CAMILA GEOVANA FERRO

SPECIES DIVERSITY AND GENETIC VARIABILITY OF BEGOMOVIRUSES AND ASSOCIATED DNA SATELLITES INFECTING NON-CULTIVATED PLANTS IN BRAZIL AND IN SPAIN

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae.*

VIÇOSA
MINAS GERAIS – BRASIL
2016

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T
F395s
2016

Ferro, Camila Geovana, 1986-
Species diversity and genetic variability of begomoviruses and associated dna satellites infecting non-cultivated plants in Brazil and in Spain / Camila Geovana Ferro. – Viçosa, MG, 2016.
x, 148f. : il. (algumas color.) ; 29 cm.

Orientador: Francisco Murilo Zerbini Júnior.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. *Begomovirus*. 2. Variabilidade genética. 3. *Sida spp.* 4. *Leonurus sibiricus*. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 571.9928

CAMILA GEOVANA FERRO

SPECIES DIVERSITY AND GENETIC VARIABILITY OF BEGOMOVIRUSES AND ASSOCIATED DNA SATELLITES INFECTING NON-CULTIVATED PLANTS IN BRAZIL AND IN SPAIN

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 28 de março de 2016.

Claudine Márcia Carvalho

Derly José Henrique da Silva

Jesús Navas Castillo

Márcio Tadeu Godinho

Francisco Murilo Zerbini Júnior
(Orientador)

“Há um tempo em que é preciso

abandonar as roupas usadas,

que já tem a forma do nosso corpo,

e esquecer os nossos caminhos,

que nos levam aos mesmos lugares.

É o tempo da travessia: e,

se não ousarmos fazê-la,

teremos ficado, para sempre,

a margem de nós mesmos.”

Existem pessoas que são essenciais para que essa travessia aconteça,

as quais tornam nossa caminhada mais significativa pela paciência, apoio, carinho e dedicação!

Em nome da minha mãe Ana Maria e meu pai Joel, dedico.

AGRADECIMENTOS

À Deus por estar presente em todos os momentos firmes e trêmulos, por me dar a segurança necessária para enfrentar os problemas e seguir adiante e principalmente por iluminar meus caminhos e permitir mais essa conquista em minha vida.

Aos meus pais, Joel e Ana Maria, pelo amor, incentivo, apoio incondicional em todos os momentos e decisões tomadas, dedicação, confiança e principalmente paciência. Vocês são responsáveis pelas minhas conquistas, sou o reflexo de vocês!

Aos meus irmãos, Lucas e Thiago, por acreditarem em mim e contribuírem para mais essa realização.

A toda a minha família pela confiança e apoio. A minha avó Maria, um obrigado especial, suas orações foram atendidas!

Ao Ricardo por estar sempre presente na minha vida, me proporcionando amor, tranquilidade, compreensão e paciência em todos os momentos, desde os difíceis até os mais memoráveis.

À Universidade Federal de Viçosa (UFV) por proporcionar a realização do doutorado em Genética e Melhoramento.

Ao Professor Francisco Murilo Zerbini, pelo suporte profissional, orientação, dedicação e amizade durante a realização deste trabalho.

Ao Professor Jesús Navas Castillo e a Elvira Fiallo Olivé por todos os ensinamentos, orientação, amizade e dedicação durante a minha trajetória tanto na Espanha, quanto no Brasil.

Aos amigos do Laboratório de Virologia Vegetal Molecular, Alison, Daniel, Tathiana, Igor, Hermano, Josiane, Angélica, Marcos, Osvaldo, Lenin, Roberto, Glória, Talita, Larissa, Diogo, Silvia, Adriana, Fábio, Murilo, Renan, Fernanda B.,

André, David, Flávia, Ana, Fernanda, Diego e Laiane pela amizade e pelo agradável ambiente de trabalho durante esses quatro anos. Em especial ao César, Márcio, Vitor e João, minha eterna gratidão! Vocês ajudaram a tornar esse sonho realidade e fizeram meus dias mais alegres.

Ao ex-funcionário Joaquim e a funcionária Patrícia, pelo apoio para a realização deste trabalho e à amizade.

As professoras Poliane e Claudine pelos ensinamentos e amizade.

Aos meus amigos de Viçosa, em especial a Dayana, Ana Lídia, Gabriela, Nívia, Haydee, Marcus e Yanna, que sempre estão ao meu lado, nos momentos de alegrias e de tristezas.

Aos amigos da vida, que fizeram parte do meu passado, do presente e seguramente estarão no meu futuro. Um agradecimento especial à Mariana, Milena, Priscila, Patrícia, Daiana, Marina, Gustavo, Rafael, Lizo, Murilo, Caroline, Fernanda W., Fernanda R., Renata, Renata L., Monique, Thaís, Eduarda, Paolo, Pedro G., Daniel, Andréia, Paulo, Daniela, Francisco, Rodrigo, Taís, Márcio e Lucas Takami.

Aos amigos, funcionários e pesquisadores do IHSM-CSIC (La Mayora). Em especial à Yazmin, Pedro, Eduardo, Amanda, Elisa, Rida e Francisco pela amizade.

A todas as amizades feitas em Málaga (Espanha), muito obrigada por me proporcionarem dias excepcionais!

Enfim, a todos que contribuíram direta ou indiretamente para conclusão deste trabalho.

BIOGRAFIA

Camila Geovana Ferro, filha de Joel Ferro e Ana Maria Bortolin Ferro, nasceu em 14 de agosto de 1986, em Rio Claro, São Paulo.

Residiu e estudou todos os anos do ensino fundamental e médio na cidade de Rio Claro, São Paulo, sendo concluído em 2003. Graduou-se em Agronomia pela Universidade Federal de Viçosa (UFV), na cidade de Viçosa, Minas Gerais, em janeiro de 2010. Ingressou, em março de 2010, no curso de pós-graduação em Genética e Melhoramento em nível de Mestrado na Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de Jaboticabal, obtendo o título de *Magister Scientiae* em fevereiro de 2012. Ingressou no mês seguinte no curso de pós-graduação em Genética e Melhoramento em nível de Doutorado na Universidade Federal de Viçosa e, em março de 2016, submeteu-se à defesa de tese.

SUMÁRIO

RESUMO

FERRO, Camila Geovana, D.Sc., Universidade Federal de Viçosa, março de 2016. **Diversidade de espécies e variabilidade genética de begomovírus e DNAs satélites infectando plantas não-cultivadas no Brasil e na Espanha.** Orientador: Francisco Murilo Zerbini Júnior. Coorientador: Alison Talis Martins Lima.

Begomovírus (gênero *Begomovirus*, família *Geminiviridae*) possuem um ou dois componentes genômicos de DNA circular de fita simples, encapsidados em partículas icosaédricas geminadas. A maioria dos begomovírus presentes no Novo Mundo (NM) são bissegmentados, enquanto no Velho Mundo (VM) prevalecem os monossegmentados, frequentemente associados a duas classes de DNAs satélites: alfassatélites e betassatélites. Recentemente, alfassatélites foram relatados em associação com begomovírus bissegmentados no NM. Membros do gênero *Begomovirus* são responsáveis por doenças em culturas economicamente importantes em todo o mundo. Begomovírus que infectam *Ipomoea* spp. (família Convolvulaceae), comumente denominados "sweepovírus", possuem organização genômica típica dos vírus monossegmentados do VM, porém são filogeneticamente distintos das demais espécies do gênero. O entendimento da dinâmica e variabilidade genética de populações virais em plantas não-cultivadas é importante para auxiliar na previsão e prevenção de doenças em plantas cultivadas. Este trabalho teve como objetivos: (*i*) estudar a diversidade de espécies de begomovírus em dois hospedeiros não-cultivados amplamente distribuídos no Brasil, *Sida* spp. e *Leonurus sibiricus*; (*ii*) determinar a estrutura e a variabilidade genética das populações de begomovírus em *Sida* spp. e *L. sibiricus*; (*iii*) determinar a variabilidade genética de deltassatélites associados a sweepovírus infectando *Ipomoea indica* no sul da Espanha, expandindo a análise para outras regiões geográficas. Para os dois primeiros objetivos, o DNA total foi extraído de plantas de *Sida* spp. e *L. sibiricus* coletadas nos estados do Rio Grande do Sul, Paraná e Mato Grosso do Sul de 2009 a 2011, e os genomas virais foram amplificados, clonados e sequenciados. A maioria das amostras de *Sida* spp. estavam infectadas pelo *Sida micrantha mosaic virus* (SiMMV). Foram encontradas também três novas espécies. A maioria absoluta das amostras de *L. sibiricus* estavam infectadas pelo *Tomato yellow spot virus* (ToYSV). Dois alfassatélites foram encontrados: *Euphorbia* yellow mosaic alphasatellite em *Sida* sp. e um novo alfassatélite em *L. sibiricus*. As populações de SiMMV e ToYSV possuem alta

variabilidade genética. Embora um elevado nível de recombinação tenha sido detectado, a dinâmica mutacional é o fator primário de diversificação. Os resultados são inconclusivos em relação a estruturação baseada em localização geográfica. Para o terceiro objetivo, DNA foi extraído de plantas de *I. indica* coletadas no sul da Espanha em 2015, e amplificação por círculo rolante (RCA), uma técnica que leva a amplificação de moléculas de ssDNA circulares sem o conhecimento prévio da sequência nucleotídica, foi usada para clonar os deltassatélites e seus sweepovírus associados, *Sweet potato leaf curl virus* (SPLCV) e *Sweet potato mosaic virus* (SPMV). Deltassatélites apresentaram baixa variabilidade de sequência, ausência de recombinação e de estruturação geográfica. Além disso, uma quimera sweepovírus-satélite com um tamanho similar ao dos deltassatélites foi detectada.

ABSTRACT

FERRO, Camila Geovana, D.Sc., Universidade Federal de Viçosa, March, 2016. **Species diversity and genetic variability of begomoviruses and associated DNA satellites infecting non-cultivated plants in Brazil and in Spain.** Advisor: Francisco Murilo Zerbini Júnior. Co-advisor: Alison Talis Martins Lima.

Begomoviruses (genus *Begomovirus*, family *Geminiviridae*) possess one or two genomic components of circular, single-stranded DNA (ssDNA) encapsidated in geminate icosahedral particles. Most begomoviruses in the New World (NW) are bipartite, while the majority in the Old World (OW) are monopartite and frequently associated with two classes of satellite DNAs: alphasatellites and betasatellites. Alphasatellites have been recently reported in association with bipartite begomoviruses in the NW. Members of the genus *Begomovirus* are responsible for important crop diseases worldwide. Begomoviruses that infect *Ipomoea* spp. (family Convolvulaceae), commonly known as "sweepoviruses", have the typical genomic organization of OW monopartite viruses but are phylogenetically distinct from all other species in the genus. Understanding the dynamics and genetic variability of viral populations in non-cultivated hosts is important for the prediction and consequent prevention of new virus diseases in cultivated plants. This work aimed to: (*i*) assess the diversity of begomoviruses in two non-cultivated hosts widely distributed in Brazil, *Sida* spp. and *Leonurus sibiricus*; (*ii*) determine the genetic structure and variability of begomovirus populations infecting *Sida* spp. and *L. sibiricus*; (*iii*) determine the genetic variability of deltasatellites associated with sweepoviruses infecting *Ipomoea indica* in Spain and expand the analysis to other geographical areas. For the first two objectives, total DNA was extracted from samples of *Sida* spp. and *L. sibiricus* collected in the states of Rio Grande do Sul, Paraná and Mato Grosso do Sul from 2009 to 2011 and viral genomes were amplified, cloned and sequenced. In *Sida* spp. the most prevalent virus was *Sida micrantha mosaic virus* (SiMMV). In addition, three new species were also detected. The vast majority of *L. sibiricus* samples were infected by *Tomato yellow spot virus* (ToYSV). Two alphasatellites were found: *Euphorbia* yellow mosaic alphasatellite in *Sida* sp. and a new alphasatellite associated with ToYSV in *L. sibiricus*. Both the SiMMV and ToYSV populations have a high degree of genetic variability. Although a high level of recombination was detected, mutational dynamics was the primary

factor of diversification. The results were inconclusive regarding genetic structuration based on geography. For the third objective, DNA was extracted from *I. indica* samples collected in southern Spain in 2015 and rolling circle amplification (RCA), a technique that allows amplification of circular ssDNA molecules without previous knowledge of nucleotide sequence, was used to clone deltasatellites and two associated sweepoviruses, *Sweet potato leaf curl virus* (SPLCV) and *Sweet potato mosaic virus* (SPMV). Deltasatellites showed low sequence variability, no evidence of recombination, and no obvious geographical structuration. Also, a sweepovirus-deltasatellite chimera with a size similar to deltasatellites was detected.

INTRODUÇÃO GERAL

A família *Geminiviridae* é composta por vírus com genoma de DNA circular de fita simples, encapsidado em partículas icosaédricas geminadas (Brown *et al.*, 2012). A família inclui os gêneros *Begomovirus*, *Becurtovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* e *Turncurtovirus*, definidos com base no tipo de inseto vetor, gama de hospedeiros, relacionamento filogenético e organização genômica (Brown *et al.*, 2012; Varsani *et al.*, 2014). O gênero *Begomovirus* é o maior na família e inclui vírus com genoma mono- ou bissegmentado, transmitidos por *Bemisia tabaci* (Hemiptera: Aleyrodidae) a plantas dicotiledôneas (Brown *et al.*, 2015). São responsáveis por sérias doenças causadas em culturas de grande importância econômica em regiões tropicais e subtropicais (Polston e Anderson, 1997; Morales e Anderson, 2001; Monci *et al.*, 2002; Briddon, 2003; Were *et al.*, 2004).

Com base em estudos filogenéticos e características do genoma, os begomovírus podem ser divididos em dois grupos: Velho Mundo (VM; Europa, África e Ásia) e Novo Mundo (NM; Américas) (Rybicki, 1994; Padidam *et al.*, 1999; Paximadis *et al.*, 1999). Begomovírus presentes no NM são predominantemente bissegmentados (dois componentes denominados DNA-A e DNA-B), com um único relato de um vírus monossegmentado nativo do NM, o *Tomato leaf deformation virus* (ToLDeV) (Melgarejo *et al.*, 2013; Sanchez-Campos *et al.*, 2013). O DNA-A contém genes envolvidos na replicação, encapsidação da progênie viral e supressão da respostas de defesa do hospedeiro (Rojas *et al.*, 2005; Hanley-Bowdoin *et al.*, 2013), e o DNA-B contém genes requeridos para o movimento intra- e intercelular na planta, determinação da gama de hospedeiros e supressão de defesas do hospedeiro

(Rojas *et al.*, 2005; Mahajan *et al.*, 2011; Hanley-Bowdoin *et al.*, 2013; Brustolini *et al.*, 2015). A grande maioria dos begomovírus que ocorrem no VM são monossegmentados e têm organização genômica semelhante à citada acima para o DNA-A dos vírus bissegmentados (Padidam *et al.*, 1996; Mansoor *et al.*, 2003), com a presença de uma ORF adicional que codifica a proteína nomeada V2/AV2, a qual está envolvida no movimento viral e na supressão do silenciamento gênico (Rybicki, 1994; Padidam *et al.*, 1996; Glick *et al.*, 2008).

Os begomovírus que infectam a batata-doce (*Ipomoea batatas*) e outras espécies da família *Convolvulaceae* (comumente denominados "sweepovírus") possuem organização genômica típica dos vírus monossegmentados do VM, porém são filogeneticamente distintos de todas as demais espécies do gênero, se agrupando basalmente aos begomovírus do VM e do NM (Lozano *et al.*, 2009; Lozano *et al.*, 2016). Nos últimos 20 anos, um grande número de sweepovírus foi identificado em vários países (Lotrakul *et al.*, 1998; Banks *et al.*, 1999; Fuentes e Salazar, 2003; Lotrakul *et al.*, 2003; Briddon *et al.*, 2006; Luan *et al.*, 2006; Miano *et al.*, 2006; Luan *et al.*, 2007; Prasanth e Hegde, 2008; Lozano *et al.*, 2009; Paprotka *et al.*, 2010a; Albuquerque *et al.*, 2011; Wasswa *et al.*, 2011; Esterhuizen *et al.*, 2012). O fato desses vírus serem comuns no NM é atribuído ao caráter de propagação vegetativa da batata-doce associado ao comércio internacional de material propagativo, que permitiu a disseminação dos sweepovírus por todo o mundo (Paprotka *et al.*, 2010a; Albuquerque *et al.*, 2012).

DNAs e RNAs satélites são agentes subvirais que não codificam proteínas relacionadas à replicação, dependendo da co-infecção de uma célula hospedeira por um vírus auxiliar para a sua multiplicação, e cujo material genético não possui identidade de sequência com o genoma do vírus auxiliar (Simon *et al.*, 2004).

O primeiro DNA satélite descoberto, nomeado ToLCV-sat, foi isolado de plantas de tomateiro infectadas com o begomovírus monossegmentado *Tomato leaf curl virus* (ToLCV) na Austrália (Dry *et al.*, 1997). O ToLCV-sat possui apenas 682 nucleotídeos, não possui ORF, contém uma região rica em adenina e duas estruturas em forma de grampo, uma contendo o nanonucleotídeo TAATATTAC (conservado nos geminivírus) e outra contendo uma sequência idêntica à do iteron do ToLCV (Saunders *et al.*, 2000). Entretanto, a grande maioria dos begomovírus monossegmentados que ocorrem no VM estão associados a dois outros tipos de DNAs satélites, denominados alfa- e betassatélites (Mansoor *et al.*, 2003; Briddon e Stanley, 2006; Zhou, 2013). Em comum, esses dois tipos de DNAs satélites possuem um genoma de ssDNA circular com aproximadamente 1300 nucleotídeos (Briddon e Stanley, 2006).

Alfassatélites (anteriormente denominados DNA-1) (Briddon *et al.*, 2004) codificam uma proteína associada à replicação (alpha-Rep) (Briddon *et al.*, 2004; Xie *et al.*, 2010; Zhou, 2013), e portanto não são DNAs satélites em sentido específico. Entretanto, necessitam de um vírus auxiliar para movimentar-se nas plantas, bem como para sua transmissão pelo inseto vetor (Mansoor *et al.*, 1999; Saunders e Stanley, 1999). Seu genoma possue uma região rica em adeninas e uma estrutura em forma de grampo semelhante à encontrada no genoma dos begomovírus. A estrutura em forma de grampo contém o nonanucleotídeo conservado TAGTATTAC, idêntica à encontrada nos nanovírus (família *Nanoviridae*) e semelhante ao nonanucleotídeo TAATATTAC conservado nos begomovírus. A proteína alpha-Rep possue similaridade de sequência significativa com as proteínas Rep de nanovírus. De fato, a região rica em adenina é a única característica que diferencia os alfassatélites do componente DNA-R dos nanovírus. É consenso que os alfassatélites se originaram

de nanovírus (Stanley, 2004; Rojas *et al.*, 2005), e acredita-se que essa região rica em adenina permanece conservada por aumentar o tamanho do genoma, de aproximadamente 1000 nucleotídeos no caso dos componentes genômicos dos nanovírus, para 1300 nucleotídeos, a metade do tamanho do genoma dos begomovírus (Briddon *et al.*, 2004). Os alfasatélites normalmente ocorrem associados a complexos begomovírus/betassatélites no VM (Xie *et al.*, 2010; Leke *et al.*, 2015), porém recentemente foram relatados em associação com begomovírus bissegmentados no NM (Brasil, Cuba e Venezuela) (Paprotka *et al.*, 2010b; Romay *et al.*, 2010; Jeske *et al.*, 2014).

Betassatélites (previamente conhecidos como DNA-β) tem sido identificados apenas no VM (Saunders *et al.*, 2000; Briddon *et al.*, 2001). Esses agentes possuem uma organização genômica contendo uma região altamente conservada (*satellite conserved region*, SCR), uma região rica em adenina, e um único gene denominado *betaC1* (Briddon *et al.*, 2008). A proteína betaC1 é multifuncional e atua como determinante de patogenicidade (Cui *et al.*, 2004), possuindo funções de supressora de silenciamento gênico transcricional (Yang *et al.*, 2011) e pós-transcricional (Li *et al.*, 2014).

Recentemente, um tipo distinto de DNA satélite foi identificado em associação a begomovírus infectando plantas da família Malvaceae em Cuba (Fiallo-Olive *et al.*, 2012a). Estes satélites possuem as mesmas características genômicas do ToLCV-sat (Fiallo-Olive *et al.*, 2012a). DNAs satélites apresentando características similares foram encontrados na Flórida (EUA), por meio de análise metagenômica de moscas-brancas adultas (*vector-enabled metagenomics*, VEM) (Ng *et al.*, 2011). Lozano *et al.* (2016) detectaram uma quimera vírus-satélite em uma planta de batata-doce, a partir da qual foram desenhados oligonucleotídeos específicos para

amplificar a sequência correspondente ao satélite. A amplificação via PCR utilizando esses oligonucleotídeos levou à detecção de DNAs satélites, também estruturalmente semelhante ao ToLCV-sat, em associação com sweepovírus em batata-doce e *Ipomoea indica* na Espanha (Lozano *et al.*, 2016). Utilizando a técnica de amplificação por círculo rolante (*rolling-circle amplification*, RCA), pequenos DNAs satélites foram detectados em plantas de *Merremia dissecta* coletadas na Venezuela (Lozano *et al.*, 2016).

Embora todos esses satélites semelhantes ao ToLCV-sat compartilhem algumas características com os betassatélites, eles não possuem o gene *betaC1*, possuem sequências que divergem significativamente da SCR, e ocorrem na ausência de betassatélites (Lozano *et al.*, 2016). Com base nessas características, Lozano *et al.* (2016) sugerem que esses satélites semelhantes ao ToLCV-sat devem ser coletivamente considerados uma classe distinta de DNAs satélites, para a qual o nome deltassatélite foi proposto.

Fiallo-Olivé *et al.* (2012a) observaram que a distribuição da variabilidade genética de deltassatélites associados a plantas não-cultivadas da família Malvaceae em Cuba é determinada por geografia e fatores do hospedeiro. Lozano *et al.* (2016) observaram que os deltassatélites associados a sweepovírus em plantas do gênero *Ipomoea*, obtidos na Espanha, formam um único grupo monofilético. Dentro desse grupo os isolados se organizam de acordo com a origem geográfica (um subgrupo contendo isolados da Espanha e outro subgrupo com isolados das Ilhas Canárias). Além disso, dentro do subgrupo espanhol existe uma subdivisão adicional: os isolados obtidos a partir de *I. indica* formam um único grupo, o qual é mais relacionado aos deltassatélites obtidos de *I. batatas* na Espanha, do que aos deltassatélites de *I. batatas* obtidos nas Ilhas Canárias (Lozano *et al.*, 2016).

Relatos recentes da associação de begomovírus do NM com alfassatélites (Paprotka *et al.*, 2010b; Romay *et al.*, 2010; Jeske *et al.*, 2014), da existência de uma terceira classe de DNAs satélites (deltassatéllites) associados a begomovírus infectando plantas cultivadas e não-cultivadas (Dry *et al.*, 1997; Ng *et al.*, 2011; Fiallo-Olivé *et al.*, 2012a; Lozano *et al.*, 2016), e da emergência de um begomovírus do Novo Mundo possuindo apenas um componente genômico (Melgarejo *et al.*, 2013) sugerem que a variabilidade genética dos begomovírus é muito maior do que se acredita.

As maiores fontes de variabilidade genética de vírus de plantas são mutação, recombinação e pseudo-recombinação (Monci *et al.*, 2002; García-Arenal *et al.*, 2003; Seal *et al.*, 2006). Para os begomovírus, assim como para outros vírus, a evolução depende primariamente da ocorrência de mutações. Os begomovírus apresentam taxas de substituição de nucleotídeos elevadas em relação a outros vírus com genoma de DNA (de fato, elas são comparáveis às de vírus com genoma composto por RNA de fita simples) (Duffy e Holmes, 2008; Duffy e Holmes, 2009). Embora a dinâmica mutacional seja o fator primário da diversificação das populações virais (Roossinck, 1997; García-Arenal *et al.*, 2003), ela não é responsável por toda a variação genética, já que outros mecanismos evolucionários, incluindo a recombinação, podem contribuir significativamente (Martin *et al.*, 2011; Lima *et al.*, 2013). A recombinação é um evento bastante comum em geminivírus (Padidam *et al.*, 1999; Lefeuvre *et al.*, 2009), e parece contribuir de forma significativa para a diversificação genética dos begomovírus, aumentando seu potencial evolucionário e adaptação local (Harrison e Robinson, 1999; Padidam *et al.*, 1999; Berrie *et al.*, 2001; Monci *et al.*, 2002). Recombinantes naturais de begomovírus foram responsáveis por epidemias severam e grandes perdas econômicas nas culturas da

mandioca na África (Zhou *et al.*, 1997; Pita *et al.*, 2001), tomateiro na Espanha (Monci *et al.*, 2002; García-Andrés *et al.*, 2007a; García-Andrés *et al.*, 2007b) e algodão e quiabo no Paquistão (Zhou *et al.*, 1998; Idris e Brown, 2002; Briddon *et al.*, 2014). Outra fonte de variabilidade genética é a pseudo-recombinação, que pode ocorrer por meio da troca entre os dois componentes genômicos de vírus bissegmentados (Andrade *et al.*, 2006).

A estrutura genética de populações de vírus de plantas refere-se à quantidade de variabilidade genética e sua distribuição dentro e entre subpopulações (García-Arenal *et al.*, 2001). Diversos estudos tem sido conduzidos para investigar os principais mecanismos responsáveis por moldar a estrutura genética e a variabilidade de populações de geminivírus em diferentes hospedeiros e regiões geográficas (Zhou *et al.*, 1997; Fondong *et al.*, 2000; Legg e Thresh, 2000; Pita *et al.*, 2001; Owor *et al.*, 2007; Lima *et al.*, 2013; Rocha *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014).

Uma elevada diversidade de espécies de begomovírus em plantas não-cultivadas tem sido observada em diversos países (Frischmuth *et al.*, 1997; Roye *et al.*, 1997; Idris *et al.*, 2003; Jovel *et al.*, 2004; Amarakoon *et al.*, 2008; Fiallo-Olive *et al.*, 2012b; Silva *et al.*, 2012; Tavares *et al.*, 2012) e algumas dessas espécies virais também infectam plantas cultivadas (Fernandes *et al.*, 2009; Fernandes *et al.*, 2011; Barbosa *et al.*, 2012; Barreto *et al.*, 2013; Rocha *et al.*, 2013; Fernandes *et al.*, 2014; Ramos-Sobrinho *et al.*, 2014), destacando a necessidade de estudar essas plantas como reservatórios de diversidade viral e como fonte de novos vírus que podem causar doenças em culturas economicamente importantes.

Este estudo teve como objetivos: (*i*) estudar a diversidade de espécies de begomovírus presentes em dois hospedeiros não-cultivados amplamente distribuídos no Brasil, *Sida* spp. e *Leonurus sibiricus*; (*ii*) determinar a estrutura genética e os

mecanismos que influenciam a variabilidade de populações de begomovírus infectando *Sida* spp. e *Leonurus sibiricus*; (*iii*) determinar a variabilidade genética de deltassatélites associados a sweepovírus infectando *Ipomoea indica*, por reamostragem das populações no sul da Espanha e expandindo as análises para outras regiões geográficas.

REFERÊNCIAS

ALBUQUERQUE, L.C.; INOUE-NAGATA, A.K.; PINHEIRO, B.; RESENDE, R.O.; MORIONES, E.; NAVAS-CASTILLO, J. Genetic diversity and recombination analysis of sweepoviruses from Brazil. **Virology Journal,** v. 9, p. 241, 2012.

ALBUQUERQUE, L.C.; INOUE-NAGATA, A.K.; PINHEIRO, B.; RIBEIRO, S.D.; RESENDE, R.O.; MORIONES, E.; NAVAS-CASTILLO, J. A novel monopartite begomovirus infecting sweet potato in Brazil. **Archives of Virology,** v. 156, p. 1291-1294, 2011.

AMARAKOON, I.I.; ROYE, M.E.; BRIDDON, R.W.; BEDFORD, I.D.; STANLEY, J. Molecular and biological characterization of *Macroptilium yellow mosaic virus* from Jamaica. **Plant Pathology,** v. 57, p. 417-426, 2008.

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,** v. 87, p. 3687-3696, 2006.

BANKS, G.; BEDFORD, I.; BEITIA, F.; RODRIGUEZ-CEREZO, E.; MARKHAM, P. A novel geminivirus of Ipomoea indica (Convolvulacae) from Southern Spain. **Plant Disease,** v. 83, p. 486-486, 1999.

BARBOSA, J.C.; ECKSTEIN, B.; BERGAMIN-FILHO, A.; REZENDE, J.A.M.; DALLAGNOL, L.J. First report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Brazil. **Plant Disease,** v. 97, p. 289-289, 2012.

BARRETO, S.S.; HALLWASS, M.; AQUINO, O.M.; INOUE-NAGATA, A.K. A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. **Phytopathology,** v. 103, p. 436-44, 2013.

BERRIE, L.C.; RYBICKI, E.P.; REY, M.E.C. Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. **Journal of General Virology,** v. 82, p. 53-8., 2001.

BRIDDON, R.; BULL, S.; BEDFORD, I. Occurrence of Sweet potato leaf curl virus in Sicily. **Plant Pathology,** v. 55, p. 286-286, 2006.

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

BRIDDON, R.W.; AKBAR, F.; IQBAL, Z.; AMRAO, L.; AMIN, I.; SAEED, M.; MANSOOR, M.S. Effects of genetic changes to the begomovirus/betasatellite complex

causing cotton leaf curl disease in South Asia post-resistance breaking. **Virus Research,** v. 186, p. 114-119, 2014.

BRIDDON, R.W.; BROWN, J.K.; MORIONES, E.; STANLEY, J.; ZERBINI, F.M.; ZHOU, X.; FAUQUET, C.M. Recommendations for the classification and nomenclature of the DNA-beta satellites of begomoviruses. **Archives of Virology,** v. 153, p. 763-781, 2008.

BRIDDON, R.W.; BULL, S.E.; AMIN, I.; MANSOOR, S.; BEDFORD, I.D.; RISHI, N.; SIWATCH, S.S.; ZAFAR, Y.; ABDEL-SALAM, A.M.; MARKHAM, P.G. Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. **Virology,** v. 324, p. 462-474, 2004.

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,** v. 285, p. 234-43, 2001.

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

BROWN, J.K.; FAUQUET, C.M.; BRIDDON, R.W.; ZERBINI, F.M.; MORIONES, E.; NAVAS-CASTILLO, J. Family *Geminiviridae*. pp. 351-373, In: KING, A.M.Q.;ADAMS, M.J.;CARSTENS, E.B. e LEFKOWITZ, E.J. (Ed.^Eds.). **Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses**. London, UK: Elsevier Academic Press. 2012.

BROWN, J.K.; ZERBINI, F.M.; NAVAS-CASTILLO, J.; MORIONES, E.; RAMOS-SOBRINHO, R.; SILVA, J.C.; FIALLO-OLIVÉ, E.; BRIDDON, R.W.; HERNANDEZ-ZEPEDA, C.; IDRIS, A.; MALATHI, V.G.; MARTIN, D.P.; RIVERA-BUSTAMANTE, R.; UEDA, S.; VARSANI, A. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. **Arch Virol,** v. 160, p. 1593-619, 2015.

BRUSTOLINI, O.J.; MACHADO, J.P.; CONDORI-APFATA, J.A.; COCO, D.; DEGUCHI, M.; LORIATO, V.A.; PEREIRA, W.A.; ALFENAS-ZERBINI, P.; ZERBINI, F.M.; INOUE-NAGATA, A.K.; SANTOS, A.A.; CHORY, J.; SILVA, F.F.; FONTES, E.P. Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. **Plant Biotechnol J,** v. 13, p. 1300-1311, 2015.

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,** v. 78, p. 13966-13974, 2004.

DRY, I.B.; KRAKE, L.R.; RIGDEN, J.E.; REZAIAN, M.A. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. **Proceedings of the National Academy of Sciences, USA,** v. 94, p. 7088-7093, 1997.

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,** 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,** v. 90, p. 1539-1547, 2009.

ESTERHUIZEN, L.L.; VAN HEERDEN, S.W.; REY, M.E.C.; VAN HEERDEN, H. Genetic identification of two sweet-potato-infecting begomoviruses in South Africa. **Archives of Virology,** v. 157, p. 2241-2245, 2012.

FERNANDES, F.R.; ALBUQUERQUE, L.C.; OLIVEIRA, C.L.; CRUZ, A.R.R.; ROCHA, W.B.; PEREIRA, T.G.; NAITO, F.Y.B.; DIAS, N.D.; NAGATA, T.; FARIA, J.C.; ZERBINI, F.M.; ARAGÃO, F.J.L.; INOUE-NAGATA, A.K. Molecular and biological characterization of a new Brazilian begomovirus, euphorbia yellow mosaic virus

(EuYMV), infecting *Euphorbia heterophylla* plants. **Archives of Virology,** v. 156, p. 2063-2069, 2011.

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,** v. 154, p. 1567-1570, 2009.

FERNANDES, N.A.N.; BOITEUX, L.S.; FONSECA, M.E.N.; GONZALES-SEGNANA, L.; KITAJIMA, E.W. Report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Paraguay and within tomato fields in Brazil. **Plant Disease,** v. 98, p. 1445, 2014.

FIALLO-OLIVÉ, E.; MARTINEZ-ZUBIAUR, Y.; MORIONES, E.; NAVAS-CASTILLO, J. A novel class of DNA satellites associated with New World begomoviruses. **Virology,** v. 426, p. 1-6, 2012a.

FIALLO-OLIVÉ, E.; NAVAS-CASTILLO, J.; MORIONES, E.; MARTINEZ-ZUBIAUR, Y. Begomoviruses infecting weeds in Cuba: Increased host range and a novel virus infecting *Sida rhombifolia*. **Archives of Virology,** v. 157, p. 141-146, 2012b.

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,** v. 81, p. 287-297, 2000.

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. **J Gen Virol,** v. 78, p. 2675-82, 1997.

FUENTES, S.; SALAZAR, L. First report of Sweet potato leaf curl virus in Peru. **Plant Disease,** v. 87, p. 98-98, 2003.

GARCÍA-ANDRÉS, 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,** v. 359, p. 302-312, 2007a.

GARCÍA-ANDRÉS, 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,** v. 365, p. 210-219, 2007b.

GARCÍA-ARENAL, F.; FRAILE, A.; MALPICA, J.M. Variability and genetic structure of plant virus populations. **Annual Review of Phytopathology,** v. 39, p. 157-86, 2001.

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

GLICK, E.; ZRACHYA, A.; LEVY, Y.; METT, A.; GIDONI, D.; BELAUSOV, E.; CITOVSKY, V.; GAFNI, Y. Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein. **Proceedings of the National Academy of Sciences of the United States of America,** v. 105, p. 157-161, 2008.

HANLEY-BOWDOIN, L.; BEJARANO, E.R.; ROBERTSON, D.; MANSOOR, S. Geminiviruses: Masters at redirecting and reprogramming plant processes. **Nature Reviews Microbiology,** v. 11, p. 777-88, 2013.

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

IDRIS, A.M.; BROWN, J.K. Molecular analysis of Cotton leaf curl virus-Sudan reveals an evolutionary history of recombination. **Virus Genes,** 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,** v. 93, p. 774-783, 2003.

JESKE, H.; KOBER, S.; SCHÄFER, B.; STROHMEIER, S. Circomics of Cuban geminiviruses reveals the first alpha-satellite DNA in the Caribbean. **Virus Genes,** v. 49, p. 312-324, 2014.

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*. **Arch Virol,** v. 149, p. 829-41, 2004.

LEFEUVRE, P.; LETT, J.M.; VARSANI, A.; MARTIN, D.P. Widely conserved recombination patterns among single-stranded DNA viruses. **Journal of Virology,** v. 83, p. 2697-2707, 2009.

LEGG, J.P.; THRESH, J.M. Cassava mosaic virus disease in East Africa: A dynamic disease in a changing environment. **Virus Research,** v. 71, p. 135-149, 2000.

LEKE, W.N.; MIGNOUNA, D.B.; BROWN, J.K.; KVARNHEDEN, A. Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. **Agriculture & Food Security,** v. 4, p. 1, 2015.

LI, F.; HUANG, C.; LI, Z.; ZHOU, X. Suppression of RNA silencing by a plant DNA virus satellite requires a host calmodulin-like protein to repress *RDR6* expression. **PLoS Pathogens,** v. 10, p. e1003921, 2014.

LIMA, A.T.M.; SOBRINHO, R.R.; GONZALEZ-AGUILERA, J.; ROCHA, C.S.; SILVA, S.J.C.; XAVIER, C.A.D.; SILVA, F.N.; DUFFY, S.; ZERBINI, F.M. Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. **Journal of General Virology,** v. 94, p. 418-431, 2013.

LOTRAKUL, P.; VALVERDE, R.; CLARK, C.; SIM, J.; DE LA TORRE, R. Detection of a geminivirus infecting sweet potato in the United States. **Plant Disease,** v. 82, p. 1253-1257, 1998.

LOTRAKUL, P.; VALVERDE, R.A.; CLARK, C.A.; FAUQUET, C. Properties of a begomovirus isolated from sweet potato [Ipomoea batatas (L.) Lam.] infected with Sweet potato leaf curl virus. **Rev Mex Fitopatol,** v. 21, p. 128-136, 2003.

LOZANO, G.; TRENADO, H.P.; FIALLO-OLIVÉ, E.; CHIRINOS, D.; GERAUD-POUEY, F.; BRIDDON, R.W.; NAVAS-CASTILLO, J. Characterization of non-coding DNA satellites associated with sweepoviruses (genus Begomovirus, Geminiviridae) - definition of a distinct class of begomovirus-associated satellites. **Frontiers in Microbiology,** v. 7, doi 10.3389/fmicb.2016.00162, 2016.

LOZANO, G.; TRENADO, H.P.; VALVERDE, R.A.; NAVAS-CASTILLO, J. Novel begomovirus species of recombinant nature in sweet potato (*Ipomoea batatas*) and *Ipomoea indica*: Taxonomic and phylogenetic implications. **Journal of General Virology,** v. 90, p. 2550-2562, 2009.

LUAN, Y.; ZHANG, J.; AN, L. First report of Sweet potato leaf curl virus in China. **Plant Disease,** v. 90, p. 1111-1111, 2006.

LUAN, Y.S.; ZHANG, J.; LIU, D.M.; LI, W.L. Molecular characterization of sweet potato leaf curl virus isolate from China (SPLCV-CN) and its phylogenetic relationship with other members of the Geminiviridae. **Virus Genes,** v. 35, p. 379-85, 2007.

MAHAJAN, N.; PARAMESWARI, C.; VELUTHAMBI, K. Severe stunting in blackgram caused by the Mungbean yellow mosaic virus (MYMV) KA27 DNA B component is

ameliorated by co-infection or post-infection with the KA22 DNA B: MYMV nuclear shuttle protein is the symptom determinant. **Virus Research,** v. 157, p. 25-34, 2011.

MANSOOR, S.; BRIDDON, R.W.; ZAFAR, Y.; STANLEY, J. Geminivirus disease complexes: An emerging threat. **Trends in Plant Science,** v. 8, p. 128-34, 2003.

MANSOOR, S.; KHAN, S.H.; BASHIR, A.; SAEED, M.; ZAFAR, Y.; MALIK, K.A.; BRIDDON, R.; STANLEY, J.; MARKHAM, P.G. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. **Virology,** v. 259, p. 190-9., 1999.

MARTIN, D.P.; BIAGINI, P.; LEFEUVRE, P.; GOLDEN, M.; ROUMAGNAC, P.; VARSANI, A. Recombination in eukaryotic single stranded DNA viruses. **Viruses,** v. 3, p. 1699-1738, 2011.

MELGAREJO, T.A.; KON, T.; ROJAS, M.R.; PAZ-CARRASCO, L.; ZERBINI, F.M.; GILBERTSON, R.L. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. **Journal of Virology,** v. 87, p. 5397-5413, 2013.

MIANO, D.W.; LABONTE, D.R.; CLARK, C.A.; VALVERDE, R.A.; HOY, M.W. First report of a begomovirus infecting sweetpotato in Kenya. **Plant Disease,** v. 90, p. 832-832, 2006.

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,** v. 303, p. 317-26, 2002.

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

NG, T.F.F.; DUFFY, S.; POLSTON, J.E.; BIXBY, E.; VALLAD, G.E.; BREITBART, M. Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies. **PLoS ONE,** v. 6, p. e19050, 2011.

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,** v. 88, p. 3154-3165, 2007.

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,** v. 224, p. 390-404, 1996.

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

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,** v. 149, p. 224-233, 2010a.

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

PAXIMADIS, M.; IDRIS, A.M.; TORRES-JEREZ, I.; VILLARREAL, A.; REY, M.E.C.; BROWN, J.K. Characterization of tobacco geminiviruses in the Old and New world. **Archives of Virology,** v. 144, p. 703-717, 1999.

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,** v. 82, p. 655-65, 2001.

POLSTON, J.E.; ANDERSON, P.K. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. **Plant Disease,** v. 81, p. 1358-1369, 1997.

PRASANTH, G.; HEGDE, V. Occurrence of Sweet potato feathery mottle virus and Sweet potato leaf curl Georgia virus on sweet potato in India. **Plant Disease,** v. 92, p. 311-311, 2008.

RAMOS-SOBRINHO, R.; XAVIER, C.A.D.; PEREIRA, H.M.D.B.; LIMA, G.S.D.A.; ASSUNÇÃO, I.P.; MIZUBUTI, E.S.G.; DUFFY, S.; ZERBINI, F.M. Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. **Journal of General Virology,** v. 95, p. 2540-2552, 2014.

ROCHA, C.S.; CASTILLO-URQUIZA, G.P.; LIMA, A.T.M.; SILVA, F.N.; XAVIER, C.A.D.; HORA-JUNIOR, B.T.; BESERRA-JUNIOR, J.E.A.; MALTA, A.W.O.; MARTIN, D.P.; VARSANI, A.; ALFENAS-ZERBINI, P.; MIZUBUTI, E.S.G.; ZERBINI, F.M. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. **Journal of Virology,** v. 87, p. 5784-5799, 2013.

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,** v. 43, p. 361-394, 2005.

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

ROOSSINCK, M.J. Mechanisms of plant virus evolution. **Annual Review of Phytopathology,** v. 35, p. 191-209, 1997.

ROYE, M.E.; MCLAUGHLIN, W.A.; NAKHLA, M.K.; MAXWELL, D.P. Genetic diversity among geminiviruses associated with the weed species *Sida* spp., *Macroptilium lathyroides*, and *Wissadula amplissima* from Jamaica. **Plant Disease,** v. 81, p. 1251-1258, 1997.

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

SANCHEZ-CAMPOS, S.; MARTINEZ-AYALA, A.; MARQUEZ-MARTIN, B.; ARAGON-CABALLERO, L.; NAVAS-CASTILLO, J.; MORIONES, E. Fulfilling Koch's postulates confirms the monopartite nature of tomato leaf deformation virus: A begomovirus native to the New World. **Virus Research,** v. 173, p. 286-293, 2013.

