MARCOS FERNANDO BASSO

CARACTERIZAÇÃO DE UM NOVO VÍRUS DE ssDNA INFECTANDO FRUTEIRAS DE CLIMA TEMPERADO, E ASPECTOS DA INTERAÇÃO MOLECULAR BEGOMOVÍRUS-HOSPEDEIRO

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

VIÇOSA
MINAS GERAIS - BRASIL
FEVEREIRO - 2015

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

T
B322c
2015

Basso, Marcos Fernando, 1982-
Caracterização de um novo vírus de ssDNA infectando fruteiras de climas temperados, e aspectos da interação molecular begomovírus-hospedeiro / Marcos Fernando Basso. – Viçosa, MG, 2015.
viii, 156f. : il. ; 29 cm.

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

1. Geminivírus. 2. Diversidade genética. 3. Interação planta-patógeno. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-graduação em Fitopatologia. II. Título.

CDD 22. ed. 571.992

MARCOS FERNANDO BASSO

CARACTERIZAÇÃO DE UM NOVO VÍRUS DE ssDNA INFECTANDO FRUTEIRAS DE CLIMA TEMPERADO, E ASPECTOS DA INTERAÇÃO MOLECULAR BEGOMOVÍRUS-HOSPEDEIRO

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

Aprovada: 27 de fevereiro de 2015.

Profa. Tatiane da Franca Silva

Profa. Elizabeth Pacheco Batista Fontes

Profa. Poliane Alfenas Zerbini

Prof. Humberto Josué de Oliveira Ramos

Prof. Francisco Murilo Zerbini Jr.

(Orientador)

AGRADECIMENTOS

Agradeço a DEUS e a Nossa Senhora Aparecida por iluminar e guiar minha vida. Aos meus pais Irineu e Delvina, meus irmãos Marcelo, Márcia e Maurício, meus sobrinhos Larissa, Tiago e Gabriela, pelo amor, incentivo e compreensão em todos os momentos da minha vida;

Aos mentores da Universidade Federal de Viçosa (UFV) e do Departamento de Fitopatologia, pela oportunidade de realização deste curso de excelência. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos;

Aos professores do Departamento de Fitopatologia e os demais em que eu tive a oportunidade de conviver e adquirir conhecimentos, pela amizade e ensinamentos. Em especial aos profs. Fabrício A. Rodrigues, Wagner C. Ottoni, Humberto J.O. Ramos e Fábio N. da Silva, também pelo ótimo exemplo de carisma.

Ao prof. Francisco Murilo Zerbini pela orientação em todas as etapas, oportunidades e ensinamentos e pelos recursos financeiros disponibilizados para que tudo fosse possível. A profa. Poliane Alfenas-Zerbini pela co-orientação e ensinamentos.

Agradeço aos professores Poliane Alfenas-Zerbini, Elizabeth P.B. Fontes, Humberto J.O. Ramos, Fábio N. da Silva, Amando Bonoto, Michihito Deguchi (LBMP-Bioagro) e ao Edvaldo Barros (NuBioMol-UFV), pela ajuda e ótimos conselhos no trabalho da MP-ToYSV.

Agradeço ao Dr. Thor V.M. Fajardo, José Cleydson F. da Silva e à profa. Elizabeth P.B. Fontes pela ajuda, incentivo, pelas ótimas idéias no trabalho de prospecção e caracterização biológica e molecular do TFDaV em fruteíras no Brasil. A Karla V.G. Ribeiro e ao Gilmar Valente (NMM-UFV) pela auxílio com a microscopia confocal e eletrônica de transmissão.

Agradeço ao José Cleydson F. da Silva, Otávio J.B. Brustolini, Régis L. Corrêa (UFRJ), Elizabeth P.B. Fontes e César Llave (CSIC-Espanha), pelo auxílio, idéias e apoio no trabalho de análise do perfil de smRNAs na interação begomovírus-N. benthamiana.

Aos meus amigos do Laboratório de Virologia Vegetal Molecular: Ana Paula O. Barros, André da S. Xavier, Álvaro J. Pereira, Angélica M. Nogueira, Amanda Bonoto, Adriana N. de Souza, Jorge Aguilera, Joyce S. Lima, Larissa G. Zanardo, Marcelo Soares, Fábio N. da Silva, Tathiana F.S. Antunes, Sheila Tavares, Igor R. Mendes, Silvia L. Carvalho, Laiane Maciel, Fernanda Pereira, Vítor B. Pinto, David D. Valente, Chaianne Balzzo, José J. AS Galvão, Sarah Cavalcante, Carolina da S. Rocha, Carolina G. Neves, Diogo F. Milanesi, Talita B. Mar, João P. Herrera, Jesús Navas-Castillo, Diego Nogueira, Diego C. Lemos, Claudine M. Carvalho, Flavia O. Souza, Camila G. Ferro, Hermano M.B. Pereira, Josiane G. Amaral, Lina M. C. Páez, Priscila F. Gomes, Renan D. Silva, Lenin P. Carrasco, Osvaldo F.L. Sandi, Rosilene Matos, Karel Ismar, Elvira Fiallo-Olivé e os demais colegas, pela amizade, ensinamentos, descontração,

companheirismo e constante apoio em todos os momentos. Em especial ao Hermano, pelo valioso auxílio no desenvolvimento dos trabalhos.

Aos amigos do Laboratório de Cultura de Tecidos (LCTII), de Biologia Molecular de Plantas (LBMP), de Virologia Vegetal (LVV) e de Microbiologia Industrial (MIND), que me salvaram sempre nas horas de aperto, pela alegria e amizade e pelos convites para o cafezinho, lanche, "happy hour", etc. Muito obrigado por dedicarem um pouco do tempo de vocês para me ajudar, sempre da melhor forma.

Aos meus admirados professores e/ou pesquisadores externos a UFV: Leandro Crestani, Thor V. M. Fajardo, Marcos F. Vanni, Marco A. Dalbó, Marcelo Eiras, Jane M. L. Gelinski, Gilvan Pio-Ribeiro, César M. Barato, Jesús Angel-Sanchez Navarro, Rodrigo G. Nogueira e Osmar Nickel, pela amizade, apoio, alegria passada e companherismo durante todo o tempo;

Meus verdadeiros amigos do dia-a-dia em Viçosa: Gizelly Mendes (super Gizeliss), Lívio Amaral (Livão, a quem considero como um irmão), Elder Oliveira, Raul Coutinho, Chaianne Balzzo, Laiane Maciel, Fernanda Pereira, Rayane Paschoalino, Farley Freitas, Adriana N. de Souza (Fiotinha), Casley Queiroz, Thaís R. Santiago (super Thaís), Ana Tiburcia, André da Silva Xavier, Denise Fernandes, Diego Batista, Eleonora R. Polanco, Carlos E. Aucique Perez, Sara Salcedo, J. Cleydson F. da Silva (brother), Alessandra Fernandes, Leonardo D. Figueiredo, Wagner Morais, Kenner Morais, Osvaldo F.L. Sandi, Carlos R. Cornelio, Deroila Marinho, Karel Ismar, Lira Pereirapaes, Priscilla A. Möller, Hiunes Murta, Michihito Deguchi, Guilherme Podestá, Tiago de Souza Leite. Valeu muito pelo companheirismo, amizade e pelos momentos de diversão compartilhados ao longo destes anos em Viçosa. Vocês são incrivelmente ótimos.

Meus amigos da graduação Djan A. Veiga, Júlio C. Vian, Wilhelm Feiten e Greice Moresco, saudades daquela nossa amizade e dos bons tempos. Ao meu orientador de iniciação científica Marco A. Dalbó, pelos ensinamentos e apoio. Meus amigos do mestrado, prof. Guilherme S. Schühli e Ricardo A. Ayub, profa. Carolina W. Galvão, Rafael M. Etto e Paulo Sérgio Ribeiro, obrigado pelos ensinamentos e amizade.

Ao imbatível time de futebol de Viçosa e melhor do mundo: "São José F.C", posso sinceramente dizer que foi muito bom fazer parte desta "família". Aos amigos da Pelada da Fitopatologia, pela parceragem dos sábados à tarde. Não consigo pensar em um dia estar longe de todos vocês, sem não me sentir triste.

A todos aqueles que dedicaram um pouco do seu tempo para me ajudar, aos que de alguma forma contribuíram direta ou indiretamente, minha sincera gratidão. Aos que me esqueci de mencionar aqui, meu pedido de desculpas e muito obrigado.

Meu sincero agradecimento a vocês!

SUMÁRIO

RESUMO

BASSO, Marcos Fernando, D.Sc. Universidade Federal de Viçosa, fevereiro de 2015. **Caracterização de um novo vírus de ssDNA infectando fruteiras de clima temperado, e aspectos da interação molecular begomovírus-hospedeiro.** Orientador: Francisco Murilo Zerbini. Co-orientadora: Poliane Alfenas-Zerbini.

Fruteiras de clima temperado são de grande importância econômica mundial e são suscetíveis a diversos patógenos transmissíveis por insetos ou nematoides. Seu caráter perene e a propagação vegetativa resultam no aparecimento de diversas doenças complexas, algumas das quais ainda etiologicamente desconhecidas. Em pomares e vinhedos brasileiros frequentemente são observados sintomas de possível origem viral, mas os agentes não foram identificados. Com o objetivo de prospectar agentes virais associados a estas culturas, 74 amostras de fruteiras foram coletadas em diferentes regiões produtoras e avaliadas por PCR e RCA. Um novo vírus altamente divergente, monossegmentado, com genoma de ssDNA circular de aproximadamente 3.400 nucleotídeos, apresentando características moleculares semelhantes às de vírus classificadas nas famílias Circo-, Nano- e Geminiviridae foi detectado em plantas de macieira, pereira e videira. A associação com os hospedeiros foi confirmada e sua infectividade comprovada por meio de inoculação com um clone correspondente ao genoma completo. A análise de sequências e suas características moleculares indicam que este novo vírus poderá pertencer a um novo gênero ou família viral. Os begomovírus (família Geminiviridae) são patógenos de grande importância econômica em diversas culturas, principalmente em regiões tropicais e subtropicais. Eles regulam, direta ou indiretamente, diversos mecanismos de defesa do hospedeiro, a exemplo da via dos pequenos RNAs (smRNAs), de forma a facilitar a infecção viral. O segundo objetivo deste trabalho foi avaliar o perfil de smRNAs em plantas de Nicotiana benthamiana infectadas pelo begomovírus Tomato yellow spot virus (ToYSV). Utilizando sequenciamento de nova geração, bibliotecas

de smRNAs de plantas de N. benthamiana, sadias e infectadas com o ToYSV, foram sequenciadas a fim de investigar sua complexidade e compreender os mecanismos regulatórios envolvidos na patogênese. Os resultados indicam que os begomovírus são alvo das vias de silenciamento de RNA (VSR) nas etapas iniciais da infecção e que suas proteínas supressoras de silenciamento atuam in trans para interferir com as VSRs. O sucesso da infecção viral e a indução de sintomas estariam assim, ao menos em parte, correlacionados com a regulação negativa das VSRs, com a mudança do perfil de smRNAs e com a regulação da expressão de diversos genes do hospedeiro que transcrevem mRNAs e miRNAs. A proteína de movimento (MP) dos begomovírus bissegmentados é responsável pelo movimento célula-a-célula, indução de sintomas e na determinação da gama de hospedeiros. O terceiro objetivo deste trabalho foi identificar proteínas de hospedeiros do ToYSV candidatas a interagirem com a MP. Complexos proteicos provenientes de plantas expressando constitutiva- ou transientemente NTAPi-MP foram purificados, seguido de identificação por espectrometria de massas. Entretanto, não foi possível encontrar proteínas que interagissem com MP do ToYSV. Ensaios de pull-down com a proteína MP fusionada a uma etiqueta de seis histidinas (6His-MP) identificaram 64 proteínas da planta candidatas a interagirem com a MP. Destas, 25 foram selecionadas para avaliação da interação pelo sistema duplo-híbrido em levedura, contudo, os resultados foram negativos para todas as 25 proteínas.

ABSTRACT

BASSO, Marcos Fernando, D.Sc. Universidade Federal de Viçosa, February, 2015. **Characterization of a new ssDNA virus infecting temperate fruit trees, and molecular aspects of the begomovirus-host interaction.** Adviser: Francisco Murilo Zerbini. Committee member: Poliane Alfenas-Zerbini.

Temperate fruit trees are of great economical importance worldwide and are susceptible to several insect or nematode-transmitted pathogens. Their perennial nature and vegetative mode of propagation result in the appearance of several complex diseases, some of which are aetiologically undetermined. In Brazilian orchards and vineyards, virus-like symptoms are often observed, but the causal agents have not been identified. With the objective of prospecting viral agents associated with these plants, 74 fruit tree samples were collected at different regions and evaluated by PCR and RCA. A novel, highly divergent virus with a monopartite circular ssDNA genome of approximately 3400 nucleotides, with molecular characteristics similar to viruses in the Circo-, Nano- and Geminiviridae families was identified in apple, pear tree and grapevine plants. Its association with the hosts was confirmed and its infectivity was proven by inoculation with a full-length genomic clone. Sequence analysis and molecular characteristics indicate that this novel virus will be classified in a new viral genus or family. Begomoviruses (Geminiviridae family) cause diseases of major economic importance in many crops, especially in tropical and subtropical regions. They regulate, directly or indirectly, several host defense mechanisms such as small RNA (smRNA) pathways so as to facilitate viral infection. The second objective of this work was to evaluate the profile of smRNAs in Nicotiana benthamiana plants infected by the begomovirus Tomato yellow spot virus (ToYSV). Using Next Generation Sequencing approaches, we sequenced smRNA libraries from healthy or ToYSV-infected N. benthamiana plants to understand their complexity and to explore the regulatory mechanisms involved in pathogenesis. The results indicate that begomoviruses are targets of RNA silencing pathways

(RSP) early in the infection and that their silencing suppressor proteins act in trans to interfere with RSP. The success of viral infection and the appearance of ToYSV-induced symptoms may be correlated, at least in part, with RSP downregulation, smRNA profile change and regulation of expression of several host genes (mRNA- and miRNA-transcribing). The movement protein (MP) of bipartite begomoviruses is responsible for cell-to-cell movement, symptom induction and determination of host range. The third objective of this work was to identify host proteins that interact with the ToYSV MP. Protein complexes were purified from plants expressing NTAPi-MP either constitutively or transiently, followed by mass spectrometry-based identification. However, no candidate host proteins were identified. Pull-down assays using the MP with a six histidine tag (6His-MP) identified 64 candidate proteins, 25 of which were selected for evaluation of the interaction using the yeast two-hybrid system. However, results were negative for all 25 proteins.

INTRODUÇÃO GERAL

Descoberta de novos vírus de ssDNA

Os vírus com genoma composto de DNA de fita simples (ssDNA) são classificados atualmente em cinco famílias: Anelloviridae, Circoviridae, Parvoviridae, Geminiviridae e Nanoviridae. As duas últimas incluem vírus que infectam plantas (King et al., 2012). Os vírus de ssDNA incluem os menores patógenos (em termos de comprimento do genoma) que infectam eucariotos. De modo geral, os vírus de ssDNA apresentam elevada variabilidade genética, com altas taxas de substituição de nucleotídeos e de recombinação (Duffy et al., 2008; Martin et al., 2011), características que podem facilitar sua emergência e disseminação.

Um grande número de espécies de vírus com genoma de ssDNA é reconhecido oficialmente pelo Comitê Internacional de Taxonomia de Vírus (ICTV). Por exemplo, a família Geminiviridae é a mais numerosa, em termos de número de espécies, dentre todos os vírus (Brown et al., 2015). Mesmo assim, estudos recentes de metagenômica indicam que uma enorme diversidade de vírus de ssDNA existe nos mais diferentes ambientes (Ng et al., 2011; Rosario et al., 2012; Terry et al., 2012; Zawar-Reza et al., 2014). Alguns desses possíveis vírus possuem características que devem classificá-los em uma das cinco famílias existentes. Entretanto, a maioria das sequências encontradas não possui similaridade com nenhum vírus conhecido, e provavelmente serão classificadas em novos gêneros e famílias (Rosario et al., 2012; Terry et al., 2012).

Um fator que provavelmente tem contribuído para a detecção de novos vírus de ssDNA, tanto em estudos de metagenômica como em trabalhos utilizando clonagem e sequenciamento convencional, é a utilização da DNA polimerase do fago phi29 em uma etapa de amplificação de DNA extraído de amostras ambientais, de animais ou de plantas (Rosario et al., 2012). Denominada amplificação por círculo rolante (rolling-circle amplification, RCA), essa técnica utiliza oligonucleotídeos de sequência aleatória durante a amplificação, eliminando qualquer viés em relação a vírus previamente conhecidos. Em plantas, o uso de RCA vem levando à detecção de um grande número de novas espécies de geminivírus em plantas cultivadas e não-cultivadas (Castillo-Urquiza et al., 2008; Rey et al., 2012; Silva et al., 2012; Tavares et al., 2012; Bernardo et al., 2013; Candresse et al., 2014), além de novos vírus de ssDNA em citrus (Loconsole et al., 2012), videiras (Krenz et al., 2012; Al-Rwahnih et al., 2013; Poojari et al., 2013) e outros hospedeiros.

O Capítulo 1 deste trabalho descreve a detecção e caracterização de um novo vírus com genoma de ssDNA em macieira, pereira e videira. Esse novo vírus apresenta baixa identidade

de sequências com outros vírus de ssDNA e provavelmente será classificado em uma nova família e gênero.

Características dos vírus pertencentes à família Geminiviridae

A família Geminiviridae é caracterizada pela morfologia de partículas icosaédricas geminadas e genoma composto por uma ou duas moléculas de ssDNA circular (Brown et al., 2012). A família é organizada em sete gêneros (Becurtovirus, Begomovirus, Curtovirus, Eragrovirus, Mastrevirus, Topocuvirus e Turncurtovirus) com base no número e organização dos componentes genômicos, relacionamento filogenético, tipo de inseto vetor (cigarrinha ou mosca-branca) e gama de hospedeiros (mono- ou dicotiledôneas) (Brown et al., 2012; Varsani et al., 2014). Os begomovírus são transmitidos pela mosca branca Bemisia tabaci e constituem um grupo de vírus de plantas de grande relevância econômica para a agricultura, devido à severidade das doenças por eles causadas em diversas culturas, principalmente em regiões tropicais e subtropicais (Polston & Anderson, 1997; Morales & Anderson, 2001; Monci et al., 2002; Briddon, 2003; Were et al., 2004). Os begomovírus podem ser divididos em dois grupos: Velho Mundo (Europa, África e Ásia) e Novo Mundo (Américas) (Rybicki, 1994; Padidam et al., 1999; Paximadis et al., 1999). A maioria dos begomovírus do Novo Mundo possuem dois componentes genômicos, denominados DNA-A e DNA-B. Os dois componentes não possuem identidade significativa de sequência, exceto em uma região com cerca de 200 nucleotídeos, denominada região comum, na qual a identidade nucleotídica entre o DNA-A e DNA-B é de aproximadamente 90%. Nesta região está localizada uma sequência capaz de formar uma estrutura em forma de grampo, contendo um nonanucleotídeo (TAATATTAC), conservado em praticamente todos os geminivírus, que constitui a origem de replicação por círculo rolante (Figura 1).

O DNA-A dos begomovírus do Novo Mundo contém cinco genes: Rep, responsável por codificar a única proteína essencial à replicação, iniciadora do mecanismo de replicação por círculo rolante (Fontes et al., 1992; Orozco et al., 1997); Trap, que codifica um fator transcricional dos genes CP e NSP e que também atua como supressora do silenciamento gênico mediado por RNA de interferência (RNAi) (Voinnet et al., 1999; Wang et al., 2005); Ren, que codifica um fator acessório (não essencial) à replicação viral (Sunter et al., 1990; Pedersen e Hanley-Bowdoin, 1994); AC4, responsável por codificar uma proteína que também está envolvida na supressão de silenciamento gênico mediado por RNAi (Vanitharani et al., 2004); e CP, que codifica a proteína capsidial, essencial também para a transmissão do vírus pelo inseto vetor (Briddon et al., 1990; Hofer et al., 1997). O DNA-B contém dois genes: NSP, que codifica a proteína responsável pelo transporte do DNA através do envelope

nuclear (núcleo-citoplasma) (Noueiry et al., 1994; Sanderfoot & Lazarowitz, 1996), e MP, que codifica a proteína envolvida no movimento célula-a-célula do vírus através dos plasmodesmas (Noueiry et al., 1994) (Figura 1).

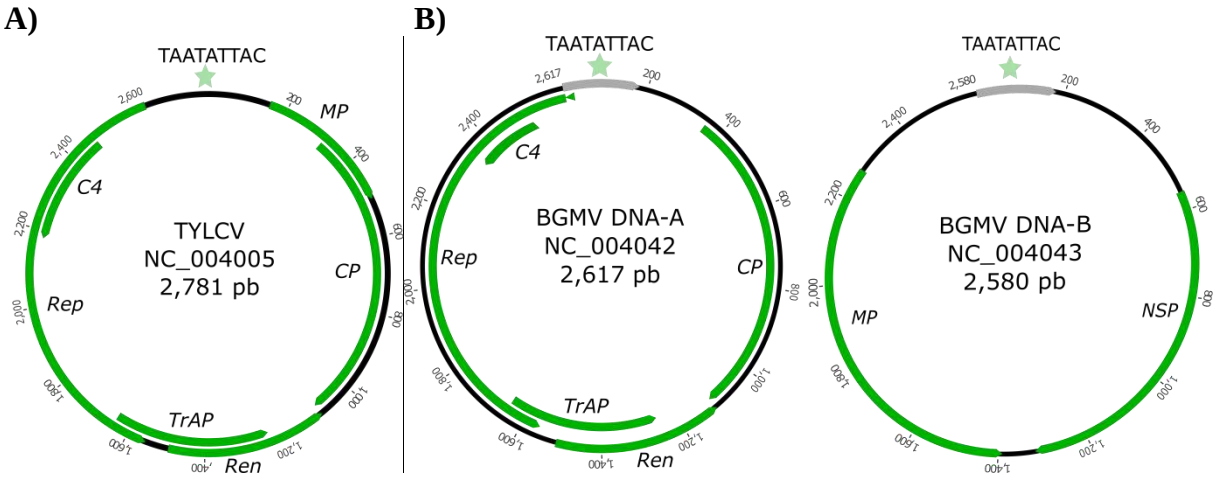


Figura 1. Organização genômica dos begomovírus **(A)** mono- e **(B)** bissegmentados. TYLCV: Tomato yellow leaf curl virus. BGMV: Bean golden mosaic virus.

O processo infeccioso dos geminivírus inicia-se no momento em que ocorre a deposição das partículas virais no espaço intracelular (por inoculação mecânica ou via inseto vetor). O genoma viral (ssDNA) é desencapsidado provavelmente de forma espontânea (Lazarowitz, 1992; Palmer & Rybicki, 1998) e direcionado ao núcleo da célula hospedeira, onde é convertido em um intermediário de fita dupla (dsDNA) que servirá de molde para síntese de mRNAs virais e de novos ssDNAs (Stenger et al., 1991; Stanley, 1995). A proteína Rep inicia o processo de replicação ao clivar o nonanucleotídeo conservado na região comum (TAATATT↓AC), funcionando como uma endonuclease sítio-específica com requerimento de estrutura e sequência (Laufs et al., 1995; Orozco & Hanley-Bowdoin, 1998).

Após a replicação viral na primeira célula infectada, é necessário que o vírus se mova para o estabelecimento da infecção sistêmica. O movimento no interior do hospedeiro pode ser dividido em dois processos: movimento célula-a-célula via plasmodesmas e movimento a longa distância ou sistêmico, no qual o vírus atinge o sistema vascular e é transportado via floema para outras partes da planta hospedeira. Para efetivar o movimento célula-a-célula, os geminivírus codificam a proteína MP. Esta se associa à membrana do retículo endoplasmático, acumulando-se nos desmotúbulos e alterando o limite de exclusão dos

plasmodesmas de forma a facilitar o transporte do genoma viral (Noueiry et al., 1994; Ward et al., 1997; Lazarowitz & Beachy, 1999). Como os geminivírus se multiplicam no núcleo da célula hospedeira, é necessária uma etapa adicional de transporte núcleo-citoplasma, que é realizada pela proteína NSP nos vírus bissegmentados (Sanderfoot et al., 1996) e pela proteína capsidial (CP) nos monossegmentados (Rojas et al., 2001).

O Tomato yellow spot virus (ToYSV) tem sido utilizado como modelo para o estudo de interações begomovírus-hospedeiro nos Laboratórios de Virologia Vegetal Molecular e de Biologia Molecular de Plantas da UFV. Apesar de ter sido isolado de tomateiro, suas características moleculares e filogenéticas são mais semelhantes às de begomovírus isolados da planta não-cultivada Sida sp., sugerindo que ele pode ter evoluído a partir de vírus que infectam essa planta (Fernandes et al., 1998; Jovel et al., 2004; Andrade et al., 2006). O ToYSV causa sintomas severos em tomateiro e em hospedeiros experimentais como Nicotiana benthamiana e N. glutinosa (Ambrozevicius et al., 2002; Andrade et al., 2006). A severidade dos sintomas pode estar associada, ao menos parcialmente, com o tropismo de tecido deste vírus, uma vez que em N. benthamiana o ToYSV é capaz de infectar além dos tecidos do floema, as células do mesofilo (Alves-Junior et al., 2009).

Os Capítulos 2 e 3 deste trabalho descrevem estudos realizados com o ToYSV com o objetivo de melhor compreender aspectos da interação molecular begomovírus-hospedeiro.

Características de smRNAs e seu papel na regulação da expressão gênica em eucariotos

A descoberta dos pequenos RNAs (smRNAs) e o entendimento detalhado de seus mecanismos regulatórios revolucionaram o campo da genética nos últimos anos. Em plantas existem duas principais classes de smRNAs regulatórios: small interfering RNAs (siRNAs) e microRNAs (miRNAs) (Chen, 2009). Ambos são gerados a partir de precursores de RNA de fita dupla (dsRNA) apresentando maior ou menor grau de pareamento entre as fitas. O processamento é realizado por proteínas denominadas Dicer. O dsRNA é degradado por essas enzimas em intervalos de 21 a 24 nucleotídeos (nt) (Hamilton e Baulcombe, 1999) e em seguida o smRNA é incorporado em um complexo proteico denominado RISC (RNA-induced silencing complex). Proteínas argonautas apresentam função chave no RISC, pois se ligam diretamente ao smRNA e promovem a regulação do mRNA alvo (Mallory et al., 2008). Na planta modelo Arabidopsis thaliana existem quatro dicers e dez argonautas. A DCL1 e a AGO1 estão preferencialmente envolvidas na geração de miRNAs (Xie et al., 2004; Baumberger & Baulcombe, 2005). Já os siRNAs podem ser gerados por qualquer uma das demais Dicers (DCL2, DCL3, DCL4) e se incorporam no RISC que podem conter diferentes argonautas (Bouche et al., 2006; Deleris et al., 2006; Fusaro et al., 2006; Mi et al., 2008).

Os genes codificadores de miRNAs podem estar localizados dentro de éxons, íntrons ou regiões intergênicas. Em geral, apresentam entre 100 e 400 nucleotídeos e são transcritos pela RNA polimerase II, gerando um precursor com estrutura secundária em forma de grampo (Bartel, 2004). Essa estrutura de fita dupla imperfeita é reconhecida pela DCL1, que a degrada e libera o chamado miRNA maduro, geralmente de 21 nt (Kurihara & Watanabe, 2008). Em seguida, uma das fitas do miRNA, chamada de fita guia, entra no RISC, enquanto a outra fita, representada como miRNA* (ou sequência estrela), provavelmente é descartada (Baumberger & Baulcombe, 2005; Qi et al., 2005). O complexo RISC é então direcionado para qualquer RNA apresentando similaridade de sequência com o miRNA incorporado. Dependendo da argonauta presente, os RNAs alvos são reprimidos por degradação direta ou inibição da tradução (Huntzinger & Izaurralde, 2011).

Ao contrário dos miRNAs, os siRNAs são gerados por precursores de RNAs apresentando uma dupla fita de RNA perfeita, ou seja, com todas ou quase todas as bases pareadas. Além disso, a origem do precursor em si também é distinta. Por exemplo, o dsRNA pode ser gerado após a transcrição de genes em mRNA (vírus com genoma de RNA e DNA) ou durante a replicação viral (vírus com genoma de RNA). Moléculas de dsRNA oriundas do genoma viral são reconhecidas e degradadas por Dicers (Bouche et al., 2006; Deleris et al., 2006; Fusaro et al., 2006; Mi et al., 2008). Os siRNAs podem ser gerados também por transposons ou estruturas em repetição invertida no genoma, estando envolvidos no processo de proteção do genoma e inclusive na formação da heterocromatina (Zhang & Zhu, 2011).

Até recentemente, a maioria dos estudos realizados sobre smRNAs foram concentrados em A. thaliana, Oryza sativa e Populus trichocarpa. Usando técnicas de sequenciamento em baixa escala, concluiu-se que a maioria dos miRNAs eram conservados, com alguns homólogos encontrados inclusive em musgos (Reinhart et al., 2002; Arazi et al., 2005). Muitos desses miRNAs conservados apresentam padrões específicos de expressão e em sua grande maioria atuam em fatores reguladores de expressão gênica envolvidos principalmente na regulação do desenvolvimento (Rubio-Somoza & Weigel, 2011). Por exemplo, a expressão de proteínas envolvidas na formação de polaridade de folha e na transição da fase vegetativa para a reprodutiva são reguladas diretamente por miRNAs ou tasiRNAs conservados (Juarez et al., 2004; Peragine et al., 2004). Consistente com esses dados, mutantes para os genes da via de miRNAs apresentam interferências significativas no desenvolvimento (Jacobsen et al., 1999; Lu & Fedoroff, 2000; Morel et al., 2002; Boutet et al., 2003).

No entanto, o advento dos métodos de sequenciamento massivo de DNA permitiu a identificação de um maior número de smRNAs. Os resultados iniciais evidenciaram claramente a existência de uma ampla gama de smRNAs não conservados, mesmo em plantas

muito próximas (Rajagopalan et al., 2006; Fahlgren et al., 2007; Ma et al., 2010), indicando que novos miRNAs podem surgir continuamente ao longo do processo evolutivo. Apenas cerca de 1/9 dos miRNAs não conservados identificados apresentam fatores de transcrição como alvos, sugerindo que esses RNAs podem estar envolvidos na regulação de mecanismos que são específicos para cada grupo taxonômico (Willmann & Poethig, 2007). De fato, miRNAs não conservados foram observados em vias de diferentes tipos de estresses, tanto bióticos quanto abióticos (Shukla et al., 2008).

smRNAs na interação molecular planta-vírus

Os smRNAs estão diretamente envolvidos na complexa rede de interação entre plantas e patógenos (Katiyar-Agarwal & Jin, 2010). Os siRNAs de origem viral, também chamados de vsiRNAs (viral small interfering RNAs), são um importante exemplo desse tipo de interação (Llave, 2010). Genomas virais podem ser degradados por Dicers gerando vsiRNAs de 18 a 28 nt. A origem exata do dsRNA viral que serve como substrato para degradação ainda é um pouco controversa. Existem evidências de que tanto intermediários replicativos ou estruturas secundárias no RNA viral podem ser os alvos da maquinaria de silenciamento. No entanto, sabe-se que em A. thaliana o dsRNA viral é degradado preferencialmente pela DCL4, o que explicaria o maior acúmulo de vsiRNAs de 21 nt na maioria das plantas analisadas até o momento (Donaire et al., 2009). No entanto, as enzimas DCL2 e DCL3 também possuem acesso ao genoma viral, gerando vsiRNAs de 22 e 23 nt, respectivamente. Os vsiRNAs produzidos são usados como guia do RISC para degradar ssRNAs do genoma invasor (Llave, 2010). No entanto, vsiRNAs também podem controlar a expressão de genes endógenos, adicionando assim, mais uma camada nessa intrincada rede regulatória (Shimura et al., 2011; Smith et al., 2011). Vírus de plantas, em contrapartida, desenvolveram proteínas que são capazes de suprimir as respostas de defesa do hospedeiro mediada por RNAi, permitindo que a infecção viral ocorra (Burgyan, 2006). Muitas destas proteínas supressoras estudadas até o momento atuam por meio de ligação aos siRNAs, evitando que eles sejam complexados ao RISC e direcionem a degradação do genoma viral (Lakatos et al., 2006). No entanto, já foram encontradas proteínas virais supressoras de RNAi atuando em praticamente todas as etapas conhecidas da via.

Sabe-se também que os vírus podem alterar drasticamente o perfil de expressão de miRNAs de seus hospedeiros (Kasschau et al., 2003; Bazzini et al., 2007; Bazzini et al., 2009; Lang et al., 2011). Proteínas virais supressoras de RNAi que atuam em etapas comuns às vias de miRNA e siRNA, em alguns casos, podem causar a desregulação das vias endógenas de smRNAs. Em acordo com essa hipótese, plantas transgênicas expressando

determinadas proteínas supressoras virais apresentam fenótipos semelhantes aos de plantas infectadas pelos vírus que codificam as respectivas proteínas. Recentemente, observou-se que a desregulação do alvo do mesmo miRNA (MIR167) é a principal causa dos sintomas induzidos por pelo menos três proteínas supressoras virais distintas em A. thaliana (Jay et al., 2011). Além disso, observou-se que vírus podem alterar a taxa de transcrição de determinados genes MIR e também induzir a produção de miRNAs (Bazzini et al., 2009). A identificação e caracterização de genes MIR expressos especificamente durante a infecção viral é de grande importância, uma vez que estes podem tanto estar atuando em benefício ou contra a infecção viral.

Interações begomovírus-hospedeiro envolvendo a proteína MP

Para o estabelecimento da infecção sistêmica é necessária à interação molecular entre fatores virais e do hospedeiro que possibilitem o movimento viral célula-a-célula e a longa distância (Sanderfoot & Lazarowitz, 1996).

A proteína MP é essencial para o movimento célula-a-célula dos begomovírus com genoma bissegmentado (Noueiry et al., 1994). A MP do Bean dwarf mosaic virus (BDMV) é responsável pelo transporte do DNA viral através dos plasmodesmas (Noueiry et al., 1994), e localiza-se ao redor do núcleo, em frações da membrana plasmática e na parede celular (Von Arnim & Stanley, 1992 ; Pascal et al., 1993). Tanto ssDNA quanto dsDNA podem ser transportados célula-a-célula, mas provavelmente dsDNA é a principal forma de movimento do BDMV (Rojas et al., 1998). Sanderfoot & Lazarowitz (1995) demonstraram que MP e NSP atuam de maneira cooperativa para o movimento do Squash leaf curl virus (SqLCV). Como o BDMV e o SqLCV apresentam tropismo de tecido diferente (o SqLCV é restrito ao floema e o BDMV infecta células do mesofilo), é de se esperar que estes begomovírus apresentem diferenças nos mecanismos de movimento célula-a-célula. Assim, foi proposto que após o transporte do DNA viral do BDMV do núcleo para o citoplasma, mediado por NSP, o DNA é transferido de NSP para MP formando um complexo MP-dsDNA, o qual se movimenta de uma célula para outra via plasmodesmas (Levy & Tzfira, 2010). Já para o SqLCV, Sanderfoot & Lazarowitz (1995) sugerem que um complexo NSP-ssDNA move-se célula-a-célula via plasmodesmas por intermédio da ligação com MP.

As MPs do BDMV e do Abutilon mosaic virus (AbMV) podem ser fosforiladas na região C-terminal por cinases associadas a parede celular, e esta modificação pós-traducional pode ser necessária para o aumento do limite de exclusão dos plasmodesmas (Levy & Tzfira, 2010). Kleinow et al. (2008) demonstraram que mutações nos três diferentes sítios de fosforilação da MP do AbMV afetam o desenvolvimento de sintomas e o acúmulo de DNA

viral. Além de modificações pós-traducionais, a interação de MP com fatores virais e do hospedeiro também é necessária. A interação entre MP e NSP foi demonstrada por meio da re-localização de NSP por MP em protoplastos de N. tabacum e também pela interação física demonstrada no sistema duplo-híbrido de levedura (Sanderfoot & Lazarowitz, 1995; Mariano et al., 2004).

Interações entre a proteína MP e as proteínas sinaptotagmina de A. thaliana (SYTA) (Lewis & Lazarowitz, 2010) e HSC70 (heat shock cognate 70 kDa) (Krenz et al., 2010) já foram relatadas. Utilizando o sistema de recrutamento SOS de levedura, Lewis & Lazarowitz (2010) demonstraram que a MP do begomovírus Cabbage leaf curl virus (CaLCuV) interage com SYTA de A. thaliana. SYTA regula o transporte célula-a-célula da MP do CaLCuV e é necessária para o movimento sistêmico e infecção de A. thaliana. Por meio de ensaio de duplo-híbrido verificou-se a interação entre HSC70 de A. thaliana e a MP do AbMV. Esta interação foi confirmada in vivo utilizando-se o ensaio de complementação de fluorescência bimolecular (BiFC) (Krenz et al., 2010). Os resultados sugerem que a interação HSC70 e MP é importante para o transporte viral e indução de sintomas pelo AbMV.

A identificação e o estudo de proteínas do hospedeiro que interagem direta ou indiretamente com a MP poderão contribuir para melhorar o entendimento do processo de infecção viral e da interação begomovírus-hospedeiro e fornecer novas estratégias de controle destas viroses, especialmente com relação a resistência do hospedeiro, seja ela natural ou derivada do patógeno.

LITERATURA CITADA

AL-RWAHNIH, M.; DAVE, A.; ANDERSON, M.M.; ROWHANI, A.; UYEMOTO, J.K.; SUDARSHANA, M.R. Association of a DNA virus with grapevines affected by red blotch disease in California. **Phytopathology,** v. 103, p. 1069-1076, 2013.

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. Synergism and negative interference during co-infection of tomato and Nicotiana benthamiana with two bipartite begomoviruses. **Virology,** v. 387, p. 257-266, 2009.

AMBROZEVICIUS, L.P.; CALEGARIO, R.F.; FONTES, E.P.B.; CARVALHO, M.G.; ZERBINI, F.M. Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. **Fitopatologia Brasileira,** v. 27, p. 372-377, 2002.

ANDRADE, E.C.; MANHANI, G.G.; ALFENAS, P.F.; CALEGARIO, R.F.; FONTES, E.P.B.; ZERBINI, F.M. Tomato yellow spot virus, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from Sida sp., forms pseudorecombinants with begomoviruses from tomato but not from Sida. **Journal of General Virology,** v. 87, p. 3687-3696, 2006.

ARAZI, T.; TALMOR-NEIMAN, M.; STAV, R.; RIESE, M.; HUIJSER, P.; BAULCOMBE, D.C. Cloning and characterization of micro-RNAs from moss. **Plant Journal,** v. 43, p. 837-848, 2005.

BARTEL, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. **Cell,** v. 116, p. 281-297, 2004.

BAUMBERGER, N.; BAULCOMBE, D.C. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. **Proceedings of the National Academy of Sciences, USA,** v. 102, p. 11928-11933, 2005.

BAZZINI, A.A.; ALMASIA, N.I.; MANACORDA, C.A.; MONGELLI, V.C.; CONTI, G.; MARONICHE, G.A.; RODRIGUEZ, M.C.; DISTEFANO, A.J.; HOPP, H.E.; DEL VAS, M.; ASURMENDI, S. Virus infection elevates transcriptional activity of miR164a promoter in plants. **BMC Plant Biology,** v. 9, p. 152, 2009.

BAZZINI, A.A.; HOPP, H.E.; BEACHY, R.N.; ASURMENDI, S. Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptoms and plant development. **Proceedings of the National Academy of Sciences, USA,** v. 104, p. 12157-12162, 2007.

BERNARDO, P.; GOLDEN, M.; AKRAM, M.; NAIMUDDIN; NADARAJAN, N.; FERNANDEZ, E.; GRANIER, M.; REBELO, A.G.; PETERSCHMITT, M.; MARTIN, D.P.; ROUMAGNAC, P. Identification and characterisation of a highly divergent geminivirus: Evolutionary and taxonomic implications. **Virus Research,** v. 177, p. 35-45, 2013.

BOUCHE, N.; LAURESSERGUES, D.; GASCIOLLI, V.; VAUCHERET, H. An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. **EMBO Journal,** v. 25, p. 3347-3356, 2006.

BOUTET, S.; VAZQUEZ, F.; LIU, J.; BECLIN, C.; FAGARD, M.; GRATIAS, A.; MOREL, J.B.; CRETE, P.; CHEN, X.M.; VAUCHERET, H. Arabidopsis HEN1: A genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. **Current Biology,** v. 13, p. 843-848, 2003.

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

BRIDDON, R.W.; PINNER, M.S.; STANLEY, J.; MARKHAM, P.G. Geminivirus coat protein gene replacement alters insect specificity. **Virology,** v. 177, p. 85-94, 1990.

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

BROWN, J.K.; ZERBINI, F.M.; NAVAS-CASTILLO, J.; MORIONES, E.; RAMOS-SOBRINHO, R.; SILVA, J.C.F.; FIALLO-OLIVE, E.; BRIDDON, R.W.; HERNÁNDEZ-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, no prelo, 2015.

BURGYAN, J. Virus induced RNA silencing and suppression: Defence and counter defence. **Journal of Plant Pathology,** v. 88, p. 233-244, 2006.

CANDRESSE, T.; FILLOUX, D.; MUHIRE, B.; JULIAN, C.; GALZI, S.; FORT, G.; BERNARDO, P.; DAUGROIS, J.H.; FERNANDEZ, E.; MARTIN, D.P.; VARSANI, A.; ROUMAGNAC, P. Appearances can be deceptive: Revealing a hidden viral infection with deep sequencing in a plant quarantine context. **PLoS One,** v. 9, p. e102945, 2014.

CASTILLO-URQUIZA, G.P.; BESERRA JR., J.E.A.; BRUCKNER, F.P.; LIMA, A.T.M.; VARSANI, A.; ALFENAS-ZERBINI, P.; ZERBINI, F.M. Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. **Archives of Virology,** v. 153, p. 1985-1989, 2008.

CHEN, X.M. Small RNAs and their roles in plant development. **Annual Review of Cell and Developmental Biology,** v. 25, p. 21-44, 2009.

DELERIS, A.; GALLEGO-BARTOLOME, J.; BAO, J.S.; KASSCHAU, K.D.; CARRINGTON, J.C.; VOINNET, O. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. **Science,** v. 313, p. 68-71, 2006.

DONAIRE, L.; WANG, Y.; GONZALEZ-IBEAS, D.; MAYER, K.F.; ARANDA, M.A.; LLAVE, C. Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. **Virology,** v. 392, p. 203-214, 2009.

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.

FAHLGREN, N.; HOWELL, M.D.; KASSCHAU, K.D.; CHAPMAN, E.J.; SULLIVAN, C.M.; CUMBIE, J.S.; GIVAN, S.A.; LAW, T.F.; GRANT, S.R.; DANGL, J.L.; CARRINGTON, J.C. High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. **PLoS One,** v. 2, p. e219, 2007.

FERNANDES, A.V.; ANDRADE, E.C.; ALFENAS, P.F.; GALVÃO, R.M.; ZERBINI, F.M.; FONTES, E.P.B. Clonagem e caracterização de novos geminivírus que infectam Sida rhombifolia. **Fitopatologia Brasileira,** v. 23, p. 317, 1998.

FONTES, E.P.B.; LUCKOW, V.A.; HANLEY-BOWDOIN, L. A geminivirus replication protein is a sequence-specific DNA binding protein. **Plant Cell,** v. 4, p. 597-608, 1992.

FUSARO, A.F.; MATTHEW, L.; SMITH, N.A.; CURTIN, S.J.; DEDIC-HAGAN, J.; ELLACOTT, G.A.; WATSON, J.M.; WANG, M.B.; BROSNAN, C.; CARROLL, B.J.; WATERHOUSE, P.M. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. **EMBO Reports,** v. 7, p. 1168-1175, 2006.

HAMILTON, A.J.; BAULCOMBE, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. **Science,** v. 286, p. 950-952, 1999.

HOFER, P.; BEDFORD, I.D.; MARKHAM, P.G.; JESKE, H.; FRISCHMUTH, T. Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. **Virology,** v. 236, p. 288-295, 1997.

HUNTZINGER, E.; IZAURRALDE, E. Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. **Nature Reviews Genetics,** v. 12, p. 99-110, 2011.

JACOBSEN, S.E.; RUNNING, M.P.; MEYEROWITZ, E.M. Disruption of an RNA helicase/RNAse III gene in Arabidopsis causes unregulated cell division in floral meristems. **Development,** v. 126, p. 5231-5243, 1999.

JAY, F.; WANG, Y.; YU, A.; TACONNAT, L.; PELLETIER, S.; COLOT, V.; RENOU, J.P.; VOINNET, O. Misregulation of AUXIN RESPONSE FACTOR 8 underlies the developmental abnormalities caused by three distinct viral silencing suppressors in Arabidopsis. **PLoS Pathogens,** v. 7, p. e1002035, 2011.

JOVEL, J.; RESKI, G.; ROTHENSTEIN, D.; RINGEL, M.; FRISCHMUTH, T.; JESKE, H. Sida micrantha mosaic is associated with a complex infection of begomoviruses different from Abutilon mosaic virus. **Archives of Virology,** v. 149, p. 829-841, 2004.

JUAREZ, M.T.; KUI, J.S.; THOMAS, J.; HELLER, B.A.; TIMMERMANS, M.C. microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. **Nature,** v. 428, p. 84-88, 2004.

KASSCHAU, K.D.; XIE, Z.; ALLEN, E.; LLAVE, C.; CHAPMAN, E.J.; KRIZAN, K.A.; CARRINGTON, J.C. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. **Developmental Cell,** v. 4, p. 205-217, 2003.

KATIYAR-AGARWAL, S.; JIN, H. Role of small RNAs in host-microbe interactions. **Annual Review of Phytopathology,** v. 48, p. 225-246, 2010.

KING, A.M.Q.; ADAMS, M.J.; CARSTENS, E.B.; LEFKOWITZ, E.J. (Eds.) **Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses**. London, UK: Elsevier Academic Press, 1327p. 2012.

KLEINOW, T.; HOLEITER, G.; NISCHANG, M.; STEIN, M.; KARAYAVUZ, M.; WEGE, C.; JESKE, H. Post-translational modifications of Abutilon mosaic virus movement protein (BC1) in fission yeast. **Virus Research,** v. 131, p. 86-94, 2008.

KRENZ, B.; THOMPSON, J.R.; FUCHS, M.; PERRY, K.L. Complete genome sequence of a new circular DNA virus from grapevine. **Journal of Virology,** v. 86, p. 7715, 2012.

KRENZ, B.; WINDEISEN, V.; WEGE, C.; JESKE, H.; KLEINOW, T. A plastid-targeted heat shock cognate 70 kDa protein interacts with the Abutilon mosaic virus movement protein. **Virology,** v. 401, p. 6-17, 2010.

