TARSIANE MARA CARNEIRO BARBOSA

NEW INSIGHTS ON BEGOMOVIRUSES: VIRUS-VIRUS INTERACTIONS IN MIXED INFECTIONS AND ESTIMATION OF GENETIC BOTTLENECKS DURING VECTOR TRANSMISSION

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

Orientador: Francisco Murilo Zerbini Júnior

Coorientadora: Angélica Maria Nogueira

VIÇOSA - MINAS GERAIS
2023

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

T
B238n
2023

Barbosa, Tarsiane Mara Carneiro, 1992-
New insights on begomoviruses: virus-virus interactions in mixed infections and estimation ofgenetic bottlenecks during vector transmission / Tarsiane Mara Carneiro Barbosa. – Viçosa, MG, 2023.
1 tese eletrônica (97 f.): il. (algumas color.).

Texto em português e inglês.
Orientador: Francisco Murilo Zerbini Júnior.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Bioquímica e Biologia Molecular, 2023.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2023.224
Modo de acesso: World Wide Web.

1. Geminivírus. 2. Vírus de plantas. 3. Mosca-branca. I. Zerbini Júnior, Francisco Murilo, 1966-. II. Universidade Federal de Viçosa. Departamento de Bioquímica e Biologia Molecular. Programa de Pós-Graduação em Bioquímica Aplicada. III. Título.

CDD 22. ed. 579.28

Bibliotecário(a) responsável: Bruna Silva CRB-6/2552

TARSIANE MARA CARNEIRO BARBOSA

NEW INSIGHTS ON BEGOMOVIRUSES: VIRUS-VIRUS INTERACTIONS IN MIXED INFECTIONS AND ESTIMATION OF GENETIC BOTTLENECKS DURING VECTOR TRANSMISSION

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

APROVADA: 31 de janeiro de 2023.

Assentimento:

gov.br Documento assinado digitalmente
TARSIANE MARA CARNEIRO BARBOSA
Data: 31/05/2023 10:59:28-0300
Verifique em https://validar.iti.gov.br

Tarsiane Mara Carneiro Barbosa
Autora

gov.br Documento assinado digitalmente
FRANCISCO MURILO ZERBINI JUNIOR
Data: 09/05/2023 01:00:04-0300
Verifique em https://validar.iti.gov.br

Francisco Murilo Zerbini Júnior
Orientador

AGRADECIMENTOS

Sou grata a Deus pela sua infinita misericórdia ter atendido as minhas preces e por não ter atendido também. Por Ele sempre saber o que é melhor para mim.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

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

Agradeço aos meus pais e meu irmão por todo amor. Por acreditar tanto em mim. Por entenderem as minhas ausências durante esses anos de estudo. Por me fazerem acreditar que através dos estudos daria uma vida melhor a nossa família.

Ao Paulo (Juninho) agradeço por ser meu alicerce, por ser a paz que minha vida tanto necessita. Sou grata por tanto amor e força recebida durante todos esses dias.

Agradeço a todos os professores e amigos que no passado ajudaram de alguma forma a minha chegada até aqui. Por mais que contatos não tenham se mantido, me recordarei sempre com carinho das tantas vezes que as mãos de vocês se estenderam para me ajudar nesta caminhada.

Ao Professor Murilo Zerbini, agradeço pela paciência, dedicação, ensinamentos, confiança, conselhos e amizade.

A Professora Poliane Alfenas Zerbini, agradeço por todo carinho, pela oportunidade imensurável de podemos trabalhar em equipe nos testes da COVID19 e as experiências compartilhadas

Ao Prof. Jorge Resende pela oportunidade, suporte e confiança.

Aos membros da banca avaliadora pelas valiosas observações para o enriquecimento da tese.

Agradeço a todos do laboratório de ecologia e evolução de vírus e do laboratório de vírus em sua antiga e atual equipe. Foram inúmeras experiência compartilhadas durante esses anos que influenciaram em meu crescimento pessoal e acadêmico. Todos a sua maneira foram importantes para a concretização desta tese.

Agradeço em especial a Aline, Baltazar, Flávia, Filipe, Rafael, Filipe, Yan, Flávia e Patrícia por terem tornado essa caminhada mais leve, pelos momentos especiais que para sempre me lembrarei com carinho, pelos incentivos constantes, companheirismo, momentos de alegria e descontração. Em especial, aos amigos Angélica, Ayane, Vinícius, Túlio, Rafael e Ruither pela

amizade, dedicação, colaboração e responsabilidades compartilhadas durante a execução deste trabalho. Minha eterna gratidão!

A todos que de alguma forma contribuíram para realização deste trabalho: Muito Obrigada!

BIOGRAFIA

Tarsiane Mara Carneiro Barbosa, filha de Maria Aparecida Carneiro Barbosa e Luiz Carlos Barbosa, nasceu em 13 de março de 1992, em João Monlevade, Minas Gerais.
Cursou o ensino fundamental nas Escola Municipal Sebastião Francisco de Ávila e na Escola Estadual Padre Osvaldo de Podestá e o ensino médio na Escola Estadual José Modesto de Ávila, na cidade de Bela Vista de Minas, Minas Gerais.

No ano de 2011, ingressou no curso de graduação em Bioquímica na Universidade Federal de Viçosa (UFV), campus Viçosa-MG, vindo a graduar-se em julho de 2016.
No ano de 2016, iniciou o mestrado no Programa de Pós-graduação em Genética e Melhoramento, na Universidade Federal de Viçosa, obtendo o título de mestre em julho de 2018.
Em agosto de 2018, iniciou o doutorado no Programa de Pós-Graduação em Bioquímica Aplicada, na Universidade Federal de Viçosa, submetendo-se à defesa de tese em 31 de janeiro de 2023.

RESUMO

BARBOSA, Tarsiane Mara Carneiro, D.Sc., Universidade Federal de Viçosa, janeiro de 2023. **Novos avanços no estudo de begomovírus: Interações vírus-vírus em infecção mista e estimativa de gargalos genéticos durante a transmissão pelo vetor.** Orientador: Francisco Murilo Zerbini Júnior. Coorientadora: Angélica Maria Nogueira.

Perdas em plantios de tomateiro causadas por um complexo de begomovírus (família *Geminiviridae*) são observadas em todas as principais regiões produtoras do Brasil. Neste trabalho, aspectos relacionados a interações de begomovírus em infecção mista e gargalos genéticos durante a transmissão pelo inseto vetor foram explorados. Estudos anteriores revelaram uma interação complexa entre os begomovírus tomato rugose mosaic virus (ToRMV) e tomato severe rugose virus (ToSRV), na qual o ToRMV interfere negativamente na infectividade e no acúmulo do ToSRV. Diferenças na proteína Rep e na sequência da região comum (RC) podem estar envolvidas na replicação preferencial dos componentes do ToRMV durante a infecção mista. Dessa forma, utilizamos clones para o DNA-A do ToSRV, construídos anteriormente no Lab. de Ecol. e Evol. de Vírus da UFV, contendo os mesmos nucleotídeos que ocorrem no ToRMV nas posições divergentes da RC e no domínio IRD da proteína Rep. A infectividade do ToSRV e o acúmulo viral em infecção simples não foram afetados por nenhuma das alterações. Na infecção mista de ToRMV com ToSRV-$A_{(ToR:RC)}$, foi observada alta infectividade e acúmulo de DNA do ToSRV, semelhante ao ToSRV selvagem. Isso não foi observado quando as plantas foram inoculadas com ToRMV e ToSRV-$A_{(ToR:IRD)}$. Esses resultados sugerem que os nucleotídeos da RC atuam como sítios de reconhecimento específicos para a ligação da proteína Rep, aumentando a taxa de replicação viral e o acúmulo de DNA viral. A emergência de begomovírus ao longo das últimas décadas está relacionada à disseminação de moscas-brancas do complexo *Bemisia tabaci* em todo o mundo. Embora a transmissão pelo vetor seja considerada um importante gargalo genético para vírus de plantas, poucos estudos tentaram estimar sua magnitude. Para avaliar a ocorrência e magnitude de gargalos genéticos durante a transmissão de begomovírus, *B. tabaci* Middle East-Asia Minor 1 (MEAM1) foi submetida a períodos de aquisição de 24 h e 48 h em plantas de tomate infectadas por ToSRV e *Euphorbia heterophylla* por Euphorbia yellow mosaic virus (EuYMV). Em seguida, os insetos foram coletados e dissecados para obtenção do mesentério (MS) e das glândulas salivares (GS). Os DNAs totais de cada órgão foram extraídos, enriquecidos por meio de RCA e sequenciados. Um maior número de leituras foi mapeada nos genomas de referência na amostra MS em comparação com as amostras GS, consistente com a maior carga viral no MS relatada na literatura. A análise de

SNPs indicou um número significativo de deleções na amostra EuYMV-MG e um alto número de substituições de nucleotídeos nas amostras GS. As estimativas da entropia de Shannon também indicaram um maior grau de variabilidade genética das amostras GS. Os resultados sugerem que o mesentério impõe um gargalo genético em *B. tabaci* MEAM1, mas que a variabilidade genética da população viral é restaurada durante a passagem pelas glândulas salivares.

Palavras-chave: Geminivírus. Vírus de plantas. Mosca branca

ABSTRACT

BARBOSA, Tarsiane Mara Carneiro, D.Sc., Universidade Federal de Viçosa, January 2023. **New insights on begomoviruses: Virus-virus interactions in mixed infections and estimation of genetic bottlenecks during vector transmission.** Advisor: Francisco Murilo Zerbini Júnior. Co-advisor:Angélica Maria Nogueira.

Yield losses in tomato crops caused by a begomovirus complex are observed in all the main producing regions of Brazil. In this work, aspects related to interactions in mixed infection and genetic bottlenecks during transmission by the insect vector were explored. Previous studies revealed a complex interaction between the begomoviruses tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus (ToSRV), in which ToRMV negatively interferes with ToSRV infectivity and accumulation. Differences in the Rep protein and in the common region (CR) sequence may be involved in the preferential replication of ToRMV components during mixed infection. Thus, we used previously constructed clones for ToSRV DNA-A containing the same nucleotides that occur in ToRMV at divergent positions in the CR and IRD domain of the Rep protein. ToSRV infectivity and viral accumulation in single infection were not affected by any of the nucleotide changes. In a mixed infection with ToRMV and ToSRV-$A_{(ToR:CR)}$, high infectivity and DNA accumulation of $ToSRV_{(ToR:CR)}$, similar to wild-type ToSRV, were observed. This was not the case when plants were inoculated with ToRMV and ToSRV-$A_{(ToR:IRD)}$. These results suggest that CR nucleotides act as specific recognition sites for Rep binding, increasing the rate of viral replication and viral DNA accumulation. The emergence of begomoviruses over the last few decades is related to the worldwide spread of whiteflies of the *Bemisia tabaci* complex. Although vector transmission is considered an important genetic bottleneck for plant viruses, few studies have attempted to estimate its magnitude. To assess the occurrence and magnitude of genetic bottlenecks during vector transmission of begomoviruses, *B. tabaci* Middle East-Asia Minor 1 (MEAM1) was subjected to acquisition periods of 24 h and 48 h in tomato plants and *Euphorbia heterophylla* infected by ToSRV and Euphorbia yellow mosaic virus (EuYMV), respectively. The insects were collected and dissected to obtain the midgut (MG) and salivary glands (SG). Total DNA of each organ was extracted, enriched through RCA and sequenced. A larger number of sequence reads mapped to EuYMV and ToSRV reference genomes in the MG sample compared to SG samples, consistent with the higher viral load in the MG reported in the literature. The analysis of SNPs indicated a significant number of deletions in the EuYMV MG sample and a high number of nucleotide substitutions in the SG samples. Estimates of Shannon's entropy also indicated a

greater degree of genetic variability of the SG samples. Our results suggest that the midgut imposes a genetic bottleneck in *B. tabaci* MEAM1, but that the genetic variability of the viral population is restored during passage through the salivary glands.

Keywords: Geminivirus. Plant Virus. Whitefly

SUMÁRIO

INTRODUÇÃO GERAL

A família *Geminiviridae* inclui vírus de plantas com genoma composto por uma ou duas moléculas de DNA circular de fita simples, encapsidadas por uma única proteína estrutural em partículas icosaédricas geminadas. A família agrega 14 gêneros (*Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrilevirus*, *Opunvirus*, *Topilevirus*, *Topocuvirus* e *Turncurtovirus*), definidos com base no tipo de inseto vetor, gama de hospedeiros, organização genômica e relações filogenéticas (Fiallo-Olive *et al.*, 2021).

De acordo com a organização genômica, distribuição geográfica e relacionamento filogenético, os begomovírus são divididos em dois grupos: Velho Mundo (VM; Europa, Ásia, África e Austrália) e Novo Mundo (NM; Américas) (Burger *et al.*, 1990; Padidam *et al.*, 1999; Paximadis *et al.*, 1999; Rybicki, 1994). A maioria dos begomovírus do NM possui genoma bissegmentado, consistindo de dois componentes denominados DNA-A e DNA-B. O DNA-A codifica proteínas associadas à replicação, supressão de respostas de defesa do hospedeiro e encapsidação (Hanley-Bowdoin *et al.*, 2013; Rojas *et al.*, 2005a), enquanto o DNA-B codifica proteínas associadas ao movimento intra- e intercelular, determinação de gama de hospedeiros e supressão de respostas de defesa (Brustolini *et al.*, 2015; Hanley-Bowdoin *et al.*, 2013; Rojas *et al.*, 2005b). Os dois componentes compartilham baixa identidade de sequência, com exceção de uma região intergênica ou região comum (RC) com aproximadamente 200 nucleotídeos na qual a identidade entre o DNA-A e o DNA-B do mesmo vírus é superior a 85% (Briddon *et al.*, 2010; Liu *et al.*, 1997).

A RC inclui uma estrutura em forma de grampo em cuja alça é encontrada uma sequência conservada de nove nucleotídeos (5'-TAATATT//AC-3') que constitui a origem de replicação viral (Laufs *et al.*, 1995). A RC também contém sequências em repetições diretas e

invertidas denominadas iterons, as quais são necessárias para o reconhecimento da origem de replicação pela proteína Rep (Eagle & Hanley-Bowdoin, 1997; Fontes *et al.*, 1994; Orozco & Hanley-Bowdoin, 1996). A sequência de nucleotídeos dos iterons é variável entre os begomovírus porém conservada entre o DNA-A e o DNA-B do mesmo vírus, e constitui um determinante da replicação específica dos dois componentes genômicos (Arguello-Astorga *et al.*, 1994; Arguello-Astorga & Ruiz-Medrano, 2001). Resíduos de aminoácidos localizados no domínio relacionado aos iterons ("iteron-related domain", IRD) da proteína Rep são importantes para o reconhecimento dos iterons, e a região que comporta esses elementos também constitui um determinante de especificidade da replicação (Arguello-Astorga & Ruiz-Medrano, 2001; Londono *et al.*, 2010; Ramos *et al.*, 2003).

Embora o reconhecimento dos iterons pela proteína Rep seja considerado vírus-específico, em alguns casos a Rep de um determinado vírus pode reconhecer os iterons presentes no DNA-B de outro vírus da mesma espécie (raramente de espécies diferentes), gerando assim pseudo-recombinantes (Andrade *et al.*, 2006; Arguello-Astorga & Ruiz-Medrano, 2001; Faria *et al.*, 1994; Garrido-Ramirez *et al.*, 2000; Gilbertson *et al.*, 1993; Rojas *et al.*, 2005b). Para que ocorra a formação de pseudo-recombinantes viáveis em infeções mistas por begomovírus é necessário que exista um certo grau de conservação dos iterons (Arguello-Astorga *et al.*, 1994; Arguello-Astorga & Ruiz-Medrano, 2001; Bull *et al.*, 2007).

Perdas em plantios de tomateiro causadas por um complexo de begomovírus são observadas em todas as principais áreas produtoras do Brasil (Duarte *et al.*, 2021). O país é um dos maiores centros de diversidade genética de begomovírus, e com poucas exceções (Fernandes *et al.*, 2014; Rodríguez-Pardina *et al.*, 2011) os begomovírus que infectam tomateiros não ocorrem em outros países neotropicais (Rocha *et al.*, 2013). Dentre os vírus desse complexo estão o tomato rugose mosaic virus (ToRMV) e o tomato severe rugose virus

(ToSRV) (Rocha *et al.*, 2013). Silva *et al.* (2014) conduziram um estudo para avaliar a dinâmica de infecções simples e mista desses dois vírus em plantas de tomateiro. Nesse estudo, foi observado um mosaico típico da infeção por esses vírus em todas as combinações possíveis entre o DNA-A e o DNA-B de cada vírus (Silva *et al.*, 2014). Embora ToRMV e ToSRV formem pseudo-recombinantes viáveis em tomateiro e possuam iterons idênticos, além de sequências de DNA-B quase idênticas (98,2% de identidade), os componentes genômicos do ToRMV foram detectados preferencialmente em relação aos do ToSRV quando três ou quatro componentes foram inoculados simultaneamente (Silva *et al.*, 2014). Ao realizar a quantificação do acúmulo viral, os componentes do ToRMV foram sempre detectados preferencialmente em relação aos componentes do ToSRV, indicando que o ToRMV interfere negativamente sobre o ToSRV e que a replicação de dois begomovírus em infecção mista não depende apenas da interação entre os iterons e a IRD (Silva *et al.*, 2014).

Apesar do ToRMV e ToSRV possuírem iterons idênticos, as RCs dos dois vírus apresentam divergências na sequência de nucleotídeos em nove posições. Além disso, as IRDs de cada vírus possuem aminoácidos distintos nas posições 2 e 3 (Thr/Pro e Arg/Lys para ToRMV/ToSRV, respectivamente). É possível que uma ou mais dessas diferenças expliquem a replicação preferencial dos componentes do ToRMV em relação aos do ToSRV em infecções mistas (Silva *et al.*, 2014).

Os begomovírus são transmitidos por moscas-brancas do complexo de espécies crípticas *Bemisia tabaci* (Hemiptera: Aleyrodidae) para plantas dicotiledôneas (Brown *et al.*, 2015; Sanger *et al.*, 1978). Após a introdução de *B. tabaci* Middle East-Asia Minor 1 (MEAM1, anteriormente classificada como *B. tabaci* biótipo B) no início da década de 1990 (Lourenção & Nagai, 1994; Melo, 1992), vários begomovírus emergiram na cultura do tomateiro no Brasil (Ribeiro *et al.*, 2003; Ribeiro *et al.*, 1998). *B. tabaci* MEAM1 é altamente polífaga e coloniza eficientemente o tomateiro, assim como muitas outras plantas cultivadas e não-cultivadas. O

complexo *B. tabaci* é constituído por pelo menos 37 espécies indistinguíveis morfologicamente (Boersma *et al.*, 1974; Lee *et al.*, 2013) e definidas geneticamente com base na sequência do gene que codifica a citocromo oxidase I mitocondrial (*mtCOI*) (De Barro *et al.*, 2011; Dinsdale *et al.*, 2010).

A transmissão dos begomovírus por *B. tabaci* ocorre de forma persistente-circulativa. A proteína capsidial (CP) é a única proteína viral essencial para aquisição e transmissão pelo vetor (Azzam *et al.*, 1994). As partículas virais são adquiridas via estilete durante a alimentação em plantas infectadas, entram no esôfago e na câmara de filtro e subsequentemente são transportadas através da parede do mesentério para a hemolinfa, onde circulam até adentrarem na glândula salivar, a partir da qual são transmitidas para novas plantas durante a alimentação do inseto (Ghanim *et al.*, 2007; Hunter *et al.*, 1998). Em pelo menos um caso, a transmissão do tomato yellow leaf curl (TYLCV) por *B. tabaci* MEAM1 e MED, o vírus é capaz de replicar nas células da glândula salivar (He *et al.*, 2020).

Apesar do papel fundamental do inseto vetor na disseminação de begomovírus, pouco se sabe sobre o efeito que a passagem pelo inseto vetor causa na dinâmica populacional do vírus (Duffy *et al.*, 2008; Ng *et al.*, 2011). Durante o processo de transmissão, apenas algumas partículas virais são efetivamente depositadas na planta (Jia *et al.*, 2018; Stafford *et al.*, 2012), caracterizando um gargalo genético severo que pode ter efeitos profundos no desenvolvimento da doença em hospedeiros individuais, bem como na evolução das populações de vírus após rodadas sucessivas de transmissão. A possível replicação viral na glândula salivar pode, por um lado, gerar variabilidade reduzindo a magnitude do gargalo genético, ou, caso a replicação ocorra a partir de um pequeno número de indivíduos, produzir um efeito fundador. Embora não existam evidências epidemiológicas que indiquem a ocorrência de um efeito fundador durante a transmissão do vírus pelo vetor, os mecanismos moleculares que modulam a ocorrência dos

gargalos genéticos durante a passagem dos begomovírus pelos diferentes órgãos do sistema digestivo de *B. tabaci* são muito pouco compreendidos.

Os objetivos deste trabalho foram: 1) Com base no trabalho de Nogueira (2018), verificar em dois novos ensaios se os nucleotídeos divergentes na RC e na região correspondente à IRD da Rep de ToSRV e ToRMV são responsáveis pela interferência negativa que ToRMV exerce sobre o ToSRV; 2) Verificar a existência de gargalos genéticos durante a transmissão de begomovírus pelo inseto vetor.

LITERATURA CITADA

Andrade, E.C.; Manhani, G.G.; Fontes, E.P.B.; Zerbini, F.M. Análise parcial dos determinantes genéticos virais responsáveis pela indução de sintomas por begomovírus em tomateiro e *Nicotiana benthamiana*. **Fitopatologia Brasileira**, v. 31, p. S225, 2006.

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

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

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

Boersma, A.; Lamblin, G.; Degand, P. [Recurrent degradation by glycosidases from Canavalia ensiformis of the polysaccharidic bond of a glycoprotein isolated from human parotid saliva]. **Comptes Rendus des Seances de l'Academie des Sciences. Serie D: Sciences Naturelles**, v. 279, p. 1031-1034, 1974.

Briddon, R.W.; Patil, B.L.; Bagewadi, B.; Nawaz-ul-Rehman, M.S.; Fauquet, C.M. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. **BMC Evolutionary Biology**, v. 10, p. 97, 2010.

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. 1593-1619, 2015.

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

Bull, S.E.; Briddon, R.W.; Sserubombwe, W.S.; Ngugi, K.; Markham, P.G.; Stanley, J. Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. **Journal of General Virology**, v. 88, p. 1624-1633, 2007.

Burger, J.T.; Brand, R.J.; Rybicki, E.P. The molecular cloning and nucleotide sequence of the 3'-terminal region of Ornitoghalum mosaic virus. **Journal of General Virology**, v. 71, p. 2527-2534, 1990.

De Barro, P.J.; Liu, S.S.; Boykin, L.M.; Dinsdale, A.B. *Bemisia tabaci*: A statement of species status. **Annual Review of Entomology**, v. 56, p. 1-19, 2011.

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

Duarte, M.F.; Fonseca, M.E.N.; Costa, H.; Fernandes, N.A.N.; Reis, A.; Boiteux, L.S.; Pereira-Carvalho, R.C. Diversity of tomato-infecting begomoviruses and spatiotemporal dynamics of an endemic viral species of the Brazilian Atlantic rain forest biome. **Virus Genes**, v. 57, p. 83-93, 2021.

Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: Patterns and determinants. **Nature Reviews Genetics**, v. 9, p. 267-276, 2008.

Eagle, P.A.; Hanley-Bowdoin, L. *Cis*-elements that contribute to geminivirus transcriptional regulation and the efficiency of DNA replication. **Journal of Virology**, v. 71, p. 6947-6955, 1997.

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

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

Fiallo-Olive, E.; Lett, J.M.; Martin, D.P.; Roumagnac, P.; Varsani, A.; Zerbini, F.M.; Navas-Castillo, J. ICTV Virus Taxonomy Profile: *Geminiviridae* 2021. **Journal of General Virology**, v. 102, p. 001696, 2021.

Fontes, E.P.B.; Eagle, P.A.; Sipe, P.S.; Luckow, V.A.; Hanley-Bowdoin, L. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. **Journal of Biological Chemistry**, v. 269, p. 8459-8465, 1994.

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

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

Gilbertson, R.L.; Hidayat, S.H.; Paplomatas, E.J.; Rojas, M.R.; Hou, Y.-H.; Maxwell, D.P. Pseudorecombination between infectious cloned DNA components of tomato mottle and bean dwarf mosaic geminiviruses. **Journal of General Virology**, v. 74, p. 23-31, 1993.

