CÉSAR AUGUSTO DINIZ XAVIER

SPECIES DIVERSITY AND GENETIC VARIABILITY OF BIPARTITE BEGOMOVIRUSES IN THE NEW WORLD

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitopatologia, para obtenção do título de *Magister Scientiae*.

VIÇOSA

MINAS GERAIS - BRASIL

2015

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

T
X3s
2015

Xavier, César Augusto Diniz, 1990-
Species diversity and genetic variability of bipartite begomoviruses in the new world / César Augusto Diniz Xavier. – Viçosa, MG, 2015.
x, 106f. : il. (algumas color.) ; 29 cm.

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

1. Begomovírus. 2. Diversidade genética. 3. Evolução. 4. Filogeografia. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-graduação em Fitopatologia. II. Título.

CDD 22. ed. 571.9928

CÉSAR AUGUSTO DINIZ XAVIER

SPECIES DIVERSITY AND GENETIC VARIABILITY OF BIPARTITE BEGOMOVIRUSES IN THE NEW WORLD

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitopatologia, para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de julho de 2015.

Poliane Alfenas Zerbini

Eduardo Seiti Gomide Mizubuti

Francisco Murilo Zerbini Junior
(Orientador)

Em nome da minha mãe Maria de Lourdes e meu pai Heleno, dedico.

"*A jornada pareceu árdua e difícil...*

O desânimo tentou se apossar por vezes...

Entretanto, o lembrar de suas faces preocupadas,

De seu trabalho, de suas orações, de seu apoio incondicional para nos dar o melhor,

Impulsionou-nos deveras para a luta.

Obrigado, muito obrigado pelo silêncio,

Quando eu reclamava e,

Obrigado também pelas suas palavras de estímulo quando eu me calava.

Nessa grande batalha, creiam-nos a vitória também é de vocês meus pais..."

AGRADECIMENTOS

A Deus por ter-me concedido mais esta graça, sempre guiando meus passos e iluminando meu caminho;

Aos meus pais, Heleno e Maria de Lourdes, pela confiança, pelo apoio incondicional nos momento de alegria e tristezas, pelo incentivo, amor e dedicação em todos os momentos da minha vida;

Aos meus irmãos, Kênia, Gustavo, Jordânia, meu sobrinho Guilherme e minha cunhada Kely, pelo apoio, por sempre acreditarem em mim e estarem presentes ao longo da minha vida;

A toda família Diniz, em especial minha avó Carmem, pela confiança e por fazerem parte da minha vida;

A Universidade Federal de Viçosa (UFV), pela oportunidade de realização deste curso;

Ao Professor Francisco Murilo Zerbini, pelo aprendizado, orientação, dedicação e amizade durante a realização deste trabalho;

Aos amigos do Laboratório de Virologia Vegetal Molecular, Alison, Daniel, Camila, Márcio, Carol, Jorge, Silvia, Diego, Talita, Osvaldo, Roberto, Daniele, Robert, Fábio, Tathiana, Renan, Fernanda, Hermano, João, Josiane, Igor, Vitor, Lenin, Adriana, Diogo, Murilo, Sarah e Joyce pela amizade e pelo agradável ambiente de trabalho. Em especial a Glória, André e Larissa minha eterna gratidão.

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

Aos amigos do Laboratório de Microbiologia Industrial, Prof[a]. Poliane, Flávia, Ana, Fernanda e Laiane pela amizade.

Aos amigos Arthur, Filipe, Diogo, Lucas, Matheus, Luiz Henrique pelo companheirismo, pelos bons momentos vividos em Viçosa, especialmente na Bagacera, ao longo desta jornada.

Agradeço a todos os professores e amigos do Departamento de Fitopatologia;

A todos que contribuíram direta ou indiretamente para a realização deste trabalho, muito obrigado.

BIOGRAFIA

César Augusto Diniz Xavier, filho de Heleno Pereira Xavier e Maria de Lourdes Diniz Xavier, nasceu em 12 de Julho de 1990, em Florestal, Minas Gerais.

Em janeiro de 2009 concluiu o ensino médio concomitante ao curso Técnico em Agropecuária na Universidade Federal de Viçosa (UFV), Campus Florestal e em março deste mesmo ano iniciou o curso de graduação em Agronomia, pela UFV, campus Viçosa, vindo a graduar em julho de 2013. Em setembro de 2013 iniciou o curso de mestrado no Programa de Pós-Graduação em Fitopatologia da UFV, defendendo sua dissertação em 27 de julho de 2015.

SUMÁRIO

RESUMO

XAVIER, César Augusto Diniz, M.Sc., Universidade Federal de Viçosa, Julho de 2015. **Diversidade de espécies e variabilidade genética de begomovírus bissegmentados do Novo Mundo.** Orientador: Francisco Murilo Zerbini Júnior.

O gênero *Begomovirus* (família *Geminiviridae*) é constituído por vírus que apresentam um ou dois componentes genômicos de DNA circular de fita simples (ssDNA), transmitidos por *Bemisia tabaci* (Homoptera: Aleyrodidae) a plantas dicotiledôneas. Estes vírus são divididos em dois grupos, begomovírus do Novo Mundo e do Velho Mundo, de acordo com organização genômica e relacionamento filogenético. Os begomovírus constituem um importante grupo de patógenos de plantas responsáveis por perdas severas em diversas culturas de importância econômica, principalmente em regiões tropicais e subtropicais. Begomovírus evoluem a taxas comparáveis as de vírus que possuem genoma de RNA. Diversos estudos determinando a estrutura genética e a dinâmica de populações de begomovírus tem sido realizados com base apenas na análise do componente DNA-A. Os objetivos deste estudo foram: (*i*) realizar a caracterização molecular e biológica de dois novos begomovírus isolados do hospedeiro não-cultivado *Sida acuta*; (*ii*) realizar uma análise comparativa da estrutura e variabilidade genética de populações de begomovírus no Novo Mundo considerando o DNA-A e o DNA-B. Para o primeiro objetivo, DNA total foi extraído de amostras *Sida acuta* coletadas em dezembro de 2011 em Viçosa, Minas Gerais, e o genoma viral foi amplificado por RCA, clonado e sequenciado. Clones infecciosos foram gerados para realização da caracterização biológica. Foram identificadas duas novas espécies de begomovírus, nomeadas Sida yellow spot virus (SiYSV) e Sida golden yellow mosaic virus (SiGYMV), que compartilham características com begomovírus do Velho Mundo e possuem uma proteína capsidial altamente divergente. Interessantemente, o SiYSV não foi transmitido por *B. tabaci* MEAM1. Para o segundo objetivo, um total de 241 sequências de DNA-B de quatro espécies de begomovírus (*Bean golden mosaic virus*, BGMV; *Blainvillea yellow spot virus*, BlYSV; *Macroptilium yellow spot virus*, MaYSV; *Tomato severe rugose virus*, ToSRV), obtidas neste estudo, além de 239 sequências de DNA-A e 16 DNA-B obtidas no GenBank, foram analizadas. O DNA-B apresentou um maior grau de variabilidade genética em relação ao DNA-A e foi mais propenso a recombinar, com a detecção de um maior número de eventos de recombinação. A análise da estrutura genética das populações

virais indicou segregação com base em origem geográfica, tanto para o DNA-A quanto para o DNA-B. Em conjunto, os resultados demonstram que embora o DNA-A e o DNA-B estejam co-evoluindo, eles se encontram sob diferentes pressões seletivas, sendo o DNA-B mais permissivo a variação.

ABSTRACT

XAVIER, César Augusto Diniz, M.Sc., Universidade Federal de Viçosa, July 2015. **Species diversity and genetic variability of bipartite begomoviruses in the New World.** Adivisor: Francisco Murilo Zerbini Júnior.

The genus *Begomovirus* (family *Geminiviridae*) includes viruses with mono- and bipartite genome of circular, single-stranded DNA (ssDNA), transmitted by *Bemisia tabaci* to dicotyledonous plants. These viruses are divided into New World (NW) and Old World (OW) groups based on genomic organization and phylogenetic relationships. The begomoviruses constitute an important group of plant pathogens responsible for severe losses in several crops of economic importance, mainly tropical and subtropical regions. Begomoviruses evolve at high rates, compared to viruses with RNA genomes. Studies to determine the genetic structure and evolutionary dynamics of populations begomoviruses has been based solely on the DNA-A. This study aimed to: (*i*) carry out the molecular and biological characterization of two new begomoviruses isolated from the non-cultivated host *Sida acuta*; (*ii*) perform a comparative analysis of the genetic structure of NW begomoviruses based on the DNA-A and DNA-B components. For the first objective, total DNA was extracted from *S. acuta* samples collected in Viçosa, state of Minas Gerais in December 2011, and the viral genome was amplified by RCA, cloned and sequenced. Infectious clones were generated to perform the biological characterization. We identified two new species of OW-like begomoviruses present in the NW, named Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV). Moreover, both viruses have a highly divergent capsid protein gene. Interestingly, SiYSV was not transmitted by *B. tabaci* MEAM1. For the second objective we analyzed a total of 241 DNA-B sequences belonging to four begomovirus species (*Bean golden mosaic virus*, BGMV; *Blainvillea yellow spot virus*, BlYSV; *Macroptilium yellow spot virus*, MaYSV; *Tomato severe rugose virus*, ToSRV), obtained in this study, in addition to 239 DNA-A and 16 DNA-B sequences retrieved from GenBank. The DNA-B was more variable and more likely to recombine compared to the DNA-A. Analysis of the genetic structure of the viral populations based on either the DNA-A or the DNA-B indicated segregation based on geographical location. Together, our results indicate that while the DNA-A and DNA-B are co-evolving, they are under different selective pressures, with the DNA-B being more permissive to variation.

GENERAL INTRODUCTION

The family *Geminiviridae* is comprised of viral species with one or two genomic components of circular, single-strand DNA (ssDNA) encapsidated in geminate icosahedral particles. The family includes the genera *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus,* classified according to host range, insect vector, genomic organization and phylogenetic relationships (Andrade et al., 2006; Rojas et al., 2005; Varsani et al., 2014).

The genus *Begomovirus* includes mono- and bipartite viruses transmitted by *Bemisia tabaci* to dicotyledonous plants (Brown et al., 2015). According to phylogenetic relationships and genome architecture, begomoviruses are divided into two major groups, named New World (NW; the Americas) and Old World (OW; Europe, Asia and Africa) (Briddon et al., 2010; Rybicki, 1994). NW begomoviruses are predominantly bipartite, with a single report of an indigenous NW monopartite virus, *Tomato leaf deformation virus* (ToLDeV) (Melgarejo et al., 2013). Begomoviruses in the OW can be monopartite or bipartite and are normally associated with satellite DNA molecules (Zhou, 2013).

The begomoviruses constitute an important group of plant pathogens responsible for severe losses in several crops of economic importance, especially in tropical and subtropical regions (Legg and Fauquet, 2004; Moriones and Navas-Castillo, 2000). Examples include cassava mosaic disease caused by *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV) in Africa (Legg and Fauquet, 2004; Sserubombwe et al., 2008) and tomato yellow leaf curl disease caused by *Tomato yellow leaf curl virus* (TYLCV) and *Tomato yellow leaf curl Sardinia virus* (TYLCSV), one of the main restrictions on tomato production in the Middle East, Mediterranean, and North and Central America (Moriones and Navas-Castillo, 2000). The begomoviruses are

widely distributed in the world, with frequent reports of epidemics, encouraging genetic and epidemiological studies (Hanley-Bowdoin et al., 2013). The main factors that have contributed to the emergence and spread of diseases caused by begomoviruses are the emergence of more aggressive viral variants and the increased dissemination of polyphagous populations of *B. tabaci* (Navas-Castillo et al., 2011).

Begomovirus populations have a high degree of genetic variability, equivalent to that of viruses with RNA genomes (Ge et al., 2007; Prasanna et al., 2010; Rocha et al., 2013). The main sources of variability of plant viruses are mutation, recombination and pseudo-recombination (García-Arenal et al., 2003; Monci et al., 2002; Seal et al., 2006b). Frequent recombination events (Padidam et al., 1999), the occurrence of pseudo-recombination between viruses with bipartite genomes (Andrade et al., 2006), and high mutation rates (Duffy and Holmes, 2008; Duffy and Holmes, 2009) are factors that promote the high variability observed for begomovirus populations.

The genetic structure of plant virus populations refers to the amount of distribution and genetic variability within and among populations (García-Arenal et al., 2001). Understanding the dynamics of variability of plant virus populations in both cultivated and non-cultivated hosts is important to understand how populations evolve, with implications for the efectiveness and durability of management stratagies (Seal et al., 2006a).

The advent of techniques for the sequence-unbiased amplification of small circular viral genomes (specially rolling circle amplification specially, RCA) (Inoue-Nagata et al., 2004) and next generation sequencing (NGS) technologies (Rosario et al., 2012) created new possibilities for the discovery of novel begomoviruses, and also of divergent ssDNA viruses (Bernardo et al., 2013; Krenz et al., 2012; Liang et al., 2015; Loconsole et al., 2012; Varsani et al., 2009). Many novel begomoviruses have been found in non-cultivated plants (Castillo-Urquiza et al., 2008; Fiallo-Olive et al., 2012; Paprotka et al., 2010; Silva

et al., 2012; Tavares et al., 2012), highlighting the need to investigate these plants as a reservoir of viral diversity and as a source of new viruses which may cause diseases in crops. In addition, the study of divergent viruses frequently found in non-cultivated plants contributes to a better understanding of viral evolution.

This study aimed to: (*i*) carry out the molecular and biological characterization of two new begomoviruses isolated from the non-cultivated host *Sida acuta*; (*ii*) perform a comparative analysis of the genetic structure of NW begomoviruses based on the DNA-A and DNA-B components of four viral populations isolated from cultivated and non-cultivated hosts.

References

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.

Bernardo, P., Golden, M., Akram, M., Naimuddin, Nadarajan, N., Fernandez, E., Granier, M., Rebelo, A.G., Peterschmitt, M., Martin, D.P., Roumagnac, P., 2013. Identification and characterisation of a highly divergent geminivirus: Evolutionary and taxonomic implications. Virus Res 177, 35-45.

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., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-Olive, 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.

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.

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.

Fiallo-Olive, 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.

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

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

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.

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.

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. J Virol Met 116, 209-211.

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

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

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. J Gen Virol, doi 10.1099/vir.0.000173.

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.

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. J Virol 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.

Moriones, E., Navas-Castillo, J., 2000. *Tomato yellow leaf curl virus*, an emerging virus complex causing epidemics worldwide. Virus Res 71, 123-134.

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

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.

Prasanna, H.C., Sinha, D.P., Verma, A., Singh, M., Singh, B., Rai, M., Martin, D.P., 2010. The population genomics of begomoviruses: global scale population structure and gene flow. Virol J 7, 220.

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 Phytopath 43, 361-394.

Rosario, K., Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. Arch Virol 157, 1851-1871.

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

Seal, S.E., Jeger, M.J., Van den Bosch, F., 2006a. Begomovirus evolution and disease management. Adv Virus Res 67, 297-316.

Seal, S.E., Van den Bosch, F., Jeger, M.J., 2006b. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. Critical Reviews in Plant Sciences 25, 23-46.

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.

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.

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.

Varsani, A., Shepherd, D.N., Dent, K., Monjane, A.L., Rybicki, E.P., Martin, D.P., 2009. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virol J 6, 36.

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

CHAPTER 1

TWO DIVERGENT BEGOMOVIRUSES FROM THE NEW WORLD WITH FEATURES RECALLING OLD WORLD BEGOMOVIRUSES

Xavier, C.A.D., Godinho, M.T., Trindade, T.A., Lima, A.T.M., Zerbini, F.M. Two divergent begomoviruses from the New World with features recalling Old World begomoviruses.

TWO DIVERGENT BEGOMOVIRUSES FROM THE NEW WORLD WITH FEATURES RECALLING OLD WORLD BEGOMOVIRUSES

C.A.D. Xavier[1*], M.T. Godinho[1*], T.A. Trindade[2], A.T.M. Lima[1,3], J.P.S. Herrera[1], 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] Dep. de Entomologia, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

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

[*]These authors contributed equally to this work

[#]Corresponding author:

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

Abstract

The genus *Begomovirus* (family *Geminiviridae*) is comprised of viruses with one or two single-stranded DNA (ssDNA) genomic components transmitted by the whitefly *Bemisia tabaci* to dicotyledonous plants, and includes important plant pathogens responsible for severe losses in many economically important crops worldwide. Begomoviruses are divided into New World (NW) and Old World (OW) groups based on genomic organization and phylogenetic relationships. In this study, we performed the biological and molecular characterization of Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV), two OW-like begomoviruses present in the NW. Total DNA was extracted from *S. acuta* samples collected in Viçosa, state of Minas Gerais, in December 2011, and the viral genome was amplified by RCA, cloned and sequenced. Infectious clones were generated to perform the biological characterization. The two genomic components of both viruses are phylogenetically related to NW viruses. Nevertheless, their DNA-A components exhibited a highly divergent 5' half, including part of the intergenic region, the *CP* gene, and an *AV2*-like gene (which is present only in OW begomoviruses). The deduced amino acid sequences of the CP and AV2-like proteins had very low identities with either NW or OW begomoviruses, having greater similarity with a divergent monopartite geminivirus recently identified in apple trees in China. The presence of conserved motifs in the CP and Rep coding regions which are characteristic of OW begomoviruses was also detected. Both viruses infected plants in the *Malvaceae* and *Solanaceae* families. Interestingly, SiYSV was not transmitted by *B. tabaci* MEAM1, a result that was not entirely unexpected considering the high level of divergence of its CP. Although NW-like begomoviruses have been found in the OW, this is the first report of an OW-like begomovirus found naturally in the NW.

Introduction

The genus *Begomovirus* (family *Geminiviridae*) is comprised of viral species with one or two genomic components of circular, single-stranded DNA (ssDNA), encapsidated in geminate icosahedral particles and transmitted by the *Bemisia tabaci* sibling species group to dicotyledonous plants (Brown et al., 2015). The begomoviruses include several plant pathogens responsible for severe losses in economically important crops such as cassava, tomato, cotton and bean, mainly in tropical and subtropical regions (Legg and Fauquet, 2004; Moriones and Navas-Castillo, 2000; Sobrinho et al., 2014; Sserubombwe et al., 2008).

According to phylogenetic relationships and genome architecture, begomoviruses are divided into two major groups (Briddon et al., 2010; Rybicki, 1994), named New World (NW - the Americas) and Old World (OW - Europe, Asia and Africa). NW begomoviruses are predominantly bipartite, with a single report of an indigenous NW monopartite virus, *Tomato leaf deformation virus* (ToLDeV) (Melgarejo et al., 2013). Begomoviruses in the OW can be monopartite or bipartite and are normally associated with satellite DNA molecules (reviewed by Zhou, 2013). The genomic components of bipartite begomoviruses are named DNA-A and DNA-B. The DNA-A encodes proteins involved in replication, encapsidation and suppression of host defenses (Hanley-Bowdoin et al., 2013; Rojas et al., 2005). The DNA-B encodes proteins involved in intracellular, cell-to-cell and long distance movement, 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). Monopartite begomoviruses have a genomic component which is homologous to the DNA-A of bipartite begomoviruses.

Several features distinguish NW and OW begomoviruses, the most outstanding being the presence in OW viruses of an additional open reading frame (ORF) which

partially overlaps the *CP* gene, named *V2* (in monopartite begomoviruses) or *AV2* (in bipartite begomoviruses). The V2/AV2 protein is involved in viral movement and gene silencing suppression (Glick et al., 2008; Padidam et al., 1996; Rybicki, 1994). Additional genomic features that distinguish NW and OW begomovirus include: (*i*) OW begomoviruses have ~100 nt larger components (Ho et al., 2014); (*ii*) NW begomoviruses have a PWRsMaGT motif in the N-terminal region of the CP, absent in OW viruses (Ha et al., 2006; Harrison et al., 2002); (*iii*) the CP N-terminal region from OW begomoviruses has two or three basic domains, typically KR, KVRRR and K/RRRR, part of a nuclear localization signal (NLS), while NW begomoviruses have only the KR domain (Ha et al., 2008); (*iv*) NW begomoviruses have a tyrosine phosphorylation site, with the consensus sequence [RK]-x(2,3)-[DE]-x(2,3)-Y, in the movement protein (MP) which is absent in OW begomoviruses.

In general, NW begomoviruses have a lower degree of genetic variability compared to OW begomoviruses (Briddon et al., 2010; Ho et al., 2014; Rybicki, 1994). This led to the proposal that NW begomoviruses evolved from ancestral OW begomoviruses after continental separation, possibly through limited introduction by the whitefly vector, followed by an explosive speciation process in host plants present in the NW (Rybicki, 1994). However, the occurrence in the OW (Vietnam) of *Corchorus yellow vein virus* (CoYVV) and *Corchorus golden mosaic virus* (CoGMV), two bipartite begomoviruses with typical characteristics of NW viruses, suggests that OW and NW begomoviruses coexisted in the OW prior to continental separation (Ha et al., 2006, 2008).

Begomoviruses are transmitted by *B. tabaci* in a persistent-circulative manner (Ghanim et al., 2007; Hunter et al., 1998). The capsid protein (CP) is the only viral protein essential for acquisition and transmission by the insect vector (Azzam et al., 1994). To determine the role of the *Bean golden mosaic virus* (BGMV) CP in systemic infection and

vector transmission, mutations were introduced in the BGMV *CP* gene. The mutants were able to infect *Phaseolus vulgaris* plants systemically, but the CP was not detected in infected plants and the progeny virus was not transmitted by *B. tabaci*, demonstrating the essentiality of the CP in the transmission process (Azzam et al., 1994). Hofer et al. (1997a) demonstrated that a chimera of the begomovirus *Abutilon mosaic virus* (not whitefly-transmitted) containing the *CP* of *Sida golden mosaic virus* could be acquired and transmitted to multiple hosts by *B. tabaci*, reinforcing the essential role of the CP in the interaction with the insect vector. The CP is also responsible for vector specificity. This was demonstrated by the replacement of the *CP* gene of the begomovirus *African cassava mosaic virus*, transmitted exclusively by *B. tabaci,* by the CP gene of the curtovirus *Beet curly top virus* (BCTV), transmitted by leafhoppers. The resulting chimera was transmitted by the leafhopper vector of BCTV (Briddon et al., 1990).

Here, we describe the molecular and biological characterization of Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV), two bipartite begomoviruses isolated from the non-cultivated plant *Sida acuta.* Although these two viruses share a number of features with OW begomoviruses, both genomic components are phylogenetically related to NW viruses. Interestingly, their CPs is highly divergent, and we hypothesized that they could be defective in particle formation. Indeed, SiYSV was not transmitted by *B. tabaci* MEAM1. Our results provide further insights into the evolution of NW begomoviruses, and suggest that alternative modes of transmission may occur in nature for these agents.

Methods

Cloning of begomovirus genomes

Foliar samples from *Sida acuta* plants showing symptoms of yellow mosaic were collected in the municipality of Viçosa, state of Minas Gerais, Brazil, on December 2011. Total DNA was extracted from press-dried samples as described by Doyle and Doyle (1987). Full-length viral genomes were amplified by rolling-circle amplification as described by Inoue-Nagata et al. (2004). Single genome-length fragments were excised with *Apa*I, *Bam*HI, *Cla*I, *Eco*RI, *Kpn*I, *Pst*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.) 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').

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 obtained and the begomoviruses with greatest similarity (as determined by the BLASTn 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 (Brown et al., 2015). Virus species names were assigned using the guidelines proposed by the *Geminiviridae* Study Group of the International Committee for Taxonomy of Viruses (ICTV) (Brown et al., 2015).

Phylogenetic analysis

Multiple sequence alignments were prepared for the full-length DNA-A and DNA-B and for the CP, Rep, MP and NSP coding sequences using the MUSCLE alignment

option in MEGA6 (Tamura et al., 2013). Phylogenetic trees were constructed using Bayesian inference performed with MrBayes v. 3.0b4 (Ronquist and Huelsenbeck, 2003), with the 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. Trees were visualized using Fig Tree (tree.bio.ed.ac.uk/software/figtree/).

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., 2010). 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 three of the analysis methods available in the program were considered reliable.

Protein prediction analysis

The program Interpro (Quevillon et al., 2005), available at the EMBOSS site (www.ebi.ac.uk/Tools/emboss/), was used to determine conserved structural domains from the observed *CP* and *AV2*-like genes.

Construction of infectious clones

Infectious clones were constructed by cloning DNA-A and DNA-B partial tandem repeats or dimers into the pKS+ vector.

Clones corresponding to the isolates BR:Vic26.2:11 and BR:Vic26.1:11, containing single full-length copies of SiYSV DNA-A and DNA-B, respectively, were cleaved with *Cla*I/*Spe*I and *Bam*HI/*Sal*I, generating fragments of 1,400 bp and 1,100 bp, respectively, containing the origin of replication. These fragments were ligated into pKS+ previously linearized with the same enzymes, generating 0.5mer DNA-A and 0.4mer DNA-B clones. Clones containing partial tandem repeats were obtained by inserting the *Spe*I fragment (1.0mer) from BR:Vic26.2:11 into *Spe*I-digested DNA-A 0.5mer, and the *Sal*I fragment (1.0mer) from BR:Vic26.1:11 into *Sal*I-digested DNA-B 0.4mer, generating 1.5mer and 1.4mer clones of the DNA-A e DNA-B, respectively.

The clone corresponding to the isolate BR:Vic25.1:11, containing a single copy of the SiGYMV DNA-A, was cleaved with *Pst*I, generating a fragment of 2,880 bp (1.0mer). This fragment was religated and amplified by rolling-circle amplification (Inoue-Nagata et al., 2004). Dimeric genomes were obtained by partial digestion with *Sac*I and ligation into pKS+, generating a DNA-A dimeric clone. The clone corresponding to the isolate BR:Vic25.1:11, containing a single copy of the SiGYMV DNA-B, was cleaved with *Sal*I, generating a 2,772 bp (1.0mer) fragment. This fragment was religated and amplified by rolling-circle amplification (Inoue-Nagata et al., 2004). Fragments of 326 (0.12mer, containing the origin of replication) and 2,772 bp (1.0mer) were generated by cleavage with *Bam*HI/*Sac*I and *Sac*I, respectively. These fragments were ligated to pKS+ previously digested with *Bam*HI/*Sac*I, generating a DNA-B 1.12mer clone.

All multimeric recombinant plasmids were confirmed by digestion with restriction enzymes and analysis of the predicted restriction pattern. The clones were preserved in 80% glycerol and stored at -80^{o}C.

An infectious clone of *Oxalis yellow vein virus* (OxYVV) were constructed using the clones corresponding to the isolate BR:Vic07:11 (obtained from *S. acuta* plants

collected at the same site) containing a single copy of the DNA-A and DNA-B. These clones were cleaved with *Cla*I and *Bam*HI, respectively, generating fragments corresponding to the full-length genome. The fragments were religated and used in infectivity assays.