KURIHARA, Y.; WATANABE, Y. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. **Proceedings of the National Academy of Sciences, USA,** v. 101, p. 12753-12758, 2008.

LAKATOS, L.; CSORBA, T.; PANTALEO, V.; CHAPMAN, E.J.; CARRINGTON, J.C.; LIU, Y.P.; DOLJA, V.V.; CALVINO, L.F.; LOPEZ-MOYA, J.J.; BURGYAN, J. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. **EMBO Journal,** v. 25, p. 2768-2780, 2006.

LANG, Q.; JIN, C.; LAI, L.; FENG, J.; CHEN, S.; CHEN, J. Tobacco microRNAs prediction and their expression infected with Cucumber mosaic virus and Potato virus X. **Molecular Biology Reports,** v. 38, p. 1523-1531, 2011.

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.

LAZAROWITZ, S.G. Geminiviruses: Genome structure and gene function. **Critical Reviews in Plant Sciences,** v. 11, p. 327-349, 1992.

LAZAROWITZ, S.G.; BEACHY, R.N. Viral movement proteins as probes for intracellular and intercellular trafficking in plants. **Plant Cell,** v. 11, p. 535-548, 1999.

LEVY, A.; TZFIRA, T. Bean dwarf mosaic virus: A model system for the study of viral movement. **Molecular Plant Pathology,** v. 11, p. 451-461, 2010.

LEWIS, J.D.; LAZAROWITZ, S.G. Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. **Proceedings of the National Academy of Sciences, USA,** v. 107, p. 2491-2496, 2010.

LLAVE, C. Virus-derived small interfering RNAs at the core of plant-virus interactions. **Trends in Plant Science,** v. 15, p. 701-707, 2010.

LOCONSOLE, G.; SALDARELLI, P.; DODDAPANENI, H.; SAVINO, V.; MARTELLI, G.P.; SAPONARI, M. Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. **Virology,** v. 432, p. 162-172, 2012.

LU, C.; FEDOROFF, N. A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. **Plant Cell,** v. 12, p. 2351-2366, 2000.

MA, Z.R.; CORUH, C.; AXTELL, M.J. Arabidopsis lyrata small RNAs: Transient MIRNA and small interfering RNA loci within the Arabidopsis genus. **Plant Cell,** v. 22, p. 1090-1103, 2010.

MALLORY, A.C.; ELMAYAN, T.; VAUCHERET, H. MicroRNA maturation and action-the expanding roles of ARGONAUTEs. **Current Opinion in Plant Biology,** v. 11, p. 560-566, 2008.

MARIANO, A.C.; ANDRADE, M.O.; SANTOS, A.A.; CAROLINO, S.M.B.; OLIVEIRA, M.L.; BARACAT-PEREIRA, M.C.; BROMMONSHENKEL, S.H.; FONTES, E.P.B. Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. **Virology,** v. 318, p. 24-31, 2004.

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

MI, S.; CAI, T.; HU, Y.; CHEN, Y.; HODGES, E.; NI, F.; WU, L.; LI, S.; ZHOU, H.; LONG, C.; CHEN, S.; HANNON, G.J.; QI, Y. Sorting of small RNAs into Arabidopsis

argonaute complexes is directed by the 5' terminal nucleotide. **Cell,** v. 133, p. 116-127, 2008.

MONCI, F.; SANCHEZ-CAMPOS, S.; NAVAS-CASTILLO, J.; MORIONES, E. A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. **Virology,** v. 303, p. 317-326, 2002.

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

MOREL, J.-B.; GODON, C.; MOURRAIN, P.; BÉCLIN, C.; BOUTET, S.; FEUERBACH, F.; PROUX, F.; VAUCHERET, H. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. **Plant Cell,** v. 14, p. 629-639, 2002.

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.

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

OROZCO, B.M.; HANLEY-BOWDOIN, L. Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. **Journal of Biological Chemistry,** v. 273, p. 24448-24456, 1998.

OROZCO, B.M.; MILLER, A.B.; SETTLAGE, S.B.; HANLEY-BOWDOIN, L. Functional domains of a geminivirus replication protein. **Journal of Biological Chemistry,** v. 272, p. 9840-9846, 1997.

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

PALMER, K.E.; RYBICKI, E.P. The molecular biology of mastreviruses. **Advances in Virus Research,** v. 50, p. 183-234, 1998.

PASCAL, E.; GOODLOVE, P.E.; WU, L.C.; LAZAROWITZ, S.G. Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. **Plant Cell,** v. 5, p. 795-807, 1993.

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

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

PERAGINE, A.; YOSHIKAWA, M.; WU, G.; ALBRECHT, H.L.; POETHIG, R.S. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. **Genes & Development,** v. 18, p. 2368-2379, 2004.

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

POOJARI, S.; ALABI, O.J.; FOFANOV, V.Y.; NAIDU, R.A. A leafhopper-transmissible DNA virus with novel evolutionary lineage in the family geminiviridae implicated in grapevine redleaf disease by next-generation sequencing. **PLoS One,** v. 8, p. e64194, 2013.

QI, Y.; DENLI, A.M.; HANNON, G.J. Biochemical specialization within Arabidopsis RNA silencing pathways. **Molecular Cell,** v. 19, p. 421-428, 2005.

RAJAGOPALAN, R.; VAUCHERET, H.; TREJO, J.; BARTEL, D.P. A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. **Genes & Development,** v. 20, p. 3407-3425, 2006.

REINHART, B.J.; WEINSTEIN, E.G.; RHOADES, M.W.; BARTEL, B.; BARTEL, D.P. MicroRNAs in plants. **Genes & Development,** v. 16, p. 1616-1626, 2002.

REY, M.E.C.; NDUNGURU, J.; BERRIE, L.C.; PAXIMADIS, M.; BERRY, S.; COSSA, N.; NUAILA, V.N.; MABASA, K.G.; ABRAHAM, N.; RYBICKI, E.P.; MARTIN, D.; PIETERSEN, G.; ESTERHUIZEN, L.L. Diversity of dicotyledenous-infecting geminiviruses and their associated DNA molecules in Southern Africa, including the South-West Indian Ocean islands. **Viruses,** v. 4, p. 1753-1791, 2012.

ROJAS, M.R.; JIANG, H.; SALATI, R.; XOCONOSTLE-CAZARES, B.; SUDARSHANA, M.R.; LUCAS, W.J.; GILBERTSON, R.L. Functional analysis of proteins involved in movement of the monopartite begomovirus, tomato yellow leaf curl virus. **Virology,** v. 291, p. 110-125, 2001.

ROJAS, M.R.; NOUEIRY, A.O.; LUCAS, W.J.; GILBERTSON, R.L. Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. **Cell,** v. 95, p. 105-113, 1998.

ROSARIO, K.; DUFFY, S.; BREITBART, M. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. **Archives of Virology,** v. 157, p. 1851-1871, 2012.

RUBIO-SOMOZA, I.; WEIGEL, D. MicroRNA networks and developmental plasticity in plants. **Trends in Plant Science,** v. 16, p. 258-264, 2011.

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

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

SANDERFOOT, A.A.; LAZAROWITZ, S.G. Cooperation in viral movement: The geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. **Plant Cell,** v. 7, p. 1185-1194, 1995.

SANDERFOOT, A.A.; LAZAROWITZ, S.G. Getting it together in plant virus movement: Cooperative interactions between bipartite geminivirus movement proteins. **Trends in Cell Biology,** v. 6, p. 353-358, 1996.

SHIMURA, H.; PANTALEO, V.; ISHIHARA, T.; MYOJO, N.; INABA, J.; SUEDA, K.; BURGYAN, J.; MASUTA, C. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. **PLoS Pathogens,** v. 7, p. e1002021, 2011.

SHUKLA, L.I.; CHINNUSAMY, V.; SUNKAR, R. The role of microRNAs and other endogenous small RNAs in plant stress responses. **Biochimica et Biophysica Acta-Gene Regulatory Mechanisms,** v. 1779, p. 743-748, 2008.

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

SMITH, N.A.; EAMENS, A.L.; WANG, M.B. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. **PLoS Pathogens,** v. 7, p. e1002022, 2011.

STANLEY, J. Analysis of African cassava mosaic virus recombinants suggest strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. **Virology,** v. 206, p. 707-712, 1995.

STENGER, D.C.; REVINGTON, G.N.; STEVENSON, M.C.; BISARO, D.M. Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling-circle replication of a plant viral DNA. **Proceedings of the National Academy of Sciences, USA,** v. 88, p. 8029-8033, 1991.

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

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

TERRY, F.F.N.; MARINE, R.; WANG, C.L.; SIMMONDS, P.; KAPUSINSZKY, B.; BODHIDATTA, L.; ODERINDE, B.S.; WOMMACK, K.E.; DELWART, E. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. **Journal of Virology,** v. 86, p. 12161-12175, 2012.

VANITHARANI, R.; CHELLAPPAN, P.; PITA, J.S.; FAUQUET, C.M. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. **Journal of Virology,** v. 78, p. 9487-9498, 2004.

VARSANI, A.; NAVAS-CASTILLO, J.; MORIONES, E.; HERNÁNDEZ-ZEPEDA, C.; IDRIS, A.; BROWN, J.K.; ZERBINI, F.M.; MARTIN, D.P. Establishment of three new

genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. **Archives of Virology,** v. 159, p. 2193-2203, 2014.

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

VON ARNIM, A.; STANLEY, J. Inhibition of African cassava mosaic virus systemic infection by a movement protein from the related geminivirus Tomato golden mosaic virus. **Virology,** v. 187, p. 555-564, 1992.

WANG, H.; BUCKLEY, K.J.; YANG, X.; BUCHMANN, R.C.; BISARO, D.M. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. **Journal of Virology,** v. 79, p. 7410-7418, 2005.

WARD, B.M.; MEDVILLE, R.; LAZAROWITZ, S.G.; TURGEON, R. The geminivirus BL1 movement protein is associated with endoplasmic reticulum-derived tubules in developing phloem cells. **Journal of Virology,** v. 71, p. 3726-3733, 1997.

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

WILLMANN, M.R.; POETHIG, R.S. Conservation and evolution of miRNA regulatory programs in plant development. **Current Opinion in Plant Biology,** v. 10, p. 503-511, 2007.

XIE, Z.; JOHANSEN, L.K.; GUSTAFSON, A.M.; KASSCHAU, K.D.; LELLIS, A.D.; ZILBERMAN, D.; JACOBSEN, S.E.; CARRINGTON, J.C. Genetic and functional diversification of small RNA pathways in plants. **PLoS Biology,** v. 2, p. e104, 2004.

ZAWAR-REZA, P.; ARGUELLO-ASTORGA, G.R.; KRABERGER, S.; JULIAN, L.; STAINTON, D.; BROADY, P.A.; VARSANI, A. Diverse small circular single-stranded DNA viruses identified in a freshwater pond on the McMurdo Ice Shelf (Antarctica). **Infection Genetics and Evolution,** v. 26, p. 132-138, 2014.

ZHANG, H.; ZHU, J.K. RNA-directed DNA methylation. **Current Opinion in Plant Biology,** v. 14, p. 142-147, 2011.

CAPÍTULO 1

A NOVEL, HIGHLY DIVERGENT ssDNA VIRUS INFECTING APPLE, PEAR AND GRAPEVINE IN BRAZIL

Basso, M.F., Silva, J.C.F., Fajardo, T.V.M., Fontes, E.P.B. & P. Zerbini, F.M (2015) A novel, highly divergent ssDNA virus infecting apple, pear and grapevine in Brazil. Submitted to Virus Research.

A novel, highly divergent ssDNA virus infecting apple, pear and grapevine in Brazil

Marcos Fernando Basso[1,2], José Cleydson Ferreira da Silva[2,3], Thor Vinícius Martins Fajardo[4], Elizabeth Pacheco Batista Fontes[2,5] & Francisco Murilo Zerbini[1,2#]

[1]Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [2]National Research Institute for Plant-Pest Interactions (INCT-IPP), Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [3]Departamento de Informática, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [4]Centro Nacional de Pesquisa Uva e Vinho (CNPUV), Embrapa Uva e Vinho, Bento Gonçalves, RS, 95700-000, Brazil; [5]Departamento de Bioquímica, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil;

[#]Corresponding author: Francisco Murilo Zerbini

Phone: (+55-31) 3899 2934 - Fax: (+55-31) 3899 2934 - E-mail: zerbini@ufv.br

ABSTRACT

Fruit trees of temperate and tropical climates are of great economical importance worldwide and several viruses have been reported affecting their productivity and longevity. Fruit trees of different Brazilian regions displaying virus-like symptoms were evaluated for infection by circular DNA viruses. Seventy-four fruit trees were sampled and a novel, highly divergent, monopartite circular ssDNA virus was cloned from apple, pear and grapevine trees. Forty-five complete viral genomes were sequenced, with a size of approx. 3.4 kb and organized into five ORFs. A large intergenic region contains a short palindromic sequence capable of forming a hairpin-like structure with the loop sequence TAGTATTAC, identical to the conserved nonanucleotide of circoviruses, nanoviruses and alphasatellites. Deduced amino acid sequences showed identities in the range of

35% with nanoviruses and alphasatellites (putative Replication-associated protein, Rep), and begomo-, curto- and mastreviruses (putative coat protein, CP, and movement protein, MP). Recombination events were not detected and phylogenetic analysis showed a relationship with circo-, nano- and geminiviruses. PCR and indirect ELISA proved the association of this novel ssDNA virus with field plants. Infectivity tests using the cloned viral genome confirmed its ability to infect apple and pear tree seedlings, but not Nicotiana benthamiana. The name "Temperate fruit decay-associated virus" (TFDaV) is proposed for this novel virus.

Keywords: ssDNA viruses; emerging viruses; novel viruses.

1. INTRODUCTION

Fruit trees of temperate and tropical climates are of great economical importance worldwide. They are susceptible to several arthropod- and graft-transmitted agents that cause several diseases, some of which are aetiologically unknown (Al Rwahnih et al., 2009; Hadidi et al., 2011). Vegetative propagation contributes to the spread of these pathogens, favoring the emergence of complex diseases (Martelli, 2012). The frequent exchange of propagative material between nurseries worldwide has contributed to their dissemination into new areas. The perennial life cycle of fruit trees also contributes to further accelerate the mixing and introduction of many viruses and subviral agents into a single plant (Hadidi et al., 2011; Rowhani et al., 2005).

Single-stranded (ss) DNA viruses represent a group of economically important plant pathogens. They are predisposed to recombination events and can present high nucleotide (nt) substitution rates, which contribute to their emergence (Lefeuvre et al., 2009; Martin et al., 2011). A method that has been used successfully for the detection of ssDNA viruses is rolling-circle amplification (RCA) (Inoue-Nagata et al., 2004). Due to its sequence-unbiased nature, this method has allowed the identification of several new ssDNA viral agents in different plant species.

In Brazilian apple and pear orchards and vineyards, symptoms of possible viral origin are frequenly observed (Basso et al., 2014). With the objective of prospecting ssDNA viruses associated with these three hosts, samples displaying virus-like symptoms were collected at different regions of the country and were evaluated by PCR and RCA. We report here the cloning and molecular characterization of a novel, highly divergent, monopartite circular ssDNA virus of approx. 3.4 kb in length, with characteristics similar to members of the Circo-, Nano- and Geminiviridae families. The name "Temperate fruit decay-associated virus" (TFDaV) is proposed for this new virus.

2. MATERIAL AND METHODS

2.1. Plant sampling

Apple, pear and grapevine samples (adult plants over two years old) displaying virus-like symptoms such as chlorotic dwarf, delay and weakening of budding and dry branches (apple and pear; Suppl. Figure S1) and shrinkage, reddening or red blistering of leaves (grapevine; Suppl. Figure S1) were collected in different Brazilian regions (Table 1). Initially, 51 samples were evaluated. Later, an additional 10 apple and 13 pear samples were collected in the same orchard where positive samples were found (Viçosa, MG) to estimate the incidence of TFDaV.

2.2. DNA extraction, viral genome amplification, cloning and sequencing

Total nucleic acids were isolated from leafy or woody tissue according to Lodhi et al. (1994). A preliminary evaluation to detect begomoviruses was performed using PCR with degenerate primers for the DNA-A (PAL1v1978 and PAR1C496) and DNA-B (PBL1v2040 and PCRc1) (Rojas et al., 1993; all primers used in this work are described in Suppl. Table S1). Circular DNAs were amplified by RCA as previously described (Inoue-Nagata et al., 2004). RCA products were

screened with different restriction enzymes and the generated fragments were excised from agarose gels, purified, cloned into a pBluescript vector (pKS, Stratagene) and sequenced at Macrogen (Seoul, South Korea).

2.3. Sequence analyses

Viral genomes were assembled using DNA Baser Sequence Assembler v3.x software (www.DnaBaser.com). The resulting contigs were screened for sequence homologies using BLAST algorithms (Altschul et al., 1990). Open reading frames (ORFs) were predicted using ORF Finder (www.ncbi.nlm.nih.gov/projects/gorf/). Conserved domains of predicted ORFs were detected using NCBI's Conserved Domain Database (NCBI) (Marchler-Bauer et al., 2011), TMHMM (Moller et al., 2001) and SMART softwares (Letunic et al., 2012). Importin α-dependent nuclear localization signals and DNA and RNA binding residues in the coat protein (CP) deduced amino acid (aa) sequence were scanned, respectively, using cNLS Mapper (Kosugi et al., 2009) and BindN softwares (Wang and Brown, 2006). Multiple sequences alignments were performed using MEGA 6 (Tamura et al., 2013). Pairwise identity score distribution analyses of the replication-associated protein (Rep) and CP amino acids sequences were performed using Sequence Demarcation Tool (SDT) v. 1.2 (Muhire et al., 2014).

2.4. Phylogenetic and recombination analysis

Phylogenetic relationships of TFDaV isolates with viruses from major gemini-, nano-, circovirus and alphasatellite lineages (Supp. Table S2) were analyzed based on Rep and CP amino acid sequences. For each sequence set, we used MUSCLE v. 3.8.31 (Edgar, 2004) to generate a multiple sequence alignment. Unrooted phylogenetic trees were constructed using Bayesian inference with MrBayes v. 3.2 (Ronquist et al., 2012) (MCMC nruns= 2, nchains= 2, ngen= 10.000.000 generations, burn-in= 2.500 samples, samplefreq= 1000, diagnfreq= 5000, printfreq=

1000) and a mixed amino acid substitution model (prset aamodel= mixed and BLOSUM as best-fit model). Trees were visualized using the FigTree v. 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) and edited using Inkscape v. 0.91.

Recombination events were searched within the TFDaV genome or in the Rep, CP and movement protein (MP) genes, using three distinct data sets: (i) complete genome sequences representing the major lineages of gemini-, nano- and circoviruses; (ii) all Rep sequences deposited in GenBank corresponding to nanoviruses and alphasatellites; and (iii) all CP and MP sequences deposited in GenBank corresponding to mastre-, curto- and begomoviruses (Suppl. Table S2). The analysis was performed with LDhat v. 2.2 (McVean et al., 2002) and RDP4 softwares (Martin et al., 2010).

2.5. PCR- and indirect ELISA-mediated detection of TFDaV

New fresh samples from the original apple, pear and grapevine plants were collected to validate PCR- and indirect ELISA-mediated detection of TFDaV. In addition, 10 apple and 13 pear symptomless samples were collected in the same orchard where infected plants were found (Viçosa, MG) and were evaluated by PCR. PCR with TFDaV-specific CP primers (Suppl. Table S1) was performed using the parameters described by Basso et al. (2010a). PCR amplicons of the expected size were purified, cloned into the pGEM-T Easy vector (Promega) and sequenced commercially (Macrogen). One of the TFDaV-infected pear trees (pear #2) was also evaluated by RCA followed by SpeI restriction analysis. The amplified and SpeI-linearized viral genome was cloned and sequenced.

The TFDaV CP gene was cloned into the pDEST17 vector (Invitrogen) and overexpressed in the E. coli pRARE2 strain. Protein identity was verified by MALDI-TOF/TOF MS analysis and the E. coli-expressed protein was used to raise a polyclonal antiserum in New Zealand rabbits according to Basso et al. (2010b). The sensitivity and specificity of the antiserum was evaluated

according to Almeida and Lima (2001). Woody tissue or petiole samples from apple, pear and grapevine plants were evaluated by indirect ELISA (Clark et al., 1986).

2.6. Infectivity assay

Plants of Nicotiana benthamiana, apple cvs. Gala and Fuji and pear cv. Max Red Bartlett (10 to 12 days after emergence) were biolistically-inoculated with the SpeI-linearized and recircularized TFDaV genome (GenBank access # KJ955447), as described by Aragão et al. (1996). Negative controls were mock-inoculated with the empty pKS vector. Inoculation was targeted to the cotyledons and to the first true leaf of apple and pear seedlings. After the development of the second or third true leaves, theinoculated plants were transplanted to soil and maintained in a greenhouse.

TFDaV systemic infection was evaluated by PCR using specific primers and by RCA followed by SpeI digestion, as described above. Samples consisted of young leaves collected at time 0 (immediately before inoculation) and at 45 and 90 days post-inoculation (dpi). PCR and RCA/SpeI-digested products of the expected sizes were purified and cloned, respectively, into the pGEM-T and pKS vectors, and sequenced. Total DNA from apple #4 (cv. Gala) was diluted 1:25 (v/v) and used as a template for a new round of RCA/SpeI digestion.

2.7. RNA silencing suppression assay and subcellular localization of CP and MP in N. benthamiana

Silencing suppression assays were performed as described previously (Johansen and Carrington, 2001) by infiltration of A. tumefaciens cultures carrying the constructs pK7WG2:GFP (expressing the reporter protein GFP), the pK7WG2:HC-Pro-PVY (expressing the known suppressor HC-Pro, as a positive control) and pK7WG2 with inserts corresponding to the complete coding regions of the CP, MP, and putative viral RNase Z protein. Negative controls were agroinfiltrated with the

pK7WG2:GFP construct only. The GFP fluorescence was monitored under UV and leaves were photographed from 3 to 8 days post-infiltration using a Nikon D1X digital camera.

To verify the subcellular localization of CP and MP, the corresponding ORFs were cloned into pK7WGF2 and transformed into A. tumefaciens C58C1. A. tumefaciens cultures (OD_{600}=0,4) carrying the appropriate DNA constructs were infiltrated into N. benthamiana leaves. The mCherry-tagged nuclear shuttle protein (NSP) from Cabbage leaf curl virus (Fontes et al., 2004) and the RFP-tagged MP of Tomato yellow spot virus (Calegario et al., 2007) were used as positive controls for nuclear and plasmodesmal localization, respectively. Foliar segments of 0,5 cm^2 (70 hours post-infiltration) were analyzed in a LSM 510 META confocal microscope (Carl Zeiss) equipped with an argon-krypton laser as excitation source and a 40x objective, and images were processed with LSM Image Browser 4 (Carl Zeiss). For fluorescence detection, the excitation wavelength and the emission wavelength were 488 nm and 505-530 nm, respectively, for GFP and 543 nm and 596-639 nm, respectively, for mCherry and RFP.

3. RESULTS

3.1. Fruit trees showing virus-like symptoms in the field were infected by a novel ssDNA virus

All 51 samples of different fruit trees showing virus-like symptoms were PCR-negative for the presence of a begomovirus (data not shown). These samples were then evaluated by RCA followed by screening with different restriction enzymes. One apple tree (cv. Eva), one pear tree (cv. Pera d'água, referred to as pear #1) and one grapevine (cv. Cannonao) yielded 3.0 and 3.4 kb fragments following digestion of the RCA product with SpeI (Figure 1A). Both fragments were excised from the gel, purified, cloned and sequenced. The 3.0 kb fragment corresponded to plant DNA (data not shown). The sequence of the 3.4 kb fragments from all three samples displayed

identity with ssDNA viruses, and conserved domains of Rep, CP and MP from ssDNA viruses were identified in the deduced amino acid sequences of the ORFs (Suppl. Table S3). These results were confirmed with newly collected samples from the same three plants.

PCR and indirect ELISA assays further confirmed the presence of TFDaV in the original apple, pear and grapevine plants (Figure 1B; Suppl. Table S4). Out of the 23 additional samples collected (10 apple, 13 pear), one young pear tree (referred to as pear #2) was PCR- and RCA-positive (Figure 1B; Table 1).

3.2. Molecular features of the ssDNA virus detected in apple, pear and grapevine

Forty-five complete viral genomes were sequenced from the four RCA-positive plants: 17 from apple, 18 from pear #1, eight from pear #2, and two from grapevine (GenBank access: KJ955447 to KJ955452 and KR134311 to KR134350). Sequence analysis revealed a genome size of 3.424-3.442 nt, containing the conserved nonanucleotide TAGTATTAC (which is also present in the genomes of nanoviruses, circoviruses and alphasatellites) positioned within an intergenic region containing a short palindromic sequence capable of forming a stem-loop structure (Figure 2A). Three iteron-like, tandem repeat sequences (putative binding sites of the Rep protein) are present in the intergenic region upstream of the stem-loop structure (Figure 2A). Sequence comparisons among TFDaV isolates indicated that the isolates from pear #1 had either a 35-nucleotide deletion in the intergenic region or an 18-nucleotide insertion upstream of the stem-loop structure (data not shown). Nevertheless, sequence identity amongst the 45 clones was greater than 99% (Suppl. Figure S2) and all subsequent analyses were performed with four clones: one from apple (KJ955447), one from pear #1 (KJ955450), one from pear #2 (KJ955451) and one from grapevine (KJ955449).

Predicted ORFs in the viral-sense strand potentially encode the MP, CP, and a putative RNase Z protein (Figure 2B; Suppl. Table S3). Two ORFs predicted in the complementary-sense

strand potentially encode Rep and an RNA-binding protein (Figure 2B; Suppl. Table S3). Amino acid sequence comparisons indicated maximum identities up to 35% with homologous proteins from nanoviruses and alphasatellites (Rep) and from mastreviruses (CP and MP) (Suppl. Table S3). A transmembrane domain was predicted in the aa sequences of the MP, which perfectly overlaps with the ones present in the MPs of nano- and mastreviruses (Suppl. Figure S3). An importin α-dependent nuclear localization signal and DNA-binding properties were predicted in the deduced aa sequence of the CP (Suppl. Figure S4). Analysis of the deduced aa sequence of the putative Rep protein revealed the presence of conserved motifs such as Motif I (virus-specific dsDNA binding), Motif II (metal-binding site), Motif III (catalytic site for DNA cleavage), Walker A and B boxes (NTP binding) and DNA helicase (Suppl. Figure S5). These motifs are similar to the ones present in the Rep proteins from circo-, nano- and geminiviruses. In addition, the sequences LEEGE and KEETRVSGPYTYG could functionally represent the LxCxE and KEEALQIIREKIP motifs found in the geminivirus Tomato golden mosaic virus (TGMV) Rep, and which are involved in binding to the plant retinoblastoma protein (pRb) (Suppl. Figure S5).

Three putative transcription regulatory sequences were predicted in the TFDaV genome, one located upstream of the Rep and stem-loop structure, one positioned upstream of the CP and MP, and another upstream of the RNase Z protein (Figure 2B, blue and green markers). Three polyadenylation signals were also predicted downstream of the CP, RNase Z and Rep ORFs (Figure 2B, red arrows).

3.3. Recombination and phylogenetic analyses

No recombination events were detected in the TFDaV sequence, independently of the data set used for the analysis (data not shown).

The phylogenetic trees based on the deduced aa sequences of the Rep and CP indicated that TFDaV isolates are located in a monophyletic clade, basal to all other ssDNA viruses and subviral

agents (Figure 3A and B). A tree based only on the conserved motifs I, II and III and the Walker A and B domains of the Rep protein displayed the same topology (data not shown).

The absence of DNA sequences with significant identities (Rep and CP pairwise identity score distribution of no more than 35%; Suppl. Figure S6) emphasize the highly divergent nature of TFDaV.

3.4. Infectivity assay

N. benthamiana plants were PCR- and RCA-negative for the presence of TFDaV at either 20 or 45 days post-inoculation. Out of 23 inoculated apple and pear seedlings, 12 (eight apple cv. Gala, three apple cv. Fuji and one pear) were PCR-positive for TFDaV at 45 dpi (Suppl. Figure S7), with one additional apple cv. Gala being PCR-positive only at 90 dpi (Suppl. Figure S7). Plants biolistically inoculated with the empty vector were all PCR-negative (Suppl. Figure S7). These plants were also evaluated by RCA at 45, 90 and 180 dpi but there was no amplification of the expected fragment size, with the exception of one apple cv. Gala (plant #4) at 180 dpi (data not shown). This 3.4 kb amplified fragment was cloned and sequenced, confirming its identity as TFDaV.

Foliar symptoms were not observed in the inoculated plants. However, growth reduction was evident at both 90 and 180 dpi (Figure 4). Attempts to visualize viral particles by transmission electron microscopy from leaf tissue or petioles at 180 dpi were unsuccessful.

3.5. Local RNA silencing suppression and cellular co-localization in N. benthamiana

The putative RNase Z, CP and MP showed no local RNA silencing suppression activity in N. benthamiana leaves (Suppl. Figure S8). The CP localized to the nuclei with a particularly strong signal in the nucleolus (Suppl. Figure S9), confirming the functionality of its nuclear localization signal and similarities with others geminivirus CPs. The MP localized to the plasma membrane,

possibly in the plasmodesmata (Suppl. Figure S9), confirming the functionality of the transmembrane domain.

4. DISCUSSION

Apple, pear and grapevines are clonally propagated by different forms of grafting. Unless quarantine measures are firmly established, the exchange of propagation materials between countries will virtually guarantee the spread of plant viruses and other graft-transmissible agents such as viroids and phytoplasmas. Indeed, plants displaying virus-like symptoms are frequently observed in Brazilian vineyards and orchards (Basso et al., 2014). However, due to mixed infections, variations in climatic and agronomic conditions, and intrinsic characteristics of these plants, it is difficult or even impossible to establish an association between field symptoms and a specific viral or subviral agent.

Until recently, viral detection strategies in plants focused on RNA viruses, for at least two reasons: (i) RNA viruses comprise the vast majority of plant viruses and include the most economically important ones; (ii) available methodologies had a strong sequence bias (e.g., use of degenerate primers in PCR-based assays) and therefore restricted the identification of divergent species. The development of sequence-unbiased methods such as RCA, and more recently the widespread availability of deep sequencing strategies, has unlocked the vault of DNA viruses with small genomes (mostly ssDNA viruses), indicating their presence in a massive scale in a variety of ecosystems (Labonte and Suttle, 2013; Rosario et al., 2012).

New ssDNA viruses have been discovered in various aquatic environments as well as in insects, plants and diatoms (Bernardo et al., 2013; Dayaram et al., 2014; Labonte and Suttle, 2013; Rosario et al., 2009; Tomaru et al., 2013) uncovering the richness and diversity of this group of viruses. Recent reports of divergent geminiviruses in citrus and grapevine (Al Rwahnih et al., 2013; Krenz et al., 2012; Loconsole et al., 2012; Poojari et al., 2013) indicated that perennial

woody plants can also support infection by DNA viruses. Nevertheless, apple and pear had not yet been reported as hosts of ssDNA viruses.

TFDaV displays low sequence identities and is only distantly phylogenetically related to other circular ssDNA viruses, suggesting that it could be classified in a new genus or family. Circo-, nano- and geminiviruses present several traits in common, strongly suggesting a common origin (Gibbs and Weiller, 1999; Niagro et al., 1998). Metagenomic studies have revealed a wide diversity of ssDNA viruses in different biomes, many of which encode Rep proteins similar to those of gemini-, nano- and (most commonly) circoviruses (Labonte and Suttle, 2013; Rosario et al., 2012). However, the origin(s) and evolutionary relationships amongst these three families of ssDNA viruses remain unknown. Interestingly, TFDaV shares common features with all three families (circoviruses: a monopartite genome, conserved domains in the Rep protein; circo- and nanoviruses: the conserved nonanucleotide at the origin of replication; geminiviruses: conserved domains in the MP and CP). Along with the absence of recombination events, which discards the possibility of TFDaV being an inter-family recombinant, these conserved features suggest that it could be a member of an ancestor group of circular ssDNA viruses.

The Rep sequence and structural similarities suggest that TFDaV replicates through a rolling-circle mechanism in a similar manner to circo-, nano- and geminiviruses. Interestingly, the presence of Rb-binding motifs suggests that, like the Rep proteins of geminiviruses, the TFDaV Rep could be involved in cell cycle regulation (Gutierrez, 2000).

The presence of an importin α-dependent NLS in the CP and its nuclear localization suggest that it can function also as a nuclear shuttle protein, facilitating the transport of viral DNA into and out of the nucleus (Rojas et al., 2001; Fontes et al., 2004). The putative MP contains a transmembrane domain commonly found in viral MPs and is localized in the plasma membrane, sugesting that it may act in viral cell-to-cell movement. None of the viral proteins tested showed local RNA silencing suppressor activity in N. benthamiana. However, since TFDaV infects

woody, perennial plants, it is possible that its supressor protein(s) are unable to suppress RNA silencing in herbaceous plants.

Infectivity assays confirmed that TFDaV is capable of sistemically infecting apple and pear, causing growth reduction. The absence of foliar symptoms could be due to the relatively short (for a perennial, woody host) period of observation of 180 dpi. Latent periods of two or three years are not uncommon in such hosts (Loconsole et al., 2012; Al Rwahnih et al., 2013; Basso et al., 2014).

Indirect ELISA, RCA and PCR-based assays were effective in detecting TFDaV in field plants. Nevertheless, viral incidence seems to be small (although our survey was limited). Additional studies are required to properly assess the incidence and the losses caused by TFDaV, as well as the possibility that it may have additional hosts.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Science and Technology in Plant-Pest Interactions, CAPES (doctoral fellowship to MFB), CNPq (grant 483607/2013-4 to FMZ) and FAPEMIG (grant APQ-00949-09 to FMZ).

REFERENCES

Al Rwahnih, M., Daubert, S., Golino, D. and Rowhani, A., 2009. Deep sequencing analysis of RNAs from a grapevine showing syrah decline symptoms reveals a multiple virus infection that includes a novel virus. Virology 387, 395-401.

Al Rwahnih, M., Dave, A., Anderson, M.M., Rowhani, A., Uyemoto, J.K. and Sudarshana, M.R., 2013. Association of a DNA virus with grapevines affected by red blotch disease in California. Phytopathology 103, 1069-1076.

Almeida, A.M.R. and Lima, J.A.A., 2001. Técnicas sorológicas aplicadas à fitovirologia. In: A.M.R. Almeida and J.A.A. Lima (Eds), Princípios e técnicas de diagnose aplicados em em fitovirologia, pp. 33-62. Sociedade Brasileira de Fitopatologia, Brasília, DF.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.

Aragão, F.J.L., Barros, L.M.G., Brasileiro, A.C.M., Ribeiro, S.G., Smith, F.D., Sanford, J.C., Faria, J.C. and Rech, E.L., 1996. Inheritance of foreign genes in transgenic bean (Phaseolus

vulgaris L.) co-transformed via particle bombardment. Theoretical and Applied Genetics 93, 142-150.

Basso, M.F., Fajardo, T.V.M., Eiras, M., Ayub, R.A. and Nickel, O., 2010a. Molecular detection and identification of virus associated with symptomatic and symptomless grapevines. Ciência Rural 40, 2249-2255.

Basso, M.F., Fajardo, T.V.M., Eiras, M., Ayub, R.A. and Nickel, O., 2010b. Production of polyclonal antiserum using recombinant coat protein of Rupestris stem pitting-associated virus. Ciência Rural 40, 2385-2388.

Basso, M.F., Fajardo, T.V.M., Pio-Ribeiro, G., Eiras, M. and Zerbini, F.M., 2014. Avanços e perspectivas no estudo das doenças virais e subvirais em videira com ênfase na realidade brasileira. Revisão Anual de Patologia de Plantas 22, 160-207.

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

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

Clark, M.F., Lister, R.M. and Bar-Joseph, M., 1986. ELISA techniques. Methods in Enzymology 118, 742-766.

Dayaram, A., Galatowitsch, M., Harding, J.S., Arguello-Astorga, G.R. and Varsani, A., 2014. Novel circular DNA viruses identified in Procordulia grayi and Xanthocnemis zealandica larvae using metagenomic approaches. Infection, Genetics and Evolution 22, 134-41.

Edgar, R.C., 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

Fontes, E.P.B., Santos, A.A., Luz, D.F., Waclawovsky, A.J. and Chory, J., 2004. The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. Genes & Development 18, 2545-2556.

Gibbs, M.J. and Weiller, G.F., 1999. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. Proceedings of the National Academy of Sciences, USA 96, 8022-8027.

Gutierrez, C., 2000. DNA replication and cell cycle in plants: Learning from geminiviruses. EMBO J 19, 792-799.

Hadidi, A., Barba, M., Candresse, T. and Jelkmann, W., eds., 2011. Virus and Virus-Like Diseases of Pome and Stone Fruits. APS Press, St. Paul, MN, USA.

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

Johansen, L.K. and Carrington, J.C., 2001. Silencing on the spot. Induction and suppression of RNA silencing in the Agrobacterium-mediated transient expression system. Plant Physiology 126, 930-8.

Kosugi, S., Hasebe, M., Tomita, M. and Yanagawa, H., 2009. Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proceedings of the National Academy of Sciences, USA 106, 10171-10176.

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

Labonte, J.M. and Suttle, C.A., 2013. Previously unknown and highly divergent ssDNA viruses populate the oceans. ISME Journal 7, 2169-2177.

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

Letunic, I., Doerks, T. and Bork, P., 2012. SMART 7: Recent updates to the protein domain annotation resource. Nucleic Acids Research 40, D302-D305.

Loconsole, G., Saldarelli, P., Doddapaneni, H., Savino, V., Martelli, G.P. and Saponari, M., 2012. Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. Virology 432, 162-72.

Lodhi, M.A., Ye, G.-N., Weeden, N.F. and Reisch, B.I., 1994. A simple and efficient method for DNA extraction from grapevine cultivars, Vitis species and Ampelopsis. Plant Molecular Biology Reporter 12, 6-13.

Marchler-Bauer, A., Lu, S.N., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z.X., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D.C., Zhang, N.G., Zheng, C.J. and Bryant, S.H., 2011. CDD: A Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research 39, D225-D229.

Martelli, G.P., 2012. Grapevine virology highlights 2010-12. Extended Abstracts, 17th Meeting of International Council of Viruses and Virus Diseases of Grapevine, 13-31.

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

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

McVean, G., Awadalla, P. and Fearnhead, P., 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160, 1231-1241.

Moller, S., Croning, M.D.R. and Apweiler, R., 2001. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 17, 646-653.

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

Niagro, F.D., Forsthoefel, A.N., Lawther, R.P., Kamalanathan, L., Ritchie, B.W., Latimer, K.S. and Lukert, P.D., 1998. Beak and feather disease virus and porcine circovirus genomes: intermediates between the geminiviruses and plant circoviruses. Archives of Virology 143, 1723-44.

Poojari, S., Alabi, O.J., Fofanov, V.Y. and Naidu, R.A., 2013. A leafhopper-transmissible DNA virus with novel evolutionary lineage in the family geminiviridae implicated in grapevine redleaf disease by next-generation sequencing. PLOS One 8(6), e64194.

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

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

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539-542.

Rosario, K., Duffy, S. and Breitbart, M., 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. Journal of General Virology 90, 2418-2424.

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

Rowhani, A., Uyemoto, J.K., Golino, D.A. and Martelli, G.P., 2005. Pathogen testing and certification of Vitis and Prunus species. Annual Review of Phytopathology 43, 261-278.

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

Tomaru, Y., Toyoda, K., Suzuki, H., Nagumo, T., Kimura, K. and Takao, Y., 2013. New single-stranded DNA virus with a unique genomic structure that infects marine diatom Chaetoceros setoensis. Scientific Reports 3, 3337.

Wang, L. and Brown, S.J., 2006. BindN: A web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. Nucleic Acids Research 34, W243-W248.

TABLE 1. Fruit tree samples displaying virus-like symptoms collected at different Brazilian regions and evaluated for the presence of ssDNA viruses.

Samples	Collection date	Sampled plants/RCA-positive plants	Location
Nectarine cv. Centenária	Oct-2012	2/0	UFV[3], Viçosa, MG
Peach cv. Chiripá	Oct-2012	2/0	UFV, Viçosa, MG
Acerola (cv. n.d.)[1]	Nov-2012	1/0	UFV, Viçosa, MG
Apple cv. Eva	Nov-2012	**3/1**	UFV, Viçosa, MG
Lemon cv. Taiti	Nov-2012	1/0	UFV, Viçosa, MG
Orange cv. Navelate	Nov-2012	1/0	UFV, Viçosa, MG
Pear cv. Pera d´água	Nov-2012	**3/1**	UFV, Viçosa, MG
Jabuticaba (cv. n.d.)	Dec-2012	1/0	UFV, Viçosa, MG
Mango cv. Tommy Atkins	Dec-2012	1/0	UFV, Viçosa, MG
Passion fruit (cv. n.d.)	Dec-2012	1/0	UFV, Viçosa, MG
Apple cv. Fuji	Jan-2013	2/0	Linha Aparecida, Videira, SC
Grapevine Muscadine (V. rotundifolia)	Jan-2013	1/0	UFV, Viçosa, MG
Grapevine rootstock VR-043-43	Jan-2013	2/0	UFV, Viçosa, MG
Papaya (cv. n.d.)	Jan-2013	1/0	UFV, Viçosa, MG
Apple (cv. n.d.)	Feb-2013	3/0	UFV, Viçosa, MG
Guava (cv. n.d.)	Feb-2013	1/0	UFV, Viçosa, MG
Grapevine cv. Alicante Bouschet (accession 1)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Alicante Bouschet (acc. 2)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Ancellotta[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Cabernet Franc (acc. 1)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Cabernet Franc (acc. 2)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Cabernet Franc (acc .3)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Cabernet Sauvignon[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Cannonao[2]	Mar-2013	**1/1**	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Chardonnay (acc. 1)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Chardonnay (acc. 2)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Carmenère[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Don Marino[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Garganega (acc. 1)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Garganega (acc. 2)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Italia[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Malvasia di Candia Aromatica[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Merlot[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Michele Palieri[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Primitivo[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Prosecco (acc. 1)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Prosecco (acc. 2)[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Regina[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Sangiovese[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Grapevine cv. Victoria[2]	Mar-2013	1/0	Embrapa Uva e Vinho, Bento Gonçalves, RS
Avocado cv. Margarida	Mar-2013	1/0	Grupo Tsuge, São Gotardo, MG
Pear cv. Pera d´água	Apr-2013	**13/1**	UFV, Viçosa, MG
Apple cv. Eva	Apr-2013	7/0	UFV, Viçosa, MG
Apple (cv. n.d.)	Apr-2013	3/0	UFV, Viçosa, MG

[1]cv. n.d.: cultivar not determined
[2]Germplasm collection maintained in quarantine greenhouse
[3]UFV: Universidade Federal de Viçosa; MG: state of Minas Gerais; RS: state of Rio Grande do Sul; SC: state of Santa Catarina.

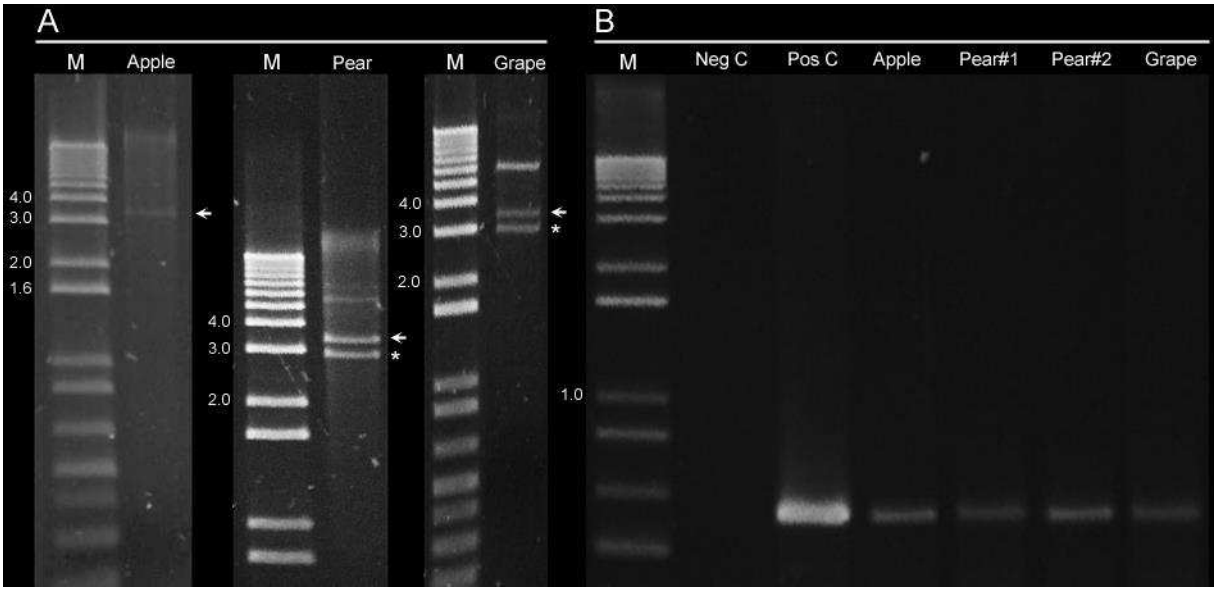


Figure 1. A. Agarose gel electrophoresis of SpeI-digested RCA products amplified from apple, pear and grapevine samples. Arrows indicate the cloned viral genome (3.4 kb) and asterisks indicate the fragment of mitochondrial DNA from the host plants. M, size marker (1 kb plus DNA ladder, in kilobases). **B.** Agarose gel electrophoresis of PCR products (736 bp) of viral detection in apple, pear and grapevine samples. M, size marker (1 kb plus DNA ladder, in kilobases). Neg C, negative control (PCR from healthy apple plant). Pos C, positive control (PCR from recombinant plasmid containing the viral genome).