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

He, Y.Z.; Wang, Y.M.; Yin, T.Y.; Fiallo-Olive, E.; Liu, Y.Q.; Hanley-Bowdoin, L.; Wang, X.W. A plant DNA virus replicates in the salivary glands of its insect vector via recruitment of host DNA synthesis machinery. **Proceedings of the National Academy of Sciences, USA**, v. 117, p. 16928-16937, 2020.

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

Jia, D.; Chen, Q.; Mao, Q.; Zhang, X.; Wu, W.; Chen, H.; Yu, X.; Wang, Z.; Wei, T. Vector mediated transmission of persistently transmitted plant viruses. **Current Opinion in Virology**, v. 28, p. 127-132, 2018.

Laufs, J.; Traut, W.; Heyraud, F.; Matzeit, G.; Rogers, S.G.; Schell, J.; Gronenborn, B. *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. **Proceedings of the National Academy of Sciences, USA**, v. 92, p. 3879-3883, 1995.

Lee, W.; Park, J.; Lee, G.S.; Lee, S.; Akimoto, S.I. Taxonomic status of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) and reassessment of the number of its constituent species. **PLoS ONE**, v. 8, p. e63817, 2013.

Liu, S.; Bedford, I.D.; Briddon, R.W.; Markham, P.G. Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. **Journal of General Virology**, v. 78, p. 1791-1794., 1997.

Londono, A.; Riego-Ruiz, L.; Arguello-Astorga, G.R. DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. **Archives of Virology**, v. 155, p. 1033-1046, 2010.

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

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

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

Nogueira, A.M. **Bipartite begomoviruses: Replication, transmission and interaction with alphasatellites.** 2018. (D.Sc. Thesis) - Department of Plant Pathology, Federal University of Viçosa, Brazil.

Orozco, B.M.; Hanley-Bowdoin, L. A DNA structure is required for geminivirus replication origin function. **Journal of Virology**, v. 70, p. 148-158, 1996.

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

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.

Ramos, P.L.; Guevara-Gonzalez, R.G.; Peral, R.; Ascencio-Ibanez, J.T.; Polston, J.E.; Arguello-Astorga, G.R.; Vega-Arreguin, J.C.; Rivera-Bustamante, R.F. *Tomato mottle Taino virus* pseudorecombines with PYMV but not with ToMoV: Implications for the delimitation of *cis*- and *trans*-acting replication specificity determinants. **Archives of Virology**, v. 148, p. 1697-1712, 2003.

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.

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

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

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

Sanger, F.; Coulson, A.R.; Friedmann, T.; Air, G.M.; Barrell, B.G.; Brown, N.L.; Fiddes, J.C.; Hutchison, C.A., 3rd; Slocombe, P.M.; Smith, M. The nucleotide sequence of bacteriophage phiX174. **Journal of Molecular Biology**, v. 125, p. 225-246, 1978.

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

Stafford, C.A.; Walker, G.P.; Ullman, D.E. Hitching a ride: Vector feeding and virus transmission. **Communicative & Integrative Biology**, v. 5, p. 43-49, 2012.

CAPÍTULO I

SPECIFIC NUCLEOTIDES IN THE COMMON REGION OF THE BEGOMOVIRUS TOMATO RUGOSE MOSAIC VIRUS (ToRMV) ARE RESPONSIBLE FOR THE NEGATIVE INTERFERENCE OVER TOMATO SEVERE RUGOSE VIRUS (ToSRV) IN MIXED INFECTION

Barbosa TMC, Nogueira AM, Quadros AFF, Orílio AF, Bigão MCJ, Xavier CAD, Ferro CG, Zerbini FM (2023) Specific nucleotides in the common region of the begomovirus tomato rugose mosaic virus (ToRMV) are responsible for the negative interference over tomato severe rugose virus (ToSRV) in mixed infection. **Virology Journal**, *in preparation*.

Specific nucleotides in the common region of the begomovirus tomato rugose mosaic virus (ToRMV) are responsible for the negative interference over tomato severe rugose virus (ToSRV) in mixed infection

Tarsiane M.C. Barbosa[1,2], Angélica M. Nogueira[1,2,3], Ayane F.F. Quadros[1,2], Anelise F. Orílio[1,2], Marcela C.J. Bigão[1,2], César A.D. Xavier[4], Camila G. Ferro[5], F. Murilo Zerbini[1,2]

[1]Dep. de Fitopatologia, [2]National Institute for Science and Technology on Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [3]Dep. de Proteção Vegetal, Universidade Estadual Paulista, Botucatu, SP, Brazil; [4]Dep. Entomology and Plant Pathology, North Carolina State University, Raleigh, North Carolina, USA; [5]Dep. Departamento de Fitopatologia e Nematologia, Universidade de São Paulo/ESALQ, Piracicaba, SP, Brazil

Abstract

Mixed infections between two or more begomoviruses are commonly found in tomato fields, and can affect disease outcomes by increasing symptom severity and viral accumulation compared to single infections. Viruses that affect tomato include tomato rugose virus (ToSRV) and tomato rugose mosaic virus (ToRMV). Previous work has shown that in mixed infections ToRMV negatively affects the infectivity and accumulation of ToSRV. ToSRV and ToRMV share a high degree of sequence identity, including *cis*-elements (iterons) in the common region (CR) and their specific recognition sites (iteron-related domain, IRD) within the *Rep* gene, which are essential for the initiation of viral replication. In this work we investigated if divergent sites in the CR and IRD could be involved in the interaction between these two begomoviruses. ToSRV DNA-A clones were constructed containing the same nucleotides as the ToRMV DNA-A at divergent positions of the CR (ToSRV-$A_{(ToR:CR)}$), the *Rep* gene IRD (ToSRV-$A_{(ToR:IRD)}$) and both regions (ToSRV-$A_{(ToR:CR+IRD)}$). ToSRV infectivity and viral accumulation in a single infection were not affected by nucleotide changes in any of the regions. In mixed inoculation of ToRMV with ToSRV-$A_{(ToR:CR),}$ high infectivity of both viruses and high DNA accumulation of ToSRV-$A_{(ToR:CR)}$, similar to wild-type ToSRV, were observed. This was not the case when plants were co-inoculated with ToRMV and ToSRV-$A_{(ToR:IRD)}$. These results suggest that the altered CR sites serve as specific recognition sites for Rep binding, increasing the rate of viral replication and viral DNA accumulation. On the other hand, the decrease in viral accumulation observed in plants inoculated with ToSRV-$A_{(ToR:CR+IRD)}$ suggests that the divergent amino acids in the IRD do not offer an advantage for ToSRV replication efficiency.

Introduction

The genus *Begomovirus* (family *Geminiviridae*) comprises plant viruses with a genome consisting of one or two single-stranded circular DNA molecules encapsidated separately by a single structural protein in twinned icosahedral particles (Fiallo-Olive et al., 2021). In bipartite begomoviruses, the genomic components are referred to as DNA-A and DNA-B (Hanley-Bowdoin et al., 2013; Hanley-Bowdoin et al., 1999; Rojas et al., 2005). Genes in the DNA-A encode proteins involved in replication, suppression of host defense responses and particle formation, while those in the DNA-B encode proteins involved in viral movement throughout the plant and suppression of host defense responses (Hanley-Bowdoin et al., 2013; Rojas et al., 2005). Each genomic component has genes in both DNA strands that form two clusters separated by an approximately 200-nt intergenic region (IR), known as the common region (CR) in bipartite viruses (Fiallo-Olive et al., 2021). The CRs of the DNA-A and DNA-B of the same virus have more than 85% sequence identity (Briddon et al., 2010).

Replication of begomoviruses occurs in the nuclei of infected cells, by rolling circle and recombination-mediated mechanisms (Gutierrez, 1999; Jeske et al., 2001; Lazarowitz, 1992). The virus-encoded Rep protein is essential for replication (Hanley-Bowdoin et al., 2013; Laufs et al., 1995). Rep is a multifunctional protein with four highly conserved motifs in its N-terminal region (Laufs et al., 1995; Rizvi et al., 2015). Motif I (FLTY) is required for binding to dsDNA (Arguello-Astorga and Ruiz-Medrano, 2001; Orozco and Hanley-Bowdoin, 1998), motif II (HLH) is involved in binding to metal ions, motif III contains the endonuclease activity responsible for cleavage at the origin of replication and binding at the resulting 3'-OH (Rizvi et al., 2015; Saunders et al., 1991; Singh et al., 2007), and motif IV (geminivirus Rep sequence, GRS), located between motifs II and III, is necessary for the initiation of rolling-circle replication and the establishment of viral infection (Nash et al., 2011).

To initiate replication, the Rep protein binds in a sequence-specific manner to the CR (Fontes et al., 1992; Laufs et al., 1995). This region contains an origin of replication (ORI) structured in a conserved stem-loop, with the loop containing the invariant nonanucleotide 5'-TAATATT//AC-3' (Hanley-Bowdoin et al., 1999) that is cleaved by Rep to generate the 3'- OH for the initiation of DNA synthesis (Gutierrez, 1999; Hanley-Bowdoin et al., 1999; Lazarowitz, 1992). The CR also contains iterated *cis*-elements (direct and inverted repeats), called iterons, which are recognized by the iteron-related domain (IRD) of Rep to promote its binding to the CR (Arguello-Astorga et al., 1994; Fontes et al., 1994). The sequence of begomovirus iterons is comprised of two guanines followed by three nucleotides (N1, N2, and N3) that vary depending on the begomovirus (Arguello-Astorga et al., 1994; Arguello-Astorga and Ruiz-Medrano, 2001). The IRD is comprised of eight amino acids and together with motif I form the functional domain of Rep (Arguello-Astorga and Ruiz-Medrano, 2001).

Generally, begomoviruses with identical iteron sequences encode Rep proteins with similar IRDs (Arguello-Astorga and Ruiz-Medrano, 2001). A correspondence between IRD and iteron sequences is found by nucleotide-amino acid pairing that occurs between the iteron nucleotide N1 and the eighth amino acid of the IRD, N2 with the sixth amino acid, and N3 with the first or third amino acids (Arguello-Astorga and Ruiz-Medrano, 2001; Ramos et al., 2003). Additional amino acids that are important for iteron binding have been identified in the motif II of Rep (Londono et al., 2010). Some of these amino acids may provide Rep with greater flexibility by allowing recognition of sequences of divergent, non-cognate iterons (Andrade et al., 2006).

Due to its site-specificity, the Rep protein efficiently recognizes cognate DNA-A and DNA-B components (which always have identical iterons). However, in the field, mixed infections between two or more begomoviruses often occur, providing the opportunity to generate new viruses by the exchange of genomic components. The viability of these reassortants (or

pseudorecombinants) will depend on the conservation of iteron sequences and of the amino acid sequence of the Rep protein within the IRD and motif II (Arguello-Astorga and Ruiz-Medrano, 2001; Ramos et al., 2003).

Tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus (ToSRV) belong to a complex of begomoviruses that infect tomato in Brazil (Rocha et al., 2013). Silva et al. (2014) demonstrated that ToRMV and ToSRV form viable pseudorecombinants in all possible combinations of single and mixed infection in tomato. However, there was a preferential detection of ToRMV DNA-A and DNA-B in relation to ToSRV DNA-A and DNA-B, and the accumulation of ToSRV in mixed infections was reduced compared to single infection. Thus, not only there is no synergism between these two viruses, but ToRMV may negatively interfere with ToSRV replication.

Comparing the CR and Rep sequences of ToRMV and ToSRV, Silva et al. (2014) found nine nucleotide differences in their CRs (although their iterons are identical) and different amino acids at positions 2 and 3 of the IRD (Thr/Pro and Arg/Lys for ToRMV/ToSRV, respectively). It is possible that one or more of these CR nucleotides and/or IRD amino acids play a role in the preferential replication of ToRMV in relation to ToSRV in a mixed infection.

In this study, we evaluated the effect of nucleotide changes in the CR and in the region corresponding to the Rep IRD on ToSRV infectivity and accumulation. Our results indicate that the differences in the CR, but not in the IRD, are responsible for the negative interference of ToRMV on ToSRV.

Material and Methods

Viral isolates and re-sequencing of isolates

Infectious clones corresponding to DNA-A and DNA-B of the viral isolates ToRMV-[BR:Ub1:96] (Fernandes et al., 2006) and ToSRV-[BR:PG1:Pep:03] (Bezerra-Agasie et al., 2006) were used. These are the same clones used in the study of Silva et al. (2014).

For confirmation of the sequences of the infectious clones, complete genome units were obtained by excision with BamHI (ToSRV DNA-A), EcoRI (ToRMV DNA-A) and KpnI (DNA-B from both viruses). After purification, these fragments were ligated into the pBLUESCRIPT-KS+ vector (Stratagene) previously linearized with the same enzymes. Recombinant plasmids were used for transformation into *Escherichia coli* by electroporation and three clones of each genomic component were sent for commercial sequencing by primer walking (Macrogen Inc, South Korea). Sequences were assembled using DNA Baser v. 3.5 (Heracle Biosoft), and organized to begin at the nicking site in the invariant nonanucleotide at the origin of replication (5'-TAATATT//AC-3'). Pairwise sequence comparisons of DNA-A and DNA-B, CRs and ORFs between ToRMV and ToSRV were performed using Sequence Demarcation Tool (SDT) v.1.2 (Muhire et al., 2014) using the MUSCLE alignment option (Edgar, 2004).

Construction of infectious clones

After re-sequencing of the ToRMV and ToSRV infectious clones, nucleotide differences were observed at eight positions in the CR, and amino acid differences were observed in two positions within the IRD of the Rep protein (Figure 1).

To verify if these differences are responsible for the reduced accumulation of ToSRV and the predominance of ToRMV over ToSRV in a mixed infection, a full-length clone of the

ToSRV DNA-A exchanging the nucleotides at the divergent positions in the CR and the IRD by those of ToRMV DNA-A (named ToSRV-$A_{(ToR:CR+IRD)}$) was synthesized using the GenPlus service from the company GenScript (www.genscript.com) (Figure 2). In the ToSRV DNA-A CR, the adenine at position 1 (Figure 1) was replaced by thymine, the cytosine at position 2 by a guanine, and the thymine at position 5 was deleted. In the *Rep* gene IRD, the first base of the fifth codon (cytosine) was replaced by adenine, and the second base of the sixth codon (adenine) was replaced by guanine. These two substitutions altered the amino acid sequence of the Rep protein (proline to threonine and lysine to arginine, respectively) (Figure 2). It is important to mention that as the Rep protein is expressed from the virion complementary strand (synthesized during viral DNA replication), the bases replaced on the virion strand of the infectious clone sequence are complementary bases, thus, in the region corresponding to the fifth codon of *Rep* in the virion strand, a guanine was replaced by thymine and in the sixth a thymine by a cytosine. The synthesized DNA-A was cloned on the pUC57 vector, and Sanger sequencing was performed to confirm that only the nucleotides of interest were inserted.

To construct the infectious clone, the complete sequence of ToSRV-$A_{(ToR:CR+IRD)}$ was subjected to *in silico* restriction analysis using the program ApE (A plasmid Editor) v. 2.0.53 (Davis and Jorgensen, 2022). The restriction enzyme SacI, which cleaves at only one point in the genome and is located closer to the origin of replication, was chosen to simulate the site that would be used for cloning the full-length ToSRV-$A_{(ToR:CR+IRD)}$ (2,593 bp). To generate the partially redundant portion of the ToSRV clone that must contain the origin of replication, we used SacI and NcoI to generate a 476 bp fragment (SacI2593 to NcoI3069). This fragment also contains the region where the ToRMV nucleotides were inserted (represented as green rectangles in Figure 2A). The presence of the same region where the ToRMV nucleotides were inserted in both the 2,593 bp and 476 bp fragments would hinder the site-specific mutagenesis process to be performed subsequently, because the primer containing the ToRMV nucleotides could

anneal in only one of two regions during PCR. Considering that for a clone to be infectious it needs to have no more than two origins of replication, the repeated region where the ToRMV nucleotides were inserted and the one immediately upstream to it were excluded from the fragment corresponding to one copy of the viral genome, generating an infectious clone with 2,829 bp instead of 3,069 (Figure 2B). After synthesis, the fragment was cloned into the pUC57 vector using the blunt end enzyme EcoRV.

Additional clones containing only the CR (ToSRV-$A_{(ToR:CR)}$) or only the IRD ToRMV nucleotides (ToSRV-$A_{(ToR:IRD)}$) were constructed by site-directed mutagenesis (also by GenScript), using ToSRV-$A_{(ToR:CR+IRD)}$ as a template. To construct ToSRV-$A_{(ToR:CR)}$, the two ToRMV nucleotides in the *Rep* gene were reversed to the original sequence. To construct ToSRV-$A_{(ToR:IRD)}$, the three ToRMV nucleotides in the CR were reversed to the original sequence (Figure 2). Confirmation of all clones was performed by Sanger sequencing.

Inoculation of plants

Tomato (*Solanum lycopersicum* cv. Santa Clara) seedlings were inoculated by biolistics (Aragão et al., 1996) using 10 μg of each genomic component (DNA-A and DNA-B) of the wild-type (wt) ToRMV and ToSRV infectious clones, and the clones ToSRV-$A_{(ToR:CR)}$, ToSRV-$A_{(ToR:IRD)}$, and ToSRV-$A_{(ToR:CR+IRD)}$ in single (ToRMV-A+B; ToSRV-$A_{(wt)}$+B; ToSRV-$A_{(ToR:CR)}$+B; ToSRV-$A_{(ToR:IRD)}$+B and ToSRV-$A_{(ToR:CR+IRD)}$+B) and mixed inoculations (ToSRV-$A_{(wt)}$+B + ToRMV-A+B; ToSRV-$A_{(ToR:CR)}$+B + ToRMV-A+B; ToSRV-$A_{(ToR:IRD)}$+B + ToRMV-A+B; and ToSRV-$A_{(ToR:CR+IRD)}$+B + ToRMV-A+B). The ToSRV wt DNA-B was always inoculated with the ToSRV clones in single and mixed inoculations. Tomato plants were also bombarded with DNA-free tungsten particles as a negative control. Ten plants were inoculated per treatment and four independent experiments were performed. All

plants were kept in a greenhouse and their phenotype was evaluated for up to 35 days after inoculation (dai).

Detection and quantification of genomic components

Total DNA from all plants was extracted at 28 days after inoculation (dai) as described by Doyle and Doyle (1987). For confirmation of single infections, the DNA-A of ToRMV and ToSRV were detected by PCR using primers ToRMV-$A_{(For)}$ (5'-CAG TAG TTG CCT TCG AAT TGA AG-3'), ToRMV-$A_{(Rev)}$ (5'-CAC GTG TAG CAA TCT CCT TAA AGG-3'), ToSRV-$A_{(For)}$ (5'-CAG TAG TTG CCC TCA AAT TGA AG-3') and ToSRV-$A_{(Rev)}$ (5'-CAC GTG TAG CAA TCT CCT TAA AGA G-3'). Reactions contained 1 µL of total DNA, 0.4 µM of each primer, 1x Reaction Buffer Complete (Celco), 0.2 mM of each dNTP and 1.25 U of Taq DNA Polymerase (5 U/µl) (Celco) in a final volume of 25 µL. PCR parameters consisted of an initial denaturation step at 95°C for 2 min, followed by 38 cycles at 95°C for 1 min, 66°C for 1 min and 72°C for 1 min, followed by a final extension at 72 °C for 10 minutes. PCR products were separated by 1% agarose gel electrophoresis and stained with ethidium bromide. Confirmation of mixed infections was performed by amplification of the complete viral genome by rolling circle amplification (RCA) (Inoue-Nagata et al., 2004) followed by cleavage with specific restriction enzymes for the detection of each component (RCA-RFLP), as described by Silva et al. (2014).

After confirmation of the plants with single and mixed infections, the positive plants were selected for analysis of ToSRV and ToRMV DNA-A accumulation. In mixed infection, only plants in which the ToSRV DNA-A was detected simultaneously with ToRMV DNA-A were selected. Accumulation was assessed by absolute quantification using quantitative real-time PCR (qPCR), performed in a StepOnePlus system (Applied Biosystems) using 100 ng of

total DNA and primers qToSRV-$A_{(For)}$ (5'-AAA GTA AAG TGA TTG TCT GTG G-3'), qToSRV-$A_{(Rev)}$ (5'-GCC GTT CAA CAA ATT GGG-3'), qToRMV-$A_{(For)}$ (5'-CTT AGA TGT ACT AGC CAT GTG G-3') and qToRMV-$A_{(Rev)}$ (5'-CCT TTA ATT TTA TTG AAA ATA ATT TTG GC-3'). All reactions were performed in technical triplicates for each biological replication. Thermocycling conditions consisted of an initial denaturation step of 95°C for 3 min followed by 40 cycles of 95°C for 10s and 66°C for 30s, with a final dissociation step. Viral accumulation was determined by interpolation of the Ct values of each tested sample within the standard curves for ToRMV and ToSRV. To construct the standard curves, serial dilutions (5 x 10^4 to 5 x 10^9 copies) were prepared from known quantities of plasmids containing a single copy of the genomic component of each virus. Quantifications of the plasmid DNA used to construct the standard curves and the total DNA samples were performed using a Nanodrop 2000c (ThermoFisher Scientific).

Sequence comparisons

Multiple sequence alignments were prepared for the CRs and the Rep of ToSRV isolates using the MUSCLE alignment option in MEGA11 (Tamura et al., 2021). Sequence logos were created using the WebLogo web server (weblogo.threeplusone.com) (Crooks et al., 2004).

Analysis of structural properties

The AlphaFold2 program (Jumper et al., 2021) was used to predict the tertiary structure of the Rep of ToRMV, ToSRV and ToSRV$_{(ToR:IRD)}$. After the 3D structure prediction, the RMSD and TM-score were calculated using the mTM-align tool (Dong et al., 2018). Structures were visualized using RCSB PDB Mol* Viewer v2.6.0.

Statistical analysis

Data were analyzed by ANOVA with a significance level of $p \leq 0.05$ followed by Fisher's least significant difference (LSD) test to separate means into homogeneous groups. The analyzes were performed using R version 4.2.2 (R Core Team, 2017). Statistical analysis of data from Nogueira (2018) was redone and evaluated together with the data generated in this work.

Results

Effects of nucleotide changes on ToSRV and ToRMV infectivity

To verify whether the nucleotide differences observed in the CR and IRD of the Rep protein are responsible for the reduction of ToSRV accumulation and for the preferential detection of ToRMV in mixed infections with ToSRV, infectious clones of ToSRV DNA-A containing the same sequence of the ToRMV DNA-A in the CR (ToSRV-$A_{(ToR:CR)}$), the IRD (ToSRV-$A_{(ToR:IRD)}$) and in both regions (ToSRV-$A_{(ToR:CR+IRD)}$) were inoculated in tomato plants. Treatments consisted of single infections, where each clone was inoculated along with wt ToSRV DNA-B, and mixed infections where the clones were inoculated along with wt ToSRV DNA-B plus wt ToRMV DNA-A and DNA-B. Symptoms started to appear at 10-14 dpi and consisted of yellow mosaic and leaf distortion in all treatments, with no differences between them (Figure 3). Thus, based on symptoms, no synergistic interaction was observed between the two viruses.

The results of the infectivity test are presented in Table 1. The overall infectivity rate of single infections was similar for all treatments in both experiments, ranging from 35% to 50% for experiment I (reproduced from Nogueira, 2018), and 35 % to 75% for experiment II. The results indicate that wt and mutant ToSRV and wt ToRMV infect tomatoes in single infection with equivalent efficiency.

Mixed infections were screened by RCA-RFLP to verify the presence of each inoculated DNA component (one representative of each DNA-A and DNA-B combination of viruses detected in each treatment is shown in Supplement Figure S1, adapted from Nogueira, 2018). As observed for single infections, the results for each treatment were consistent in the two experiments, with only a few discrepancies observed for detection of wt DNA-B components in the ToSRV-$A_{(ToR:CR)}$ treatment. It is also noteworthy that the wt ToSRV DNA-B component was detected at a much lower percentage than the wt ToRMV DNA-B component, in all treatments (Table 1).

