AYANE FERNANDA FERREIRA QUADROS

COMPOSITION OF BEGOMOVIRUS POPULATIONS IN CULTIVATED AND NON-CULTIVATED HOSTS DETERMINED BY HIGH-THROUGHPUT SEQUENCING

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*.

Orientador: Francisco Murilo Zerbini Júnior

VIÇOSA - MINAS GERAIS
2019

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

Q1c
2019

Quadros, Ayane Fernanda Ferreira, 1994-
Composition of begomovirus populations in cultivated and non-cultivated hosts determined by high-throughput sequencing [recurso eletrônico] / Ayane Fernanda Ferreira Quadros. – Viçosa, MG, 2019.
54 f.: il. (algumas color.).

Texto em inglês.
Orientador: Francisco Murilo Zerbini Junior.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.
Disponível em: https://www.locus.ufv.br/

1. Begomovírus. 2. Tomate. 3. *Sida* sp. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-Graduação em Fitopatologia. II. Título.

CDD 22. ed. 579.28

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB6 2523

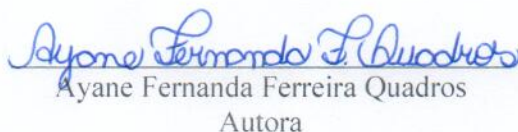
AYANE FERNANDA FERREIRA QUADROS

COMPOSITION OF BEGOMOVIRUS POPULATIONS IN CULTIVATED AND NON-CULTIVATED HOSTS DETERMINED BY HIGH-THROUGHPUT SEQUENCING

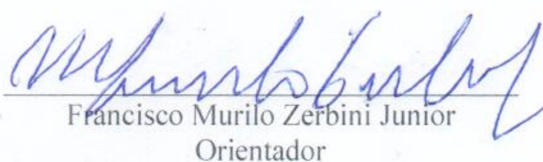
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitopatologia, para obtenção do título de *Magister Scientiae*.

APROVADA: 31 de agosto de 2019.

Assentimento:



Ayane Fernanda Ferreira Quadros
Autora



Francisco Murilo Zerbini Junior
Orientador

AKNOWLEDGEMENTS

A Deus, pela vida, por todas as oportunidades concedidas e por estar presente em TODOS os momentos me dando força e sabedoria necessária para enfrentar todas as dificuldades.

Aos meus pais e a meu irmão, Guilherme, meus maiores incentivados, agradeço pelo amor incondicional, pelo carinho, dedicação, paciência e incentivo.

A minha família, que é meu porto seguro em todos os momentos e torce sempre por mim.

A Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Fitopatologia pela oportunidade em realizar o mestrado e conduzir esta pesquisa.

Ao professor Francisco Murilo Zerbini Júnior, por toda a orientação, paciência, dedicação, ensinamentos, confiança, conselhos e amizade.

A professora Poliane Alfenas Zerbini pela oportunidade, suporte e confiança.

Aos amigos do Laboratório de Virologia Vegetal Molecular, Anelise, Roberta, Osvaldo, Baltazar, Patrícia, Lucas, Fernanda e César pelos incentivos constantes, companheirismo, momentos de alegria e descontração e por tornarem o ambiente tão agradável. Em especial, aos amigos, Monique, Tarsiane, Angélica, João Paulo, Rafael e Tales pela amizade, dedicação, colaboração e responsabilidades compartilhadas durante os dois anos que passamos em Viçosa. Minha eterna gratidão!

Aos meus amigos, em especial a Izabel Batista e Amarildo Junior, pelo apoio que serviram de motivação a todos os passos desta caminhada, se tornando fundamental na conquista desde objetivo.

A todos os demais amigos e colegas, que participaram direta ou indiretamente da minha formação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

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

ABSTRACT

QUADROS, Ayane Fernanda Ferreira, M.Sc., Universidade Federal de Viçosa, July, 2019. **Composition of begomovirus populations in cultivated and non-cultivated hosts determined by high-throughput sequencing**. Adviser: Francisco Murilo Zerbini Júnior.

The genus *Begomovirus* (family *Geminiviridae*) includes single-stranded DNA plant viruses transmitted by whiteflies. Begomoviruses are among the most damaging plant pathogens, causing epidemics in economically important crops worldwide. Tomato-infecting begomoviruses emerged in Brazil in the early 1990's following the introduction of *Bemisia tabaci* Middle East-Asia Minor 1 (MEAM1, previously known as *B. tabaci* biotype B). Several lines of evidence indicate that these viruses evolved from indigenous viruses infecting non-cultivated hosts. However, tomato-infecting viruses are only rarely found in non-cultivated hosts, and vice-versa. It is possible that viral populations in a given host are composed primarily of viruses which are better adapted to this host, but also include a very small proportion of viruses which are poorly adapted. Then, after transfer to a different host by the whitefly vector, the composition of the viral population shifts rapidly, with the viruses which are better adapted to the new host becoming predominant. To test this hypothesis, we collected tomato and *Sida* sp. plants, growing next to each other, at two locations (Coimbra and Florestal, both in Minas Gerais state, Brazil), in 2014 and 2018. Viral infection was confirmed by by polymerase chain reaction (PCR) using specific primers. Total DNA from one tomato and one *Sida* sp. sample from each location and year (except Florestal/2018) were sequenced on the Illumina HiSeq 2000 (2014) and NovaSeq 6000 (2018) platforms. Following a highly stringent set of criteria (*E*-value cutoff of 1.10^{-10}), reads were mapped using BLAST*n* to a data set including all New World begomoviruses. The reads were classified as (i) *Tomato severe rugose virus* (ToSRV), (ii) *Sida micrantha mosaic virus* (SiMMV) and (iii) *Sida common mosaic virus* (SiCmMV), when the first three hits were isolates of these species, or (iv) begomovirus, when the first three hits included isolates of different species. To resolve the results at the analysis interface and to minimize false positives, mapping was performed using reference sequences from the samples collected. For the 2014 samples, >98% of the reads from *Sida* sp. mapped to SiMMV, but 0.01% of the reads mapped to ToSRV. Conversely, >99% of the reads from tomato mapped to ToSRV, with 0.001% mapping to SiMMV. For the 2018 samples, the results indicate an increase in the diversity of viruses infecting *Sida* sp. (including SiMMV, SiCmMV and Abutilon viruses), but 0.1% of the reads mapped to ToSRV. These results are consistent with the hypothesis that viral

populations in a given host are composed primarily of the virus that is most adapted to this host but also includes a very small proportion of viruses that are less adapted.

Keywords: Begomovirus. Tomato. *Sida* sp.

RESUMO

QUADROS, Ayane Fernanda Ferreira Quadros, M.Sc., Universidade Federal de Viçosa, julho de 2019. **Composição de populações de begomovírus em hospedeiros cultivados e não-cultivados determinada por meio de sequenciamento de alto desempenho**. Orientador: Francisco Murilo Zerbini Júnior.

O gênero *Begomovirus* inclui vírus de plantas de DNA de fita simples, transmitidos por moscas brancas. Os begomovírus estão entre os fitopatógenos mais importantes, causando epidemias em culturas economicamente importantes em todo o mundo. Os begomovírus que infectam o tomateiro emergiram no Brasil no início da década de 1990, após a introdução de *Bemisia tabaci* Middle EastAsia Minor 1 (MEAM1, anteriormente conhecida como *B. tabaci* biótipo B). Várias linhas de evidência indicam que esses vírus evoluíram a partir de vírus nativos que infectam hospedeiros não-cultivados. No entanto, os vírus que infectam tomateiro são raramente encontrados em hospedeiros não-cultivados, e vice-versa. É possível que as populações virais de um determinado hospedeiro sejam compostas primariamente por vírus mais adaptados a esse hospedeiro, mas também incluam uma proporção muito pequena de vírus que são mal adaptados. Após a transferência para um hospedeiro diferente pelo inseto vetor, a composição da população viral mudaria rapidamente, com os vírus que são melhor adaptados para o novo hospedeiro tornando-se predominantes. Para testar esta hipótese, coletamos plantas de tomateiro e *Sida* sp., crescendo lado a lado, em dois locais (Coimbra e Florestal, ambos no estado de Minas Gerais), em 2014 e 2018. A infecção viral foi confirmada por por reação em cadeia da polimerase (PCR) usando primers específicos. DNA total de uma amostra de tomateiro e uma de *Sida* sp. de cada local e ano (exceto Florestal/2018) foram sequenciadas nas plataformas Illumina HiSeq 2000 (2014) e NovaSeq 6000 (2018). Seguindo critérios altamente rigorosos (*E-value* de $1,10^{-10}$), as leituras foram mapeadas utilizando BLAST*n* e um conjunto de dados incluindo todos os begomovírus do Novo Mundo. As leituras foram como (i) *Tomato severe rugose virus* (ToSRV), (ii) *Sida micrantha mosaic virus* (SiMMV) e (iii) *Sida common mosaic virus* (SiCmMV), quando os três primeiros *hits* consistiram de isoladoss dessas espécies, e (iv) begomovírus, quando os três primeirso *hits* consistiram de isolados de diferentes espécies. Para resolver os resultados na interface de análise e minimizar falsos positivos, o mapeamento foi realizado usando sequências de referência das amostras coletadas. Nas amostras coletadas em 2014, >90% das leituras de *Sida* sp. foram mapeadas para SiMMV, e 0,01% foram mapeadas para ToSRV. Por outro lado, >90% das leituras de tomateiro foram mapeadas para ToSRV, com 0,001% mapeadas para SiMMV. Nas amostras coletadas em 2018,

os resultados indicaram um aumento na diversidade de vírus em *Sida* sp. (incluindo SiMMV, SiCmMV e vírus de Abutilon), porém 0,1% das leituras foram mapeadas para ToSRV. Estes resultados são consistentes com a hipótese de que as populações virais em um determinado hospedeiro são compostas principalmente do vírus que é mais adaptado a este hospedeiro, mas também inclui uma proporção muito pequena de vírus menos adaptados.

Palavras-chave: Begomovírus. Tomateiro. *Sida* sp.

SUMMARY

INTRODUCTION 11

Abstract 26

Introduction 26

Material and Methods 28

Sample collection 28

Viral detection 28

Sequencing 29

Quality control and filtering 29

Matching reads to begomovirus sequences 30

Mapping reads to a begomovirus reference genome 30

Recombination analysis 30

Results 30

Virus identification 31

Matching reads to begomovirus sequences 31

Mapping reads to begomovirus reference genomes 33

Recombination analysis 34

Discussion 34

References 38

INTRODUCTION

The *Geminiviridae* family includes plant viruses with a genome composed of one or two molecules of single stranded, circular DNA, encapsidated by a single structural protein in geminated, quasi-icosahedral particles. The family is composed of the genera *Begomovirus, Becurtovirus, Capulavirus, Curtovirus, Eragrovirus, Grablovirus, Mastrevirus, Topocuvirus* and *Turncurtovirus*, defined based on the type of insect vector, host range, genomic organization and phylogenetic relationships (Brown *et al.*, 2012; Varsani *et al.*, 2017; Zerbini *et al.*, 2017).

The genus *Begomovirus* includes viruses transmitted by whiteflies of the *Bemisia tabaci* cryptic species complex (Homoptera: Aleyrodidae) to dicotyledonous plants (Brown *et al.*, 2015). Begomoviruses cause severe diseases in economically important crops, especially in tropical and subtropical regions (Polston & Anderson, 1997; Monci *et al.*, 2002; Briddon, 2003; Legg & Fauquet, 2004; Morales, 2006; Navas-Castillo *et al.*, 2011; Rocha *et al.*, 2013; Inoue-Nagata *et al.*, 2016; Rojas *et al.*, 2018).

Based on phylogenetic analysis and genomic features, begomoviruses are divided into two groups: Old World (OW; Europe, Africa, Asia and Oceania) and New World (NW; the Americas) (Rybicki, 1994; Padidam *et al.*, 1999; Paximadis *et al.*, 1999). Begomoviruses in the New World are mostly bipartite, with two genomic components named DNA-A and DNA-B, with a few exceptions (Melgarejo *et al.*, 2013; Vu *et al.*, 2015; Macedo *et al.*, 2017; Romay *et al.*, 2019). Circular DNA satellites (alphasatellites) have also been described in association with NW bipartite begomoviruses (Mansoor *et al.*, 2003; Zhou *et al.*, 2003; Paprotka *et al.*, 2010; Fiallo-Olivé *et al.*, 2012a; Zhou, 2013; Lozano *et al.*, 2016; Ferro *et al.*, 2017; Mar *et al.*, 2017a).