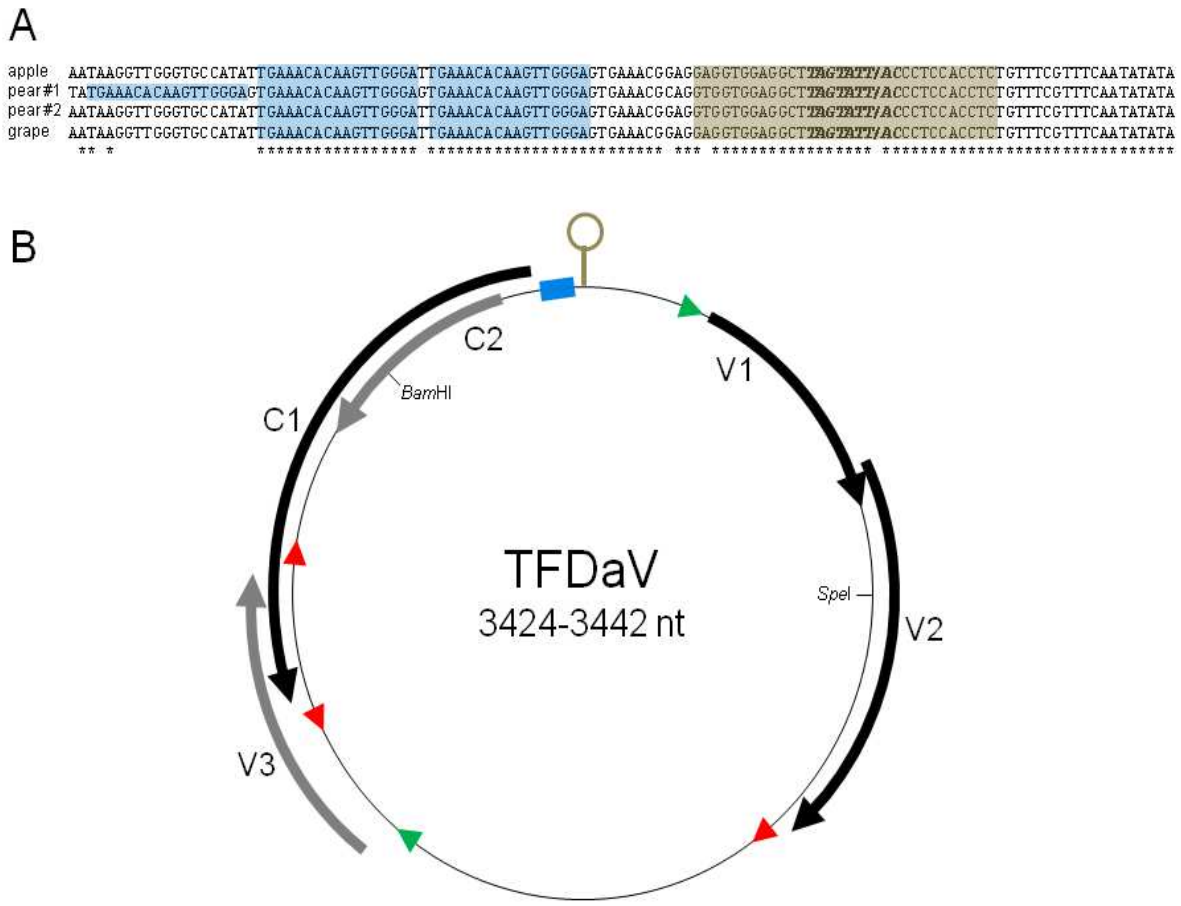


Figure 2. Genomic features of Temperate fruit decay-associated virus (TFDaV). **A.** Nucleotide sequence alignment of the region comprising the putative origin of replication. The sequence forming the stem-loop structure containing the conserved nonanucleotide (in bold italics) is higlighted in grey. Iteron-like sequences which may function as Rep binding sites to promote rolling-circle replication are highlighted in blue. **B.** Genomic organization of TFDaV indicating the five predicted ORFs (black and grey arrows), the stem-loop structure, the iteron-like sequences (blue rectangle), promoter regions (green arrowheads) and polyadenylation signals (red arrowheads). V1, movement protein (MP, nt 220-456); V2, coat protein (CP, nt 589-1303); V3, putative RNase Z (nt 2228-2663); C1, replication-associated protein (Rep, nt 3374-2429); C2, putative RNA-binding protein (nt 3310-2920).

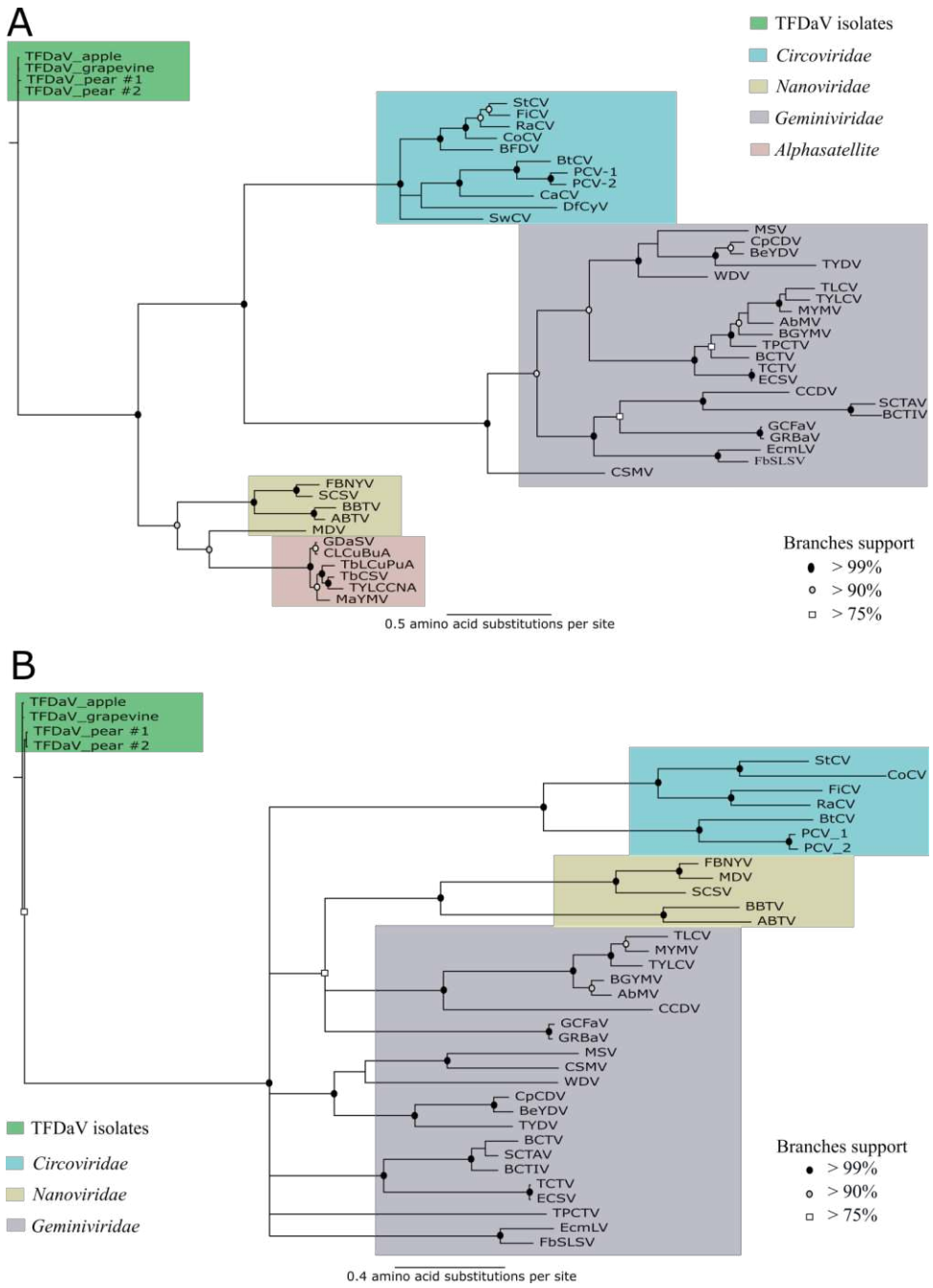


Figure 3. Phylogenetic reconstruction trees describing the evolutionary relationships between TFDaV isolates (apple, pear #1, pear #2 and grapevine), circo-, nano-, geminiviruses and alphasatellites replication associated protein (**A**) and coat protein (**B**) amino acid sequences. Phylogenetic trees were constructed using Bayesian inference with MrBayes v.3.2 and BLOSUM as best-fit model. Branches with a filled dot have >99% posterior probability support, empty dot have >90% posterior probability support, whereas those with a rectangle empty have >75% posterior probability support for those nodes. Branch length is proportional to number of amino acid substitutions per site. GenBank accession number is available in Table S2.

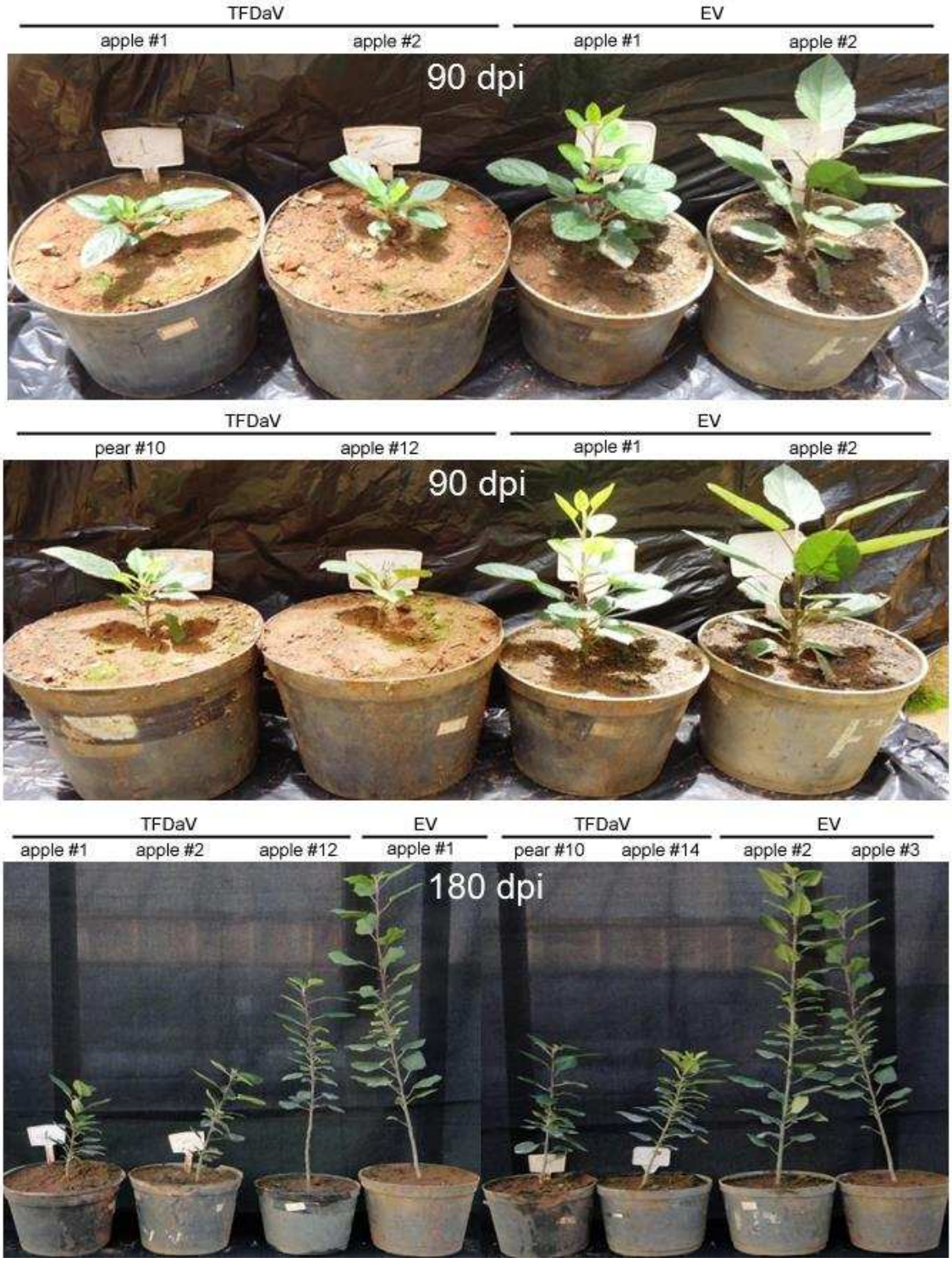


Figure 4. Results of infectivity assay with TFDaV in apple and pear plants. Growth reduction observed in apple and pear plants following biolistic inoculation with cloned TFDaV, at 90 and 180 days post- inoculation (dpi). TFDaV, apple and pear plants inoculated with cloned TFDaV. EV, apple plants inoculated with the empty vector (negative control). Results of PCR-based viral detection in the inoculated plants are presented in Suppl. Figure S7.

Supplementary Table S1. Primers used for TFDaV detection and cloning.

Name	Sequence (5´- 3´)	Tm (ºC)	Amplicon size (bp)	Amplified region
APG (F)	GAGGAGACGCCCTCTGAAGC	54	601	partial CP[1]
APG (R)	CATAACCCAGAGCAATGCG			
APG_RNaseZ (F)	GATCGGATCCATATGCTAAGCGACTCG	54	427	RNase Z
APG_RNaseZ (R)	GATCCTCGAGTCAGGTAAAGATTATG			
APG_CP (F)	GATCGGATCCATATGGCGATTCGGTATCC	53	736	complete CP
APG_CP (R)	GATCCTCGAGTTACAAACACTTGAAG			
APG_MP (F)	GATCGTCGACATGGATCACCTAG	54	492	MP
APG_MP (R)	GATCGAGCTCTCATACTTTCGCCTTGC			

[1]CP, coat protein; MP, movement protein

Supplementary Table S2. Viral sequences used in phylogenetic, pairwise identity score distribution and recombination analyses.

Family	Genus	Species	Acronym	Access number
Circoviridae	Circovirus	Starling circovirus	StCV	NC_008033
		Columbid circovirus	CoCV	NC_002361
		Finch circovirus	FiCV	NC_008522
		Raven circovirus	RaCV	NC_008375
		Bat circovirus	BtCV	NC_021206
		Porcine circovirus 1	PCV-1	NC_013774
		Porcine circovirus 2	PCV-2	NC_005148
		Beak and feather disease virus	BFDV	NC_001944
		Swan circovirus	SwCV	ABU48445, ABU48446
		Canine circovirus	CaCV	NC_020904
	Gyrovirus	Human gyrovirus	HGyV	NC_015630
		Chicken anemia virus	CAV	NC_001427
		Gyrovirus GyV3	GyV3	NC_017091
	Cyclovirus[1]	Dragonfly cyclovirus	DfCyV	HQ638060
Nanoviridae	Nanovirus	Faba bean necrotic yellows virus	FBNYV	NC_003567, NC_003566, NC_003565, NC_003564, NC_003563, NC_003562, NC_003561, NC_003560, NC_003559, NC_003558
		Milk vetch dwarf virus	MDV	NC_003646, NC_003647, NC_003645, NC_003644, NC_003643, NC_003642, NC_003641, NC_003640, NC_003639, NC_003638, NC_003648
		Subterranean clover stunt virus	SCSV	NC_003819, NC_003818, NC_003817, NC_003816, NC_003815, NC_003814, NC_003813, NC_003812
	Babuvirus	Banana bunchy top virus	BBTV	NC_003479, NC_003476, NC_003477, NC_003475, NC_003474, NC_003473
		Abaca bunchy top virus	ABTV	NC_010319, NC_010318, NC_010317, NC_010316, NC_010315, NC_010314
Geminiviridae	Mastrevirus	Maize streak Reunion virus	MSV	NC_017917
		Wheat dwarf virus	WDV	NC_003326
		Chickpea chlorotic dwarf virus	CpCDV	NC_011058
		Chloris striate mosaic virus	CSMV	NC_001466
		Tobacco yellow dwarf virus	TYDV	NC_003822
		Bean yellow dwarf virus	BeYDV	NC_003493
	Begomovirus	Tomato leaf curl Iran virus	TLCV	NC_005842
		Bean golden yellow mosaic virus	BGYMV	NC_004042
		Abutilon mosaic virus	AbMV	NC_001928
		Malvastrum yellow mosaic virus	MYMV	NC_008559
		Tomato yellow leaf curl virus	TYLCV	NC_004005
	Curtovirus	Beet curly top virus	BCTV	NC_001412
	Becurtovirus	Spinach curly top Arizona virus	SCTAV	HQ443515
		Beet curly top Iran virus	BCTIV	EU273818
	Turncurtovirus	Turnip curly top virus	TCTV	GU456685
	Eragrovirus	Eragrostis curvula streak virus	ECSV	FJ665631
	Topocuvirus	Tomato pseudo-curly top virus	TPCTV	NC_003825
	unassigned	Grapevine cabernet franc-associated virus	GCFaV	JX559642
		Grapevine red blotch-associated virus	GRBaV	KC896623
		Citrus chlorotic dwarf associated virus	CCDV	NC_018151
		Euphorbia caput-medusae latent virus	EcmLV	HF921459
		French bean severe leaf curl virus	FbSLSV	JX094280
Alphasatellite		Gossypium darwinii symptomless alphasatellite	GDaSA	EU384623
		Cotton leaf curl Burewala alphasatellite	CLCuBuA	FN658730
		Tobacco leaf curl PUSA alphasatellite	TbLCuPuA	HQ180392
		Tobacco curly shoot alphasatellite	TbCSA	FN678903
		Tomato yellow leaf curl China alphasatellite	TYLCCNA	AM749494
		Malvastrum yellow mosaic alphasatellite	MaYMA	NC_008561
		TFDaV apple	TFDaV	KJ955447
		TFDaV pear #1	TFDaV	KJ955450
		TFDaV grapevine	TFDaV	KJ955449
		TFDaV pear #2	TFDaV	KJ955451

[1] Genus Cyclovirus was recently proposed in the family Circoviridae.

Supplementary Table S3. Results of BLASTp/x analysis of the proteins encoded by TFDaV.

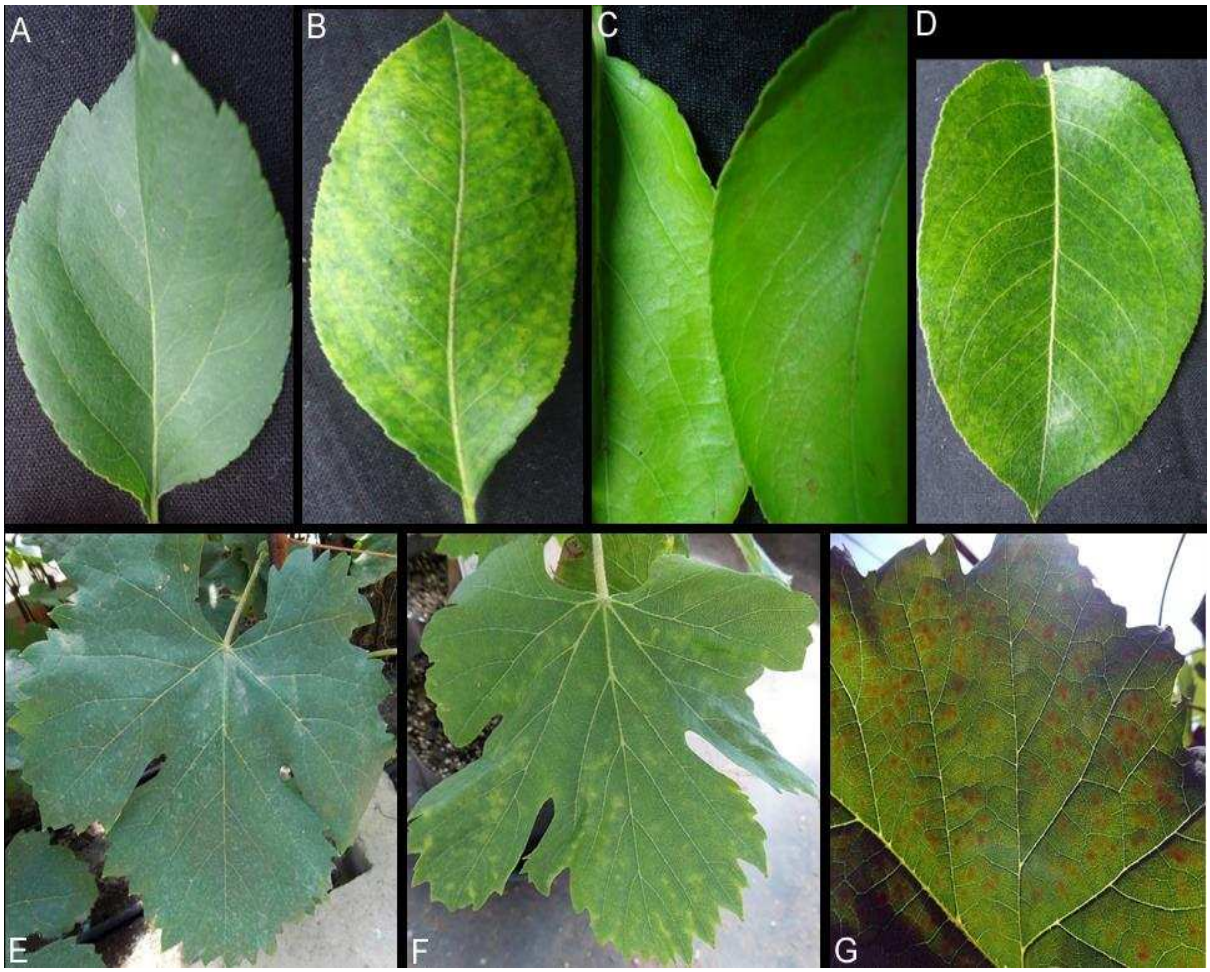
ORF	Homologous protein (GenBank accession number and genus)	Coverage[1] (%)	Identity[1] (%)
V1 or MP-Like	movement protein - Wheat dwarf India virus (YP_006273068, Mastrevirus)	91/91	36/36
	movement protein - Dragonfly-associated mastrevirus (YP_007004036, Mastrevirus)	82/82	39/39
	movement protein - Chickpea chlorotic dwarf virus (CBX20817, Mastrevirus)	84/84	44/44
	movement protein - Sweetpotato symptomless mastrevirus 1 (ACN56753, Mastrevirus)	66/66	44/44
	putative movement protein - Sugar beet mastrevirus SK-11 (ABV82710, Mastrevirus)	44/44	60/60
	V2 protein - Chickpea redleaf virus (YP_004046661, Mastrevirus)	61/61	40/40
	V1 protein - Bean yellow dwarf virus (NP_612219, Mastrevirus)	43/43	62/62
	movement protein - Wheat dwarf virus (AEZ00921, Mastrevirus)	75/75	38/38
	movement protein - Maize streak Reunion virus (AFJ15132, Mastrevirus)	85/85	32/32
	movement protein - Sporobolus striate mosaic virus 2 (YP_006666529, Mastrevirus)	75/75	35/35
V2 or CP-Like	coat protein - Chickpea chlorosis Australia virus (AFD63000, Mastrevirus)	81/74	25/30
	coat protein - Tobacco yellow dwarf virus-A (AFD63088, Mastrevirus)	67/74	28/30
	coat protein - Digitaria ciliaris striate mosaic virus (YP_006666534, Mastrevirus)	81/81	26/26
	coat protein - Beet mild curly top virus (ABW74010, Curtovirus)	94/94	26/26
	coat protein - Tomato chino La Paz virus (ABC74540, Begomovirus)	93/85	26/45
	coat protein - Tomato severe leaf curl virus (AEP68824, Begomovirus)	93/85	26/45
	coat protein - Paspalum dilatatum striate mosaic virus (AFN80656, Mastrevirus)	70/70	27/27
V3 or RNaseZ-Like	ribonuclease Z protein - Tannerella sp. (ZP_09338067)	53/53	30/30
C1 or Rep-Like	Rep protein - Faba bean necrotic stunt virus (ACX50511, Nanovirus)	97/97	33/33
	Rep protein - Malvastrum yellow mosaic alphasatellite (CAJ85969, Begomovirus-associated alphasatellite)	99/99	35/35
	Rep protein - Milk vetch dwarf virus (NP_619769, Nanovirus)	95/95	34/34
	Rep protein - Subterranean clover stunt virus (NP_620694, Nanovirus)	97/97	34/33
	Rep protein - Gossypium darwinii symptomless alphasatellite (ACC78692, Begomovirus-associated alphasatellite)	98/98	33/33
	Rep protein - Cotton leaf curl Burewala alphasatellite (CBJ19312, Begomovirus-associated alphasatellite)	98/98	33/33
	Rep protein - Banana bunchy top virus (ACA34591, Babuvirus)	93/93	32/32
C2	valine-tRNA ligase (ZP_06752514.1, Parascardovia denticolens)	85/85	67/67
	hypothetical protein PTSG (EGD82448.1, Salpingoeca sp.)	38/38	38/38
	kinase subdomain-containing protein (EGD75615.1, Salpingoeca sp.)	38/38	38/38
	zinc finger protein 385C-like (XP_001508885.2, Ornithorhynchus anatinus)	59/59	32/32

[1]Nucleotide sequence/deduced amino acid sequence.

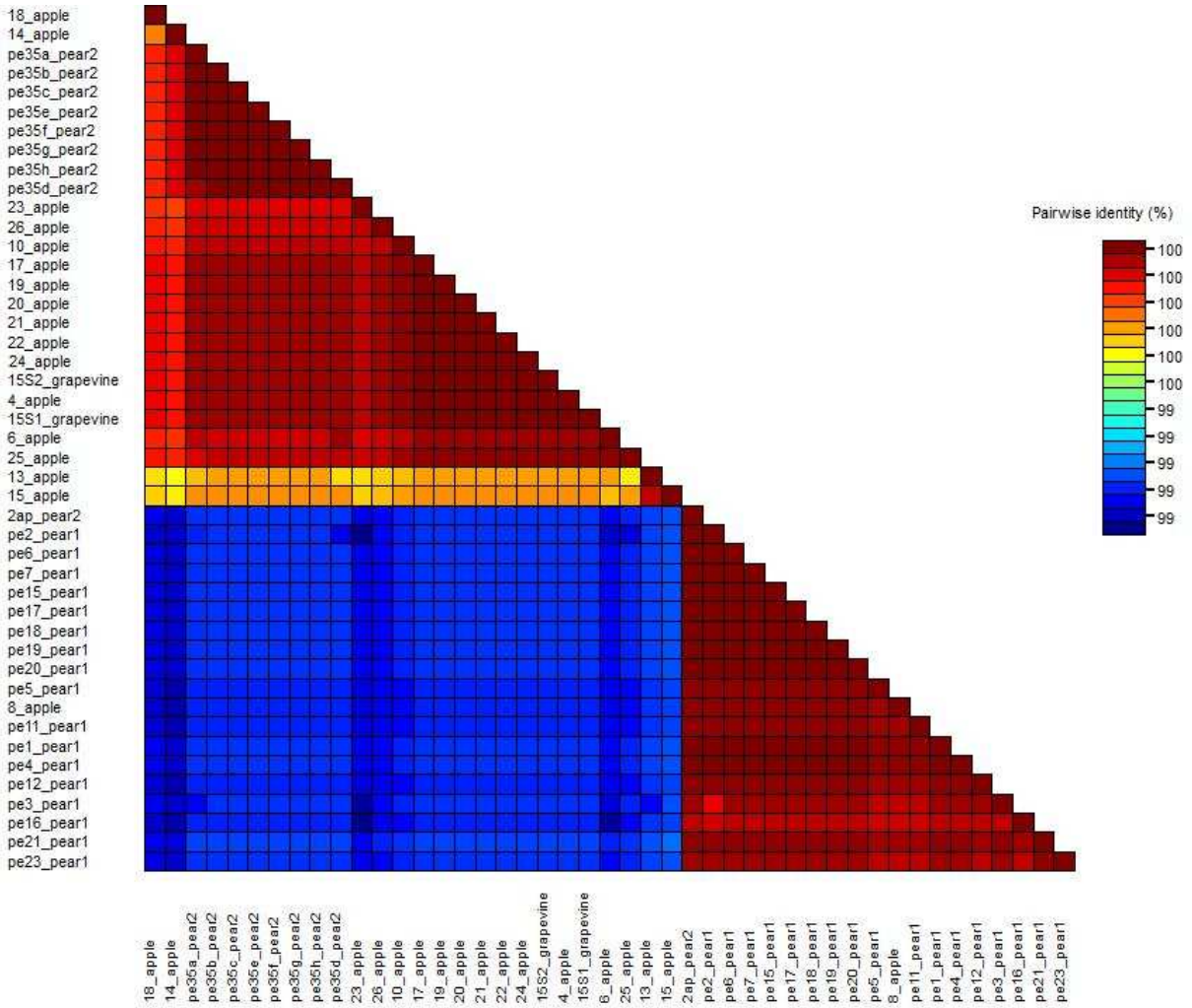
Supplementary Table S4. Results of indirect ELISA using the polyclonal antiserum against the TFDaV coat protein.

Samples	**A_{405}**	**Result**
Extraction buffer	0.302	Negative
E. coli-expressed CP (0,52 ng)	1.141	Positive
E. coli-expressed CP (0,128 ng)	1.389	Positive
E. coli -expressed CP (0,235 ng)	1.456	Positive
Healthy grapevine cv. Paulsen 1103 (wood tissue)	0.287	Negative
Apple cv. Eva (wood tissue)	0,725	Positive
Pear cv. Pera d'água (plant #1, wood tissue)	0,625	Positive
Pear cv. Pera d'água (plant #2, wood tissue)	0,352	Negative
Grapevine cv. Cannonao (wood tissue)	0,335	Negative
Grapevine cv. Cannonao (petiole tissue)	0,412	Negative
Healthy Nicotiana benthamiana (leaf tissue)	0.315	Negative
ToYSV[1]-infected N. benthamiana (leaf tissue)	0,298	Negative
Healthy Sida sp. virus-free (leaf tissue)	0,275	Negative
SiYMV-infected Sida sp. (leaf tissue)	0,284	Negative
Healthy Euphorbia heterophylla (leaf tissue)	0,301	Negative
EuYMV-infected E. heterophylla (leaf tissue)	0,321	Negative

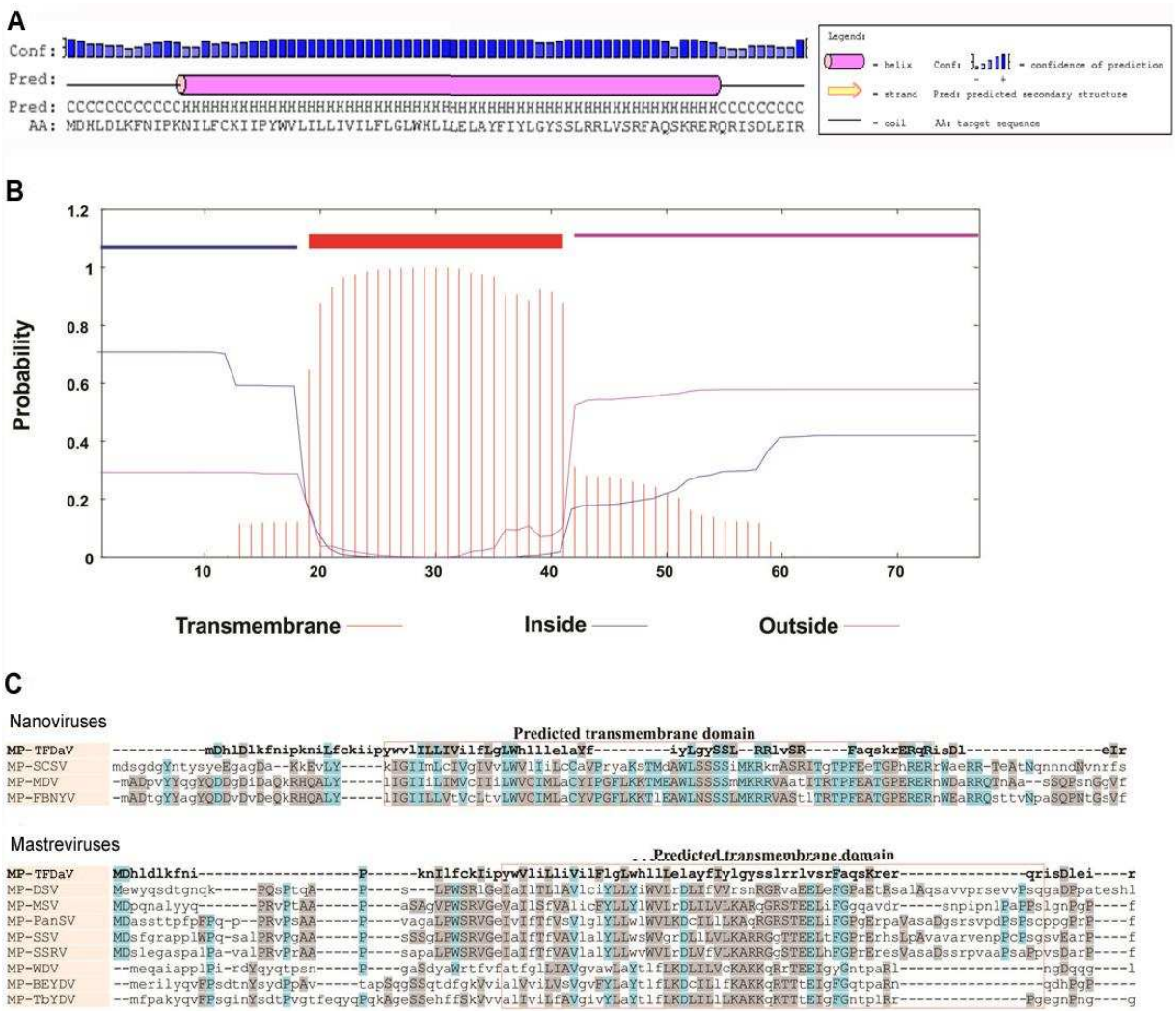
[1]ToYSV, Tomato yellow spot virus; SiYMV, Sida yellow mosaic virus; EuYMV, Euphorbia yellow mosaic virus. The three viruses are classified in the genus Begomovirus.



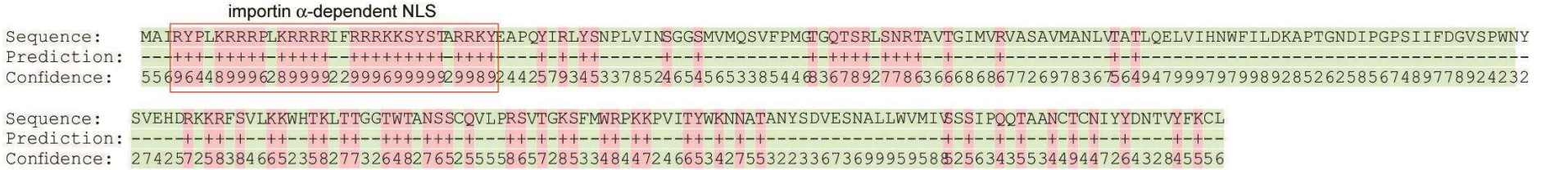
Supplementary Figure S1. Virus-like symptoms in leaves of apple tree, pear tree #1 and grapevine, later shown to be infected with TFDaV. **A, B.** Asymptomatic and symptomatic apple leaves, respectively; **C, D.** Asymptomatic and symptomatic pear leaves, respectively; **E.** Asymptomatic grapevine leaf; **F, G.** Symptomatic grapevine leaves at the beginning and at the end of the growing season, respectively.



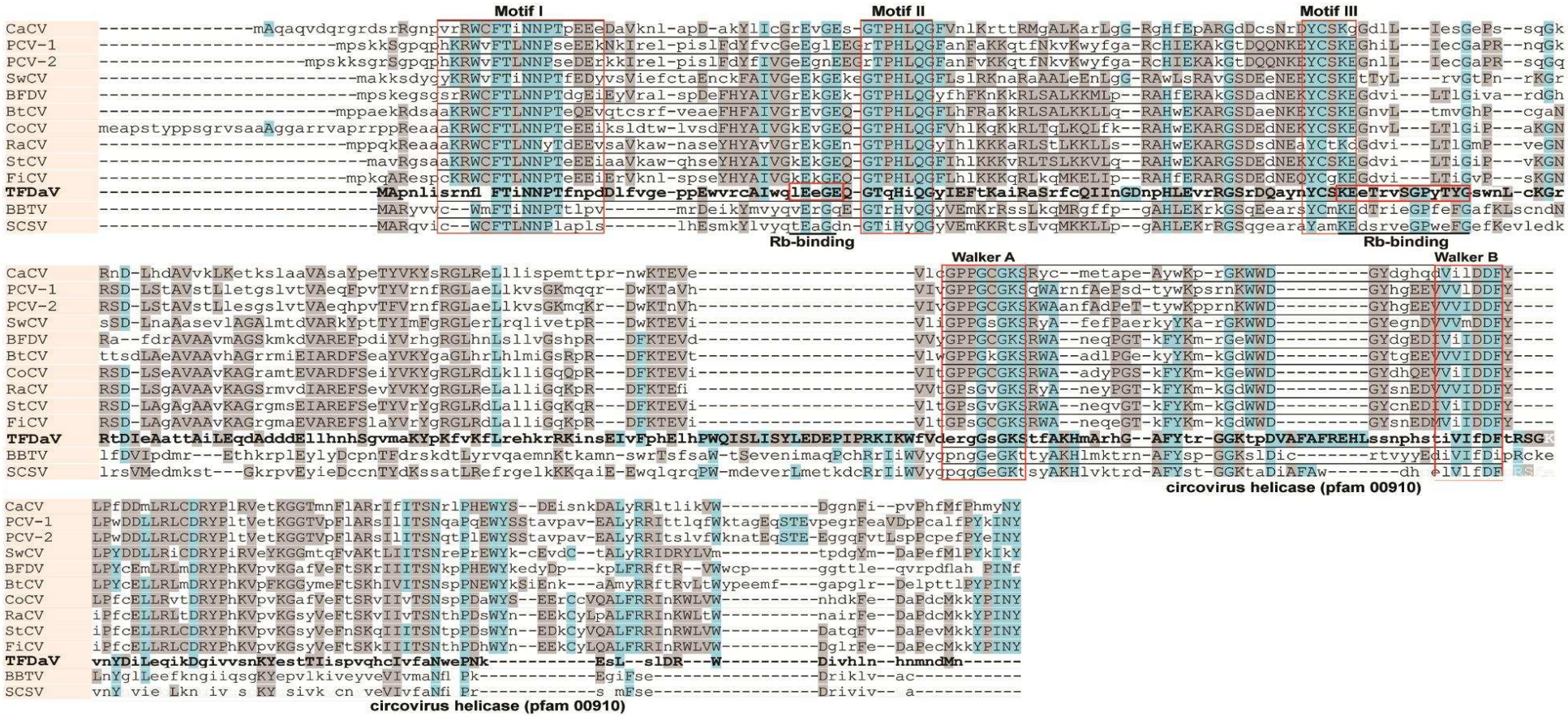
Supplementary Figure S2. Pairwise sequence identities between the 45 clones of TFDaV.



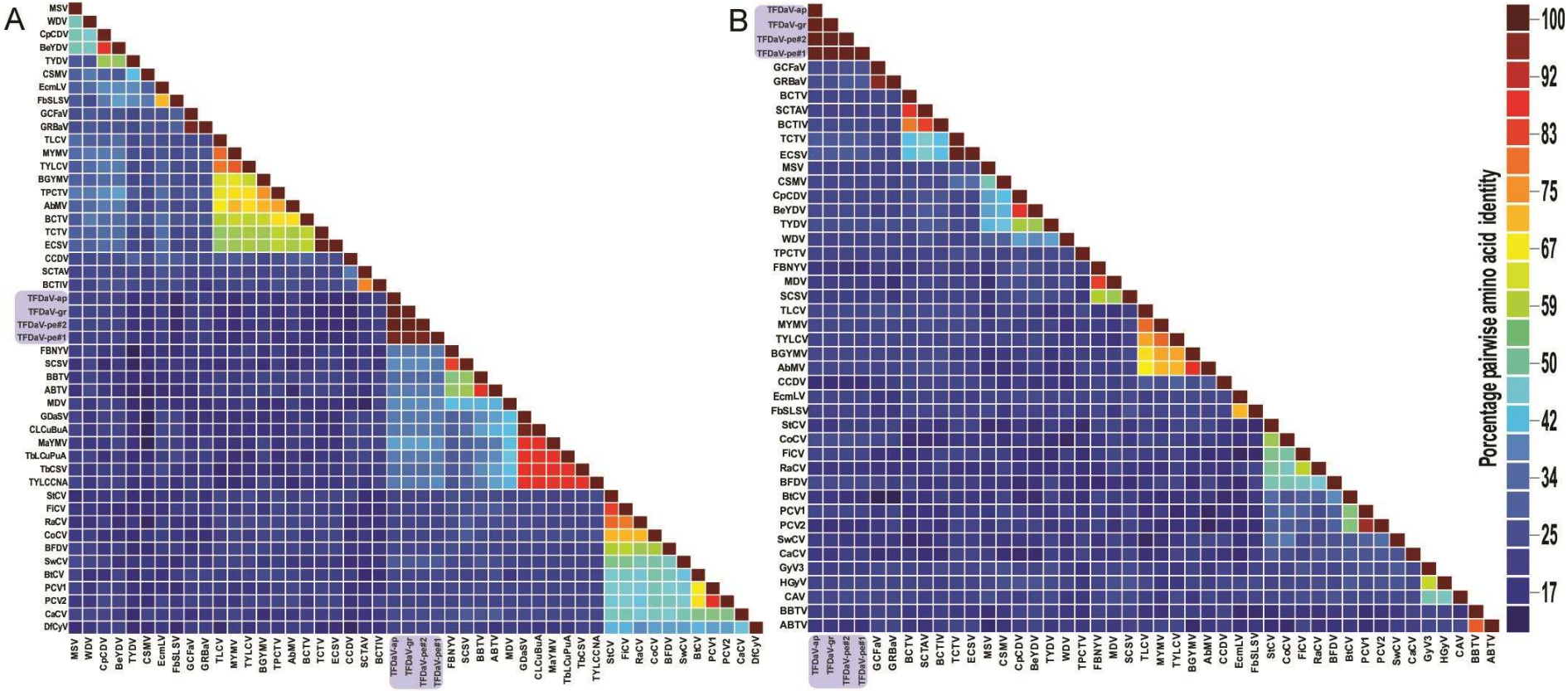
Supplementary Figure S3. Structural analysis of the movement protein. **A.** Predicted secondary structure. **B.** Predicted transmembrane domain. Red: transmembrane region; blue: inner (cytoplasmic) domain; pink: outter (exterior) domain. **C.** Amino acid sequence alignment of movement proteins from TFDaV and from selected nanoviruses and mastreviruses (see Suppl. Table S2 for GenBank access numbers).



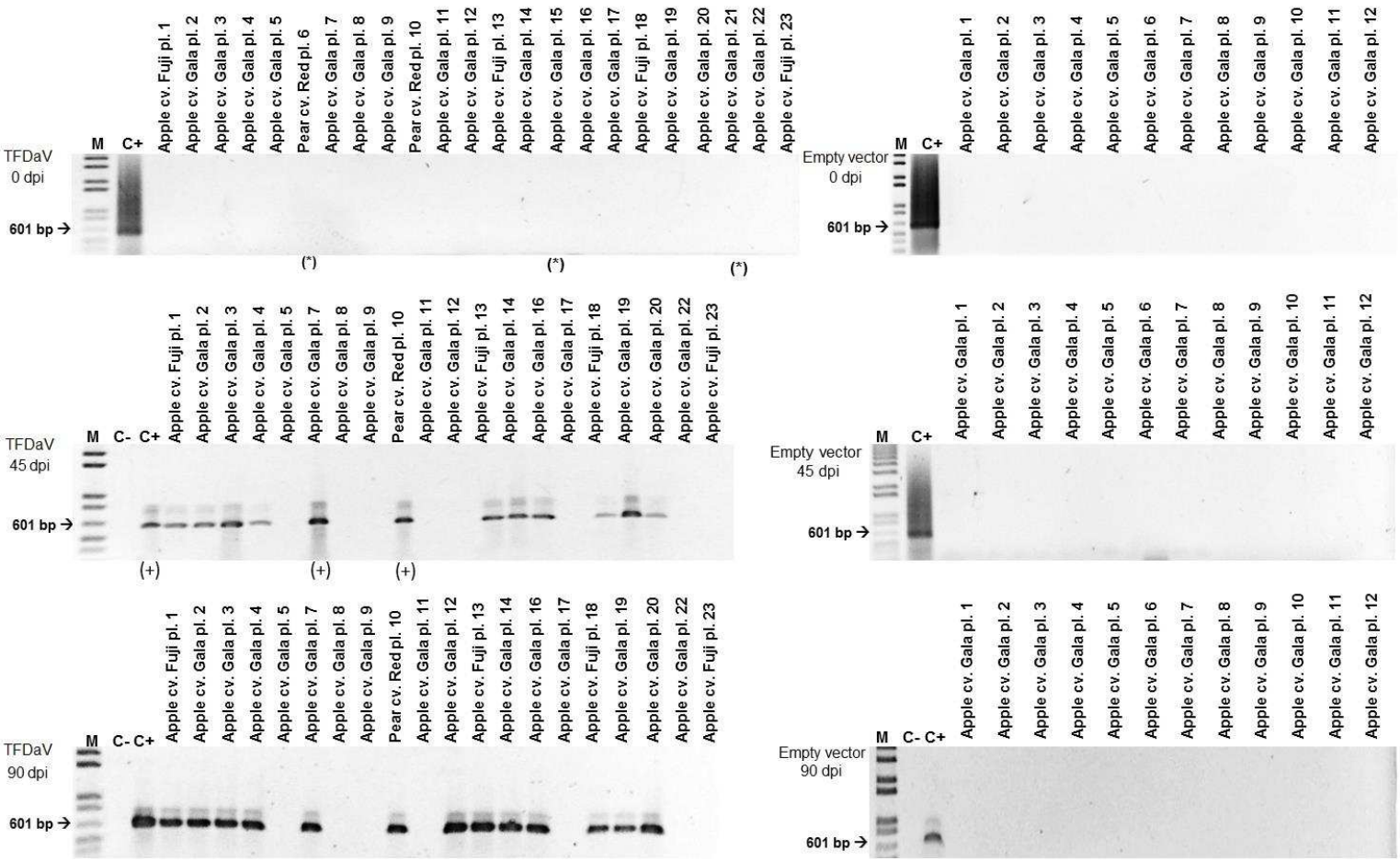
Supplementary Figure S4. TFDaV coat protein amino acid sequence. The predicted importin α-dependent nuclear localization signals (NLS) is in the red box. The confidence level of the analysis is based on a scale from 0 to 9.