Considering the DNA-A infectivity and combining the results of the two experiments, when wt ToRMV was inoculated with wt ToSRV, the ToRMV DNA-A (henceforth referred to as ToR-A) was detected at a higher rate compared to the ToSRV DNA-A (ToS-A) (37.5% and 27.5% for ToR-A and ToS-A, respectively). However, in most plants the DNA-A component of each virus was detected alone, with only a small number of plants infected by the two components (4 plants with ToS-A+ToR-A vs. 7 plants with ToS-A alone and 11 with ToR-A alone). Equivalent results were obtained with the ToSRV-$A_{(ToR:IRD)}$ clone, with 27.5% and 17.5% of the plants infected with ToR-A and ToS-$A_{(ToR:IRD)}$, respectively, but with both components being detected in only 1 plant vs. 6 plants with ToS-$A_{(ToR:IRD)}$ alone and 10 plants with ToR-A alone. Strikingly, for the ToSRV$_{(ToR:CR)}$ clone, the ToR-A and ToS-$A_{(ToR:CR)}$ components were detected in a similar number of plants (45% and 50%, respectively), and most plants were infected by the two components (15 plants with both components vs. 5 with ToS-$A_{(CR)}$ alone and 3 with ToR-A alone). With the ToSRV$_{(ToR:CR+IRD)}$ clone, both components were also detected at similar rates (28.9% and 26.3% for ToR-A and ToS-$A_{(ToR:CR+IRD)}$, respectively), with 6 plants infected by the two components, 4 with ToS-$A_{(ToR:CR+IRD)}$ alone and 5 with ToR-A alone (Table 1). Together, these results indicate that the nucleotide differences in the CR, but

not the amino acid differences in the IRD, are responsible for the prevalence of ToRMV components over ToSRV components.

Effects of nucleotide changes on ToSRV and ToRMV accumulation

To verify the accumulation of ToRMV and ToSRV DNA-A in single and mixed infections, quantitative real-time PCR (qPCR) was performed using virus-specific DNA-A primers. For the analysis of plants with mixed infections, quantification of viral accumulation was performed only for plants in which DNA-A components of both viruses were confirmed to be present. Since this was the case for only one plant in the ToSRV-$A_{(IRD)}$ + ToRMV treatment (Table 1), this treatment was not included in the analysis.

In single infections, no differences in ToS-A accumulation were observed for the wt ToSRV, ToSRV-$A_{(ToR:CR)}$, ToSRV-$A_{(ToR:IRD)}$ and ToSRV-$A_{(ToR:CR+IRD)}$ treatments (Figure 4A), confirming that all three ToSRV DNA-A clones are capable of replicating at equivalent, wt levels. ToRMV was also quantified in single infection and in mixed infection with ToSRV, with no differences in its accumulation among treatments (Figure 4B). When wt ToSRV was quantified in mixed infection with ToRMV, the accumulation of ToSRV was significantly reduced (Figure 4C), confirming the previous reported by Silva et al. (2014). Strikingly, in the plants inoculated with ToSRV-$A_{(ToR:CR)}$ and ToRMV, the accumulation of ToS-$A_{(ToR:CR)}$ was statistically equivalent to that of wt ToSRV in single infection, and corresponded to an 80% increase in relation to wt ToSRV in mixed infection with ToSRV (Figure 4C). The accumulation of ToSRV-$A_{(ToR:CR+IRD)}$ in mixed infected plants corresponded to 50% that of wt ToSRV in single infection but, interestingly, was statistically equivalent to that of ToSRV$_{(ToR:CR)}$ (Figure 4C). The results of viral DNA accumulation are consistent with the infectivity assay (Table

1), indicating that the nucleotide differences in the CR are responsible for the negative interference of ToRMV on ToSRV.

Nucleotide and amino acid differences among ToSRV isolates

We compared ToSRV-[BR:PG1:Pep:03] to a total of 144 ToSRV isolates from different hosts to see how common the observed variations in the CR and IRD are among isolates of the virus. Analyzing the IRD nucleotide sequence, we found that only five of these isolates contained a single variable nucleotide, and that only two out of these five isolates lack cytosine as the first base of the fifth Rep codon (like ToSRV-[BR:PG1:Pep:03] does) (Suppl. Figure S2). For the CR region, it was observed that the adenine found in position 1 of ToSRV-[BR:PG1:Pep:03] is also present in six other isolates obtained from beans, soybeans, and tomatoes (BR:ITA1274:14, BR:SOITA1014:14, BR:DF607, BR:Pip1792:03, BR:G2, BR:G3). All other isolates have a thymine at this position, which is the same as what is found in ToRMV. At position 2, the ToSRV-[BR:PG1:Pep:03] isolate has a cytosine, but every other isolate, including the ToRMV isolate, has a guanine. At position 5, 3.5% of the isolates contain cytosine, 45.1% of the isolates resemble ToRMV without a nucleotide in this region and 51.4% of the isolates resemble ToSRV-[BR:PG1:Pep:03] (Suppl. Figure S2).

Analysis of structural properties

To investigate possible conformational differences in the structure of the Rep protein of $ToSRV_{(ToR:IRD)}$ and $ToSRV_{(ToR:CR+IRD)}$, an *in silico* prediction was performed using the AlphaFold2 program. As mutations in the CR would not affect the structure of any virus-encoded protein, this was not considered in the modeling (Suppl. Figure S3). All models generated by the program had high confidence metrics. The similarity measure provided by the TM-score

value was 0.62 between wt ToRMV and ToSRV, and the same value was observed for the comparison between ToRMV and ToSRV$_{(ToR:IRD)}$. The TM-score value for ToSRV$_{(ToR:IRD)}$ in relation to wt ToSRV was 0.96. Another widely used measure to express structural similarity is the root mean square distance value (RMSD) calculated between equivalent atoms in the two structures. An RMSD equal to 0 means identical structures, and the greater the difference between the structures, the greater the value. The values of ToRMV were 1.618 Å in relation to wt ToSRV and 1.812 Å for ToSRV$_{(ToR:IRD)}$. The RMSD between ToSRV$_{(ToR:IRD)}$ and wt ToSRV was 1.647 Å.

The ToSRV$_{(ToR:IRD)}$ clone was evaluated with single changes: ToSRV$_{(Pro>Thr)}$ and ToSRV$_{(Lys>Arg)}$. ToSRV$_{(Pro>Thr)}$ had a TM-score of 0.62, 0.94 and 0.98, and RMDS of 1.569 Å, 1.841Å and 0.898 Å in relation to ToRMV, wt ToSRV and ToSRV$_{(ToR:IRD)}$, respectively. ToSRV$_{(Lys>Arg)}$ yielded TM-scores of 0.62, 0.97 and 0.97, and RMDS of 1,360 Å, 1,352 Å and 1,162 Å in relation to ToRMV, wt ToSRV and ToSRV$_{(ToR:IRD)}$, respectively.

Discussion

The host range of viruses classified in the same family often overlaps considerably. Moreover, viruses classified in the same genus usually share the same type of vector (such as whiteflies for begomoviruses). Thus, mixed infections by begomoviruses (genus *Begomovirus*) in the same plant are very common in nature (Martin and Elena, 2009; Moreno and López-Moya, 2020; Renteria-Canett et al., 2011). In this co-infection process, viral fitness can be altered due to virus-virus or virus-plant interactions that may occur and influence viral replication and accumulation (Lecoq et al., 2011; Power, 1996). Nevertheless, the current understanding of the dynamics of mixed infections is limited, since the mechanisms that control the differences in viral accumulation and, consequently, the possible predominance of one virus over the other are unknown. In this study we analyzed the interaction between ToSRV and ToRMV to better

understand co-infection aspects. Our results indicate that the apparently minor differences in the common region (CR) sequence of these two begomoviruses can strongly influence their interaction.

ToRMV and ToSRV have a high level of nucleotide sequence identity (98.2%). Ribeiro et al. (2007) and Silva et al. (2014) suggested that ToRMV is a recombinant whose parental viruses are ToSRV and tomato chlorotic mottle virus (ToCMoV). Recombination analysis revealed that the ToRMV DNA-A contains a portion of the CR, including the iterons, and nearly all of the *Rep* gene derived from ToSRV, and that ToRMV captured the ToSRV DNA-B.

Recombination and pseudorecombination are frequent and relevant mechanisms for the generation of genetic variability in begomoviruses (Andrade et al., 2006; Lefeuvre et al., 2009; Padidam et al., 1999; Rojas et al., 2005). Recombination sites are not randomly distributed in the genome: the origin of replication and the 5'-terminal portion of the Rep gene are recombination hot spots, frequently exchanged during replication (Hanley-Bowdoin et al., 1999; Lefeuvre et al., 2007). The formation of viable pseudorecombinants is common among isolates of the same virus (Faria et al., 1994; Garrido-Ramirez et al., 2000; Stanley et al., 1985; von Arnim and Stanley, 1992) and can also occur between distinct viruses that exhibit high identity in the CR sequences of the heterologous components (Andrade et al., 2006; Chakraborty et al., 2008; Gilbertson et al., 1993; Höfer et al., 1997; Ramos et al., 2003; Unseld et al., 2000). Thus, the viral factors involved in replication can be transferred (by recombination) or shared (by pseudorecombination) between different viruses.

ToSRV clones containing the same nucleotides as ToRMV at divergent positions in the CR and IRD were constructed to determine whether these regions could improve ToSRV infectivity and replication efficiency when in mixed infection with ToRMV. Indeed, both parameters were improved in the case of ToSRV-$A_{(ToR:CR)}$, but not $ToSRV_{(ToR:IRD)}$. ToSRV-$A_{(ToR:CR+IRD)}$ also had improved infectivity and accumulation, although not as significantly as

ToSRV$_{(ToR:CR)}$. Thus, the presence of T and G at CR positions 1 and 2 and a deletion at position 5 (numbered according to the alignment of Figure 1A) seems to confer an adaptive advantage to ToSRV isolates in tomato plants when in mixed infection with ToRMV.

Recognition of the iterons by the Rep protein is a key event for binding of the Rep/REn complex to the origin of replication, and thus for the initiation of viral replication (Arguello-Astorga et al., 1994; Fontes et al., 1994; Fontes et al., 1992). According to the proposed model of recognition between Rep and the iterons (Arguello-Astorga and Ruiz-Medrano, 2001), the sixth and eighth amino acids of the IRD (which are the same for ToRMV and ToSRV) control the binding efficiency of Rep to the viral genome, and the second and third amino acids (which diverge between ToRMV and ToSRV) are not supposed to dictate preference for specific iteron sequences. However, this model was based entirely on *in silico* analyses and functional analyzes were not performed.

The amino acids that determine the high binding affinity of Rep to iterons have been little explored so far (Avalos-Calleros et al., 2021). The exchange of only one amino acid residue at position 10 between Rep proteins from two tomato leaf curl New Delhi virus (ToLCNDV) isolates (equivalent to IRD position 9 in ToRMV and ToSRV) resulted in a shift in replication specificity and an increase in viral DNA accumulation (Chatterji et al., 1999). The IRD domain is contained within a β1 mini sheet structure composed of a cluster of amino acid residues that recognize dsDNA, as proposed by Campos-Olivas et al. (2002) when analyzing the Rep of tomato yellow leaf curl virus (TYLCV). In β-sheet structures, hydrophobic amino acids are recurrent, whereas proline and charged amino acids (Arg, Lys, Glu, and Asp) are underrepresented (Kim et al., 2016). In ToRMV, the IRD has Thr and Arg amino acids at positions 2 and 3 of the domain, while ToSRV has Pro and Lys. Thus, it was not unreasonable to assume that the replacement of Pro by Thr in ToSRV$_{(ToR:IRD)}$ would have improved the recognition of the iterons, while the replacement of Lys by Arg would have a neutral effect. The fact

that an improvement was not observed is puzzling. On the contrary, the phenotype of ToSRV$_{(ToR:CR+IRD)}$ suggests that the IRD mutations had an adverse effect in the IRD in terms of iteron recognition. In addition, we observed that the mutations did not lead to conformational changes in the Rep protein, thus only influencing recognition affinity.

One hypothesis to explain these observation is that the ToSRV Rep could be more flexible in its ability to recognize iterons, whereas the ToRMV Rep would be more specific. This hypothesis arises from analyzes of the N-terminal region of the Rep of other begomoviruses that infect tomato, such as ToCMoV, tomato yellow spot virus (ToYSV) and tomato golden vein virus (TGMV). Despite having distinct IRD sequences, these viruses have the same amino acids as ToSRV at these divergent positions in the IRD (Andrade et al., 2006). Another possibility is that the ToRMV-derived CR sequence may favor replication, while the ToRMV-derived IRD sequence provides a disadvantage. To verify whether the ToRMV Rep IRD amino acids confer greater binding specificity, infectivity assays with ToRMV clones in which amino acids in the divergent IRD sites were replaced with the ToSRV-encoded amino acids should be conducted. It would also be interesting to determine if mixed infection with ToYSV or ToCMoV has the same effect negative on ToSRV, and whether the adaptive changes observed with the ToSRV$_{(ToR:CR)}$ mutant also occur in mixed infection with these viruses.

Complex interactions are observed in the coexistence of two or more viruses and can lead to changes in a number of characteristics, including transmission efficiency (Wintermantel et al., 2008), the severity of symptoms (Renteria-Canett et al., 2011), breakdown of resistance in cultivated hosts (García-Cano et al., 2006; Singh et al., 2016), cell tropism (Alves-Junior et al., 2009; Wege and Siegmund, 2007), and viral DNA titer (Martin and Elena, 2009; Wang et al., 2009). Our results of viral load quantification indicated that ToSRV accumulation is drastically reduced in mixed infections with ToRMV. Interestingly, ToRMV also negatively interferes with the accumulation of ToYSV during the initial stages of infection in tomato (Alves-

Junior et al., 2009). A reduced rate of viral accumulation may be due to competition for host resources in mixed infections, resulting in less accumulation of one or both viruses compared to single infections (Singhal et al., 2021). On the other hand, beneficial effects of mixed infection involving begomoviruses have been reported for pepper plants infected with pepper huasteco yellow vein virus (PHYVV) and pepper golden mosaic virus (PepGMV) (Renteria-Canett et al., 2011) and in cassava infected with African cassava mosaic virus (ACMV) and East African cassava mosaic virus (EACMV) (Vanitharani et al., 2004). In both cases, mixed infection resulted in higher viral concentrations of both viruses and more severe symptoms compared to single infections. Despite differences in viral accumulation, synergism was not identified in our investigation, with similar symptoms observed in single and mixed infections.

According to Martin and Elena (2009), different individuals in a population might be viewed as unique participants in a biological game. The winner will be the one with the best strategy (highest fitness) and, as a result, will have the highest frequency in the population. Based on the infectivity of the genomic components and the accumulation of viral DNA, we propose that the divergent nucleotides in the ToRMV CR make this virus a better competitor in the mixed infection with ToSRV due to more efficient binding of Rep to the viral DNA.

Given the prevalence of ToSRV over ToRMV in the field, the disadvantage presented by ToSRV in comparison to ToRMV in mixed infections is somewhat contradictory (Fernandes et al., 2008; Inoue-Nagata et al., 2016; Rocha et al., 2013). Interestingly, most ToSRV field isolates share the same nucleotides as ToRMV at the three CR positions examined in our study. Therefore, it is possible that the isolate utilized in the construction of the infectious clone (BR:PG1:Pep:03) is an atypical ToSRV isolate. To determine whether the negative effect exerted by ToRMV prevails, interaction studies between ToRMV and other ToSRV isolates that do not contain nucleotide differences at these positions would be helpful.

As a final observation, the isolate ToSRV-[BR:PG1:Pep:03] was obtained from pepper rather than tomato. The differences observed in the CR sequence of ToSRV-[BR:PG1:Pep:03] may be the result of viral adaptation to a new host, and this adaptation may have generated a fitness cost. Although ToSRV naturally infects tomato and pepper, infection in tomato is much more prevalent and severe. New studies comparing ToSRV isolates from various hosts may additional insights into viral evolution in terms of adaptation to new hosts and potential fitness effects.

References

Alves-Junior, M., Alfenas-Zerbini, P., Andrade, E.C., Esposito, D.A., Silva, F.N., Cruz, A.C.F., Ventrella, M.C., Otoni, W.C., Zerbini, F.M., 2009. Synergism and negative interference during co-infection of tomato and *Nicotiana benthamiana* with two bipartite begomoviruses. Virology 387, 257-266.

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

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

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

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

Avalos-Calleros, J.A., Pastor-Palacios, G., Bolaños-Martínez, O.C., Mauricio-Castillo, A., Gregorio-Jorge, J., Martínez-Marrero, N., Bañuelos-Hernández, B., Méndez-Lozano, J., Arguello-Astorga, G.R., 2021. Two strains of a novel begomovirus encoding Rep proteins with identical β1 strands but different β5 strands are not compatible in replication. Arch Virol 166, 1691-1709.

Bezerra-Agasie, I.C., Ferreira, G.B., Ávila, A.C., Inoue-Nagata, A.K., 2006. First report of *Tomato severe rugose virus* in chili pepper in Brazil. Plant Dis 90, 114.

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

Campos-Olivas, R., Louis, J.M., Clerot, D., Gronenborn, B., Gronenborn, A.M., 2002. The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. Proc Natl Acad Sci USA 99, 10310-10315.

Chakraborty, S., Vanitharani, R., Chattopadhyay, B., Fauquet, C.M., 2008. Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct

species of begomovirus associated with severe tomato leaf curl disease in India. J Gen Virol 89, 818-828.

Chatterji, A., Padidam, M., Beachy, R.N., Fauquet, C.M., 1999. Identification of replication specificity determinants in two strains of tomato leaf curl virus from New Delhi. J Virol 73, 5481-5489.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res 14, 1188-1190.

Davis, M.W., Jorgensen, E.M., 2022. ApE, A plasmid Editor: A freely available DNA manipulation and visualization program. Front Bioinform 2, 818619.

Dong, R., Peng, Z., Zhang, Y., Yang, J., 2017. mTM-align: an algorithm for fast and accurate multiple protein structure alignment. Bioinformatics 34, 1719-1725.

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

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

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

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

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

Fiallo-Olive, E., Lett, J.M., Martin, D.P., Roumagnac, P., Varsani, A., Zerbini, F.M., Navas-Castillo, J., 2021. ICTV Virus Taxonomy Profile: *Geminiviridae* 2021. J Gen Virol 102, 001696.

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

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

García-Cano, E., Resende, R.O., Fernandez-Muñoz, R., Moriones, E., 2006. Synergistic interaction between Tomato chlorosis virus and Tomato spotted wilt virus results in breakdown of resistance in tomato. Phytopathology 96, 1263-1269.

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

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

Gutierrez, C., 1999. Geminivirus DNA replication. Cell Mol Life Sci 56, 313-329.

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.

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S., Robertson, D., 1999. Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Plant Sci 18, 71-106.

Höfer, P., Engel, M., Jeske, H., Frischmuth, T., 1997. Host range limitation of a pseudorecombinant virus produced by two distinct bipartite geminiviruses. Mol Plant Microbe In 10, 1019-1022.

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 Met 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.

Jeske, H., Lutgemeier, M., Preiss, W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon mosaic virus*. EMBO J 20, 6158-6167.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583-589.

Kim, D.N., Jacobs, T.M., Kuhlman, B., 2016. Boosting protein stability with the computational design of β-sheet surfaces. Protein Sci 25, 702-710.

Laufs, J., Traut, W., Heyraud, F., Matzeit, G., Rogers, S.G., Schell, J., Gronenborn, B., 1995. *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. Proc Natl Acad Sci USA 92, 3879-3883.

Lazarowitz, S.G., 1992. Geminiviruses: Genome structure and gene function. Crit Rev Plant Sci 11, 327-349.

Lecoq, H., Fabre, F., Joannon, B., Wipf-Scheibel, C., Chandeysson, C., Schoeny, A., Desbiez, C., 2011. Search for factors involved in the rapid shift in watermelon mosaic virus (WMV) populations in South-eastern France. Virus Res 159, 115-123.

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

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

Londono, A., Riego-Ruiz, L., Arguello-Astorga, G.R., 2010. DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. Arch Virol 155, 1033-1046.

Martin, S., Elena, S.F., 2009. Application of game theory to the interaction between plant viruses during mixed infections. J Gen Virol 90, 2815-2820.

Moreno, A.B., López-Moya, J.J., 2020. When viruses play team sports: Mixed infections in plants. Phytopathology 110, 29-48.

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

Nash, T.E., Dallas, M.B., Reyes, M.I., Buhrman, G.K., Ascencio-Ibanez, J.T., Hanley-Bowdoin, L., 2011. Functional analysis of a novel motif conserved across geminivirus Rep proteins. J Virol 85, 1182-1192.

Orozco, B.M., Hanley-Bowdoin, L., 1998. Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. J Biol Chem 273, 24448-24456.

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

Power, A.G., 1996. Competition between viruses in a complex plant-pathogen system. Ecology 77, 1004-1010.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramos, P.L., Guevara-Gonzalez, R.G., Peral, R., Ascencio-Ibanez, J.T., Polston, J.E., Arguello-Astorga, G.R., Vega-Arreguin, J.C., Rivera-Bustamante, R.F., 2003. *Tomato mottle Taino virus* pseudorecombines with PYMV but not with ToMoV: Implications for the delimitation of *cis*- and *trans*-acting replication specificity determinants. Arch Virol 148, 1697-1712.

Renteria-Canett, I., Xoconostle-Cazares, B., Ruiz-Medrano, R., Rivera-Bustamante, R.F., 2011. Geminivirus mixed infection on pepper plants: Synergistic interaction between PHYVV and PepGMV. Virol J 8.

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

Rizvi, I., Choudhury, N.R., Tuteja, N., 2015. Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication. Arch Virol 160, 375-387.

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

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

Saunders, K., Lucy, A., Stanley, J., 1991. DNA forms of the geminivirus African cassava mosaic virus are consistent with a rolling circle mechanism of replication. Nucleic Acids Res 19, 2325-2330.

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

Singh, A.K., Kushwaha, N., Chakraborty, S., 2016. Synergistic interaction among begomoviruses leads to the suppression of host defense-related gene expression and breakdown of resistance in chilli. Appl Microbiol Biotechnol 100, 4035-4049.

Singh, D.K., Islam, M.N., Choudhury, N.R., Karjee, S., Mukherjee, S.K., 2007. The 32 kDa subunit of replication protein A (RPA) participates in the DNA replication of Mung bean yellow mosaic India virus (MYMIV) by interacting with the viral Rep protein. Nucleic Acids Res 35, 755-770.

Singhal, P., Nabi, S.U., Yadav, M.K., Dubey, A., 2021. Mixed infection of plant viruses: Diagnostics, interactions and impact on host. J Plant Dis Protect 128, 353-368.

Stanley, J., Townsend, R., Curson, S.J., 1985. Pseudorecombinants between cloned DNAs of two isolates of Cassava Latent Virus. J Gen Virol 66, 1055-1061.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38, 3022-3027.

Unseld, S., Ringel, M., Konrad, A., Lauster, S., Frischmuth, T., 2000. Virus-specific adaptations for the production of a pseudorecombinant virus formed by two distinct bipartite geminiviruses from Central America. Virology 274, 179-188.

Vanitharani, R., Chellappan, P., Pita, J.S., Fauquet, C.M., 2004. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. J Virol 78, 9487-9498.

von Arnim, A., Stanley, J., 1992. Determinants of tomato golden mosaic virus symptom development located on DNA B. Virology 186, 286-293.

Wang, J.B., Turina, M., Medina, V., Falk, B.W., 2009. Synergistic interaction between the potyvirus, *Turnip mosaic virus* and the crinivirus, *Lettuce infectious yellows virus* in plants and protoplasts. Virus Res 144, 163-170.

Wege, C., Siegmund, D., 2007. Synergism of a DNA and an RNA virus: Enhanced tissue infiltration of the begomovirus Abutilon mosaic virus (AbMV) mediated by Cucumber mosaic virus (CMV). Virology 357, 10-28.

Wintermantel, W.M., Cortez, A.A., Anchieta, A.G., Gulati-Sakhuja, A., Hladky, L.L., 2008. Co-infection by two criniviruses alters accumulation of each virus in a host-specific manner and influences efficiency of virus transmission. Phytopathology 98, 1340-1345.

Table 1. Infectivity of tomato severe rugose virus (ToSRV) in single infection and in mixed infection with tomato rugose mosaic virus (ToRMV).