The DNA-A of bipartite begomoviruses contains five genes, encoding the replication-associated protein (Rep), which is the initiator of rolling circle replication and has nucleic acid binding, endonuclease and ATPase activities (Fontes *et al.*, 1992; Orozco *et al.*, 1997); the trans-activating protein (TrAP), a transcriptional factor of the CP and NSP genes and also a suppressor of post-transcriptional gene silencing (Voinnet *et al.*, 1999; Wang *et al.*, 2005); the replication-enhancer protein (Ren), an accessory factor required for optimal replication (Sunter *et al.*, 1990; Pedersen & Hanley-Bowdoin, 1994); the AC4 protein, which is involved in gene silencing suppression (Vanitharani *et al.*, 2004); and the coat protein (CP), responsible for encapsidation of the ssDNA genome and also essential for insect transmission (Briddon *et al.*, 1990; Hofer *et al.*, 1997a). The DNA-B contains two genes, encoding the nuclear shuttle protein (NSP) and the movement protein (MP), which are involved in intra- and intercellular viral movement, respectively (Noueiry *et al.*, 1994; Sanderfoot *et al.*, 1996).

Brazil is a center of genetic diversity of begomoviruses, with reports of their detection since the 1950's (Costa & Bennett, 1950). The first report of a disease caused by a begomovirus was the golden mosaic of beans (*Phaseolus* spp.) (Costa, 1965). The virus was named as *Bean golden mosaic virus* (BGMV), and the disease acquired major economic importance during the 1970's after the expansion of soybean cultivation, an excellent host for the whitefly vector (Menten & Roston, 1980; De Fazio, 1985). The occurrence of tomato golden mosaic, caused by *Tomato golden mosaic virus* (TGMV), was reported in the mid-1970's (Matyis *et al.*, 1975). However, this disease never caused economic damage, probably because *B. tabaci* New World (NW, previously known as *B. tabaci* biotype A), the only whitefly species present in Brazil at the time, colonizes tomato with very low efficiency (Bedford *et al.*, 1994; Ribeiro *et al.*, 2003). However, following the introduction of *B. tabaci* Middle East-Asia Minor 1 (MEAM1, previously known as *B. tabaci* biotype B) in the early 1990's (Melo, 1992; Lourenção & Nagai, 1994), several begomoviruses emerged in tomato crops throughout the country (Ribeiro *et al.*,

1998; Ribeiro *et al.*, 2003). *B. tabaci* MEAM1 is highly polyphagous and feeds on tomato, as well as many other crops and non-cultivated plants. This vector drove the emergence of new species of tomato-infecting begomoviruses in Brazil and in many other regions of the world (Gilbertson *et al.*, 2015).

Identification of a begomovirus species requires the use of molecular tools (RCA, polymerase chain reaction [PCR] and complete genome sequencing), and in the last decades a large number of new begomoviruses infecting tomato plants have been characterized in Brazil (Ribeiro *et al.*, 2003; Fernandes *et al.*, 2006; Calegario *et al.*, 2007; Cotrim *et al.*, 2007; Ribeiro *et al.*, 2007; Castillo-Urquiza *et al.*, 2008; Fernandes *et al.*, 2008; Fernandes *et al.*, 2009; Silva *et al.*, 2011; Albuquerque *et al.*, 2012; Silva *et al.*, 2012; Tavares *et al.*, 2012; Rocha *et al.*, 2013; Macedo *et al.*, 2017; Quadros *et al.*, 2019). Despite the great diversity of viruses detected, they are not equally dispersed throughout the country. Studies in the major tomato producing regions revealed that two begomoviruses are prevalent: *Tomato severe rugose virus* (ToSRV) and *Tomato mottle leaf curl virus* (ToMoLCV). ToSRV is prevalent in the midwest and southeastern states (Goiás, Minas Gerais, São Paulo and the Federal District), while ToMoLCV is most common in the northeastern states (Bahia, Ceará, Pernambuco, Piauí) and the northern region of Minas Gerais (Inoue-Nagata *et al.*, 2016).

The emergence of a new virus depends on its adaptation to a new host and/or vector, so that efficient transmission within the new host population is ensured (Elena *et al.*, 2014). Adaptation to a new environment, such as a new host, requires that sufficient viral genetic variation is available for selection to operate (García-Arenal & Zerbini, 2019). The viruses that emerged in Brazil are unique and most likely originated from the large and genetically diverse reservoir of indigenous begomoviruses infecting non-cultivatd (wild) plants (Rocha *et al.*, 2013). Non-cultivated plants of many families have been reported as hosts of begomoviruses in several countries in the Americas (Frischmuth *et al.*, 1997; Roye *et al.*, 1997; Faria & Maxwell,

1999; Fernandes *et al.*, 1999; Idris *et al.*, 2003; Jovel *et al.*, 2004; Assunção *et al.*, 2006; Amarakoon *et al.*, 2008; Castillo-Urquiza *et al.*, 2008; Barbosa *et al.*, 2009; Silva *et al.*, 2011; Silva *et al.*, 2012; Tavares *et al.*, 2012; Barreto *et al.*, 2013; Pinto *et al.*, 2016; Ferro *et al.*, 2017; Passos *et al.*, 2017a; Passos *et al.*, 2017b) and *Sida* spp. has been described as the most abundant natural reservoirs for begomoviruses in several regions of the world (Frischmuth *et al.*, 1997; Hofer *et al.*, 1997b; Roye *et al.*, 1997; Echemendía *et al.*, 2004; Jovel *et al.*, 2004; Xiong *et al.*, 2005; Guo & Zhou, 2006; Das *et al.*, 2008; Fiallo-Olivé *et al.*, 2010; Fiallo-Olivé *et al.*, 2012b; Tavares *et al.*, 2012; Ferro *et al.*, 2017). Some of these begomoviruses also infect cultivated hosts (Barreto *et al.*, 2013; Rocha *et al.*, 2013; Ferro *et al.*, 2017) and begomoviruses originally described in *Sida* spp. (e.g. *Sida mottle virus*, SiMoV; *Sida micrantha mosaic virus*, SiMMV; *Sida common mosaic virus*, SiCmMV) have already been found naturally infecting tomato (Calegario, 2004; Castillo-Urquiza *et al.*, 2007; Cotrim *et al.*, 2007).

Although all available evidence indicates that tomato-infecting begomoviruses in Brazil originated from indigenous viruses in non-cultivated hosts, the viruses that predominate in tomato, such as ToSRV and ToMoLCV, are only rarely found in non-cultivated hosts, and vice-versa. This could be explained by the rapid evolution of such indigenous viruses after they are transferred to a new (cultivated) host by the insect vector. This rapid evolution, driven by the high nucleotide substitution rates of begomoviruses (Duffy & Holmes, 2008; Duffy & Holmes, 2009) and their propensity to recombine and reassort (Lefeuvre & Moriones, 2015), would originate the new species detected in crops. The predominance of certain viruses in tomato would be due to differences in adaptation to this host, or differences in efficiency of transmission by the insect vector.

Several observations support this hypothesis. First, all begomoviruses reported so far in tomato in Brazil are indigenous to the country, and except for neighboring Argentina (Rodríguez-Pardina *et al.*, 2011) are restricted to the country. Second, the biological

characterization of a number of tomato-infecting begomoviruses (e.g. ToSRV; *Tomato chlorotic mottle virus*, ToCMoV; *Tomato rugose mosaic virus*, ToRMV; *Tomato yellow spot virus*, ToYSV; *Tomato yellow vein streak virus*, ToYVSV) confirm that they can also infect weeds such as *Euphorbia heterophylla*, *Datura stramonium*, *Nicandra physaloides*, *Solanum nigrum* and *Sida* spp. (Fernandes *et al.*, 2006; Calegario *et al.*, 2007; Ribeiro *et al.*, 2007; Barreto *et al.*, 2013; Ferro *et al.*, 2017). Third, begomoviruses originally found in weeds, such as SiMoV (Fernandes *et al.*, 1999), SiMMV (Jovel *et al.*, 2004) and SiCmMV (Sande, 2014) have been found infecting tomato under field conditions (Calegario, 2004; Castillo-Urquiza *et al.*, 2007; Cotrim *et al.*, 2007; Castillo-Urquiza *et al.*, 2010). Fourth, evidence of recombination and reassortment has been obtained for tomato-infecting begomoviruses (Andrade *et al.*, 2006; Inoue-Nagata *et al.*, 2006; Ribeiro *et al.*, 2007).

However, the absence of an evolutionary relationship between the begomoviruses found in non-cultivated and cultivated plants, demonstrated in several in-depth studies (Silva *et al.*, 2012; Rocha *et al.*, 2013; Ramos-Sobrinho *et al.*, 2014; Sande, 2014; Mar *et al.*, 2017b), does not support this hypothesis.

As an alternative hypothesis, it is possible that viral populations in a given host are composed primarily of viruses which are better adapted to this host, but also include a small proportion of viruses which are poorly adapted that would not be easily detected by use of traditional molecular cloning and diagnostics tools. The composition of the virus population would change rapidly after transfer to a different host, with the viruses which are better adapted to the new host becoming predominant (and therefore being detected at high frequency). In this context, the recent advances of high-throughput sequencing technologies (HTS) open new possibilities to detect viruses which are present at extremely low concentration in a host.

Advances in HTS technologies have revolutionized the detection and discovery of new viruses (Maree *et al.*, 2018). The main feature of these technologies is their ability to produce

a huge amount of data, in some cases over one billion readings per round of sequencing, providing fast, accurate, and relatively low-cost genomic information (Barba *et al.*, 2014). The use of HTS can help in the understanding of the mechanisms involved in the emergence of new viruses in cultivated plants, as well as in the investigation of the role of non-cultivated hosts in the occurrence of epidemics in cultivated hosts.

The objective of this work was to test the hypothesis that viral populations in a given host are composed primarily of the virus that is better adapted to this host but also includes a very small proportion of viruses that are poorly adapted.

Literature Cited

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

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

Andrade, E.C.; Manhani, G.G.; Alfenas, P.F.; Calegario, R.F.; Fontes, E.P.B.; Zerbini, F.M. *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from *Sida* sp., forms pseudorecombinants with begomoviruses from tomato but not from *Sida*. **Journal of General Virology,** v. 87, p. 3687-3696, 2006.

Assunção, I.P.; Listik, A.F.; Barros, M.C.S.; Amorim, E.P.R.; Silva, S.J.C.; Izael, O.S.; Ramalho-Neto, C.E.; Lima, G.S.A. Genetic diversity of begomoviruses in invasive species of Northeastern Brazil (in Portuguese). **Planta Daninha,** v. 24, p. 239-244, 2006.

Barba, M.; Czosnek, H.; Hadidi, A. Historical perspective, development and applications of next-generation sequencing in plant virology. **Viruses,** v. 6, p. 106-136, 2014.

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

Barreto, S.S.; Hallwass, M.; Aquino, O.M.; Inoue-Nagata, A.K. A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. **Phytopathology,** v. 103, p. 436-44, 2013.

Bedford, I.D.; Briddon, R.W.; Brown, J.K.; Rosell, R.C.; Markham, P.G. Geminivirus transmission and biological characterization of *Bemisia tabaci* (Gennadius) biotypes from different geographical regions. **Annals of Applied Biology,** v. 125, p. 311-325, 1994.

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

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

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

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. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. **Archives of Virology,** v. 160, p. 1593-619, 2015.

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

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

Castillo-Urquiza, G.P.; Alfenas-Zerbini, P.; Beserra-Junior, J.E.A.; Mizubuti, E.S.G.; Varsani, A.; Martin, D.P.; Zerbini, F.M. Genetic structure of tomato-infecting begomovirus populations in two tomato-growing regions of Southeastern Brazil. Program and Abstracts, 6th International Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop, 2010, Guanajuato, Mexico. November 7-12th.

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

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

Costa, A.S. Three whitefly-transmitted virus diseases of beans in São Paulo, Brazil. **Plant Protection Bulletin F.A.O.,** v. 13, p. 2-12, 1965.

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

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

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

De Fazio, G. O mosaico dourado do feijoeiro no Brasil. **Fitopatologia Brasileira,** v. 10, p. 41-48, 1985.

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

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

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

Elena, S.F.; Fraile, A.; Garcia-Arenal, F. Evolution and emergence of plant viruses. **Advances in Virus Research,** v. 88, p. 161-191, 2014.

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

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

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

Fernandes, F.R.; Cruz, A.R.R.; Faria, J.C.; Zerbini, F.M.; Aragão, F.J.L. Three distinct begomoviruses associated with soybean in central Brazil. **Archives of Virology,** v. 154, p. 1567-1570, 2009.