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, USA,** v. 97, p. 6890-5, 2000.

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,** v. 264, p. 142-152, 1999.

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,** v. 25, p. 23-46, 2006.

SILVA, S.J.C.; CASTILLO-URQUIZA, G.P.; HORA-JUNIOR, B.T.; ASSUNÇÃO, I.P.; LIMA, G.S.A.; PIO-RIBEIRO, G.; MIZUBUTI, E.S.G.; ZERBINI, F.M. Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. **Plant Pathology,** v. 61, p. 457-467, 2012.

SIMON, A.E.; ROOSSINCK, M.J.; HAVELDA, Z. Plant virus satellite and defective interfering RNAs: New paradigms for a new century. **Annual Review of Phytopathology,** v. 42, p. 415-437, 2004.

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

TAVARES, S.S.; RAMOS-SOBRINHO, R.; GONZALEZ-AGUILERA, J.; LIMA, G.S.A.; ASSUNÇÃO, I.P.; ZERBINI, F.M. Further molecular characterization of weed-associated begomoviruses in Brazil with an emphasis on *Sida* spp. **Planta Daninha,** v. 30, p. 305-315, 2012.

VARSANI, A.; NAVAS-CASTILLO, J.; MORIONES, E.; HERNÁNDEZ-ZEPEDA, C.; IDRIS, A.; BROWN, J.K.; ZERBINI, F.M.; MARTIN, D.P. Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. **Archives of Virology,** v. 159, p. 2193-2203, 2014.

WASSWA, P.; OTTO, B.; MARUTHI, M.; MUKASA, S.; MONGER, W.; GIBSON, R. First identification of a sweet potato begomovirus (sweepovirus) in Uganda: characterization, detection and distribution. **Plant Pathology,** v. 60, p. 1030-1039, 2011.

WERE, H.K.; WINTER, S.; MAISS, E. Viruses infecting cassava in Kenya. **Plant Disease,** v. 88, p. 17-22, 2004.

XIE, Y.; WU, P.J.; LIU, P.; GONG, H.R.; ZHOU, X.P. Characterization of alphasatellites associated with monopartite begomovirus/betasatellite complexes in Yunnan, China. **Virology Journal,** v. 7, p. 178, 2010.

YANG, X.L.; XIE, Y.; RAJA, P.; LI, S.Z.; WOLF, J.N.; SHEN, Q.T.; BISARO, D.M.; ZHOU, X.P. Suppression of methylation-mediated transcriptional gene silencing by beta C1-SAHH protein interaction during geminivirus-betasatellite infection. **PLoS Pathogens,** v. 7, p. e1002329, 2011.

ZHOU, X. Advances in understanding begomovirus satellites. **Annual Review of Phytopathology,** v. 51, p. 357-381, 2013.

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,** 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,** v. 79, p. 915-923, 1998.

CAPÍTULO 1

THE EVER INCREASING DIVERSITY OF BEGOMOVIRUSES INFECTING NON-CULTIVATED HOSTS: NEW SPECIES FROM *Sida* spp. AND *Leonurus sibiricus*, PLUS TWO NEW WORLD ALPHASATELLITES

Ferro, C.G., Silva, J.P., Xavier, C.A.D., Godinho, M.T., Lima, A.T.M., Mar, T.B., Lau, D., Zerbini, F.M. The ever increasing diversity of begomoviruses infecting non-cultivated hosts: new species from *Sida* spp. and *Leonurus sibiricus*, plus two New World alphasatellites. Annals of Applied Biology, *submitted*.

THE EVER INCREASING DIVERSITY OF BEGOMOVIRUSES INFECTING NON-CULTIVATED HOSTS: NEW SPECIES FROM *Sida* spp. AND *Leonurus sibiricus*, PLUS TWO NEW WORLD ALPHASATELLITES

C.G. Ferro[1], J.P. Silva[1], C.A.D. Xavier[1], M.T. Godinho[1], A.T.M. Lima[1*], Mar, T.B.[1], D. Lau[2], F.M. Zerbini[1#]

[1]Dep. de Fitopatologia/BIOAGRO and National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[2]Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Trigo, Passo Fundo, RS 99001-970, Brazil

[*]Present address: Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, MG 38400-902, Brazil

[#]Corresponding author:

Phone: (+55-31) 3899-2935; Fax: (+55-31) 3899-2240; E-mail: zerbini@ufv.br

Summary

Begomoviruses (whitefly-transmitted, single-stranded DNA plant viruses) are among the most damaging pathogens causing epidemics in economically important crops worldwide. Besides cultivated plants, many weed and wild hosts act as virus reservoirs where recombination may occur, resulting in new species. The aim of this study was to further characterize the diversity of begomoviruses infecting two major weed genera, *Sida* and *Leonurus*. Total DNA was extracted from samples collected in the states of Rio Grande do Sul, Paraná and Mato Grosso do Sul during the years 2009 to 2011. Viral genomes were enriched by rolling circle amplification (RCA), linearised into unit length genomes using various restriction enzymes, cloned and sequenced. A total of 78 clones were obtained: 37 clones from *Sida* spp. plants and 41 clones from *Leonurus sibiricus* plants. Sequence analysis indicated the presence of six bipartite begomovirus species and two alphasatellites. In *Sida* spp. plants we found *Sida micrantha mosaic virus* (SiMMV), *Euphorbia yellow mosaic virus* (EuYMV), and three isolates that represent new species, for which the following names are proposed: Sida chlorotic mottle virus (SiCMoV), Sida bright yellow mosaic virus (SiBYMV) and Sida golden yellow spot virus (SiGYSV), an Old World-like begomovirus. *L. sibiricus* plants had a lower diversity of begomoviruses compared to *Sida* spp., with only *Tomato yellow spot virus* (ToYSV) and EuYMV (for the first time detected infecting plants of the genus *Leonurus*) detected. Two satellite DNA molecules were found: Euphorbia yellow mosaic alphasatellite, for the first time detected infecting plants of the genus *Sida,* and a new alphasatellite associated with ToYSV in *L. sibiricus*. These results constitute further evidence of

the high species diversity of begomoviruses in non-cultivated hosts, particularly *Sida* spp.

Keywords: begomovirus, biodiversity, geminivirus, recombination, satellite DNA

Introduction

The family *Geminiviridae* is comprised of plant viruses with one or two genomic components of circular, single-strand DNA (ssDNA) encapsidated in geminate particles (Brown *et al.*, 2012). The family is divided into seven genera (*Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus*) according to the type of insect vector, host range, genome organization and phylogeny (Brown *et al.*, 2012; Varsani *et al.*, 2014). The genus *Begomovirus* is the largest in the family and includes mono- and bipartite viruses transmitted by *Bemisia tabaci* (Hemiptera: Aleyrodidae) to dicotyledonous plants (Brown *et al.*, 2015).

Based on phylogenetic analysis and genomic features, begomoviruses are broadly divided into two groups: Old World (OW; Europe, Africa, Asia and Oceania) and New World (NW; the Americas) (Rybicki, 1994; Paximadis *et al.*, 1999; Padidam *et al.*, 1999). Begomoviruses in the New World are mostly bipartite (DNA-A and DNA-B), except for *Tomato leaf deformation virus* (ToLDeV), an indigenous NW monopartite virus (Melgarejo *et al.*, 2013). The DNA-A contains genes involved in replication, encapsidation of viral progeny and suppression of host defenses (Rojas *et al.*, 2005; Hanley-Bowdoin *et al.*, 2013), and the DNA-B contains genes required for intra- and intercellular movement in the plant, host range determination and suppression of host defenses (Hanley-Bowdoin *et al.*, 2013; Rojas *et al.*, 2005; Mahajan *et al.*, 2011; Brustolini *et al.*, 2015). The majority of the begomoviruses that occur in the Old World (OW) are monopartite, with a genomic organization similar to the DNA-A of bipartite viruses (Mansoor *et al.*, 2003; Padidam *et al.*, 1996) and the presence of an additional open reading frame (ORF) which partially overlaps the *cp* gene, named *v2* in monopartite viruses or *av2* in

bipartite viruses. The V2/AV2 protein is involved in viral movement and gene silencing suppression (Rybicki, 1994; Padidam *et al.*, 1996; Glick *et al.*, 2008). Begomoviruses in the OW are generally associated with satellite DNA molecules (Zhou, 2013).

Brazil is a begomovirus diversity hotspot, with reports of their detection dating back to the 1950's (Costa, 1955; Costa and Bennett, 1950). Begomoviruses are limiting factors for common bean and tomato production (Faria *et al.*, 2000; Zerbini *et al.*, 2005), and a large number of new species of tomato-infecting begomoviruses has been identified in the country (Calegario *et al.*, 2007; Ribeiro *et al.*, 2007; Fernandes *et al.*, 2006; Ribeiro *et al.*, 2003; Castillo-Urquiza *et al.*, 2008; Fernandes *et al.*, 2008; Albuquerque *et al.*, 2012). The advent of techniques for the unbiased amplification of circular DNA genomes (specially rolling circle amplification, RCA (Inoue-Nagata *et al.*, 2004), created new possibilities for the discovery of novel begomoviruses, and also of divergent ssDNA viruses (Loconsole *et al.*, 2012; Krenz *et al.*, 2012; Basso *et al.*, 2015).

Non-cultivated species of the families Asteraceae, Caparaceae, Euphorbiaceae, Fabaceae, Labiatae, Malvaceae, Solanaceae and Sterculiaceae have been reported as hosts of many begomoviruses in Brazil and in several other countries in the Americas (Faria and Maxwell, 1999; Assunção *et al.*, 2006; Tavares *et al.*, 2012; Castillo-Urquiza *et al.*, 2008; Silva *et al.*, 2012; Barbosa *et al.*, 2009; Barreto *et al.*, 2013; Silva *et al.*, 2011; Frischmuth *et al.*, 1997; Roye *et al.*, 1997; Fernandes *et al.*, 1999; Idris *et al.*, 2003; Jovel *et al.*, 2004; Amarakoon *et al.*, 2008). There is evidence that some of these begomoviruses from non-cultivated hosts can be transmitted to cultivated species by the insect vector and by grafting (Castillo-Urquiza *et al.*, 2007; Cotrim *et al.*, 2007; Barreto *et al.*, 2013; Silva *et al.*, 2010;

Rocha *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Arnaud *et al.*, 2007), highlighting the need to investigate these plants as reservoirs of viral diversity and as a source of new viruses which may cause diseases in crops.

Materials and Methods

During the years 2009 to 2011, leaf samples from *Sida* spp. (Malvaceae) and *Leonurus sibiricus* (Lamiaceae) plants displaying symptoms of yellow mosaic and leaf distortion and/or infestation by *B. tabaci* were collected in the states of Rio Grande do Sul (*n*=27), Paraná (*n*=33) and Mato Grosso do Sul (*n*=10). Total DNA was extracted from press-dried samples as described by Doyle and Doyle (1987). Full-length viral circular genomes were enriched by rolling-circle amplification (RCA) as described by Inoue-Nagata *et al.* (2004). Unit length genomes fragments were excised with *Apa*I, *Bam*HI, *Cla*I, *Eco*RI, *Hind*III, *Kpn*I, *Sac*I, *Sal*I or *Spe*I and ligated into the pBLUESCRIPT-KS+ (pKS+) plasmid vector (Stratagene), previously cleaved with the same enzyme. Viral inserts were sequenced commercially (Macrogen Inc., Seoul, South Korea) by primer walking. All genome sequences were organized to begin at the nicking site in the invariant nonanucleotide at the origin of replication (5'-TAATATT//AC-3').

Pairwise sequence comparisons were performed using Sequence Demarcation Tool (SDT) v.1.2 (Muhire *et al.*, 2014) using the MUSCLE alignment option (Edgar, 2004). Multiple sequence alignments were obtained using the MUSCLE algorithm implemented in MEGA6 (Tamura *et al.*, 2013). Phylogenetic analyses were performed with the sequences of the closest begomoviruses determined by BLASTn comparison of the clones generated in this study and the sequences deposited in Genbank, plus some begomovirus sequences of the New and Old World.

Phylogenetic trees were constructed using Bayesian inference performed with MrBayes v. 3.0b4 (Ronquist and Huelsenbeck, 2003), with the nucleotide substitution model selected by MrModeltest v. 2.2 (Nylander, 2004) in the Akaike Information Criterion (AIC). The analysis was run for 10 million generations, excluding the first 2,000,000 generations as burn-in. The trees were visualized in FigTree v.1.3.1 (tree.bio.ed.ac.uk/software/figtree/). Recombination analysis was performed with Recombination Detection Program (RDP) v.4.5.1 (Martin *et al.*, 2010) using default settings and a Bonferroni-corrected *P*-value cutoff of 0.05. The same data set used for the phylogenetic analysis was used for recombination analysis. Only the recombination events detected by more than four of the seven tests implemented in RDP were considered to be reliable.

Results and Discussion

A total of 70 samples samples (43 *Sida* spp. samples, 27 *L. sibiricus* samples) were collected and 61 were preliminarily positive for the presence of a begomovirus, based on the detection of an ~2,600-bp band after digestion of the RCA products with restriction enzymes (data not shown). A total of 78 clones were obtained: 37 clones from *Sida* spp. plants and 41 clones from *L. sibiricus* plants (Table 1; these include DNA-A, DNA-B and satellite DNA clones). BLASTn analysis and pairwise sequence comparisons indicated the presence of six begomovirus species and two alphasatellites (Table 1; Suppl. Figure S1 and S4).

Pairwise sequence comparisons of cloned genome sequences with those deposited in GenBank indicated that *Sida micrantha mosaic virus* (SiMMV) was the predominant begomovirus infecting *Sida* spp. plants. Out of 15 DNA-A clones obtained from this host, 10 correspond to SiMMV isolates, as well as 19 out of the

21 DNA-B clones (Table 1). The DNA-A sequences share >96.3% nucleotide identity amongst themselves and 96.3% to 97.2% identity with SiMMV (accession number FN436003), and the DNA-B sequences share >88.0% nucleotide identity amongst themselves and 89.1% to 93.3% identity with SiMMV (FN436004) (Suppl. Figure S1). Bayesian phylogenetic trees based on either the DNA-A or DNA-B placed these isolates in a monophyletic branch together with SiMMV (Figure 1).

Two DNA-A sequences (BR:Trm531.2:10 and BR:Caa691:10) obtained from *Sida* spp. samples #531 and #691 displayed 97.0% and 97.2% nt sequence identity to *Euphorbia yellow mosaic virus* (EuYMV; FJ619507), respectively (Table 1; Suppl. Figure S1). This virus has already been found infecting *Sida santaremnensis* in Minas Gerais (Tavares *et al.*, 2012).

The *Sida* sample #531 actually had a mixed infection. A virus representing a new species (BR:Trm531.1:10) based on the criteria of <91% nucleotide sequence identity for the DNA-A, recently updated by the *Geminiviridae* Study Group of the ICTV (Brown *et al.*, 2015), was cloned from this sample, for which the name Sida chlorotic mottle virus (SiCMoV) is proposed. Pairwise sequence comparisons of the DNA-A sequence (2601 nt) with those deposited in GenBank indicated a maximum nucleotide sequence identity of 81.5% with *Tomato dwarf leaf virus* (ToDfLV, JN564749) (Suppl. Figure S1). A DNA-B was detected in the sample but has not yet been cloned. SiCMoV is placed in a monophyletic branch with ToDfLV, *Tomato chino La Paz virus* (ToChLPV), *Tomato leaf deformation virus* (ToLDeV), *Tomato golden mosaic virus* (TGMV) and *Sida mosaic Bolivia virus 1* (SiMBoV1) and *Abutilon mosaic Bolivia virus* (AbMBoV) (Figure 1A). This close relationship is consistent with the pairwise sequence identity analysis and with previously reported data (Medina and Lambertini, 2012; Marquez-Martin *et al.*, 2011; Melgarejo *et al.*,

2013). Interestingly, although no recombination events were detected for this virus, it clustered with viruses of the EuYMV group (including BR:Trm531.2:10 obtained from the same sample) in a *cp* nt sequence tree, but with SiMBoV1, ToDfLV and *Tomato yellow spot virus* (ToYSV) in a *rep* nt sequence tree (Suppl. Figure S2), suggesting a recombinant origin. Mixed infections by different begomoviruses are common in non-cultivated hosts (Alabi *et al.*, 2008; García-Andrés *et al.*, 2006; Monde *et al.*, 2010), facilitating recombination events among distantly related begomoviruses which may contribute to the frequent emergence of new species.

A virus corresponding to a second new species was cloned from the *Sida* spp. sample #720 (BR:Tac720:10). Nucleotide sequence identity between the common regions (CR) of the DNA-A and DNA-B was 94.7%, and the two components have identical iterons (TGGGG), indicating that they constitute a cognate pair. Both the DNA-A (2692 nt) and the DNA-B (2656 nt) show the highest nucleotide sequence identity with SiMMV (86% and 75.5%, respectively) (Suppl. Figure S1). The name Sida bright yellow mosaic virus (SiBYMV) is proposed for this new species. Analysis with the RDP4 program detected one strongly supported recombination event in the DNA-A, with SiMMV (FN436003) and an unknown virus as the putative parents (Table 2). Phylogenetic reconstruction based on the DNA-A placed this isolate in a monophyletic branch with SiMMV isolates, occupying a basal position in the clade (Figure 1A). The recombination event has strong phylogenetic support: the *rep* tree places BR:Tac720:10 in a monophyletic branch (99% posterior probability) with the SiMMV isolates, while the *cp* tree places the isolate in a monophyletic branch (99% posterior probability) with *Abutilon mosaic Brazil virus* (AbMBV, JF694480) (Suppl. Figure S2). Phylogenetic analysis based on the DNA-B placed this isolate in a monophyletic branch with SiMMV, *Tomato rugose mosaic*

virus (ToRMV, AF291706) and *Tomato severe rugose virus* (ToSRV, KC004086), occupying a basal position in the clade (Figure 1B). One recombination event was detected, with BR:Tol1075:11 and *Bean dwarf mosaic virus* (BDMV, M88180) as the isolates inferred to be most closely related to the parents (Table 2). The recombination event has good phylogenetic support: BR:Tac720:10 groups with different begomoviruses in the *mp* and *nsp* nt sequence trees (Suppl. Figure S3). Recombination is a common event among geminiviruses (Padidam *et al.*, 1999; Lefeuvre *et al.*, 2009) and contributes greatly to the genetic diversity of begomoviruses, increasing their evolutionary potential and local adaptation (Harrison and Robinson, 1999; Padidam *et al.*, 1999; Berrie *et al.*, 2001; Monci *et al.*, 2002). A number of natural begomovirus recombinants have been responsible for severe diseases and great economic losses in cassava in East Africa (Zhou *et al.*, 1997; Pita *et al.*, 2001), tomatoes in Spain (Monci *et al.*, 2002; García-Andrés *et al.*, 2006; García-Andrés *et al.*, 2007a; García-Andrés *et al.*, 2007b), and cotton and okra in Pakistan (Zhou *et al.*, 1998; Idris and Brown, 2002; Briddon *et al.*, 2014).

A DNA-A component (BR:Sab889:10) was cloned from the *Sida* spp. sample #889 (Table 1; a DNA-B was detected in sample #889 but has not yet been cloned). Pairwise comparisons indicated the highest nucleotide sequence identities of 72.1% with *Tomato mottle leaf curl virus* (ToMoLCV, JF803249) and 73.8% with two unpublished sequences obtained from *Sida acuta* samples in our laboratory (BR:Vic25.4:11 and BR:Vic26.1:11; Xavier, 2015). Thus, BR:Sab889:10 constitutes the third new virus species detected in the *Sida* spp. samples, for which the name Sida golden yellow spot virus (SiGYSV) is proposed.

Strikingly, the DNA-A components of BR:Sab889:10, BR:Vic25.4:11 and BR:Vic26.1:11 have a length within the range of OW begomoviruses (2813, 2828

and 2828 nt, respectively) and contain an *av2*-like gene (which is present only in OW begomoviruses). The region encompassing part of the CR and the *av2*-like and *cp* genes of these three components (approximately 1,100 nt) has very low similarity to any other begomovirus sequence in GenBank. The deduced amino acid sequences of their CP and AV2-like proteins were further analyzed with BLAST*p* and with the program Interpro. The analysis performed with Interpro indicated the presence of a domain related to geminivirus CPs, despite the divergence of the *cp* gene (*data not shown*). No functional domains were predicted in the AV2-like protein. BLAST*p* analysis with the CP and AV2 proteins detected only a very low similarity with a highly divergent monopartite geminivirus recently described in China infecting apple trees, named "apple geminivirus" (AGV) (Liang *et al.*, 2015). The BR:Sab889:10 CP shares 27% identity (90% coverage, E value $2e^{-12}$) with the AGV CP and the AV2-like protein shares 43% identity (83% coverage, E value $1e^{-15}$) with the putative V2 protein of AGV.

Analysis with the RDP4 program detected two strongly supported recombination events in BR:Sab889:10, with *Sida yellow leaf curl virus* (SiYLCV, KC706539), *Sida mosaic Bolivia virus 1* (SiMBoV1, HM585441) and two unknowns viruses identified as putative parents (Table 2). In the phylogenetic tree, SiGYSV was placed in a cluster with BR:Vic25.4:11, BR:Vic26.1:11 and ToMoLCV (Figure 1A), which is consistent with the pairwise identity analysis. A *cp* nt sequence tree was constructed with a data set including four highly divergent geminiviruses (Citrus chlorotic dwarf-associated virus, CCDaV; *Euphorbia caput-medusae* latent virus, EcmLV; Grapevine red blotch associated virus, GRBaV; apple geminivirus, AGV), one topocuvirus (*Tomato pseudo-curly top virus*, TPCTV), BR:Vic25.4:11 and BR:Vic26.1:11, and seven NW and OW begomoviruses (Figure 2). BR:Sab889:10

clusters with BR:Vic25.4:11, BR:Vic26.1:11 and AGV, reflecting the BLAST*p* analysis in which a similarity was found among the CPs of these four viruses. Studies are in progress in our laboratory to characterize these divergent, OW-like begomovirus species.

The vast majority of the DNA-A clones obtained from *L. sibiricus* samples corresponded to *Tomato yellow spot virus* (ToYSV) isolates (15 out of 16 DNA-A clones and 19 out of 20 DNA-B clones; Table 1). Pairwise sequence comparisons indicated that the DNA-A sequences share >93.4% sequence identity with each other and 92.3% to 95.5% nt identity with the sequence of ToYSV (DQ336350), and the DNA-B sequences share >91.9% nucleotide identity amongst themselves and 91.0% to 92.3% nt identity with ToYSV (DQ336351) (Table 1; Suppl. Figure S1). The Bayesian phylogenetics trees based on the DNA-A and DNA-B components placed these isolates in monophyletic clusters with ToYSV (Figure 1).

The cloned DNA-A and DNA-B (BR:Arm713:10) from *L. sibiricus* sample #713 have 97.2% and 95.3% nt sequence identity with EuYMV (FJ619507 and FJ619508), respectively (Table 1; Suppl. Figure S1). EuYMV has been found in many non-cultivated hosts (Fernandes *et al.*, 2011; Rocha *et al.*, 2013; Tavares *et al.*, 2012; Silva *et al.*, 2012; Barreto *et al.*, 2013), but to our knowledge, this is the first report of EuYMV infecting plants of the genus *Leonurus*.

Alphasatellite DNA molecules were cloned from samples #18 and #1095 (*Sida* spp. and *L. sibiricus*, respectively) (Table 1). Isolate BR:Cha18:09 (1338 nt) showed the highest nucleotide sequence identity (93.2%) to *Euphorbia* yellow mosaic alphasatellite (EuYMA, FN436008) (Suppl. Figure S4) and a close phylogenetic relationship with this isolate (Figure 3).

From sample #1095, five alphasatellite clones (BR:Dou1095.1:11 to BR:Dou1095.5:11; all 1367 nt) as well as begomovirus DNA-A and DNA-B components were cloned (Table 1). Pairwise sequence comparisons indicated that the DNA-A and DNA-B nucleotide sequences showed the highest identities with ToYSV (94.6% and 93.0%, respectively) (Suppl. Figure S1). Alphasatellite sequences shared >99.9% nucleotide identity amongst themselves and 82.3% to 82.4% identity with the sequence of EuYMA (Suppl. Figure S4). The sequences showed typical features of alphasatellite molecules, containing one ORF (*alpha-Rep*) potentially encoding a Rep protein with 313 amino acids (*data not shown*). The deduced amino acid sequences of the ORF display 86.9% identity with the *Cleome* leaf crumple alphasatellite (ClLCrA) alpha-Rep protein. The sequences also contain an A-rich region located immediately downstream of the ORF (coordinates 1115-1222, with a 57% adenine content) and a predicted hairpin structure containing, within the loop, the nonanucleotide TAGTATTAC, which is conserved in alphasatellites (Zhou, 2013).

Phylogenetic analysis showed that the isolates grouped most closely with EuYMA (Figure 3), consistent with the pairwise identity analysis. According to the proposed demarcation threshold of 83% nucleotide sequence identity for alphasatellites (Mubin *et al.*, 2009; Mubin *et al.*, 2012), this indicates that the five clones from *L. sibiricus* sample #1095 represent a distinct alphasatellite, for which the name Leonurus yellow spot alphasatellite (LeYSA) is proposed.

Originally thought to be restricted to the OW, alphasatellites have recently been found in association with bipartite begomoviruses in Brazil (state of Mato Grosso do Sul), Cuba and Venezuela (Paprotka *et al.*, 2010; Jeske *et al.*, 2014; Romay *et al.*, 2010). Our results extend the geographical range of alphasatellites in

South America, suggesting that these molecules may be widespread in the continent, and also their host range, with the first report of EuYMA infecting plants of the genus *Sida* and the detection of a new alphasatellite, LeYSA, in *L. sibiricus*.

Sida spp. are arguably the most abundant natural reservoirs for begomoviruses in several regions of the world (Jovel *et al.*, 2004; Frischmuth *et al.*, 1997; Fiallo-Olivé *et al.*, 2010; Hofer *et al.*, 1997; Roye *et al.*, 1997; Echemendía *et al.*, 2004; Fiallo-Olivé *et al.*, 2012; Xiong *et al.*, 2005; Guo and Zhou, 2006; Das *et al.*, 2008). For example, in a recent survey, Tavares *et al.* (2012) reported the occurrence of nine begomoviruses (including four new species) in 57 *Sida* spp. samples collected in the Brazilian states of Minas Gerais and Alagoas. Here, out of 43 *Sida* spp. samples, we found two previously described begomoviruses, three new species, and also (for the first time) an alphasatellite, further emphasizing the tremendous diversity of begomoviruses and associated DNA satellites naturally infecting *Sida* spp.

Conversely, *L. sibiricus* harbors a much lower diversity of begomoviruses, with ToYSV as the causative agent of infection in almost all samples, corroborating with published data (Fernandes *et al.*, 2014). ToYSV was first reported infecting tomato plants in the state of Minas Gerais (Ambrozevicius *et al.*, 2002), and was later reported in bean and soybean plants in northwestern Argentina (Rodríguez-Pardina *et al.*, 2011). *L. sibiricus* is a widely distributed plant in Brazil, and seems to be the main natural reservoir, as well as a potential source of inoculum, of ToYSV to bean, soybean and tomato crops, as previously noted also by Barbosa *et al.* (2012).

Acknowledgements

This work was funded by grants from CNPq (483607/2013-4) and Fapemig (CAG-APQ-02037-13) to FMZ. CGF was the recipient of a CNPq doctoral fellowship. The authors wish to thank Elvira Fiallo-Olivé, Eduardo S.G. Mizubuti and Jesus Navas-Castillo for helpful discussions.

References

Alabi O. J., Ogbe F. O., Bandyopadhyay R., Kumar P. L., Dixon A. G. O., Hughes J. D., Naidu R. A. (2008) Alternate hosts of African cassava mosaic virus and East African cassava mosaic Cameroon virus in Nigeria. *Archives of Virology,* **153**, 1743-1747.

Albuquerque L. C., Varsani A., Fernandes F. R., Pinheiro B., Martin D. P., Ferreira P. T. O., Lemos T. O., Inoue-Nagata A. K. (2012) Further characterization of tomato-infecting begomoviruses in Brazil. *Archives of Virology,* **157**, 747-752.

Amarakoon I. I., Roye M. E., Briddon R. W., Bedford I. D., Stanley J. (2008) Molecular and biological characterization of *Macroptilium yellow mosaic virus* from Jamaica. *Plant Pathology,* **57**, 417-426.

Ambrozevicius L. P., Calegario R. F., Fontes E. P. B., Carvalho M. G., Zerbini F. M. (2002) Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. *Fitopatologia Brasileira,* **27**, 372-377.

Arnaud L. S. E. P., Santos C. D. G., Lima J. A. A., Feitosa F. A. A. (2007) Predominance of begomoviruses in tomatoes in the region of Serra de Ibiapaba, Ceará state, and their detection in weeds (in Portuguese). *Fitopatologia Brasileira,* **32**, 241-246.

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. (2006) Genetic diversity of begomoviruses in invasive species of Northeastern Brazil (in Portuguese). *Planta Daninha,* **24**, 239-244.

Barbosa J. C., Barreto S. S., Inoue-Nagata A. K., Reis M. S., Firmino A. C., Bergamin-Filho A., Rezende J. A. M. (2009) Natural infection of *Nicandra physaloides* by *Tomato severe rugose virus* in Brazil. *Journal of General Plant Pathology,* **75**, 440-443.

Barbosa J. C., Eckstein B., Bergamin-Filho A., Rezende J. A. M., Dallagnol L. J. (2012) First report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Brazil. *Plant Disease,* **97**, 289-289.

Barreto S. S., Hallwass M., Aquino O. M., Inoue-Nagata A. K. (2013) A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. *Phytopathology,* **103**, 436-444.

Basso M. F., Silva J. C., Fajardo T. V., Fontes E. P., Zerbini F. M. (2015) A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. *Virus Research,* **210**, 27-33.

Berrie L. C., Rybicki E. P., Rey M. E. C. (2001) Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. *Journal of General Virology,* **82**, 53-58.

Briddon R. W., Akbar F., Iqbal Z., Amrao L., Amin I., Saeed M., Mansoor M. S. (2014) Effects of genetic changes to the begomovirus/betasatellite complex causing cotton leaf curl disease in South Asia post-resistance breaking. *Virus Research,* **186**, 114-119.

Brown J. K., Fauquet C. M., Briddon R. W., Zerbini F. M., Moriones E., Navas-Castillo J. (2012) Family *Geminiviridae*. *In: Virus Taxonomy. Ninth Report of the International*

Committee on Taxonomy of Viruses, pp. 351-373 Eds A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London, UK: Elsevier Academic Press.

Brown J. K., Zerbini F. M., Navas-Castillo J., Moriones E., Ramos-Sobrinho R., Silva J. C., Fiallo-Olivé E., Briddon R. W., Hernandez-Zepeda C., Idris A., Malathi V. G., Martin D. P., Rivera-Bustamante R., Ueda S., Varsani A. (2015) Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Archives of Virology*, **160**, 1593-1619.

Brustolini O. J., Machado J. P., Condori-Apfata J. A., Coco D., Deguchi M., Loriato V. A., Pereira W. A., Alfenas-Zerbini P., Zerbini F. M., Inoue-Nagata A. K., Santos A. A., Chory J., Silva F. F., Fontes E. P. (2015) Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. *Plant Biotechnology Journal*, **13**, 1300-1311.

Calegario R. F., Ferreira S. S., Andrade E. C., Zerbini F. M. (2007) Characterization of *Tomato yellow spot virus*, (ToYSV), a novel tomato-infecting begomovirus from Brazil. *Pesquisa Agropecuaria Brasileira,* **42**, 1335-1343.

Castillo-Urquiza G. P., Beserra Jr. J. E. A., Bruckner F. P., Lima A. T. M., Varsani A., Alfenas-Zerbini P., Zerbini F. M. (2008) Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. *Archives of Virology*, **153**, 1985-1989.

Castillo-Urquiza G. P., Beserra Junior J. E. A., Alfenas-Zerbini P., Varsani A., Lima A. T. M., Barros D. R., Zerbini F. M. (2007) Genetic diversity of begomoviruses infecting tomato in Paty do Alferes, Rio de Janeiro state, Brazil. *Virus Reviews and Research,* **12**, 233.

Costa A. S. (1955) Studies on *Abutilon* mosaic in Brazil. *Phytopathologische Zeitschrift,* **24**, 97-112.

Costa A. S., Bennett C. W. (1950) Whitefly transmitted mosaic of *Euphorbia prunifolia*. *Phytopathology*, **40**, 266-283.

Cotrim M. A., Krause-Sakate R., Narita N., Zerbini F. M., Pavan M. A. (2007) Genetic diversity ot tomato-infecting begomoviruses in Central São Paulo state (in Portuguese). *Summa Phytopathologica,* **33**, 300-303.

Das S., Paul S., Ghosh R., Palit P., Acharyya S., Das A., Mir J. I., Ghosh S. K., Roy A. (2008) Occurrence of a begomovirus and satellite DNA associated with yellow vein disease of *Sida* sp. from eastern India and its computational analyses. *Indian Journal of Virology*, **19**, 119-119.

Doyle J. J., Doyle J. L. (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochemical Bulletin,* **19**, 11-15.

Echemendía A. L., Ramos P. L., Díaz L., Peral R., Fuentes A., Pujol M., González G. (2004) First report of *Sida golden yellow vein virus* infecting *Sida* species in Cuba. *Plant Pathology,* **53**, 234.

Edgar R. C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics,* **5**, 1-19.

Faria J. C., Bezerra I. C., Zerbini F. M., Ribeiro S. G., Lima M. F. (2000) Current status of geminivirus diseases in Brazil (in Portuguese). *Fitopatologia Brasileira,* **25**, 125-137.

Faria J. C., Maxwell D. P. (1999) Variability in geminivirus isolates associated with *Phaseolus* spp. in Brazil. *Phytopathology,* **89**, 262-268.

Fernandes A. V., Galvão R. M., Machado J. J., Zerbini F. M., Fontes E. P. B. (1999) Cloning and molecular characterization of A components of two new *Sida rhombifolia*-infecting geminiviruses. *Virus Reviews and Research,* **4**, 148.

Fernandes F. R., Albuquerque L. C., Giordano L. B., Boiteux L. S., Ávila A. C., Inoue-Nagata A. K. (2008) Diversity and prevalence of Brazilian bipartite begomovirus species associated to tomatoes. *Virus Genes,* **36**, 251-258.

Fernandes F. R., Albuquerque L. C., Oliveira C. L., Cruz A. R. R., Rocha W. B., Pereira T. G., Naito F. Y. B., Dias N. D., Nagata T., Faria J. C., Zerbini F. M., Aragão F. J. L., Inoue-Nagata A. K. (2011) Molecular and biological characterization of a new

Brazilian begomovirus, euphorbia yellow mosaic virus (EuYMV), infecting *Euphorbia heterophylla* plants. *Archives of Virology*, **156**, 2063-2069.

Fernandes J. J., Carvalho M. G., Andrade E. C., Brommonschenkel S. H., Fontes E. P. B., Zerbini F. M. (2006) Biological and molecular properties of *Tomato rugose mosaic virus* (ToRMV), a new tomato-infecting begomovirus from Brazil. *Plant Pathology*, **55**, 513-522.

Fernandes N. A. N., Boiteux L. S., Fonseca M. E. N., Gonzales-Segnana L., Kitajima E. W. (2014) Report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Paraguay and within tomato fields in Brazil. *Plant Disease*, **98**, 1445.

Fiallo-Olivé E., Martinez-Zubiaur Y., Moriones E., Navas-Castillo J. (2010) Complete nucleotide sequence of *Sida golden mosaic Florida virus* and phylogenetic relationships with other begomoviruses infecting malvaceous weeds in the Caribbean. *Archives of Virology*, **155**, 1535-1537.

Fiallo-Olivé E., Navas-Castillo J., Moriones E., Martinez-Zubiaur Y. (2012) Begomoviruses infecting weeds in Cuba: Increased host range and a novel virus infecting *Sida rhombifolia*. *Archives of Virology*, **157**, 141-146.

Frischmuth T., Engel M., Lauster S., Jeske H. (1997) Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted, *Sida*-infecting bipartite geminiviruses in Central America. *Journal of General Virology*, **78**, 2675-2682.

García-Andrés S., Accotto G. P., Navas-Castillo J., Moriones E. (2007a) 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*, **359**, 302-312.

García-Andrés S., Monci F., Navas-Castillo J., Moriones E. (2006) Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: Evidence for the presence of a new virus species of recombinant nature. *Virology*, **350**, 433-442.

García-Andrés S., Tomas D. M., Sanchez-Campos S., Navas-Castillo J., Moriones E. (2007b) Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. *Virology*, **365**, 210-219.

Glick E., Zrachya A., Levy Y., Mett A., Gidoni D., Belausov E., Citovsky V., Gafni Y. (2008) Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 157-161.

Guo X. J., Zhou X. P. (2006) Molecular characterization of a new begomovirus infecting *Sida cordifolia* and its associated satellite DNA molecules. *Virus Genes*, **33**, 279-285.

Hanley-Bowdoin L., Bejarano E. R., Robertson D., Mansoor S. (2013) Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology*, **11**, 777-788.

Harrison B. D., Robinson D. J. (1999) Natural genomic and antigenic variation in white-fly transmitted geminiviruses (begomoviruses). *Annual Review of Phytopathology*, **39**, 369-398.

Hofer P., Engel M., Jeske H., Frischmuth T. (1997) Nucleotide sequence of a new bipartite geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica. *Virology*, **78**, 1785-1790.

Idris A. M., Brown J. K. (2002) Molecular analysis of Cotton leaf curl virus-Sudan reveals an evolutionary history of recombination. *Virus Genes*, **24**, 249-256.

Idris A. M., Hiebert E., Bird J., Brown J. K. (2003) Two newly described begomoviruses of *Macroptilium lathyroides* and common bean. *Phytopathology*, **93**, 774-783.

Inoue-Nagata A. K., Albuquerque L. C., Rocha W. B., Nagata T. (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi 29 DNA polymerase. *Journal of Virological Methods*, **116**, 209-211.

Jeske H., Kober S., Schäfer B., Strohmeier S. (2014) Circomics of Cuban geminiviruses reveals the first alpha-satellite DNA in the Caribbean. *Virus Genes*, **49**, 312-324.

Jovel J., Reski G., Rothenstein D., Ringel M., Frischmuth T., Jeske H. (2004) *Sida micrantha* mosaic is associated with a complex infection of begomoviruses different from *Abutilon mosaic virus*. *Archives of Virology,* **149**, 829-841.

Krenz B., Thompson J. R., Fuchs M., Perry K. L. (2012) Complete genome sequence of a new circular DNA virus from grapevine. *Journal of Virology,* **86**, 7715.

Lefeuvre P., Lett J. M., Varsani A., Martin D. P. (2009) Widely conserved recombination patterns among single-stranded DNA viruses. *Journal of Virology,* **83**, 2697-2707.

Liang P., Navarro B., Zhang Z., Wang H., Lu M., Xiao H., Wu Q., Zhou X., Di Serio F., Li S. (2015) Identification and characterization of a novel geminivirus with monopartite genome infecting apple trees. *Journal of General Virology,* **96**, 2411-2420.

Loconsole G., Saldarelli P., Doddapaneni H., Savino V., Martelli G. P., Saponari M. (2012) Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family *Geminiviridae*. *Virology,* **432**, 162-172.

Mahajan N., Parameswari C., Veluthambi K. (2011) Severe stunting in blackgram caused by the Mungbean yellow mosaic virus (MYMV) KA27 DNA B component is ameliorated by co-infection or post-infection with the KA22 DNA B: MYMV nuclear shuttle protein is the symptom determinant. *Virus Research,* **157**, 25-34.

Mansoor S., Briddon R. W., Zafar Y., Stanley J. (2003) Geminivirus disease complexes: An emerging threat. *Trends in Plant Science,* **8**, 128-134.

Marquez-Martin B., Aragon-Caballero L., Fiallo-Olivé E., Navas-Castillo J., Moriones E. (2011) Tomato leaf deformation virus, a novel begomovirus associated with a severe disease of tomato in Peru. *European Journal of Plant Pathology,* **129**, 1-7.

Martin D. P., Lemey P., Lott M., Moulton V., Posada D., Lefeuvre P. (2010) RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics,* **26**, 2462-2463.

Medina C. G. V., Lambertini P. M. L. (2012) Tomato dwarf leaf virus, a New World begomovirus infecting tomato in Argentina. *Archives of Virology,* **157**, 1975-1980.

Melgarejo T. A., Kon T., Rojas M. R., Paz-Carrasco L., Zerbini F. M., Gilbertson R. L. (2013) Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. *Journal of Virology,* **87**, 5397-5413.

Monci F., Sanchez-Campos S., Navas-Castillo J., Moriones E. (2002) 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,* **303**, 317-326.

Monde G., Walangululu J., Winter S., Bragard C. (2010) Dual infection by cassava begomoviruses in two leguminous species (Fabaceae) in Yangambi, Northeastern Democratic Republic of Congo. *Archives of Virology,* **155**, 1865-1869.

Mubin M., Akhtar S., Amin I., Briddon R. W., Mansoor S. (2012) Xanthium strumarium: a weed host of components of begomovirus-betasatellite complexes affecting crops. *Virus Genes,* **44**, 112-119.

Mubin M., Briddon R. W., Mansoor S. (2009) Diverse and recombinant DNA betasatellites are associated with a begomovirus disease complex of *Digera arvensis*, a weed host. *Virus Research,* **142**, 208-212.

Muhire B. M., Varsani A., Martin D. P. (2014) SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One,* **9**, e108277.

Nylander J. A. A. (2004) MrModeltest v2. *In: Program distributed by the author. Evolutionary Biology Centre, Uppsala University*.

Padidam M., Beachy R. N., Fauquet C. M. (1996) The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. *Virology,* **224**, 390-404.

Padidam M., Sawyer S., Fauquet C. M. (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology,* **265**, 218-224.

Paprotka T., Metzler V., Jeske H. (2010) The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. *Virology,* **404**, 148-157.

Paximadis M., Idris A. M., Torres-Jerez I., Villarreal A., Rey M. E. C., Brown J. K. (1999) Characterization of tobacco geminiviruses in the Old and New world. *Archives of Virology*, **144**, 703-717.

Pita J. S., Fondong V. N., Sangare A., Otim-Nape G. W., Ogwal S., Fauquet C. M. (2001) Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology*, **82**, 655-665.

Ramos-Sobrinho R., Xavier C. A. D., Pereira H. M. B., Lima G. S. A., Assunção I. P., Mizubuti E. S. G., Duffy S., Zerbini F. M. (2014) Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. *Journal of General Virology*, **95**, 2540-2552.

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. (2003) Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. *Archives of Virology*, **148**, 281-295.

Ribeiro S. G., Martin D. P., Lacorte C., Simões I. C., Orlandini D. R. S., Inoue-Nagata A. K. (2007) Molecular and biological characterization of *Tomato chlorotic mottle virus* suggests that recombination underlies the evolution and diversity of Brazilian tomato begomoviruses. *Phytopathology*, **97**, 702-711.

