Annals of Applied Biology ISSN 0003-4746

RESEARCH ARTICLE

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[1,3], T.B. Mar[1], D. Lau[2] & F.M. Zerbini[1]

1 Dept de Fitopatologia/BIOAGRO and National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil
2 Embrapa Trigo, Rodovia BR285 km 294, Passo Fundo, RS 99050-970, Brazil
3 Present address: Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, MG 38400-902, Brazil

Keywords
Begomovirus; geminivirus: recombination; plant virology; satellite DNA.

Correspondence
F.M. Zerbini, Dept de Fitopatologia/BIOAGRO and National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil.
Email: zerbini@ufv.br

Received: 3 March 2016; revised version accepted: 3 August 2016; published online: 20 January 2017.

doi:10.1111/aab.12329

Abstract

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 characterise 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 2009–2011. 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.

Introduction

The family *Geminiviridae* comprises 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 organisation 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; Padidam *et al.*, 1999; Paximadis *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 (Rojas *et al.*, 2005; Mahajan *et al.*, 2011; Hanley-Bowdoin *et al.*, 2013; Brustolini *et al.*, 2015). The majority of the begomoviruses that occur in the OW are monopartite, with a genomic organisation similar to the DNA-A of bipartite viruses (Padidam *et al.*, 1996; Mansoor *et al.*, 2003) 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 1950s (Costa & Bennett, 1950; Costa, 1955). 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 (Ribeiro *et al.*, 2003; Fernandes *et al.*, 2006; Calegario *et al.*, 2007; Ribeiro *et al.*, 2007; 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 (Krenz *et al.*, 2012; Loconsole *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 (Frischmuth *et al.*, 1997; Roye *et al.*, 1997; Faria and Maxwell, 1999; Fernandes *et al.*, 1999; Idris *et al.*, 2003; Jovel *et al.*, 2004; Assunção *et al.*, 2006; Amarakoon *et al.*, 2008; Castillo-Urquiza *et al.*, 2008; Barbosa *et al.*, 2009; Silva *et al.*, 2011; Silva *et al.*, 2012; Tavares *et al.*, 2012; Barreto *et al.*, 2013). 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 (Arnaud *et al.*, 2007; Castillo-Urquiza *et al.*, 2007; Cotrim *et al.*, 2007; Silva *et al.*, 2010; Barreto *et al.*, 2013; Rocha *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014), 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 & 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 were excised with *Apa*I, *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Sac*I, *Sal*I or *Spe*I and ligated into the pBLUESCRIPT-KS+ (pKS+) plasmid vector (Stratagene, San Diego, CA, USA), previously cleaved with the same enzyme. Viral inserts were sequenced commercially (Macrogen Inc., Seoul, South Korea) by primer walking. All genome sequences were organised 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 & 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 visualised 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 those 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 (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 approximately 2600-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. samples and 41 clones from *L. sibiricus* samples (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; Figs. 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 (nt) identity amongst themselves and 96.3% to 97.2% identity with SiMMV (accession number FN436003), and the DNA-B sequences share >88.0% nt identity amongst themselves and 89.1–93.3% identity with SiMMV (FN436004) (Fig. S1). Bayesian phylogenetic trees based on either the DNA-A or DNA-B placed these isolates in a monophyletic branch together with SiMMV (Fig. 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; Fig. 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% nt 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 nt sequence identity of 81.5% with *Tomato dwarf leaf virus* (ToDfLV, JN564749) (Fig. 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), *Sida mosaic Bolivia virus 1* (SiMBoV1) and *Abutilon mosaic Bolivia virus* (AbMBoV; Fig. 1A). This close relationship is consistent with the pairwise sequence identity analysis and with previously reported data (Márquez-Martín *et al.*, 2011; Medina & Lambertini, 2012; 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 (Fig. S2), suggesting a recombinant origin. Mixed infections by different begomoviruses are common in non-cultivated hosts (García-Andrés *et al.*, 2006; Alabi *et al.*, 2008; 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 nt sequence identity with SiMMV (86% and 75.5%, respectively; Fig. 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 (Fig. 1A). The recombination event has strong phylogenetic support: the *rep* nt sequence tree places BR:Tac720:10 in a monophyletic branch (99% posterior probability) with the SiMMV isolates, while the *cp* nt sequence tree places the isolate in a monophyletic branch (99% posterior probability) with *Abutilon mosaic Brazil virus* (AbMBV, JF694480; Fig. 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 (Fig. 1B). One recombination event was detected, with BR:Tol1075:11 and *Bean dwarf mosaic virus* (BDMV, M88180) as the putative 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 (Fig. S3). Recombination is a common event among geminiviruses (Padidam *et al.*, 1999; Lefeuvre *et al.*, 2009) and contributes greatly to their evolutionary potential and local adaptation (Harrison & 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.*, 2007*a*,*b*), and cotton and okra in Pakistan (Zhou *et al.*, 1998; Idris and Brown, 2002; Briddon *et al.*, 2014).