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

Ferro, C.G.; Silva, J.P.; Xavier, C.a.D.; Godinho, M.T.; Lima, A.T.M.; Mar, T.B.; Lau, D.; Zerbini, F.M. The ever increasing diversity of begomoviruses infecting non-cultivated hosts: new species from *Sida* spp. and *Leonurus sibiricus*, plus two New World alphasatellites. **Annals of Applied Biology,** v. 170, p. 204-218, 2017.

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

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

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

Fontes, E.P.B.; Luckow, V.A.; Hanley-Bowdoin, L. A geminivirus replication protein is a sequence-specific DNA binding protein. **Plant Cell,** v. 4, p. 597-608, 1992.

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

García-Arenal, F.; Zerbini, F.M. Life on the edge: geminiviruses at the interface between crops and wild plant hosts. **Annual Review of Virology,** v. 6, 2019.

Gilbertson, R.L.; Batuman, O.; Webster, C.G.; Adkins, S. Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. **Annual Review of Virology,** v. 2, p. 67-93, 2015.

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

Hofer, P.; Bedford, I.D.; Markham, P.G.; Jeske, H.; Frischmuth, T. Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. **Virology,** v. 236, p. 288-95, 1997a.

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

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

Inoue-Nagata, A.K.; Lima, M.F.; Gilbertson, R.L. A review of geminivirus diseases in vegetables and other crops in Brazil: current status and approaches for management. **Horticultura Brasileira,** v. 34, p. 8-18, 2016.

Inoue-Nagata, A.K.; Martin, D.P.; Boiteux, L.S.; Giordano, L.D.; Bezerra, I.C.; De Avila, A.C. New species emergence via recombination among isolates of the Brazilian tomato-infecting begomovirus complex. **Pesquisa Agropecuaria Brasileira,** v. 41, p. 1329-1332, 2006.

Jovel, J.; Reski, G.; Rothenstein, D.; Ringel, M.; Frischmuth, T.; Jeske, H. *Sida micrantha* mosaic is associated with a complex infection of begomoviruses different from *Abutilon mosaic virus*. **Archives of Virology,** v. 149, p. 829-41, 2004.

Lefeuvre, P.; Moriones, E. Recombination as a motor of host switches and virus emergence: Geminiviruses as case studies. **Current Opinion in Virology,** v. 10, p. 14-9, 2015.

Legg, J.; Fauquet, C. Cassava mosaic geminiviruses in Africa. **Plant Molecular Biology,** v. 56, p. 585-599, 2004.

Lourenção, A.L.; Nagai, H. Surtos populacionais de *Bemisia tabaci* no Estado de São Paulo. **Bragantia,** v. 53, p. 53-59, 1994.

Lozano, G.; Trenado, H.P.; Fiallo-Olive, E.; Chirinos, D.; Geraud-Pouey, F.; Briddon, R.W.; Navas-Castillo, J. Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, *Geminiviridae*) - definition of a distinct class of begomovirus-associated satellites. **Frontiers in Microbiology,** v. 7, p. 162, 2016.

Macedo, M.A.; Albuquerque, L.C.; Maliano, M.R.; Souza, J.O.; Rojas, M.R.; Inoue-Nagata, A.K.; Gilbertson, R.L. Characterization of tomato leaf curl purple vein virus, a new monopartite New World begomovirus infecting tomato in Northeast Brazil. **Archives of Virology,** v. 163, p. 737-743, 2017.

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

Mar, T.B.; Mendes, I.R.; Lau, D.; Fiallo-Olive, E.; Navas-Castillo, J.; Alves, M.S.; Zerbini, F.M. Interaction between the New World begomovirus *Euphorbia yellow mosaic virus* and

its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. **Journal of General Virology,** v. 98, p. 1552-1562, 2017a.

Mar, T.B.; Xavier, C.a.D.; Lima, A.T.M.; Nogueira, A.M.; Silva, J.C.F.; Ramos-Sobrinho, R.; Lau, D.; Zerbini, F.M. Genetic variability and population structure of the New World begomovirus *Euphorbia yellow mosaic virus*. **Journal of General Virology,** v. 98, p. 1537-1551, 2017b.

Maree, H.J.; Fox, A.; Al Rwahnih, M.; Boonham, N.; Candresse, T. Application of HTS for routine plant virus diagnostics: state of the art and challenges. **Frontiers in Plant Science,** v. 9, p. 1082, 2018.

Matyis, J.C.; Silva, D.M.; Oliveira, A.R.; Costa, A.S. Purificação e morfologia do vírus do mosaico dourado do tomateiro. **Summa Phytopathologica,** v. 1, p. 267-275, 1975.

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

Melo, P.C.T. **Mosca branca ameaça produção de hortaliças**. Campinas, SP, Brazil: Asgrow do Brasil Sementes Ltda., Technical Bulletin 1992.

Menten, J.O.; Roston, A.J. Mosaico dourado e a necessidade da regionalização da cultura do feijoeiro no Estado de São Paulo. **Revista da Agricultura,** v. 55, p. 287-300, 1980.

Monci, F.; Sanchez-Campos, S.; Navas-Castillo, J.; Moriones, E. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and *Tomato yellow leaf curl virus* exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. **Virology,** v. 303, p. 317-26, 2002.

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

Navas-Castillo, J.; Fiallo-Olivé, E.; Sánchez-Campos, S. Emerging virus diseases transmitted by whiteflies. **Annual Review of Phytopathology,** v. 49, p. 219-248, 2011.

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

Orozco, B.M.; Miller, A.B.; Settlage, S.B.; Hanley-Bowdoin, L. Functional domains of a geminivirus replication protein. **Journal of Biological Chemistry,** v. 272, p. 9840-6, 1997.

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

Paprotka, T.; Metzler, V.; Jeske, H. The first DNA 1-like alphasatellites in association with New World begomoviruses in natural infections. **Virology,** v. 404, p. 148-157, 2010.

Passos, L.S.; Rodrigues, J.S.; Soares, É.C.S.; Silva, J.P.; Zerbini, F.M.; Araújo, A.S.F.; Beserra, J.E.A. Complete genome sequence of a new bipartite begomovirus infecting *Macroptilium lathyroides* in Brazil. **Archives of Virology,** v. 162, p. 3551-3554, 2017a.

Passos, L.S.; Teixeira, J.W.M.; Teixeira, K.J.M.L.; Xavier, C.a.D.; Zerbini, F.M.; Araújo, A.S.F.; Beserra, J.E.A. Two new begomoviruses that infect non-cultivated malvaceae in Brazil. **Archives of Virology,** v. 162, p. 1795-1797, 2017b.

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

Pedersen, T.J.; Hanley-Bowdoin. Molecular characterization of the AL3 protein encoded by a bipartite geminivirus. **Virology,** v. 202, p. 1070-1075, 1994.

Pinto, V.B.; Silva, J.P.; Fiallo-Olivé, E.; Navas-Castillo, J.; Zerbini, F.M. Novel begomoviruses recovered from *Pavonia* sp. in Brazil. **Archives of Virology,** v. 161, p. 735-739, 2016.

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

Quadros, A.F.F.; Silva, J.P.; Xavier, C.a.D.; Zerbini, F.M.; Boari, A.J. Two new begomoviruses infecting tomato and Hibiscus sp. in the Amazon region of Brazil. **Archives of virology**, p. 1-5, 2019.

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

Ribeiro, S.G.; Ambrozevicius, L.P.; Ávila, A.C.; Bezerra, I.C.; Calegario, R.F.; Fernandes, J.J.; Lima, M.F.; Mello, R.N.; Rocha, H.; Zerbini, F.M. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. **Archives of Virology,** v. 148, p. 281-295, 2003.

Ribeiro, S.G.; Ávila, A.C.; Bezerra, I.C.; Fernandes, J.J.; Faria, J.C.; Lima, M.F.; Gilbertson, R.L.; Zambolim, E.M.; Zerbini, F.M. Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. **Plant Disease,** v. 82, p. 830, 1998.

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

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

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

Rojas, M.R.; Macedo, M.A.; Maliano, M.R.; Soto-Aguilar, M.; Souza, J.O.; Gilbertson, R.L.; Briddon, R.W.; Kenyon, L.A.; Bustamante, R.F.R.; Zerbini, F.M.; Adkins, S.; Legg, J.P.; Kvarnheden, A.; Wintermantel, W.M.; Sudarshana, M.R.; Peterschmitt, M.; Lapidot, M.; Martin, D.P.; Moriones, E.; Inoue-Nagata, A.K. World management of geminiviruses. **Annual Review of Phytopathology,** v. 56, p. 637-677, 2018.

Romay, G.; Geraud-Pouey, F.; Chirinos, D.T.; Mahillon, M.; Gillis, A.; Mahillon, J.; Bragard, C. Tomato twisted leaf virus: a novel indigenous New World monopartite begomovirus infecting tomato in Venezuela. **Viruses,** v. 11, p. 327, 2019.

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

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

Sande, O.F.L. **Estrutura genética de populações do begomovírus *Tomato severe rugose virus* (ToSRV)**. 2014. 44 Dissertação M.S. (Mestrado em Fitopatologia). Dep. de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG.

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

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

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

Sunter, G.; Hartitz, M.D.; Hormuzdi, S.G.; Brough, C.L.; Bisaro, D.M. Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for eficient DNA replication. **Virology,** v. 179, p. 69-77, 1990.

Tavares, S.S.; Ramos-Sobrinho, R.; Gonzalez-Aguilera, J.; Lima, G.S.A.; Assunção, I.P.; Zerbini, F.M. Further molecular characterization of weed-associated begomoviruses in Brazil with an emphasis on *Sida* spp. **Planta Daninha,** v. 30, p. 305-315, 2012.

Vanitharani, R.; Chellappan, P.; Pita, J.S.; Fauquet, C.M. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. **Journal of Virology,** v. 78, p. 9487-98, 2004.

Varsani, A.; Roumagnac, P.; Fuchs, M.; Navas-Castillo, J.; Moriones, E.; Idris, A.; Briddon, R.W.; Rivera-Bustamante, R.F.; Zerbini, F.M.; Martin, D.P. *Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. **Archives of Virology,** v. 162, p. 1819-1831, 2017.

Voinnet, O.; Pinto, Y.M.; Baulcombe, D.C. Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. **Proceedings of the National Academy of Sciences, USA,** v. 96, p. 14147-14152, 1999.

Vu, S.; Melgarejo, T.; Chen, L.-F.; Souza, J.; Macedo, M.; Inoue-Nagata, A.K.; Gilbertson, R.L. Evidence that *Tomato mottle leaf curl virus* from Northeastern Brazil is an indigenous New World monopartite begomovirus. **Phytopathology,** v. 105 (Suppl. 4), p. S4.143, 2015.

Wang, H.; Buckley, K.J.; Yang, X.; Buchmann, R.C.; Bisaro, D.M. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. **Journal of Virology,** v. 79, p. 7410-7418, 2005.

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

Zerbini, F.M.; Briddon, R.W.; Idris, A.; Martin, D.P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Varsani, A.; Ictv Consortium. ICTV Virus Taxonomy Profile: *Geminiviridae*. **Journal of General Virology,** v. 98, p. 131-133, 2017.

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

Zhou, X.; Xie, Y.; Tao, X.; Zhang, Z.; Li, Z.; Fauquet, C.M. Characterization of DNAbeta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. **Journal of General Virology,** v. 84, p. 237-47, 2003.

COMPOSITION OF BEGOMOVIRUS POPULATIONS IN CULTIVATED AND NON-CULTIVATED HOSTS DETERMINED BY HIGH-THROUGHPUT SEQUENCING

Quadros AFF, Ferro CG, Godinho MT, Xavier CAD, Rezende RR, Nogueira AM, Zerbini FM. Composition of begomovirus populations in cultivated and non-cultivated hosts determined by high-throughput sequencing. Virology Journal, *in preparation*.

Composition of begomovirus populations in cultivated and non-cultivated hosts determined by high-throughput sequencing

Ayane F. F. Quadros[1], Camila G. Ferro[1], Márcio T. Godinho[1], César A. D. Xavier[1], Rafael R. de Rezende[2], Angélica M. Nogueira[1], P. Alfenas-Zerbini[2]; F. Murilo Zerbini[1]*

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

[2]Dep. de Microbiologia, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil

Ayane F. F. Quadros and Camila G. Ferro contributed equally to this article.