Rocha C. S., Castillo-Urquiza G. P., Lima A. T. M., Silva F. N., Xavier C. A. D., Hora-Junior B. T., Beserra-Junior J. E. A., Malta A. W. O., Martin D. P., Varsani A., Alfenas-Zerbini P., Mizubuti E. S. G., Zerbini F. M. (2013) Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. *Journal of Virology*, **87**, 5784-5799.

Rodríguez-Pardina P. E., Hanada K., Laguna I. G., Zerbini F. M., Ducasse D. A. (2011) Molecular characterisation and relative incidence of bean- and soybean-infecting begomoviruses in northwestern Argentina. *Annals of Applied Biology*, **158**, 69-78.

Rojas M. R., Hagen C., Lucas W. J., Gilbertson R. L. (2005) Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. *Annual Review of Phytopathology*, **43**, 361-394.

Romay G., Chirinos D., Geraud-Pouey F., Desbiez C. (2010) Association of an atypical alphasatellite with a bipartite New World begomovirus. *Archives of Virology*, **155**, 1843-1847.

Ronquist F., Huelsenbeck J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Roye M. E., McLaughlin W. A., Nakhla M. K., Maxwell D. P. (1997) Genetic diversity among geminiviruses associated with the weed species *Sida* spp., *Macroptilium lathyroides*, and *Wissadula amplissima* from Jamaica. *Plant Disease*, **81**, 1251-1258.

Rybicki E. P. (1994) A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Archives of Virology*, **139**, 49-77.

Silva A. K. F., Santos C. D. G., Nascimento A. K. Q. (2010) Begomovirus transmission from weeds to tomato by the whitefly. *Planta Daninha*, **28**, 507-514.

Silva S. J. C., Castillo-Urquiza G. P., Hora-Junior B. T., Assunção I. P., Lima G. S. A., Pio-Ribeiro G., Mizubuti E. S. G., Zerbini F. M. (2012) Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. *Plant Pathology*, **61**, 457-467.

Silva S. J. C., Castillo-Urquiza G. P., Hora-Júnior B. T., Assunção I. P., Lima G. S. A., Pio-Ribeiro G., Mizubuti E. S. G., Zerbini F. M. (2011) High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed *Cleome affinis* in northeastern Brazil. *Archives of Virology*, **156**, 2205-2213.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729.

Tavares S. S., Ramos-Sobrinho R., Gonzalez-Aguilera J., Lima G. S. A., Assunção I. P., Zerbini F. M. (2012) Further molecular characterization of weed-associated

begomoviruses in Brazil with an emphasis on *Sida* spp. *Planta Daninha*, **30**, 305-315.

Varsani A., Navas-Castillo J., Moriones E., Hernández-Zepeda C., Idris A., Brown J. K., Zerbini F. M., Martin D. P. (2014) Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. *Archives of Virology*, **159**, 2193-2203.

Xavier C. A. D. (2015) *Species diversity and genetic variability of bipartite begomoviruses in the New World.* M.Sc. Dissertation, Universidade Federal de Viçosa.

Xiong Q., Guo X. J., Che H. Y., Zhou X. P. (2005) Molecular characterization of a distinct Begomovirus and its associated satellite DNA molecule infecting Sida acuta in China. *Journal of Phytopathology*, **153**, 264-268.

Zerbini F. M., Andrade E. C., Barros D. R., Ferreira S. S., Lima A. T. M., Alfenas P. F., Mello R. N. (2005) Traditional and novel strategies for geminivirus management in Brazil. *Australasian Plant Pathology*, **34**, 475-480.

Zhou X. (2013) Advances in understanding begomovirus satellites. *Annual Review of Phytopathology*, **51**, 357-381.

Zhou X., Liu Y., Calvert L., Munoz C., Otim-Nape G. W., Robinson D. J., Harrison B. D. (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, **78**, 2101-2111.

Zhou X., Liu Y., Robinson D. J., Harrison B. D. (1998) 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*, **79**, 915-923.

Table 1. Begomovirus and alphasatellite sequences reported in this study.

Sample code	Sampling date	Location	Geographical coordinates		Host	Enzyme[1]			Isolate name	GenBank access number
						DNA-A	DNA-B	Satellite		
Sida micrantha mosaic virus (SiMMV)										
CF48	03/19/09	Santo Angelo, RS	S28 22' 54,70"	W54 18' 17,24"	*Sida* sp.	*Eco*RI			BR:Saa48:09	KX348158
CF65	03/20/09	São Miguel das Missões, RS	S28 29' 35,59"	W54 33' 37,15"	*Sida* sp.	*Spe*I			BR:Smm65:09	KX348159
CF69	03/20/09	Santo Antonio das Missões, RS	S28 29' 42,70"	W55 25' 18,00"	*Sida* sp.		*Spe*I		BR:Sam69:09	KX348195
CF115	03/20/09	Panambi, RS	S28 18' 24,45"	W53 29' 21,38"	*Sida* sp.	*Eco*RI	*Spe*I		BR:Pan115:09	KX348160 (A) KX348192 (B)
CF547	03/24/10	São Miguel das Missões, RS	S28 29' 35,59"	W54 33' 37,15"	*Sida* sp.		*Spe*I		BR:Smm547:10	KX348196
CF556	03/25/10	Cruz Alta, RS	S28 36' 10,89"	W53 39' 25,34"	*Sida* sp.		*Sal*I		BR:Cra556:10	KX348194
CF662	03/25/10	Realeza, PR	S25 40' 45,00"	W53 33' 09,00"	*Sida* sp.	*Eco*RI	*Spe*I		BR:Rea662:10	KX348156 (A) KX348186 (B)
CF679	06/08/10	Marechal Candido Rondom, PR	S24 30' 59,50"	W54 04' 37,20"	*Sida* sp.		*Apa*I		BR:Mcr679:10	KX348197
CF698	06/09/10	Dourados, MS	S22 17' 58,00"	W54 49' 14,10"	*Sida* sp.		*Hin*dIII		BR:Dou698:10	KX348191
CF704	06/09/10	Laguna Carapã, MS	S22 25' 35,60"	W55 21' 30,00"	*Sida* sp.		*Sal*I		BR:Lac704:10	KX348193
CF732	06/10/10	Umuarama, PR	S23 50' 12,30"	W53 17' 23,60"	*Sida* sp.	*Eco*RI			BR:Umu732:10	KX348164
CF755	06/10/10	Marialva,PR	S23 30' 05,00"	W51 47' 08,00"	*Sida* sp.		*Sal*I		BR:Mar755:10	KX348198
CF799	08/11/10	São Domingos, PR	S24 00' 51,00"	W51 30' 37,00"	*Sida* sp.		*Apa*I		BR:Sad799:10	KX348190
CF822	08/24/10	Chapada, RS	S28 01' 08,80"	W53 05' 52,31"	*Sida* sp.	*Spe*I	*Spe*I		BR:Cha822:10	KX348161 (A) KX348199 (B)
CF832	08/25/10	Santo Angelo, RS	S28 22' 54,70"	W54 18' 17,23"	*Sida* sp.		*Spe*I		BR:Saa832:10	KX348202
CF876	10/06/10	São Miguel das Missões, RS	S28 23' 59,00"	W54 39' 52,00"	*Sida* sp.	*Sac*I	*Spe*I		BR:Smm876:10	KX348155 (A) KX348200 (B)
CF895	10/07/10	Tapatuba, RS	S29 03' 57,00"	W54 43' 53,00"	*Sida* sp.		*Sal*I		BR:Tap895:10	KX348189
CF926	03/15/11	Chapada, RS	S28 01' 08,80"	W53 05' 52,31"	*Sida* sp.	*Spe*I	*Spe*I		BR:Cha926:11	KX348162 (A) KX348201 (B)
CF949	03/16/11	São Miguel Das Missões, RS	S28 29' 35,59"	W54 33' 37,15"	*Sida* sp.		*Sal*I		BR:Smm949:11	KX348203
CF1116	06/08/11	Tacuru, RS	S23 38' 17,60"	W54 58' 17,70"	*Sida* sp.		*Apa*I		BR:Tac1116:11	KX348188

CF1120	06/09/11	Guaíra, PR	S24 04' 57,00"	W54 10' 04,00"	*Sida* sp.		*Apa*I	BR:Gua1120:11	KX348204
CF1121	06/09/11	Umuarama, PR	S23 50' 12,30"	W53 17' 23,60"	*Sida* sp.	*Spe*I		BR:Umu1121:11	KX348157
CF1128	06/09/11	Janiópolis, PR	S24 08' 02,00"	W52 47' 22,00"	*Sida* sp.		*Bam*HI	BR:Jan1128:11	KX348187
CF1142	06/09/11	Marialva, PR	S23 30' 09,40"	W51 47' 42,90"	*Sida* sp.	*Eco*RI		BR:Mar1142:11	KX348163
Euphorbia yellow mosaic virus (EuYMV)									
CF531	03/24/10	Três de Maio, RS	S27 45' 39,44"	W54 15' 42,87"	*Sida* sp.	*Hin*dIII		BR:Trm531.2:10	KX348180
CF691	06/08/11	Caarapo, MS	S22 26' 42,50"	W54 49' 49,50"	*Sida* sp.	*Sal*I		BR:Caa691:10	KX348181
CF713	06/09/10	Aral Moreira, MS	S22 46' 19,70"	W55 24' 43,90"	*Leonurus sibiricus*	*Apa*I	*Kpn*I	BR:Arm713:10	KX348182 (A) KX348224 (B)
Sida chlorotic mottle virus (SiCMoV)									
CF531	03/24/10	Três de Maio, RS	S27 45' 39,44"	W54 15' 42,87"	*Sida* sp.	*Apa*I		BR:Trm531.1:10	KX348183
Sida bright yellow mosaic virus (SiBYMV)									
CF720	06/09/10	Tacuru, MS	S23 38' 17,60"	W54 58' 17,70"	*Sida* sp.	*Spe*I	*Cla*I	BR:Tac720:10	KX348184 (A) KX348225 (B)
Sida golden yellow spot virus (SiGYSV)									
CF889	06/10/10	São Borja, RS	S28 57' 30,00"	W55 32' 30,00"	*Sida* sp.	*Apa*I		BR:Sab889:10	KX348185
Tomato yellow spot virus (ToYSV)									
CF724	06/10/10	Guaíra, PR	S24 08' 35,30"	W54 14' 53,20"	*Sida* sp.		*Apa*I	BR:Gua724:10	KX348226
CF661	06/07/10	Realeza, PR	S25 40' 45,00"	W53 33' 09,00"	*L. sibiricus*	*Spe*I	*Kpn*I	BR:Rea661:10	KX348169 (A) KX348214 (B)
CF673	06/08/10	Toledo, PR	S24 46' 45,00"	W53 40' 41,00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Tol673:10	KX348174 (A) KX348223 (B)
CF677	06/08/10	Marechal Candido Rondom, PR	S24 30' 59,50"	W54 04' 37,20"	*L. sibiricus*	*Apa*I		BR:Mcr677:10	KX348170
CF697	06/09/10	Dourados, MS	S22 17' 58,00"	W54 49' 14,10"	*L. sibiricus*	*Apa*I	*Hin*dIII	BR:Dou697:10	KX348171 (A) KX348205 (B)
CF703	06/09/10	Laguna Carapã, MS	S22 25' 35,60"	W55 21' 30,00"	*L. sibiricus*	*Apa*I		BR:Lac703:10	KX348177
CF735	06/10/10	Janiópolis, PR	S24 08' 02,00"	W52 47' 22,00"	*L. sibiricus*		*Apa*I	BR:Jan735:10	KX348206
CF739	06/10/10	Engenheiro Beltrão, PR	S23 42' 12,60"	W52 08' 54,60"	*L. sibiricus*	*Sac*I	*Hin*dIII	BR:Egb739:10	KX348168 (A) KX348207 (B)
CF775	08/10/10	Cascavel, PR	S24 52' 06,00"	W53 20' 28,00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Cas775:10	KX348172 (A) KX348222 (B)

CF776	08/10/10	Cafelândia, PR	S24 39' 21,30"	W53 13' 07,40"	*L. sibiricus*	*Apa*I	*Apa*I		BR:Caf776:10	KX348175 (A) KX348221 (B)
CF779	08/10/10	Ubiratã, PR	S24 33' 07,00"	W53 00' 34,00"	*L. sibiricus*		*Hin*dIII		BR:Ubi779:10	KX348208
CF784	08/10/10	Campo Mourão, PR	S23 57' 14,90"	W52 20' 59,30"	*L. sibiricus*	*Sac*I	*Kpn*I		BR:Cam784:10	KX348167 (A) KX348209 (B)
CF793	08/11/10	Londrina, PR	S23 26' 20,30"	W51 08' 17,90"	*L. sibiricus*		*Apa*I		BR:Lon793:10	KX348207
CF796	08/11/10	São Domingos, PR	S24 00' 51,00"	W51 30' 37,00"	*L. sibiricus*		*Kpn*I		BR:Sad796:10	KX348219
CF802	08/11/10	Ivaiporã, PR	S24 18' 37,00"	W51 43' 24,00"	*L. sibiricus*		*Apa*I		BR:Iva802:10	KX348218
CF1024	04/26/11	Francisco Beltrão, PR	S26 00' 09,00"	W52 56' 36,00"	*L. sibiricus*		*Apa*I		BR:Frb1024:11	KX348210
CF1058	06/06/11	Pato Branco, PR	S26 11' 47,44"	W52 49' 23,87"	*L. sibiricus*	*Apa*I			BR:Pab1058:11	KX348179
CF1067	06/06/11	Ampere, PR	S25 57' 06,90"	W53 24' 25,00"	*L. sibiricus*	*Apa*I			BR:Amp1067:11	KX348173
CF1075	06/07/11	Toledo, PR	S24 46' 46,00"	W53 40' 41,00"	*L. sibiricus*		*Apa*I		BR:Tol1075:11	KX348211
CF1077	06/07/11	Nova Mercedes, PR	S24 30' 54,70"	W54 07' 0,22"	*L. sibiricus*		*Apa*I		BR:Nom1077:11	KX348217
CF1083	06/07/11	Guaíra, PR	S24 14' 04,00"	W54 11' 57,00"	*L. sibiricus*	*Spe*I	*Apa*I		BR:Gua1083:11	KX348166 (A) KX348216 (B)
CF1095	06/08/11	Dourados, MS	S22 17' 58,00"	W54 49' 14,10"	*L. sibiricus*	*Apa*I	*Apa*I		BR:Dou1095:11	KX348165 (A) KX348212 (B)
CF1111	06/08/11	Aral Moreira, MS	S22 46' 19,70"	W55 24' 43,90"	*L. sibiricus*		*Hin*dIII		BR:Arm1111:11	KX348213
CF1130	06/09/11	Araruna, PR	S24 03' 50,00"	W52 33' 52,00"	*L. sibiricus*	*Apa*I			BR:Ara1130:11	KX348176
CF1135	06/09/11	Sertanópolis, PR	S23 50' 25,00"	W52 18' 19,00"	*L. sibiricus*	*Apa*I	*Apa*I		BR:Ser1135:11	KX348178 (A) KX348215 (B)
Euphorbia yellow mosaic alphasatellite										
CF18	03/19/09	Chapada, RS	S28 01' 08,80"	W53 05' 52,31'	*Sida* sp.			*Eco*RI	BR:Cha18:09	KX348227
Leonurus yellow spot alphasatellite										
CF1095	06/08/11	Dourados, MS	S22 17' 58,00"	W54 49' 14,10"	*L. sibiricus*			*Eco*RI	BR:Dou1095.1:11	KX348228
									BR:Dou1095.2:11	KX348229
									BR:Dou1095.3:11	KX348230
									BR:Dou1095.4:11	KX348231
									BR:Dou1095.5:11	KX348232

[1]Enzyme used for releasing genome-length DNA components after rolling-circle amplification and subsequent cloning into the plasmid vector pKS+.

Table 2. Recombination events detected in the Sida bright yellow mosaic virus (SiBYMV-BR:Tac720:10) and Sida golden yellow spot virus (SiGYSV -BR:Sab889:10) genomes, based on a data set including the nucleotide sequences of the begomoviruses described in this study and additional begomoviruses[1] from the New World (NW) and Old World (OW).

Event[1]	Recombinant	Recombination breakpoints[2]		Parents		Method[3]	*P-value*
		Begin	End	Minor	Major		
	SiGYSV (DNA-A)						
1	BR:Sab889:10	180	1178	Unknown	SiMBoV1 (HM585441)	**R**GBMCS3	5.066 X 10^{-38}
2	BR:Sab889:10	1448	1889	SiYLCV (KC706539)	Unknown	RMC**3**	2.524 X 10^{-7}
	SiBYMV (DNA-A)						
3	BR:Tac720:10	193	1906	Unknown	SiMMV (FN436003)	RGBMCS**3**	7.891 X 10^{-28}
	SiBYMV (DNA-B)						
4	BR:Tac720:10	2562	587	BR:Tol1075:11	BDMV (M88180)	**R**BMCS	1.725 X 10^{-6}

[1]See Table S1 in the supplemental material for the complete data set and full virus names.

[2]Recombination breakpoint coordinates are according to the first nucleotide after the cleavage site at the origin of replication, increasing clockwise.

[3]Recombination events and their putative parental viruses were identified using the the Rdp (R), Geneconv (G), Boostcan (B), Maxichi (M), Chimaera (C), Siscan (S) and 3Seq (3) modules in RDP4.

[4]The reported *P*-values are for the programs indicated in bold in the "Method" column and are the lowest *P*-values calculated for the region in question.

Figure legends

Figure 1. Bayesian phylogenetic trees based on the complete DNA-A **(A)** and DNA-B **(B)** nucleotide sequences of the begomoviruses obtained from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) in this study plus begomoviruses from the New World and Old World (see Suppl. Table S1 for full names and GenBank access numbers). BR:Vic25.4:11 and BR:Vic26:11 (blue) are unpublished sequences determined in our laboratory (Xavier, 2015). The OW begomoviruses *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *Tomato yellow leaf curl virus* (TYLCV) and *Watermelon chlorotic stunt virus* (WmCSV) were used as outgroups. Nodes with posterior probability values between 0.50 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles.

Figure 2. Bayesian phylogenetic tree based on the nucleotide sequences of the *cp* genes of Sida golden yellow spot virus (SiGYSV, in red). The unrooted *cp* tree includes four highly divergent geminiviruses (apple geminivirus, AGV; citrus chlorotic dwarf associated virus, CCDaV; *Euphorbia caput-medusae* latent virus, ECMLV; and grapevine red-blotch associated virus, GRBaV), one topocuvirus (*Tomato pseudo-curly top virus*, TPCTV), the BR:Vic25.4:11 and BR:Vic26:11 sequences (blue) determined in our laboratory, in addition to seven New World and Old World begomoviruses (see Suppl. Table S1 for full names and GenBank access numbers). Nodes with posterior probability values between 0.60 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles.

Figure 3. Bayesian phylogenetic tree based on the alphasatellites sequences described in this study from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) and the most closely related alphasatellites (see Suppl. Table S2 for full names and GenBank access numbers). Nodes with posterior probability values between 0.50 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles.

Supporting Information

Supplementary Table S1. Geminivirus sequences used for pairwise sequence comparisons, phylogenetic analysis and detection of recombination events.

Species	Acronym	Access number		Origin
		DNA-A	DNA-B	
Abutilon mosaic Brazil virus	AbMBV	JF694480	JF694483	NW
Abutilon mosaic Bolivia virus	AbMBoV	HM585445	HM585446	NW
Abutilon mosaic virus	AbMV	LN611623	LN611625	NW
Bean dwarf mosaic virus	BDMV	M88179	M88180	NW
Bean golden mosaic virus	BGMV	KJ939853		NW
Blainvillea yellow spot virus	BlYSV	KC706520	KC706525	NW
Cleome leaf crumple virus	CleLCrV	FN435999	FN436000	NW
Corchorus yellow spot virus	CoYSV	DQ875868	DQ875869	NW
Cotton leaf crumple virus	CLCrV	AY742220	AY742221	NW
Euphorbia mosaic virus	EuMV	JQ963887	JQ963888	NW
Euphorbia yellow mosaic virus	EuYMV	FJ619507	FJ619508	NW
Tomato yellow spot virus	ToYSV	DQ336350	DQ336351	NW
Okra mottle virus	OMoV	EU914817	EU914818	NW
Oxalis yellow vein virus	OxYVV	KM887907	n.a.[1]	NW
Passionfruit severe leaf distortion virus	PSLDV	FJ972767	FJ972768	NW
Pepper golden mosaic virus	PepGMV	AY928514	AY928515	NW
Cabbage leaf curl virus	CabLCV	DQ406672	DQ406673	NW
Sida golden mosaic virus	SiGMV	AF049336	AF039841	NW
Sida micrantha mosaic virus	SiMMV	FN436003	FN436004	NW
Sida mosaic Bolivia virus 1	SiMBoV	HM585441	HM585442	NW
Sida mottle virus	SiMoV	AJ557450	AJ557454	NW
Sida yellow leaf curl virus	SiYLCV	KC706539	n.a.	NW
Sida yellow mosaic virus	SiYMV	JX871369	n.a.	NW
Squash leaf curl virus	SLCuV	M38183	M38182	NW
Tomato chino La Paz virus	ToChLPV	AY339618	n.a.	NW
Tomato chlorotic mottle virus	ToCMoV	KC706557	KC706562	NW
Tomato common mosaic virus	ToCmMV	KC706570	KC706595	NW
Tomato dwarf leaf virus	ToDfLV	JN564749	JN564750	NW
Tomato golden mosaic virus	TGMV	JF694488	JF694489	NW
Tomato golden leaf distortion virus	ToGLDV	HM357456	n.a.	NW
Tomato leaf deformation virus	ToLDeV	GQ334472	m.	NW
Tomato mild mosaic virus	ToMlMV	KC706611	KC706613	NW
Tomato mottle leaf curl virus	ToMoLCV	JF803249	JF803264	NW
Tomato rugose mosaic virus	ToRMV	AF291705	AF291706	NW
Tomato severe rugose virus	ToSRV	KC706625	KC004086	NW
Tomato yellow spot virus	ToYSV	DQ336350	DQ336351	NW
Tomato mottle virus	ToMoV	AY965900	AY965901	NW

Tomato pseudo-curly top virus	TPCTV	X84735	m.	
African cassava mosaic virus	ACMV	FN668378	FN668379	OW
Ageratum yellow vein virus	AYVV	KJ744212	m.	OW
Apple geminivirus	AGV	NC026760	m.	OW
Citrus chlorotic dwarf associated virus	CCDaV	NC018151	m.	OW
Corchorus golden mosaic virus	CoGMV	FJ463902	FJ463901	OW
Corchorus yellow vein virus	CoYVV	AY727903	AY727904	OW
Euphorbia caput-medusae latent virus	EcmLV	HF921477	m.	OW
East African cassava mosaic virus	EACMV	HE979774	HE979778	OW
Grapevine red-blotch associated virus	GRBaV	Q901105	m.	OW
Tomato yellow leaf curl virus	TYLCV	JX128099	m.	OW
Watermelon chlorotic stunt virus	WmCSV	KJ939448	KJ939447	OW

[1] Not available.
[2] Monopartite geminivirus.

Supplementary Table S2. Alphasatellite sequences used for pairwise sequence comparisons, phylogenetic analysis and detection of recombination events.

Species	**Acronym**	**Access number**
Cleome leaf crumple alphasatellite	ClLCrA	FN436007
Croton yellow vein mosaic alphasatellite	CYVMA	FN658711
Cuban alphasatellite 1	CuA1	HE806451
Euphorbia yellow mosaic alphasatellite	EuYMA	FN436008
Melon chlorotic mosaic alphasatellite	MeCMA	HM163578
Papaya leaf curl alphasatellite	PLCA	HQ668024

Supplementary Figure S1. Pairwise sequence identity matrices of the DNA-A **(A)** and DNA-B **(B)**, between the begomoviruses obtained from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) in this study and additional begomovirus from the Americas. BR:Vic25.4:11 and BR:Vic26:11 (blue) are unpublished sequences determined in our laboratory (Xavier, 2015).

Supplementary Figure S2. Bayesian phylogenetic trees based on the nucleotide sequences of the the *cp* **(A)** and *rep* **(B)** genes of the begomoviruses obtained from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) in this study plus begomoviruses from the New World and Old World (see Suppl. Table S1 for full names and GenBank access numbers). BR:Vic25.4:11 and BR:Vic26:11 (blue) are unpublished sequences determined in our laboratory (Xavier, 2015). The OW begomoviruses *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *Tomato yellow leaf curl virus* (TYLCV) and *Watermelon chlorotic stunt virus* (WmCSV) were used as outgroups. Nodes with posterior probability values between 0.50 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles.

Suplementary Figure S3. Bayesian phylogenetic trees based on the nucleotide sequences of the *mp* **(A)** and *nsp* **(B)** genes of the begomoviruses obtained from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) in this study plus begomoviruses from the New World and Old World (see Suppl. Table S1 for full names and GenBank access numbers). The OW begomoviruses *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV) and *Watermelon chlorotic stunt virus* (WmCSV) were used as outgroups. Nodes with posterior probability

values between 0.50 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles.

Supplementary Figure S4. Pairwise sequence identity matrices between the alphasatellites sequences described in this study from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) and the most closely related alphasatellites (see Suppl. Table S2 for full names and GenBank access numbers).

Figure 1A

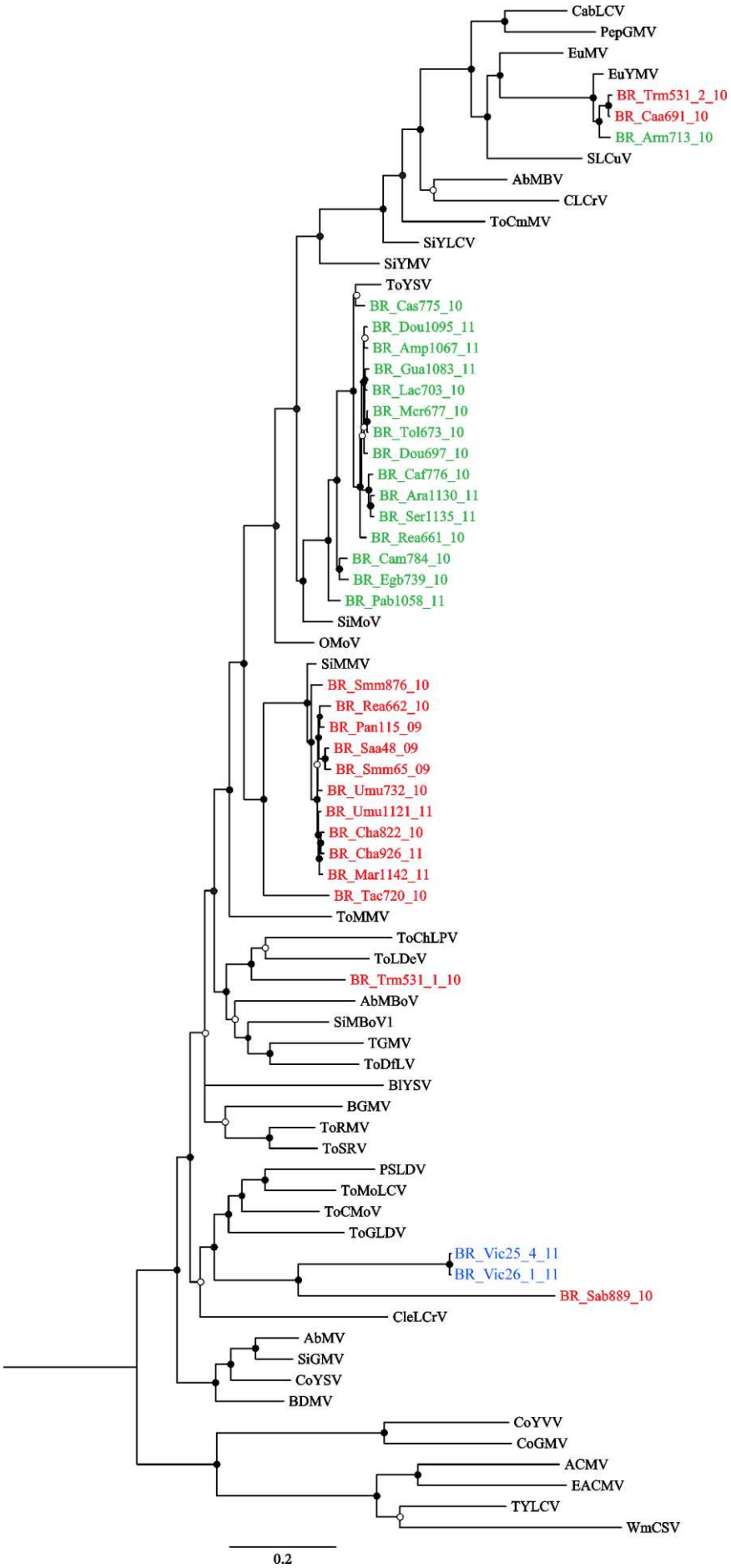

CabLCV
PepGMV
EuMV
EuYMV
BR_Trm531_2_10
BR_Caa691_10
BR_Arm713_10
SLCuV
AbMBV
CLCrV
ToCmMV
SiYLCV
SiYMV
ToYSV
BR_Cas775_10
BR_Dou1095_11
BR_Amp1067_11
BR_Gua1083_11
BR_Lac703_10
BR_Mcr677_10
BR_Tol673_10
BR_Dou697_10
BR_Caf776_10
BR_Ara1130_11
BR_Ser1135_11
BR_Rea661_10
BR_Cam784_10
BR_Egb739_10
BR_Pab1058_11
SiMoV
OMoV
SiMMV
BR_Smm876_10
BR_Rea662_10
BR_Pan115_09
BR_Saa48_09
BR_Smm65_09
BR_Umu732_10
BR_Umu1121_11
BR_Cha822_10
BR_Cha926_11
BR_Mar1142_11
BR_Tac720_10
ToMMV
ToChLPV
ToLDeV
BR_Trm531_1_10
AbMBoV
SiMBoV1
TGMV
ToDfLV
BlYSV
BGMV
ToRMV
ToSRV
PSLDV
ToMoLCV
ToCMoV
ToGLDV
BR_Vic25_4_11
BR_Vic26_1_11
BR_Sab889_10
CleLCrV
AbMV
SiGMV
CoYSV
BDMV
CoYVV
CoGMV
ACMV
EACMV
TYLCV
WmCSV
0.2

Figure 1B

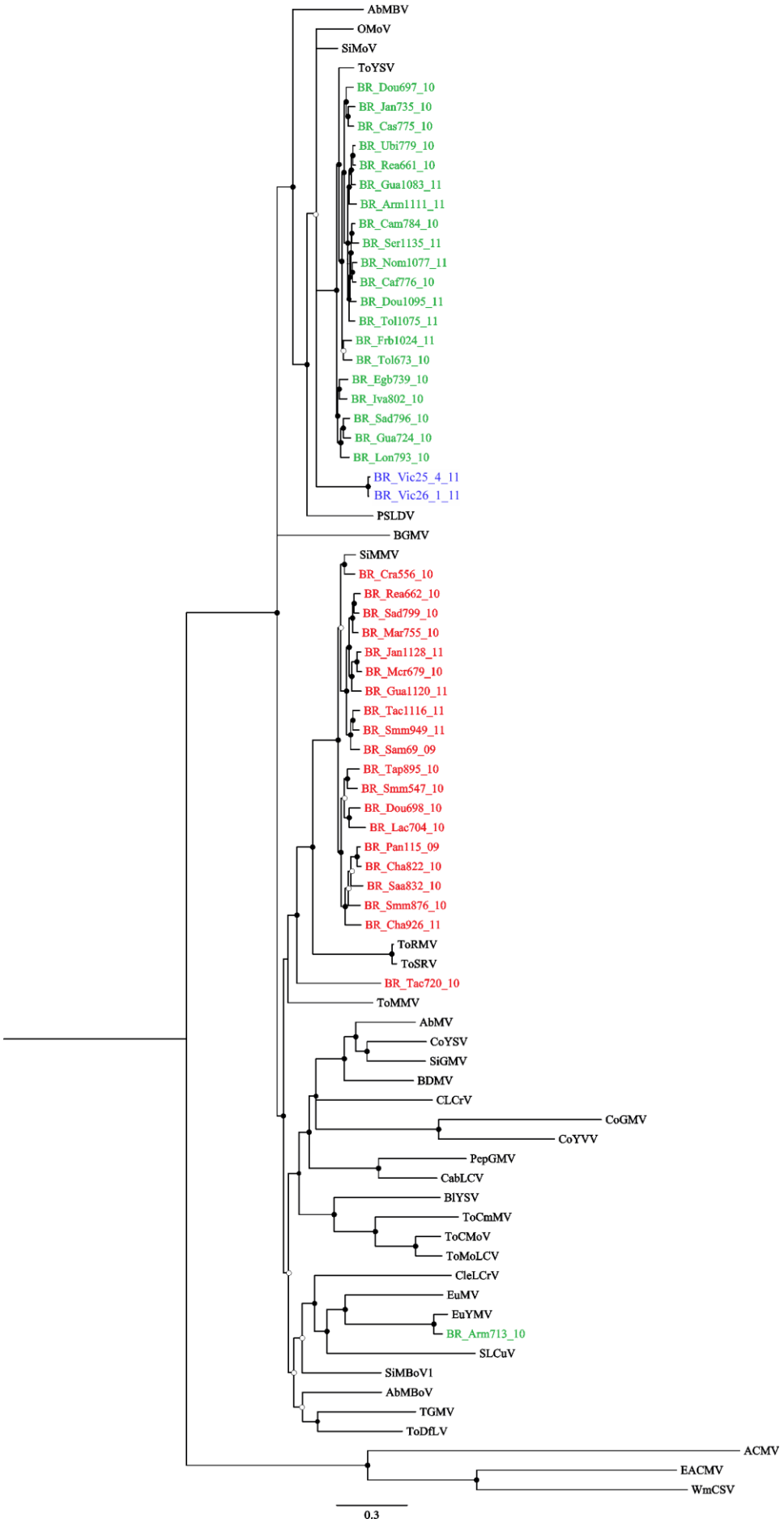

AbMBV
OMoV
SiMoV
ToYSV
BR_Dou697_10
BR_Jan735_10
BR_Cas775_10
BR_Ubi779_10
BR_Rea661_10
BR_Gua1083_11
BR_Arm1111_11
BR_Cam784_10
BR_Ser1135_11
BR_Nom1077_11
BR_Caf776_10
BR_Dou1095_11
BR_Tol1075_11
BR_Frb1024_11
BR_Tol673_10
BR_Egb739_10
BR_Iva802_10
BR_Sad796_10
BR_Gua724_10
BR_Lon793_10
BR_Vic25_4_11
BR_Vic26_1_11
PSLDV
BGMV
SiMMV
BR_Cra556_10
BR_Rea662_10
BR_Sad799_10
BR_Mar755_10
BR_Jan1128_11
BR_Mcr679_10
BR_Gua1120_11
BR_Tac1116_11
BR_Smm949_11
BR_Sam69_09
BR_Tap895_10
BR_Smm547_10
BR_Dou698_10
BR_Lac704_10
BR_Pan115_09
BR_Cha822_10
BR_Saa832_10
BR_Smm876_10
BR_Cha926_11
ToRMV
ToSRV
BR_Tac720_10
ToMMV
AbMV
CoYSV
SiGMV
BDMV
CLCrV
CoGMV
CoYVV
PepGMV
CabLCV
BlYSV
ToCmMV
ToCMoV
ToMoLCV
CleLCrV
EuMV
EuYMV
BR_Arm713_10
SLCuV
SiMBoV1
AbMBoV
TGMV
ToDfLV
ACMV
EACMV
WmCSV
0.3

Figure 2

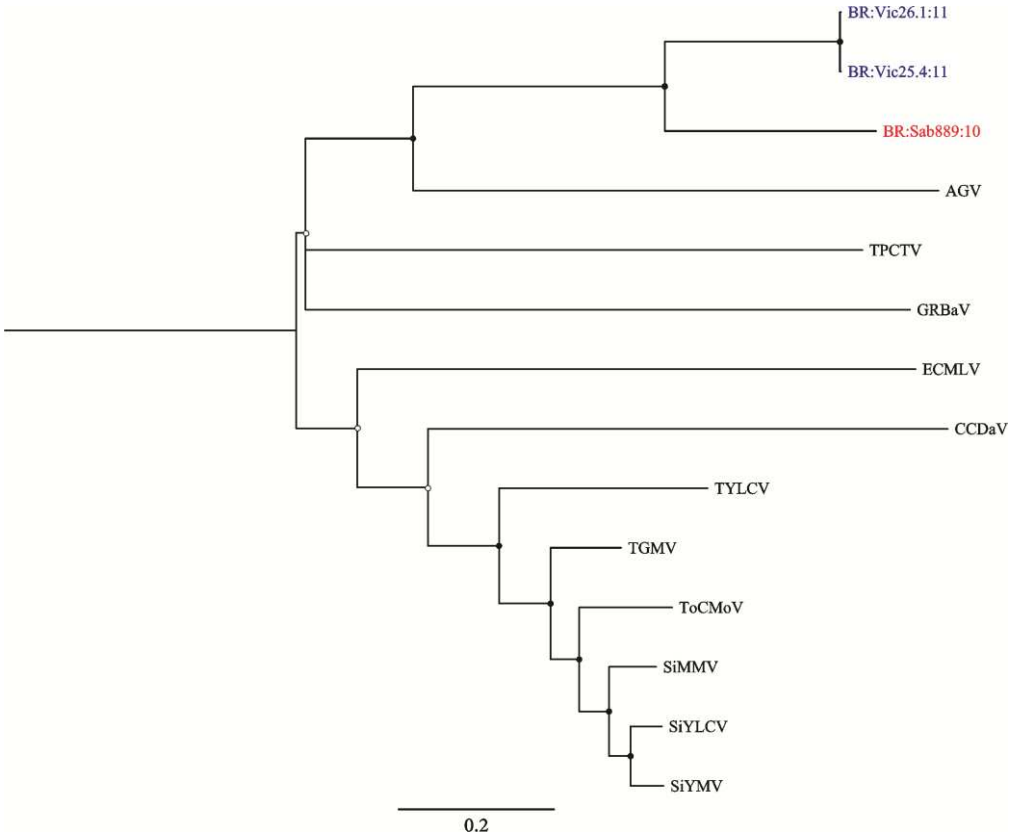
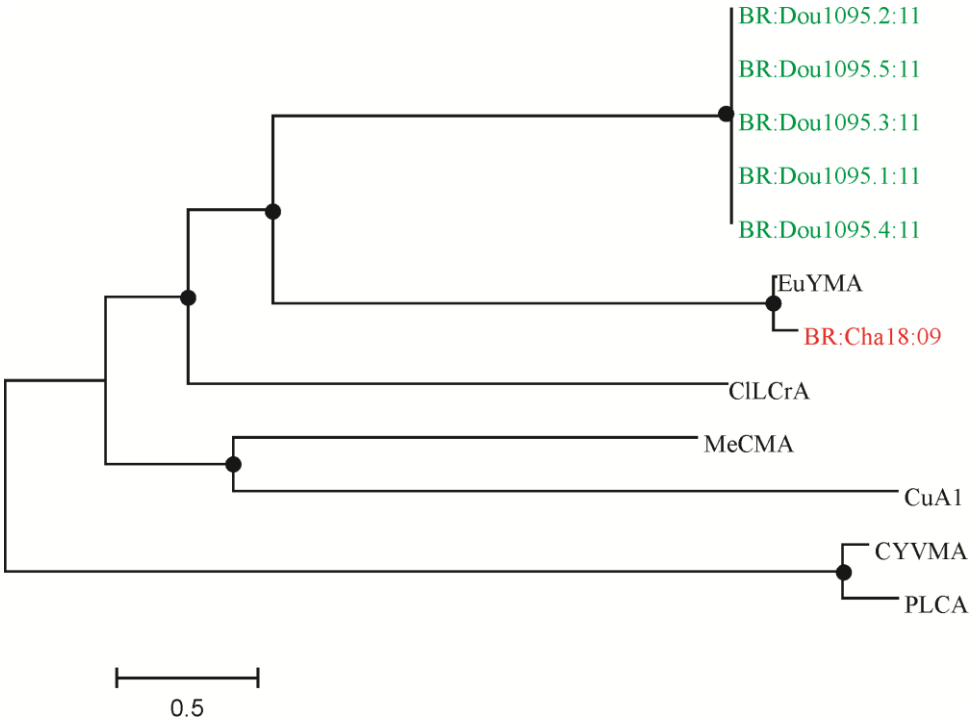
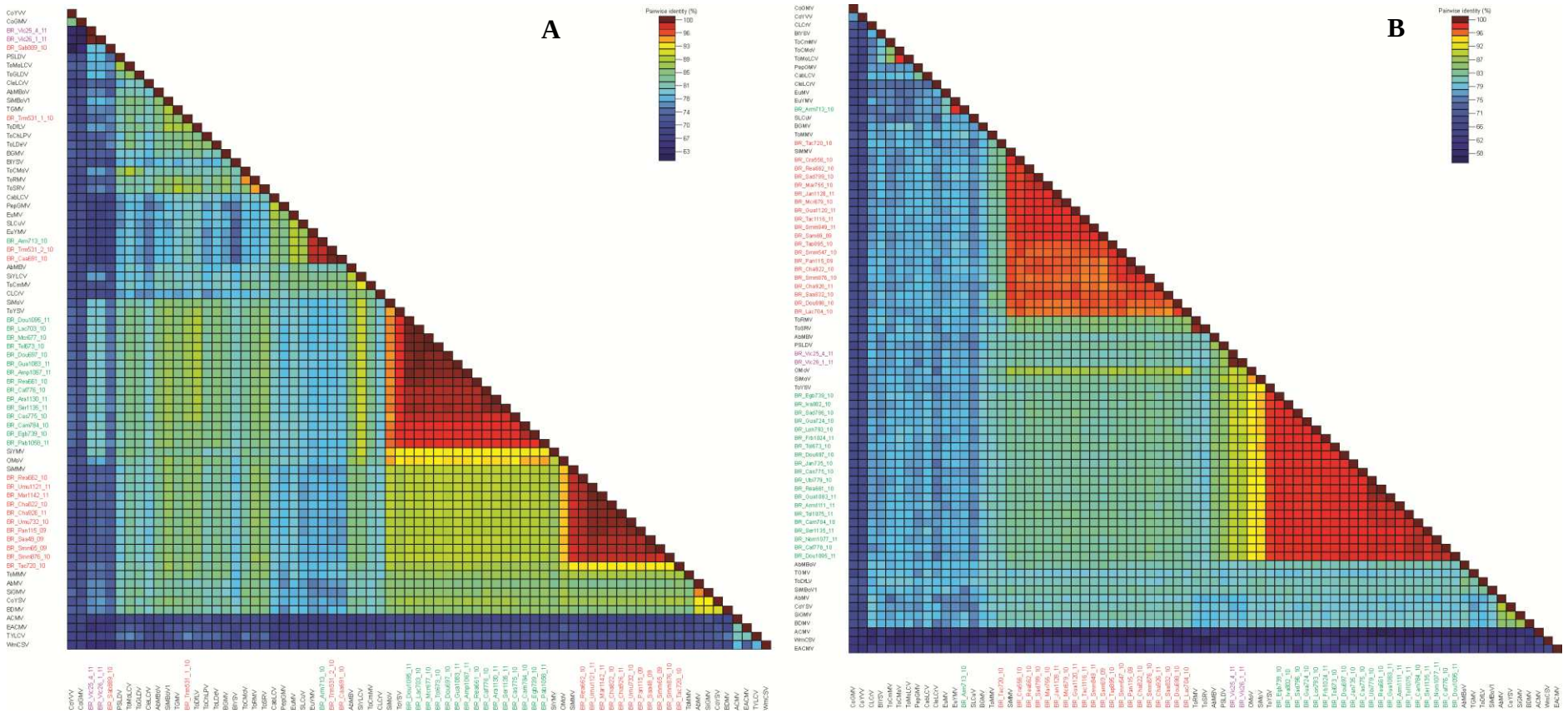