Treatment[*]	**Genomic component**	**Infected plants**[$]	
		Exp. I+II[§]	**Exp. III+IV**
$ToSRV_{(wt)}$	$ToS\text{-}A_{(wt)}$[#]	7/20 (35)	9/20 (45)
ToRMV	ToR-A[#]	7/20 (35)	7/20 (35)
$ToSRV\text{-}A_{(ToR:CR)}$	$ToS\text{-}A_{(ToR:CR)}$[#]	7/20 (35)	12/19 (63)
$ToSRV\text{-}A_{(ToR:IRD)}$	$ToS\text{-}A_{(ToR:IRD)}$[#]	8/20 (40)	15/20 (75)
$ToSRV\text{-}A_{(ToR:CR+IRD)}$	$ToS\text{-}A_{(ToR:CR+IRD)}$[#]	10/20 (50)	10/20 (50)
$ToSRV\text{-}A_{(ToR:CR)}$ + ToRMV	$ToS\text{-}A_{(ToR:CR)}$+ToR-A[&]	7/20 (35)	8/20 (40)
	$ToS\text{-}A_{(ToR:CR)}$[¶]	1/20 (5)	4/20 (15)
	ToR-A[¶]	1/20 (5)	2/20 (10)
	ToS-B+ToR-B[&]	0/20 (0)	9/20 (45)
	ToS-B[¶]	0/20 (0)	0/20 (0)
	ToR-B[¶]	9/20 (45)	5/20 (25)
$ToSRV\text{-}A_{(ToR:IRD)}$ + ToRMV	$ToS\text{-}A_{(ToR:IRD)}$+ToR-A[&]	0/20 (0)	1/20 (5)
	$ToS\text{-}A_{(ToR:IRD)}$[¶]	3/20 (15)	3/20 (15)
	ToR-A[¶]	6/20 (30)	4/20 (20)
	ToS-B+ToR-B[&]	1/20 (5)	0/20 (0)
	ToS-B[¶]	1/20 (5)	0/20 (0)
	ToR-B[¶]	7/20 (35)	5/20 (25)
$ToSRV\text{-}A_{(ToR:CR+IRD)}$ + ToRMV	$ToS\text{-}A_{(ToR:CR+IRD)}$+ToR-A[&]	3/20 (15)	3/18 (17)
	$ToS\text{-}A_{(ToR:CR+IRD)}$[¶]	3/20 (15)	1/18 (6)
	ToR-A[¶]	3/20 (15)	2/18 (11)
	ToS-B+ToR-B[&]	1/20 (5)	2/18 (11)
	ToS-B[¶]	0/20 (0)	0/18 (0)
	ToR-B[¶]	8/20 (40)	7/18 (35)
$ToSRV_{(wt)}$ + ToRMV	ToS-A+ToR-A[&]	2/20 (10)	2/20 (10)
	ToS-A[¶]	4/20 (20)	3/20 (15)
	ToR-A[¶]	4/20 (20)	7/20 (35)
	ToS-B+ToR-B[&]	3/20 (15)	1/20 (5)
	ToS-B[¶]	1/20 (5)	0/20 (0)
	ToR-B[¶]	6/20 (30)	6/20 (30)

[*] For simplicity, treatments are named based on the DNA-A; wild-type DNA-B components were always inoculated together with their cognate DNA-A components.

[$] Number of infected plants/number of inoculated plants (percentage of infected plants), confirmed by PCR with virus-specific primers (single infections) or by RCA-RFLP (mixed infections). Four independent treatments were carried out, which were added as indicated. See Supplementary table S1 for results of each of the four experiments.

[§] Results refer to the work of Nogueira (2018).

[#] Plants inoculated with each virus in single infection.

[&] Plants inoculated with both viruses in which the DNA-A or DNA-B of both viruses were detected.

[¶] Plants inoculated with both viruses in which the DNA-A or DNA-B of only one of the two viruses was detected.

Figure legends

Figure 1. A. Alignment of the nucleotide sequences of the common region (CR) of the DNA-A and DNA-B of tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus (ToSRV). The arrowheads and numbers indicate the eight divergent positions among the CRs of the four components. Asterisks represent conserved nucleotides among the aligned sequences. The horizontal arrows indicate the direction of the iterons. The TATA box and the G-box are boxed. The conserved nonanucleotide at the origin of replication is indicated in light gray. **B.** Alignment of the amino acid sequences of the N-terminal portion of the Rep proteins of ToRMV and ToSRV. Divergent amino acids are highlighted in black. The iteron-related domain (IRD), specificity determinants (SPDs) and protein motifs are boxed with dashed lines. **C.** Alignment of the nucleotide sequences of the region corresponding to the IRDs of ToRMV and ToSRV. Identical nucleotides are indicated by an asterisk.

Figure 2. Schematics of the construction of the ToSRV clones containing the ToRMV nucleotides at the divergent positions in the CR and the *Rep* gene IRD. Altered nucleotides in the CR and IRD are indicated by the green and yellow boxes, respectively. In the CR (green box), the numbers 1, 2 and 5 represent divergent positions and underlined bases indicate the changes. In the IRD (yellow box), the numbers represent the amino acids encoded by the fifth and sixth codons of the *Rep* gene, and underlined bases indicate the changes. **A**. Representation of the ToSRV DNA-A infectious clone containing a full-length copy of the DNA-A (2,593 nt, gray line) plus a 476-bp tandem fragment (black line) containing the origin of replication. **B**. Representation of the ToSRV DNA-A infectious clone after the deletion of the repeated region (dotted line). The nonanucleotide at the origin of replication (TAATATTAC) is indicated. The red and blue arrows represent the viral genes and the direction in which they are transcribed (viral and complementary sense, respectively).

Figure 3. Symptoms in tomato plants biolistically inoculated with tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus ($ToSRV_{(wt)}$, $ToSRV\text{-}A_{(ToR:CR)}$, $ToSRV\text{-}A_{(ToR:IRD)}$ and $ToSRV\text{-}A_{(ToR:CR+IRD)}$) infectious clones in single and mixed infections, as indicated.

Figure 4. Accumulation of tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus ($ToSRV_{(wt)}$, $ToSRV\text{-}A_{(ToR:CR)}$, $ToSRV\text{-}A_{(ToR:IRD)}$ and $ToSRV\text{-}A_{(ToR:CR+IRD)}$) DNA-A components in single and mixed infection in tomato plants. Absolute quantification of viral DNA was performed at 28 days post-inoculation. Boxplots correspond to viral accumulation, presented as the logarithm of the number of molecules. Dots indicate outliers. Means were compared using Fisher's Least Significant Difference (LSD) test. Letters represent statistically significant differences ($p<0.05$) and error bars indicate the standard deviation.

Supplementary Figure S1. Specific detection of DNA-A and DNA-B components of tomato rugose mosaic virus (ToRMV) and tomato severe rugose virus (ToSRV) in mixed infection by RCA-RFLP. Digestion with HindIII enables the detection of ToSRV DNA-A generating a fragment of approximately 2,600 nucleotides (nt). Digestion with XhoI enables the detection of ToRMV DNA-A generating a fragment of approximately 2,600 nt. Digestion with BglII+SacII enables the detection and differentiation of DNA-B components generating fragments of 1,800 and 800 nt for ToSRV DNA-B and approximately 2,600 nt for ToRMV DNA-B. Detection of ToRMV and ToSRV in single infection shown as controls. M, size marker (1 kb plus DNA ladder, Kasvi).

Supplementary Figure S2. Weblogo describing the alignment of multiple common region (CR) and Rep partial nucleotide sequences of tomato severe rugose virus (ToSRV) isolates. **A.** Partial alignment of CR nucleotide sequences including the divergent positions 1-5. **B.** Partial

alignment of Rep nucleotide sequences including the IRD domain delimited by dashed lines. The asterisks correspond to the altered nucleotides for the construction of the ToSRV-$A_{(ToR:IRD)}$ and ToSRV-$A_{(ToR:CR+IRD)}$ clones. The y-axis represents the bit score, where 2 means 100% conservation. The x-axis displays the nucleotide position in the multiple sequence alignment. The height of the letters indicates the level of conservation at that specific position.

Supplementary Figure S3. Rep protein tertiary strucutre prediction by Alphafold. **A.** Tomato rugose mosaic virus (ToRMV) Rep protein (orange). **B.** Tomato severe rugose virus **(**ToSRV) wild type (wt) Rep protein (blue). **C.** $ToSRV_{(ToR:IRD)}$ Rep protein (purple). **D.** Overlay of ToRMV and $ToSRV_{(ToR:IRD)}$ Rep proteins. **E.** Overlay of wt ToSRV and $ToSRV_{(ToR:IRD)}$ Rep proteins. **F.** Overlay of ToRMV, wt ToSRV and $ToSRV_{(ToR:IRD)}$ Rep proteins. The sites where nucleotides were changed in $ToSRV_{(ToR:IRD)}$ are highlighted in green.

Figure 1

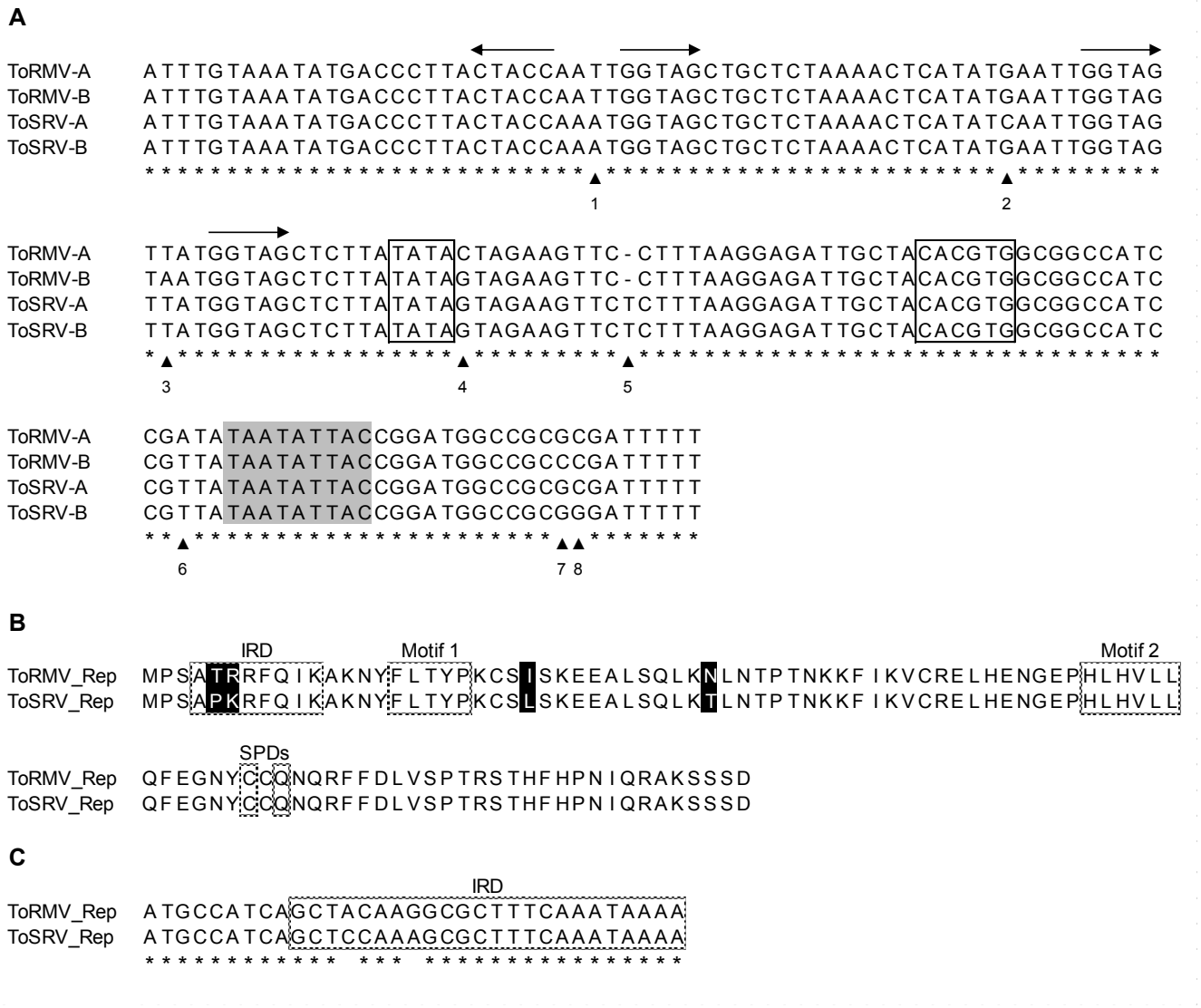

A
ToRMV-A ATTTGTAAATATGACCCTTACTACCAATTGGTAGCTGCTCTAAAACTCATATGAATTGGTAG
ToRMV-B ATTTGTAAATATGACCCTTACTACCAATTGGTAGCTGCTCTAAAACTCATATGAATTGGTAG
ToSRV-A ATTTGTAAATATGACCCTTACTACCAAATGGTAGCTGCTCTAAAACTCATATCAATTGGTAG
ToSRV-B ATTTGTAAATATGACCCTTACTACCAAATGGTAGCTGCTCTAAAACTCATATGAATTGGTAG
1
2
ToRMV-A TTATGGTAGCTCTTATATACTAGAAGTTC-CTTTAAGGAGATTGCTACACGTGGCGGCCATC
ToRMV-B TAATGGTAGCTCTTATATAGTAGAAGTTC-CTTTAAGGAGATTGCTACACGTGGCGGCCATC
ToSRV-A TTATGGTAGCTCTTATATAGTAGAAGTTCTCTTTAAGGAGATTGCTACACGTGGCGGCCATC
ToSRV-B TTATGGTAGCTCTTATATAGTAGAAGTTCTCTTTAAGGAGATTGCTACACGTGGCGGCCATC
3
4
5
ToRMV-A CGATATAATATTACCGGATGGCCGCGCGATTTTT
ToRMV-B CGTTATAATATTACCGGATGGCCGCCCGATTTTT
ToSRV-A CGTTATAATATTACCGGATGGCCGCGCGATTTTT
ToSRV-B CGTTATAATATTACCGGATGGCCGCGGGATTTTT
6
7 8
B
IRD
Motif 1
Motif 2
ToRMV_Rep MPSATRRFQIKAKNYFLTYPKCSISKEEALSQLKNLNTPTNKKFIKVCRELHENGEPHLHVLL
ToSRV_Rep MPSAPKRFQIKAKNYFLTYPKCSLSKEEALSQLKTLNTPTNKKFIKVCRELHENGEPHLHVLL
SPDs
ToRMV_Rep QFEGNYCCQNQRFFDLVSPTRSTHFHPNIQRAKSSSD
ToSRV_Rep QFEGNYCCQNQRFFDLVSPTRSTHFHPNIQRAKSSSD
C
IRD
ToRMV_Rep ATGCCATCAGCTACAAGGCGCTTTCAAATAAAA
ToSRV_Rep ATGCCATCAGCTCCAAAGCGCTTTCAAATAAAA

Figure 2

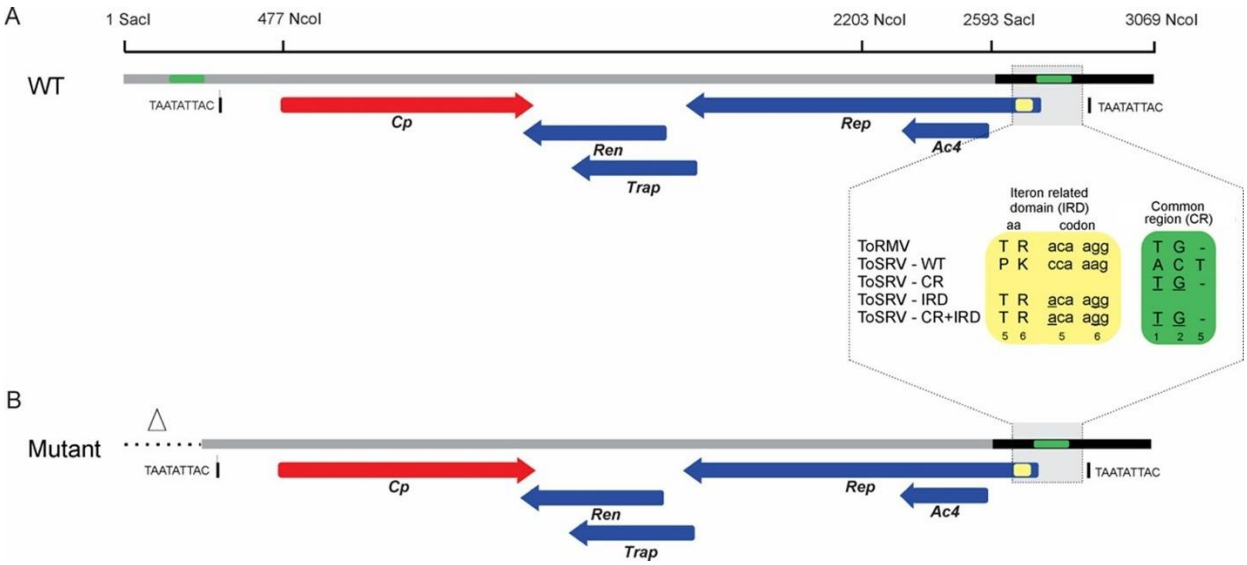
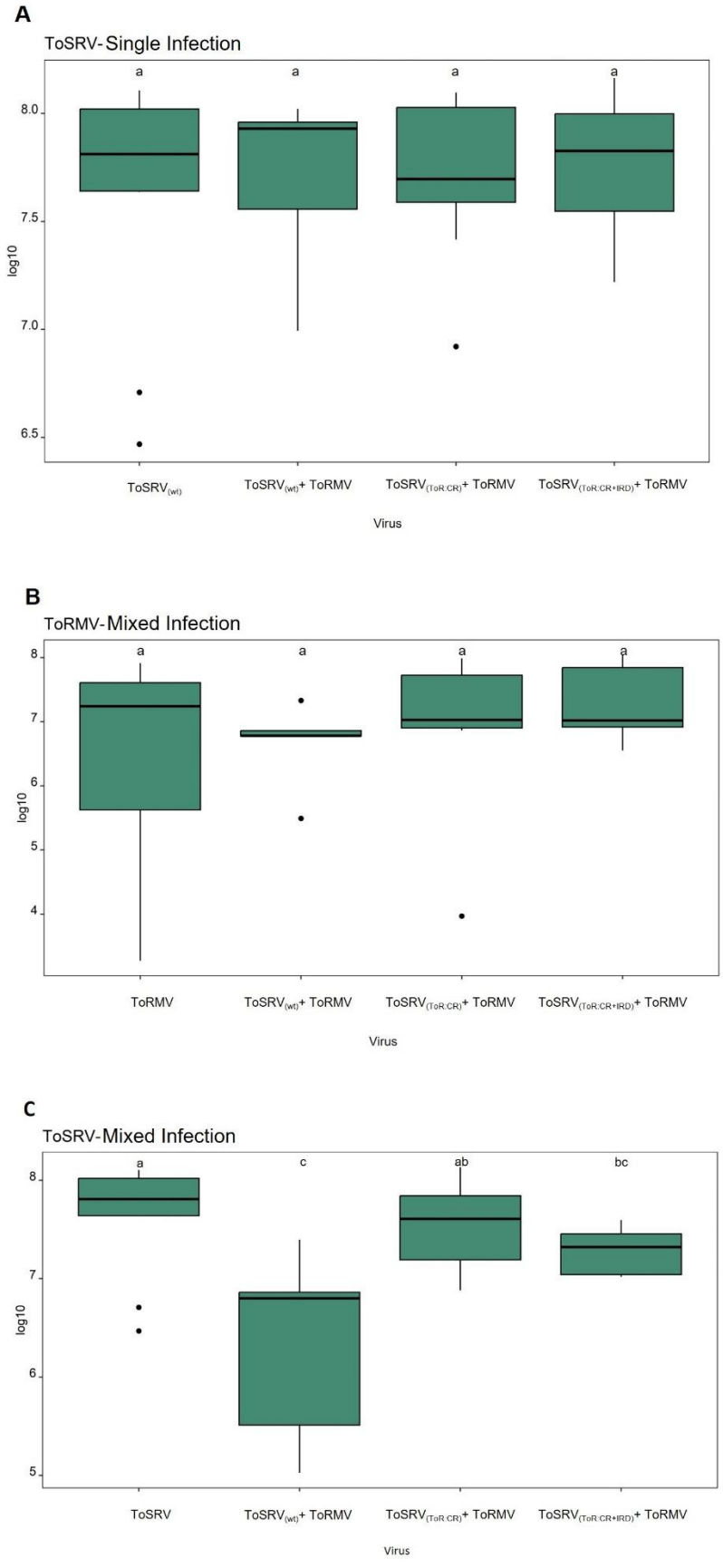

A
1 SacI
477 NcoI
2203 NcoI
2593 SacI
3069 NcoI
WT
TAATATTAC
Cp
Ren
Trap
Rep
Ac4
TAATATTAC
Iteron related domain (IRD)
aa
codon
Common region (CR)
ToRMV
ToSRV - WT
ToSRV - CR
ToSRV - IRD
ToSRV - CR+IRD
T R aca agg
P K cca aag
T R aca agg
T R aca agg
T G -
A C T
T G -
T G -
5 6 5 6
1 2 5
B
Mutant
TAATATTAC
Cp
Ren
Trap
Rep
Ac4
TAATATTAC

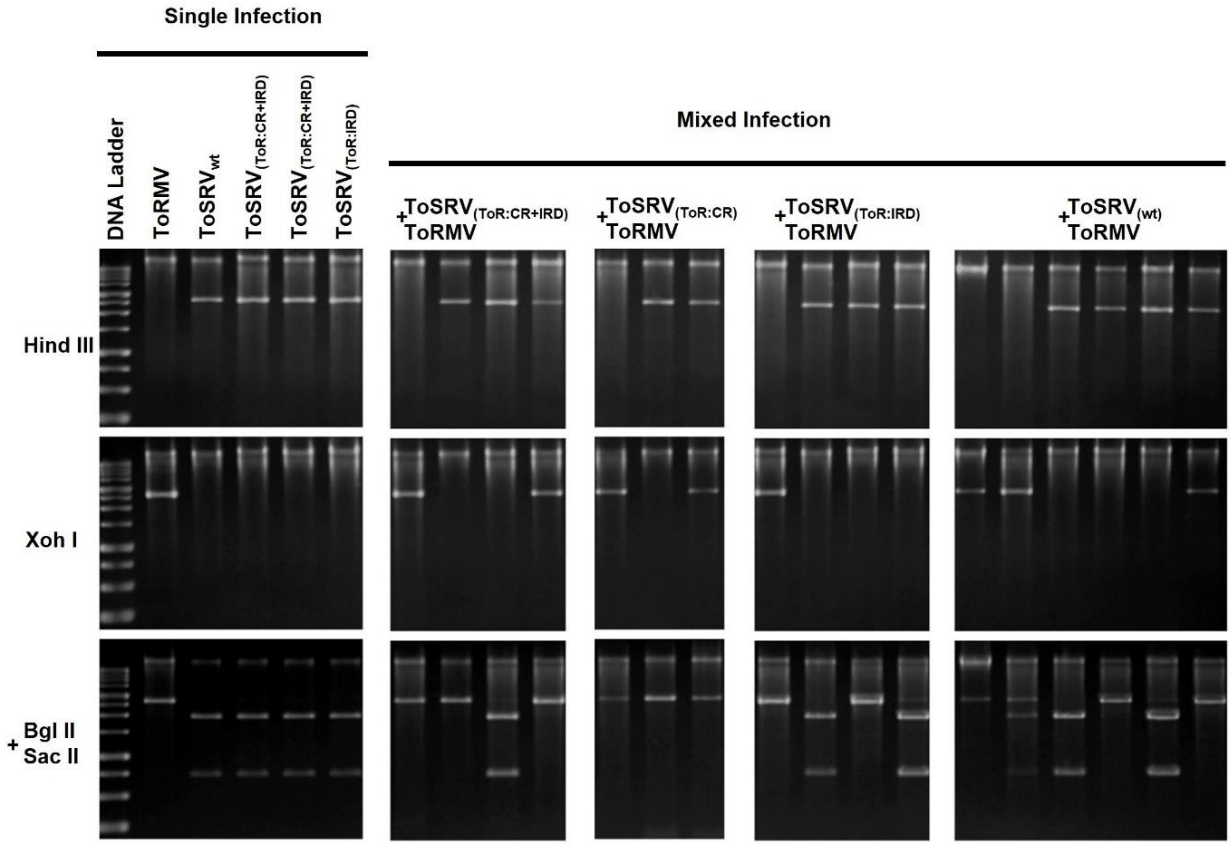
Figure 3


Mock
ToRMV
ToSRV wt
ToRMV+ToSRV wt
ToSRV (ToR:CR)
ToSRV (ToR:IRD)
ToSRV (ToR:CR+IRD)
ToRMV+ToSRV (ToR:CR)
ToRMV+ToSRV (ToR:IRD)
ToRMV+ToSRV (ToR:CR+IRD)