Table 1 Begomovirus and alphasatellite sequences reported in this study

Sample Code	Sampling Date	Location	Geographical Coordinates		Host	Enzyme[a]			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* spp.	*Eco*RI			BR:Saa48:09	KX348158
CF65	03/20/09	São Miguel das Missões, RS	S28 29′ 35.59″	W54 33′ 37.15″	*Sida* spp.	*Spe*I			BR:Smm65:09	KX348159
CF69	03/20/09	Santo Antonio das Missões, RS	S28 29′ 42.70″	W55 25′ 18.00″	*Sida* spp.		*Spe*I		BR:Sam69:09	KX348195
CF115	03/20/09	Panambi, RS	S28 18′ 24.45″	W53 29′ 21.38″	*Sida* spp.	*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* spp.		*Spe*I		BR:Smm547:10	KX348196
CF556	03/25/10	Cruz Alta, RS	S28 36′ 10.89″	W53 39′ 25.34″	*Sida* spp.		*Sal*I		BR:Cra556:10	KX348194
CF662	03/25/10	Realeza, PR	S25 40′ 45.00″	W53 33′ 09.00″	*Sida* spp.	*Eco*RI	*Spe*I		BR:Rea662:10	KX348156 (A) KX348186 (B)
CF679	06/08/10	Marechal Candido Rondon, PR	S24 30′ 59.50″	W54 04′ 37.20″	*Sida* spp.		*Apa*I		BR:Mcr679:10	KX348197
CF698	06/09/10	Dourados, MS	S22 17′ 58.00″	W54 49′ 14.10″	*Sida* spp.		*Hind*III		BR:Dou698:10	KX348191
CF704	06/09/10	Laguna Carapã, MS	S22 25′ 35.60″	W55 21′ 30.00″	*Sida* spp.		*Sal*I		BR:Lac704:10	KX348193
CF732	06/10/10	Umuarama, PR	S23 50′ 12.30″	W53 17′ 23.60″	*Sida* spp.	*Eco*RI			BR:Umu732:10	KX348164
CF755	06/10/10	Marialva,PR	S23 30′ 05.00″	W51 47′ 08.00″	*Sida* spp.		*Sal*I		BR:Mar755:10	KX348198
CF799	08/11/10	São Domingos, PR	S24 00′ 51.00″	W51 30′ 37.00″	*Sida* spp.		*Apa*I		BR:Sad799:10	KX348190
CF822	08/24/10	Chapada, RS	S28 01′ 08.80″	W53 05′ 52.31″	*Sida* spp.	*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* spp.		*Spe*I		BR:Saa832:10	KX348202
CF876	10/06/10	São Miguel das Missões, RS	S28 23′ 59.00″	W54 39′ 52.00″	*Sida* spp.	*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* spp.		*Sal*I		BR:Tap895:10	KX348189
CF926	03/15/11	Chapada, RS	S28 01′ 08.80″	W53 05′ 52.31″	*Sida* spp.	*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* spp.		*Sal*I		BR:Smm949:11	KX348203
CF1116	06/08/11	Tacuru, RS	S23 38′ 17.60″	W54 58′ 17.70″	*Sida* spp.		*Apa*I		BR:Tac1116:11	KX348188
CF1120	06/09/11	Guaíra, PR	S24 04′ 57.00″	W54 10′ 04.00″	*Sida* spp.		*Apa*I		BR:Gua1120:11	KX348204
CF1121	06/09/11	Umuarama, PR	S23 50′ 12.30″	W53 17′ 23.60″	*Sida* spp.	*Spe*I			BR:Umu1121:11	KX348157
CF1128	06/09/11	Janiópolis, PR	S24 08′ 02.00″	W52 47′ 22.00″	*Sida* spp.		*Bam*HI		BR:Jan1128:11	KX348187
CF1142	06/09/11	Marialva, PR	S23 30′ 09.40″	W51 47′ 42.90″	*Sida* spp.	*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* spp.	*Hind*III			BR:Trm531.2:10	KX348180
CF691	06/08/11	Caarapo, MS	S22 26′ 42.50″	W54 49′ 49.50″	*Sida* spp.	*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* spp.	*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* spp.	*Spe*I	*Cla*I		BR:Tac720:10	KX348184 (A) KX348225 (B)