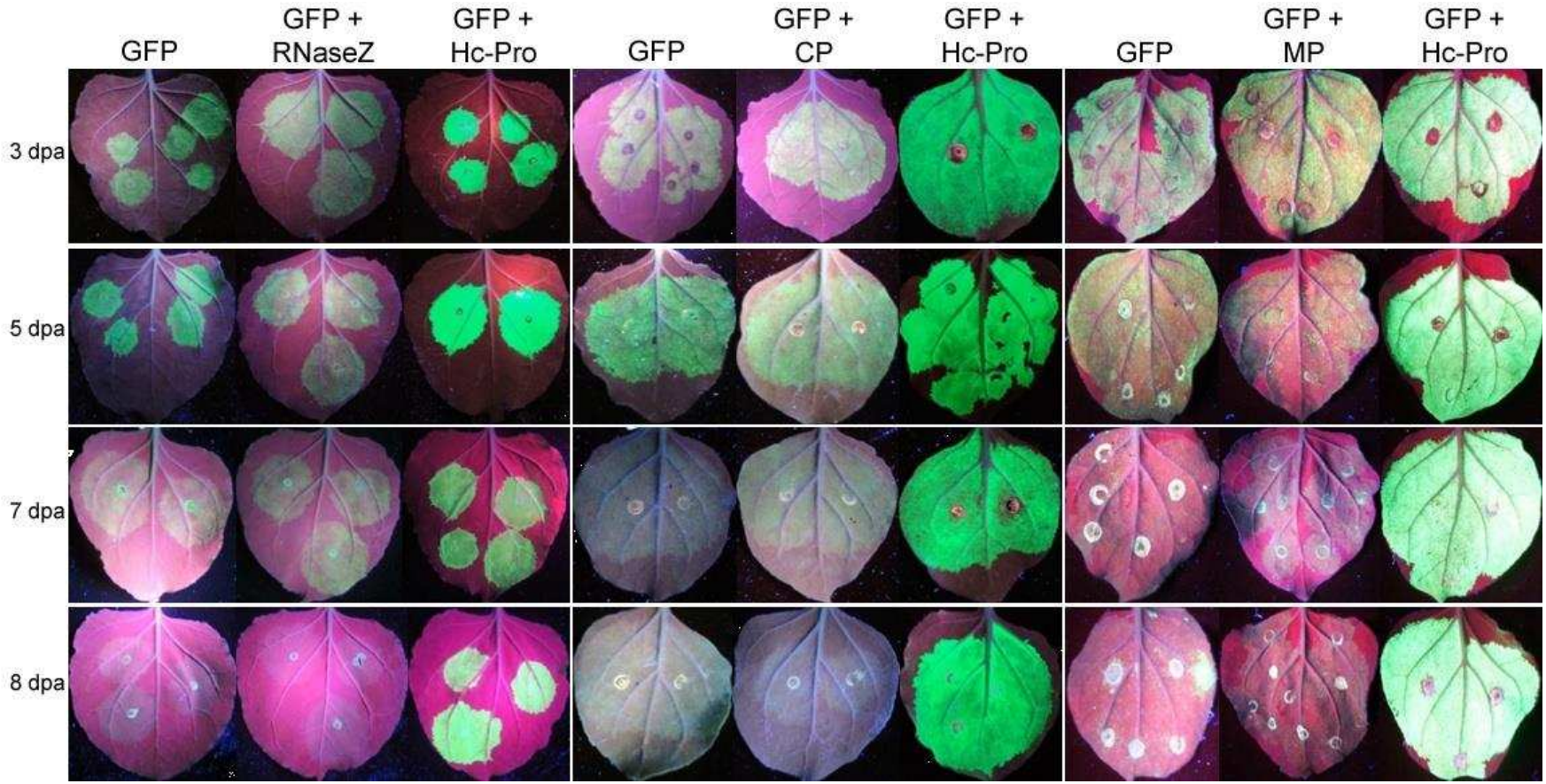
Supplementary Figure S5. Amino acid sequence alignment of the replication-associated proteins (Rep) from TFDaV and from viruses from the Nanoviridae and Circoviridae families (see Suppl. Table S2 for GenBank access numbers). Motifs I, II and III, which are typically found in rolling-circle replication initiator proteins, are indicated by red boxes as well as Walker A and Walker B boxes (dNTP-binding), a helicase domain conserved in circoviruses and two putative Rb-binding motifs.



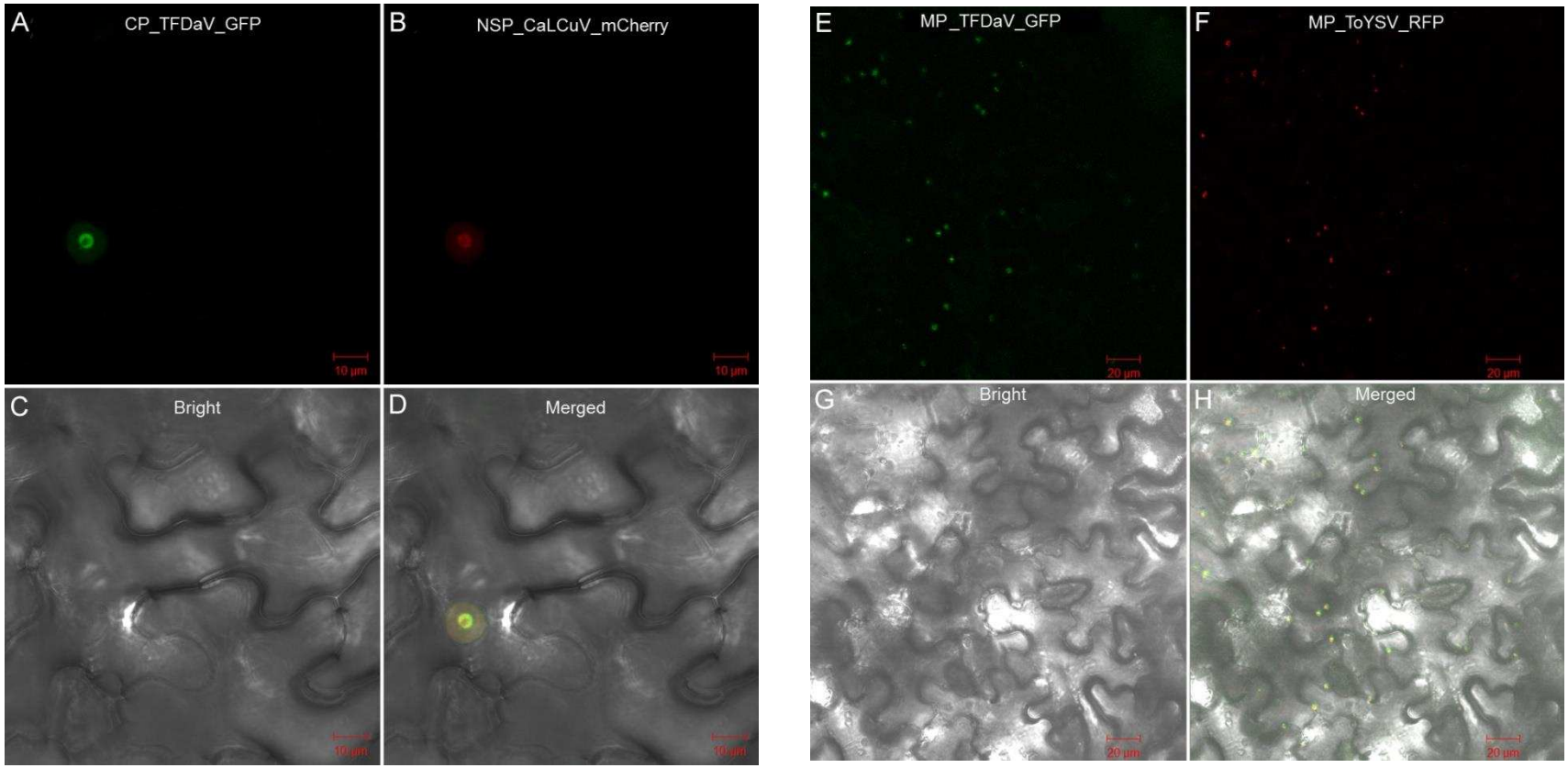
Supplementary Figure S6. Pairwise amino acid sequence identities between (**A**) Rep protein and (**B**) CP of TFDaV with those from other ssDNA viruses (see Suppl. Table S2 for GenBank access numbers).



Supplementary Figure S7. Results of infectivity assay. Agarose gel electrophoresis of PCR products for the detection of TFDaV in biolistically inoculated apple and pear plants. Samples for DNA isolation were composed of young leaves collected at 0 (immediately before inoculating), 45 and 90 days post-inoculation (dpi). The negative control was comprised of leaves from plants biolistically inoculated with the empty vector. Amplicons marked with (+) were cloned and sequenced to confirm viral identity. Plants marked with (*) died after inoculation.



Supplementary Figure S8. Local RNA silencing suppression assay in N. benthamiana. Leaves were infiltrated with mixtures of A. tumefaciens cultures carrying binary vectors for transient overexpression of GFP, HC-Pro from Lettuce mosaic virus (positive control), TFDaV putative RNase Z protein, coat protein (CP) and movement protein (MP). GFP fluorescence was monitored under long-wavelength UV and leaves were photographed at 3, 5, 7 and 8 days post agroinfiltration (dpa).



Supplementary Figure S9. Subcellular localization of TFDaV coat protein (CP) and movement protein (MP) in N. benthamiana leaves. **A-D.** Leaves co-agroinfiltrated with GFP-tagged TFDaV CP and Cabbage leaf curl virus mCherry-tagged nuclear shuttle protein (NSP, positive control for nuclear localization). **E-H.** Leaves co-infiltrated with GFP-tagged TFDaV MP and Tomato yellow spot virus mCherry-tagged MP (positive control for plasma membrane localization). Images were taken with a LSM 510 META confocal microscope (Carl Zeiss) attached to an Axiocam camera.

CAPÍTULO 2

PROFILE OF SMALL RNAS FROM Nicotiana benthamiana INFECTED WITH THE BEGOMOVIRUS Tomato yellow spot virus

Basso, M.F., Silva, J.C.F., Brustolini, O.J.B., Fontes, E.P.B., Corrêa, R.L., Llave, C & Zerbini, F.M. Profile of small RNAs from Nicotiana benthamiana infected with the begomovirus Tomato yellow spot virus. Em preparação.

Profile of small RNAs from Nicotiana benthamiana infected with the begomovirus Tomato yellow spot virus

Marcos Fernando Basso[1], José Cleydson Ferreira da Silva[2], Otávio José Bernardes Brustolini[2], Elizabeth Pacheco Batista Fontes[2,3], Régis Lopes Corrêa[4], César Llave[5] and Francisco Murilo Zerbini[1,3]

[1]Dep. de Fitopatologia, [2]National Research Institute for Plant-Pest Interactions (INCT-IPP) and [3]Dep. de Bioquímica e Biologia Molecular, Universidade Federal de Viçosa, 36570-900, Viçosa, MG, Brazil; [4]Dep. de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, 21944-970, Rio de Janeiro, RJ, Brazil; [5]Departamento de Biología de Plantas, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, 28040 Madrid, Spain.

ABSTRACT

Begomoviruses replicate their DNA genomes inside the host cell nucleus and are targets for gene silencing. To counteract, they encode silencing suppressor proteins that affect silencing pathways, causing enhanced susceptibility and changing small (sm)RNA profiles of the host. Changes in the expression of MIR genes, processing and accumulation of miRNAs and their consequences in the differential regulation of host miRNA targets have also been suggested. Using NGS approaches we sequenced libraries of smRNAs from Nicotiana benthamiana, healthy or infected with the begomovirus Tomato yellow spot virus (ToYSV), to understand their complexity and to explore the regulatory mechanisms involved in pathogenesis. The results indicate that begomoviruses are targets of RNA silencing pathways (RSP) early in the infectious cycle and that their silencing suppressor proteins act in trans to interfere with RSP. The success of viral infection and the appearance of ToYSV-induced symptoms may be correlated, at least in part, with RSP downregulation, siRNAs profile change, MIR genes

down- or upregulation, and transcriptional/posttranscriptional regulation of several host genes. Most ToYSV-regulated MIR genes have genes encoding transcription factors and defense or stress response proteins as targets. The implications of MIR genes and their targets' deregulation are discussed.

INTRODUCTION

RNA silencing pathways (RSPs) in plants control genomic stability and several biological processes through the regulation of gene expression at the transcriptional and posttranscriptional levels, as well as being important defense mechanisms against viruses and other invasive nucleic acids (Hamilton & Baulcombe, 1999; Tuschl et al., 1999; Hammond et al., 2000). RSPs are based on the production of different classes of small RNAs (smRNAs). Depending on their biogenesis and mode of action, smRNAs are grouped into at least eight groups: small interfering RNAs (siRNAs), microRNAs (miRNAs), heterochromatic siRNAs (hc-siRNAs), repeat-associated siRNAs (ra-siRNAs), trans-acting siRNAs (ta-siRNAs), natural antisense siRNAs (nat-siRNAs), transfer RNA-derived fragments (tRFs) and virus-derived siRNAs (vsiRNAs) (reviewed by Voinnet, 2009).

The two most abundant classes are siRNAs and miRNAs. They are derived from single- or double-stranded RNA precursors that are processed by RNaseIII enzymes named Dicer-like (DCL1-4) in plants (Chapman & Carrington, 2007). Then, one strand is degraded and the other is loaded into Argonaute (AGO 1-10)-containing complexes known as RNA-induced silencing complexes (RISC), which scan and catalyze the endonucleotic cleavage of endogenous RNA molecules, viral RNAs or any coding RNA that presents sequence identity with the si/miRNA (a process know as posttranscriptional gene silencing, PTGS). The siRNAs can act also in the induction of DNA methylation, termed transcriptional gene silencing (TGS) (reviewed by Chapman & Carrington, 2007, and Voinnet, 2009).

vsiRNAs can be 18 to 28 nt in length, and are generated by the action of the RNA silencing machinery upon either viral dsRNA intermediates of genome replication, or dsRNA secondary structures along single-stranded viral genomes or transcripts. Once incorporated into RISC, vsiRNAS guide the sequence-specific degradation of cognate viral genomes and also of endogenous sequence-homologous RNAs (Mlotshwa et al., 2008; Wang et al., 2010; Li et al., 2014).

Geminiviruses are a group of single-stranded DNA viruses that cause important economic losses worldwide (Rojas et al., 2005). Seven genera are currently recognized (Varsani et al., 2014). All include viruses with monopartite genomes transmitted by leafhoppers, except the genus Begomovirus, whose members are whitefly-transmitted and have either monopartite (a single genomic DNA of approximately 2.9 kb) or bipartite genomes (two DNA components, each approximately 2.6 kb, referred to as DNA-A and DNA-B) (Brown et al., 2012). The begomoviruses are the most numerous and geographically widespread amongst the geminiviruses (Morales & Anderson, 2001; Monci et al., 2002; Briddon, 2003; Were et al., 2004). Their genomes are replicated in the nuclei of host cells by a rolling-circle mechanism using a dsDNA replicative form, which associates with histones to form a viral minichromosome and is also the template for bidirectional transcription (Hanley-Bowdoin et al., 2013).

Geminiviruses can be inducers and targets of different RSPs. Overlapping geminiviral transcripts are processed to generate vsiRNAs that act in PTGS or are loaded into AGO4-containing RISC to direct TGS of the geminiviral minichromosome (Bisaro, 2006). To counteract these host defenses, begomoviruses encode the Rep, TrAP and AC4 proteins to suppress PTGS and TGS (Vanitharani et al., 2004; Rodríguez-Negrete et al., 2013; Hanley-Bowdoin et al., 2013).

TrAP (transcriptional activator protein, also known as AC2 or AL2) is a nuclear protein which trans-activates transcription of the viral CP and NSP genes (Hanley-Bowdoin et al.,

2013). In begomovirus-infected Arabidopsis thaliana, TrAP increases the expression of Werner exonuclease-like 1 (Wel1), which may function as an endogenous negative regulator of RSPs (Trinks et al., 2005). It was also demonstrated that TrAP interacts with adenosine 5'-phosphotransferase (ADK), a nucleoside kinase that catalyzes the synthesis of 5'-AMP from adenosine and ATP. ADK plays an important role in sustaining the methyl cycle required for S-adenosyl methionine biosynthesis and, consequently, for de novo methylation of viral genomes. Thus, it is suggested that TrAP downregulates ADK activity to counteract the plant's TGS mechanism which acts to reduce transcription from the viral minichromosome (Moffatt et al., 2002; Wang et al., 2003 and 2005; Baliji et al., 2010). TrAP also interacts with SNF1 kinase (Sucrose nonfermenting 1), a global regulator of metabolism, gene expression and growth, which is activated by 5'-AMP in response to a wide variety of cellular stresses. The metabolic alterations mediated by SNF1 are a component of innate antiviral defenses (Hao et al., 2003). Finally, TrAP interacts with the promoter and induces the overexpression of Arabidopsis thaliana regulator of gene silencing-calmodulin-like protein (At-rgsCaM), an endogenous suppressor of RNA silencing, leading to increased host susceptibility. It was shown that TrAP also interacts with At-rgsCaM protein into nucleus, altering its subcellular localization and suggesting that this interaction prevents autophagic degradation of TrAP (Chung et al., 2014). The global protein profile of N. benthamiana in response to the expression of the TrAP gene reveals a cellular imbalance of proteins involved in metabolic and photosynthetic processes, modulation of defense-related and oxidative stress proteins, and transcriptional processes (Carmo et al., 2013).

Rep (also termed AC1, AL1 or C1) is essential for viral rolling circle replication and, additionally interferes with the host DNA cytosine methylation machinery and suppresses TGS by reduction of the expression of the plant DNA methyltransferases, Methyltransferase 1 (MET1) and Chromomethylase 3 (CMT3) (Rodríguez-Negrete et al., 2013).

AC4 (also termed C4) is a cytoplasmatic protein associated with viral movement, induction of symptoms and suppression of PTGS, acting downstream of siRNA biogenesis through the sequestration of single stranded siRNAs, preventing their incorporation into the RISC (Vanitharani et al., 2004; Hanley-Bowdoin et al., 2013).

Despite the large amount of information available regarding begomovirus-host molecular interactions, there is still limited knowledge on the effects of viral infection on RSPs, such as changes in the smRNAs profile in the host and consequent effects in gene expression. Next-generation sequencing (NGS) has become a routine experimental technology that has been used with great success in recent years to evaluate smRNA profiles in plants (Lee et al., 2009; Silva et al., 2011; Thieme et al., 2012; Ye et al., 2012; Zhang et al., 2012; Loss-Morais et al., 2013). siRNA profiles can indicate which regions of the viral genome are hot and cold spots for Dicer processing. More interestingly, the number of reads that match with annotated MIR genes (specially those whose targets are known) allows the discovery of quantitative changes in gene expression levels between virus-infected and healthy plants, providing valuable clues on molecular virus-host interactions (Silva et al., 2011; Romanel et al., 2012; Visser et al., 2014).

The begomovirus Tomato yellow spot virus (ToYSV) infects both cultivated and non-cultivated plants inducing severe symptoms (Andrade et al., 2006; Calegario et al., 2007). Nicotiana benthamiana has been considered one of best models for the study of molecular plant-pathogen interactions as it is susceptible to several fungi, oomycetes, bacteria and viruses (including ToYSV), can be easily transformed or used for the transient expression of proteins in VIGS-based approaches, and possesses homologous genes to other agronomically important solanaceous plants (Goodin et al., 2008; Carmo et al., 2013). Using NGS approaches, we determined the smRNA profile from healthy and ToYSV-infected N. benthamiana plants to understand and explore the regulatory mechanisms involved in begomovirus pathogenesis.

MATERIAL AND METHODS

Quantification of ToYSV load in N. benthamiana. Wild-type N. benthamiana plants were biolistically inoculated with a ToYSV infectious clone (DNA-A and DNA-B; Andrade et al., 2006). Mock-inoculated plants were inoculated with tungsten particles alone. At 5, 7, 10 and 15 days post inoculation (dpi), systemically infected leaves from ToYSV- and mock-inoculated plants were collected and total DNA was isolated as described by Doyle & Doyle (1987). Each treatment (ToYSV and mock) was composed of five individual plants, and these five plants were sampled at the different time points as described above. Total DNA was quantified by spectrophotometry (Nanodrop ND-1000; Thermo Fisher Scientific) and the final concentration was adjusted to 50 ng/µl.

Viral load in infected plants was measured by quantitative PCR (qPCR) from total DNA on a CFX96 Real-Time PCR detection system (Bio-Rad), using Fast SYBR Green Master Mix (Applied Biosystems) following the manufacturer's instructions and 2 µM of forward and reverse primers (qToY3; Suppl. Table S1) in a final volume of 10 µl. The amplification conditions were 20 seg at 95°C for initial denaturation followed by 40 cycles of denaturation at 95°C for 3 seg and annealing/extension at 60°C for 30 seg. All reactions were performed using 5 independent biological samples for each treatment and each sample was analyzed in triplicate wells. The standard curve was obtained by serial dilution of plasmid DNA containing the full-length ToYSV genome (10^1 to 10^6 copies per reaction) and regression analysis of Ct values of each one of three replicates of a given dilution in relation to the log of plasmid DNA amount in each dilution (Suppl. Figure S1). Viral load was determined by interpolation of Ct values of each sample evaluated (Rutledge & Cote, 2003).

Relative quantification was performed using the same reaction components and parameters described above for absolute quantification. The values were normalized against N. benthamiana histone H2b (Xavier, 2012), whose expression was previously shown to be

unaltered in ToYSV- and mock-inoculated plants (data not shown). Each qPCR run included also a no-template control (without DNA).

The specificity of primers used in qPCR assays was confirmed by the amplification of a single fragment in conventional PCR using cDNA from mock- or ToYSV-infected N. benthamiana (Suppl. Figure S2) and by the appearance of a single peak in the melting curve (data not shown).

NbWEL2 gene expression in mock- and ToYSV-infected N. benthamiana. Wild-type N. benthamiana plants were biolistically inoculated with ToYSV as described above. Systemically infected leaves from mock- or ToYSV-inoculated plants were collected at 5, 7, 10 and 15 dpi and total RNA was isolated with the Plant RNA Purification Reagent (Invitrogen) following the manufacturer's instructions. The RNA was quantified by spectrophotometry and its integrity was verified by agarose gel electrophoresis. It was then treated with RNase-free DNase I (Promega), and 2 µg of DNase-treated RNA were used as a template for cDNA synthesis using 100 µM of oligo-$dT_{(18)}$ primer and SuperScript III Reverse Transcriptase (Invitrogen), according to the manufacturer's instructions. The cDNAs were quantified by spectrophotometry, their final concentration was adjusted to 100 ng/µl and 2 µl were analyzed by qPCR using 2 µM of each specific primer pair (NbWEL2a, NbWEL2b and H2b; Suppl. Table S1). The qPCR assays were performed and analyzed as described above for relative quantification. The specificity of primers used was confirmed by the amplification of a single fragment in conventional PCR using cDNA from mock- or ToYSV-infected N. benthamiana (Suppl. Figure S3) and by the appearance of a single peak in the melting curve (data not shown).

Isolation of smRNAs, libraries construction and sequencing. Wild-type N. benthamiana plants were biolistically inoculated with ToYSV as described above. Systemically infected

leaves from mock- or ToYSV-inoculated plants were harvested at 5 and 10 dpi and smRNAs (10-50 bp) were isolated using the mirVana miRNA Isolation Kit (Ambion), following the manufacturer's instructions. The amount and quality of smRNAs were measured using, respectively, a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific; Suppl. Table S2) and an Agilent 2100 Bioanalyzer (Agilent Technologies). Systemic infection in ToYSV-inoculated plants and the healthy status of mock-inoculated plants were confirmed at 5 dpi by conventional PCR using degenerate primers for the DNA-A (PAL1v1978 and PAR1C496) and DNA-B (PBL1v2040 and PCRc1) of begomoviruses (Rojas et al., 1993) (data not shown). Two independent biological experiments (R1 and R2) were conducted, resulting in eight smRNAseq libraries (library 1: Mock-5dpi-R1; library 2: Mock-10dpi-R1; library 3: ToYSV-5dpi-R1; library 4: ToYSV-10dpi-R1; library 5: Mock-5dpi-R2; library 6: Mock-10dpi-R2; library 7: ToYSV-5dpi-R2; library 8: ToYSV-10dpi-R2). Each library was comprised of one set of 12 plants, with the same plants sampled at 5 and 10 dpi.

Libraries 1-4 were sequenced at Fasteris Life Science Co. (Geneva, Switzerland), and libraries 5-8 were sequenced at Macrogen (Seoul, South Korea), in both cases using the Illumina HiSeq 2000 platform. The libraries were constructed using the TruSeq small RNA SBS kit v. 3 (Illumina) according to the manufacturer's instructions. smRNAs of 15-30 nt were size-selected on denaturing PAGE. The 3' IDT miRNA cloning linker (Integrated DNA Technologies) and 5' Illumina adapters were linked to smRNAs with T4 RNA ligase. Adapter-linked smRNAs were again purified and submitted to reverse transcription. The cDNA was purified on 6% Novex TBE PAGE (Life Technologies) to remove empty adapters and submitted to PCR with specific primers (adapters) that contain one library-specific index sequence (barcode). PCR products were again purified and quantified by qPCR, size-checked in an Agilent 2100 Bioanalyzer and the index sequence was verified by Sanger sequencing. High-throughput sequencing was performed as single-end and multiplexed for 50 cycles plus 7 cycles to read the indexes. The libraries were demultiplexed with CASAVA v. 1.8.4

(Illumina) allowing one mismatch in the index selection and generating output .fastq format files.

Data mining and bioinformatics analysis. Sequence quality was evaluated by measurement of the percentage of bases having a base quality greater than or equal to 20 (Q20) and 30 (Q30) using the FASTX Toolkit version 0.0.6 (hannonlab.cshl.edu/fastx_toolkit/). Then, 3' adapters, primer sequences and reads with undetermined bases were trimmed with Prinseq software (Schmieder & Edwards, 2011). Reads of <18 and >28 nt were filtered with FASTX Toolkit. The quality of sequences was again confirmed with FastQC software (Andrews, 2010) (Suppl. Figure S4). All aligments were performed using Bowtie 1.1.1 (Langmead et al., 2009), allowing no mismatches. The alignment files were converted to .bam format using SAMtools (Li et al., 2009) and the files .sam format were visualized with Tablet software (Milne et al., 2013).

The filtered reads from each library were matched against each ToYSV genomic component (ToYSV-A: NC_007726 and ToYSV-B: NC_007727). The amount of matching reads and their distribution by genome position were calculated using locally developed Perl scripts, normalized in reads per 1 million (RPM) and the total number of filtered reads for each library was used for normalization. The read distribution maps were generated with R 2.7.1 software (R Foundation for Statistical Computing). The same reads matching ToYSV genomes were also aligned (allowing no mismatches) to N. benthamiana transcriptome (transcriptome assembly v. 3; Nakasugi et al., 2013), in eukaryotic tRNAs (genomic tRNA database) and rRNAs (SILVA database; Quast et al., 2013).

Differential expression of known MIR genes. The filtered libraries (reads of 18 to 28 nt; no matches to rRNA, tRNA or ToYSV genome) were matched with known Viridiplantae mature miRNA and hairpin MIR genes deposited into miRBase database release 21 (Kozomara &

Griffiths-Jones, 2014), using Bowtie 1.1.1 software (Langmead et al., 2009) allowing two mismatches. The amount of matched reads was calculated using locally developed Perl scripts. The expression level of a MIR gene was represented by the number of occurrences of its respective read/sequence in each library. To generate a conservative list, we filtered this data set by considering only miRNAs with read counts ≥10 in all samples/libraries (miRNAs with fewer reads did not exhibit dramatic expression changes). The differential expression of MIR genes was calculated using two approaches. Initially, the IDEG6 software (Romualdi et al., 2003; available at telethon.bio.unipd.it/bioinfo/IDEG6_form/) was used, with the libraries normalized in terms of reads per 1 million (RPM; RPM = number of mapped miRNA reads / number of clean sample reads × 10^6). The total number of filtered reads for each library was used for normalization (Table 1). Statistical analysis (Fisher exact test and Audic and Claverie test) were performed with p-value cutoff <0.05 and Bonferroni correction. The second approach used the DEGseq Bioconductor R package version 1.20.0 (Wang et al., 2010; available at www.bioconductor.org/packages/release/bioc/html/DEGseq.html), implementing Fisher's exact test and determining differential expression using binomial or Poisson distribution. The libraries were normalized by the "median" method and statistical analysis was performed with a p-value and two q-values (adjusted p-values with BH and ST methods) (Benjamini & Hochberg, 1995; Storey & Tibshirani, 2003). We considered MIR genes to be differentially expressed and biologically significant only when the p-value and the two q-values were <0.05, and the fold change was >0.8 or <-0.8 between two contrasting treatments (ToYSV or mock and 5 or 10 dpi) for each gene. Only differentially expressed MIR genes supported by both statistical tests and approaches (IDEG6 and DEGseq) were considered as strong candidates to being over- or downregulated during ToYSV infection. The normalized MIR gene expression data will be submitted to NCBI Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo/).

Three additional packages were also tested: edgeR Bioconductor R package v. 3.8.5 (Robinson et al., 2010; available at www.bioconductor.org/packages/release/bioc/html/edgeR.html), baySeq Bioconductor R package v. 2.0.50 (Hardcastle & Kelly, 2010; available at www.bioconductor.org/packages/release/bioc/html/baySeq.html) and DESeq Bioconductor R package v. 1.18.0 (Anders & Huber, 2010; available at bioconductor.org/packages/release/bioc/ html/DESeq.html). However, DEGseq was the one that provided the better adjustment to our smRNAseq data. The normalization method used in IDEG6 (RPM) is also widely used. Using the two statistical aproaches (Fisher exact test or Audic and Claverie test) implemented in IDEG6, no significant differences were observed between DEGseq and IDEG6 (Suppl. Tables S2 and S3).

Identification of putative novel MIR genes and differential expression. Novel MIR genes were identified in the N. benthamiana draft genome assembly v. 1.0.1 (available at 06/03/2015 in ftp.solgenomics.net/genomes/Nicotiana_benthamiana/assemblies/annotation) from mock and ToYSV libraries (reads of 18 to 28 nt; unmatched to rRNA, tRNA, known mature and hairpin MIR genes from miRBase) using the miRCat pipeline and the UEA sRNA toolkit (Stocks et al., 2012). miRCat annotates putative novel MIR genes through non-comparative prediction methods based on expressed miRNA sequences and stable hairpin structures. It evaluates several features, including the number of read stacks in the locus, the number of reads mapping anti-sense to the locus, the size of bulges in the candidate miRNA duplex, the number/fraction of paired nucleotides in the duplex and in the hairpin, and the energetic stability of the hairpin. miRCat was run with the following parameters: (i) minimum smRNA abundance of five reads, (ii) minimum size of 18 nt and maximum 28 nt, (iii) maximum number of 10 genome hits, (iv) at least 68% of the nucleotides centered around the miRNA must be involved in base-pairing, (v) the hairpin must be at least 75 nt in length, (vi) no more than three consecutive mismatches between miRNA and MIR gene, (vii) at least 50%

of bases in the hairpin should be paired, (viii) at least 10 reads matched without mismatches to each of the two possible mature miRNAs derived from the hairpin precursor, (ix) the most abundant reads from each arm of the precursor must pair in the mature miRNA duplex with 0-4 nt overhangs at their 3' ends, (x) at least 50% of reads mapping to each arm of the hairpin precursor must have the same 5' end, (xi) p-value cutoff <0.05. The remaining criteria were run with default parameters. The resulting secondary structure was filtered by miRCat based on: (i) the predicted secondary structure must contain a highly negative minimal free energy (MFE) and high minimal free energy index (MFEI) values (>0.6); (ii) presence of miRNA* (star sequence); and (iii) the miRNA and miRNA* are both >20 and <25 nt. The MFEI was calculated manually as MFEI = [(MFE/length of the precursor RNA sequence)×100] /(G+C)%]. Only reads with an MFEI value >0.6 were considered to be miRNAs candidates. The folding structures of the precursors of the novel miRNA within the MIR gene identified were predicted using the UEA sRNA toolkit RNA/Folding Annotation (Stocks et al., 2012), which uses the Vienna Package to obtain the secondary structure of a precursor sequence and highlights miRNA/MIR gene sequences on the hairpin structure.

These putative novel MIR genes were filtered and assumed as promising when they were present in the respective libraries corresponding to two replications of each treatment. Their differential expression was measured as described above. These sets were considered as high confidence miRNA/MIR gene sequences and will be deposited in the miRBase database (Kozomara & Griffiths-Jones, 2014).

Prediction of vsiRNAs, known and novel miRNA targets. Known and putative novel MIR genes, which were identified as differentially expressed by viral infection (p-value and q-values <0.05 and fold change >0.8 or <-0.8) were used to predict potential miRNA targets. The vsiRNAs targets were indentified by vsiRNAs directly aligned in the N. benthamiana transcriptome

(ftp://ftp.solgenomics.net/genomes/Nicotiana_benthamiana/annotation/Niben101/) assembly v. 1.0.1 using Bowtie 1.1.1 software (Langmead et al., 2009) allowing one mismatch, and only targets with more than 100 reads matched per mRNA were represented. Known and novel mature miRNAs targets were identified using Plant Small RNA Target Analysis Server (Dai & Zhao, 2011; plantgrn.noble.org/psRNATarget/) with the same N. benthamiana transcriptome version described above and the following parameters (i) maximum expectation less than 2.0; (ii) 20 bp of length for complementarity scoring; (iii) target accessibility equal to 25; (iv) flanking length around target site for target accessibility analysis of 17 bp upstream and 13 bp downstream; and (v) range of central mismatch leading to translational inhibition of 9-11 nt. Next, identified targets were subjected to gene ontology analysis. We extracted the targets GO annotations and these were subjected to AgBase toolkit (McCarthy et al., 2006). A database (Niben ToYSV smallRNAseq Database) was developed to facilitate information access to the up- and down-regulated MIR genes and their targets and will soon be available.

Characterization of siRNA candidates. The remaining smRNAs of 18 to 28 nts from each library (unmapped to ToYSV, rRNA, tRNA and known and novel miRNAs) were classified as siRNA candidates. The smRNAs aligment against specific databases was performed using Bowtie 1.1.1 software (Langmead et al., 2009) allowing one mismatch, and the amount was calculated using locally developed Perl scripts. To search for siRNAs candidates to retroelements, A. thaliana, Oryza sativa japonica, Solanum lycopersicum, Sorghum bicolor and Zea mays retrotransposons, retrieved from the MASiVEdb database, were used as references (Bousios et al., 2012). Non-coding RNAs (lncRNA, snoRNA and snRNA) were retrieved from the Plant Non-coding RNA Database (Yi et al., 2015; structuralbiology.cau.edu.cn/PNRD/download.php). Miniature inverted-repeat transposable elements (MITEs) were retrieved from the Plant MITE Database (Chen et al., 2013; pmite.hzau.edu.cn/). Repetitive DNA elements from Solanaceae, Brassicacea, Fabaceae and

Poaceae were retrieved from the TIGR Plant Repeat Databases (Ouyang & Buell, 2004; plantrepeats.plantbiology.msu.edu/downloads.html). To discriminate smRNAs from mRNA degradation products, the remaining smRNAs were aligned against the N. benthamiana transcriptome assembly v. 1.0.1.

RESULTS

ToYSV load and level of NbWEL2 gene expression. ToYSV-inoculated N. benthamiana plants showed the first symptoms at 3 to 4 dpi with all plants symptomatic at 5 dpi (Figure 1A). The typical symptoms were severe yellow mosaic, leaf rolling and reduction in growth and in the leaf area, reaching the greatest severity at 12 to 15 dpi (Figure 1B). ToYSV load was evaluated in systemically infected leaves at 5, 7, 10 and 15 dpi. Viral load increased progressively until reaching a peak at 10 dpi, followed by a small decrease at 15 dpi (Figure 2). The NbWEL2 gene expression level was also evaluated at 5, 7, 10 and 15 dpi, with higher accumulation at 5 and 10 dpi (Figure 3), correlating with the viral load and suggesting downregulation of the RSP. Based on these observations, N. benthamiana plants at 5 and 10 dpi were selected for evaluation of their smRNA profile.

Characterization of smRNAseq libraries. To characterize the smRNA profile produced during ToYSV infection, total smRNAs obtained from ToYSV- and mock-inoculated plants (5 and 10 dpi) were isolated and sequenced using the HiSeq 2000 Illumina platform. To analyze these libraries we followed the workflow presented in Figure 4. A large amount of filtered reads were obtained from the eight libraries and all showed high quality per base (Table 1 and Suppl. Figure S4). Of the total redundants reads, approximately 10% were unique reads (Table 1). Total reads of 21, 22 and 24 nt were the most prevalent in all libraries (Figure 5A) and unique reads of 23 and 24 nt were the most prevalent in all libraries (Figure

5B), with 24 nt reads (total and unique) being 2 to 5 times more abundant than the others (Figure 5A, B). A relatively high amount of smRNAs in both the ToYSV and mock libraries were derived from tRNAs, in contrast with insignificant amounts of rRNAs (Table 1).

To characterize the viral-derived smRNAs produced during ToYSV infection, reads ranging from 18 to 28 nt were mapped against the ToYSV DNA-A and DNA-B components. Only sequences showing no mismatches were regarded as bona fide ToYSV vsiRNAs in the ToYSV libraries. A considerable amount of vsiRNAs were identified (Figure 5C-F) covering almost the entire sequence of the genome. However, their distribution along the genome was not uniform, with hotspots in coding regions (mainly the CP and NSP genes) and with the DNA-B having almost twice as many matched reads compared to the DNA-A (Table 1; Figure 5G, H). These vsiRNAs hotspots might be generated by the large expression levels of these viral genes. Alternatively, these regions may have structural characteristics that make them more accessible to DCL ribonucleases. Temporally, a higher accumulation of total or unique 18 to 28 nt vsiRNAs was observed for each 5 dpi replication compared to 10 dpi (Table 1; Figure 5C-H). Equivalent amounts of sense and antisense vsiRNAs were found in the ToYSV libraries, suggesting that vsiRNAs are derived from the cleavage of dsRNA (Figure 5E, 5F).

To further characterize the libraries, the frequencies of redundant and unique 18-28 nt vsiRNAs were analyzed and in all cases 21 and 22 nt vsiRNAs were the most abundant, (Figure 5C, D). These results suggest that N. benthamiana homologues of DCL1 and DCL2 may be the predominant DCLs involved in their biogenesis, and that all DCLs contribute to the generation of vsiRNAs, but with distinct substrate affinities. Additionally, we compared the overall profile of smRNAs between the ToYSV and mock libraries and observed that 21 and 22 nt reads were predominant only in the pool of vsiRNAs. In both libraries, the 24 nt class was the most abundant among the endogenous smRNAs, followed by the 21, 22 and 23 nt class (Figure 5A, B). However, the infected plants displayed decreased levels of the 24 nt

class as compared to mock-inoculated plants (Figure 5B), indicating that ToYSV infection may decrease the production of endogenous smRNAs. The high accumulation of the endogenous 24 nt smRNAs is consistent with smRNA profiles observed in other plants, indicating that the N. benthamiana RNA silencing machinery responsible for the biogenesis of endogenous or vsiRNAs does not tend to produce 22 nt sequences, and that the high levels of 21 and 22 nt vsiRNAs observed is the result of antiviral RNA silencing mechanisms or a specific ToYSV-host interaction.

In Arabidopsis thaliana, the 5'-terminal nucleotide determines the preference of siRNAs for loading in AGO-containing RISC. For example, AGO1 has a 5'-terminal nucleotide preference for uracil, AGO2 and AGO4 for adenine, and AGO5 for cytosine (Mi et al., 2008; Takeda et al., 2008). The distribution of 5'-terminal nucleotides was determined for the sequenced vsiRNAs. For all four predominant types of ToYSV vsiRNAs characterized (21 to 24 nt), we observed a uniform distribution with a tendency to uracil being the most commonly occurring nucleotide and citosine the least common (Figure 5I), suggesting that ToYSV vsiRNAs can be potentially loaded to multiple AGO-containing complexes.

Novel MIR genes in N. benthamiana. A search for novel miRNA candidates was carried out using the miRCat pipeline mapped to the N. benthamiana genome using filtered reads from the miRBase database release 21 (Kozomara & Griffiths-Jones, 2014). For the identification of novel MIR genes, we selected only miRNA precursors that exhibited miRNA and miRNA* and that meet all criteria for annotation of plant MIR genes. Finally, 1,777 novel MIR genes (Nb5dpi-Mock-R1: 875, Nb5dpi-Mock-R2: 627, Nb10dpi-Mock-R1: 551, Nb10dpi-Mock-R2: 988, Nb5dpi-ToYSV-R1: 618, Nb5dpi-ToYSV-R2: 787, Nb10dpi-ToYSV-R1: 688 and Nb10dpi-ToYSV-R2: 982) were predicted based on the N. benthamiana draft genome, with approx. 4% being differentially expressed due to ToYSV infection. The length of putative novel MIR gene precursors ranged from 75 to 200 nt. Their hairpin, miRNA/miRNA*

sequences, genome locus and results of all evaluated parameters are available in Suppl. File S1. The majority of the novel miRNAs are 24 nt in length and the miRNAs* are 22 to 23 nt.

Differential expression of known and novel MIR genes. To identify known miRNAs, 18 to 28 nt smRNAs from both ToYSV and mock libraries were used to search against the miRBase database release 21 (Kozomara & Griffiths-Jones, 2014). Using the IDEG6 software we observed 1.008 (Fisher exact test) and 954 (Audic and Claverie test) differentially expressed known MIR genes between both contrasts (Suppl. Files S2 and S3), supported by a p-value <0.05. We did not observe a significative diference between the two statistical tests (Suppl. Files S2 and S3). This approach has been used with smRNAseq data for differential expression studies (Thiebaut et al., 2014; Ma et al., 2015). However, we observed a tendency of recent works to use approaches based on adjusted p-values with BH and ST methods, and cutoffs from <0.05 to <0.01. Therefore, other statistical packages were evaluated and only MIR genes scored in multiple approaches were considered promising. Using the DEGseq package, MIR genes were only considered to be differentially expressed when their fold change was >0.8 or <-0.8. Independently of the data normalization method used (none, loess or median) we verified no significantly alterations in our dataset, and therefore opted for the "median" method. The results obtained using the DEGseq package were similar to the ones obtained using IDEG6, although MIR genes were detected in a statistically reduced amount (Suppl. Files S4 to S6).

To evaluate the expression level of novel MIR genes we used the same approaches and criteria described above for expression analysis of known MIR genes, but using only the DEGseq Bioconductor R package. The identification of miRNA targets was performed using the psRNATarget package (Dai & Zhao, 2011), allowing us to identify a greater number of miRNAs targets. We also observed that the majority of novel and known MIR genes might

have more than one mRNA target, and in some cases it was predicted that they can act also by inhibiting translation of the target mRNA (Suppl. Files S4 to S5).

Then, we examined the overall expression profile among all treatments. Using the pvclust R package (Shimodaira, 2002), the hierarchical clustering via multiscale bootstrap resampling method was employed to obtain the clusters from the normalized treatment data sets (Suppl. Figure S5). The dispersion of differentially expressed MIR genes in each contrast is shown in the form of MA-plots, and the statistically significant ones are indicated in red. The box-plots of read counts for each MIR gene, the scatterplots comparing the number of reads for each MIR gene using an ordinary log2 transformation, and the histograms of the number of reads for MIR genes from 5dpi_ToYSV vs 5dpi_Mock, 10dpi_ToYSV vs 10dpi_Mock, 10dpi_Mock vs 5dpi_Mock and 10dpi_ToYSV vs 5dpi_ToYSV contrasts are presented in Suppl. Figure S6.

In the 5dpi_ToYSV vs 5dpi_Mock contrast we detected 124 differentially expressed MIR genes (p-value and the two q-values <0.05 and fold change >0.8 or <-0.8), of which 31% were down-regulated (Suppl. File S4). From these, 16 are known MIR genes belonging to 6 different families (MIR166, 390, 398, 482, 3627 and 5181) and 108 are novel MIR genes. The MIR166 and MIR390 families, which regulate HD-ZIPIII transcription factors and TAS transcripts, respectivelly, were the most abundant in this contrast. The miRNA targets with greater differential expression in this contrast were chaperone protein DnaJ, MYB transcription factors and other transcription factors, ethylene receptor, aminotransferase and DNA-binding protein. MYB and DnaJ proteins are two example of host factors already confirmed as differentially expressed by influence of viral infection (Xavier, 2012; Zorzatto et al., 2015). These potential targets were assigned based on Gene Ontology annotations (Ashburner et al., 2000). Based on molecular function the targets fell largely into 24 categories, with the three most over-represented being nucleotide binding and catalytic activity (Figure 7).

In the 10dpi_ToYSV vs 10dpi_Mock contrast we observed 375 differentially expressed MIR genes, of which 54% were downregulated. From these, 117 are novel MIR genes and 258 are known MIR genes belonging to 52 MIR families, with the MIR156, 166, 171, 319, 393, 396, 397, 398 and 403 families being the most abundant (Suppl. File S5). In A. thaliana, these nine MIR families regulate, respectively: (i) SPL transcription factors; (ii) HD-ZIPIII transcription factors; (iii) SCL transcription factors; (iv) TCP and MYB transcription factors; (v) TIR1, AFB1, 2 and 3 auxin signaling or receptors; (vi) bHLH transcription factors and GRF factors; (vii) drought stress; (viii) responses to biotic and abiotic stresses; (ix) PTGS of AGO2 and AGO3. The targets of these differentially expressed MIR genes included MYB, GRAS, B3, BZIP and others transcription factors, kinases, proteins involved in defense responses and stress- and ubiquitin-related proteins. Based on molecular function the targets fell largely into 23 categories, with DNA binding, catalytic activity and transferase activity being the most representative (Figure 7).

Overall, 65 differentially expressed MIR genes are exclusive of the 5dpi_ToYSV library (with 18 downregulated MIR genes), 326 are exclusive of 10dpi_ToYSV (22 downregulated) and 59 are common between these two libraries (Suppl. File S6). A heatmap of gene expression of these 59 MIR genes is shown in Figure 6. Some of the differentially expressed MIR genes at 5dpi_ToYSV are similarly regulated at 10dpi_ToYSV, while others are down- or up-regulated from 5 to 10 dpi (4 MIR genes downregulated at 5 dpi are upregulated at 10 dpi, and 7 MIR genes upregulated at 5 dpi are downregulated at 10 dpi). For example, a MIR gene targeting a MYB transcription factor is downregulated at 5 dpi but upregulated at 10 dpi, in association with a large number of MIR genes (one MIR gene at 5 dpi and 22 MIR genes at 10 dpi) that present this same target.

vsiRNAs targets. We observed that several vsiRNAs are matched simultaneously also in the N. benthamiana transcriptome. Some of these display a high amount of reads that matched

host mRNAs. For example, at 5 dpi, 7,165 20-22 nt reads per million (or normalized vsiRNAs) matched the NbS00001875g0029.1 mRNA, a protein with unknown function (Suppl. File S8). This suggests that the vsiRNAs can also act in the PTGS of host mRNAs. At 5 dpi we observed that 616 host mRNAs can be indirectly regulated by vsiRNAs, while at 10 dpi this number was reduced to 114 (Suppl. File S9). The targets of vsiRNAs include several proteins with a kinase-like domain, aminotransferases, transcription factors, dnaJ or proteins with a J-like domain, ribosomal proteins and DNA-binding proteins (Suppl. Files S8 and S9). This suggests that the defense response mounted by the host is more intense at 5 dpi or early in the infectious process, and is then attenuated at 10 dpi due to viral suppression of RNA silencing. These potential targets were assigned based on Gene Ontology annotations. With respect to cellular component, the targets fell largely into 11 (5 dpi) and 7 (10 dpi) categories, with the four most over-represented being membrane, celullar component, cell, cytoplasm and intracelular (5 dpi), and membrane, endoplasmatic reticulum, cytoplasm, cellular component and cell (10 dpi) (Figure 8). Based on molecular function the targets fell largely into 18 (5 dpi) and 13 (10 dpi) categories, with the five most over-represented being protein binding, transferase activity, binding, catalytic activity and hydrolase activity (5 dpi), and transferase activity, protein binding, binding, hydrolase activity and nucleic acid binding (10 dpi) (Figure 8). These results indicate that vsiRNAs can act in the downregulation of several host genes, with greater coverage early in the viral infection, and suggest that several vsiRNAs may be directly or indirectly involved in host susceptibility induction through post-transcricional regulation of target genes. Biochemical validation of target degradation will provide stronger evidence for this hypothesis.

siRNAs candidates. The remaining smRNAs of 18 to 28 nt (unmatched to ToYSV, miRNA, rRNA and tRNA) were characterized as siRNA candidates. Interestingly, the siRNAs profile between at eight libraries was similar. In all libraries, the fractions with more

aligned reads were the transcriptome and non-coding RNAs (lncRNA, snoRNA and snRNA), with less reads aligned to repetitive DNA elements and MITEs (Suppl. File S10 and Figure 9). The amount of siRNA candidates between treatments (mock vs ToYSV-infected and 5 vs 10 dpi) did not present significant statistical difference (data not shown).