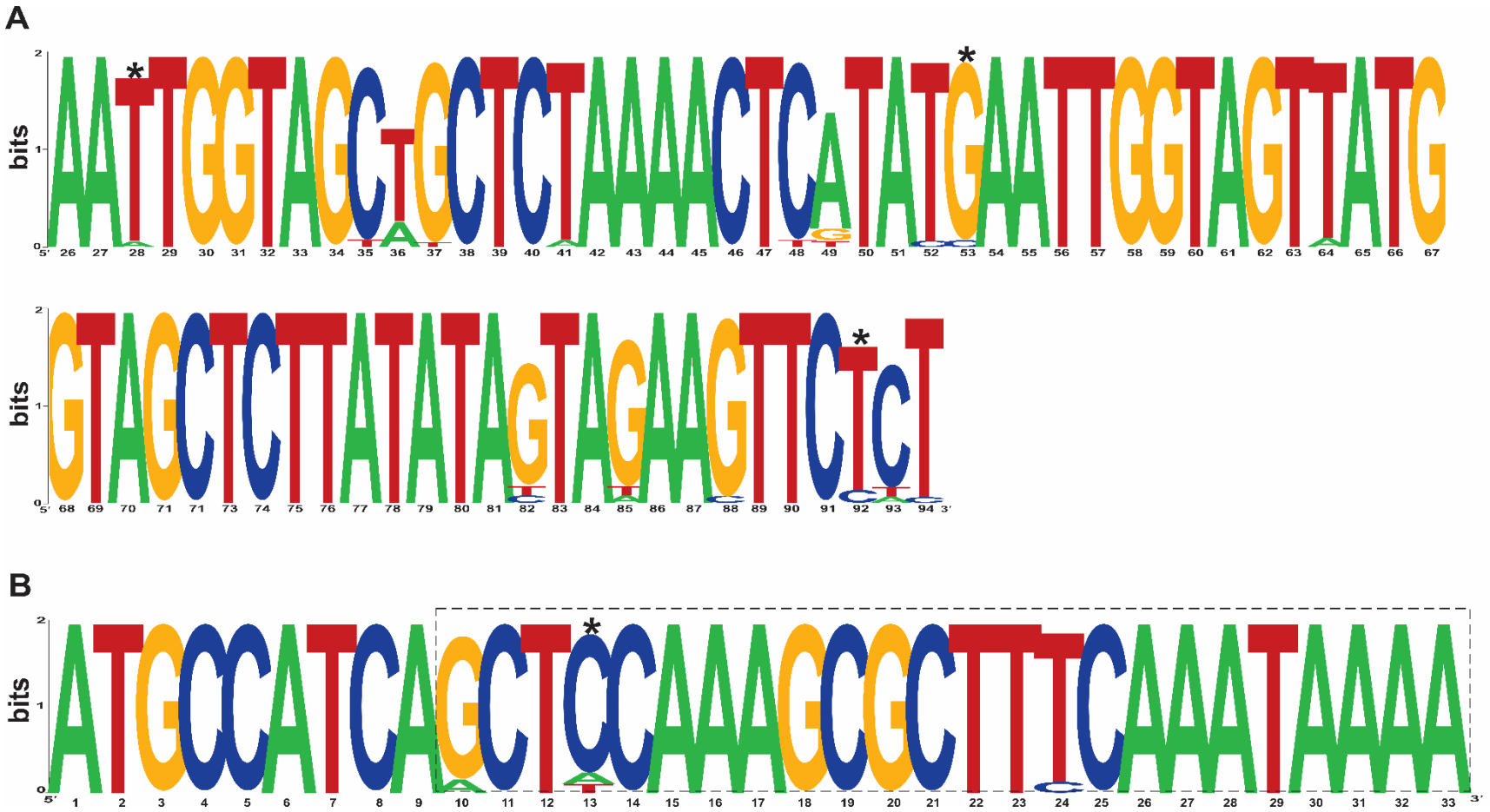
Figure 4


A
ToSRV-Single Infection
a
a
a
a
8.0
7.5
7.0
6.5
log10
ToSRV(wt)
ToSRV(wt)+ ToRMV
ToSRV(ToR:CR)+ ToRMV
ToSRV(ToR:CR+IRD)+ ToRMV
Virus
B
ToRMV-Mixed Infection
a
a
a
a
8
7
6
5
4
log10
ToRMV
ToSRV(wt)+ ToRMV
ToSRV(ToR:CR)+ ToRMV
ToSRV(ToR:CR+IRD)+ ToRMV
Virus
C
ToSRV-Mixed Infection
a
c
ab
bc
8
7
6
5
log10
ToSRV
ToSRV(wt)+ ToRMV
ToSRV(ToR:CR)+ ToRMV
ToSRV(ToR:CR+IRD)+ ToRMV
Virus

Supplementary Figure S1



Supplementary Figure S2



Supplementary Figure S3

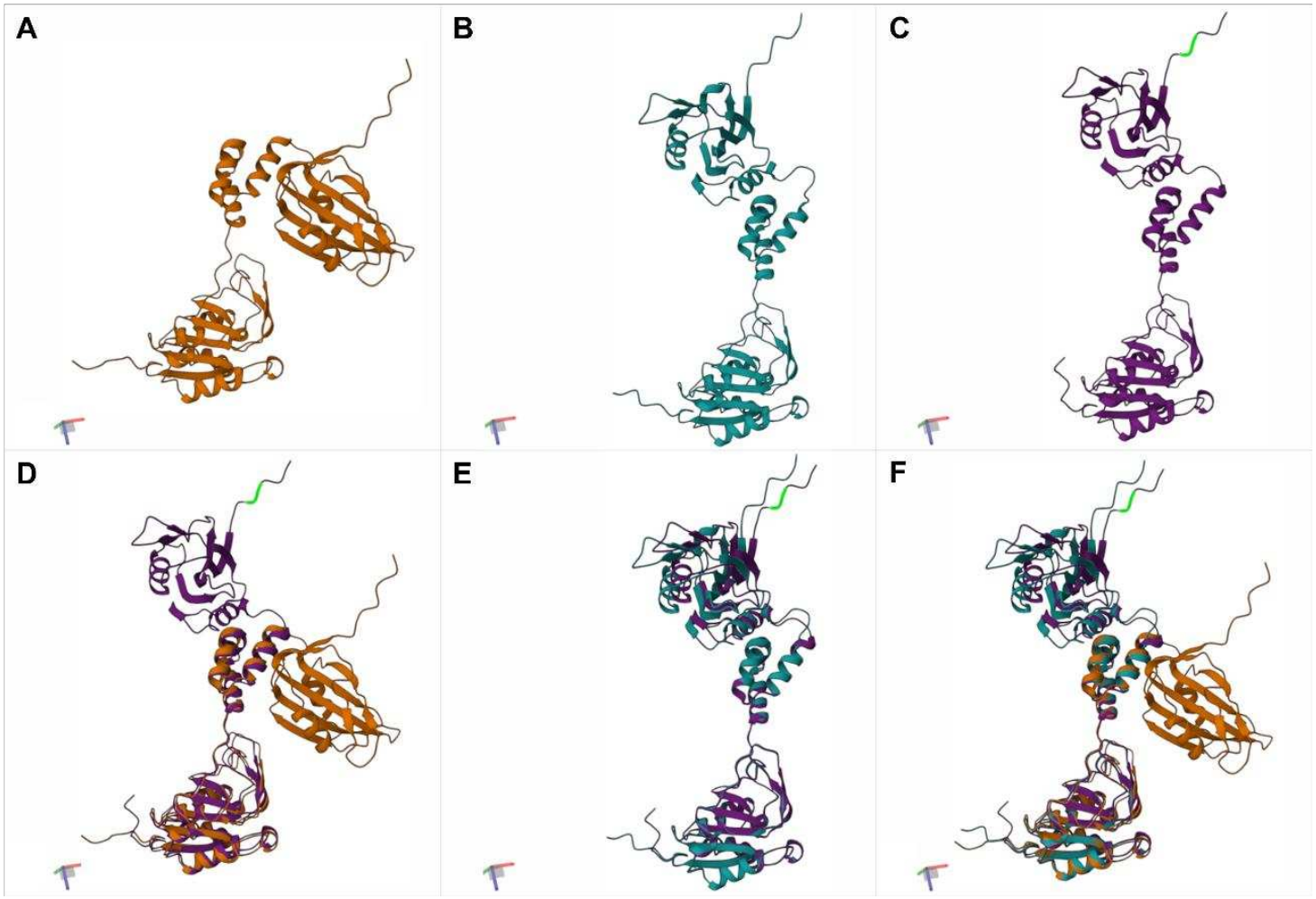


Table S1. Infectivity of tomato severe rugose virus (ToSRV) clones in single infection and in mixed infection with tomato rough mosaic virus (ToRMV).

Treatment*	Genomic component	Infected plants$			
		Exp. I§	Exp. II§	Exp. III	Exp. IV
ToSRV-$A_{(wt)}$	ToS-$A_{(wt)}$#	3/10 (30)	4/10 (40)	4/10 (40)	5/10 (50)
ToRMV-A	ToR-A#	4/10 (40)	3/10 (30)	3/10 (30)	4/10 (40)
ToSRV-$A_{(ToR:CR)}$	ToS-$A_{(ToR:CR)}$#	6/10 (60)	1/10 (10)	7/9 (78)	5/10 (50)
ToSRV-$A_{(ToR:IRD)}$	ToS-$A_{(ToR:IRD)}$#	6/10 (60)	2/10 (20)	10/10 (100)	5/10 (50)
ToSRV-$A_{(ToR:CR+IRD)}$	ToS-$A_{(ToR:CR+IRD)}$#	7/10 (70)	3/10 (30)	7/10 (70)	3/10 (30)
ToSRV-$A_{(ToR:CR)}$/ToRMV-A	ToS-$A_{(ToR:CR)}$+ToR-A&	2/10 (20)	5/10 (50)	7/10 (78)	1/10 (10)
	ToS-$A_{(ToR:CR)}$¶	0/10 (0)	1/10 (10)	2/10 (10)	2/10 (10)
	ToR-A¶	0/10 (0)	1/10 (10)	1/10 (10)	1/10 (10)
	ToS-B+ToR-B&	0/10 (0)	0/10 (0)	7/10 (35)	2/10 (20)
	ToS-B¶	0/10 (0)	0/10 (0)	0/10 (0)	0/10 (0)
	ToR-B¶	2/10 (20)	7/10 (70)	3/10 (20)	2/10 (20)
ToSRV-$A_{(ToR:IRD)}$/ToRMV-A	ToS-$A_{(ToR:IRD)}$+ToR-A&	0/10 (0)	0/10 (0)	1/10 (10)	0/10 (0)
	ToS-$A_{(ToR:IRD)}$¶	2/10 (20)	1/10 (10)	1/10 (10)	2/10 (20)
	ToR-A¶	3/10 (30)	3/10 (30)	1/10 (10)	3/10 (30)
	ToS-B+ToR-B&	1/10 (10)	0/10 (0)	0/10 (0)	0/10 (0)
	ToS-B¶	0/10 (0)	1/10 (10)	0/10 (0)	0/10 (0)
	ToR-B¶	4/10 (40)	3/10 (30)	1/10 (10)	4/10 (40)
ToSRV-$A_{(ToR:CR+IRD)}$/ToRMV-A	ToS-$A_{(ToR:CR+IRD)}$+ToR-A&	2/10 (20)	1/10 (10)	3/9 (33)	0/9 (0)
	ToS-$A_{(ToR:CR+IRD)}$¶	1/10 (10)	2/10 (20)	0/9 (0)	1/9 (11)
	ToR-A¶	1/10 (10)	2/10 (20)	1/9 (11)	1/9 (11)
	ToS-B+ToR-B&	1/10 (10)	0/10 (0)	2/9 (22)	0/9 (0)
	ToS-B¶	0/10 (0)	0/10 (0)	0/9 (0)	0/9 (0)
	ToR-B¶	3/10 (30)	5/10 (50)	3/9 (30)	4/9 (44)
ToSRV-$A_{(wt)}$/ToRMV-A	ToS-$A_{(wt)}$+ToR-A&	2/10 (20)	0/10 (0)	2/10 (20)	0/10 (0)
	ToS-A(¶	2/10 (20)	2/10 (20)	1/10 (10)	2/10 (20)
	ToR-A¶	1/10 (10)	3/10 (30)	4/10 (40)	3/10 (70)
	ToS-B+ToR-B&	1/10 (10)	2/10 (20)	1/10 (10)	0/10 (0)
	ToS-B¶	0/10 (0)	1/10 (10)	0/10 (0)	0/10 (0)
	ToR-B¶	4/10 (40)	2/10 (20)	4/10 (40)	2/10 (20)

* For simplicity, treatments are named based on the DNA-A; wild-type DNA-B components were always inoculated together with their cognate DNA-A components.

$ Number of infected plants/number of inoculated plants (percentage of infected plants), confirmed by PCR with virus-specific primers (single infections) or RCA-RFLP (mixed infections). Four independent treatments were carried out, which are presented in Table 1 as Exp. I (Exp. I+II) and Exp. II (Exp. III+IV).

§ Results from Nogueira (2018).

Plants inoculated with each virus in single infection.

& Plants inoculated with both viruses in which DNA-A or DNA-B of both viruses were detected.

¶ Plants inoculated with both viruses in which DNA-A or DNA-B of only one of the two viruses was detected..

CAPÍTULO 2

ASSESSMENT OF GENETIC BOTTLENECKS DURING WHITEFLY TRANSMISSION OF BEGOMOVIRUSES

Barbosa, T.M.C., Quadros, A.F.F., Bello, V.H., Nogueira, A.M., Morgan, T., Rezende, R.R., Rezende, J.A.M., Zerbini, F.M. (2023) Assessment of genetic bottlenecks during whitefly transmission of begomoviruses. **Virology**, *in preparation*.

Assessment of genetic bottlenecks during whitefly transmission of begomoviruses

Tarsiane M. C. Barbosa[1,2], Ayane F. F. Quadros[1,2], Vinicius Henrique Bello[4], Angélica M. Nogueira[5], Túlio Morgan[3], Rafael Reis de Rezende[3], Jorge A.M. Rezende[4], F. Murilo Zerbini[1,2]

[1]Dep. de Fitopatologia, [2]National Institute for Science and Technology on Plant-Pest Interactions, [3]Dep. de Microbiologia, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [4]Dep. Departamento de Fitopatologia e Nematologia, Universidade de São Paulo/ESALQ, Piracicaba, SP, Brazil; [5]Dep. de Proteção Vegetal, Universidade Estadual Paulista, Botucatu, SP, Brazil

Abstract

When a vector insect feeds on an infected plant it acquires viral particles, which in the case of persistent-circulative transmission will travel through the insect's digestive system reaching the intestinal wall before being delivered to the hemolymph. After reaching the hemolymph, they are directed to the salivary glands from where they will be ejected through the salivary canal into the plant phloem during the insect's feeding. Although vector transmission is considered an important genetic bottleneck for plant viruses, few studies have attempted to estimate its magnitude. Such studies can provide information on how much of the viral diversity can be transferred to another plant during the transmission process, with important implications for the evolution of plant virus populations. In this work we used the begomoviruses Euphorbia yellow mosaic virus (EuYMV) and tomato severe rugose virus (ToSRV) to evaluate the occurrence and magnitude of genetic bottlenecks during transmission by their vector, the whitefly *Bemisia tabaci* Middle East-Asia Minor 1 (MEAM1). We dissected whiteflies after a 48 h acquisition-access period in infected plants, and performed high-throughput sequencing of the viral populations in the midgut (MG) and salivary glands (SG). A larger number of reads mapped to the EuYMV and ToSRV reference genomes in the MG sample compared to SG samples, consistent with the higher viral load in the MG reported in the literature. The analysis of SNPs indicated a significant number of deletions in the EuYMV MG sample and a high number of nucleotide substitutions in the SG samples. Estimates of Shannon's entropy also indicated a greater degree of genetic variability of the SG samples. Despite its limitations, our results suggest that the midgut imposes a genetic bottleneck in *B. tabaci* MEAM1, but that the genetic variability of the viral population is restored during passage through the salivary glands.

Introduction

Transmission is a fundamental step in a viral disease cycle and a key element in understanding the epidemiology of plant viruses, which usually rely on insect vectors to reach new hosts and establish a new infection. Most insect vectors of plant viruses belong to the order Hemiptera (including aphids, leafhoppers and whiteflies), and have piercing-sucking mouthparts (Hull, 2014). Transmission by hemipteran insect vectors can be classified into four modes: non-persistent, semi-persistent, persistent-circulative, and persistent-propagative (Whitfield et al., 2015). Each mode of transmission has certain characteristics concerning the time of acquisition of the virus by the vector when feeding in infected plants, retention in the vector, and inoculation when feeding on healthy plants (Brault et al., 2010; Whitfield et al., 2015).

In persistent and semi-persistent transmission, viruses have only a transient relationship with the insect vector, i.e., they interact only with the mouthparts (mostly the stylet) and/or foregut, have short retention times in the vector, and acquisition and inoculation access periods (AAP and IAP, respectively) are usually in the range of minutes. In persistent-circulative transmission, as the name implies, the virus circulates through the insect's body and interacts with vector cellular factors, with AAP and IAP in the range of hours to days. Moreover, in this mode of transmission, there may also exist a latent period between acquisition and inoculation, which can last from one to several days (Bragard et al., 2013; Gutiérrez et al., 2012; Moury et al., 2007).

Acquisition of virus particles occurs via the stylet when feeding on infected plants; following acquisition, they enter the esophagus, the filter chamber, and are transported through the intestinal wall into the hemolymph after crossing multiple barriers formed by membranes such as epithelial cells and the apical plasma membrane. After reaching the hemolymph, they circulate and pass through the basal lamina of the salivary glands, and then the virus is egested along with the saliva, from the salivary channel to the phloem, during insect feeding (Czosnek

et al., 2017; Ghanim et al., 2007; Hunter et al., 1998). The multiple barriers present in the insect body may serve as bottlenecks for transmission, but there is little experimental evidence in this direction. Studies of transmission bottlenecks may provide information about how much of the viral diversity of one host can be passed on to another during the transmission process, with important implications for the evolution of plant virus populations (McCrone and Lauring, 2018).

Genetic bottlenecks occur when population size is strongly reduced, so that a small number of individuals in a population colonize a new host. The individuals transferred to the new host are called founders, and the number of these founders in relation to the number of individuals in the previous generation is the size of the genetic bottleneck (Zwart and Elena, 2015). These founders come from a small sample of the ancestral population, and the number of viruses capable of replicating and actually contributing to the formation of the next generation determines the effective population size (Charlesworth, 2009).

Transmission bottlenecks that are severe enough to reduce the effective population size will increase genetic drift and reduce the action of selection, thus being a barrier to adaptive evolution (McCrone and Lauring, 2018). Furthermore, the repeated occurrence of genetic bottlenecks will stochastically reduce the genetic variability of the population (Gutiérrez et al., 2012). In the context of viral infections, genetic bottlenecks are important because highly fit genotypes may not be available when genetic variability in the population is low, reducing the overall adaptability of the population (Zwart and Elena, 2015). On the other hand, strong bottlenecks can eliminate defective particles that hinder viral replication (McCrone and Lauring, 2018). Since composition of the founder population after a bottleneck is completely random, both effects may occur over successive generations (Zwart and Elena, 2015).

Although vector transmission is considered an important genetic bottleneck for plant viruses, few studies have attempted to estimate their size and impact on virus evolution. One

such study demonstrating the occurrence of bottlenecks in the non-persistent transmission of cucumber mosaic virus (CMV; *Bromoviridae*) by aphids was conducted by Ali et al. (2006). In this study, 12 restriction enzyme marker mutants were used to inoculate young pumpkin plants. The reduced number of mutants detected relative to the inoculum indicated that transmission by aphids caused a bottleneck in the viral population. Hall et al. (2001) showed that a bottleneck occurs in the transmission of wheat streak mosaic virus (WSMV; *Potyviridae*) by the wheat mite *Aceria tosichella*, although it was not determined whether this occurs at the acquisition or inoculation steps. The occurrence of a narrow bottleneck in the transmission of potato virus Y (PVY; *Potyviridae*) by aphids using infectious and non-infectious but equally transmissible virus variants was investigated by Moury et al. (2007). A low number of virus particles were transmitted when compared to the census population.

The advent of high-throughput sequencing technologies has facilitated genetic analysis of viral populations, and these tools are being used to elucidate the structure of transmissible populations. Brackney et al. (2011) examined the genetic diversity of West Nile virus (WNV; *Flaviviridae*) populations during infection in the midgut, hemolymph, and salivary secretions of the mosquito *Culex pipiens quinquefasciatus*. The results indicated that the insect's anatomical barriers did not cause a genetic bottleneck in the virus population. This and other examples found in the literature provide detailed data on genetic bottlenecks within the insect vector, but do not involve plant viruses.

The vectors for the transmission of begomoviruses are whiteflies of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) (Navas-Castillo et al., 2011). *Bemisia tabaci* is a cryptic species complex comprising at least 39 species, two of which are of major economic importance as highly invasive pests: *B. tabaci* Mediterranean (MED) and *B. tabaci* Middle East-Asia Minor 1 (MEAM1), formerly known as Q and B biotypes, respectively (Vyskočilová et al., 2018). In Brazil, the emergence of begomoviruses in tomato (*Solanum lycopersicum*) is attributed to the

introduction of *B. tabaci* MEAM1 in the early 1990's (Ribeiro et al., 1998). *B. tabaci* MEAM1 is highly polyphagous and efficiently colonizes tomato as well as many other cultivated and non-cultivated plants and can efficiently transmit begomoviruses (Gilbertson et al., 2015). Tomato severe rugose virus (ToSRV) is widely distributed in tomato crops in the central-southern regions of Brazil (Inoue-Nagata et al., 2016) and its widespread incidence could be due to highly efficient transmission by *B. tabaci* MEAM1 (Macedo et al., 2015). In addition to its natural host, ToSRV has been reported to infect several other cultivated and non-cultivated plants, implying that a variety of plants may serve as reservoirs for ToSRV and sources of inoculum (Barbosa et al., 2009; Barreto et al., 2013; Macedo et al., 2015).

Numerous non-cultivated plants have been reported to host begomoviruses. *Euphorbia heterophylla* (wild poinsettia; *Euphorbiaceae*) is a highly invasive non-cultivated plant often associated with soybean and other crop plants in Brazil and Paraguay, where its center of origin is located (Wilson, 1981). *E. heterophylla* is often infected by the begomovirus Euphorbia yellow mosaic virus (EuYMV), showing a bright yellow mosaic (Fernandes et al., 2011; Mar et al., 2017b). Interestingly, *E. heterophylla* can also be infected by ToSRV under experimental conditions, although virus titer is low (Barreto et al., 2013).

A low degree of genetic variability has been reported for both EuYMV and ToSRV populations (Mar et al., 2017b; Xavier et al., 2021), which is unusual for begomoviruses (García-Arenal and Zerbini, 2019; Rocha et al., 2013). Whether this low variability is related to genetic bottlenecks due to vector transmission is unknown. Thus, EuYMV and ToSRV constitute two good systems to evaluate the occurrence and magnitude of genetic bottlenecks during vector transmission.

In this study, we dissected whiteflies after a 48h AAP in EuYMV-infected *E. heterophylla* and ToSRV-infected tomato, and sequenced the viral populations in the midgut and salivary glands. Our results suggest that the midgut imposes a genetic bottleneck in *B.*

tabaci MEAM1, but that the genetic variability of the viral population may be restored during passage through the salivary glands.

Material and Methods

Viral isolates, plant inoculations and viral detection

Infectious clones of the begomoviruses Euphorbia yellow mosaic virus (EuYMV) BR:Cha510:10 (GenBank access numbers KY559518 and KY559617 for the DNA-A and DNA-B, respectively; Mar et al., 2017a) and tomato severe rugose virus (ToSRV) BR:PG1:Pep:03 (DQ207749 and EF534708 for the DNA-A and DNA-B, respectively; Silva et al., 2014) were used in the experiments. Tomato (*Solanum lycopersicum*) cv. Santa Clara and *E. heterophylla* seedlings were biolistically inoculated (Aragão et al., 1996) using 10 μg of each genomic component of ToSRV and EuYMV, respectively. Plants were also bombarded with DNA-free gold particles as a negative control. All plants were kept in a greenhouse and their phenotype was evaluated at 28 days after inoculation (dpi).

Detection of genomic components

After 28 dpi, the total DNA of all plants was extracted according to Doyle and Doyle (1987). Infection was confirmed by PCR using the universal begomovirus primers PALv1978 and PAR1c496 (Rojas et al., 1993). Reactions contained 1 μL of total DNA, 0.4 μM of each primer, 1x Reaction Buffer Complete (Celco), 0.2 mM of each dNTP, and 1.25 U of Taq DNA Polymerase (5 U/μl) (Celco) in a final volume of 25 μL. PCR parameters consisted of an initial denaturation step at 95°C for 2 min, followed by 38 cycles at 95°C for 1 min, 66°C for 1 min, and 72°C for 1 min, followed by a final extension at 72°C for 10 minutes. PCR products were

separated by 1% agarose gel electrophoresis and stained with ethidium bromide. After confirming the positive plants, four of them were selected per treatment of each experiment for the whitefly acquisition assay.