*Corresponding author: Francisco Murilo Zerbini

Phone: (+55-31) 3616-2423; E-mail: zerbini@ufv.br

Abstract

Begomoviruses (single-stranded DNA viruses transmitted by whiteflies) are economically important plant pathogens causing epidemics worldwide. Tomato-infecting begomoviruses emerged in Brazil in the 1990's following the introduction of *Bemisia tabaci* Middle East-Asia Minor 1. It is believed that these viruses evolved from indigenous viruses infecting non-cultivated hosts. However, tomato-infecting viruses are rarely found in non-cultivated hosts, and vice-versa. It is possible that viral populations are composed primarily of viruses which are well adapted to the host, but also include a small proportion of poorly adapted viruses. Following transfer to a different host, the viral population could shift rapidly, with the viruses which are better adapted to the new host predominating. To test this hypothesis, we collected tomato and *Sida* plants growing next to each other at two locations in 2014 and 2018. Total DNA from one tomato and one *Sida* sample from each location and year were sequenced using HTS. Reads were mapped following a highly stringent set of criteria. For the 2014 samples, >98% of the *Sida* reads mapped to SiMMV, but 0.01% of the reads mapped to ToSRV. Conversely, >99% of the tomato reads mapped to ToSRV, with 0.001% mapping to SiMMV. For the 2018 samples, 40 % of the *Sida* reads mapped to three *Sida*-infecting viruses, and 0.1% of the reads mapped to ToSRV. These results are consistent with the hypothesis that viral populations are composed primarily of the virus that is most adapted to the host but also include a small proportion of viruses that are poorly adapted.

Introduction

Geminiviruses (family *Geminiviridae*) have genomes comprised of one or two molecules of circular, single strand DNA (ssDNA) encapsidated by a single structural protein into twinned quasi-icosahedral particles. The family is composed by the genera *Begomovirus*, *Becurtovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus*, defined based on the type of insect vector, host range, genomic organization and phylogenetic relationships (Brown et al., 2012; Brown et al., 2015; Varsani et al., 2017; Zerbini et al., 2017). The genus *Begomovirus* includes viruses transmitted by whiteflies of the *Bemisia tabaci* cryptic species complex to dicotyledonous plants. Begomoviruses are widespread in all tropical and subtropical regions of the world, and cause severe diseases in a number of economically relevant crops (Briddon, 2003; Inoue-Nagata et

al., 2016; Legg and Fauquet, 2004; Monci et al., 2002; Morales, 2006; Navas-Castillo et al., 2011; Polston and Anderson, 1997; Rocha et al., 2013; Rojas et al., 2018).

Based on phylogenetic relationships and genome characteristics, begomoviruses can be divided into two major groups: Old World (OW; Europe, Africa and Asia) and New World (NW; the Americas) (Padidam et al., 1999; Paximadis et al., 1999; Rybicki, 1994). The majority of NW begomoviruses are bipartite, with two genomic components named DNA-A and DNA-B. The DNA-A contains genes involved in replication, transcriptional control, suppression of host defenses and encapsidation of the viral progeny. The DNA-B contains genes required for intra- and intercellular movement in the plant and suppression of host defenses (Hanley-Bowdoin et al., 2013; Rojas et al., 2005)

Reports of the begomovirus detection in Brazil date back to the 1950's (Costa, 1955; Costa and Bennett, 1950). Begomoviruses are limiting factors for common bean and tomato production (Faria et al., 2000; Zerbini et al., 2005). Tomato-infecting begomoviruses emerged in the country in the mid-1990's following the introduction of *Bemisia tabaci* Middle East-Asia Minor 1 (MEAM1) (Ribeiro et al., 2003; Ribeiro et al., 1998; Rocha et al., 2013). This vector drove the emergence of begomoviruses in Brazil and in many other regions of the world (Gilbertson et al., 2015).

Non-cultivated species of many families have been reported as hosts of begomoviruses in several countries in the Americas (Amarakoon et al., 2008; Assunção et al., 2006; Barbosa et al., 2009; Barreto et al., 2013; Castillo-Urquiza et al., 2008; Faria and Maxwell, 1999; Fernandes et al., 1999; Ferro et al., 2017a; Ferro et al., 2017b; Fiallo-Olivé et al., 2010; Fiallo-Olivé et al., 2012; Frischmuth et al., 1997; Idris et al., 2003; Jovel et al., 2004; Pinto et al., 2016; Roye et al., 1997; Silva et al., 2012; Silva et al., 2011; Tavares et al., 2012). Although all available evidence indicates that tomato-infecting begomoviruses in Brazil originated from indigenous viruses in non-cultivated hosts, the viruses that predominate in tomato are only rarely found in non-cultivated hosts, and vice-versa (Barreto et al., 2013; Calegario, 2004; Castillo-Urquiza et al., 2007; Cotrim et al., 2007; Ferro et al., 2017a; Rocha et al., 2013). This could be explained by the rapid evolution of such indigenous viruses after they are transferred to a new (cultivated) host by the insect vector. This rapid evolution, driven by the high nucleotide substitution rates of begomoviruses (Duffy and Holmes, 2008; Duffy and Holmes, 2009) and their propensity to recombine and reassort (Lefeuvre and Moriones, 2015), would originate the new species detected in crops. The predominance of certain viruses in tomato

would be due to differences in adaptation to this host, or differences in efficiency of transmission by the insect vector.

However, the absence of an evolutionary relationship between the begomoviruses found in non-cultivated and cultivated plants, demonstrated in several in-depth studies (Mar et al., 2017; Ramos-Sobrinho et al., 2014; Rocha et al., 2013; Sande, 2014; Silva et al., 2012), does not support this hypothesis for the emergence of these viruses.

As an alternative hypothesis, it is possible that begomovirus populations in non-cultivated hosts are a mixture of viruses which are well adapted and poorly adapted, present at high and low concentrations, respectively. Following transfer to a different host by the insect vector, the composition of the viral population would shift rapidly, with the viruses that are better adapted to the new host becoming predominant. The use of high-throughput sequencing technologies (HTS) should allow the detection of viruses which are present at very low concentrations in the host.

The objective of this work was to test the hypothesis that viral populations in a given host are composed primarily of the virus that is best adapted to this host but also includes a very small proportion of viruses that are poorly adapted.

Material and Methods

Sample collection

In 2014, 20 tomato (*Solanum lycopersicum*) and 8 *Sida* sp. plants growing next to each other were collected in Coimbra (S20°51′27.720″, W42°50′7.440″) and 15 tomato plants and 8 *Sida* sp. plants, as well as whiteflies, were collected in Florestal (S19º 53' 22", W44º 25' 57"), both in Minas Gerais state, Brazil. In 2018, 19 tomato and 9 *Sida* sp. plants growing side by side, and whiteflies, were collected in Coimbra (Figure 1).

Viral detection

Total DNA from all collected plants was extracted (Doyle and Doyle, 1987). To verify the presence of begomovirus(es), the DNA was used as a template for rolling-circle amplification (Inoue-Nagata et al., 2004) followed by digestion with restriction enzyme MspI to select samples with digestion pattern compatible with a single infection (2014 samples), or PCR with virus-specific primers (2018 samples). Specific primers (Table 1) were designed based on the DNA-A sequences of *tomato severe rugose virus* (ToSRV; GenBank accession

number KC004070) and *sida micrantha mosaic virus* (SiMMV; KC706537), both viruses were identified and cloned into samples collected in 2014. The primers were designed based on genomic regions that would allow only one of the viruses to be applied. Universal primers PAL1v1978 e PAR1c496 (Rojas et al., 1993) and species-specific primers were used to confirm that the samples were infected with a begomovirus and only one virus, respectively. RCA digestion with restriction enzyme MspI (Promega) were performed as per manufacturer recommendations. PCR reactions were performed using GoTaq DNA Polymerase (Promega) in a final volume of 25 µL. The PCR parameters consisted of an initial denaturing step at 95°C for 5 min, followed by 34 cycles at 94°C for 1 min, 60°C for 1 min and 72°C for 1.5 min, followed by a final extension at 72°C for 5 min. RCA and PCR products were separated by 1% agarose gel electrophoresis and stained with ethidium bromide.

Sequencing

Sida sp. plant in which only SiMMV was detected, and tomato plant in which only ToSRV was detected, collected at each location and year, were selected for sequencing. DNA of whitefly pool collected at each site was sent for sequencing (data not shown). DNA extractions were performed using the DNeasy Plant Mini kit (Qiagen), except for the DNA of Sida sp. collected in 2014 extracted with Doyle and Doyle (1987). The extracted DNA was purified using the DNA Clean & Concentrator-5 kit (Zymo Research), which retains DNA molecules shorter than 23 kbp. The resulting DNA, obtained without any amplification steps, was used for high-throughput sequencing (HTS). DNA from the 2014 samples was sequenced in the Illumina HiSeq 2000 platform, and DNA from the 2018 samples was sequenced in the Illumina NovaSeq 6000 platform, both at Macrogen Inc. (Seoul, South Korea).

Quality control and filtering

Sequenced libraries were received for data processing and bioinformatics. Reads were paired-end with a size of 101 bases each (2014 samples) and 151 bases each (2018 samples). Quality control was initially performed with FastQC v. 0.11.5 (Andrews, 2010), to check the adapters and the quality distribution of the bases without any prior computer processing of the libraries. Reads were trimmed using with Trimmomatic v. 0.36 (Bolger et al., 2014) and were filtered for a quality score greater than 30, corresponding to an accuracy of 99.99%, and thus obtaining high homogeneity in all reads. The first ten to fifteen nucleotides were trimmed to prevent bias.

Matching reads to begomovirus sequences

Initially, the identity of the reads was determined by Blast*n* against all New World begomovirus sequences in GenBank Taxonomy Browser (www.ncbi.nlm.nih.gov). The *E*-value cutoff used was 1.10^{-10}, an extremely strict criterion, to minimize false positives in the begomovirus reads. For each read, the first three hits were recorded and three files were prepared, containing the reads for which the first three hits matched to (i) ToSRV, (ii) SiMMV (2014 samples) or *Sida common mosaic virus* (SiCmMV; 2018 samples) and (iii) any other begomovirus, without distinction between DNA-A or DNA-B. Reads that had the first three hits matching to more than one virus were placed in the two respective files (ToSRV and SiMMV/SiCmMV; ToSRV and begomovirus; SiMMV/SiCmMV and begomovirus), or in all three files. Venn diagrams were generated to represent the number of viral reads placed in each file.

Mapping reads to a begomovirus reference genome

To confirm the viral identity of the reads and to prove that the reads actually found belonged to the detected viruses, the reads were mapped to reference sequences for ToSRV (BR:Coi70.2:13 for the samples from Coimbra, and KC004070 for the sample from Florestal), SiMMV (KC706537 for the samples from Coimbra, and KC706532 for the sample from Florestal) and SiCmMV (KC706531, for the samples from Coimbra). Reads were mapped using Bowtie v. 2.3.5.1 (Langmead and Salzberg, 2012), using very-sensitive-local alignment (Same as: -D 20 -R 3 -N 0 -L 20 -i S,1,0.50). The mapped reads were viewed using the Tablet program (Milne et al., 2009).

Recombination analysis

Recombination analysis was performed with Recombination Detection Program (RDP) v.4.5.1 (Martin et al., 2015) using default settings and a Bonferroni-corrected *P*-value cutoff of 0.05. Sequences of ToSRV and SIMMV (2014 samples) or SiCmMV (2018 samples), used in the mapping, were selected for the recombination analysis. Only those recombination events detected by more than four of the seven tests implemented in RDP were considered to be reliable.

Results

Virus identification

Samples collected in 2014 (28 from Coimbra, 23 from Florestal) were initially screened for the presence of begomoviruses using RCA. All RCA-positive tomato samples (nine from Coimbra, 10 from Florestal) had an *Msp*I cleavage pattern indicative of single infection by *Tomato severe rugose virus* (ToSRV). Samples C14T (from Coimbra) and F14T (from Florestal) were selected for HTS. RCA-positive *Sida* samples (six from Coimbra and four from Florestal) displayed different *Msp*I cleavage patterns, with evidence of single and mixed infections. Samples C14S and F14S (from Coimbra and Florestal, respectively), which displayed the pattern indicative of single infection by *Sida micrantha mosaic virus* (SiMMV), were selected for HTS.