A high amount of reads were considered as unannotated smRNAs, possibly because the N. benthamiana genome has not yet been completely sequenced. In fact, the poor annotation of the N. benthamiana draft genome and transcriptome, even for the latest version, has hampered the identification of a higher amount of MIR genes possibly regulated by viral infection and their annotated putative targets. Many N. benthamiana proteins remain with unknown or uncharacterized functions.

DISCUSSION

The begomoviruses are one of the largest and most economically important groups of plant viruses and cause several diseases in major crops worldwide, with significant economic losses in several dicotyledonous plants (Morales & Anderson, 2001; Monci et al., 2002; Briddon, 2003; Were et al., 2004).

Begomovirus-host interactions involve several complex mechanisms orchestrated in perfect connection, which depend on the interaction between viral and host partners (Sakamoto et al., 2012; Brustolini et al., 2015; Zorzatto et al., 2015). RNAi-induced transcriptional or post-transcripcional silencing is an important antiviral mechanism in plants (Hamilton & Baulcombe, 1999; Tuschl et al., 1999; Hammond et al., 2000; Chapman & Carrington, 2007; Voinnet, 2009), processing viral transcripts into viral small interfering RNAs (vsiRNAs) and using these vsiRNAs to drive antiviral silencing. To counteract, begomoviruses encode multifuncional proteins (AC4, TrAP and Rep) that act in suppression of these RNAi-induced silencing pathways (RSPs; Bisaro et al., 2006; Rodríguez-Negrete et

al., 2013). Additionally, vsiRNAs also can act on the regulation of the expression of endogenous genes (Shimura et al., 2011; Smith et al., 2011). This interference in RSPs plays an important role in symptom development, attenuation of host defense responses and, ultimately, in the success of the viral infection.

The analysis of the N. benthamiana smRNA profile in response to ToYSV infection confirms that begomoviruses are targeted by PTGS and TGS and can also interfere and regulate directly or indirectly several host factors. We observed several MIR gene families differentially regulated between infected and uninfected plants, indicating that ToYSV infection can disrupt global gene regulatory networks. One of the main explanations for the deregulation of MIR genes and alterations on smRNA profile on infected plants might be related to the action of viral silencing suppressor proteins (Vanitharani et al., 2004; Amin et al., 2011; Rodríguez-Negrete et al., 2013; Liu et al., 2014). In begomovirus-infected A. thaliana, TrAP increases the expression of the Wel1 gene, which may function as an endogenous negative regulator of RNA silencing (Trinks et al., 2005). AtWel-1 is located within a cluster of seven genes coding for highly homologous proteins differing in size due to short deletions and short or long insertions or duplications. Another gene from this cluster (At3g12440) is also induced by TrAP, albeit rather weakly (2.7 fold), whereas At3g12470 and At3g12420 are not induced (Trinks et al., 2005). In N. benthamiana, nine putative NbWel genes are annotated. We selected and monitor NbWel2 gene expression during ToYSV infection and confirmed its up-regulation with a tendency to correlate its greatest expression level with the highest viral accumulation. TrAP also interacts with and downregulates ADK, interfering with the TGS pathway (Moffatt et al., 2002; Wang et al., 2003 and 2005; Baliji et al., 2010). In addition, TrAP interacts with the promoter and induces the overexpression of At-rgsCaM, which acts also as an endogenous suppressor of RNA silencing (Chung et al., 2014). Similarly, AC4 acts in the suppression of RNA silencing but downstream of siRNA biogenesis, preventing their incorporation into the RISC (Vanitharani et al., 2004; Hanley-

Bowdoin et al., 2013). Rep and TrAP also interfere with the host's DNA methylation machinery and TGS, suggesting epigenetic deregulation. These punctual interferences in host RSPs allow the virus to be successful in establishing a systemic infection.

RSPs are organized to control various endogenous process based on TGS and PTGS of host genes. Viral interference with these processes results in severe consequences for the plant. Symptom severity might be positively correlated with a larger amount of viral DNA accumulation, possibly as a consequence of deregulation of RSPs and MIR genes. Such RSP disruption, as well as interference in other factors, leads to a reorganization of the host's defense response. It is known that viruses can affect MIR gene-based pathways in two ways: (i) by alterating the rate of transcription or (ii) by changing the accumulation of the miRNA, interfering directly in the regulation of its targets (Bazzini et al, 2007; Bazzini et al, 2009; Silva et al., 2011; Romanel et al., 2012).

Some MIR gene families are known to participate in distinct and overlapping regulatory networks such as disease resistance, development and stress response (Mallory & Vaucheret, 2006). It has been observed that regulation of a single target for a specific miRNA (MIR167) is the main cause of the symptoms induced by at least three different viral suppressors in A. thaliana (Jay et al., 2011). Another example is MIR159, which is regulated upon viral infection by several viruses, with the degree of up-regulation corresponding with symptom severity in most cases (Naqvi et al., 2010; Amin et al., 2011). We verified that MIR159 was not differentially expressed at 5 dpi, but was down-regulated at 10 dpi. Its putative targets are MYB transcription factors, which possibly act in antiviral defense reponses (Zorzatto et al., 2015). Other differentially expressed MIR gene families were also detected in infected plants: (i) MIR319, which targets members of the TCP family of transcription factors (Nag et al., 2009); (ii) MIR403, which acts in miRNA pathways (Mallory & Vaucheret, 2006); (iii) MIR390, whose targets are primary TAS transcripts encoding ta-siRNAs (Marin et al., 2010); (iv) MIR166, which spatially restricts the expression of class III homeodomain leucine zipper

(HD-ZIPIII) transcription factors; (v) MIR396, which targets seven growth-regulating factor (GRF) genes and the basic helix-loop-helix transcription factor 74 (bHLH74) gene in A. thaliana (Bao et al., 2014); (vi) MIR403, which has been implicated in targeting of AGO2 and AGO3 mRNAs in A. thaliana (Harvey et al., 2011); and (vii) MIR156, which acts by repressing the expression of functionally distinct SPL transcription factors (Wu et al., 2009). We observed that most of these MIR genes were up-regulated at 5 dpi and down-regulated at 10 dpi in infected plants. The miRNA and vsiRNA targets identified are correlated, at least partially, with transcriptomics of the interaction between host and geminiviruses (Carmo et al., 2013; Miozzi et al., 2014; Brustolini et al., 2015).

In summary, we detected differentially expressed genes belonging to 70 different MIR families. The up- or down-regulation of MIR genes and their putative miRNAs targets suggests that many host factors can be regulated by ToYSV to increase the susceptibility of N. benthamiana. Furthermore, several additional putative miRNA targets were regulated during ToYSV infection which may be involved in hitherto unknown antiviral defense responses. Several MIR genes had their expression increased or decreased at 5 dpi and 10 dpi in ToYSV-infected plants. Several transcription factors, proteins involved in defense responses, binding or signaling, among other target proteins of differentially expressed MIR genes, had their regulator miRNAs identified. The vsiRNAs can act in the regulation of many host mRNAs to increase susceptibility to viral infection, and possibly acting as markers for activation of host defense mechanisms. Validation of MIR gene expression by qPCR and of its targets through qPCR or degradome analysis will help us to understand better the molecular interaction mechanisms of begomoviruses with their hosts.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Science and Technology in Plant-Pest Interactions, CAPES (doctoral fellowship to MFB), CNPq (grant 483607/2013-4 to FMZ) and FAPEMIG (grant APQ-00949-09 to FMZ).

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM & Sherlock G (2000) Gene Ontology: tool for the unification of biology. Nature Genetics 25:25-29.

Allen E, Xie Z, Gustafson AM & Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121:207-221.

Amin I, Patil BL, Briddon RW, Mansoor S & Fauquet CM (2011) Comparison of phenotypes produced in response to transient expression of genes encoded by four distinct begomoviruses in Nicotiana benthamiana and their correlation with the levels of developmental miRNAs. Virology Journal 8:1-24.

Anders S & Huber W (2010) Differential expression analysis for sequence count data. Genome Biology 11:R106.

Andrade EC, Manhani GG, Alfenas PF, Calegario RF, Fontes EPB & Zerbini FM (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. Journal of General Virology 87:3687-3696.

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Antunes TFS (2011) Viral determinants involves in the differential adaptation of two begomovirus in tomato and Nicotiana benthamiana. MSc Dissertation, Dep. de Bioquímica Agrícola. Universidade Federal de Viçosa, Viçosa, MG. 73p.

Baliji S, Lacatus G & Sunter G (2010) The interaction between geminivirus pathogenicity proteins and adenosine kinase leads to increased expression of primary cytokinin responsive genes. Virology 402:238-247.

Bazzini AA, Almasia NI, Manacorda CA, Mongelli VC, Conti G, Maroniche GA, Rodriguez MC, Distefano AJ, Hopp HE, Del Vas M & Asurmendi S (2009) Virus infection elevates transcriptional activity of miR164a promoter in plants. BMC Plant Biology 9:152.

Bazzini AA, Hopp HE, Beachy RN & Asurmendi S (2007) Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. Proceedings of the National Academy of Sciences, USA 104:12157-12162.

Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B 57:289-300.

Bisaro DM (2006) Silencing suppression by geminivirus proteins. Virology 344:158-168.

Briddon RW (2003) Cotton leaf curl disease, a multicomponent begomovirus complex. Molecular Plant Pathology 4:427-434.

Bousios A, Minga E, Kalitsou N, Pantermali M, Tsaballa A & Darzentas N (2012) MASiVEdb: the Sirevirus Plant Retrotransposon Database. BMC Genomics 13:1-10.

Brown JK, Fauquet CM, Briddon RW, Zerbini FM, Moriones E, Navas-Castillo J (2012) Family Geminiviridae. pp. 351-373, In: King AMQ, Adams MJ, Carstens EBE & Lefkowitz EJ (Ed.). Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. London, UK: Elsevier Academic Press, 2012.

Brustolini OJ, Machado JP, Condori-Apfata JA, Coco D, Deguchi M, Loriato VA, Pereira WA, Alfenas-Zerbini P, Zerbini FM, Inoue-Nagata AK, Santos AA, Chory J, Silva FF & Fontes EP (2015) Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. Plant Biotechnology Journal doi: 10.1111/pbi.12349.

Carmo LST, Resendo RO, Silva LP, Ribeiro SG & Mehta A (2013) Identification of host proteins modulated by the virulence factor AC2 of Tomato chlorotic mottle virus in Nicotiana benthamiana. Proteomics 13:1947-1960.

Chapman EJ & Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. Nature Reviews Genetics 8:884-896.

Chen, J., Hu, Q., Zhang, Y., Lu, C. and Kuang, H. (2013) P-MITE: a database for plant miniature inverted-repeat transposable elements. Nucleic Acids Research 42:D1176-D1181.

Chung HY, Lacatus G & Sunter G (2014) Geminivirus AL2 protein induces expression of, and interacts with, a calmodulin-like gene, an endogenous regulator of gene silencing. Virology 460:108-118.

Dai X & Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Research 39:W155-W159.

Doyle JJ & Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15.

McCarthy FM, Wang N, Magee GB, Nanduri B, Lawrence ML, Camon EB, Barrell DG, Hill DP, Dolan ME, Williams WP, Luthe DS, Bridges SM & Burgess SC (2006) AgBase: a functional genomics resource for agriculture. BMC Genomics 7:229.

Goodin MM, Zaitlin D, Naidu RA & Lommel SA (2008) Nicotiana benthamiana: its history and future as a model for plant-pathogen interactions. Molecular Plant-Microbe Interactions 21:1015-1026.

Hardcastle TJ & Kelly KA (2010) BaySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 11:1-14.

Hamilton AJ & Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950-951.

Hammond SM, Bernstein E, Beach D & Hannon GJ (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404:293-296.

Hanley-Bowdoin L, Bejarano ER, Robertson D & Mansoor S (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. Nature Reviews Microbiology 11:777-788.b

Hao L, Wang H, Sunter G & Bisaro DM (2003) Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. Plant Cell 15:1034-1048.

Jay F, Wang Y, Yu A, Taconnat L, Pelletier S, Colot V, Renou JP & Voinnet O (2011) Misregulation of AUXIN RESPONSE FACTOR 8 underlies the developmental abnormalities caused by three distinct viral silencing suppressors in Arabidopsis. PLoS Pathogens 7:e1002035.

Jones-Rhoades MW, Bartel DP & Bartel B (2006) MicroRNAs and their regulatory roles in plants. Annual Review of Plant Biology 57:19-53.

Kozomara A & Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Research 42:D68-D73.

Langmead B, Trapnell C, Pop M & Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10:R25.

Lee YS, Shibata Y, Malhotra A & Dutta A (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes & Development 23:2639-2649.

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 25:2078-2079.

Li F, Huang C, Li Z & Zhou X (2014) Suppression of RNA silencing by a plant DNA virus satellite requires a host calmodulin-like protein to repress RDR6 expression. PLoS Pathogens 10:e1003921:1-19.

Liu, Chung HY, Lacatus G, Baliji S, Ruan J & Sunter G (2014) Altered expression of Arabidopsis genes in response to a multifunctional geminivirus pathogenicity protein. BMC Plant Biology 14:302

Loss-Morais G, Waterhouse PM & Margis R (2013) Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets. Biology Direct 8:1-5.

Ma X, Xin Z, Wang Z, Yang Q, Guo S, Guo X, Cao L, Lin T (2015) Identification and comparative analysis of differentially expressed miRNAs in leaves of two wheat (Triticum aestivum L.) genotypes during dehydration stress. BMC Plant Biology 15:21.

Mallory AC & Vaucheret H (2006) Functions of microRNAs and related small RNAs in plants. Nature Genetics 38:S31-S36.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ & Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. Cell 133:116-127.

Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD & Marshall D (2013) Using Tablet for visual exploration of second-generation sequencing data. Briefings in Bioinformatics 14:193-202.

Miozzi L, Napoli C, Sardo L & Accotto GP (2014) Transcriptomics of the interaction between the monopartite phloem-limited geminivirus tomato yellow leaf curl Sardinia virus and Solanum lycopersicum highlights a role for plant hormones, autophagy and plant immune system fine tuning during infection. PLoS One 9:e89951.

Mlotshwa S, Pruss GJ & Vance V (2008) Small RNAs in viral infection and host defense. Trends in Plant Science 13:375-382.

Moffatt BA, Stevens YY, Allen MS, Snider JD, Pereira LA, Todorova MI, Summers PS, Weretilnyk EA, Martin-McCaffrey L & Wagner C (2002) Adenosine kinase deficiency is associated with developmental abnormalities and reduced transmethylation. Plant Physiology 128:812-821.

Monci F, Sanchez-Campos S, Navas-Castillo J & Moriones E (2002) A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303:317-326.

Morales FJ & Anderson PK (2001) The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Archives of Virology 146:415-441.

Ouyang S & Buell CR (2004) The TIGR Plant Repeat Databases: a collective resource for identification of repetitive sequences in plants. Nucleic Acids Research 32:D360-D363.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & r Glockner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41:D590-D596.

Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Rodríguez-Negrete E, Lozano-Durán R, Piedra-Aguilera A, Cruzado L, Bejarano ER & Castillo AG (2013) Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytologist 199:464-475.

Rojas MR, Gilbertson RL, Russell DR & Maxwell DP (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminivirus. Plant Disease 77:4:340-347.

Rojas MR, Hagen C, Lucas WJ & Gilbertson RL (2005) Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. Annual Review of Phytopathology 43:361-394.

Romanel E, Silva TF, Corrêa RL, Farinelli L, Hawkins JS, Schrago CEG, Vaslin MFS (2012) Global alteration of microRNAs and transposon-derived small RNAs in cotton (Gossypium hirsutum) during Cotton leafroll dwarf polerovirus (CLRDV) infection. Plant Molecular Biology 80:443-460.

Romualdi C, Bortoluzzi S, D'Alessi F & Danieli GA (2003) IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. Physiological Genomics 12:159-162.

Rutledge RG & Cote C (2003) Mathematics of quantitative kinetic PCR and the application of standard curves. Nucleic Acids Research 31:e93:1-6.

Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51:492-508.

Schmieder R & Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863-864.

Silva TF, Romanel EAC, Andrade RRS, Farinelli L, Osteras M, Deluen C, Corrêa RL, Schrago CEG & Vaslin MFS (2011) Profile of small interfering RNAs from cotton plants infected with the polerovirus Cotton leafroll dwarf virus. BMC Molecular Biology 12:1-12.

Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T & Moulton V (2012) The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. Bioinformatics 28:2059-2061.

Storey J & Tibshirani R (2003) Statistical significance for genomewide studies. PNAS USA 100:9440-9445.

Thieme CJ, Schudoma C, May P & Walther D (2012) Give It AGO: The search for miRNA-argonaute sorting signals in Arabidopsis thaliana indicates a relevance of sequence positions other than the 5′-position alone. Frontiers in Plant Scince 3:1-15.

Takeda A, Iwasaki S, Watanabe T, Utsumi M & Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiology 49:493-500.

Trinks D, Rajeswaran R, Shivaprasad PV, Akbergenov R, Oakeley EJ, Veluthambi K, Hohn T & Pooggin MM (2005) Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. Journal of Virology 79:2517-2527.

Tuschl T, Zamore PD, Lehmann R, Bartel DP & Sharp PA (1999) Targeted mRNA degradation by double-stranded RNA in vitro. Genes & Development 13:3191-3197.

Vanitharani R, Chellappan P, Pita JS & Fauquet CM (2004) Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. Journal of Virology 78:9487-9498.

Varsani A, Navas-Catillo J, Mariones E, Hernández-Zepeda C, Idris A, Brown JK, Zerbini FM & Martin DP (2014) Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. Archives of Virology 159:2193-2203.

Visser M, Maree HJ, Rees DJG & Burger JT (2014) High-throughput sequencing reveals small RNAs involved in ASGV infection. BMC Genomics 15:1-10.

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136:669-687.

Xavier AS (2012) Effect of gene silencing of dnaJ and TCTP in the infection of tomato and Nicotiana benthamiana by the potyvirus Pepper yellow mosaic virus (PepYMV). MSc Dissertation, Dep. de Fitopatologia. Universidade Federal de Viçosa, Viçosa, MG. 74p.

Yi X, Zhang Z, Ling Y, Xu W & Su Z (2015) PNRD: a plant non-coding RNA database. Nucleic Acids Research 43:D982-D989.

Ye R, Wang W, Iki T, Wu Y, Ishikawa M, Zhou X & Qi Y (2012) Cytoplasmic assembly and selective nuclear import of Arabidopsis ARGONAUTE4/siRNA complexes. Molecular Cell l 46:859-870.

Wang L, Feng Z, Wang X & Zhang X (2010) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26:136-138.

Wang H, Buckley KJ, Yang X, Buchmann RC & Bisaro DM (2005) Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. Journal of Virology 79:7410-7418.

Wang H, Hao L, Shung C-Y, Sunter G & Bisaro DM (2003) Adenosine Kinase is inactivated by geminivirus AL2 and L2 proteins. Plant Cell 15:3020-3032.

Wang X-B, Wu Q, Ito T, Cillo F, Li W-X, Chen X, Y J-L & Ding S-W (2010) RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. Proceedings of the National Academy of Sciences, USA 107:484-489.

Were HK, Winter S & Maiss E (2004) Viruses infecting cassava in Kenya. Plant Disease 88:17-22.

Zorzatto C, Machado JP, Lopes KV, Nascimento KJ, Pereira WA, Brustolini OJ, Reis PA, Calil IP, Deguchi M, Sachetto-Martins G, Gouveia BC, Loriato VA, Silva MA, Silva FF, Santos AA, Chory J, Fontes EP (2015) NIK1-mediated translation suppression functions as a plant antiviral immunity mechanism. Nature doi: 10.1038/nature14171

Table 1. Summary of results obtained after computational data mining of smRNA libraries from the interaction Nicotiana benthamina infected with Tomato yellow spot virus (ToYSV), at 5 and 10 days post-inoculation (dpi).

Parameters	**5dpi-Mock-R1**	**5dpi-ToYSV-R1**	**10dpi-Mock-R1**	**10dpi-ToYSV-R1**	**5dpi-Mock-R2**	**5dpi-ToYSV-R2**	**10dpi-Mock-R2**	**10dpi-ToYSV-R2**
Total number reads in each smRNAseq libraries	43.043.745	20.072.686	27.102.684	28.428.364	33.747.974	40.163.941	44.803.828	52.624.670
Clusters that fulfill the default Illumina quality criteria (Q20 %)	89,72	90,52	90,46	89,95	98,67	98,77	98,82	98,49
Bases (passed filter) with a quality score greater or equal to 30 (Q30 %)	94,99	94,91	94,39	94,52	96,76	97,00	97,20	96,34
Number of filtered and trimmed reads of 18-28 nt	31.291.645	14.940.632	19.749.226	19.364.051	11.807.414	25.776.223	28.884.076	26.266.555
Number of unique RPM[1] reads of 18-28 nt	152.147	172.469	153.120	193.467	265.240	168.852	310.374	232.794
Total number of ToYSV genome-matched RPM reads of 18-28 nt	12	19.588	10	44	90	353.319	192	48.192
Number of unique ToYSV genome-matched RPM reads of 18-28 nt	und.[2]	1.005	und.	31	und.	1.426	und.	1.057
Number of ToYSV-A/B genome-matched RPM reads	6/2	16.820/ 2.766	7/2	25/18	143/921	303.250/ 50.068	143/49	31.035/ 17.175
Number of ToYSV genome-matched unique RPM reads	und.	1.005	und.	31	und.	1.424	und.	1.057
Percentage of ToYSV-unmatched, tRNA-matched reads of 18-28 nt	1.71	1.49	1.61	2.86	1.59	1.44	1.50	1.94
Percentage of ToYSV-unmatched, rRNA-matched reads of 18-28 nt	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Number of tRNA-matched RPM reads of 18-28 nt	17.100	14.900	16.100	28.600	15.900	14.400	15.000	19.400
Number of rRNA-matched RPM reads of 18-28 nt	< 100	< 100	< 100	< 100	< 200	< 100	< 100	< 100

[1]RPM: reads per 1 million (normalization).
[2]und.: undetermined;

A

B

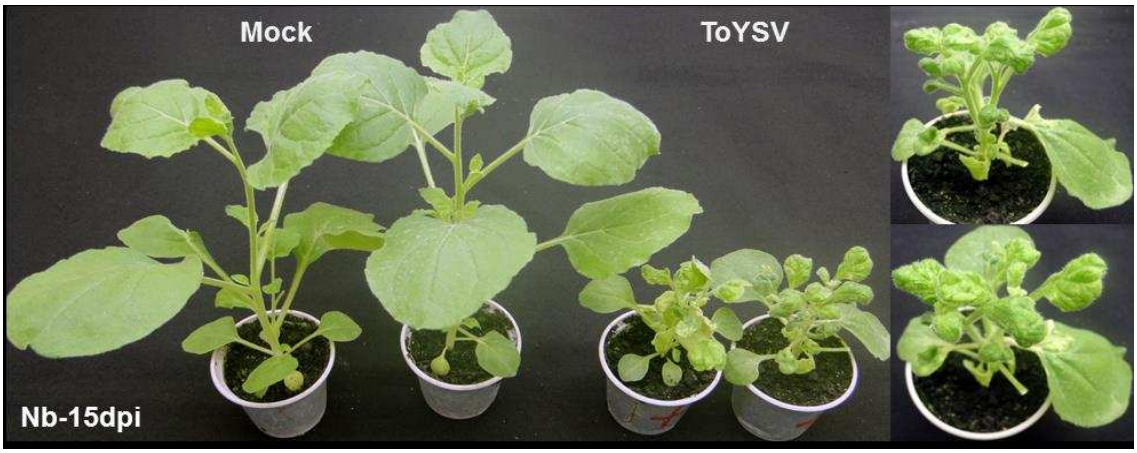
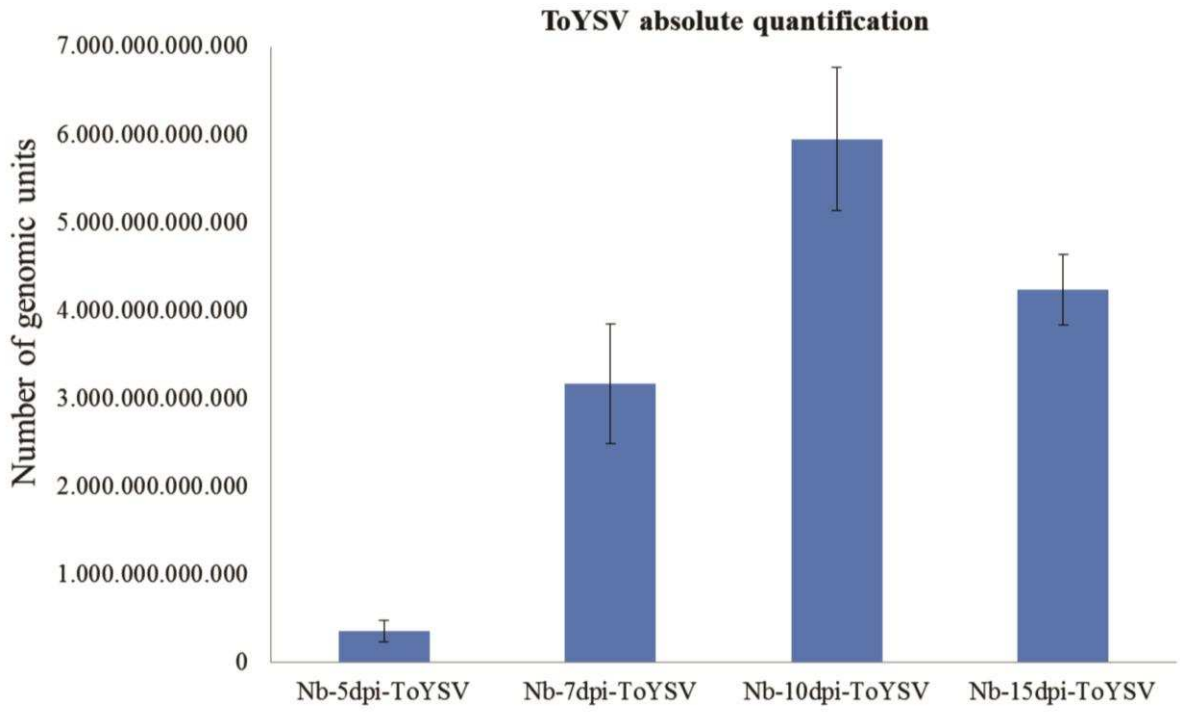


Figure 1. A. Mock- or ToYSV-inoculated Nicotiana benthamiana plants at 0, 5, 7, 10 and 15 days post inoculation (dpi). **B.** Symptoms in ToYSV-inoculated N. benthamiana plants at 15 dpi.

A



B

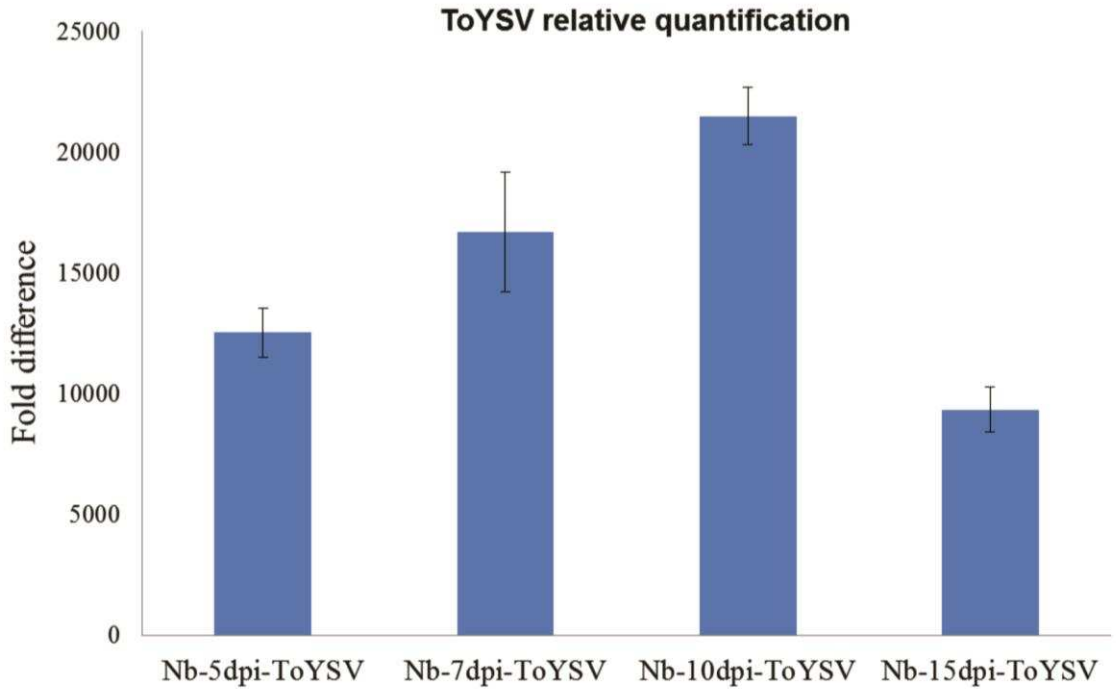
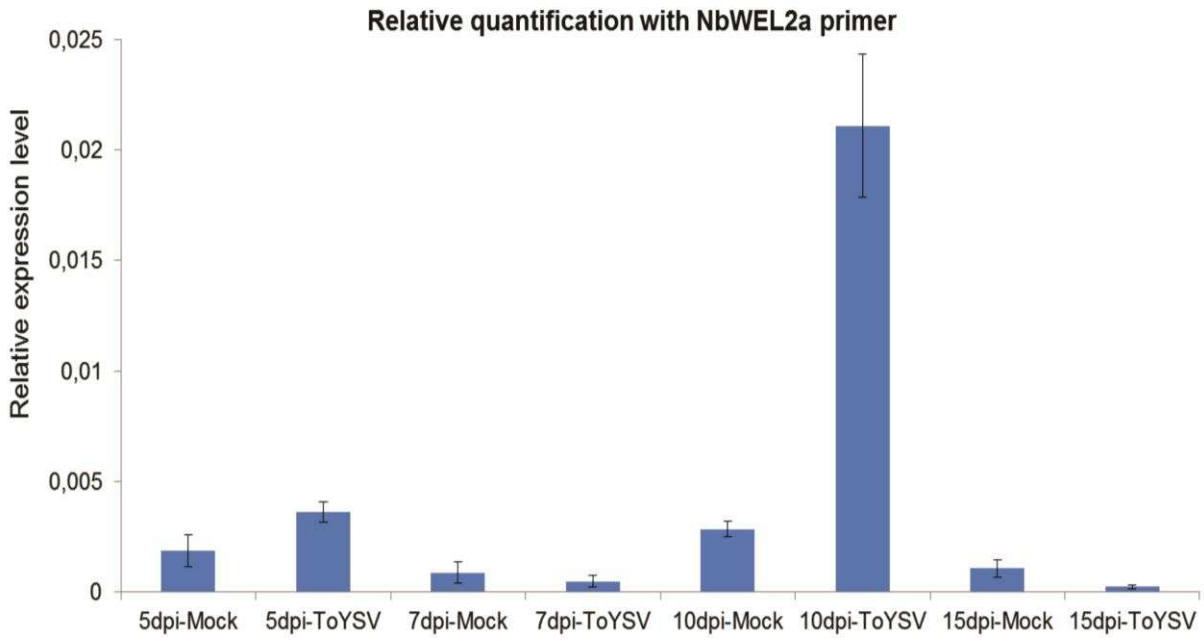


Figure 2. Quantification of ToYSV load by real-time PCR. **A.** Absolute and **B.** Relative ToYSV load in Nicotiana benthamiana at 5, 7, 10 and 15 days post inoculation (dpi).

A



B

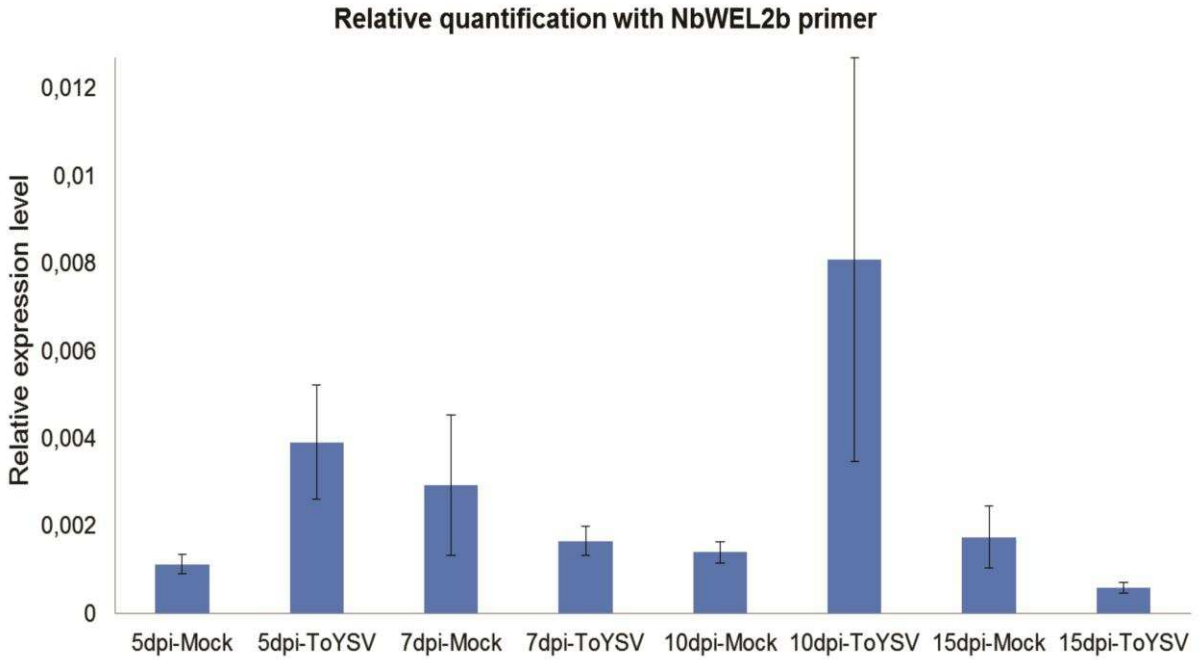


Figure 3. Relative quantification of the Werner syndrome-like exonuclease 2 (NbWEL2) gene of ToYSV-infected Nicotiana benthamiana at 5, 7, 10 and 15 days post inoculation (dpi), using **A.** NbWEL2a and **B.** NbWEL2b primers.

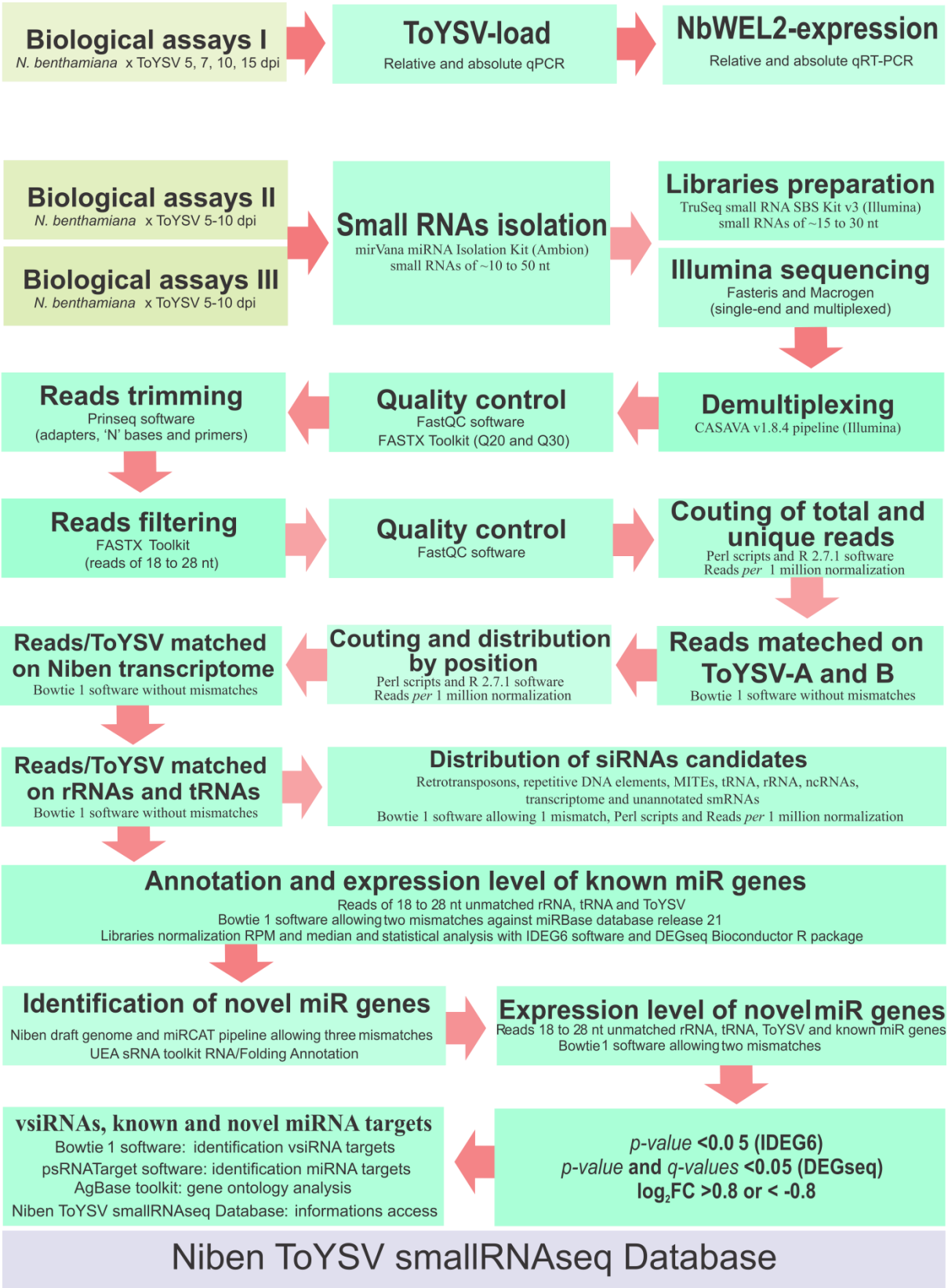
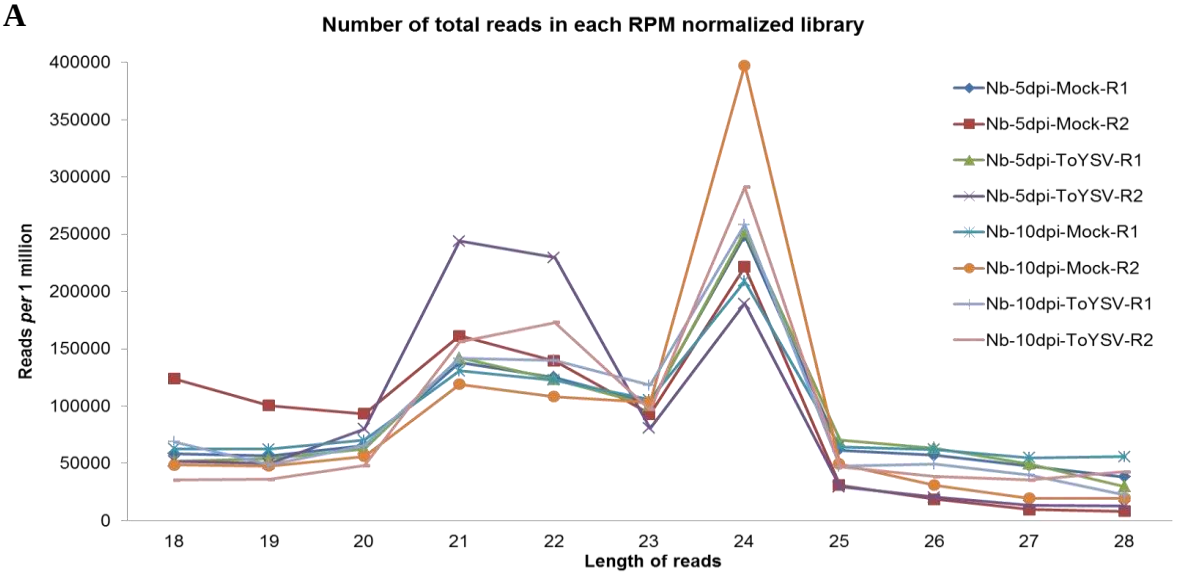


Figure 4. Workflow of the biological assays and smRNA sequence analysis.

A



B

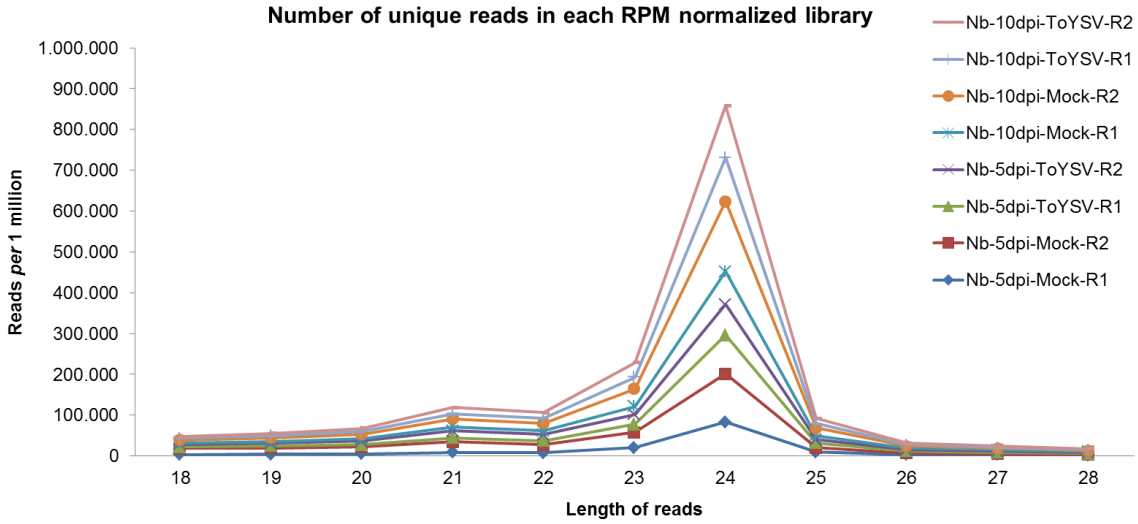
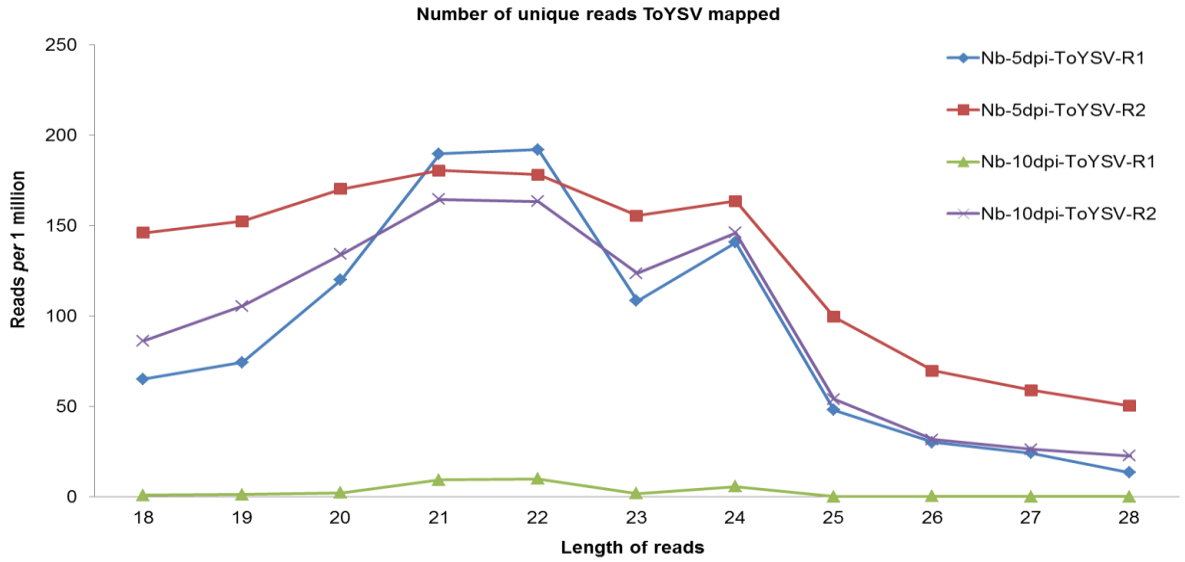


Figure 5. Size distribution of smRNA reads normalized as reads per 1 million (RPM). The total number of filtered reads for each library (Table 1) were used for normalization. **A.** Total 18-28 nucleotides (nt) reads in each library. **B.** Unique 18-28 nt reads in each library. **C.** Total 18-28 nt reads matched to the ToYSV genome. **D.** Unique 18-28 nt reads matched to the ToYSV genome. **E.** Total forward 18-28 nt reads matched to the ToYSV genome. **F.** Total reverse 18-28 nt reads matched to the ToYSV genome. **G, H.** Genome view of ToYSV DNA-A and DNA-B-derived 21-24 nt vsiRNAs recovered from infected N. benthamiana at 5 and 10 days post inoculation (dpi), normalized as RPM. On top, schematic representation of ToYSV DNA-A and DNA-B components. R1 and R2, biological replications 1 and 2, respectively. **I.** Percentage of vsiRNAs identified based on the 5'-terminal nucleotide. A, adenine; U, uracil; C, cytosine and G, guanine.

Figure 5 (cont.)