BR:Vic26.1:11
BR:Vic25.4:11
BR:Sab889:10
AGV
TPCTV
GRBaV
ECMLV
CCDaV
TYLCV
TGMV
ToCMoV
SiMMV
SiYLCV
SiYMV
0.2

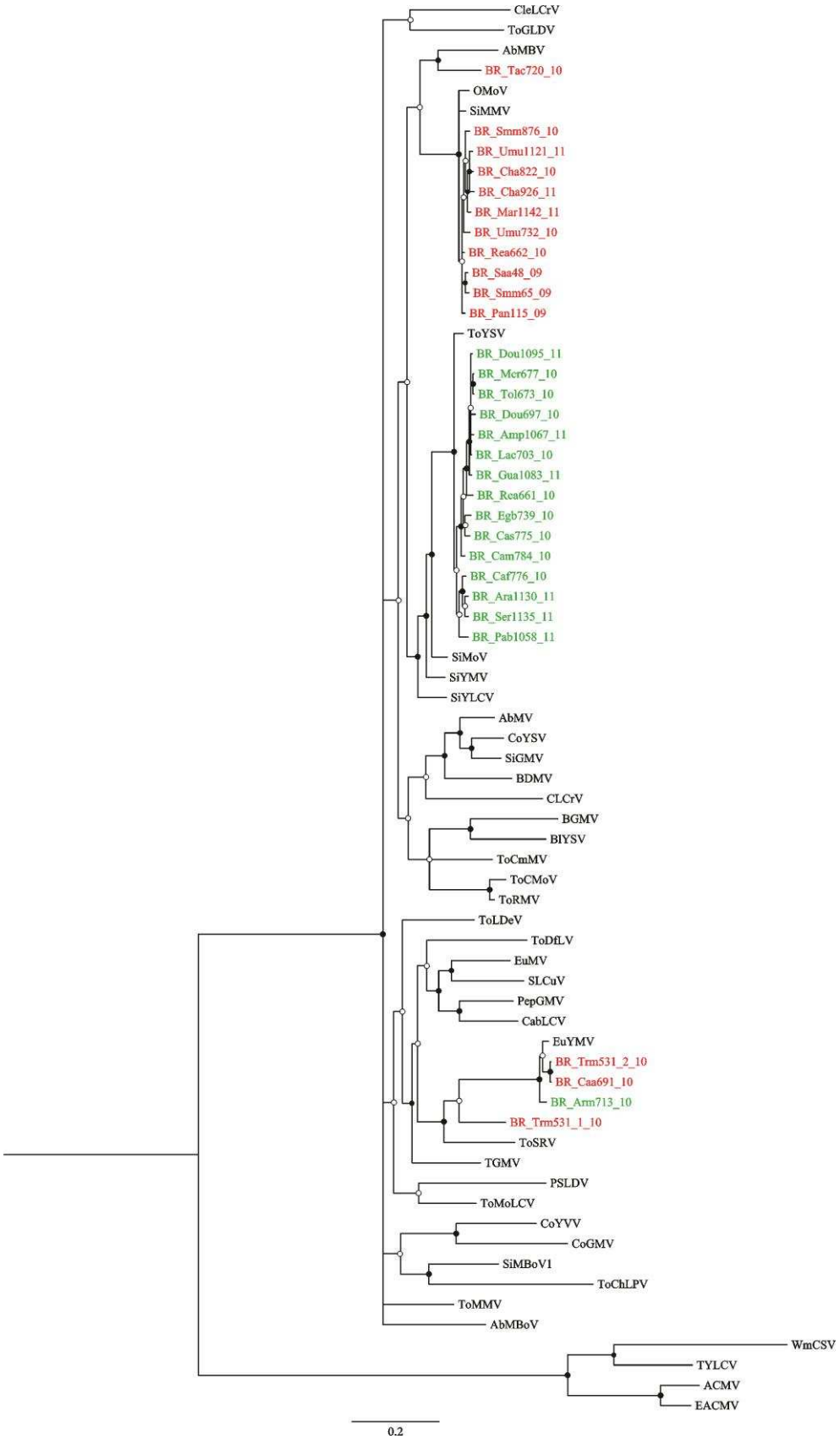
Figure 3


BR:Dou1095.2:11
BR:Dou1095.5:11
BR:Dou1095.3:11
BR:Dou1095.1:11
BR:Dou1095.4:11
EuYMA
BR:Cha18:09
ClLCrA
MeCMA
CuA1
CYVMA
PLCA
0.5

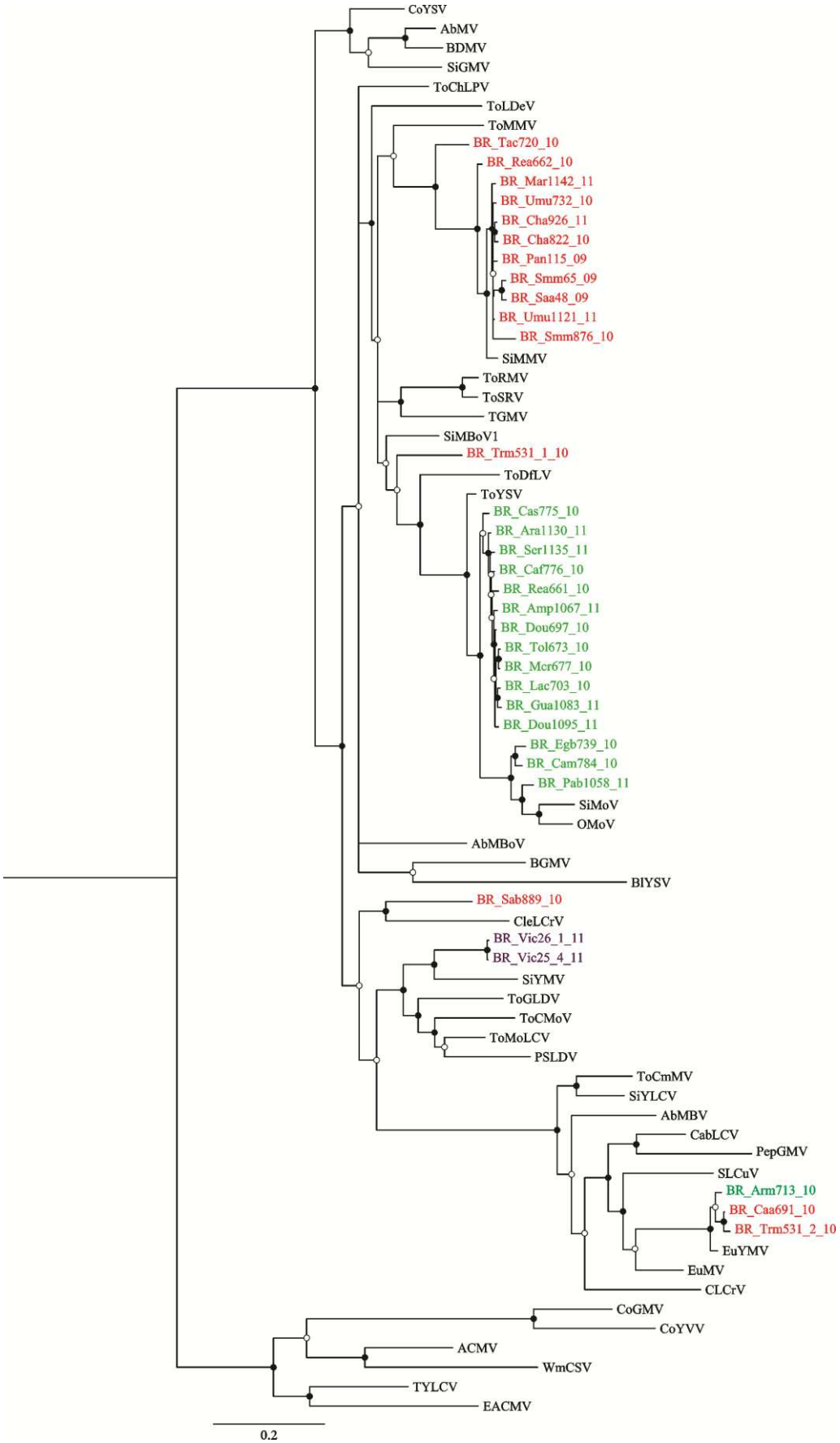
Supplementary Figure S1



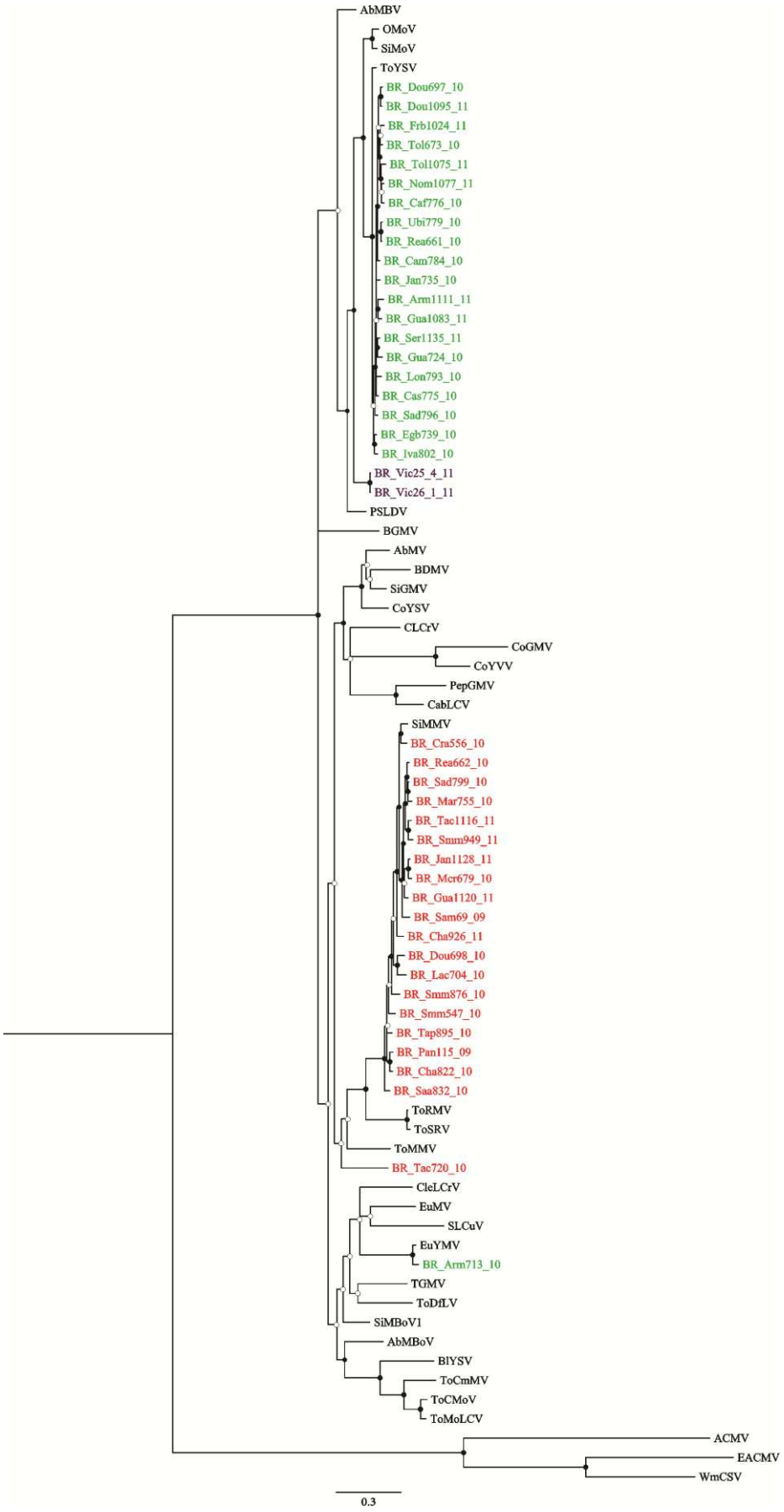
Supplementary Figure S2A



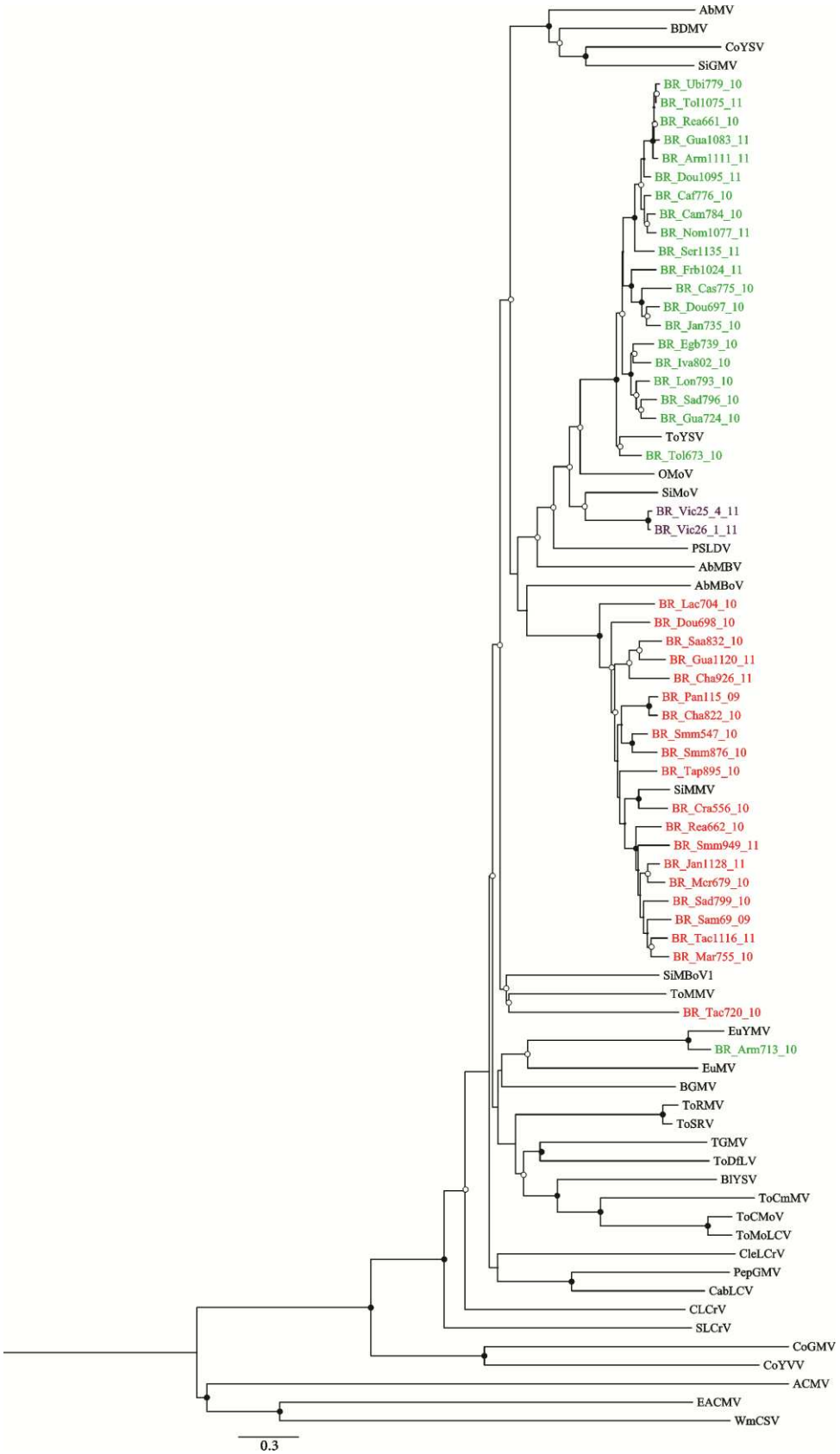
Supplementary Figure S2B



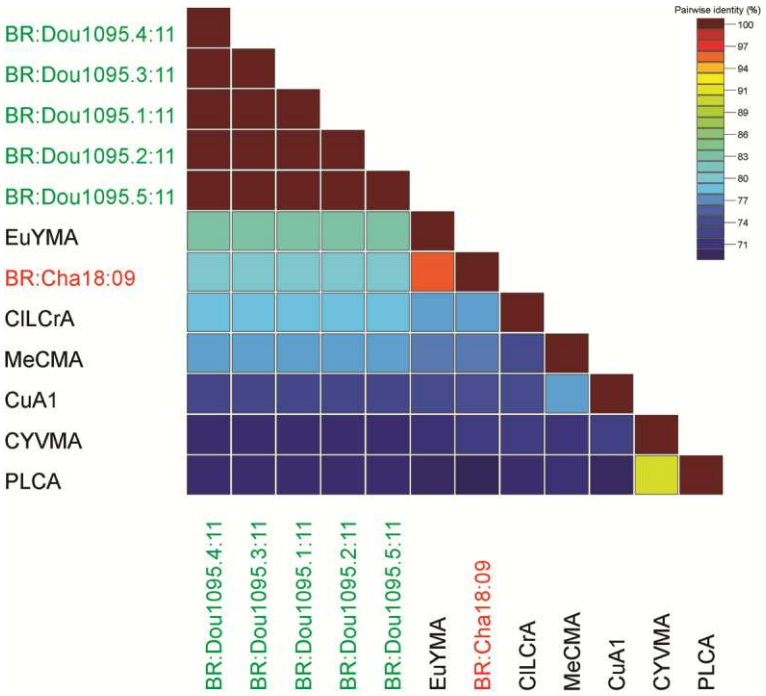
Supplementary Figure S3A



Supplementary Figure S3B



Supplementary Figure S4



CAPÍTULO 2

GENETIC STRUCTURE AND VARIABILITY OF TWO BEGOMOVIRUS POPULATIONS INFECTING NON-CULTIVATED HOSTS

Ferro, C.G., Silva, J.P., Pinto, V. B., Pereira, H.M.B., Xavier, C.A.D., Godinho, M.T., Lima, A.T.M., Lau, D., Zerbini, F.M. Genetic structure and variability of two begomovirus populations infecting non-cultivated hosts. Journal of General Virology, *in preparation*.

Genetic structure and variability of two begomovirus populations infecting non-cultivated hosts

C.G. Ferro[1], J.P. Silva[1], V.B. Pinto[1], H.M.B. Pereira[1], C.A.D. Xavier[1], M.T. Godinho[1], A.T.M. Lima[1,2], D. Lau[3], F.M. Zerbini[1#]

[1]Dep. de Fitopatologia/BIOAGRO and National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[2]Present address: Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, MG 38400-902, Brazil

[3]Present address: Embrapa Trigo, Passo Fundo, RS 99001-970, Brazil

[#]Corresponding author:

Phone: (+55-31) 3899-2935; Fax: (+55-31) 3899-2240; E-mail: zerbini@ufv.br

Abstract

Begomoviruses are responsible for serious diseases in several crops of great economic importance, especially in tropical and subtropical regions. Non-cultivated plants act as begomovirus reservoirs, enabling the occurrence of mixed infections where recombination can occur. A high species diversity of begomoviruses is found in non-cultivated plants, and begomoviruses originally described in these plants have already been found naturally infecting cultivated plants. Previous work indicated a higher degree of genetic variability in populations of begomoviruses infecting non-cultivated hosts compared to crop-infecting begomoviruses. Understanding the dynamics and genetic variability of viral populations is important to assist on the prediction and consequent prevention of new virus diseases in cultivated plants. In this work, full-length begomovirus DNA-A and DNA-B components of *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV) were sequenced from samples of the non-cultivated hosts *Sida* spp. and *Leonurus sibiricus*, respectively, collected in the states of Rio Grande do Sul (RS), Paraná (PR) and Mato Grosso do Sul (MS) between 2009 and 2011. We found a high genetic variability for both begomovirus populations, with the DNA-B being more permissible to the variation than the DNA-A. Recombination partly explains the high genetic variability found for these two viruses, but mutational dynamics was the primary factor in the diversification of both viral populations.

Introduction

The *Geminiviridae* family includes viruses whose genomes are comprised of one or two molecules of circular, single strand DNA (ssDNA) encapsidated by a single structural protein into twinned icosahedral particles. Geminiviruses are currently divided into seven genera (*Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus*) on the basis of their genome organization, phylogenetic relationships, host range, and insect vectors (Brown *et al.*, 2012; Varsani *et al.*, 2014). The genus *Begomovirus* includes viruses with mono- and bipartite genomes transmitted by whiteflies of the *Bemisia tabaci* cryptic species complex (Homoptera: Aleyrodidae) to dicotyledonous plants (Brown *et al.*, 2015).

Begomoviruses are broadly divided into two groups: Old World (OW; Europe, Africa, Asia and Oceania) and New World (NW; the Americas), based on phylogenetic studies and genomic features (Padidam *et al.*, 1999; Paximadis *et al.*, 1999; Rybicki, 1994). The DNA-A contains five open reading frames (ORFs) which encode proteins with functions in viral replication (Rep, Ren), suppression of host defenses (Trap) and particle formation (CP), while the two ORFs in the DNA-B encode a movement protein (MP) and a nuclear shuttle protein (NSP) which are involved in intra- and intercellular movement in the plant, host range determination and suppression of host defenses (Brustolini *et al.*, 2015; Hanley-Bowdoin *et al.*, 2013; Mahajan *et al.*, 2011; Rojas *et al.*, 2005).

Begomoviruses are responsible for severe diseases in several crops of great economic importance, especially in tropical and subtropical regions (Briddon, 2003; Legg & Fauquet, 2004; Monci *et al.*, 2002; Morales, 2006; Polston & Anderson, 1997; Rocha *et al.*, 2013). This fact has been attributed to the ability of these viruses

to evolve rapidly, enabling the emergence of new species or more aggressive variants, together with the emergence of polyphagous insect vector populations with greater reproductive capacity (Gilbertson *et al.*, 2015; Navas-Castillo *et al.*, 2011).

Virus evolution depends firstly on the occurrence of mutations. Begomoviruses evolve at rates which are comparable to those of ssRNA viruses, with estimates of substitution rates in the order of 10^{-3}-10^{-4} substitutions/site/year (Duffy & Holmes, 2008; Duffy & Holmes, 2009). Although mutational dynamics is a primary factor in the diversification of viral populations (García-Arenal *et al.*, 2003; Roossinck, 1997), it is not responsible for all the genetic variation, since other mechanisms that generate variability, including recombination, might contribute significantly (Lima *et al.*, 2013; Martin *et al.*, 2011). Recombination is a common event in geminiviruses (Lefeuvre *et al.*, 2009; Padidam *et al.*, 1999) and contributes greatly to their genetic diversity, increasing their potential for adaptation to different hosts and environmental conditions (Berrie *et al.*, 2001; Lefeuvre & Moriones, 2015; Monci *et al.*, 2002). A number of natural begomovirus recombinants have been responsible for severe diseases and great economic losses in cassava in East Africa (Pita *et al.*, 2001; Zhou *et al.*, 1997) and cotton and okra in Pakistan (Briddon *et al.*, 2014). Another source of genetic variability is pseudo-recombination (or reassortment), which may occur by the exchange of genomic components between bipartite viruses (Andrade *et al.*, 2006; Garrido-Ramirez *et al.*, 2000; Gilbertson *et al.*, 1993).

The genetic structure of plant virus populations refers to the amount of genetic variability and its distribution within and among subpopulations (García-Arenal *et al.*, 2001). Several studies have been conducted to investigate the main mechanisms which are responsible for shaping the genetic structure and variability of

geminivirus populations in different hosts and geographic regions (Fondong *et al.*, 2000; Legg & Thresh, 2000; Lima *et al.*, 2013; Owor *et al.*, 2007; Pita *et al.*, 2001; Rocha *et al.*, 2013; Zhou *et al.*, 1997). Comparative analyses of begomovirus populations found in cultivated and non-cultivated plants indicated that populations infecting non-cultivated plants have a higher degree of genetic variability compared to those present in cultivated plants (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013). Some of these studies have suggested that recombination explains the greater genetic variability of begomoviruses in non-cultivated hosts (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014).

A high species diversity of begomoviruses in non-cultivated plants, especially in *Macroptilium* and *Sida* spp., has been observed in studies conducted in several countries in the Americas (Amarakoon *et al.*, 2008; Fiallo-Olivé *et al.*, 2012; Frischmuth *et al.*, 1997; Idris *et al.*, 2003; Jovel *et al.*, 2004; Roye *et al.*, 1997; Silva *et al.*, 2012; Tavares *et al.*, 2012). Many of the viral species found in these hosts also infect cultivated plants (Barbosa *et al.*, 2012; Barreto *et al.*, 2013; Fernandes *et al.*, 2011; Fernandes *et al.*, 2009; Fernandes *et al.*, 2014; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013), reinforcing the role of non-cultivated plants as virus reservoirs and the importance of characterizing begomovirus populations infecting these hosts. An understanding of the dynamics and genetic variability of viral populations in non-cultivated plants can be helpful for the prediction and consequent prevention of new virus diseases in cultivated plants.

We have carried out a study to obtain detailed information on the genetic structure and on the mechanisms which influence the genetic variability of begomovirus populations infecting two ubiquitous non-cultivated plants in Brazil. We hypothesized, based on previous studies, that begomovirus populations obtained

from these hosts would present a high degree of genetic variability and would be structured by geographical region. Indeed, we found a high genetic variability for begomovirus populations in both hosts. Although a high level of recombination was detected, mutational dynamics was the primary factor in the diversification of viral populations. The results are inconclusive regarding genetic structuration based on geographical location.

Results

A total of 31 DNA-A clones and 41 DNA-B clones were obtained from the *Sida* spp. and *Leonurus sibiricus* samples. A number of these clones corresponded to new begomovirus species and to DNA satellites, and were described in detail in Chapter 1. Pairwise sequence comparisons of 10 DNA-A clones from *Sida* spp. with those deposited in GenBank indicated that the sequences share >96.3% nucleotide identity amongst themselves and 96.3% to 97.2% identity with *Sida micrantha mosaic virus* (SiMMV) (accession number FN436003), and 19 DNA-B clones share >88.0% nucleotide sequence identity amongst themselves and 89.1% to 93.3% identity with SiMMV (FN436004) (Figure 1; Suppl. Table S1). Pairwise sequence comparisons of the 15 DNA-A clones obtained from *L. sibiricus* plants indicated that the sequences share >93.4% sequence identity with each other and 94.6% to 99.0% identity with *Tomato yellow spot virus* (ToYSV) (JX863082), and the 19 DNA-B clones share >91.9% nucleotide identity amongst themselves and 92.3% to 94.4% identity with ToYSV (JX513953) (Figure 1). A further 24 DNA-A and 14 DNA-B sequences were retrieved from GenBank (Suppl. Table S2). The final data sets consisted of 25 DNA-A and 30 DNA-B sequences of SiMMV, and 24 DNA-A and 22 DNA-B sequences of ToYSV.

The SiMMV population is more variable than the ToYSV population

The SiMMV population showed greater genetic variability for both DNA components compared to the ToYSV population (Tables 1 and 2). In general, DNA-B sequences were more variable than DNA-A sequences in the two data sets analyzed. The nucleotide sequences of the *Rep*, *CP*, *MP* and *NSP* genes and of the DNA-A intergenic region (IR-A), DNA-B large intergenic region (LIR-B) and short intergenic region (SIR-B) were analyzed separately to verify whether the distribution of variability is evenly distributed along the genome. Non-coding regions were more variable compared to coding regions. The SIR-B and LIR-B were more variable compared to IR-A, and the SIR-B was more variable than the LIR-B for both populations. In general, the *MP* and *NSP* genes were more variable than the *CP* and *Rep* genes, reflecting the greater variability of the DNA-B compared to the DNA-A. Also, within each component, *Rep* was more variable than *CP*, and *NSP* was more variable than *MP* (Tables 1 and 2).

Phylogenetic analysis

In the SiMMV DNA-A phylogenetic tree, isolates from Goiás (GO) and Rio de Janeiro (RJ) formed one clade, most Bolivian (BO) isolates formed a second clade, and a third clade includes isolates from Mato Grosso do Sul (MS), Paraná (PR) and Rio Grande do Sul (RS) plus one additional isolate from BO (Figure 2A). Within this third clade there are two well-defined subclades, one with isolates from MS, and another with isolates from PR and RS, suggesting geographical subdivision (Figure 2A). Isolates BR:Rea662:10 (PR) and BR:Smm876:10 (RS) grouped with isolates of PR and RS, but are more distantly related to them. Phylogenetic trees

based on the nucleotide sequences of the *CP* and *Rep* genes were constructed and they are not congruent (Figure 3), suggesting the existence of recombination events among the isolates.

The SiMMV DNA-B tree showed three major clades. The first clade includes isolates from RS, MS and PR, subdivided into three well-defined minor clades separating the isolates from each state and therefore suggesting a relationship by geography. However, this relationship is not reflected when the other two clades are examined and compared with the first one. The second clade includes isolates from MS, GO e RJ, and the third incudes isolates from RS and BO (Figure 2B). Thus, isolates from RS and MS are each split into two clades, together with isolates from different regions (Figure 2B). Phylogenetic trees based on the nucleotide sequences of the *MP* and *NSP* genes were highly incongruent. The MP tree shows few well supported clades (Figure 4A), and the NSP tree shows three clades with significant genetic distance between them (two with 100% bootstrap support; Figure 4B).

The ToYSV DNA-A tree showed four major clades: one clade with isolates from PR and Argentina (AR); a second clade with isolates from PR; a third with one isolate from PR and two from Minas Gerais (MG); and a fourth clade with isolates from Paraguay (PY), PR and MS (Figure 5A). In the ToYSV DNA-B tree there were two major clades, one with isolates from PR, MG and AR and the second with isolates from PR and MS. This second clade is subdivided into three smaller clades, all three including isolates from both PR and MS (Figure 5B). Phylogenetic trees based on the nucleotide sequences of the *CP*, *Rep*, *MP* and *NSP* genes were constructed, however the CP/Rep and MP/NSP trees were incongruent between each other (Figure 6 and 7), suggesting the existence of recombination events in both DNA components.

Genetic structure of SiMMV and ToYSV populations

Due to the indication of a possible geographical subdivision among isolates observed in the SiMMV DNA-A tree, and in spite of the small number of isolates from each region, population subdivision tests were performed. The subdivision test performed for the SiMMV total population indicated high genetic differentiation for both the DNA-A and DNA-B components, with a higher differentiation for the DNA-A compared to the DNA-B (Fst = 0.42929 and 0.32271, respectively) (Table 4). The tests contrasting the DNA-A of isolates from different regions showed a high genetic differentiation in all cases, except for the isolates from PR and RS (Fst = 0.02513; Table 4), corroborating the phylogeny results. Isolates from GO and RJ showed a smaller degree of genetic differentiation compared to isolates from other regions, corroborating the phylogenetic analysis in which GO and RJ isolates grouped in the same clade without a clear separation between them (Table 4 and Figure 2A). Together, the SiMMV DNA-A phylogenetic tree and the population subdivision tests are indicative of geographical structuring, with four subpopulations: (*i*) GO and RJ; (*ii*) BO; (*iii*) MS; (*iv*) PR and RS. However, due to the small number of isolates from each region, these results should be considered preliminary. Likewise, the population subdivision tests for the DNA-B also indicated high genetic differentiation for most pairwise comparisons (Table 4), except for the isolates from MS and RS (Fst = 0.03542; Table 4). This also reflected the phylogenetic analysis, in which isolates from these two states were split into different clades, while isolates from the other regions clustered together (Figure 2B). Thus, although the DNA-B analysis indicates the existence of three subpopulations (GO/RJ, BO and PR) which were also infered by the DNA-A analysis, it does not support subdivision of isolates from MS and RS.

Population subdivision test performed for the ToYSV total population indicated high genetic differentiation for the DNA-A and low genetic differentiation for the DNA-B (Fst = 0.58686 and Fst= 0.03882, respectively). For the DNA-B, PR and MS isolates comprise a single population. Although the population subdivision tests contrasting the DNA-A of isolates from different regions indicated a high genetic differentiation for most of the comparisons (Table 4), this is not supported in the phylogeny (Figure 5A), mostl likely due to the strong bias in terms of isolates from PR (16 out of the 24 isolates comprising the data set) and very small numbers of isolates from the other regions (one from PY, two from AR, two from MG, three from MS).

SiMMV is most prone to recombination than ToYSV

To investigate possible recombination events, full-length DNA-A and DNA-B sequences of SiMMV and ToYSV were analyzed using the RDP4 program. A higher number of putative recombination events were detected in both SiMMV components compared with ToYSV (Table 3). For both viruses, the DNA-B was most prone to recombination than the DNA-A, with a higher number of putative events (Table 3). For SiMMV, 7 and 15 single recombination events were detected for the DNA-A and DNA-B, respectively, whereas for ToYSV, 5 and 7 events were detected for the DNA-A and DNA-B, respectively (Table 3).

In the ToYSV DNA-A, most recombination events involved breakpoints located within the *Rep* gene and the IR-A, whereas for SiMMV the breakpoints were mostly located in the *Rep* and *CP* genes. In the ToYSV DNA-B, most recombination events involved breakpoints within the *MP* gene and the LIR-B. For the SiMMV DNA-B, in contrast, most of the breakpoints are located within the *NSP* gene and the

LIR-B (Table 3). These results corroborate with the conflicting CP/Rep and MP/NSP phylogenies for both viruses. Isolates that showed evidence of a shared recombination event in the SiMMV Rep sequence, for example, were readily identified in well supported clades observed on the SiMMV Rep-based phylogenetic tree (Table 3 and Figure 3B). There were no geographical patterns associated with recombination events.

Relative contribution of mutation and recombination to the genetic variability of SiMMV and ToYSV populations

All substitutions over the branches in midpoint-rooted ML trees were mapped, and those branches which were associated with well-supported recombination events were determined as described by Lima *et al.* (2013).

For SiMMV, the total number of substitutions was higher in the NSP and MP trees compared to the Rep and CP trees (Table 5). In the CP tree a total of 252 substitutions were counted (Table 5), and 41 substitutions were assigned to three recombination events (events 3, 4 and 6 in Table 3; showed in red, blue and pink, respectively, in Figure 3A). Event 4 accounted for the largest individual contribution amongst the unique recombination events (37 substitutions) and was exclusively detected in isolate BR:Pda43:05 (KC706536) (Figure 3A). As expected, a considerably higher number of substitutions (565) was observed for the Rep ML tree (Table 5), and 82 substitutions were assigned to recombination events 1, 5, 6 and 7 (showed in yellow, orange, pink and purple, respectively, in Figure 3B). Sixty-two and 11 substitutions, respectively, mapped to events 1 [shared by isolates BR:780Si4a:08 (JX415194), BR:780Si4b:08 (JX415195) and BR:GO:Luz780:03 (JX415187)] and 5 (shared by 9 isolates from Paraná and Rio Grande do Sul). One

substitution mapped to event 6 [shared by isolates BO:CF1:07 (HM585437) and BO:CF2:07 (HM585439)] and 8 mapped to event 7 (BR:Sag3:Soy:08, FJ686693) (Figure 4B).

A total of 692 substitutions were counted in the SiMMV MP tree (Table 5), with just 17 substitutions being assigned to three recombination events (events 7, 10 and 13 in Table 3; showed in green, brown and light pink, respectively, in Figure 4A). All three events were detected in unique isolates: event 7 in BR:Saa832:10 (9 substitutions), event 10 in BR:Sad799:10 (1 substitution) and event 13 in BR:Smm876:10 (7 substitutions) (Figure 4A). In the NSP tree, a total of 819 substitutions were counted (Table 5), and 60 substitutions were assigned to 5 recombination events (events 3, 5, 8, 10 and 11 in Table 3; showed in blue, orange, grey, brown and pink, respectively, in Figure 4B). Events 3, 5 and 10 were detected, respectively, in isolates BR:okra:5157:07 (EU908734; 5 substitutions), BR:Saa832:10 (9 substitutions) and BR:Sad799:10 (7 substitutions). Events 8 [shared by isolates BO:BoVi:2007 (HM585432), BO:CF1:07 (HM585438) and BO:CF2:07 (HM585440)] and 11 (shared by BR:Gua1120:11 and BR:Saa832:10) are associated with 27 and 12 substitutions, respectively (Figure 4B).

Isolate SiMMV-[BR:Saa832:10] is highly recombinant, with three well supported recombination events detected in its DNA-B. Two of these events overlap, with breakpoints in the *NSP* gene (one of them detected only in this isolate) and the third has breakpoints in the *MP* gene (also detected only in this isolate). This isolate contributed with 50% of the substitutions attributed to recombination in the DNA-B tree.

Many well supported recombination events were detected in the intergenic regions of the DNA-A and DNA-B, but as only the coding regions were being

analyzed, they were not counted and for the events in which the exact location of the breakpoints was variable and frequently ended in the intergenic region instead of the *Rep*, *CP*, *MP* and *NSP* genes, only those substitutions that occurred within coding regions were added to η_r. Nevertheless, for the four SiMMV genes analyzed (*CP*, *Rep*, *MP* and *NSP*), η_m was >84% (Table 5), indicating that most of the genetic variability is explained by mutational dynamics rather than recombination.

In all ToYSV trees, the total number of substitutions was lower than in the SiMMV trees. The ToYSV Rep tree had the highest number of substitutions (393; Table 5). In spite of the *Rep* gene having the largest total number of substitutions, less than 1% of them were assigned to recombination events (only 3 substitutions). Conversely, the *CP* gene had the lowest number of substitutions (127) but the highest number of substitutions assigned to recombination (but even then, only 6 substitutions; Table 5). In the Rep tree, all three substitutions were assigned to recombination event 2 (indicated in green in Figure 6B). In the CP tree, the 6 substitutions were assigned to recombination events 4 and 5 (showed in yellow and orange, respectively, in Figure 6A), with three substituions assigned to each event.

In the MP tree, a total of 283 substitutions were counted (Table 5) and only 13 substitutions were assigned to two recombination events (events 1 and 3 in Table 3; showed in blue and pink, respectively, in Figure 7A). Both events were detected in unique isolates: event 1 in BR:Dou1095:11 (3 substitutions) and event 3 in BR:Arm1111:11 (10 substitutions) (Figure 7A). A total of 320 substitutions were counted in the NSP tree (Table 5) and only 11 substitutions were assigned to a unique recombination event (event 2) shared by isolates BR:Cas775:10, BR:Dou697:10 and BR:Jan735:10) (indicated in red in Figure 7B).

As observed for SiMMV, for the four ToYSV genes analyzed (*CP*, *Rep*, *MP* and *NSP*), η_m was much higher than η_r (>95%; Table 5), indicating that mutation, rather than recombination, is the main driver of genetic variability.

Discussion

It it currently accepted that non-cultivated hosts act as a reservoir of begomovirus diversity and as a source of inoculum which may contribute to epidemics in cultivated hosts (Barreto *et al.*, 2013; Castillo-Urquiza *et al.*, 2008; Idris *et al.*, 2003; Jovel *et al.*, 2004). Studies conducted in Brazil and in a number of other countries indicate a high species diversity begomoviruses in non-cultivated plants (Amarakoon *et al.*, 2008; Barbosa *et al.*, 2012; Fernandes *et al.*, 2014; Fiallo-Olivé *et al.*, 2012; Frischmuth *et al.*, 1997; Idris *et al.*, 2003; Jovel *et al.*, 2004; Roye *et al.*, 1997; Silva *et al.*, 2012; Tavares *et al.*, 2012), and comparative analyses of begomovirus populations found in cultivated and non-cultivated plants indicate that the viral populations infecting non-cultivated plants have a higher degree of genetic variability compared to those present in cultivated plants (Lima *et al.*, 2013; Rocha *et al.*, 2013).

Our analyses of the genetic variability of two begomovirus populations obtained from non-cultivated hosts, SiMMV from *Sida* spp. and ToYSV from *Leonurus sibiricus*, indicate a high degree of variability in both genomic components for the two viruses. Rocha *et al.* (2013) found a much lower degree of variability for four tomato-infecting begomoviruses compared to that of a begomovirus from a non-cultivated host (*Blainvillea yellow spot virus*, BlYSV) sampled in south-eastern Brazil. The genetic variability indices for the SiMMV and ToYSV populations

observed here, for both components, are much higher than those determined by Rocha *et al.* (2013) for BlYSV.

The SiMMV population, which is more recombination-prone, was also more variable, for both components, than the ToYSV population. The genetic variability of SiMMV was similar to that observed for *Macroptilium yellow spot virus* (MaYSV), a begomovirus that infects both non-cultivated and cultivated hosts (*Macroptilium lathyroides* and common bean) (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013).

ToYSV is a virus that was first described infecting tomato plants (Ambrozevicius *et al.*, 2002) and has also been detected in bean and soybean (Rodríguez-Pardina *et al.*, 2011), but the results suggest that its original host may be a non-cultivated plants (either *L. sibiricus* or an unidentified species). Even though the genetic variability of ToYSV was lower than that of SiMMV, it is much higher than that observed for tomato-infecting viruses (Lima *et al.*, 2013; Rocha *et al.*, 2013), and equivalent to the variability observed for begomoviruses described in other non-cultivated plants (Rocha *et al.*, 2013; Silva *et al.*, 2012). Phylogenetically, ToYSV isolates are grouped with begomoviruses that infect primarily non-cultivated plants (Andrade *et al.*, 2006; Calegario *et al.*, 2007; Celli *et al.*, 2014; Rodríguez-Pardina *et al.*, 2011), and it has often been found in natural conditions infecting *Leonurus sibiricus* (Barbosa *et al.*, 2012; Fernandes *et al.*, 2014). Together, these results strongly indicate that ToYSV is a virus from non-cultivated plants that is capable of, eventually, infecting cultivated plants. On this regard, it has an unusually wide host range, including *L. sibiricus* (family Lamiaceae) and cultivated plants from at least two botanical families (Solanaceae and Fabaceae). Although it is only rarely

detected in tomatoes in Brazil (Rocha *et al.*, 2013), it is significantly widespread in bean and soybean in northwestern Argentina (Rodríguez-Pardina *et al.*, 2011).

