, Mariana, Márcio, Mercedes, Marcos, Pedro, Renan, Roberto, Sílvia, e em especial a Carolina, Danielle, Fábio, Glória, Jorge, Joyce, Sheila e Tathiana;

A todos os professores do Curso de Pós-Graduação em Fitopatologia da Universidade Federal Rural de Pernambuco pelos ensinamentos transmitidos;

A todos os amigos da turma da Pós-Graduação em Fitopatologia da Universidade Federal Rural de Pernambuco da pela agradável convivência durante o curso;

Aos meus amigos e familiares, por terem suportado e entendido minha ausência e por acreditarem que não existe distância para uma verdadeira amizade;

A todos aqueles que de alguma forma contribuíram para realização deste trabalho;

SUMÁRIO

SUMÁRIO... viii
RESUMO... ix
ABSTRACT... xi
CAPÍTULO I - Introdução Geral... 01
Referências Bibliográficas... 34
CAPÍTULO II – Species diversity, phylogeny and genetic structure of begomovirus populations infecting leguminous weeds in Northeastern Brazil........... 58
Summary... 60
Introduction... 61
Results... 63
Discussion 66
Methods... 71
References... 73
CAPÍTULO III – Genetic structure of a begomovirus population infecting the ubiquitous weed *Cleome affinis* in Northeastern Brazil ………………………………… 95
Abstract... 97
Introduction... 98
Materials and Methods... 99
Results... 102
Discussion... 105
References... 101
CONCLUSÕES GERAIS... 125

RESUMO

A incidência e severidade de doenças causadas por begomovírus (família *Geminiviridae*) têm aumentado rapidamente em muitas áreas do mundo, incluindo o Brasil, onde são fatores limitantes à produção de feijão e tomateiro. Begomovírus são também associados a uma ampla gama de plantas daninhas e silvestres, as quais em alguns casos podem atuar como fonte de inóculo para plantas cultivadas. Acredita-se que begomovírus que infectam plantas daninhas podem ser transferidos horizontalmente para plantas cultivadas, e que no novo hospedeiro eles podem evoluir rapidamente por meio de recombinação e pseudo-recombinação, dando origem a novas espécies. Atuando como reservatórios, estas plantas podem desempenhar um importante papel nas epidemias virais em várias culturas. O estudo de epidemias de vírus de plantas é grandemente facilitado quando uma abordagem baseada em genética de populações é empregada. O primeiro passo para estudar populações virais é definir sua estrutura genética, o que se refere ao seu grau de variabilidade genética. O conhecimento da dinâmica da variabilidade genética é essencial para entender o potencial das populações para evoluir, o que afeta diretamente a durabilidade de estratégias de manejo da doença baseadas na resistência do hospedeiro. Estudos para entender a estrutura genética e dinâmica de populações de begomovírus em plantas daninhas e possíveis efeitos sobre epidemias em espécies cultivadas são escassos. Dessa forma, o objetivo desse estudo foi determinar a diversidade e estrutura genética de begomovírus que infectam plantas daninhas no Nordeste do Brasil, como passo para avaliar seu papel como reservatório de begomovírus. Plantas daninhas pertencentes às famílias Fabaceae e Capparaceae com sintomas típicos de infecção por begomovírus foram coletadas nos estados de Alagoas (AL), Bahia (BA), Paraíba (PB), Pernambuco (PE) e Sergipe (SE) de maio de 2005 a julho de 2010. Um total de 59 amostras de fabáceas, incluindo 42 amostras de *Macroptilium* spp., e 23 amostras de *Cleome affinis* (fam. Capparaceae) foram coletadas. DNA total foi extraído a partir das amostras e genomas completos dos begomovírus foram amplificados e clonados por amplificação por círculo rolante. Os clones foram completamente sequenciados e as sequências foram usadas para comparações com begomovírus previamente descritos, para análise filogenética e para determinação da estrutura genética das populações virais.

Comparações de sequências indicaram a presença de seis begomovírus em fabáceas (cinco em *Macroptilium* spp.), incluindo quatro representando novas espécies. As características das sequências indicam que todas as novas espécies são begomovírus bissegmentados típicos do Novo Mundo que agruparam com begomovírus brasileiros na árvore filogenética. Em contraste, apenas uma espécie de begomovírus foi encontrada infectando plantas de *Cleome affinis*, sugerindo um baixo grau de diversidade de espécies nessa hospedeira. Filogenia reticulada foi usada para detectar possíveis eventos de recombinação nas populações begomovírus em fabáceas e em *C. affinis*. Esses prováveis eventos de recombinação foram confirmados por análise no programa RDP3. Foram detectados eventos de recombinação ocorrendo naturalmente nas populações de Macroptilium yellow spot virus (MaYSV) e Cleome leaf crumple virus (ClLCrV). A análise da estrutura genética das populações de MaYSV e ClLCrV indica um alto grau de variabilidade genética em ambos os casos. Mutação e recombinação são importantes processos envolvidos na alta variabilidade genética encontrada nas populações desses vírus. Em conjunto, os resultados sugerem que *Macroptilium* spp. e *Cleome affinis* podem constituir importantes reservatórios de begomovírus.

ABSTRACT

The incidence and severity of diseases caused by begomoviruses has increase rapidly in many areas of the world, including Brazil, where they are limiting factors to tomato and common bean production. Begomoviruses are also associated with a wide range of weed plants which in some cases act as inoculum sources for cultivated plants. It is believed that begomoviruses infecting weed hosts can be horizontally transferred to crop plants and that in the new host they will rapidly evolve by recombination and pseudorecombination, giving rise to novel species. Acting as reservoirs these plants can play a relevant role in viral epidemics in several crops species. The study of plant virus epidemics is greatly facilitated when a population genetics approach is employed. The first step to study viral population is to define their genetic structure, which refers to their degree of variability. Knowledge of the dynamics of genetic variability is essential to understand the potential of the population to evolve, which directly affects the durability of disease management strategies based on the deployment of resistance genes. Studies to understand the genetic structure and dynamics of begomovirus populations in wild reservoirs and the possible effects on epidemics in crop species are scarce. Thus, the aim of this study was to determine the species diversity and population genetic structure of begomoviruses infecting weeds in Northeastern Brazil, as a step towards assessing their role as begomovirus reservoirs. Weed samples belonging to the family Fabaceae and Capparaceae displaying typical symptoms of begomovirus infection were collected in Alagoas (AL), Bahia (BA), Paraíba (PB), Pernambuco (PE) and Sergipe (SE) states from May/2005 to July/2010. A total of 59 leguminous weeds including 42 samples of *Macroptilium* spp. and 23 samples of *Cleome affinis* (fam. Caparaceae) were collected. Total DNA was extracted from the samples and full-length begomovirus genomes were amplified and cloned by rolling circle amplification. Clones were completely sequenced and the sequences were used for comparisons with previously described begomoviruses, for phylogenetic analysis and for the determination of the genetic structure of viral populations. Sequence comparisons indicated the presence of six begomoviruses in leguminous weeds (five in *Macroptilium* spp.), four of them representing novel species. Sequence features indicate that all four novel species are typical New World, bipartite begomoviruses which clustered with Brazilian begomoviruses in the phylogenetic tree.

In contrast, only one begomovirus was found infecting *C. affinis*, suggesting low species diversity in this host phylogenetic reticulate analysis was used to detected possible recombination events in begomovirus populations in leguminous weeds and *C. affinis*. Putative recombination events were confirmed by RDP3 package analysis. We detected recombination events in Macroptilium yellow spot virus (MaYSV) and Cleome leaf crumple virus (ClLCrV) populations. Analysis of the genetic structure of these virus populations indicates a high degree of genetic variability in both cases. Mutation and recombination are important processes involved in the high genetic variability found in MaYSV and ClLCrV populations. Together, these results suggest that *Macroptilium* spp. and *Cleome affinis* can be important begomovirus reservoirs.

CAPÍTULO I

Introdução Geral

INTRODUÇÃO GERAL

1. Família Geminiviridae

Os vírus pertencentes à família *Geminiviridae* apresentam genoma composto de DNA de fita simples (ssDNA) circular encapsidado em um capsídeo icosaédrico geminado. A família é dividida em quatro gêneros: *Mastrevirus*, *Curtovirus*, *Topocuvirus* e *Begomovirus*, com base no tipo de inseto vetor, gama de hospedeiros, organização genômica e relacionamento filogenético (STANLEY et al., 2005). O gênero *Mastrevirus* inclui os geminivírus com um componente genômico, transmitidos por diversas cigarrinhas (Homoptera: Cicadellidae) a plantas monocotiledôneas. A espécie-tipo é o *Maize streak virus* (MSV), um vírus economicamente importante para a cultura do milho (*Zea mays*). No gênero *Curtovirus* estão os geminivírus com um componente genômico, transmitidos por diversas cigarrinhas (Hemiptera: Cicadellidae) a plantas dicotiledôneas. O *Beet severe curly top virus* (BSCTV) é a espécie-tipo e mais importante economicamente. O gênero *Topocuvirus* possui uma única espécie, o *Tomato pseudo-curly top virus* (TPCTV), com um componente genômico, transmitida pela cigarrinha *Micrutalis malleifera* (Homoptera: Auchenorrhyncha) a plantas dicotiledôneas. O gênero *Begomovirus* engloba espécies com um ou dois componentes genômicos, transmitidas pela mosca-branca *Bemisia tabaci* (Homoptera: Aleyrodidae) a plantas dicotiledôneas (STANLEY et al., 2005). A espécie-tipo é o *Bean golden yellow mosaic virus* (BGYMV) (FAUQUET et al., 2008).

Os begomovírus do “Velho Mundo” (Europa, Ásia e África) possuem em sua maioria um componente genômico (monossegmentados), e estão frequentemente associados a moléculas de ssDNA circular conhecidas como DNA β (betassatélites) e DNA-1 (alfassatélites) (BRIDDON, 2003; BRIDDON; STANLEY, 2006). Os

betassatélites contêm uma ORF, βC1, que codifica uma proteína responsável pela indução de sintomas e que atua como supressora do silenciamento gênico pós-transcricional (CUI et al., 2004; CUI et al., 2005; BRIDDON; STANLEY, 2006). Os alfassatélites são semelhantes ao componente genômico denominado DNA-R dos nanovírus, os quais contêm uma ORF que codifica uma proteína associada à replicação (Rep), seguida de uma região rica em adenina e uma estrutura em forma de grampo que inclui a origem de replicação (IDRIS et al., 2005). Os alfassatélites podem replicar autonomamente mas, requerem um vírus auxiliar para infecção sistêmica da planta e transmissão por inseto (SAUNDERS; STANLEY, 1999; SAUNDERS et al., 2000; SAUNDERS; BEDFORD; STANLEY, 2002). Recentemente, alfassatélites foram identificados no Brasil e na Venezuela, associados aos begomovírus bissegmentados *Cleome leaf crumple virus* (ClLCrV), *Euphorbia mosaic virus 1*(EuMV) e Melon chlorotic mosaic virus (MeCMV), sendo esses os primeiros relato de alfassatélites associados a begomovírus ocorrendo naturalmente no "Novo Mundo" (Américas) (PAPROTKA; METZLER; JESKE, 2010c; ROMAY et al., 2010). Os begomovírus do "Novo Mundo" possuem dois componentes genômicos (bissegmentados), denominados DNA-A e DNA-B, cada um com aproximadamente 2600 nucleotídeos (Figura 1). Os dois componentes genômicos de uma mesma espécie viral não possuem identidade entre as suas sequências, exceto por uma região com aproximadamente 200 nucleotídeos denominada região comum (RC), que inclui a origem de replicação (HANLEY-BOWDOIN et al., 1999).

O DNA-A dos begomovírus bissegmentados pode codificar de quatro a seis proteínas: uma proteína associada à replicação, Rep ("replication-associated protein"), iniciadora do mecanismo de replicação por círculo rolante, com propriedade de ligação a ácidos nucléicos, endonuclease e ATPase (FONTES; LUCKOW; HANLEY-

BOWDOIN, 1992; OROZCO et al., 1997); uma proteína transativadora, TrAP ("trans-activating protein"), fator transcricional dos genes *CP* e *NS* e que também atua como supressora do silenciamento gênico (SUNTER; BISARO, 1992; VOINNET; PINTO; BAUCOMBE, 1999; WANG et al., 2005); a proteína Ren ("replication-enhancer protein"), fator acessório da replicação viral (SUNTER et al., 1990; PEDERSEN; HANLEY-BOWDOIN, 1994); e a proteína capsidial ("coat protein", CP), que além de formar o capsídeo viral é essencial para a transmissão do vírus pelo inseto vetor (BRIDDON et al., 1990; HÖFER et al., 1997a). O gene *AV2* ("pre-coat") está presente apenas nos begomovírus do "Velho Mundo", e atua no movimento do vírus na planta (PADIDAM; BEACHY; FAUQUET, 1996). O gene *AC4* codifica uma proteína supressora de silenciamento gênico (VANITHARANI et al., 2004). O DNA B codifica a proteína MP ("movement protein"), envolvida no movimento célula-a-célula do vírus por meio do aumento do limite de exclusão dos plasmodesmas (NOUEIRY; LUCAS; GILBERTSON, 1994), e a proteína NSP ("nuclear shuttle protein"), responsável pelo transporte do DNA através do envelope nuclear (NOUEIRY; LUCAS; GILBERTSON, 1994; SANDERFOOT; INGHAM; LAZAROWITZ, 1996).

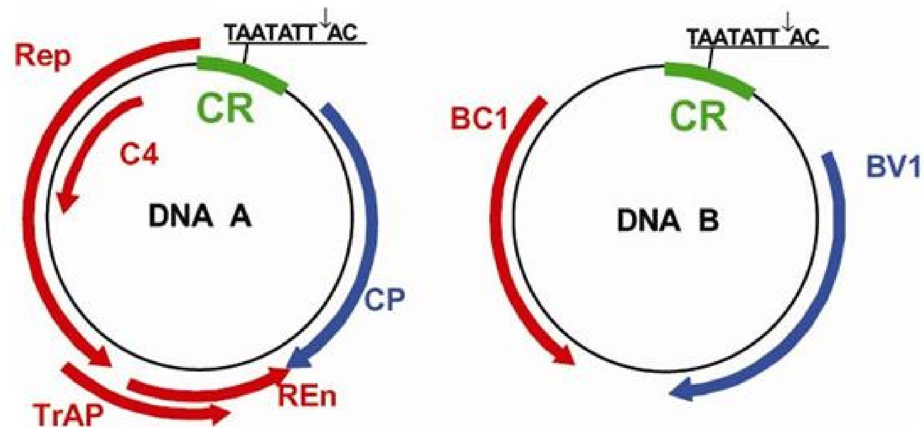


Figura 1. Representação esquemática do genoma do *Bean golden yellow mosaic virus* (BGYMV), espécie-tipo do gênero *Begomovirus.* Os círculos representam o genoma viral, com dois componentes (DNA-A e DNA-B) de aproximadamente 2.600 nucleotídeos cada. Uma sequência de aproximadamente 200 nucleotídeos, denominada região comum (CR), contém a origem de replicação viral, com uma estrutura em forma de grampo e uma sequência invariável de nove nucleotídeos (TAATATT↓AC), conservada em todos os membros da família *Geminiviridae*. A seta (↓) indica o sítio de início da replicação do DNA viral por círculo rolante. As setas azuis e vermelhas indicam os genes virais e a direção em que ocorre a transcrição (viral e complementar, respectivamente). Reproduzido de (GUTIERREZ et al., 2004).

2. Replicação viral

No processo de infecção dos geminivírus, as partículas virais são inoculadas na planta pelo inseto vetor e o genoma viral (ssDNA) se desassocia de forma espontânea do capsídeo (LAZAROWITZ, 1992; PALMER; RYBICKI, 1998). No interior da célula o ssDNA viral é transportado para o núcleo, onde é convertido em um intermediário de fita dupla (dsDNA) denominado forma replicativa (RF). A maneira como esta conversão ocorre não é conhecida, no entanto evidências indiretas, como a necessidade de desestabilização local do dsDNA para o iniciação da replicação por círculo rolante em procariotos por "strand-nicking enzimes" indicam que é realizada por fatores do hospedeiro. A RF serve como molde para síntese dos novos componentes genômicos e

também para a transcrição dos genes virais. O genoma viral é replicado via mecanismo de círculo rolante semelhante ao utilizado pelos bacteriófagos φX174 e M13, utilizando a RF como molde (STENGER et al., 1991; STANLEY, 1995).

A origem de replicação (*ori*) está localizada na região intergênica comum entre os dois componentes genômicos. A sequência da *ori* é conservada entre componentes de um mesmo vírus, porém variável entre espécies, com exceção de uma região de aproximadamente 30 nucleotídeos conservada entre todas as espécies (DAVIES et al., 1987; LAZAROWITZ, 1992). Nesta região se localiza uma sequência repetida e invertida composta predominantemente por guanina e citosina, formando uma estrutura conservada em forma de grampo ("structurally-conserved element", SCE), com uma sequência invariável (5'-TAATATTAC-3') encontrada em todos geminivírus, que constitui o domínio funcional da origem de replicação (HEYRAUD-NITSCHKE et al., 1995; OROZCO; HANLEY-BOWDOIN, 1998). É nesse nonanucleotídeo que ocorre a clivagem (TAATATT↓AC) que inicia o processo de replicação por círculo rolante (FONTES et al., 1994; LAUFS et al., 1995). A clivagem é realizada pela proteína Rep, que atua como endonuclease sítio-específica com requerimento de estrutura e sequência (LAUFS et al., 1995; OROZCO; HANLEY-BOWDOIN, 1998). Na região comum encontram-se as sequências específicas para ligação da proteína Rep (FONTES; LUCKOW; HANLEY-BOWDOIN, 1992; FONTES et al., 1994) e regiões promotoras da RNA polimerase tipo II de plantas, responsável pela transcrição dos genes virais (HANLEY-BOWDOIN et al., 1999).

O sítio de ligação de REP ao DNA viral está localizado entre a caixa TATA do gene *Rep* e a SCE (OROZCO; HANLEY-BOWDOIN, 1998), sendo constituído por duas sequências em repetição direta e pelo menos uma repetição invertida denominadas "iterons" (ARGÜELLO-ASTORGA et al., 1994). A ligação de Rep aos iterons é

essencial para o início da replicação. Após a ligação de Rep ao DNA viral e estabilização do complexo formado por Rep, Ren e fatores do hospedeiro, a proteína Rep cliva o nonanucleotídeo localizado na SCE, dando início à replicação por círculo rolante (GUTIERREZ, 1999). O reconhecimento pela proteína Rep é considerado vírus-específico (ARGÜELLO-ASTORGA et al., 1994; HARRISON; ROBINSON, 1999; RAMOS et al., 2003), de modo que só inicia a replicação de DNAs cognatos. O domínio funcional de Rep foi mapeado na sua região N-terminal, e este inclui o domínio de ligação a DNA, conservado em todas as proteínas Rep (JUPIN, 1995; GLADFELTER et al., 1997; CHATTERJI et al., 1999). Uma vez que o reconhecimento e ligação aos iterons por Rep é específico, foi proposto que esta depende da sequência de nucleotídeos dos iterons e dos aminoácidos de um motivo conservado na proteína Rep denominado domínio relacionado aos iterons ("iteron-related domain", IRD) (ARGUELLO-ASTORGA; RUIZ-MEDRANO, 2001). Porém, a replicação do DNA-B do *Tomato yellow spot virus* (ToYSV) pela Rep do *Tomato golden mosaic virus* (TGMV) indica que a interação entre os aminoácidos do IRD e os iterons não é a única forma de reconhecimento da origem de replicação, uma vez que tanto os iterons quanto os IRDs são diferentes entre esses dois vírus (ANDRADE et al., 2006b). Além disso, a ausência de iterons nos DNAs satélites associados a begomovírus é uma evidência adicional de que outros fatores afetam o reconhecimento da origem de replicação pela proteína Rep (LIN et al., 2003; STANLEY, 2004).

3. Movimento do vírus na planta

O movimento do vírus no interior do hospedeiro pode ser dividido em dois processos: movimento célula-a-célula via plasmodesmas, e movimento a longa distância, no qual o vírus atinge o sistema vascular e é transportado sistemicamente para

toda a planta. Para esse fim, a partir do DNA-B dos begomovírus bissegmentados são codificadas duas proteínas relacionadas ao movimento viral, NSP e MP. Como os begomovírus replicam no núcleo da célula hospedeira, necessitam de uma etapa adicional de transporte do núcleo para o citoplasma, a qual é realizada pela proteína NSP (PALMER; RYBICKI, 1998). Já a proteína MP associa-se à membrana celular e altera o limite de exclusão dos plasmodesmas, viabilizando o transporte do genoma viral (NOUEIRY; LUCAS; GILBERTSON, 1994). Estas duas proteínas atuam de forma cooperativa para mediar o tráfego intra- e intercelular do DNA viral (SANDERFOOT; LAZAROWITZ, 1995), permitindo ao vírus infectar sistemicamente o hospedeiro.

Os estudos sobre o movimento viral na planta tem como base a interação física entre as proteínas de movimento MP e NSP (ROJAS et al., 2005b). A interação direta das proteínas MP e NSP *in vitro* foi demonstrada para o TGMV, utilizando o sistema duplo-híbrido de levedura (MARIANO et al., 2004). A interação *in vivo* entre NSP e MP do *Cabbage leaf curl virus* (CaLCuV) foi recentemente demonstrada, também utilizando-se o sistema duplo- híbrido levedura. Nestes estudos foi identificada uma GTPase citoplasmática designada NIG (NSP-interacting GTPase), que interage com NSP de begomovírus *in vitro* e *in vivo* e promove o transporte da proteína viral do núcleo para o citoplasma, onde ela é redirecionada para a superfície da célula para interagir com MP (CARVALHO et al., 2008).

Dois modelos tem sido propostos para explicar o movimento intracelular de begomovírus (LEVY; TZFIRA, 2010). No primeiro modelo, denominado "couple-skating" (KLEINOW et al., 2008), NSP transporta ssDNA ou dsDNA do núcleo para a periferia da célula e, no citoplasma, MP atua nos plasmodesmas para facilitar o movimento célula-a-célula do complexo NSP-DNA (SANDERFOOT; LAZAROWITZ, 1995; FRISCHMUTH et al., 2004; 2007; KLEINOW et al., 2008). No segundo modelo,

denominado “relay-race”, NSP inicialmente transporta o dsDNA do núcleo para o citoplasma. No citoplasma, o dsDNA se associa a MP, e o complexo MP-dsDNA se movimenta célula-a-célula através dos plasmodesmas (NOUEIRY; LUCAS; GILBERTSON, 1994; ROJAS et al., 1998).

Seguindo o movimento célula-a-célula, o vírus atinge os plasmodesmas associados ao tecido vascular e então inicia-se o movimento a longa distância. O movimento viral a longa distância é passivo, acompanhando o fluxo de fotoassimilados dos tecidos fonte para os tecidos dreno através do sistema vascular. A grande maioria dos vírus é transportada via floema na forma de partícula completa, atingindo, a partir do ponto de penetração, primeiramente as raízes, em seguida as folhas jovens e posteriormente a planta toda, estabelecendo uma infecção sistêmica (JEFFREY; POOMA; PETTY, 1996).

Para mastrevírus, curtovírus e begomovírus monossegmentados, a proteína CP é necessária para os movimentos célula-a-célula e a longa distância (ROJAS et al., 2001; GAFNI; EPEL, 2002). Além da CP, as proteínas V1 e C4 também são necessárias para o movimento de begomovírus monossegmentados. No caso de *Tomato yellow leaf curl virus* (TYLCV), a CP é responsável pelo transporte do DNA do núcleo para o citoplasma, funcionando como uma proteína análoga a NSP dos begomovírus bissegmentados, e o movimento célula-a-célula através do plasmodesma é mediado pelas proteínas C4 e/ou V1 (ROJAS et al., 2001; 2005b). Recentemente, foi demonstrado que a proteína C4 do curtovírus *Beet severe curly top virus* (BSCVT) é capaz de se ligar de forma não específica a ssDNA e a dsDNA, é essencial para o desenvolvimento de sintomas, e quando expressa em plantas infectadas com mutantes deficientes para C4 pode complementar *in trans* o movimento sistêmico. Em conjunto,

esses dados sugerem o envolvimento de C4 no movimento desse vírus (CHEN et al., 2010).

A proteína CP é dispensável para o estabelecimento da infecção sistêmica de begomovírus bissegmentados na maioria dos casos já estudados (ROJAS et al., 2005a). Tanto MP quanto NSP reconhecem o DNA viral de maneira específica com relação à forma e comprimento (ROJAS et al., 1998; GILBERTSON et al., 2003), o que elimina a necessidade da proteína capsidial para o movimento a longa distância. Raras exceções, como o begomovírus bissegmentado *Tomato chlorotic mottle virus* (ToCMoV), são capazes inclusive de infectar sistemicamente alguns hospedeiros na ausência do DNA-B cognato (GALVÃO et al., 2003; FONTENELLE et al., 2007).

4. Evolução dos geminivírus

Os geminivírus podem ter evoluído a partir de um replicon primitivo de DNA extracromossomal, presente em procariotos ou em ancestrais primitivos das plantas (ROJAS et al., 2005b). Evidências indiretas, como características conservadas com as proteínas iniciadoras da replicação de replicons de procariotos e eucariotos contemporâneos (ILYINA; KOONIN, 1992; CAMPOS-OLIVAS et al., 2002), presença de mRNAs policistrônicos, e a capacidade dos geminivírus de replicarem em *Agrobacterium tumefaciens* (RIDGEN et al., 1996; SELTH; RANDLES; REZAIAN, 2002), apóiam esta hipótese. Durante a co-evolução com seus hospedeiros, estes replicons de DNA teriam adquirido novos genes por meio de recombinação com o DNA do hospedeiro ou com outros replicons revisado por ROJAS et al., 2005b.

Estudos filogenéticos propõem que os geminivírus são derivados de um ancestral comum que possuía apenas um componente, infectava monocotiledôneas e era transmitido por cigarrinhas (RYBICKI, 1994; ROJAS et al., 2005b). Comparações de

sequências de espécies do gênero *Mastrevirus* com aquelas dos gêneros *Curtovirus* e *Begomovirus* demonstraram que as primeiras são mais divergentes entre si, o que sugere que os mastrevírus evoluíram por um período de tempo mais longo. O processo evolutivo levou à capacidade de infecção de plantas dicotiledôneas e em seguida à transmissão pela mosca-branca, uma vez que existem mastrevírus (transmitidos por cigarrinhas) que infectam dicotiledôneas, mas até o presente não foram encontrados geminivírus transmitidos por mosca-branca que infectem monocotiledôneas. Esse ancestral dos begomovírus modernos possuía apenas um componente. A aquisição do segundo componente teria ocorrido antes da separação dos continentes, uma vez que os begomovírus bissegmentados são encontrados tanto no "Velho Mundo" como no "Novo Mundo". Análises filogenéticas dos betassatélites e seus begomovírus associados sugerem que o satélite e o vírus auxiliar co-evoluíram como consequência do isolamento geográfico e adaptação ao hospedeiro (ZHOU et al., 2003; ROJAS et al., 2005b). Membros do gênero *Curtovirus* seriam derivados de antigas recombinações entre mastrevírus e begomovírus, resultando na aquisição da CP de um mastrevírus ancestral transmitido por uma cigarrinha primitiva, ao passo que um begomovírus teria contribuído com os genes associados à replicação (RYBICKI, 1994; PADIDAM; BEACHY; FAUQUET, 1995). Outro evento de recombinação foi identificado para o TPCTV, o único membro do gênero *Topocuvirus,* que teria surgido após recombinação entre um curtovírus ancestral e um vírus que não possui semelhança com nenhum outro geminivírus, o que sugere que outros geminivírus, não relacionados com vírus pertencentes aos quatro gêneros atualmente reconhecidos, podem estar presentes no campo (BRIDDON et al., 1996). De fato, tais vírus têm sido recentemente identificados e caracterizados (YAZDI; HEYDARNEJAD; MASSUMI, 2008; VARSANI et al., 2009; BRIDDON et al., 2010a).

Atualmente, com base em análises filogenéticas do componente DNA-A de 212 espécies, os begomovírus estão classificados em sete diferentes grupos de acordo com sua origem geográfica ou planta hospedeira (PADIDAM; BEACHY; FAUQUET, 1995; FAUQUET et al., 2008). Os begomovírus do "Velho Mundo" segregam em grupos originados na África, Índia, Ásia e Japão. Entretanto, um número crescente de vírus, os quais são referidos como "outsiders", não se encaixa nesses grupos baseados em região geográfica ou hospedeira. Esses vírus são originários da Indochina, Indonésia e Austrália. Begomovírus do "Novo Mundo" formam grupos de acordo com a origem (América Central ou do Sul). Duas espécies originárias do Vietnam isoladas de *Corchorus* sp. são relacionadas aos begomovírus do "Novo Mundo", e formam um grupo referido como "corchovirus" (HA et al., 2006; 2008). Dois grupos de vírus, um infectando leguminosas originárias da Índia e Sudeste da Ásia ("legumovirus") e outro composto de vírus isolados de *Ipomoea* spp., particularmente batata-doce (*I. batatas*) originários da América, Ásia e Europa ("sweepovirus"), são distintos e basais a todos os demais begomovírus. Esta posição anômala desses begomovírus reflete uma história evolutiva distinta. Para os "legumovirus" foi sugerido que isto seja devido ao isolamento genético de suas espécies hospedeiras (QAZI et al., 2007).

Um cenário atual para a evolução da família *Geminiviridae* foi proposto por (NAWAZ-UL-REHMAN; FAUQUET, 2009). Nesse, plasmídeos que replicam em algas vermelhas e outras formas de vida mais primitivas conseguiram adquirir novos genes, tornando-se mais independentes de seu hospedeiro e assim capazes de infectar plantas, provavelmente em primeiro lugar monocotiledôneas, como um pré-mastrevírus. Esta evolução deve ter coincidido com a aquisição da transmissão por insetos. Em algum momento eles passaram a infectar dicotiledôneas, mas ainda tinham o mesmo tipo de vetor, as cigarrinhas. Com a aquisição de novos genes tornou-se um pré-

monossegmentado, transmitido pela mosca-branca. Esse begomovírus monossegmentado teve a capacidade de capturar outras moléculas, adquirindo então um alfassatélite a partir de um pré-nanovírus ou betassatélite de uma fonte desconhecida. Por recombinação entre um begomovírus monossegmentado que infecta dicotiledôneas e um mastrevírus foram formados híbridos que deram origem aos ancestrais dos curtovírus e topocuvírus. Em um período posterior, um monossegmentado conseguiu capturar um ancestral do que hoje é o componente B, e esta combinação de dois componentes foi extremamente bem sucedida ao ponto de begomovírus bissegmentados serem os únicos presentes no continente americano, seguindo a deriva dos continentes que aconteceu há cerca de 125 milhões de anos atrás (NAWAZ-UL-REHMAN; FAUQUET, 2009).

Briddon et al. (2010b) demonstraram por meio de análises filogéticas e exaustivas comparações duas a duas dos componente DNA-A e DNA-B de begomovírus, que estas moléculas de fato tem histórias evolutivas diferentes. O DNA-B apresenta grande variação genética quando comparado ao DNA-A. Esse fato pode ser atribuído à menor quantidade de funções codificadas pelo DNA-B, sendo assim mais permissivo à variação, evoluindo exclusivamente em resposta ao hospedeiro (o DNA-A deve manter a interação com o vetor). Uma explicação adicional é que o DNA-B teria uma origem distinta do DNA-A, surgido inicialmente como um satélite que foi capturado pelo seu progenitor monossegmentado e posteriormente evoluído para se tornar parte integral do genoma. A situação atual com satélites associados aos begomovírus gera algumas pistas (por exemplo, a capacidade de infectar com eficiência algumas hospedeiras e fornecimento de função adicional de movimento) para os processos e pressões de seleção que devem ter levado à "domesticação" de um

progenitor selvagem do DNA-B (NAWAZ-UL-REHMAN; FAUQUET, 2009; NAWAZ-UL-REHMAN et al., 2009; BRIDDON et al., 2010b).

5. Variabilidade e estrutura genética de populações de begomovírus

Populações de geminivírus, incluindo os begomovírus, possuem um elevado grau de diversidade genética. A ocorrência de eventos frequentes de recombinação (PADIDAM; SAWYER; FAUQUET, 1999b), a ocorrência de pseudo-recombinação entre vírus com genoma bissegmentado (ANDRADE et al., 2006a), e a alta taxa de mutação (DUFFY; HOLMES, 2008; 2009) contribuem para esse elevado grau de diversidade. Mutação, recombinação e pseudo-recombinação são as principais fontes de variabilidade genética de vírus em plantas (GARCÍA-ARENAL; FRAILE; MALPICA, 2003; SEAL; VAN DEN BOSCH; JEGER, 2006b).

5.1. Mutação

Assim como para todos os vírus, a evolução dos geminivírus depende primariamente de mutações. Há evidências de que a rápida evolução dos geminivírus é, ao menos em parte, dirigida por processos mutacionais que agem especificamente sobre ssDNA (HARKINS et al., 2009). O impacto das mutações pontuais tem sido estudado nesse grupo de vírus. Sob diferentes condições de seleção, como presença de um efeito gargalo (população inicial pequena do vírus, período curto de aquisição pelo vetor), transferências sucessivas entre hospedeiros sem emprego do vetor, e inoculação em plantas resistentes, isolados de MSV apresentaram alta frequência de mutação, da ordem de 10^{-4} e 10^{-5} (ISNARD et al., 1998). Resultados similares foram obtidos num experimento controlado de análise da taxa de variabilidade genética do begomovírus *Tomato yellow leaf curl China virus* (TYLCCNV) onde foi encontrada uma frequência

média de mutação de $3{,}5x10^{-4}$ e $5{,}3x10^{-4}$ após 60 dias de infecção em *N. benthamiana* e tomateiro (*Solanum lycopersicon*), respectivamente (GE et al., 2007). Uma série de experimentos de evolução a longo prazo (de 6 a 32 anos) também revelaram alta frequência de mutação, entre 2 e $3x10^{-4}$, para MSV e *Sugarcane streak Réunion virus* (SSRV), sugerindo que mastrevírus provavelmente não co-divergem com seus hospedeiros (HARKINS et al., 2009). Estes resultados discordam com a hipótese de aparente co-divergência entre alguns mastrevírus e seus hospedeiros, o que implicaria em taxas de substituições de apenas 10^{-8} subs/sítio/ano na natureza (WU et al., 2008).