Samples collected in 2018 PCR-positive with the universal primers (six from tomato and four from *Sida*) were further analyzed with specific primers (Suppl. Figure S1A). A positive result was obtained from the tomato sample C18T using the ToSRV-specific primers, but not with the SiMMV-specific primers, were selected (Suppl. Figure S1B). The fragment was sequenced, and the sequence obtained showed 100% identity with ToSRV DNA-A. The *Sida* sample C18S was positive using the SiMMV DNA-A-specific primers, but not with the ToSRV DNA-A-specific primers, were selected (Suppl. Figure S1C). The fragment was sequenced, and the sequence obtained showed 84% of identity with SiMMV and 99% with SiCmMV. These two samples were selected for HTS.

Matching reads to begomovirus sequences

In the tomato plant collected in Coimbra in 2014 (C14T), 31,902 reads matched to the ToSRV genome (*ie*, they had ToRSV isolates as the first three hits in the Blast*n* analysis), while only 40 reads matched to the SiMMV genome and 13 reads matched to other begomoviruses (Figure 2A). A total of 372 reads matched to ToSRV or other begomoviruses (*ie*, the first three hits in the Blast*n* analysis included isolates of ToSRV or any begomovirus other than SiMMV), and four reads matched to SiMMV or other begomoviruses. Reads that matched to both ToSRV and SiMMV, or to ToSRV, SiMMV and any other begomovirus were not found (Figure 2A.

In the tomato sample collected at the same location in 2018 (C18T), no reads that matched to SiMMV or SiCmMV were detected, but 1,518 reads were detected that matched to ToSRV and other begomoviruses (Figure 2B). Therefore, it is possible to conclude that SiMMV or SiCmMV were not present in this plant. We detected 9,794 reads matching to ToSRV and

three reads that matched to other begomoviruses. No reads matched to ToSRV, SiCmMV and other begomoviruses (Figure 2B).

In the tomato plant collected at Florestal in 2014 (F14T), the Blast*n* analyses matched 39 reads to SiMMV, 21,689 reads to ToSRV, and four reads to other begomoviruses (Figure 2C). A total of 3,455 reads matched to ToSRV and other begomoviruses, and 83 reads matched to ToSRV or SiMMV. No reads matched to SiMMV or other begomoviruses, or to ToSRV, SiMMV and other begomoviruses (Figure 2C).

Together, the results clearly indicate that the viral population in tomato plants consisted overwhelmingly of ToSRV (>90% of the reads matched unambiguously to this virus), but also included a small contingent (9.9%) of individuals that belong to other begomovirus species, some of which (0.1%) could be classified as SiMMV (Suppl. Figure S2).

In *Sida* plants, as expected, we found a predominance of reads matching viruses commonly detected in this host. The *Sida* plant collected at Coimbra in 2014 (C14S) had only eight reads that matched to ToSRV, while 34,438 reads matched to SiMMV and 206 matched to other begomoviruses (Figure 3A). We also detected 154 reads that could be from ToSRV or SiMMV, seven reads that could be from ToSRV or other begomoviruses, 2,784 reads that could be from SiMMV or other begomoviruses, and four reads that matched to ToSRV, SiMMV and other begomoviruses (Figure 3A).

In the *Sida* sample collected in Coimbra at 2018 (C18S), 3,127 reads matched to the SiCmMV genome, seven reads matched to the ToSRV genome, and 4,464 matched to other begomoviruses (Figure 3B). In this sample, 40 reads matched to the SiMMV genome. We also detected 186 reads that matched to SiCmMV and ToSRV, 108 reads that matched to ToSRV and other begomoviruses, and 2,784 reads that matched to SiCmMV and other begomoviruses. No reads matched to ToSRV, SiCmMV and other begomoviruses (Figure 3B). While in the *Sida* sample only 0.1% of the reads matched to ToSRV, 40% of the reads matched to SiCmMV and the other reads (58%) matched to others begomovirus.

In the *Sida* sample collected in Florestal at 2014 (F14S), three reads matched to ToSRV, 29,059 reads matched to SiMMV, and 481 reads matched to other begomoviruses (Figure 2C). Eight reads matched to ToSRV and SiMMV, two reads matched to ToSRV and other begomoviruses, and 2,351 reads matched to SiMMV and other begomoviruses. One read matched to ToSRV, SiMMV and other begomoviruses (Figure 3C). Together, the results obtained from *Sida* plants clearly indicate that the viral population consisted predominantly of viruses normally found in these plants (>87% of the reads matched unambiguously to either

SiMMV or SiCmMV), but (as in tomato) also included a small contingent of individuals that belong to other begomovirus species, some of which (0.1%) could be classified as ToSRV (Suppl. Figure S2).

Mapping reads to begomovirus reference genomes

In the tomato sample C14T, 23,248 reads (mismatches %=0.7) and 46,548 reads (mismatches %=1.0) mapped to the reference DNA-A and DNA-B ToSRV genomes, respectively. In the tomato sample C18T, 11,997 (mismatches %=1.6) and 12,777 (mismatches %=3.2) reads mapped to the reference DNA-A and DNA-B ToSRV genomes, respectively. In the tomato sample F14T, 16,795 (mismatches %=2.2) reads and 36,342 (mismatches %=4.2) reads mapped to the reference DNA-A and DNA-B ToSRV genomes, respectively Genome coverage for the DNA-A was complete for samples C14T, C18T and F14T (Figure 2D; Suppl. Figure S3).

In the *Sida* sample C14S, 7,718 (mismatches %=10.6) and 2,780 (mismatches %=11.5), respectively, mapped to the reference DNA-A and DNA-B ToSRV genomes. In the *Sida* sample C18S, 2,304 (mismatches %=14.5) mapped to the reference DNA-A ToSRV genome, and 513 reads (mismatches %=12.6) mapped to the ToSRV DNA-B genome. In the sample F14S, 4,423 (mismatches %=11.2) and 1702 (mismatches %=12.2) reads, respectively, mapped to the reference DNA-A and DNA-B ToSRV genomes. Genome coverage for the DNA-A was 84% for samples the sample C14T, and 82% for the sample C18S and 74% for the sample F14T (Figure 2D; Suppl. Figure S3).

Using SiMMV as reference, 39,266 (mismatches %=2.2) and 37,663 reads (mismatches %=4.6) from the *Sida* sample C14S mapped to the DNA-A and DNA-B genomes respectively. In the *Sida* sample F14S, 17,994 reads (mismatches %=2.4) and 44,459 reads (mismatches %=5.5) mapped to the SiMMV DNA-A and DNA-B, respectively (Figure 3D; Suppl. Figure S4). Genome coverage for the DNA-A was complete for sample C14S and F14S. Genome coverage for the DNA-B was complete for the sample C14S and aprox. 90.44% for the sample F14S (Figure 3D; Suppl. Figure S3). In the *Sida* sample C18S 3,192 reads (mismatches %=10) mapped to SiMMV DNA-A, and 1,074 reads (mismatches %=14.1) mapped to the DNA-B (data not shown). When the reads from this sample were mapped to a reference SiCmMV genome we observed coverage of 97%, with 7,284 reads (mismatches %=3.4) mapping to the DNA-A (Figure 3D; Suppl. Figure S4). No mapping was performed for the DNA-B, since there are no SiCmMV DNA-B sequences available.

In the tomato sample C14T, 6,498 (mismatches %=10.5) and 6,559 (mismatches %=10.9) reads mapped to SiMMV DNA-A and DNA-B, respectively. In the tomato sample F14T, 4,093 reads (mismatches %=10.1) mapped to the SiMMV DNA-A reference genome, and 4,614 reads (mismatches %=10.7) mapped to the DNA-B. Genome coverage for the DNA-A was 94% for sample C14T and 68% for the sample F14T. Genome coverage for the DNA-B was 82% for the sample C14T and aprox. 60% for the sample F14T. In tomato sample C18T, 4,132 (mismatches %=13.4) and 2,278 (mismatches %=14.4) reads mapped to SiMMV DNA-A and DNA-B, respectively.

Recombination analysis

Recombination analysis was performed to infer if any reads mapped to recombinant regions in the viral genomes, which could interfere with virus identification. Recombination analysis of the four viruses to which the reads matched revealed one recombination event between the DNA-A of SiMMV and SiCmMV (Table 2A), and one recombination event between the DNA-B of ToSRV and SiMMV (Table 2B). Since no reads mapped to the recombinant regions, it can be concluded that the virus identification was not affected by recombination events that may have occurred between these viruses.

Discussion

In this work, we tested the hypothesis that begomovirus populations in non-cultivated hosts are a mixture of well-adapted and poorly-adapted viruses, with the composition of the viral population shifting rapidly after transfer to a new host, with the viruses that are better adapted to the new host becoming predominant.

Using high-throughput sequencing, tomato and *Sida* samples growing side by side in fields in the municipality of Coimbra (MG) in 2014 showed that in tomato plants 98.7% of the reads matched to ToSRV, but 0.12% to SiMMV, while in 91.6% of the reads from *Sida* matched to SiMMV, but 0.02% matched to ToSRV. For the samples collected in Florestal (MG), also in 2014 in tomato plants 85.3% of the reads matched to ToSRV and 0.15% to SiMMV, while 91.1% of the reads from Sida matched to SiMMV and 0.01% to ToSRV.

Another sampling was carried out in the municipality of Coimbra in 2018, to confirm if the results obtained would be observed after 4 years of the first sampling. The analyses showed that in tomato plants 86.6% of the reads matched with ToSRV and that the remaining 13.4%

matched to a specific virus and it was not possible to identify which virus species was associated. While in the *Sida* sample only 0.1% of the reads matched to ToSRV, 40% of the reads matched to SiCmMV, and the other reads (58%) matched to others begomovirus. This result shows the variation of the viral population that can occur in non-cultivated plants, but a small amount of reads from the less adapted virus is still observed.

Begomoviruses originally found in weeds can be found infecting tomato crop under field conditions (Calegario, 2004; Castillo-Urquiza et al., 2010; Castillo-Urquiza et al., 2007; Cotrim et al., 2007; Jovel et al., 2004; Sande, 2014). From this observation (and also due to the fact that the tomato-infecting begomoviruses found in Brazil have never been detected elsewhere) it was hypothesized that the tomato-infecting begomoviruses evolved from the viruses present in non-cultivated plants (Castillo-Urquiza et al., 2007; Ribeiro et al., 2003). However, the analysis of begomovirus populations found in cultivated and non-cultivated plants indicated that these viral populations are geographically structured, and did not indicate an ancestral relationship between virus of non-cultivated and cultivated hosts (Lima et al., 2013; Mar et al., 2017; Ramos-Sobrinho et al., 2014; Rocha et al., 2013; Rodelo-Urrego et al., 2015).

Tavares et al., 2012 characterized the diversity of begomovirus infecting weeds in Brazil, with a special focus on *Sida* spp., to acess their importance as source of new virus for cultivated plants. Their results suggest that plants of *Sida* spp. could be infected by heterogeneous begomovirus populations with rare variants that would not be easily detected by RCA-based cloning. After the transmission of these heterogeneous populations to tomato plants by the insect vector, the rare variants would become predominant (and therefore would be detected at high frequency) due to better adaptation to the new host. Our results corroborate this idea and reinforces this hypothesis.

Studies with the objective of molecular characterization of begomovirus isolate that infects wild and weedy plants revealed a great diversity of species (Castillo-Urquiza et al., 2008; Silva et al., 2012; Tavares et al., 2012). Although there is little data on the variability of begomovirus populations in non-cultured hosts, currently available data suggest that it is higher than that observed in begomoviruses infecting cultivated plants (Lima et al., 2013; Rocha et al., 2013; Silva et al., 2012; Silva et al., 2011; Wyant et al., 2011). Probably this great diversity present in these populations would be related to the substitution of the predominant virus species in the plant of *Sida* collected in Coimbra.

Geminiviruses were identified from different plant species and sampled in the natural ecosystems around cultivated areas (agroecological interface) in the Northern-Pacific region of

Mexico. The analysis of the reads revealed DNA signatures corresponding to 52 crop-infecting and 35 non-cultivated-infecting geminiviruses which, curiously, were present in different plant species. In our results we observed that that more than 99% of the readings corresponded to the genus *Begomovirus*, including viruses adapted to crop species in different plant families, and the remaining 1% corresponded to the other members of the *Geminiviridae* family. In different plant species, in the five agroecological regions included in this study, were observed quadruple infections with TYLCV (*Tomato yellow leaf curl virus*), PHYVV (*Pepper huasteco yellow vein virus*), SiMSiV (*Sida mosaic Sinaloa virus*), RhGMV (*Rhynchosia golden mosaic virus*) and RhGMSV (*Rhynchosia golden mosaic Sinaloa virus*) with 31.25%. (Rodríguez-Negrete et al., 2019). These results show the great diversity of viruses that can simultaneously infect a single plant, including both cultivated and non-cultivated plant viruses, and that such diversity would probably not be observed using other techniques (such as PCR and RCA). The existence of several plant species harboring different viruses in multiple infections represents an important melting pot for geminiviruses to evolve in different directions triggered by vector and environmental factors. These results also suggest that a number of widely distributed non-cultivated plants in North-Pacific Mexico represent a reservoir for begomoviruses (Rodríguez-Negrete et al., 2019).