Our results further indicate that the DNA-B is more variable than the DNA-A for both SiMMV and ToYSV, in accordance to previous studies (Briddon *et al.*, 2010; Rocha *et al.*, 2013). This could be explained by two hypotheses: *i*) the DNA-B has a different origin and evolutionary history in relation to the DNA-A, possibly evolving from a satellite molecule that was captured by a monopartite begomovirus (Briddon *et al.*, 2010; Nawaz-Ul-Rehman & Fauquet, 2009); *ii*) the DNA-B can tolerate a greater number of mutations since it does not contain overlapping genes. The DNA-A encodes four overlapping genes (*Rep*, *AC4*, *Trap* and *Ren*), which are involved in several *cis*- and *trans*-interactions (Rep-Rep, Rep-iterons, Rep-Ren, Trap-promoter) that could be negatively affected by nucleotide substitutions (Briddon *et al.*, 2010). It is noteworthy that the two hypotheses are not mutually exclusive, with the combination of them leading to greater genetic variability in the DNA-B compared to the DNA-A.

The genetic structure of begomovirus populations is determined by processes that generate genetic variability (mutation and recombination) and the interaction between adaptation to host species and to the vector, and is influenced by the geographical distribution of the hosts, vectors and other begomoviruses (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013). Studies have shown that populations of Brazilian begomovirus are highly recombinant (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013) and that recombination explains the greater variability of begomoviruses in non-cultivated hosts (Lima *et al.*, 2013). Segregation of tomato-infecting begomoviruses by geographical region was observed in Brazil, with different viruses being prevalent in different growing areas (Rocha *et*

al., 2013), and evidence of structuring by both host and geography was found for BGMV populations (Ramos-Sobrinho *et al.*, 2014). In contrast, for MaYSV populations, no evidence of structure by geographical region or host was found (Lima *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014). In our study, despite the high Fst values for both SiMMV and ToYSV populations, support for structuration by geographical region was inconclusive. We detected many well supported recombination events, and the phylogeny reflected, in part, the clustering of isolates that shared the same recombination event, with no geographical patterns associated with these recombination events. For example, in the SiMMV DNA-A and Rep trees, the same clade was observed with isolates from PR and RS that share the same recombination event (event 2 in Table 3). Recombination is known to reduce the accuracy of the phylogenetic signal (Posada & Crandall, 2002), and therefore recombinant isolates should be removed from the data set prior to phylogenetic analysis. However, SiMMV and ToYSV populations are both highly recombinant, with isolates showing recombination events throughout the entire genome, which prevented us from excluding recombinant isolates of the analysis.

Recombination has been extensively studied in begomoviruses as one of the main mechanisms contributing to the diversification of their populations (Lefeuvre *et al.*, 2007; Lima *et al.*, 2013; Monci *et al.*, 2002; Ramos-Sobrinho *et al.*, 2014; Rocha *et al.*, 2013). A complex recombination pattern was found for the SiMMV and ToYSV populations. In the DNA-A, most recombination events detected in ToYSV involved breakpoints located within the *Rep* gene and the IR-A, whereas in SiMMV the breakpoints were located mostly in the *Rep* and *CP* genes. Furthermore, in most recombination events involving the *Rep* gene, the breakpoints were located in the 5-terminal portion, and most of the recombination events involving the *CP* gene had

initial breakpoints located in the central portion of the gene and spanned its entire 3'-terminal portion. Recombination breakpoints are known to be non-randomly distributed in monopartite begomoviruses and in the DNA-A of bipartite begomoviruses, with hot spots in the *Rep* gene 5'-terminal portion, the 5'-end of the common region and the central portion of *CP* gene (Lefeuvre *et al.*, 2007; Prasanna & Rai, 2007).

For both populations, the DNA-B was shown to be more prone to recombination than the DNA-A. This could be explained by its lowest organizational complexity when compared to the DNA-A, as was demonstrated that tolerance to recombination in a given region of the genome is correlated with the maintenance of intragenome interactions (protein-protein/protein-DNA), which restrict recombination events that could deregulate these interactions (Martin *et al.*, 2005). Most of the recombination events in the DNA-B occurred in the LIR-B, and this non-coding region showed greater genetic variability than all the coding regions. For both populations, our results show a higher genetic variability in the *Rep* gene and the LIR-B, which are also the regions where we observed the largest number of recombination events.

Our results suggest that, as much as recombination partly explains the high genetic variability of the SiMMV and ToYSV populations, both populations are clearly dominated by mutational diversification, as observed by Lima *et al.* (2013) for MaYSV. However, these estimates should be interpreted as the minimal relative contribution of recombination, as some of the variation that is attributed to mutation is likely due to recombination that was not statistically detectable. Moreover, Lima *et al.* (2013) demonstrated that the relative contribution of recombination to the total variability does not necessarily correlate with the number of detectable

recombination events. Indeed, in the SiMMV CP tree, 41 substitutions were assigned to three recombination events but a unique event accounted for 37 substitutions. Our results indicate that the highest relative contribution of recombination for both populations was in the *CP* gene, which is also the least variable. Moreover, the DNA-B of both populations has greater genetic variability and a greater number of recombination events, but >93% of the nucleotide substitutions are attributed to mutation instead of recombination.

In this study we found a high degree of genetic variability for two populations of begomoviruses infecting two ubiquitous non-cultivated hosts. For both populations, the DNA-B was more permissible to genetic variability than the DNA-A. Phylogenetic analysis portrayed the clustering of isolates that shared the same recombination events, with no geographical patterns associated with these recombination events. Our results indicate that most of the genetic variability is due to mutations, although recombination may play a significant role for certain regions of the genome.

Methods

Sample collection and cloning of begomovirus genomes

Foliar samples from *Sida* spp. and *Leonurus sibiricus* showing symptoms of yellow mosaic and leaf distortion and/or infestation by *B. tabaci* were collected in the states of Rio Grande do Sul (RS), Paraná (PR) and Mato Grosso do Sul (MS) between 2009 and 2011 (these are the same samples described in Chapter 1). Total DNA extraction, RCA-based amplification and cloning of full-length viral genomes were performed as described in Chapter 1. Viral inserts were sequenced commercially (Macrogen Inc., Seoul, South Korea) by primer walking. All genome

sequences were organized to begin at the nicking site in the invariant nonanucleotide at the origin of replication (5'-TAATATT//AC-3').

Sequence comparisons and phylogenetic analysis

Sequences were initially analyzed with the BLAST*n* algorithm (Altschul *et al.*, 1990) to determine the viral species with which they shared greatest similarity. Pairwise comparisons between all DNA-A sequences and the begomoviruses with greatest similarity (as determined by the BLAST*n* search) were performed with the program SDT v. 1.2 (Muhire *et al.*, 2014) using the MUSCLE alignment option (Edgar, 2004). Species demarcation was based on a threshold of 91% nucleotide sequence identity for the full length DNA-A proposed by the *Geminiviridae* Study Group of the International Committee for Taxonomy of Viruses (ICTV) (Brown *et al.*, 2015).

Multiple sequence alignments were prepared for the full-length DNA-A and DNA-B using the MUSCLE alignment option in MEGA6 (Tamura *et al.*, 2013). DNA-A and DNA-B trees were constructed using Bayesian inference performed with MrBayes v. 3.0b4 (Ronquist & Huelsenbeck, 2003), with the nucleotide substitution model selected by MrModeltest v. 2.2 (Nylander, 2004) in the Akaike Information Criterion (AIC). The analyses were carried out running 10,000,000 generations, excluding the first 2,000,000 generations as burn in. Maximum likelihood (ML) trees were inferred for nucleotide sequences of the *CP*, *Rep*, *MP* and *NSP* genes using PAUP* v. 4.0 (Swofford, 2003). The best-fit model of nucleotide substitution was determined using MODELTEST (Posada & Crandall, 1998) by the Akaike Information Criterion. The heuristic ML search was initiated with a neighbour-joining tree, with optimization using the tree-bisection-reconnection

algorithm. The robustness of each individual branch was estimated from 2,000 nearest neighbour interchange bootstrap replicates. Trees were visualized using FigTree v. 1.3.1 (tree.bio.ed.ac.uk/software/figtree/).

Genetic structure and variability indices

Partitioning of genetic variability and inferences about population structure were based on Wright's F fixation index (Wright, 1951), calculated using DnaSP v. 5.10 (Rozas *et al.*, 2003). Descriptors of molecular variability [haplotype diversity (Hd) and average pairwise number of nucleotide differences per site (nucleotide diversity, π)] were estimated for each population/subpopulation using DnaSP v. 5.10.

Recombination analysis

Recombination analysis was performed using the Rdp, Geneconv, Bootscan, Maximum Chi Square, Chimaera, SisterScan and 3Seq methods implemented in Recombination Detection Program (RDP) v. 4 (Martin *et al.*, 2015). Alignments were scanned with default settings for the different methods. Statistical significance was inferred by *p*-values lower than a Bonferroni-corrected cut-off of 0.05. Only recombination events detected by at least four of the analysis methods available in the program were considered reliable.

Relative contribution of recombination and mutation for the genetic variability of ToYSV and SiMMV populations

To discriminate the relative contributions of mutation and recombination to the genetic variability of the SiMMV and ToYSV populations, we used the phylogeny-based method developed by Lima *et al.* (2013). Groups of sequences were

identified which descended from a shared recombination event (based on RDP4 analysis), and frequently these formed clades on midpoint-rooted CP and Rep ML trees. The sequence of the ancestral node of each of these clades reflected the recombination event. In the case of recombination events associated with only one sequence, the event was assigned to the branch leading to the corresponding tip. Using PAUP* (Swofford, 2003), all substitutions that occurred over each ML tree were identified and used to calculate η, the total number of substitutions over the phylogeny. We then looked at the location within the gene of substitutions that occurred on branches associated with recombination. Substitutions that occurred in the region likely introduced by recombination were added to $\eta_{recombination}$ (η_r), the total number of substitutions on the phylogeny likely due to recombination. All other substitutions, including those on branches associated with recombination but outside of the region implicated in the recombination event, were added to $\eta_{mutation}$ (η_m).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

Amarakoon, I. I., Roye, M. E., Briddon, R. W., Bedford, I. D. & Stanley, J. (2008). Molecular and biological characterization of *Macroptilium yellow mosaic virus* from Jamaica. *Plant Pathol* **57**, 417-426.

Ambrozevicius, L. P., Calegario, R. F., Fontes, E. P. B., Carvalho, M. G. & Zerbini, F. M. (2002). Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. *Fitopatol Bras* **27**, 372-377.

Andrade, E. C., Manhani, G. G., Alfenas, P. F., Calegario, R. F., Fontes, E. P. B. & Zerbini, F. M. (2006). *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*. *J Gen Virol* **87**, 3687-3696.

Barbosa, J. C., Eckstein, B., Bergamin-Filho, A., Rezende, J. A. M. & Dallagnol, L. J. (2012). First report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Brazil. *Plant Dis* **97**, 289-289.

Barreto, S. S., Hallwass, M., Aquino, O. M. & Inoue-Nagata, A. K. (2013). A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. *Phytopathology* **103**, 436-444.

Berrie, L. C., Rybicki, E. P. & Rey, M. E. C. (2001). Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. *J Gen Virol* **82**, 53-58.

Briddon, R. W. (2003). Cotton leaf curl disease, a multicomponent begomovirus complex. *Mol Plant Pathol* **4**, 427-434.

Briddon, R. W., Akbar, F., Iqbal, Z., Amrao, L., Amin, I., Saeed, M. & Mansoor, M. S. (2014). Effects of genetic changes to the begomovirus/betasatellite complex causing cotton leaf curl disease in South Asia post-resistance breaking. *Virus Res* **186**, 114-119.

Briddon, R. W., Patil, B. L., Bagewadi, B., Nawaz-ul-Rehman, M. S. & Fauquet, C. M. (2010). Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evol Biol* **10**, 97.

Brown, J. K., Fauquet, C. M., Briddon, R. W., Zerbini, F. M., Moriones, E. & Navas-Castillo, J. (2012). Family *Geminiviridae*. In *Virus Taxonomy Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 351-373. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. London, UK: Elsevier Academic Press.

Brown, J. K., Zerbini, F. M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J. C., Fiallo-Olivé, E., Briddon, R. W., Hernandez-Zepeda, C., Idris, A., Malathi, V. G., Martin, D. P., Rivera-Bustamante, R., Ueda, S. & Varsani, A. (2015). Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch Virol* **160**, 1593-1619.

Brustolini, O. J., Machado, J. P., Condori-Apfata, J. A., Coco, D., Deguchi, M., Loriato, V. A., Pereira, W. A., Alfenas-Zerbini, P., Zerbini, F. M., Inoue-Nagata, A. K., Santos, A. A., Chory, J., Silva, F. F. & Fontes, E. P. (2015). Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. *Plant Biotechnol J* **13**, 1300-1311.

Calegario, R. F., Ferreira, S. S., Andrade, E. C. & Zerbini, F. M. (2007). Characterization of *Tomato yellow spot virus*, (ToYSV), a novel tomato-infecting begomovirus from Brazil. *Pesq Agropec Bras* **42**, 1335-1343.

Castillo-Urquiza, G. P., Beserra Jr., J. E. A., Bruckner, F. P., Lima, A. T. M., Varsani, A., Alfenas-Zerbini, P. & Zerbini, F. M. (2008). Six novel begomoviruses

infecting tomato and associated weeds in Southeastern Brazil. *Arch Virol* **153**, 1985-1989.

Celli, M. G., Perotto, M. C., Martino, J. A., Flores, C. R., Conci, V. C. & Pardina, P. R. (2014). Detection and identification of the first viruses in chia (Salvia hispanica). *Viruses* **6**, 3450-3457.

Duffy, S. & Holmes, E. C. (2008). Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. *J Virol* **82**, 957-965.

Duffy, S. & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *J Gen Virol* **90**, 1539-1547.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinf* **5**, 1-19.

Fernandes, F. R., Albuquerque, L. C., Oliveira, C. L., Cruz, A. R. R., Rocha, W. B., Pereira, T. G., Naito, F. Y. B., Dias, N. D., Nagata, T., Faria, J. C., Zerbini, F. M., Aragão, F. J. L. & Inoue-Nagata, A. K. (2011). Molecular and biological characterization of a new Brazilian begomovirus, euphorbia yellow mosaic virus (EuYMV), infecting *Euphorbia heterophylla* plants. *Arch Virol* **156**, 2063-2069.

Fernandes, F. R., Cruz, A. R. R., Faria, J. C., Zerbini, F. M. & Aragão, F. J. L. (2009). Three distinct begomoviruses associated with soybean in central Brazil. *Arch Virol* **154**, 1567-1570.

Fernandes, N. A. N., Boiteux, L. S., Fonseca, M. E. N., Gonzales-Segnana, L. & Kitajima, E. W. (2014). Report of *Tomato yellow spot virus* infecting *Leonurus sibiricus* in Paraguay and within tomato fields in Brazil. *Plant Dis* **98**, 1445.

Fiallo-Olivé, E., Navas-Castillo, J., Moriones, E. & Martinez-Zubiaur, Y. (2012). Begomoviruses infecting weeds in Cuba: Increased host range and a novel virus infecting *Sida rhombifolia*. *Arch Virol* **157**, 141-146.

Fondong, V. N., Pita, J. S., Rey, M. E. C., Kochko, A., Beachy, R. N. & Fauquet, C. M. (2000). Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *J Gen Virol* **81**, 287-297.

Frischmuth, T., Engel, M., Lauster, S. & Jeske, H. (1997). Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted, *Sida*-infecting bipartite geminiviruses in Central America. *J Gen Virol* **78**, 2675-2682.

García-Andrés, S., Tomas, D. M., Sanchez-Campos, S., Navas-Castillo, J. & Moriones, E. (2007). Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. *Virology* **365**, 210-219.

García-Arenal, F., Fraile, A. & Malpica, J. M. (2001). Variability and genetic structure of plant virus populations. *Annu Rev Phytopathol* **39**, 157-186.

García-Arenal, F., Fraile, A. & Malpica, J. M. (2003). Variation and evolution of plant virus populations. *Int Microbiol* **6**, 225-232.

Garrido-Ramirez, E. R., Sudarshana, M. & Gilbertson, R. L. (2000). *Bean golden yellow mosaic virus* from Chiapas, Mexico: Characterization, pseudorecombination with other bean-infecting geminiviruses and germ plasm screening. *Phytopathology* **90**, 1224-1232.

Gilbertson, R. L., Batuman, O., Webster, C. G. & Adkins, S. (2015). Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. *Annu Rev Virol* **2**, 67-93.

Gilbertson, R. L., Hidayat, S. H., Paplomatas, E. J., Rojas, M. R., Hou, Y.-H. & Maxwell, D. P. (1993). Pseudorecombination between infectious cloned DNA components of tomato mottle and bean dwarf mosaic geminiviruses. *J Gen Virol* **74**, 23-31.

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D. & Mansoor, S. (2013). Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nat Rev Microbiol* **11**, 777-788.

Idris, A. M., Hiebert, E., Bird, J. & Brown, J. K. (2003). Two newly described begomoviruses of *Macroptilium lathyroides* and common bean. *Phytopathology* **93**, 774-783.

Jovel, J., Reski, G., Rothenstein, D., Ringel, M., Frischmuth, T. & Jeske, H. (2004). *Sida micrantha* mosaic is associated with a complex infection of begomoviruses different from *Abutilon mosaic virus*. *Arch Virol* **149**, 829-841.

Lefeuvre, P., Lett, J. M., Varsani, A. & Martin, D. P. (2009). Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol* **83**, 2697-2707.

Lefeuvre, P., Martin, D. P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., Becker, N., Reynaud, B. & Lett, J. M. (2007). Begomovirus 'melting pot' in the south-west Indian Ocean islands: Molecular diversity and evolution through recombination. *J Gen Virol* **88**, 3458-3468.

Lefeuvre, P. & Moriones, E. (2015). Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. *Curr Opin Virol* **10**, 14-19.

Legg, J. & Fauquet, C. (2004). Cassava mosaic geminiviruses in Africa. *Plant Mol Biol* **56**, 585-599.

Legg, J. P. & Thresh, J. M. (2000). Cassava mosaic virus disease in East Africa: A dynamic disease in a changing environment. *Virus Res* **71**, 135-149.

Lima, A. T. M., Sobrinho, R. R., Gonzalez-Aguilera, J., Rocha, C. S., Silva, S. J. C., Xavier, C. A. D., Silva, F. N., Duffy, S. & Zerbini, F. M. (2013). Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. *J Gen Virol* **94**, 418-431.

Mahajan, N., Parameswari, C. & Veluthambi, K. (2011). Severe stunting in blackgram caused by the Mungbean yellow mosaic virus (MYMV) KA27 DNA B component is ameliorated by co-infection or post-infection with the KA22 DNA B: MYMV nuclear shuttle protein is the symptom determinant. *Virus Res* **157**, 25-34.

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P. & Varsani, A. (2011). Recombination in eukaryotic single stranded DNA viruses. *Viruses* **3**, 1699-1738.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* **1**, vev003.

Martin, D. P., van der Walt, E., Posada, D. & Rybicki, E. P. (2005). The evolutionary value of recombination is constrained by genome modularity. *PLoS Genet* **1**, e51.

Monci, F., Sanchez-Campos, S., Navas-Castillo, J. & Moriones, E. (2002). 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* **303**, 317-326.

Morales, F. J. (2006). History and current distribution of begomoviruses in Latin America. *Adv Virus Res* **67**, 127-162.

Muhire, B. M., Varsani, A. & Martin, D. P. (2014). SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* **9**, e108277.

Navas-Castillo, J., Fiallo-Olivé, E. & Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annu Rev Phytopathol* **49**, 219-248.

Nawaz-Ul-Rehman, M. S. & Fauquet, C. M. (2009). Evolution of geminiviruses and their satellites. *Febs Let* **583**, 1825-1832.

Nylander, J. A. A. (2004). MrModeltest v2. In *Program distributed by the author Evolutionary Biology Centre, Uppsala University*.

Owor, B. E., Martin, D. P., Shepherd, D. N., Edema, R., Monjane, A. L., Rybicki, E. P., Thomson, J. A. & Varsani, A. (2007). Genetic analysis of *Maize streak virus* isolates from Uganda reveals widespread distribution of a recombinant variant. *J Gen Virol* **88**, 3154-3165.

Padidam, M., Sawyer, S. & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218-224.

Paximadis, M., Idris, A. M., Torres-Jerez, I., Villarreal, A., Rey, M. E. C. & Brown, J. K. (1999). Characterization of tobacco geminiviruses in the Old and New world. *Arch Virol* **144**, 703-717.

Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S. & Fauquet, C. M. (2001). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J Gen Virol* **82**, 655-665.

Polston, J. E. & Anderson, P. K. (1997). The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. *Plant Dis* **81**, 1358-1369.

Posada, D. & Crandall, K. A. (1998). MODELTEST: Testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Posada, D. & Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogeny estimation. *J Mol Evol* **54**, 396-402.

Prasanna, H. C. & Rai, M. (2007). Detection and frequency of recombination in tomato-infecting begomoviruses of South and Southeast Asia. *Virol J* **4**, 111.

Ramos-Sobrinho, R., Xavier, C. A. D., Pereira, H. M. B., Lima, G. S. A., Assunção, I. P., Mizubuti, E. S. G., Duffy, S. & Zerbini, F. M. (2014). Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. *J Gen Virol* **95**, 2540-2552.

Rocha, C. S., Castillo-Urquiza, G. P., Lima, A. T. M., Silva, F. N., Xavier, C. A. D., Hora-Junior, B. T., Beserra-Junior, J. E. A., Malta, A. W. O., Martin, D. P., Varsani, A., Alfenas-Zerbini, P., Mizubuti, E. S. G. & Zerbini, F. M. (2013). Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. *J Virol* **87**, 5784-5799.

Rodríguez-Pardina, P. E., Hanada, K., Laguna, I. G., Zerbini, F. M. & Ducasse, D. A. (2011). Molecular characterisation and relative incidence of bean- and soybean-infecting begomoviruses in northwestern Argentina. *Ann Appl Biol* **158**, 69-78.

Rojas, M. R., Hagen, C., Lucas, W. J. & Gilbertson, R. L. (2005). Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. *Annu Rev Phytopathol* **43**, 361-394.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Roossinck, M. J. (1997). Mechanisms of plant virus evolution. *Annu Rev Phytopathol* **35**, 191-209.

Roye, M. E., McLaughlin, W. A., Nakhla, M. K. & Maxwell, D. P. (1997). Genetic diversity among geminiviruses associated with the weed species *Sida* spp., *Macroptilium lathyroides*, and *Wissadula amplissima* from Jamaica. *Plant Dis* **81**, 1251-1258.

Rozas, J., Sánchez-Del Barrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP: DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Rybicki, E. P. (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch Virol* **139**, 49-77.

Silva, S. J. C., Castillo-Urquiza, G. P., Hora-Junior, B. T., Assunção, I. P., Lima, G. S. A., Pio-Ribeiro, G., Mizubuti, E. S. G. & Zerbini, F. M. (2012). Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. *Plant Pathol* **61**, 457-467.

Swofford, D. L. (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725-2729.

Tavares, S. S., Ramos-Sobrinho, R., Gonzalez-Aguilera, J., Lima, G. S. A., Assunção, I. P. & Zerbini, F. M. (2012). Further molecular characterization of weed-associated

begomoviruses in Brazil with an emphasis on *Sida* spp. *Planta Daninha* **30**, 305-315.

Varsani, A., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Brown, J. K., Zerbini, F. M. & Martin, D. P. (2014). Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. *Arch Virol* **159**, 2193-2203.

Wright, S. (1951). The genetical structure of populations. *Ann Eugen* **15**, 323-354.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J. & Harrison, B. D. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J Gen Virol* **78**, 2101-2111.

Table 1. Genetic variability based on the DNA-A of the begomoviruses *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV).

Population	No. of sequences	H[1]	Hd	DNA-A π	Rep π	CP π	IR-A π
SiMMV (Total)	25	23	0.99 (± 0.02)	0.059 (± 0.00433)	0.073 (± 0.00470)	0.033 (± 0.00435)	0.074 (± 0.00618)
Bolívia	4	4	1.00 (± 0.18)	0.061 (± 0.01455)	0.065 (± 0.01815)	0.043 (± 0.00761)	0.082 (± 0.02140)
Goiás	6	4	0.80 (± 0.17)	0.051 (± 0.01027)	0.064 (± 0.01223)	0.019 (± 0.00714)	0.071 (± 0.01353)
Paraná	4	4	1.00 (± 0.18)	0.017 (± 0.00445)	0.019 (± 0.00674)	0.015 (± 0.00289)	0.028 (± 0.00664)
Mato Grosso do Sul	2	2	1.00 (± 0.50)	0.023 (± 0.01177)	0.026 (± 0.01300)	0.020 (± 0.01022)	0.041 (± 0.02049)
Rio de Janeiro	3	3	1.00 (± 0.27)	0.034	0.031 (± 0.01006)	0.045 (± 0.01859)	0.024 (± 0.00760)
Rio Grande do Sul	6	6	1.00 (± 0.09)	0.026 (± 0.00404)	0.029 (± 0.00648)	0.018 (± 0.00213)	0.041 (± 0.00591)
ToYSV (Total)	23	23	1.00 (± 0.01)	0.042 (± 0.00380)	0.055 (± 0.00606)	0.022 (± 0.00178)	0.070 (± 0.00628)
Argentina	2	2	1.00 (± 0.50)	0.045 (± 0.02264)	0.072 (± 0.03621)	0.013 (± 0.00661)	0.070
Paraná	16	16	1.00 (± 0.02)	0.039	0.050 (± 0.00475)	0.020 (± 0.00148)	0.069 (± 0.00535)
Mato Grosso do Sul	3	3	1.00 (± 0.27)	0.0085(± 0.00236)	0.009 (± 0.00254)	0.005 (± 0.00212)	0.012 (± 0.00454)
Minas Gerais	2	2	1.00 (± 0.50)	0.003 (± 0.00131)	0.002 (± 0.00093)	0.001 (± 0.00066)	0.005 (± 0.00259)

[1]H, number of haplotypes; Hd, haplotype diversity; π, average pairwise number of nucleotide differences per site (nucleotide diversity)

Table 2. Genetic variability based on the DNA-B of the begomoviruses *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV).

Population	No. of sequences	H[1]	Hd	DNA-B π	NSP π	MP π	LIR π	SIR π
SiMMV (Total)	30	30	1.00 (± 0.009)	0.103 (± 0.00407)	0.097 (± 0.00584)	0.068 (± 0.00245)	0.136 (± 0.00580)	0.208 (± 0.00908)
Bolívia	4	4	1.00 (± 0.177)	0.106 (± 0.02141)	0.102 (± 0.01980)	0.074 (± 0.01359)	0.134 (± 0.03309)	0.135 (± 0.02623)
Goiás	2	2	1.00 (± 0.500)	0.043 (± 0.02132)	0.041 (± 0.02075)	0.0078(± 0.00397)	0.078 (± 0.03913)	0.026 (± 0.01282)
Paraná	6	6	1.00 (± 0.096)	0.055 (± 0.00622)	0.050 (± 0.00636)	0.033 (± 0.00449)	0.074 (± 0.00876)	0.112 (± 0.01865)
Mato Grosso do Sul	4	4	1.00 (± 0.177)	0.102 (± 0.01884)	0.092 (± 0.01742)	0.067 (± 0.01209)	0.136 (± 0.02531)	0.200 (± 0.04050)
Rio de Janeiro	3	3	1.00 (± 0.272)	0.064 (± 0.02027)	0.056 (± 0.02064)	0.034 (± 0.00935)	0.085 (± 0.02852)	0.211 (± 0.06621)
Rio Grande do Sul	11	11	1.00 (± 0.039)	0.088 (± 0.00479)	0.069 (± 0.00495)	0.059 (± 0.00386)	0.122 (± 0.00706)	0.196 (± 0.01644)
ToYSV (Total)	20	20	1.00 (± 0.014)	0.057 (± 0.00317)	0.052 (± 0.00377)	0.034 (± 0.00181)	0.084 (± 0.00571)	0.123 (± 0.00893)
Paraná	17	17	1.00 (± 0.020)	0.056 (± 0.00298)	0.047 (± 0.00359)	0.033 (± 0.00182)	0.082 (± 0.00603)	0.139 (± 0.00997)
Mato Grosso do Sul	3	3	1.00 (± 0.272)	0.046 (± 0.01326)	0.044 (± 0.01407)	0.030 (± 0.01168)	0.062 (± 0.01779)	0.115 (± 0.04474)

[1]H, number of haplotypes; Hd, haplotype diversity; π, average pairwise number of nucleotide differences per site (nucleotide diversity)

Table 3. Recombination events detected in the DNA-A and DNA-B of *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV) populations.

Event	Recombinant	Recombination breakpoints[1]		Parents[2]		Method	*P-value*[3]
		Begin	End	Minor	Major		
ToYSV(DNA-A)							
1	KJ742419	2015	2554	JX513952	FJ538207	RGBMCS**3**	2.645 X 10^{-19}
2	JX863082	2327	2636	JX513952	BR:Dou1095:11	RGMC**S**3	6.016 X 10^{-12}
	JX863081						
	JQ429791						
	BR:Egb739:10						
	BR:Cam784:10						
3	JX513952	2135	2530	Unknown	BR:Lac703:10	RGMC**S**3	3.432 X 10^{-18}
	BR:Pab1058:11						
	KJ742419						
4	JX513952	254	1401	BR:Lac703:10	BR:Pab1058:11	RGMCS**3**	4.041 X 10^{-09}
5	BR:Ara1130:11	2616	754	Unknown	BR:Tol673:10	RMC**S**	3.693 X 10^{-07}
	BR:Caf776:10						
	BR:Ser1135:11						
ToYSV(DNA-B)							
1	BR:Dou1095:11	1511	44	BR:Dou697:10	BR:Caf776:10	RGBMCSP**3**	1.493 X 10^{-19}
2	BR:Cas775:10	398	1305	Unknown	BR:Gua1083:11	RGBMCS3	8.897 X 10^{-12}
	BR:Dou697:10						
	BR:Jan735:10						

3	BR:Arm1111:11	1745	2471	Unknown	BR:Gua1083:11	**M**CSP3	3.200 X 10^{-08}
4	JX513953	2445	122	BR:Rea661:10	BR:Sad796:10	RGBMCS3	5.664 X 10^{-08}
5	BR:Tol1075:11 BR:Ubi779:10 BR:Arm1111:11	(?) 2522	487	BR:Cam784:10	BR:Rea661:10	BMC**S**3	1.311 X 10^{-05}
6	BR:Rea661:10 BR:Gua1083:11	(?)399	(?)52	BR:Dou697:10	BR:Nom1077:11	MCS**3**	2.001 X 10^{-03}
7	BR:Cam784:10 BR:Dou1095:11 BR:Nom1077:11 BR:Caf776:10	(?)2133	(?)1315	BR:Cas775:10	Unknown	MC**S**3	8.259 X 10^{-04}
SiMMV (DNA-A)							
1	JX415194 JX415195 JX415187	2676	2129	EU908733	Unknown	RGMCS**3**	4.902 X 10^{-17}
2	BR:Smm65:09 BR:Smm876:10 BR:Rea662:10 BR:Umu1121:11 BR:Saa48:09 BR:Pan115:09 BR:Cha822:10 BR:Cha926:11 BR:Mar1142:11 BR:Umu732:10 HM585433	633	1460	HM357459	Unknown	RGMCS**3**	7.927 X 10^{-11}

	FN436005						
	FN436003						
3	HM357459	652	1027	KC706535	KC706536	RGMCS3	5.802 X 10^{-12}
4	KC706536	678	989	Unknown	KC706535	**R**GMCS3	9.054 X 10^{-10}
5	BR:Smm876:10	(?)1872	2241	HM357459	BR:Rea662:10	RMC**S**3	3.083 X 10^{-16}
	BR:Umu1121:11						
	BR:Saa48:09						
	BR:Smm65:09						
	BR:Pan115:09						
	BR:Cha822:10						
	BR:Cha926:11						
	BR:Mar1142:11						
	BR:Umu732:10						
6	HM585439	52	(?)1600	Unknown	KC706537	**R**M**S**3	2.549 X 10^{-06}
	HM585437						
7	FJ686693	1616	1899	Unknown	EU908733	**R**GBMC3	4.626 X 10^{-05}
SiMMV (DNA-B)							
1	BR:Sam69:09	2492	118	BR:Tap895:10	BR:Rea662:10	**R**GBMC3	1.254 X 10^{-15}
	BR:Tac1116:11						
	BR:Smm949:11						
2	KC706534	2544	2644	Unknown	KC706532	R**G**BMCS3	2.359 X 10^{-10}
	BR:Tap895:10						
	BR:Dou698:10						
	BR:Lac704:10						
	BR:Cra556:10						
	BR:Sam69:09						

	BR:Smm547:10						
3	EU908734	2585	697	BR:Lac704:10	FJ686694	RB**M**C3	4.655 X 10^{-11}
4	BR:Cha822:10	2550	211	BR:Rea662:10	Unknown	**R**GBMCS	7.461 X 10^{-07}
	BR:Pan115:09						
5	BR:Saa832:10	781	1389	BR:Gua1120:11	BR:Pan115:09	RGBM**C**S	5.753 X 10^{-05}
6	KC706532	1023	1371	Unknown	BR:Lac704:10	RGB**M**CS3	2.723 X 10^{-06}
	KC706534						
	KC706533						
7	BR:Saa832:10	2079	(?)2541	BR:Rea662:10	BR:Pan115:09	RBMS**3**	9.479 X 10^{-08}
8	HM585440	231	843	Unknown	FN436006	RBMC**S**	2.967 X 10^{-14}
	HM585438						
	HM585432						
9	EU908734	(?)863	1440	Unknown	FN436006	RGBMC**S**3	3.783 X 10^{-07}
	BR:Lac704:10						
	KC706534						
	KC706533						
	FJ686694						
10	BR:Sad799:10	856	1526	FN436006	BR:Rea662:10	GB**M**C	2.141 X 10^{-06}
11	BR:Gua1120:11	796	1002	BR:Cha926:11	BR:Mar755:10	RGBMC**S**3	1.725 X 10^{-05}
	BR:Saa832:10						
12	HM585434	254	521	Unknown	FN436006	RMC3	1.862 X 10^{-04}
	BR:Smm876:10						
13	BR:Smm876:10	1815	2133	BR:Mar755:10	BR:Pan115:09	RG**B**MC	9.802 X 10^{-05}
14	HM585434	(?) 1430	(?) 2531	Unknown	BR:Dou698:10	BM**S**3	1.484 X 10^{-05}
	BR:Smm547:10						

15	FN436006 BR:Tac1116:11	(?)2586	229	BR:Cha926:11	EU908734	RG**C**3	4.145 X 10^{-04}

[1]Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increase clockwise. (?), Breakpoints could not be accurately located.
[2]Recombination events and their putative parental viruses were identified using the Rdp (R), Geneconv (G), Boostcan (B), Maxichi (M), Chimaera (C), Siscan (S) and 3Seq (3), modules in RDP4.
[3]The reported *p*-value is from the method in bold and is the lowest *p*-value calculated for the featured event.

Table 4. Results of subdivision tests performed for populations of *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV).

Population	F_{st}[1]	
	DNA-A	**DNA-B**
SiMMV		
Mato Grosso do Sul/Bolívia	0.41	0.19
Mato Grosso do Sul/Goiás	0.52	0.32
Mato Grosso do Sul/Paraná	0.36	0.17
Mato Grosso do Sul/Rio de Janeiro	0.54	0.21
Mato Grosso do Sul/Rio Grande do Sul	0.33	**0.04**
Goiás/Rio de Janeiro	0.24	0.27
Goiás/Bolívia	0.31	0.43
Goiás/Paraná	0.55	0.58
Goiás/Rio Grande do Sul	0.51	0.45
Rio de Janeiro/Bolívia	0.28	0.37
Rio de Janeiro/Paraná	0.59	0.50
Rio de Janeiro/Rio Grande do Sul	0.53	0.36
Bolívia/Paraná	0.45	0.36
Bolívia/Rio Grande do Sul	0.41	0.23
Paraná/Rio Grande do Sul	**0.03**	0.20
Total	0.43	0.32
ToYSV		
Argentina/Paraná	0.36	
Argentina/Minas Gerais	0.68	
Argentina/Mato Grosso do Sul	0.62	
Paraná/Minas Gerais	0.61	
Paraná/Mato Grosso do Sul	0.26	**0.04**
Minas Gerais/Mato Grosso do Sul	0.89	
Total	0.59	**0.04**

[1]Values from 0 to 0.05 indicate little genetic diferentiation; from 0.05 to 0.15, moderate differentiation; from 0.15 to 0.25, great differentiation; >0.25, high differentiation. Values highlighted in bold correspond to subpopulations that have low to moderate genetic differentiation between them and therefore are considered to comprise a single population.

Table 5. Relative contribution of mutation and recombination to the genetic variability of *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV) populations.

Gene	SiMMV			ToYSV		
	η_μ	η_r	η_{total}	η_μ	η_r	η_{total}
CP	211 (84%)	41 (16%)	252	121 (95%)	6 (5%)	127
Rep	483 (85%)	82 (15%)	565	390 (99%)	3 (1%)	393
MP	675 (98%)	17 (2%)	692	270 (95%)	13 (5%)	283
NSP	759 (93%)	60 (7%)	819	309 (97%)	11 (3%)	320

Supplementary Table S1. Begomovirus sequences reported in this study.

Sample code	Sampling date	Location	Geographical coordinates		Host	Enzyme[1]		Isolate name	GenBank access number
						DNA-A	DNA-B		
Sida micrantha mosaic virus (SiMMV)									
CF48	03/19/09	Santo Ângelo, RS	S28 22' 54.70"	W54 18' 17.24"	*Sida* sp.	*Eco*RI		BR:Saa48:09	
CF65	03/20/09	São Miguel das Missões, RS	S28 29' 35.59"	W54 33' 37.15"	*Sida* sp.	*Spe*I		BR:Smm65:09	
CF69	03/20/09	Santo Antônio das Missões, RS	S28 29' 42.70"	W55 25' 18.00"	*Sida* sp.		*Spe*I	BR:Sam69:09	
CF115	03/20/09	Panambi, RS	S28 18' 24.45"	W53 29' 21.38"	*Sida* sp.	*Eco*RI	*Spe*I	BR:Pan115:09	
CF547	03/24/10	São Miguel das Missões, RS	S28 29' 35.59"	W54 33' 37.15"	*Sida* sp.		*Spe*I	BR:Smm547:10	
CF556	03/25/10	Cruz Alta, RS	S28 36' 10.89"	W53 39' 25.34"	*Sida* sp.		*Sal*I	BR:Cra556:10	
CF662	03/25/10	Realeza, PR	S25 40' 45.00"	W53 33' 09.00"	*Sida* sp.	*Eco*RI	*Spe*I	BR:Rea662:10	
CF679	06/08/10	Marechal Cândido Rondon, PR	S24 30' 59.50"	W54 04' 37.20"	*Sida* sp.		*Apa*I	BR:Mcr679:10	
CF698	06/09/10	Dourados, MS	S22 17' 58.00"	W54 49' 14.10"	*Sida* sp.		*Hin*dIII	BR:Dou698:10	
CF704	06/09/10	Laguna Carapã, MS	S22 25' 35.60"	W55 21' 30.00"	*Sida* sp.		*Sal*I	BR:Lac704:10	
CF732	06/10/10	Umuarama, PR	S23 50' 12.30"	W53 17' 23.60"	*Sida* sp.	*Eco*RI		BR:Umu732:10	
CF755	06/10/10	Marialva, PR	S23 30' 05.00"	W51 47' 08.00"	*Sida* sp.		*Sal*I	BR:Mar755:10	
CF799	08/11/10	São Domingos, PR	S24 00' 51.00"	W51 30' 37.00"	*Sida* sp.		*Apa*I	BR:Sad799:10	
CF822	24/08/10	Chapada, RS	S28 01' 08.80"	W53 05' 52.31"	*Sida* sp.	*Spe*I	*Spe*I	BR:Cha822:10	
CF832	25/08/10	Santo Ângelo, RS	S28 22' 54.70"	W54 18' 17.23"	*Sida* sp.		*Spe*I	BR:Saa832:10	
CF876	10/06/10	São Miguel das Missões, RS	S28 23' 59.00"	W54 39' 52.00"	*Sida* sp.	*Sac*I	*Spe*I	BR:Smm876:10	
CF895	10/07/10	Tapatuba, RS	S29 03' 57.00"	W54 43' 53.00"	*Sida* sp.		*Sal*I	BR:Tap895:10	
CF926	03/15/11	Chapada, RS	S28 01' 08.80"	W53 05' 52.31"	*Sida* sp.	*Spe*I	*Spe*I	BR:Cha926:11	
CF949	03/16/11	São Miguel Das Missões, RS	S28 29' 35.59"	W54 33' 37.15"	*Sida* sp.		*Sal*I	BR:Smm949:11	
CF1116	06/08/11	Tacuru, RS	S23 38' 17.60"	W54 58' 17.70"	*Sida* sp.		*Apa*I	BR:Tac1116:11	

CF1120	06/09/11	Guaíra, PR	S24 04' 57.00"	W54 10' 04.00"	*Sida* sp.		*Apa*I	BR:Gua1120:11
CF1121	06/09/11	Umuarama, PR	S23 50' 12.30"	W53 17' 23.60"	*Sida* sp.	*Spe*I		BR:Umu1121:11
CF1128	06/09/11	Janiópolis, PR	S24 08' 02.00"	W52 47' 22.00"	*Sida* sp.		*Bam*HI	BR:Jan1128:11
CF1142	06/09/11	Marialva, PR	S23 30' 09.40"	W51 47' 42.90"	*Sida* sp.	*Eco*RI		BR:Mar1142:11
Tomato yellow spot virus (ToYSV)								
CF661	06/07/10	Realeza, PR	S25 40' 45.00"	W53 33' 09.00"	*Leonurus sibiricus*	*Spe*I	*Kpn*I	BR:Rea661:10
CF673	06/08/10	Toledo, PR	S24 46' 45.00"	W53 40' 41.00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Tol673:10
CF677	06/08/10	Marechal Cândido Rondon, PR	S24 30' 59.50"	W54 04' 37.20"	*L. sibiricus*	*Apa*I		BR:Mcr677:10
CF697	06/09/10	Dourados, MS	S22 17' 58.00"	W54 49' 14.10"	*L. sibiricus*	*Apa*I	*Hin*dIII	BR:Dou697:10
CF703	06/09/10	Laguna Carapã, MS	S22 25' 35.60"	W55 21' 30.00"	*L. sibiricus*	*Apa*I		BR:Lac703:10
CF735	06/10/10	Janiópolis, PR	S24 08' 02.00"	W52 47' 22.00"	*L. sibiricus*		*Apa*I	BR:Jan735:10
CF739	06/10/10	Engenheiro Beltrão, PR	S23 42' 12.60"	W52 08' 54.60"	*L. sibiricus*	*Sac*I	*Hin*dIII	BR:Egb739:10
CF775	08/10/10	Cascavel, PR	S24 52' 06.00"	W53 20' 28.00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Cas775:10
CF776	08/10/10	Cafelândia, PR	S24 39' 21.30"	W53 13' 07.40"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Caf776:10
CF779	08/10/10	Ubiratã, PR	S24 33' 07.00"	W53 00' 34.00"	*L. sibiricus*		*Hin*dIII	BR:Ubi779:10
CF784	08/10/10	Campo Mourão, PR	S23 57' 14.90"	W52 20' 59.30"	*L. sibiricus*	*Sac*I	*Kpn*I	BR:Cam784:10
CF793	08/11/10	Londrina, PR	S23 26' 20.30"	W51 08' 17.90"	*L. sibiricus*		*Apa*I	BR:Lon793:10
CF796	08/11/10	São Domingos, PR	S24 00' 51.00"	W51 30' 37.00"	*L. sibiricus*		*Kpn*I	BR:Sad796:10
CF802	08/11/10	Ivaiporã, PR	S24 18' 37.00"	W51 43' 24.00"	*L. sibiricus*		*Apa*I	BR:Iva802:10
CF1024	26/04/11	Francisco Beltrão, PR	S26 00' 09.00"	W52 56' 36.00"	*L. sibiricus*		*Apa*I	BR:Frb1024:11
CF1058	06/06/11	Pato Branco, PR	S26 11' 47.44"	W52 49' 23.87"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Pab1058:11
CF1067	06/06/11	Ampere, PR	S25 57' 06.90"	W53 24' 25.00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Amp1067:11
CF1075	06/07/11	Toledo, PR	S24 46' 46.00"	W53 40' 41.00"	*L. sibiricus*		*Apa*I	BR:Tol1075:11
CF1077	06/07/11	Nova Mercedes, PR	S24 30' 54.70"	W54 07' 0.22"	*L. sibiricus*		*Apa*I	BR:Nom1077:11
CF1083	06/07/11	Guaíra, PR	S24 14' 04.00"	W54 11' 57.00"	*L. sibiricus*	*Spe*I	*Apa*I	BR:Gua1083:11
CF1095	06/08/11	Dourados, MS	S22 17' 58.00"	W54 49' 14.10"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Dou1095:11
CF1111	06/08/11	Aral Moreira, MS	S22 46' 19.70"	W55 24' 43.90"	*L. sibiricus*	*Apa*I	*Hin*dIII	BR:Arm1111:11

CF1130	06/09/11	Araruna, PR	S24 03' 50.00"	W52 33' 52.00"	*L. sibiricus*	*Apa*I		BR:Ara1130:11
CF1135	06/09/11	Sertanópolis, PR	S23 50' 25.00"	W52 18' 19.00"	*L. sibiricus*	*Apa*I	*Apa*I	BR:Ser1135:11

[1]Enzyme used for releasing genome-length DNA components after rolling-circle amplification and subsequent cloning into the plasmid vector pKS+.