Duffy; Holmes (2008; 2009) realizaram análises estruturadas no tempo de isolados de TYLCV e *East African cassava mosaic virus* (EACMV), para estimar a taxa de evolução dessas espécies de begomovírus na natureza. Taxas de mutação para o TYLCV foram estimadas em $2{,}88x10^{-4}$ subs/sítio/ano para o genoma completo (DUFFY; HOLMES, 2008). A região que codifica a proteína CP apresentou uma taxa maior ($4{,}63x10^{-4}$ subs/sítio/ano) e a região intergênica (não codificadora) apresentou uma taxa ainda maior ($1{,}56x10^{-4}$ subs/sítio/ano). Entretanto, as mutações observadas foram na maioria sinônimas, sugerindo que as altas taxas de mutação observadas refletem mais uma rápida dinâmica mutacional do que uma frequência de evolução adaptativa. Para o EACMV as taxas foram estimadas em $1{,}6x10^{-3}$ e $1{,}33x10^{-4}$ subs/sítio/ano para o DNA-A e DNA-B, respectivamente (DUFFY; HOLMES, 2008). A região que codifica a proteína CP apresentou $1{,}37x10^{-3}$ subs/sítio/ano e a região que codifica a proteína associada à replicação mostrou $1{,}24x10^{-3}$ subs/sítio/ano. As regiões codificadoras presentes no DNA-B, ORFs BV1 e BC1, apresentaram $2{,}77x10^{-4}$ e $3{,}45x10^{-4}$, respectivamente. Contudo, os autores validaram esses altos níveis de heterogeneidade apenas para o DNA-A e a ORF AV1. Foi observado então que as taxas de evolução indicadas para essas duas espécies de begomovírus, entre 10^{-3} e 10^{-5},

corroboram em geral aquelas determinadas experimentalmente para MSV (ISNARD et al., 1998; HARKINS et al., 2009) e TYLCCNV (GE et al., 2007).

Erros de incorporação de nucleotídeos durante a replicação viral também contribuem para a diversidade genética. Estudos de bactérias e sistemas animais indicaram que a taxa de mutação dos vírus de dsDNA e ssDNA diferem significativamente (DUFFY; SHAKELTON; HOLMES, 2008). Taxas de mutação para fagos bacterianos, poliomavírus e papillomavírus, com genoma composto de dsDNA, são da ordem de 10^{-7} a 10^{-8} subs/sítio/ano (DRAKE, 1991; HOLMES, 2004; RANEY; DELONGCHAMP; VALENTINE, 2004). Em contraste, altas taxas de mutação (10^{-4}) foram relatadas para parvovírus e circovírus (vírus de ssDNA) (GALLIAN et al., 2002; BIAGINI, 2004). Semelhante aos geminivírus, os parvovírus e circovírus replicam seu genoma via mecanismo de círculo rolante, sugerindo que os altos níveis de heterogeneidade relatados para begomovírus e mastrevírus podem refletir erros de replicação (ARGUELLO-ASTORGA et al., 2004). Foi sugerido que os mecanismos de correção de erro associados à replicação de DNA em eucariotos não sejam eficientes na replicação por círculo rolante e, ou na replicação de ssDNA (VAN DER WALT et al., 2008).

Mutantes para a proteína Rep do TGMV e do CaLCuV que não permitem a interação com a proteína pRB, inoculados em protoplastos de fumo (*Nicotiana tabacum*) e em plantas *N. benthamiana*, apresentaram até 100% de frequência de reversão de mutações, evidenciando a capacidade de populações de geminivírus de evoluir rapidamente para alterar mudanças deletérias em seu genoma (ARGUELLO-ASTORGA et al., 2007).

5.2. Recombinação

Recombinação é o processo pelo qual segmentos de uma fita de DNA ou RNA tornam-se incorporados na fita de um indivíduo diferente durante o mecanismo de replicação (PADIDAM; BEACHY; FAUQUET, 1999a). A recombinação é um evento bastante comum em geminivírus (PADIDAM; BEACHY; FAUQUET, 1999a; LEFEUVRE et al., 2007b), e parece contribuir grandemente para a diversificação genética dos begomovírus, aumentando seu potencial evolutivo e adaptação local (HARRISON; ROBINSON, 1999; PADIDAM; SAWYER; FAUQUET, 1999b; BERRIE; RYBICKI; REY, 2001; MONCI et al., 2002). A elevada frequência de recombinação nesse grupo de vírus pode ser em parte explicada pela existência de uma possível estratégia de replicação dependente de recombinação (RDR) (JESKE; LUTGEMEIER; PREISS, 2001; PREISS; JESKE, 2003) em adição à replicação por círculo rolante (RCR) (SAUNDERS; BEDFORD; STANLEY, 2001), e pela ocorrência frequente de infecções mistas (TORRES-PACHECO et al., 1996; SANZ et al., 2000; PITA et al., 2001; RIBEIRO et al., 2003; GARCIA-ANDRES et al., 2006; DAVINO et al., 2009), com a evidência de infecção do mesmo núcleo da célula por mais de um begomovírus (MORILLA et al., 2004).

Eventos de recombinação têm sido diretamente implicados na emergência de novas doenças e epidemias em plantas cultivadas. Essas incluem a epidemia devastadora do mosaico da mandioca (*Manihot esculenta*), causada pelo recombinante EACMV na Uganda e países vizinhos (ZHOU et al., 1997; PITA et al., 2001); as epidemias do complexo TYLCV na Bacia Ocidental do Mediterrâneo, com o surgimento dos recombinantes *Tomato yellow leaf curl Málaga virus* (TYLCMalV) e *Tomato yellow leaf curl Axarquía virus* (TYLCAxV) nos campos de tomate na Espanha (MONCI et al., 2002; GARCIA-ANDRES et al., 2006; 2007a; 2007b); e as epidemias

de *Cotton leaf curl virus* (CLCuV) no Paquistão causadas por um complexo de espécies incluindo diversos begomovírus recombinantes (ZHOU et al., 1998; IDRIS; BROWN, 2002).

A emergência frequente de novas espécies de geminivírus devido a eventos de recombinação foi demonstrada por meio de análise de conversão gênica (PADIDAM; BEACHY; FAUQUET, 1999a). Embora na época o número de genomas completos sequenciados fosse pequeno, os autores analisaram todas as combinações dois-a-dois possíveis, e identificaram 420 fragmentos recombinantes tanto entre espécies como entre gêneros da família *Geminiviridae*.

Os mecanismos precisos que controlam a recombinação em begomovírus permanecem desconhecidos (PADIDAM; BEACHY; FAUQUET, 1999a). No entanto, é conhecido que sítios recombinantes não são uniformemente distribuídos ao longo do genoma, com a existência de sítios frequentes ("hot spots") e não-frequentes ("cold spots") (STANLEY, 1995; FAUQUET et al., 2005; GARCIA-ANDRES et al., 2007b; LEFEUVRE et al., 2007b). Análises bioinformáticas para detectar vírus recombinantes ocorrendo naturalmente revelaram que a origem de replicação viral é um sítio frequente de recombinação (GUTIERREZ, 1999; HANLEY-BOWDOIN et al., 1999). A comparação de sequência de begomovírus mono e bissegmentados depositadas no GenBank até maio de 2006 (123 e 116 sequências, respectivamente) indicou que a região do gene *Rep* que codifica a porção N-terminal da proteína Rep, assim como a região intergênica adjacente (RC), são frequentemente intercambiadas durante a replicação. Também foram identificados sítios frequentes de recombinação localizados na região intergênica entre os genes *CP* e *Ren* (LEFEUVRE et al., 2007b).

A análise comparativa da distribuição de sítios de recombinação dentro do genoma de diversas famílias de vírus de ssDNA novamente sugeriu a distribuição não

aleatória dos sítios e também uma tendência significativa para estes caírem tanto fora como na periferia dos genes. Além disso, foi particularmente observado que poucos sítios de recombinação foram encontrados dentro de genes que codificam proteínas estruturais, a exemplo da proteína capsidial (LEFEUVRE et al., 2007a). Esses resultados sugerem que a seleção natural agindo contra vírus que expressam proteínas recombinantes é o principal determinante na distribuição não aleatória dos sítios de recombinação na maioria das famílias de vírus de ssDNA.

Eventos de recombinação também têm sido relatados entre begomovírus e DNA satélites, e entre diferentes moléculas de betassatélites (BRIDDON et al., 2001; SAUNDERS; BEDFORD; STANLEY, 2001; BRIDDON et al., 2003; NAWAZ-UL-REHMAN; FAUQUET, 2009; NAWAZ-UL-REHMAN et al., 2009).

5.3. Pseudo-recombinação

A existência de dois componentes genômicos na maioria dos begomovírus promove um mecanismo alternativo, conhecido como pseudo-recombinação, pelo qual a troca de material genético pode ocorrer sem necessidade de recombinação intermolecular, ocorrendo apenas a troca de componentes genômicos entre dois vírus distintos (GILBERTSON et al., 1993b; SUNG; COUTTS, 1995; ANDRADE et al., 2006a); revisado por (ROJAS et al., 2005b). A ocorrência natural de pseudo-recombinantes no campo foi verificada no México, em tomateiros infectados pelo *Chino del tomate virus* (CdTV) (PAPLOMATAS et al., 1994).

Experimentos com pseudo-recombinação são ferramentas úteis no estudo de funções de genes e podem revelar relações filogenéticas, como é o caso da mistura de componentes genômicos do BGYMV e do *Bean golden mosaic virus* (BGMV), que possuem identidade inferior a 75% em suas sequências de nucleotídeos e não formam

pseudo-recombinantes infecciosos (GILBERTSON et al., 1993a). Por outro lado, pseudo-recombinantes formados a partir da mistura de componentes genômicos de dois isolados de BGYMV mostraram-se infecciosos. Quando inoculada, a mistura formada a partir de DNA-A do isolado da Guatemala (BGYMV-GA) e DNA-B do isolado da República Dominicana (BGYMV-DR) foi capaz de induzir os mesmos sintomas apresentados pelos parentais, enquanto o pseudo-recombinante recíproco induziu sintomas atenuados e tardios. Esses resultados demonstram que geminivírus com regiões comuns suficientemente similares podem formar pseudo-recombinantes infecciosos, mas ressaltam que frequentemente os pseudo-recombinantes recíprocos apresentam diferenças na eficiência de replicação e infecção sistêmica (FARIA et al., 1994). Esse fato foi também observado para o *African cassava mosaic virus* (ACMV) (STANLEY et al., 2005) e TGMV (VON ARNIM; STANLEY, 1992).

A especificidade da ligação da proteína Rep aos iterons é considerada a principal determinante da formação de pseudo-recombinantes viáveis entre diferentes espécies/estirpes de begomovírus (ARGÜELLO-ASTORGA et al., 1994; EAGLE; OROZCO; HANLEY-BOWDOIN, 1994; FONTES et al., 1994; CHATTERJI et al., 1999; ANDRADE et al., 2006a; BULL et al., 2007). Outro fator importante é a conservação da sequência de aminoácidos da proteína Rep, especialmente os três aminoácidos do IRD que estariam envolvidos diretamente na ligação aos iterons (ARGUELLO-ASTORGA; RUIZ-MEDRANO, 2001; RUIZ-MEDRANO; XOCONOSTRE-CAZARES; LUCAS, 2001). A viabilidade de pseudo-recombinantes indica que fatores envolvidos na replicação e movimento são intercambiáveis entre espécies altamente relacionadas, ou entre estirpes de uma mesma espécie. A assimetria entre pseudo-recombinantes recíprocos indica que a pseudo-recombinação entre

begomovírus é um fenômeno complexo que envolve interações entre fatores do vírus e do hospedeiro (HILL et al., 1998).

Embora a pseudo-recombinação seja comum entre estirpes de uma mesma espécie de begomovírus, a formação de pseudo-recombinantes viáveis entre espécies distintas é mais difícil. Um pseudo-recombinante foi obtido entre o DNA-A do *Abutilon mosaic virus* (AbMV) e o DNA-B do *Sida golden mosaic Costa Rica virus* (SiGMCRV), porém o pseudo-recombinante recíproco não foi infeccioso (HOFER et al., 1997b). Similarmente, um pseudo-recombinante viável foi formado pelo DNA-A de um isolado de *Sida golden mosaic virus* (SiGMV) de Honduras (SiGMV-[Ho_{yv}]) e o DNA-B do SiGMCRV (UNSELD et al., 2000). Entretanto, dentre os pseudo-recombinantes recíprocos formados pelo DNA-A do SiGMCRV combinado ao DNA-B de três isolados de SiGMV-[Ho_{yv}] que possuíam pequenas diferenças na composição de nucleotídeos, apenas um mostrou-se viável, porém pouco eficiente, e não foi capaz de infectar a planta a partir da qual foi originalmente isolado (UNSELD et al., 2000). Pseudo-recombinantes infecciosos entre o DNA-A do CdTV e o DNA-B do BGYMV foram formados apesar da baixa identidade da região comum (68%), porém o pseudo-recombinante recíproco não foi infeccioso quando inoculado em feijoeiro (*Phaseolus vulgaris*) (GARRIDO-RAMIREZ; SUDARSHANA; GILBERTSON, 2000).

Um pseudo-recombinantes produzido entre o DNA-A do *Tomato mottle virus* (ToMoV) e o DNA-B do *Bean dwarf mosaic virus* (BDMV), embora infeccioso, apresentou acúmulo reduzido do DNA-B e induziu sintomas atenuados em *N. benthamiana* (GILBERTSON et al., 1993b; HOU; GILBERTSON, 1996). Entretanto, após três passagens mecânicas sucessivas nesse hospedeiro, os sintomas tornaram-se idênticos aos produzidos pelo ToMoV e o nível do DNA-B tornou-se igual ao do DNA-A. A análise das regiões comuns dos DNAs-A e -B do pseudo-recombinante comprovou

a ocorrência de recombinação intermolecular na região comum do BDMV, que foi substituída quase que totalmente pela região comum do DNA-A do ToMoV (HOU; GILBERTSON, 1996). Assim, o DNA-B passou a ser reconhecido com 100% de eficiência pela proteína Rep do ToMoV. Esse resultado evidencia a importância da pseudo-recombinação na evolução de geminivírus e em sua adaptação a novos hospedeiros.

A formação de pseudo-recombinantes viáveis não depende somente da relação filogenética e conservação dos iterons, já que pseudo-recombinantes infecciosos foram formados entre o DNA-A do TGMV e o DNA-B do ToYSV, que possuem iterons distintos. Além disso, a assimetria na formação do pseudo-recombinante recíproco sugere que a proteína Rep do TGMV tem maior versatilidade em termos de reconhecimento de componentes de DNA heterólogos comparada à do ToYSV (ANDRADE et al., 2006a).

5.4. Estrutura genética de populações de geminivírus

A estrutura genética de populações de vírus de plantas refere-se à quantidade de variabilidade genética e a sua distribuição dentro e entre subpopulações (GARCÍA-ARENAL; FRAILE; MALPICA, 2001). Definir a estrutura genética é o primeiro passo para se estudar as populações virais, pois a estrutura genética reflete a história evolutiva e o potencial da população para evoluir (PINEL et al., 2003; MORENO et al., 2004; FONT et al., 2007). O entendimento da dinâmica da variabilidade de populações é necessário para entender como as populações evoluem, bem como as implicações para a durabilidade de medidas de manejo da virose (SEAL; JEGER; VAN DEN BOSCH, 2006a). Para a maior parte dos objetivos, a genética de populações fornece a ferramenta mais conveniente para estimar a diversidade genética de populações de patógenos. Os

principais mecanismos evolutivos que afetam a variabilidade das populações são seleção, deriva genética ao acaso, migração, mutação e recombinação (HARTL; CLARK, 2007). Quantificar a contribuição de cada mecanismo é importante e constitui o objetivo de vários estudos de biologia de populações de vírus de plantas (BULL et al., 2006; WANG; HUANG; COOPER, 2006; GARCIA-ANDRES et al., 2007a).

Diversos estudos já foram realizados como objetivo de investigar a estrutura genética de populações de geminivírus em diversos hospedeiros e em diferentes regiões geográficas. Recentemente, com o advento da técnica de amplificação por círculo rolante do genoma viral completo ("rolling circle amplification", RCA) (INOUE-NAGATA et al., 2004), novas possibilidades foram criadas para a análise de populações virais em escala genômica (HAIBLE; KOBER; JESKE, 2006), e alguns trabalhos nesse sentido já foram publicados (OWOR et al., 2007b; CASTILLO-URQUIZA, 2008; VARSANI et al., 2008; HARKINS et al., 2009; VARSANI et al., 2009).

Diversos trabalhos realizados ao longo das décadas de 1990 e 2000 avaliaram a estrutura populacional de begomovírus infectando mandioca na África Sub-Sahariana e no Sub-Continente Indiano. Nos países dessas regiões, a mandioca pode ser infectada por sete espécies de begomovírus (FAUQUET; FARGETTE, 1990; LEGG; RAYA, 1993; FARGETTE; THRESH; OTIM-NAPE, 1994) (curiosamente, não existem relatos de begomovírus que infectam mandioca no Brasil, o centro de origem e diversidade genética desta cultura). Os estudos realizados demonstraram um elevado grau de variabilidade genética da população viral em diversos países. A ocorrência frequente de infecções mistas facilita a ocorrência de pseudo-recombinação e recombinação, e em pelo menos dois casos foi demonstrada a emergência de novas espécies como consequência direta desses mecanismos (ZHOU et al., 1997; FONDONG et al., 2000).

Na Tanzânia, todas as sete espécies de begomovírus descritas que infectam mandioca já foram relatadas (NDUNGURU et al., 2005). Diversos eventos de recombinação foram detectados entre as estirpes TZ1 e TZ7 do *East African cassava mosaic Cameron virus* (EACMCV). A análise das sequências indicou que as duas estirpes têm a mesma origem local e, portanto, não foram introduzidas recentemente. A variabilidade genética da população viral foi analisada também com base no DNA-B, o que também indicou a existência de diversos eventos de recombinação. Os resultados indicam que a região central do continente africano é um centro de diversidade genética de begomovírus (NDUNGURU et al., 2005).

Além dos begomovírus que infectam a mandioca, a África também é o centro de origem dos mastrevírus que infectam gramíneas (PALMER; RYBICKI, 1998). Um estudo recente utilizando RCA analisou a estrutura genética da população viral em Uganda, um dos países mais afetados pelo estriado do milho causado pelo MSV (OWOR et al., 2007a). Amostras foram coletadas em 155 locais cobrindo todo o país. Inicialmente, fragmentos do genoma viral foram amplificados via PCR e a variabilidade foi analisada por meio de PCR-RFLP. Um total de 49 variantes foram identificados a partir de 391 isolados virais. A partir dessas 49 variantes, um total de 62 genomas completos foram sequenciados, e uma origem recombinante foi demonstrada para 52 desses genomas. Entretanto, um único recombinante, denominado MSV-A(1)UgIII, estava presente em infecção simples em mais de 60% das amostras infectadas em todo o país. Os autores concluíram que, embora a ocorrência de recombinação entre mastrevírus seja tão ou mais frequente em comparação com os begomovírus, o MSV deve estar sujeito a gargalos que limitam a variabilidade genética das populações naturais (OWOR et al., 2007a).

Font et al. (2007) determinaram a estrutura e variabilidade genética de populações de *Tomato yellow leaf curl Sardinia virus* (TYLCSV) e TYLCV em plantas de tomateiro em seis regiões da Espanha (Andaluzia, Ilhas Canárias, Lanzarote, Levante, Majorca e Murcia) entre os anos de 1997 e 2001. A análise de PCR-RFLP do gene da proteína capsidial e da região comum de 358 isolados revelou a presença de 14 haplótipos, e eventos de recombinação foram identificados na região comum. Em todas as regiões geográficas, exceto em Murcia, as populações eram compostas de um haplótipo predominante com uma baixa diversidade genética (<0,0180), ou estavam evoluindo para esta condição. Em Murcia, houve mudanças na predominância de haplótipos. O haplótipo I (TYLCSV) era predominante em 1997, mas sua frequência decresceu em 1998, com o aumento correspondente do haplótipo III (TYCLV) de modo que ambos haplótipos apresentaram frequências semelhantes. Em 1999, o haplótipo II surgiu e rapidamente tornou-se predominante na população. Esses resultados sugerem que a seleção negativa ocorreu de forma acentuada nessas populações. No entanto, o surgimento de haplótipos altamente adaptados se dispersando na população indica que seleção positiva também estava ocorrendo.

No Brasil, Castillo-Urquiza (2008), estudando duas populações de begomovírus que infectam tomateiro, *Tomato yellow vein streak virus* (ToYVSV) e *Tomato common mosaic virus* (ToCmMV) na região Sudeste do Brasil (municípios de Coimbra, MG e Paty do Alferes, RJ), observou maior variabilidade genética na população de ToCmMV. Demonstrou ainda que entre subpopulações de ToCmMV em Coimbra e Paty de Alferes havia maior variabilidade na subpopulação localizada em Coimbra.

A análise de uma população de BGMV infectando fava (*Phaseolus lunatus*) no estado de Alagoas, região Nordeste do Brasil, indicou uma alta taxa de variabilidade genética, significativamente maior que a observada para as duas populações de

begomovírus que infectam tomateiro no sudeste brasileiro (RAMOS-SOBRINHO et al., 2010).

6. Diversidade de begomovírus infectando plantas cultivadas e invasoras no Brasil

Durante as duas últimas décadas, begomovírus têm emergido como um dos principais patógenos de plantas, particularmente nas regiões tropicais e subtropicais no mundo, causando severas perdas econômicas (MORALES, 2006). No Brasil, as culturas mais severamente afetadas são o feijoeiro e tomateiro (FARIA; MAXWELL, 1999; ZERBINI et al., 2005). Embora existam relatos de infecção por begomovírus em outras culturas importantes como a soja (*Glycine max*) (MELLO; ALMEIDA; ZERBENI, 2000; MELLO et al., 2002) e o pimentão (*Capsicum annum*) (NOZAKI et al., 2005), esses ocorrem esporadicamente nas áreas de cultivo, não sendo considerados fatores limitantes à produção.

Begomovírus que infectam feijoeiro (*Phaseolus* spp.) são distribuídos através das Américas, sendo sua incidência um fator limitante para a produtividade dessa cultura. A diversidade genética de begomovírus que infectam feijoeiro é baixa, com apenas quatro espécies descritas: *Bean calico mosaic virus* (BcaMV), *Bean dwarf mosaic virus* (BDMV), BGMV e BGYMV (FAUQUET et al., 2008). Foi demonstrado também que isolados brasileiros de BGMV apresentam um baixo grau de variabilidade genética, o que não é comum para begomovírus (FARIA; MAXWELL, 1999). No entanto, estudos realizados em populações de BGMV infectando fava (*P. lunatus*) demonstraram que a variabilidade genética dentro dessa espécie é alta (SILVA, 2006; RAMOS-SOBRINHO et al., 2010).

Apesar da ocorrência frequente de BGMV em feijoeiro, infecções de begomovírus em soja não são comuns no Brasil. Ocorrências esporádicas têm sido

relatadas desde 1980, com a detecção de BGMV, *Sida mottle virus* (SiMoV) e duas possíveis novas espécies em amostras coletadas na região Sudeste (MELLO et al., 2002); e BGMV, *Sida micrantha mosaic virus* (SiMMV) e *Okra mottle virus* (OMoV) na região Centro-Oeste do país (FERNANDES et al., 2009). Este cenário está em contraste com a Argentina, onde a infecção de soja por três begomovírus distintos, incluindo o SiMoV, é frequente na região Noroeste, causando perdas moderadas a severas na produção (RODRÍGUEZ-PARDINA; ZERBINI; DUCASSE, 2006).

Uma situação oposta é observada para begomovírus que infectam solanáceas, a exemplo do tomateiro e do pimentão, onde um grande número de espécies tem sido descritas, e a variabilidade genética entre os isolados de uma determinada espécie é normalmente muito alta (RIBEIRO et al., 2003; CASTILLO-URQUIZA et al., 2008; FERNANDES et al., 2008).

O primeiro relato de begomovírus em tomateiro no Brasil foi feito na década de 1970 (COSTA; OLIVEIRA; SILVA, 1975). O vírus foi caracterizado e denominado TGMV. Além do TGMV, cinco outros vírus transmitidos por mosca-branca foram identificados, porém sem causar danos de importância econômica (MATYIS et al., 1975). Isso provavelmente ocorria porque o biótipo A de *B. tabaci*, o único presente no país naquela época, coloniza o tomateiro com baixa eficiência (BEDFORD et al., 1994). No entanto, no início da década de 1990 um complexo de begomovírus surgiu em tomateiro no Brasil, coincidindo com a introdução e disseminação do biótipo B de *B. tabaci* (AMBROZEVICIUS et al., 2002; RIBEIRO et al., 2003). Desde então, cinco espécies de begomovírus já foram descritas: ToCMoV, ToYSV, ToYVSV, *Tomato rugose mosaic virus* (ToRMV) e *Tomato severe rugose virus* (ToSRV) (FARIA; MAXWELL, 1999; FERNANDES et al., 2006; CALEGARIO et al., 2007; RIBEIRO et al., 2007). Além dessas, três novas espécies tentativas (Tomato commom mosaic virus,

ToCmMV; Tomato leaf distortion virus, ToLDV; Tomato mild mosaic virus, ToMlMV) foram identificadas com base na sequência do genoma completo (CASTILLO-URQUIZA et al., 2008), e seis outras foram descritas a partir de sequências parciais (RIBEIRO et al., 2003; FERNANDES et al., 2008). Algumas dessas espécies encontram-se amplamente distribuídas pelo país, enquanto outras estão restritas a certas regiões. Por exemplo, o ToSRV já foi relatado nos estados de Goiás, Minas Gerais, Pernambuco, Rio de Janeiro, Santa Catarina e São Paulo (REZENDE et al., 1997; LIMA et al., 2006; CASTILLO-URQUIZA et al., 2007; COTRIM et al., 2007; FERNANDES et al., 2008). Por outro lado, o ToYSV foi relatado apenas em Minas Gerais (CALEGARIO et al., 2007).

Levantamentos realizados para acessar a diversidade de begomovírus em tomateiro indicam que determinadas espécies tornaram-se prevalentes em diferentes regiões do país (CASTILLO-URQUIZA et al., 2007; COTRIM et al., 2007; CASTILLO-URQUIZA, 2008; FERNANDES et al., 2008). O sequenciamento direto de fragmentos de PCR de amostras de tomateiro coletadas na região central do estado de São Paulo nos anos de 2003 e 2004 revelou como espécie predominante o ToRSV, presente em 50% das amostras analisadas. O ToYVSV e o SiMoV também estavam presentes (COTRIM et al., 2007). A mesma estratégia foi utilizada para identificar begomovírus em amostras de tomateiro coletadas entre 2002 e 2004 no Distrito Federal e nos estados da Bahia, Goiás, Minas Gerais, Pernambuco e São Paulo. Verificou-se a presença do ToSRV em 61% das amostras, além do ToYVSV, Tomato mottle leaf curl virus (ToMoLCV) e duas possíveis novas espécies (FERNANDES et al., 2008).

Nos anos de 2005 e 2007 foi realizado um estudo sobre a diversidade de begomovírus em duas importantes regiões produtoras de tomate no Sudeste do Brasil, Paty do Alferes (RJ) e Coimbra (MG). A análise de sequências do genoma completo do

DNA-A revelou que em Paty do Alferes o ToYVSV era o vírus predominante, encontrado em 56,4% das amostras analisadas, seguido pelo ToCmMV. Já em Coimbra o ToCmMV foi o único vírus encontrado infectando tomateiro (CASTILLO-URQUIZA, 2008).

Acredita-se que a emergência dos begomovírus que infectam tomateiro no Brasil seja resultado da transferência horizontal de vírus nativos que infectam plantas silvestres ou invasoras pelo biótipo B da mosca-branca. Uma vez presentes no novo hospedeiro, esses vírus evoluiram rapidamente via recombinação e pseudo-recombinação, dando origem às espécies atualmente detectadas no campo. A predominância de algumas espécies poderia ser devido a diferenças na adaptação ao tomateiro ou diferenças na eficiência de transmissão pelo vetor (CASTILLO-URQUIZA et al., 2008).

Três observações corroboram essa hipótese. Em primeiro lugar, todas as espécies de begomovírus detectadas até o presente em tomateiro no Brasil são de ocorrência restrita ao país. Em segundo lugar, a caracterização biológica de algumas espécies (ToRMV, ToCMoV e ToYSV) confirmou que plantas daninhas como *Nicandra physaloides*, *Solanum nigrum* e *Datura stramonium* são hospedeiras (FERNANDES et al., 2006; CALEGARIO et al., 2007; RIBEIRO et al., 2007). Por fim, begomovírus originalmente encontrados em plantas silvestres/daninhas, como o SiMoV e o SimMV, já foram encontrados infectando naturalmente o tomateiro (CASTILLO-URQUIZA et al., 2007; COTRIM et al., 2007).

A presença de diversas espécies no campo, todas transmitidas pelo mesmo inseto vetor, torna comum a ocorrência de infecções mistas, com dois ou mais vírus presentes simultaneamente na mesma planta, aumentando a probabilidade da ocorrência de eventos de recombinação e pseudo-recombinação, o que pode levar ao surgimento de

espécies melhor adaptadas ao hospedeiro (PITA et al., 2001; MONCI et al., 2002; ANDRADE et al., 2006a; INOUE-NAGATA et al., 2006; RIBEIRO et al., 2007). Evidências de recombinação e pseudo-recombinação já foram encontradas em associação ao complexo de begomovírus infectando o tomateiro no Brasil. Galvão et al. (2003) e Ribeiro et al. (2007) sugeriram que os isolados MG-Bt1 e BA-Se1 do ToCMoV possuem origem recombinante. A formação de pseudo-recombinantes viáveis entre clones infecciosos do TGMV (DNA-A) e ToYSV (DNA-B), e entre o ToYSV (DNA-A) e o Tomato crinkle leaf yellow virus (ToCrLYV), já foi demonstrada (ANDRADE et al., 2006a). Além disso, foi sugerida a presença de um pseudo-recombinante ocorrendo naturalmente entre o ToRMV e um novo vírus (FERNANDES et al., 2006).

PAPROTKA et al. (2010a) estudaram a diversidade genética de begomovírus presentes em acessos de batata-doce naturalmente infectados no Banco de Germoplasma brasileiro. Nesse estudo foram identificadas duas novas espécies, Sweet potato golden vein-associated virus (SPGVaV) e Sweet potato mosaic-associated virus (SPMaV), além de três novos isolados e vários variantes do *Sweet potato leaf curl virus* (SPLCV). A comparação de sequências dos begomovírus encontrados nesses acessos revelou a presença de “footprints” de recombinação em seus genomas, ressaltando o risco do surgimento de novos begomovírus no material propagado vegetativamente no Banco de Germoplasma.

Além das plantas cultivadas, muitas espécies silvestres e/ou invasoras têm sido relatadas como hospedeiras de begomovírus em vários países, incluindo o Brasil (IDRIS et al., 2003; JOVEL et al., 2004; VARSANI et al., 2009; FIALLO-OLIVE et al., 2010; MUBIN et al., 2010). As espécies comumente relatadas como hospedeiras pertencem às famílias Malvaceae, Euphorbiaceae e Fabaceae (MORALES; ANDERSON, 2001).

Alguns estudos demonstraram que begomovírus provenientes de plantas invasoras podem ser transmitidos para espécies cultivadas pelo inseto vetor ou mediante inoculação via extrato vegetal tamponado (FRISCHMUTH et al., 1997; FARIA et al., 2000; MORALES; ANDERSON, 2001; CASTILLO-URQUIZA et al., 2007; COTRIM et al., 2007).

No Brasil, já se realizaram estudos com o objetivo de caracterizar molecularmente isoalados de begomovírus que infectam plantas silvestres e daninhas, sobretudo em associação às culturas do feijoeiro e do tomateiro (RIBEIRO et al., 1998; FARIA; MAXWELL, 1999; CASTILLO-URQUIZA et al., 2008). Os resultados desses estudos revelaram que, a exemplo do que ocorre com plantas cultivadas, a diversidade genética é alta entre os isolados de begomovírus que infectam plantas invasoras (AMBROZEVICIUS et al., 2002; CALEGARIO, 2004; CASTILLO-URQUIZA, 2008).

O SiMoV, obtido de plantas de *Sida rhombifolia* coletadas em Viçosa, MG (FERNANDES et al., 1999), foi encontrado em plantas de tomateiro na Zona Metalúrgica no estado de Minas Gerais (CALEGARIO, 2004).

Na Serra do Ibiapaba, CE, amostras assintomáticas de plantas invasoras de sete famílias botânicas e 18 espécies vegetais foram avaliadas por ELISA e PCR para infecção por begomovírus. Espécies de plantas daninhas pertencentes às famílias Amaranthaceae (*Amaranthus deflexus, A. spinosus*, *A. viridis*), Asteraceae (*Acanthospermum hispidum*, *Ageratum conyzoides*, *Bidens pilosa*), Euphorbiaceae (*Euphorbia heterophylla*) e Rubiaceae (*Borreria capitata*) foram identificadas como hospedeiras naturais de begomovírus (SANTOS; GONÇALVES; OLIVEIRA, 2003; ARNAUD et al., 2007).

Silva; Santos; Nascimento (2010), realizaram ensaios de inoculação por mosca-branca e enxertia com o objetivo de observar a transmissão de begomovírus a partir de

tomateiros infectados para quatro espécies de plantas invasoras (*Amaranthus spinosus*, *A. viridis*, *Ageratum conizoydes* e *B. pilosa*) e verificação de seu retorno para o tomateiro. Os resultados indicaram que o vetor transmitiu eficientemente o vírus para as quatro espécies. Por enxertia, apenas *B. pilosa* foi infectada. Esses resultados demonstram que as espécies invasoras são hospedeiras alternativas dos begomovírus de tomateiro presentes na região da Serra de Ibiapaba e, em condições de campo, na presença do vetor, podem constituir importantes fontes de inóculo para essa cultura. No entanto, as espécies de begomovírus infectando estas plantas não foram identificadas.