Table 1 Continued

Sample Code	Sampling Date	Location	Geographical Coordinates		Host	Enzyme[a] DNA-A	Enzyme[a] DNA-B	Enzyme[a] Satellite	Isolate Name	GenBank Access Number
CF677	06/08/10	Marechal Candido Rondon, PR	S24 30′ 59.50″	W54 04′ 37.20″	*L. sibiricus*	*Apa*I			BR:Mcr677:10	KX348170
Sida golden yellow spot virus (SiGYSV)										
CF889	06/10/10	São Borja, RS	S28 57′ 30.00″	W55 32′ 30.00″	*Sida* spp.	*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* spp.		*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)
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* spp.			*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

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

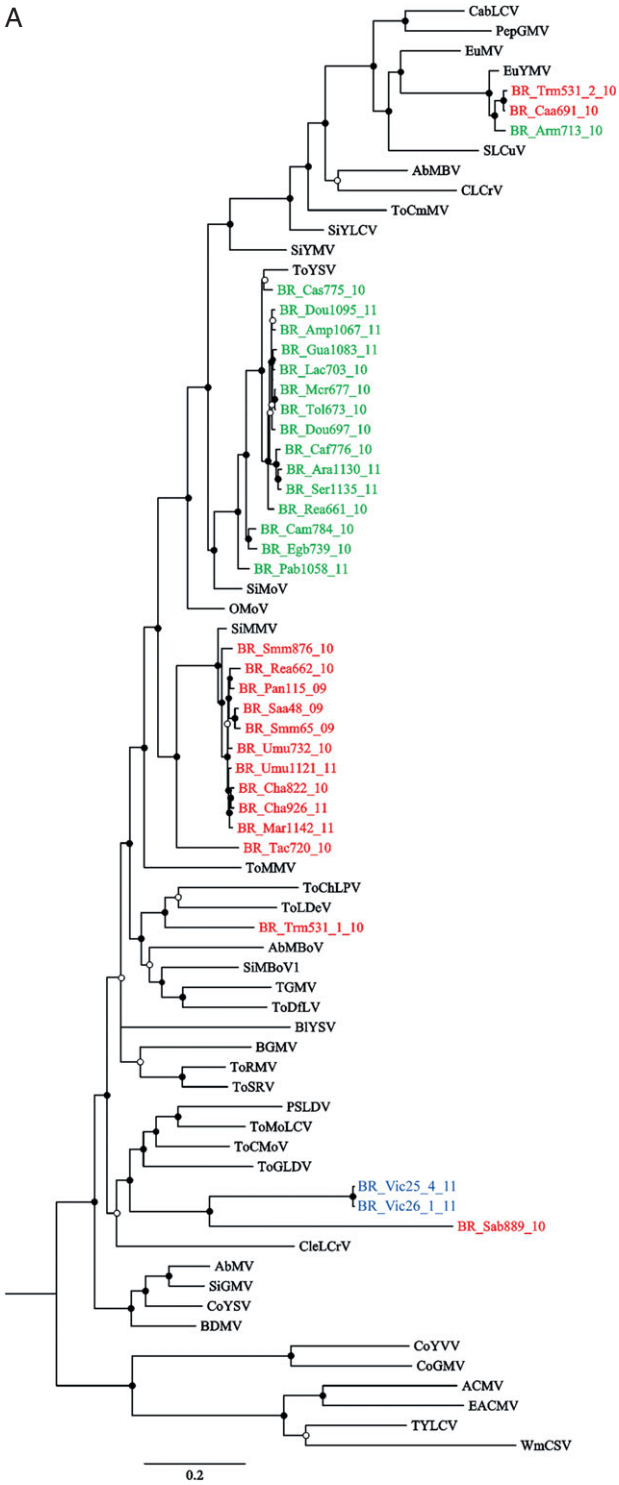

Figure 1 Bayesian phylogenetic trees based on the complete DNA-A (A) and DNA-B (B) nucleotide sequences of the begomoviruses obtained from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) in this study plus begomoviruses from the New World and Old World (see Table S1 for full names and GenBank 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.