Supplementary Table S2. Begomovirus sequences retrieved from GenBank.

Species	Acronym	Host	Access number		Location
			DNA- A	DNA-B	
Tomato yellow spot virus	ToYSV	Bean and soybean	FJ538207		Argentina
Tomato yellow spot virus	ToYSV	Chia *(Salvia hispanica)*	KJ742419	KJ742420	Argentina
Tomato yellow spot virus	ToYSV	Tomato	DQ336350	DQ336351	Minas Gerais
Tomato yellow spot virus	ToYSV	Tomato	KC706628		Minas Gerais
Tomato yellow spot virus	ToYSV	*Leonurus sibiricus*	KC683374		Paraguay
Tomato yellow spot virus	ToYSV	*L. sibiricus*	JX513952	JX513953	Paraná
Tomato yellow spot virus	ToYSV	*L. sibiricus*	JX863082		Paraná
Tomato yellow spot virus	ToYSV	*L. sibiricus*	JX863081		Paraná
Tomato yellow spot virus	ToYSV	*L. sibiricus*	JQ429791		Paraná
Sida micrantha mosaic virus	SiMMV	*Sida rhombifolia*	HM585439	HM585440	Bolivia
Sida micrantha mosaic virus	SiMMV	*S. rhombifolia*	HM585437	HM585438	Bolivia
Sida micrantha mosaic virus	SiMMV	*S. micrantha*	HM585433	HM585434	Bolivia
Sida micrantha mosaic virus	SiMMV	*S. rhombifolia*	HM585431	HM585432	Bolivia
Sida micrantha mosaic virus	SiMMV	*S. santaremnensis*	JX415195		Goiás
Sida micrantha mosaic virus	SiMMV	*S. santaremnensis*	JX415194		Goiás
Sida micrantha mosaic virus	SiMMV	*S. santaremnensis*	JX415187		Goiás
Sida micrantha mosaic virus	SiMMV	Tomato	KC706537	KC706534	Rio de Janeiro
Sida micrantha mosaic virus	SiMMV	Tomato	KC706536	KC706533	Rio de Janeiro
Sida micrantha mosaic virus	SiMMV	Tomato	KC706535	KC706532	Rio de Janeiro
Sida micrantha mosaic virus	SiMMV	*S. micrantha*	FN436005	FN436006	Mato Grosso do Sul
Sida micrantha mosaic virus	SiMMV	*S. rhombifolia*	FN436003	FN436004	Mato Grosso do Sul
Sida micrantha mosaic virus	SiMMV	Soybean	FJ686693	FJ686694	Goiás
Sida micrantha mosaic virus	SiMMV	Okra	EU908733	EU908734	Goiás
Sida micrantha mosaic virus	SiMMV	Bean	HM357459		Goiás

Figure legends

Figure 1. Pairwise nucleotide sequence identity matrices of the DNA-A **(A)** and DNA-B **(B)** of *Sida micrantha mosaic virus* (SiMMV) and *Tomato yellow spot virus* (ToYSV) isolates obtained from *Sida* spp. (red) and *Leonurus sibiricus* (green) plants in this study, with additional begomovirus sequences retrieved from Genbank.

Figure 2. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (**A**) and DNA-B (**B**) nucleotide sequences of *Sida micrantha mosaic virus* (SiMMV) isolates. Nodes with posterior probability values between 0.50 and 0.89 are indicated by empty circles, and nodes with values equal or greater than 0.90 are indicated by filled circles. Colors indicate sampling locations.

Figure 3. Midpoint-rooted maximum likelihood trees based on the nucleotide sequences of the *CP* (**A**) and *Rep* (**B**) genes of *Sida micrantha mosaic virus* (SiMMV) isolates. Nodes with bootstrap values equal or higher than 80% are indicated by filled circles, and those with values lower than 80 and higher 50% by empty circles. The unique recombination events detected within the *CP* (events 3, 4 and 6 in Table 3, indicated in red, blue and pink, respectively) and *Rep* genes (events 1, 5, 6 and 7 in Table 3, indicated in yellow, orange, pink and purple, respectively) are shown as diagrams close to the branches where the substitutions due to each recombination event were mapped. Isolates showed in differents colors indicate sampling locations.

Figure 4. Midpoint-rooted maximum likelihood trees based on the nucleotide sequences of the *MP* (**A**) and *NSP* (**B**) genes of *Sida micrantha mosaic virus* (SiMMV) isolates. Nodes with bootstrap values equal or higher than 80% are indicated by filled circles, and those with values lower than 80 and higher 50% by empty circles. The unique recombination events detected within the *NSP* (events 3, 5, 8, 10 and 11 in Table 3, indicated in blue, orange, grey, brown and pink, respectively) and *MP* genes (events 7, 10 and 13 in Table 3, indicated in green, brown, and light pink, respectively) are shown as diagrams close to the branches where the substitutions due to each recombination event were mapped. Isolates showed in differents colors indicate sampling locations.

Figure 5. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (**A**) and DNA-B (**B**) nucleotide sequences of *Tomato yellow spot virus* (ToYSV) isolates. Nodes with posterior probability values between 0.50 and 0.89 are indicated by empty circles, and nodes with values equal or greater than 0.90 are indicated by filled circles. Colors indicate sampling locations.

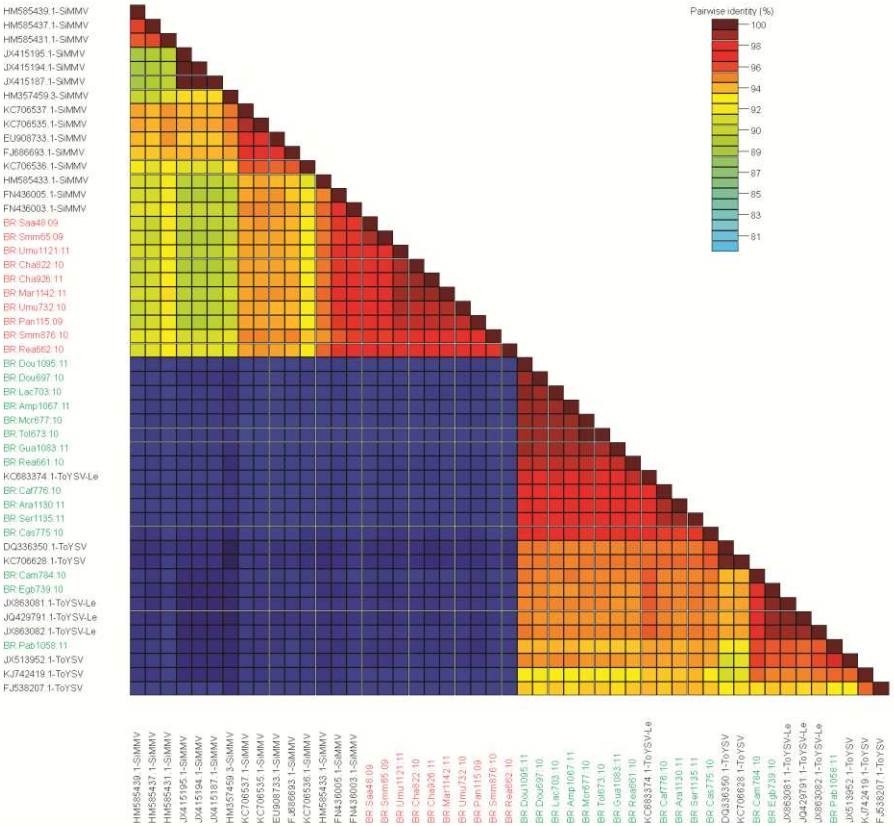
Figure 6. Midpoint-rooted maximum likelihood trees based on the nucleotide sequences of the *CP* (**A**) and *Rep* (**B**) genes of *Tomato yellow spot virus* (ToYSV) isolates. Nodes with bootstrap values equal or higher than 80% are indicated by filled circles, and those with values lower than 80 and higher 50% by empty circles. The unique recombination events detected within the *CP* (events 4 and 5 in Table 3, indicated in yellow and orange, respectively) and *Rep* genes (event 2 in Table 3, indicated in green) are shown as diagrams close to the branches where the

substitutions due to each recombination event were mapped. Isolates showed in differents colors indicate sampling locations.

Figure 7. Midpoint-rooted maximum likelihood trees based on the nucleotide sequences of the *MP* (**A**) and *NSP* (**B**) genes of *Tomato yellow spot virus* (ToYSV) isolates. Nodes with bootstrap values equal or higher than 80% are indicated by filled circles, and those with values lower than 80 and higher 50% by empty circles. The unique recombination events detected within the *NSP* (event 2 in Table 3, indicated in red) and *MP* genes (events 1 and 3 in Table 3, indicated in blue and pink, respectively) are shown as diagrams close to the branches where the substitutions due to each recombination event were mapped. Isolates showed in different colors indicate sampling locations.

Figure 1

A


Pairwise identity (%)
100
98
96
94
92
90
89
87
85
83
81
HM585439.1-SiMMV
HM585437.1-SiMMV
HM585431.1-SiMMV
JX415195.1-SiMMV
JX415194.1-SiMMV
JX415187.1-SiMMV
HM357459.3-SiMMV
KC706537.1-SiMMV
KC706535.1-SiMMV
EU908733.1-SiMMV
FJ686693.1-SiMMV
KC706536.1-SiMMV
HM585433.1-SiMMV
FN436005.1-SiMMV
FN436003.1-SiMMV
BR:Saa48:09
BR:Smm65:09
BR:Umu1121:11
BR:Cha822:10
BR:Cha926:11
BR:Mar1142:11
BR:Umu732:10
BR:Pan115:09
BR:Smm876:10
BR:Rea662:10
BR:Dou1095:11
BR:Dou697:10
BR:Lac703:10
BR:Amp1067:11
BR:Mcr677:10
BR:Tol673:10
BR:Gua1083:11
BR:Rea661:10
KC683374.1-ToYSV-Le
BR:Caf776:10
BR:Ara1130:11
BR:Ser1135:11
BR:Cas775:10
DQ336350.1-ToYSV
KC706628.1-ToYSV
BR:Cam784:10
BR:Egb739:10
JX863081.1-ToYSV-Le
JQ429791.1-ToYSV-Le
JX863082.1-ToYSV-Le
BR:Pab1058:11
JX513952.1-ToYSV
KJ742419.1-ToYSV
FJ538207.1-ToYSV

B

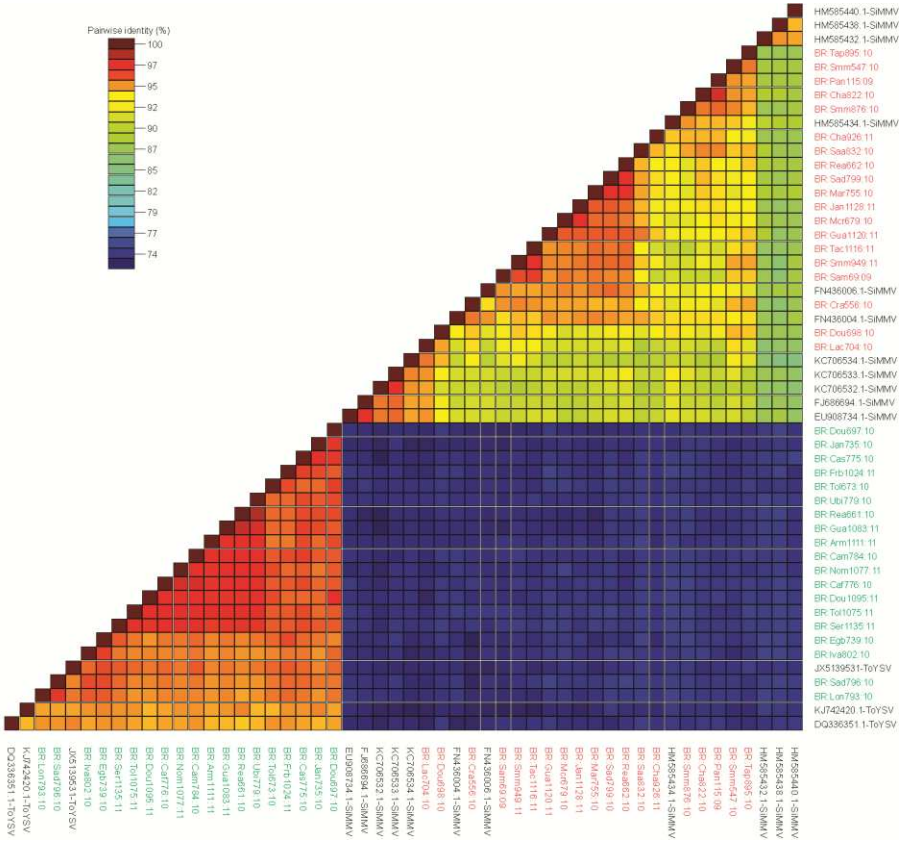

Pairwise identity (%)
100
97
95
92
90
87
85
82
79
77
74
HM585440.1-SiMMV
HM585438.1-SiMMV
HM585432.1-SiMMV
BR:Tap895:10
BR:Smm547:10
BR:Pan115:09
BR:Cha822:10
BR:Smm876:10
HM585434.1-SiMMV
BR:Cha926:11
BR:Saa832:10
BR:Rea662:10
BR:Sad799:10
BR:Mar755:10
BR:Jan1128:11
BR:Mcr679:10
BR:Gua1120:11
BR:Tac1116:11
BR:Smm949:11
BR:Sam69:09
FN436006.1-SiMMV
BR:Cra556:10
FN436004.1-SiMMV
BR:Dou698:10
BR:Lac704:10
KC706534.1-SiMMV
KC706533.1-SiMMV
KC706532.1-SiMMV
FJ686694.1-SiMMV
EU908734.1-SiMMV
BR:Dou697:10
BR:Jan735:10
BR:Cas775:10
BR:Frb1024:11
BR:Tol673:10
BR:Ubi779:10
BR:Rea661:10
BR:Gua1083:11
BR:Arm1111:11
BR:Cam784:10
BR:Nom1077:11
BR:Caf776:10
BR:Dou1095:11
BR:Tol1075:11
BR:Ser1135:11
BR:Egb739:10
BR:Iva802:10
JX513953.1-ToYSV
BR:Sad796:10
BR:Lon793:10
KJ742420.1-ToYSV
DQ336351.1-ToYSV

Figure 2

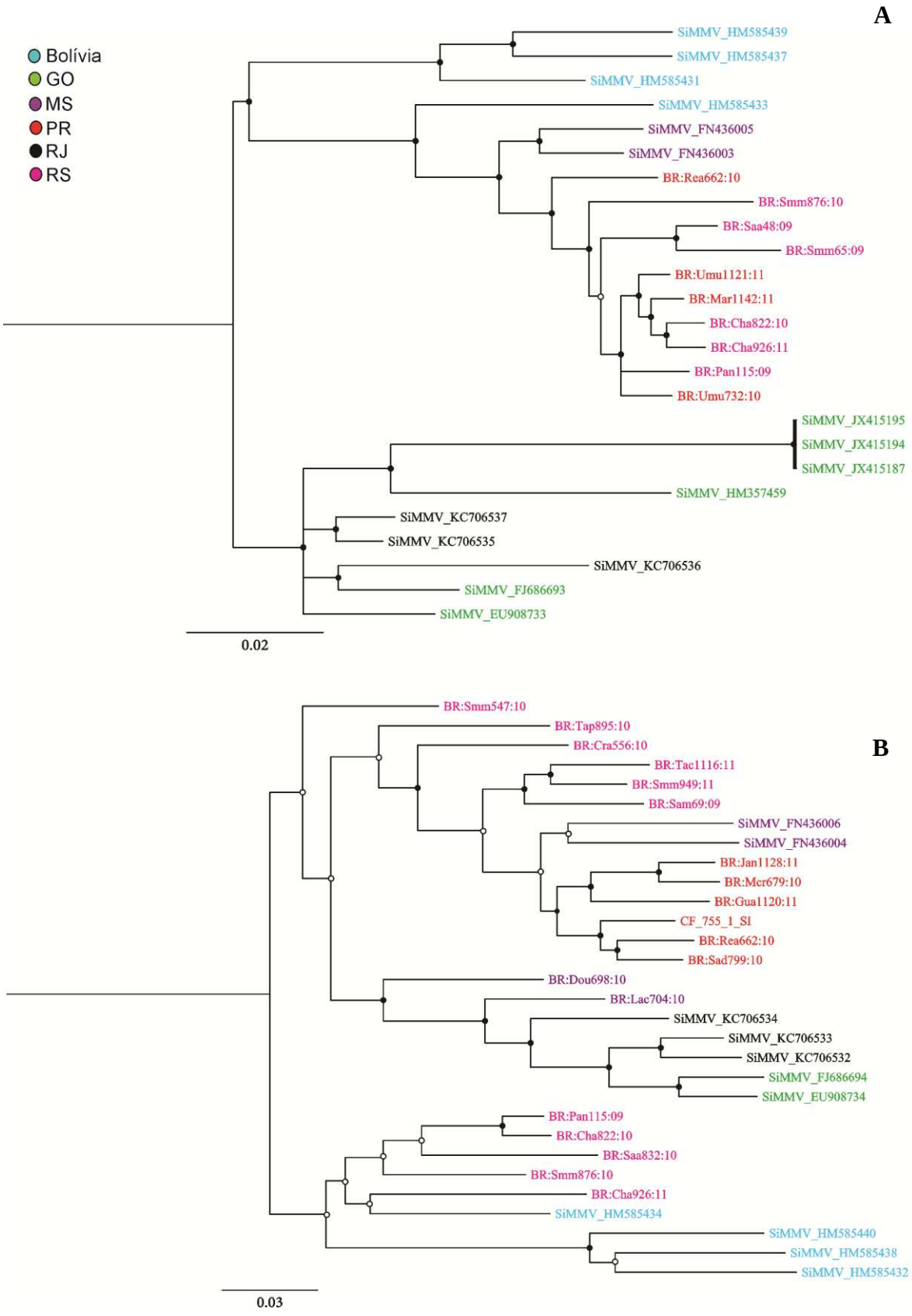

A
Bolívia
GO
MS
PR
RJ
RS
SiMMV_HM585439
SiMMV_HM585437
SiMMV_HM585431
SiMMV_HM585433
SiMMV_FN436005
SiMMV_FN436003
BR:Rea662:10
BR:Smm876:10
BR:Saa48:09
BR:Smm65:09
BR:Umu1121:11
BR:Mar1142:11
BR:Cha822:10
BR:Cha926:11
BR:Pan115:09
BR:Umu732:10
SiMMV_JX415195
SiMMV_JX415194
SiMMV_JX415187
SiMMV_HM357459
SiMMV_KC706537
SiMMV_KC706535
SiMMV_KC706536
SiMMV_FJ686693
SiMMV_EU908733
0.02
B
BR:Smm547:10
BR:Tap895:10
BR:Cra556:10
BR:Tac1116:11
BR:Smm949:11
BR:Sam69:09
SiMMV_FN436006
SiMMV_FN436004
BR:Jan1128:11
BR:Mcr679:10
BR:Gua1120:11
CF_755_1_SI
BR:Rea662:10
BR:Sad799:10
BR:Dou698:10
BR:Lac704:10
SiMMV_KC706534
SiMMV_KC706533
SiMMV_KC706532
SiMMV_FJ686694
SiMMV_EU908734
BR:Pan115:09
BR:Cha822:10
BR:Saa832:10
BR:Smm876:10
BR:Cha926:11
SiMMV_HM585434
SiMMV_HM585440
SiMMV_HM585438
SiMMV_HM585432
0.03

Figure 3

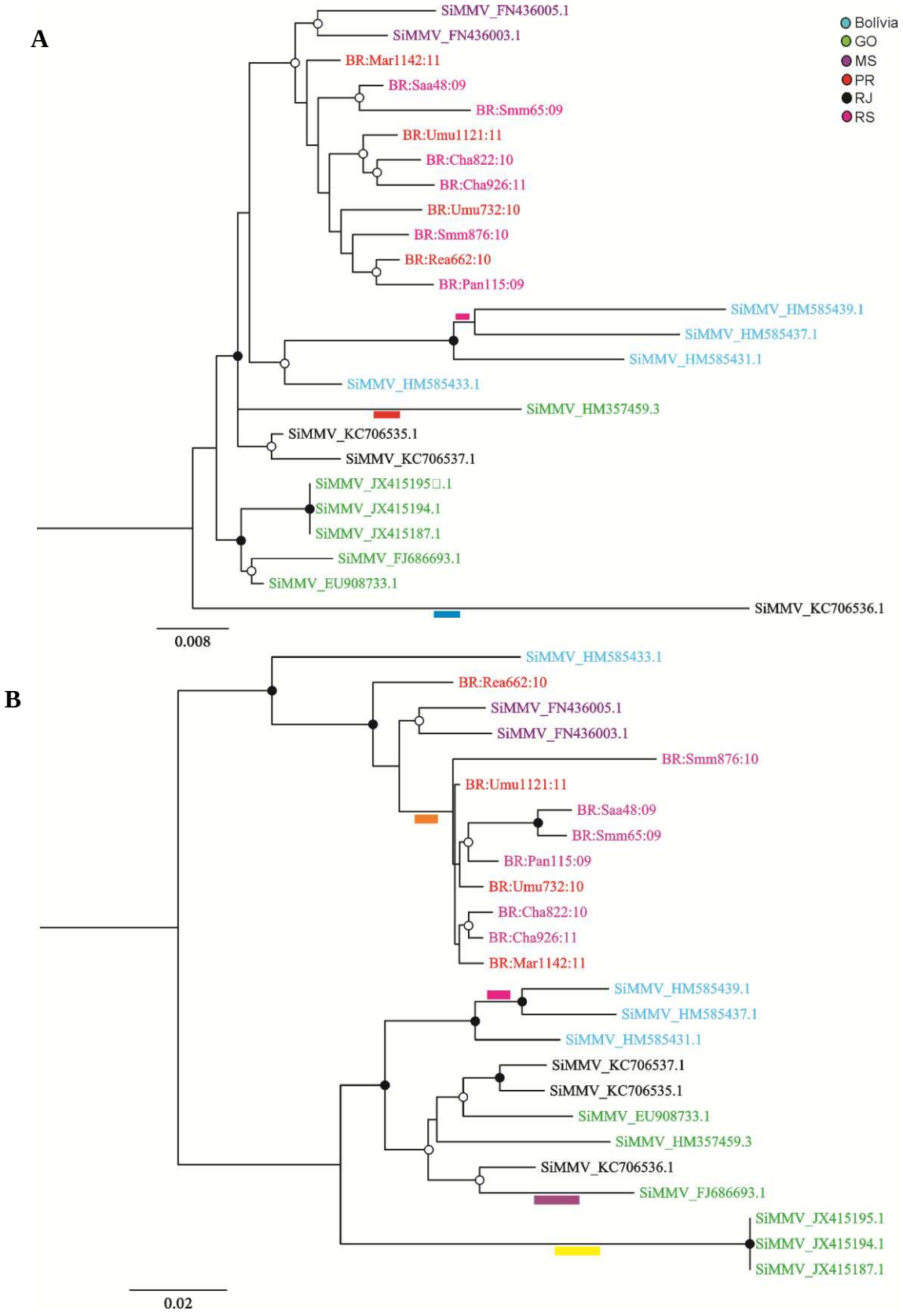

A
SiMMV_FN436005.1
SiMMV_FN436003.1
BR:Mar1142:11
BR:Saa48:09
BR:Smm65:09
BR:Umu1121:11
BR:Cha822:10
BR:Cha926:11
BR:Umu732:10
BR:Smm876:10
BR:Rea662:10
BR:Pan115:09
SiMMV_HM585439.1
SiMMV_HM585437.1
SiMMV_HM585431.1
SiMMV_HM585433.1
SiMMV_HM357459.3
SiMMV_KC706535.1
SiMMV_KC706537.1
SiMMV_JX415195□.1
SiMMV_JX415194.1
SiMMV_JX415187.1
SiMMV_FJ686693.1
SiMMV_EU908733.1
SiMMV_KC706536.1
0.008
Bolívia
GO
MS
PR
RJ
RS
B
SiMMV_HM585433.1
BR:Rea662:10
SiMMV_FN436005.1
SiMMV_FN436003.1
BR:Smm876:10
BR:Umu1121:11
BR:Saa48:09
BR:Smm65:09
BR:Pan115:09
BR:Umu732:10
BR:Cha822:10
BR:Cha926:11
BR:Mar1142:11
SiMMV_HM585439.1
SiMMV_HM585437.1
SiMMV_HM585431.1
SiMMV_KC706537.1
SiMMV_KC706535.1
SiMMV_EU908733.1
SiMMV_HM357459.3
SiMMV_KC706536.1
SiMMV_FJ686693.1
SiMMV_JX415195.1
SiMMV_JX415194.1
SiMMV_JX415187.1
0.02

Figure 4

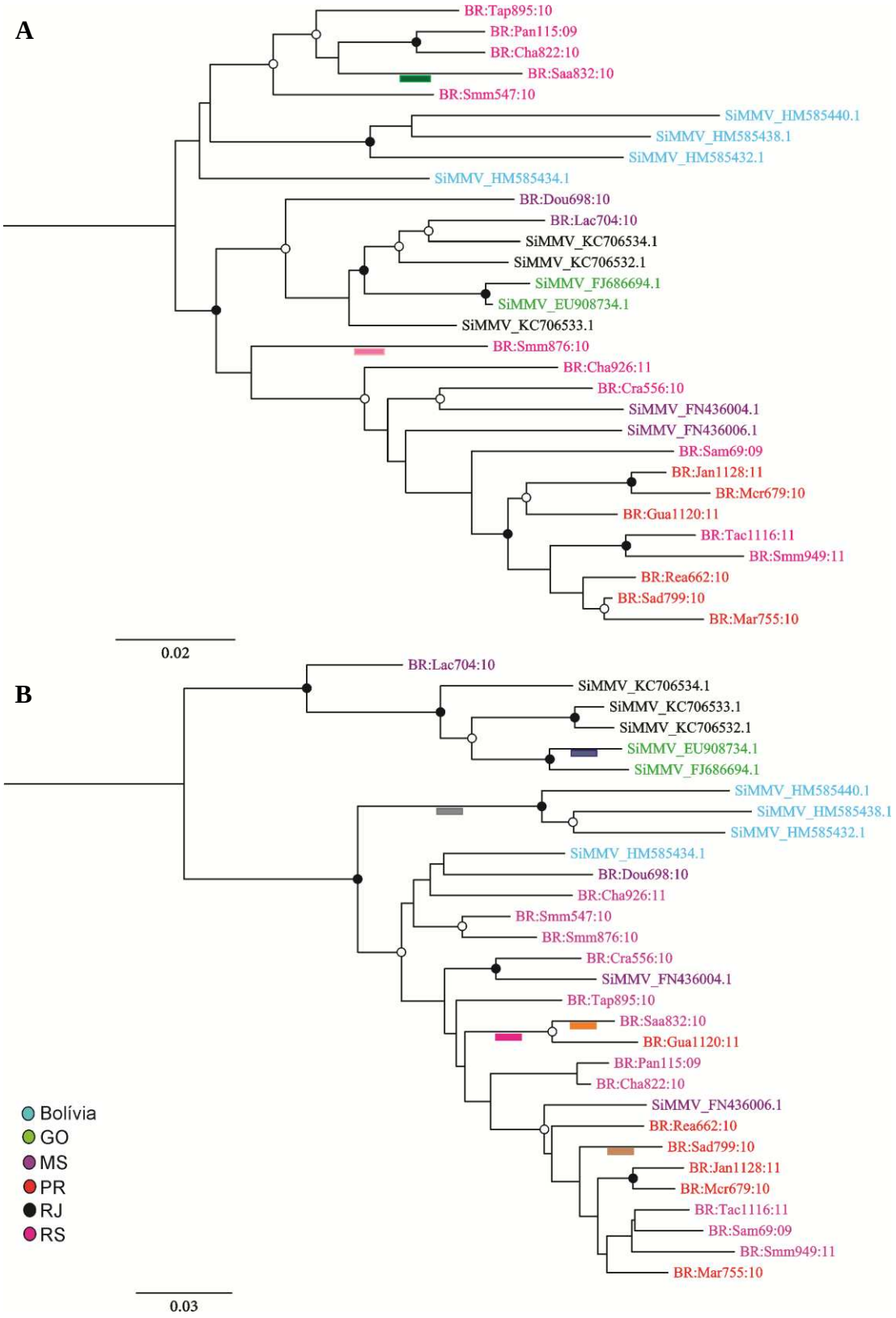

A
BR:Tap895:10
BR:Pan115:09
BR:Cha822:10
BR:Saa832:10
BR:Smm547:10
SiMMV_HM585440.1
SiMMV_HM585438.1
SiMMV_HM585432.1
SiMMV_HM585434.1
BR:Dou698:10
BR:Lac704:10
SiMMV_KC706534.1
SiMMV_KC706532.1
SiMMV_FJ686694.1
SiMMV_EU908734.1
SiMMV_KC706533.1
BR:Smm876:10
BR:Cha926:11
BR:Cra556:10
SiMMV_FN436004.1
SiMMV_FN436006.1
BR:Sam69:09
BR:Jan1128:11
BR:Mcr679:10
BR:Gua1120:11
BR:Tac1116:11
BR:Smm949:11
BR:Rea662:10
BR:Sad799:10
BR:Mar755:10
0.02
B
BR:Lac704:10
SiMMV_KC706534.1
SiMMV_KC706533.1
SiMMV_KC706532.1
SiMMV_EU908734.1
SiMMV_FJ686694.1
SiMMV_HM585440.1
SiMMV_HM585438.1
SiMMV_HM585432.1
SiMMV_HM585434.1
BR:Dou698:10
BR:Cha926:11
BR:Smm547:10
BR:Smm876:10
BR:Cra556:10
SiMMV_FN436004.1
BR:Tap895:10
BR:Saa832:10
BR:Gua1120:11
BR:Pan115:09
BR:Cha822:10
SiMMV_FN436006.1
BR:Rea662:10
BR:Sad799:10
BR:Jan1128:11
BR:Mcr679:10
BR:Tac1116:11
BR:Sam69:09
BR:Smm949:11
BR:Mar755:10
Bolívia
GO
MS
PR
RJ
RS
0.03

Figure 5.

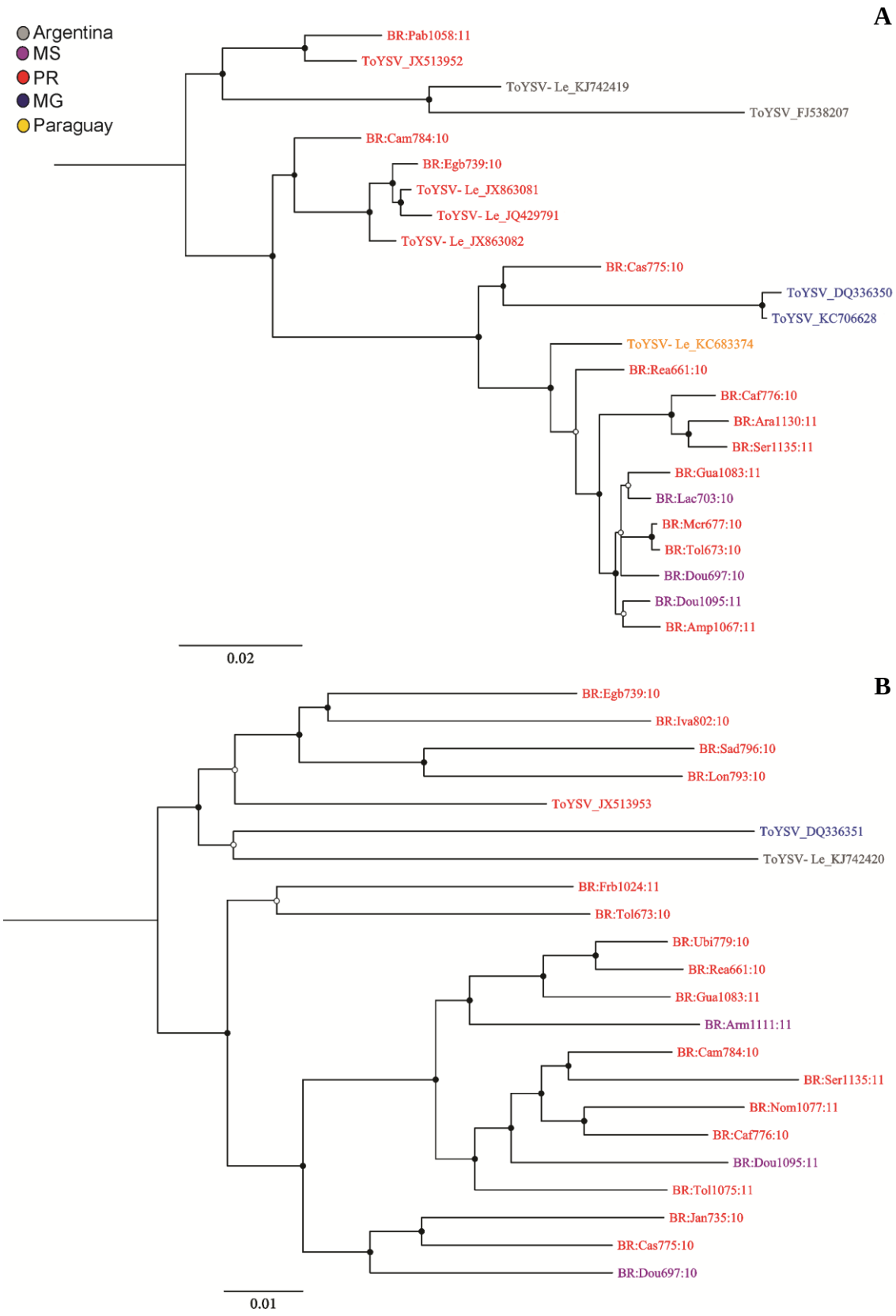

A
Argentina
MS
PR
MG
Paraguay
BR:Pab1058:11
ToYSV_JX513952
ToYSV- Le_KJ742419
ToYSV_FJ538207
BR:Cam784:10
BR:Egb739:10
ToYSV- Le_JX863081
ToYSV- Le_JQ429791
ToYSV- Le_JX863082
BR:Cas775:10
ToYSV_DQ336350
ToYSV_KC706628
ToYSV- Le_KC683374
BR:Rea661:10
BR:Caf776:10
BR:Ara1130:11
BR:Ser1135:11
BR:Gua1083:11
BR:Lac703:10
BR:Mcr677:10
BR:Tol673:10
BR:Dou697:10
BR:Dou1095:11
BR:Amp1067:11
0.02
B
BR:Egb739:10
BR:Iva802:10
BR:Sad796:10
BR:Lon793:10
ToYSV_JX513953
ToYSV_DQ336351
ToYSV- Le_KJ742420
BR:Frb1024:11
BR:Tol673:10
BR:Ubi779:10
BR:Rea661:10
BR:Gua1083:11
BR:Arm1111:11
BR:Cam784:10
BR:Ser1135:11
BR:Nom1077:11
BR:Caf776:10
BR:Dou1095:11
BR:Tol1075:11
BR:Jan735:10
BR:Cas775:10
BR:Dou697:10
0.01

Figure 6

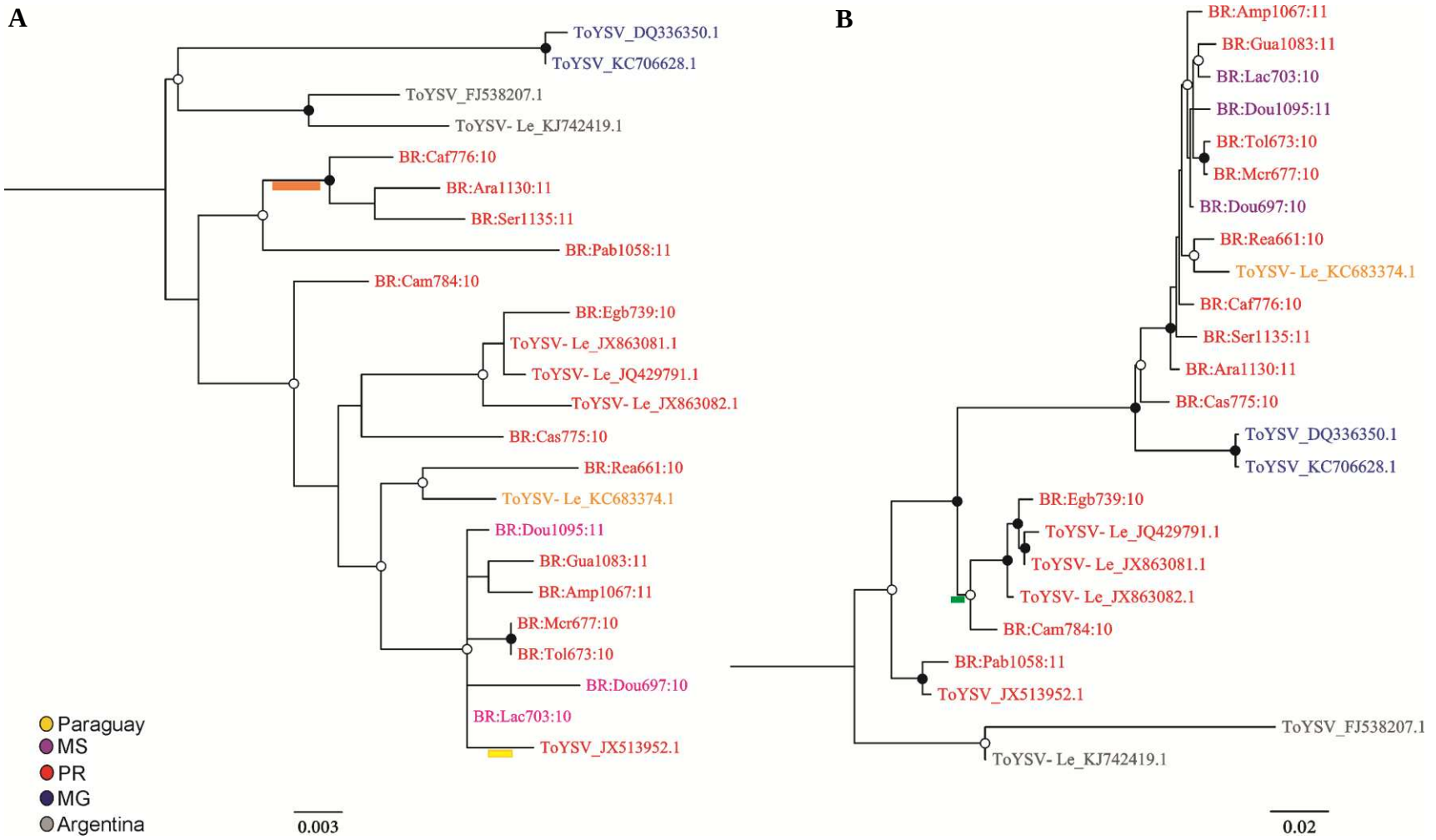
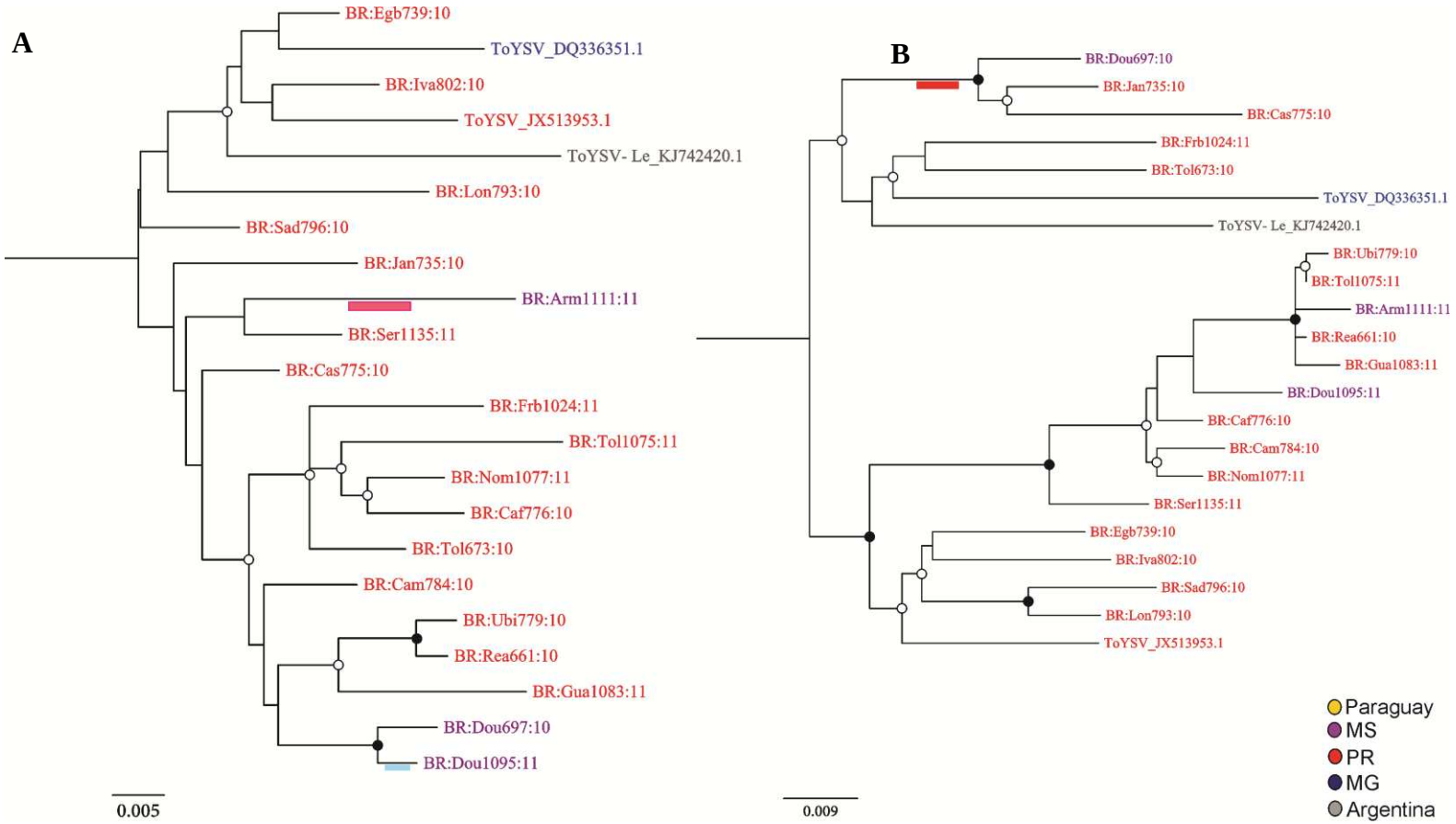