Plantas daninhas coletadas em municípios dos estados de Alagoas, Bahia e Pernambuco, com sintomas de mosaico amarelo, deformação do limbo foliar e redução do crescimento, foram avaliadas para a presença de begomovírus via PCR (ASSUNÇÃO et al., 2006). A infecção viral foi confirmada em *Cleome affinis* (Capparaceae), *Cnidoscolus urens* (Euphorbiaceae), *Desmodium* sp., *Macroptilium lathyroides* (Fabaceae), *Herissantia crispa*, *Sidastrum micranthum, S. rhombifolia*, *Sida spinosa* (Malvaceae), *Triumfetta semitriloba* e *Waltheria indica* (Sterculiaceae). Padrões distintos de clivagem obtidos em análise de PCR-RFLP sugeriram a existência de um alto grau de variabilidade genética (ASSUNÇÃO et al., 2006). Entretanto, as espécies de begomovírus infectando estas plantas não foram identificadas.

Castillo-Urquiza et al. (2008) analisaram a presença de begomovírus em tomateiro e plantas invasoras associadas à cultura. Foram encontradas seis novas espécies, três provenientes do tomateiro e três provenientes das invasoras *Blainvillea rhomboidea* (*Blainvillea yellow spot virus*, BlYSV), *Sida rhombifolia* (*Sida yellow mosaic virus*, SiYMV) e *Sida micrantha* (*Sida common mosaic virus*, SiCmMV).

A partir de material foliar de plantas sintomáticas pertencentes às famílias Malvaceae, Euphorbiaceae e Capparaceae, coletadas no município de Miranda (Mato

Grosso do Sul) foram identificadas duas novas espécies de begomovírus, Cleome leaf crumple virus (ClLCrV), obtido de *Cleome affinis*, e Sida mosaic Brazil virus (SiMBV). Além disso, foram encontrados dois alfassatélites associados ao *Euphorbia mosaic virus* (Euphorbia mosaic virus Mato Grosso do Sul-associated DNA1) e ao ClLCrV (Cleome leaf crumple virus-associated DNA1). Este foi o primeiro relato de alfassatélites ocorrendo naturalmente no Novo Mundo (PAPROTKA; METZLER; JESKE, 2010c).

Um novo begomovírus, Abutilon mosaic Brazil virus (AbMBV), foi identificado infectando *Abutilon* sp. no estado da Bahia. Análises filogenéticas demonstraram que ambos os componentes genômicos são distintos da espécie clássica, *Abutilon mosaic virus* (ABMV) originária do oeste da Índia. Além disso, inoculação via biobalística comprovou sua transmissão para *Malva parviflora*, a qual desenvolveu sintomas característicos de clareamento de nervuras e mosaico (PAPROTKA; METZLER; JESKE, 2010b).

O objetivo desse estudo foi determinar a diversidade e estrutura genética de populações de begomovírus que infectam plantas daninhas (famílias Capparaceae e Fabaceae), no Nordeste do Brasil, como um passo para acessar sua importância como reservatórios naturais e fontes de inóculo desses vírus.

Referências Bibliográficas

AMBROZEVICIUS, L.P.; CALEGARIO, R.F.; FONTES, E.P.B.; CARVALHO, M.G.; ZERBINI, F.M. Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. **Fitopatologia Brasileira**, Brasília, v. 27, p. 372-377, 2002.

ANDRADE, E.C.; MANHANI, G.G.; ALFENAS, P.F.; CALEGARIO, R.F.; FONTES, E.P.B.; ZERBINI, F.M. *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from *Sida* sp., forms pseudorecombinants with begomoviruses from tomato but not from *Sida*. **Journal of General Virology**, London, v. 87, p. 3687-3696, 2006a.

ANDRADE, E.C.; MANHANI, G.G.; FONTES, E.P.B.; ZERBINI, F.M. Análise parcial dos determinantes genéticos virais responsáveis pela indução de sintomas por begomovírus em tomateiro e *Nicotiana benthamiana*. **Fitopatologia Brasileira**, Brasília, v. 31, p. S225, 2006b.

ARGUELLO-ASTORGA, G.; ASCENCIO-IBANEZ, J.T.; DALLAS, M.B.; OROZCO, B.M.;HANLEY-BOWDOIN, L. High-frequency reversion of geminivirus replication protein mutants during infection. **Journal of Virology**, Washington, v. 81, p. 11005-11015, 2007.

ARGUELLO-ASTORGA, G.; LOPEZ-OCHOA, L.; KONG, L.J.; OROZCO, B.M.; SETTLAGE, S.B.; HANLEY-BOWDOIN, L. A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. **Journal of Virology**, Washington, v. 78, p. 4817-4826, 2004.

ARGÜELLO-ASTORGA, G.R.; GUEVARA-GONZÁLEZ, R.G.; HERRERA--ESTRELLA, L.R.; RIVERA-BUSTAMANTE, R.F. Geminivirus replication origins have a group-specific organization of interative elements: a model for replication. **Virology**, New York, v. 203, p. 90-100, 1994.

ARGUELLO-ASTORGA, G.R.; RUIZ-MEDRANO, R. An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of

potential interacting amino acid-base pairs by a comparative approach. **Archives of Virology**, New York, v. 146, p. 1465-1485, 2001.

ARNAUD, L.S.E.P.; SANTOS, C.D.G.; LIMA, J.A.A.; FEITOSA, F.A.A. Predominância de begomovírus em tomateiros na região produtora da Serra de Ibiapaba, Ceará, e sua detecção natural em plantas daninhas. **Fitopatologia Brasileira**, Brasília, v. 32, p. 241-246, 2007.

ASSUNÇÃO, I.P.; LISTIK, A.F.; BARROS, M.C.S.; AMORIM, E.P.R.; SILVA, S.J.C.; IZAEL, O.S.; RAMALHO-NETO, C.E.; LIMA, G.S.A. Diversidade genética de begomovírus que infectam plantas invasoras na Região Nordeste. **Planta Daninha**, Rio de Janeiro, v. 24, p. 239-244, 2006.

BEDFORD, I.D.; BRIDDON, R.W.; BROWN, J.K.; ROSELL, R.C.; MARKHAM, P.G. Geminivirus transmission and biological characterization of *Bemisia tabaci* (Gennadius) biotypes from different geographical regions. **Annals of Applied Biology**, Warwick, v. 125, p. 311-325, 1994.

BERRIE, L.C.; RYBICKI, E.P.; REY, M.E. Complete nucleotide sequence and host range of South African cassava mosaic virus: further evidence for recombination amongst begomoviruses. **Journal of General Virology**, London, v. 82, p. 53-58., 2001.

BIAGINI, P. **Human circoviruses.** In: (Ed.). Veterinary Microbiology. Netherlands, v.98, 2004. p.95-101.

BRIDDON, R.W. Cotton leaf curl disease, a multicomponent begomovirus complex. **Molecular Plant Pathology**, Bristol, v. 4, p. 427-434, 2003.

BRIDDON, R.W.; BEDFORD, I.D.; TSAI, J.H.; MARKHAM, P.G. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. **Virology**, New York, v. 219, p. 387-394, 1996.

BRIDDON, R.W.; BULL, S.E.; AMIN, I.; IDRIS, A.M.; MANSOOR, S.; BEDFORD, I.D.; DHAWAN, P.; RISHI, N.; SIWATCH, S.S.; ABDEL-SALAM, A.M.; BROWN, J.K.; ZAFAR, Y.; MARKHAM, P.G. Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. **Virology**, New York, v. 312, p. 106-121, 2003.

BRIDDON, R.W.; HEYDARNEJAD, J.; KHOSROWFAR, F.; MASSUMI, H.; MARTIN, D.P.; VARSANI, A. Turnip curly top virus, a highly divergent geminivirus infecting turnip in Iran. **Virus Research**, Amsterdam, v. 152, p. 169-175, 2010a.

BRIDDON, R.W.; MANSOOR, S.; BEDFORD, I.D.; PINNER, M.S.; SAUNDERS, K.; STANLEY, J.; ZAFAR, Y.; MALIK, K.A.; MARKHAM, P.G. Identification of DNA components required for induction of cotton leaf curl disease. **Virology**, New York, v. 285, p. 234-243, 2001.

BRIDDON, R.W.; PATIL, B.L.; BAGEWADI, B.; NAWAZ-UL-REHMAN, M.S.; FAUQUET, C.M. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. **BMC Evolutionary Biology**, London, v. 10, p., 2010b.

BRIDDON, R.W.; PINNER, M.S.; STANLEY, J.; MARKHAM, P.G. Geminivirus coat protein gene replacement alters insect specificity. **Virology**, New York, v. 177, p. 85-94, 1990.

BRIDDON, R.W.; STANLEY, J. Subviral agents associated with plant single-stranded DNA viruses. **Virology**, New York, v. 344, p. 198-210, 2006.

BULL, S.E.; BRIDDON, R.W.; SSERUBOMBWE, W.S.; NGUGI, K.; MARKHAM, P.G.; STANLEY, J. Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. **Journal of General Virology**, London, v. 87, p. 3053-3065, 2006.

BULL, S.E.; BRIDDON, R.W.; SSERUBOMBWE, W.S.; NGUGI, K.; MARKHAM, P.G.; STANLEY, J. Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. **Journal of General Virology**, v. 88, p. 1624-1633, 2007.

CALEGARIO, R.F. Caracterização do isolado de begomovírus MG-Bi2, um possível membro da espécie *Sida micrantha mosaic virus* (SimMV). (Tese M.S.). Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, 2004. 48 p.

CALEGARIO, R.F.; FERREIRA, S.S.; ANDRADE, E.C.; ZERBINI, F.M. Characterization of *Tomato yellow spot virus*, (ToYSV), a novel tomato-infecting

begomovirus from Brazil. **Pesquisa Agropecuária Brasileira**, Rio de Janeiro, v. 42, p. 1335-1343, 2007.

CAMPOS-OLIVAS, R.; LOUIS, J.M.; CLEROT, D.; GRONENBORN, B.; GRONENBORN, A.M. 1H, 13C, and 15N assignment of the N-terminal, catalytic domain of the replication initiation protein from the geminivirus TYLCV. **Journal of Biomolecular NMR**, Zürich, v. 24, p. 73-74., 2002.

CARVALHO, C.M.; FONTENELLE, M.R.; FLORENTINO, L.H.; SANTOS, A.A.; ZERBINI, F.M.; FONTES, E.P.B. A novel nucleocytoplasmic traffic GTPase identified as a functional target of the bipartite geminivirus nuclear shuttle protein. **Plant Journal**, Oxford, v. 55, p. 869-880, 2008.

CASTILLO-URQUIZA, G.P. Diversidade e estrutura genética de begomovírus em duas regiões produtoras de tomate do Sudeste do Brasil. (Tese D.S.). Dep. de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, 2008. 107 p.

CASTILLO-URQUIZA, G.P.; BESERRA JR., J.E.A.; BRUCKNER, F.P.; LIMA, A.T.M.; VARSANI, A.; ALFENAS-ZERBINI, P.; ZERBINI, F.M. Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. **Archives of Virology**, New York, v. 153, p. 1985-1989, 2008.

CASTILLO-URQUIZA, G.P.; BESERRA JUNIOR, J.E.A.; ALFENAS-ZERBINI, P.; VARSANI, A.; LIMA, A.T.M.; BARROS, D.R.; ZERBINI, F.M. Genetic diversity of begomoviruses infecting tomato in Paty do Alferes, Rio de Janeiro state, Brazil. **Virus Reviews and Research**, Belo Horizonte, v. 12, p. 233, 2007.

CHATTERJI, A.; PADIDAM, M.; BEACHY, R.N.; FAUQUET, C.M. Identification of replication specificity determinants in two strains of tomato leaf curl virus from New Delhi. **Journal of Virology**, Washington, v. 73, p. 5481-5489, 1999.

CHEN, H.; ZHANG, Z.H.; TENG, K.L.; LAI, J.B.; ZHANG, Y.Y.; HUANG, Y.L.; LI, Y.; LIANG, L.M.; WANG, Y.Q.; CHU, C.C.; GUO, H.S.; XIE, Q. Up-regulation of LSB1/GDU3 affects geminivirus infection by activating the salicylic acid pathway. **Plant Journal**, Oxford, v. 62, p. 12-23, 2010.

COSTA, A.S.; OLIVEIRA, A.R.; SILVA, D.M. Transmissão mecânica do mosaico dourado do tomateiro. **Revista da Sociedade Brasileira de Fitopatologia**, Brasília, v. 6, p. 147, 1975.

COTRIM, M.A.; KRAUSE-SAKATE, R.; NARITA, N.; ZERBINI, F.M.; PAVAN, M.A. Diversidade genética de begomovírus em cultivos de tomateiro no Centro-Oeste Paulista. **Summa Phytopathologica**, Jaguariuna, v. 33, p. 300-303, 2007.

CUI, X.F.; LI, G.X.; WANG, D.W.; HU, D.W.; ZHOU, X.P. A begomovirus DNA beta-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. **Journal of Virology**, Washington, v. 79, p. 10764-10775, 2005.

CUI, X.F.; TAO, X.R.; XIE, Y.; FAUQUET, C.M.; ZHOU, X.P. A DNA beta associated with Tomato Yellow Leaf Curl China Virus is required for symptom induction. **Journal of Virology**, Washington, v. 78, p. 13966-13974, 2004.

DAVIES, J.W.; STANLEY, J.; DONSON, J.; MULLINEAUX, P.M.; BOULTON, M.I. Structure and replication of geminivirus genomes. **Journal of Cell Science**, London, v. 7, p. 95-107, 1987.

DAVINO, S.; NAPOLI, C.; DELLACROCE, C.; MIOZZI, L.; NORIS, E.; DAVINO, M.; ACCOTTO, G.P. Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. **Virus Research**, Amsterdam, v. 143, p. 15-23, 2009.

DRAKE, J.W. A constant rate of spontaneous mutation in DNA-based microbes. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 88, p. 7160-7164, 1991.

DUFFY, S.; HOLMES, E.C. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. **Journal of Virology**, Washington, v. 82, p. 957-965, 2008.

DUFFY, S.; HOLMES, E.C. Validation of high rates of nucleotide substitution in geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses. **Journal of General Virology**, London, v. 90, p. 1539-1547, 2009.

DUFFY, S.; SHACKELTON, L.A.; HOLMES, E.C. Rates of evolutionary change in viruses: Patterns and determinants. **Nature Reviews Genetics**, London, v. advanced online publication, p., 2008.

EAGLE, P.A.; OROZCO, B.M.; HANLEY-BOWDOIN. A DNA sequence required for geminivirus replication also mediates transcriptional regulation. **Plant Cell**, Rockville, v. 6, p. 1157-1170, 1994.

FARGETTE, D.; THRESH, J.M.; OTIM-NAPE, G.W. The epidemiology of African cassava mosaic geminivirus: Reversion and the concept of equilibrium. **Tropical Science**, London, v. 34, p. 123-133, 1994.

FARIA, J.C.; BEZERRA, I.C.; ZERBINI, F.M.; RIBEIRO, S.G.; LIMA, M.F. Situação atual das geminiviroses no Brasil. **Fitopatologia Brasileira**, Brasília, v. 25, p. 125-137, 2000.

FARIA, J.C.; GILBERTSON, R.L.; HANSON, S.F.; MORALES, F.J.; AHLQUIST, P.G.; LONIELLO, A.O.; MAXWELL, D.P. Bean golden mosaic geminivirus type II isolates from the Dominican Republic and Guatemala: Nucleotide sequences, infectious pseudorecombinants, and phylogenetic relationships. **Phytopathology**, Saint Paul, v. 84, p. 321-329, 1994.

FARIA, J.C.; MAXWELL, D.P. Variability in geminivirus isolates associated with *Phaseolus* spp. in Brazil. **Phytopathology**, Saint Paul, v. 89, p. 262-268, 1999.

FAUQUET, C.; FARGETTE, D. African cassava mosaic virus: etiology, epidemiology and control. **Plant Disease**, Palo Alto, v. 74, p. 404-411, 1990.

FAUQUET, C.M.; BRIDDON, R.W.; BROWN, J.K.; MORIONES, E.; STANLEY, J.; ZERBINI, F.M.; ZHOU, X. Geminivirus strain demarcation and nomenclature. **Archives of Virology**, New York, v. 153, p. 783-821, 2008.

FAUQUET, C.M.; SAWYER, S.; IDRIS, A.M.; BROWN, J.K. Sequence analysis and classification of apparent recombinant begomoviruses infecting tomato in the nile and Mediterranean basins. **Phytopathology**, Saint Paul, v. 95, p. 549-555, 2005.

FERNANDES, A.V.; GALVÃO, R.M.; MACHADO, J.J.; ZERBINI, F.M.; FONTES, E.P.B. Cloning and molecular characterization of A components of two new *Sida rhombifolia*-infecting geminiviruses. **Virus Reviews and Research**, Belo Horizonte, v. 4, p. 148, 1999.

FERNANDES, F.R.; ALBUQUERQUE, L.C.; GIORDANO, L.B.; BOITEUX, L.S.; ÁVILA, A.C.; INOUE-NAGATA, A.K. Diversity and prevalence of Brazilian bipartite begomovirus species associated to tomatoes. **Virus Genes**, Norwell, v. 36, p. 251-258, 2008.

FERNANDES, F.R.; CRUZ, A.R.R.; FARIA, J.C.; ZERBINI, F.M.; ARAGÃO, F.J.L. Three distinct begomoviruses associated with soybean in central Brazil. **Archives of Virology**, New York, v. 154, p. 1567-1570, 2009.

FERNANDES, J.J.; CARVALHO, M.G.; ANDRADE, E.C.; BROMMONSCHENKEL, S.H.; FONTES, E.P.B.; ZERBINI, F.M. Biological and molecular properties of *Tomato rugose mosaic virus* (ToRMV), a new tomato-infecting begomovirus from Brazil. **Plant Pathology**, Oxford, v. 55, p. 513-522, 2006.

FIALLO-OLIVE, E.; NAVAS-CASTILLO, J.; MORIONES, E.; MARTINEZ-ZUBIAUR, Y. Two novel begomoviruses belonging to different lineages infecting Rhynchosia minima. **Archives of Virology**, New York, v. 155, p. 2053-2058, 2010.

FONDONG, V.N.; PITA, J.S.; REY, M.E.C.; KOCHKO, A.; BEACHY, R.N.; FAUQUET, C.M. Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. **Journal of General Virology**, London, v. 81, p. 287-297, 2000.

FONT, M.I.; RUBIO, L.; MARTINEZ-CULEBRAS, P.V.; JORDA, C. Genetic structure and evolution of natural populations of viruses causing the tomato yellow leaf curl disease in Spain. **Virus Research**, Amsterdam, v. 128, p. 43-51, 2007.

FONTENELLE, M.R.; LUZ, D.F.; GOMES, A.P.; FLORENTINO, L.H.; ZERBINI, F.M.; FONTES, E.P. Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus *Tomato chlorotic mottle virus*. **Virus Research**, Amsterdam, v. 126, p. 262-267, 2007.

FONTES, E.P.B.; EAGLE, P.A.; SIPE, P.S.; LUCKOW, V.A.; HANLEY-BOWDOIN, L. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. **Journal of Biological Chemistry**, Bethesda, v. 269, p. 8459-8465, 1994.

FONTES, E.P.B.; LUCKOW, V.A.; HANLEY-BOWDOIN, L. A geminivirus replication protein is a sequence-specific DNA binding protein. **Plant Cell**, Rockville, v. 4, p. 597-608, 1992.

FRISCHMUTH, S.; KLEINOW, T.; ABERLE, H.J.; WEGE, C.; HULSER, D.; JESKE, H. Yeast two-hybrid systems confirm the membrane association and oligomerization of BC1 but do not detect an interaction of the movement proteins BC1 and BV1 of Abutilon mosaic geminivirus. **Archives of Virology**, New York, v. 149, p. 2349-2364, 2004.

FRISCHMUTH, S.; WEGE, C.; HULSER, D.; JESKE, H. The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast. **Protoplasma**, New York, v. 230, p. 117-123, 2007.

FRISCHMUTH, T.; ENGEL, M.; LAUSTER, S.; JESKE, H. Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted, Sida-infecting bipartite geminiviruses in Central America. **Journal of General Virology**, London, v. 78, p. 2675-2682, 1997.

GAFNI, Y.; EPEL, B.L. The role of host and viral proteins in intra and inter-cellular trafficking of geminiviruses. **Physiological and Molecular Plant Pathology**, London, v. 60, p. 231-241, 2002.

GALLIAN, P.; BIAGINI, P.; ATTOUI, H.; CANTALOUBE, J.F.; DUSSOL, B.; BERLAND, Y.; DE MICCO, P.; DE LAMBALLERIE, X. High genetic diversity revealed by the study of TLMV infection in French hemodialysis patients. **Journal of Medical Virology**, New York, v. 67, p. 630-635, 2002.

GALVÃO, R.M.; MARIANO, A.C.; LUZ, D.F.; ALFENAS, P.F.; ANDRADE, E.C.; ZERBINI, F.M.; ALMEIDA, M.R.; FONTES, E.P.B. A naturally occurring recombinant DNA-A of a typical bipartite begomovirus does not require the cognate

DNA-B to infect *Nicotiana benthamiana* systemically. **Journal of General Virology**, London, v. 84, p. 715-726, 2003.

GARCIA-ANDRES, S.; ACCOTTO, G.P.; NAVAS-CASTILLO, J.; MORIONES, E. Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. **Virology**, New York, v. 359, p. 302-312, 2007a.

GARCIA-ANDRES, S.; MONCI, F.; NAVAS-CASTILLO, J.; MORIONES, E. Begomovirus genetic diversity in the native plant reservoir Solanum nigrum: Evidence for the presence of a new virus species of recombinant nature. **Virology**, New York, v. 350, p. 433-442, 2006.

GARCIA-ANDRES, S.; TOMAS, D.M.; SANCHEZ-CAMPOS, S.; NAVAS-CASTILLO, J.; MORIONES, E. Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. **Virology**, New York, v. 365, p. 210-219, 2007b.

GARCIA-ARENAL, F.; FRAILE, A.; MALPICA, J.M. Variability and genetic structure of plant virus populations. **Annual Review of Phytopathology**, Palo Alto, v. 39, p. 157-186, 2001.

GARCIA-ARENAL, F.; FRAILE, A.; MALPICA, J.M. Variation and evolution of plant virus populations. **International Microbiology**, Madrid, v. 6, p. 225-232, 2003.

GARRIDO-RAMIREZ, E.R.; SUDARSHANA, M.; GILBERTSON, R.L. *Bean golden yellow mosaic virus* from Chiapas, Mexico: Characterization, pseudorecombination with other bean-infecting geminiviruses and germ plasm screening. **Phytopathology**, Saint Paul, v. 90, p. 1224-1232, 2000.

GE, L.M.; ZHANG, J.T.; ZHOU, X.P.; LI, H.Y. Genetic structure and population variability of tomato yellow leaf curl China virus. **Journal of Virology**, Washington, v. 81, p. 5902-5907, 2007.

GILBERTSON, R.L.; FARIA, J.C.; AHLQUIST, P.; MAXWELL, D.P. Genetic diversity in geminiviruses causing bean golden mosaic disease: the nucleotide sequence

of the infectious cloned DNA components of a Brazilian isolate of bean golden mosaic geminivirus. **Phytopathology**, Saint Paul, v. 83, p. 709-715, 1993a.

GILBERTSON, R.L.; HIDAYAT, S.H.; PAPLOMATAS, E.J.; ROJAS, M.R.; HOU, Y.-H.; MAXWELL, D.P. Pseudorecombination between infectious cloned DNA components of tomato mottle and bean dwarf mosaic geminiviruses. **Journal of General Virology**, London, v. 74, p. 23-31, 1993b.

GILBERTSON, R.L.; SUDARSHANA, M.; JIANG, H.; ROJAS, M.R.; LUCAS, W.J. Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: Insights into DNA trafficking. **Plant Cell**, Rockville, v. 15, p. 2578-2591, 2003.

GLADFELTER, H.J.; EAGLE, P.A.; FONTES, E.P.B.; BATTS, L.; HANLEY-BOWDOIN, L. Two domains of the AL 1 protein mediate geminivirus origin recognition. **Virology**, New York, v. 239, p. 186-197, 1997.

GUTIERREZ, C. Geminivirus DNA replication. **Cellular and Molecular Life Sciences**, Basel, v. 56, p. 313-329, 1999.

GUTIERREZ, C.; RAMIREZ-PARRA, E.; CASTELLANO, M.M.; SANZ-BURGOS, A.P.; LUQUE, A.; MISSICH, R. Geminivirus DNA replication and cell cycle interactions. **Veterinary Microbiology**, Amsterdam, v. 98, p. 111-119, 2004.

HA, C.; COOMBS, S.; REVILL, P.; HARDING, R.; VU, M.; DALE, J. Corchorus yellow vein virus, a New World geminivirus from the Old World. **Journal of General Virology**, London, v. 87, p. 997-1003, 2006.

HA, C.; COOMBS, S.; REVILL, P.; HARDING, R.; VU, M.;DALE, J. 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. In: (Ed.). Journal Of General Virology. England, v.89, 2008. p.312-326.

HAIBLE, D.; KOBER, S.; JESKE, H. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. **Journal of Virological Methods**, Amsterdam, v. 135, p. 9-16, 2006.

HANLEY-BOWDOIN, L.; SETTLAGE, S.B.; OROZCO, B.M.; NAGAR, S.; ROBERTSON, D. Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. **Critical Reviews in Plant Sciences**, Boca Raton, v. 18, p. 71-106, 1999.

HARKINS, G.W.; DELPORT, W.; DUFFY, S.; WOOD, N.; MONJANE, A.L.; OWOR, B.E.; DONALDSON, L.; SAUMTALLY, S.; TRITON, G.; BRIDDON, R.W.; SHEPHERD, D.N.; RYBICKI, E.P.; MARTIN, D.P.; VARSANI, A. Experimental evidence indicating that mastreviruses probably did not co-diverge with their hosts. **Virology Journal**, London, v. 6, p.1-14, 2009.

HARRISON, B.D.; ROBINSON, D.J. Natural genomic and antigenic variation in white-fly transmitted geminiviruses (begomoviruses). **Annual Review of Phytopathology**, Palo Alto, v. 39, p. 369-398, 1999.

HARTL, D.L.; CLARK, A.G. **Principles of population genetics**. Sunderland: Sinauer Associates, 2007, 221p.

HEYRAUD-NITSCHKE, F.; SCHUMACHER, S.; LAUFS, J.; SCHAEFER, S.; SCHELL, J.; GRONENBORN, B. Determination of the origin cleavage and joining domain of geminivirus Rep proteins. **Nucleic Acids Research**, Oxford, v. 23, p. 910-916, 1995.

HILL, J.E.; STRANDBERG, J.O.; HIEBERT, E.; LAZAROWITZ, S.G. Asymmetric infectivity of pseudorecombinants of cabbage leaf curl virus and squash leaf curl virus: Implications for bipartite geminivirus evolution and movement. **Virology**, New York, v. 250, p. 283-292, 1998.

HÖFER, P.; BEDFORD, I.D.; MARKHAM, P.G.; JESKE, H.; FRISCHMUTH, T. Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. **Virology**, New York, v. 236, p. 288-295, 1997a.

HÖFER, P.; ENGEL, M.; JESKE, H.; FRISCHMUTH, T. Nucleotide sequence of a new bipartite geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica. **Virology**, New York, v. 78, p. 1785-1790, 1997b.

HOLMES, E.C. The phylogeography of human viruses. **Molecular Ecology**, Oxford, v. 13, p. 745-756, 2004.

HOU, Y.M.; GILBERTSON, R.L. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. **Journal of Virology**, Washington, v. 70, p. 5430-5436, 1996.

IDRIS, A.M.; BRIDDON, R.W.; BULL, S.E.; BROWN, J.K. Cotton leaf curl Gezira virus-satellite DNAs represent a divergent, geographically isolated Nile Basin lineage: predictive identification of a satDNA REP-binding motif. **Virus Research**, Amsterdan, v. 109, p. 19-32, 2005.

IDRIS, A.M.; BROWN, J.K. Molecular analysis of Cotton leaf curl virus-Sudan reveals an evolutionary history of recombination. **Virus Genes**, Norwell, v. 24, p. 249-256., 2002.

IDRIS, A.M.; HIEBERT, E.; BIRD, J.; BROWN, J.K. Two newly described begomoviruses of *Macroptilium lathyroides* and common bean. **Phytopathology**, Saint Paul, v. 93, p. 774-783, 2003.

ILYINA, T.V.; KOONIN, E.V. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. **Nucleic Acids Research**, Oxford, v. 20, p. 3279-3285, 1992.

INOUE-NAGATA, A.K.; ALBUQUERQUE, L.C.; ROCHA, W.B.; NAGATA, T. A simple method for cloning the complete begomovirus genome using the bacteriophage phi 29 DNA polymerase. **Journal of Virological Methods**, Amsterdam, v. 116, p. 209-211, 2004.

INOUE-NAGATA, A.K.; MARTIN, D.P.; BOITEUX, L.S.; GIORDANO, L.D.; BEZERRA, I.C.; DE AVILA, A.C. New species emergence via recombination among isolates of the Brazilian tomato infecting Begomovirus complex. **Pesquisa Agropecuária Brasileira**, Rio de Janeiro, v. 41, p. 1329-1332, 2006.

ISNARD, M.; GRANIER, M.; FRUTOS, R.; REYNAUD, B.; PETERSCHMITT, M. 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. **Journal of General Virology**, London, v. 79, p. 3091-3099., 1998.

JEFFREY, J.L.; POOMA, W.; PETTY, I.T. Genetic requirements for local and systemic movement of tomato golden mosaic virus in infected plants. **Virology**, New York, v. 223, p. 208-218., 1996.

JESKE, H.; LUTGEMEIER, M.; PREISS, W. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. **EMBO Journal**, Oxford, v. 20, p. 6158-6167, 2001.

JOVEL, J.; RESKI, G.; ROTHENSTEIN, D.; RINGEL, M.; FRISCHMUTH, T.; JESKE, H. *Sida micrantha* mosaic is associated with a complex infection of begomoviruses different from *Abutilon mosaic virus*. **Archives of Virology**, New York, v. 149, p. 829-841, 2004.

JUPIN, I., HEROCOURT, F., BENZ, B., GRONENBORN, B. DNA replication specificity of TYLCV geminivirus is mediated by the amino-terminal 116 amino acids of the Rep protein. **FEBS Letters**, Amsterdam, v. 262, p. 116-120, 1995.

KLEINOW, T.; HOLEITER, G.; NISCHANG, M.; STEIN, M.; KARAYAVUZ, M.; WEGE, C.; JESKE, H. Post-translational modifications of Abutilon mosaic virus movement protein (BC1) in fission yeast. **Virus Research**, Amsterdam, v. 131, p. 86-94, 2008.

LAUFS, J.; TRAUT, W.; HEYRAUD, F.; MATZEIT, G.; ROGERS, S.G.; SCHELL, J.; GRONENBORN, B. *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 92, p. 3879-3883, 1995.

LAZAROWITZ, S.G. Geminiviruses: Genome structure and gene function. **Critical Reviews in Plant Sciences**, Boca Raton, v. 11, p. 327-349, 1992.

LEFEUVRE, P.; LETT, J.M.; REYNAUD, B.; MARTIN, D.P. Avoidance of protein fold disruption in natural virus recombinants. **PLoS Pathogen**, Cambridge, v. 3, p. e181, 2007a.

LEFEUVRE, P.; MARTIN, D.P.; HOAREAU, M.; NAZE, F.; DELATTE, H.; THIERRY, M.; VARSANI, A.; BECKER, N.; REYNAUD, B.; LETT, J.M. Begomovirus 'melting pot' in the south-west Indian Ocean islands: Molecular diversity and evolution through recombination. **Journal of General Virology**, v. 88, p. 3458-3468, 2007b.

LEGG, J.P.; RAYA, M.D. Survey of cassava virus diseases in Tanzania. **International Journal of Pest Management**, London, v., p. 17-23, 1993.

LEVY, A.; TZFIRA, T. Bean dwarf mosaic virus: a model system for the study of viral movement. **Molecular Plant Pathology**, Bristol, v. 11, p. 451-461, 2010.

LIMA, A.T.M.; PEREIRA, C.O.; ALFENAS, P.F.; PAULA, M.B.; MELLO, R.N.; ZERBINI, F.M. Primeiro relato de infecção pelo geminivírus *Tomato severe rugose virus* (ToSRV) em tomateiro no estado de Santa Catarina. **Fitopatologia Brasileira**, Brasília, v. 31(Suplemento), p. S224, 2006.

LIN, B.; AKBAR BEHJATNIA, S.A.; DRY, I.B.; RANDLES, J.W.; REZAIAN, M.A. High-affinity Rep-binding is not required for the replication of a geminivirus DNA and its satellite. **Virology**, New York, v. 305, p. 353-363, 2003.

MARIANO, A.C.; ANDRADE, M.O.; SANTOS, A.A.; CAROLINO, S.M.B.; OLIVEIRA, M.L.; BARACAT-PEREIRA, M.C.; BROMMONSHENKEL, S.H.; FONTES, E.P.B. Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. **Virology**, New York, v. 318, p. 24-31, 2004.

MATYIS, J.C.; SILVA, D.M.; OLIVEIRA, A.R.; COSTA, A.S. Purificação e morfologia do vírus do mosaico dourado do tomateiro. **Summa Phytopathologica**, Jaguariuna, v. 1, p. 267-275, 1975.

MELLO, R.N.; ALMEIDA, A.M.R.; ZERBINI, F.M. Detection and identification of geminiviruses infecting soybean and associated weeds in Brazil. **Fitopatologia Brasileira**, Brasília, v. 25, p. 444, 2000.

MELLO, R.N.; COTRIM, M.A.A.; LOPES, E.F.; MOREIRA, A.G.; CONTIN, F.S.; FONTES, E.P.B.; ALMEIDA, A.M.R.; ZERBINI, F.M. Survey of begomoviruses associated with soybean and identification of *Sida mottle virus* (SiMoV) infecting this

crop in Brazil. **Virus Reviews and Research**, Belo Horizonte, v. 7(Supplement), p. 157, 2002.

MONCI, F.; SANCHEZ-CAMPOS, S.; NAVAS-CASTILLO, J.; MORIONES, E. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and *Tomato yellow leaf curl virus* exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. **Virology**, New York, v. 303, p. 317-326, 2002.

MORALES, F.J. History and current distribution of begomoviruses in Latin America. **Advances in Virus Research**, San Diego, v. 67, p. 127-162, 2006.

MORALES, F.J.; ANDERSON, P.K. The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. **Archives of Virology**, New York, v. 146, p. 415-441, 2001.