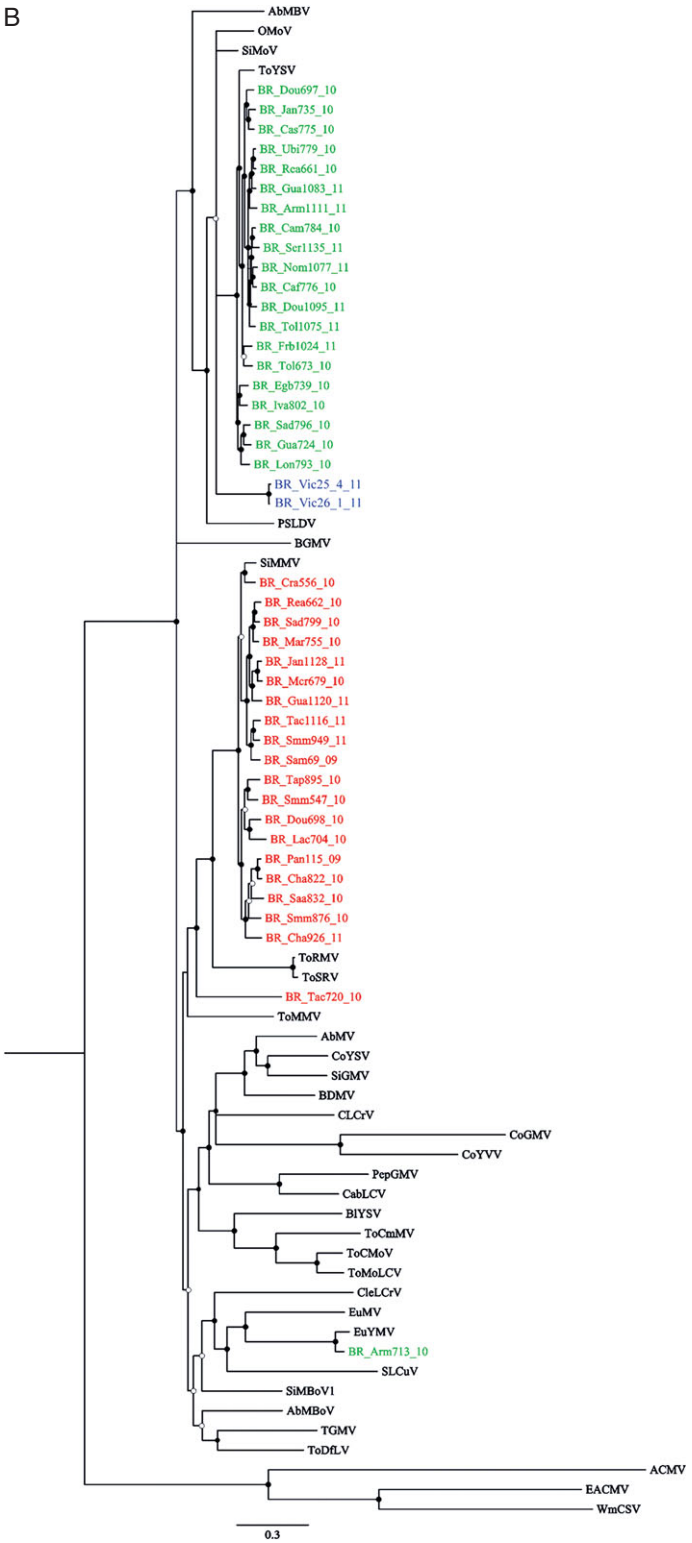

Figure 1 Continued.

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 from the New World (NW) and Old World (OW)

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

[a]See Table S1 for the complete data set and full virus names.
[b]Recombination breakpoint coordinates are according to the first nucleotide after the cleavage site at the origin of replication, increasing clockwise.
[c]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.
[d]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.