Host range assay

The host range assay was conducted using plants belonging to the families Chenopodiaceae, Fabaceae, Malvaceae and Solanaceae (Table 4). Cotyledons (Fabaceae) or young plants (Chenopodiaceae, Malvaceae and Solanaceae) with two to four pairs of fully expanded leaves were inoculated by biolistics (Aragão et al., 1996) with the SiYSV and SiGYMV infectious clones. Tungsten particles coated with 2 µg of each DNA component were delivered at 40 psi or 50 psi, depending on the plant species. As negative controls, three plants of each species were inoculated with tungsten particles without DNA. Following inoculation, the plants were kept in a greenhouse and observed for symptom development until 40 days post-inoculation (dpi). Total DNA was extracted as described by Doyle and Doyle (1987) at 28 dpi and used as a template for PCR using the DNA-B universal primer pair PBL1v2040/PCRc1 (Rojas et al., 1993) for confirmation of viral infection. The PCR reaction was performed using GoTaq Colorless Master Mix (Promega), following the manufacturer's instructions. The PCR program consisted of 30 cycles (94°C for 1 min, 52°C for 1 min and 72°C for 1 min) with a final extension at 72°C for 10 min.

Whitefly transmission assay

Colonies of aviruliferous *B. tabaci* Middle East-Asia Minor 1 (MEAM1; previously classified as *B. tabaci* biotype B) were maintained in cabbage plants (*Brassica oleracea*) inside insect-proof cages kept in a greenhouse. Adult whiteflies were randomly collected

and transferred to individual *Sida acuta* plants infected with OxYVV (used as a positive control) or SiYSV. At the end of a 72-h acquisition access period (AAP), thirty whiteflies were collected and transferred to one healthy *S. acuta* plant with three pairs of fully expanded leaves, where they were maintained for a 72-h inoculation access period (IAP). Six plants were inoculated for each virus. As a negative control, two *S. acuta* plants were exposed to a 72-h IAP with thirty whiteflies collected directly from the aviruliferous colony. After whitefly elimination with insecticide application (acetamiprid) the plants were transferred to a greenhouse and observed for symptom development until 40 dpi. At this same time, leaf samples from each plant were collected and used for total DNA extraction (Doyle and Doyle, 1987). The DNA was used as template for PCR amplification using DNA-A-specific primers for SiYSV (SiY-2F and SiY2-R) and SiAMV (SiA-1F and SiA-1R), yielding fragments with 1,112 and 1,568 pb, respectively. The PCR reaction was performed using GoTaq Colorless Master Mix (Promega), following the manufacturer's instructions. The PCR program consisted of an initial denaturation at 94°C for 10 min followed by 35 cycles at 94°C for 1 min, 60°C for 1 min and 72°C for 2 min with a final extension at 72°C for 10 min.

Results

Two novel begomoviruses infecting the non-cultivated host Sida acuta

A total of 50 samples were collected, and 47 were preliminarily positive for the presence of a begomovirus based on the detection of a 2,600-bp band after digestion of the RCA products with restriction enzymes (data not shown). Twenty-three complete genomic components were cloned from seven samples, including 15 DNA-A and eight DNA-B components (Table 1). From these, 12 DNA-A and five DNA-B components were randomly chosen for further analysis.

Pairwise comparisons were performed with the most closely related viruses (based on BLAST*n* analysis), plus four OW begomoviruses: *Ludwigia yellow vein virus* and *Sauropus leaf curl virus*, which have a small portion of the genome similar to the *S. acuta* viruses, and *Corchorus golden mosaic virus* and *Corchorus yellow vein virus*, two OW begomoviruses that have features of NW viruses. The results (Figure 1) indicated that the 12 DNA-A clones can be divided into two groups, with within-group identities of 96-100% and 99.6-99.7%, and between-group identities of 87-90.5%. Clones in either group have the highest sequence identity (74-76%) with *Sida yellow mosaic virus* (SiYMV). Therefore, based on the criteria established by the *Geminiviridae* study group of the ICTV (Brown et al., 2015), these two groups correspond to two new begomoviruses species, for which the names Sida golden yellow mosaic virus (SiGYMV) and Sida yellow spot virus (SiYSV) are proposed (Table 1). The five clones corresponding to the DNA-B shared 89.1% to 99.8% sequence identity among them. All five clones have the highest sequence identity (81-83%) with Oxalis yellow vein virus (OxYVV) DNA-B. Four clones, obtained from samples from which SiYSV DNA-A components were cloned, were closely related to each other, with identities of 99.2-99.8%. The remaining clone, obtained from a sample from which SiYGMV was cloned, was distinct from the other four. To verify that these DNA-A and DNA-B clones obtained from the same samples were indeed cognate components, their common region (CR) sequences were compared. The CRs of SiYSV DNA-A components and of the DNA-B clones obtained from the same samples displayed 92-95% identity and their iteron sequences are identical (TGTA; Suppl. Fig. S1). The CRs of SiGYMV DNA-A components and of the DNA-B clone obtained from the same sample displayed 93-94% identity, and also had identical iterons (GGTA; Suppl. Figure S1). Therefore, the DNA-A and DNA-B components cloned from the same samples were considered to be cognate DNA components.

SiYSV and SiGYMV have features of OW begomoviruses

Analyzing the full-length genomes of SiYSV and SiGYMV, we detected a number of features which are typical of OW begomoviruses: (*i*) the DNA-A and DNA-B components of both viruses have lengths in the range of OW begomoviruses (DNA-A, 2,810-2,828 nucleotides and DNA-B, 2,745-2,772 nucleotides; compare with the sizes of the DNA-A and DNA-B components of Oxalis yellow vein virus, also cloned from *S. acuta* samples collected at the same location and on the same date) (Table 1); (*ii*) the DNA-A of both viruses contains an *AV2*-like gene in the virion-sense strand (Figure 2); (*iii*) conserved motifs that are characteristic of OW begomoviruses could be found in the deduced amino acid sequences of their proteins (Tables 2 and 3). The conserved motifs include the KVRRR motif in the N-terminal region of the CP, present in both viruses (Table 2), and the iteron-related domain (IRD) PKRFQI in the Rep protein with the corresponding iteron GGTA (Arguello-Astorga and Ruiz-Medrano, 2001) present in the SiGYMV isolates (Table 3; Suppl. Fig. S1). Furthermore, SiGYMV has an iteron organization resembling that of OW begomoviruses, with three repeats (two direct and one inverted) located upstream and one direct repeat located downstream of the *Rep* gene TATA box (Suppl. Fig. S1). Conversely, SiYSV has a NW-like iteron organization, with all repeats located upstream of the *Rep* gene TATA box (Suppl. Fig. S1). However, it has atypical iteron and IRD sequences (TGTA and AKKIQL, respectively; Suppl. Fig. S1).

Although the most striking features that distinguish OW and NW begomoviruses are located in the DNA-A component, Ho et al. (2014) recently reported the presence of a tyrosine phosphorylation site in the movement protein (MP) of NW bipartite begomoviruses, absent in OW begomoviruses. We detect the presence of this tyrosine phosphorylation site in the MP of both SiYSV and SiGYMV (data not shown).

SiYSV and SiGYMV have a highly divergent DNA-A 5' half

Altough the genomic organization of both species is typical of begomoviruses, all 12 DNA-A clones have a region of approximately 1,100 nt with very low similarity to other begomoviruses. This region includes the *CP* and *AV2*-like genes, as well as part of the common region. The genomic organization of one representative SiYSV clone and the pairwise identities of different regions of the genome are shown in Figure 2. The deduced amino acid sequences of the CP and AV2-like proteins were further analyzed with BLAST*p* and the program Interpro. Despite the divergence of the *CP* gene, the analysis performed with Interpro indicated the presence of a domain related to geminivirus coat proteins. No functional domains were predicted in the AV2-like protein (data not shown). Interestingly, BLAST*p* analysis failed to detect any similarity with the CP and AV2 proteins of other begomoviruses. Only a very low similarity was found with a highly divergent monopartite geminivirus recently described in China infecting apple trees, named "apple geminivirus" (AGV) (Liang et al., 2015). The SiYSV CP shares 26% identity (99% coverage, E value $4e^{-9}$), and the SiGYMV CP shares 25% identity (99% coverage, E value $2e^{-8}$) with the AGV CP. The SiYV AV2-like protein shares 33% identity (76% coverage, E value $4e^{-12}$), and the SiGYMV AV2-like protein shares 39% identity (54% coverage, E value $4e^{-12}$) with the putative V2 protein of AGV.

SiGYMV and SiYSV are phylogenetically related to NW begomoviruses

The Bayesian phylogenetics trees based on both the DNA-A and DNA-B divide the SiYSV and SiGYMV isolates into two distinct clades, corresponding to the two species (Figure 3). Interestingly, and in spite of their OW begomovirus features, SiGYMV and SiYSV isolates cluster with NW begomoviruses in the phylogenetic trees (Figure 3). The DNA-A components were closely related to *Sida yellow mosaic virus* (SiYMV), forming a

monophyletic branch with begomoviruses present in Brazil, Central America and the Caribbean. The DNA-B components are closely related to Oxalis yellow vein virus (OxYVV), grouping with begomoviruses reported in Brazil infecting cultivated (tomato, okra) and non-cultivated (*Sida micrantha*) hosts. A CP-based tree constructed with a different 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) and one topocuvirus (*Tomato pseudo-curly top virus*, TPCTV), along with seven NW begomoviruses and one OW begomovirus, segregated the SiGYMV and SiYSV isolates from both OW and NW begomoviruses, clustering them with AGV (Suppl. Fig. S2). However, the Rep-based tree clustered both species with NW begomoviruses (Suppl. Fig. 2).

The interspecific CP- and Rep-based trees including only SiYSV and SiGYMV isolates were not congruent (data not shown), but the topological differences between them were due to a strongly supported recombination event near the CP gene of SiGYMV-[BR:Vic43.1:11], mentioned below. Without this clone, the clustering of isolates in the CP and Rep trees is identical (data not shown), mimicking the clustering found in the full-length genome tree.

Detection of interspecies recombination events

Analysis with the RDP4 program detected a strongly supported interspecies recombination event in DNA-A of the SiGYMV isolate BR:Vic43.1:11 (lowest *p*-value 3.717×10^{-36} with SiScan). The major (SiYSV-[BR:Vic43.2:11]) and minor (SiGYMV-[BR:Vic25.1:11]) parents have 90.6% and 96.7% identity, respectively, with BR:Vic43.1:11. The portion that corresponds to BR:Vic25.1:11 is between nucleotide

positions 122 to 1,142 in the alignment and includes the entire *CP* and *AV2*-like genes as well as 16 nucleotides of the common region.

Recombination analysis based on the DNA-B detected a well-supported event shared by all SiYSV and SiGYMV isolates (lowest *p*-value 1.821×10^{-15} with MaxChi), having as putative major parent Oxalis yellow vein virus (OxYVV-[BR:Vic07:11]) and an unknown minor parent. The portion that corresponds to OxYVV is between nucleotide positions 293 to 2,308 and includes the entire *MP* and *NSP* genes and the small intergenic region, as well as 296 nucleotides from the large intergenic region.

Experimental hosts of SiYSV and SiGYMV

To determine whether SiYSV and SiGYMV have the potential to infect hosts outside the *Malvaceae* family, several plants species belonging to three families, in addition to four species in the *Malvaceae* family, were inoculated experimentally by biolistics. Both viruses infected *Sida acuta*, the host from which they were isolated, confirming the infectivity of the clones (Table 4). Both viruses also infected *Sida rhombifolia* (Table 4), a non-cultivated species in the *Malvaceae* family, as well as two plants species in the *Solanaceae* family: SiYSV infected *Nicotiana benthamiana* and SiGYMV infected *Nicotiana tabacum* 'TNN'.

SiYSV caused yellow mosaic and leaf distortion in *S. acuta* and SiGYMV caused mild but clear symptoms of yellow mosaic in this host (Figure 4). Both viruses caused severe symptoms in *S. rhombifolia*, including leaf distortion, yellow mosaic and dwarfing in the case of SiYSV, and yellow mosaic, rugosity and leaf curling in the case of SiGYMV (Figure 4). Infection of *N. benthamiana* by SiYSV and of *N. tabacum* by SiGYMV were symptomless (Table 4).

SiYSV is not transmitted by B. tabaci MEAM1

The CPs of both SiSYV and SiYGMV are highly divergent (Figure 5), and we hypothesized that these viruses may be defective in the formation of particles and consequently may be transmitted with low efficiency (or not transmitted) by the whitefly vector. To test this hypothesis, a trasmission experiment was conducted using *B. tabaci* MEAM1. A high density of viruliferous whiteflies was used, together with long (72 hours) AAP/IAP, to assure that even a low efficiency of transmission would be detected. Oxalis yellow vein virus (OxYVV), also isolated from the *S. acuta* samples (Table 1) and which has a typical begomovirus CP, was used as a positive control. Interestingly, SiYSV was not transmitted by *B. tabaci* MEAM1. In turn, OxYVV was transmitted with 100% efficiency (Figure 6).

Discussion

The recent 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 (Bernardo et al., 2013; Krenz et al., 2012; Loconsole et al., 2012; Varsani et al., 2009). Many of these novel begomoviruses have been found in non-cultivated plants (Castillo-Urquiza et al., 2008; Fiallo-Olive et al., 2012; Paprotka et al., 2010; Silva et al., 2012; Tavares et al., 2012), highlighting the need to investigate these plants as a reservoir of viral diversity and as a source of new viruses which may cause diseases in crops. In addition, the study of divergent viruses often found in non-cultivated plants contributes to a better understanding of viral evolution.

Sida yellow spot virus and Sida golden yellow vein virus were both isolated from the non-cultivated host *Sida acuta*. Malvaceous hosts in general, and *Sida* spp. in

particular, are notorious begomovirus hosts in the Americas (Andrade et al., 2006; Frischmuth et al., 1997; Hofer et al., 1997b; Roye et al., 1997), and several new begomovirus species have been described infecting *Sida* spp. in Brazil (Jovel et al., 2004; Tavares et al., 2012). However, while all previously described species are typical NW begomoviruses, SiYSV and SiGYMV have several features recalling OW begomoviruses: (*i*) the presence of an *AV2*-like gene; (*ii*) a longer genome, in the range of OW begomoviruses; (*iii*) the presence of the KVRRR motif in the CP; (*iv*) the absence of the (P/S)WRxMxGT motif in the CP; (*v*) the presence of an iteron (GGTA) and its respective IRD (PKRFQI) found in several OW begomoviruses but in only one NW begomovirus. Furthermore, the two viruses have highly divergent capsid proteins, displaying only a distant similarity with the CP of a geminivirus recently described infecting apple in China. This is specially striking since the CP is the most conserved protein among begomoviruses. Interestingly, SiYSV was not transmitted by *Bemisi tabaci* MEAM1, suggesting that it may be defective in particle formation or that it may have a distinct vector.

It is generally accepted that NW begomoviruses arose more recently than the OW viruses, diverging after continental separation (Rybicki, 1994). However, not only have NW-like viruses been found in the OW (CoYVV and CoGMV), but now also OW-like viruses have been found in the NW. Thus, it seems that the divergence of these two major clades predates continental separation. Further work is necessary to assess the distribution of SiGYMV and SiYSV, as well as the presence of additional OW-like viruses in other non-cultivated hosts in the Americas.

Diversification of begomoviruses is greatly accelerated by frequent recombination (Lefeuvre et al., 2009; Padidam et al., 1999; Rocha et al., 2013), particularly in the context of adaptation to new host species and vector biotypes (Berrie et al., 2001; Monci et al., 2002; Padidam et al., 1999; Sanz et al., 2000). Confirming previous reports of

frequent interspecies recombination among Brazilian begomoviruses (Inoue-Nagata et al., 2006; Ribeiro et al., 2007; Rocha et al., 2013), we found evidence of interspecies recombination involving both DNA components of SiGYMV and SiYSV. Besides the recombination events detected by RDP4 (whose results are strongly dependent on the quality of the alignment, limiting the analysis to viruses with a medium or high degree of similarity), the *CP* gene is an obvious candidate to having a recombinant origin: the CP is phylogenetically related to the recently described "apple geminivirus" (AGV), a divergent monopartite geminivirus identified in China infecting apple (Liang et al., 2015 19506). It is tempting to speculate that SiYSV and SiGYMV arose as a result of a recombination between a NW begomovirus and a highly divergent OW geminivirus. A divergent geminiviruses has recently been reported in Brazil, coincidentally also infecting apple (and also pear and grapevine) (Basso et al., 2015), however its CP does not display any significant similarity with those of SiYSV and SiYGMV (data not shown). Of course the true diversity of geminiviruses (and of circular ssDNA viruses in general) is only starting to be truly sampled, and it is perfectly possible that additional divergent viruses will be detected in Brazil and in other countries in the Americas in the near future, providing support for our hypothesis on the origin of SiYSV and SiGYMV.

The protein encoded by the *AV2* gene is involved in symptom development, viral movement and viral DNA accumulation (Padidam et al., 1996; Rojas et al., 2001). The *AV2* gene overlaps the *CP* gene region encoding the N-terminal motif present only in OW viruses, suggesting that functions normally attributed to this CP N-terminal motif and to AV2 are encoded by the DNA-B component, which is present in all but one NW virus. Curiously, we could not find a predicted structure for the AV2-like proteins from SiGYMV and SiYSV with the algorithms used. Out of the several highly divergent geminiviruses that have been recently discovered and characterized, *Euphorbia caput-medusae* latent

virus (EcmLV) (Bernardo et al., 2013) and apple geminivirus (AGV) (Liang et al., 2015) have highly divergent V2 proteins, with no significant degrees of identity with any other known protein. Interestingly, the AV2-like protein of SiYSV and SiGYMV has similarity to the AGV V2. AGV and EcmLV experimentally infected *Nicotiana bethamiana* and *Solanum lycopersicum*, suggesting that their V2 proteins are functional in the movement process (Bernardo et al., 2013; Liang et al., 2015). Therefore, and in spite of the presence of a DNA-B component for the two *S. acuta* viruses, it is not unreasonable to suppose that the AV2 protein encoded by SiYSV and SiGYMV may be functional. The functional characterization of these proteins is necessary to confirm this hypothesis.

Besides their OW-like features, both SiGYMV and SiYSV have a highly divergent CP, with very low identity with other begomoviruses. This is unexpected considering that the CP is the most conserved protein among begomoviruses, being essential not only for particle formation but also for vector specificity and transmission (Azzam et al., 1994; Briddon et al., 1990; Hofer et al., 1997a). Although the deduced amino acid sequences of the SiYSV and SiGYMV CPs include a gemini-like coat protein domain, it is possible that these viruses could be defective in particle formation (and therefore also in vector transmission).

We conducted a vector transmission assay to verify the possible influence of their highly divergent CP in the whitefly's ability to acquire and transmit SiYSV and SiGYMV. Indeed, SiYSV was not transmitted, suggesting either that (*i*) this virus is defective in particle formation, (*ii*) that it may be transmitted by another insect vector, or (*iii*) that it may have an alternative mode of natural transmission. We are testing the first hypothesis by attempting to purify the viruses from *S. acuta* plants and viasualizing geminate viral particles in the transmission electron microscope.

The known geminivirus vectors include different species of leafhoppers (Hemiptera: Cicadellidae), one species of treehopper (Hemiptera: Membracidae), and species of the *Bemisia tabaci* sibling species group (Hemiptera: Aleyrodidae) (Dinsdale et al., 2010; Rojas et al., 2005). However, Roumagnac et al. (2015) have recently demonstrated that a divergent geminivirus named Alfalfa leaf curl virus (ALCV, related to EcmLV) is transmitted by the aphid *Aphis craccivora*, the first report of an aphid-transmitted geminivirus. Also, (Kim et al., 2015) recently reported the transmission of a begomovirus, *Sweet potato leaf curl virus* (SPLCV) by the seed of sweet potato, the first report of a seed-transmitted geminivirus. Together, these reports highlight the gaps in our knowledge of geminivirus biology, and should serve as a reminder that every possibility should be taken into consideration when studying divergent viruses. Therefore, the possibilities that SiYSV and SiGYMV may have a different type of insect vector or may be seed-transmitted in *S. acuta* are also being currently investigated.

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.

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.

Aragão, F.J.L., Barros, L.M.G., Brasileiro, A.C.M., Ribeiro, S.G., Smith, F.D., Sanford, J.C., Faria, J.C., Rech, E.L., 1996. Inheritance of foreign genes in transgenic bean (*Phaseolus vulgaris L.*) co-transformed via particle bombardment. Theor Appl Genet 93, 142-150.

Arguello-Astorga, G.R., Ruiz-Medrano, R., 2001. An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. Arch Virol 146, 1465-1485.

Azzam, O., Frazer, J., De La Rosa, D., Beaver, J.S., Ahlquist, P.G., Maxwell, D.P., 1994. Whitefly transmission and efficient ssDNA accumulation of bean golden mosaic geminivirus require functional coat protein. Virology 204, 289-296.

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 Res.

Bernardo, P., Golden, M., Akram, M., Naimuddin, Nadarajan, N., Fernandez, E., Granier, M., Rebelo, A.G., Peterschmitt, M., Martin, D.P., Roumagnac, P., 2013. Identification and characterisation of a highly divergent geminivirus: Evolutionary and taxonomic implications. Virus Res 177, 35-45.

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., 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.

Briddon, R.W., Pinner, M.S., Stanley, J., Markham, P.G., 1990. Geminivirus coat protein gene replacement alters insect specificity. Virology 177, 85-94.

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-Olive, 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., doi 10.1111/pbi.12349.

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.

Dinsdale, A., Cook, L., Riginos, C., Buckley, Y.M., De Barro, P., 2010. Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. Annals of the Entomological Society of America 103, 196-208.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochem Bull 19, 11-15.

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

Fiallo-Olive, 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.

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.

Ghanim, M., Sobol, I., Ghanim, M., Czosnek, H., 2007. Horizontal transmission of begomoviruses between *Bemisia tabaci* biotypes. Arthropod-Plant Inte 1, 195-204.

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.

Ha, C., Coombs, S., Revill, P., Harding, R., Vu, M., Dale, J., 2006. Corchorus yellow vein virus, a New World geminivirus from the Old World. J Gen Virol 87, 997-1003.

Ha, C., Coombs, S., Revill, P., Harding, R., Vu, M., Dale, J., 2008. Molecular characterization of begomoviruses and DNA satellites from Vietnam: Additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J Gen Virol 89, 312-326.

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.

Harrison, B.D., Swanson, M.M., Fargette, D., 2002. Begomovirus coat protein: serology, variation and functions. Physiol Mol Plant Pathol 60, 257-271.

Ho, E.S., Kuchie, J., Duffy, S., 2014. Bioinformatic analysis reveals genome size reduction and the emergence of tyrosine phosphorylation site in the movement protein of New World bipartite begomoviruses. PLoS One 9, e111957.

Hofer, P., Bedford, I.D., Markham, P.G., Jeske, H., Frischmuth, T., 1997a. Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. Virology 236, 288-295.

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

Hunter, W.B., Hiebert, E., Webb, S.E., Tsai, J.H., Polston, J.E., 1998. Location of geminiviruses in the whitefly Bemisia tabaci (Homoptera : Aleyrodidae). Plant Dis 82, 1147-1151.

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. J Virol Met 116, 209-211.

Inoue-Nagata, A.K., Martin, D.P., Boiteux, L.S., Giordano, L.D., Bezerra, I.C., de Avila, A.C., 2006. New species emergence via recombination among isolates of the Brazilian tomato-infecting begomovirus complex. Pesqui Agropecu Bras 41, 1329-1332.

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.

Kim, J., Kil, E.J., Kim, S., Seo, H., Byun, H.S., Park, J., Chung, M.N., Kwak, H.R., Kim, M.K., Kim, C.S., Yang, J.W., Lee, K.Y., Choi, H.S., Lee, S., 2015. Seed transmission of Sweet potato leaf curl virus in sweet potato (Ipomoea batatas). Plant Pathol, doi 10.1111/ppa.12366.

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

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.

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

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. J Gen Virol, doi 10.1099/vir.0.000173.

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 Res 157, 25-34.

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.

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. J Virol 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.

Moriones, E., Navas-Castillo, J., 2000. *Tomato yellow leaf curl virus*, an emerging virus complex causing epidemics worldwide. Virus Res 71, 123-134.

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., 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.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez , R., 2005. InterProScan: protein domains identifier. Nucleic Acids Res 33, W116-W120.

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. J Virol 87, 5784-5799.

Rojas, M.R., Gilbertson, R.L., Russell, D.R., Maxwell, D.P., 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis 77, 340-347.

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 Phytopath 43, 361-394.

Rojas, M.R., Jiang, H., Salati, R., Xoconostle-Cazares, B., Sudarshana, M.R., Lucas, W.J., Gilbertson, R.L., 2001. Functional analysis of proteins involved in movement of the monopartite begomovirus, tomato yellow leaf curl virus. Virology 291, 110-125.

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

Roumagnac, P., Granier, M., Bernardo, P., Deshoux, M., Ferdinand, R., Galzi, S., Fernandez, E., Julian, C., Abt, I., Filloux, D., Mesleard, F., Varsani, A., Blanc, S., Martin, D.P., Peterschmitt, M., 2015. Alfalfa leaf curl virus: An aphid-transmitted geminivirus. J Virol., doi 10.1128/JVI.00453-15.

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.

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

Sanz, A.I., Fraile, A., García-Arenal, F., Zhou, X., Robinson, D.J., Khalid, S., Butt, T., Harrison, B.D., 2000. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. J Gen Virol 81, 1839-1849.

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.

Sobrinho, R.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. J Gen Virol 95, 2540-2552.

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.

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., Shepherd, D.N., Dent, K., Monjane, A.L., Rybicki, E.P., Martin, D.P., 2009. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virol J 6, 36.

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

Table 1. Begomovirus sequences reported in this study.

Sample code	Clone		Isolate	Enzyme[1]		Size (nt)		Species[2]	GenBank access number	
	DNA-A	DNA-B		DNA-A	DNA-B	DNA-A	DNA-B		DNA-A	DNA-B
25D	VIC25D_1P	VIC25D_4SI	BR:Vic25.1:11	*Pst*I	*Sal*I	2810	2772	SiGYMV		
	VIC25D_2A		BR:Vic25.2:11	*Apa*I		2813		SiGYMV		
	VIC25D_2P		BR:Vic25.3:11	*Pst*I		2810		SiGYMV		
43D	VIC43D_1S		BR:Vic43.1:11	*Spe*I		2826		SiGYMV		
25D	VIC25D_1S	VIC25D_5SI	BR:Vic25.4:11	*Spe*I	*Sal*I	2828	2755	SiYSV		
26D	VIC26D_3S	VIC26D_2SI	BR:Vic26.1:11	*Spe*I	*Sal*I	2828	2745	SiYSV		
	VIC26D_4S		BR:Vic.26.2:11	*Spe*I		2828		SiYSV		
39D	VIC39D_1C	VIC39D_1SI	BR:Vic39.1:11	*Cla*I	*Sal*I	2828	2756	SiYSV		
	VIC39D_1P	VIC39D_6SI	BR:Vic39.2:11	*Pst*I	*Sal*I	2828	2757	SiYSV		
	VIC39D_1S		BR:Vic39.3:11	*Spe*I		2828		SiYSV		
	VIC39D_5S		BR:Vic39.4:11	*Spe*I		2828		SiYSV		
43D	VIC43D_5S		BR:Vic43.2:11	*Spe*I		2828		SiYSV		
03D	VIC03D_1C	VIC03D_2B	BR:Vic03:11	*Cla*I	*Bam*HI	2659	2613	OxYVV		
07D	VIC07D_1C	VIC07D_2B	BR:Vic07:11	*Cla*I	*Bam*HI	2661	2610	OxYVV		
19D	VIC19D_2C	VIC19D_1B	BR:Vic19:11	*Cla*I	*Bam*HI	2658	2612	OxYVV		

[1]Enzyme used for releasing genome-length DNA components after rolling-circle amplification and subsequent cloning into the plasmid vector pKS+.
[2]SiGYMV, Sida golden yellow mosaic virus; SiYSV, Sida yellow spot virus; OxYVV, Oxalis yellow vein virus.