Whitefly colony maintenance and acquisition of ToSRV and EuYMV by adult insects

B. tabaci Middle East-Asia Minor 1 (MEAM1) were reared on cabbage plants grown in insect-proof, temperature-controlled greenhouses. Virus-infected tomato and *E. heterophylla* plants were placed in insect-proof wooden cages, and whiteflies were placed for virus acquisition. Approximately 100 whiteflies were caged with infected plants for 24 h and 48 h acquisition access periods (AAP). The insects were then collected with a mouth aspirator and used for dissection.

Dissection of individual whitefly organs

After AAPs of 24 h and 48 h, the insects were collected and maintained at 4ºC for 20 minutes to stop their movement. The insects were dissected on a glass slide under a binocular microscope (×40). Dissections were performed in TE buffer pH 8.0 with or without 1% bromophenol blue, as described by Ghanim et al. (1998). The isolated tissues were placed in a PCR tube containing 20 µl of TE buffer and a metallic wire fixed to a plastic rod was used for dissection. To obtain the midgut, the abdomen was separated from the thorax at the connection between them, and its contents were expelled by gently pushing the abdomen. To isolate the salivary glands, the prothorax was separated from the mesothorax and abdomen. The salivary glands were removed from the body in a 1% bromophenol blue solution so that the organ could absorb the dye, for easier visualization and dissection.

DNA extraction, virus enrichment and sequencing

Approximately 30 units from each organ were pooled and used for DNA extraction with the PureLink Viral RNA/DNA Mini kit (Invitrogen) followed by RCA using the phi29 DNA Polymerase Amplification kit (New England Biolabs). RCA products were checked on 1% agarose gels and sent out for sequencing in the Illumina NovaSeq 6000 platform at Macrogen, Inc. (Seoul, South Korea).

Quality control and filtering

The sequenced libraries corresponding to each organ were received and data processing was performed followed by bioinformatics analysis. Reads were paired with a size of 151 nucleotides each. The quality of the reads was checked using FastQC v. 0.11.5 (bioinformatics.babraham.ac.uk/projects/fastqc). The reads were trimmed using Trimmomatic v. 0.36 (Bolger et al., 2014) and filtered for a quality index >30, corresponding to a precision of 99.99%, resulting in high homogeneity in all readings.

Mapping of readings and analysis of SNPs

Bowtie version 2.2.8 (Langmead and Salzberg, 2012) with default parameters was used to map the reads from the reference genomes corresponding to the ToSRV and EuYMV clones. The output files were converted from sequence alignment map (SAM) to binary alignment map (BAM) using SAMtools. Then, the Tablet program (Milne et al., 2010) was used to visualize the mapped reads. Three SNP callers were used to perform site prediction of SNPs from alignments with the reference DNA sequences: BCFTOOLS v. 1.16 (Li et al., 2009), Samtools v. 1.6 (Li et al., 2009) and Freebayes v. 1.3.1 (Garrison and Marth, 2012).

Shannon's entropy estimate

To assess the viral diversity of the sequencing data, the DiversiTools program (josephhughes.github.io/DiversiTools/) was used to calculate the derivation of the reference sequence based on Shannon's entropy:

$$S = \frac{1}{n}\sum_{i=1}^{n}[((fi * \log(fi)) + (1 - fi) * \log(1 - fi))) * -1]$$

where S = total deviation from the reference sequence, N = total number of sites in the sequence, and f*i* = sum of frequency variations at each site.

Results

Detection of ToSRV and EuYMV in inoculated plants prior to vector acquisition

To verify the occurrence and magnitude of genetic bottlenecks during vector transmission, tomato and *E. heterophilla* plants were inoculated with infectious clones of the begomoviruses Euphorbia yellow mosaic virus (EuYMV) and tomato severe rugose virus (ToSRV). Symptoms on inoculated plants began to appear at 10-14 days post-inoculation (dpi). Symptom in tomato plants consisted of a mild mosaic and leaf distortion, whereas *E. heterophylla* showed yellow mosaic, vein chlorosis and leaf distortion (Suppl. Figure S1). Infections were confirmed by PCR at 28 dpi, and positive plants were selected for the whitefly acquisition assay.

Viral acquisition by* B. tabaci *MEAM1 and organ dissection

Acquisition access periods of 24 hours and 48 hours were performed, and the salivary glands and intestines were extracted from the insects (Suppl. Figure S2), with two replications of each treatment being performed. Then, DNA extraction was carried out with a kit for the

specific enrichment of viral DNA. Subsequently, the viral DNA was amplified by RCA. After amplification and quality checking, the samples were sequenced on the Illumina NovaSeq 6000 platform. Although we performed two AAPs, only the 48-hour treatments were sequenced. The 24-hour treatments were stored at -80^{o}C for future analysis.

Sequence quality and read mapping

Between 13 and 15 million reads were obtained for the different libraries, with an average length of 150 bp per read. To achieve homogeneity across all reads, we filtered all libraries for a quality score of Q30, which corresponds to an accuracy of 99.99% (Figures 1-4). The Q30 quality threshold was 90.9% and 82.8% for samples of midgut (MG) and salivary glands (SG), respectively, from insects that acquired EuYMV (Suppl. Table S1). For insects that acquired ToSRV, it was 72.2% for the SG. Unfortunately, the ToSRV MG sample could not be sequenced and will need to be repeated in the future.

In the ToSRV-SG samples, 445,987 reads (mismatches = 2.5%) and 392,698 reads (mismatches = 4.3%) were mapped to the ToSRV DNA-A and DNA-B reference genomes, respectively (Table 1). For the EuYMV-SG sample, 215,836 readings (mismatches = 5.8%) and 754,417 readings (mismatches = 0.4%) were mapped to the EuYMV DNA-A and DNA-B reference genomes, respectively (Table 1), and for the EuYMV-MG sample, 1,028,123 readings (mismatches = 0.3%) and 757,471 readings (mismatches = 0.4%) were mapped to the DNA-A and DNA-B, respectively (Table 1). All mappings showed 100% coverage (Figures 1-4).

Detection of SNPs by different call pipelines and entropy analysis

The number of SNPs occurring in the viral genome during passage through the vector's body was assessed using three different variant callers: Samtools, BCFtool and Freebayes (Figure 5 and Suppl. Tables S2-S4).

Samtools detected a very low number of SNPs in all treatments (Figure 5 and Suppl. Table S2). No SNPs were detected in ToSRV-SG DNA-B and in the two components of EuYMV-MG. Only three SNPs were detected in EuYMV-SG DNA-A, at positions 1,060 and 1,480 (Vf: 33.33% and 69.52%, respectively) in the region were the *Ren* and *TrAP* genes overlap, and at position 1,920 in the region where the *Rep* and *TrAP* genes overlap. Also, only three SNPs were detected in the EuYMV-SG DNA-B, at positions 524 and 647 (Vf: 14.08% and 8.82%, respectively) in the *NSP* gene, and 1787 (Vf: 13.4%) in the *MP* gene. All observed variations were nucleotide substitutions.

When variant calling was performed with BCFtools, all observations were deletions. We observed a low number of these deletions for ToSRV-SG (Figure 5), with five in the DNA-A between positions 1609 to 1613, all located in the *Rep* gene. In the DNA-B, five deletions were also observed, but none of them were found in coding regions. For EuYMV-SG we did not detect any variation with BCFtools (Figure 5). Strikingly, a total of 324 deletions were found in the EuYMV-MG DNA-A, and 151 in the DNA-B. Of these, 66 were in the *TrAP* gene, 45 were in the region where *Rep* and *TrAP* overlap, and 132 in the region where *Rep* and *AC4* overlap. In the DNA-B, 121 of the 151 observed deletions were in the *MP* gene.

With Freebayes, 30 SNPs were observed in the ToSRV-SG DNA-A (Figure 5 and Suppl. Table S3). Seventeen of these variations were distributed along the DNA-A genes (Vf: *CP*: 9-92%; *REn*: 70-100%; *TrAP*: 78-100 %; *Rep*: 5-92%; *AC4*: 10-93%) (Suppl. Table S3), with two of them being nucleotide substitutions. A total of 90 SNPs were detected in the ToSRV-SG DNA-B, with 54 distributed in the coding regions (Vf: *NSP*: 6-78%; *MP*: 5-99%)

(Suppl. Table S3). SNPs were distributed over all genes in the EuYMV-SG mapping. A total of 116 variations were observed in the DNA-A, of which 105 were in coding regions (Vf: *CP*: 5-41%; *REn*: 5-42%; *TrAP*: 5-68%; *Rep*: 5-80 %; *AC4*: 7-56%). All variations were nucleotide substitutions (Suppl. Table S4A). For the DNA-B, 122 variations were detected and 84 of these were located in coding regions (Vf: *NSP*: 5-79%; *MP*: 5-91%). Of these 84 variations, we observed in position 2007 an insertion event located in the MP region (T > C,G; Vf: 6.8%) (Suppl. Table S4B). No variations were detected in EuYMV-MG using Freebayes (Figure 5).

In addition to analyzing SNPs, we estimated Shannon's entropy, a second metric for estimating viral diversity based on the level of "chaos" in a population (Shannon, 1948). EuYMV-SG had much higher entropy compared to EuYMV-MG for both the DNA-A and DNA-B (Figure 6). Similar values were observed for EuYMV-SG and ToSRV-SG DNA-B, with a slightly lower value observed for the ToSRV-SG DNA-A (Figure 6).

Discussion

Although genetic bottlenecks play an important role in the evolutionary dynamics of viruses, few studies have focused on determining aspects of their occurrence during vector transmission. The transmission of begomoviruses by *B. tabaci* occurs in a persistent-circulatory manner, in which, after being acquired, viral particles travel throughout the insect's digestive system, hemolymph and eventually accumulate in the salivary glands to be subsequently transmitted. The insect organs may act as barriers for viral particles and thus constitute genetic bottlenecks. To test this hypothesis, we sequenced the viral population in the midgut and salivary glands of *B. tabaci* MEAM1 after a 48 h AAP in begomovirus-infected tomato and *E. heterophylla* plants.

The first major challenge encountered was to obtain large enough amounts of viable virus for HTS. We pooled 30 organs, proceed to DNA extraction and attempted to sequence

this material without an amplification step using the Illumina MiSeq platform. However, the reads obtained mapped exclusively to the whitefly genome. The vast majority of the studies to detect and identify geminiviruses by HTS have used rolling circle amplification (RCA) as an amplification step prior to sequencing (Bornancini et al., 2020; Dayaram et al., 2012; Hagen et al., 2012; Idris et al., 2014). We have avoided this in our previous studies (Pinto et al., 2021; Quadros et al., 2023) as RCA can lead to amplification bias (Gallet et al., 2017). However, some studies have shown that DNA accumulation of the begomoviruses tomato yellow leaf curl virus (TYLCV) and *tomato yellow leaf curl Sardinia virus* (TYLCSV) in the body of *B. tabaci* reaches a maximum load of $6x10^8$ genomic units after 24 h and 48 h AAPs (Czosnek et al., 2001; Zeidan and Czosnek, 1991), which may be too low for HTS. Furthermore, these works evaluated the entire body of the insect, where we intended to evaluate individual organs, which would certainly lead to an even lower amount of extracted DNA, even using an extraction kit aimed at enrichment of viral DNA. Thus, we decided to carry out an amplification step (RCA) and use the Illumina HiSeq 6000 platform, which yields a much higher output than the MiSeq platform.

Mapping to the reference genomes showed that a larger number of reads mapped in the midgut (MG) sample than in the salivary gland (SG) samples, with no major differences between the mapped reads for each genome component. This was somewhat expected since previous studies show that the highest concentration of viral DNA occurs in the MG. For example, Wei et al. (2014), using conventional PCR, detected higher amounts of tomato yellow leaf curl China virus (TYLCCNV) in the MG compared to the SG, saliva and hemolymph of *B. tabaci* MEAM1 and MED after a 48 h AAP.

The detection of SNPs can be useful for evaluating the genetic variability of a population in relation to the original sequence. The three variant calling programs used here, Samtools, BCFtools, and Freebayes, have different characteristics and can generate different data but are

all capable of extracting useful and complementary information. Several works have mentioned the use of more than one variant calling method to increase the sensitivity of the analysis (Cheng et al., 2014; Gézsi et al., 2015; Liu et al., 2013; O'Rawe et al., 2013; Yu and Sun, 2013). Thus, it was not surprising that we obtained discrepant data from each one of the programs.

When the insect feeds on an infected plant, the virus passes through the food canal in the stylet and reaches the foregut and intestines (midgut and hindgut). As mentioned above, the midgut is the site where there is higher accumulation and longer retention of viral DNA (Pakkianathan et al., 2015; Wei et al., 2014). So, after acquisition, a large number of viral particles reach the MG and these can remain there for a relatively long time (Czosnek et al., 2017). Nevertheless, autophagic pathways and other defense pathways can be activated in the MG that can lead to viral degradation (Geng et al., 2018). We analyzed our samples 48 h after acquisition, corresponding to the period in which autophagy begins (Wang et al., 2016). These observations could explain, first, the large number of reads mapped in the MG, and second, the large number of deletions in the viral genome observed for the MG sample. These hypotheses are supported by the fact that TYLCV DNA could be detected by PCR in the MG after a 5-10 min AAP (Czosnek et al., 2002). Unfortunately, the ToSRV MG library was not sequenced and therefore we were unable to validate these observations for a second begomovirus.

After traveling from the food canal to the midgut and hemolymph, the next stop for the virus are the salivary glands, and only a small number of viruses are able to cross the hemolymph and reach the salivary glands (Czosnek et al., 2017). SNP analysis indicated that a much higher number of nucleotide substitutions and insertions occurred in the salivary glands compared to the midgut. In addition to higher variation in the SG, it is noted that there was a difference between the viruses, with ToSRV presenting a larger number of SNPs compared to EuYMV.

Shannon's entropy is a variability metric where zero indicates no variability (all individuals have the same nucleotide at each alignment position) and one indicates maximum variability (the four nucleotides occur at the same proportion at each alignment position). We obtained strikingly similar entropy values for EuYMV-SG DNA-A and DNA-B (0.16 for both components) and for ToSRV-SG DNA-B (0.152), with a slightly lower value for ToSRV-SG DNA-A (0.09). Entropy values for EuYMV-MG were much lower (0.025 and 0.031 for the DNA-A and DNA-B, respectively). Although both viruses have low genetic variability, EuYMV has higher variability when compared to ToSRV, and for both viruses the DNA-B is more variable than the DNA-A (Xavier et al., 2021). These results are consistent with the higher number of nucleotide substitutions observed in the SG samples, and indicate a larger degree of genetic variability of the viral population in the SG compared to the MG. Under these conditions, a high number of diverse viral genomes could be transmitted to the next host. However, it has been demonstrated that the occurrence of mutations in the CP hinders the interaction of viral particles with the principal salivary gland (PSG) (Czosnek et al., 2017). To answer the question of whether the viral population that is present in the SG can be efficiently transmitted to establish a new infection, transmission assays with the same clones used in this work need to be conducted and the viral population in the newly infected plants needs to be sequenced.

A number of studies have shown that there may be population segregation, causing selection within the vector, which can restrict the viral diversity to be transmitted to the recipient (McCrone and Lauring, 2018). This is because the elimination of viral particles may occur due to the action of the insect's immune system, and only a small number of viral particles may be able to cross the barriers imposed by the organs (Zwart and Elena, 2015). Under this perspective, genetic bottlenecks could actually have a positive effect as they would eliminate defective particles that hinder viral replication (McCrone and Lauring, 2018). As a greater number of deletion mutants were observed in the MG, the midgut may indeed be acting as an important

barrier to the elimination of defective genomes. At the same time, we observed the occurrence of a significantly high number of nucleotide substitutions for the viral population in the SG, with no substitutions being observed in the MG population. Thus, genetic variability seems to be restored in the SG. Begomovirus replication in the vector has been a controversial topic for decades (Czosnek et al., 2001; Rubinstein and Czosnek, 1997; Sánchez-Campos et al., 2016). A recent study indicated that replication does take place but only in the salivary glands during transmission of TYLCV by *B. tabaci* MEAM1 and MED (He et al., 2020). Thus, it is not unreasonable to consider that EuYMV and ToSRV could be replicating in the SG, which could explain the genetic variability observed in the SG samples.

We performed viral DNA extraction from both the PSG and the accessory salivary gland (ASG) together. It has been shown that the interaction between the viral CP and specific cells of the PSG can determine which particles are able to be transmitted (Wei et al., 2014). We attempted to obtain the two SGs separately, however, due to the small size of the PSG and of the ASG ducts, we were unable to extract them separately. Finally, it should be noted that neither the inoculum source plants nor the recipient plants were sequenced in our study, which could have provide additional evidence for the hypotheses presented here.

Despite its limitations, our results provide evidence that the midgut barrier acts as a genetic bottleneck during begomovirus transmission by whiteflies, and also that genetic variability is somehow restored during viral passage through the salivary glands (maybe due to viral replication in that organ). To the best of our knowledge, this is the first work to use HTS in whitefly organs to assess viral genetic variability. Our future studies will build on the hypothesis presented here, to confirm the genetic bottleneck at the midgut and to investigate how genetic variability is restored in the salivary glands.

References

Ali, A., Li, H., Schneider, W.L., Sherman, D.J., Gray, S., Smith, D., Roossinck, M.J., 2006. Analysis of genetic bottlenecks during horizontal transmission of cucumber mosaic virus. J Virol 80, 8345-8350.

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

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.

Bornancini, V.A., Irazoqui, J.M., Flores, C.R., Vaghi Medina, C.G., Amadio, A.F., López Lambertini, P.M., 2020. Reconstruction and characterization of full-length begomovirus and alphasatellite genomes infecting pepper through metagenomics. Viruses 12, 202.

Brackney, D.E., Pesko, K.N., Brown, I.K., Deardorff, E.R., Kawatachi, J., Ebel, G.D., 2011. West Nile virus genetic diversity is maintained during transmission by *Culex pipiens quinquefasciatus* mosquitoes. PLoS ONE 6, e24466.

Bragard, C., Caciagli, P., Lemaire, O., Lopez-Moya, J.J., MacFarlane, S., Peters, D., Susi, P., Torrance, L., 2013. Status and prospects of plant virus control through interference with vector transmission. Annu Rev Phytopathol 51, 177-201.

Brault, V., Uzest, M., Monsion, B., Jacquot, E., Blanc, S., 2010. Aphids as transport devices for plant viruses. C R Biol 333, 524-538.

Charlesworth, B., 2009. Effective population size and patterns of molecular evolution and variation. Nat Rev Genet 10, 195-205.

Cheng, A.Y., Teo, Y.Y., Ong, R.T., 2014. Assessing single nucleotide variant detection and genotype calling on whole-genome sequenced individuals. Bioinformatics 30, 1707-1713.

Czosnek, H., Ghanim, M., Ghanim, M., 2002. The circulative pathway of begomoviruses in the whitefly vector *Bemisia tabaci*-insights from studies with *Tomato yellow leaf curl virus*. Ann Appl Biol 140, 215-232.

Czosnek, H., Ghanim, M., Morin, S., Rubinstein, G., Fridman, V., Zeidan, M., 2001. Whiteflies: vectors, and victims (?), of geminiviruses. Adv Virus Res 57, 291-322.

Czosnek, H., Hariton-Shalev, A., Sobol, I., Gorovits, R., Ghanim, M., 2017. The incredible journey of begomoviruses in their whitefly vector. Viruses 9, 273.

Dayaram, A., Opong, A., Jäschke, A., Hadfield, J., Baschiera, M., Dobson, R.C., Offei, S.K., Shepherd, D.N., Martin, D.P., Varsani, A., 2012. Molecular characterisation of a novel cassava associated circular ssDNA virus. Virus Res 166, 130-135.

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

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

Gallet, R., Fabre, F., Michalakis, Y., Blanc, S., 2017. The number of target molecules of the amplification step limits accuracy and sensitivity in ultradeep-sequencing viral population studies. J Virol 91, e00561-00517.

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, 411-433.

Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. arXiv, 1207.3907.

Geng, L., Qian, L.X., Shao, R.X., Liu, Y.Q., Liu, S.S., Wang, X.W., 2018. Transcriptome profiling of whitefly guts in response to tomato yellow leaf curl virus infection. Virol J 15, 14.

Gézsi, A., Bolgár, B., Marx, P., Sarkozy, P., Szalai, C., Antal, P., 2015. VariantMetaCaller: automated fusion of variant calling pipelines for quantitative, precision-based filtering. BMC Genomics 16, 875.

Ghanim, M., Morin, S., Zeidan, M., Czosnek, H., 1998. Evidence for transovarial transmission of tomato yellow leaf curl virus by its vector, the whitefly *Bemisia tabaci*. Virology 240, 295-303.

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

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.

Gutiérrez, S., Michalakis, Y., Blanc, S., 2012. Virus population bottlenecks during within-host progression and host-to-host transmission. Curr Opin Virol 2, 546-555.

Hagen, C., Frizzi, A., Gabriels, S., Huang, M.Y., Salati, R., Gabor, B., Huang, S.S., 2012. Accurate and sensitive diagnosis of geminiviruses through enrichment, high-throughput sequencing and automated sequence identification. Arch Virol 157, 907-915.

Hall, J.S., French, R., Hein, G.L., Morris, T.J., Stenger, D.C., 2001. Three distinct mechanisms facilitate genetic isolation of sympatric *Wheat streak mosaic virus* lineages. Virology 236, 230-236.

He, Y.Z., Wang, Y.M., Yin, T.Y., Fiallo-Olive, E., Liu, Y.Q., Hanley-Bowdoin, L., Wang, X.W., 2020. A plant DNA virus replicates in the salivary glands of its insect vector via recruitment of host DNA synthesis machinery. Proc Natl Acad Sci USA 117, 16928-16937.

Hull, R., 2014. Plant Virology, 5th Ed. ed. Elsevier Academic Press, San Diego, USA.

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

Idris, A., Al-Saleh, M., Piatek, M.J., Al-Shahwan, I., Ali, S., Brown, J.K., 2014. Viral metagenomics: analysis of begomoviruses by illumina high-throughput sequencing. Viruses 6, 1219-1236.

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.

Langmead, B., Salzberg, S., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9, 357.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25.

Liu, X., Han, S., Wang, Z., Gelernter, J., Yang, B.Z., 2013. Variant callers for next-generation sequencing data: a comparison study. PLoS ONE 8, e75619.

Macedo, M.A., Michereff, M., Navas-Castillo, J., Inoue-Nagata, A.K., 2015. Host range and whitefly transmission efficiency of *Tomato severe rugose virus* and *Tomato golden vein virus* in tomato plants. Trop Plant Pathol 40, 405-409.

Mar, T.B., Mendes, I.R., Lau, D., Fiallo-Olive, E., Navas-Castillo, J., Alves, M.S., Zerbini, F.M., 2017a. Interaction between the New World begomovirus *Euphorbia yellow mosaic virus* and its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. J Gen Virol 98, 1552-1562.

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., 2017b. Genetic variability and population structure of the New World begomovirus *Euphorbia yellow mosaic virus*. J Gen Virol 98, 1537-1551.

McCrone, J.T., Lauring, A.S., 2018. Genetic bottlenecks in intraspecies virus transmission. Curr Opin Virol 28, 20-25.

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., Marshall, D., 2010. Tablet - Next Generation Sequence assembly visualization. Bioinformatics 26, 401-402.

Moury, B., Fabre, F., Senoussi, R., 2007. Estimation of the number of virus particles transmitted by an insect vector. Proc Natl Acad Sci USA 104, 17891-17896.

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

O'Rawe, J., Jiang, T., Sun, G., Wu, Y., Wang, W., Hu, J., Bodily, P., Tian, L., Hakonarson, H., Johnson, W.E., Wei, Z., Wang, K., Lyon, G.J., 2013. Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. Genome Med 5, 28.

Pakkianathan, B.C., Kontsedalov, S., Lebedev, G., Mahadav, A., Zeidan, M., Czosnek, H., Ghanim, M., 2015. Replication of Tomato Yellow Leaf Curl Virus in Its Whitefly Vector, Bemisia tabaci. J Virol 89, 9791-9803.

Pinto, V.B., Quadros, A.F.F., Godinho, M.T., Silva, J.C., Alfenas-Zerbini, P., Zerbini, F.M., 2021. Intra-host evolution of the ssDNA virus tomato severe rugose virus (ToSRV). Virus Res 292, 198234.

Quadros, A.F.F., Ferro, C.G., de Rezende, R.R., Godinho, M.T., Xavier, C.A.D., Nogueira, A.M., Alfenas-Zerbini, P., Zerbini, F.M., 2023. Begomovirus populations in single plants are complex and may include both well-adapted and poorly-adapted viruses. Virus Res 323, 198969.