C



D

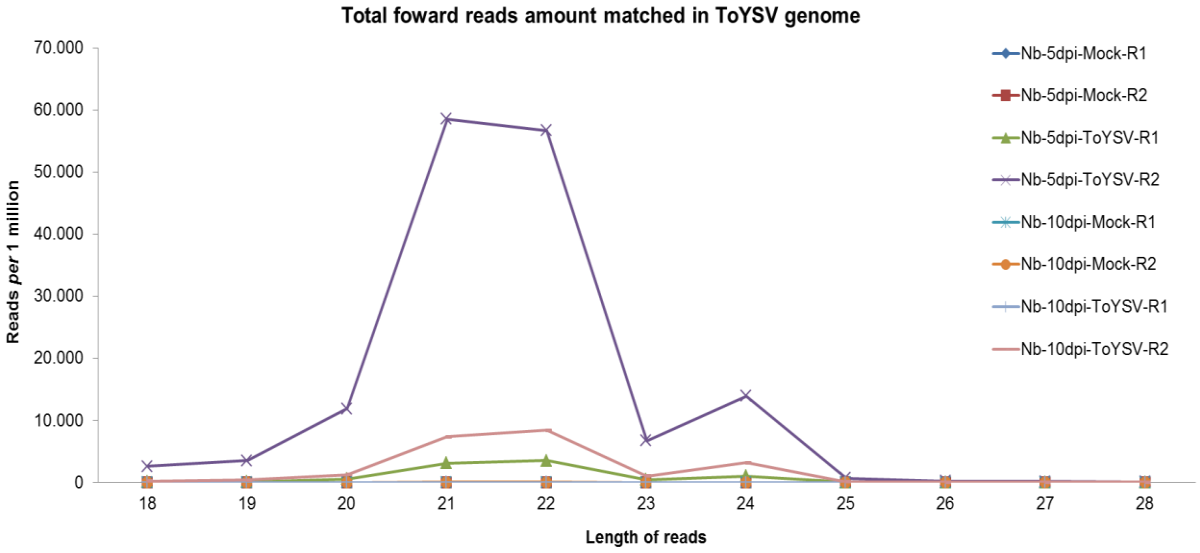
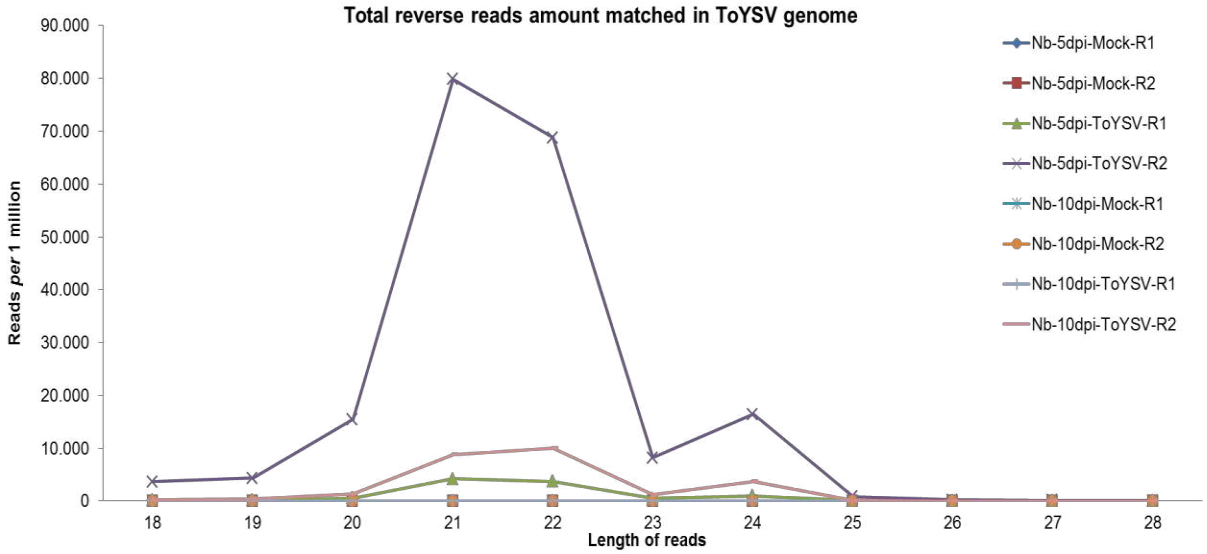


Figure 5 (cont.)

E



F

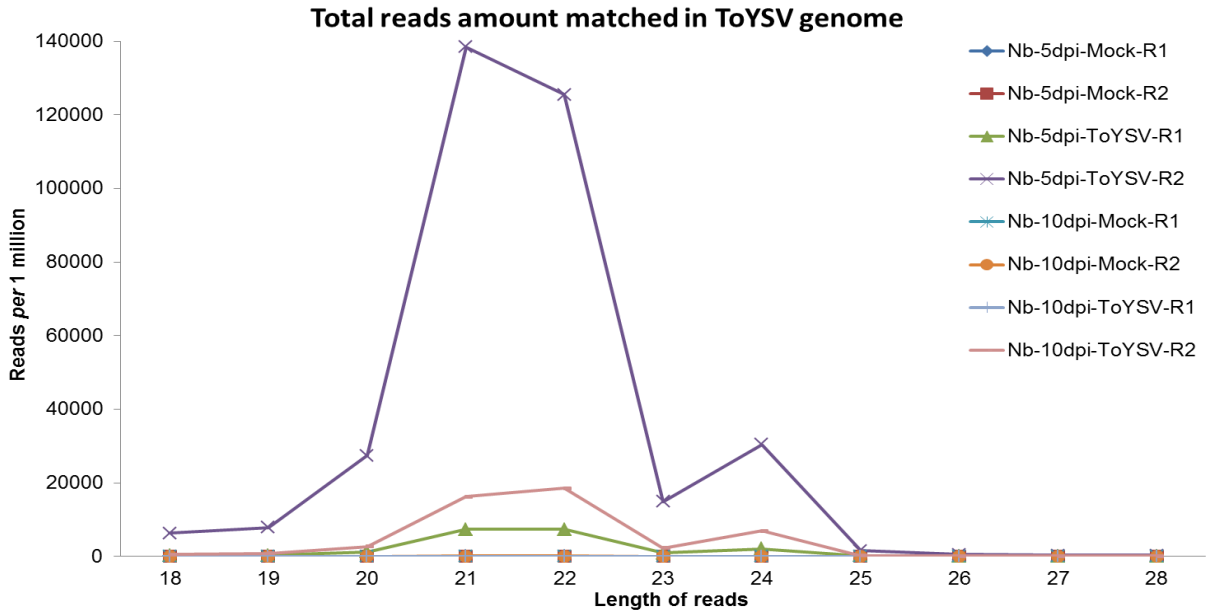


Figure 5 (cont.)

G

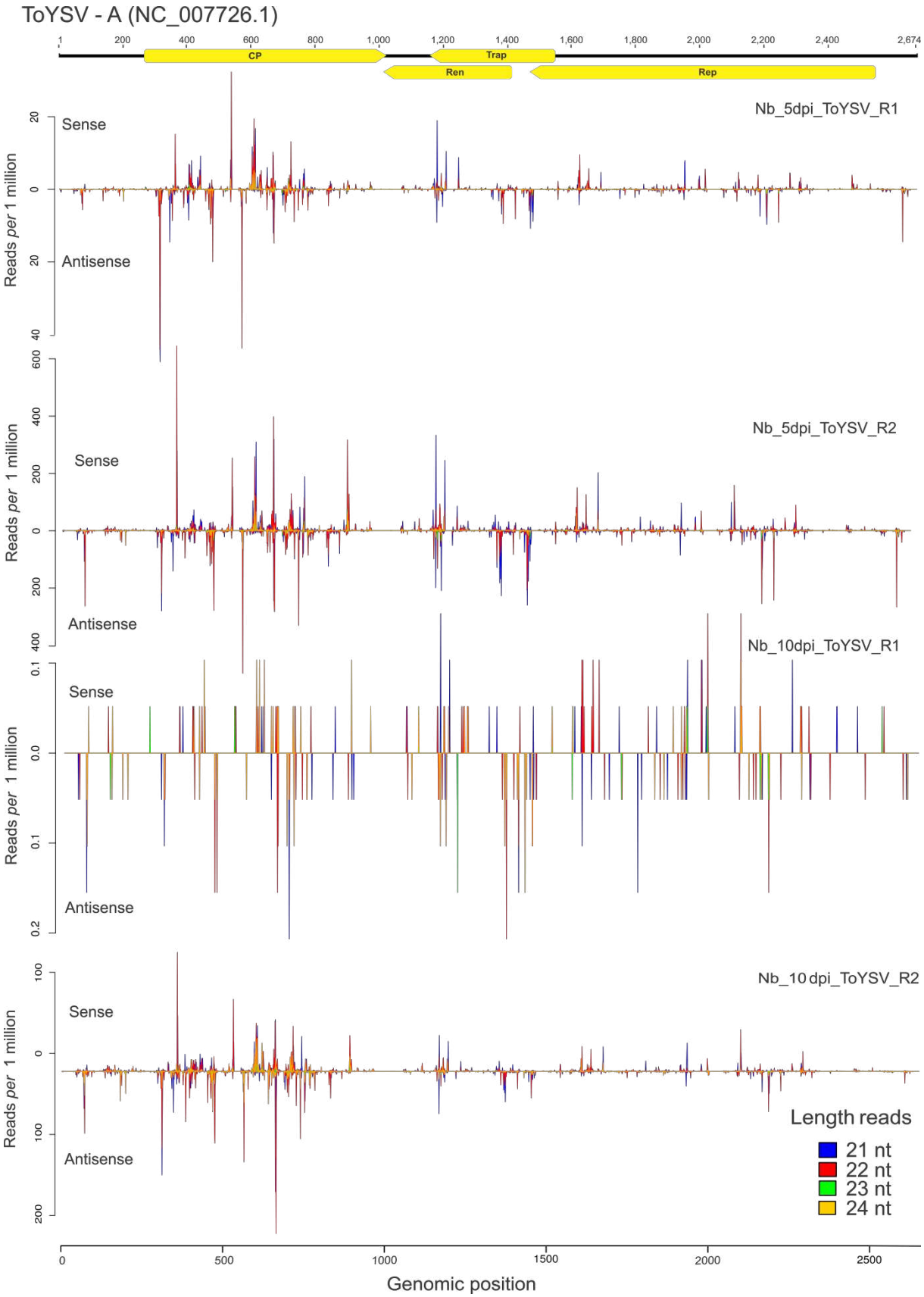


Figure 5 (cont.)

H

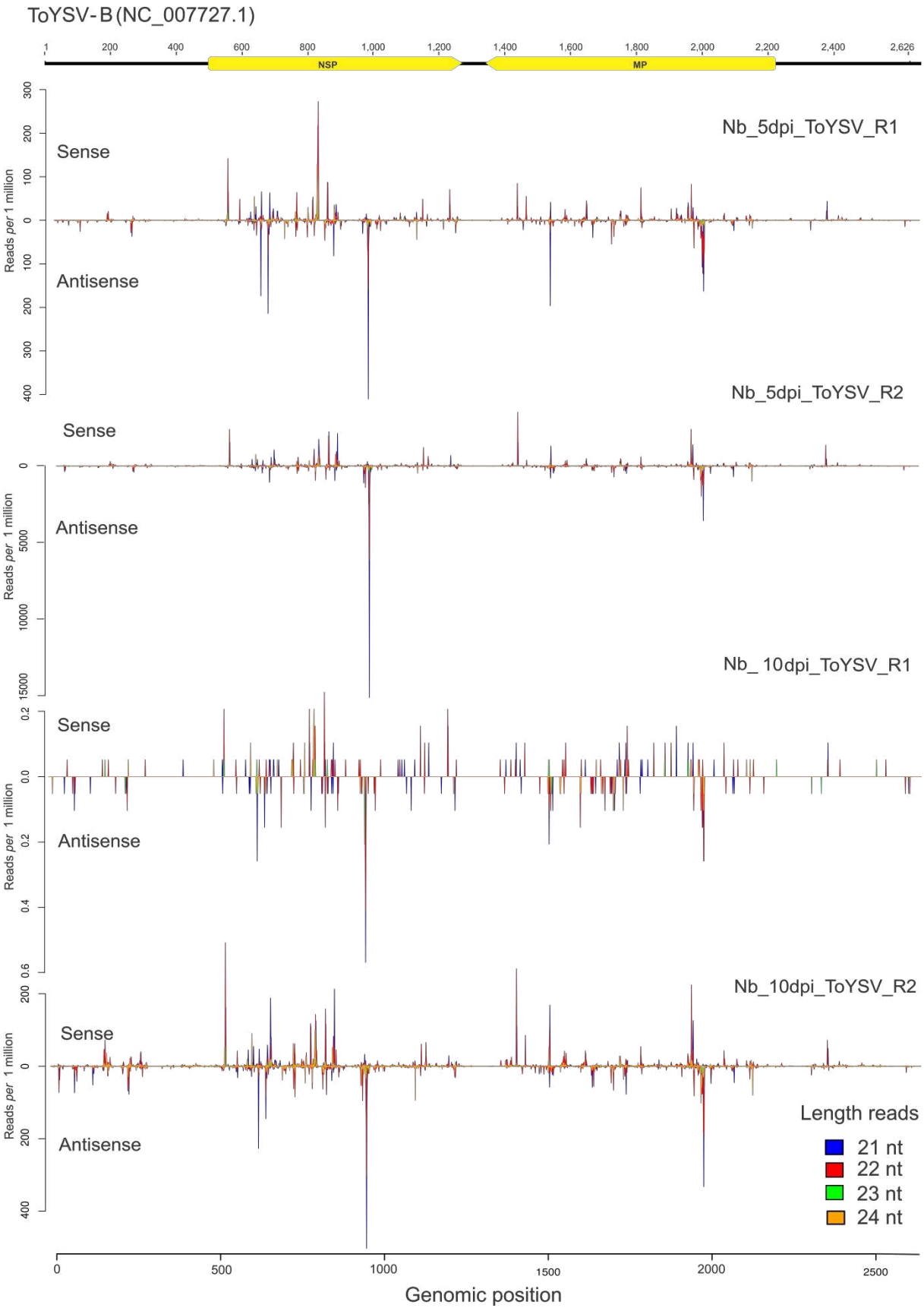
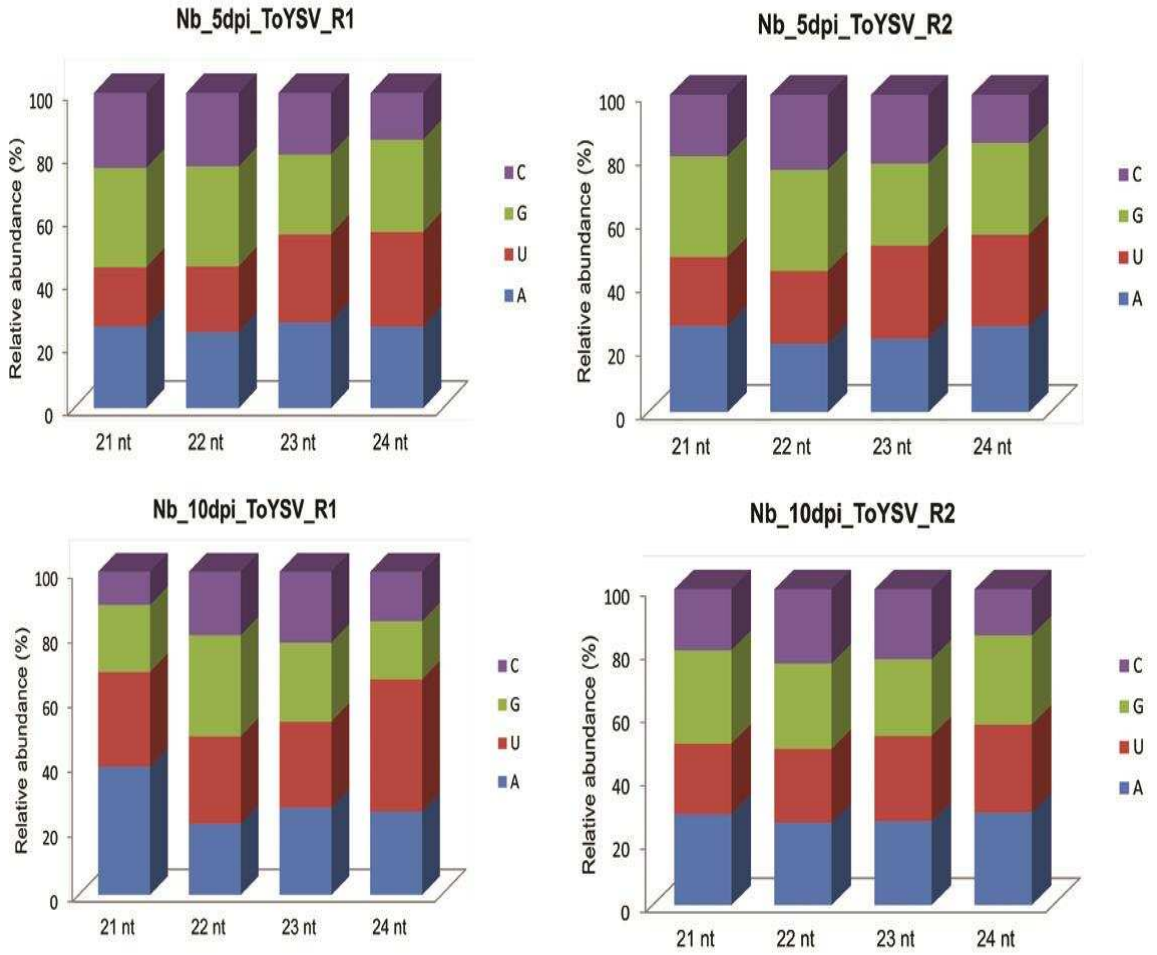


Figure 5 (cont.)

I



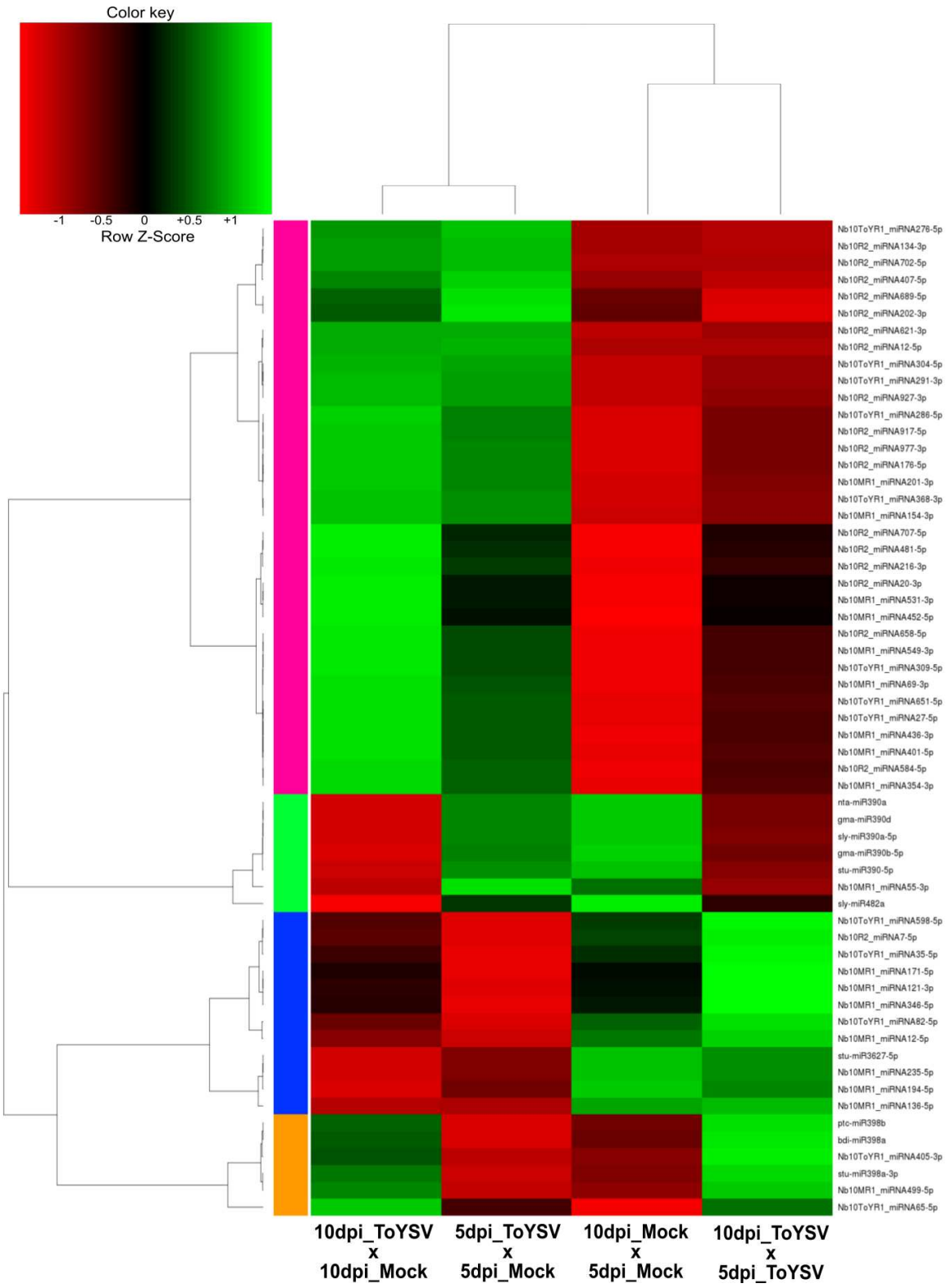


Figure 6. Heatmap of 59 differentially expressed genes common between 5dpi_ToYSV vs 5dpi_Mock and 10dpi_ToYSV vs 10dpi_Mock contrasts. Green shade indicates higher expression, red shade indicates lower expression and intermediate color indicates that MIR gene expression does not change at a statistically significant level. Color key indicates the intensity associated with normalized expression values.

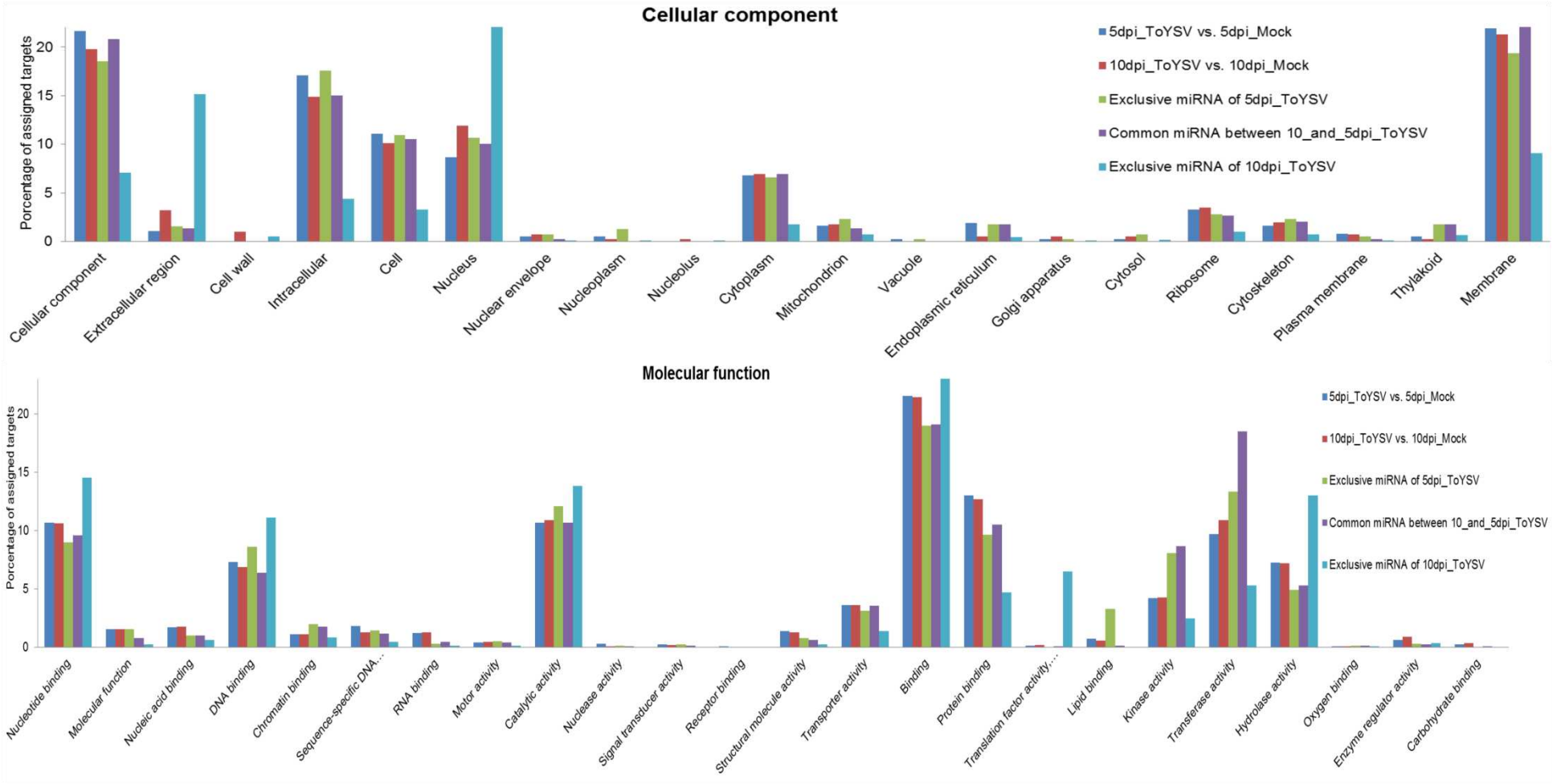
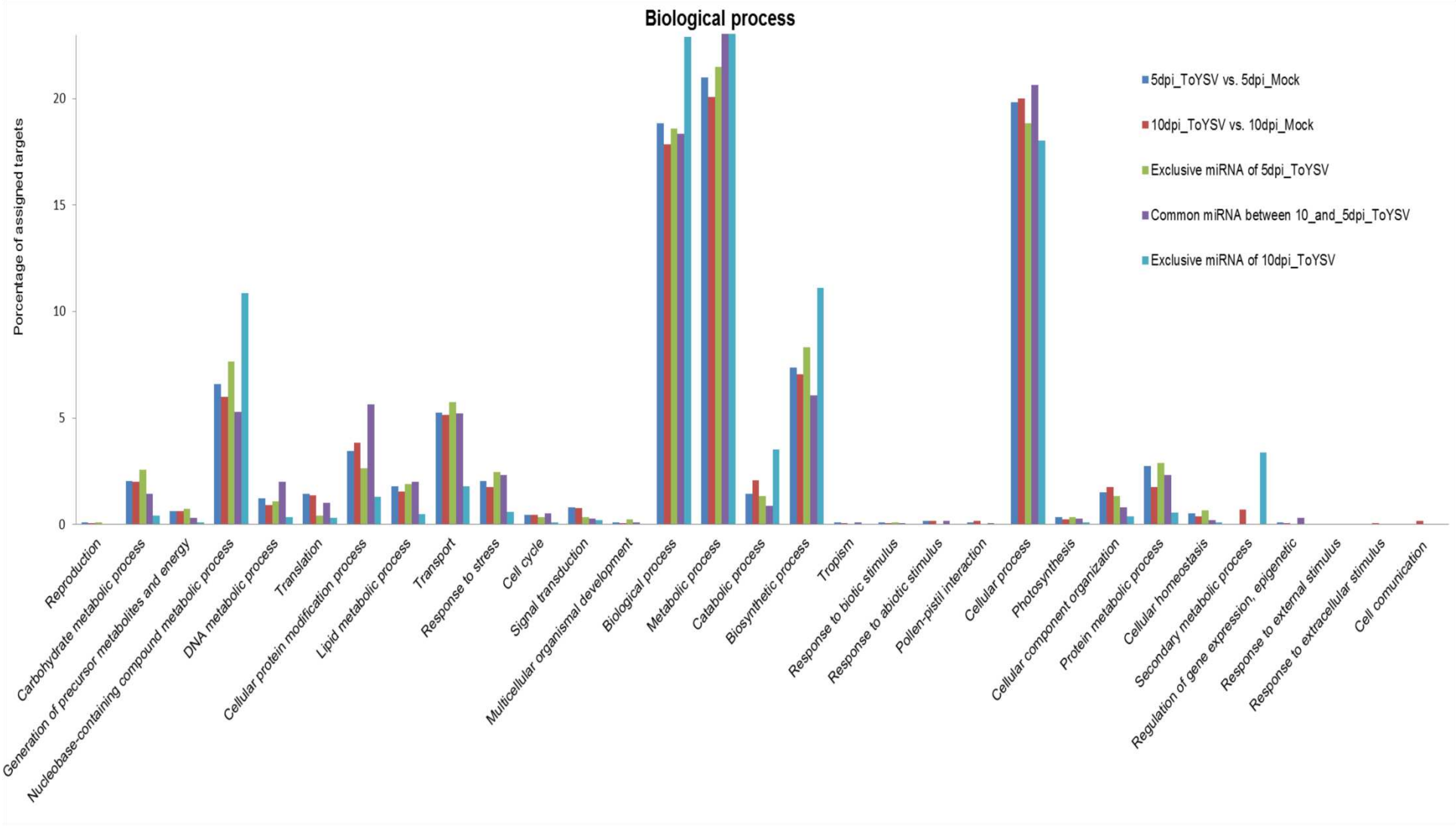


Figure 7. Gene ontology of the miRNAs target genes at 5dpi_ToYSV vs 5dpi_Mock, 10dpi_ToYSV vs 10dpi_Mock and 10dpi_ToYSV vs 5dpi_ToYSV contrasts, filtered against the 10dpi_Mock vs 5dpi_Mock contrast. Categorization of miRNA-target genes was performed according to the cellular component, molecular function and biological process categories. Bars indicate the percentage of total annotated targets genes mapping to GO terms.

Figure 7 (cont.)



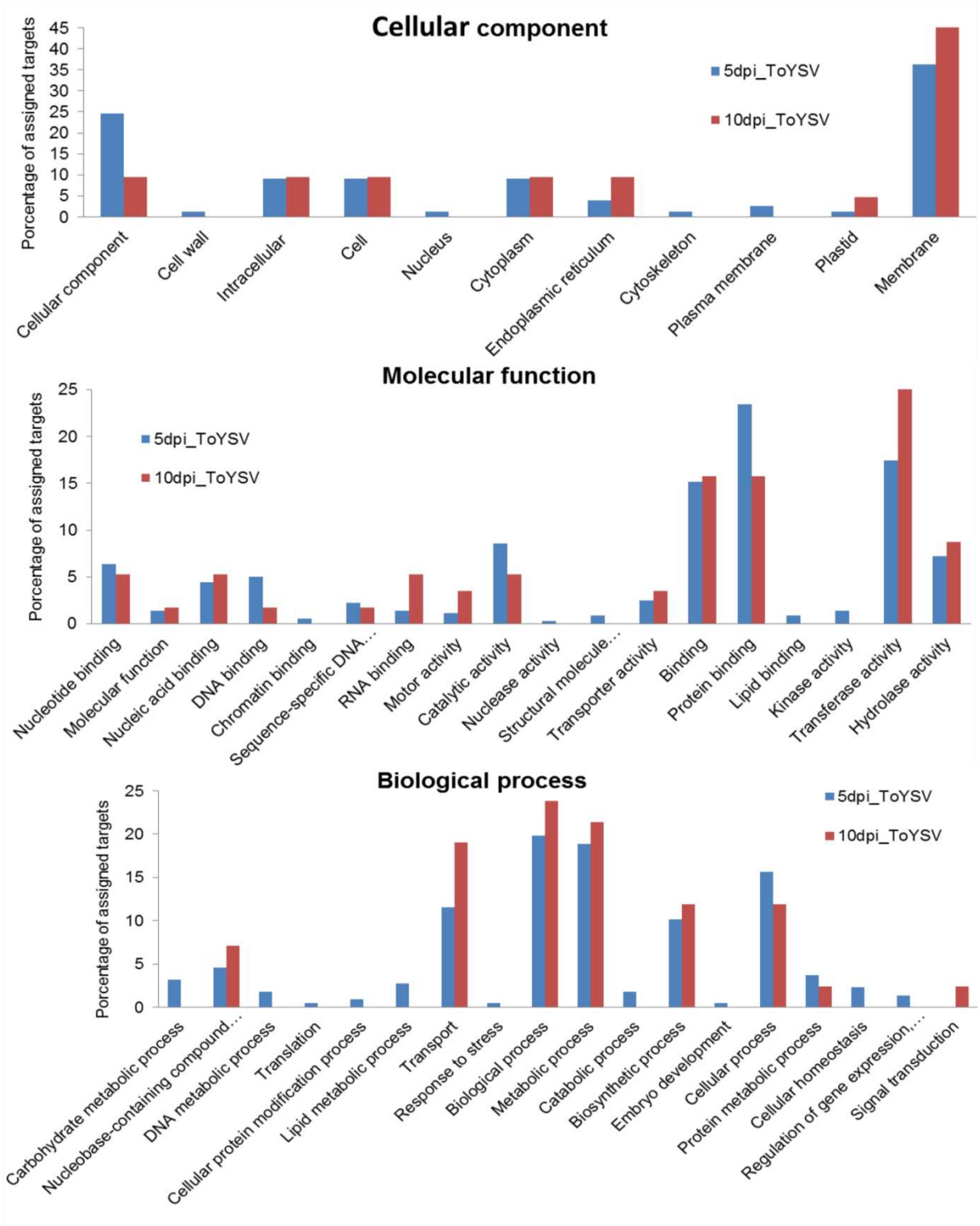


Figure 8. Gene ontology of the vsiRNAs target genes at 5dpi (left column) and 10 dpi (right column). Categorization of miRNA-target genes was performed according to the cellular component, molecular function and biological process categories. Bars indicate the percentage of total annotated targets genes mapping to GO terms. From these N. benthamiana genes identified as vsiRNAs target, 161 (5dpi) and 29 (10 dpi) genes do not have Gene Ontology annotation.

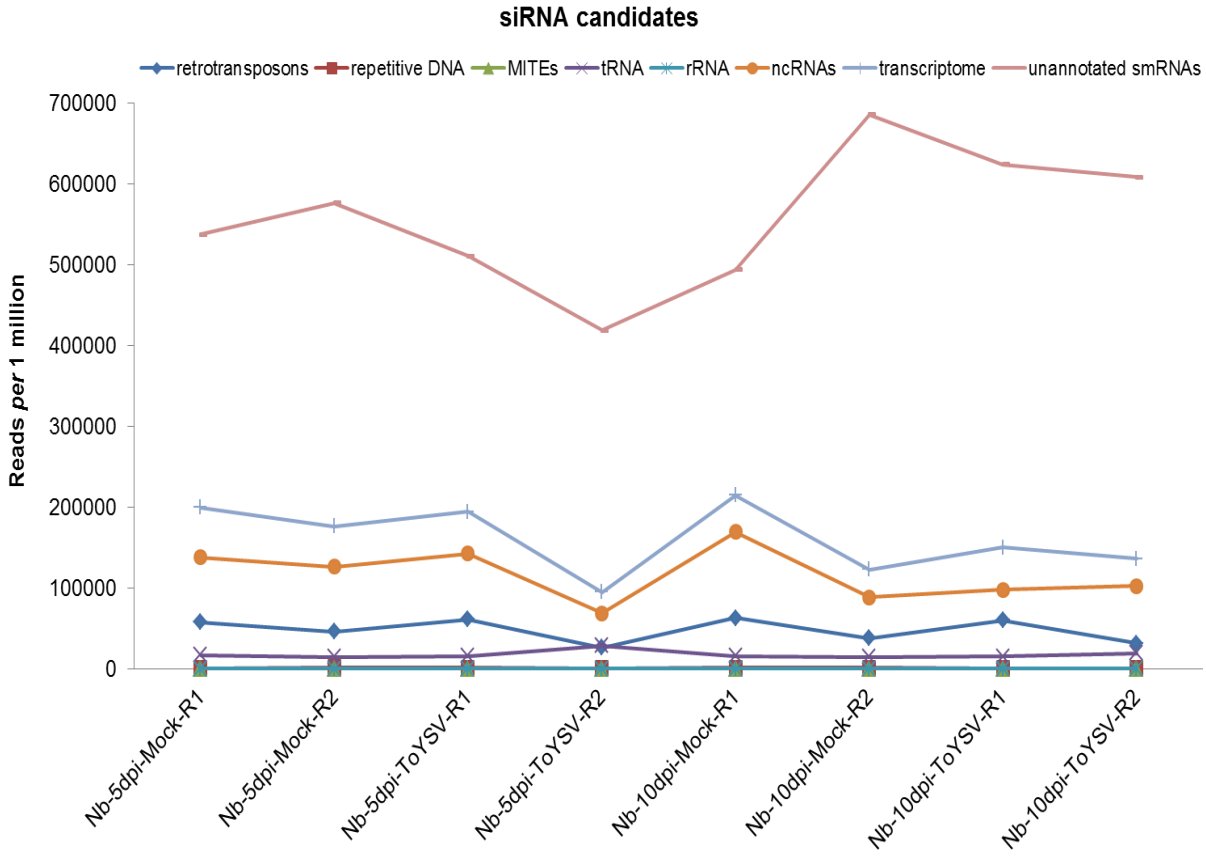


Figure 9. Distribution of siRNA candidates normalized in reads per 1 million (RPM) aligned in tandem. Total number of 18-28 nt reads in each library (Table 1) was used for normalization.

Suppl. Table S1. Sequence and amplicon characteristics of H2b, qToY3, NbWEL2 (Werner syndrome-like exonuclease 2 of Nicotiana benthamiana) primer pairs used for quantitative PCR.

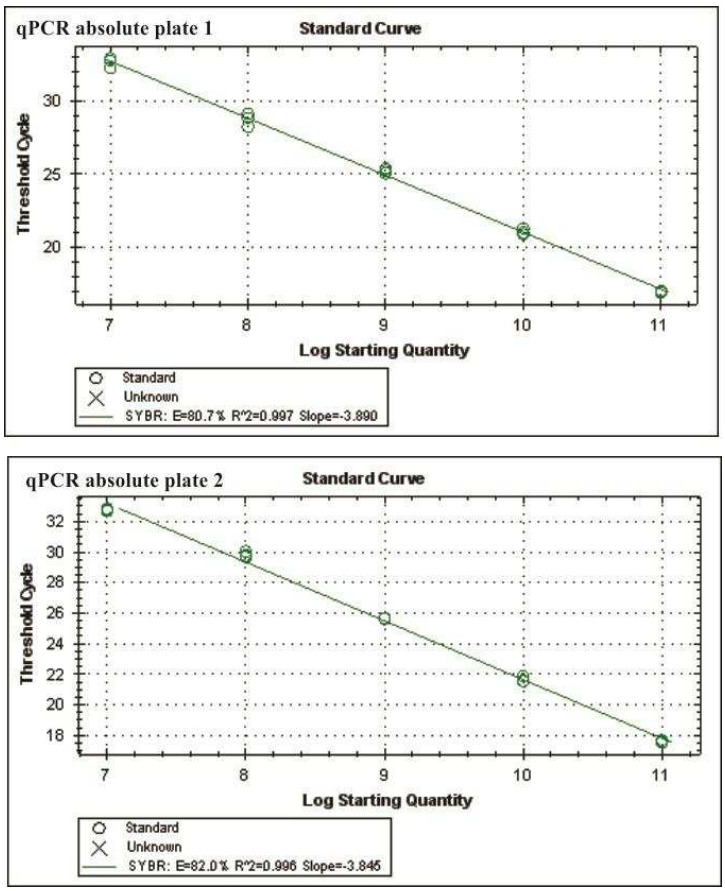
Primer pair	Gene	Forward primer sequence (5'- 3')	Reverse primer sequence (5'-3')	Size (bp)	Locus accession number	Reference
H2b	H2b histone	GAGAAGCTGGCTCAGGAATC	CACAAGCCTCACAGCAGTCT	93	EF189156	Xavier, 2012
qToY3	ToYSV DNA-A	CCACGATTTTAAAGCTGCATTCT	CAATCCTGGTGAGGGAGTCAGT	70	NC_007726	Antunes, 2011
NbWEL2a	NbWEL2	CAACCAGTTGCTCTTCTCCA	GGGCATGGAATCTTTGTGTA	78	NbS00007831g0002	This work
NbWEL2b	NbWEL2	AATTGTTCGATGAAATGGCA	TTCAATGGTTGATAATGGCAA	82	NbS00007831g0002	This work

.

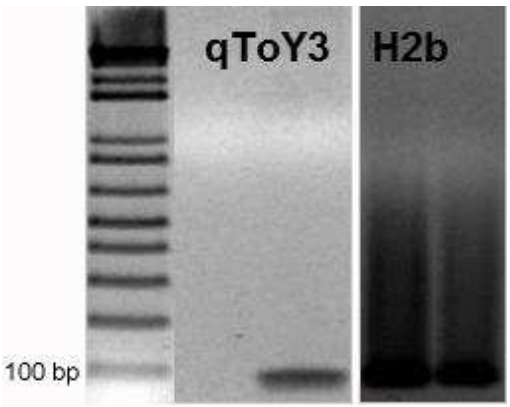
Suppl. Table S2. Quantification of smRNAs isolated from mock- and ToYSV-inoculated N. benthamiana plants at 5 and 10 days post inoculation. R1 and R2, experimental replications 1 and 2, respectively.

Library	Concentration (ng/µl)	A260/280	Total amount (µg)
5dpi-Mock-R1	71,6	1,98	6,87
5dpi-ToYSV-R1	68,2	1,95	6,55
10dpi-Mock-R1	66,9	1,91	6,42
10dpi-ToYSV-R1	67,6	2,00	6,49
5dpi-Mock-R2	71,6	1,92	6,87
5dpi-ToYSV-R2	68,2	1,98	6,55
10dpi-Mock-R2	66,9	2,01	6,42
10dpi-ToYSV-R2	67,6	1,96	6,49

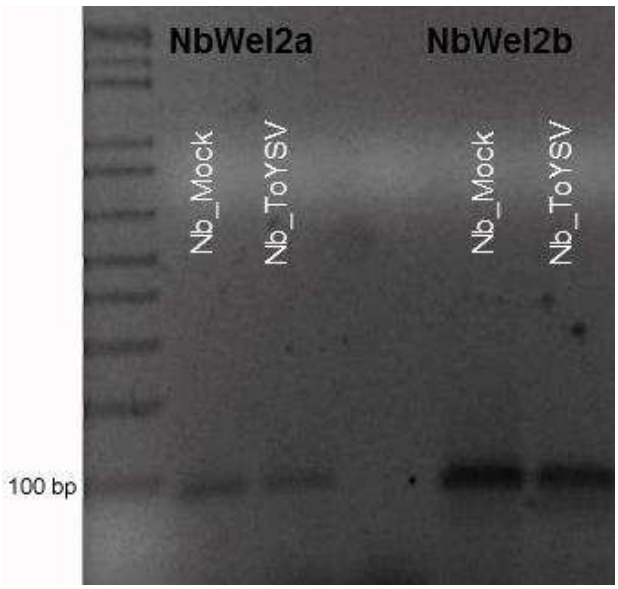
Suppl. Figure S1. Standard curve for absolute quantification of ToYSV load.



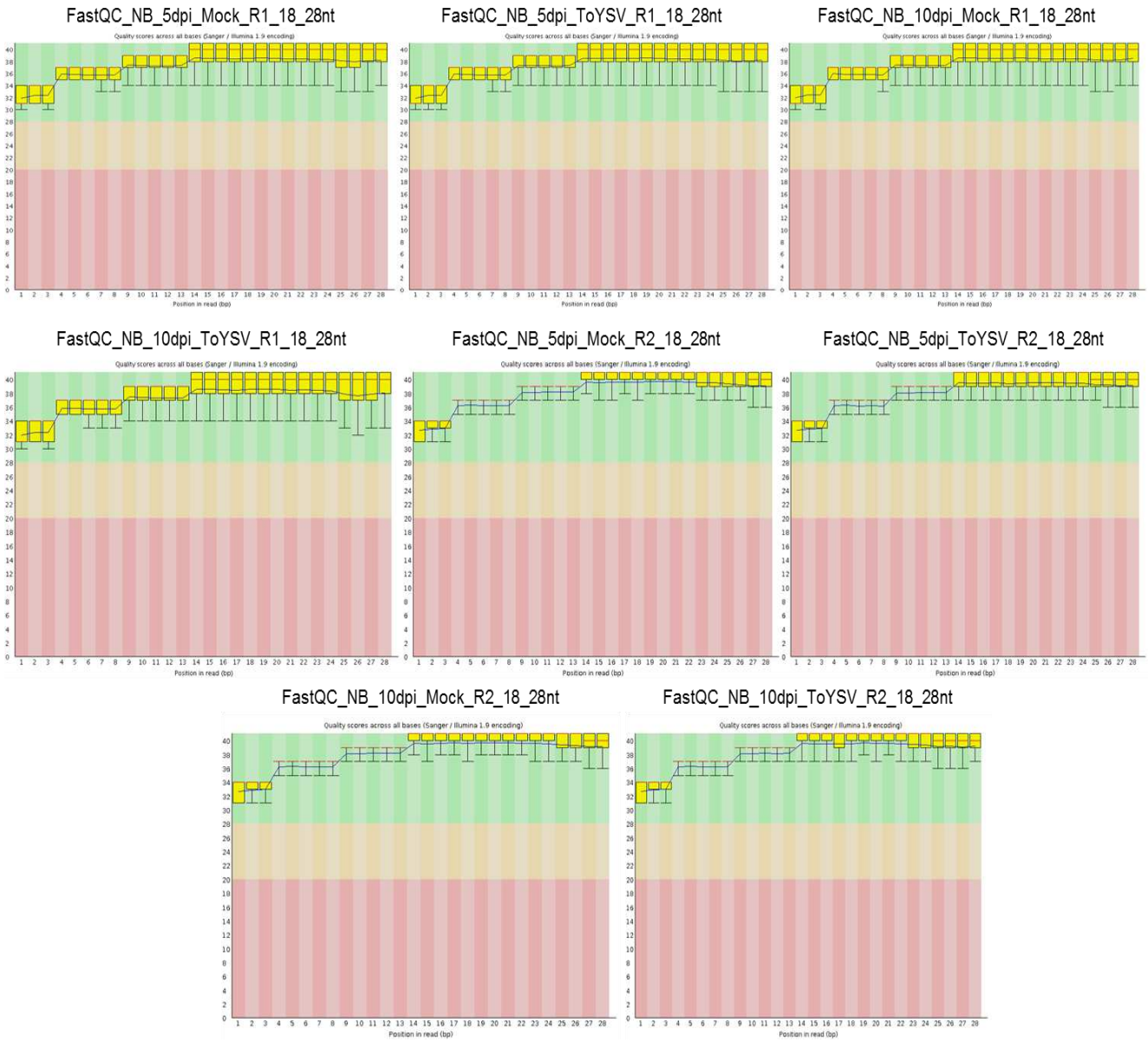
Suppl. Figure S2. Specificity tests for the qToY3 and H2b primers.