Table 2. Comparison of deduced amino acid sequences from the N-terminal region of the capsid protein of Sida golden yellow mosaic virus (SiGYMV) and Sida yellow spot virus (SiYSV) isolates determined in this study, with begomoviruses from the Old and New Worlds.

Virus	N-terminal CP amino acid sequence	Type	Origin
SiGYMV-[BR:Vic25.1:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiGYMV-[BR:Vic25.2:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiGYMV-[BR:Vic25.3:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiGYMV-[BR:Vic43.1:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic25.4:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic26.1:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic.26.2:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic39.1:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic39.2:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic39.3:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic39.4:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
SiYSV-[BR:Vic43.2:11]	MDYRKRKRSYLTPLTPDQIAKRFRQIQSGIGVRRRLSYPYVPQGIPRLPVKSR	OW	NW (this study)
Corchorus yellow vein mosaic virus	MPKRPLDMSISTPGWKVRRRLNFDSPYSARAAVPIVPVISKSR	OW	OW
Ludwigia yellow vein virus	MSKRPADIVISTPVSKVRRRLNFDSPGMNRAAARTVLG	OW	OW
Tomato yellow leaf curl virus	MSKRPGDIIISTPVSKVRRRLNFDSPYSSRAAVPIVQGTNKRRSW	OW	OW
African cassava mosaic virus	MSKRPGDIIISTPVSKVRRRLNFDSPYSSRAAVPIVQGTNKRRSW	OW	OW
Indian cassava mosaic virus	MSKRPGDIIISTPGSKVRRRLNFDSPYRNRATAPTVHVTNRKRAW	OW	OW
Loofa yellow mosaic virus	MSKRPADIIISTPGSKVRRRLNFDSPYSSRAAVPTVRVTKRQSW	OW	OW
Squash leaf curl China virus	MSKRPADIIISTPASKVRRRLNFDSPYVSRAVVPIARVTKGKAW	OW	OW
Corchorus yellow vein virus	MPKRDAPWRLMAGTSKVSRSSNYSPRGGVSDSGSYLPRRFSRASL	NW	OW
Sida mottle virus	MPKRDPSWRQMAGTSKVSRSSNFSPRGGIGPKFNKASEW	NW	NW
Cabbage leaf curl virus	MPKRDAPWRSMAGTSKVSRNANYSPRAGMIHKFDKAAAW	NW	NW
Squash leaf curl virus	MVKRDAPWRLMAGTSKVSRSANFSPREGMGPKFNKAAAW	NW	NW
Tomato golden mottle virus	MPKRDAPWRLMGGTSKVSRSFNQVSRTGTGPKFDKAHAW	NW	NW
Macroptilium mosaic Puerto Rico virus	MPKRDAPWRSSAGTSKVSRNLNYSPGGGPKSNRANAW	NW	NW
Rhynchosia golden mosaic virus	MPKRDAPWRLSAGTSKVSRSANYSPGGGMGPKSNRANAW	NW	NW

Table 3. Comparison of deduced amino acid sequences of the N-terminal region of the Rep protein of Sida golden yellow mosaic virus (SiGYMV) isolates determined in this study, with begomoviruses from the Old and New Worlds.

Virus	N-terminal Rep amino acid sequence	Iteron	Origin
SiGYMV-[BR:Vic25.1:11]	MPPPKRFQINAKNYFLTY	GGTA	NW (this study)
SiGYMV-[BR:Vic25.2:11]	MPPPKRFQINAKNYFLTY	GGTA	NW (this study)
SiGYMV-[BR:Vic25.3:11]	MPPPKRFQINAKNYFLTY	GGTA	NW (this study)
SiGYMV-[BR:Vic43.1:11]	MPPPKRFQINAKNYFLTY	GGTA	NW (this study)
Ageratum yellow vein virus	MAPPKRFQINAKNYFLTY	GGTA	OW
Indian cassava mosaic virus	MSPPKRFQINAKNYFLTY	GGTA	OW
Bean golden mosaic virus	MPPPKRFKINAKNYFLTY	GGAG	NW
Abutilon mosaic virus	MPPPKKFRIQAKNYFLTY	GGAG	NW

Table 4. Results of biolistic inoculation of Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV).

Family	Species and cultivar	SiYSV		SiGMV	
		Infected/ inoculated	Symptoms[1]	Infected/ inoculated	Symptoms[1]
Solanaceae	*Nicotiana benthamiana*	1/12	as	0/6	-
	Nicotiana glutinosa	0/5	-	0/6	-
	Nicotiana rustica	0/14	-	0/6	-
	Nicotiana sylvestris	0/6	-	0/6	-
	Nicotiana debneyi	0/14	-	0/6	-
	Nicotiana tabacum 'TNN'	0/6	-	1/6	as
	Nicotiana tabacum 'Xanthi'	0/5	-	0/3	-
	Nicotiana tabacum 'Havana'	0/5	-	0/6	-
	Nicotiana tabacum 'White Burley'	0/6	-	0/6	-
	Solanum lycopersicum 'Santa Clara'	0/12	-	0/6	-
	Solanum lycopersicum 'Rutgers'	0/6	-	-	-
	Capsicum annum 'Ikeda'	0/14	-	0/5	-
Fabaceae	*Phaseolus vulgaris* 'Ouro Negro'	0/6	-	0/6	-
	Phaseolus vulgaris 'Pérola'	0/6	-	0/6	-
Chenopodiaceae	*Chenopodium quinoa*	0/8	-	-	-
Malvaceae	*Sida acuta*	10/19	ym, ld	4/18	mym
	Sida rhombifolia	1/8	ld, ym, dw	2/6	ld, mym, ru, lc
	Sidastrum micranthum	0/6	-	0/6	-
	Gossipium sp.	0/13	-	0/6	-

[1]Symptoms in systemically infected leaves, thirty days after inoculation: as, asymptomatic; ld, leaf distortion; ym, yellow mosaic; mym, mild yellow mosaic; dw, dwarfing; ru, rugosity; lc, leaf curl.

Supplementary Table S1. Geminivirus sequences retrieved from GenBank.

Species	Acronym	Access number		Origin
		DNA-A	DNA-B	
Abutilon mosaic Brazil virus	AbMBV	JF694480.1	JF694483.1	NW
Abutilon mosaic Bolivia virus	AbMBoV	HM585445.1	HM585446.1	NW
Abutilon mosaic virus	AbMV	LN611623.1	LN611625.1	NW
Bean dwarf mosaic virus	BDMV	M88179.1	M88180.1	NW
Bean golden mosaic virus	BGMV	KJ939853.1	n.d.[1]	NW
Blainvillea yellow spot virus	BlYSV	KC706520.1	KC706525.1	NW
Cleome leaf crumple virus	ClLCrV	FN435999.1	FN436000.1	NW
Corchorus yellow spot virus	CoYSV	DQ875868.1	DQ875869.1	NW
Cotton leaf crumple virus	CLCrV	AY742220.1	AY742221.1	NW
Euphorbia mosaic virus	EuMV	JQ963887.1	JQ963888.1	NW
Euphorbia yellow mosaic virus	EuYMV	FJ619507.1	FJ619508.2	NW
Leonurus mosaic virus	LeMV	JX863082.1	n.a.[2]	NW
Okra mottle virus	OMoV	EU914817.1	EU914818.1	NW
Oxalis yellow vein virus	OxYVV	KM887907.1	n.a.	NW
Passionfruit severe leaf distortion virus	PSLDV	FJ972767.1	FJ972768.1	NW
Pepper golden mosaic virus	PepGMV	AY928514.1	AY928515.1	NW
Rhynchosia golden mosaic Yucatan virus	RhGMYuV	GQ352453.1	FJ792608.1	NW
Sida golden mosaic virus	SiGMV	GQ357649.1	EF537046.1	NW
Sida micrantha mosaic virus	SiMMV	FN557522.1	FN557523.1	NW
		FN436003.1	FN436004.1	
Sida mosaic Bolivia virus	SiMBoV	HM585441.1	HM585442.1	NW
Sida mottle virus	SiMoV	JX871378.1	n.a.	NW
Sida yellow leaf curl virus	SiYLCV	KC706539.1	n.a.	NW
Sida yellow mosaic virus	SiYMV	JX871369.1	n.a.	NW
Squash leaf curl virus	SLCV	M38183.1	M38182.1	NW
Tomato chlorotic mottle virus	ToCMoV	KC706557.1	KC706562.1	NW
Tomato common mosaic virus	ToCmMV	KC706570.1	KC706595.1	NW
Tomato golden mosaic virus	TGMV	JF694488.1	JF694489.1	NW
Tomato mild mosaic virus	ToMiMV	KC706611.1	KC706613.1	NW
Tomato rugose mosaic virus	ToRMV	AF291705.1	AF291706.1	NW
Tomato severe rugose virus	ToSRV	KC706625.1	KC004086.1	NW
Tomato yellow spot virus	ToYSV	JX513952.1	JX513953.1	NW
Tomato mottle virus	ToMoV	AY965900.1	AY965901.2	NW
Tomato pseudo-curly top virus	TPCTV	X84735.1	-[3]	
African cassava mosaic virus	ACMV	FN668378.1	FN668379.1	OW
Ageratum yellow vein virus	AYVV	KJ744212.1	-	OW
Apple geminivirus	AGV	NC026760.1	-	OW
Citrus chlorotic dwarf associated virus	CCDaV	NC018151.1	-	OW
Corchorus golden mosaic virus	CoGMV	FJ463902.1	FJ463901.1	OW
Corchorus yellow vein virus	CoYVV	AY727903.1	AY727904.1	OW
Euphorbia caput-medusae latent virus	EcmLV	HF921477.1	-	OW
East African cassava mosaic virus	EACMV	HE979774.1	HE979778.1	OW
Grapevine red-blotch associated virus	GRBaV	Q901105.2	-	OW

Suppl. Table S1 (cont.)

Indian cassava mosaic virus	ICMV	NC001932.1	-	OW
Ludwigia yellow vein virus	LuYVV	DQ641708.1	-	OW
Sauropus leaf curl virus	SaLCV	JN809827.1	-	OW
Tomato yellow leaf curl virus	TYLCV	JX128099.1	-	OW
Watermelon chlorotic stunt virus	WmCSV	KJ939448.1	KJ939447.1	OW

[1] Sequence not yet deposited in GenBank.
[2] Not available.
[3] Monopartite geminivirus.

Figure legends

Figure 1. Pairwise sequence identity matrices of the DNA-A and DNA-B of Sida yellow spot virus (SiYSV, in red) and Sida golden yellow mosaic virus (SiGYMV, in green) isolates with the most closely related begomoviruses.

Figure 2. Genome organization of Sida yellow spot virus (SiYSV) and pairwise nucleotide sequence identities of the virion-sense and complementary-sense genes with the most closely related begomoviruses.

Figure 3. Bayesian phylogenetic tree based on the complete DNA-A (A) and DNA-B (B) nucleotide sequences of Sida yellow spot virus (SiYSV, in green) and Sida golden yellow mosaic virus (SiGYMV, in red) isolates, plus begomovirus from the New World (NW) and Old World (OW) (see Table S1 for full names and GenBank access numbers). 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 0.60 and 0.80 are indicated by empty circles, nodes with values between 0.81 and 0.89 are indicated by half-filled circles and those with values equal to or greater than 0.90 are indicated by filled circles. Isolates marked in blue correspond to Oxalis yellow vein virus (OyVV).

Figure 4. Symptoms induced in *Sida acuta* (A-E) and *Sida rhombifolia* (F-J) plants, biolistically inoculated with infectious clones of Sida yellow spot virus (SiYSV) (B, C, G and H) and Sida golden yellow mosiac virus (SiGYMV) (D, E, I and J). Mock-inoculated plants (A and F) correspond to negative controls. Images were obtained 40 days after inoculation. Images B and C, D and E, G and H, I and J, correspond to the same plants, at different magnifications.

Figure 5. Alignment of the deduced amino acid sequences of the capsid protein (CP) of Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV) isolates with the begomoviruses *Bean golden mosaic virus* (BGMV), *Sida micrantha mosaic virus* (SiMMV), *Tomato severe rugose virus* (ToSRV), Oxalis yellow vein virus (OxYVV) and *Tomato golden mosaic virus* (TGMV). Amino acid residues that are identical among all sequences are highlighted in black, and those that are similar are highlighted in gray. Amino acid residues that are identical only between SiYSV and SiGYMV sequences are highlighted in blue. Amino acid residues that are only identical among OxYVV, SiMMV, ToSRV, TGMV and BGMV sequences are highlighted in green. The isolates BR:Vic.26.2:11 and BR:Vic25.1:11 were used to represent SiYSV and SiGYMV, respectively, since the intraspecific identity of the *CP* gene is 100%.

Figure 6. PCR-based virus detection in *Sida acuta* plants inoculated with Sida yellow spot virus (SiYSV) or Oxalis yellow vein virus (OxYVV) (positive control for transmission) using whiteflies (*Bemisia tabaci* MEAM1). M, size marker (1Kb plus DNA ladder, Invitrogen); +, PCR positive control (total DNA from plant infected with SiYSV or OxYVV); -, negative control (no DNA); Mock, plants inoculated with aviruliferous whiteflies.

Supplementary Figure S1. Alignment of the intergenic region (IR) of the DNA-A and DNA-B of Sida yellow spot virus (SiYSV) and Sida golden yellow mosaic virus (SiGYMV) isolates, showing putative iterons (gray boxes; arrows indicate orientation) located near the TATA-box of the *Rep* gene (box with solid line and bold nucleotides). The nonanucleotide is highlighted in bold and underlined. Red nucleotides differ from the "core" iteron sequence.

Supplementary Figure S2. Bayesian phylogenetic tree based on the nucleotide sequences of the *CP* (A) and *Rep* (B) genes of Sida yellow spot virus (SiYSV, in green) and Sida golden

yellow mosaic virus (SiGYMV, in red). The unrooted *CP*-based tree includes four highly divergent geminivirus, Apple geminivirus (AGV), Citrus chlorotic dwarf associated virus (CCDaV), *Euphorbia caput-medusae* latent virus (ECMLV) and Grapevine red-blotch associated virus (GRBaV), and one curtovirus, *Tomato pseudo-curly top virus* (TPCTV), in addition to seven NW and OW begomoviruses. The *Rep*-based tree includes NW and OW begomoviruses. 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 0.60 and 0.89 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles. Isolates marked in blue correspond to Oxalis yellow vein virus (OxYVV).

Suplementary Figure S3. Bayesian phylogenetic tree based on the nucleotide sequences of the *MP* (A) and *NSP* (B) genes of Sida yellow spot virus (SiYSV, in green) and Sida golden yellow mosaic virus (SiGYMV, in red) isolates, plus begomovirus from the New World (NW) and Old World (OW). 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 0.60 and 0.80 are indicated by empty circles and those with values equal to or greater than 0.90 are indicated by filled circles. Isolates marked in blue correspond to Oxalis yellow vein virus (OxYVV).

Figure 1

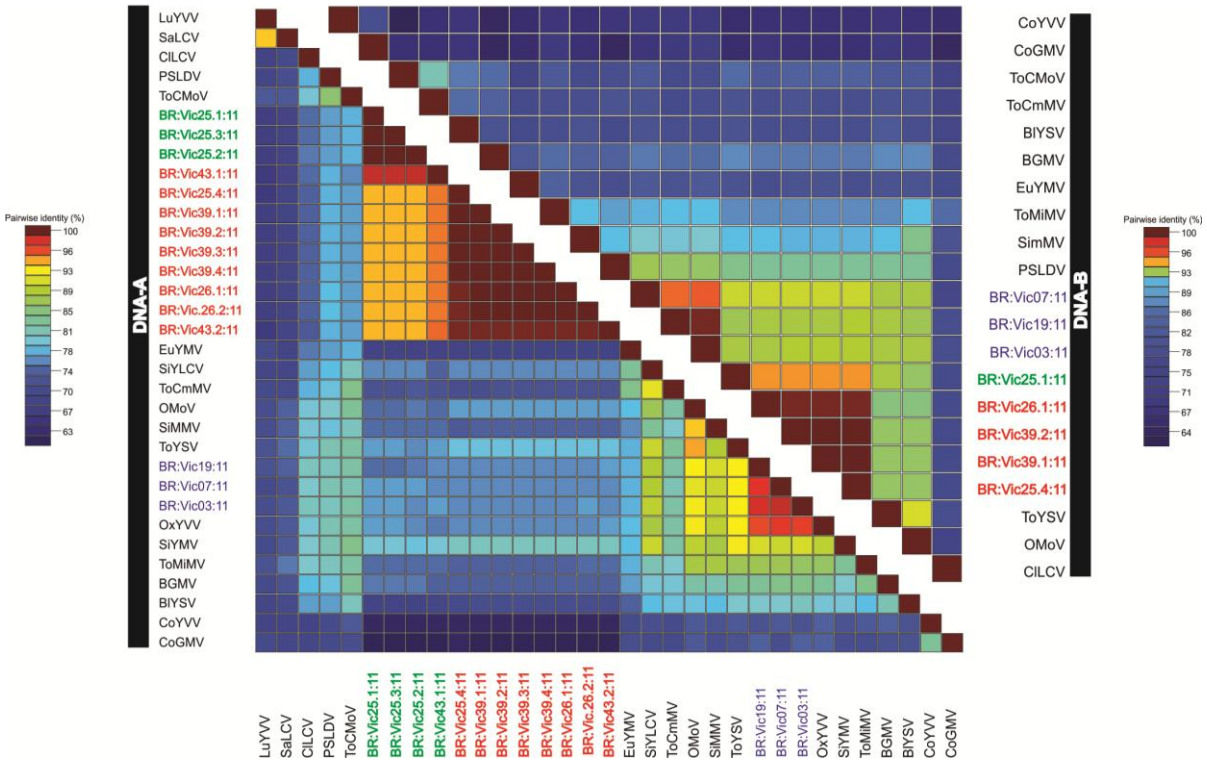

Pairwise identity (%)
100
96
93
89
85
81
78
74
70
67
63
DNA-A
LuYVV
SaLCV
CILCV
PSLDV
ToCMoV
BR:Vic25.1:11
BR:Vic25.3:11
BR:Vic25.2:11
BR:Vic43.1:11
BR:Vic25.4:11
BR:Vic39.1:11
BR:Vic39.2:11
BR:Vic39.3:11
BR:Vic39.4:11
BR:Vic26.1:11
BR:Vic.26.2:11
BR:Vic43.2:11
EuYMV
SiYLCV
ToCmMV
OMoV
SiMMV
ToYSV
BR:Vic19:11
BR:Vic07:11
BR:Vic03:11
OxYVV
SiYMV
ToMiMV
BGMV
BlYSV
CoYVV
CoGMV
DNA-B
CoYVV
CoGMV
ToCMoV
ToCmMV
BlYSV
BGMV
EuYMV
ToMiMV
SimMV
PSLDV
BR:Vic07:11
BR:Vic19:11
BR:Vic03:11
BR:Vic25.1:11
BR:Vic26.1:11
BR:Vic39.2:11
BR:Vic39.1:11
BR:Vic25.4:11
ToYSV
OMoV
CILCV
Pairwise identity (%)
100
96
93
89
86
82
78
75
71
67
64

Figure 2

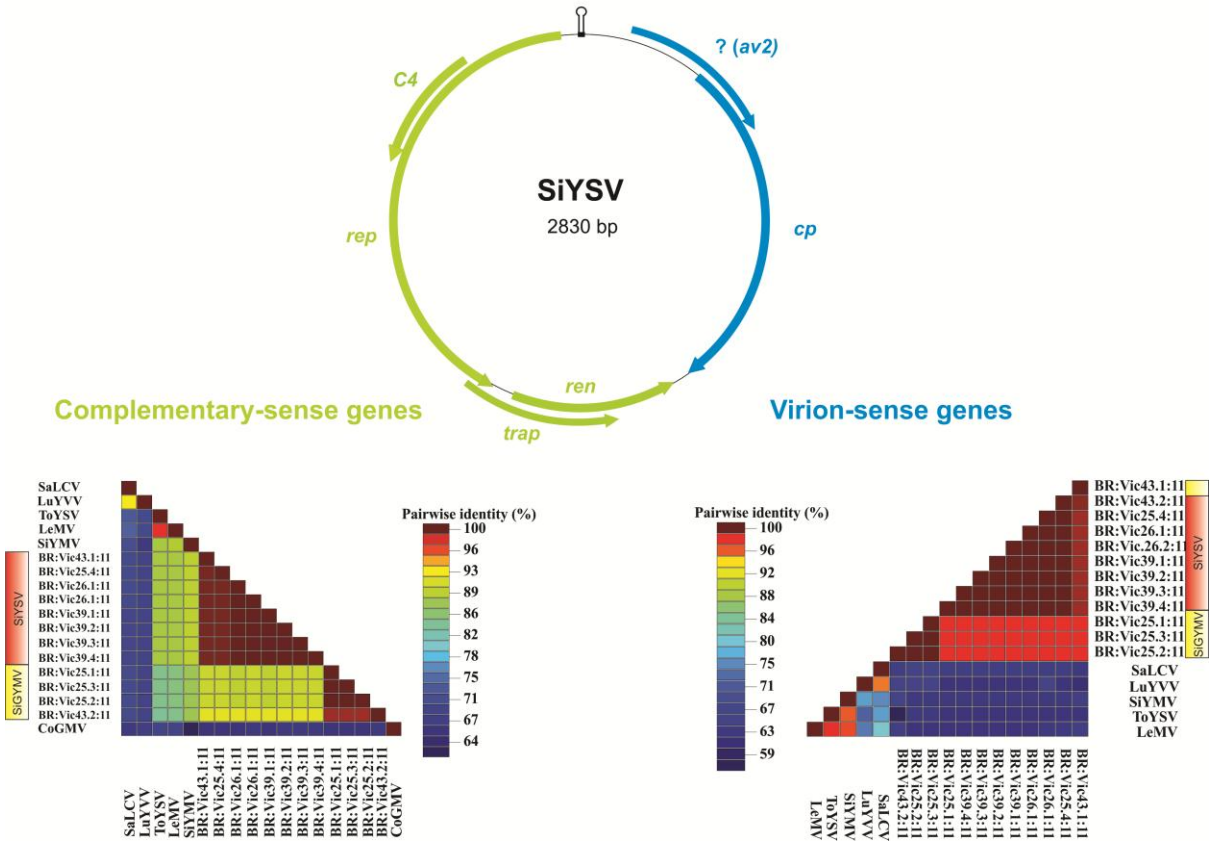

? (av2)
C4
SiYSV
2830 bp
cp
rep
ren
trap
Complementary-sense genes
Virion-sense genes
Pairwise identity (%)
100
96
93
89
86
82
78
75
71
67
64
SaLCV
LuYVV
ToYSV
LeMV
SiYMV
BR:Vic43.1:11
BR:Vic25.4:11
BR:Vic26.1:11
BR:Vic26.1:11
BR:Vic39.1:11
BR:Vic39.2:11
BR:Vic39.3:11
BR:Vic39.4:11
BR:Vic25.1:11
BR:Vic25.3:11
BR:Vic25.2:11
BR:Vic43.2:11
CoGMV
SiYSV
SiGYMV
Pairwise identity (%)
100
96
92
88
84
80
75
71
67
63
59
BR:Vic43.1:11
BR:Vic43.2:11
BR:Vic25.4:11
BR:Vic26.1:11
BR:Vic.26.2:11
BR:Vic39.1:11
BR:Vic39.2:11
BR:Vic39.3:11
BR:Vic39.4:11
BR:Vic25.1:11
BR:Vic25.3:11
BR:Vic25.2:11
SaLCV
LuYVV
SiYMV
ToYSV
LeMV
SiYSV
SiGYMV