The other studies related to identification and characterization of Geminiviruses using high-throughput sequencing, such as (Shakir et al., 2019) and (Rodríguez-Negrete et al., 2019) included an enrichment step for viral sequences through the use of rolling circle amplification. The use of this technique may lead to an amplification bias since viral titer can influence the activity of the phi29 DNA polymerase, and thus may not portray the actual composition of the viral population. Therefore, in our analysis, no amplification steps were used because we expected to observe the actual composition of each virus in the plants. By mapping the reads using the Bowtie program, it was possible to observe that despite the low number of reads that were mapped, the poorly adapted viruses had unique reads, proving that the reads belong to the virus investigated.

A number of studies suggest that the level of genetic diversity seems to be an intrinsic property of the virus, regardless of host variety or coexistence in blending infection with other viruses (Mar et al., 2017; Rocha et al., 2013; Silva et al., 2012). "Sealed container" and "mixing vessels" are two patterns of plant-begomovirus interactions in wild hosts, which play different roles in emergencies. *Sida* spp. they are "mixing vessels" that allow efficient infection with various begomoviruses, possibly similarly adapted to the host, resulting in a high frequency of

mixed infections. The life cycle of these plants suggests a relationship between life expectancy, chance of virus infection and potential as a reservoir for crops, this would be a good topic for future studies. The hosts of "mixing vessels" would be centers of virus transmission and diversification, acting as efficient reservoirs for begomovirus emergence (García-Arenal and Zerbini, 2019).

In agreement with this hypothesis, *Macroptilium yellow spot virus* (MaYSV) has recently emerged in bean crops, and its genetic diversity is similar in the cultivated and in non-cultivated host (Ramos-Sobrinho et al., 2014). However, to date, no begomoviruses have emerged on malvaceous crops (such as cotton) in Brazil, again indicating the difficulty of establishing general models for the emergence of the virus (García-Arenal and Zerbini, 2019).

To act as a good reservoir of viruses, alternative host species must maintain viruses that produce little or no symptoms in that reservoir plant. This allows the alternative host plant to survive and the virus to be transmitted over generations. A second reason could be the complete resistance of these wild species to the viruses of the second epidemic. Alternative hosts may maintain less virulent viruses but may have complete resistance to more virulent species (Shakir et al., 2019). However, the presence of such viruses in non-cultivated plants, as alternative hosts in nature, increases the chances of viral evolution through recombination events or other mechanisms (Rodríguez-Negrete et al., 2019).

Conclusions

Our results are consistent with the hypothesis that both well adapted and poorly adapted viruses are present in different hosts, and that the composition of the viral population shifts after vector-mediated transfer from one host to another.

Biological studies can elucidate under experimental conditions whether the adaptability of these viruses in the host is really responsible for composition of the viral population observed through the use of high-throughput sequencing technologies.

References

Amarakoon, I.I., Roye, M.E., Briddon, R.W., Bedford, I.D., Stanley, J., 2008. Molecular and biological characterization of *Macroptilium yellow mosaic virus* from Jamaica. Plant Pathol 57, 417-426.
Andrews, S., 2010. FastQC: A quality control tool for high throughput sequence data, Reference Source.
Assunção, I.P., Listik, A.F., Barros, M.C.S., Amorim, E.P.R., Silva, S.J.C., Izael, O.S., Ramalho-Neto, C.E., Lima, G.S.A., 2006. Genetic diversity of begomoviruses in invasive species of Northeastern Brazil (in Portuguese). Planta Daninha 24, 239-244.
Barbosa, J.C., Barreto, S.S., Inoue-Nagata, A.K., Reis, M.S., Firmino, A.C., Bergamin, A., Rezende, J.A.M., 2009. Natural infection of *Nicandra physaloides* by *Tomato severe rugose virus* in Brazil. J Gen Plant Pathol 75, 440-443.
Barreto, S.S., Hallwass, M., Aquino, O.M., Inoue-Nagata, A.K., 2013. A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. Phytopathology 103, 436-444.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.
Briddon, R.W., 2003. Cotton leaf curl disease, a multicomponent begomovirus complex. Mol Plant Pathol 4, 427-434.
Brown, J.K., Fauquet, C.M., Briddon, R.W., Zerbini, F.M., Moriones, E., Navas-Castillo, J., 2012. Family *Geminiviridae*, in: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, UK, pp. 351-373.
Brown, J.K., Zerbini, F.M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R., Silva, J.C., Fiallo-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.
Calegario, R.F., 2004. Caracterização do isolado de begomovírus MG-Bi2, um possível membro da espécie *Sida micrantha mosaic virus* (SimMV), Departamento de Fitopatologia. Universidade Federal de Viçosa, Viçosa, MG, p. 48.
Castillo-Urquiza, G.P., Alfenas-Zerbini, P., Beserra-Junior, J.E.A., Mizubuti, E.S.G., Varsani, A., Martin, D.P., Zerbini, F.M., 2010. Genetic structure of tomato-infecting begomovirus populations in two tomato-growing regions of Southeastern Brazil, Program and Abstracts, 6th International Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop, Guanajuato, Mexico.
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.
Castillo-Urquiza, G.P., Beserra Junior, J.E.A., Alfenas-Zerbini, P., Varsani, A., Lima, A.T.M., Barros, D.R., Zerbini, F.M., 2007. Genetic diversity of begomoviruses infecting tomato in Paty do Alferes, Rio de Janeiro state, Brazil. Virus Reviews and Research 12, 233.
Costa, A.S., 1955. Studies on *Abutilon* mosaic in Brazil. Phytopathol Z 24, 97-112.
Costa, A.S., Bennett, C.W., 1950. Whitefly transmitted mosaic of *Euphorbia prunifolia*. Phytopathology 40, 266-283.

Cotrim, M.A., Krause-Sakate, R., Narita, N., Zerbini, F.M., Pavan, M.A., 2007. Genetic diversity ot tomato-infecting begomoviruses in Central São Paulo state (in Portuguese). Summa Phytopathologica 33, 300-303.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochemical Bulletin 19, 11-15.
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., Bezerra, I.C., Zerbini, F.M., Ribeiro, S.G., Lima, M.F., 2000. Current status of geminivirus diseases in Brazil (in Portuguese). Fitopatol Bras 25, 125-137.
Faria, J.C., Maxwell, D.P., 1999. Variability in geminivirus isolates associated with *Phaseolus* spp. in Brazil. Phytopathology 89, 262-268.
Fernandes, A.V., Galvão, R.M., Machado, J.J., Zerbini, F.M., Fontes, E.P.B., 1999. Cloning and molecular characterization of A components of two new *Sida rhombifolia*-infecting geminiviruses. Virus Reviews and Research 4, 148.
Ferro, C.G., Silva, J.P., Xavier, C.A.D., Godinho, M.T., Lima, A.T.M., Mar, T.B., Lau, D., Zerbini, F.M., 2017a. The ever increasing diversity of begomoviruses infecting non-cultivated hosts: new species from *Sida* spp. and *Leonurus sibiricus*, plus two New World alphasatellites. Ann Appl Biol 170, 204-218.
Ferro, M.M.M., Ramos-Sobrinho, R., Silva, J.T., Assunção, I.P., Lima, G.S.A., 2017b. Genetic structure of populations of the begomoviruses Tomato mottle leaf curl virus and Sida mottle Alagoas virus infecting tomato (Solanum lycopersicum) and Sida spp., respectively. Trop Plant Pathol 42, 39-45.
Fiallo-Olivé, E., Martinez-Zubiaur, Y., Moriones, E., Navas-Castillo, J., 2010. Complete nucleotide sequence of *Sida golden mosaic Florida virus* and phylogenetic relationships with other begomoviruses infecting malvaceous weeds in the Caribbean. Arch Virol 155, 1535-1537.
Fiallo-Olivé, E., Navas-Castillo, J., Moriones, E., Martinez-Zubiaur, Y., 2012. Begomoviruses infecting weeds in Cuba: Increased host range and a novel virus infecting *Sida rhombifolia*. 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.
García-Arenal, F., Zerbini, F.M., 2019. Life on the edge: geminiviruses at the interface between crops and wild plant hosts. Annu Rev Virol 6.
Gilbertson, R.L., Batuman, O., Webster, C.G., Adkins, S., 2015. Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. Annu Rev Virol 2, 67-93.
Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D., Mansoor, S., 2013. Geminiviruses: Masters at redirecting and reprogramming plant processes. Nat Rev Microbiol 11, 777-788.
Idris, A.M., Hiebert, E., Bird, J., Brown, J.K., 2003. Two newly described begomoviruses of *Macroptilium lathyroides* and common bean. Phytopathology 93, 774-783.
Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B., Nagata, T., 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. J Virol Methods 116, 209-211.
Inoue-Nagata, A.K., Lima, M.F., Gilbertson, R.L., 2016. A review of geminivirus diseases in vegetables and other crops in Brazil: current status and approaches for management. Hortic Bras 34, 8-18.

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.
Langmead, B., Salzberg, S., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357.
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.
Mar, T.B., Xavier, C.A.D., Lima, A.T.M., Nogueira, A.M., Silva, J.C.F., Ramos-Sobrinho, R., Lau, D., Zerbini, F.M., 2017. Genetic variability and population structure of the New World begomovirus *Euphorbia yellow mosaic virus*. J Gen Virol 98, 1537-1551.
Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol 1, vev003.
Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., Marshall, D., 2009. Tablet—next generation sequence assembly visualization. Bioinformatics 26, 401-402.
Monci, F., Sanchez-Campos, S., Navas-Castillo, J., Moriones, E., 2002. A natural recombinant between the geminiviruses *Tomato yellow leaf curl Sardinia virus* and *Tomato yellow leaf curl virus* exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303, 317-326.
Morales, F.J., 2006. History and current distribution of begomoviruses in Latin America. Adv Virus Res 67, 127-162.
Navas-Castillo, J., Fiallo-Olivé, E., Sánchez-Campos, S., 2011. Emerging virus diseases transmitted by whiteflies. Annu Rev Phytopathol 49, 219-248.
Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218-224.
Paximadis, M., Idris, A.M., Torres-Jerez, I., Villarreal, A., Rey, M.E.C., Brown, J.K., 1999. Characterization of tobacco geminiviruses in the Old and New world. Arch Virol 144, 703-717.
Pinto, V.B., Silva, J.P., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F.M., 2016. Novel begomoviruses recovered from *Pavonia* sp. in Brazil. Arch Virol 161, 735-739.
Polston, J.E., Anderson, P.K., 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. Plant Dis 81, 1358-1369.
Ramos-Sobrinho, R., Xavier, C.A.D., Pereira, H.M.B., Lima, G.S.A., Assunção, I.P., Mizubuti, E.S.G., Duffy, S., Zerbini, F.M., 2014. Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. J Gen Virol 95, 2540-2552.
Ribeiro, S.G., Ambrozevicius, L.P., Ávila, A.C., Bezerra, I.C., Calegario, R.F., Fernandes, J.J., Lima, M.F., Mello, R.N., Rocha, H., Zerbini, F.M., 2003. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. Arch Virol 148, 281-295.
Ribeiro, S.G., Ávila, A.C., Bezerra, I.C., Fernandes, J.J., Faria, J.C., Lima, M.F., Gilbertson, R.L., Zambolim, E.M., Zerbini, F.M., 1998. Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. Plant Dis 82, 830.
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.
Rodelo-Urrego, M., García-Arenal, F., Pagán, I., 2015. The effect of ecosystem biodiversity on virus genetic diversity depends on virus species: A study of chiltepin-infecting begomoviruses in Mexico. Virus Evol 1, vev004.
Rodríguez-Negrete, E.A., Morales-Aguilar, J.J., Domínguez-Duran, G., Torres-Devora, G., Camacho-Beltrán, E., Leyva-López, N.E., Voloudakis, A.E., Bejarano, E.R., Méndez-Lozano, J., 2019. High-throughput sequencing reveals differential begomovirus species diversity in non-cultivated plants in Northern-Pacific Mexico. Viruses 11, 594.
Rojas, A., Kvarnheden, A., Marcenaro, D., Valkonen, J.P.T., 2005. Sequence characterization of *Tomato leaf curl Sinaloa virus* and *Tomato severe leaf curl virus*: Phylogeny of New World begomoviruses and detection of recombination. Arch Virol 150, 1281-1299.
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., Macedo, M.A., Maliano, M.R., Soto-Aguilar, M., Souza, J.O., Briddon, R.W., Kenyon, L.A., Bustamante, R.F.R., Zerbini, F.M., Adkins, S., Legg, J.P., Kvarnheden, A., Wintermantel, W.M., Sudarshana, M.R., Peterschmitt, M., Lapidot, M., Martin, D.P., Moriones, E., Inoue-Nagata, A.K., Gilbertson, R.L., 2018. World management of geminiviruses. Annu Rev Phytopathol 56, 637-677.
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.
Sande, O.F.L., 2014. Estrutura genética de populações do begomovírus *Tomato severe rugose virus* (ToSRV), Dep. de Fitopatologia. Universidade Federal de Viçosa, Viçosa, MG, p. 44.
Shakir, S., Zaidi, S.S.-A., Farooq, M., Amin, I., Scheffler, J., Scheffler, B., Shah Nawaz-ul-Rehman, M., Mansoor, S., 2019. Non-cultivated cotton species (*Gossypium* spp.) act as a reservoir for cotton leaf curl begomoviruses and associated satellites. Plants 8, 127.
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.
Silva, S.J.C., Castillo-Urquiza, G.P., Hora-Júnior, B.T., Assunção, I.P., Lima, G.S.A., Pio-Ribeiro, G., Mizubuti, E.S.G., Zerbini, F.M., 2011. High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed *Cleome affinis* in northeastern Brazil. Arch Virol 156, 2205-2213.
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., Roumagnac, P., Fuchs, M., Navas-Castillo, J., Moriones, E., Idris, A., Briddon, R.W., Rivera-Bustamante, R.F., Zerbini, F.M., Martin, D.P., 2017. *Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. Arch Virol 162, 1819-1831.
Wyant, P.S., Gotthardt, D., Schafer, B., Krenz, B., Jeske, H., 2011. The genomes of four novel begomoviruses and a new *Sida micrantha mosaic virus* strain from Bolivian weeds. Arch Virol 156, 347-352.