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 nt 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 1100 nt) has very low similarity to any other begomovirus. The deduced amino acid sequences of their CP and AV2-like proteins were further analysed with BLASTp 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. BLASTp 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% amino acid (aa) identity (90% coverage, *E* value $2e^{-12}$) with the AGV CP and the AV2-like protein shares 43% aa 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 unknown 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 (Fig. 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 (Fig. 2). BR:Sab889:10 clusters with BR:Vic25.4:11, BR:Vic26.1:11 and AGV, reflecting the BLASTp analysis in which a similarity was found among the CPs of these four viruses. Studies are in progress in our laboratory to characterise 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% nt 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% nt identity amongst themselves and 91.0% to 92.3% nt identity with ToYSV (DQ336351) (Table 1; Fig. S1). The Bayesian phylogenetics trees based on the DNA-A and DNA-B components placed these isolates in clusters with ToYSV (Fig. 1).

The cloned DNA-A and DNA-B (BR:Arm713:10) from *L. sibiricus* sample #713 have 97.2% and 95.3%

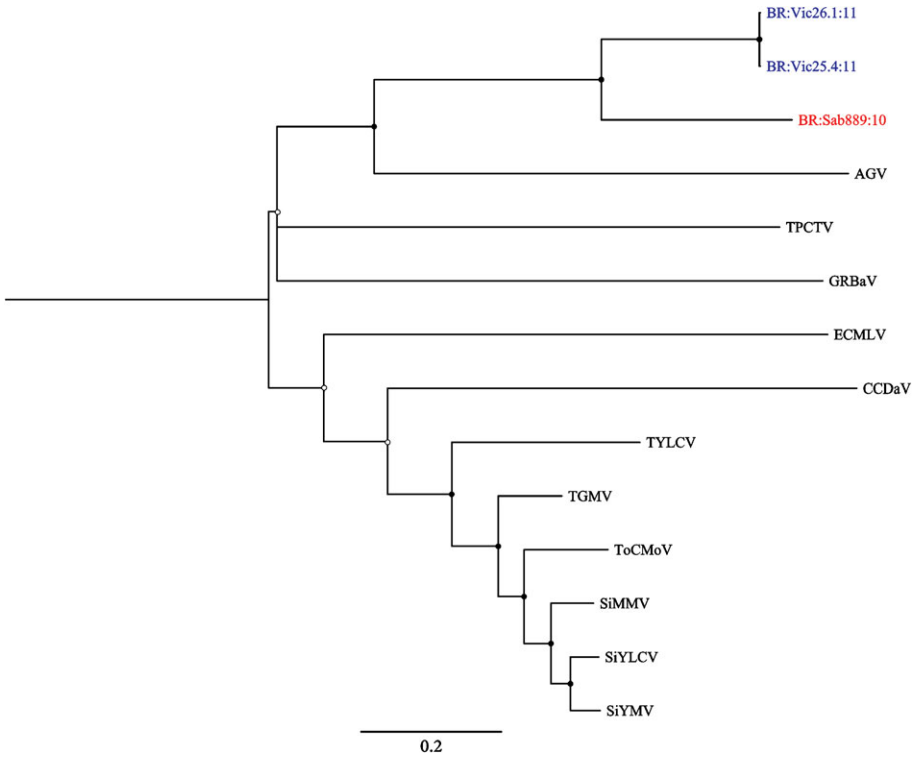

Figure 2 Bayesian phylogenetic tree based on the nucleotide sequences of the *cp* genes of Sida golden yellow spot virus (SiGYSV, in red and selected geminiviruses). The unrooted *cp* tree includes four highly divergent geminiviruses (apple geminivirus, AGV; citrus chlorotic dwarf associated virus, CCDaV; *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 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.

nt sequence identity with EuYMV (FJ619507 and FJ619508), respectively (Table 1; Fig. S1). EuYMV has been found sporadically in some non-cultivated hosts (Fernandes *et al.*, 2011; Silva *et al.*, 2012; Tavares *et al.*, 2012; Barreto *et al.*, 2013; Rocha *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 nt sequence identity (93.2%) to *Euphorbia yellow mosaic alphasatellite* (EuYMA, FN436008) (Fig. S4) and a close phylogenetic relationship with this isolate (Fig. 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 nt sequence comparisons indicated that the DNA-A and DNA-B sequences showed the highest identities with ToYSV (94.6% and 93.0%, respectively; Fig. S1). Alphassatellite sequences shared >99.9% nt identity amongst themselves and 82.3% to 82.4% nt identity with the sequence of EuYMA (Fig. 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 (Fig. 3), consistent with the pairwise identity analysis. According to the proposed demarcation threshold of 83% nt sequence identity for alphasatellites (Mubin *et al.*, 2009, 2012), the five clones from *L. sibiricus* sample #1095 represent a distinct alphasatellite, for which the name Leonurus yellow spot alphasatellite (LeYSA) is proposed.