Figure 3

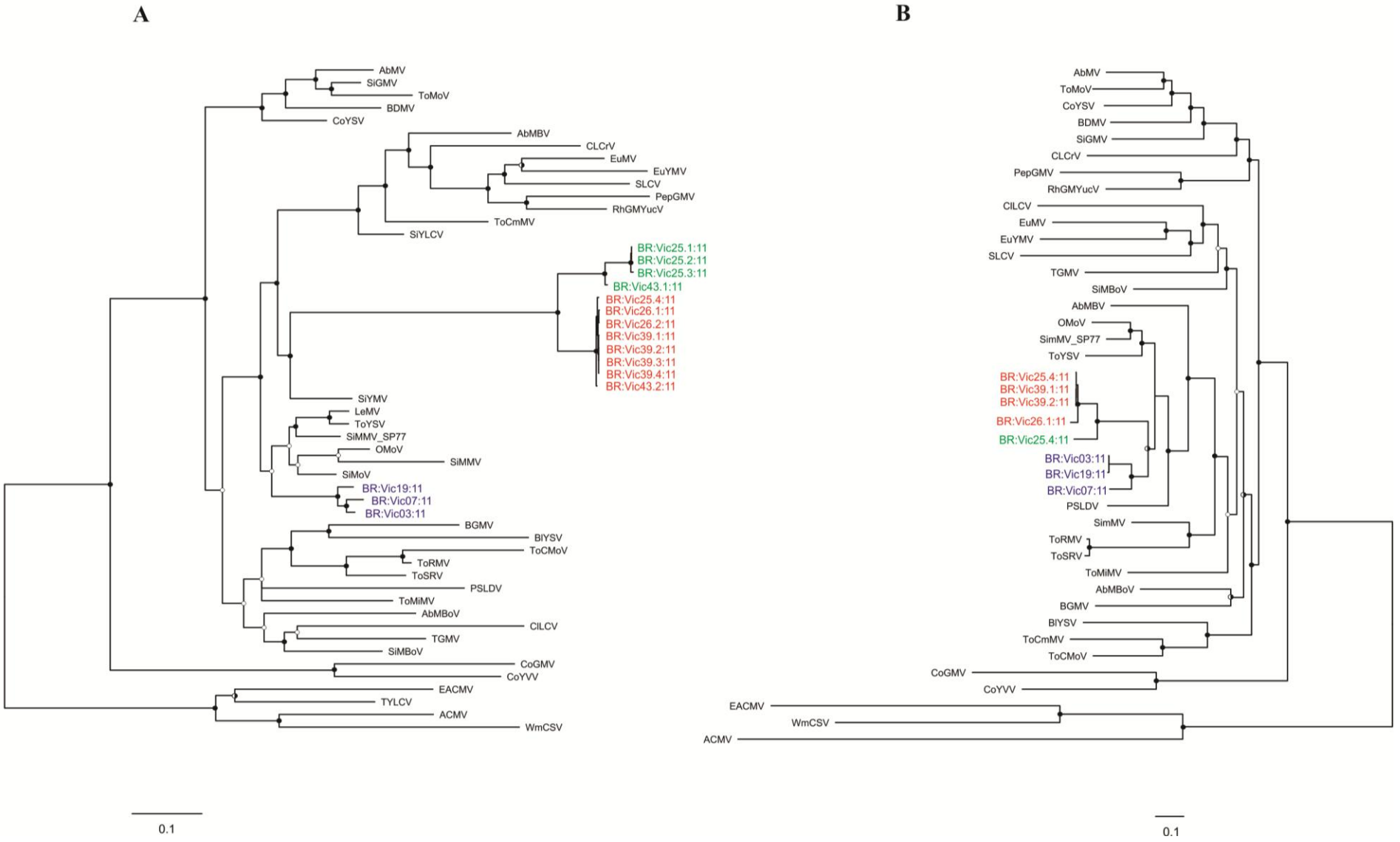

A
AbMV
SiGMV
ToMoV
BDMV
CoYSV
AbMBV
CLCrV
EuMV
EuYMV
SLCV
PepGMV
RhGMYucV
ToCmMV
SiYLCV
BR:Vic25.1:11
BR:Vic25.2:11
BR:Vic25.3:11
BR:Vic43.1:11
BR:Vic25.4:11
BR:Vic26.1:11
BR:Vic26.2:11
BR:Vic39.1:11
BR:Vic39.2:11
BR:Vic39.3:11
BR:Vic39.4:11
BR:Vic43.2:11
SiYMV
LeMV
ToYSV
SiMMV_SP77
OMoV
SiMMV
SiMoV
BR:Vic19:11
BR:Vic07:11
BR:Vic03:11
BGMV
BlYSV
ToCMoV
ToRMV
ToSRV
PSLDV
ToMiMV
AbMBoV
CiLCV
TGMV
SiMBoV
CoGMV
CoYVV
EACMV
TYLCV
ACMV
WmCSV
0.1
B
AbMV
ToMoV
CoYSV
BDMV
SiGMV
CLCrV
PepGMV
RhGMYucV
CiLCV
EuMV
EuYMV
SLCV
TGMV
SiMBoV
AbMBV
OMoV
SimMV_SP77
ToYSV
BR:Vic25.4:11
BR:Vic39.1:11
BR:Vic39.2:11
BR:Vic26.1:11
BR:Vic25.4:11
BR:Vic03:11
BR:Vic19:11
BR:Vic07:11
PSLDV
SimMV
ToRMV
ToSRV
ToMiMV
AbMBoV
BGMV
BlYSV
ToCmMV
ToCMoV
CoGMV
CoYVV
EACMV
WmCSV
ACMV
0.1

Figure 4

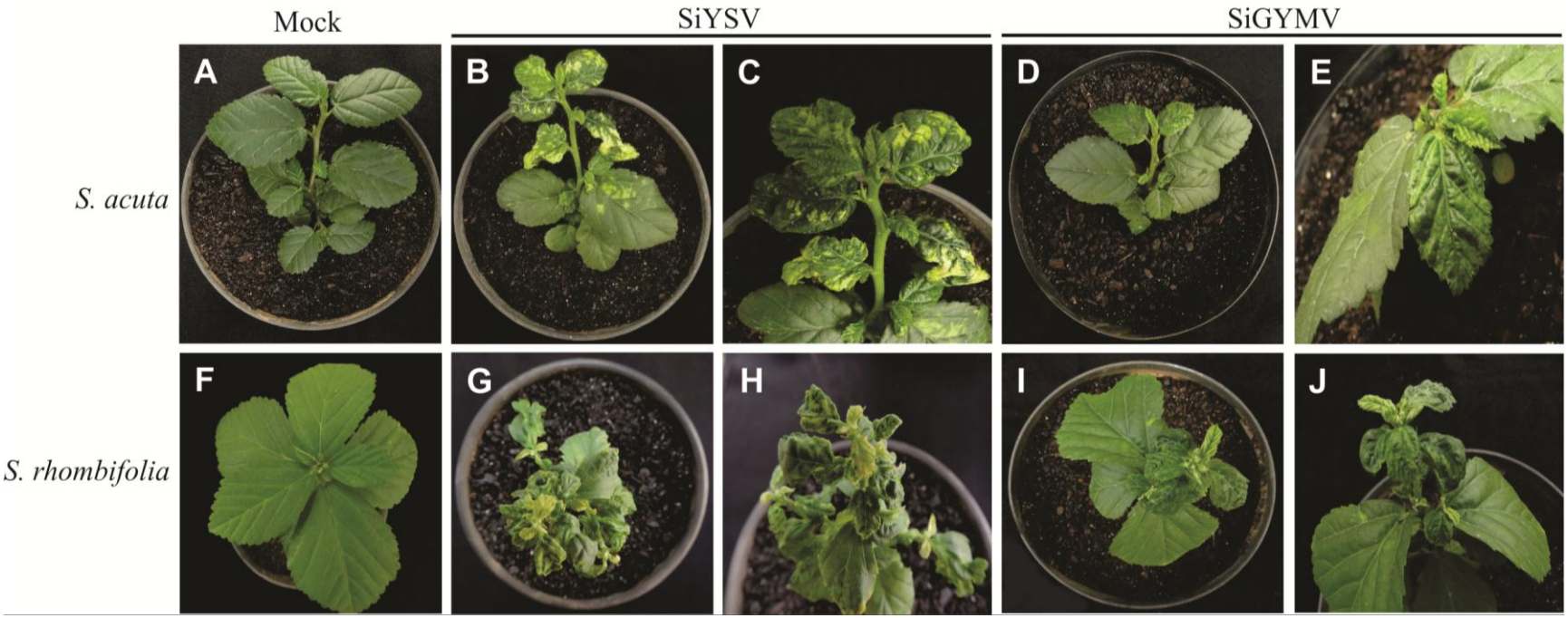


Figure 5

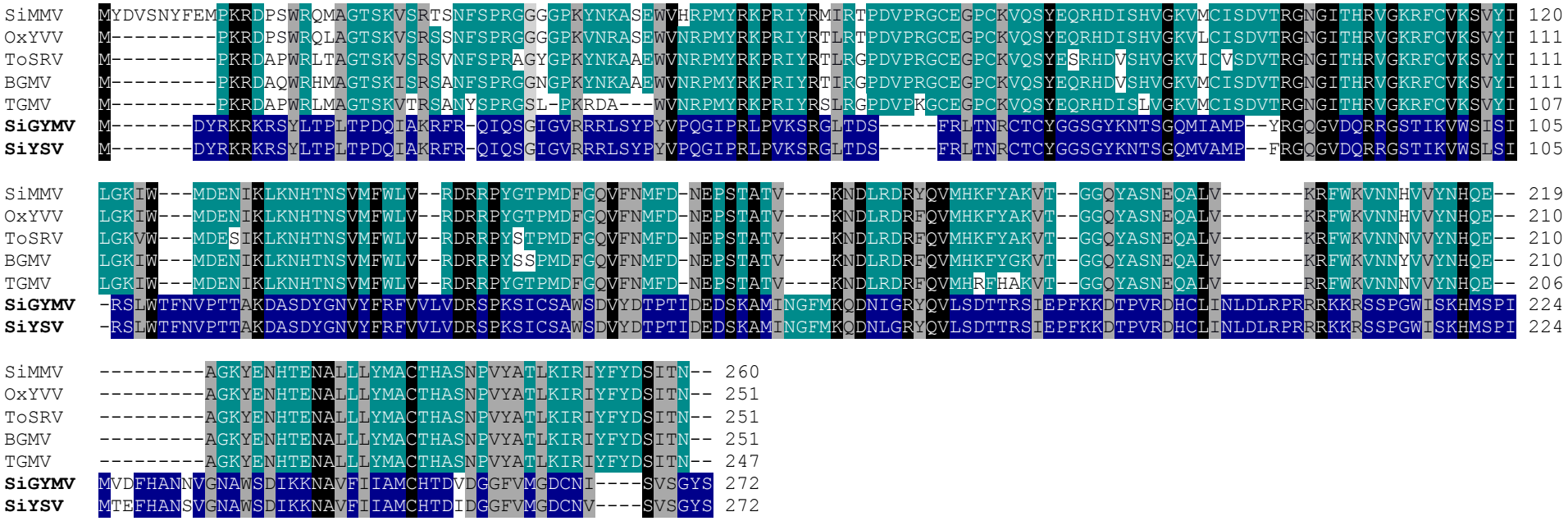
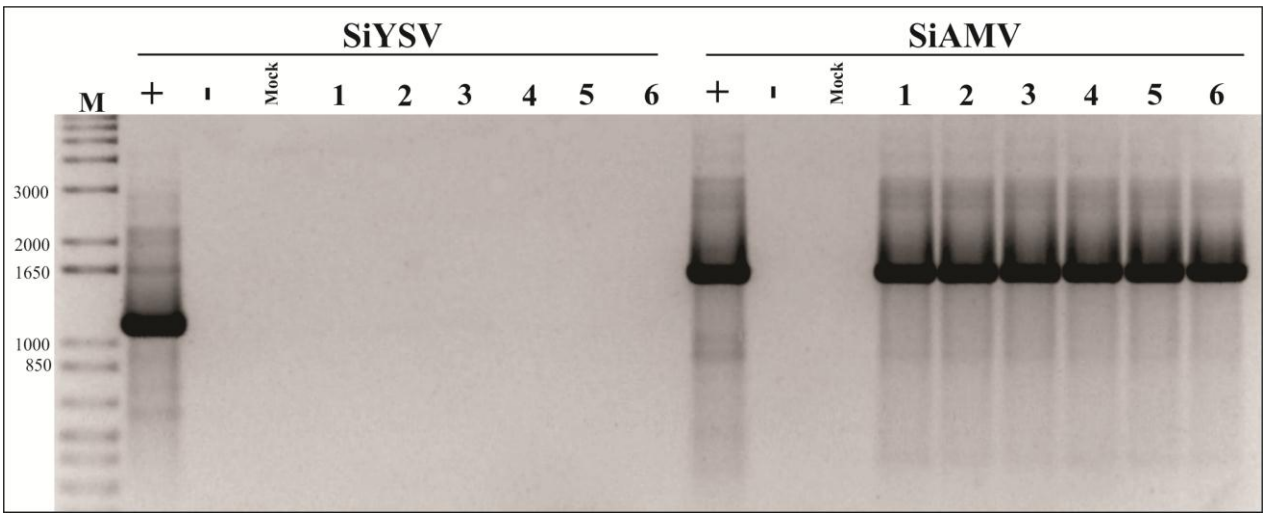
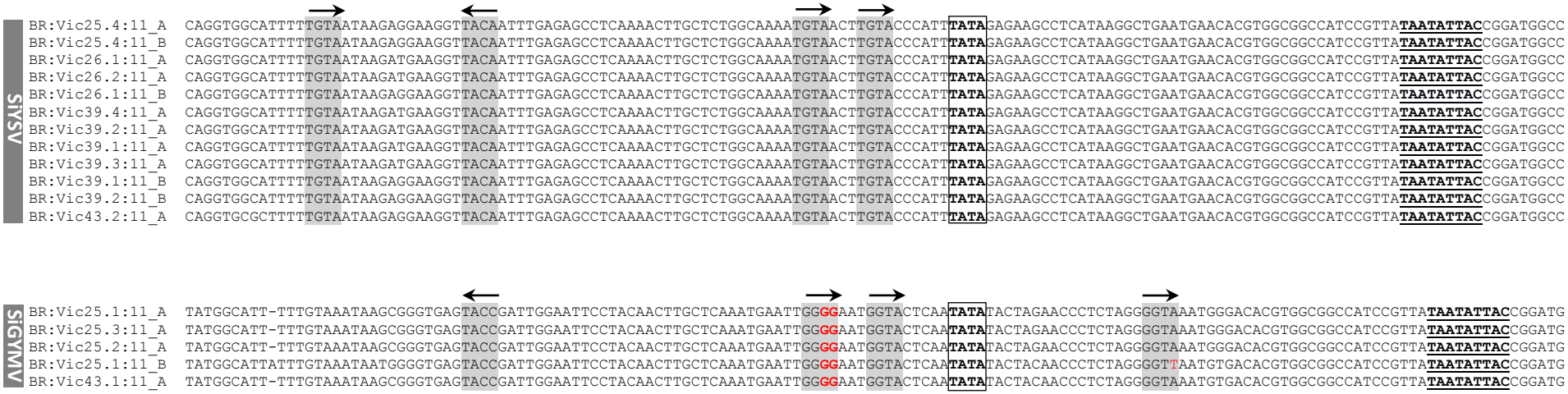


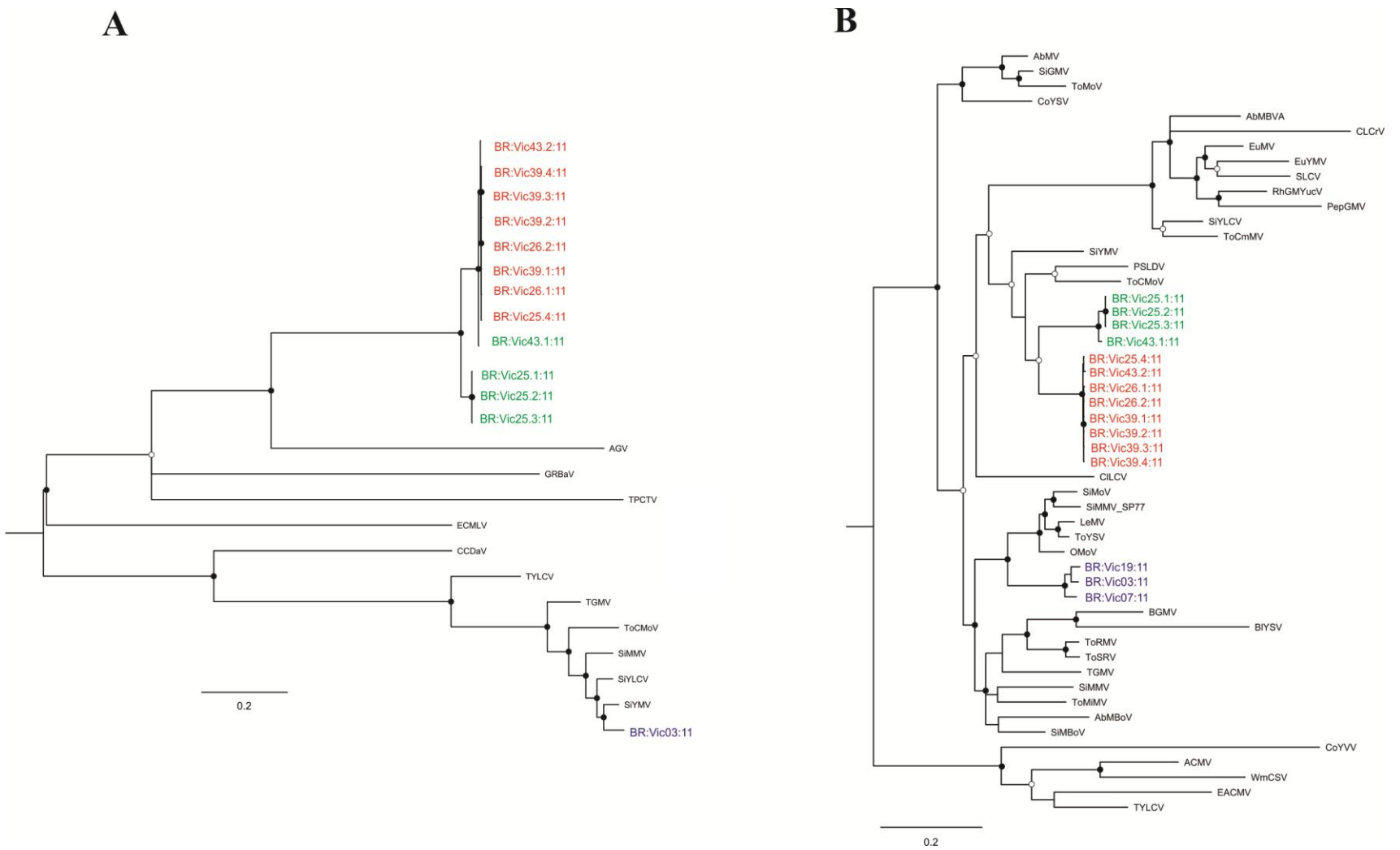
Figure 6


SiYSV
SiAMV
M
+
-
Mock
1
2
3
4
5
6
+
-
Mock
1
2
3
4
5
6
3000
2000
1650
1000
850

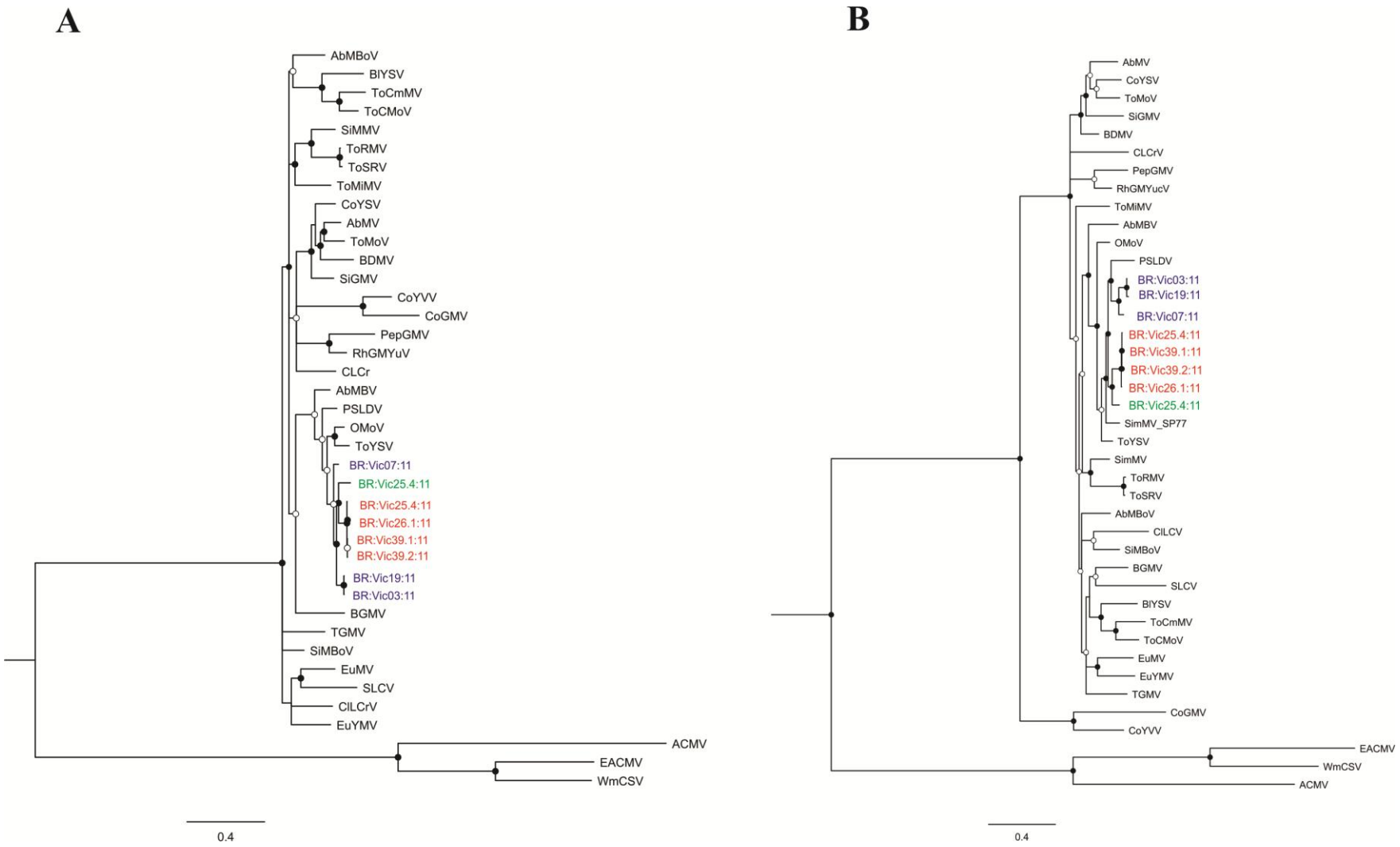
Supplementary Figure S1



Suplementary Figure S2



Supplementary Figure S3



CHAPTER 2

COMPARATIVE ANALYSYS OF BIPARTITE BEGOMOVIRUS GENETIC VARIABILITY BASED ON THE DNA-A AND DNA-B COMPONENTS

Xavier, C.A.D., Godinho, M.T., Silva, J.C., Ramos-Sobrinho, R., Sande, O.F.L., Lima, A.T.M., Zerbini, F.M. Comparative analysis of bipartite begomovirus genetic variability based on the DNA-A and DNA-B components.

COMPARATIVE ANALYSYS OF BIPARTITE BEGOMOVIRUS GENETIC VARIABILITY BASED ON THE DNA-A AND DNA-B COMPONENTS

C.A.D. Xavier[1,2], Márcio T. Godinho[1,2], José C. Silva[2], Roberto Ramos-Sobrinho[1,3],Osvaldo F.L. Sande[1,4], Alison T.M. Lima[1,5], F.M. Zerbini[1,2#]

[1]Dep. de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[2]National Research Institute for Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[3]Present address: Setor de Fitossanidade/CECA, Universidade Federal de Alagoas, Rio Largo, AL 57100-000, Brazil

[4]Present address:

[5]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

Abstract

Knowledge of how pathogens evolve and what mechanisms operate in this process is critical to adopt control strategies that are efficient and durable. The genus *Begomovirus* (family *Geminiviridae*) is comprised of viruses with a mono- or bipartite single-stranded DNA (ssDNA) genome transmitted by *B. tabaci*, and includes important plant pathogens responsible for severe losses in many economically important crops worldwide. Although there are several studies addressing evolutionary aspects shaping the structure of begomovirus populations, these studies are based on analysis of the DNA-A component. Here, we performed a comparative analysis of the evolution of DNA-A and DNA-B in New World (NW) begomoviruses based on populations of four viral species infecting cultivated and non-cultivated hosts. Our results demonstrate that the DNA-B, as well as the DNA-A, segregates based on geographical origin. In most datasets analyzed, the DNA-B was more variable than the DNA-A. The exception was *Macroptilium yellow spot virus* (MaYSV), for which the DNA-A was more variable than the DNA-B due to a recombination event at the interface between the *Rep* gene 5' region and the intergenic region. Nevertheless, the DNA-B was most prone to recombination than the DNA-A, with a higher number of events. Interestingly, we detected small ORFs in the complementary-sense strand of the large intergenic region of the DNA-B of several MaYSV isolates. These ORFs are homologous to the *Rep* gene located in the DNA-A, indicating occurrence of intercomponent recombination events. Our results indicate the two DNA components of New World begomoviruses have similar evolutionary histories. The higher degree of genetic variability of the DNA-B may reflect weaker selection pressures due to the fact the functions encoded by its proteins can, to some extent, be provided by the proteins encoded by the DNA-A.

Introduction

The family *Geminiviridae* is comprised of viral species with one or two genomic components of circular, single-strand DNA (ssDNA) encapsidated in geminate icosahedral particles. The family includes the genera *Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus*, classified according to host range, insect vector, genomic organization and phylogenetic relationships (Andrade et al., 2006; Rojas et al., 2005; Varsani et al., 2014). The genus *Begomovirus* includes mono- and bipartite viruses transmitted by *Bemisia tabaci* to dicotyledonous plants (Brown et al., 2015).

Begomoviruses in the New World (NW) are bipartite, except for *Tomato leaf deformation virus* (ToLDeV), an indigenous NW monopartite virus (Melgarejo et al., 2013). The DNA-A of NW begomoviruses encodes five genes: *CP* (capsid protein) in the virion-sense strand, *Rep* (replication-associated protein), *TrAP* (trans-activating protein), *Ren* (replication enhancer protein) and *AC4* in the complementary-sense strand. The DNA-B encodes two genes, *NSP* (nuclear shuttle protein) in the virion-sense strand and *MP* (movement protein) in the complementary-sense strand (Rojas et al., 2005). The two components share very low sequence identity, except for a small region of approximately 200 nucleotides known as the common region (CR). The CR includes the replication origin (*ori*), a conserved stem loop structure with the invariable nonanucleotide TAATATT//AC, and conserved repeat sequences (iterons) that are specifically recognized by Rep (Arguello-Astorga et al., 1994; Fontes et al., 1994; Lazarowtiz et al., 1992). The CR is important for allowing both components to be replicated by Rep, since Rep proteins show high specificity for their cognate *ori* (Lazarowtiz et al., 1992).

Strong evidence suggests that the geminiviruses, as well as other ssDNA viruses, may have evolved from extrachromosomal DNA replicons present in prokaryotes, with the subsequent acquisition of a capsid protein from RNA viruses through a recombination event

(Koonin and Ilyina, 1992; Krupovic, 2013; Wang et al., 2013). Recent evidence suggests that the exchange of genetic material between DNA and RNA viruses may occur, supporting this hypothesis (Diemer and Stedman, 2012). Several key evolutionary events marked the evolution of geminiviruses, culminating with the emergence of bipartite genomes, currently represented by viruses classified in the genus *Begomovirus* (Krupovic, 2013; Rojas et al., 2005).

It is believed that the DNA-B could have originated from a satellite molecule captured by the monopartite progenitor of all begomoviruses (Briddon et al., 2010). Possibly, this association provided greater adaptability to the monopartite progenitor, and consequently was maintained over the evolutionary process. Indeed, this is confirmed by the relationship between DNA-A and DNA-B components of bipartite begomoviruses, both necessary for the virus to infect the plant systemically and be transmitted under natural conditions (Stanley, 1983). The association of an ancestral monopartite virus with the early progenitor of the DNA-B, coupled with the loss of the *V2* gene (present only in Old World begomoviruses) may have been the key events that highlighted the emergence of bipartite begomoviruses in the NW. Recently, it was shown that the NSP of NW begomoviruses is involved in the suppression of a defense mechanism of the plant based on the overall suppression of translation (Zorzatto et al., 2015). Thus, in addition to the movement function, other functions may have contributed to the maintenance of the DNA-B component during the evolutionary process.

Interestingly, a number of reports exist of bipartite begomoviruses being capable of establishing a systemic infection, following artificial inoculation, in the absence of the DNA-B (Briddon et al., 2010; Fontenelle et al., 2007; Galvão et al., 2003; Saunders et al., 2002). These viruses may represent evolutionary intermediates between mono- and bipartite begomoviruses.