MORENO, I.M.; MALPICA, J.M.; DIAZ-PENDON, J.A.; MORIONES, E.; FRAILE, A.; GARCIA-ARENAL, F. Variability and genetic structure of the population of watermelon mosaic virus infecting melon in Spain. **Virology**, New York, v. 318, p. 451-460, 2004.

MORILLA, G.; KRENZ, B.; JESKE, H.; BEJARANO, E.R.;WEGE, C. Tête à tête of tomato yellow leaf curl virus and tomato yellow leaf curl sardinia virus in single nuclei. **Journal of Virology**, Washington, v. 78, p. 10715-10723, 2004.

MUBIN, M.; SHAHID, M.S.; TAHIR, M.N.; BRIDDON, R.W.; MANSOOR, S. Characterization of begomovirus components from a weed suggests that begomoviruses may associate with multiple distinct DNA satellites. **Virus Genes**, Norwell, v. 40, p. 452-457, 2010.

NAWAZ-UL-REHMAN, M.S.; FAUQUET, C.M. Evolution of geminiviruses and their satellites. **FEBS Letters**, Amsterdam, v. 583, p. 1825-1832, 2009.

NAWAZ-UL-REHMAN, M.S.; MANSOOR, S.; BRIDDON, R.W.; FAUQUET, C.M. Maintenance of an Old World Betasatellite by a New World Helper Begomovirus and Possible Rapid Adaptation of the Betasatellite. **Journal of Virology**, Washington, v. 83, p. 9347-9355, 2009.

NDUNGURU, J.; LEGG, J.; AVELING, T.; THOMPSON, G.; FAUQUET, C. 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. **Virology Journal**, London, v. 2, p. 21, 2005.

NOUEIRY, A.O.; LUCAS, W.J.; GILBERTSON, R.L. Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. **Cell**, Cambridge, v. 76, p. 925-932, 1994.

NOZAKI, D.N.; KRAUSE-SAKATE, R.; HASEGAWA, J.M.; CESAR, M.A.; DZIUBA, P.H.; PAVAN, M.A. Ocorrência de *Tomato severe rugose virus* em pimentão (*Capsicum annuum* L.) no estado de São Paulo. **Fitopatologia Brasileira**, Brasília, v. 30 (Suplemento), p. S189, 2005.

OROZCO, B.M.; HANLEY-BOWDOIN, L. Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. **Journal of Biological Chemistry**, Bethesda, v. 273, p. 24448-24456, 1998.

OROZCO, B.M.; MILLER, A.B.; SETTLAGE, S.B.; HANLEY-BOWDOIN, L. Functional domains of a geminivirus replication protein. **Journal of Biological Chemistry**, v. 272, p. 9840-9846, 1997.

OWOR, B.E.; MARTIN, D.P.; SHEPHERD, D.N.; EDEMA, R.; MONJANE, A.L.; RYBICKI, E.P.; THOMSON, J.A.; VARSANI, A. Genetic analysis of *Maize streak virus* isolates from Uganda reveals widespread distribution of a recombinant variant. **Journal of General Virology**, London, v. 88, p. 3154-3165, 2007a.

OWOR, B.E.; SHEPHERD, D.N.; TAYLOR, N.J.; EDEMA, R.; MONJANE, A.L.; THOMSON, J.A.; MARTIN, D.P.; VARSANI, A. Successful application of FTA Classic Card technology and use of bacteriophage phi29 DNA polymerase for large-scale field sampling and cloning of complete maize streak virus genomes. **Journal of Virological Methods**, Amsterdam, v. 140, p. 100-105, 2007b.

PADIDAM, M.; BEACHY, R.N.; FAUQUET, C.M. Classification and identification of geminiviruses using sequence comparisons. **Journal of General Virology**, London, v. 76, p. 249-263, 1995.

PADIDAM, M.; BEACHY, R.N.; FAUQUET, C.M. The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. **Virology**, New York, v. 224, p. 390-404, 1996.

PADIDAM, M.; BEACHY, R.N.; FAUQUET, C.M. A phage single-stranded DNA (ssDNA) binding protein complements ssDNA accumulation of a geminivirus and interferes with viral movement. **Journal of Virology**, Washington, v. 73, p. 1609-1616, 1999a.

PADIDAM, M.; SAWYER, S.; FAUQUET, C.M. Possible emergence of new geminiviruses by frequent recombination. **Virology**, New York, v. 265, p. 218-224, 1999b.

PALMER, K.E.;RYBICKI, E.P. The molecular biology of mastreviruses. **Advances in Virus Research**, San Diego, v. 50, p. 183-234, 1998.

PAPLOMATAS, E.J.; PATEL, V.P.; HOU, Y.M.; NOUEIRY, A.O.; GILBERTSON, R.L. Molecular characterization of a new sap-transmissible bipartite genome geminivirus infecting tomatoes in Mexico. **Phytopathology**, Saint Paul, v. 84, p. 1215-1224, 1994.

PAPROTKA, T.; BOITEUX, L.S.; FONSECA, M.E.N.; RESENDE, R.O.; JESKE, H.; FARIA, J.C.; RIBEIRO, S.G. Genomic diversity of sweet potato geminiviruses in a Brazilian germplasm bank. **Virus Research**, Amsterdam, v. 149, p. 224-233, 2010a.

PAPROTKA, T.; METZLER, V.; JESKE, H. The complete nucleotide sequence of a new bipartite begomovirus from Brazil infecting Abutilon. **Archives of Virology**, New York, v. 155, p. 813-816, 2010b.

PAPROTKA, T.; METZLER, V.; JESKE, H. The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. **Virology**, New York, v. 404, p. 148-157, 2010c.

PEDERSEN, T.J.; HANLEY-BOWDOIN. Molecular characterization of the AL3 protein encoded by a bipartite geminivirus. **Virology**, New York, v. 202, p. 1070-1075, 1994.

PINEL, A.; ABUBAKAR, Z.; TRAORE, O.; KONATE, G.; FARGETTE, D. Molecular epidemiology of the RNA satellite of Rice yellow mottle virus in Africa. **Archives of Virology**, New York, v. 148, p. 1721-1733, 2003.

PITA, J.S.; FONDONG, V.N.; SANGARE, A.; OTIM-NAPE, G.W.; OGWAL, S.; FAUQUET, C.M. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. **Journal of General Virology**, London, v. 82, p. 655-665, 2001.

PREISS, W.; JESKE, H. Multitasking in replication is common among geminiviruses. **Journal of Virology**, v. 77, p. 2972-2980, 2003.

QAZI, J.; ILYAS, M.; MANSOOR, S.; BRIDDON, R.W. Legume yellow mosaic viruses: Genetically isolated begomoviruses. In: (Ed.). Molecular Plant Pathology. England, v.8, 2007. p.343-348.

RAMOS-SOBRINHO, R.; SILVA, S.J.C.; SILVA, T.A.L.; RIBEIRO, S.G.; LIMA, G.S.A.; ASSUNÇÃO, I.P.; ZERBINI, F.M. Genetic structure of a population of the begomovirus *Bean golden mosaic virus* (BGMV) that infects lima bean (*Phaseolus lunatus* L.) in the state of Alagoas, Brazil. Program and Abstracts, 6th International Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop. Guadalajara, México. November 7-12, 2010, 2010. p.

RAMOS, P.L.; GUEVARA-GONZALEZ, R.G.; PERAL, R.; ASCENCIO-IBANEZ, J.T.; POLSTON, J.E.; ARGUELLO-ASTORGA, G.R.; VEGA-ARREGUIN, J.C.; RIVERA-BUSTAMANTE, R.F. Tomato mottle Taino virus pseudorecombines with PYMV but not with ToMoV: Implications for the delimitation of cis- and trans-acting replication specificity determinants. **Archives of Virology**, New York, v. 148, p. 1697-1712, 2003.

RANEY, J.L.; DELONGCHAMP, R.R.; VALENTINE, C.R. Spontaneous mutant frequency and mutation spectrum for gene A of phiX174 grown in E. coli. **Environmental and Molecular Mutagenesis**, New York, v. 44, p. 119-127, 2004.

REZENDE, W.L.; MILITÃO NETO, V.; GOULART, L.R.; GIOVANINI, M.P.; JULIATTI, F.C.; FERNANDES, J.J. Infecção mista em plantas de tomate infectadas

por geminivírus, detectada por meio de LIS-SSCP-PCR. **Fitopatologia Brasileira**, Brasília, v. 22, p. 338, 1997.

RIBEIRO, S.G.; AMBROZEVICIUS, L.P.; ÁVILA, A.C.; BEZERRA, I.C.; CALEGARIO, R.F.; FERNANDES, J.J.; LIMA, M.F.; MELLO, R.N.; ROCHA, H.; ZERBINI, F.M. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. **Archives of Virology**, New York, v. 148, p. 281-295, 2003.

RIBEIRO, S.G.; ÁVILA, A.C.; BEZERRA, I.C.; FERNANDES, J.J.; FARIA, J.C.; LIMA, M.F.; GILBERTSON, R.L.; ZAMBOLIM, E.M.; ZERBINI, F.M. Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. **Plant Disease**, Saint Paul, v. 82, p. 830, 1998.

RIBEIRO, S.G.; MARTIN, D.P.; LACORTE, C.; SIMÕES, I.C.; ORLANDINI, D.R.S.; INOUE-NAGATA, A.K. Molecular and biological characterization of *Tomato chlorotic mottle virus* suggests that recombination underlies the evolution and diversity of Brazilian tomato begomoviruses. **Phytopathology**, Saint Paul, v. 97, p. 702-711, 2007.

RIDGEN, J.E.; DRY, I.B.; KRAKE, L.R.; REZAIAN, M.A. Plant virus DNA replication processes in *Agrobacterium*: insight into the origins of the geminiviruses? **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, p. 10280-10284, 1996.

RODRÍGUEZ-PARDINA, P.E.; ZERBINI, F.M.; DUCASSE, D.A. Genetic diversity of begomoviruses infecting soybean, bean and associated weeds in Northwestern Argentina. **Fitopatologia Brasileira**, Brasília, v. 31, p. 342-348, 2006.

ROJAS, A.; KVARNHEDEN, A.; MARCENARO, D.; VALKONEN, J.P.T. Sequence characterization of Tomato leaf curl Sinaloa virus and Tomato severe leaf curl virus: Phylogeny of New World begomoviruses and detection of recombination. **Archives of Virology**, New York, v. 150, p. 1281-1299, 2005a.

ROJAS, M.R.; HAGEN, C.; LUCAS, W.J.; GILBERTSON, R.L. Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. **Annual Review of Phytopathology**, Palo Alto, v. 43, p. 361-394, 2005b.

ROJAS, M.R.; JIANG, H.; SALATI, R.; XOCONOSTLE-CAZARES, B.; SUDARSHANA, M.R.; LUCAS, W.J.; GILBERTSON, R.L. Functional analysis of proteins involved in movement of the monopartite begomovirus, tomato yellow leaf curl virus. **Virology**, New York, v. 291, p. 110-125, 2001.

ROJAS, M.R.; NOUEIRY, A.O.; LUCAS, W.J.; GILBERTSON, R.L. Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. **Cell**, Cambridge, v. 95, p. 105-113, 1998.

ROMAY, G.; CHIRINOS, D.; GERAUD-POUEY, F.; DESBIEZ, C. Association of an atypical alphasatellite with a bipartite New World begomovirus. **Archives of Virology**, New York, v. 155, p. 1843-1847, 2010.

RUIZ-MEDRANO, R.; XOCONOSTLE-CAZARES, B.; LUCAS, W.J. The phloem as a conduit for inter-organ communication. **Current Opinion in Plant Biology**, Cambridge, v. 4, p. 202-209., 2001.

RYBICKI, E.P. A phylogenetic and evolutionary justification for three genera of Geminiviridae. **Archives of Virology**, New York, v. 139, p. 49-77, 1994.

SANDERFOOT, A.A.; INGHAM, D.J.; LAZAROWITZ, S.G. A viral movement protein as a nuclear shuttle. The geminivirus BR1 movement protein contains domains essential for interaction with BL1 and nuclear localization. **Plant Physiology**, Washington, v. 110, p. 23-33, 1996.

SANDERFOOT, A.A.; LAZAROWITZ, S.G. Cooperation in viral movement: The geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. **Plant Cell**, Rockville, v. 7, p. 1185-1194, 1995.

SANTOS, C.D.G.; GONÇALVES, M.F.B.; OLIVEIRA, O.R. Detecção por ELISA de begomovírus em plantas daninhas presentes em áreas produtoras de tomateiro no estado do Ceará. **Fitopatologia Brasileira**, Brasília, v. 28 (Supl.), p. 252, 2003.

SANZ, A.I.; FRAILE, A.; GARCÍA-ARENAL, F.; ZHOU, X.; ROBINSON, D.J.; KHALID, S.; BUTT, T.; HARRISON, B.D. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. **Journal of General Virology**, London, v. 81, p. 1839-1849, 2000.

SAUNDERS, K.; BEDFORD, I.D.; BRIDDON, R.W.; MARKHAM, P.G.; WONG, S.M.; STANLEY, J. A unique virus complex causes *Ageratum yellow vein disease*. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 97, p. 6890-6895, 2000.

SAUNDERS, K.; BEDFORD, I.D.; STANLEY, J. Pathogenicity of a natural recombinant associated with ageratum yellow vein disease: implications for geminivirus evolution and disease aetiology. **Virology**, New York, v. 282, p. 38-47, 2001.

SAUNDERS, K.; BEDFORD, I.D.; STANLEY, J. Adaptation from whitefly to leafhopper transmission of an autonomously replicating nanovirus-like DNA component associated with ageratum yellow vein disease. **Journal of General Virology**, London, v. 83, p. 907-913, 2002.

SAUNDERS, K.; STANLEY, J. A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: Evidence for interfamilial recombination between plant DNA viruses. **Virology**, New York, v. 264, p. 142-152, 1999.

SEAL, S.E.; JEGER, M.J.; VAN DEN BOSCH, F. Begomovirus evolution and disease management. **Advances in Virus Research**, San Diego, v. 67, p. 297-316, 2006a.

SEAL, S.E.; VAN DEN BOSCH, F.; JEGER, M.J. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. **Critical Reviews in Plant Sciences**, Boca Raton, v. 25, p. 23-46, 2006b.

SELTH, L.A.; RANDLES, J.W.; REZAIAN, M.A. Agrobacterium tumefaciens supports DNA replication of diverse geminivirus types. **FEBS Letters**, Amsterdam, v. 516, p. 179-182., 2002.

SILVA, S.J.C. Detecção, caracterização molecular e diversidade genética de begomovirus que infectam fava (*Phaseolus lunatus* L.). Dep. de Agronomia, Universidade Federal de Alagoas, Rio Largo, AL, 2006. 87 p.

SILVA, A.K.F; SANTOS, C.D.G.; NASCIMENTO, A.K.Q. Transmissão de begomovírus de plantas daninhas para tomateiros pela mosca-branca. **Planta Daninha**, Rio de Janeiro, v.28, n.3, p. 507-514, 2010.

STANLEY, J. Analysis of African cassava mosaic virus recombinants suggest strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. **Virology**, New York, v. 206, p. 707-712, 1995.

STANLEY, J. Subviral DNAs associated with geminivirus disease complexes. **Veterinary Microbiology**, Amsterdam, v. 98, p. 121-129, 2004.

STANLEY, J.; BISARO, D.M.; BRIDDON, R.W.; BROWN, J.K.; FAUQUET, C.M.; HARRISON, B.D.; RYBICKI, E.P.;STENGER, D.C. Family *Geminiviridae*. In: Fauquet, C.M.; Mayo, M.A.; Maniloff, J.; Desselberger, U. Ball, L.A. (Ed.). Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. San Diego: Elsevier Academic Press, 2005. p.301-326.

STENGER, D.C.; REVINGTON, G.N.; STEVENSON, M.C.;BISARO, D.M. Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling-circle replication of a plant viral DNA. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 88, p. 8029-8033, 1991.

SUNG, Y.K.; COUTTS, R.H. Pseudorecombination and complementation between potato yellow mosaic geminivirus and tomato golden mosaic geminivirus. **Journal of General Virology**, London, v. 76, p. 2809-2815., 1995.

SUNTER, G.; BISARO, D.M. Transactivation of geminivirus AR1 and BR2 gene expression by the viral AL2 gene product occurs at the level of transcription. **Plant Cell**, Rockville, v. 4, p. 1321-1331, 1992.

SUNTER, G.; HARTITZ, M.D.; HORMUZDI, S.G.; BROUGH, C.L.; BISARO, D.M. Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for eficient DNA replication. **Virology**, New York, v. 179, p. 69-77, 1990.

TORRES-PACHECO, I.; GARZÓN-TIZNADO, J.A.; BROWN, J.K.; BECERRA-FLORA, A.; RIVERA-BUSTAMANTE, R. Detection and distribution of geminiviruses in Mexico and the Southern United States. **Phytopathology**, Saint Paul, v. 86, p. 1186-1192, 1996.

UNSELD, S.; RINGEL, M.; KONRAD, A.; LAUSTER, S.; FRISCHMUTH, T. Virus-specific adaptations for the production of a pseudorecombinant virus formed by two distinct bipartite geminiviruses from Central America. **Virology**, New York, v. 274, p. 179-188, 2000.

VAN DER WALT, E.; MARTIN, D.P.; VARSANI, A.; POLSTON, J.E.; RYBICKI, E.P. Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. **Virology Journal**, London, v. 5, p. -, 2008.

VANITHARANI, R.; CHELLAPPAN, P.; PITA, J.S.;FAUQUET, C.M. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. **Journal of Virology**, Washington, v. 78, p. 9487-9498, 2004.

VARSANI, A.; SHEPHERD, D.N.; DENT, K.; MONJANE, A.L.; RYBICKI, E.P.; MARTIN, D.P. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. **Virology Journal**, London, v. 6, p. -, 2009.

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.; LEFEUVRE, P.; LETT, J.M.; MARTIN, D.P. Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. **Journal of General Virology**, London, v. 89, p. 2063-2074, 2008.

VOINNET, O.; PINTO, Y.M.; BAULCOMBE, D.C. Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 96, p. 14147-14152, 1999.

VON ARNIM, A.; STANLEY, J. Inhibition of *African cassava mosaic virus* systemic infection by a movement protein from the related geminivirus *Tomato golden mosaic virus*. **Virology**, New York, v. 187, p. 555-564, 1992.

WANG, H.; BUCKLEY, K.J.; YANG, X.; BUCHMANN, R.C.; BISARO, D.M. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. **Journal of Virology**, Washington, v. 79, p. 7410-7418, 2005.

WANG, H.; HUANG, L.F.; COOPER, J.I. Analyses on mutation patterns, detection of population bottlenecks, and suggestion of deleterious-compensatory evolution among members of the genus Potyvirus. **Archives of Virology**, New York, v. 151, p. 1625-1633, 2006.

WU, B.; MELCHER, U.; GUO, X.; WANG, X.; FAN, L.; ZHOU, G. Assessment of codivergence of mastreviruses with their plant hosts. **BMC Evolutionary Biology**, London, v. 8, p. 335, 2008.

YAZDI, H.R.B.; HEYDARNEJAD, J.; MASSUMI, H. Genome characterization and genetic diversity of beet curly top Iran virus: a geminivirus with a novel nonanucleotide. **Virus Genes**, Norwell, v. 36, p. 539-545, 2008.

ZERBINI, F.M.; ANDRADE, E.C.; BARROS, D.R.; FERREIRA, S.S.; LIMA, A.T.M.; ALFENAS, P.F.; MELLO, R.N. Traditional and novel strategies for geminivirus management in Brazil. **Australasian Plant Pathology**, Perth, v. 34, p. 475-480, 2005.

ZHOU, X.; LIU, Y.; CALVERT, L.; MUNOZ, C.; OTIM-NAPE, G.W.; ROBINSON, D.J.;HARRISON, B.D. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. **Journal of General Virology**, London, v. 78, p. 2101-2111, 1997.

ZHOU, X.; LIU, Y.; ROBINSON, D.J.; HARRISON, B.D. Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. **Journal of General Virology**, London, v. 79, p. 915-923, 1998.

ZHOU, X.; XIE, Y.; TAO, X.; ZHANG, Z.; LI, Z.; FAUQUET, C.M. Characterization of DNAbeta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. **Journal of General Virology**, London, v. 84, p. 237-247, 2003.

CAPÍTULO II

Species diversity, phylogeny and genetic structure of begomovirus populations infecting leguminous weeds in Northeastern Brazil

1 **Species diversity, phylogeny and genetic structure of begomovirus populations infecting**
2 **leguminous weeds in Northeastern Brazil**
3

4 Sarah. J. C. Silva[1,3], Gloria. P. Castillo-Urquiza[1], Braz T. Hora Júnior[1], Iraildes P. Assunção[2],
5 Gaus S. A. Lima[2,3], Gilvan Pio-Ribeiro[3], Eduardo S. G. Mizubuti[1] and Francisco M. Zerbini[1*]
6
7 [1]Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG,
8 36570-000, Brazil.
9 [2]Departamento of Fitossanidade/CECA, Universidade Federal de Alagoas, Rio Largo, AL,
10 57100-000, Brazil.
11 [3]Departamento de Fitopatologia, Universidade Federal Rural de Pernambuco, Recife, PE,
12 52171-900, Brazil.
13
14 *To whom correspondence should be addressed. Email: zerbini@ufv.br
15
16

17 Silva, S.J.C., Castillo-Urquiza, G.P., Hora Júnior, B.T., Assunção, I.P., Lima, G.S.A., Pio-
18 Ribeiro, G., Mizubuti, E.S.G. & Zerbini, F.M. Species diversity, phylogeny and genetic
19 structure of begomovirus populations infecting leguminous weeds in Northeastern Brazil.
20 Journal of General Virology, *submitted*
21
22
23
24
25
26
27
28

29 **Summary**

30 Begomoviruses are whitefly-transmitted plant viruses with a circular, ssDNA genome.
31 Begomovirus diseases are a serious constraint to crop yields in most tropical and subtropical
32 regions of the world. In Brazil, begomoviruses affect mostly common bean and tomato
33 production. Weeds are considered to be begomovirus reservoirs as well as primary inoculum
34 sources for epidemics in crop plants. Although a number of studies have investigated the genetic
35 diversity of crop-infecting begomoviruses, such studies are lacking for begomoviruses infecting
36 weeds. We have carried out a survey of leguminous weeds (family Fabaceae) in four states of
37 the Brazilian Northeast. A total of 59 samples were collected, and 26 full-length begomovirus
38 genomes were amplified using rolling-circle amplification, cloned and sequenced. Sequence
39 analysis indicated the presence of six distinct viruses, including four novel species.
40 *Macroptilium lathyroides* was revealed as a common host for several of these viruses, and could
41 act as a mixing vessel from which recombinant viruses could emerge. Phylogenetic analysis
42 indicated that five of the viruses cluster with other Brazilian begomoviruses, but one of them
43 (Euphorbia yellow mosaic virus, EuYMV) clusters with viruses from other countries in Central
44 and South America. Strong evidence of recombination was found among isolates of
45 Macroptilium yellow spot virus (MaYSV). The genetic structure of the MaYSV population
46 indicates a high degree of genetic variability. Our results indicate that leguminous weeds are
47 reservoirs of several begomoviruses, and could play a significant role in begomovirus epidemics
48 both as inoculum sources and as sources of emerging novel viruses.

49

50 **Key words:** geminivirus, *Macroptilium*, recombination, MaYSV

51

52 Introduction

53 Viruses belonging to the family *Geminiviridae* have a genome comprised of circular
54 ssDNA molecules encapsidated in a twinned icosahedral capsid. Based on their genome
55 organization, host range and insect vectors, geminiviruses are classified into four different
56 genera: *Mastrevirus*, *Topocuvirus*, *Curtovirus* and *Begomovirus* (Stanley *et al.*, 2005).
57 Begomoviruses (whitefly-transmitted geminiviruses) constitute one of the most economically
58 important groups of plant viruses due to their high incidence and the severity of diseases they
59 cause in vegetable and field crops throughout tropical and subtropical regions of the world
60 (Briddon & Markham, 2001; Morales & Anderson, 2001). In South America, begomoviruses
61 are limiting factors to tomato (*Solanum lycopersicum*), common bean (*Phaseolus vulgaris*) and,
62 to a lesser extent, sweet and hot pepper (*Capsicum* spp.) production (Morales, 2006). The most
63 severely affected crops in Brazil are beans and tomatoes (Faria & Maxwell, 1999; Zerbini *et al.*,
64 2005). In beans (*P. vulgaris* and *P. lunatus*), golden mosaic caused by *Bean golden mosaic virus*
65 (BGMV) has been an important disease since the 1970's, and its dissemination has been
66 attributed to the increase in soybean cultivation (Costa, 1976). In tomatoes, the emergence of
67 begomovirus-associated diseases coincided with the introduction and spread of the B biotype of
68 *Bemisia tabaci* during the mid-1990's (Melo, 1992; Ribeiro *et al.*, 1998).

69 Begomoviruses are also associated with a wide range of weed species, which in some
70 cases act as primary inoculum sources for crop plants (Assunção *et al.*, 2006; Frischmuth *et al.*,
71 1997). Most of the weed species commonly reported as hosts belong to the families
72 Euphorbiaceae, Fabaceae, Malvaceae and Solanaceae (Morales & Anderson, 2001). Surveys
73 carried out to identify weed-associated viruses in Brazil indicate that, similarly to what is
74 observed for begomoviruses in crops, the species diversity of begomoviruses infecting weeds is
75 very high (Ambrozevicius *et al.*, 2002; Assunção *et al.*, 2006; Castillo-Urquiza *et al.*, 2008;
76 Paprotka *et al.*, 2010b). However, information on the genetic variability of begomoviruses in
77 wild and weed hosts is lacking.

78 It is believed that begomoviruses infecting wild and weed hosts in Brazil have been
79 horizontally transferred to crop plants, and that in the new host they rapidly evolved by
80 recombination an pseudorecombination, give rising to novel species (Castillo-Urquiza *et al.*,
81 2008; Fernandes *et al.*, 2009). Four independent lines of evidence give support to this
82 hypothesis. First, all begomoviruses reported so far in crops in Brazil are indigenous to the
83 country, and have never been reported elsewhere (except for neighboring Argentina; Rodríguez-
84 Pardina *et al.*, 2010). Second, the biological characterization of a number of crop-infecting
85 begomoviruses (*eg*, *Bean golden mosaic virus*, BGMV; *Tomato chlorotic mottle virus*,
86 ToCMoV; *Tomato rugose mosaic virus*, ToRMV; *Tomato yellow spot virus*, ToYSV; and
87 *Tomato yellow vein streak virus*, ToYVSV) indicated that weeds such as *Datura stramonium*,
88 *Macroptilium lathyroides*, *Nicandra physaloides* and *Solanum nigrum* are hosts (Albuquerque *et*
89 *al.*, 2010; Calegario *et al.*, 2007; Chagas *et al.*, 1981; Fernandes *et al.*, 2006; Ribeiro *et al.*,
90 2007). Third, begomoviruses originally detected in wild/weed plants, such as *Sida mottle virus*
91 (SiMoV) and *Sida micrantha mosaic virus* (SiMMV), have been found naturally infecting crop
92 species (Castillo-Urquiza *et al.*, 2010; Castillo-Urquiza *et al.*, 2007; Cotrim *et al.*, 2007).
93 Fourth, strong evidence of recombination and pseudorecombination events has been obtained
94 for the viruses which are prevalent in crop species such as tomato and common bean (Andrade
95 *et al.*, 2006; Inoue-Nagata *et al.*, 2006; Ribeiro *et al.*, 2007; Silva *et al.*, 2010).
96 Wild and weed hosts, whether indigenous or introduced, can also act as a reservoir of a
97 large number of plant viruses, and therefore may play a relevant role in viral epidemics in
98 several crops species (Seal *et al.*, 2006). The study of plant virus epidemics is greatly facilitated
99 when a population genetics approach is employed (Scherm *et al.*, 2006). The first step to study
100 viral population is to define their genetic structure, which refers to the degree of variability and
101 its distribution within and among subpopulations (Garcia-Arenal *et al.*, 2001). Knowledge of the
102 dynamics of genetic variability is essential to understand how populations evolve, with obvious
103 implications for the durability of disease management strategies (Seal *et al.*, 2006).

104 The purpose of this study was to characterize begomovirus populations infecting
105 leguminous weeds (family Fabaceae), as a step towards assessing their role as begomovirus
106 reservoirs in Northeastern Brazil.
107
108 **Results**
109 *Sequence comparisons and phylogenetic analysis*
110 Weed samples belonging to the genera *Canavalia*, *Calopogonium*, *Centrosema* and
111 *Macroptilium* (all in the family Fabaceae), displaying typical symptoms of begomovirus
112 infection (Figure 1), were collected in four states of Northeastern Brazil from May/2005 to
113 July/2010. A total of 59 samples were collected: 42 from AL, one from PB, nine from PE and
114 seven from SE (Supplementary Table S1). All 59 samples tested positive for the presence of a
115 begomovirus by PCR. A total of 19 full length DNA-A components were cloned, as well as 7
116 DNA-B components (Table 1). BLAST analysis and pairwise sequence comparisons of the
117 DNA-A clones indicated the presence of six begomovirus species (Table 1; Supplementary
118 Table S2). Clone SF114 from *Macroptilium atropurpureus* corresponds to an isolate of
119 *Euphorbia yellow mosaic virus* (EuYMV), with 97% identity with EuYMV (FJ619507). Clones
120 SF116, SF117 and SF129 obtained from *M. lathyroides* corresponds to isolates of BGMV, with
121 89-90% identity with the type isolate from common bean (M88686). One clone, SF102 from
122 *Macroptilium lathyroides,* represents a novel species which is most closely related to ToCMoV
123 (AF490004, 86% identity), for which the name *Macroptilium* yellow net virus (MaYNV) is
124 proposed. Clone SJC115 from *Centrosema brasilianum* also corresponds to a new species, most
125 closely related to ToYSV (DQ336350, 79% identity) and for which the name *Centrosema*
126 yellow spot virus (CenYSV) is proposed. A third new species is represented by clones SF118,
127 SK139, SF146, SJ160, SK161, SK162, SJ168, SK169, SK172, SJH173, SJ174 and SJ176, from
128 *Calopogonium mucunoides*, *Canavalia* sp. and *M. lathyroides*, which is also mostly closely to
129 ToCMoV (75-80% identity) and for which the name *Macroptilium* yellow spot virus (MaYSV)
130 is proposed. A fourth novel species is represented by clone SK175 from *M. lathyroides*, which

131 is most closely related to BGMV (M88686, 85% identity) and for which the name *Macroptilium*
132 yellow vein virus (MaYVV) is proposed.

133 Isolates of the four novel species display <85% sequence identity amongst themselves
134 (Supplementary Table S2). The genomes of all four novel species showed a typical bipartite,
135 New World begomovirus organization, with five ORFs in the DNA-A and two in the DNA-B.
136 The common regions (CR) have the conserved nonanucleotide (5'TAATATT/AC3') as part of a
137 stem-loop in the origin of replication. Cognate DNA-A and DNA-B components have identical
138 iterons, but the iterons are different among the six species: GGTG/GGTG for MaYNV and
139 MaYVV, GGAGT/GGAGT for CenYSV, GGAG/GGAG for MaYSV (data not shown).

140 A phylogenetic tree based on the complete DNA-A nucleotide sequence of the
141 begomoviruses from leguminous weeds and other Brazilian begomoviruses was constructed
142 used Bayesian inference, with the nucleotide substitution model GTR+I+G (Figure 2). The
143 weed-infecting begomoviruses were placed in three major monophyletic clusters within the tree.
144 The first cluster, with 98% Bayesian posterior probability (Bpp), includes the EuYMV and
145 BGMV isolates (SF114, SF116, SF117 and SF129), the novel species CenYSV (SJC115) and
146 MaYVV (SFK175), and other bean-, tomato- and weed-infecting begomoviruses. Within this
147 major cluster, CenYSV grouped with tomato- and weed-infecting begomoviruses, and MaYVV
148 grouped with BGMV. The second major cluster, with 100% Bpp, includes the new species
149 MaYSV (SF146 plus 11 additional clones) and *Blainvillea yellow spot virus* (BlYSV), also a
150 weed-infecting begomovirus. The third major cluster, with 100% Bpp, comprises the novel
151 species MaYNV (SF102) and ToCMoV.