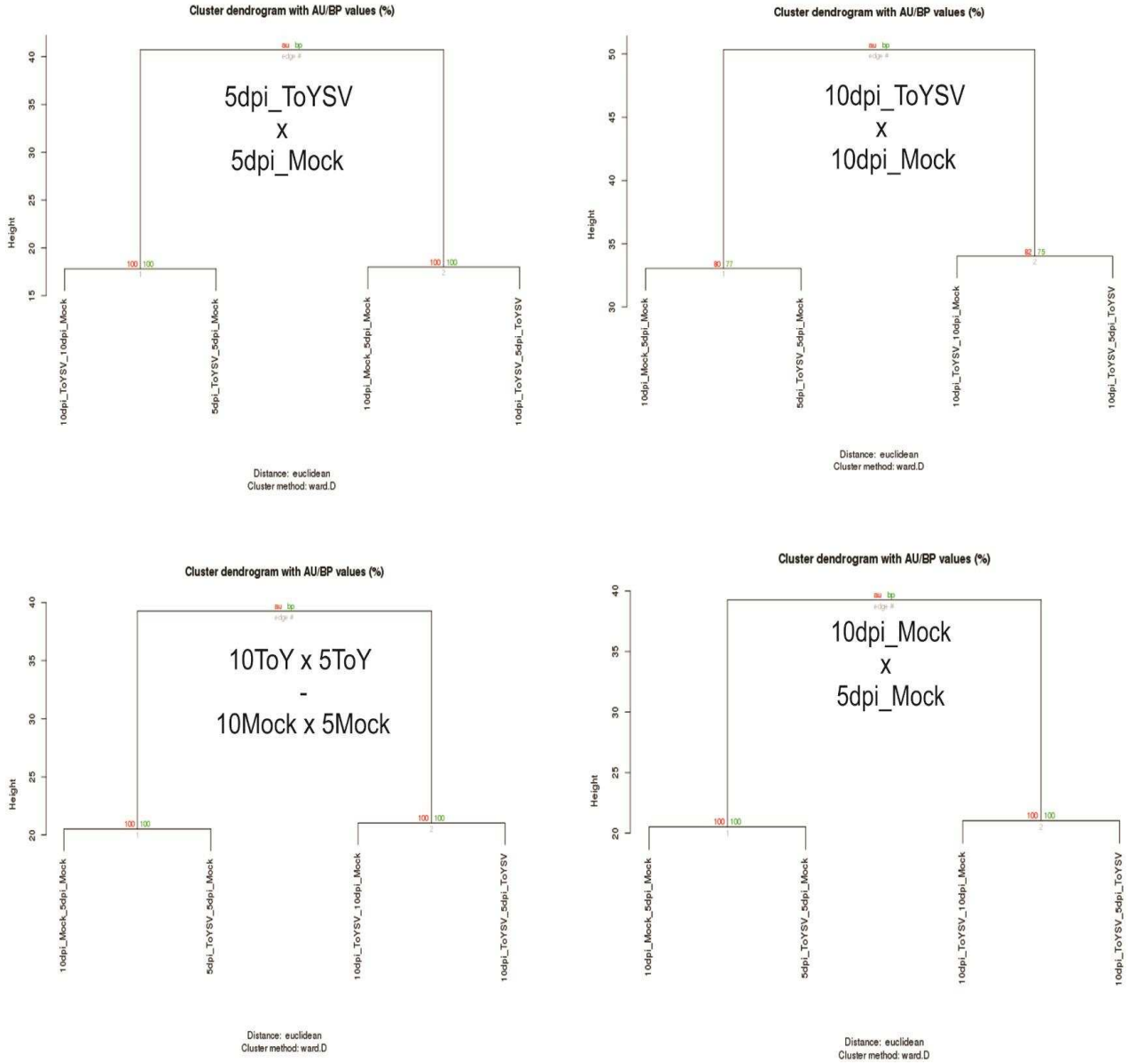
Suppl. Figure S3. Specificity tests for the NbWel2a and NbWel2b primers.



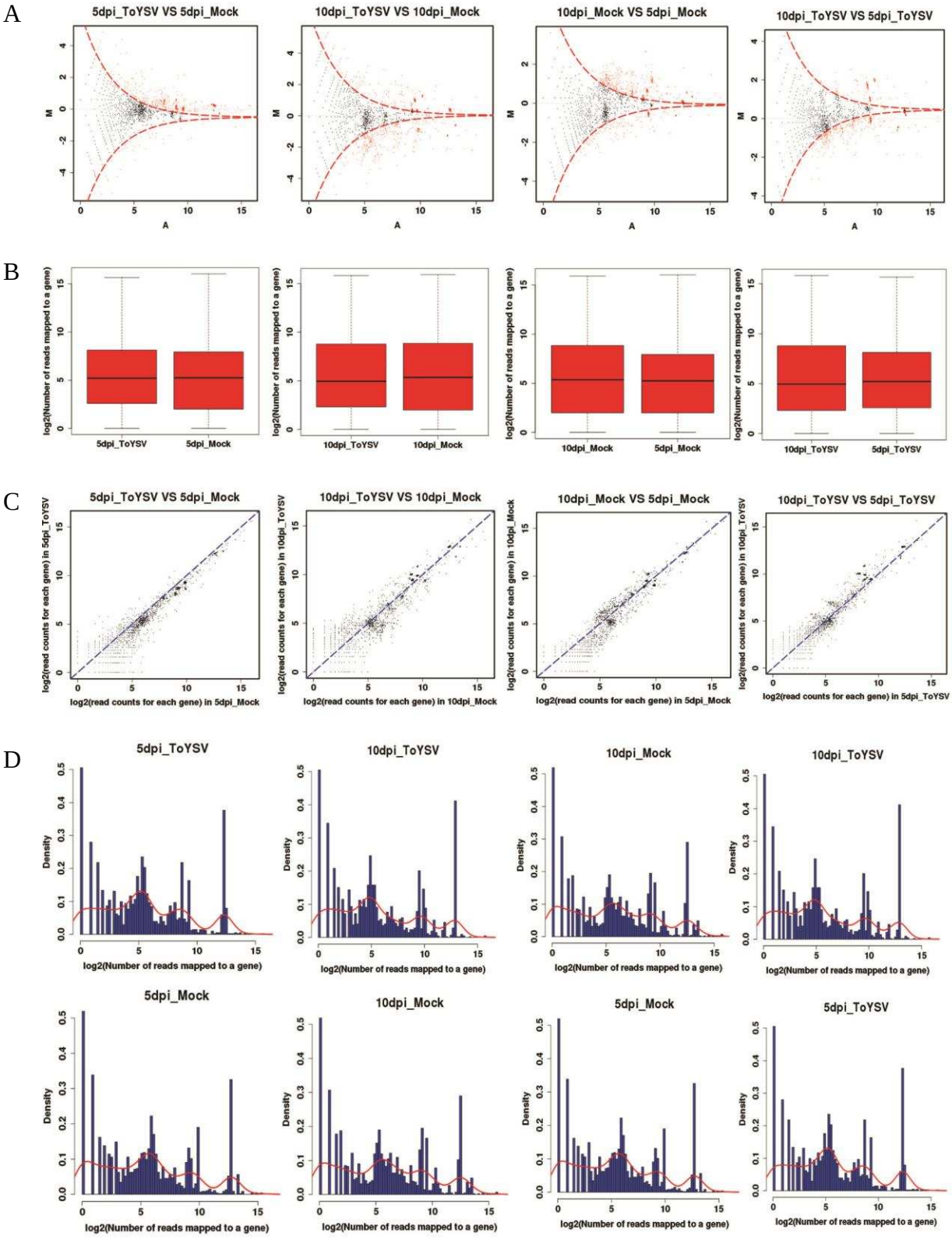
Suppl. Figure S4. Qualities of 18-28 nt sequences from each library, verified with FastQC software.



Suppl. Figure S5. Hierarchical clustering comparison between 5dpi_ToYSV vs 5dpi_Mock, 10dpi_ToYSV vs 10dpi_Mock, 10dpi_ToYSV vs 5dpi_ToYSV contrasts, filtered against 10dpi_Mock vs 5dpi_Mock and 10dpi_Mock vs 5dpi_Mock constrasts. The red numbers indicate approximately unbiased p-values and green numbers are bootstrap probability values.



Suppl. Figure S6. A. MA-plots of differentially expressed genes. Red points indicate adjusted p-values <0.05. **B.** Box-plots of read counts for each gene. **C.** Scatterplot comparing the number of reads for each gene using an ordinary log2 transformation. **D.** Histograms of the number of reads for genes from 5dpi_ToYSV vs 5dpi_Mock, 10dpi_ToYSV vs 10dpi_Mock, 10dpi_Mock vs 5dpi_Mock and 10dpi_ToYSV vs 5dpi_ToYSV contrasts.



Suppl. File S1. Novel MIR genes identified from the eight N. benthamiana libraries.

Suppl. File S2. Differentially expressed MIR genes detected with the IDEG6 software and Fisher's exact test.

Suppl. File S3. Differentially expressed MIR genes detected with the IDEG6 software and Audic Claverie's test.

Suppl. File S4. Differentially expressed known MIR genes from the 5dpi_ToYSV vs 5dpi_Mock contrast and their putative targets.

Suppl. File S5. Differentially expressed known MIR genes from the 10dpi_ToYSV vs 10dpi_Mock contrast and their putative targets.

Suppl. File S6. Differentially expressed known and novel MIR genes from 10dpi_ToYSV vs 5dpi_ToYSV contrast filtered against the 10dpi_Mock vs 5dpi_Mock contrast, and their putative targets.

Suppl. File S7. Gene ontology of the miRNAs target genes at 5dpi_ToYSV vs 5dpi_Mock, 10dpi_ToYSV vs 10dpi_Mock and 10dpi_ToYSV vs 5dpi_ToYSV contrast filtered against the 10dpi_Mock vs 5dpi_Mock contrast. Categorization of miRNA-target genes was performed according to the cellular component, molecular function and biological process categories. Bars indicate the percentage of total annotated targets genes mapping to GO terms.

Suppl. File S8. Identification of vsiRNAs targets. The vsiRNAs generated at 5dpi (18-28 nt reads, unmatched to rRNA and tRNA) matched simultaneously to ToYSV DNA-A or DNA-B components and to the Nicotiana benthamiana transcriptome. These potential targets were assigned based on Gene Ontology annotations (cellular component, molecular function and biological process).

Suppl. File S9. Identification of vsiRNAs targets. The vsiRNAs generated at 10dpi (18-28 nt reads, unmatched to rRNA and tRNA) matched simultaneously to ToYSV DNA-A or DNA-B components and to the Nicotiana benthamiana transcriptome.

Suppl. File S10. Distribution of siRNA candidates normalized in reads per 1 million (RPM) aligned in tandem. Total number of 18-28 nr reads in each library (Table 1) was used for normalization.

CAPÍTULO 3

BUSCA POR PROTEÍNAS DO HOSPEDEIRO CANDIDATAS A INTERAGIREM COM A PROTEÍNA DE MOVIMENTO DO BEGOMOVÍRUS Tomato yellow spot virus

Basso MF, Pereira HMB, Silva FN, Ramos HJO, Fontes EPB, Alfenas-Zerbini P & Zerbini FM. Busca por proteínas do hospedeiro candidatas a interagirem com a proteína de movimento do begomovírus Tomato yellow spot virus.

INTRODUÇÃO

A família Geminiviridae é caracterizada pela morfologia de partículas icosaédricas geminadas e genoma composto por DNA de fita simples circular (Hanley-Bowdoin et al., 1999). A família é organizada em sete gêneros (Becurtovirus, Begomovirus, Curtovirus, Eragrovirus, Mastrevirus, Topocuvirus e Turncurtovirus) com base no número de componentes genômicos (mono- ou bissegmentados), organização genômica, relacionamento filogenético, tipo de inseto vetor (cigarrinha ou mosca-branca) e gama de hospedeiros (mono- ou dicotiledôneas) (Brown et al., 2012; Varsani et al., 2014).

Os vírus pertencentes ao gênero Begomovirus são de grande importância econômica, principalmente em regiões tropicais e subtropicais (Briddon, 2003; Monci et al., 2002; Morales & Anderson, 2001; Were et al., 2004). Infectam plantas dicotiledôneas e são transmitidos naturalmente pela "mosca-branca" Bemisia tabaci (Homoptera: Aleyrodidae) (Brown et al., 2012). Com base em estudos filogenéticos e características genômicas, os begomovírus são divididos em dois grupos: "Velho Mundo" (Europa, Ásia e África), caracterizados por serem na maioria dos casos monossegmentados e estarem associados a DNAs satélites (alfa- e betassatélites) (Mansoor et al., 2003), e "Novo Mundo" (as Américas), em sua maioria bissegmentados. Os dois componentes genômicos dos begomovírus bissegmentados são denominados DNA-A (que contém genes envolvidos na replicação, supressão de respostas de defesa da planta e encapsidação do genoma viral) e DNA-B (que contém genes envolvidos no movimento viral intra- e intercelular) (Rojas et al., 2005). Com exceção da região intergênica (aprox. 200 nucleotídeos) denominada região comum (RC), os dois componentes não apresentam identidade nucleotídica.

Após a replicação viral na primeira célula infectada, é necessário que o vírus se mova para o estabelecimento da infecção sistêmica. O movimento no interior do hospedeiro pode ser dividido em dois processos: movimento célula-a-célula via plasmodesmas e movimento a longa distância ou sistêmico, no qual o vírus atinge o sistema vascular e é transportado via

floema para outras partes da planta hospedeira. Para efetivar o movimento célula-a-célula, os begomovírus bissegmentados codificam uma proteína de movimento (movement protein; MP). Esta se associa à membrana do retículo endoplasmático, acumulando-se nos desmotúbulos e plasmodesmas e alterando o seu limite de exclusão de forma a facilitar o transporte do genoma viral (Ward et al., 1997; Noueiry et al., 1994; Lazarowitz & Beachy, 1999). Como os begomovírus se multiplicam no núcleo da célula hospedeira, é necessária uma etapa adicional de transporte núcleo-citoplasma, que é realizada pela proteína de transporte nuclear (nuclear shuttle protein; NSP) (Sanderfoot et al., 1996). Estas duas proteínas de movimento, atuando de maneira cooperativa, permitem que o vírus alcance o sistema vascular. A interação entre MP e NSP foi demonstrada por meio da re-localização de NSP por MP em protoplastos de N. tabacum e pela interação física no sistema duplo-híbrido de levedura (Sanderfoot & Lazarowitz, 1995; Mariano et al., 2004).

A MP dos begomovírus é classificada na superfamília 30K, na qual o transporte do genoma viral, na forma de nucleoproteína, é guiado pelos desmotúbulos que se movimentam através dos plasmodesmas por meio da ligação específica entre MP e o genoma viral (Rojas et al., 1998; Melcher, 2000). Em um primeiro modelo proposto, denominado couple-skating, NSP transporta o ssDNA ou dsDNA do núcleo para o citoplasma, e em seguida este complexo tem o movimento célula-a-célula mediado por MP (Sanderfoot & Lazarowitz, 1995; Frischmuth et al., 2007; Kleinow et al., 2008). No segundo modelo, denominado relay-race, NSP transfere o dsDNA do núcleo para o citoplasma e em seguida o dsDNA associa-se a MP e movimenta-se célula-a-célula através dos plasmodesmas (Noueiry et al., 1994; Rojas et al., 1998).

Para o estabelecimento da infecção sistêmica é necessária também a interação molecular entre fatores virais e do hospedeiro que possibilitem o movimento viral célula-a-célula e à longa distância (Sanderfoot & Lazarowitz, 1996). No entanto, há poucos trabalhos que demonstram interação direta da MP com fatores do hospedeiro (Krenz et al., 2010; Lewis &

Lazarowitz, 2010; Balmant, 2011). A identificação e o estudo de proteínas do hospedeiro que interagem direta ou indiretamente com esta proteína poderão contribuir para melhorar o entendimento do processo de infecção viral e da interação begomovírus-hospedeiro e fornecer novas estratégias de controle destas viroses, especialmente com relação a resistência do hospedeiro, seja ela natural ou derivada do patógeno.

O método TAP (tandem affinity purification) constitui uma ferramenta para a identificação in vivo de interações diretas ou indiretas entre proteínas (Rohila et al., 2004, 2006, 2009; Rubio et al., 2005; Dufresne et al., 2008). A etiqueta NTAP, fusionada na extremidade N-terminal da proteína alvo, é composta de um domínio peptídico de ligação a calmodulina (que se liga a esferas de calmodulina-agarose na presença de íons de cálcio), um sítio de clivagem proteolítica (AcTEV, originado do potyvírus Tobacco etch virus) e dois domínios da proteína A de Staphylococcus aureus (que se ligam a esferas de IgG-agarose), precedidos por um íntron que liga a proteína alvo, ambos em tandem e sob o controle do promotor 35S do Cauliflower mosaic virus. O método permite a purificação de proteínas complexadas a uma proteína alvo etiquetada, reduzindo assim a presença de proteínas "contaminantes" (não ligadas à etiqueta ou à proteína alvo) e facilitando a posterior identificação por espectrometria de massa (MS). Em plantas, o método TAP tem sido utilizado principalmente em Arabidopsis thaliana (Rubio et al., 2005; Van Leene et al., 2007; Dufresne et al., 2008; Thivierge et al., 2008), N. benthamiana (Rohila et al., 2004) e Oryza sativa (Rohila et al., 2006). Outro método que tem sido amplamente utilizado para estudos de interação proteína-proteína e purificação de complexos protéicos por afinidade é o pull-down baseado na proteína alvo fusionada a uma etiqueta de seis histidinas (6His) ou de glutationa S-transferase (GST) (McGarry et al., 2003).

Estas duas metodologias tem possibilitado a identificação e caracterização de diversas interações proteína-proteína e desta forma podem, teoricamente, ser empregadas para a identificação de proteínas do hospedeiro que interagem com a MP de begomovírus. Este

trabalho teve como objetivos identificar e caracterizar proteínas candidatas a interagirem com a MP do begomovírus Tomato yellow spot virus (ToYSV), utilizando o sistema TAP e a técnica de pull-down com etiqueta 6His.

MATERIAL E MÉTODOS

Transformação genética de Arabidopsis thaliana. As construções pNTAPi-MP (MP-ToYSV íntegra) e pNTAPi-GFP foram utilizadas anteriormente por Silva (2011). A construção pNTAPi-GFP foi utilizada como controle para interações proteína-proteína não específicas.

Células de A. tumefaciens estirpe C58C1 foram transformadas com os vetores binários pNTAPi-MP ou pNTAPi-GFP pelo método de choque térmico (Sambrook & Russel, 2001). Colônias isoladas de transformantes foram incubadas por 12-15 horas a 28^{o}C em 5 ml de meio LB contendo os antibióticos gentamicina, espectinomicina e rifampicina nas concentrações finais de 50, 100 e 50 µg/ml, respectivamente. Posteriormente, 100 µl deste pré-inóculo foram transferidos para 250 ml de meio LB contendo os mesmos antibióticos e estas foram cultivadas até atingirem a fase estacionária (OD_{600} aproximadamente igual a 2,0). As suspensões bacterianas foram centrifugadas a 4.000 g durante 10 minutos a 4ºC, o sobrenadante foi descartado e o pellet ressuspendido em uma solução de sacarose 5% acrescido de Silwett L-77 (Momentive) 0,05% de forma a ajustar OD_{600} para 0,8.

Plantas de A. thaliana ecótipo Columbia (Col-0) cultivadas em substrato comercial (Plantmax) foram mantidas em câmara de crescimento sob condições controladas de temperatura (22^{o}C) e fotoperíodo (16 horas de luz e 8 horas de escuro). As plantas foram transformadas por meio de submersão das inflorescências na suspensão de A. tumefaciens carreando os respectivos vetores binários descritos acima (Clough & Bent, 1998; Martinez-Trujillo et al., 2004). As sementes provenientes das plantas transformadas foram coletadas e semeadas em substrato. Aos 15 dias após a germinação, foram pulverizadas semanalmente

durante quatro semanas com o herbicida glufosinato de amônio (Finale, Bayer Crop Science) na concentração de 0,00578%. Linhagens T2 pré-selecionadas com herbicida foram avaliadas por meio de PCR com oligonucletídeos específicos para os genes MP ou GFP. A confirmação da expressão das proteínas MP ou GFP foi realizada por Western blot utilizando anti-soro policlonal anti-MP-ToYSV (Silva, 2011) e anti-GFP (Invitrogen), conforme descrito por Basso (2010). As plantas confirmadas como transgênicas e superexpressando as proteínas de interesse foram propagadas em casa-de-vegetação.

Transformação genética de tomateiro (Solanum lycopersicum). Sementes de tomateiro cv. Moneymaker foram esterilizadas por imersão em etanol 70% durante um minuto, seguido de incubação sob agitação em solução de hipoclorito de sódio (1,25% de cloro ativo) acrescido de Tween-20 a 0,1%, durante 20 minutos. As sementes foram enxaguadas quatro vezes em água Milli-Q autoclavada e após secagem em papel de filtro estéril foram germinadas em frascos magenta contendo meio de cultura MS meia força (Murashige & Skoog, 1962) acrescido de mio-inositol a 50 mg/l, sacarose a 10 g/l, vitaminas MS a 1 ml/l e ágar a 6 g/l. Os frascos contendo as sementes foram mantidos por dez dias em sala de cultura com controle de temperatura (25°C ±2ºC), irradiação (24 µmol/m^2/s) e fotoperíodo (16 horas de luz e 8 horas de escuro).

Aos 15 a 20 dias após a germinação, os hipocótilos foram seccionados em fragmentos de aproximadamente 1 cm e pré-cultivados durante 24 horas em placas de Petri contendo o meio de pré-cultivo composto por sais MS força total (Murashige & Skoog, 1962) acrescido de vitaminas de Nitsch & Nitsch (1969) a 1 ml/l, zeatina a 2 mg/l, sacarose a 30 g/l, mio-inositol a 100 mg/l, acetoseringona a 100 µM/l e ágar a 6 g/l. Após o pré-cultivo foi realizado o co-cultivo destes explantes com as culturas de *A. tumefaciens* carreando as construções pNTAPi-MP ou pNTAPi-GFP. As culturas foram cultivadas a 28°C em meio LB acrescido dos antibióticos apropriados, conforme descrito acima. Em seguida foram submetidas a centrifugação a 4.000 g durante 10 minutos a 4°C. As células foram ressuspendidas em meio

MS líquido 0,2% acrescido de mio-inositol a 400 mg/l, tiamina-HCl a 400 mg/l e sacarose a 20 g/l, ajustando-se a OD_{600} para 0,4. Os explantes foram imersos durante 5 minutos nas suspensões de A. tumefaciens e em seguida secos em papel de filtro estéril e transferidos novamente para o mesmo meio de pré-cultivo contendo dois discos de papel filtro estéreis dispostos sobre o meio. Os explantes foram mantidos por mais dois dias em sala de cultura nas mesmas condições de temperatura, luz e fotoperíodo descritas anteriormente.

Após o co-cultivo, os explantes foram transferidos para meio de regeneração seletivo contendo sais MS, mio-inositol a 100 mg/l, vitaminas de Nitsch & Nitsch a 1 ml/l, sacarose a 20 g/l, zeatina a 2 mg/l, ágar a 6 g/l, timentina a 100 mg/l e glufosinato de amônio a 600 mg/l. Os explantes foram mantidos nesse meio durante 20 dias em sala de cultura nas mesmas condições descritas acima.

Decorrida a etapa de regeneração, os explantes foram transferidos para o meio de elongamento (sais MS, mio-inositol a 100 mg/l, vitaminas de Nitsch & Nitsch a 1 ml/l, sacarose a 20 g/l, zeatina a 1 mg/l, ágar a 6 g/l, timentina a 100 mg/l e glufosinato de amônio a 600 mg/l) e mantidos durante 30 dias em sala de cultura. Brotações diferenciadas e elongadas foram seccionadas, transferidas para meio de enraizamento composto de sais MS, mio-inositol a 100 mg/l, vitaminas MS a 1 ml/l, sacarose a 20 g/l, glicina a 2 mg/l, ágar a 6 g/l, ácido indol-3-acético a 200 µg/l e timentina a 100 mg/l, e mantidas em sala de cultura nas mesmas condições descritas acima.

Os explantes enraizados e com a parte aérea desenvolvida foram pré-aclimatados em copos descartáveis contendo água Milli-Q autoclavada e cobertos por sacos plásticos de forma a manter a umidade. Os sacos plásticos foram gradativamente abertos a cada dois dias para trocas gasosas e diminuição da umidade relativa do ar interior. As plantas foram mantidas em sala de cultura por oito a dez dias e, em seguida, estas foram transferidas para vasos individuais contendo substrato orgânico autoclavado e mantidas em casa-de-vegetação sob telado sombrite 50% até a completa aclimatação. A confirmação da inserção do tDNA, bem

como a expressão dos transgenes NTAPi-MP ou NTAPi-GFP foram realizadas por meio de PCR com oligonucleotídeos específicos para a MP ou GFP, ELISA indireto (Almeida & Lima, 2001) e Western blot utilizando anti-soros policlonais, conforme descrito acima.

Ensaios de expressão transiente em folhas de Nicotiana benthamiana. Células de A. tumefaciens estirpe C58C1 carreando os vetores binários pNTAPi-MP ou pNTAPi-GFP foram cultivadas até OD_{600} igual a 1,2, conforme descrito acima. A suspensão bacteriana foi centrifugada a 4.000 g durante 10 minutos e o pellet foi ressuspendido em 5 ml de meio de agro-infiltração (MES 100 mM pH 5.5, $MgCl_2$ 10 mM e acetoseringona 200 mM). Em seguida foi novamente centrifugado a 4.000 g por 10 minutos e o pellet ressuspendido em meio de agro-infiltração, ajustando-se a OD_{600} para 1,0.

A agroinfiltração em folhas totalmente expandidas de N. benthamiana sadias e infectadas, 10 dias após a inoculação mecânica do ToYSV (isolado ToYSV-[Bi2]; Calegario et al., 2007) foi realizada com seringa descartável. Amostras destas folhas foram coletadas 48 horas após a agroinfiltração (Silva, 2011) e a confirmação da expressão transiente das proteínas recombinantes foi realizada por meio de Western blot utilizando-se anti-soros policlonais anti-MP e anti-GFP. Para cada tratamento foram coletadas sub-amostras de 60 g de folhas agroinfiltradas com as construções pNTAPi-MP ou pNTAPi-GFP (sadias ou infectadas), que foram cortadas em tiras e mantidas em PBS 1x contendo 1% de formaldeído a 2ºC. Em seguida, foram infiltradas a vácuo durante 30 minutos (Rohila et al., 2004). Esta reação foi bloqueada em solução resfriada de glicina 300 mM durante 30 minutos e, em seguida, estas foram enxaguadas com PBS 1x e o material vegetal armazenado a -80^{o}C para posterior extração de proteínas pelo método TAP.

Purificação dos complexos protéicos por afinidade em tandem (TAP). A purificação por afinidade em tandem dos complexos protéicos NTAPi-MP ou NTAPi-GFP obtidos a partir da expressão transiente ou constitutiva foi realizada inicialmente conforme descrito por Rohila et

al. (2004) (Figura 1). Cada amostra foi composta de 60 g de tecido foliar proveniente das plantas transgênicas (expressão constitutiva) ou selvagens (expressão transiente), sadias ou infectadas com o ToYSV. Estas foram maceradas em almofariz contendo nitrogênio líquido e em seguida homogeneizadas em 180 ml de tampão de extração [Tris-HCl 20 mM pH 8,0, NaCl 150 mM, IGEPAL 0,1%, EDTA 2,5 mM, benzamidina 2 mM, β-mercaptoetanol 10 mM, PMSF 1 mM, coquetel inibidor de proteases (Sigma) 1%, leupeptina 10 µM e 3-4-di-cloro-isocomarina 10 µM]. Em seguida a suspensão foi filtrada em três camadas de gaze e centrifugada duas vezes a 30.000 g por 20 minutos a 4ºC, coletando-se o sobrenadante. O sobrenadante final foi incubado sob agitação lenta em uma solução de 150 µl contendo esferas de IgG Sepharose Fast Flow (GE Healthcare) por 2 horas a 4^{o}C. Em seguida, foi centrifugado a 200 g por 3 minutos e as esferas foram enxaguadas com 30 ml de tampão de lavagem (Tris-HCl 20 mM pH 8.0, NaCl 150 mM, IGEPAL 0,1% e EDTA 2,5 mM). Após esta etapa de lavagem as esferas foram transferidas para uma coluna de cromatografia por afinidade (Poly-Prep, BioRad) e enxaguadas três vezes com 10 ml de tampão de clivagem (Tris-HCl 10 mM pH 8,0, NaCl 150 mM, IGEPAL 0,1%, EDTA 0,5 mM e DTT 1 mM). Os complexos protéicos foram eluídos com a adição de 135 µl de tampão de clivagem contendo 1 µM de inibidor de protease E-64 (Sigma) e 150 U de protease AcTEV (Invitrogen), incubando-se a reação por 1 hora a 16ºC.

Aos complexos protéicos recuperados foram adicionados três volumes de tampão de ligação a calmodulina (Tris-HCl 10 mM pH 8,0, NaCl 150 mM, IGEPAL 0,1%, acetato de magnésio 1 mM, imidazole 1mM, $CaCl_2$ 2 mM e beta-mercaptoetanol 10 mM). Em seguida foram adicionados 100 µl de solução contendo esferas Calmodulin Sepharose 4B (GE Healthcare), incubando-se com agitação lenta por 1 hora a 4ºC. Esta suspensão foi transferida para uma coluna de cromatografia de afinidade (Poly-Prep) e as esferas foram enxaguadas com 30 ml de tampão de ligação a calmodulina. Os complexos protéicos ligados às esferas foram eluídos com adição de 90 µl de tampão de eluição (Tris-HCl 10 mM pH 8.0, NaCl 150

mM, IGEPAL 0,1%, acetato de magnésio 1 mM, beta-mercaptoetanol 10 mM, imidazole 1 mM e EGTA 2 mM).

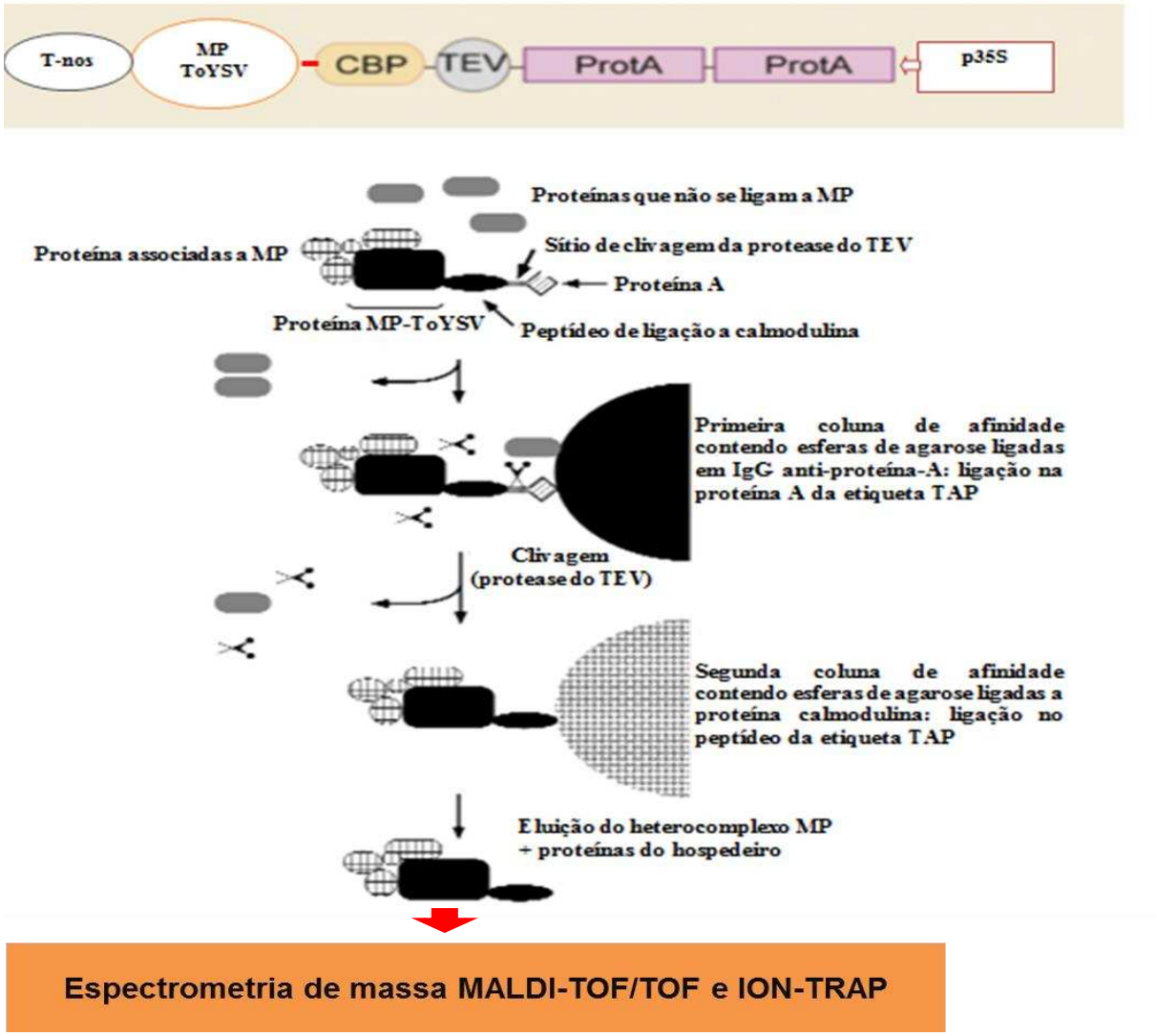


Figura 1. Fluxograma da purificação dos heterocomplexos protéicos através do método de purificação por afinidade em tandem (TAP).

A purificação dos complexos protéicos NTAPi-MP ou NTAPi-GFP também foi realizada com modificações no método TAP originalmente descrito por Rohila et al. (2004). A primeira modificação foi feita no tampão de eluição, com a ausência de IGEPAL. Posteriomente retirou-se o IGEPAL de todas as etapas seguintes.

Na tentaviva de obter-se um heterocomplexo protéico mais estável, o procedimento de purificação foi finalizado antecipadamente em duas etapas independentes: na etapa posterior à ligação da etiqueta TAP às esferas de agarose contendo a calmodulina (parada 1), e na etapa posterior à ligação da etiqueta TAP às esferas de agarose ligadas em IgG anti-proteína-A (parada 2). Após ligação (esferas + calmodulina ou IgG anti-proteína-A + etiqueta TAP + proteínas do hospedeiro) procedeu-se à centrifugação a 1.000 g por 20 segundos. O sobrenadante foi cuidadosamente descartado e as esferas de agarose contendo o heterocomplexo foram enxagadas três vezes com tampão de ligação a calmodulina (parada 1) ou tampão de lavagem (parada 2). O heterocomplexo protéico foi ressuspendido em tampão de eluição sem IGEPAL e eluído por aquecimento a 90ºC durante 3 minutos. Após centrifugação a 13.200 g o sobrenadante foi coletado, preparado e analisado por MS.

Etapas adicionais de purificação dos complexos protéicos ou das frações peptídicas provenientes tanto da purificação pelo método TAP original quanto do método modificado foram acrescentadas na tentativa de aumentar a pureza e melhorar a qualidade das análises por MS. Para isso foram utilizadas colunas de cromatografia de fase estacionária Strata C18-E 100 mg/3 ml (fase reversa) e Strata SCX 500 mg/3 ml (troca catiônica) (Phenomenex) ou ponteiras de micropipeta ZipTip com resina C18 (Millipore), conforme as instruções do fabricante.

Purificação dos complexos protéicos por pull-down. A proteína MP do ToYSV fusionada a uma cauda de seis histidinas (6His) superexpressa em Escherichia coli e purificada (Silva, 2011) foi utilizada. Inicialmente, colunas de cromatografia de afinidade (Poly-Prep) contendo resina Ni-NTA Agarose (Qiagen) foram equilibradas com água Milli-Q autoclavada e posteriormente com tampão N ($NaHCO_3$ 100 mM pH 9,0). Nestas foram adicionadas 70 µg de MP-6His diluída em 1 ml de tampão N, e as colunas foram enxaguada com 10 ml do mesmo tampão. As proteínas totais provenientes de plantas de N. benthamiana, A. thaliana Col-0 e tomateiro, sadias ou infectadas com o ToYSV, foram maceradas em nitrogênio

líquido e tampão de extração [Tris-HCl 20 mM pH 8,0, NaCl 150 mM, EDTA 2,5 mM, benzamidina 2 mM, β-mercaptoetanol 10 mM, PMSF 1 mM, coquetel inibidor de protease (Sigma) 1%, leupeptina 10 µM e 3-4-di-cloro-isocomarina 10 µM], na proporção 1:10 (peso de folha: volume de tampão). Após centrifugação, foram transferidos 10 ml deste extrato protéico foram transferidos para a coluna de cromatografia contendo a proteína MP-ToYSV em fase estacionária e em seguida esta foi enxaguada com 15 volumes de tampão N acrescido de 0,1% de SDS. Os complexos protéicos foram eluídos com 3 ml de tampão N acrescido de SDS 0,1% e imidazole 250 mM.

Precipitação dos complexos protéicos e digestão tríptica. Os complexos protéicos purificados pelos dois métodos descritos acima foram precipitados com a adição de ácido tricloroacético (TCA, Sigma) na proporção de 1 µl de TCA para 4 µl de amostra, mantidos a -20ºC durante 6-8 horas e em seguida centrifugados a 20.000 g por 15 minutos a 4ºC. O pellet foi ressuspendido em acetona (inicialmente 100% e em seguida 80%) e centrifugado a 20.000 g por 10 minutos a 4ºC. O pellet foi desidratado em estufa a 37ºC, ressuspendido em 100 µl de bicarbonato de amônio 25 mM pH 8,8 e homogeneizado por sonicação. A digestão tríptica dos complexos protéicos foi realizada conforme descrito por Shevchenko et al. (2006) e Kinter & Sherman (2000).

Identificação das proteínas por espectrometria de massa. As análises de MS foram realizadas conforme descrito por Rohila et al. (2004 e 2006) utilizando espectrômetros de massas MALDI-TOF/TOF e LC/ION-TRAP. Para as análises em espectrômetro de massa LC/ION-TRAP, os peptídeos trípticos foram solubilizados em 30 µl de água grau MS (Sigma) contendo 0,1 % de ácido fórmico e injetados no espectrômetro LC-MS utilizando um nanoAquity UPLC System (Waters), contendo uma trap column (Waters) e uma coluna de capilaridade ProteCol C_{18}5-Å 300 µm x 100-mm (SGE Analytical Science) operando com razão de fluxo de 3,0 µl/min. Os tampões de fase móvel utilizados para o gradiente foram

água acrescida de 0,1% de ácido fórmico e acetonitrila acrescida de 0,1% de ácido fórmico. Os compostos eluídos foram injetados em um espectômetro de massas ION-TRAP (Bruker Daltonics).

A aquisição de dados de ambos os instrumentos LC-MS foi gerenciada utilizando-se o Hystar Package (Bruker) e os espectros foram processados utilizando Data Analysis Package (Bruker) com as configurações padrões para proteômica. O espectrômetro de massas operou em modo Auto-MSn, coletando espectro MS2 para os íons mais intensos e excluindo íons de carga simples. O tempo de varredura foi ajustado para 0,5 segundos para a digitalização do escaneamento e cada espectro foi registrado cinco vezes para permitir a coleta do número máximo de espectros MS2 durante a análise. Um arquivo com formato mascot generic format (.mgf) com uma lista de picos (peak list) foi gerada com o Data Analysis Package (Bruker). A identificação das proteínas foi realizada pela busca nos bancos de dados públicos de proteínas do NCBI e SOL Genomics. Os dados do MS/MS ion search foram utilizados para a busca de proteínas pelo software MASCOT versão 2.2.07 (Matrix Science), em um sistema de busca local, considerando os seguintes paramêtros: tolerância de 0,1 Da para os íons, uma clivagem perdida, estados de carga 2, 3 e 4 e protease tripsina. As modificações variáveis nos peptídeos e aminoácidos foram carbamidometilação da cisteína, oxidação da metionina e deaminação da asparagina e glutamina. A identificação dos peptídeos e proteínas foi estatisticamente avaliada com os algoritmos PeptideProphet e ProteinProphet utilizando-se o Scaffold package (Proteome Software, Inc.) para 90% de probabilidade.

Para o método MALDI-TOF/TOF, a ionização dos peptídeos para a fase gasosa foi feita em MALDI (Ultraflex 3, Bruker Daltonics) e o método de detecção das massas foi obtido por detector do tipo TOF/TOF. Um volume de 2 µl de cada amostra de peptídeos foi homogeneizado em 2 µl de matriz α-ciano-4-hidroxicinâmico (3 mg/ml de ácido α-ciano-4-hidroxicinâmico em 50% acetonitrila e 0,1% ácido trifluoracético) ou ácido 2,5-di-hidróxi-benzóico (10 mg/ml de ácido 2,5-di-hidróxibenzóico em acetonitrila e 0,1% de ácido

trifluoracético) (Sigma), sendo em seguida espotadas em triplicata em placa de aço inoxidável de 96 spots. Após secagem da preparação peptídeo-matriz, estes foram analisados em MALDI para a obtenção das listas de massas dos peptídeos e seus fragmentos. Uma lista de picos foi gerada, e os mais intensos foram fragmentados pelo método LIFT para gerar o espectro MS2. Cada espectro foi detectado em modo linear positivo e calibrado externamente utilizando-se uma mistura padrão de peptídeos. A identificação dos peptídeos foi realizada utilizando-se os softwares MASCOT e Scaffold, conforme descrito acima.

Clonagem do gene MP do ToYSV sem o domínio transmembrana. O mapeamento do domínio transmembrana presente na MP-ToYSV foi baseado no domínio transmembrana previamente caracterizado na MP do Abutilon mosaic virus (AbMV) (Zhang et al., 2002) e por meio de alinhamento das sequências de aminoácidos. O gene MP do ToYSV (isolado ToYSV-[Bi2]; Calegario et al., 2007) foi amplificado por PCR convencional em duas etapas. Para amplificação da região anterior ao domínio transmembrana (365 pb) utilizaram-se os oligonucleotídeos MPEcoRI(Forward) (5'-CGC GGA ATT CGT TGA AAT GGA TTC TCA G-3') e MPNotI(Reverse) (5'-ACG CGG CCA ATC TTA TCT AAA CAG TTC-3'). Para amplificar a região posterior ao domínio transmembrana (421 bp) utilizaram-se os oligonucleotídeos MPNotI(Forward) (5'-ACG CGG CCC GGT ATT GGG TCT TT-3') e MPXhoI(Reverse) (5'-CGC GCT CGA GCC CTC TCT TAA CTC AAT G-3'). Os produtos da amplificação foram purificados, submetidos à restrição com as enzimas apropriadas e clonados no vetor de entrada pENTR11 (Invitrogen). A integridade e a fase de leitura desta construção foram confirmadas por sequenciamento (Macrogen, Inc.). Os sítios de clonagem attL1 e attL2 presentes neste vetor foram utilizados para subclonar o gene MP sem domínio transmembrana (MPΔtmd) no vetor binário pNTAPi (attR1 e attR2) por meio de recombinação homóloga mediada por LR clonase (Invitrogen), gerando a construção pNTAPi-MPΔtmd.

Clonagem e análise in vitro da interação entre MP-ToYSV e proteínas do hospedeiro. Entre as 64 proteínas de plantas isoladas por pull-down e identificadas por MS (Tabela 1), foram selecionadas 25 com base em sua função biológica, localização celular e identificação em mais de um tipo de hospedeiro (Tabela 2). Para a clonagem destes genes procedeu-se inicialmente ao isolamento de RNA total das plantas (A. thaliana Col-0, N. benthamiana e tomateiro sadias ou infectadas com o ToYSV) com o Concert Plant RNA Reagent (Invitrogen), conforme as instruções do fabricante. Ensaios de RT-PCR foram conduzidos com oligonucleotídeos específicos desenhados com base nas sequências disponíveis nos bancos de dados do TAIR (The Arabidopsis Information Resource) e SGN (SOL Genomics Network). O método de clonagem foi baseado na inserção de sítios de restrição nos oligonucleotídeos de forma a auxiliar a subclonagem dos genes no vetor de entrada (pENTR11) e a posterior transferência por recombinação homóloga (Gateway LR Clonase II Enzyme mix, Invitrogen) para vetores de duplo híbrido (pDEST22 e pDEST32; Invitrogen). A identidade, integridade e orientação dos genes clonados foram confirmadas por sequenciamento (Macrogen, Inc.).

Para os testes de interação proteína-proteína no sistema duplo-híbrido de levedura utilizou-se Saccharomyces cerevisiae estirpe AH109 (Clontech), que possui os genes repórteres ADE2, HIS3, LacZ e MEL1 controlados por sequências promotoras da transcrição GAL4 ou MEL1. Os 25 genes de plantas selecionados foram clonados tanto no vetor contendo o domínio de ativação (pDEST22) quanto no vetor com o domínio de ligação (pDEST32). Da mesma forma, a MP-ToYSV íntegra ou sem domínio transmembrana foram clonadas em ambos os vetores. A co-transformação sequencial das células competentes de levedura foi realizada pelo método de acetato de lítio e polietilenoglicol acrescido de 2,5 µg de DNA de esperma de salmão, seguido de choque-térmico. Os transformantes simples e duplos foram selecionados após incubação a 28ºC em meio SD (Synthetic Dropout) com deficiência em leucina e/ou triptofano. Dentre os duplos transformantes, três colônias foram

selecionadas, cultivadas em meio SD líquido sem leucina e triptofano até atingir OD_{600} igual a 1,0 e plaqueadas em diluições seriadas (OD_{600} igual a 1,0, 0,1 e 0,01) neste mesmo meio sólido com deficiência em leucina, triptofano e histidina e suplementado com 5 mM de 3-amino-1,2,4-triazole (3AT; Sigma). As placas foram incubadas a 28ºC durante 3 a 4 dias, e o crescimento celular foi avaliado por meio de comparação com os controles positivo e negativo. Como controle negativo foi utilizado a MP (com ou sem domínio transmembrana) clonada nos vetores de duplo-híbrido em combinação com os respectivos vetores vazios. Como controle positivo foi utilizado o gene responsável por codificar uma proteína receptora do tipo kinase, NSP-interacting kinase (NIK), previamente caracterizada por Santos et al. (2010) e gentilmente cedido pela Profa. Elizabeth P.B. Fontes (UFV).

Os genes candidatos a interagirem com a MP e selecionados por duplo-híbrido foram avaliados por ensaios de quantificação da atividade de β-galactosidase, determinada conforme descrito por Amberg et al. (2005). Após a permeabilização das células de levedura, utilizou-se orto-nitropfenil-β-galactosídeo (4 mg/ml) (Sigma) como substrato para medir a atividade da enzima. Os controles utilizados nestes ensaios foram os mesmos descritos anteriormente, acrescido de controles compostos dos genes de plantas acima selecionados, clonados nos vetores de duplo-híbrido em combinação com os respectivos vetores vazios.