Diseases caused by begomoviruses have recently emerged as major constraints to the production of crops of importance economic, mainly in tropical and subtropical regions (Legg and Fauquet, 2004; Mansoor et al., 2003; Moriones and Navas-Castillo, 2000). 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 (Navas-Castillo et al., 2011).

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 and Holmes, 2008; Duffy and Holmes, 2009). Their propensity to recombine also contributes greatly to their genetic diversity, increasing their potential of adaptation to different hosts and environmental conditions (Lefeuvre and Moriones, 2015; Monci et al., 2002). The exchange of genomic components (pseudo-recombination or reassortment) may also occur between bipartite viruses. Although NW begomoviruses are predominantly bipartite, there are few examples of pseudo-recombination between them (Andrade et al., 2006; Faria et al., 1994; Garrido-Ramirez et al., 2000; Gilbertson et al., 1993; Hou and Gilbertson, 1996; Silva et al., 2014), unlike in the OW, where it has been demonstrated the pseudo-recombination is a frequent event (Briddon et al., 2010), playing an important role in the evolution and diversification process of begomoviruses. Most of the reported cases of pseudo-recombination involve DNA components of different strains of the same begomovirus, although a few cases have been reported of pseudo-recombination between distinct species (Andrade et al., 2006; Hou and Gilbertson, 1996). In one interesting case, pseudo-recombination was followed by a recombination event in which the DNA-A CR of one virus replaced the DNA-B CR of the other virus, enabling the DNA-B to be replicated at a higher rate and thus increasing the virulence of the pseudo-recombinant (Hou and Gilbertson, 1996).

A number of studies have been conducted to determine the main mechanisms which are responsible for shaping the genetic structure and variability of begomovirus populations (Lima et al., 2013; Rocha et al., 2013; Sobrinho et al., 2014). However, all these studies were based on analysis of the DNA-A component. Understanding how the variability is distributed within and among populations is necessary to understand how populations evolve. This knowledge is of great practical importance so that control strategies can be effective and durable (Acosta-Leal et al., 2011). Considering the important roles played by the DNA-B in the infection process, evolutionary studies which take this component into account are needed.

Here, we performed a comparative analysis of the genetic structure of the DNA-A and DNA-B components of NW begomoviruses, based on populations of four NW species infecting cultivated and non-cultivated hosts. Considering the possible distinct origins and functions of the DNA-A and DNA-B, we hypothesized that the two components evolve under different evolutionary pressures. However, our results indicate that, although the DNA-B has a higher variability and is more prone to recombination than the DNA-A, it is under similar evolutionary constraints.

Methods

Cloning of DNA-B components

Begomovirus DNA-B clones were obtained from total DNA from the samples collected for the studies of Lima et al. (2013), Rocha et al. (2013) and Sobrinho et al. (2014), which have been stored -80^{o}C. Total DNA was used as a template for rolling-circle amplification viral of genomes as described by Inoue-Nagata et al. (2004). To facilitate the cloning of DNA-B components, restriction analysis of the previously cloned DNA-As was performed using Ape 2.0 Plasmid Editor, and only restriction enzymes that do not cleave the

cognate DNA-A component were chosen. Single genome-length fragments were excised with *Apa*I, *Bam*HI, *Bgl*II, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sac*I, *Spe*I or *Xba*I, and ligated into the pBLUESCRIPT-KS+ (pKS+) plasmid vector (Stratagene), previously cleaved with the same enzyme. Viral inserts were sequenced commercially (Macrogen Inc.) by primer walking. Full-length begomovirus genomes were assembled using Geneious v. 8.1 (Kearse et al., 2012).

Multiple sequence alignments and phylogenetic analysis

All genome sequences were organized to begin at the nicking site in the invariant nonanucleotide at the origin of replication (5'-TAATATT//AC-3'). Multiple sequence alignments were prepared for the full-length DNA-A and DNA-B and for the *CP*, *Rep*, *MP* and *NSP* genes using the MUSCLE option in MEGA6 (Tamura et al., 2013). Phylogenetic trees were constructed using Bayesian inference performed with MrBayes v. 3.0b4 (Ronquist and Huelsenbeck, 2003), with the 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 and excluding the first 2,000,000 generations as burn-in. Trees were visualized using Fig Tree (tree.bio.ed.ac.uk/software/figtree/).

Genetic structure and variability indices

Inferences about population genetic structure were based on Wright's fixation index (Wright, 1951) calculated using DnaSP v.5.10 (Rozas et al., 2003). Descriptors of molecular variability [average pairwise number of nucleotide differences per site (nucleotide diversity, π) and haplotype diversity (Hd)] were estimated for each population/subpopulation using DnaSP v. 5.10. Nucleotide diversity was also calculated on a 100-nucleotide sliding window with a step size of 10 nucleotides across the *CP*, *Rep*, *MP* and *NSP* genes, the intergenic region of the DNA-A (IR-A) and the large intergenic region of the DNA-B (LIR-B). Mutation

frequencies were calculated by dividing the total number of mutations by the total number of sequenced nucleotides.

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 different methods were considered reliable.

Results

A total of 241 DNA-B sequences were obtained. In addition, 239 DNA-A and 16 DNA-B sequences were retrieved from GenBank. Detailed information on the samples and the corresponding DNA-A and DNA-B sequences are presented in the Table 1 and Suppl. Tables S1 and S2. The final data sets consisted of 117 DNA-A and 123 DNA-B sequences of *Bean golden mosaic virus* (BGMV), 28 DNA-A and 37 DNA-B sequences of *Blainvillea yellow spot virus* (BlYSV), 17 DNA-A and 20 DNA-B sequences of *Macroptilium yellow spot virus* (MaYSV) and 77 DNA-A and 77 DNA-B sequences of *Tomato severe rugose virus* (ToSRV) (Table 1).

The DNA-B is more variable than the DNA-A

The DNA-B was more variable than the DNA-A in most data sets analyzed. The exception was MaYSV, for which the DNA-A was more variable than the DNA-B (Tables 2 and 3). BlYSV and MaYSV populations showed greater genetic variability for both

components compared to populations of ToSRV and BGMV. The nucleotide diversity calculated for all BGMV isolates was similar to that observed for BlYSV and MaYSV, which can be attributed to the larger number of isolates sampled in a larger area. Indeed, when the nucleotide diversity was calculated for each BGMV subpopulation, with similar numbers of sequences and sampled areas to the BlYSV and MaYSV populations, the nucleotide diversity was much lower (Tables 2 and 3).

The nucleotide sequences of the *Rep*, *CP*, *MP* and *NSP* genes and of the IR-A, LIR-B and SIR-B were analyzed separately to verify whether the distribution of variability is evenly distributed throughout the genome. Non-coding regions were more variable compared to coding regions, except for the MaYSV Rep, which was the most variable region among all (Tables 2 and 3; Figure 1). The SIR-B and LIR-B were more variable compared to IR-A, and the SIR-B was more variable than the LIR-B. In general, the *MP* and *NSP* genes were more variable than the *CP* and *Rep* genes (Tables 2 and 3).

Nucleotide diversity was also calculated on a sliding window across the *Rep*, *CP*, *MP*, *NSP*, IR-A and LIR-B (Figure 2). For MaYSV, the region encompassing the 5' portion of the *Rep* gene and the IR-A was more variable than any other region of the genome, a pattern which is not reproduced in the other species. In fact, the greater variability of the DNA-A compared to the DNA-B in the case of MaYSV can be explained due to the high variability of this region. The CP showed an even distribution of variability for BlYSV, MaYSV and ToSRV. However, for BGMV, the CP N-terminal region was more variable compared to the C-terminal region, possibly due to a recombination event located this region shared by isolates collected at Murici and Palmeira dos Índios (Suppl. Tab. S3). The C-terminal region of MP showed a higher variability for BlYSV and BGMV. NSP in general had an even distribution of variation along the coding region, except for BlYSV, which showed a slightly more variable N-terminal region. The LIR-B has a 5' half which is much more variable than

the 3' half. This is especially true for MaYSV, due to the presence in this region of small ORFs which are homologous to the N-terminal region of the Rep protein.

The DNA-B is most prone to recombination than the DNA-A

To investigate possible recombination events, the DNA-A of BlYSV and ToSRV, and the DNA-B of BGMV, BlYSV, MaYSV and ToSRV were analyzed using RDP4. Recombination data for the DNA-A of BGMV and MaYSV was obtained from Sobrinho et al. (2014). The DNA-B was most prone to recombination than the DNA-A, with a higher number of putative recombination events (Figure 3; Suppl. Table S3). For BlYSV and MaYSV, 8 and 11 single recombination events, respectively, were detected for the DNA-B, and 1 and 3 events for the DNA-A (Figure 3; Suppl. Tab. S3). ToSRV and BGMV were less prone to recombination compared to BlYSV and MaYSV, but even then their DNA-B has a greater number of unique recombination events compared to the DNA-A (Figure 3; Suppl. Tab. S3). Interestingly, most recombination events detected in BlYSV and MaYSV involved breakpoints located within the *MP* gene and the LIR-B. In contrast, for ToSRV and BGMV most of the breakpoints are located within the *NSP* gene and the LIR-B (Figure 3). Also, a much higher number of breakpoints are located within the LIR-B (31/50) compared to the SIR-B (1/50).

Truncated Rep ORFs present in the DNA-B of MaYSV

Interestingly, we detected small ORFs in the complementary-sense strand of the LIR-B of several MaYSV isolates. These ORFs are homologous to the *Rep* gene located in the DNA-A, and thus were named truncated Rep (tRep) (Figure 4A; Suppl. Tab. S4). The tRep ORFs were found in ten isolates from three different regions, ruling out the possibility of a cloning artifact. Based on their common features, the tRep ORFs were grouped into three

types (tRep I, II and III; Suppl. Tab. S4). All tRep ORFs were located upstream of the start codon and downstream of the promoter of the *MP* gene, suggesting that they may be transcribed and translated (Figure 4B). Although their size varied from 35 to 95 amino acids, they all included the conserved Motif I (FLTYP) and the iteron related domain (Figure 4A). An analysis of the LIR-B from the other three begomoviruses failed to identify any truncated ORFs.

The DNA-B segregates based on geographical origin

Phylogenetic trees based on full-lenght nucleotide sequences of BGMV DNA-A and DNA-B are congruent (Figure 5), with strong evidence for population structuring by geography. Although three different hosts were sampled, it was not possible to verify whether there is structuring based on the host, since we did not obtains isolates from the same geographical region collected from different hosts. Both DNA-A and DNA-B trees showed four major clades: (*i*) isolates collected in Murici (Alagoas state, AL), (*ii*) Palmeiras dos Índios (AL), (*iii*) Florestal (Minas Gerais state, MG) and (*iv*) Unaí (MG), Cristalina (Goiás state, GO), Santo Antônio de Goiás (GO) and Paranoá (DF). One isolate from Murici (BR: Mur11:11) clustered with isolates from Palmeira dos Índios, suggesting that migration can occur between these subpopulations. Population subdivision test performed based on Wright's fixation index indicated a strong genetic differentiation between subpopulations sampled from these locations (Table 4), in accordance with results observed in the phylogenetic trees.

The ToSRV DNA-A and DNA-B phylogenetic trees are not congruent (Figure 6). In the DNA-A tree, isolates collected from Florestal formed two clades, with isolates collected in 2008 segregated from isolates collected in 2014. Isolates collected in Coimbra (MG) and Carandaí (MG) formed a third clade (Figure 6A). The DNA-B tree showed four well-defined clades: (*i*) Florestal 2008, (*ii*) Florestal 2014, (*iii*) Carandaí and (*iv*) Coimbra, with two

isolates from Carandaí placed in the Coimbra clade (Figure 6B). Population subdivision test performed for ToSRV isolates from Carandaí and Coimbra indicated higher genetic differentiation for the DNA-B ($Fst_{Car/Coi,DNA-B} = 0.30625$) compared to the DNA-A ($Fst_{Car/Coi,DNA-A} = 0.21654$), corroborating the phylogeny results. Isolate BR:Flo01:14 grouped together with isolates from Coimbra (Figure 6B) as a result of recombination event between isolates BR:Flo18:14 and BR:Coi183:13 (Figure 3), from Florestal and Coimbra, respectively.

The MaYSV DNA-A and DNA-B trees were highly incongruent (Figure 7). The MaYSV DNA-A tree shows three well supported clades, while the DNA-B tree shows several poorly supported clades, possibly reflecting the complex recombination pattern (Figure 3). There is no evidence of structuring either by geography or by host. To minimize the effect of recombination, phylogenetic trees based on the CP, Rep, MP and NSP deduced amino acid sequences were constructed, however all trees were incongruent (*data not shown*). Population subdivision test was not performed due to the small number of isolates from each region and host.

BGMV in AL: emergence of a new species ?

The two BGMV subpopulations from AL are separated from the subpopulations present in MG, GO and DF by very long branches in both the DNA-A and DNA-B trees (Figure 5), indicating a high level of divergence. Pairwise sequence comparisons between all BGMV DNA-A sequences shows a wide variation in identity among these isolates (Suppl. Fig. S1A). Isolates collected in AL (Murici and Palmeira dos Índios) have identities ranging from 96-97%, while isolates collected in DF, MG and GO have identities ranging from 97-99%. However, isolates from AL have identities ranging from 90.7-91.8% with isolates from DF, MG and GO, which is very close to the species demarcation threshold for begomoviruses (Brown et al., 2015). Pairwise comparisons between DNA-B sequences show a similar pattern

(Suppl. Fig. S1B). In addition, analysis of the iterons and the IRD, located in the intergenic region (IR) and in the Rep protein, respectively, indicates that the subpopulations from AL have different iterons and IRDs from the subpopulation from MG, GO and DF (Suppl. Table S5), suggesting that these can be independently replicating lineages.

Discussion

In accordance to previous studies (Briddon et al., 2010; Rocha et al., 2013), our results indicate that the DNA-B is generally more variable than the DNA-A. A number of hypotheses have been presented to explain this greater variability. First, it has been suggested that 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 means that a much higher proportion of nucleotide substitutions will lead to amino acid changes. Coupled to the fact that the proteins encoded by the DNA-A are involved in several *cis* and *trans* interactions (Rep-Rep, Rep-iteron, Rep-Ren, Trap-promoter) which could be negatively affected by changes in their amino acid sequences, this component would be under much stricter evolutionary constraints (Briddon et al., 2010). A second explanation is the distinct origin of the DNA-B, possibly from a satellite molecule captured by a monopartite begomovirus, as suggested by Briddon et al. (2010). Third, the DNA B has two intergenic (non-coding) regions comprising approximately 35% of its length, in contrast to the DNA-A in which the intergenic region encompasses only 13% of its length. On this regard, our results show that the SIR-B is more variable, followed by the LIR-B and the IR-A. The SIR-B does not have *cis*-elements involved in replication and transcription, thus being more permissive to variation. On the contrary, both the LIR-B and the IR-A contain importants *cis*-elements involved in the replication and acting as promoter regions of genes, therefore being under greater selective pressure for maintenance of these elements. The LIR-B is approximately

threefold longer than the IR-A, thus supporting a greater accumulation of variation in absolute terms. Finally, it must be pointed out that these three aspects are not mutually exclusive, *ie*, the most likely explanation for the higher variability of the DNA-B is the combination of these three factors.

Contrary to this general observation, the MaYSV DNA-A is more variable than its DNA-B. It has been shown that the DNA-A of this virus has a complex pattern of recombination, with breakpoints located at the interface between the 5' region of the *Rep* gene and the IR-A (Lima et al., 2013; Sobrinho et al., 2014). Indeed, this region was the most variable of the genome, which certainly contributes to the greater variability of the DNA-A. Although the DNA-B component showed a greater number of recombination events, most events are intraspecific, providing little contribution to variability. 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. Our results demonstrate that although in most cases the DNA-B is more variable than the DNA-A, this is not an absolute rule.

Recombination has been extensively studied in begomoviruses, being considered one of the main mechanisms that contribute to their diversification (Lefeuvre et al., 2007; Lima et al., 2013; Monci et al., 2002; Rocha et al., 2013; Sobrinho et al., 2014). Our results clearly demonstrate a greater propensity of the DNA-B to recombine compared to the DNA-A. It was demonstrated that tolerance to recombination of a given region of the genome is correlated with the degree of similarity between the regions exchanged and the complexity of inter- and intra- genomic interactions (protein-protein/protein-DNA) established by the exchanged region (Martin et al., 2005). The greatest propensity of the DNA-B to recombine may be explained by its lowest organizational complexity (only two non-overlaping genes, compared to four overlapping genes in the DNA-A) and the lowest number of inter- and intragenomic

interactions of its encoded proteins. Therefore, the DNA-B would be more permissive to the exchange of genomic regions without disrupting inter- and intra-molecular interactions.

Gregorio-Jorge et al. (2010) identified small fragments of *Rep* gene-derived sequence, ranging from 35 to 51 nt in length, located in the viral-sense strand of the DNA-B intergenic region of *Euphorbia mosaic virus* (from Mexico), *Euphorbia yellow mosaic virus* (form Brazil) and *Tomato mild yellow leaf curl Aragua virus* (from Venezuela). The authors suggested that these short Rep sequences can be involved in the posttranscriptional regulation of the cognate *Rep* gene. We detected truncated Rep (tRep) ORFs in the complementary-sense strand of the LIR-B of several MaYSV isolates, ranging from 35 to 95 amino acids (considerably longer than the short regions identified by (Gregorio-Jorge et al., 2010). The presence of these tRep ORFs is a strong indication of inter-component recombination. The tRep ORFs were located downstream to the *MP* gene promoter, suggesting that they can be expressed. Moreover, they include the coding region for Motif I and the IRD, which are involved in the process of recognition and binding of Rep to the origin of replication (Arguello-Astorga et al., 1994). If the tRep ORFs are actually expressed, one possibility is that the truncated proteins can compete with Rep for binding at the origin of replication, negatively interfering in the replication process. Another possibility is that it may interfere with the translation of MP, due to their location upstream of the MP initiation codon. These possibilities need to be tested experimentally.

DNA-A-based studies have shown that begomovirus populations segregate based on geographical location (Briddon et al., 2010; Rocha et al., 2013). Our results demonstrate this to be the case for the DNA-B as well. For BGMV populations, this segregation was perfectly symmetrical for both components, yielding 'mirror image' phylogenetic trees. This geographical segregation, although not an absolute rule (for example, the ToSRV DNA-B, but not the DNA-A, segregated based on geography), can be explained by the requisite

relationship between DNA-A and DNA-B components of bipartite begomoviruses, both being necessary for the virus to establish a systemic infection and be transmitted under natural conditions (Stanley, 1983). Thus, it is expected that both components will co-evolve as a single entity. In contrast, it has been shown that betassatellites segregate based on host range (Briddon and Stanley, 2006). In general, the helper begomovirus does not depend on the betasatellite to infect the plant, thus explaining the non co-evolutionary relationship between helper begomovirus and betasatellite (Briddon and Stanley, 2006; Zhou, 2013).

Subdivision observed within begomovirus populations can be attributed primarily to geographical isolation, preventing gene flow between populations and contributing to the speciation process. Indeed, we have evidence that the BGMV population present in AL is at the limit of becoming a new species. It is thus possible that these isolates may have distinct biological characteristics compared to a typical BGMV population. In accordance with this hypothesis, Sobrinho et al. (2014) showed a temporal change in the prevalence of begomoviruses infecting legume hosts in northeastern Brazil, with the emergence of MaYSV, predominantly infecting lima bean and bean in samples collected in 2011. BGMV was isolated only in lima bean, suggesting that this emerging variant may be less adapted compared to MaYSV. This hypothesis (as well as the relative fitness of the variant BGMV compared to the typical BGMV) needs be tested.

Here, we present a study comparing the genetic structure of populations of begomovirus in NW. Our results showed that the DNA-A and DNA-B have similar evolutionary histories, albeit under different selective pressures, with the DNA-B being more permissive to variation.

References

Acosta-Leal, R., Duffy, S., Xiong, Z., Hammond, R.W., Elena, S.F., 2011. Advances in plant virus evolution: Translating evolutionary insights into better disease management. Phytopathology 101, 1136-1148.

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.

Arguello-Astorga, G.R., Guevara-González, R.G., Herrera--Estrella, L.R., Rivera-Bustamante, R.F., 1994. Geminivirus replication origins have a group-specific organization of interative elements: a model for replication. Virology 203, 90-100.

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.

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

Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-Olive, 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.

Diemer, G.S., Stedman, K.M., 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. Biology Direct 7, 13.

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.

Faria, J.C., Gilbertson, R.L., Hanson, S.F., Morales, F.J., Ahlquist, P.G., Loniello, A.O., Maxwell, D.P., 1994. Bean golden mosaic geminivirus type II isolates from the Dominican Republic and Guatemala: Nucleotide sequences, infectious pseudorecombinants, and phylogenetic relationships. Phytopathology 84, 321-329.

Fontenelle, M.R., Luz, D.F., Gomes, A.P., Florentino, L.H., Zerbini, F.M., Fontes, E.P., 2007. Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus *Tomato chlorotic mottle virus*. Virus Res 126, 262-267.

Fontes, E.P.B., Eagle, P.A., Sipe, P.S., Luckow, V.A., Hanley-Bowdoin, L., 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. J Biol Chem 269, 8459-8465.

Galvão, R.M., Mariano, A.C., Luz, D.F., Alfenas, P.F., Andrade, E.C., Zerbini, F.M., Almeida, M.R., Fontes, E.P.B., 2003. A naturally occurring recombinant DNA-A of a typical bipartite begomovirus does not require the cognate DNA-B to infect *Nicotiana benthamiana* systemically. J Gen Virol 84, 715-726.

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., 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.

Gregorio-Jorge, J., Bernal-Alcocer, A., Banuelos-Hernandez, B., Alpuche-Solis, A.G., Hernandez-Zepeda, C., Moreno-Valenzuela, O., Frias-Trevino, G., Arguello-Astorga, G.R., 2010. Analysis of a new strain of *Euphorbia mosaic virus* with distinct replication specificity unveils a lineage of begomoviruses with short Rep sequences in the DNA-B intergenic region. Virol J 7, 275.

Hou, Y.M., Gilbertson, R.L., 1996. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. J Virol 70, 5430-5436.

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. J Virol Met 116, 209-211.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647-1649.

Koonin, E.V., Ilyina, T.V., 1992. Geminivirus replication proteins are related to prokariotic plasmid rolling circle DNA replication initiator proteins. J Gen Virol 73, 2763-2766.

Krupovic, M., 2013. Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. Current Opinion in Virology 3, 578-586.

Lazarowtiz, S.G., Wu, L.C., Rogers, S.G., Elmer, J.S., 1992. Sequence-specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. Plant Cell 4, 799-809.

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.

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.

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

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

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. J Virol 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.

Moriones, E., Navas-Castillo, J., 2000. *Tomato yellow leaf curl virus*, an emerging virus complex causing epidemics worldwide. Virus Res 71, 123-134.

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

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

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 Phytopath 43, 361-394.

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

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.

Saunders, K., Salim, N., Mali, V.R., Malathi, V.G., Briddon, R., Markham, P.G., Stanley, J., 2002. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. Virology 293, 63-74.

Silva, F.N., Lima, A.T.M., Rocha, C.S., Castillo-Urquiza, G.P., Alves, M., Zerbini, F.M., 2014. Recombination and pseudorecombination driving the evolution of the begomoviruses *Tomato severe rugose virus* (ToSRV) and *Tomato rugose mosaic virus* (ToRMV): two recombinant DNA-A components sharing the same DNA-B. Virol J 11, 66.

Sobrinho, R.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. J Gen Virol 95, 2540-2552.

Stanley, J., 1983. Infectivity of the cloned geminivirus genome requires sequences from both DNAs. Nature 305, 643-645.

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.

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.

Wang, W.-C., Hsu, Y.-H., Lin, N.-S., Wu, C.-Y., Lai, Y.-C., Hu, C.-C., 2013. A novel prokaryotic promoter identified in the genome of some monopartite begomoviruses. PloS one 8, e70037.

Wright, S., 1951. The genetical structure of populations. Annals of Eugenics 15, 323-354.

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

Zorzatto, C., Machado, J.P., Lopes, K.V., Nascimento, K.J., Pereira, W.A., Brustolini, O.J., Reis, P.A., Calil, I.P., Deguchi, M., Sachetto-Martins, G., Gouveia, B.C., Loriato, V.A., Silva, M.A., Silva, F.F., Santos, A.A., Chory, J., Fontes, E.P., 2015. NIK1-mediated translation suppression functions as a plant antiviral immunity mechanism. Nature 520, 679-682.

Table 1. DNA-A and DNA-B sequences analyzed in this study.

Species	State	Host	No of sequences			Total number of sequences by species	
			DNA-A (retrieved from GenBank)	DNA-B			
				This study	GenBank	DNA-A	DNA-B
BGMV[1]	Distrito Federal (DF)	*Phaseolus vulgaris*	30	30	-		
	Góias (GO)	*Phaseolus vulgaris*	29	27	-		
	Minas Gerais(MG)	*Phaseolus vulgaris*	16	19	-	117	123
		Macroptilium lathyroides	13	13	-		
	Alagoas (AL)	*Phaseolus lunatus*	29	34	-		
ToSRV	Minas Gerais(MG)	*Solanum lycopersicum*	77	68	9	77	77
BlYSV	Minas Gerais(MG)	*Blainvillea rhomboidea*	28	30	7	28	37
MaYSV	Alagoas (AL)	*Phaseolus vulgaris*	3	4	-		
		Phaseolus lunatus	9	12	-	17	20
		Macroptilium lathyroides	5	4	-		
Total			239	241	16		

[1]BGMV, *Bean golden mosaic virus*; ToSRV, *Tomato severe rugose virus*, BlYSV, *Blainvillea yellow spot virus*; MaYSV, *Macroptilium yellow spot virus*.

Table 2. Genetic variability indices based on the DNA-A of the begomoviruses *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV), *Macroptilium yellow spot virus* (MaYSV) and *Tomato severe rugose virus* (ToSRV).