152 A second phylogenetic tree based on the complete DNA-A sequences of viruses from
153 leguminous weeds and other begomoviruses from the Americas was constructed
154 (Supplementary Figure S1). The viruses within this tree clustered into four major groups.
155 Clusters 1 and 4 comprised only non-Brazilian viruses. Cluster 2 includes EuYMV, two weed-
156 infecting viruses from Brazil (*Sida yellow leaf curl virus*, SiYLCV, and *Tomato common mosaic*
157 *virus*, ToCmMV), plus several viruses from other countries in the Americas. The other five
158 begomoviruses from leguminous weeds were grouped in Cluster 3, which mainly comprises

159 Brazilian begomoviruses that infect bean, tomato, passionfruit and weeds. Within this cluster,
160 the different viruses were grouped identically as in the tree containing only Brazilian viruses.
161
162 *Recombination analysis*
163 Since we found several new begomovirus species infecting weeds in the four sampling
164 areas, but also found previously described species, including one (BGMV) which was described
165 more than 40 years ago, we wanted to investigate whether recombination events contributed to
166 the emergence of the new species. Therefore, we used neighbor-net/reticulate network analysis
167 to detect possible recombination events. Phylogenetic relationships inferred by neighbor-net
168 analysis based on a data set consisting of all Brazilian begomoviruses, including our viruses
169 from leguminous weeds, reveled clear evidence of multiple recombination events (Figure 3A).
170 Strong evidence for recombination was found in cluster I, containing the 12 MaYSV clones.
171 Weaker evidence was observed in clusters II, III, IV and V. These results were confirmed using
172 a second data set comprised only of the viruses from leguminous weeds (Figure 3B). Evidence
173 for recombination was again obtained when the analysis was restricted to the 12 MaYSV clones,
174 and was reinforced by phylogenetic inconsistency observed for SJ160, SJ168 and SK172, which
175 always grouped separately from the others nine isolates (Figure 3C).
176 To further investigate these putative recombination signals, the same three sets of
177 sequences were analyzed using the RDP3 package. This analysis identified many unique
178 recombination signals. To omit unreliable signals we selected only recombination events
179 supported by at least four different methods. A strongly supported recombination event was
180 detected involving MaYSV clones SF118 and SF146, with breakpoints at the CP and Rep
181 coding regions. This event was detected with all three data sets (Supplementary Tables S3, S4,
182 and S5), with BlYSV and MaYNV (SF102) identified as putative parents when all Brazilian
183 begomoviruses were included in the analysis (Supplementary Table 5). A recombination event
184 also with breakpoints at the CP and Rep was identified for MaYSV clones SJ160, SJ168 and
185 SK172 with the three data sets (Supplementary Tables S3, S4 and S5), with possible parents
186 varying depending on the data set: SK162 when only MaYSV isolates were analyzed

187 (Supplementary Table S3), CenYSV (SJC115) when all viruses from leguminous weeds were
188 included (Supplementary Table S4), and CenYSV and BlYSV when all Brazilian
189 begomoviruses were included (Supplementary Table S5). A recombination event in the Rep
190 region was detected for the three BGMV isolates (SF116, SF117, SF129) and MaYVV
191 (SK175), with one of the parents identified as *Sida Brazil virus* (SiBV) (Supplementary Table
192 S5). A recombination event was observed in the Rep region of MaYSV clones SF139, SK161,
193 SK162, SK169, SJH173, SJ174 and SJ176, with BlYSV identified as one of the parents
194 (Supplementary Table S5).
195
196 *Genetic structure of the MaYSV population*
197 The MaYSV population has a high degree of genetic variability, characterized by
198 genetic descriptors with considerably higher values than those observed for two populations of
199 tomato-infecting begomoviruses from Southeastern Brazil (Table 2).
200 Neutrality tests were used to assess for evidence of selection or demographic forces
201 acting on the MaYSV population. The four ORFs encoded by the DNA-A (Rep, Trap, Ren and
202 CP) varied in this regard. Negative values were obtained, but were not statistically supported,
203 for Tajima's D, Fu and Li's D and Fu and Li's F for Ren, Trap and CP (Table 3). The Rep ORF
204 showed positive values for these three tests, confirming the hypothesis of neutrality. The values
205 of dN/dS <1 for all ORFs are indicative of purifying selection acting on this population.
206
207 **Discussion**
208 The incidence and severity of diseases caused by geminiviruses has increased
209 dramatically in many areas of the world, including Brazil, due to the explosion of *Bemisia*
210 *tabaci* populations (Morales, 2006). The efficient dissemination and high poliphagy of the B
211 biotype of *B. tabaci* has enabled the transmission of indigenous begomoviruses to new
212 cultivated hosts, and the emergence of novel recombinant variants arising from mixed infections
213 (Ribeiro *et al.*, 2007). The important role that weeds and wild plants have played as sources of
214 begomoviruses for tomato and other important crops in Brazil is becoming increasingly clear. In

215 this study we investigated the species diversity and genetic structure of begomovirus
216 populations infecting leguminous weeds in Northeastern Brazil to determine the significance of
217 these hosts as begomovirus reservoirs.

218 In this study, six different begomoviruses were found out of 19 DNA-A clones: an
219 isolate of EuYMV obtained from *Macroptilium atropurpureum*, three BGMV isolates from
220 *Macroptilium lathyroides*, and four new species, one of them infecting *Centrosema brasilianum*
221 and three infecting *Canavalia* sp., *Calopogonium mucunoides*, *M. lathyroides* and *M.*
222 *atropurpureum*. This result indicates a high species diversity of begomoviruses infecting
223 leguminous weeds in Brazil, similarly to what has been observed for malvaceous and
224 solanaceous weed species (Castillo-Urquiza *et al.*, 2008; Jovel *et al.*, 2004; Paprotka *et al.*,
225 2010b). Furthermore, it indicates that *Macroptilium* spp. harbor many distinct begomoviruses,
226 and therefore may act as "mixing vessels" in which recombinant viruses may arise at high
227 frequency. *M. lathyroides* has been reported as a host of distinct begomoviruses in Central
228 America and the Caribbean (Idris *et al.*, 2003), altough it had been previously ruled out as an
229 inoculum source for begomovirus epidemics in Jamaica (Roye *et al.*, 1999). Our results indicate
230 that MaYSV, one of the new species, is capable of infecting at least three weed species (besides
231 *M. lathyroides*, it also infects *Calopogonium mucunoides* and *Canavalia* sp.). MaYSV was
232 detected in 12 (out of 17) samples collected in three different states, and therefore seems to be
233 the most common begomovirus in leguminous weeds in Northeastern Brazil. However, it will
234 be necessary to conclude the analysis of all 59 collected samples in order to confirm this
235 assumption.

236 Phylogenetic analyses based on DNA-A sequences begomoviruses from the Americas
237 showed that the four new species cluster with Brazilian viruses. The twelve isolates that
238 represent the species MaYSV formed a monophyletic group with another weed-infecting
239 begomovirus, BlYSV obtained from *Blainvillea rhomboidea* (Castillo-Urquiza *et al.*, 2008).
240 MaYNV and CenYSV grouped with tomato-infecting begomoviruses. MaYVV and three
241 BGMV isolates clustered with BGMV. Interestingly, EuYMV was placed in a group comprising
242 viruses from Mexico, Central and South America, including *Sida yellow leaf curl virus*

243 (SiYLCV), *Tomato common mosaic virus* (ToCmMV) and *Abutilon Brazil virus* (AbBV) which
244 have also been obtained from samples collected in Brazil (Castillo-Urquiza *et al.*, 2008;
245 Paprotka *et al.*, 2010a). Therefore, contrary to earlier beliefs, the Brazilian begomoviruses do
246 not collectively form a distinct and well separated monophyletic group relative to other viruses
247 from the Americas. The continent-wide phylogeographical mixing of begomovirus species in
248 South America is in fact reminiscent of that seen in African begomoviruses (Bull *et al.*, 2006;
249 Lefeuvre *et al.*, 2007b). In the past, South American begomoviruses also apparently segregated
250 into crop- and weed-infecting clades (Rojas *et al.*, 2005), but in the current scenario it is now
251 clear that most "crop-infecting" clades also contain an assortment of "weed-infecting" viruses
252 (Albuquerque *et al.*, 2010).

253 Accumulating evidence suggests that recombination is a common and important source
254 of genetic diversity in Brazilian begomoviruses (Galvão *et al.*, 2003; Inoue-Nagata *et al.*, 2006;
255 Ribeiro *et al.*, 2007). Recombinant begomoviruses have been directly implicated in the
256 emergence of new diseases and epidemics on crops in many countries (Garcia-Andres *et al.*,
257 2007a; Garcia-Andres *et al.*, 2006; Garcia-Andres *et al.*, 2007b; Lefeuvre *et al.*, 2010; Pita *et*
258 *al.*, 2001, Monci, 2002 #3748). Neighbor-net analysis indicated the presence of strong
259 recombination signals among the begomoviruses infecting leguminous weeds, particularly for
260 the MaYSV isolates. These results were confirmed using RDP3. Our analysis revealed that
261 recombination is a common event among begomoviruses in leguminous weeds. We found
262 strong evidence that MaYSV isolates SF118 and SF146 are recombinants, with BlYSV and
263 MaYNV as parents. A similarly strong evidence for recombination was found for MaYSV
264 isolates SJ160, SJ168 and SK172, for which the parents were identified as BlYSV and
265 CenYSV. BlYSV was also identified one of the parents of MaYSV isolates SK139, SK161,
266 SK162, SK169, SJH173, SJ174 and SJ176. The close relationship between MaYSV and BlYSV
267 was confirmed by phylogenetic analysis, in which these two viruses formed a group with 100%
268 Bpp. Interestingly, BlYSV has been found, so far, only in the weed *Blainvillea rhomboidea*,
269 from the family Asteraceae (Castillo-Urquiza *et al.*, 2008). It remains to be demonstrated
270 whether MaYSV and BlYSV share a common host.

271 The recombination events detected occurred primarily in the CP and Rep coding
272 regions. However, one recombination breakpoint was found in the common region (CR) of
273 MaYSV isolates SF118 and SF146. The CR is well characterized as a 'hot spot' of
274 recombination (Padidam *et al.*, 1999). Although coding regions are generally less susceptible to
275 recombination (Lefeuvre *et al.*, 2007a), the begomovirus CP and Rep coding regions have been
276 demonstrated to be recombination hot spots (Garcia-Andres *et al.*, 2007b; Lefeuvre *et al.*,
277 2007b).

278 Although this is the first report of the species MaYSV, it appears to be widely
279 distributed in the Brazilian Northeast, having been detected in the states of Alagoas, Paraíba and
280 Sergipe. Determination of the genetic structure of the MaYSV population demonstrated that
281 genetic variability is very high, with each isolate representing a single haplotype. This high
282 diversity is further demonstrated by high rates of nucleotide diversity, haplotype diversity and
283 mutation. These values were considerably higher than those observed for two populations of
284 tomato-infecting begomoviruses from Southeastern Brazil (Castillo-Urquiza *et al.*, 2010), and
285 were similar to those observed for a BGMV population obtained from lima bean (*Phaseolus*
286 *lunatus*) samples collected in Alagoas state (Ramos-Sobrinho *et al.*, 2010). Therefore, it seems
287 that viruses infecting weed/wild hosts have a greater degree of genetic variability compared to
288 viruses infecting crop species.

289 As with all viruses, the evolution of begomoviruses depends primarily on mutations.
290 There is evidence that the rapid evolution of geminiviruses is, at least in part, driven by
291 mutational processes acting specifically on ssDNA (Harkins *et al.*, 2009). High mutations rates,
292 similar to those observed for RNA viruses, have been estimated for the begomoviruses *Tomato*
293 *yellow leaf curl China virus* (TYLCCNV), *Tomato yellow leaf curl virus* (TYLCV), *East*
294 *African cassava mosaic virus* (EAMCV) and for the mastrevirus *Maize streak virus* (MSV)
295 (Duffy & Holmes, 2008; Duffy & Holmes, 2009; Ge *et al.*, 2007; Harkins *et al.*, 2009).
296 However, it has been shown that Brazilian BGMV isolates have an unusually low (for
297 begomoviruses) degree of genetic variability (Faria & Maxwell, 1999). This study was
298 conducted before RCA greatly simplified the cloning of full-length begomovirus genomes and

299 DNA sequencing technologies became widely available at a low cost. Therefore, a limited
300 number of isolates was completely sequenced, possibly underestimating the true genetic
301 variability of the virus. Indeed, our own studies conducted with a BGMV population infecting
302 lima bean showed that the variability within this species is high (Ramos-Sobrinho *et al.*, 2010).

303 Neutrality tests were performed to assess whether there was evidence of selection or
304 demographic forces acting on the MaYSV population. The negative values obtained for
305 Tajima's D, Fu and Li's *D** and Fu and Li's *F** tests were not statistically supported. The dN/dS
306 ratio was used to quantify selection pressures acting on protein-coding regions of the MaYSV
307 population. This measure quantifies selection pressures by comparing the rate of substitutions at
308 silent sites (dS), which are presumed neutral, to the rate of substitutions at non-silent sites (dN),
309 which possibly are undergoing a process of selection. The dN/dS ratio is expected to exceed
310 unity when natural selection promotes changes in the protein sequence (diversifying selection),
311 whereas a ratio less than unity is expected if natural selection suppresses protein changes
312 (purifying selection) (Yang & Bielawski, 2000). We found dN/dS < 1 values for MaYSV,
313 indicating the occurrence of purifying selection. Purifying selection and population expansion
314 were concluded to be the major evolutionary forces acting on ToYVSV and ToCmMV in
315 tomato (Castillo-Urquiza *et al.*, 2008). These results suggest that the MaYSV population may be
316 under the influence of purifying selection or underwent a recent expansion, so that the
317 occurrence of mutations is not sufficient to fully explain its genetic variability, and reinforce the
318 possible influence of additional evolutionary forces such as migration and recombination upon
319 the population.

320 Our findings indicate that leguminous weeds such as *Macroptilium lathyroides*, *M.*
321 *atropurpureum*, *Canavalia sp.* and *Centrosema brasilianum* constitute important reservoirs of
322 begomovirus species. *Macroptilium* spp. may also act as a mixing vessel that facilitates the
323 emergence of novel viruses by recombination. This hypothesis is reinforced by the detection of
324 recombination events in the MaYSV population. We conclude that recombination as well as
325 mutation is an important evolutionary process in the genetic diversification of the MaYSV
326 population. Additional studies are necessary to demonstrate that weed species play an active

327 role in begomovirus epidemics in crop plants, either by acting as primary inoculum sources or
328 as a continuous source of novel viruses, which could disrupt management strategies based on
329 the deployment of resistance genes.

330

331 **Methods**

332 *Sample colletion, processing and storage*

333 Surveys of leguminous weeds were carried out in locations throughout the states of
334 Alagoas (AL), Paraíba (PB), Pernambuco (PE) and Sergipe (PE) (Figure 4). Plants displaying
335 symptoms of mosaic, yellowing and stunting typical of begomovirus infection were
336 preferentially collected. Samples were desiccated by pressing and stored at -80ºC.

337

338 *DNA amplification and cloning*

339 DNA extraction was carried out from dried leaves according to Doyle & Doyle (1987).
340 To confirm the presence of begomoviruses, PCR was carried out using universal primers for
341 members of the genus (Rojas *et al.*, 1993). Full length viral genomes were amplified from PCR-
342 positive samples by rolling-circle amplification (RCA) (Inoue-Nagata *et al.*, 2004), cloned in
343 pBLUESCRIPT KS + (Stratagene) after monomerization with the restriction enzymes *Bam*H I,
344 *Cla* I, *Eco*R I, *Hin*d III, *Kpn* I, *Pst* I, *Sac* I or *Spe* I, and sequenced at Macrogen Inc. (Seoul,
345 South Korea) by primer walking.

346

347 *Sequence comparisons and phylogenetic analysis*

348 DNA-A nucleotide sequences were initially submitted to a BLAST search for
349 preliminary species assignment based on the 89% threshold level established by the
350 *Geminiviridae* Study Group of the ICTV (Fauquet *et al.*, 2008). Additional pairwise nucleotide
351 sequence comparisons were made with DNAMan version 4.0. using the Optimal Alignment
352 option with the following parameters: Ktuple = 2, Gap penalty = 7, Gap open = 10, Gap
353 extension = 5. Nucleotide sequences of begomoviruses used in the recombination and
354 phylogenetic analyses (see Supplementary Table S6 for the viruses and GenBank accession

355 numbers used in the analyses) were aligned using the Muscle module in Mega 5.0 (Tamura *et*
356 *al.*, 2007). Phylogenetic analysis was performed using Bayesian inference and Markov chain
357 Monte Carlo simulation implemented in MrBayes ver 3.0 (Ronquist & Huelsenbeck, 2003).
358 Bayesian analysis was conducted on the aligned data set after the nucleotide substitution model
359 was determined by MrModeltest v. 2.2 (Nylander, 2004). The Markov Chain Monte Carlo
360 (MCMC) analysis of four chains started with a heating parameter of 0.1 from a random tree
361 topology and lasted 5,000,000 generations. Trees were saved each 100 generations, resulting in
362 50,000 saved trees. Burn-in was set at 1,250,000 generations after which the likelihood values
363 were stationary, leaving 37,000 trees from which the 50% majority rule consensus trees and
364 posterior probabilities were calculated.
365
366 *Recombination analysis*
367 Phylogenetic network analysis for evidence of recombination was performed with the
368 Neighbor-Net method implemented in the program SplitsTree4 (Huson & Bryant, 2006).
369 Analysis of potential recombination events was carried out using the Recombination Detection
370 Program (RDP) ver. 3.0 (Martin *et al.*, 2010) using default parameters.
371
372 *Genetic structure of the MaYSV population*
373 The main descriptors of genetic variability were quantified: number of polymorphic
374 sites, total number of mutations (η), average number of nucleotide differences (k), nucleotide
375 diversity (π), number of haplotypes, haplotype diversity (Hd), Watterson's estimate of the
376 population mutation rate based on the total number of segregating sites (Theta-W) and on the
377 total number of mutations (Theta-Eta). Four types of neutrality tests were used to test the
378 hypothesis of occurrence of selection in populations: Tajima's D, Fu and Li's D* and F* and the
379 test based on the number of synonymous (Ds) and non-synonymous (Dns) substitutions with the
380 Pamilo-Bianchi-Li (PBL) model. These analyses were performed using the program DnaSP
381 version 5 (Rozas *et al.*, 2003).
382

383

384 **References**

385

386 Albuquerque, L. C., Martin, D. P., Avila, A. C. & Inoue-Nagata, A. K. (2010). Characterization
387 of tomato yellow vein streak virus, a begomovirus from Brazil. *Virus Genes* 40, 140-
388 147.
389 Ambrozevicius, L. P., Calegario, R. F., Fontes, E. P. B., Carvalho, M. G. & Zerbini, F. M.
390 (2002). Genetic diversity of begomoviruses infecting tomato and associated weeds in
391 Southeastern Brazil. *Fitopatol Bras* 27, 372-377.
392 Andrade, E. C., Manhani, G. G., Alfenas, P. F., Calegario, R. F., Fontes, E. P. B. & Zerbini, F.
393 M. (2006). *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with
394 a closer relationship to viruses from *Sida* sp., forms pseudorecombinants with
395 begomoviruses from tomato but not from *Sida*. *J Gen Virol* 87, 3687-3696.
396 Assunção, I. P., Listik, A. F., Barros, M. C. S., Amorim, E. P. R., Silva, S. J. C., Izael, O. S.,
397 Ramalho-Neto, C. E. & Lima, G. S. A. (2006). Diversidade genética de begomovírus
398 que infectam plantas invasoras na Região Nordeste. *Planta Daninha* 24, 239-244.
399 Briddon, R. W. & Markham, P. G. (2001). Complementation of bipartite begomovirus
400 movement functions by topocuviruses and curtoviruses. *Arch Virol* 146, 1811-1819.
401 Bull, S. E., Briddon, R. W., Sserubombwe, W. S., Ngugi, K., Markham, P. G. & Stanley, J.
402 (2006). Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. *J*
403 *Gen Virol* 87, 3053-3065.
404 Calegario, R. F., Ferreira, S. S., Andrade, E. C. & Zerbini, F. M. (2007). Characterization of
405 *Tomato yellow spot virus*, (ToYSV), a novel tomato-infecting begomovirus from Brazil.
406 *Braz J Agric Res* 42, 1335-1343.
407 Castillo-Urquiza, G. P., Alfenas-Zerbini, P., Beserra-Junior, J. E. A., Mizubuti, E. S. G.,
408 Varsani, A., Martin, D. P. & Zerbini, F. M. (2010). Genetic structure of tomato-
409 infecting begomovirus populations in two tomato-growing regions of Southeastern
410 Brazil. In *Program and Abstracts, 6th International Geminivirus Symposium and 4th*
411 *International ssDNA Comparative Virology Workshop*. Guanajuato, Mexico.
412 Castillo-Urquiza, G. P., Beserra Jr., J. E. A., Bruckner, F. P., Lima, A. T. M., Varsani, A.,
413 Alfenas-Zerbini, P. & Zerbini, F. M. (2008). Six novel begomoviruses infecting tomato
414 and associated weeds in Southeastern Brazil. *Arch Virol* 153, 1985-1989.
415 Castillo-Urquiza, G. P., Beserra Junior, J. E. A., Alfenas-Zerbini, P., Varsani, A., Lima, A. T.
416 M., Barros, D. R. & Zerbini, F. M. (2007). Genetic diversity of begomoviruses infecting
417 tomato in Paty do Alferes, Rio de Janeiro state, Brazil. *Virus Rev Res* 12, 233.
418 Chagas, C. M., Barradas, M. M. & Vicente, M. (1981). Espécies hospedeiras do vírus do
419 mosaico dourado do feijoeiro. *Arq Inst Biol São Paulo* 48, 123-127.
420 Costa, A. S. (1976). Whitefly-transmitted plant diseases. *Annu Rev Phytopathol* 14, 429-440.
421 Cotrim, M. A., Krause-Sakate, R., Narita, N., Zerbini, F. M. & Pavan, M. A. (2007).
422 Diversidade genética de begomovírus em cultivos de tomateiro no Centro-Oeste
423 Paulista. *Summa Phytopathol* 33, 300-303.
424 Doyle, J. J. & Doyle, J. L. (1987). A rapid DNA isolation procedure for small amounts of fresh
425 leaf tissue. *Phytochem Bull* 19, 11-15.
426 Duffy, S. & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution
427 in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. *J Virol* 82, 957-
428 965.
429 Duffy, S. & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in
430 geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses. *J Gen*
431 *Virol* 90, 1539-1547.
432 Faria, J. C. & Maxwell, D. P. (1999). Variability in geminivirus isolates associated with
433 *Phaseolus* spp. in Brazil. *Phytopathology* 89, 262-268.
434 Fauquet, C. M., Briddon, R. W., Brown, J. K., Moriones, E., Stanley, J., Zerbini, F. M. & Zhou,
435 X. (2008). Geminivirus strain demarcation and nomenclature. *Arch Virol* 153, 783-821.

436 Fernandes, F. R., Cruz, A. R. R., Faria, J. C., Zerbini, F. M. & Aragão, F. J. L. (2009). Three
437 distinct begomoviruses associated with soybean in central Brazil. *Arch Virol* 154, 1567-
438 1570.
439 Fernandes, J. J., Carvalho, M. G., Andrade, E. C., Brommonschenkel, S. H., Fontes, E. P. B. &
440 Zerbini, F. M. (2006). Biological and molecular properties of *Tomato rugose mosaic*
441 *virus* (ToRMV), a new tomato-infecting begomovirus from Brazil. *Plant Pathol* 55,
442 513-522.
443 Frischmuth, T., Engel, M., Lauster, S. & Jeske, H. (1997). Nucleotide sequence evidence for the
444 occurrence of three distinct whitefly-transmitted, Sida-infecting bipartite geminiviruses
445 in Central America. *J Gen Virol* 78, 2675-2682.
446 Galvão, R. M., Mariano, A. C., Luz, D. F., Alfenas, P. F., Andrade, E. C., Zerbini, F. M.,
447 Almeida, M. R. & Fontes, E. P. B. (2003). A naturally occurring recombinant DNA-A
448 of a typical bipartite begomovirus does not require the cognate DNA-B to infect
449 *Nicotiana benthamiana* systemically. *J Gen Virol* 84, 715-726.
450 Garcia-Andres, S., Accotto, G. P., Navas-Castillo, J. & Moriones, E. (2007a). Founder effect,
451 plant host, and recombination shape the emergent population of begomoviruses that
452 cause the tomato yellow leaf curl disease in the Mediterranean basin. *Virology* 359,
453 302-312.
454 Garcia-Andres, S., Monci, F., Navas-Castillo, J. & Moriones, E. (2006). Begomovirus genetic
455 diversity in the native plant reservoir Solanum nigrum: Evidence for the presence of a
456 new virus species of recombinant nature. *Virology* 350, 433-442.
457 Garcia-Andres, S., Tomas, D. M., Sanchez-Campos, S., Navas-Castillo, J. & Moriones, E.
458 (2007b). Frequent occurrence of recombinants in mixed infections of tomato yellow leaf
459 curl disease-associated begomoviruses. *Virology* 365, 210-219.
460 Garcia-Arenal, F., Fraile, A. & Malpica, J. M. (2001). Variability and genetic structure of plant
461 virus populations. *Annu Rev Phytopathol* 39, 157-186.
462 Ge, L. M., Zhang, J. T., Zhou, X. P. & Li, H. Y. (2007). Genetic structure and population
463 variability of tomato yellow leaf curl China virus. *J Virol* 81, 5902-5907.
464 Harkins, G. W., Delport, W., Duffy, S., Wood, N., Monjane, A. L., Owor, B. E., Donaldson, L.,
465 Saumtally, S., Triton, G., Briddon, R. W., Shepherd, D. N., Rybicki, E. P., Martin, D. P.
466 & Varsani, A. (2009). Experimental evidence indicating that mastreviruses probably did
467 not co-diverge with their hosts. *Virology J* 6, -.
468 Huson, D. H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary
469 studies. *Mol Biol Evol* 23, 254-267.
470 Idris, A. M., Hiebert, E., Bird, J. & Brown, J. K. (2003). Two newly described begomoviruses
471 of *Macroptilium lathyroides* and common bean. *Phytopathology* 93, 774-783.
472 Inoue-Nagata, A. K., Albuquerque, L. C., Rocha, W. B. & Nagata, T. (2004). A simple method
473 for cloning the complete begomovirus genome using the bacteriophage phi 29 DNA
474 polymerase. *J Virol Met* 116, 209-211.
475 Inoue-Nagata, A. K., Martin, D. P., Boiteux, L. S., Giordano, L. D., Bezerra, I. C. & de Avila,
476 A. C. (2006). New species emergence via recombination among isolates of the Brazilian
477 tomato infecting Begomovirus complex. *Braz J Agric Res* 41, 1329-1332.
478 Jovel, J., Reski, G., Rothenstein, D., Ringel, M., Frischmuth, T. & Jeske, H. (2004). *Sida*
479 *micrantha* mosaic is associated with a complex infection of begomoviruses different
480 from *Abutilon mosaic virus*. *Arch Virol* 149, 829-841.
481 Lefeuvre, P., Lett, J. M., Reynaud, B. & Martin, D. P. (2007a). Avoidance of protein fold
482 disruption in natural virus recombinants. *PLoS Pathog* 3, e181.
483 Lefeuvre, P., Martin, D. P., Harkins, G., Lemey, P., Gray, A. J. A., Meredith, S., Lakay, F.,
484 Monjane, A., Lett, J. M., Varsani, A. & Heydarnejad, J. (2010). The spread of tomato
485 yellow leaf curl virus from the Middle East to the world. *PLoS Pathog* 6, e1001164.
486 Lefeuvre, P., Martin, D. P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A.,
487 Becker, N., Reynaud, B. & Lett, J. M. (2007b). Begomovirus 'melting pot' in the south-
488 west Indian Ocean islands: Molecular diversity and evolution through recombination. *J*
489 *Gen Virol* 88, 3458-3468.

490 Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefeuvre, P. (2010). RDP3: A
491 flexible and fast computer program for analyzing recombination. *Bioinformatics* 26,
492 2462-2463.
493 Melo, P. C. T. (1992). Mosca branca ameaça produção de hortaliças. Campinas, SP, Brazil:
494 Asgrow do Brasil Sementes Ltda., Technical Bulletin.
495 Morales, F. J. (2006). History and current distribution of begomoviruses in Latin America. *Adv*
496 *Virus Res* 67, 127-162.
497 Morales, F. J. & Anderson, P. K. (2001). The emergence and dissemination of whitefly-
498 transmitted geminiviruses in Latin America. *Arch Virol* 146, 415-441.
499 Nylander, J. A. A. (2004). MrModeltest v2. *Program distributed by the author. Evolutionary*
500 *Biology Centre, Uppsala University*.
501 Padidam, M., Sawyer, S. & Fauquet, C. M. (1999). Possible emergence of new geminiviruses
502 by frequent recombination. *Virology* 265, 218-224.
503 Paprotka, T., Metzler, V. & Jeske, H. (2010a). The complete nucleotide sequence of a new
504 bipartite begomovirus from Brazil infecting Abutilon. *Arch Virol* 155, 813-816.
505 Paprotka, T., Metzler, V. & Jeske, H. (2010b). The first DNA 1-like alpha satellites in
506 association with New World begomoviruses in natural infections. *Virology* 404, 148-
507 157.
508 Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S. & Fauquet, C. M. (2001).
509 Recombination, pseudorecombination and synergism of geminiviruses are determinant
510 keys to the epidemic of severe cassava mosaic disease in Uganda. *J Gen Virol* 82, 655-
511 665.
512 Ramos-Sobrinho, R., Silva, S. J. C., Silva, T. A. L., Ribeiro, S. G., Lima, G. S. A., Assunção, I.
513 P. & Zerbini, F. M. (2010). Genetic structure of a population of the begomovirus *Bean*
514 *golden mosaic virus* (BGMV) that infects lima bean (*Phaseolus lunatus* L.) in the state
515 of Alagoas, Brazil. In *Program and Abstracts, 6th International Geminivirus*
516 *Symposium and 4th International ssDNA Comparative Virology Workshop*.
517 Guadalajara, México.
518 Ribeiro, S. G., Bezerra, I. C., Resende, R. O., Lima, M. F., Resende, L. V. & Ávila, A. C.
519 (1998). New tomato geminiviruses in mixed infections in Brazil. In *2nd Internation*
520 *Worshop on Bemisia and Geminiviruses*, pp. Abstract P-63. San Juan, Porto Rico.
521 Ribeiro, S. G., Martin, D. P., Lacorte, C., Simões, I. C., Orlandini, D. R. S. & Inoue-Nagata, A.
522 K. (2007). Molecular and biological characterization of *Tomato chlorotic mottle virus*
523 suggests that recombination underlies the evolution and diversity of Brazilian tomato
524 begomoviruses. *Phytopathology* 97, 702-711.
525 Rodríguez-Pardina, P. E., Hanada, K., Laguna, I. G., Zerbini, F. M. & Ducasse, D. A. (2010).
526 Molecular characterisation and relative incidence of bean- and soybean-infecting
527 begomoviruses in northwestern Argentina. *Ann Appl Biol* 158, 69-78.
528 Rojas, M. R., Gilbertson, R. L., Russell, D. R. & Maxwell, D. P. (1993). Use of degenerate
529 primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses.
530 *Plant Dis* 77, 340-347.
531 Rojas, M. R., Hagen, C., Lucas, W. J. & Gilbertson, R. L. (2005). Exploiting chinks in the
532 plant's armor: Evolution and emergence of geminiviruses. *Annu Rev Phytopathol* 43,
533 361-394.
534 Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under
535 mixed models. *Bioinformatics* 19, 1572-1574.
536 Roye, M. E., Spence, J., McLaughlin, W. A. & Maxwell, D. P. (1999). The common weed
537 *Macroptilium lathyroides* is not a source of crop-infecting geminiviruses from Jamaica.
538 *Trop Agric* 76, 256-262.
539 Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP: DNA
540 polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496-
541 2497.
542 Scherm, H., Ngugi, H. K. & Ojiambo, P. S. (2006). Trends in theoretical plant epidemiology.
543 *Eur J Pl Pathol* 115, 61-73.

544 Seal, S. E., Van den Bosch, F. & Jeger, M. J. (2006). Factors influencing begomovirus evolution
545 and their increasing global significance: Implications for sustainable control. *Crit Rev*
546 *Pl Sci* 25, 23-46.
547 Silva, F. N., Lima, A. T. M., Rocha, C. S., Alves-Júnior, M., Hallwass, M., Inoue-Nagata, A. K.
548 & Zerbini, F. M. (2010). Recombination and pseudorecombination driving the
549 evolution of *Tomato severe rugose virus* and *Tomato rugose mosaic virus*: Two distinct
550 DNA-As sharing the same DNA-B. In *Program and Abstracts, 6th International*
551 *Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop*.
552 Guanajuato, Mexico.
553 Stanley, J., Bisaro, D. M., Briddon, R. W., Brown, J. K., Fauquet, C. M., Harrison, B. D.,
554 Rybicki, E. P. & Stenger, D. C. (2005). Family *Geminiviridae*. In *Virus Taxonomy*
555 *Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 301-326.
556 Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San
557 Diego: Elsevier Academic Press.
558 Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary
559 Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24, 1596-1599.
560 Yang, Z. & Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Tr*
561 *Ecol Evol* 15, 496-503.
562 Zerbini, F. M., Andrade, E. C., Barros, D. R., Ferreira, S. S., Lima, A. T. M., Alfenas, P. F. &
563 Mello, R. N. (2005). Traditional and novel strategies for geminivirus management in
564 Brazil. *Aust Pl Pathol* 34, 475-480.

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

Table 1. Full-length clones corresponding to bipartite begomovirus DNA-A and DNA-B obtained from samples of leguminous weeds collected in the Brazilian Northeastern states of Alagoas (AL), Paraíba (PB), Pernambuco (PE) and Sergipe (SE)

Sample code	Location	Host	Clones		Species Assignment*
			DNA-A	DNA-B	
102F	Murici, AL	*Macroptilium lathyroides*	SF102A	SF102B	MaYNV[†] (new)
114F	Caruaru, PE	*M. atropurpureum*	SF114A	SF114B	EuYMV
115F	Caruaru, PE	*Centrosema brasilianum*	SJC115A		CenYSV (new)
116F	Caruaru, PE	*M. lathyroides*	SF116A		BGMV
117F	Caruaru, PE	*M. lathyroides*	SF117A		BGMV
118F	Barra de Santana, PB	*M. lathyroides*	SF118A		MaYSV (new)
120F	Santana do Mundaú, AL	*M. lathyroides*		SF120B	n.a.[‡]
129F	Maceió, AL	*M. lathyroides*	SF129A		BGMV
139F	Cedro, SE	*M. lathyroides*	SF139A		MaYSV
145F	Messias, AL	*M. lathyroides*		SF145B	n.a.
146F	Maceió, AL	*M. lathyroides*	SF146A		MaYSV
148F	Maceió, AL	*M. atropurpureum*		SJ148B	n.a.
152F	Quipapá, PE	*M. atropurpureum*		SK152B	n.a.
160F	Batalha, AL	*M. lathyroides*	SJ160A		MaYSV
161F	Água das Flores, AL	*M. lathyroides*	SK161A		MaYSV
162F	Água das Flores, AL	*M. lathyroides*	SK162A		MaYSV
168F	Piranhas, AL	*Calopogonium mucunoides*	SJ168A		MaYSV
169F	Delmiro Gouveia, AL	*Calopogonium mucunoides*	SK169A		MaYSV
171F	Delmiro Gouveia, AL	*M. lathyroides*		SK175B	n.a.
172F	Inhapi, AL	*M. lathyroides*	SK172A		MaYSV
173F	Inhapi, AL	*Canavalia sp.*	SJH173A		MaYSV
174F	Palmeira dos Índios, AL	*M. lathyroides*	SJ174A		MaYSV
175F	Maceió, AL	*M. lathyroides*	SK175A		MaYVV (new)
176F	Maceió – AL	*M. lathyroides*	SJ176A		MaYSV

[*]Species assignment based on the ICTV-established criteria of 89% nucleotide sequence identity for the full-length DNA-A (Fauquet *et al.*, 2008).
[†]MaYNV, *Macroptilium* yellow net virus; EuYMV, *Euphorbia* yellow mosaic virus; CenYSV, *Centrosema* yellow spot virus; BGMV, Bean golden mosaic virus; MaYSV, *Macroptilium* yellow spot virus; MaYVV, *Macroptilium* yellow vein virus.
[‡]n.a., not assigned, since the cognate DNA-A was not cloned.