A
ToYSV_DQ336350.1
ToYSV_KC706628.1
ToYSV_FJ538207.1
ToYSV- Le_KJ742419.1
BR:Caf776:10
BR:Ara1130:11
BR:Ser1135:11
BR:Pab1058:11
BR:Cam784:10
BR:Egb739:10
ToYSV- Le_JX863081.1
ToYSV- Le_JQ429791.1
ToYSV- Le_JX863082.1
BR:Cas775:10
BR:Rea661:10
ToYSV- Le_KC683374.1
BR:Dou1095:11
BR:Gua1083:11
BR:Amp1067:11
BR:Mcr677:10
BR:Tol673:10
BR:Dou697:10
BR:Lac703:10
ToYSV_JX513952.1
Paraguay
MS
PR
MG
Argentina
0.003
B
BR:Amp1067:11
BR:Gua1083:11
BR:Lac703:10
BR:Dou1095:11
BR:Tol673:10
BR:Mcr677:10
BR:Dou697:10
BR:Rea661:10
ToYSV- Le_KC683374.1
BR:Caf776:10
BR:Ser1135:11
BR:Ara1130:11
BR:Cas775:10
ToYSV_DQ336350.1
ToYSV_KC706628.1
BR:Egb739:10
ToYSV- Le_JQ429791.1
ToYSV- Le_JX863081.1
ToYSV- Le_JX863082.1
BR:Cam784:10
BR:Pab1058:11
ToYSV_JX513952.1
ToYSV_FJ538207.1
ToYSV- Le_KJ742419.1
0.02

Figure 7


A
BR:Egb739:10
ToYSV_DQ336351.1
BR:Iva802:10
ToYSV_JX513953.1
ToYSV- Le_KJ742420.1
BR:Lon793:10
BR:Sad796:10
BR:Jan735:10
BR:Arm1111:11
BR:Ser1135:11
BR:Cas775:10
BR:Frb1024:11
BR:Tol1075:11
BR:Nom1077:11
BR:Caf776:10
BR:Tol673:10
BR:Cam784:10
BR:Ubi779:10
BR:Rea661:10
BR:Gua1083:11
BR:Dou697:10
BR:Dou1095:11
0.005
B
BR:Dou697:10
BR:Jan735:10
BR:Cas775:10
BR:Frb1024:11
BR:Tol673:10
ToYSV_DQ336351.1
ToYSV- Le_KJ742420.1
BR:Ubi779:10
BR:Tol1075:11
BR:Arm1111:11
BR:Rea661:10
BR:Gua1083:11
BR:Dou1095:11
BR:Caf776:10
BR:Cam784:10
BR:Nom1077:11
BR:Ser1135:11
BR:Egb739:10
BR:Iva802:10
BR:Sad796:10
BR:Lon793:10
ToYSV_JX513953.1
0.009
Paraguay
MS
PR
MG
Argentina

CAPÍTULO 3

DIVERSITY OF DELTASATELLITES ASSOCIATED WITH SWEEPOVIRUSES INFECTING *Ipomoea indica* IN SOUTHERN SPAIN

Ferro, C.G., Fiallo-Olivé, E., Zerbini, F.M., Navas-Castillo, J. **Diversity of deltasatellites associated with sweepoviruses infecting *Ipomoea indica* in southern Spain.**

Diversity of deltasatellites associated with sweepoviruses infecting *Ipomoea indica* in southern Spain

C. G. Ferro[1,2,3], E. Fiallo-Olivé[1,2,3], F. M. Zerbini[2,3], J. Navas-Castillo[1,2,3#]

[1]Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga - Consejo Superior de Investigaciones Científicas (IHSM-UMA-CSIC), 29750 Algarrobo-Costa, Málaga, Spain

[2]Dep. de Fitopatologia/BIOAGRO and [3]National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[#]Corresponding author:

Phone: (+34) 952548990; Fax: (+34) 952552677; E-mail: jnavas@eelm.csic.es

Abstract

Begomoviruses (whitefly-transmitted geminiviruses) possess one or two circular single-stranded DNA genomic components encapsidated in twinned quasi-icosahedral (geminate) virions. Most begomoviruses in the New World (NW) are bipartite, while the majority in the Old World (OW) are monopartite. They are responsible for many economically important crop diseases worldwide. Begomoviruses that infect *Ipomoea* spp. (family Convolvulaceae), commonly known as sweepoviruses, have the typical genomic organization of OW monopartite viruses but are phylogenetically distinct from all other species in the genus. Monopartite begomoviruses have been frequently found in association with two classes of satellite molecules, betasatellites and alphasatellites. The first DNA satellite discovered, referred to as ToLCV-sat, has a region with similarity to the satellite common region (SCR) of betasatellites, an adenine-rich (A-rich) region, a predicted stem-loop structure containing the nonanucleotide TAATATTAC, and a secondary stem-loop. Recently, non-coding ToLCV-sat-like satellites associated with sweepoviruses have been described from Spain and Venezuela. These and other small non-coding circular DNA satellites associated with begomoviruses have been proposed to be included in a new class of DNA satellites, deltasatellites. Available sequences of deltasatellites of sweet potato (*I. batatas*) and blue morning glory (*I. indica*) plants from Spain were obtained through PCR amplification using a single pair of primers. In this work, the genetic variability of deltasatellites associated with sweepoviruses infecting *I. indica* was studied by resampling the southern Spain populations where they were first detected, and expanding the sampling to other geographical areas. Rolling circle amplification (RCA), a technique that allows

amplification of circular ssDNA molecules without previous knowledge of nucleotide sequence, was used to clone deltasatellites and associated sweepoviruses, *Sweet potato leaf curl virus* (SPLCV) and *Sweet potato mosaic virus* (SPMV). Deltasatellites showed low sequence variability, no evidence for recombination, and no obvious geographical structuration. Also, a sweepovirus-deltasatellite chimera with a size similar to deltasatellites was detected.

Introduction

Begomoviruses (genus *Begomovirus*, family *Geminiviridae*) possess one or two genomic components of circular, single-stranded DNA (ssDNA) encapsidated in twinned quasi-icosahedral (geminate) virions (Brown et al., 2012) that are transmitted in nature by whiteflies (Hemiptera: Aleyrodidae) of the *Bemisia tabaci* complex to dicotyledonous plants (Brown et al., 2015; Navas-Castillo et al., 2011). Begomoviruses are responsible for many economically important crop diseases worldwide (Briddon, 2003; Legg and Fauquet, 2004; Monci et al., 2002; Morales and Anderson, 2001; Polston and Anderson, 1997; Sserubombwe et al., 2008; Trenado et al., 2011; Were et al., 2004). The sweepoviruses constitute a group of begomoviruses which infect sweet potato (*Ipomoea batatas*) and other species of the family Convolvulaceae (Lozano et al., 2009, 2016). Sweepoviruses group in a cluster which is basal to the main phylogenetic groups in the genus, the Old World (OW) and the New World (NW) begomoviruses (Lozano et al., 2009). In recent years, a number of sweepoviruses have been identified in various parts of the world (Albuquerque et al., 2011; Banks et al., 1999; Briddon et al., 2006; Esterhuizen et al., 2012; Fuentes and Salazar, 2003; Lotrakul et al., 1998; Lotrakul et al., 2003; Lozano et al., 2009; Luan et al., 2006; Luan et al., 2007; Miano et al., 2006; Paprotka et al., 2010a; Prasanth and Hegde, 2008; Wasswa et al., 2011). The sweepoviruses have the typical genomic organization of monopartite begomoviruses from the OW. The virion-sense strand contains two open reading frames (ORFs) which encode the coat protein (CP), with functions in particle formation and transmission by *B. tabaci* (Rojas et al., 2005), and the V2 protein, which is involved in viral movement and suppression of gene silencing (Glick et al., 2008; Padidam et al., 1996; Rybicki, 1994). The

complementary-sense strand encodes proteins involved in viral replication (Rep, REn) and a protein that encodes a transcriptional factor of the *CP* gene and which also acts as a suppressor of gene silencing (TrAP) (Rojas et al., 2005). The intergenic region (IR) contains a predicted stem-loop structure with the nonanucleotide TAATATTAC, conserved among the members of the family *Geminiviridae*, and repeated short sequence motifs (known as iterons) around the TATA box of the Rep promoter that are Rep binding sites and, together with the stem-loop structure, form the origin of virion-sense DNA replication (Rojas et al., 2005; Zhou, 2013).

Many begomoviruses from the OW are associated with satellite DNA molecules. DNA satellites are subviral molecules which lack genes encoding proteins with functions related to replication, depend on the co-infection with a helper virus for their multiplication, and their genetic material does not have significant sequence similarity to the helper virus genome (Briddon et al., 2012). They either encode non-structural proteins or are non-coding, and are encapsidated by CP of the helper virus. Replication of the satellite interferes with replication of the helper virus and may affect disease symptoms, ranging from attenuation to exacerbation depending on the satellite, the helper virus and the host plant (Hull, 2002; Fiallo-Olivé et al., 2012).

Monopartite begomoviruses have been frequently found in association with two classes of DNA satellites referred to as alphasatellites and betasatellites (Briddon and Stanley, 2006; Zhou, 2013). Alphasatellites (initially known as DNA-1) (Briddon et al., 2004) are approximately half the size of begomovirus DNAs and encode a rolling-circle replication (RCR) initiator protein (alpha-*Rep*) (Briddon et al., 2004; Xie et al., 2010; Zhou, 2013). They are not strictly satellites since they are capable of autonomous replication in plant cells, but require a helper begomoviruses

for movement in plants as well as for insect transmission (Mansoor et al., 1999; Saunders and Stanley, 1999). Commonly, alphasatellites occur associated with begomovirus/betasatellites complexes in the OW (Leke et al., 2015; Xie et al., 2010), but recently they have been found in association with NW bipartite begomoviruses in Brazil, Cuba and Venezuela (Jeske et al., 2014; Paprotka et al., 2010b; Romay et al., 2010). Betasatellites (previously known as DNA-β) have so far only been identified in the OW and are generally associated with monopartite begomoviruses (Briddon et al., 2001; Saunders et al., 2000). Betasatellite genomes are typically half the length of their helper begomoviruses and they have a highly conserved genome organization consisting of: *i*) a region known as the satellite conserved region (SCR) that is highly conserved among all betasatellites, *ii*) a single gene, *betaC1*, which is conserved both in sequence and position and is encoded in the complementary-sense strand DNA, and *iii*) an A-rich region (Briddon et al., 2003; Zhou et al., 2003). Most betasatellites increase the severity of symptoms caused by the helper virus (Sivalingam and Varma, 2012; Zhou, 2013), likely due to the transcriptional and post transcriptional gene silencing suppression activity of the betaC1 protein (Yang et al., 2011).

The first DNA satellite discovered, referred to as ToLCV-sat, was isolated from tomato plants infected with the monopartite begomovirus *Tomato leaf curl virus* (ToLCV) originating from Australia (Dry et al., 1997). The satellite is a non-coding molecule about a quarter the size of begomovirus genomic components (682 nucleotides) with no significant sequence similarity with the helper begomovirus. ToLCV-sat contains a sequence with similarity to the satellite conserved region (SCR) of betasatellites, an A-rich sequence and two predicted stem-loop structures, the conserved stem-loop conserved in the begomoviruses that contains the

nonanucleotide TAATATTAC, and a secondary stem-loop containing a sequence identical to the predicted iteron sequence of ToLCV within the loop. ToLCV-sat depends on the helper virus for replication and, based upon sequence and structural similarities with betasatellites, it has been suggested that it originated as a defective betasatellite (Saunders et al., 2000).

Recently, a novel class of DNA satellites associated with begomoviruses has been identified infecting malvaceous plants in Cuba (Fiallo-Olivé et al., 2012). These satellites contain the same genome features of ToLCV-sat: a genome of about a quarter the size of begomovirus components, lack of coding capacity, a stem-loop containing the conserved nonanucleotide TAATATTAC, a putative secondary stem-loop structure situated close to begomovirus iteron-like sequences, an A-rich region, and a short region with high sequence identity with the betasatellite SCR (Fiallo-Olivé et al., 2012). Using a vector-enabled metagenomics (VEM) approach, similar satellites have also been identified in whiteflies collected in Florida (Ng et al., 2011).

A novel class of DNA satellites in association with several distinct sweepoviruses has been detected in sweet potato and *Ipomoea indica* plants in Spain and *Merremia dissecta* plants in Venezuela (Lozano et al., 2016). These DNA satellites are structurally similar to ToLCV-sat, with a region with similarity to the betasatellite SCR, an A-rich region, a predicted stem-loop structure containing the nonanucleotide TAATATTAC, and a secondary predicted stem-loop. The name deltasatellites has been proposed for this class of small non-coding DNA satellites that include ToLCV-sat, the satellites associated with begomoviruses from malvaceous hosts in Cuba, and those found associated with sweepoviruses (Lozano et al., 2016).

The sweepovirus-associated DNA satellites characterized from sweet potato and *I. indica* in Spain seem to cluster according to geographic origin (either mainland Spain or the Canary Islands) and plant host (either sweet potato or *I. indica*), although the number of currently available sequences is small. These satellites were amplified by PCR using a pair of abutting primers designed to anneal at the satellite-like sequence contained in a virus-satellite chimera found in a sweet potato sample (Lozano et al., 2016). The advent of rolling circle amplification (RCA) using phage phi29 DNA polymerase, that allows amplification of the complete genome of ssDNA viruses without previous knowledge of nucleotide sequence, has facilitated the characterization of new ssDNA viruses and subviral molecules (Haible et al., 2006; Inoue-Nagata et al., 2004). In this work, RCA was used to reveal the actual genetic diversity of deltasatellites associated with sweepoviruses infecting *I. indica* plants by resampling the southern Spain populations, where they were detected by the first time, and expanding the analysis to other geographical areas. The sweepoviruses present in the samples that contained satellites were also cloned and sequenced to define the range of sweepoviruses which can act as helper viruses. Variability of satellites was lower than that found in helper sweepovirus genomes, with no evidence of recombination. A sweepovirus-satellite chimera with a similar size to deltasatellites was also characterized in this study.

Materials and Methods

Sample collection and cloning

Leaf samples from *I. indica* plants (Figure 1) were collected in southern Spain (Cádiz, Málaga, Granada and Murcia provinces) in 2015 (Figure 2). Total DNA was

extracted from leaf tissue using a CTAB-based purification method (Haible et al., 2006). Circular ssDNA (sweepovirus genomes and associated deltasatellites) was amplified by RCA using the TempliPhi DNA Amplification Kit (GE Healthcare). Amplified products were initially digested with the four-base cutter restriction enzyme *Hpa*II, to screen for putative sweepovirus-infected samples. RCA products of the selected samples were digested with a set of six-base cutter restriction enzymes to identify those that cleave the viral and satellite components at a single site. RCA products digested with *Nco*I, putatively corresponding to sweepovirus genomes, were cloned into the pGEM-T-Easy vector (Promega), while those corresponding putatively to DNA satellites were digested with *Pst*I and cloned into pBluescript II SK(+) (Stratagene). Cloned inserts were sequenced commercially (Macrogen Inc.) and sequences were assembled with SeqMan, part of the Lasergene sequence analysis package (DNAStar, Inc.).

Sequence comparisons, phylogenetic analysis and genetic variability

Sequences were initially analyzed with the BLAST*n* algorithm (Altschul et al., 1990) for similarity searches of the GenBank database. Sweepovirus and satellite sequences were aligned with MUSCLE (Edgar, 2004) and pairwise comparisons between all the sequences obtained and the sequences with the greatest similarity (as determined by the BLAST*n* search) were performed with the program Sequence Demarcation Tool (SDT) v. 1.2 (Muhire et al., 2014). Multiple sequence alignments were prepared using the MUSCLE alignment option in MEGA6 before phylogenetic inference by the Neighbor-Joining method (Tamura et al., 2013). The average pairwise number of nucleotide differences per site (nucleotide diversity, π) was estimated using DnaSP v. 5.10 (Rozas et al., 2003).

Recombination analysis

Recombination analysis was performed using the Rdp, Geneconv, Bootscan, Maximum Chi Square, Chimaera, SisterScan and 3Seq methods implemented in Recombination Detection Program (RDP) v. 4 (Martin et al., 2015). Alignments were scanned with default settings for the different methods. Statistical significance was inferred by *p*-values lower than a Bonferroni-corrected cut-off of 0.05. Only recombination events detected by at least four of the analysis methods available in the program were considered reliable.

Results

Presence of sweepoviruses and associated deltasatellites infecting I. indica *in southern Spain*

Fifty-nine out of 65 *I. indica* samples collected from four provinces in southern Spain were putatively infected by begomoviruses based on the detection of a ~2,800-bp band after digestion of the RCA products with *Nco*I. Out of these 59 samples, 47 were possibly infected by deltasatellites, due the presence of a ~700-bp band obtained after digestion of the RCA products with *Pst*I (data not shown). From these samples, 31 full-length sweepovirus genomes and 92 satellite DNA molecules were obtained (Table 1).

Pairwise comparisons indicated that the sweepovirus sequences were subdivided into two groups. One of the groups contained 26 sequences that share >88.3% nucleotide identity amongst them, and pairwise comparisons with the most closely related sweepoviruses showed that the sequences were 94.5% to 98.9% identical with the previously reported isolates of *Sweet potato leaf curl virus*

(SPLCV). The other group consisted of five sequences (LM7572, LM7842, LM7844, LM7286 and LM7805) that shared >94.2% nucleotide identity amongst them and 91.2% to 95.0% identity with *Sweet potato mosaic virus* (SPMV) (Figure 3).

Sequencing of the *Pst*I clones showed that they were deltasatellites, with one exception which resulted to be a sweepovirus-deltasatellite chimera (see below). The overall nucleotide identity between the 92 deltasatellite clones ranged from 89.1 to 100%. These sequences shared 93.3% to 99.1% identity with the deltasatellites previously detected in southern Spain from *I. indica* and *I. batatas* (Lozano et al., 2016) (Figure 4). The names of the isolates retrieved from Genbank are shown in Supplementary Tables 1 and 2. All satellites possess the genomic structure previously described by Lozano et al. (2016): a region with similarity to the SCR of betasatellites, an A-rich sequence, the conserved stem-loop structure containing the nonanucleotide TAATATTAC, and a secondary stem-loop.

Genetic variability and recombination of sweepoviruses and their associated deltasatellites

The sweepovirus population obtained in this work showed a genetic variability (Table 2) similar to that observed for other begomovirus populations (Duffy and Holmes, 2008, 2009; Ge et al., 2007; Lima et al., 2013; Ramos-Sobrinho et al., 2014; Rocha et al., 2013). The nucleotide sequences of the *CP*, *V2*, *Rep*, *TrAP*, *REn* and *C4* genes and that of the IR were analyzed separately to ascertain whether the distribution of variability was evenly distributed throughout the genome. In general, the non-coding IR was more variable compared to coding regions, and the *Rep* and *REn* (SPMV) and *Rep* and *V2* (SPLCV) genes were more variable than the other genes (Table 2).

To investigate possible recombination events, the full-length genomes of the sweepoviruses detected in this study plus all the sweepoviruses described in Spain and other closer related were analyzed using the RDP4 program (Martin et al., 2015). Twenty-one unique recombination events were detected (Table 3), and most of them involved breakpoints located within the *Rep* gene and the IR (Table 3).

Deltasatellite sequences showed lower variability than the sweepovirus sequences (Table 2). The isolates obtained from Granada and Málaga showed greater genetic variability (0.02846 and 0.03903, respectively) than isolates from Cádiz and Murcia (0.015 and 0.017, respectively) (Table 2). There was no evidence for recombination in the deltasatellite sequences after analyzing a set of 108 isolates [92 obtained in this study plus 16 from *I. indica* and *I. batatas* reported by Lozano et al. (2016)] (data not shown).

Phylogenetic relationships of the deltasatellites and their helper sweepoviruses

Deltasatellite sequences obtained in this study were aligned with deltasatellite sequences available in GenBank, and a phylogenetic tree was produced (Figure 5). Deltasatellites characterized in this work appeared to form several clusters with low phylogenetic signal, without obvious relation to geographical origin (Figure 5). The tree shows two clades (with 100% bootstrap support): one major clade containing Canary Islands, Cádiz, Granada, Málaga and Murcia isolates, subdivided into smaller groups with low phylogenetic support; and another clade occupying a basal position in relation to the rest of deltasatellites associated with sweepoviruses found in Spain, containing three isolates (LM7773, LM7604 and LM7605). Within the larger clade, isolates from Canary Islands plus two isolates from Málaga (LM7543 and LM7544) formed one well-defined clade, while the other isolates from Cádiz, Granada, Málaga

and Murcia formed several smaller clades not supported by bootstrap analysis (Figure 5).

The sweepovirus phylogenetic tree showed three well-supported clades (Figure 6). A major well-supported clade contains the SPLCV isolates obtained in this study from Málaga, Granada and Cádiz, some of the SPLCV isolates from Canary Islands and Málaga obtained by Lozano et al. (2009), and the SPMV isolates obtained in this study. Most sequences from this study grouped with SPLCV, except five isolates (LM7572, LM7842, LM7844, LM7286 and LM7805) which formed a well-supported clade with SPMV isolates (Figure 6). This close relationship is consistent with the pairwise sequence identity analysis (Figure 4). This major clade was subdivided into four well-supported smaller groups: two clades containing isolates from Málaga, a third clade containing isolates from Málaga, Granada and Canary Islands, and a fourth clade containing the SPMV sequences from Málaga and Cádiz. The other major clade contained SPLCV and *Sweet potato leaf curl Canary virus* (SPLCCaV) isolates from the Canary Islands, with the exception of one SPLCV isolate from Málaga (Lozano et al., 2009). The last major well-supported clade contained representative sequences from other sweepovirus species (Figure 6).

Detection of a sweepovirus-deltasatellite chimera

One of the clones obtained with an insert with the typical size of deltasatellites was shown to correspond to a sweepovirus-deltasatellite chimera (Figure 7). This clone was obtained from sample ii26, collected in the province of Málaga. The insert of this clone (LM7556) is 699 bp in length, and BLAST*n* analysis showed significant identity with available sequences of sweepoviruses for only about 50% of the length (coordinates 444-98). This region has the IR containing a stem-

loop with the conserved nonanucleotide TAATATTAC. This fragment showed 89.0% identity to the corresponding region of the genome of the sweepovirus SPLCV that was isolated from the same sample (isolate LM7861). BLAST*n* analysis of the remaining 344 bp (coordinates 99-443) showed 97% identity to a region of the deltasatellite that was isolated from the same sample (isolate LM7629). This DNA fragment contained the A-rich region and half of the conserved stem-loop structure of the deltasatellite (Figure 7).

Discussion

Lozano et al. (2016) found satellites associated with sweepoviruses in Spain and Venezuela, similar in structure to ToLCV-sat, i.e., they contain a region with similarity to the SCR of betasatellites, an A-rich region, a predicted stem-loop structure containing the nonanucleotide TAATATTAC, and a secondary stem-loop. The name deltasatellite was proposed for the members of this new class of small non-coding satellites, which includes in addition to ToLCV-sat and sweepovirus-associated satellites, those associated to NW bipartite begomoviruses (Fiallo-Olivé et al., 2012). Lozano et al. (2016) obtained, by PCR amplification, 16 satellite sequences showing low genetic variability, the overall nucleotide identity amongst them ranging from 94.6 to 99.9%. The low genetic diversity of deltasatellite sequences could have been the result of the use of specific primers for cloning, together with the small number of clones obtained.

In this work we have studied the diversity of deltasatellites and associated sweepoviruses infecting *I. indica* plants by resampling the southern Spain populations where they were detected by the first time and expanding the analysis to

other geographical areas. We used the RCA technique (which allows amplification without previous knowledge of nucleotide sequence) which can reveal the actual variability present in natural populations. However, the observed genetic variability was not higher in the deltasatellite sequences obtained in this study. Furthermore, there was no evidence for recombination in the deltasatellite sequences.

This study also provided insights into the nature of the helper sweepoviruses. Most of the sweepovirus isolates belonged to the species SPLCV. In addition, two isolates from Cádiz and three from Málaga were shown to belong to the species SPMV. SPMV was previously reported only from Brazil (Paprotka et al., 2010a) and South Africa (Esterhuizen et al., 2012), and this is the first report of this virus in Spain. SPMV isolates reported here were shown to be recombinants involving genome fragments from SPLCV.

SPLCV and SPMV populations reported in this work showed a genetic variability equivalent to that reported for other begomovirus populations (Duffy and Holmes, 2008, 2009;Ge et al., 2007; Lima et al., 2013; Ramos-Sobrinho et al., 2014; Rocha et al., 2013). Numerous studies have shown that populations of geminiviruses are highly recombinant (García-Andrés et al., 2007; Lefeuvre et al., 2009; Lefeuvre et al., 2007b; Lima et al., 2013; Martin et al., 2011; Ramos-Sobrinho et al., 2014; Rocha et al., 2013; van der Walt et al., 2009) and that recombination could explain the higher variability of begomoviruses infecting wild hosts in relation to those infecting crops (Lima et al., 2013). We found 21 unique recombination events in the sweepovirus genomes, and most of the recombination breakpoints occur in the IR and in the middle of the *Rep* gene. Similarly, two recombination breakpoints (events 13 and 15) involving the interface between genes *CP* and *REn* were identified. Lefeuvre et al. (2007a) described the presence of recombination cold-spots within the

V2 gene and the third quarter of the *CP* gene of begomoviruses. Here, however, we detected the occurrence of recombination breakpoints in the first half of the *CP* gene and within the *V2* gene. Albuquerque et al. (2012) also detected recombination breakpoints in the first half of the *CP* gene. These results are consistent with those obtained from geminivirus recombination analyses, which show that the *Rep* gene, the IR and the interface between genes *CP* and *REn* are recombination hot spots (Albuquerque et al., 2012; García-Andrés et al., 2007; Lefeuvre et al., 2007a; Lefeuvre et al., 2009; Lefeuvre et al., 2007b). Some sweepovirus recombination events have been previously described (Albuquerque et al., 2012; Lozano et al., 2009; Zhang and Ling, 2011) and most were observed in our study.

In this study we found a sweepovirus-deltasatellite chimera in an *I. indica* plant with the typical size of a deltasatellite. This molecule was the result of a recombination event between an isolate of SPLCV closely related to isolate LM7861 and a deltasatellite closely related to LM7629, both present in the same *I. indica* sample. Lozano et al. (2016) also found a similar, but longer, recombinant chimeric molecule. This chimera was 2598 bp in length and about 70% of its length had identity with an isolate of the sweepovirus SPLCCaV that was obtained from the same sample (the region contained the *V2* and *CP* genes, a truncated *REn* gene and the complete IR including the iterons) and the remaining was found to correspond to most of the sequence of a sweepovirus-associated deltasatellite.

Due to the vegetative propagation of many species of the genus *Ipomoea*, including *I. indica* and sweet potato, virus accumulation and perpetuation is a common phenomenon which allows frequent mixed infections and facilitates the occurrence of recombination (Valverde et al., 2007). This can be a rapid process to create new genomes with adaptive advantages, which could accelerate their

evolution and favor the expansion of the host range and, therefore, the emergence of novel diseases (Berrie et al., 2001; Lefeuvre and Moriones, 2015; Monci et al., 2002). The same process that occurs to originate viral recombinants may have generated the virus-satellite chimera found in this study.

Phylogenetic grouping by geographic origin and by the plant host from which they were isolated was observed for the deltasatellites associated with NW bipartite begomoviruses in Cuba (Fiallo-Olivé et al., 2012). Lozano et al. (2016), studying the phylogenetic relationships of the deltasatellites associated to sweepoviruses in Spain, observed that, although they are closely related, they could be grouped in two clusters related to geographic origin (either mainland Spain or the Canary Islands). Within the satellites from mainland Spain, the isolates from *I. indica* grouped together. The present study did not reveal a clear grouping of deltasatellites according to geographical origin or host plant (Figure 5). However, two of the new isolates from Málaga grouped with isolates previously described from Canary Islands within a well-defined cluster. Absence of phylogenetic differentiation are in agreement with the low nucleotide diversity data.

References

Albuquerque, L.C., Inoue-Nagata, A.K., Pinheiro, B., Resende, R.O., Moriones, E., Navas-Castillo, J., 2012. Genetic diversity and recombination analysis of sweepoviruses from Brazil. Virol J 9, 241.

Albuquerque, L.C., Inoue-Nagata, A.K., Pinheiro, B., Ribeiro, S.D., Resende, R.O., Moriones, E., Navas-Castillo, J., 2011. A novel monopartite begomovirus infecting sweet potato in Brazil. Arch Virol 156, 1291-1294.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol 215, 403-410.

Banks, G., Bedford, I., Beitia, F., Rodriguez-Cerezo, E., Markham, P., 1999. A novel geminivirus of *Ipomoea indica* (Convolvulacae) from Southern Spain. Plant Dis 83, 486-486.

Berrie, L.C., Rybicki, E.P., Rey, M.E.C., 2001. Complete nucleotide sequence and host range of South African cassava mosaic virus: Further evidence for recombination amongst begomoviruses. J Gen Virol 82, 53-58.

Briddon, R., Bull, S., Bedford, I., 2006. Occurrence of *Sweet potato leaf curl virus* in Sicily. Plant Pathol 55, 286-286.

Briddon, R.W., 2003. Cotton leaf curl disease, a multicomponent begomovirus complex. Molecular Plant Pathol 4, 427-434.

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., 2003. Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. Virology 312, 106-121.

Briddon, R.W., Bull, S.E., Amin, I., Mansoor, S., Bedford, I.D., Rishi, N., Siwatch, S.S., Zafar, Y., Abdel-Salam, A.M., Markham, P.G., 2004. Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. Virology 324, 462-474.

Briddon, R.W., Mansoor, S., Bedford, I.D., Pinner, M.S., Saunders, K., Stanley, J., Zafar, Y., Malik, K.A., Markham, P.G., 2001. Identification of DNA components required for induction of cotton leaf curl disease. Virology 285, 234-243.

Briddon, R.W., Stanley, J., 2006. Subviral agents associated with plant single-stranded DNA viruses. Virology 344, 198-210.

Brown, J.K., Fauquet, C.M., Briddon, R.W., Zerbini, F.M., Moriones, E., Navas-Castillo, J., 2012. Family *Geminiviridae*, in: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, UK, pp. 351-373.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-Olivé, E., Briddon, R.W., Hernandez-Zepeda, C., Idris, A., Malathi, V.G., Martin, D.P., Rivera-Bustamante, R., Ueda, S., Varsani, A., 2015. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Arch Virol 160, 1593-1619.

Dry, I.B., Krake, L.R., Rigden, J.E., Rezaian, M.A., 1997. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. Proc Natl Acad Sci, USA 94, 7088-7093.

Duffy, S., Holmes, E.C., 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus *Tomato yellow leaf curl virus*. J Virol 82, 957-965.

Duffy, S., Holmes, E.C., 2009. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J Gen Virol 90, 1539-1547.

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinf 5, 1-19.

Esterhuizen, L.L., van Heerden, S.W., Rey, M.E.C., van Heerden, H., 2012. Genetic identification of two sweet-potato-infecting begomoviruses in South Africa. Arch Virol 157, 2241-2245.

Fiallo-Olivé, E., Martinez-Zubiaur, Y., Moriones, E., Navas-Castillo, J., 2012. A novel class of DNA satellites associated with New World begomoviruses. Virology 426, 1-6.

Fuentes, S., Salazar, L., 2003. First report of *Sweet potato leaf curl virus* in Peru. Plant Dis 87, 98-98.

García-Andrés, S., Tomas, D.M., Sanchez-Campos, S., Navas-Castillo, J., Moriones, E., 2007. Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. Virology 365, 210-219.

Ge, L.M., Zhang, J.T., Zhou, X.P., Li, H.Y., 2007. Genetic structure and population variability of *Tomato yellow leaf curl China virus*. J Virol 81, 5902-5907.

Glick, E., Zrachya, A., Levy, Y., Mett, A., Gidoni, D., Belausov, E., Citovsky, V., Gafni, Y., 2008. Interaction with host SGS3 is required for suppression of RNA silencing by *Tomato yellow leaf curl virus* V2 protein. Proc Natl Acad Sci, USA 105, 157-161.

Haible, D., Kober, S., Jeske, H., 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. J Virol Meth 135, 9-16.

Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B., Nagata, T., 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. J Virol Meth 116, 209-211.

Jeske, H., Kober, S., Schäfer, B., Strohmeier, S., 2014. Circomics of Cuban geminiviruses reveals the first alpha-satellite DNA in the Caribbean. Virus Genes 49, 312-324.

Lefeuvre, P., Lett, J.M., Reynaud, B., Martin, D.P., 2007a. Avoidance of protein fold disruption in natural virus recombinants. PLoS Pathog 3, e181.

Lefeuvre, P., Lett, J.M., Varsani, A., Martin, D.P., 2009. Widely conserved recombination patterns among single-stranded DNA viruses. J Virol 83, 2697-2707.

Lefeuvre, P., Martin, D.P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., Becker, N., Reynaud, B., Lett, J.M., 2007b. Begomovirus 'melting pot' in the

south-west Indian Ocean islands: molecular diversity and evolution through recombination. J Gen Virol 88, 3458-3468.

Lefeuvre, P., Moriones, E., 2015. Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. Curr Opin Virol 10, 14-19.

Legg, J., Fauquet, C., 2004. Cassava mosaic geminiviruses in Africa. Plant Mol Biol 56, 585-599.

Leke, W.N., Mignouna, D.B., Brown, J.K., Kvarnheden, A., 2015. Begomovirus disease complexes: emerging threats to vegetable production systems of West and Central Africa. Agriculture & Food Security 4, 1.

Lima, A.T.M., Sobrinho, R.R., Gonzalez-Aguilera, J., Rocha, C.S., Silva, S.J.C., Xavier, C.A.D., Silva, F.N., Duffy, S., Zerbini, F.M., 2013. Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. J Gen Virol 94, 418-431.

Lotrakul, P., Valverde, R., Clark, C., Sim, J., De La Torre, R., 1998. Detection of a geminivirus infecting sweet potato in the United States. Plant Dis 82, 1253-1257.

Lotrakul, P., Valverde, R.A., Clark, C.A., Fauquet, C., 2003. Properties of a begomovirus isolated from sweet potato [*Ipomoea batatas* (L.) Lam.] infected with *Sweet potato leaf curl virus*. Rev Mex Fitopatol 21, 128-136.

Lozano, G., Trenado, H.P., Fiallo-Olivé, E., Chirinos, D., Geraud-Pouey, F., Briddon, R.W., Navas-Castillo, J., 2016. Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, *Geminiviridae*) - definition of a distinct class of begomovirus-associated satellites. Frontiers in Microbiology 7, doi 10.3389/fmicb.2016.00162.

Lozano, G., Trenado, H.P., Valverde, R.A., Navas-Castillo, J., 2009. Novel begomovirus species of recombinant nature in sweet potato (*Ipomoea batatas*) and *Ipomoea indica*: taxonomic and phylogenetic implications. J Gen Virol 90, 2550-2562.

Luan, Y., Zhang, J., An, L., 2006. First report of *Sweet potato leaf curl virus* in China. Plant Dis 90, 1111-1111.

Luan, Y.S., Zhang, J., Liu, D.M., Li, W.L., 2007. Molecular characterization of *Sweet potato leaf curl virus* isolate from China (SPLCV-CN) and its phylogenetic relationship with other members of the *Geminiviridae*. Virus Genes 35, 379-385.

Mansoor, S., Khan, S.H., Bashir, A., Saeed, M., Zafar, Y., Malik, K.A., Briddon, R., Stanley, J., Markham, P.G., 1999. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259, 190-199.

Martin, D.P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.Y., Dijoux, B., Vincent, C., Reynaud, B., Lett, J.M., 2011. Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. PLoS Pathog 7, e1002203.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol 1, vev003.

Miano, D.W., LaBonte, D.R., Clark, C.A., Valverde, R.A., Hoy, M.W., 2006. First report of a begomovirus infecting sweetpotato in Kenya. Plant Dis 90, 832-832.

Monci, F., Sanchez-Campos, S., Navas-Castillo, J., Moriones, E., 2002. 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 303, 317-326.

Morales, F.J., Anderson, P.K., 2001. The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Arch Virol 146, 415-441.

Muhire, B.M., Varsani, A., Martin, D.P., 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9, e108277.

Navas-Castillo, J., Fiallo-Olivé, E., Sánchez-Campos, S., 2011. Emerging virus diseases transmitted by whiteflies. Annu Rev Phytopathol 49, 219-248.

Ng, T.F.F., Duffy, S., Polston, J.E., Bixby, E., Vallad, G.E., Breitbart, M., 2011. Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies. PLoS One 6, e19050.

Padidam, M., Beachy, R.N., Fauquet, C.M., 1996. The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. Virology 224, 390-404.

Paprotka, T., Boiteux, L.S., Fonseca, M.E.N., Resende, R.O., Jeske, H., Faria, J.C., Ribeiro, S.G., 2010a. Genomic diversity of sweet potato geminiviruses in a Brazilian germplasm bank. Virus Res 149, 224-233.

Paprotka, T., Metzler, V., Jeske, H., 2010b. The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. Virology 404, 148-157.

Polston, J.E., Anderson, P.K., 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. Plant Dis 81, 1358-1369.

Prasanth, G., Hegde, V., 2008. Occurrence of *Sweet potato feathery mottle virus* and Sweet potato leaf curl Georgia virus on sweet potato in India. Plant Dis 92, 311-311.

Ramos-Sobrinho, R., Xavier, C.A.D., Pereira, H.M.B., Lima, G.S.A., Assunção, I.P., Mizubuti, E.S.G., Duffy, S., Zerbini, F.M., 2014. Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. J Gen Virol 95, 2540-2552.

Rocha, C.S., Castillo-Urquiza, G.P., Lima, A.T.M., Silva, F.N., Xavier, C.A.D., Hora-Junior, B.T., Beserra-Junior, J.E.A., Malta, A.W.O., Martin, D.P., Varsani, A., Alfenas-Zerbini, P., Mizubuti, E.S.G., Zerbini, F.M., 2013. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. J Virol 87, 5784-5799.

Rojas, M.R., Hagen, C., Lucas, W.J., Gilbertson, R.L., 2005. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu Rev Phytopathol 43, 361-394.

Romay, G., Chirinos, D., Geraud-Pouey, F., Desbiez, C., 2010. Association of an atypical alphasatellite with a bipartite New World begomovirus. Arch Virol 155, 1843-1847.

Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP: DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Rybicki, E.P., 1994. A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. Arch Virol 139, 49-77.

Saunders, K., Bedford, I.D., Briddon, R.W., Markham, P.G., Wong, S.M., Stanley, J., 2000. A unique virus complex causes *Ageratum yellow vein disease*. Proc Natl Acad Sci, USA 97, 6890-6895.

Saunders, K., Stanley, J., 1999. A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: Evidence for interfamilial recombination between plant DNA viruses. Virology 264, 142-152.

Sivalingam, P.N., Varma, A., 2012. Role of betasatellite in the pathogenesis of a bipartite begomovirus affecting tomato in India. Arch Virol 157, 1081-1092.

Sserubombwe, W.S., Briddon, R.W., Baguma, Y.K., Ssemakula, G.N., Bull, S.E., Bua, A., Alicai, T., Omongo, C., Otim-Nape, G.W., Stanley, J., 2008. Diversity of begomoviruses associated with mosaic disease of cultivated cassava (*Manihot esculenta* Crantz) and its wild relative (*Manihot glaziovii* Mull. Arg.) in Uganda. J Gen Virol 89, 1759-1769.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30, 2725-2729.

Trenado, H.P., Orilio, A.F., Marquez-Martin, B., Moriones, E., Navas-Castillo, J., 2011. Sweepoviruses cause disease in sweet potato and related *Ipomoea* spp.: Fulfilling Koch's postulates for a divergent group in the genus *Begomovirus*. PLoS One 6.

Valverde, R.A., Clark, C.A., Valkonen, J.P., 2007. Viruses and virus disease complexes of sweetpotato. Plant Vir 1, 116-126.

van der Walt, E., Rybicki, E.P., Varsani, A., Polston, J.E., Billharz, R., Donaldson, L., Monjane, A.L., Martin, D.P., 2009. Rapid host adaptation by extensive recombination. J Gen Virol 90, 734-746.

Wasswa, P., Otto, B., Maruthi, M., Mukasa, S., Monger, W., Gibson, R., 2011. First identification of a sweet potato begomovirus (sweepovirus) in Uganda: characterization, detection and distribution. Plant Pathol 60, 1030-1039.

Were, H.K., Winter, S., Maiss, E., 2004. Viruses infecting cassava in Kenya. Plant Dis 88, 17-22.

Xie, Y., Wu, P.J., Liu, P., Gong, H.R., Zhou, X.P., 2010. Characterization of alphasatellites associated with monopartite begomovirus/betasatellite complexes in Yunnan, China. Virol J 7:178.

Yang, X.L., Xie, Y., Raja, P., Li, S.Z., Wolf, J.N., Shen, Q.T., Bisaro, D.M., Zhou, X.P., 2011. Suppression of methylation-mediated transcriptional gene silencing by betaC1-SAHH protein interaction during geminivirus-betasatellite infection. PLoS Pathog 7: e1002329.

Zhang, S.C., Ling, K.S., 2011. Genetic diversity of sweet potato begomoviruses in the United States and identification of a natural recombinant between *Sweet*

potato leaf curl virus and *Sweet potato leaf curl Georgia virus*. Arch Virol 156, 955-968.