Population	No. of sequences	H[1]	Hd	DNA-A π	Rep π	CP π	IR-A π
BGMV (Total)	117	110	0.998 (± 0.002)	0.03948 (± 0.00296)	0.03051 (± 0.00246)	0.03758 (± 0.00267)	0.09406 (± 0.00786)
DF/GO/Unaí (MG)	75	75	1.000 (± 0.002)	0.00665 (± 0.00019)	0.00436 (± 0.00024)	0.00582 (± 0.00024)	0.01149 (± 0.00092)
Florestal (MG)	13	10	0.923 (± 0.069)	0.00182 (± 0.00047)	0.00085 (± 0.00021)	0.00204 (± 0.00057)	0.00233 (± 0.00070)
Murici (AL)	12	10	0.955 (± 0.057)	0.00340 (± 0.00190)	0.00283 (± 0.00151)	0.00355 (± 0.00149)	0.00696 (± 0.00518)
Palmeira dos Índios (AL)	17	16	0.993 (± 0.002)	0.00173 (± 0.00024)	0.00129 (± 0.00028)	0.00150 (± 0.00037)	0.00391 (± 0.00068)
ToSRV(Total)[2]	77	65	0.993 (± 0.004)	0.01076 (± 0.00084)	0.00963 (± 0.00095)	0.01084 (± 0.00110)	0.01995 (± 0.00158)
Carandaí (MG)	16	15	0.992 (± 0.025)	0.00339 (± 0.00038)	0.00198 (± 0.00039)	0.00180 (± 0.00045)	0.01183 (± 0.00134)
Coimbra (MG)	37	29	0.977 (± 0.025)	0.00753 (± 0.00051)	0.00680 (± 0.00056)	0.00525 (± 0.00044)	0.01740 (± 0.00155)
Florestal 2008 (MG)	5	5	1.000 (± 0.126)	0.00802 (± 0.00156)	0.00812 (± 0.00176)	0.00767 (± 0.00150)	0.01176 (± 0.00242)
Florestal 2014 (MG)	16	13	0.967 (± 0.036)	0.00315 (± 0.00041)	0.00278 (± 0.00040)	0.00306 (± 0.00033)	0.00540 (± 0.00130)
BlYSV (MG)	28	24	0.989 (± 0.004)	0.03687 (± 0.00411)	0.03386 (± 0.00372)	0.03353 (± 0.00350)	0.05872 (± 0.00774)
MaYSV (AL)	17	17	1.000 (± 0.020)	0.06380 (± 0.00627)	0.10741 (± 0.01300)	0.02100 (± 0.00144)	0.06721 (± 0.00567)

[1] H, haplotype number; Hd, haplotype diversity; π, average pairwise number of nucleotide differences per site (nucleotide diversity)
[2] Includes two isolates from São Paulo (SP) state and one isolate from GO.

Table 3. Genetic variability indices based on the DNA-B of the begomoviruses *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV), *Macroptilium yellow spot virus* (MaYSV) and *Tomato severe rugose virus* (ToSRV).

Population	No. of sequences	H[1]	Hd	DNA-B π	NSP π	MP π	LIR π	SIR π
BGMV (Total)	123	104	0.997 (± 0.002)	0.05459 (± 0.00312)	0.03942 (± 0.00239)	0.04062 (± 0.00231)	0.08146 (± 0.00471)	0.11623 (± 0.00730)
DF/GO/Unaí (MG)	76	63	0.994 (± 0.003)	0.01358 (± 0.00269)	0.00947 (± 0.00193)	0.00857 (± 0.00187)	0.02171 (± 0.00461)	0.02908 (± 0.00358)
Florestal (MG)	13	11	0.962 (± 0.003)	0.00245 (± 0.00110)	0.00020 (± 0.00017)	0.00105 (± 0.00020)	0.00267 (± 0.00055)	0.00200 (± 0.00180)
Murici (AL)	15	14	0.990 (± 0.028)	0.01659 (± 0.00749)	0.01266 (± 0.00613)	0.01772 (± 0.00765)	0.01489 (± 0.00635)	0.05544 (± 0.02926)
Palmeira dos Índios (AL)	19	17	0.982 (± 0.026)	0.00618 (± 0.00099)	0.00405 (± 0.00079)	0.00480 (± 0.00076)	0.00970 (± 0.00210)	0.00881 (± 0.00258)
ToSRV (Total)[2]	77	75	0.999 (± 0.002)	0.02183 (± 0.00169)	0.0107 0 (± 0.00180)	0.01130 (± 0.00120)	0.03895 (± 0.00221)	0.05255 (± 0.00344)
Carandaí (MG)	18	18	1.000 (± 0.019)	0.00501 (± 0.00056)	0.00171 (± 0.00044)	0.00271 (± 0.00059)	0.00994 (± 0.00121)	0.01135 (± 0.00213)
Coimbra (MG)	28	28	1.000 (± 0.010)	0.00970 (± 0.00146)	0.00540 (± 0.00114)	0.00616 (± 0.00086)	0.01761 (± 0.00278)	0.08415 (± 0.00934)
Florestal 2008 (MG)	5	5	1.000 (± 0.126)	0.01296 (± 0.00317)	0.01608 (± 0.00420)	0.00658 (± 0.00225)	0.01656 (± 0.00386)	0.01679 (± 0.00504)
Florestal 2014 (MG)	23	20	0.988 (± 0.016)	0.00853 (± 0.00188)	0.00306 (± 0.00071)	0.00513 (± 0.00090)	0.01677 (± 0.00389)	0.17464 (± 0.06626)
BlYSV (MG)	37	35	0.997 (±0.007)	0.06922 (± 0.00213)	0.07366 (± 0.00497)	0.04809 (± 0.00122)	0.07782 (± 0.00223)	0.13013 (± 0.00933)
MaYSV (AL)	20	18	0.989 (± 0.019)	0.04857 (± 0.00193)	0.03058 (± 0.00127)	0.02843 (± 0.00139)	0.08182 (± 0.00562)	0.12080 (± 0.00822)

[1] H, haplotype number; Hd, haplotype diversity; π, average pairwise number of nucleotide differences per site (nucleotide diversity).
[2] Includes two isolates from São Paulo (SP) state and one isolate from GO.

Table 4. Subdivision test performed for populations of *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV) and *Tomato severe rugose virus* (ToSRV).

Population	F_{st}[1]	
	DNA-A	**DNA-B**
BGMV[2]		
Florestal/Murici	0.96988	0.90504
Florestal/Palmeira dos Índios	0.97986	0.95853
Florestal/Unaí	0.85721	0.78163
Florestal/Cristalina	0.85763	0.85652
Florestal/Santo Antônio de Góias	0.85935	0.92402
Florestal/Paranoá	0.85650	0.90125
Palmeira dos Índios/Unaí	0.95338	0.84364
Palmeira dos Índios/Cristalina	0.95543	0.90548
Palmeira dos Índios/Santo Antônio de Góias	0.95532	0.94298
Palmeira dos Índios/Paranoá	0.95539	0.93062
Unaí/Cristalina	0.20225	**0.00194**
Unaí/Santo Antônio de Góias	**-0.0176**	**0.04472**
Unaí/Paranoá	**0.11506**	**0.02346**
Cristalina/Santo Antônio de Góias	0.21261	**0.06170**
Cristalina/Paranoá	**0.05476**	**0.01667**
Santo Antônio de Góias/Paranoá	**0.07350**	**0.05131**
Murici/Palmeira dos Índios	0.69456	0.68059
Murici/Unaí	0.94331	0.78179
Murici/Cristalina	0.94530	0.84673
Murici/Santo Antônio de Góias	0.94521	0.88492
Murici/Paranoá	0.94518	0.87251
ToSRV[3]		
Carandaí/Coimbra	0.21654	0.30625
Carandaí/Florestal 2008	0.74857	0.72810
Carandaí/Florestal 2014	0.68358	0.66796
Coimbra/Florestal 2008	0.68951	0.77453
Coimbra/Florestal 2014	0.58432	0.66796
Florestal 2008/Florestal 2014	0.77385	0.78318
BlYSV[3]		
Coimbra/Viçosa	**0.04571**	**0.14919**

[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 among them and therefore are considered to comprise a single population.

[2]Florestal, Unaí: Minas Gerais (MG) state; Murici, Palmeira dos Índios: Alagoas (AL) state; Cristalina, Santo Antônio de Goiás: Góias (GO) state; Paranoá: Distrito Federal (DF).

[3]All in MG state.

Figure legends

Figure 1. Mutation frequencies for complete nucleotide sequences (DNA-A and DNA-B), deduced amino acid sequences (CP, capsid protein; Rep, replication-associated protein; MP, movement protein; NSP, nuclear shuttle protein), and nucleotide sequences of the DNA-A intergenic region (IR-A), DNA-B large intergenic region (LIR-B) and DNA-B small intergenic region (SIR-B), for the populations of *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV), *Macroptilium yellow spot virus* (MaYSV) and *Tomato severe rugose virus* (ToSRV). (A) Mutation frequencies determined including all isolates for the same species. (B) Mutation frequencies determinated for subpopulations, according to the results of the subdivision test (Table 2).

Figure 2. Average pairwise number of nucleotide differences per site (nucleotide diversity, π) calculated on a 100-nucleotide sliding window with a step size of 10 nucleotides across the capsid protein (*CP*), replication-associated protein (*Rep*), movement protein (*MP*) and nuclear shuttle protein (*NSP*) genes, the DNA-A intergenic region (IR-A) and the DNA-B large intergenic region (LIR-B), for populations of *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV), *Macroptilium yellow spot virus* (MaYSV) and *Tomato severe rugose virus* (ToSRV).

Figure 3. Recombination events detected in the DNA-B of *Bean golden mosaic virus* (BGMV), *Tomato severe rugose virus* (ToSRV), *Macroptilium yellow spot virus* (MaYSV) and *Blainvillea yellow spot virus* (BlYSV) populations. Recombination breakpoint coordinates are according to the first nucleotide after the cleavage site at the origin of replication, increasing clockwise. The genome at the top of the figure corresponds to the

schematic representations of sequences below. Regions highlighted in black correspond to the donated minor parent portion, while the remaining portion corresponds to the major parent sequence. Breakpoints that could not be accurately located are indicated by (?), and the regions are filled with a gradient. 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. The reported *p*-value is from the method in bold and is the lowest *p*-value calculated for the featured event. LIR, large intergenic region; SIR, small intergenic region.

Figure 4. Truncated Rep (*t*Rep) ORFs present in the DNA-B of *Macroptilium yellow spot virus* (MaYSV) isolates. (A) Schematic representation of the DNA-B, with genes indicated by arrows (*MP* in the viral sense, *NSP* in the complementary sense). *t*Rep regions (types I, II and III) filled in black represent homologous regions to the Rep protein in the cognate DNA-A (Suppl. Tab. S3). Amino acid alignments represent the Rep N-terminal region of the cognate DNA-A and *t*Rep (identical residues in black, similar residues in gray). Numbers to the right of the alignment represent *t*Rep length and the length of the homologous region of the Rep protein in the cognate DNA-A. (B) Nucleotide sequence alignment of the DNA-B large intergenic region (LIR) containing *cis*-acting elements. The nonanucleotide at the origin of replication is highlighted in bold and underlined. Putative Rep-binding elements (iterons) are shaded in gray and their orientation is indicated by black arrows. Putative *cis*-acting elements found in eukaryotic promoters located in the LIR are marked by dotted boxes located at the right side of the *t*Rep start codon (nucleotides in bold and marked with asterisks).

Figure 5. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (A) and DNA-B (B) of *Bean golden mosaic virus* (BGMV) populations. Nodes with posterior

probability values between 0.80 and 0.89 are indicated by filled circles, nodes with values between 0.90 and less than 0.94 are indicated by half-filled circles, and those with values equal to or greater than 0.95 are indicated by empty circles. Colors indicate sampling locations (MG, Minas Gerais; GO, Góias; AL, Alagoas; DF, Federal District).

Figure 6. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (A) and DNA-B (B) of *Tomato severe rugose virus* (ToSRV) populations. Nodes with posterior probability values between 0.80 and 0.89 are indicated by filled circles, nodes with values between 0.90 and less than 0.94 are indicated by half-filled circles, and those with values equal to or greater than 0.95 are indicated by empty circles. Colors indicate sampling locations (MG, Minas Gerais; GO, Goiás; SP, São Paulo).

Figure 7. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (A) and DNA-B (B) of *Macroptilium yellow spot virus* (MaYSV) populations. Nodes with posterior probability values between 0.80 and 0.89 are indicated by filled circles, nodes with values between 0.90 and less than 0.94 are indicated by half-filled circles, and those with values equal to or greater than 0.95 are indicated by empty circles. Colors indicate sampling locations (AL, Alagoas).

Figure 8. Midpoint-rooted Bayesian phylogenetic trees based on the complete DNA-A (A) and DNA-B (B) of *Blainvillea yellow spot virus* (BlYSV) populations. Nodes with posterior probability values between 0.80 and 0.89 are indicated by filled circles, nodes with values between 0.90 and less than 0.94 are indicated by half-filled circles, and those with values equal to or greater than 0.95 are indicated by empty circles. Colors indicate sampling locations (MG, Minas Gerais).

Supplementary Figure S1. Pairwise sequence identity matrices for the DNA (A) and DNA-B (B) of BGMV isolates. Neighbor-joining phylogenetic trees based on the complete DNA-A and DNA-B nucleotide sequences are showed in the lefth side of the corresponding matrices. Branch colors represent isolates collected in Murici, AL (red), Palmeira dos Índios, AL (blue), GO/DF/Unaí, MG (green) and Florestal, MG (orange).

Figure 1

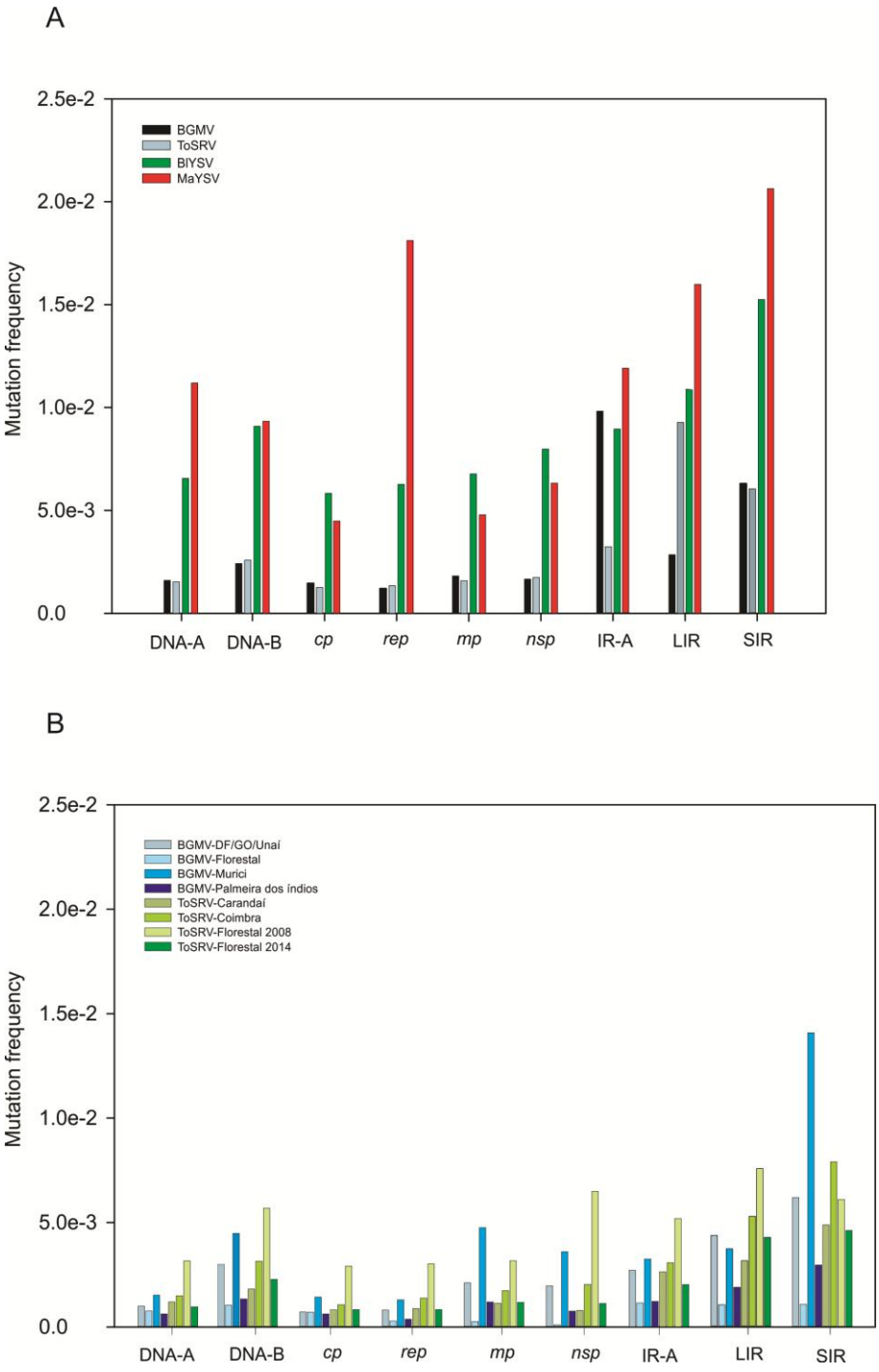

A
2.5e-2
2.0e-2
1.5e-2
1.0e-2
5.0e-3
0.0
Mutation frequency
BGMV
ToSRV
BlYSV
MaYSV
DNA-A
DNA-B
cp
rep
mp
nsp
IR-A
LIR
SIR
B
2.5e-2
2.0e-2
1.5e-2
1.0e-2
5.0e-3
0.0
Mutation frequency
BGMV-DF/GO/Unaí
BGMV-Florestal
BGMV-Murici
BGMV-Palmeira dos índios
ToSRV-Carandaí
ToSRV-Coimbra
ToSRV-Florestal 2008
ToSRV-Florestal 2014
DNA-A
DNA-B
cp
rep
mp
nsp
IR-A
LIR
SIR

Figure 2

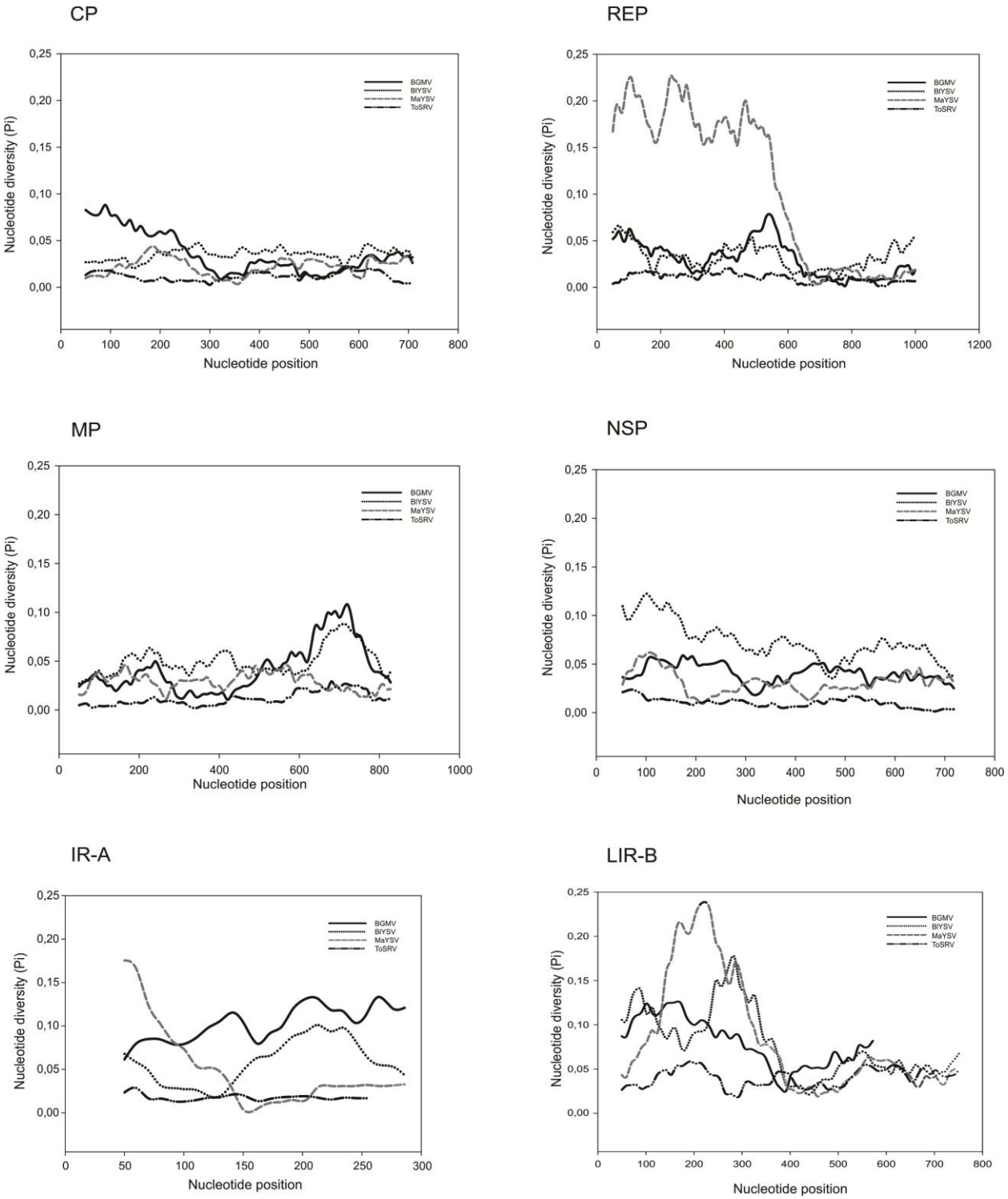
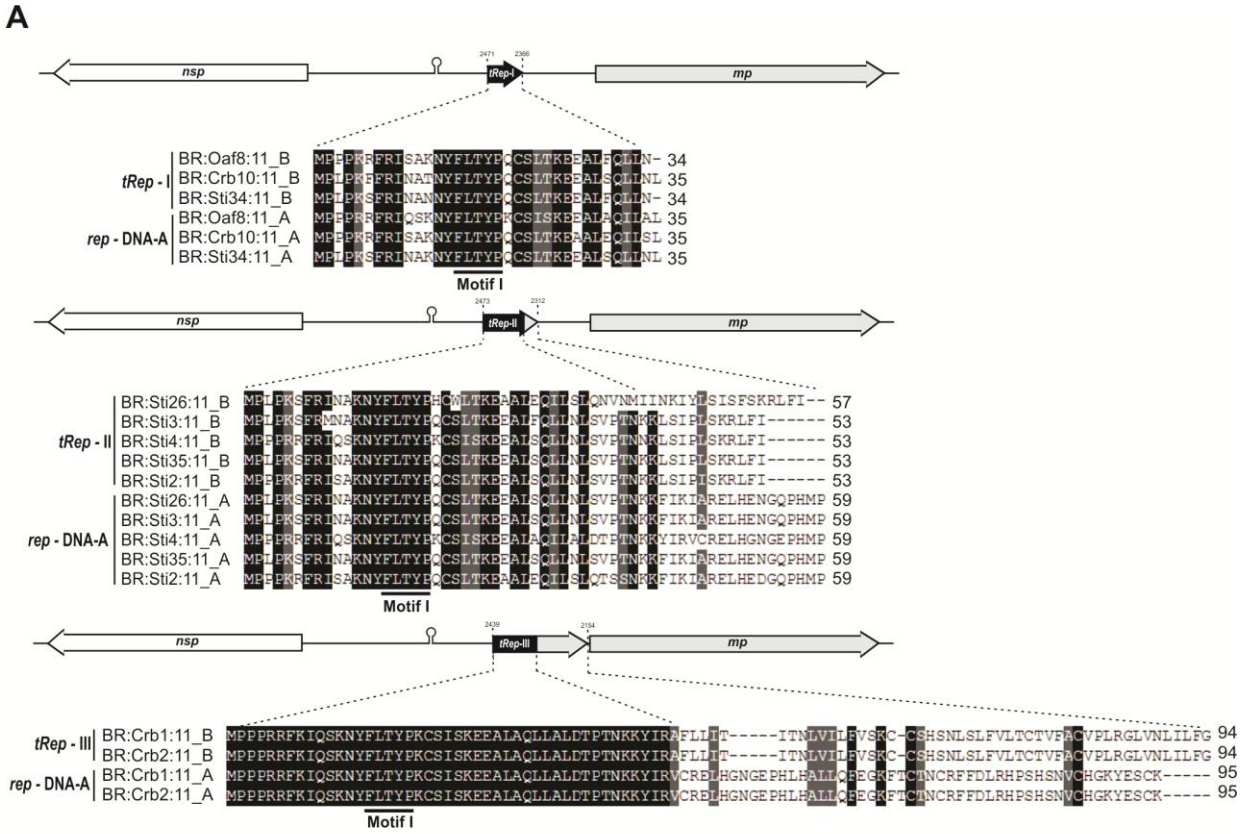

CP
REP
MP
NSP
IR-A
LIR-B
Nucleotide diversity (Pi)
Nucleotide position
BGMV
BlYSV
MaYSV
ToSRV

Figure 3

Event	Recombinant	Recombination breakpoints		Parents		Method	*P-value*
		Begin	End	Minor	Major		
BGMV							
1	BR:Cri15:12	53	180	Unknown	BR:Cri14:12	GMCS**3**	5.147×10^{-30}
2	BR:Mur12.2:11, BR:Mur12.1:11	544	2422	BR:Mur3:11	Unknown	RGMCS**3**	1.390×10^{-18}
3	BR:Cri13.2:12	1013	2435	BR:Una10.1:12	BR:Cri4:12	MC**S**3	2.727×10^{-11}
4	BR:Una16:12	988	2368	BR:Una10.1:12	BR:Una2:12	MCS**3**	1.629×10^{-03}
ToSRV							
1	BR:Flo01:14	1276	229	BR:Flo18:14	BR:Coi183:13	GM**C**3	2.514×10^{-14}
2	Sumare	2535	339	Baccatum-9	BR:Flo:14	GMCS3	3.764×10^{-05}
BIYSV							
1	BR:Vic07:10, BR:Vic04.1:10, BR:Coi145:14, BR:Coi150:14, BR:Coi150:14, BR:Coi154:14, BR:Coi155:14, BR:Vic09:10, BR:Coi164:14, BR:Coi165:14, BR:Coi32.1:13, BR:Coi33:13, BR:Coi25:07, BR:Vic08:10, BR:Vic18:10, BR:Vic21:10, BR:Vic26c:10, BR:Coi36:13 BR:Coi37.1:13, BR:Vic26s:10	2572	1864	BR:Vic13:10	Unknown	RGBMC**S**3	7.272×10^{-23}
2	BR:Coi142.1:14, BR:Coi151:14, BR:Coi152.1:14	(?) 2571	464	BR:Coi158s:14	BR:Vic07:10	RGMC**S**3	9.284×10^{-08}
3	BR:Coi148:14, BR:Coi158b:14, BR:Coi158s:14	1908	2485	BR:Vic11:10	BR:Coi152.2:14	**R**GMCS3	9.111×10^{-10}
4	BR:Coi157:14	1364	2512	BR:Coi160.2:14	BR:Coi35:13	RGM**C**S3	4.015×10^{-23}
5	BR:Vic13:10, BR:Vic11:10	1742	2509	BR:Coi34:13	BR:Coi142.2:14	RMC**S**3	8.291×10^{-29}
6	BR:Coi142.2:14, BR:Coi35:13, BR:Coi147:14	(?) 1800	2334	BR:Vic09:10	BR:Coi34:13	GMC**3**	7.176×10^{-05}
7	BR:Coi154:14 BR:Coi155:14	(?) 1863	2389	BR:Vic07:9 BR:Vic07:10	BR:Vic18:10 Unknown	RG**S**3	5.084×10^{-08}
8	BR:Coi158s:14, BR:Coi158b:14	141	468	Unknown	BR:Coi34:13	RCS**3**	4.420×10^{-05}
MaYSV							
1	BR:Crb10:11	2237	2601	BR:Crb2:11	BR:Sti3:11	RGMCS**3**	8.580×10^{-25}
2	BR:Sti34:11 BR:Sti4:11	1452	2507	BR:Oaf8:11 BR:Oaf8:11	BR:Oaf27:11 BR:Oaf28:11	RGCS**3**	4.989×10^{-13}
3	BR:Sti29:11, BR:Sti4:11	2600	2427	Unknown	BR:Oaf24:11	RGBMC**S**3	1.718×10^{-23}
4	BR:Sti10:11, BR:Sti26:11	1876	2439	BR:Sti2:11	BR:Oaf24:11	RMCS**3**	3.440×10^{-12}
5	BR:Sti26:11, BR:Sti9.1:11, BR:Sti9.3:11, BR:Sti9.2:11	(?) 2467	219	BR:Oaf27:11	BR:Sti2:11	RGMC**S**3	3.309×10^{-08}
6	BR:Sti35:11, BR:Sti3:11, BR:Oaf28:11	1198	2450	BR:Oaf24:11	BR:Sti2:11	RGBMC**S**3	3.064×10^{-13}
7	BR:Oaf8:11	162	2197	BR:Sti2:11	BR:Crb2:11	RGMCS**3**	1.865×10^{-13}
8	BR:Oaf24:11, BR:Oaf25.1:11, BR:Oaf25.2:11	(?) 83	673	BR:Sti2:11	BR:Sti4:11	RG**M**3	1.913×10^{-07}
9	BR:Oaf27:11	406	(?) 1451	BR:Sti10:11	BR:Sti34:11	RGCS**3**	5.571×10^{-10}
10	BR:Oaf27:11	2444	82	BR:Oaf24:11	BR:Crb2:11	RGMC**S**3	1.930×10^{-08}
11	BR:Sti9.1:11, BR:Sti9.2:11	(?) 216	531	BR:Sti2:11	BR:Sti9.3:11	RGMC**3**	7.065×10^{-06}