Table 2. Genetic structure of a population of *Macroptilium* yellow spot virus (MaYSV) obtained from leguminous weeds in Northeastern Brazil

Number of sequences	Genome size	*s*[*]	Eta[†]	*k*[‡]	π[§]	h	Hd[¶]	θ-w[#]	θ-Eta[*]
10	2658	402	419	150,177	0,0572	10	1,0	0,0537	0,0542

[*] Total number of segregating sites.
[†] Total number of mutations.
[‡] Average number of nucleotide differences between sequences (Tajima's estimate of the population mutation rate, θ).
[§] Nucleotide diversity.

[||] Haplotype number.
[¶] Haplotype diversity.
[#] Watterson's estimate of the population mutation rate based on the total number of segregating sites.
[*] Watterson's estimate of the population mutation rate based on the total number of mutations.

Table 3. Results of the different neutrality tests for each open reading frame (ORF) in the DNA-A of viral isolates comprising a population of *Macroptilium* yellow spot virus (MaYSV) obtained from leguminous weeds in Northeastern Brazil

ORF[*]	Tajima's *D*	Fu and Li's *D*	Fu and Li's *F*	dN/dS
Rep	0.8439 (ns)[†]	0.6996 (ns)	0.8319 (ns)	0.7631
Trap	-0.0477 (ns)	-0.3255 (ns)	-0.2892 (ns)	0.4545
REn	-0.0518 (ns)	-0.3323 (ns)	-0.2960 (ns)	0.2132
CP	-0.7991 (ns)	-1.4288 (ns)	-1.4366 (ns)	0.0643

[*]Rep, Replication-associated protein; Trap, Trans-activating protein; Ren, Replication enhancer protein; CP, Coat protein.

[†]ns, not significant values at $p < 0.10$

Supplementary Table S1. Location, year of collection and host species of the leguminous weed samples collected in four Northeastern Brazilian states from 2005 to 2010

Collection site	Year	Host	Sample code
Alagoas state			
Viçosa	2009	*Macroptilium lathyroides*	101F
Murici	2009	*Macroptilium lathyroides*	102F
Rio Largo	2009	*Senna sp.*	103F
Arapiraca	2009	*Senna sp.*	104F
Arapiraca	2009	*Crotalaria sp.*	105F
Messias	2010	*Macroptilium lathyroides*	107F
União dos Palmares	2010	*Macroptilium lathyroides*	108F
União dos Palmares	2005	*Macroptilium lathyroides*	119F
Santana do Mundaú	2005	*Macroptilium lathyroides*	120F
Messias	2005	*Macroptilium lathyroides*	124F
Arapiraca	2005	*Macroptilium lathyroides*	125F
Arapiraca	2005	*Macroptilium lathyroides*	126F
Maceió	2010	*Macroptilium lathyroides*	129F
Rio Largo	2010	*Calopogonium mucunoides*	131F
Chã Preta	2010	*Macroptilium lathyroides*	134F
Rio Largo	2009	*Senna sp.*	135F
Rio Largo	2010	*Calopogonium mucunoides*	137F
Porto Calvo	2010	*Macroptilium lathyroides*	138F
Messias	2010	*Macroptilium lathyroides*	145F
Maceió	2009	*Macroptilium lathyroides*	146F
Maceió	2010	*Macroptilium atropurpureum*	148F
Maceió	2010	*Macroptilium lathyroides*	149F
Maceió	2010	*Calopogonium mucunoides*	150F
Marechal Deodoro	2010	*Macroptilium lathyroides*	153F
Flexeiras	2010	*Macroptilium atropurpureum*	154F
Murici	2010	*Macroptilium lathyroides*	155F
Murici	2010	*Macroptilium lathyroides*	156F
Jaramataia	2010	*Macroptilium atropurpureum*	158F
Jaramataia	2010	*Macroptilium lathyroides*	159F
Batalha	2010	*Macroptilium lathyroides*	160F
Água das Flores	2010	*Macroptilium lathyroides*	161F
Água das Flores	2010	*Macroptilium lathyroides*	162F
São José da Tapera	2010	unknown	163F
São José da Tapera	2010	unknown	164F
São José da Tapera	2010	*Macroptilium lathyroides*	165F
Piranhas	2010	*Calopogonium mucunoides*	168F
Delmiro Gouveia	2010	*Calopogonium mucunoides*	169F
Delmiro Gouveia	2010	*Macroptilium lathyroides*	171F
Inhapi	2010	*Macroptilium lathyroides*	172F
Inhapi	2010	*Canavalia sp.*	173F
Palmeira dos Índios	2010	*Macroptilium lathyroides*	174F
Maceió	2010	*Macroptilium lathyroides*	175F
Maceió	2010	*Macroptilium lathyroides*	176F
Paraíba state			
Barra de Santana	2009	*Macroptilium lathyroides*	118F

Supplementary Table S2 (cont.)

Pernambuco state			
Ribeirão	2009	*Calopogonium mucunoides*	112F
Ribeirão	2009	*Calopogonium mucunoides*	113F
Caruaru	2009	*Macroptilium atropurpureum*	114F
Caruaru	2009	*Centrosema brasilianum*	115F
Caruaru	2010	*Macroptilium lathyroides*	116F
Caruaru	2010	*Macroptilium lathyroides*	117F
Goiana	2010	*Calopogonium mucunoides*	133F
Caruaru	2010	*Mimosa caesalpiniaefolia*	151F
Quipapá	2010	*Macroptilium atropurpureum*	152F
Sergipe state			
Neópolis	2009	*Macroptilium lathyroides*	109F
Neópolis	2010	*Senna sp.*	110F
Estância	2010	*Macroptilium lathyroides*	111F
Neópolis	2009	*Macroptilium lathyroides*	130F
Cedro	2009	*Macroptilium lathyroides*	139F
Aquibadã	2009	*Macroptilium lathyroides*	140F
Aquibadã	2009	*Macroptilium lathyroides*	141F

Supplementary Table S2. Percent identities between the complete DNA-A nucleotide sequences of the six begomovirus species detected in leguminous weeds in four states of Northeastern Brazil

	BGMV*	EuYMV	ToCMoV	ToYSV	SF102	SF114	SJC115	SF116	SF117	SF118	SF129	SK139	SF146	SJ160	SK161	SK162	SJ168	SK169	SK172	SJH173	SJ174	SK175	SJ176
BGMV	-	70	76	77	78	71	76	89	89	79	90	80	80	79	79	79	79	79	79	80	80	85	79
EuYMV		-	70	69	71	97	72	69	70	69	70	69	69	71	69	69	71	71	70	71	70	71	69
ToCMoV			-	74	86	69	72	74	74	76	74	75	77	80	77	77	80	77	79	78	77	75	78
ToYSV				-	75	71	79	76	76	75	76	74	74	77	74	74	77	74	77	74	74	75	75
SF102					-	71	72	76	71	81	78	76	81	79	76	78	79	76	79	78	79	77	79
SF114						-	72	70	70	70	71	70	69	71	70	70	71	70	71	70	70	71	69
SJC115							-	75	75	75	76	74	76	76	75	75	76	75	76	74	76	76	76
SF116								-	98	78	95	78	78	79	78	78	79	78	79	78	78	85	78
SF117									-	78	95	78	79	79	78	78	79	78	79	78	78	85	78
SF118										-	78	95	98	88	96	95	88	95	88	95	95	80	95
SF129											-	79	79	79	78	78	79	78	79	78	78	85	79
SK139												-	95	89	97	96	89	96	89	97	96	80	96
SF146													-	87	95	95	88	95	88	95	95	80	95
SJ160														-	90	90	98	90	98	89	90	79	89
SK161															-	97	90	98	90	98	97	80	97
SK162																-	90	99	90	99	99	80	96
SJ168																	-	90	98	89	90	78	89
SK169																		-	90	99	99	80	96
SK172																			-	90	91	79	89
SJH173																				-	99	79	96
SJ174																					-	80	96
SK175																						-	80
SJ176																							-

*BGMV, *Bean golden mosaic virus* (M88686); EuYMV, *Euphorbia yellow mosaic virus* (FJ619507); ToCMoV, *Tomato chlorotic mottle virus* (AF490004); ToYSV, *Tomato yellow spot virus* (DQ336350).

BGMV isolates are highlited in blue; EuYMV isolates are highlighted in red; *Macroptilium* yellow spot virus (MaYSV) isolates are highlighted in green.

Supplementary Table S3. Putative recombination events detected among isolates of *Macroptilium* yellow spot virus (MaYSV) infecting leguminous weeds in Northeastern Brazil

Clone/isolate	Parents	Breakpoints		*P*-value						
		Initial	Final	R[‡]	G	B	M	C	S	3S
SF118	unknown[*]	99[†]	1041	4.526×10^{-15}	1.780×10^{-14}	–[§]	2.985×10^{-13}	3.700×10^{-7}	–	2.057×10^{-21}
SF146	unknown	91	1073	4.526×10^{-15}	1.780×10^{-14}	–	2.985×10^{-13}	3.700×10^{-7}	–	2.057×10^{-21}
SJ160	SK162	398	1097	1.563×10^{-04}	–	–	5.294×10^{-05}	9.944×10^{-4}	–	4.574×10^{-3}
SJ168	SK162	242	1094	1.563×10^{-04}	–	–	5.294×10^{-05}	9.944×10^{-4}	–	4.574×10^{-3}
SK172	SK162	242	1094	1.563×10^{-04}	–	–	5.294×10^{-05}	9.944×10^{-4}	–	4.574×10^{-3}
SJ176	unknown	132	1144	4.526×10^{-15}	1.780×10^{-14}	–	2.985×10^{-13}	3.700×10^{-7}	–	2.057×10^{-21}

[*]When only the major parent is indicated, the minor parent has not been identified. "Unknown", neither parent identified.

[†]Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.

[‡]R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SisScan; 3S, 3SEQ.

[§] –, no recombination event detected.

Supplementary Table S4. Putative recombination events detected among begomoviruses infecting leguminous weeds in Northeastern Brazil

Clone/isolate	Parents	Breakpoints		*P*-value						
		Initial	Final	R[‡]	G	B	M	C	S	3S
SF114	SJ176[*]	606[†]	986	4.797×10^{-02}	–[§]	6.320×10^{-04}	4.825×10^{-02}	1.892×10^{-02}	6.204×10^{-03}	–
SJC115	SF114	1652	1925	4.436×10^{-03}	–	9.666×10^{-03}	3.365×10^{-02}	–	1.389×10^{-03}	–
SF116	SJC115	1753	2170	7.731×10^{-07}	1.393×10^{-02}	2.141×10^{-05}	1.252×10^{-04}	2.694×10^{-03}	1.821×10^{-10}	–
SF117	SJC115	1755	2172	7.731×10^{-07}	1.393×10^{-02}	2.141×10^{-05}	1.252×10^{-04}	2.694×10^{-03}	1.821×10^{-10}	–
SF118	SF102	439	903	1.237×10^{-19}	7.041×10^{-05}	1.838×10^{-19}	7.437×10^{-05}	5.060×10^{-09}	3.652×10^{-05}	7.535×10^{-06}
SF129	SJC115	1755	2172	7.731×10^{-07}	1.393×10^{-02}	2.141×10^{-05}	1.252×10^{-04}	2.694×10^{-03}	1.821×10^{-10}	–
SF146	SF102	410	902	1.237×10^{-19}	7.041×10^{-05}	1.838×10^{-19}	7.437×10^{-05}	5.060×10^{-09}	3.652×10^{-05}	7.535×10^{-06}
SJ160	SJC115	2127	2414	1.898×10^{-20}	–	1.127×10^{-17}	4.238×10^{-08}	4.393×10^{-05}	5.064×10^{-08}	2.529×10^{-09}
SJ168	SJC115	2143	2413	1.898×10^{-20}	–	1.127×10^{-17}	4.238×10^{-08}	4.393×10^{-05}	5.064×10^{-08}	2.529×10^{-09}
SK172	SJC115	2110	2413	1.898×10^{-20}	–	1.127×10^{-17}	4.238×10^{-08}	4.393×10^{-05}	5.064×10^{-08}	2.529×10^{-09}
SK175	SJC115	1765	2282	7.731×10^{-07}	1.393×10^{-02}	2.141×10^{-05}	1.252×10^{-04}	2.694×10^{-03}	1.821×10^{-10}	–

[*] When only the major parent is indicated, the minor parent has not been identified.

[†] Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.

[‡] R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SisScan; 3S, 3SEQ.

[§] –, no recombination event detected.

Supplementary Table S5. Putative recombination events detected among Brazilian begomoviruses, including the viruses infecting leguminous weeds in Northeastern Brazil

Clone/isolate	Parents	Breakpoints		*P*-value						
		Initial	Final	R[‡]	G	B	M	C	S	3S
SF114	Unknown[*]	1841[†]	2124	1.209 x 10^{-22}	–[§]	2.386 x 10^{-17}	3.149 x 10^{-11}	3.013 x 10^{-09}	–	2.780 x 10^{-03}
SJC115	Unknown	2150	2385	3.057 x 10^{-07}	6.318 x 10^{-07}	3.128 x 10^{-05}	4.548 x 10^{-05}	–	1.247 x 10^{-04}	1.555 x 10^{-04}
SF116	SiBV	1950	2537	1.203 x 10^{-07}	1.268 x 10^{-02}	7.559 x 10^{-05}	1.560 x 10^{-07}	1.577 x 10^{-05}	2.450 x 10^{-09}	–
SF117	SiBV	1745	2545	1.203 x 10^{-07}	1.268 x 10^{-02}	7.559 x 10^{-05}	1.560 x 10^{-07}	1.577 x 10^{-05}	2.450 x 10^{-09}	–
SF118	SF102	457	902	5.712 x 10^{-11}	1.828 x 10^{-04}	2.730 x 10^{-11}	6.196 x 10^{-07}	3.788 x 10^{-07}	8.405 x 10^{-05}	5.568 x 10^{-11}
	BlYSV	1787	2576	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SF129	SiBV	1941	2539	1.203 x 10^{-07}	1.268 x 10^{-02}	7.559 x 10^{-05}	1.560 x 10^{-07}	1.577 x 10^{-05}	2.450 x 10^{-09}	–
SK139	SF129	498	588	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1822	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SF146	SF102	438	825	5.712 x 10^{-11}	1.828 x 10^{-04}	2.730 x 10^{-11}	6.196 x 10^{-07}	3.788 x 10^{-07}	8.405 x 10^{-05}	5.568 x 10^{-11}
	BlYSV	1822	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SJ160	SF129	481	609	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x10^{-01}	–	2.177 x 10^{-22}	–	–
SK161	SF129	462	590	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1786	2475	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SK162	SF129	498	588	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1784	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SJ168	SF129	480	608	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x10^{-01}	–	2.177 x 10^{-22}	–	–
SK169	SF129	498	588	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1784	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SK172	SF129	480	608	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-22}	–	–
SJ173	SF129	437	1064	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1784	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SJ174	SF129	498	588	2.112 x 10^{-02}	1.256 x 10^{-02}	7.598 x 10^{-01}	–	2.177 x 10^{-02}	–	–
	BlYSV	1784	2448	2.251 x 10^{-08}	1.214 x 10^{-02}	7.886 x 10^{-07}	2.547 x 10^{-07}	1.032 x 10^{-05}	1.969 x 10^{-13}	6.282 x 10^{-04}
SK175	SiBV	1790	2572	1.203 x 10^{-07}	1.268 x 10^{-02}	7.559 x 10^{-05}	1.560 x 10^{-07}	1.577 x 10^{-05}	2.450 x 10^{-09}	–

[*] When only the major parent is indicated, the minor parent has not been identified. "Unknown", neither parent identified.

[†] Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.

[‡] R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SisScan; 3S, 3SEQ.

[§] –, no recombination event detected.

Supplementary Table S6. Begomoviruses used in pairwise sequence comparisons, phylogenetic and recombination analyses

Virus	Acronym	GenBank access # (DNA-A)
From Brazil		
Abutilon Brazil virus	AbBV	NC_014138
Bean golden mosaic virus	BGMV	M88686
Blainvillea yellow spot virus	BlYSV	EU710756
Cleome leaf crumple virus	ClLCrV	FN35999
Euphorbia yellow mosaic virus	EuYMV	FJ619507
Nicandra deforming necrosis virus	NDNV	n.a.
Okra mottle virus	OmoV	NC_011181
Passionfruit severe leaf distortion virus	PSLDV	NC_012786
Sida common mosaic virus	SiCmMV	EU710751
Sida mosaic Brazil virus	SiMBV	FN436001
Sida micrantha mosaic virus	SiMMV	NC_005330
Sida mottle virus	SiMoV	NC_004637
Sida yellow leaf curl virus	SiYLCV	EU710750
Sida yellow mosaic virus	SiYMV	NC_004639
Soybean blistering mosaic virus	SoBlMV	EF016486
Tomato chlorotic mottle virus	ToCMoV	AF490004
Tomato common mosaic virus	ToCmMV	NC_010835
Tomato golden mosaic virus	TGMV	NC_001507
Tomato mild mosaic virus	ToMlMV	EU710752
Tomato rugose mosaic virus	ToRMV	NC_002555
Tomato severe rugose virus	ToSRV	NC_009607
Tomato yellow spot virus	ToYSV	DQ336350
Tomato yellow vein streak virus	ToYVSV	NC_010949
From other countries in the Americas		
Abutilon mosaic virus	AbMV	NC_001928
Bean calico mosaic virus	BCaMV	NC_003504
Bean dwarf mosaic virus	BDMV	NC_001931
Bean golden yellow mosaic virus	BGYMV	NC_001439
Cabbage leaf curl virus	CaLCuV	NC_033866
Chino del tomate virus	CdTV	AF101476
Cotton leaf curl virus	CLCrV	NC_004580
Corchorus yellow spot virus	CoYSV	NC_008492
Curcubit leaf crumple virus	CuLCrV	NC_002984
Desmodium leaf distortion virus	DesLDV	NC_008494
Dicliptera yellow mosaic virus	DiYMV	NC_003856
Dicliptera yellow mosaic Cuba virus	DiYMCUV	AJ549960
Euphobia mosaic virus - Yucatan Peninsula	EUMV_YP	NC_008304
Macroptillium golden mosaic virus	MaGMV	NC_010952
Macroptillium mosaic Puerto Rico virus	MaMPR	NC_004097
Macroptilium yellow mosaic Florida virus	MaYMFV	NC_004009

Supplementary Table S1 (cont.)		
Macroptilium yellow mosaic virus	MaYMV	NC_010647
Melon chlorotic leaf curl virus	MCLCuV	NC_003865
Merremia mosaic virus	MeMV	NC_007965
Okra yellow mosaic Mexico virus	OYMMV	NC_014066
Okra yellow mottle Iguala virus	OYMoIV	AY751753
Pepper golden mosaic virus	PepGMV	NC_004101
Pepper huasteco yellow vein virus	PHYVV	NC_001359
Potato yellow mosaic Panama virus	PYMPV	NC_002048
Potato yellow mosaic virus	PYMV	NC_001934
Rhyncosia golden mosaic Sinaloa virus	RhGMSV	DQ406672
Rhyncosia golden mosaic virus	RhGMV	NC_010294
Rhyncosia rugose golden mosaic virus	RhRGMV	HM236360
Sida golden mosaic Costa Rica virus	SGMCRV	NC_004657
Sida golden mosaic Honduras virus	SGMHV	NC_004659
Sida golden mosaic virus	SGMV	NC_002046
Sida golden yellow vein virus	SiGYVV	NC_004635
Sida yellow mosaic Yucatan virus	SiYMYuV	NC_008779
Sida yellow vein virus	SiYVV	NC_004661
Squash leaf curl virus	SqLCV	NC_001936
Squash mild leaf curl virus	SqMLCV	NC_004645
Tomato Chino La Paz virus	ToChLPV	NC_005843
Tomato golden motlle virus	ToGMoV	NC_008058
Tobacco leaf curl Cuba virus	TLCCUV	AM050143
Tomato mosaic Havana virus	ToMHV	NC_003867
Tomato mottle Taino virus	ToMoTV	NC_001828
Tomato mottle virus	ToMoV	NC_001938
Tomato mild yellow leaf curl Aragua virus	ToMYLCAV	NC_009490
Tomato yellow leaf distortion virus	ToYLDV	FJ174698
Tomato yellow margin leaf curl virus	ToYMLCV	AY508998
Tomato severe leaf curl virus	ToSLCV	NC_004642
Tobacco yellow crinkle virus	TYCV	NC_011402
Wissadula golden mosaic virus	WGMV	NC_010948
Outgroup		
Tomato leaf curl New Delhi virus	TLCNDV	NC_004611

Figure 1. Symptoms in *Macroptilium* spp. infected by three novel begomoviruses. **A.** Reticulate yellow mosaic and growth reduction symptoms in the plant from which isolate SF102 (*Macroptilium* yellow net virus, MaYNV) was obtained. **B**. Yellow mosaic and vein banding symptoms in the plant from which isolate SK175 (Macroptilium yellow vein virus, MaYVV) was obtained. **C.** Yellow spot symptoms in the plant from which isolate SF146 (*Macroptilium* yellow spot virus, MaYSV) was obtained.

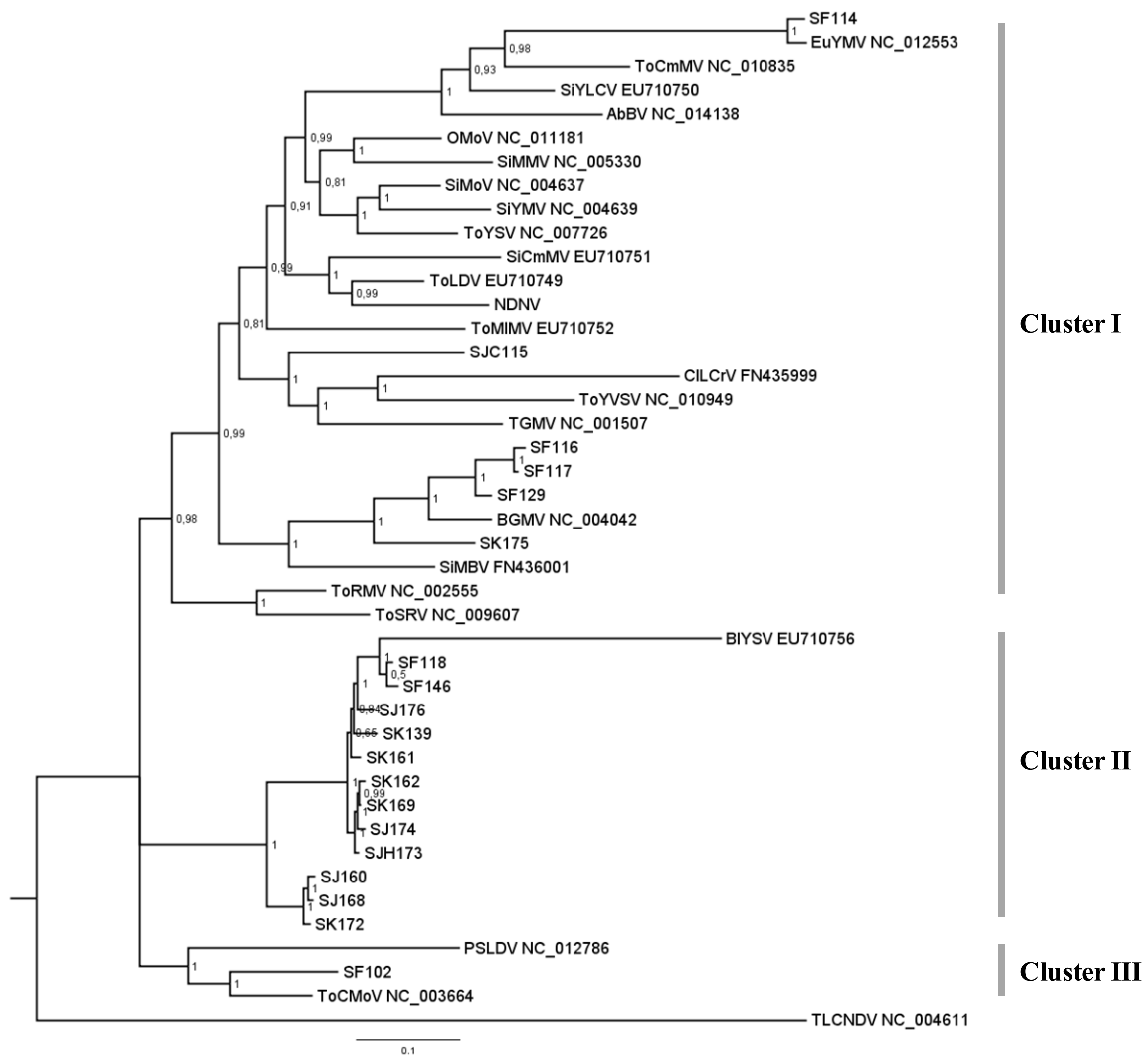


Figure 2. Bayesian 50% majority rule consensus tree of begomoviruses from leguminous weeds and other Brazilian begomoviruses (see Supplementary Table S1 for full virus names). Numbers at the nodes indicate Bayesian posterior probabilities.

A

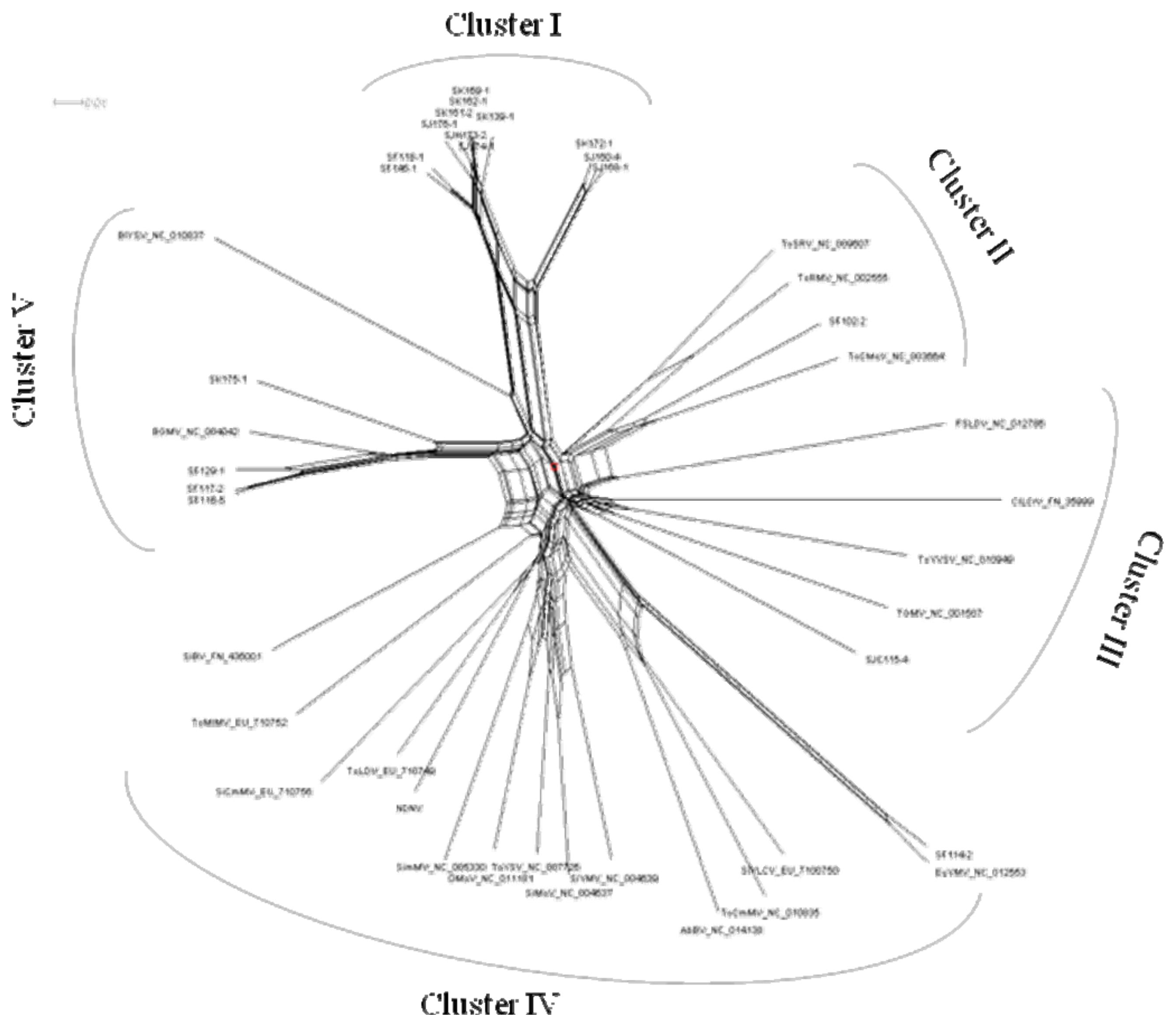


Figure 3. Phylogenetic evidence for recombination among (A) all Brazilian begomoviruses, inclusing those described in this work, (B) begomoviruses infecting leguminous weeds in Northeastern Brazil, and (C) a population of MaYSV obtained from leguminous weeds in Northeastern Brazil. Neighbor Net network analysis was performed using SplitsTree4. Formation of a reticular network rather than a single bifurcated tree is suggestive of recombination.

Fig. 3 (cont.)

B

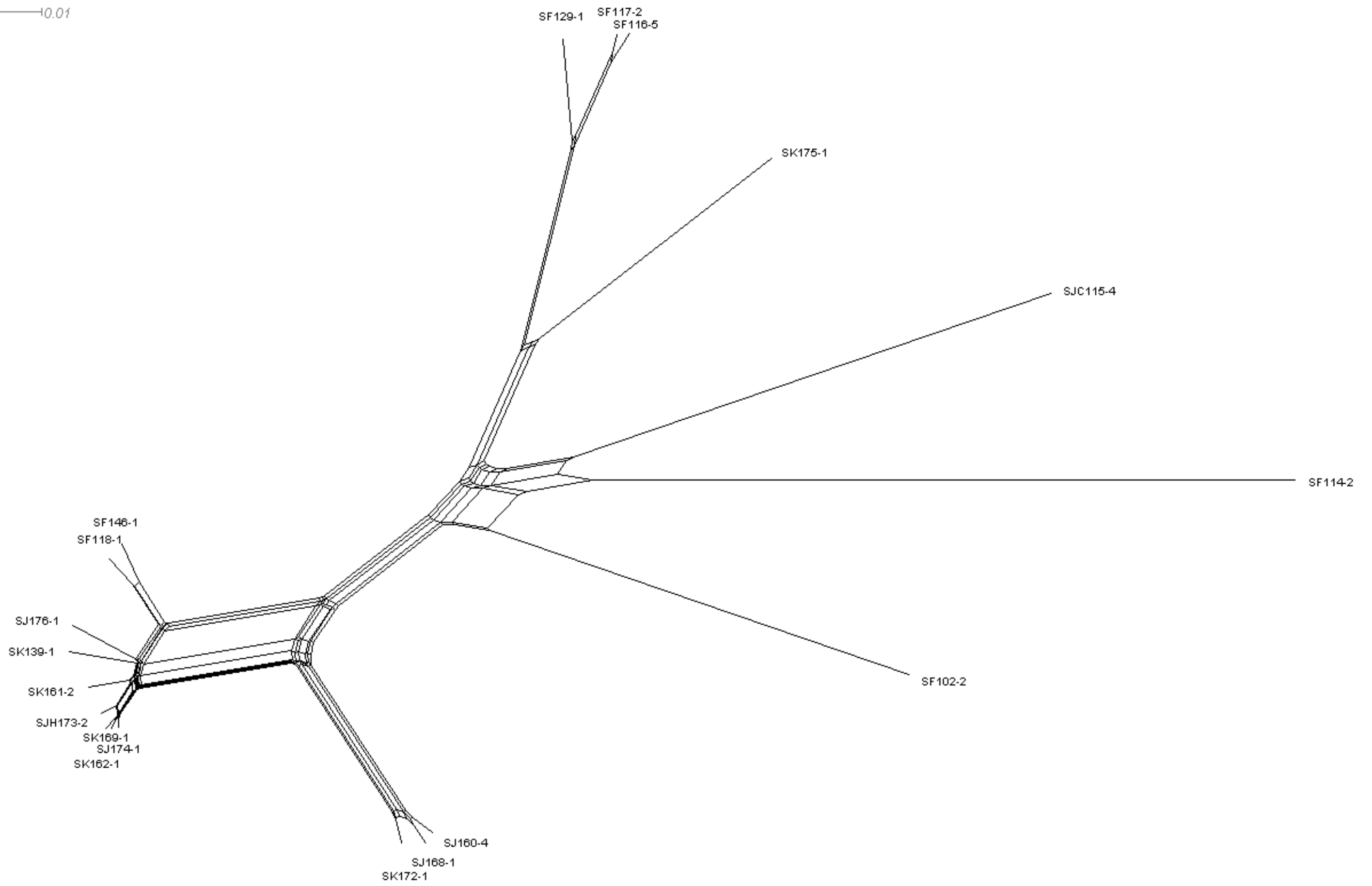
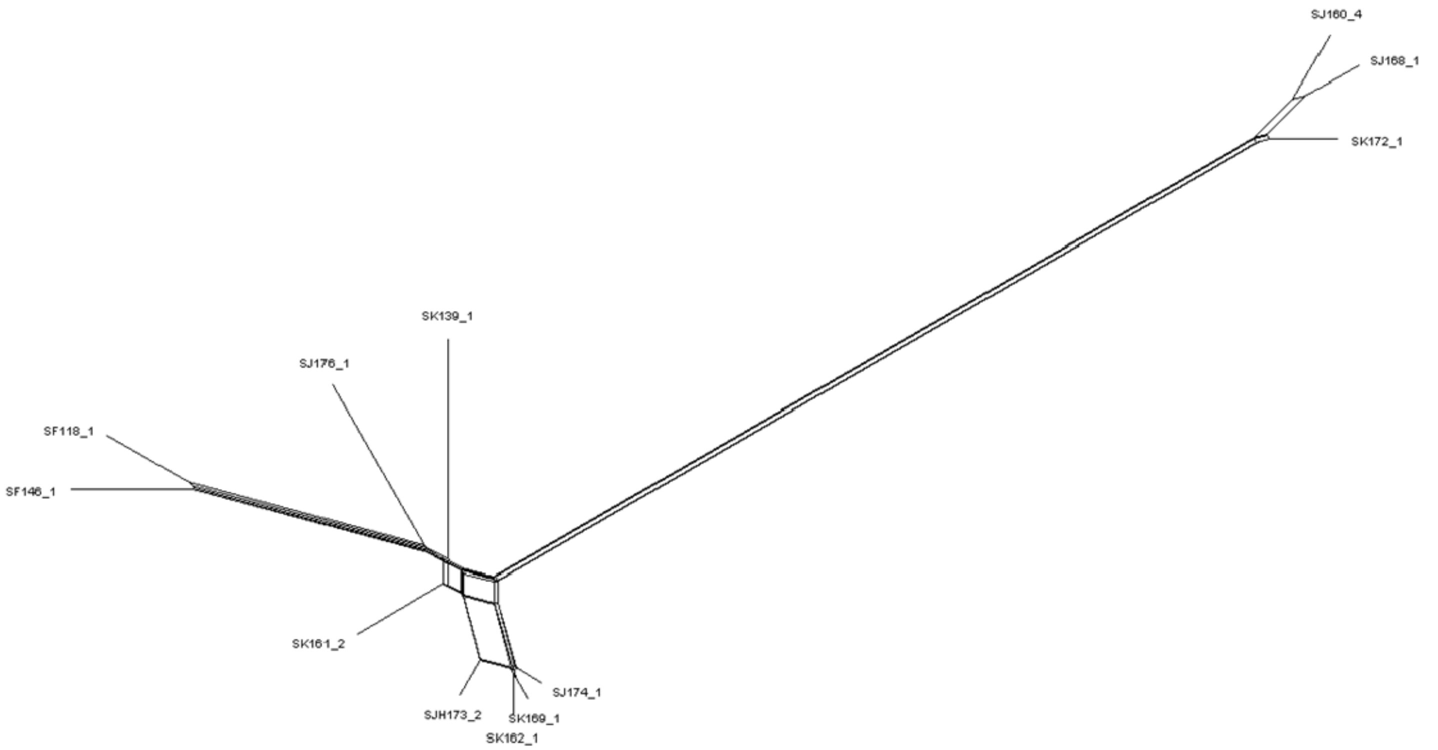


Fig. 3 (cont.)