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.

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.

Rubinstein, G., Czosnek, H., 1997. Long-term association of tomato yellow leaf curl virus with its whitefly vector *Bemisia tabaci*: Effect on the insect transmission capacity, longevity and fecundity. Virology 78, 2683-2689.

Sánchez-Campos, S., Rodriguez-Negrete, E.A., Cruzado, L., Grande-Pérez, A., Bejarano, E.R., Navas-Castillo, J., Moriones, E., 2016. Tomato yellow leaf curl virus: No evidence for replication in the insect vector Bemisia tabaci. Sci Rep 6, 30942.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal 27, 379-423.

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

Vyskočilová, S., Tay, W.T., van Brunschot, S., Seal, S., Colvin, J., 2018. An integrative approach to discovering cryptic species within the *Bemisia tabaci* whitefly species complex. Sci Rep 8, 10886.

Wang, L.L., Wang, X.R., Wei, X.M., Huang, H., Wu, J.X., Chen, X.X., Liu, S.S., Wang, X.W., 2016. The autophagy pathway participates in resistance to tomato yellow leaf curl virus infection in whiteflies. Autophagy 12, 1560-1574.

Wei, J., Zhao, J.J., Zhang, T., Li, F.F., Ghanim, M., Zhou, X.P., Ye, G.Y., Liu, S.S., Wang, X.W., 2014. Specific cells in the primary salivary glands of the whitefly *Bemisia tabaci* control retention and transmission of begomoviruses. J Virol 88, 13460-13468.

Whitfield, A.E., Falk, B.W., Rotenberg, D., 2015. Insect vector-mediated transmission of plant viruses. Virology 480, 278-289.

Wilson, A.K., 1981. *Euphorbia heterophylla*: A review of distribution, importance and control. Trop Pest Manag 27, 32-38.

Xavier, C.A.D., Godinho, M.T., Mar, T.B., Ferro, C.G., Sande, O.F.L., Silva, J.C., Ramos-Sobrinho, R., Nascimento, R.N., Assuncao, I., Lima, G.S.A., Lima, A.T.M., Zerbini, F.M., 2021. Evolutionary dynamics of bipartite begomoviruses revealed by complete genome analysis. Mol Ecol 15, 3747-3767.

Yu, X., Sun, S., 2013. Comparing a few SNP calling algorithms using low-coverage sequencing data. BMC Bioinformatics 14, 274.

Zeidan, M., Czosnek, H., 1991. Acquisition of tomato yellow leaf curl virus by the whitefly Bemisia tabaci. J Gen Virol 72, 2607-2614.

Zwart, M.P., Elena, S.F., 2015. Matters of size: genetic bottlenecks in virus infection and their potential impact on evolution. Annu Rev Virol 2, 161-179.

Table 1. Total number of mapped reads for each genomic component of tomato severe rugose virus (ToSRV) and Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary gland and midgut of *B. tabaci* MEAM1. In addition to the number of reads for each component, the ratio between DNA-A and DNA-B reads was calculated. Mapping was done using Bowtie2. In parentheses is the percentage of mismatches for each mapping.

	Salivary Glands			**Midgut**		
Virus	**DNA-A**	**DNA-B**	**DNA-A/ DNA-B**	**DNA-A**	**DNA-B**	**DNA-A/ DNA-B**
EuYMV	215,836 (5.8 %)	754,417 (0.4%)	0.286	1,028,123 (0.3%)	757,471 (0.4%)	1.357
ToSRV	445,987 (2.5%)	392,698 (4.3%)	1.135	n.d.*	n.d.	n.d.

* n.d., not done.

Figure legends

Figure 1. Mapping (upper panels), coverage (middle panels), and mapping of SNPs (lower panel) in the DNA-A of tomato severe rugose virus (ToSRV) sampled in the salivary glands of *B. tabaci* MEAM1 after a 48 h acquisition period.

Figure 2. Mapping (upper panels), coverage (middle panels), and mapping of SNPs (lower panel) in the DNA-B of tomato severe rugose virus (ToSRV) sampled in the salivary glands of *B. tabaci* MEAM1 after a 48 h acquisition period.

Figure 3. Mapping (upper panels), coverage (middle panels), and SNPs mapping (lower panel) in the DNA-A of Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary glands and midgut of *B. tabaci* MEAM1 sfter a 48 h acquisition period.

Figure 4. Mapping (upper panels), coverage (middle panels), and SNPs mapping (lower panel) in the DNA-B of Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary glands and midgut of *B. tabaci* MEAM1 after a 48 h acquisition period.

Figure 5. Number of SNPs detected by different calling pipelines in the DNA-A and DNA-B of tomato severe rugose virus (ToSRV) and Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary glands and midgut of *B. tabaci* MEAM1 after a 48 h acquisition period.

Figure 6. Estimates of Shannon's entropy for the DNA-A and DNA-B of tomato severe rugose virus (ToSRV) and Euphorbia yellow mosaic virus (EuYMV) genome, sampled in the salivary glands and midgut of *B. tabaci* MEAM1 after a 48 h acquisition period.

Figure 1

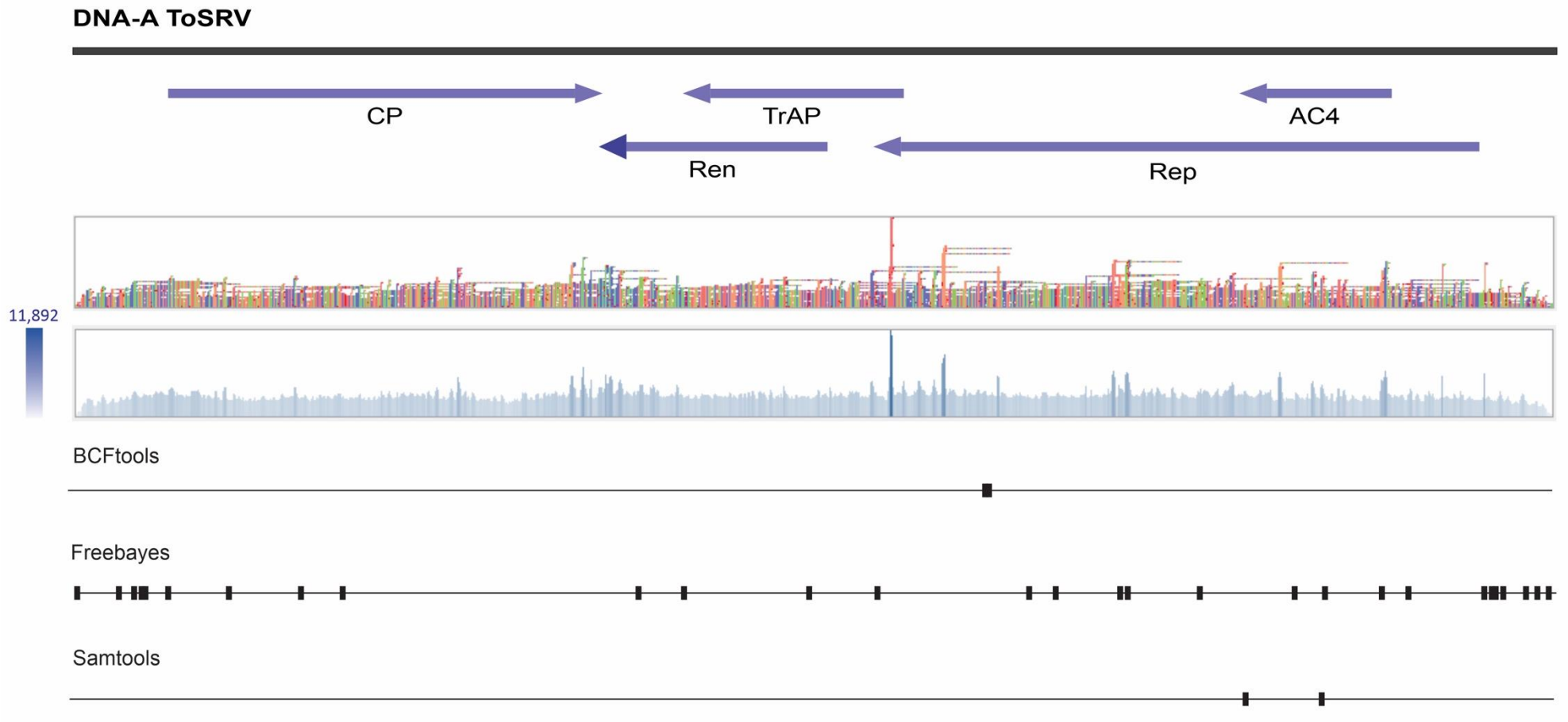

DNA-A ToSRV
CP
TrAP
AC4
Ren
Rep
11,892
BCFtools
Freebayes
Samtools

Figure 2

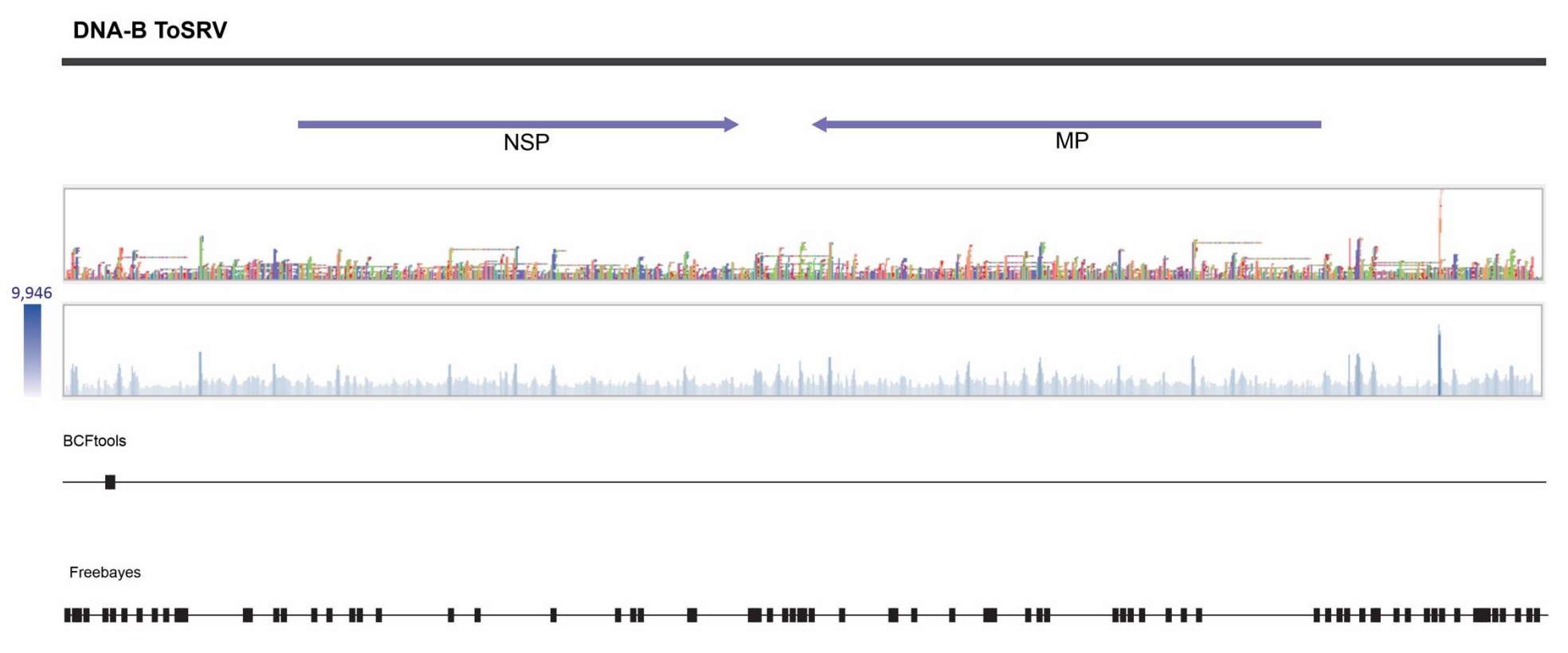


Figure 3

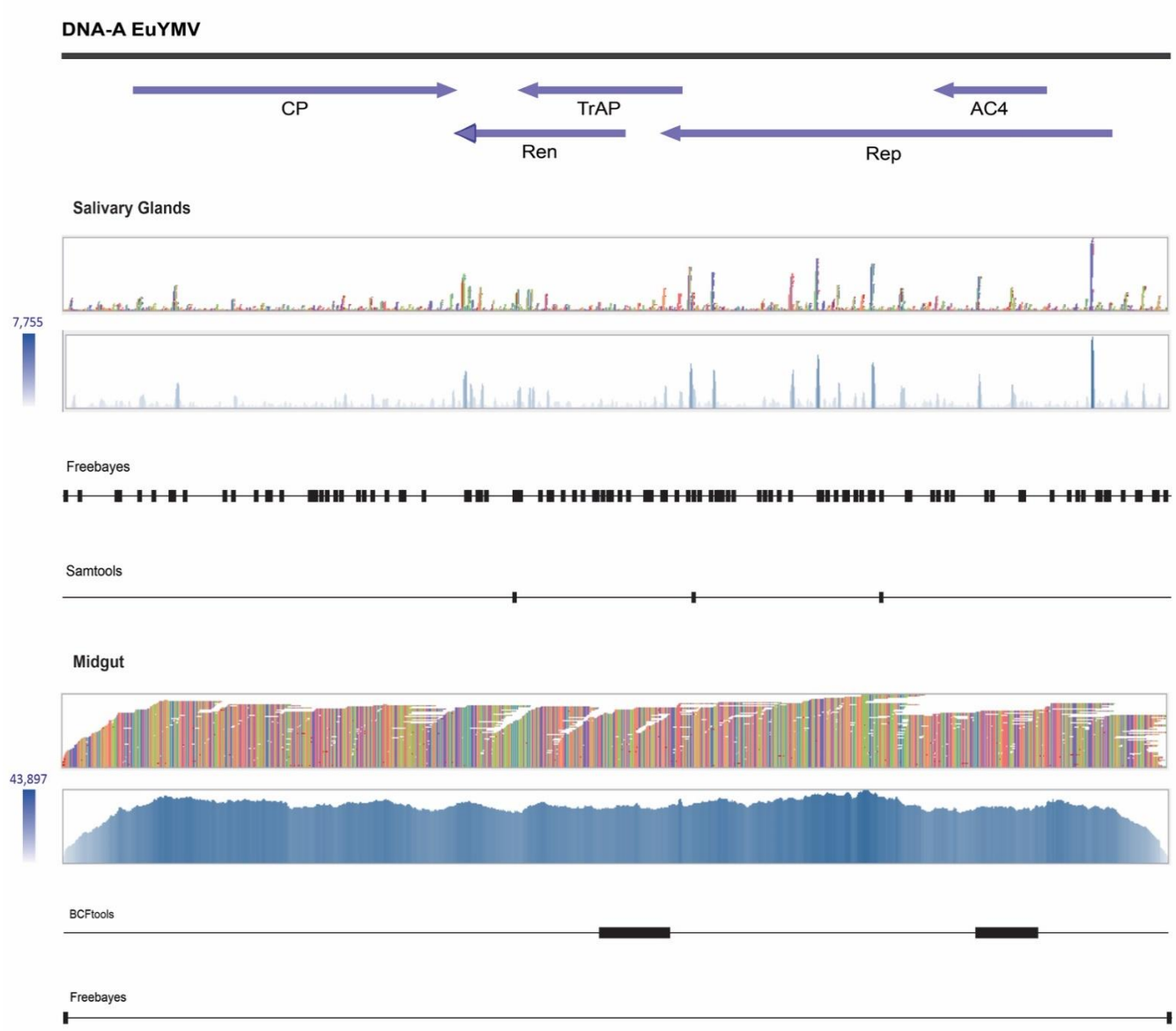

DNA-A EuYMV
CP
TrAP
AC4
Ren
Rep
Salivary Glands
7,755
Freebayes
Samtools
Midgut
43,897
BCFtools
Freebayes

Figure 4

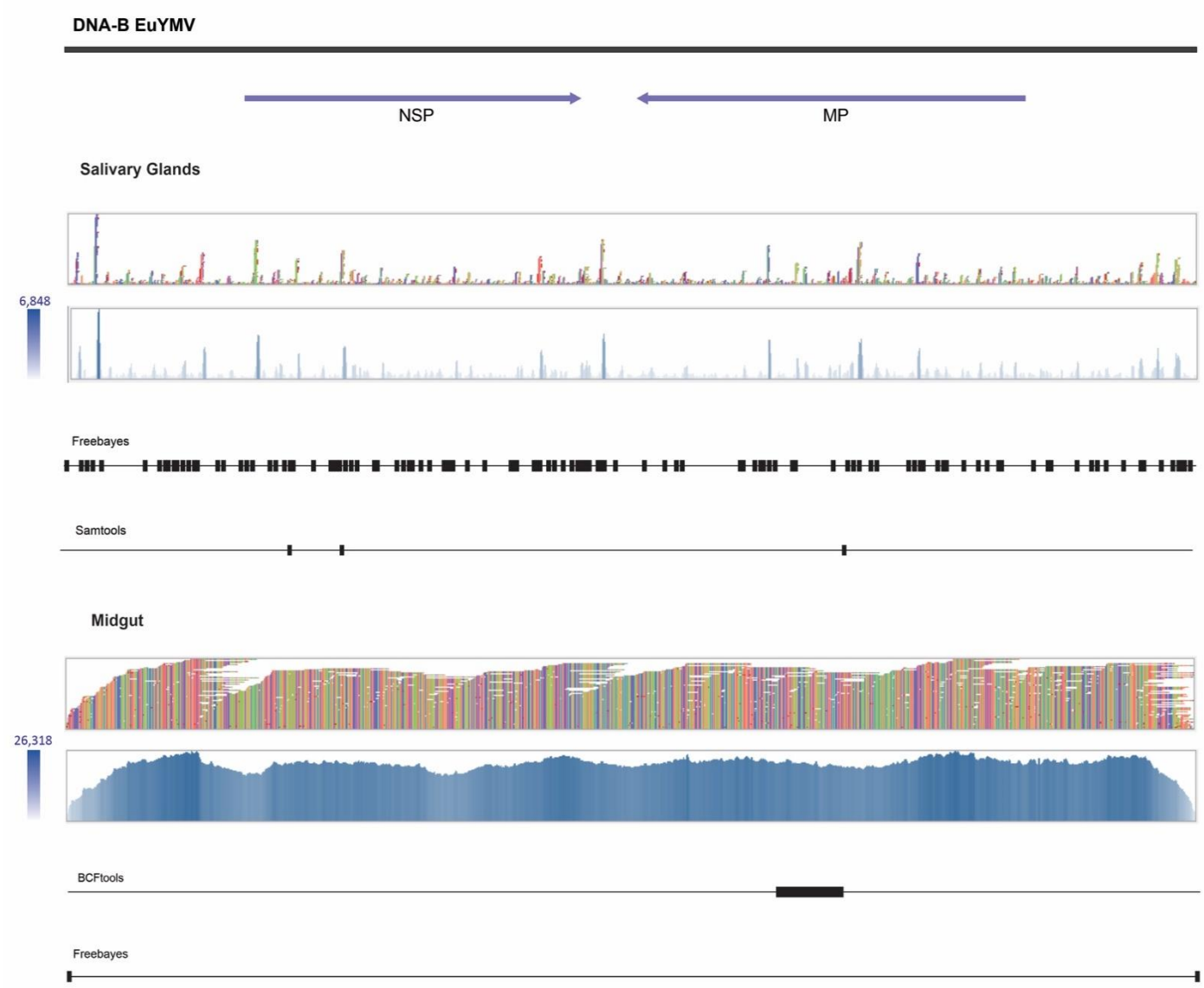

DNA-B EuYMV
NSP
MP
Salivary Glands
6,848
Freebayes
Samtools
Midgut
26,318
BCFtools
Freebayes

Figure 5

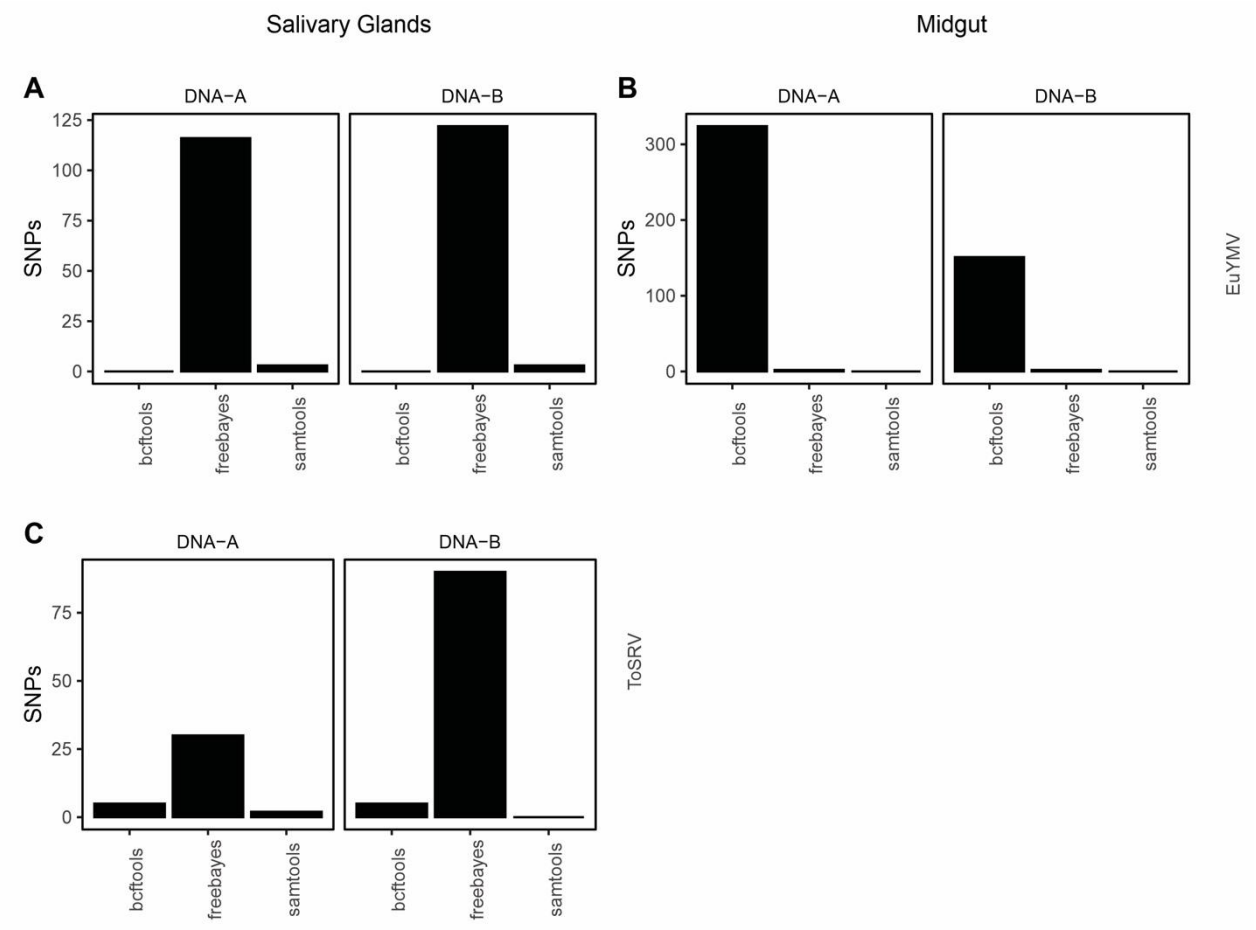
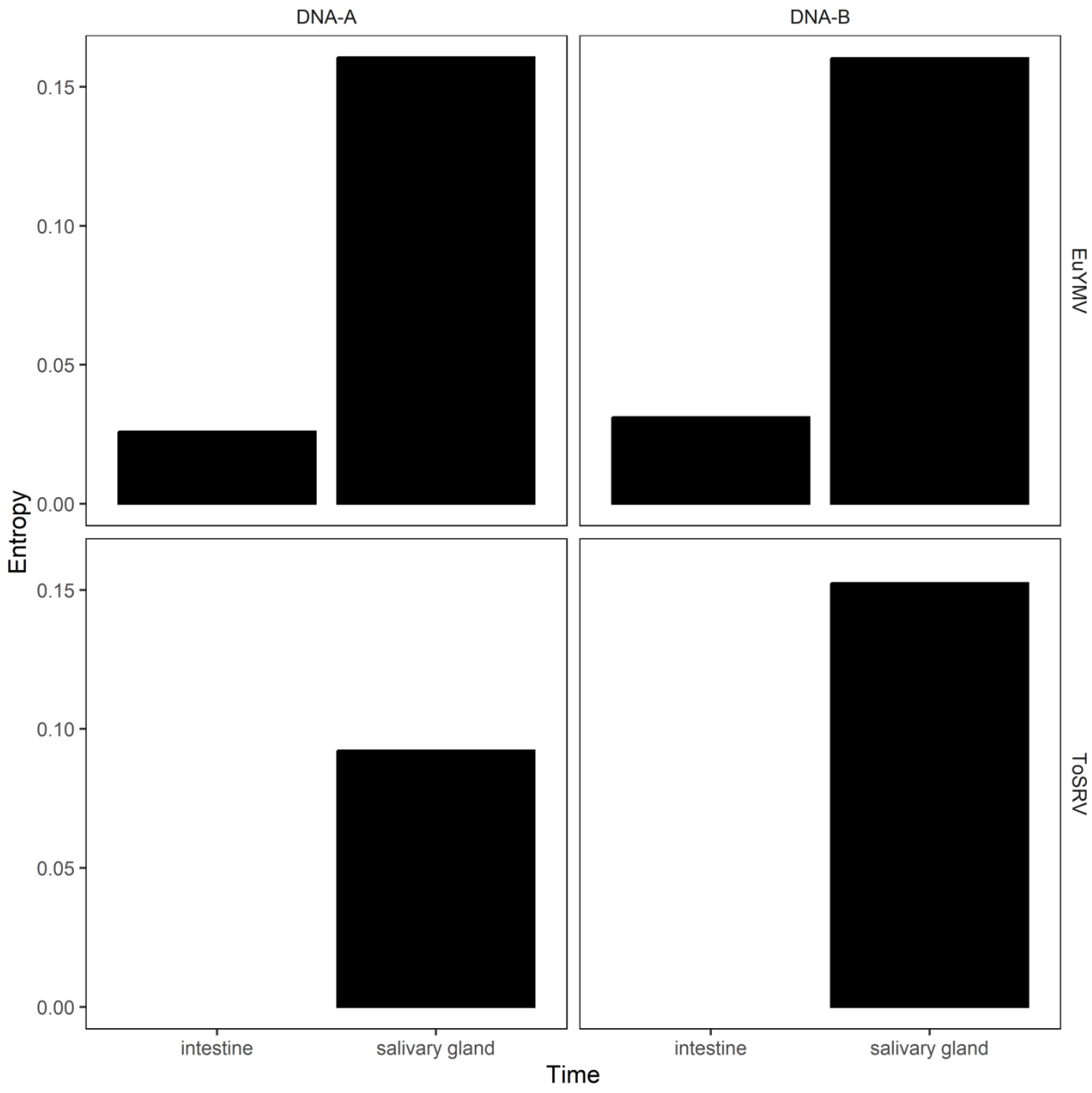

Salivary Glands
Midgut
A
B
C
DNA−A
DNA−B
SNPs
bcftools
freebayes
samtools
EuYMV
ToSRV

Figure 6


DNA-A
DNA-B
EuYMV
ToSRV
Entropy
0.15
0.10
0.05
0.00
intestine
salivary gland
Time

Supplementary Table S1. Raw data statistics of sequenced libraries. EuYMV, Euphorbia yellow mosaic virus. ToSRV, tomato severe rugose virus. MG, midgut, SG, salivary glands.