Figure 4


A
nsp
tRep-I
mp
tRep - I
rep - DNA-A
BR:Oaf8:11_B
BR:Crb10:11_B
BR:Sti34:11_B
BR:Oaf8:11_A
BR:Crb10:11_A
BR:Sti34:11_A
Motif I
tRep-II
tRep - II
BR:Sti26:11_B
BR:Sti3:11_B
BR:Sti4:11_B
BR:Sti35:11_B
BR:Sti2:11_B
BR:Sti26:11_A
BR:Sti3:11_A
BR:Sti4:11_A
BR:Sti35:11_A
BR:Sti2:11_A
tRep-III
tRep - III
BR:Crb1:11_B
BR:Crb2:11_B
BR:Crb1:11_A
BR:Crb2:11_A

B

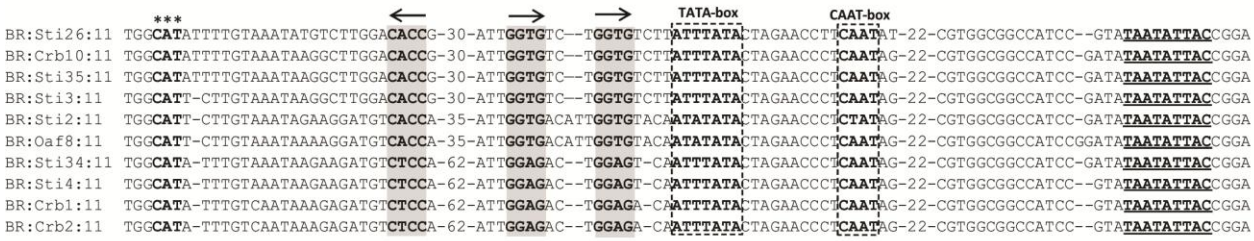

TATA-box
CAAT-box
BR:Sti26:11
BR:Crb10:11
BR:Sti35:11
BR:Sti3:11
BR:Sti2:11
BR:Oaf8:11
BR:Sti34:11
BR:Sti4:11
BR:Crb1:11
BR:Crb2:11

Figure 5

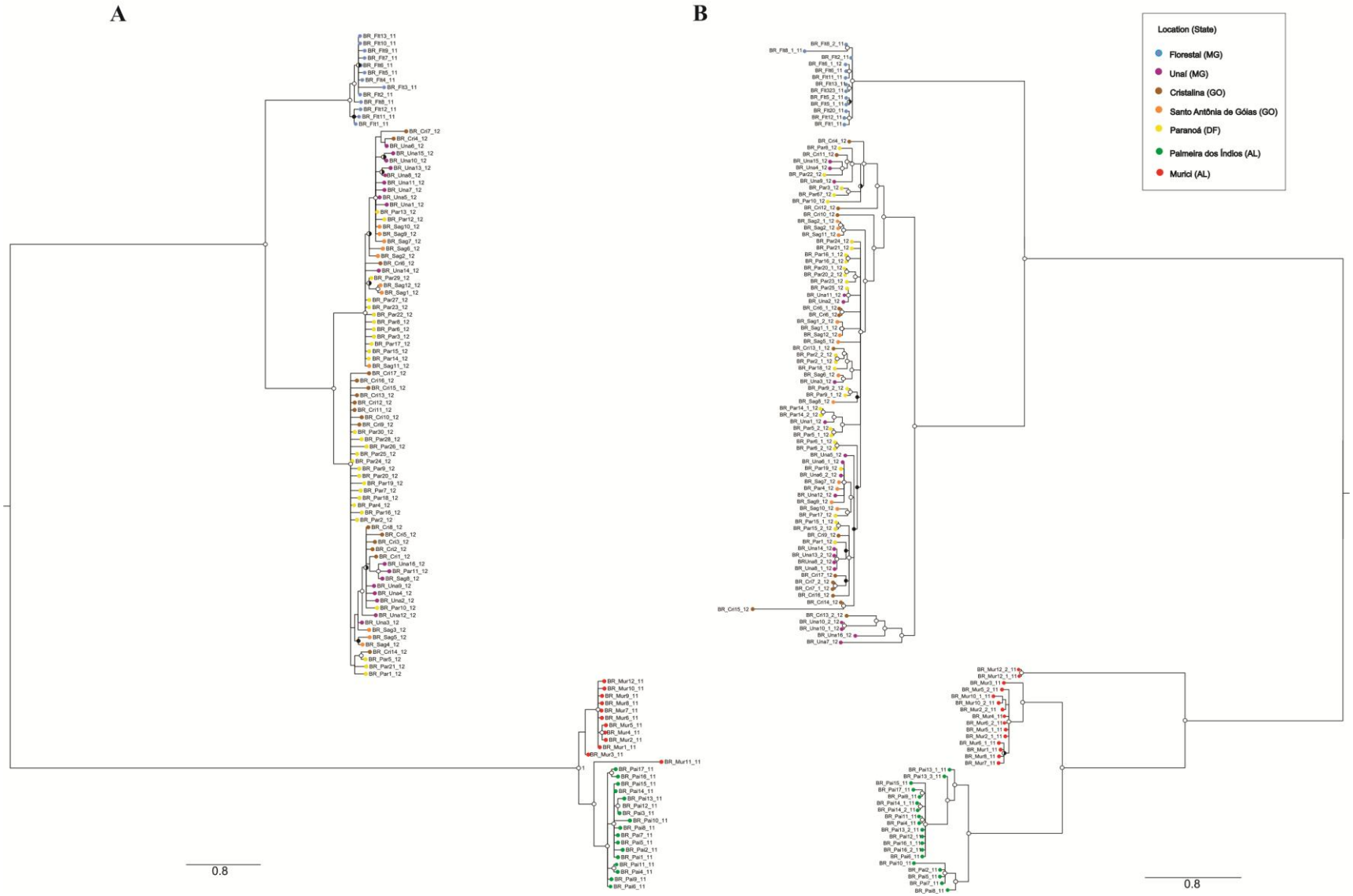

A
B
Location (State)
Florestal (MG)
Unaí (MG)
Cristalina (GO)
Santo Antônia de Góias (GO)
Paranoá (DF)
Palmeira dos Índios (AL)
Murici (AL)
0.8
0.8

Figure 6

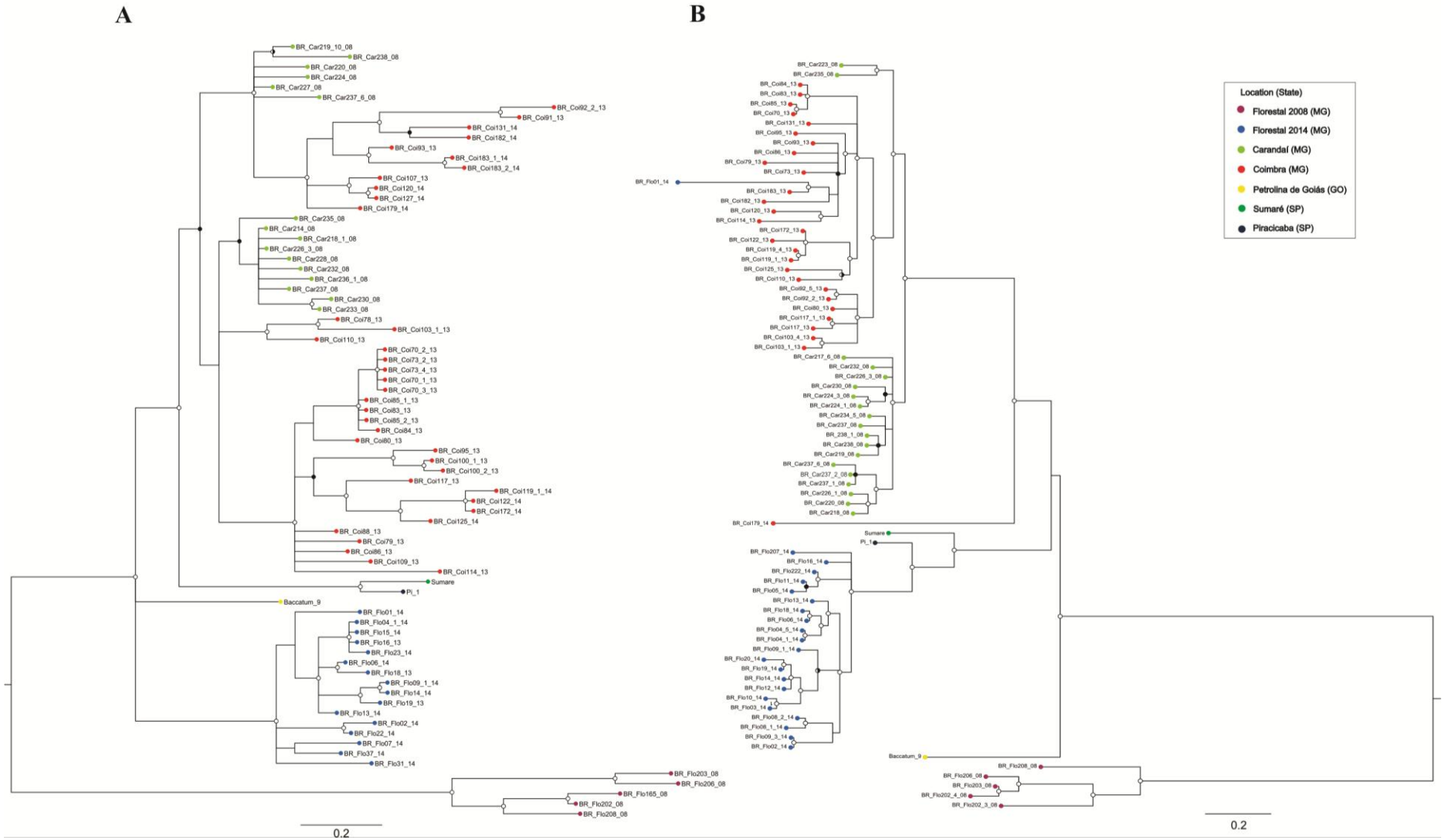

A
B
Location (State)
Florestal 2008 (MG)
Florestal 2014 (MG)
Carandaí (MG)
Coimbra (MG)
Petrolina de Goiás (GO)
Sumaré (SP)
Piracicaba (SP)
0.2
0.2

Figure 7

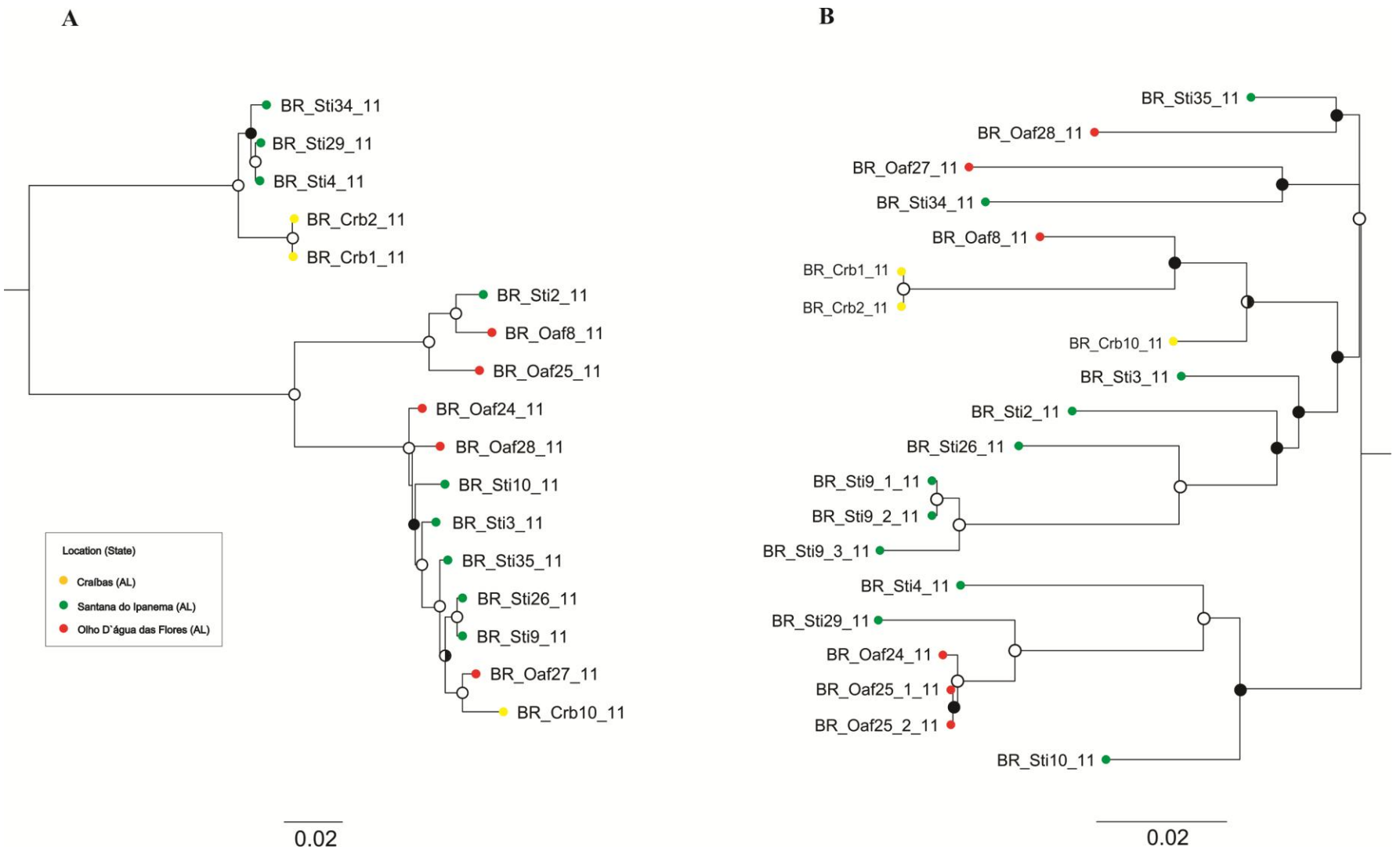
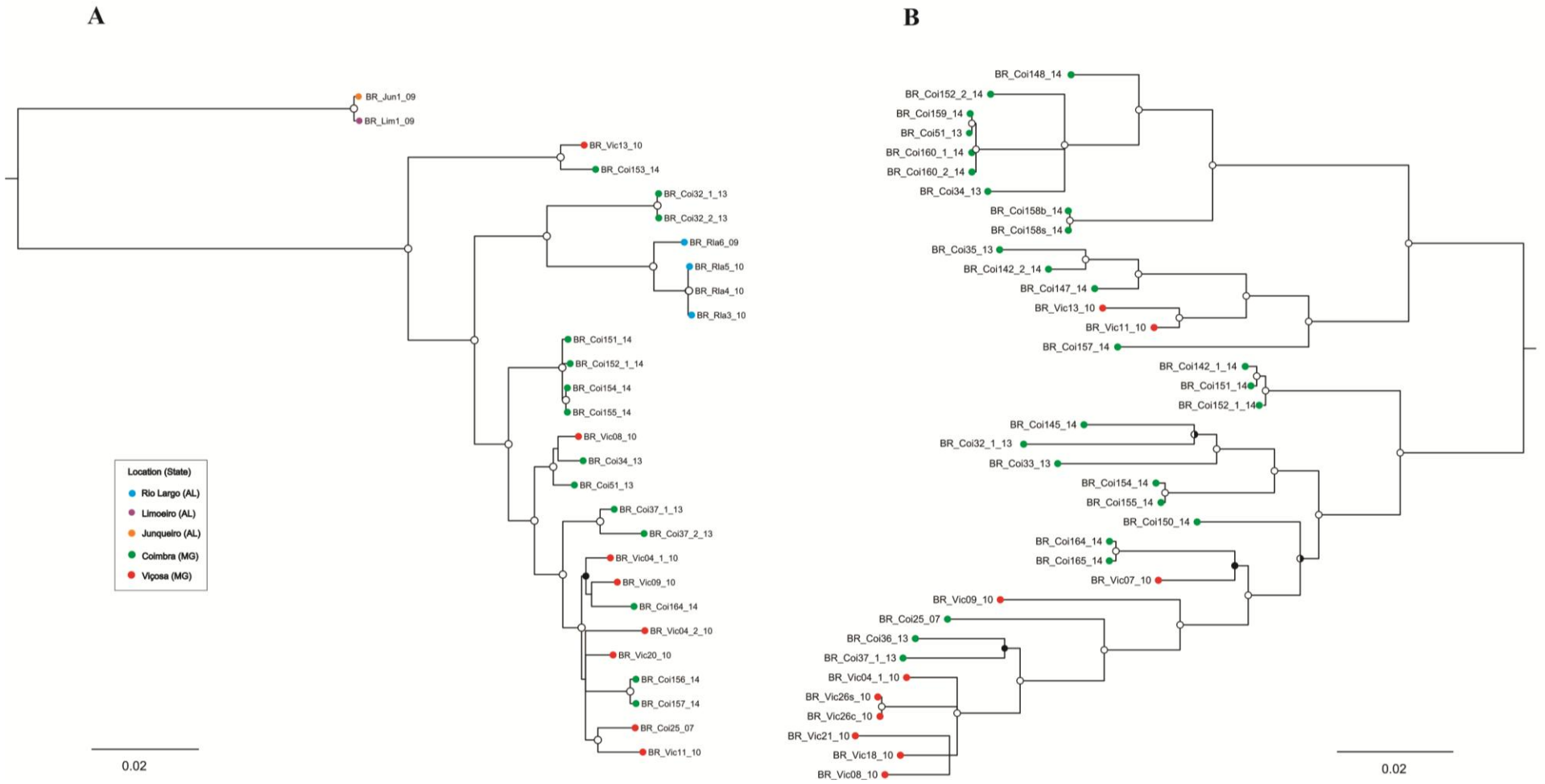
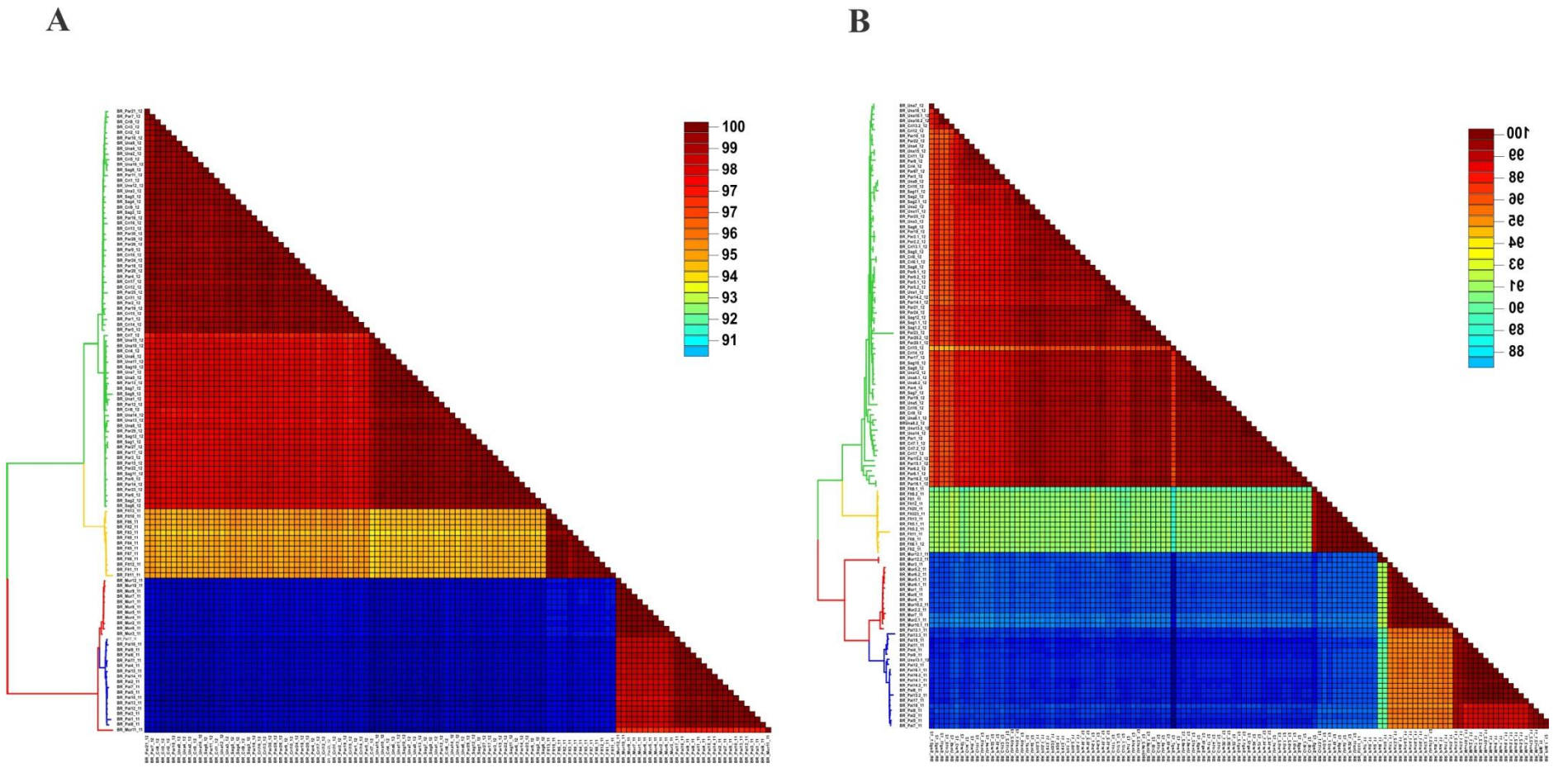

A
BR_Sti34_11
BR_Sti29_11
BR_Sti4_11
BR_Crb2_11
BR_Crb1_11
BR_Sti2_11
BR_Oaf8_11
BR_Oaf25_11
BR_Oaf24_11
BR_Oaf28_11
BR_Sti10_11
BR_Sti3_11
BR_Sti35_11
BR_Sti26_11
BR_Sti9_11
BR_Oaf27_11
BR_Crb10_11
Location (State)
Craíbas (AL)
Santana do Ipanema (AL)
Olho D'água das Flores (AL)
0.02
B
BR_Sti35_11
BR_Oaf28_11
BR_Oaf27_11
BR_Sti34_11
BR_Oaf8_11
BR_Crb1_11
BR_Crb2_11
BR_Crb10_11
BR_Sti3_11
BR_Sti2_11
BR_Sti26_11
BR_Sti9_1_11
BR_Sti9_2_11
BR_Sti9_3_11
BR_Sti4_11
BR_Sti29_11
BR_Oaf24_11
BR_Oaf25_1_11
BR_Oaf25_2_11
BR_Sti10_11
0.02

Figure 8


A
BR_Jun1_09
BR_Lim1_09
BR_Vic13_10
BR_Coi153_14
BR_Coi32_1_13
BR_Coi32_2_13
BR_Rla6_09
BR_Rla5_10
BR_Rla4_10
BR_Rla3_10
BR_Coi151_14
BR_Coi152_1_14
BR_Coi154_14
BR_Coi155_14
BR_Vic08_10
BR_Coi34_13
BR_Coi51_13
BR_Coi37_1_13
BR_Coi37_2_13
BR_Vic04_1_10
BR_Vic09_10
BR_Coi164_14
BR_Vic04_2_10
BR_Vic20_10
BR_Coi156_14
BR_Coi157_14
BR_Coi25_07
BR_Vic11_10
Location (State)
Rio Largo (AL)
Limoeiro (AL)
Junqueiro (AL)
Coimbra (MG)
Viçosa (MG)
0.02
B
BR_Coi148_14
BR_Coi152_2_14
BR_Coi159_14
BR_Coi51_13
BR_Coi160_1_14
BR_Coi160_2_14
BR_Coi34_13
BR_Coi158b_14
BR_Coi158s_14
BR_Coi35_13
BR_Coi142_2_14
BR_Coi147_14
BR_Vic13_10
BR_Vic11_10
BR_Coi157_14
BR_Coi142_1_14
BR_Coi151_14
BR_Coi152_1_14
BR_Coi145_14
BR_Coi32_1_13
BR_Coi33_13
BR_Coi154_14
BR_Coi155_14
BR_Coi150_14
BR_Coi164_14
BR_Coi165_14
BR_Vic07_10
BR_Vic09_10
BR_Coi25_07
BR_Coi36_13
BR_Coi37_1_13
BR_Vic04_1_10
BR_Vic26s_10
BR_Vic26c_10
BR_Vic21_10
BR_Vic18_10
BR_Vic08_10
0.02

Supplementary Figure S1



Supplementary Table S1. Begomovirus sequences reported in this study.