Zerbini, F.M., Andrade, E.C., Barros, D.R., Ferreira, S.S., Lima, A.T.M., Alfenas, P.F., Mello, R.N., 2005. Traditional and novel strategies for geminivirus management in Brazil. Australas Plant Path 34, 475-480.
Zerbini, F.M., Briddon, R.W., Idris, A., Martin, D.P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Varsani, A., ICTV Consortium, 2017. ICTV Virus Taxonomy Profile: *Geminiviridae*. J Gen Virol 98, 131-133.

Table 1. Features of the primers used for PCR detection of *Tomato severe rugose virus* (ToSRV), *Sida micrantha mosaic virus* (SiMMV) and *Sida common mosaic virus* (SiCmMV).

Primer		**Sequence (5'-3')**	**Amplicon size**	**Genome position**
ToSRV-AF	Forward	ACTGGGATCCCCTAAGCGTGATGCC C	500 pb	nt 2227
ToSRV-AR	Reverse	GTCAAAGCTTTTAATTAGTAATCGA AT		nt 2872
SiMMV-AF[1]	Forward	CCAAGGTCACTGGTGGTCAA	800 pb	nt 803
SiMMV-AR[1]	Reverse	CACGAGGACCAGAGTGTACC		nt 1302

[1]Also used to amplify SiCmMV

Table 2. Putative recombination events detected among ToSRV and SiMMV, based on the full-length DNA-A **(A)** and DNA-B **(B)**.

A

Event	**Recombinant**	**Recombination Breakpoints**[a]		**Parents**		**Method**[b]	**P-value**[c]
		Begin	End	Minor	Major		
1	EU710751 - *Sida common mosaic virus*	25	64	KC706537 - *Sida micrantha mosaic virus* ToSRV Florestal (BR:Flo165:08)	Unknown	GBMC	7,85x10^{+06}

B

Event	**Recombinant**	**Recombination Breakpoints**[a]		**Parents**		**Method**[b]	**P-value**[c]
		Begin	End	Minor	Major		
1	ToSRV Coimbra (BR:Coi103.1:13)	84	128	KC706532 - *Sida micrantha mosaic virus*	Unknown	R**G**MC	2,94x10-^{01}

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

[b]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.

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

* The actual breakpoint position is undetermined (it was most likely overprinted by a subsequent recombination event).

Figure legends

Figure 1. Tomato (green elipse) and *Sida* sp. (yellow circle) plants growing side by side in the municipality of Coimbra, Minas Gerais in 2018.

Figure 2. Venn diagrams indicating the number of reads in samples C14T (**A),** C18T (**B**) and F14T (**C**) and for which the first three hits in the Blast*n* analysis matched ToSRV, SiMMV/SiCmMV or any other begomovirus. The double intersects indicate the number of reads for which the first three hits included ToSRV or SiMMV/SiCmMV, ToSRV or any begomovirus other than SiMMV/SiCmMV, and SiMMV/SiCmMV or any begomovirus other than ToSRV. The triple intersect indicates the number of reads for which the first three hits included ToSRV, SiMMV/SiCmMV and any other begomovirus. **D.** Mapping of reads from tomato and *Sida* sp. samples against a reference ToSRV DNA-A genome. The reads are reflected by four columns separated by different colors (top image). Coverage of reads against a reference ToSRV DNA-A genome. The number to the left and the color variation in the figure refers to the number of reads mapped at each position, ranges from 0 (white) to highest number of aligned reads (dark blue) (bottom image).

Figure 3. Venn diagrams indicating the number of reads in samples C14S **(A)**, C18S **(B)** and F14S **(C)** and for which the first three hits in the Blast*n* analysis matched ToSRV, SiMMV/SiCmMV or any other begomovirus. The double intersects indicate the number of reads for which the first three hits included ToSRV or SiMMV/SiCmMV, ToSRV or any begomovirus other than SiMMV/SiCmMV, and SiMMV/SiCmMV or any begomovirus other than ToSRV. The triple intersect indicates the number of reads for which the first three hits included ToSRV, SiMMV/SiCmMV and any other begomovirus. **D.** Mapping of reads from Sida sp. and tomato samples against a reference SiMMV/SiCmMV DNA-A genome. The reads

are reflected by four columns separated by different colors (top image). Coverage of reads against a reference SiMMV/SiCmMV DNA-A genome. The number to the left and the color variation in the figure refers to the number of reads mapped at each position, ranges from 0 (white) to highest number of aligned reads (dark blue) (bottom image).

Figure 1.

Figure 2

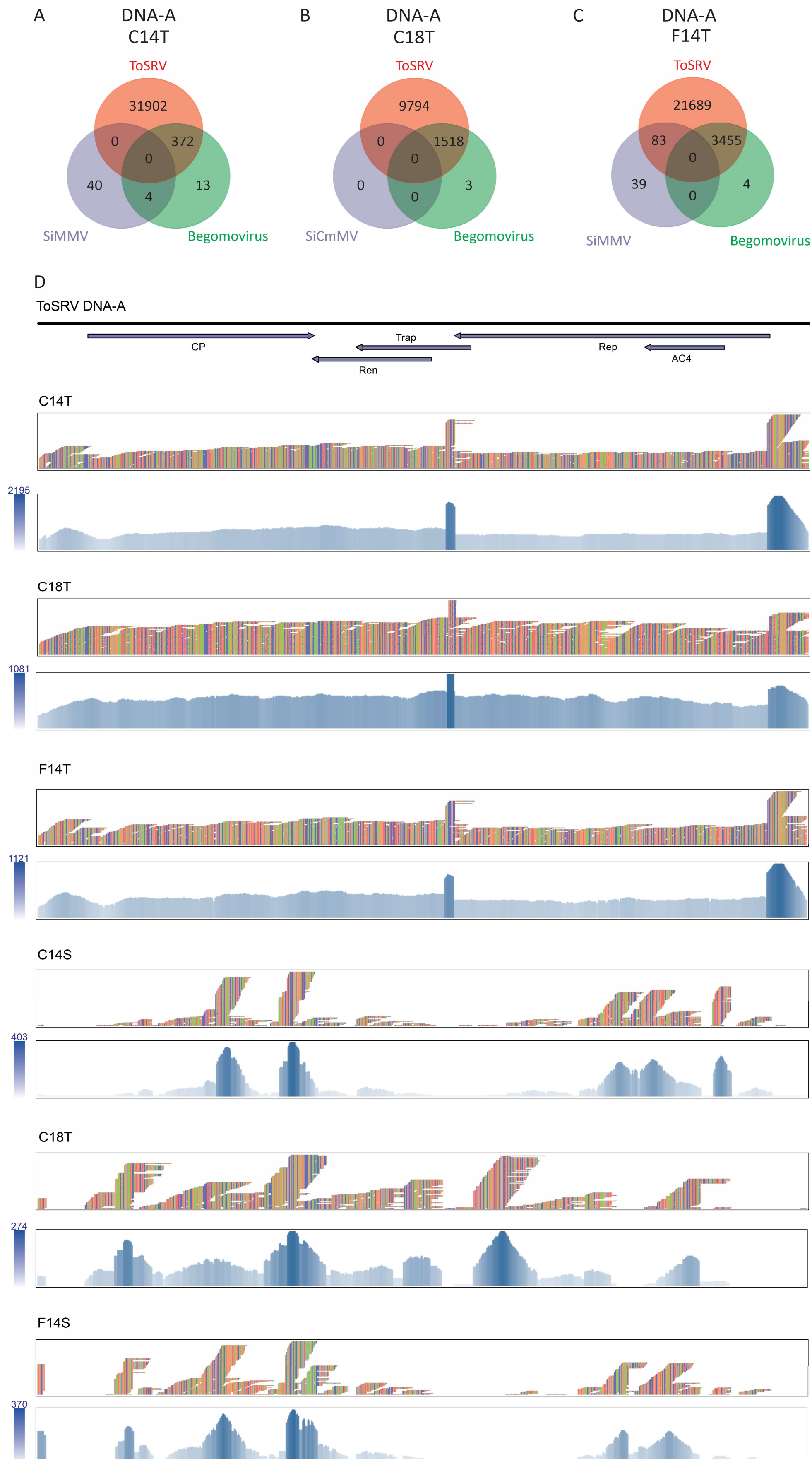
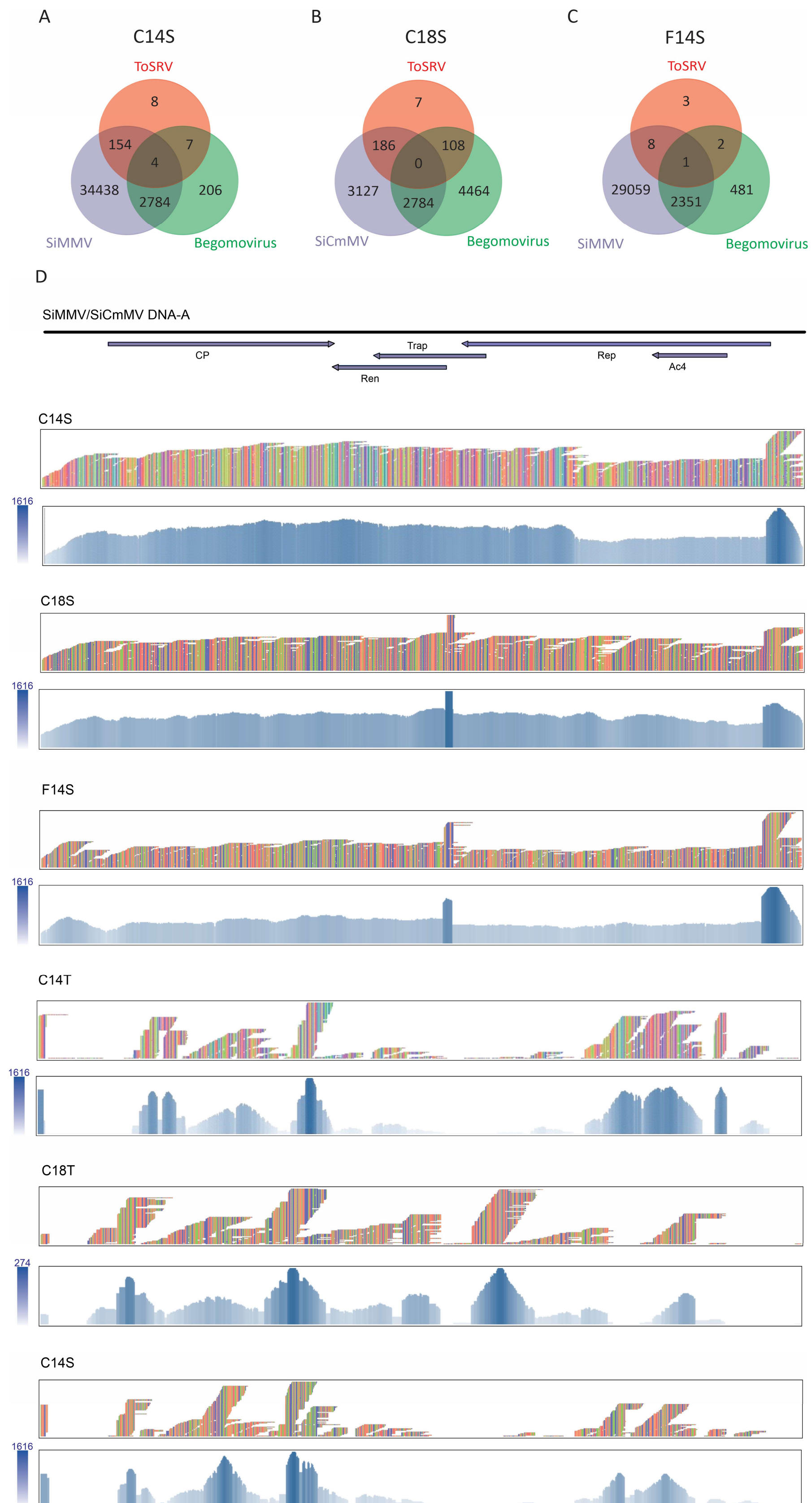
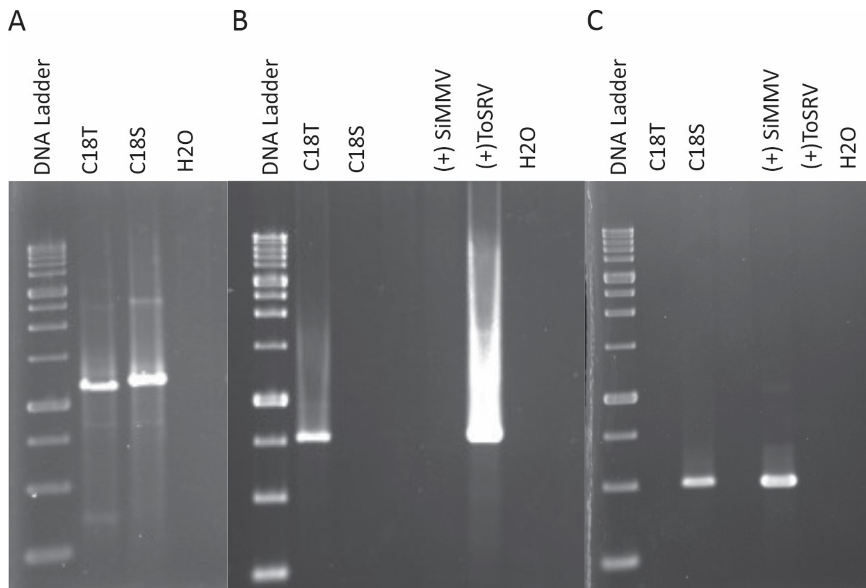