Sample ID	Total bases	Total reads	GC(%)	AT(%)	Q20(%)	Q30(%)
EuYMV-MG	2,121,926,896	14,052,496	58.5	41.5	95.6	90.9
EuYMV-SG	2,082,018,804	13,788,204	35.7	64.3	90.8	82.8
ToSRV-SG	2,327,625,740	15,414,740	45.9	54.1	82.5	72.3

Supplementary Table S2. Variant frequency mapped from the SNP analysis by the Samtools program for tomato severe rugose virus (ToSRV) and Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary glands (SG) of *B. tabaci* MEAM1 after a 48 h acquisition period. Only the variants present in coding regions are listed. Ref., sequence of the reference genome. Chg., observed change.

Virus and organ	**Comp.**	**Position**	**Ref.**	**Chg.**	**Polymorphism type**	**Gene**	**Variant frequency (%)**
ToSRV-SG	DNA-A*	2,062	G	A	SNP (transition)	*AC4*	59.72
		2,194	G	A	SNP (transition)	*AC4*	37.07
EuYMV-SG	DNA-A	1,060	C	A	SNP (transversion)	*REn/TrAP*	33.33
		1,480	T	C	SNP (transition)	*REn/TrAP*	69.52
		1,920	A	G	SNP (transition)	*TrAP/Rep*	19.17
	DNA-B	524	G	A	SNP (transition)	*NSP*	14.08
		647	C	A	SNP (transversion)	*NSP*	8.82
		1,787	C	A	SNP (transversion)	*MP*	13.40

* no SNPs were detected in the ToSRV-SG DNA-B.

Supplementary Table S3. Variant frequency mapped from the SNP analysis by the Freebayes program for tomato severe rugose virus (ToSRV) sampled in the salivary glands of *B. tabaci* MEAM1 after a 48 h acquisition period. Only variants present in coding regions are listed. Ref., sequence of the reference genome. Chg., observed change.

Comp.	Position	Ref.	Chg.	Polymorphism type	Gene	Variant frequency (%)
DNA-A	272; 395; 473	T	C	SNP (transition)	*CP*	94.91; 79.41; 92.99
	989	A	G	SNP (transition)	*REn*	70.41
	1,071	G	C	SNP (transversion)	*REn/ TrAP*	87.98
	1,287	C	A	SNP (transversion)	*REn/ TrAP*	85.97
	1,405	A	G	SNP (transversion)	*Trap*	78.09
	1,672	A	G	SNP (transition)	*Rep*	80.91
	1,720	T	C	SNP (transition)	*Rep*	87.44
	1,834	A	T	SNP (transversion)	*Rep*	78.63
	1,847	G	T	SNP (transversion)	*Rep*	52.07
	1,972	C	A	SNP (transversion)	*Rep*	5.23
	2,140	A	G	SNP (transition)	*Rep/AC4*	92.94
	2,194	G	A	SNP (transversion)	*Rep/AC4*	11.79
	2,294	C	G	SNP (transversion)	*Rep/AC4*	10.82
DNA-B	432; 1,094	G	T	SNP (transversion)	*NSP*	6.87; 20.97
	459	G	C	SNP (transversion)	*NSP*	56.64
	497	A	G	SNP (transition)	*NSP*	78.07
	512; 718; 1,084	A	T	SNP (transversion)	*NSP*	71.61; 43.27; 31.55
	544; 958	T	C	SNP (transition)	*NSP*	72.54; 57.47
	667	C	T	SNP (transition)	*NSP*	52.10
	847	T	G	SNP (transversion)	*NSP*	23.44
	988	T	A	SNP (transversion)	*NSP*	6.91
	1,001	G	A	SNP (transition)	*NSP*	42.33
	1,057	A	C	SNP (transversion)	*NSP*	71.76
	1,350; 1,542; 1,673; 1,689; 1,826	T	A	SNP (transversion)	*MP*	5.23; 58.96; 10.26; 7.84; 6.49
	1,437	TGGAC	CGGAT	substitution	*MP*	99.02
	1,442; 1,847	G	C	SNP (transversion)	*MP*	6.29; 5.33
	1,476; 1,705; 1,839; 2,249	A	G	SNP (transition)	*MP*	92.1; 76.68; 6.19; 20.0
	1,600;1,918; 2,277	A	C	SNP (transversion)	*MP*	43.70; 69.24; 27.72
	1,603; 1,704; 2,255	G	A	SNP (transition)	*MP*	43.70; 69.24; 5.31
	1,612	T	G	SNP (transversion)	*MP*	76.57
	1,693; 2,170; 2,285	T	C	SNP (transition)	*MP*	47.80; 45.11; 95,6
	1,872; 1,970; 2,274	A	T	SNP (transversion)	*MP*	7.60; 11.42; 7.84
	1,942	C	A	SNP (transversion)	*MP*	68.21

Supplementary Table S4. Variant frequency mapped from the SNP analysis by the Freebayes program for Euphorbia yellow mosaic virus (EuYMV) sampled in the salivary glands (SG) of *B. tabaci* MEAM1 after a 48 h acquisition period. Only variants present in coding regions are listed. Results are presented in separate tables for the DNA-A **(A)** and DNA-B **(B)**. Ref., sequence of the reference genome. Chg.,: observed change.

A

Position	Ref.	Chg.	Polymorphism type	Gene	Variant frequency (%)
220; 583	A	C	SNP (transversion)	*CP*	5.13; 7.69
257; 261; 515; 590	C	A	SNP (transversion)	*CP*	6.89; 28.23; 8.1; 32.94
263; 809; 946	T	A	SNP (transversion)	*CP*	15.38; 11.9; 24.55
291	C	G	SNP (transversion)	*CP*	13.51
383	G	A	SNP (transition)	*CP*	5.13
402; 493; 610; 757	A	T	SNP (transversion)	*CP*	23.83; 18.03; 5.63; 23.42
454; 730	T	C	SNP (transition)	*CP*	5.63; 23.42
484	G	C	SNP (transversion)	*CP*	5.13
596; 764	C	T	SNP (transition)	*CP*	7.14; 7.36
623	T	G	SNP (transversion)	*CP*	15.38
625; 643; 659; 794	G	T	SNP (transversion)	*CP*	9.46; 40.91; 6.56; 13.33
698; 707	T	G	SNP (transversion)	*CP*	6.98; 19.44
847; 952	A	G	SNP (transition)	*CP*	9.52; 12.56
958	A	C	SNP (transversion)	*REn*	5.32
972	A	G	SNP (transition)	*REn*	36.07
981	T	C	SNP (transition)	*REn*	8.33
992	A	T	SNP (transversion)	*Ren*	8.75
1,060	C	A	SNP (transversion)	*REn*	42.00
1,068; 1,074	T	A	SNP (transversion)	*REn*	9.62; 33.27
1,119	G	A	SNP (transition)	*REn/TrAP*	17.33
1,140; 1,175	C	A	SNP (transversion)	*REn/TrAP*	36.11; 16.67

1,148; 1,281	G	T	SNP (transversion)	*REn/TrAP*	11.76; 5.13
1,203	C	T	SNP (transition)	*REn/TrAP*	9.09
1,223; 1,248	T	G	SNP (transversion)	*REn/TrAP*	5.97; 11.86
1,251	A	G	SNP (transition)	*REn/TrAP*	14.75
1,268; 1,329	A	C	SNP (transversion)	*REn/TrAP*	5.46 ;15.29
1,289	C	G	SNP (transversion)	*REn/TrAP*	5.71
1,305	T	A	SNP (transversion)	*REn/TrAP*	9.6
1,367; 1,377	T	A	SNP (transversion)	*TrAP*	68.75; 10.14
1,378	A	C	SNP (transversion)	*TrAP*	6.06
1,404; 1,464	C	G	SNP (transversion)	*TrAP/Rep*	12.5; 29.41
1,417	A	C	SNP (transversion)	*TrAP/Rep*	5.66
1,439	G	A	SNP (transition)	*TrAP/Rep*	5.08
1,466	T	G	SNP (transversion)	*TrAP/Rep*	11.39
1,480	T	C	SNP (transition)	*TrAP/Rep*	53.55
1,495; 1,520	T	A	SNP (transversion)	*Rep*	64.41; 70.69
1,531; 1,534, 1,576; 1,646	G	A	SNP (transition)	*Rep*	16.0; 5.08; 6.98; 27.78
1,543; 1,545, 1,890; 2,422	A	T	SNP (transversion)	*Rep*	57.14; 6.03; 6.19
1,563; 1,825, 2,452	G	T	SNP (transversion)	*Rep*	23.21; 7.42; 35.53
1,634; 1,657, 1,980	C	G	SNP (transversion)	*Rep*	9.52; 5.22; 26.67
1,682; 1,780, 1,987; 2,434	T	A	SNP (transversion)	*Rep*	6.6; 30.66; 24.69; 20.75
1,706; 1,723, 1,771; 1,890	A	T	SNP (transversion)	*Rep*	5.45; 48.46; 15.52; 6.19
1,791; 1,810	C	T	SNP (transition)	*Rep*	22.35; 10.0
1,832	A	C	SNP (transversion)	*Rep*	38.71
1,840; 1,898, 1,920	A	G	SNP (transition)	*Rep*	16.22; 25.19; 80.30
1,862	G	C	SNP (transversion)	*Rep*	5.00
1,871	C	A	SNP (transversion)	*Rep*	66.67

2,442	T	C	SNP (transition)	*Rep*	42.70
2,041	G	C	SNP (transversion)	*Rep/AC4*	8.78
2,054	C	A	SNP (transversion)	*Rep/AC4*	11.35
2,055; 2,252, 2,389	C	T	SNP (transition)	*Rep/AC4*	36.36; 16.80; 22.22
2,074	G	A	SNP (transversion)	*Rep/AC4*	56.60
2,088; 2,251, 2,377	C	G	SNP (transversion)	*Rep/AC4*	9.54; 24.46; 20.00
2,168; 2,244, 2,355	T	A	SNP (transversion)	*Rep/AC4*	9.80; 7.76; 9.88
2,176	G	T	SNP (transversion)	*Rep/AC4*	35.71
2,180	T	G	SNP (transversion)	*Rep/AC4*	14.58
2,321	T	C	SNP (transition)	*Rep/AC4*	10.53

Suppl. Table 4 (cont.)

B

Position	Ref.	Chg.	Polymorphism type	Gene	Variant frequency (%)
431; 483; 756; 834; 1,081; 1,170; 1,191	T	A	SNP (transversion)	*NSP*	30.95; 2.41; 18.18; 5.0; 47.89; 12.95; 6.52
475; 647; 790; 791; 865; 1,072	C	A	SNP (transversion)	*NSP*	7.27; 51.69; 6.38; 23.40; 5.17; 71.25
505; 654; 667	C	T	SNP (transition)	*NSP*	10.56; 8.06; 25.93
517; 631; 524; 571	G	C	SNP (transversion)	*NSP*	20.59; 22.73; 79.35; 42.57
574; 716; 1,118	T	G	SNP (transversion)	*NSP*	12.59; 11.64; 55.56
610	T	C	SNP (transition)	*NSP*	5.77
614; 619; 816; 879; 907; 1,136	G	A	SNP (transition)	*NSP*	11.11; 41.56; 65.0; 28.89; 13.33; 14.81
621; 960; 1,026; 1,104; 1,176; 1,197	A	T	SNP (transversion)	*NSP*	44.10; 5.0; 14.29; 18.23; 42.95; 9.52
709; 1157	A	G	SNP (transition)	*NSP*	15.07; 7.32
777; 886; 919; 1,020; 1,035; 1,088; 1,186	G	T	SNP (transversion)	*NSP*	25.0; 23.40; 11.29; 7.46; 32.26; 40.30; 15.65
797	T	G	SNP (transversion)	*NSP*	5.93
1,323; 1,852	G	C	SNP (transversion)	*MP*	51.94; 18.0
1,371; 1,393; 1,546; 1,570; 1,595; 1,808; 1,813; 2,051	C	A	SNP (transversion)	*MP*	7.06; 8,0; 12.60; 28.57; 10.17; 14.83; 13.11; 18.18
1,409; 1,584; 1,660; 1,922; 1,986; 2,127	G	T	SNP (transversion)	*MP*	15.64; 62.22; 14.81; 8.13; 71.91; 22.60
1,538; 1,999; 2,135	A	C	SNP (transversion)	*MP*	47.22; 14.81; 11.99
1,575, 1,618, 1,800; 1,836	T	G	SNP (transversion)	*MP*	8.84; 11.49; 5.0; 23.76

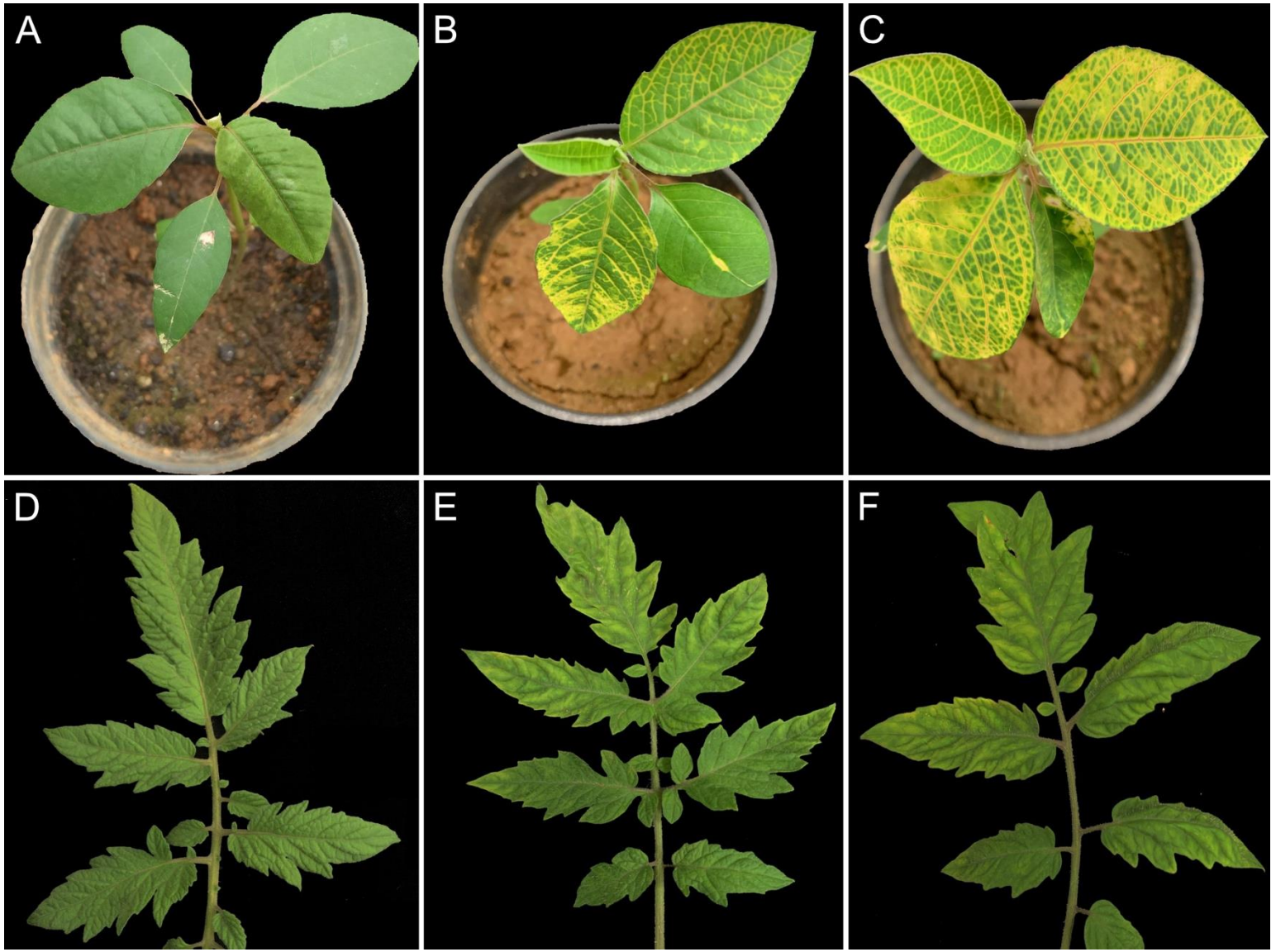
1,603, 1,787, 1,957	C	T	SNP (transition)	*MP*	43.22; 14.58; 8.39
1,668; 1,750; 2,078	T	A	SNP (transversion)	*MP*	90.65; 15.79; 12.75
1,812	C	G	SNP (transversion)	*MP*	5.49
1,938; 1,946	T	C	SNP (transition)	*MP*	56.58; 7.73
1,958; 2,101	A	T	SNP (transversion)	*MP*	17.97; 25.30
2,007	T	C, G	insertion	*MP*	6.75

Supplementary figure legends

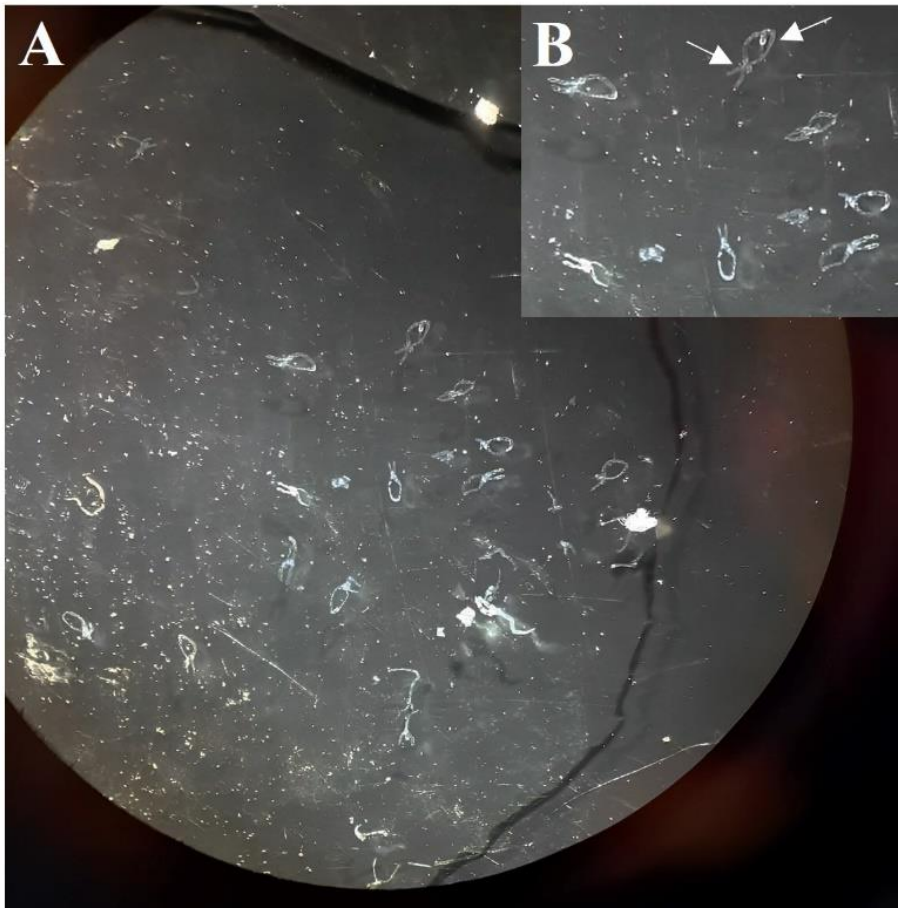
Suppl. Figure S1. Symptoms in *Euphorbia heterophylla* infected with Euphorbia yellow mosaic virus (EuYMV) **(B, C)** and in tomato plants infected with tomato severe rugose virus (ToSRV) (**E, F**). **A** and **D** are mock-inoculated *E. heterophylla* and tomato plants, respectively, used as negative controls. All images were obtained 28 days after inoculation.

Suppl. Figure S2. **A.** Dissected midguts of *B. tabaci* MEAM1. The image was obtained with a stereoscopic magnifier (Leica M125, Wetzlar, Germany) at 100x magnification. **B.** Cropping of the region in **(A)** with the highest number of organs. Arrows indicate structures of the ascending and descending midgut (loop-shaped structure) and cecum.

Suppl. Figure S1



Suppl. Figure S2



GENERAL CONCLUSIONS

- Nucleotide changes in three sites in the common region (CR) of the ToSRV-[BR:PG1:Pep:03] isolate increased ToSRV-$A_{(ToR:CR)}$ infectivity and viral accumulation when in mixed infection with ToRMV, suggesting that these sites act as specificity determinants for Rep binding.

- Amino acid substitutions (Pro > Thr and Lys > Arg) in the Rep protein iteron-related domain (IRD) provided a replication disadvantage of ToSRV-$A_{(ToR:CR+IRD)}$ and ToSRV-$A_{(ToR:IRD)}$, likely by negatively affecting iteron recognition.

- Divergent nucleotides in the ToRMV CR make this virus a better competitor in mixed infection with ToSRV, maybe due to the more efficient binding of Rep to viral DNA.

- The midgut (MG) is where there is greater detection of begomovirus DNA accumulation and retention compared to the salivary glands (SG).

- The detection of a large number of viral genome deletions observed for the MG sample compared to the SG samples could be due to activation of autophagic pathways and insect immunity. Thus, the midgut may be acting as an important barrier to the elimination of defective genomes.

- The greater detection of nucleotide substitutions and insertions occurred in the SG, and the greater number of SNPs for ToSRV in relation to EuYMV indicated a greater degree of genetic variability of the viral population in the SG, indicating that genetic variability may have been restored in this organ.