Sample code	Date of colection	Location	Geographical coordinates		Host	Enzyme	Isolate name	GenBank access number
Blainvillea yellow spot virus (BlYSV)								
HV4	May, 2010	Viçosa, MG	-	-	*Blainvillea rhomboidea*	SacI	BR:Vic04.1:10	
HV9	May, 2010	Viçosa, MG	-	-	*Blainvillea rhomboidea*	SacI	BR:Vic09:10	
HV13	May, 2010	Viçosa, MG	-	-	*Blainvillea rhomboidea*	SacI	BR:Vic13:10	
OS32	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	ClaI	BR:Coi32.1:13	
OS33	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	SacI	BR:Coi33:13	
OS34	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	SacI	BR:Coi34:13	
OS35	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	SacI	BR:Coi35:13	
OS36	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	SacI	BR:Coi36:13	
OS37	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	XbaI	BR:Coi37.1:13	
OS51	July, 2013	Coimbra, MG	S20°51'29,9''	W042°51'42,2"	*Blainvillea rhomboidea*	SacI	BR:Coi51:13	
OS142	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi142.1:14	
						SacI	BR:Coi142.2:14	
OS143	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi143:14	
OS145	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi145:14	
OS147	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi147:14	
OS148	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi148:14	
OS150	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi150:14	
OS151	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi151:14	
OS152	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi152.1:14	
						SacI	BR:Coi152.2:14	
OS154	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi154:14	
OS155	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi155:14	
OS157	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi157:14	
OS158	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi158s:14	
						BamHI	BR:Coi158b:14	
OS159	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi159:14	
OS160	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi160.1:14	
						SacI	BR:Coi160.2:14	
OS164	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi164:14	
OS165	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Blainvillea rhomboidea*	SacI	BR:Coi165:14	

Suppl. Table S1 (cont.)

Macroptilium yellow spot virus (MaYSV)							
RC18	July, 2011	Craíbas, AL	S09°40'37.9"	W036°46'37.8"	*Phaseolus vulgaris*	PstI	BR:Crb1:11
RC19	July, 2011	Craíbas, AL	S09°40'39.1"	W036°46'38.4"	*Phaseolus vulgaris*	PstI	BR:Crb2:11
RC29	July, 2011	Craíbas, AL	S09°40'39.8"	W036°46'38.4"	*Phaseolus vulgaris*	PstI	BR:Crb10:11
RC77	July, 2011	Olho D'água das Flores, AL	S09°32'28.1"	W037°17'26.8"	*Phaseolus vulgaris*	PstI	BR:Oaf8:11
46AL	July, 2011	Olho D'água das Flores, AL	S09°32'27.8"	W037°17'22.1"	*Macroptilium lathyroides*	PstI	BR:Oaf28:11
61AL	July, 2011	Olho D'água das Flores, AL	S09°32'26.9"	W037°17'17.6"	*Macroptilium lathyroides*	PstI	BR:Oaf27:11
RC94	July, 2011	Olho D'água das Flores, AL	S09°32'59.4"	W037°18'43.5"	*Macroptilium lathyroides*	PstI	BR:Oaf24:11
RC95	July, 2011	Olho D'água das Flores, AL	S09°32'59.2"	W037°18'43.0"	*Macroptilium lathyroides*	PstI	BR:Oaf25.:11
						PstI	BR:Oaf25.2:11
97AL	July, 2011	Santana do Ipanema, AL	S09°23'25.1"	W037°12'46.9"	*Phaseolus lunatus*	PstI	BR:Sti34:11
99AL	July, 2011	Santana do Ipanema, AL	S09°23'24.7"	W037°12'48.0"	*Phaseolus lunatus*	BglII	BR:Sti2:11
100AL	July, 2011	Santana do Ipanema, AL	S09°23'24.8"	W037°12'48.0"	*Phaseolus lunatus*	SpeI	BR:Sti3:11
102AL	July, 2011	Santana do Ipanema, AL	S09°23'24.9"	W037°12'48.0"	*Phaseolus lunatus*	BglII	BR:Sti4:11
105AL	July, 2011	Santana do Ipanema, AL	S09°23'24.9"	W037°12'48.3"	*Phaseolus lunatus*	PstI	BR:Sti35:11
111AL	July, 2011	Santana do Ipanema, AL	S09°23'24.9"	W037°12'47.5"	*Phaseolus lunatus*	BglII	BR:Sti9.1:11
						BglII	BR:Sti9.2:11
						SpeI	BR:Sti9.3:11
112AL	July, 2011	Santana do Ipanema, AL	S09°23'24.9"	W037°12'47.6"	*Phaseolus lunatus*	PstI	BR:Sti10:11
128AL	July, 2012	Santana do Ipanema, AL	S09°23'24.9"	W037°12'47.6"	*Phaseolus lunatus*	ApaI	BR:Sti26:11
152AL	July, 2011	Santana do Ipanema, AL	S09°23'24.8"	W037°12'48.6"	*Phaseolus lunatus*	XbaI	BR:Sti29:11
Bean golden mosaic virus (BGMV)							
MG-146	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una1:12
MG-147	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una2:12
MG-148	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una3:12
MG-149	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una4:12
MG-150	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una5:12
MG-152	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una6.1:12
						SpeI	BR:Una6.2:12
MG-153	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una7:12
MG-154	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una8.1:12
						SpeI	BRUna8.2:12
MG-155	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una9:12

Suppl. Table S1 (cont.)

MG-156	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una10.1:12
						SpeI	BR:Una10.2:12
MG-157	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una11:12
MG-158	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una12:12
						SpeI	BR:Una13.2:12
MG-162	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una14:12
MG-163	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una15:12
MG-165	June, 2012	Unaí, MG	S16°43'02.2"	W046°45'14"	*Phaseolus vulgaris*	SpeI	BR:Una16:12
GO-2	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag11:12
GO-3	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag1.1:12
						SpeI	BR:Sag1.2:12
GO-4	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag2:12
						SpeI	BR:Sag2.3:12
GO-9	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag7:12
GO-15	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag5:12
GO-22	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag6:12
GO-25	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag12:12
GO-26	June, 2013	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag8:12
GO-27	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag9:12
GO-29	June, 2012	Santo Antônio de Goiás, GO	S16°30'22.5"	W049°17'06.8"	*Phaseolus vulgaris*	SpeI	BR:Sag10:12
GO-63	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par14.1:12
						HindIII	BR:Par14.2:12
GO-64	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par1:12
GO-65	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par2.1:12
						SpeI	BR:Par2.2:12
GO-66	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par15.1:12
						HindIII	BR:Par15.2:12
GO-67	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	HindIII	BR:Par67:12
GO-68	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par16.1:12
						HindIII	BR:Par16.2:12
GO-70	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par3:12
GO-71	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par17:12
GO-72	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	SpeI	BR:Par4:12

Suppl. Table S1 (cont.)

GO-73	June, 2012	Paranoá, DF	S16°00'47"	W047°33'21.5"	*Phaseolus vulgaris*	HindIII	BR:Par5.1:12
						SpeI	BR:Par5.2:12
GO-75	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par6.1:12
						HindIII	BR:Par6.2:12
GO-76	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par18:12
GO-78	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par19:12
GO-79	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par8:12
GO-80	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par20.1:12
						HindIII	BR:Par20.2:12
G0-81	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par21:12
GO-82	June, 2013	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par22:12
GO-85	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par23:12
GO-88	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par9.1:12
						SpeI	BR:Par9.2:12
GO-89	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par10:12
GO-91	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par24:12
GO-92	June, 2012	Paranoá, DF	S15°57'57.4"	W047°31'35.8"	*Phaseolus vulgaris*	SpeI	BR:Par25:12
GO-200	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri9:12
GO-201	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri10:12
GO-204	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri11:12
GO-207	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri12:12
GO-214	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri4:12
GO-221	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri6:12
						SpeI	BR:Cri6.1:12
GO-222	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri7.1:12
						SpeI	BR:Cri7.2:12
GO-227	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri13.1:12
						SpeI	BR:Cri13.2:12
GO-228	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri14:12
GO-236	June, 2013	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri15:12
GO-238	June, 2012	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri16:12
GO241	June, 2013	Cristalina, GO	S17°05'29.2"	W047°37'01.4"	*Phaseolus vulgaris*	SpeI	BR:Cri17:12
159AL	June, 2013	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	Spe I	BR:Pai2:11

Suppl. Table S1 (cont.)

162AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai4:11
163AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai5:11
164AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai6:11
165AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai7:11
166AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai8:11
167AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai9:11
168AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai10:11
169AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SacI	BR:Pai11:11
170AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai12:11
173AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai13.1:11
						SpeI	BR:Pai13.2:11
						SpeI	BR:Pai13.3:11
174AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SpeI	BR:Pai14.1:11
						SpeI	BR:Pai14.2:11
176AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SacI	BR:Pai15:11
177AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SacI	BR:Pai16.1:11
						SpeI	BR:Pai16.2:11
178AL	July, 2011	Palmeira dos Índios, AL	S09°27'11.3"	W036°36'34.3"	*Phaseolus lunatus*	SacI	BR:Pai17:11
254AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur1:11
255AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur6.1:11
						SpeI	BR:Mur6.2:11
257AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SacI	BR:Mur7:11
258AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur8:11
262AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur10.1:11
						SpeI	BR:Mur10.2:11
265AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SacI	BR:Mur2.1:11
						SpeI	BR:Mur2.2:11
269AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SacI	BR:Mur3:11
272AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur12.1:11
						SpeI	BR:Mur12.2:11
274AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur4:11
275AL	July, 2011	Murici, AL	S09°17'10.0"	W035°57'43.3"	*Phaseolus lunatus*	SpeI	BR:Mur5.1:11
						SpeI	BR:Mur5.2:11

Suppl. Table S1 (cont.)

310MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt1:11
311MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt2:11
321MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	HindIII	BR:Flt20:11
323MG		Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt323:11
325MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	HindIII	BR:Flt5.1:11
						XbaI	BR:Flt5.2:11
326MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	HindIII	BR:Flt6:11
						XbaI	BR:Flt6.1:12
330MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt8.1:11
						XbaI	BR:Flt8.2:11
332MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt13:11
337MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt11:11
338MG	March, 2011	Florestal, MG	-	-	*Macroptilium lathyroides*	XbaI	BR:Flt12:11
Tomato severe rugose virus (ToSRV)							
DV202	July, 2008	Florestal, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Flo202.3:08
						KpnI	BR:Flo202.4:08
DV203	July, 2008	Florestal, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Flo203:08
DV206	July, 2008	Florestal, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Flo206:08
DV208	July, 2008	Florestal, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Flo208:08
DV218	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car218:08
DV219	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car219:08
DV220	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car220:08
DV224	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car224.1:08
						KpnI	BR:Car224.3:08
DV226	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car226.1:08
						KpnI	BR:Car226.3:08
DV230	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car230:08
DV232	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car232:08
DV237	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car237.1:08
						KpnI	BR:Car237.2:08
DV238	July, 2008	Carandaí, MG	-	-	*Solanum lycopersicum*	KpnI	BR:Car238.1:08
OS70	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi70:13
OS73	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi73:13

Suppl. Table S1 (cont.)

OS79	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi79:13
OS80	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi80:13
OS83	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi83:13
OS84	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi84:13
OS85	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi85:13
OS86	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi86:13
OS92	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi92.2:13
						KpnI	BR:Coi92.5:13
OS93	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi93:13
OS95	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi95:13
OS103	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi103.1:13
						KpnI	BR:Coi103.4:13
OS110	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi110:13
OS114	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi114:13
OS117	July, 2013	Coimbra, MG	S20°51'43,2"	W042°51'27"	*Solanum lycopersicum*	KpnI	BR:Coi117:13
						KpnI	BR:Coi117.1:13
OS119	February, 2014	Coimbra, MG	S20°36'39,3 "	W042°25'58,9"	*Solanum lycopersicum*	KpnI	BR:Coi119.1:13
						KpnI	BR:Coi119.4:13
OS120	February, 2014	Coimbra, MG	S20°36'39,3"	W042°25'58,9"	*Solanum lycopersicum*	KpnI	BR:Coi120:14
OS122	February, 2014	Coimbra, MG	S20°36'39,3"	W042°49'16,9"	*Solanum lycopersicum*	KpnI	BR:Coi122:14
OS125	February, 2014	Coimbra, MG	S20°50'51 "	W042°52'41,2"	*Solanum lycopersicum*	KpnI	BR:Coi125:14
OS131	February, 2014	Coimbra, MG	S20°50'49,4"	W042°52'41,9"	*Solanum lycopersicum*	KpnI	BR:Coi131:14
OS172	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Solanum lycopersicum*	KpnI	BR:Coi172:14
OS179	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Solanum lycopersicum*	KpnI	BR:Coi179:14
OS182	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Solanum lycopersicum*	KpnI	BR:Coi182:14
OS183	February, 2014	Coimbra, MG	S20°50'58,3"	W042°53'14,4"	*Solanum lycopersicum*	KpnI	BR:Coi183:14
OS201	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo01:14
OS202	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo02:14
OS203	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo03:14
OS204	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo04.1:14
						KpnI	BR:Flo04.5:14
OS205	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo05:14
OS206	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo06:14

Suppl. Table S1 (cont.)

OS207	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo07:14
OS208	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo08.1:14
						KpnI	BR:Flo08.2:14
OS209	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo09.1:14
						KpnI	BR:Flo09.3:14
OS210	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo10:14
OS211	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo11:14
OS212	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo12:14
OS213	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo13:14
OS214	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo14:14
OS216	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo16:14
OS218	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo18:14
OS219	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo19:14
OS220	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo20:14
OS222	June, 2014	Florestal, MG	S19°55'55,8"	W044°23'52,4"	*Solanum lycopersicum*	KpnI	BR:Flo22:14

Supplementary Table S2. Begomoviruses sequences retrieved from GenBank.

Isolate	Genbank access number		Location, state	Date of collection	Host	Reference
	DNA-A	DNA-B				
Blainvillea yellow spot virus (BlYSV)						
BR:Vic26s:10		KC706529.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic26c:10		KC706528.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic21:10		KC706527.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic18:10		KC706526.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic11:10	KC706520.1	KC706525.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic08:10	KC706518.1	KC706524.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic07:10		KC706523.1	Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Coi25:07	EU710756.1	EU710757.1	Coimbra, MG	July, 2007	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic04.1:10	KC706516.1		Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic04.2:10	KC706517.1		Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic13:10	KC706521.1		Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic09:10	KC706519.1		Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Vic20:10	KC706522.1		Viçosa, MG	May, 2010	*Blainvillea rhomboidea*	Rocha et al., 2013
BR:Jun1:09	JX871394.1		Junqueiro, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Lim1:09	JX871393.1		Limoeiro, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Rla6:09	JX871392.1		Rio Largo, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Rla5:10	JX871391.1		Rio Largo, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Rla4:10	JX871390.1		Rio Largo, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Rla3:10	JX871389.1		Rio Largo, AL	November, 2009	*Blainvillea rhomboidea*	Tavares et al., 2012
BR:Coi32.1:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi32.2:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi34:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi37.1:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi37.2:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi51:13			Coimbra, MG	July, 2013	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi151:14			Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi152.1:14			Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi153:14			Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi154:14			Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi155:14			Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014

Suppl. Table S2 (cont.)

BR:Coi156:14		Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi157:14		Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
BR:Coi164:14		Coimbra, MG	February, 2014	*Blainvillea rhomboidea*	Sande, 2014
Bean golden mosaic virus (BGMV)					
BR:Cri17:12	KJ939853.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri16:12	KJ939852.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri15:12	KJ939851.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri14:12	KJ939850.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri13:12	KJ939849.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri8:12	KJ939848.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri7:12	KJ939847.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri5:12	KJ939845.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri6:12	KJ939846.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri4:12	KJ939844.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri12:12	KJ939844.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri11:12	KJ939842.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri10:12	KJ939841.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri9:12	KJ939840.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri3:12	KJ939839.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri2:12	KJ939838.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Cri1:12	KJ939837.1	Cristalina, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una16:12	KJ939836.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una15:12	KJ939835.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una14:12	KJ939834.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una12:12	KJ939832.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una13:12	KJ939833.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una11:12	KJ939831.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una10:12	KJ939830.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una9:12	KJ939829.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una8:12	KJ939828.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una7:12	KJ939827.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una5:12	KJ939825.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una6:12	KJ939826.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014

Suppl. Table S2 (cont.)

BR:Una4:12	KJ939824.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una3:12	KJ939823.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una2:12	KJ939822.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Una1:12	KJ939821.1	Unaí, MG	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par13:12	KJ939820.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par12:12	KJ939819.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par30:12	KJ939818.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par29:12	KJ939817.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par28:12	KJ939816.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par11:12	KJ939815.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par27:12	KJ939814.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par26:12	KJ939813.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par25:12	KJ939812.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par24:12	KJ939811.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par10:12	KJ939810.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par9:12	KJ939809.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par23:12	KJ939808.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par22:12	KJ939807.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par21:12	KJ939806.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par20:12	KJ939805.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par8:12	KJ939804.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par19:12	KJ939803.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par7:12	KJ939802.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par18:12	KJ939801.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par6:12	KJ939800.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par5:12	KJ939799.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par3:12	KJ939796.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par17:12	KJ939797.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par4:12	KJ939798.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par16:12	KJ939795.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par15:12	KJ939794.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par2:12	KJ939793.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Par1:12	KJ939792.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014

Suppl. Table S2 (cont.)

BR:Par14:12	KJ939791.1	Paranoá, DF	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag12:12	KJ939790.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag11:12	KJ939789.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag10:12	KJ939788.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag9:12	KJ939787.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag8:12	KJ939786.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag7:12	KJ939785.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag6:12	KJ939784.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag5:12	KJ939783.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag4:12	KJ939782.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag3:12	KJ939781.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag2:12	KJ939780.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Sag1:12	KJ939779.1	Santo Antônio de Goiás, GO	June, 2012	*Phaseolus vulgaris*	Sobrinho et al., 2014
BR:Flt13:11	KJ939778.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt12:11	KJ939777.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt11:11	KJ939776.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt10:11	KJ939775.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt9:11	KJ939774.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt8:11	KJ939773.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt7:11	KJ939772.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt6:11	KJ939771.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt5:11	KJ939770.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt4:11	KJ939769.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt3:11	KJ939768.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt2:11	KJ939767.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Flt1:11	KJ939766.1	Florestal, MG	March, 2011	*Macroptilium lathyroides*	Sobrinho et al., 2014
BR:Mur12:11	KJ939765.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur11:11	KJ939764.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur10:11	KJ939763.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur9:11	KJ939762.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur8:11	KJ939761.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur7:11	KJ939760.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur6:11	KJ939759.1	Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014

Suppl. Table S2 (cont.)

BR:Mur5:11	KJ939758.1		Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur4:11	KJ939757.1		Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur3:11	KJ939756.1		Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur2:11	KJ939755.1		Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Mur1:11	KJ939754.1		Murici, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai17:11	KJ939753.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai16:11	KJ939752.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai15:11	KJ939751.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai14:11	KJ939750.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai13:11	KJ939749.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai12:11	KJ939748.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai11:11	KJ939747.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai10:11	KJ939746.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai9:11	KJ939745.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai8:11	KJ939744.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai7:11	KJ939743.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai6:11	KJ939742.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai5:11	KJ939741.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai4:11	KJ939740.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai3:11	KJ939739.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai2:11	KJ939738.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
BR:Pai1:11	KJ939737.1		Palmeira dos Índios, AL	July, 2011	*Phaseolus lunatus*	Sobrinho et al., 2014
Tomato severe rugose virus (ToSRV)						
BR:Car238:08	KC004089	KC706627.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car237.6:08	KC004088	KC706626.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car235:08	KC004086	KC706625.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car237:08	KC706620	KC706624.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car234.5:08		KC706623.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car223:08		KC706622.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car217.6:08		KC706621.1	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
Baccatum-9	DQ207749.1	EF534708.1	Petrolina de Goiás, GO	November, 2003	*Capsicum baccatum*	Bezerra-Agasie et al., 2006
Sumare	EU086591.2	GU358449.1	Sumaré, SP	February, 2007	*Nicandra physaloides*	Barbosa et al., 2007
Pi-1	HQ606467.1	HQ606468.1	Piracicaba, SP	October, 2007	*Solanum lycopersicum*	Barbosa et al., 2011

Suppl. Table S2 (cont.)

BR:Coi70.1:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi70.2:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi70.3:13		Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi73.2:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi73.4:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi78:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi79:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi80:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi83:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi84:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi85.1:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi85.2:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi86:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi88:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi91:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi92.2:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi93:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi95:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi100.1:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi100.2:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi103.1:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi107:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi109:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi110:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi114:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi117:13	-	Coimbra, MG	July, 2013	*Solanum lycopersicum*	Sande, 2014
BR:Coi119.1:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi120:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi122:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi125:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi127:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi131:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi172:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014

Suppl. Table S2 (cont.)

BR:Coi179:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi182:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi183.1:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Coi183.2:14	-	Coimbra, MG	February, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo01:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo02:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo04.1:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo06:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo07:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo09.1:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo13:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo14:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo15:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo16:13	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo18:13	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo19:13	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo22:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo23:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo31:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Flo37:14	-	Florestal, MG	June, 2014	*Solanum lycopersicum*	Sande, 2014
BR:Car214:08	KC004075	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car218.1:08	KC004076	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car219.10:08	KC004077	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car220:08	KC004078	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car224:08	KC004079	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car226.3:08	KC004080	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car227:08	KC004081	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car228:08	KC004082	Carandaí, MG	July, 2008	*Sida* sp.	Rocha et al., 2013
BR:Car230:08	KC004083	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car232:08	KC004084	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car233:08	KC004085	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Car236.1:08	KC004087	Carandaí, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Flo165:08	KC004070	Florestal, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013

Suppl. Table S2 (cont.)

BR:Flo202:08	KC004071	Florestal, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Flo203:08	KC004072	Florestal, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Flo206:08	KC004073	Florestal, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013
BR:Flo208:08	KC004074	Florestal, MG	July, 2008	*Solanum lycopersicum*	Rocha et al., 2013

Supplementary Table S3. Recombination events detected in the DNA-A of *Bean golden mosaic virus* (BGMV), *Blainvillea yellow spot virus* (BlYSV) and *Macroptilium yellow spot virus* (MaYSV) populations.

Event	Recombinant isolate (s)	Recombination breakpoints[2]		Parents		Method[3]	*p*-value[4]
		Begin	End	Minor	Major		
BGMV[1]							
2	BR:Mur12:11, BR:Mur11:11, BR:Mur10:11, BR:Mur9:11, BR:Mur8:11, BR:Mur7:11, BR:Mur6:11, BR:Mur5:11, BR:Mur4:11, BR:Mur3:11, BR:Mur2:11, BR:Mur1:11, BR:Pai17:11, BR:Pai16:11, BR:Pai15:11 , BR:Pai14:11, BR:Pai13:11, BR:Pai12:11, BR:Pai11:11, BR:Pai10:11, BR:Pai9:11, BR:Pai8:11, BR:Pai7:11, BR:Pai6:11, BR:Pai5:11, BR:Pai4:11, BR:Pai3:11, BR:Pai2:11, BR:Pai1:11	410	1115	BR:Rec1:05	BR:Car4:10	RGBMCS3	$1,870 \times 10^{-13}$
BlYSV							
1	BR:Vic04.2:10, BR:Coi25:07, BR:Vic04.1:10, BR:Vic08:10, BR:Vic09:10, BR:Vic11:10, BR:Vic20:10, BR:Coi34:13, BR:Coi37.1:13, BR:Coi37.2:13, BR:Coi51:13, BR:Coi151:14, BR:Coi152.1:14, BR:Coi154:14, BR:Coi155:14, BR:Coi156:14, BR:Coi157:14, BR:Coi164:14	2403	2628	BR:Vic13:10	BR:Coi32.2:13	RGMCS3	$2,057 \times 10^{-07}$
MaYSV[1]							
1	BR:Sti35:11, BR:Sti26:11, BR:Sti10:11, BR:Sti9:11, BR:Sti3:11, BR:Sti2:11, BR:Oaf25:11, BR:Oaf24:11, BR:Oaf27:11, BR:Oaf28:11, BR:Oaf8:11, BR:Crb10:11	2618	1660	BR:Pdi1:10	Unknown	RGBMCS3	$6,86 \times 10^{-59}$
4	BR:Oaf8:11 BR:Oaf25:11	1975	2157	Unknown Unknown	BR:Crb20:11 BR:Crb14:11	RGBMCS3	$1,88 \times 10^{-23}$
5	BR:Sti2:11	(?) 1661	(?) 1974	BR:Crb10:11	BR:Oaf25:11	RGBS3	$7,78 \times 10^{-19}$
7	BR:Oaf28:11 BR:Crb10:11 BR:Sti35:11, BR:Sti26:11, BR:Sti9:11, BR:Sti3:11, BR:Oaf24:11, BR:Oaf27:11,	2386	(?) 2617	Unknown Unknown Unknown	BR:Oaf11:11 BR:Oaf22:11 BR:Sti20:11	GMCS	$2,99 \times 10^{-20}$

[1]Data obtained from Ramos-Sobrinho et al., 2014.
[2]Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increase clockwise. (?), Breakpoints could not be accurately located.
[3]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.
[4]The reported *p*-value is from the method in bold and is the lowest *p*-value calculated for the featured event.

Supplementary Table S4. Features of the tRep ORFs located in DNA-B of Macroptilium yellow spot virus (MaYSV) isolates.

Isolate	Type	Length (nt/aa)	Location in the genome[1]		Coverage (%aa)[2]	Identity (% nt/aa)	E value (nt/aa)
			Start	Stop			
BR:Oaf8:11	I	105/35	2471	2366	100	91/88	$4e^{-31}/6e^{-12}$
BR:Crb10:11	I	108/36	2472	2364	100	97/94	$2e^{-41}/1e^{-13}$
BR:Sti34:11	I	105/35	2468	2363	100	98/94	$6e^{-41}/1e^{-12}$
BR:Sti2:11	II	162/54	2473	2312	96	92/80	$2e^{-44}/4e^{-18}$
BR:Sti3:11	II	162/54	2470	2308	81	97/93	$3e^{-54}/9e^{-12}$
BR:Sti4:11	II	162/54	2473	2312	79	90/79	$3e^{-34}/7e^{-17}$
BR:Sti26:11	II	174/58	2486	2312	87	90/68	$6e^{-32}/3e^{-11}$
BR:Sti35:11	II	162/54	2470	2308	81	99/98	$6e^{-57}/5e^{-20}$
BR:Crb1:11	III	285/95	2439	2154	47	99/100	$1e^{-61}/1e^{-23}$
BR:Crb2:11	III	285/95	2439	2154	47	99/100	$1e^{-61}/1e^{-23}$

[1]Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.
[2]Percentage of tRepORF in amino acids, which is homologous to the Rep ORF in the cognate DNA-A. Analyses performed using the BLASTn and BLASTp algorithms.

Supplementary Table S5. Conserved sequences involved in the replication process located in the Rep protein and intergenic region of DNA-A and DNA-B.

Population[1]	Iteron[2]	IRD[3]
Murici (AL)	GGTG/GG**G**G	FKIN
Palmeira (AL)	GG**G**G/GG**G**G	FKI**S**
MG/GO/DF	GGTG/GGTG	FKIN

[1]AL, Alagoas; MG, Minas Gerais; GO, Goiás; DF, Distrito Federal
[2]Conserved sequences that are specifically recognized by Rep, located in the intergenic region of the DNA-A and DNA-B.
[3]Iteron related domain, located in the N-terminal region of Rep.