Zhou, X., 2013. Advances in understanding begomovirus satellites. Annu Rev Phytopathol 51, 357-381.

Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z., Fauquet, C.M., 2003. Characterization of DNA beta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. J Gen Virol 84, 237-247.

Table 1. Sweepovirus and associated deltasatellite sequences reported in this study.

Province	Sample code	Sampling date	Geographical coordinates	Sweepovirus	Satellite	Sweepovirus clones	Satellite clones	Symptoms
Cádiz	Ii-56	09/15/15	36° 04'N 5° 30'W	+	+	LM7842		-
	Ii-57	09/15/15	36° 04'N 5° 30'W	+	+		LM7785 LM7786	-
	Ii-58	09/15/15	36° 04'N 5° 30'W	+	+	LM7844	LM7787 LM7788	-
Málaga	Ii-01	02/03/15	36° 45.444'N 4° 02.654'W	+	-			-
	Ii-02	02/03/15	36° 45.452'N 4° 02.670'W	+	-			-
	Ii-03	02/03/15	36° 45.453'N 4° 02.687'W	+	+	LM7275	LM7315 LM7316	-
	Ii-04	02/03/15	36° 44.794'N 4° 03.014'W	+	-			-
	Ii-05	02/03/15	36° 44.800'N 4° 03.013'W	+	-			-
	Ii-06	02/03/15	36° 44.053'N 4° 06.920'W	+	+	LM7317	LM7277 LM7278	-
	Ii-07	02/03/15	36° 44.041'N 4° 06.927'W	+	+	LM7283	LM7281 LM7366	YV, LC
	Ii-08	02/03/15	36° 44.084'N 4° 06.930'W	+	+	LM7286	LM7284 LM7285	-
	Ii-09	02/03/15	36° 46.002'N 4° 06.455'W	+	+	LM7327 LM7328	LM7287 LM7288	YV
	Ii-10	02/03/15	36° 45.994'N 4° 06.460'W	+	+	LM7294	LM7291 LM7293	YV, LC
	Ii-11	02/03/15	36° 45.990'N 4° 06.463'W	+	+	LM7295	LM7367 LM7368 LM7369	-
	Ii-12	02/03/15	36° 47.620'N 4° 07.252'W	+	+	LM7297	LM7296 LM7318	YV, LC
	Ii-13	02/03/15	36° 47.640'N 4° 07.254'W	+	+	LM7300	LM7298 LM7299	-
	Ii-14	02/03/15	36° 47.954'N 4° 07.369'W	+	+	LM7303	LM7301 LM7302	-
	Ii-15	02/03/15	36° 47.952'N 4° 07.362'W	-	-			-
	Ii-16	02/03/15	36° 49.881'N 4° 07.555'W	-	-			-
	Ii-17	02/03/15	36° 49.896'N 4° 07.538'W	-	-			-
	Ii-18	05/22/15	36° 44.748'N 4° 02.216'W	+	+		LM7539 LM7540	-
	Ii-19	05/22/15	36° 44.629'N 4° 02.164'W	+	+		LM7541 LM7542	YV

Ii-20	05/22/15	36° 44.631'N 4° 01.113'W	+	+		LM7543 LM7544	-
Ii-21	05/22/15	36° 43.967'N 3° 56.509'W	+	+	LM7568	LM7545 LM7546	LC
Ii-22	05/22/15	36° 43.963'N 3° 56.510'W	+	+	LM7569	LM7547 LM7548 LM7549	LC
Ii-23	05/22/15	36° 44.162'N 3° 55.942'W	+	+	LM7570	LM7550 LM7551 LM7552	-
Ii-24	05/22/15	36° 44.189'N 3° 55.878'W	-	-			-
Ii-25	05/22/15	36° 44.191'N 3° 55.870'W	+	+	LM7606	LM7553 LM7554	-
Ii-26	05/22/15	36° 44.668'N 3° 52.526'W	+	+	LM7861	LM7629 LM7556	-
Ii-27	05/22/15	36° 44.660'N 3° 52.539'W	+	+	LM7572	LM7557 LM7558 LM7559	-
Ii-28	05/22/15	36° 44.710'N 3° 52.531'W	+	+	LM7607	LM7560 LM7561	-
Ii-29	05/22/15	36° 44.883'N 3° 52.261'W	-	-			-
Ii-30	05/22/15	36° 44.877'N 3° 52.248'W	+	+	LM7573	LM7580 LM7581	-
Ii-31	05/22/15	36° 44.880'N 3° 52.259'W	+	+	LM7608	LM7845 LM7583	-
Ii-32	05/22/15	36° 45.466'N 3° 51.732'W	+	+		LM7584	-
Ii-33	05/22/15	36° 45.468'N 3° 51.750'W	+	+		LM7586 LM7630	-
Ii-34	05/22/15	36° 45.291'N 3° 51.200'W	+	+	LM7609	LM7587 LM7634	-
Ii-35	05/22/15	36° 45.294'N 3° 51.115'W	+	+	LM7865	LM7588 LM7589 LM7590	-
Ii-36	05/22/15	36° 44.448'N 3° 45.639'W	+	+		LM7591 LM7592	-
Ii-37	05/22/15	36° 44.279'N 3° 45.639'W	+	-			-
Ii-38	05/22/15	36° 44.374'N 3° 41.753'W	+	+	LM7610	LM7594 LM7595	-
Ii-39	05/22/15	36° 44.367'N 3° 41.736'W	+	+		LM7596 LM7597	-
Ii-40	05/22/15	36° 44.775'N 3° 36.233'W	+	+	LM7627	LM7598 LM7599	-
Ii-41	05/22/15	36° 44.807'N 3° 36.210'W	+	+		LM7600 LM7601 LM7602	-
Ii-42	05/22/15	36° 42.543'N 3° 29.545'W	+	+	LM7628	LM7603 LM7682	-
Ii-43	05/22/15	36° 42.486'N 3° 29.541'W	+	+	LM7681	LM7604 LM7605	-
Ii-44	05/22/15	36° 49.914'N 3° 31.106'W	+	-			-
Ii-45	08/04/15	36° 45.728'N 3° 53.302'W	+	+		LM7773	-

	Ii-46	08/04/15	36° 45.692'N 3° 53.302'W	-	-			-
	Ii-47	08/04/15	36° 45.839'N 3° 53.291'W	+	+	LM7805	LM7774 LM7790	-
	Ii-48	08/04/15	36° 45.858'N 3° 53.307'W	+	+		LM7776 LM7777	-
	Ii-64	10/13/15	36° 44.686'N 4° 5.275'W	+	+	LM7868	LM7853	-
	Ii-65	10/13/15	36° 44.686'N 4° 5.275'W	+	+		LM7854 LM7856	-
Granada	Ii-49	08/04/15	36° 44.182'N 3° 40.806'W	+	+		LM7778	-
	Ii-50	08/04/15	36° 44.612'N 3° 33.341'W	+	+		LM 7759 LM7791	-
	Ii-51	08/04/15	36° 43.562'N 3° 32.320'W	+	+	LM7806	LM7760	-
	Ii-52	08/04/15	36° 44'N 3° 32'W	+	+		LM7761	-
	Ii-53	08/04/15	36° 44'N 3° 32'W	+	+		LM7780 LM7781 LM7782	-
	Ii-54	09/08/15	36° 27'N 5° 04'W	+	-			-
	Ii-55	09/08/15	36° 27'N 5° 04'W	+	-			-
Murcia	Ii-59	10/07/15	37°55'58.2"N 1°07'54.4"W	+	-			-
	Ii-60	10/07/15	38°03'13.0"N 1°13'03.2"W	+	-			-
	Ii-61	10/07/15	38°03'13.0"N 1°13'03.2"W	+	+		LM7851 LM7852	LC
	Ii-62	10/07/15	38°03'13.0"N 1°13'03.2"W	+	-			YV
	Ii-63	10/07/15	38°03'13.0"N 1°13'03.2"W	+	-			-

+, presence of sweepoviruses or satellites; -, absence of sweepoviruses, satellites or symptoms; YV, yellow veins; LC, leaf curling

Table 2. Genetic variability of the sweepoviruses and associated deltasatellite genomes characterized in this study.

Population/genomic region	**No. of sequences**	**Nucleotide diversity, π**
Deltasatellite (Total)	92	0.03715 (± 0.00113)
Cádiz	4	0.01777 (± 0.00810)
Granada	8	0.02846 (± 0.00470)
Málaga	78	0.03903 (± 0.00115)
Murcia	2	0.01560 (± 0.00780)
Sweepovirus (Total)	31	0.07065 (± 0.00436)
Sweet potato mosaic virus (SPMV)	3	0.05688 (± 0.01996)
CP		0.02614 (± 0.00850)
Rep		0.07062 (± 0.02732)
V1		0.03899 (± 0.01293)
REn		0.07763 (± 0.02324)
TrAP		0.04889 (± 0.01692)
C4		0.04822 (± 0.01416)
IR		0.06136 (± 0.02532)
Sweet potato leaf curl virus (SPLCV)	28	0.06832 (± 0.00484)
CP		0.03525 (± 0.00202)
Rep		0.09444 (± 0.00847)
V1		0.07641 (± 0.00306)
REn		0.01929 (± 0.00447)
TrAP		0.02256 (± 0.00284)
C4		0.02319 (± 0.00271)
IR		0.12566 (± 0.00942)

Table 3. Recombination events detected in the sweepovirus genomes analyzed in this study.

Event	Recombinant	Recombination breakpoints[1]		Parents[2]		Method	*P-value*[3]
		Begin	End	Minor	Major		
1	LM7681_ii43	2758	2186	SPLCV_EU839578	SPLCCV_EF456742	RGBMCS**3**	3.598 X 10^{-43}
	LM7568_ii21						
	LM7470_ii23						
	LM7569_ii22						
	SPLCV_EU839576						
2	LM7573_ii30	2692	2178	LM7861_ii26	SPLCCV_EF456742	RGBMCS**3**	7.343 X 10^{-83}
3	LM7606_ii25	2230	2729	LM7865_ii35	LM7608_ii31	RGBMCS**3**	7.527 X 10^{-37}
	LM7294_ii10						
	LM7295_ii11						
	LM7327_ii9						
4	LM7805_ii47	2763	2168	LM7609_ii34	SPMV_FJ969831	GBMS**3**	2.535 X 10^{-49}
5	LM7286_ii8	49	2138	LM7607_ii28	SPMV_FJ969831	RGBMS**3**	9.735 X 10^{-48}
6	SPLCV_EU839576	374	2186*	LM7303_ii14	SPLCV_EU839579	RGBMC**S**3	5.916 X 10^{-39}
	LM7568_ii21						
	LM7470_ii23						
	LM7610_ii38						
	LM7681_ii43						
	LM7569_ii22						
	SPLCV_EU839578						
7	LM7865_ii35	508	1626	LM7608_ii31	SPLCV_EU839579	RGBMS**3**	4.715 X 10^{-47}
8	LM7607_ii28	2233	2697	LM7865_ii35	LM7609_ii34	RGBMCS**3**	1.018 X 10^{-32}
	SPLCV_EU839577						
9	LM7842_ii56	96	1927	LM7608_ii31	SPMV_FJ969831	RGBM**C**S3	2.723 X 10^{-22}
	LM7844_ii58						
10	LM7607_ii28	87	477*	SPLCV_EF456746	7868_ii64	**R**GBMCS3	8.219 X 10^{-21}
	LM7286_ii8						
	LM7573_ii30						
	LM7609_ii34						
	LM7805_ii47						
	LM7861_ii26						
	SPLCV_FJ151200						
11	LM7610_ii38	419*	1624	LM7627_ii40	SPLCV_EU839579	GBM**3**	6.929 X 10^{-28}
12	SPLCV_EU839578	2220*	2698	SPLCCNV_DQ512731	SPLCCV_EF456745	RGBMCS**3**	2.126 X 10^{-25}
	LM7610_ii38						
	LM7865_ii35						
13	LM7572_ii27	482*	1154	SPLCV_FJ151200	SPMV_FJ969831	**R**GMCS3	2.111 X 10^{-16}
	LM7275_ii3						
14	LM7627_ii40	163	321	SPLCV_EU839579	LM7297_ii12	**R**GBMC3	1.509 X 10^{-19}
15	LM7470_ii23	1080	1166*	SPLCV_EU839579	LM7861_ii26	**R**GMC3	8.116 X 10^{-11}
	LM7842_ii56						
16	LM7294_ii10	38*	243	SPMV_FJ969831	LM7606_ii25	**R**GBMCS3	7.785 X 10^{-11}

	LM7295_ii11						
	LM7327_ii9						
17	LM7295_ii11	322*	492	SPLCV_EU839579	LM7608_ii31	**R**GBMCS3	1.762 X 10^{-11}
18	LM7327_ii9	1868	1988	SPMV_FJ969831	LM7568_ii21	**R**GB3	1.368 X 10^{-06}
19	LM7569_ii22	373*	504*	LM7865_ii35	LM7608_ii31	RG**B**S3	9.727 X 10^{-05}
	LM7568_ii21						
20	LM7572_ii27	1322	1785	LM7861_ii26	SPMV_FJ969831	**R**BMC3	4.518 X 10^{-05}
21	LM7572_ii27	222	480*	LM7806_ii51	SPMV_FJ969831	**R**GBS3	1.164 X 10^{-06}

[1]Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise. (*), Breakpoints could not be accurately located.

[2]Recombination events and their putative parental viruses were identified using the Rdp (R), Geneconv (G), Boostcan (B), Maxichi (M), Chimaera (C), Siscan (S) and 3Seq (3), modules in RDP4.

[3]The reported *p*-value is from the method in bold and is the lowest *p*-value calculated for the featured event.

Supplementary Table S1. Sweepovirus sequences retrieved from GenBank and used for pairwise sequence comparisons and phylogenetic analysis.

Sweepovirus	Acronym	Host	Location	Genbank access number	Reference
Sweet potato leaf curl Canary virus	SPLCCV	*Ipomoea batatas*	Canary Islands	EF456742	Lozano et al. (2009)
Sweet potato leaf curl Canary virus	SPLCCV	*Ipomoea batatas*	Canary Islands	EF456745	Lozano et al. (2009)
Sweet potato leaf curl Canary virus	SPLCCV	*Ipomoea batatas*	Canary Islands	EU856365	Lozano et al. (2009)
Sweet potato leaf curl Canary virus	SPLCCV	*Ipomoea batatas*	Canary Islands	FJ529203	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EF456746	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EF456741	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EF456743	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EF456744	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EU856364	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Canary Islands	EU856366	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea batatas*	Málaga	EU839579	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea indica*	Málaga	FJ151200	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea indica*	Málaga	EU839576	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea indica*	Málaga	EU839577	Lozano et al. (2009)
Sweet potato leaf curl virus	SPLCV	*Ipomoea indica*	Málaga	EU839578	Lozano et al. (2009)
Sweet potato leaf curl China virus	SPLCCNV	*Ipomoea batatas*	China	DQ512731	Luan et al. (2007)
Sweet potato leaf curl Henan virus	SPLCHnV	*Ipomoea batatas*	China-Henan	KC907406	Liu et al. (2014)
Sweet potato leaf curl Sichuan virus 1	SPLCSiV-1	*Ipomoea batatas*	China-Sichuan	KC488316	Liu et al. (2013)
Sweet potato leaf curl Sichuan virus 2	SPLCSiV-2	*Ipomoea batatas*	China-Sichuan	KF156759	Unpublished
Sweet potato leaf curl Georgia virus	SPLCGV	*Ipomoea batatas*	Unite States-Georgia	AF326775	Lotrakul et al. (2003)
Sweet potato leaf curl South Carolina virus	SPLCSCV	*Ipomoea batatas*	United States-South Carolina	HQ333144	Zhang et al. (2011)
Sweet potato leaf curl Uganda virus	SPLCUV	*Ipomoea batatas*	Uganda	FR751068	Wasswa et al. (2011)
Sweet potato mosaic virus	SPMV	*Ipomoea batatas*	Brazil-Brasilia	FJ969831	Paprotka et al. (2010)
Sweet potato leaf curl Sao Paulo virus	SPLCSPV	*Ipomoea batatas*	Brazil-São Paulo	HQ393477	Albuquerque et al. (2012)

Supplementary Table S2. Satellite DNA sequences retrieved from GenBank and used for pairwise sequence comparisons and phylogenetic analysis.

Satellite DNA	Host	Location	Genbank access number	Reference
SBG32	*Ipomoea batatas*	Canary Islands	FJ914391.1	Lozano et al. (2016)
SBG51	*Ipomoea batatas*	Canary Islands	FJ914390.1	Lozano et al. (2016)
SBG57	*Ipomoea batatas*	Canary Islands	FJ914397.1	Lozano et al. (2016)
SBG58	*Ipomoea batatas*	Canary Islands	FJ914398.1	Lozano et al. (2016)
SBG59	*Ipomoea batatas*	Canary Islands	FJ914403.1	Lozano et al. (2016)
SBG52	*Ipomoea batatas*	Málaga	FJ914392.1	Lozano et al. (2016)
SBG54	*Ipomoea batatas*	Málaga	FJ914394.1	Lozano et al. (2016)
SBG53	*Ipomoea batatas*	Málaga	FJ914393.1	Lozano et al. (2016)
SBG55	*Ipomoea batatas*	Málaga	FJ914395.1	Lozano et al. (2016)
SBG56	*Ipomoea batatas*	Málaga	FJ914396.1	Lozano et al. (2016)
SBGB3-5	*Ipomoea batatas*	Málaga	FJ914404.1	Lozano et al. (2016)
SBGB3-6	*Ipomoea batatas*	Málaga	FJ914405.1	Lozano et al. (2016)
SI3C-5	*Ipomoea indica*	Málaga	FJ914400.1	Lozano et al. (2016)
SI3C-3	*Ipomoea indica*	Málaga	FJ914399.1	Lozano et al. (2016)
SI3D-11	*Ipomoea indica*	Málaga	FJ914401.1	Lozano et al. (2016)
SI3D-12	*Ipomoea indica*	Málaga	FJ914402.1	Lozano et al. (2016)
1764E13	*Merremia dissecta*	Venezuela	KF716173.1	Lozano et al. (2016)
IN-Cb1	*Croton bonplandianus*	India	AJ968684.1	Unpublished
PH-Mc1	*Malvastrum coromandelianum*	Philippines	KF433066.1	Unpublished
ToLCV- sat	Tomato	Australia	U74627.1	Dry et al. (1997)
177H1	*Malvastrum coromandelianum*	Cuba	JN986808.1	Fiallo-Olivé et al. (2012)
404N1	*Sidastrum micranthum*	Cuba	JN819495.1	Fiallo-Olivé et al. (2012)

Figure legends

Figure 1. Plants of blue morning glory (*Ipomoea indica*).

Figure 2. Sampling sites in southern Spain (Cádiz, Granada, Málaga and Murcia). Samples containing sweepoviruses (yellow markers), sweepoviruses plus deltasatellites (purple markers) and not containing virus (pink markers) are indicated in the image.

Figure 3. Pairwise sequence identity matrices of the sweepoviruses obtained in this work and the additional closely related sweepoviruses, obtained with the program Sequence Demarcation Tool (SDT) v. 1.2.

Figure 4. Pairwise sequence identity matrices of the deltasatellites described in this study plus the most closely related deltasatellites, obtained with the program Sequence Demarcation Tool (SDT) v. 1.2.

Figure 5. Phylogenetic tree showing the relationship among the deltasatellite molecules described in this work and other deltasatellites. Full names of the sequences retrieved from GenBank are found in Supplementary Table 2. The tree was constructed using the Neighbor-Joining method with MEGA 6. Only bootstrap values higher than 60% are shown. Samples were obtained from Cádiz (red circles), Canary Islands (green circles), Granada (purple circles), Málaga (yellow circles) and Murcia (blue circles). Some plants showed symptoms of yellow vein (YV, blue squares) and leaf curling (LC, pink squares).

Figure 6. Phylogenetic tree showing the relationships among the sweepoviruses obtained in this work and other sweepoviruses. Full names of the sequences retrieved from GenBank are found in Supplementary Table 1. The tree was constructed using the Neighbor-Joining method with MEGA 6. Only bootstrap values higher than 60% are shown. Samples were obtained from Cádiz (red circles), Canary Islands (green circles), Granada (purple circles) and Málaga (yellow circles).

Figure 7. Schematic representation of the chimeric sweepovirus-DNA satellite molecule identified in blue morning glory sample ii26. The stem-loop containing the nanonucleotide TAATATTAC, the regions corresponding to the satellite sequence (blue color) and to the sweepovirus sequence (purple color), and the recombination sites are indicated.

Figure 1.

Figure 2.


Murcia
Andaluzia
ibraltar
Sweepoviruses
Sweepoviruses plus satellite
No virus

Figure 3.

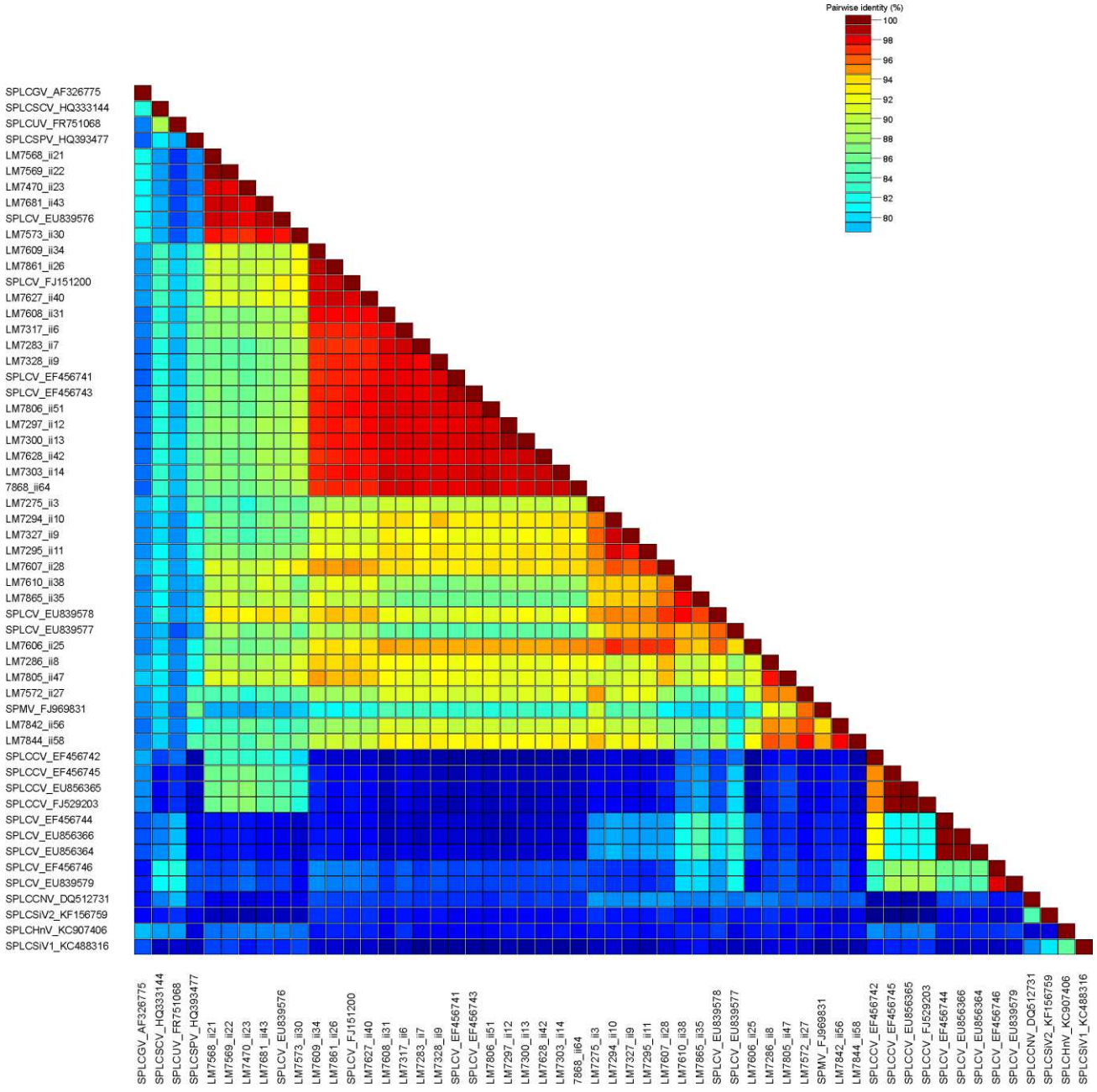

Pairwise identity (%)
100
98
96
94
92
90
88
86
84
82
80
SPLCGV_AF326775
SPLCSCV_HQ333144
SPLCUV_FR751068
SPLCSPV_HQ393477
LM7568_ii21
LM7569_ii22
LM7470_ii23
LM7681_ii43
SPLCV_EU839576
LM7573_ii30
LM7609_ii34
LM7861_ii26
SPLCV_FJ151200
LM7627_ii40
LM7608_ii31
LM7317_ii6
LM7283_ii7
LM7328_ii9
SPLCV_EF456741
SPLCV_EF456743
LM7806_ii51
LM7297_ii12
LM7300_ii13
LM7628_ii42
LM7303_ii14
7868_ii64
LM7275_ii3
LM7294_ii10
LM7327_ii9
LM7295_ii11
LM7607_ii28
LM7610_ii38
LM7865_ii35
SPLCV_EU839578
SPLCV_EU839577
LM7606_ii25
LM7286_ii8
LM7805_ii47
LM7572_ii27
SPMV_FJ969831
LM7842_ii56
LM7844_ii58
SPLCCV_EF456742
SPLCCV_EF456745
SPLCCV_EU856365
SPLCCV_FJ529203
SPLCV_EF456744
SPLCV_EU856366
SPLCV_EU856364
SPLCV_EF456746
SPLCV_EU839579
SPLCCNV_DQ512731
SPLCSIV2_KF156759
SPLCHnV_KC907406
SPLCSIV1_KC488316

Figure 4.

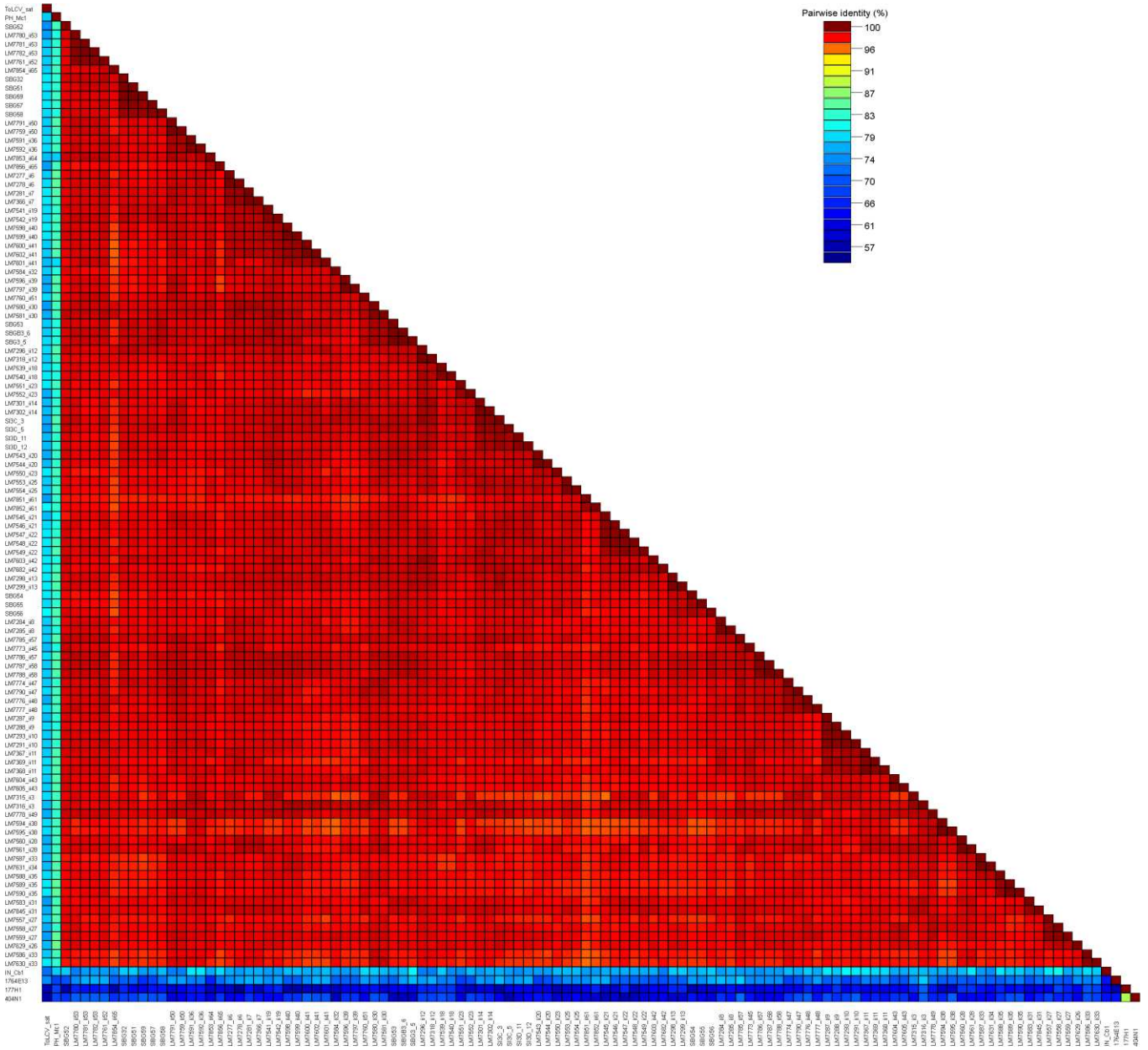


Figure 5.

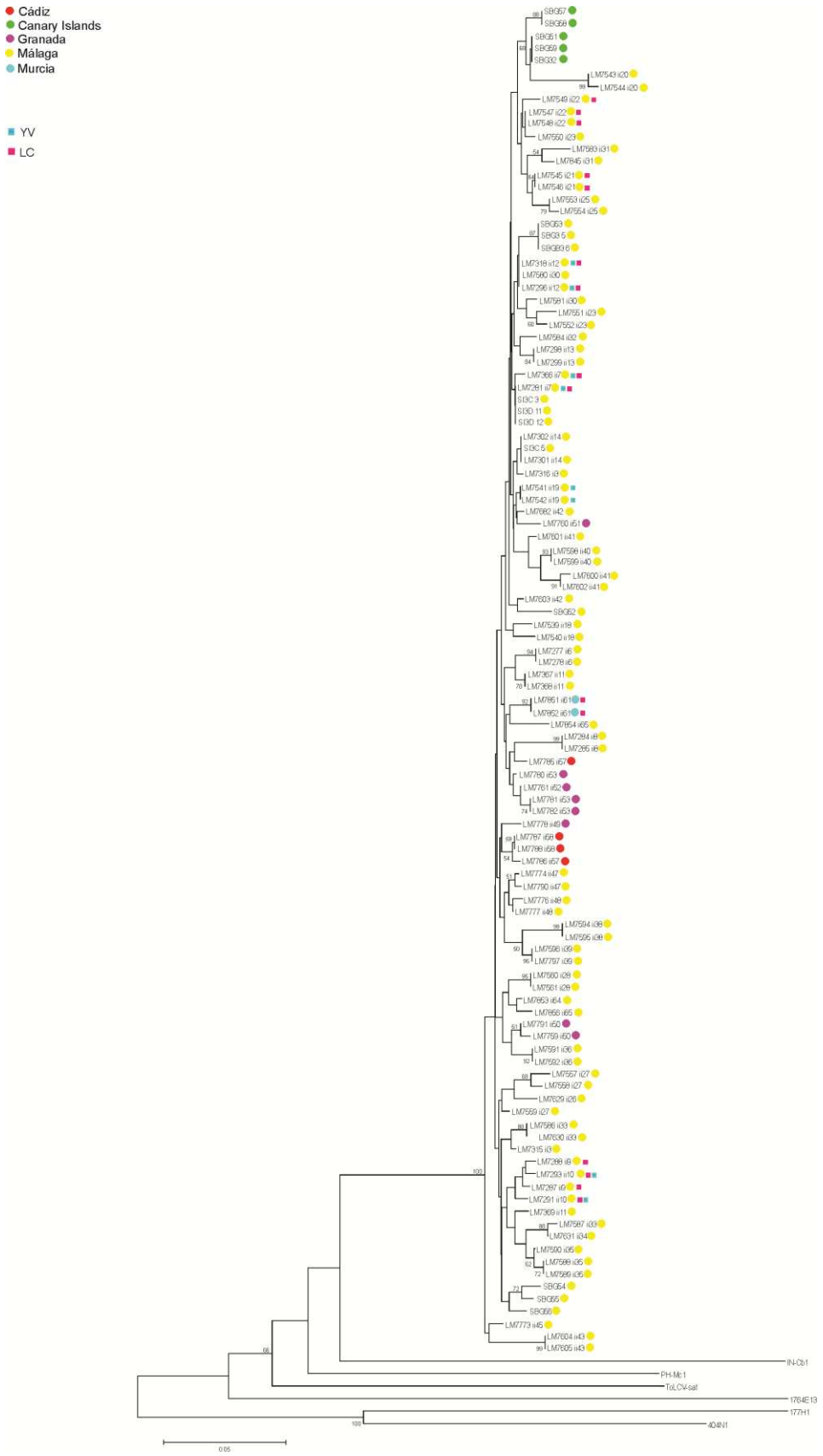

Cádiz
Canary Islands
Granada
Málaga
Murcia
YV
LC
IN-Cb1
PH-Mc1
ToLCV-sat
1764E13
177H1
404N1
0.05

Figure 6.

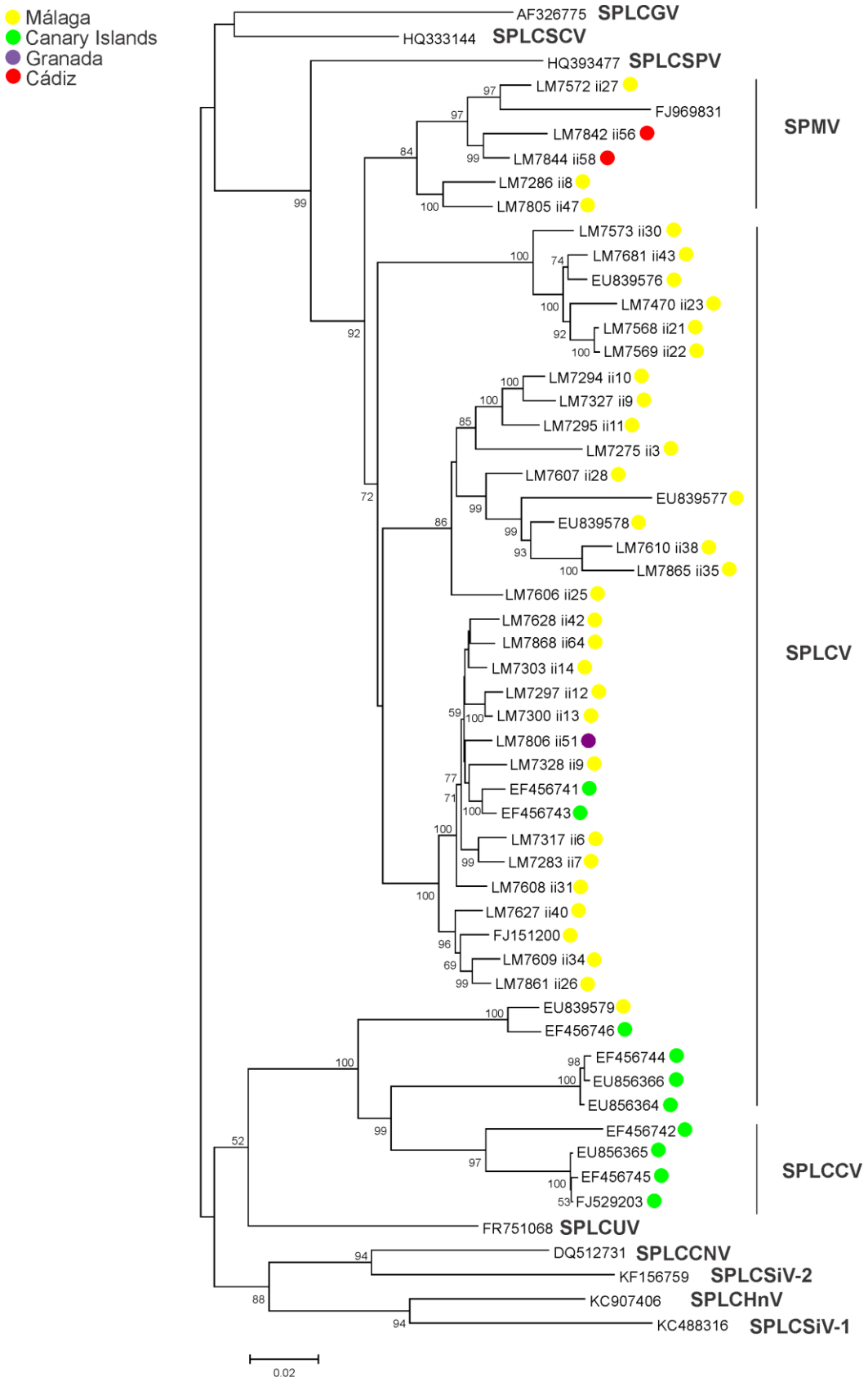
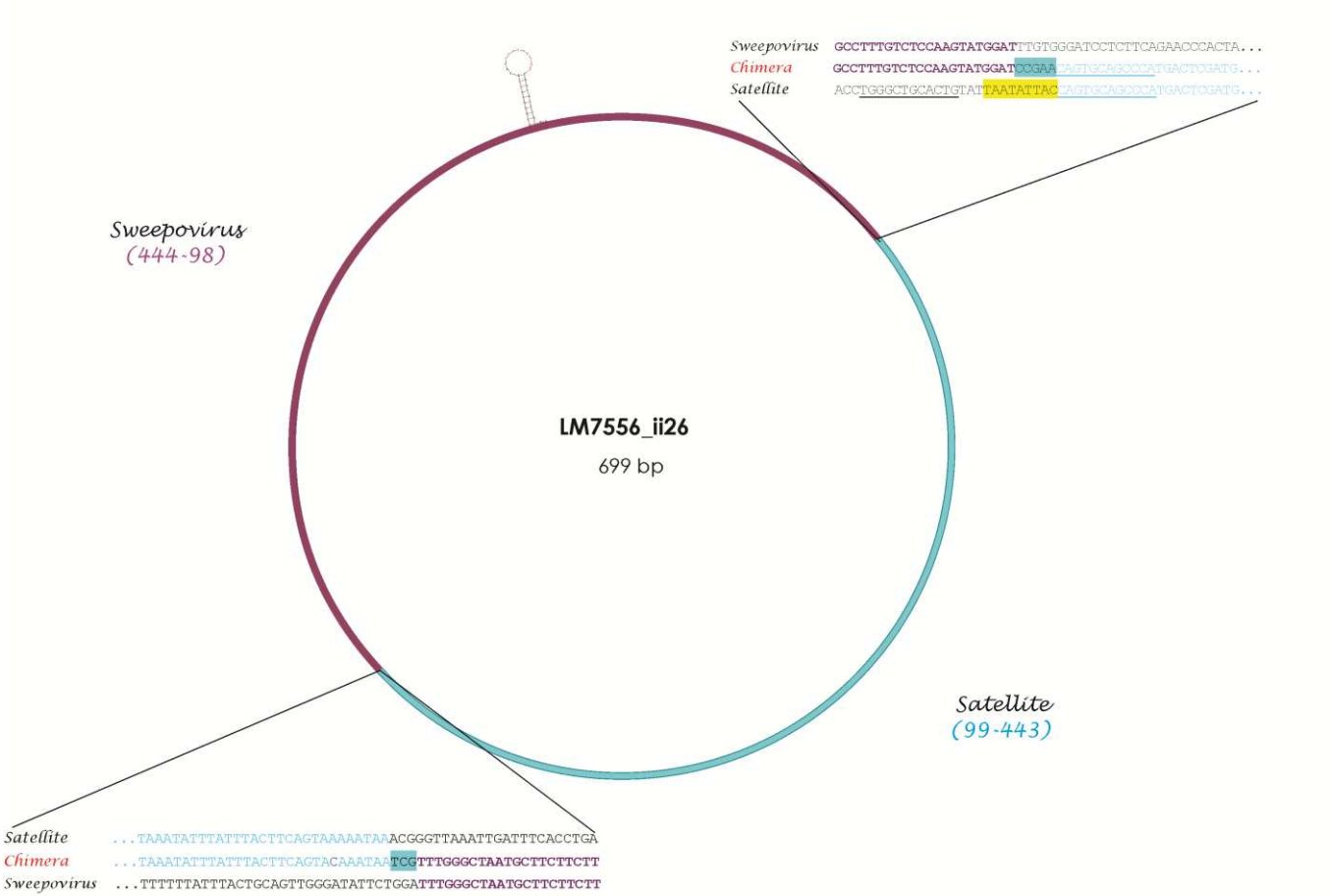

Málaga
Canary Islands
Granada
Cádiz
AF326775 SPLCGV
HQ333144 SPLCSCV
HQ393477 SPLCSPV
LM7572 ii27
FJ969831
LM7842 ii56
LM7844 ii58
LM7286 ii8
LM7805 ii47
SPMV
LM7573 ii30
LM7681 ii43
EU839576
LM7470 ii23
LM7568 ii21
LM7569 ii22
LM7294 ii10
LM7327 ii9
LM7295 ii11
LM7275 ii3
LM7607 ii28
EU839577
EU839578
LM7610 ii38
LM7865 ii35
LM7606 ii25
LM7628 ii42
LM7868 ii64
LM7303 ii14
LM7297 ii12
LM7300 ii13
LM7806 ii51
LM7328 ii9
EF456741
EF456743
LM7317 ii6
LM7283 ii7
LM7608 ii31
LM7627 ii40
FJ151200
LM7609 ii34
LM7861 ii26
SPLCV
EU839579
EF456746
EF456744
EU856366
EU856364
EF456742
EU856365
EF456745
FJ529203
SPLCCV
FR751068 SPLCUV
DQ512731 SPLCCNV
KF156759 SPLCSiV-2
KC907406 SPLCHnV
KC488316 SPLCSiV-1
0.02

Figure 7.


Sweepovirus GCCTTTGTCTCCAAGTATGGATTTGTGGGATCCTCTTCAGAACCCACTA...
Chimera GCCTTTGTCTCCAAGTATGGATCCGAAAAGTGCAGCCCATGACTCGATG...
Satellite ACCTGGGCTGCACTGTATTAATATTACAAGTGCAGCCCATGACTCGATG...
Sweepovirus
(444-98)
LM7556_ii26
699 bp
Satellite
(99-443)
Satellite ...TAAATATTTATTTACTTCAGTAAAAATAAACGGGTTAAATTGATTTCACCTGA
Chimera ...TAAATATTTATTTACTTCAGTACAAATAATCGTTTGGGCTAATGCTTCTTCTT
Sweepovirus ...TTTTTTATTTACTGCAGTTGGGATATTCTGGATTTGGGCTAATGCTTCTTCTT