C

0.01



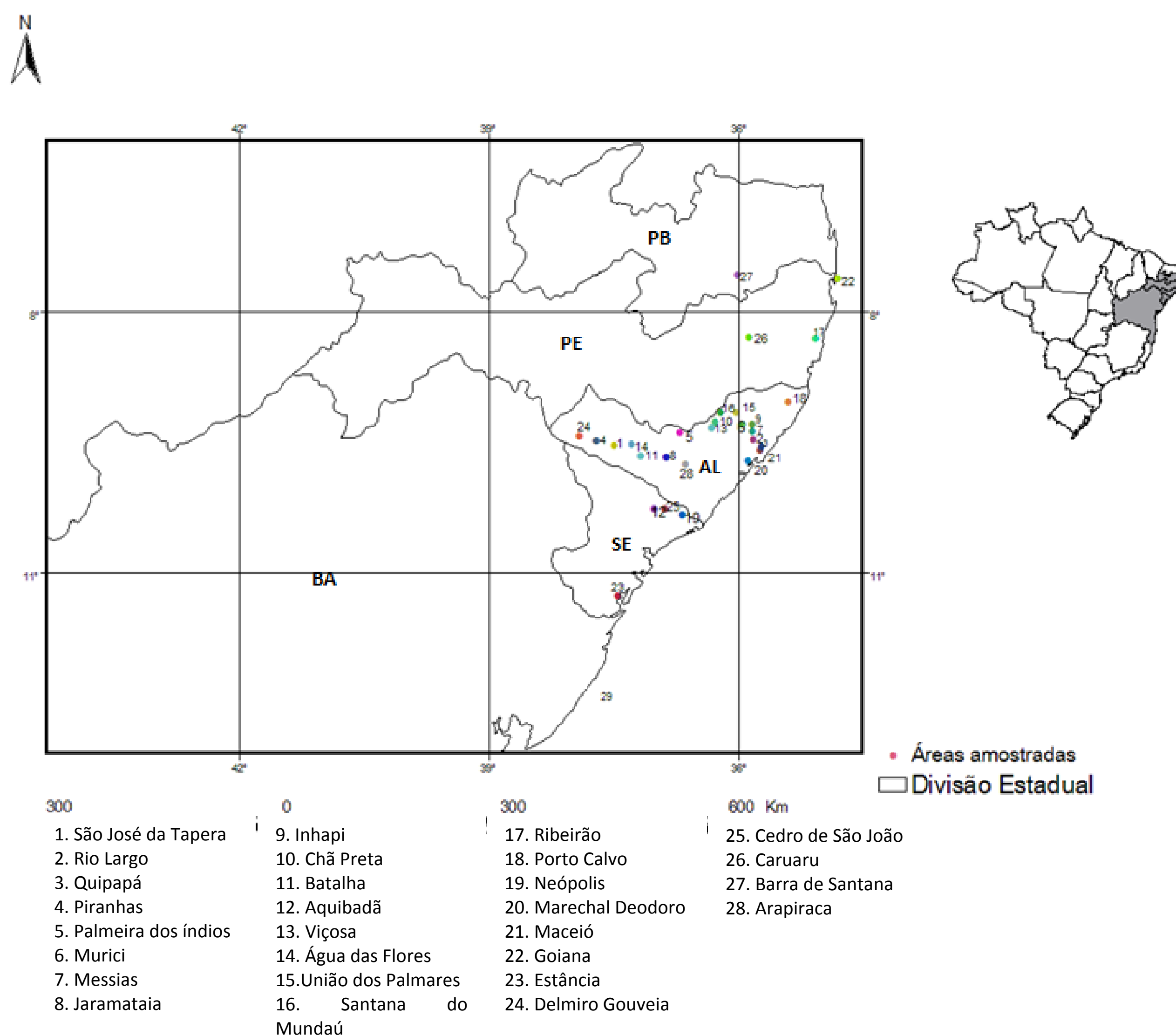
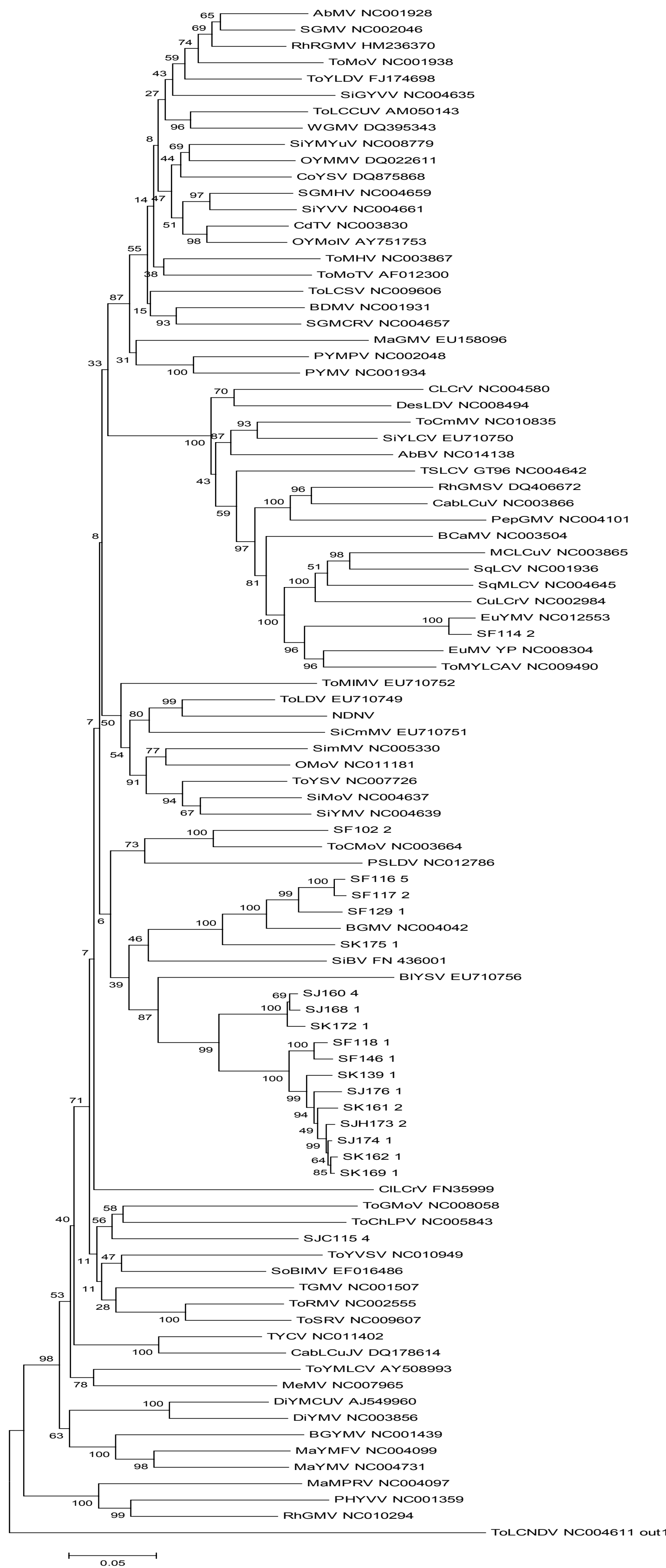


Figure 4. Geographical map of the Brazilian Northeastern states of Alagoas (AL), Paraíba (PB), Pernambuco (PE) and Sergipe (SE), indicating the locations where samples of leguminous weeds were collected. Numbers represent the different collection sites.



Supplementary Figure S1. Neighbor-joining tree based on the complete DNA-A nucleotide sequences of begomoviruses from the Americas, including the viruses infecting leguminous weeds in Northeastern Brazil.

CAPÍTULO III

Genetic structure of a begomovirus population infecting the ubiquitous weed *Cleome affinis* in Northeastern Brazil

1 **Genetic structure of a begomovirus population infecting the ubiquitous weed**
2 ***Cleome affinis* in Northeastern Brazil**
3
4 S. J. C. da Silva[1,3], G. P. Castillo-Urquiza[1], B. T. Hora Júnior[1], I. P. Assunção[2], G. S. A.
5 Lima[2,3], G. Pio-Ribeiro[3] , E. S. G. Mizubuti[1] and F. M. Zerbini[1*]
6
7 [1]Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa,
8 MG, 36570-000, Brazil.
9 [2]Departamento of Fitossanidade/CECA, Universidade Federal de Alagoas, Rio Largo,
10 AL, 57100-000, Brazil.
11 [3]Departamento de Fitopatologia, Universidade Federal Rural de Pernambuco, Recife,
12 PE, 52171-900, Brazil.
13
14
15 *To whom correspondence should be addressed. Email: zerbini@ufv.br
16
17
18 Silva, S.J.C., Castillo-Urquiza, G.P., Hora Júnior, B.T., Assunção, I.P., Lima, G.S.A.,
19 Pio-Ribeiro, G., Mizubuti, E.S.G. & Zerbini, F.M. Genetic structure of a begomovirus
20 population infecting the ubiquitous weed *Cleome affinis* in Northeastern Brazil, plant
21 pathology, *submitted*
22
23
24
25

26 **Abstract**

27 Begomoviruses are circular single-stranded DNA viruses with twinned incomplete
28 icosahedra particle morphology transmitted by whitefly. The incidence and severity of
29 diseases caused by begomoviruses has increase rapidly in many areas of the world,
30 including Brazil, where these are limiting factors to tomato and common bean
31 production. Begomoviruses are also associated with a wide range of weed plants which
32 in some cases act as inoculum sources for cultivated plants. *Cleome affinis* (family
33 Capparaceae) is a weed which is frequently associated with lima bean, commom bean
34 and other important leguminous crops. Samples of *C. affinis* showing mosaic, yellowing
35 and growth reduction were collected in the states of Alagoas, Bahia, Paraíba,
36 Pernambuco, and Sergipe, Northeastern Brazil. Sequences analysis of fourteen full-
37 length DNA-A viral genomes revealed that only one begomovirus species was found
38 infecting *C. affinis* with 91-96% identity with an isolate of the Cleome leaf crumple
39 virus (ClLCrV) from Mato Grosso do Sul. In a phylogenetic tree fourteen ClLCrV form
40 a basal group relative to all other Brazilian begomoviruses. RDP3 analysis showed
41 strong evidence of multiple recombination events among the ClLCrV isolates and other
42 begomoviruses from Brazil. High degree of genetic variability was found in the ClLCrV
43 population infecting *C. affinis* in Northeastern Brazil. Despite ClLCrV to be the only
44 species found in the collected samples, each clone represents a distinct isolate of the
45 same virus suggesting that *C. affinis* may act as a potential inoculum source or, more
46 likely, as a source of novel viruses for crop plants.

47

48 Key words: geminivirus, recombination, ClLCrV

49

50

Introduction

Geminiviruses (family *Geminiviridae*) have circular, single-stranded (ss) DNA genomes that are packaged within twinned quasi-isometric virions. Geminiviruses are divided into four genera, *Mastrevirus*, *Topocuvirus*, *Curtovirus* and *Begomovirus*, based on genome organization and biological properties, the most important being the type of insect vector (either whitefly, leafhopper or treehopper) and host range (either mono- or dicotyledonous hosts) (Fauquet et al., 2008). Begomoviruses (whitefly-transmitted geminiviruses) cause serious diseases in a number of economically important crops, mostly in tropical and subtropical regions (Rojas et al., 2005). Over the last four decades, agricultural intensification and the emergence and prevalence of a new and more aggressive biotype of the insect vector (*Bemisia tabaci* biotype B) have facilitated an increase in begomovirus populations and their expansion to new plant hosts throughout tropical and subtropical regions of the Americas (Morales & Anderson, 2001). This has contributed to the emergence of new and more virulent viruses, producing an increase in frequency and severity of disease (Hanssen et al., 2010, Hagen et al., 2008, Jones, 2009). In Brazil, begomoviruses are limiting factors to tomato and common bean production (Faria et al., 2000, Zerbini et al., 2005). In beans (*Phaseolus vulgaris* and *P. lunatus*), golden mosaic caused by *Bean golden mosaic virus* (BGMV) has been an important disease since the 1970's, and its dissemination has been attributed to the increase in soybean (*Glycine max*) cultivation (Costa, 1975, Costa, 1976). In tomatoes, the emergence of begomoviruses-associated diseases coincided with the introduction and spread of the B biotype of *Bemisia tabaci* (Melo, 1992, Ribeiro et al., 1998).

Weeds are considered as reservoirs of begomoviruses that infect crop plants, as well as sources of novel recombinant viruses due to mixed infections (Ilyas et al., 2010,

76 Graham et al., 2010, Castillo-Urquiza et al., 2008). Some of the economically important
77 begomoviruses in crop plants are closely related to begomoviruses found in weeds
78 (Andrade et al., 2006, Jovel et al., 2004, Ilyas et al., 2010). Similarly to what is
79 observed for begomoviruses in crops, the genetic diversity of begomoviruses infecting
80 weeds is very high, with a particularly high species diversity in *Sida* spp. (family
81 Malvaceae) (Frischmuth et al., 1997, Hofer et al., 1997, Castillo-Urquiza et al., 2008,
82 Guo & Zhou, 2006, Fiallo-Olive et al., 2010, Ambrozevicius et al., 2002, Assunção et
83 al., 2006). For example, the *Sida micrantha mosaic virus* complex consists of at least
84 three bipartite begomoviruses (Jovel et al., 2004).

85 Weed species, either indigenous or introduced, acting as reservoirs, can play an
86 important role in the emergence of plant virus epidemics (Seal et al., 2006). The
87 characterization of weed-infecting begomovirus is therefore, important for elucidating
88 their ecological and evolutionary behavior (Assunção et al., 2006). However, studies to
89 understand the genetic structure and dynamics of begomovirus populations in wild
90 reservoirs and the potential effects on cultivated species are scarce and less detailed
91 (Roye et al., 1999, Roye et al., 1997, Sanz et al., 2000, Garcia-Andres et al., 2006).

92 In this report we examine the begomovirus population present in *Cleome affinis,*
93 a weed that belongs to the family Capparaceae and which is frequently associated with
94 common bean (*Phaseolus vulgaris*) and lima bean (*P. lunatus*) crops in Northeastern
95 Brazil, as a step towards assessing their diversity and role as begomoviruses reservoirs.

96

97 **Material and Methods**

98 *Sample colletion*

99 Twenty-three samples of *Cleome affinis* were collected during the years of 2007
100 to 2010 in the states of Alagoas (AL), Bahia (BA), Paraíba (PA), Pernambuco (PE) and

101 Sergipe (SE), all in Northeastern Brazil (Table 1). Plants displaying symptoms of
102 mosaic, yellowing and growth reduction typical of begomovirus infection were
103 preferentially collected. Samples were desiccated by pressing and stored at -80ºC.
104 *DNA amplification and cloning*
105 Total DNA was extracted according to (Doyle & Doyle, 1987). To confirm the
106 presence of begomoviruses, PCR was carried out using universal primers for members
107 of the genus (Rojas et al., 1993). Full length viral genomes were amplified from PCR-
108 positive samples by rolling-circle amplification (RCA) (Inoue-Nagata et al., 2004),
109 cloned in pBLUESCRIPT KS + (Stratagene) after monomerization with the restriction
110 enzymes *Cla* I, *Hin*d III or *Pst* I, and sequenced at Macrogen Inc. (Seoul, South Korea)
111 by primer walking.
112
113 *Sequence comparisons and phylogenetic analysis*
114 DNA-A nucleotide sequences were submitted to a BLAST search for
115 preliminary species assignment based on the 89% threshold level established by the
116 *Geminiviridae* Study Group of the ICTV (Fauquet et al., 2008). Additional nucleotide
117 pairwise comparisons were performed with DNAMan version 4.0 (Lynnon Co.) using
118 the Optimal Alignment option with the following parameters: Ktuple = 2, Gap penalty =
119 7, Gap open = 10, Gap extension = 5.
120 Nucleotide sequences of begomoviruses used in the recombination and
121 phylogenetic analyses (see Supplementary Table S1 for the viruses and GenBank
122 accession numbers used in the analyses) were aligned using the Muscle module in Mega
123 5.0 (Tamura et al., 2007). Phylogenetic analysis was performed using Bayesian
124 inference and Markov chain Monte Carlo simulation implemented in MrBayes ver 3.0
125 (Ronquist & Huelsenbeck, 2003). Bayesian analysis was conducted on the aligned

126 dataset after MrModeltest v. 2.2 (Nylander, 2004) was used to determine the nucleotide
127 substitution model models. The Markov Chain Monte Carlo (MCMC) analysis of four
128 chains started with a heating parameter of 0.1 from a random tree topology and lasted
129 5,000,000 generations. Trees were saved each 100 generations, resulting in 50,000
130 saved trees. Burn-in was set at 1,250,000 generations after which the likelihood values
131 were stationary, leaving 37,000 trees from which the 50% majority rule consensus trees
132 and posterior probabilities were calculated.
133
134 *Recombination analysis*
135 Phylogenetic network analysis for evidence of recombination was performed
136 using the neighbour-net method implemented in SplitsTree4 (Huson & Bryant, 2006).
137 Additional analyses of potential recombination events and identification of putative
138 parental sequences were carried out using the Recombination Detection Program (RDP)
139 ver. 3.0 (Martin et al., 2010). Recombination events detected by at least four of the
140 analysis methods available in the program were considered thrustworthy. Alignments
141 were scanned using default settings for each analysis method using a Bonferroni-
142 corrected *p* value cutoff of 0.05.
143
144 *Genetic structure of viral populations*
145 The main descriptors of genetic variability were quantified using the program
146 DnaSP version 5 (Rozas, 2009): number of polymorphic sites, total number of
147 mutations (η), average number of nucleotide differences (k), nucleotide diversity (π),
148 number of haplotypes, haplotype diversity (Hd), number of segregrating sites,
149 Watterson's estimate of the population mutation rate based on the total number of
150 segregating sites (Theta-W) and on the total number of mutations (Theta-Eta). Four

151 types of neutrality tests were used to test the hypothesis of occurrence of selection in the
152 population: Tajima's D, Fu and Li's D* and F*, and the test based on the number of
153 synonymous (Ds) and non-synonymous (Dns) substitutions with the Pamilo-Bianchi-Li
154 (PBL) model.

155

156 **Results**

157 A total of 23 samples of *Cleome affinis* showing mosaic, yellowing and growth
158 reduction were collected: 11 from Alagoas, one from Bahia, two from Paraíba, two from
159 Pernambuco, six from Sergipe, and one from an unknown location (Table 1). All 23
160 samples tested positive for the presence of a begomovirus by PCR with universal
161 primers (data not shown). Fourteen full-length DNA-A viral genomes were cloned
162 (Table 1). No evidence of the presence of alphasatellites, or of any other kind of DNA
163 satellite, was obtained. Pairwise sequence comparisons showed that all fourteen clones
164 corresponded to isolates of *Cleome leaf crumple virus* (ClLCrV), displaying 91-96%
165 identity with a recently described ClLCrV isolate from Mato Grosso do Sul, Brazil
166 (FN435999) (Supplementary Table S2). In fact, clone SC215 is the only one showing
167 91% identity with ClLCrV, with the remaining 13 clones displaying >95% identity
168 (Supplementary Table S2). This suggests that SC215 represents a distinct strain of
169 ClLCrV.

170 Phylogenetic reconstruction based on the complete DNA-A nucleotide
171 sequences of the 14 ClLCrV isolates and 22 aditional Brazilian begomoviruses was
172 conducted using Bayesian inference, with the nucleotide substitution model GTR+I+G.
173 Strikingly, ClLCrV isolates form a cluster with two tomato-infecting begomoviruses
174 (Figure 1).

175 A phylogenetic tree based on the complete DNA-A sequences of the *C. affinis*
176 isolates, plus additional sequences of begomoviruses from Brazil and from the
177 Americas was constructed using the neighbor-joining method (Figure 2). The sequences
178 within the tree clustered into five major groups. Clusters I and V includes only non-
179 Brazilian begomoviruses. Cluster II comprises viruses from Central and South America,
180 plus four additional begomoviruses infecting tomato and weeds from Brazil (*Abutilon*
181 *Brazil virus*, AbMV; *Euphobia yellow mosaic virus*, EuYMV; *Sida yellow leaf curl*
182 *virus*, SiYLCV; and *Tomato common mosaic virus*, ToCmMV). Cluster III includes
183 mostly Brazilian begomoviruses that infecting bean okra, passion fruit, soybean,
184 tomatoes and other weeds. All ClLCrV isolates from *C. affinis* grouped with the original
185 ClLCrV isolate from Mato Grosso do Sul (FN435999) in cluster IV, which is placed at
186 a basal position relative to other Brazilian begomoviruses.
187
188 *Recombination analysis*
189 Analisys of nucleotide sequences revealed phylogenetic inconsistency between
190 the DNA components of ClLCrV. DNA-A formed a cluster with several tomato-
191 infecting begomoviruses from Brazil, whereas the DNA-B clustered with EuMV from
192 Brazil and Central America on a separated branch. The placing of the DNA-A and of
193 the DNA-B in separate branches of the respective trees suggests an ancient
194 pseudorecombination event during the evolution of ClLCrV, something which was also
195 proposed by (Paprotka et al., 2010). Therefore, to further investigate this hypothesis,
196 neighbor-net analysis was used to infer phylogenetics relationships among ClLCrV
197 isolates and all Brazilian begomoviruses. The analysis revealed clear evidence of
198 several recombination events (Figure 3A). Strong evidence for recombination was
199 found in cluster I, represented by the 14 ClLCrV isolates and the isolate from Mato

200 Grosso do Sul (FN435999). Recombination events were less evident in other clusters
201 (II, III, IV and IV). These results were corroborated when the analysis was restricted to
202 begomoviruses from *C. affinis*, and was also reinforced by the phylogenetic
203 inconsistency observed for SC215 and SC226, which grouped separately from the other
204 twelve isolates (Figure 3B).

205 The same groups of sequences were analyzed using the RDP3 package with the
206 aim of investigating these putative recombination signals. To avoid the detection of
207 unreliable signals, we selected only events supported by at least four different methods.
208 Analysis RDP3 including all Brazilian begomoviruses revealed that a weak
209 recombination event was detected for the 14 ClLCrV isolates, with breakpoints within
210 the Rep coding region (Tables 2). In this event *Tomato yellow spot virus* (ToYSV) was
211 identified as one of the putative parents (Table 2). An additional recombination event
212 was observed within the Rep gene for SC215 when Brazilian begomoviruses were
213 added, with SC216 identified as one of the parents (Table 2). Another strong
214 recombination event was detected involving SC215 and 226, with breakpoints at the
215 common region (CR), CP and Rep coding regions, with SC201 isolate identified as one
216 of the putative parents (Tables 2).

217

218 *Genetic structure of the* ClLCrV *population*

219 The analysis of genetic descriptors demonstrated that the ClLCrV population has
220 a high degree of genetic variability, which is considerably higher than those observed
221 for two populations of tomato infecting begomoviruses from Southeastern Brazil (Table
222 3).

223 Evidence of selection or demographic forces acting on the ClLCrV population
224 were assessed by four different neutrality tests. The four ORFs encoded by the DNA-A

225 (Rep, Trap, Ren and CP) varied in this regard. Significant probability for rejecting the
226 hypothesis of neutrality was found for the Rep ORF (Table 4), indicating that this
227 genomic region is potentially under purifying selection. Negative values were obtained,
228 but were not statistically supported, for Tajima's D, Fu and Li's *D** and Fu and Li's *F**
229 for Ren, Trap and CP (Table 4). However, the values of dN/dS <1 for all ORFs are
230 indicative of purifying selection acting on this population.

231

232 **Discussion**

233 *Cleome affinis* is classified in the family Capparaceae, and is frequently
234 associated with lima bean (*Phaseolus lunatus*), common bean (*Phaseolus vulgaris*) and
235 other important leguminous plants in Northeastern Brazil. Recently, a new begomovirus
236 species, *Cleome leaf crumple virus* (ClLCrV) was found infecting this weed in the state
237 of Mato Grosso do Sul (Paprotka et al., 2010). An unusual feature of this particular
238 isolate was its association with an alphasatellite molecule (Cleome leaf crumple virus-
239 associated DNA1), the first time that DNA satellites of any kind were detected in
240 association with begomoviruses the New World (Paprotka et al., 2010). A careful
241 examination of the RCA products obtained from our *C. affinis* samples (including
242 digestion with 4-base cutter restriction enzymes) failed to indicate the presence of
243 alphasatellites or of any other kind of DNA satellite.

244 The fact that every collected sample was infected by a begomovirus suggests
245 that *C. affinis* may act as a potential inoculum source or, more likely, as a source of
246 novel viruses for crop plants, considering that every clone obtained represented an
247 isolate of the same virus (ClLCrV). Sequence analysis of the fourteen isolates obtained
248 from *C. affinis* indicated 91-96% identity with the ClLCrV isolate from Mato Grosso
249 do Sul. The ICTV guidelines propose a demarcation threshold of 89% DNA-A sequence

250 identity for begomovirus species, and 94% for their strains (Fauquet et al., 2008). Clone
251 SC215 from Atalaia (AL) showed 91% identity with ClLCrV, suggesting that this
252 isolate represents a distinct strain.

253 Phylogenetic analysis using Bayesian inference method revealed that ClLCrV
254 isolates (including the one from Mato Grosso do Sul) form a group with *Tomato golden*
255 *mosaic virus* (TGMV) and Tomato yellow vein streak virus (ToYVSV) two tomato-
256 infecting begomoviruses. The our Neighbor-joining phylogenetic tree based on
257 Brazilian and American begomovirus sequences placed ClLCrV in a basal group
258 relative to all other Brazilian begomovirus, suggestive of an ancestral origin for this
259 virus. However, the discordance of these results may be due to differences in
260 methodology. A phylogenetic tree based on DNA-A sequences of several Brazilian and
261 a number of South American begomoviruses using Bayesian inference method placed
262 ClLCrV in a cluster with several tomato-infecting begomoviruses from Brazil (Paprotka
263 et al., 2010). Despite include a large number of virus and a considerably longer running
264 time (5,000,000 generations), our Bayesian inference analysis confirm the results
265 findings for (Paprotka et al. 2010), which are consistent with a ClLCrV Latin America
266 origin.

267 Phylogentic inconsistency among ClLCrV DNA-A and DNA-B components
268 lead to the hypothesis that an ancient pseudorecombination event is involved in the
269 origin of this virus (Paprotka et al., 2010). We found evidence of multiple
270 recombination events among the ClLCrV isolates and other begomoviruses from Brazil.
271 Recombination signals were particularly strong for clones SC215 and SC226, which
272 always clustered separately of the other ClLCrV isolates in phylogenetic trees.
273 Recombination breakpoints were identified primarily in the Rep coding region, a known
274 hot spot for recombination among geminiviruses (Lefeuvre et al., 2009, Lefeuvre et al.,

275 2007). It is interesting, though, that ClLCrV seems to be restricted to *C. affinis*, and also
276 seems to be the only begomovirus associated with this host. Parent identification in
277 recombination analysis is obviously limited by the data set used, and it is possible that
278 the true viruses involved in these recombination events are either ancestral viruses
279 which no longer exist, or unknown viruses infecting distinct, unidentified hosts.
280 Therefore, despite recombination frequently resulting in local adaptation, at least in this
281 specific virus-host system it seems to be acting on the viral population without an
282 obvious effect on its evolution.

283 In contrast to the low diversity of species found infecting the host, the analysis
284 of population genetic structure of ClLCrV revealed high genetic variability, which was
285 represented by the presence of unique haplotypes and high rates of nucleotide diversity,
286 haplotype diversity and mutation. These values were considerably higher than those
287 observed for two populations of tomato-infecting begomoviruses from Southeastern
288 Brazil (Castillo-Urquiza et al., 2010), and were similar to those observed for a BGMV
289 population obtained from lima bean (*Phaseolus lunatus*) samples collected in Alagoas
290 state (Ramos-Sobrinho et al., 2010). Therefore, it seems that viruses infecting
291 weed/wild hosts have a greater degree of genetic variability compared to viruses
292 infecting crop species.

293 Since mutation is the initial source of variation, much effort has been devoted to
294 determining spontaneous mutation rates in plant virus. High mutations rates, similar to
295 those observed for RNA viruses, have been estimated for the begomoviruses *Tomato*
296 *yellow leaf curl China virus* (TYLCCNV), *Tomato yellow leaf curl virus* (TYLCV),
297 *East African cassava mosaic virus* (EAMCV) and for the mastrevirus *Maize streak*
298 *virus* (MSV) (Ge et al., 2007, Duffy & Holmes, 2008, Duffy & Holmes, 2009, Harkins
299 et al., 2009). Reports about mutation rates in weed plants are scarce, although was

300 observed that Tobacco leaf curl geminivirus (TLCV) infecting *Eupatorium makinoi* also
301 revealed high mutation rates (Ooi et al., 1997).

302 Evolutionary forces acting on the ClLCrV population were evaluated using four
303 distinct neutrality tests. The negative values obtained for Tajima's D, Fu and Li's *D** and
304 Fu and Li's *F** tests were not statistically supported for the REn, Trap and CP ORFs.
305 However, Tajima's D, Fu and Li D* and F* test statistics were significant and negative
306 for the Rep ORF, indicating that this genomic region is potentially selection. In protein
307 coding sequences, selection pressures can be more accurately identified by the ratio of
308 nonsynonymous (amino-acid replacement) and synonymous (silent) substitution rates,
309 dN and dS, respectively. The dN/dS ratio (ω) is expected to exceed unity when natural
310 selection promotes changes in the protein sequence (diversifying selection), whereas a
311 ratio less than unity is expected if natural selection suppress protein changes (purifying
312 selection) (Yang & Bielawski, 2000) (Kimura, 1983). The calculation of the ratio ω for
313 each gene (Rep, Trap, REn and CP) was less than 1, which indicates purifying selection
314 acting, especially for the Rep gene (Table 4). These findings confirm the results from
315 the neutrality tests of strong purifying selection acting on the Rep gene. As Rep encodes
316 an essential replication protein, purifying selection can be act to preserve protein
317 function, although Rep appears to be under positive selection. Purifying selection and
318 population expansion were concluded to be the major evolutionary forces acting
319 on TLCV in *Eupatorium makinoi* (Yahara et al., 1998), on ToYVSV and ToCmMV in
320 tomato (Castillo-Urquiza et al., 2008), *Tomato spotted wilt virus* (TSWV) in peanut
321 (Kaye et al., 2011). These results suggest that the ClLCrV population may be under the
322 influence of purifying selection or under a recent expansion, so that the occurrence of
323 mutations is not sufficient to fully explain its genetic variability, and reinforce the

324 possible influence of additional evolutionary forces such as migration and
325 recombination upon the population.

326 Our results suggest that *C. affinis* may act as a potential inoculum source, or as
327 source of novel viruses for crop plants. This fact was confirmed by intensive detection
328 of inter and intra-specific recombination events in the ClLCrV. Together this results
329 indicating that mutation and recombination are important evolutionary process in the
330 genetic variability of the ClLCrV population.

331

332 **References**

333 Ambrozevicius LP, Calegario RF, Fontes EPB, Carvalho MG, Zerbini FM, 2002.
334 Genetic diversity of begomoviruses infecting tomato and associated weeds in
335 Southeastern Brazil. *Fitopatologia Brasileira* **27**, 372-7.

336 Andrade EC, Manhani GG, Alfenas PF, Calegario RF, Fontes EPB, Zerbini FM, 2006.
337 *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with a closer
338 relationship to viruses from *Sida* sp., forms pseudorecombinants with begomoviruses
339 from tomato but not from *Sida*. *Journal of General Virology* **87**, 3687-96.

340 Assunção IP, Listik AF, Barros MCS*, et al.*, 2006. Diversidade genética de
341 begomovírus que infectam plantas invasoras na Região Nordeste. *Planta Daninha* **24**,
342 239-44.

343 Castillo-Urquiza GP, Alfenas-Zerbini P, Beserra-Junior JEA*, et al.* Genetic structure of
344 tomato-infecting begomovirus populations in two tomato-growing regions of
345 Southeastern Brazil. *Proceedings of the Program and Abstracts, 6th International*
346 *Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop*,
347 *2010*. Guanajuato, Mexico.