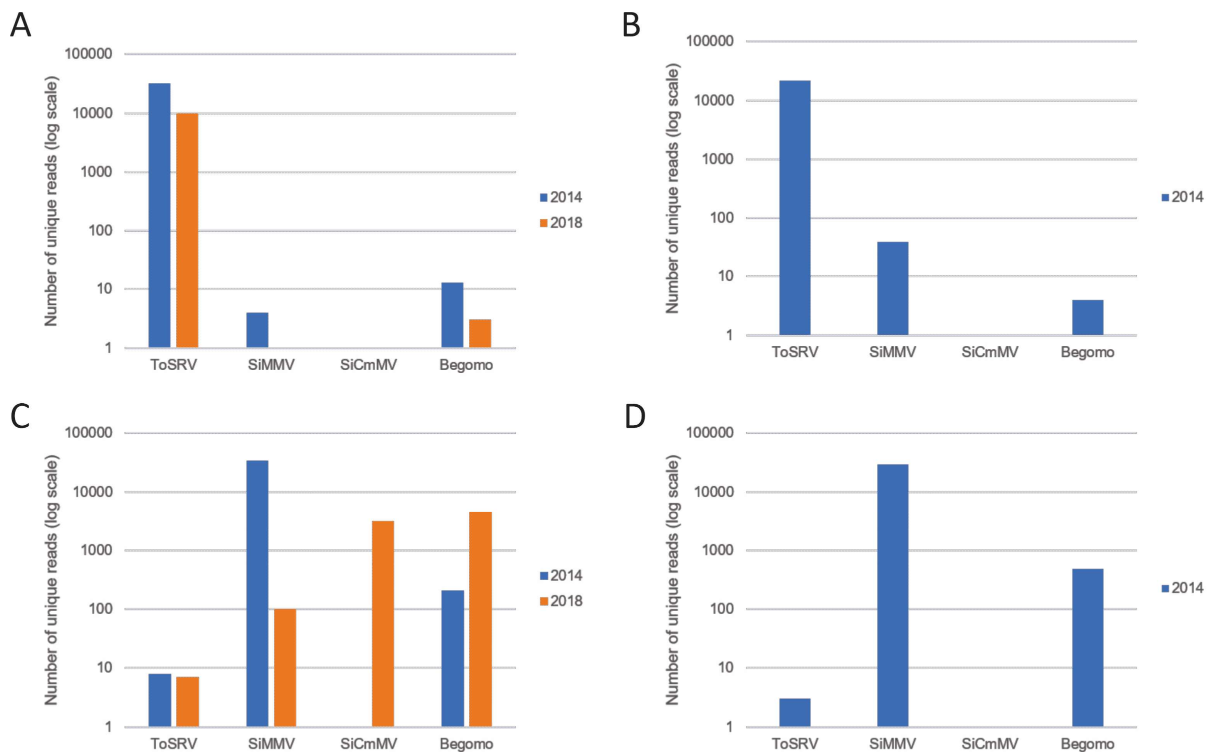
Figure 3



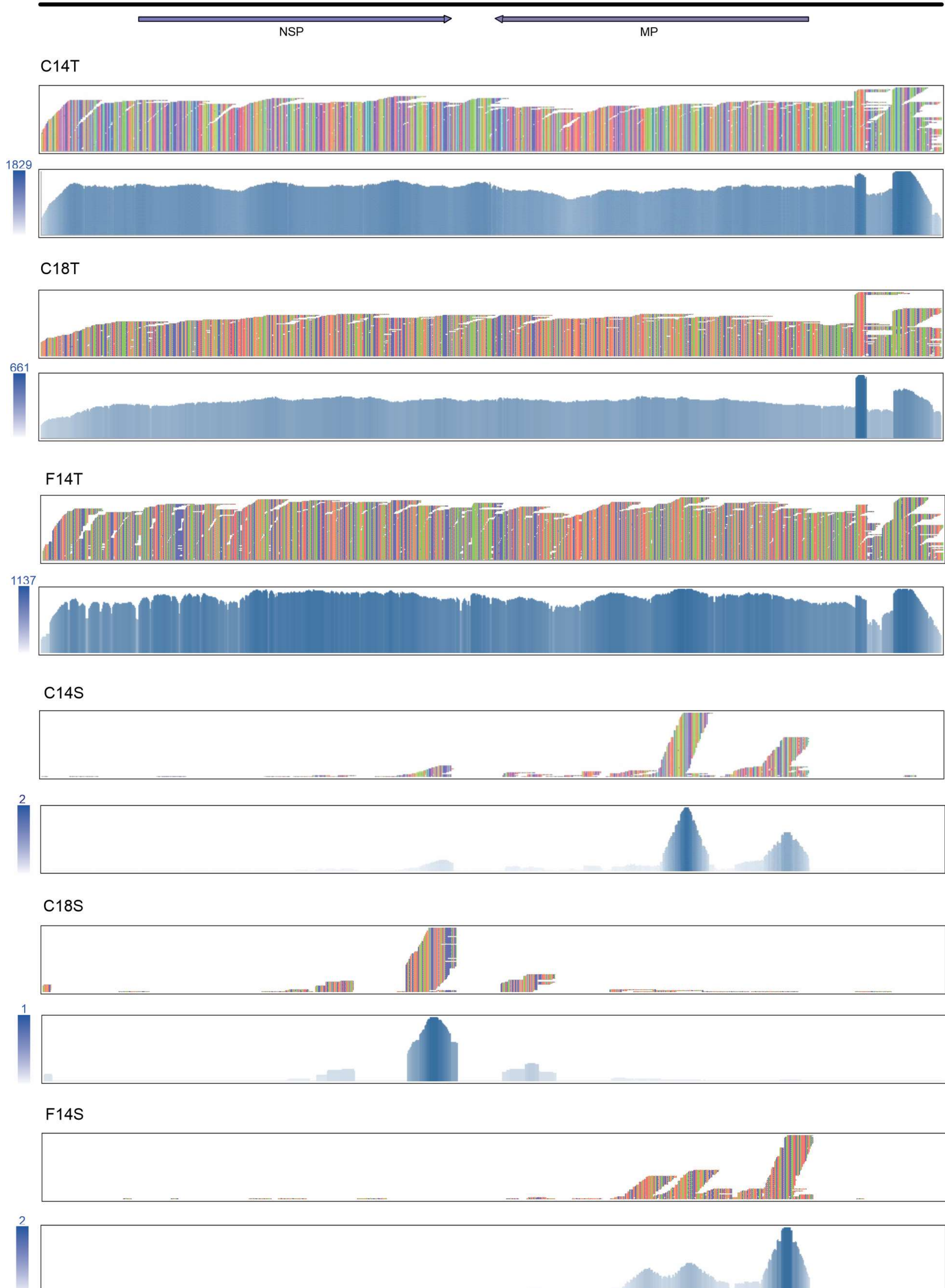
Suppl. Figure S1. A. PCR-positive sample detection using universal primers PAL1v1978 e PAR1c496. **B.** PCR-positive sample detection using specific primers ToSRV-AF/R **C.** PCR-positive sample detection using specific primers SiMMV-F/R. M, size marker (1 kb plus DNA ladder, Kasvi).



Suppl. Figure S2. A. Number of unique reads (log scale) from the tomato samples collected in Coimbra that matched to *Tomato severe rugose virus* (ToSRV), *Sida micrantha mosaic virus* (SiMMV), *Sida common mosaic virus* (SiCmMV), or any other begomovirus. **B.** Number of unique reads (log scale) from the tomato sample collected in Florestal that matched to *Tomato severe rugose virus* (ToSRV), *Sida micrantha mosaic virus* (SiMMV), *Sida common mosaic virus* (SiCmMV), or any other begomovirus. **C.** Number of unique reads (log scale) from the *Sida* samples collected in Coimbra that matched to *Tomato severe rugose virus* (ToSRV), *Sida micrantha mosaic virus* (SiMMV), *Sida common mosaic virus* (SiCmMV), or any other begomovirus. **D.** Number of unique reads (log scale) from the *Sida* sample collected in Florestal that matched to *Tomato severe rugose virus* (ToSRV), *Sida micrantha mosaic virus* (SiMMV), *Sida common mosaic virus* (SiCmMV), or any other begomovirus.



Suppl. Figure S3. Mapping of reads from tomato and *Sida* sp. samples against a reference ToSRV DNA-B genome. The reads are reflected by four columns separated by different colors (top image). Coverage of reads against a reference ToSRV DNA-B genome. The number to the left and the color variation in the figure refers to the number of reads mapped at each position, ranges from 0 (white) to highest number of aligned reads (dark blue) (bottom image).


DNA-B ToSRV
NSP
MP
C14T
1829
C18T
661
F14T
1137
C14S
2
C18S
1
F14S
2

Suppl. Figure S4. Mapping of reads from *Sida* sp. and tomato samples against a reference SiMMV DNA-B genome. The reads are reflected by four columns separated by different colors (top image). Coverage of reads against a reference SiMMV DNA-B genome. The number to the left and the color variation in the figure refers to the number of reads mapped at each position, ranges from 0 (white) to highest number of aligned reads (dark blue) (bottom image).