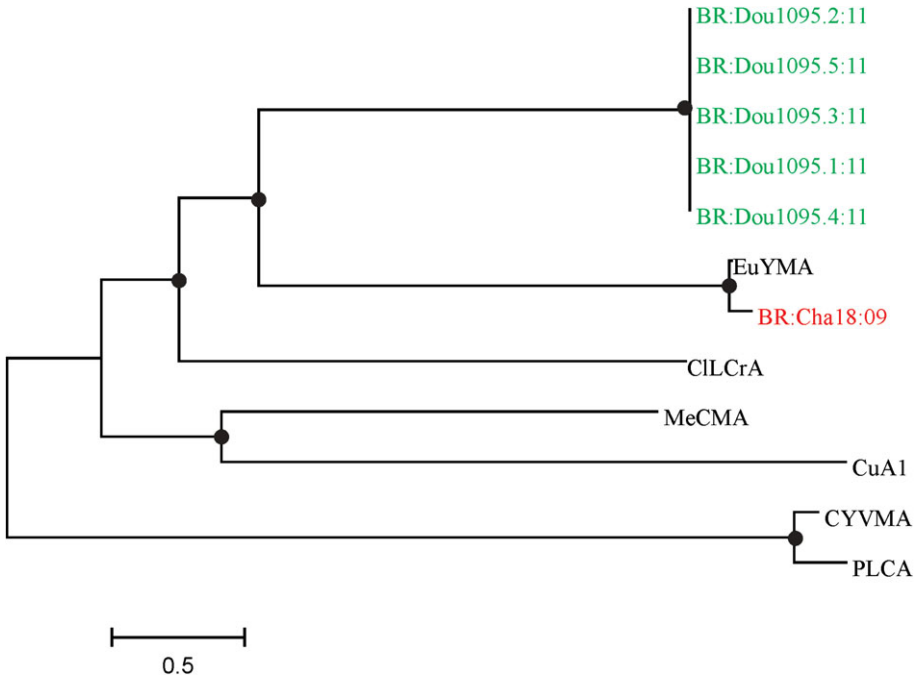

Figure 3 Bayesian phylogenetic tree based on the alphasatellites sequences described in this study from *Sida* spp. (red) and *Leonurus sibiricus* plants (green) and the most closely related alphasatellites (see Table S2 for full names and GenBank access numbers). Nodes with posterior probability values 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.

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; Romay *et al.*, 2010; Jeske *et al.*, 2014). 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 (Frischmuth *et al.*, 1997; Hofer *et al.*, 1997; Roye *et al.*, 1997; Echemendía *et al.*, 2004; Jovel *et al.*, 2004; Xiong *et al.*, 2005; Guo & Zhou, 2006; Das *et al.*, 2008; Fiallo-Olivé *et al.*, 2010; Fiallo-Olivé *et al.*, 2012). 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 emphasising the tremendous diversity of begomoviruses and associated DNA satellites naturally infecting *Sida* spp.

Conversely, *L. sibiricus* harbours 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 F. M. Z. C. G. F. 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., Oliveira Ferreira P.D.T., 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 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**, 289–298.

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 and 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.F., Fiallo-Olivé E., Briddon R.W., Hernández-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.B., Machado J.P.B., Condori-Apfata J.A., Coco D., Deguchi M., Loriato V.A.P., Pereira W.A., Alfenas-Zerbini P., Zerbini F.M., Inoue-Nagata A.K., Santos A.A., Chory J., Silva F.F., Fontes E.P.B. (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*, a novel tomato-infecting begomovirus in Brazil. *Pesquisa Agropecuária Brasileira*, **42**, 1335–1343.

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.

Castillo-Urquiza G.P., Beserra J.E.A. Jr., 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.

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

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.

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 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 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 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., 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., 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., 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 alphasatellite 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.

Márquez-Martín B., Aragón-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., 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.

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.

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**, 148157.

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.D.B., Lima G.S.D.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-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.

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.

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.

Supporting Information

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

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

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

Fig. 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).

Fig. S2. Bayesian phylogenetic trees based on the nucleotide sequences of 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 Table S1, Supporting Information, 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.

Fig. 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 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.

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 Table S2 for full names and GenBank access numbers).