348 Castillo-Urquiza GP, Beserra Jr. JEA, Bruckner FP*, et al.*, 2008. Six novel
349 begomoviruses infecting tomato and associated weeds in Southeastern Brazil. *Archives*
350 *of Virology* **153**, 1985-9.

351 Costa AS, 1975. Increase in the populational density of *Bemisia tabaci*, a threat to
352 widespread virus infection of legume crops in Brazil. In: Bird J, Maramorosch K, eds.
353 *Tropical Diseases of Legumes.* New York: Academic Press, 171.

354 Costa AS, 1976. Whitefly-transmitted plant diseases. *Annual Review of Phytopathology*
355 **14**, 429-40.

356 Doyle JJ, Doyle JL, 1987. A rapid DNA isolation procedure for small amounts of fresh
357 leaf tissue. *Phytochemical Bulletin* **19**, 11-5.

358 Duffy S, Holmes EC, 2008. Phylogenetic evidence for rapid rates of molecular
359 evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*.
360 *Journal of Virology* **82**, 957-65.

361 Duffy S, Holmes EC, 2009. Validation of high rates of nucleotide substitution in
362 geminiviruses: Phylogenetic evidence from East African cassava mosaic viruses.
363 *Journal of General Virology* **90**, 1539-47.

364 Faria JC, Bezerra IC, Zerbini FM, Ribeiro SG, Lima MF, 2000. Situação atual das
365 geminiviroses no Brasil. *Fitopatologia Brasileira* **25**, 125-37.

366 Fauquet CM, Briddon RW, Brown JK*, et al.*, 2008. Geminivirus strain demarcation and
367 nomenclature. *Archives of Virology* **153**, 783-821.

368 Fiallo-Olive E, Martinez-Zubiaur Y, Moriones E, Navas-Castillo J, 2010. Complete
369 nucleotide sequence of Sida golden mosaic Florida virus and phylogenetic relationships
370 with other begomoviruses infecting malvaceous weeds in the Caribbean. *Archives of*
371 *Virology* **155**, 1535-7.

372 Frischmuth T, Engel M, Lauster S, Jeske H, 1997. Nucleotide sequence evidence for the
373 occurrence of three distinct whitefly-transmitted, *Sida*-infecting bipartite geminiviruses
374 in Central America. *J Gen Virol* **78**, 2675-82.

375 Garcia-Andres S, Monci F, Navas-Castillo J, Moriones E, 2006. Begomovirus genetic
376 diversity in the native plant reservoir Solanum nigrum: Evidence for the presence of a
377 new virus species of recombinant nature. *Virology* **350**, 433-42.

378 Ge LM, Zhang JT, Zhou XP, Li HY, 2007. Genetic structure and population variability
379 of tomato yellow leaf curl China virus. *Journal of Virology* **81**, 5902-7.

380 Graham AP, Martin DP, Roye ME, 2010. Molecular characterization and phylogeny of
381 two begomoviruses infecting *Malvastrum americanum* in Jamaica: Evidence of the
382 contribution of inter-species recombination to the evolution of malvaceous weed-
383 associated begomoviruses from the Northern Caribbean. *Virus Genes* **40**, 256-66.

384 Guo XJ, Zhou XP, 2006. Molecular characterization of a new begomovirus infecting
385 Sida cordifolia and its associated satellite DNA molecules. *Virus Genes* **33**, 279-85.

386 Hagen C, Rojas MR, Sudarshana MR*, et al.*, 2008. Biology and molecular
387 characterization of Cucurbit leaf crumple virus, an emergent cucurbit-infecting
388 Begomovirus in the Imperial Valley of California. *Plant Disease* **92**, 781-93.

389 Hanssen IM, Lapidot M, Thomma B, 2010. Emerging viral diseases of tomato crops.
390 *Molecular Plant-Microbe Interactions* **23**, 539-48.

391 Harkins GW, Delport W, Duffy S*, et al.*, 2009. Experimental evidence indicating that
392 mastreviruses probably did not co-diverge with their hosts. *Virology Journal* **6**, -.

393 Hofer P, Engel M, Jeske H, Frischmuth T, 1997. Nucleotide sequence of a new bipartite
394 geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica. *Virology*
395 **78**, 1785-90.

396 Huson DH, Bryant D, 2006. Application of phylogenetic networks in evolutionary
397 studies. *Molecular Biology and Evolution* **23**, 254-67.

398 Ilyas M, Qazi J, Mansoor S, Briddon RW, 2010. Genetic diversity and phylogeography
399 of begomoviruses infecting legumes in Pakistan. *Journal of General Virology* **91**, 2091-
400 101.

401 Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T, 2004. A simple method for
402 cloning the complete begomovirus genome using the bacteriophage phi 29 DNA
403 polymerase. *Journal of Virological Methods* **116**, 209-11.

404 Jones RAC, 2009. Plant virus emergence and evolution: Origins, new encounter
405 scenarios, factors driving emergence, effects of changing world conditions, and
406 prospects for control. *Virus Research* **141**, 113-30.

407 Jovel J, Reski G, Rothenstein D, Ringel M, Frischmuth T, Jeske H, 2004. *Sida*
408 *micrantha* mosaic is associated with a complex infection of begomoviruses different
409 from *Abutilon mosaic virus*. *Arch Virol* **149**, 829-41.

410 Kaye AC, Moyer JW, Parks EJ, Carbone I, Cubeta MA, 2011. Population genetic
411 analysis of Tomato spotted wilt virus on peanut in North Carolina and Virginia.
412 *Phytopathology* **101**, 147-53.

413 Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University
414 Press.Cambrigde, 367 p.

415 Lefeuvre P, Lett JM, Varsani A, Martin DP, 2009. Widely conserved recombination
416 patterns among single-stranded DNA viruses. *Journal of Virology* **83**, 2697-707.

417 Lefeuvre P, Martin DP, Hoareau M*, et al.*, 2007. Begomovirus 'melting pot' in the
418 south-west Indian Ocean islands: Molecular diversity and evolution through
419 recombination. *J Gen Virol* **88**, 3458-68.

420 Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P, 2010. RDP3: A
421 flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**,
422 2462-3.

423 Melo PCT, 1992. Mosca branca ameaça produção de hortaliças. In. Campinas, SP,
424 Brazil: Asgrow do Brasil Sementes Ltda., Technical Bulletin.

425 Morales FJ, Anderson PK, 2001. The emergence and dissemination of whitefly-
426 transmitted geminiviruses in Latin America. *Archives of Virology* **146**, 415-41.

427 Nylander JAA, 2004. *MrModeltest v2.* Program distributed by the author.Evolutionary
428 Biology Centre, Uppsala University.

429 Ooi K, Ohshita S, Ishii I, Yahara T, 1997. Molecular phylogeny of geminivirus
430 infecting wild plants in Japan. *Journal of Plant Research*, 247-57.

431 Paprotka T, Metzler V, Jeske H, 2010. The first DNA 1-like alpha satellites in
432 association with New World begomoviruses in natural infections. *Virology* **404**, 148-57.

433 Ramos-Sobrinho R, Silva SJC, Silva TAL*, et al.* Genetic structure of a population of the
434 begomovirus *Bean golden mosaic virus* (BGMV) that infects lima bean (*Phaseolus*
435 *lunatus* L.) in the state of Alagoas, Brazil. *Proceedings of the Program and Abstracts,*
436 *6th International Geminivirus Symposium and 4th International ssDNA Comparative*
437 *Virology Workshop*, *2010*. Guadalajara, México.

438 Ribeiro SG, Ávila AC, Bezerra IC*, et al.*, 1998. Widespread occurrence of tomato
439 geminiviruses in Brazil, associated with the new biotype of the whitefly vector. *Plant*
440 *Disease* **82**, 830.

441 Rojas MR, Gilbertson RL, Russell DR, Maxwell DP, 1993. Use of degenerate primers
442 in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant*
443 *Disease* **77**, 340-7.

444 Rojas MR, Hagen C, Lucas WJ, Gilbertson RL, 2005. Exploiting chinks in the plant's
445 armor: Evolution and emergence of geminiviruses. *Annual Review of Phytopathology*
446 **43**, 361-94.

447 Ronquist F, Huelsenbeck JP, 2003. MrBayes 3: Bayesian phylogenetic inference under
448 mixed models. *Bioinformatics* **19**, 1572-4.

449 Roye ME, McLaughlin WA, Nakhla MK, Maxwell DP, 1997. Genetic diversity among
450 geminiviruses associated with the weed species *Sida* spp., *Macroptilium lathyroides*,
451 and *Wissadula amplissima* from Jamaica. *Plant Disease* **81**, 1251-8.

452 Roye ME, Spence J, McLaughlin WA, Maxwell DP, 1999. The common weed
453 *Macroptilium lathyroides* is not a source of crop-infecting geminiviruses from Jamaica.
454 *Tropical Agriculture* **76**, 256-62.

455 Rozas J, 2009. DNA sequence polymorphism analysis using DnaSP. *Methods in*
456 *Molecular Biology* **537**, 337-50.

457 Sanz AI, Fraile A, García-Arenal F*, et al.*, 2000. Multiple infection, recombination and
458 genome relationships among begomovirus isolates found in cotton and other plants in
459 Pakistan. *Journal of General Virology* **81**, 1839-49.

460 Seal SE, Van den Bosch F, Jeger MJ, 2006. Factors influencing begomovirus evolution
461 and their increasing global significance: Implications for sustainable control. *Critical*
462 *Reviews in Plant Sciences* **25**, 23-46.

463 Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: Molecular Evolutionary
464 Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**,
465 1596-9.

466 Yahara T, Ooi K, Oshita S, Ishii I, Ikegami M, 1998. Molecular evolution of a host-
467 range gene in geminiviruses infecting asexual populations of Eupatorium makinoi.
468 *Genes Genet Syst* **73**, 137-41.

469 Yang Z, Bielawski JP, 2000. Statistical methods for detecting molecular adaptation.
470 *Trends in Ecology and Evolution* **15**, 496-503.

471 Zerbini FM, Andrade EC, Barros DR*, et al.*, 2005. Traditional and novel strategies for
472 geminivirus management in Brazil. *Australasian Plant Pathology* **34**, 475-80.
473
474
475
476

Table 1. Location, year of collection and full-length begomovirus clones obtained from *Cleome affinis* samples collected in five Northeastern Brazilian states from 2007 to 2010.

Collection site	Year of collection	Sample code	Clones (DNA-A)	Species assignment[a]
Alagoas				
Paripueira	2009	SC201	SC201A	ClLCrV[b]
Maragogi	2009	SC202	SC202A	ClLCrV
São Miguel dos Campos	2009	SC203	SC203A	ClLCrV
Maceió	2010	SC205	SC205A	ClLCrV
Rio Largo	2010	SC206		
Atalaia	2007	SC215	SC215A	ClLCrV
Rio Largo	2007	SC216	SC216A	ClLCrV
Maceió	2010	SC217		
Arapiraca	2010	SC219		
Maceió	2010	SC220		
Joaquim Gomes	2010	SC224		
Bahia				
Costa do Sauípe	2010	SC207	SC207A	ClLCrV
Paraíba				
Alhandra	2010	SC225		
Alhandra	2010	SC226	SC226A	ClLCrV
Pernambuco				
Limoeiro	2010	SC214		
Goiana	2010	SC218	SC218A	ClLCrV
Sergipe				
Indiaroba	2010	SC208	SC208A	ClLCrV
Neópolis	2009	SC209	SC209A	ClLCrV
Japoatã	2009	SC210	SC210A	ClLCrV
Neópolis	2009	SC212	SC212A	ClLCrV
Neópolis	2009	SC213	SC213A	ClLCrV
Aquibadã	2009	SC221		
Unknown	2009	SC223		

[a]Species assignment based on the ICTV-established criteria of 89% nucleotide sequence identity for the full-length DNA-A (Fauquet et al., 2008).

[b]ClLCrV, *Cleome leaf crumple virus*.

Table 2. Putative recombination events detected among Brazilian begomoviruses, including the viruses infecting *Cleome affinis* in Northeastern Brazil.

Clone/isolate	Parents	Breakpoints		*P*-value						
		Initial	Final	R[c]	G	B	M	C	S	3S
SC201	ToYSV[a]	2486[b]	2601	$3.962X10^{-03}$	–[d]	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC202	ToYSV	2190	2661	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC203	ToYSV	2201	2659	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC205	ToYSV	2211	2674	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC207	ToYSV	2290	2652	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC208	ToYSV	2207	2664	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC209	ToYSV	2189	2660	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC210	ToYSV	2201	2663	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC212	ToYSV	2212	2660	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC213	ToYSV	2201	2663	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC215	SC201	18	1620	$2.101x10^{-22}$	$6.076x10\text{-}^{12}$	$2.623x10^{-13}$	$1.683x10^{-19}$	$2.080x10^{-12}$	$4.572x10^{-25}$	$6.068x10^{-24}$
	SC216	1658	2007	$4.552X10^{-08}$	$2.701X10^{-08}$	$2.009X10^{-11}$	$2.646X10^{-03}$	$3.136x10^{-05}$	$1.581x10^{-16}$	–
	ToYSV	2205	2663	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC216	ToYSV	2201	2661	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC218	ToYSV	2201	2661	$3.962X10^{-03}$	–	$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–
SC226	SC201	17	1584	$2.101x10^{-22}$	$6.076x10\text{-}^{12}$	$2.623x10^{-13}$	$1.683x10^{-19}$	$2.080x10^{-12}$	$4.572x10^{-25}$	$6.068x10^{-24}$
	ToYSV	2112	2660	$3.962X10^{-03}$		$3.290x10^{-02}$	$1.277x10^{-05}$	$2.106x10^{-04}$	–	–

[a] When only the major parent is indicated, the minor parent has not been identified. "Unknown", neither parent identified.

[b] Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.

[c] R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SisScan; 3S, 3SEQ. ‘’

[d] –, no recombination event detected.

Table 3. Genetic structure of the *Cleome leaf crumple virus* (ClLCrV) population obtained from *C. affinis* samples collected in five states of Northeastern Brazil.

Population	Sequence number	Genome Size	s[a]	Eta[b]	k[c]	π[d]	h[e]	Hd[f]	θw[g]	θ-Eta[h]
ClLCrV	14	2756	253	267	51.758	0.0191	14	1.0000	0.02944	0.03107

[a] Total number of segregating sites.
[b] Total number of mutations.
[c] Average number of nucleotide differences between sequences (Tajima's estimate of the population mutation rate, θ).
[d] Nucleotide diversity.
[e] Haplotype number.
[f] Haplotype diversity.
[g] Watterson's estimate of the population mutation rate based on the total number of segregating sites.
[h] Watterson's estimate of the population mutation rate based on the total number of mutations.

Table 4. Results of the different neutrality tests for each open reading frame (ORF) in the DNA-A of isolates of *Cleome leaf crumple virus* (ClLCrV) obtained from *Cleome affinis* samples collected in five states of Northeastern Brazil.

ORF[a]	Tajima's *D*	Fu and Li's *D**	Fu and Li's *F**	dN/dS
Rep	-1.8653*	-2. 5503**	-2.71489*	0.0228
Trap	-0.4477 (ns)[b]	-1.4769 (ns)	-1.3752 (ns)	0.0887
REn	-1.3826 (ns)	-1.8803 (ns)	-1.9997 (ns)	0.3171
CP	-0.0972 (ns)	-1.0095 (ns)	-0.8731 (ns)	0.2124

[a] Rep, Replication-associated protein; Trap, Tans-activating protein; Ren, Replication enhancer protein; CP, Coat protein.
[b] ns, not significant at $p > 0,10$
* significant at $p < 0,05$
** siginificant at $p < 0,02$

Supplementary Table S1. Begomoviruses used in pairwise sequence comparisons, phylogenetic and recombination analyses.

Virus	Acronym	GenBank access # (DNA-A)
From Brazil		
Abutilon Brazil virus	AbBV	NC_014138
Bean golden mosaic virus	BGMV	M88686
Blainvillea yellow spot virus	BlYSV	EU710756
Cleome leaf crumple virus	ClLCrV	FN435999
Euphorbia yellow mosaic virus	EUYMV	NC_012553
Nicandra deforming necrosis virus	NDNV	n.a.
Okra mottle virus	OmoV	NC_011181
Passionfruit severe leaf distortion virus	PSLDV	NC_012786
Sida common mosaic virus	SiCmMV	EU710751
Sida mosaic Brazil virus	SiMBV	FN436001
Sida micrantha mosaic virus	SiMMV	NC_005330
Sida mottle virus	SiMoV	NC_004637
Sida yellow leaf curl virus	SiYLCV	EU710750
Sida yellow mosaic virus	SiYMV	NC_004639
Soybean blistering mosaic virus	SoBlMV	EF016486
Tomato chlorotic mottle virus	ToCMoV	NC_003664
Tomato common mosaic virus	ToCmMV	NC_010835
Tomato golden mosaic virus	TGMV	NC_001507
Tomato mild mosaic virus	ToMlMV	EU710752
Tomato rugose mosaic virus	ToRMV	NC_002555
Tomato severe rugose virus	ToSRV	NC_009607
Tomato yellow spot virus	ToYSV	NC_007726
Tomato yellow vein streak virus	ToYVSV	NC_010949
From other countries in the Americas		
Abutilon mosaic virus	AbMV	NC_001928
Bean calico mosaic virus	BCaMV	NC_003504
Bean dwarf mosaic virus	BDMV	NC_001931
Bean golden yellow mosaic virus	BGYMV	NC_001439
Cabbage leaf curl virus	CaLCuV	NC_033866
Chino del tomate virus	CdTV	NC_003830
Cotton leaf curl virus	CLCrV	NC_004580
Corchorus yellow spot virus	CoYSV	NC_008492
Curcubit leaf crumple virus	CuLCrV	NC_002984
Desmodium leaf distortion virus	DesLDV	NC_008494
Dicliptera yellow mosaic virus	DiYMV	NC_003856
Dicliptera yellow mosaic Cuba virus	DiYMCUV	AJ549960
Euphobia mosaic virus - Yucatan Peninsula	EUMV_YP	NC_008304
Macroptillium golden mosaic virus	MaGMV	NC_010952
Macroptillium mosaic Puerto Rico virus	MaMPR	NC_004097

Macroptilium yellow mosaic Florida virus	MaYMFV	NC_004009

Supplementary Table S1 (cont.)

Macroptilium yellow mosaic virus	MaYMV	NC_010647
Melon chlorotic leaf curl virus	MCLCuV	NC_003865
Merremia mosaic virus	MeMV	NC_007965
Okra yellow mosaic Mexico virus	OYMMV	NC_014066
Okra yellow mottle Iguala virus	OYMoIV	AY751753
Pepper golden mosaic virus	PepGMV	NC_004101
Pepper huasteco yellow vein virus	PHYVV	NC_001359
Potato yellow mosaic Panama virus	PYMPV	NC_002048
Potato yellow mosaic virus	PYMV	NC_001934
Rhyncosia golden mosaic Sinaloa virus	RhGMSV	DQ406672
Rhyncosia golden mosaic virus	RhGMV	NC_010294
Rhyncosia rugose golden mosaic virus	RhRGMV	HM236360
Sida golden mosaic Costa Rica virus	SGMCRV	NC_004657
Sida golden mosaic Honduras virus	SGMHV	NC_004659
Sida golden mosaic virus	SGMV	NC_002046
Sida golden yellow vein virus	SiGYVV	NC_004635
Sida yellow mosaic Yucatan virus	SiYMYuV	NC_008779
Sida yellow vein virus	SiYVV	NC_004661
Squash leaf curl virus	SqLCV	NC_001936
Squash mild leaf curl virus	SqMLCV	NC_004645
Tomato Chino La Paz virus	ToChLPV	NC_005843
Tomato golden motlle virus	ToGMoV	NC_008058
Tobacco leaf curl Cuba virus	TLCCUV	AM050143
Tomato mosaic Havana virus	ToMHV	NC_003867
Tomato mottle Taino virus	ToMoTV	NC_001828
Tomato mottle virus	ToMoV	NC_001938
Tomato mild yellow leaf curl Aragua virus	ToMYLCAV	NC_009490
Tomato yellow leaf distortion virus	ToYLDV	FJ174698
Tomato yellow margin leaf curl virus	ToYMLCV	AY508998
Tomato severe leaf curl virus	ToSLCV	NC_004642
Tobacco yellow crinkle virus	TYCV	NC_011402
Wissadula golden mosaic virus	WGMV	NC_010948
Outgroups		
Tomato leaf curl New Delhi virus	TLCNDV	NC_004611

Supplementary Table S2. Percent identities between the complete DNA-A nucleotide sequences of the 14 *Cleome leaf crumple virus* (ClLCrV) isolates detected in samles of *Cleome affinis* in Northeastern Brazil.

	ClLCrV	**SC201**	**SC202**	**SC203**	**SC205**	**SC207**	**SC208**	**SC209**	**SC210**	**SC212**	**SC213**	**SC215**	**SC216**	**SC218**	**SC226**
ClLCrV[a]	–														
SC201	96	–													
SC202	95	99	–												
SC203	95	99	99	–											
SC205	95	99	99	99	–										
SC207	95	99	99	99	99	–									
SC208	96	99	99	98	99	99	–								
SC209	95	99	99	99	99	99	99	–							
SC210	95	99	99	99	99	99	99	99	–						
SC212	96	99	99	99	99	99	99	99	99	–					
SC213	96	99	99	98	99	99	99	99	99	99	–				
SC215	91	94	94	94	94	94	94	94	94	94	94	–			
SC216	96	99	99	99	99	99	99	99	99	99	99	99	–		
SC218	95	99	99	99	99	99	99	99	99	99	99	99	99	–	
SC226	94	98	98	97	98	98	98	98	98	98	98	98	98	98	–

[a]ClLCrV isolate obtained from a *C. affinis* sample from Mato Grosso do Sul (FN435999) (Paprotka et al., 2010).

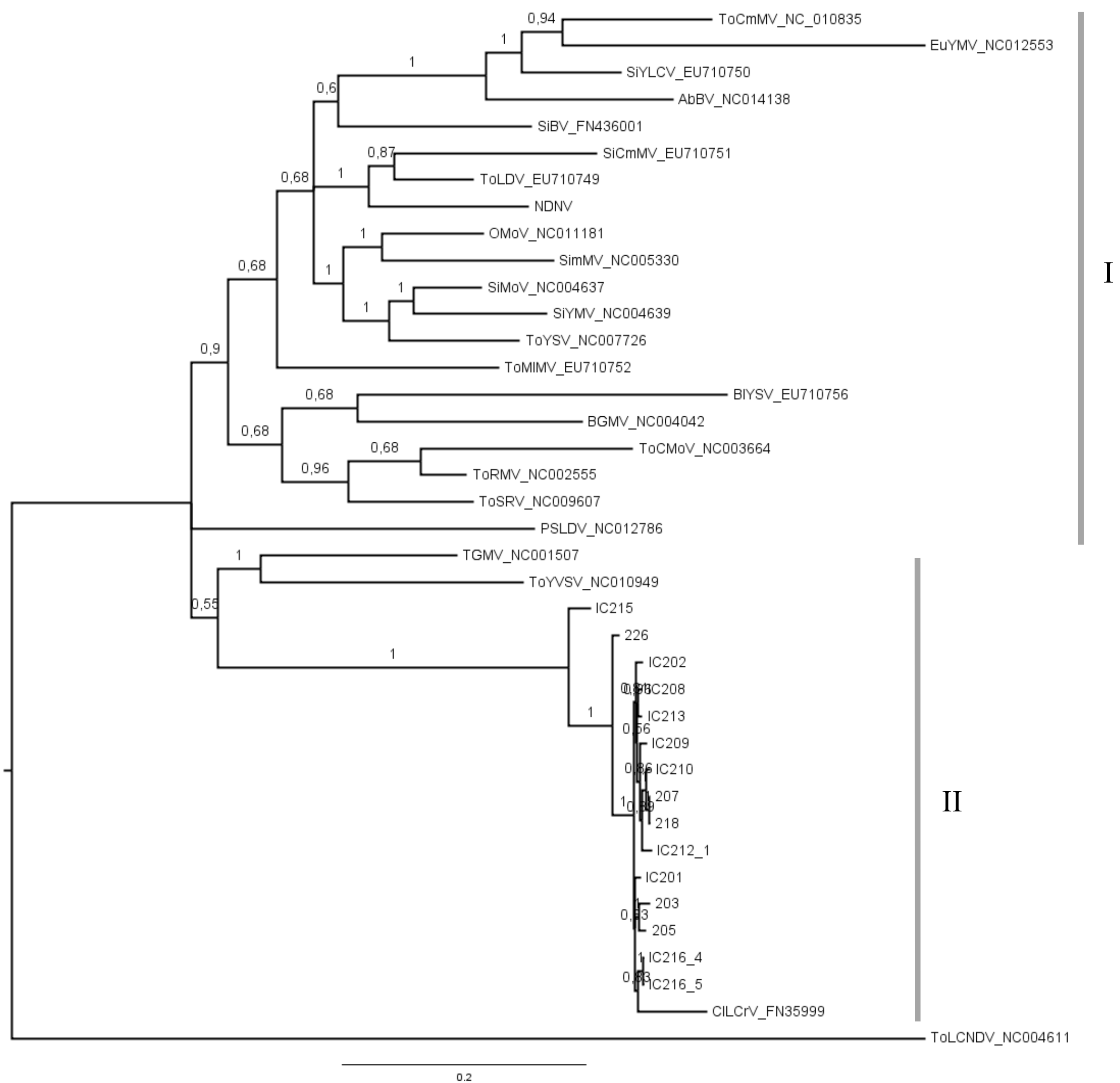


Figure 1. Bayesian 50% majority rule consensus tree of begomoviruses from *Cleome affinis* and other Brazilian begomoviruses (see Supplementary Table S1 for full virus names). Numbers at the nodes indicate Bayesian posterior probabilities. Cluster II includes all 14 ClLCrV isolates obtained in this study, plus the isolate from Mato Grosso do Sul (FN435999).

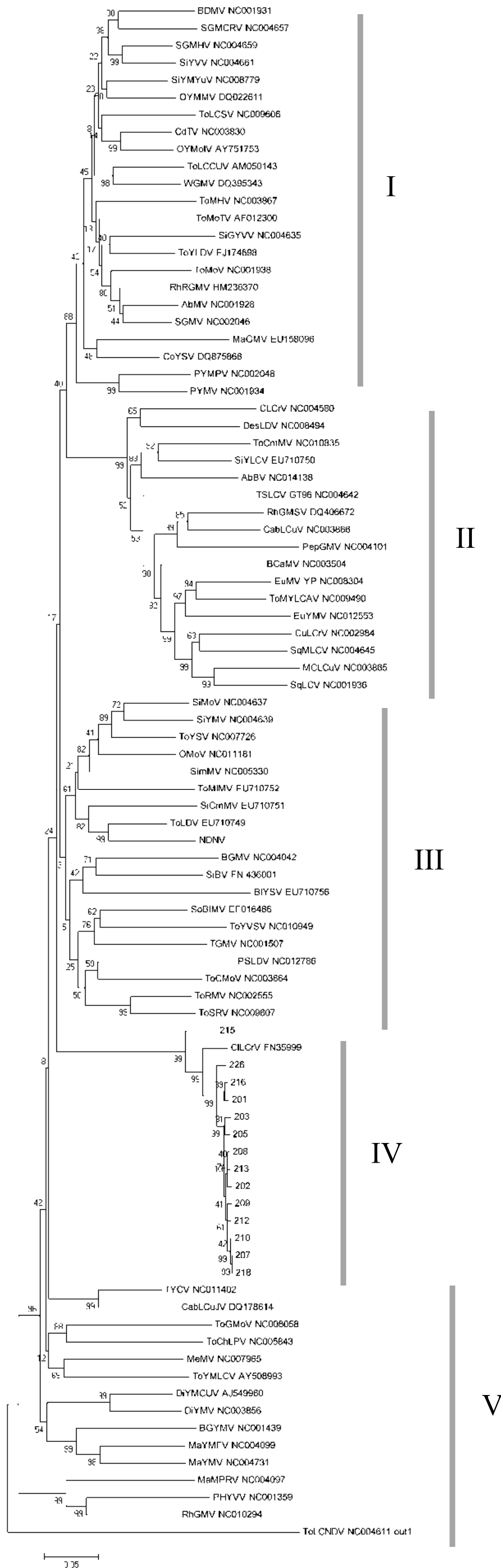


Figure 2. Neighbor-joining tree based on the complete DNA-A nucleotide sequences of begomoviruses from the Americas, including the viruses infecting *Cleome affinis* in Northeastern Brazil.Cluster IV includes all 14 ClLCrV isolates obtained in this study, plus the isolate from Mato Grosso do Sul (FN435999).

A

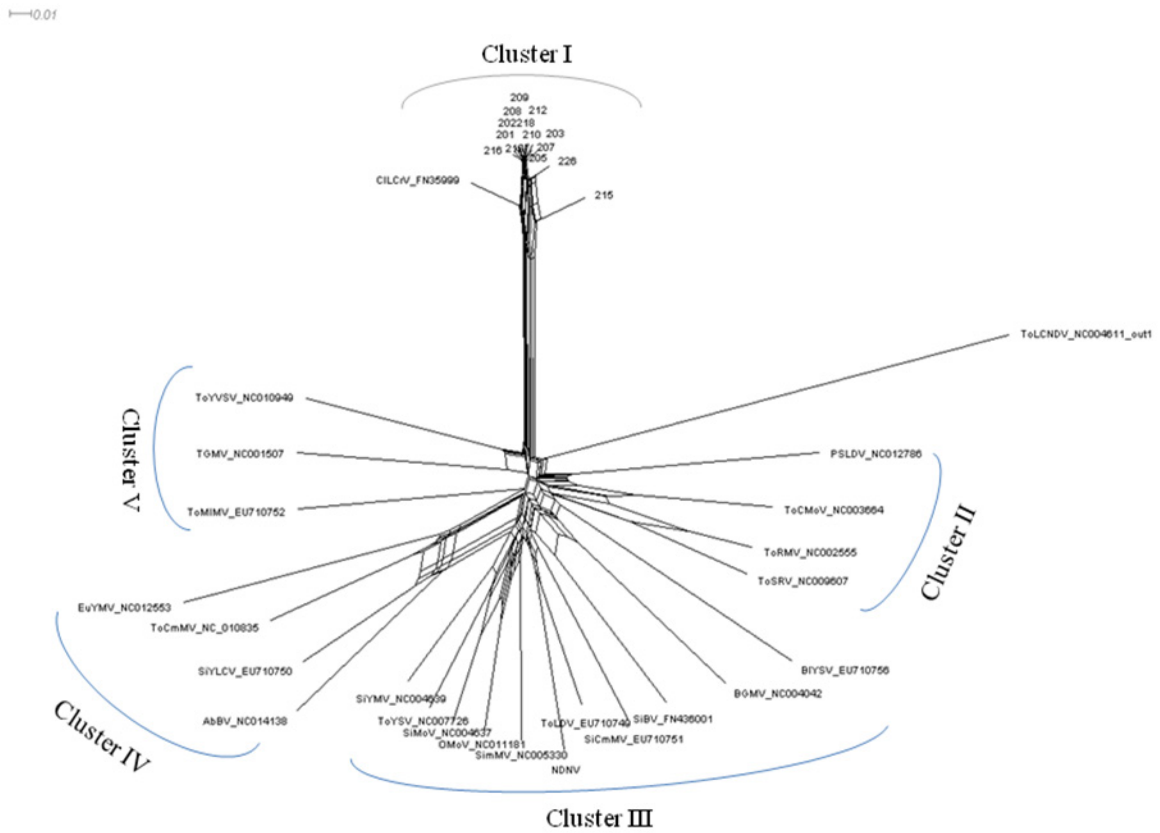
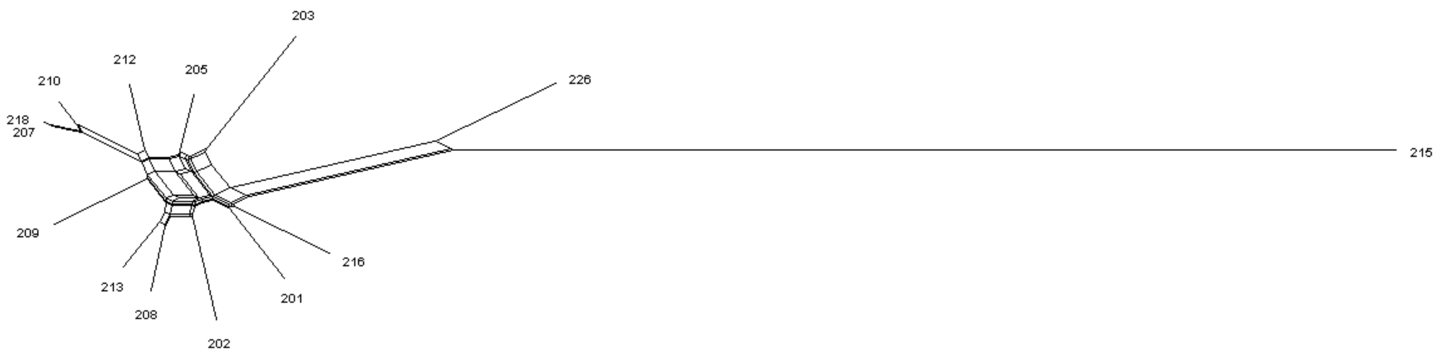


Figure 3. Phylogenetic evidence for recombination among (**A**) all Brazilian begomoviruses, including the ones describes in this work, and (**B**) a population of *Cleome leaf crumple virus* (ClLCrV) obtained from samples of *Cleome affinis* collected in five different states of Northeastern Brazil. Neighbor-net analysis was performed using SplitsTree4. Formation of a reticular network rather than a single bifurcated tree is indicative of recombination.

Figure 3 (cont.)

B



CONCLUSÕES GERAIS

CONCLUSÕES GERAIS

- Alto grau de diversidade de espécies de begomovírus que infectam plantas daninhas da família Fabaceae foi observado no nordeste do Brasil, onde quatro novas espécies foram encontradas.

- *Macroptilium lathyroides* foi revelado como hospedeira comum para diferentes begomovírus, e esta pode atuar como reservatórios a partir da qual vírus recombinantes podem surgir.

- Alta variabilidade genética foi encontrada para as populações de MaYSV e ClLCrV, infectando plantas das famílias Fabaceae e Capparaceae, respectivamente, podendo estas constituírem importantes fontes de novos vírus para planta cultiváveis.