RESULTADOS

Transformação genética de Arabidopsis thaliana e tomateiro. Por meio de PCR convencional e Western blot confirmou-se que quatro plantas de A. thaliana expressavam constitutivamente NTAPi-MP (Figura 2A, B) e sete plantas expressavam NTAPi-GFP (dados não mostrados). Estas plantas transgênicas foram propagadas em câmara-de-crescimento e a geração T2 de cada tratamento foi também avaliada por Western blot.

A transformação genética de plantas de tomateiro também foi confirmada por PCR e Western blot (Figura 2C, D). As plantas transgênicas aclimatadas foram mantidas em casa-de-vegetação e propagadas vegetativamente por estaquia.

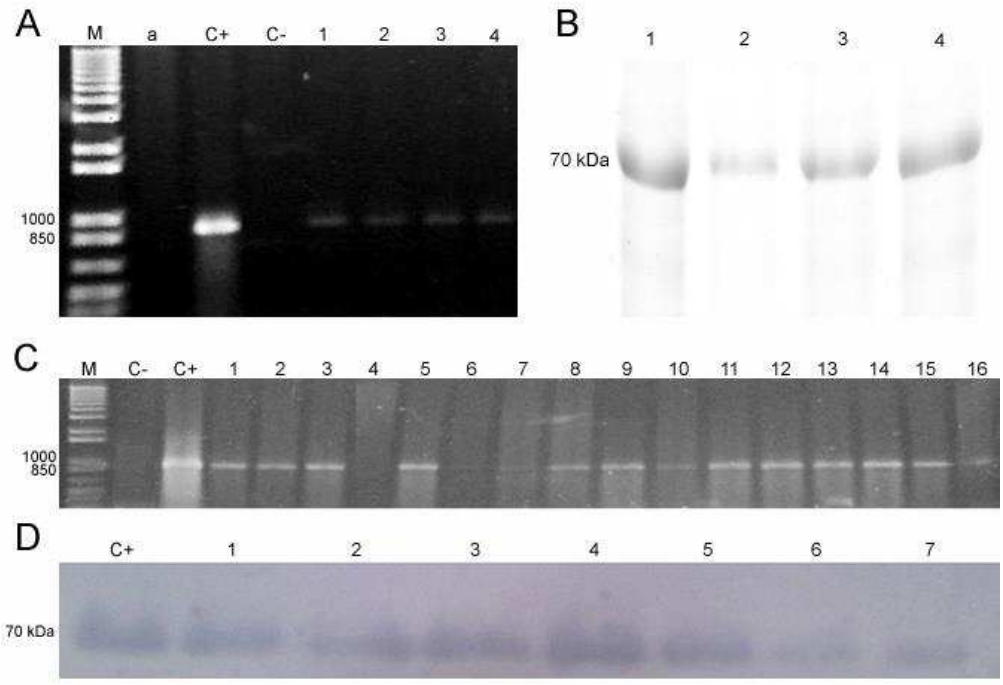


Figura 2. Transformação genética de A. thaliana e tomateiro para expressão constitutiva da etiqueta NTAPi-MP. **A.** PCR com oligonucleotídeos específicos para MP (882 pb). M, marcador de tamanho (1 kb plus DNA ladder, em pb); a, reação sem DNA (controle negativo); C+, amplificação a partir de DNA do clone contendo o gene MP (controle positivo); C-, planta não transformada; 1-4, plantas transformadas. **B.** Western blot utilizando anti-soro policlonal anti-MP-ToYSV para confirmação da transformação de A. thaliana. 1-4, plantas transformadas. **C.** PCR com oligonucleotídeos específicos para MP (882 pb). M, marcador de tamanho (1 kb plus DNA ladder, em pb); C+, amplificação a partir de DNA do clone contendo o gene MP (controle positivo); C-, planta não transformada; 1-16, plantas transformadas. **D.** Western blot utilizando anti-soro policlonal anti-MP-ToYSV para confirmação da transformação de tomateiro. C+, proteína MP expressa em E. coli (controle positivo); 1-7, plantas transformadas.

Expressão transiente em folhas de Nicotiana benthamiana. Os ensaios de expressão transiente das construções pNTAPi-MP e pNTAPi-GFP foram conduzidos em folhas completamente expandidas de N. benthamiana sadias ou infectadas com o ToYSV. As plantas foram agroinfiltradas 10 dias após a inoculação (dpi) com o vírus. Após 48 horas da agroinfiltração (12 dpi), tempo previamente determinado como de maior expressão transiente dessas construções (Silva, 2011), as folhas agroinfiltradas foram coletadas e a expressão das proteínas MP e GFP foi confirmada por Western blot (Figura 3A). As folhas agroinfiltradas foram estocadas a -80ºC. Após a coleta das folhas, as plantas inoculadas com o ToYSV foram mantidas em casa-de-vegetação e a infecção viral foi confirmada com base na observação visual de sintomas nas folhas sistêmicas (não-inoculadas).

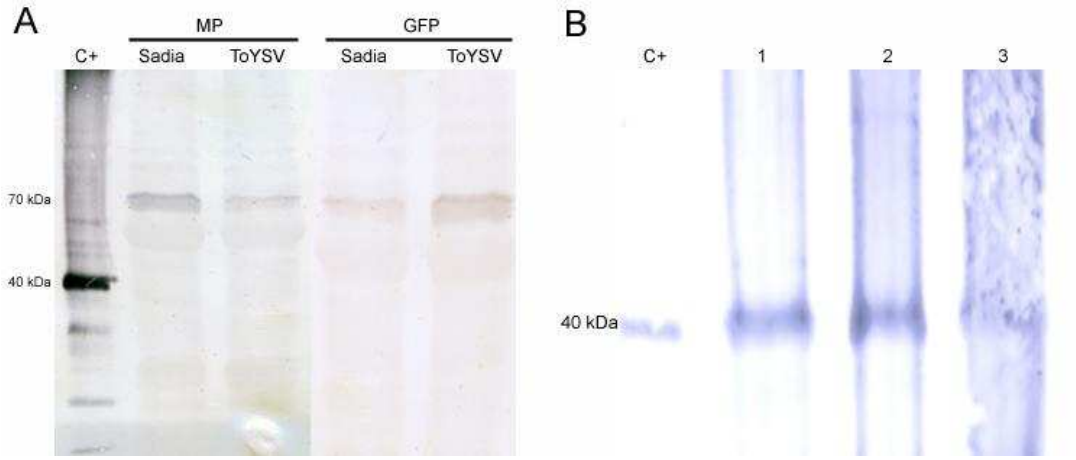


Figura 3. Expressão transiente em N. benthamiana de NTAPi-MP e NTAPi-GFP. **A.** Western blot a partir do extrato de proteínas totais provenientes de folhas de N. benthamiana sadias ou infectadas com ToYSV, agroinfiltradas com as construções pNTAPi-MP e pNTAPi-GFP**. B.** Purificação dos complexos protéicos por pull-down. Western blot utilizando anti-soro policlonal anti-MP-ToYSV para confirmação da presença da MP no extrato final de proteínas totais. C+, MP-ToYSV expressa em E. coli (controle positivo); 1, 2, 3. Proteínas totais eluídas provenientes da interação MP-ToYSV com proteínas totais de N. benthamiana, A. thaliana e tomateiro, respectivamente.

Tentativas de purificação dos complexos protéicos por TAP. A partir de plantas transgênicas (*A. thaliana* e tomateiro) expressando constitutivamente e *N. benthamiana* expressando transientemente a etiqueta NTAPi-MP e utilizando-se o método de purificação TAP original (Rohila et al., 2004) ou com modificações e adaptações ao longo do protocolo TAP, não foram identificadas proteínas dos hospedeiros complexadas nestas etiquetas. Em alguns casos, nas duas paradas ao longo do protocolo TAP a partir de amostras expressando transientemente as etiquetas, foram detectadas apenas as proteínas MP ou GFP, contudo em baixas concentrações. Aparentemente a expressão constitutiva das etiquetas, em ambos os hospedeiros, foi baixa em comparação com a expressão transiente, de modo a fornecer quantidade suficiente de heterocomplexos ao ponto de serem purificados pelo método TAP e identificados por MS.

A presença de contaminantes, tanto a partir de material vegetal expressando constitutivamente quanto transientemente a proteína etiquetada, reduziu significativamente a sensibilidade de detecção pelos dois métodos de MS utilizados (MALDI-TOF/TOF e LC/ION-TRAP). Na tentativa de eliminar estes contaminantes das frações peptídicas foram utilizadas colunas cromatográficas de fase reversa e de troca catiônica, contudo, além de eliminar as impurezas, ocorreu também a redução da concentração dos peptídeos e as proteínas etiquetadas não foram detectadas por MS.

Purificação dos complexos protéicos por pull-down. Utilizando-se a proteína 6His-MP-ToYSV como isca e proteínas totais dos hospedeiros *A. thaliana*, *N. benthamiana* e tomateiro como presas, foi possível isolar e identificar por MS 64 proteínas candidatas a interagirem com a MP-ToYSV (Tabela 1). Entre estas 64 proteínas, 25 foram selecionadas para avaliação in vitro com base em sua função biológica, localização celular, maior números de peptídeos identificados por MS e identificação em mais de uma espécie hospedeira.

As proteínas rubisco e queratina também foram detectadas entre as proteínas purificadas e identificadas por espectrometria de massa. A presença dessas proteínas pode ser explicada devido a sua alta concentração na planta e consequente ligação inespecífica às esferas de agarose (rubisco), e à contaminação durante o manuseio das amostras (queratina).

Tabela 1. Proteínas de plantas hospedeiras (Arabidopsis thaliana, Nicotiana benthamiana e tomateiro) do Tomato yellow spot virus (ToYSV) candidatas a interagirem com a MP viral, obtidas por meio de pull-down com a proteína MP fusionada a uma etiqueta de seis histidicnas (6His-MP-ToYSV) e identificadas por espectrometria de massa (ION-TRAP MS). Análise das sequências peptídicas ("protein fingerprinting") com base nos bancos de dados do NCBI e SOL Genomics.

Proteína candidata	Planta hospedeira	Massa molecular (kDa)	Cobertura	No. de peptídeos identificados	Identidade	Espectro do cromatograma	Banco de dados	Sequência da proteína candidata e peptídeos identificados por MS (em negrito e maiúsculas)
Cobalamin synthesis protein P (CobW-like)	N. benthamiana	51	15%	6	100%	Excelente	SOL Genomics	maslsidiattfmglakrqtpqfyglrtavfptlwktskcqrfslstssssrisvaintsskpnrfsrgftvsatstsapqsedsdiltkipqdnr**IPATIITGFLGSGK**ttllnhiltadhgkriavieneygevdidgslvaaktagaeeimmlnngclcctvrgdlvr**MIAELVSR**kkgkfdhiviettglanpapiiqtfyaedqvfndvk**LDGVVTLVDSK**hvgfhldetkpk**GVVNEAVEQIAYADR**iiinktdlvgdsevssliqrikninkmahlkrtqfgk**VDLDYVLGIGGFDLER**iessvgaegskedhtghdhdhdhdhhhhhdhdhdhehdhkhehhdhhhshdhthdpgvssvsivcegtldlekanmwlgtllmersddiyrmkgllsvegmderfvfqgvhdifqgspdrpwksdeprtnk**IVFIGK**nldakeleegfkaclt*
PTAC17 plastid transcriptionally active 17	A. thaliana	50	2%	1	75%	Bom	NCBI	matlstleiattflsftaprssaafnyrfssaavsvlsrprattvsvrttpsffyspvvrrqrfssvsasatqtedsdvttkippdnripatiitgflgsgkttllnhiltrdhgkriavienefgevdidgslvasksigaedivmlnngclcctvrgdlvrmigelvntkkgkfdhiviettglanpapiiqtfyaeeeifndvkldgvvtlvdakharlhldevkpegvvneaveqiayadriivnktdlvgeaelgsvvqrik**TINSMAQMTR**tkygnvdldyvlgiggfdleriessvneddkgdhhdhdhdhhhdhnhdhdhhhhdghdhhhhshdhthdpgvssvsivcegsldlekanmwlgtllmersediyrmkgllsvhtmeerfvfqgvhdifqgspdrlwgreeervnkivfigknlnreelekgfkacli*
AGT (alanine:glyoxylate aminotransferase)	A. thaliana	44	38%	8	100%	Excelente	NCBI	mdymygpgr**HHLFVPGPVNIPEPVIR**amnrnnedyrspaipaltktlledvkkifkttsgtpflfpttgtgawesaltntlspgdrivsfligqfsllwidqqkrlnfnvdvvesdwgqganlqvlasklsqdenhtik**AICIVHNETATGVTNDISAVR**tlldhykhpalllvdgvssicaldfr**MDEWGVDVALTGSQKALSLPTGLGIVCASPK**aleatktskslkvffdwndylkfyklgtywpytpsiqllyglr**AALDLIFEEGLENIIAR**harlgkatrlaveawglknctqkeewisntvtavmvpphidgseivrrawqr**YNLSLGLGLNK**vagkvfr**IGHLGNVNELQLLGCLAGVEMILKDVGYPVVMGSGVAAASTYLQHHIPLIPSRI***
mtLPD1 (mitochondrial lipoamide dehydrogenase 1)	N. benthamiana A. thaliana Tomateiro	54	21%	2 2 3	100%	Excelente	NCBI	mamaslarrkayfltrnlsnsptdalrfsfslsrgfassgsdendvviigggpggyvaaikasqlglkttciekr**GALGGTCLNVGCIPSK**allhsshmyheakhsfanhgikvssvevdlpamlaqkdnavknltrgieglfkknkvtyvkgygk**FISPNEVSVETIDGGNTIVK**gkhiivatgsdvk**SLPGITIDEK**kivsstgalslsevpkklivigagyiglemgsvwgrlgsevtvvefagdivpsmdgeirkqfqrslekqkmkfmlktkvvsvdsssdgvk**LTVEPAEGGEQSILEADVVLVSAGR**tpftsgldlekigvetdkagrilvndrflsnvpgvyaigdvipgpmlahkaeedgvacvefiagkhghvdydkvpgvvythpevasvgkteeqlkkegvsyr**VGKFPFMANSR**ak**AIDNAEGLVK**iladketdkilgvhimapnageliheavlainydassediar**VCHAHPTMSEALK**eaamatydkpihi*
Aminotransferase-like protein Gamma-aminobutyrate transaminase-like protein	A. thaliana	51	15%	4	100%	Excelente	NCBI	mnsgdvfilsrskghdmlapftagwqsadldplviaksegsyvyddtgkkyldslaglwctalggnepr**LVSAAVEQLNTLPFYHSFWNR**ttkpsldlak**VLLEMFTANK**makafftsggsdandtqvklvwyynnalgrpekkkfiarkksyhgstlisaslsglpplhqnfdlpapfvlhtdcphywrfhlpgeteeefstrlaknledliikegpetigafiaepvmgaggvipppatyfekvqavvkkydilfiadevicafgrlgtmfgcdkynikpdlvtlak**ALSSAYMPIGAILMSQEVADVINSHSSK**lgvfshgftysghpvscavaiealkiykernipeyvakvaprfqdgvk**AFASGSPIIGETR**gtglilgtefvdnkspnepfppewgvgaffgaecqkhgmlvrvagdgilmsppliispeeidelisiygkalkateekvkelkaqhkk*
Serine-glyoxylate aminotransferase Aminotransferase, class V/Cysteine desulfurase"	A. thaliana Tomateiro	44	4%	5 1	100% 98%	Bom	SOL Genomics	mdyvngpgrnhlfvpgpvnipeqvlramnrnnedyrspavpaltktlledvkkifkttsgtpflfpttgtgawesaltntlspgdrtvsfligqfsllwidqqkrlnfdvdvvesdwgqganlevlaskiaedkshtikavcivhnetatgvtnnlatvrkildhyqhpalflvdgvssicaldfrmdewgvdvaltgsqk**ALSLPTGLGIVCASPK**aleasktaksvrvffdwsdylkfyklgtywpytpsiqllyglraaldllfeegldnviarharlgkatrlaveawglknctqkeewfsdtvtavvvpphiesseivkrawkrynlslglglnkvagkvfrighlgnlnelqllgclagvemvlkdigypvkfgsgvaaasaflqnstplipsri*

Tabela 1 (cont.)

L-threonine 3-dehydrogenase "Alcohol dehydrogenase, zinc-containing, conserved site"	N. benthamiana	38	19%	6	100%	Excelente	SOL Genomics	mgkggsdenmaawllgvntlkiqpfnlpalgphdvrvrmkavgicgsdvhylktmrcadfvvkepmvighecagiieevggevktlv pgdrvalepgiscwrcnlckegr**YNLCPEMK**ffatppvhgslanqvvhpadlcfklpddisleegamceplsvgvhacrr**ANV GPETNILVLGAGPIGLVTLLAAR**afgapr**IVIVDVDDYR**lsvakklgaddivkvsiniqdvatdieniqkamg ggidasfdcagfnktmstalgatrpggkvclvgmghhemtvpltpaaar**EVDVIGIFRYKNTWPLCLEFLR**sgkidv kplithrfgfsqeeveeafetsarggdaikvmfnl*
Sorbitol related enzyme	N. benthamiana	38	14%	3	100%	Excelente	NCBI	mgkggsdenmaawllgvntlkiqpfnlpalgphdvrvrmkavgicgsdvhylktmrcadfvvkepmvighecagiieevggevktlv pgdrvalepgiscwrcnlckegrynlcpemkffatppvhgslanqvvhpadlcfklpddisleegamceplsvgvhacrr**ANVGPE TNILVLGAGPIGLVTLLAAR**afgapr**IVIVDVDDYR**lsvakklgaddivkvsiniqdvatdieniqkamgggida sfdcagfnktmstalgatrpggkvclvgmghhemtvpltpaaarevdvigifr**YKNTWPLCLEFLR**sgkidvkplithrfgfs qeeveeafetsarggdaikvmfnl*
CBL (cystathionine beta-lyase)	A. thaliana	50	8%	2	100%	Excelente	NCBI	mtsslslhssfvpsfadlsdrglisknsptsvsiskvptwekkqisnrnsfklncvmeksvdgqthstvnnttdslntmnikeeasvstllvnl dnkfdpfdamstplyqtatfkqpsaiengpydytrsgnptrdalesllakldkadr**AFCFTSGMAALSAVTHLIK**ngeeiv agddvyggsdrllsqvvprsgvvvkrvnttkldevaaaigpqtklvwlesptnprqqisdirkisemahaqgalvlvdnsimspvlsrple lgadivmhsatkfiaghsdvmagvlavkgeklakevyflqnsegsglapfdcwlclrgiktmalriekqqenarkiamylsshprvkkv yyaglpdhpghhlhfsqakgagsvfsfitgsvalskhlvettkyfsiavsfgsvkslismpcfmshasipaevreargltedlvr**ISAGIE DVDDLISDLDIAFK**tfpl*
Nodulin-related	A. thaliana	20	13%	2	100%	Excelente	NCBI	mdfftdqvkkkfsdkkpessdpepnhnknkpghteptthkpghgeptthkpvsntdptthrpatnaelmasak**IVAEAAQAAA R**hesdkldkak**VAGATADILDAASR**ygkldeksgvgqylekaeqylhkyetshshsstggtgshgnvgghgggagapaak kedeksgggghgfgdyakmaqgfmk*
Chain A, crystal structure of a m-loop deletion variant of Ment in the cleaved conformation	Tomateiro	44	3%	1	89%	Excelente	NCBI	mggshhhhhhgmasmtggqqmgr**DLYDDDDKDR**wgseleismeqvsasignftvdlfnklnetnrdkniffspwsissalal tylaakgstaremaevlhfeqaenihsgfkelltafnkprnnyslrsanriyvektyallptylqlskkyykaepqkvnfktapeqsrkeint wvekqteskiknllssddvkattrlilvnaiyfkaewevkfqaektsiqpfrlsknkskpvkmmymrdtfpvlimekmnfkmielpyv krelsmfillpddikdgttgleqlereltyerlsewadskmmtetlvdlhlpkfsledridlrdtlrnmgmttafttnadfrgmtdkkdlaiskv ihqsfvavdekgteavaataviis*
ASR4 (Fragment) ABA/WDS induced protein	Tomateiro	33	11%	2	100%	Excelente	SOL Genomics	maeekkhhfgglfnhhknkeedtpiekttyeettyedsek**TSTYGDNTYGEK**tsygddtygkktttygddnkysek**TSYG DDTYDEKTNTYGDENK**ygektsysegddnkygektsyggdtygekptsyggdntygektsygggdenkygektsygek asygggddnkygektsygneeggygggvgayssetttnyeenddsgtktsedykeekkhhkhleelgglgavaagafalhekhkaekd pehahkhkieeeiaavaavgaggfafhehhqkkeakeeeeeaegkkkhhff*
Thiazole biosynthetic enzyme Thiamine biosynthesis Thi4 protein	N. benthamiana	38	10%	2	100%	Excelente	SOL Genomics	mastlassvisktnfidthkssfygvpissqtrlkivkstpqnmsvsmsadasppydlgsfsfnpikesivaremtrrymtdmityadtdv vivgagsaglscayelsknpnvqvaileqsvspgggawlggqlfsamvvrkpahlflnelgidydeqdnyvvikhaalftstimskllarp nvklfnavatedlivkngrvggvvtnwslvsqnhdtqscmdpnvmeakivvsscghdgpmgatgvkrlrsigminsvpgmkaldm naaedaivrltr**EVVPGMIVTGMEVAEIDGAPRMGPTFGAMMISGQK**aahlalralglpnaldgtaetsvlp elmlaaadeaeiada*
THI1; protein homodimerization	A. thaliana	37	14%	3	100%	Excelente	NCBI	Maaiastlslsstkpqrlfdssfhgsaisaapisiglkprsfsvrattagydlnaftfdpikesivsremtrrymtdmityaetdvvvvgagsag lsaayeisknpnvqvaiieqsvspgggawlggqlfsamivrkpahlfldeigvaydeqdtyvvvkhaalftstimskllarpnvklfnava aedlivkgnrvggvvtnwalvaqnhhtqscmdpnvmeakivvsscghdgpfgatgvkrlksigmidhvpgmk**ALDMNTAE DAIVR**ltr**EVVPGMIVTGMEVAEIDGAPRMGPTFGAMMISGQK**agqlalkalglpnaidgtlvgnlsp elvlaaadsaetvda*
Putative 1-acyl-sn-glycerol-3-phosphate acyltransferase	Tomateiro	36	3%	1	88%	Baixa	NCBI	mvdflpgflkgglsfllylintivlcsaiflvaffk**FVIPIPWFTR**qcdrlligisslwitinsltskffnriewdvrgvsdldpngwylvi snhqswvdilvlqtvfnrripflkfflkkeliwvpflglawwaldfpfmqryskkflekhphlrgndlkrtrracekfktipvsvmnfvegtr fselkhsaqgsefksllkpkaggvafvlgamgdclhkvvnvtivypdgadsfwgfisgrvrrivldvevlpvtsemtgdffnnstfrddfct winrlwkekdrkilairdglrdskstvnqg*
Lipid-associated family protein (PLAT-plant-stress domain-containing protein)	A. thaliana	20	10%	1	75%	Bom	NCBI	mmprrdvlflslllviatvsavaladdeadcvytfflrtgstfkagtdsiisarvydkygdyigir**NLEAWGGLMGPGYNYY ER**gnldifsgkapclpspvcslnltsdgsgdhhgwyvnyvevttagvhakcsyqsfdveqwlasdtspyelsavrnncpvslresvgrv gseirktlswiv*

Tabela 1 (cont.)

Dihydrolipoyl dehydrogenase Pyridine nucleotide-disulphide oxidoreductase, class I	A. thaliana	53	2%	1	98%	Excelente	SOL Genomics	maigslarrkastilssrylkysfslsrgfasgsdendvvvigggpggyvaaikaaqlglkttciekrgtlggtclnvgcipskallhsshmyh eaqhsfashgvkfssvevdlpammaqkdkavagltrgieglfkknkvnyvkgygkflspsevsvdtveggntivkgkniiiatgsdvk**S LPGLTIDEK**rivsstgalalteipkklvvigagyiglemgsvwgrlgsevtvvefasdivptmdgevrkqfqrslekqkmkfmlkt kvvsvdtvgdsvkltlepaaggeqttleadvvlvsagrvpftsglgldkigvetdkagrilvnerfasnvpgvyaigdvipgpmlahkaee dgvacvefiagkeghvdydlvpgvcythpevasvgkteeqvkalgvdyrvgkfpflansrakaiddaegivkviaekgtdkilgvhims snageliheavlalnygassediartchahptmsealkeaamatydkpihm*
hypothetical protein LOC100167534 [Acyrthosiphon pisum]	Tomateiro	214	1%	1	87%	Excelente	NCBI	matevgsdialiqdedmlrkmwqqtedfgkkkeirmrmyklreqrlkefyttgemiqtettskrsstrthtesitdqgfmtmktkeirdse sptedyqrqsnanknyyvsskidesdgnrgiveieqasslkpneeytqnslqtsmsssts tkwkssqttisvseevqk**AEVNLNSK**f ivneqnnedsqltydtsynnsdssnidskvtesisttnvdnikrddnnkysnvhqkveetytsngeqfepvsenisnadsnfientyisnis snksehfeiatdnkiinelqkldsylstqqsttdsstksnsivsnddkiikkvedkvidtskehiegqyitthqhsyqpprisvdlspshdafa kslrstpersspspsrerisperrfksaspekyraspeksgspskslnystrrrlsstqtkglskkrsitptrgekydssddsdcsgathgtydky kgcdakktlfkeekkitstskygqkssstspvrrekspgyssegsvgkevrksspnvknsshesspersafkpeqvcisveqvdtekttqp tssndinksprdkspsksvrkenktkqdefieneksnvvkqsniindvitikpkekspvkeikpvtkdtkkyptvkssenisytsdiksvtp skktistvpssipsptkkvvpstlektpskknlihkdskdtllnktvkkdlsrdkmdikitrsdsktlihknskeritpkmvvkkpshellvd kkttissrqkvtspetktktletikilekkdsknklkpsispqssqkklvlvksiaetkivdskkttgniyskripslnvtkprthtspksdinetp trnelppdnfesdtdlddstnkpkyiknissnssstssssdeddedrddirkiqdidnirieaeeeygkkmtnkdallnvivqlppssressp eysarfgqpycsvsddaslpryadvvsepedandyslhsnrydvvtdldeesnvtvadrvskflnnvnkqeeiqkteipqspqavrktk qifesiakgkteetdindviksetmnikneevhdlskinntpssvltrkisgasdyktrkdffenkksndisekvthltptkitrsssikdrrasf etkidkktplkdktnvprtkspesksssppdrttktpikisktetfdvkveeinrsrrlsgsksvkdrtatfekneptqklykenkttfrhqvpntp lntskisetriktatvtdtktadrvlskrvrspekpsatdsiiktnestskkvpskspdrkitetpkrnstakspdrqitetpkrnstakspdrqitet pkrnstakspdrqiteapkrnsmtkspersaliktkedtsenrstrdtissssrqratittaktefgvttknsstpkttlttttqsisiqnaddeveie eifdlhiletmlekavgydrrrriraqirvvkrqletnsgkatntttrvtknfqdvlpaktqkqsnvtvektenvkkfevkkersvspqktyan pignsrspspqktlrsvvetrtplptqkavrtienersnssqktvisvekvqplssqkitpiekvrspspqkitstekvrspspkkisvvqkqq dsrfnkneksrnstitststtvtkttkvkseqvsssstnlyqeqnvtdlitssygvgptddngrplfglralrktnanqsvinglltflrcallsltlv vsspetdssssvhyeaesrdtqasdqknrsfglralqsentytvqessstvyqtasdqkggvdqlitdhkgkplfglkalqsigkneeepiyd dmpeppvspqlkdlvlkhekhakessklesqprqkpkakfrdsfildtkdeglystfeengftdpneitslrsiikkseddqnnetivksks qstvfqsksvmrsdgsgnvsvtqdflkg*
Intronic ORF	Tomateiro	37	7%	1	89%	Baixa	NCBI	lyiwpiwyeiinlyfaicwkllliikykfklitiyisllynnilyklailvkiliyynnnqqitnnmikiikryksylvgisetirknsdldkk**DI KFNQWLAGLIDGDGYLGVSK**agytsceitvaledekalrqiqnkfggsikirsgikairyrlhnkegmiklinaingnirn tkrlpqlhkvcsilnlpiiepinltinnswfigffdadgtinyyfknnypqltisvsnkkyidlkpfmdifggniyydksnngcykwsiqskt dilffinsyvkfnssrttkfnklmlcskyydlinikawlifnknfnske*
RNA-binding glycine-rich protein-1b RNA recognition motif, glycine rich protein	Tomateiro	12	11%	1	98%	Bom	SOL Genomics	maaeveyscfvgglawattdrtlsdafstygevvdskiindretgrsrgfgfvtfkdeksmkdaisgmngqeldgrnitvneaqar**GSG GGGGGYGGGR**regggggygggggyggdry*
DDT domain-containing protein WSTF/Acf1/Cbp146	Tomateiro	93	1%	1	98%	Bom	SOL Genomics	mifvrmfnhiearlssnsrrkvslwrelnniyreiesigkhvpnhhrflslppstsntrkttrisnnpyphnlllnlglkmplfkr**KPFAL AEKPK**dlkpnelvfqvrftkeifrdygeylrrinlyrhrvwtcktgkhnltyeealvsekkaaeevqkfpeelvapvlrdvqfsmlslk dlgdavaqklqgclsegselygrkndhvysckiervvkdgektryevawldkyermpedtaideedlircklrfsraflksfirestyrsipw vlherlakkhgiptdppddllkdqfsmqdgvvvvnrkrkksedsetkengfqdpyeigredklkaqstrypiddllvqytefdkqlaerpp pcrefnvpmecvgdhlmvwdfctsfgrllhlspfsledferavcqkgsnivlitechsallrlllkdnseysiavqkkrpklkitlitwteylsd fleligiveltnhvatikrghyglldikaklailrelvcrvletdyfkekldediekqyalaatrreeileesrkkredhlkiqsngkeatkgrgns sdtgsdnhlrengdmpssngkqtspskhslenseseltisslknskkrkvdvknstakmnasskiasdklikdegketlenrsmdqraah kmrkneikehlenrskeqrkeylereiekrvirtnplgkdrdynrylffrrdgrifvessdslewgyyssqeeldaligslnvkgereralkkq lenlyhkislelqkrskeaqkaetddadvrrstrvrappgdnpalaflkyvnkwked*

Tabela 1 (cont.)

Acriflavin resistance protein	A. thaliana	112	2%	1	64%	Média	NCBI	mnlsapfirrpvatillslaltlsglfayrllpvaalptvdfptisvtaqlpgaspesmaasvatplikefstipaiatisatnyqgftsitiefelsrni dqaaadvqaaitrtlrrlpveltippsfrklnpadapvillavssdttplptldafaqtvispslstitgvgqvfvygsqkfavriqfdpnvlaargi gvdevqaaisnantstpvgvvegrgqnltiqtntqlmnadafrdiivavrngrpvrlgevarvidsvennriaswkdgrralvlavqrqpda ntvevvdrvramlprfeeqlggagtiavvndrsvsireavhdvqftlltiglvvvviylflgrlaatlipsiavpisiiatfgamyllgysidnisl lgltlavglvvddaivmlenivrhveqgekpfeaalkgageigftilsitvslvavfipvllmggvvgrifnefaivvtiaivasavisvtltpml aar**LPAGAAGHGAGAEGHGAR**rnlfergfarlqagyergvdlclrhqgavmlvflgsvaltaamfaaipkgffppedigq lqiatearqdisfeavaalqheaerviarnpavahvvnragsngftgtlnqgaffvelvergarpplartvaelrrdlaevpglsafitpvqnlrl ggrqsksqyqyvvqgldraelerwtvrltdamaqdratldratftgvtndlqnaalqatitvdtdkaqalgitsdqirqtlytgfgtrqistiygt adsyqvitefdprlqwtadrldeirlraasgklvplsavatvartpgplsinqlgllpavtisfdvppgvalgqavsriaeikqslglpgtlsttfa gtaqvfqdalanqglllaaavltiylvlgilyesfihpitiltglpaaavgalgalqlfgmdlsviaiigvlmligivkknaimmidvalvlqreq gwtpraaireacllrfrpimmttaaavmgtlpiaighgasaelrqplgvavvggllvsqlltlfitpvlylamdrlasgttrlllalrrgrlrgprtg rlpae*
Acetyl-CoA carboxylase biotin carboxylase subunit	N. benthamiana	49	4%	1	56%	Baixa	NCBI	mirkilianrgeialrvlrackelgiktvavhstadadamhvrladesvcigpppvrdsylsiqqiiaaceitgadavhpgygflsenakfad vleahkitfigptaahirimgdkieakktakklgipvvpgsdgaiteeaealriaheigypviikasaggggrgmk**VVRSEQEFAI ALETTR**seaaaafgddavyiekylekprhieiqvigdgagnaihlgerdcslqrrhqkiweeatspalneierkkigdivanacaqlgy rgagtieflyengkfyfiemntrlqvehpvteaitgidlvheqihiasgcrlsvvqedicfsghaiecrinaedpinftpspgtithfhtpgglgi rvdsgaysgycippyydsmigklivhgrtrlecmmrlrraldefvvdgikttlplfrdlinnqdiadgnynihwlekylssqstssds*
ATP-dependent RNA helicase A-like protein Helicase-associated region	Tomateiro	116	1%	1	95%	Excelente	SOL Genomics	msvrllnhphssallrfhfcntslrsslqrsysfsrsvmayrpnyrggrrggggrsgggrgggggrgggggrggeqrwwdpvwraerlrqq aaemevmnenewwgkmeqfkrggeqemvirrnfsrddqqklsdmayqlelyfhaynkgkalvaskvplpsyradlderhgstqkei rmsteieervgnllsssqdavsagtssstsgtsakllskavettkpklsieddiatkrlnvelkqkqektrgsekvkemisfreklpafkvksef meavannqvlvvsgetgcgkttqlpqfileeeisslrgvdcniictqpr**RISAISVAAR**issergdslgdtvgyqirleakrsaqtrllfc ttgvllrrlvqdpdltgvshllvdeihergmnedflliilrdllprrpdlrlilmsatinaelfskyfrdaptihipgltypvaelfledvlektrylik seadnfqgnsrrrmrqqdskrdpltdlfedvdigshykgysmttrqsleawsgslldlglveasieyicrcegegailvflsgwdeisklldk ikannflgdarkflvlplhgsmptvnqreifdrppantrkivlatniaessitiddvvyvidcgkaketsydalnklacllpswiskasahqrr gragrvqpgvcyrlypklihdamaqyqlpeilrtplqelclhikslqfgaiesflakalqppdalsvhnaiellktigalddteeltplgrhlctlp ldpnigkmllmgsifqclnpaltiaaalahrdpfvlpinrkeeadaakrsfagdscsdhiallkafegwkdakrygkertfcwenflspvtlq mmedmrnqfidllsdigfvdksrgakayneysndlemvcailcaglypnvvqckrrgkrtafytkevgkvdihpasvnasvhlfplpyl vysekvktssiyirdstnisdysllmfggnltpsksgdgiemlggylhfsasksvldlikklrveldkilkrkieephfdvsvegkgvvaavv ellhsqdiry*
Plant neutral invertase	Tomateiro	16	9%	1	95%	Bom	SOL Genomics	mk**NGALSLAVTLRIR**dpwsyhnggswptllwqftlacikmgrpelaqkavdlaekrlsadhwpeyydtrhgrfigkqarlc qtwtiagyltskmllqnpdmasklfwnedyellencvcalrpngrrkcsrsaarsqvgl*
FIP1 protein	Tomateiro	23	5%	1	95%	Excelente	SOL Genomics	mkdliggnmvsvpmssavcsferqpkrllslsscqdhipsatskilsklkqlsispkltetvkgklslgakilqvgglekifkqkfsvrddek llkvsqcylsttagpiagllfistdkiafcskrsik**LSSPTGKSLR**irykvsipiskinkakecgnmekpsqkyiqlvteddyefwfm gffnnqktlrylqqaistftnfnqlrr*
Hypothetical protein	Tomateiro	12	11%	1	60%	Baixa	NCBI	mspaaclpfpsswdrlvfplcrpcsrslaadsqlrrsprprrvsvptr**HARTVSCASSVR**tsvhpvlvlassrrssgcrhprltvrpr dspvgarpcaplarrapvrrrs*
LIM domain protein Zinc finger, LIM-type Zinc finger-like protein, C2H2-type	Tomateiro Tomateiro	22 27	5% 3%	1 1	95% 95%	Excelente Excelente	SOL Genomics SOL Genomics	mafagttqkcmacdktvylvdkltadnriyhkacfrchhckgtlklgnynsfegvlycrphfdqlfkqtgsldksfegtpkivkpqklidse kpqvak**VTSMFGGTR**ekcfgckntvyptekvsvngtpyhkncfkcshggcvislsnyiahegrlyckhhhiqlikekgnlskle gdhetipaittevtaes* masfggttqkckacektvylvdqltadnkvyhkacfrchhckgtlklsnynsfegvlycrphfdqlfkmtgsldksfegapktvrersadq gsnsk**VSSLFGGTQDK**cvackktvyplekvavdgtsyhrpcfkcshggcvispsnyvahehrlycrhhhtqlfkergnfsqme dhekn* mektkdketqdfmnvesfsqlpfmrptkekaairlfgkellgttirhedqsieihdsvgdtteinnnnrkfecqyccrnfptsqalgghqna hkrerqhakraqyakyqynayfnsalyntsstssttsgglyvnnnghshyysqqinrdinenhr**SLSALWR**vphvhhssissssnsy spnnvsnfvrpivynnvngdlkkintksssislтrfgyelkegvhqdhvsldlhl*

Tabela 1 (cont.)

RXT3-like protein C1259.07 Histone deacetylation protein Rxt3	Tomateiro	92	1%	1	95%	Bom	SOL Genomics	msgtpnkrphedggnggssnhsyssapkyshddsgafpkvmssgtpeyhasfdvgqnarmpkiqrtessrdadrrspvlpmyrvsscpvvshpdhsvasenrlepkevnkdvkvenrdakseirelyqgtksdkddrfenraddgkdiknsrdtypeykgdvktdk**DRFSGVSWK**dpkeqtrgkrypdlpvpvgnmdpwhasrthgaaeigkevsnsenrdfakvreavaenkmdlkgddkykdkerkrkegkhrewgerdkerndcrnnlqlgnstsdnkellkeereserwekerndlskdkdrpkdwekdhakrevwngverevlqsekevidvpgktnepenstveqkkqkdhdnwkntdrdgserrkerdtdlegerpekrvrchdkepeegdldtegggerereafnygvqqrkrmsrprgspmanrdprfrshthenegsqvkhdvsavnyrvgecmpeliklwkeyesskadeasdsspsdptleiripaehvsatnrqvrggqlwgtdiytndsdlvavlmhtgycrttaspllptitelratirvlppqncyistlrnnvrsrawgaavgcsyriercsvvkkgggtidlepclthsstleptlapvavertmttraaasnalrqqrfvrevtiqfnlcnepwlkysisvvadkglkkalftssrlkkgevlylethskryelcfsgekmvkattslmhemdvdkpqshnlhmangeknqvngentmvdmfrlsrckkplpqklmqsvgiplplehvevleenleweniqwsqtgvwiagkeyplttrahflspn*
NADH dehydrogenase (ubiquinone) flavoprotein 1	A. thaliana	54	2%	1	95%	Bom	SOL Genomics	m**APIKGLLSLQR**talaqrsserwgqygrlfstqaastastpqptpppppperthfgglkdedriftnlyglhdpylkgamkrgdwyrtkdlvikgsdwivnemkksglrgrggagfpsglkwsfmpkttdgrpsylvvnadesepgtckdreimrhdphkllegcliagvgmraraayiyirgeyvnerlslqkarqeayeagllgknacgsgydfdvyihfgagayicgeetalleslegkqgkprlkppfpanaglygcpttvtnvetvavsptilrrgpewfasfgrknnagtklfcisghvnkpctveeemsislkellerhcggvrggwdnllavipggssvpllpknicedvlmdfdalkavqsglgtaavivmdkstdvvdaiarlsyfykhescgqctpcregtgwlwmimermkvgnakleeidmlqevtkqieghticalgdaaawpvqglirhfrpelerrirehaerelqqaaaa*
Afadin- and alpha-actinin-binding protein	Tomateiro	45	2%	1	95%	Bom	SOL Genomics	mpptdndydlrrastppsgmgisecnfadinnlehcakylnqtlvtfgfpasldlfahdpvsiartcncvyallqqrqrdiefrestneqrqrllsdisrleakverlesqlqvkdreiatitreeakataalk**AQIDKLQK**erdefqrmvlgnqqvrtqqiyemkkkekeyiklqerlnqvmmekkkesrsgmeimnllqkegrqrgtwngkktdndfykkivdayeaknqelaaenadlrallrsmqadmreflnapngssrqsssserldtdhlqsplggrtdvfdlplhmardqieeslrnkmasikermgqlqdaqkgaevtseatereleleaqlvearsiiqeqasimskhltksekprrlsghmnserdllisspteg l*
Os07g0175100 protein (Fragment) Homeodomain-like containing protein	Tomateiro	57	2%	1	95%	Excelente	SOL Genomics	meissflfpnqedypspnffsffqdfdfpttdtlapivaeplpkkqrvddfdfdleqvveqqgslksvedilnkflgfdkeedktnlkqftnqnqsvfdfsnqqptgliminekvkpmastnnkrsrqgsaefistseeesqpqqqrrlwvkdrskawweqcnspdfpeeefkkafrvsratfdmiceelesvvtkkdtmlrqaipvrqrvavciwr**LATGEPLR**evskrfglgistchklvlevctaikgvlmpkfvqwpnedykmneiksefqmlsgmpnvggsiytthvpiiapkvsvaayfnkrhternqktsysvtvqgvvdpkgvftdvcigwpgsmsddkvleksalyqranrgqlkdtwvvgnsgyplmdwvlapytrqnltwtqhafnekvgevqkvakeafmrmkarwsclrkrtevklqdlpvvlgaccvlhnicemrgeqlnpglrfdlfddemvpenivrsmnavqardqiahqllhqnhagtnfl*
V-type proton ATPase V1/A1 complex, subunit E	Tomateiro	28	3%	1	95%	Excelente	SOL Genomics	mndadvsk**QIQQMVR**firqeaeekaneisvsaeeefnieklqlveaekkkirqeyerkekqvdvrkkieysmqlnasrikvlqaqddlvntmkeaaakellnvshhehgiidsilhhhhggykkllhdlivqsllrlkepcvllrcrkhdvhlvehvlegvkeeyaekasvhqpeiivdeihlppapshhnmhgpscsggvvlasrdgkivcentldarlevvfrkklpeirkclfgqvaa*
Cysteine synthase A	Tomateiro	38	5%	1	95%	Baixa	SOL Genomics	maslvrrrfysseasfvqrlrdlpkylpgtkiktqvsqligktplvylnkvsegcgayiavkqemmqptssikdrpafamindaekkglitpgkttlieptsgnmgismafmaamkgykmiltmpsytslerr**VTMRAFGADLVITDPTK**gmggtikkaydllestpnaymlqqfsnpantqvhfettgpeiwedtqgnvdifvmgigsggtvsgvgqylksknpnvkiygiepaesnvlnggkpgpheitgngvgfkpdildmdvmeevlmvsseesvnmarelalkeglmvgissgantvaalrlanrpenkgklivtihpsfgerylssvlyedlrkeaqnmqpvsvd*
Glyceraldehyde-3-phosphate dehydrogenase	N. benthamiana	44	4%	1	94%	Excelente	SOL Genomics	masaalsvanssvhvsknkgsfseftglrtssavpfgvgwktnvdllslvayqtsvisgagknnnnnsrvveaklkvaingfgrigrnflrcwhgrkdspldviaindsggvkqashlllkydstlgifdadvkpvgtdgisvdgkiikvvsdrnpinlpwkdlgvdlviegtgvfvdregagkhiqagakkvlitapgkgdiptyvvgvnadlyspdesiisnascttnclapfvkvldqkfgiikgtmtttthsytgdqrlldashrdlrrar**AAALNIVPTSTGAAK**avalvlpslkgklngialrvptpnvsvvdlvvqvskktfaeevnaafresadnelagilsvcdeplvsvdfrcsdvsstvdasltmvmgddmvkviawydnewgysqrvvdlanivanqwi*
Pistil extensin like protein (Fragment) Pollen Ole e 1 allergen and extension	N. benthamiana	27	5%	1	94%	Excelente	SOL Genomics	mdfvptkslgfilaflitsftvlgqndagivaselaphslvssppveapkphkgghhhhlkhhsqppaspppshsspppvipttppahsptkspsppvkppahspvkppahppvkppsplpvrkfvgvrgvvyckackyrgvdtllgaspiqgavvklacnntkyhltslgttdkngfffiqpkwlttagyhkckvflakspkaecsvptnfhngqsgamlipappspmtlsepevk**LFNVGPFAFEPSK**klpcktl*

Tabela 1 (cont.)

Serine/threonine-protein kinase TEL1 PIK-related kinase, FAT	N. benthamiana	441	-1%	1	94%	Excelente	SOL Genomics	mspiqdfeqhsrhlyeadlpiqtrlqmamevrdsleithtgeylnflkcyfrafsgvlyhitkpqfsdnpehklrnivveilnrlphsevlrpf vqellkvamhvlttdneengliciriifdllrnfrpslenevqpfldfvckiyqnfratvsyffesgamavppppmptssvsslgeadvkpm evsdqmstsngyfgagqlnpstrsfkivtesplvvmflfqlygrlvqtniphllplmvsaisvagpekvpphlkthfielkgaqvktvsflty llksfadyikpheesicksivnllvtcsdsvsirkellvalkhvlgtdfkrglfplidtlleervlvgtgracfetlrplaysllaeivhhvrgdlslsq lsriiylfssnmhdaslslsihttcarlmlnlvepifekgvdqqsmdearillgrildafvgkfntfkrtipqlleegedvkgrstlrsklelpvqa vlnlqvpvehskevndckhliktlvmgmktiiwsithahlprsqvsastqgtppqvlssastsssvpqpfkgmredevwkasgvlksgv hclalfkekeeeremihlfsqilaimeprdlmdmfslcmpelfecmisntqlvhifssllqapkvfrpfadvlvnflvsskldvlkhpdspa aklvlhlfrflfgavakapsdcerilqphvhvimetcmknatevekpigylqllrtmfralaggkfelllrdlismlqlclsmllallegpnged mrelllelcltlparlssllpylprlmkplvmclkgsddlvslglrtlefwidslnpdflepsmanvmsevilalwshlrpapypwggkslqll gklggrnrrflkeplaleckenpehglrviltfepstpflvpldrcislavaavmqrsaivdsfyrkqalkflrvclssqlnlpgsatddgftsrm lstllvssvdpswrrsetsdikadlgvktktqllaersvfkillmtiiaasaepdlhdskddyvinvcrhfaiifhiessaahgtlsaapvgasvls sssivsaksrystssnlkeldplifldalvdvladenrlhakaalnalnvfaetllflarsk**HSDVLMSR**ggpatpmmvsspsmspvy spppsvrvpvfeqllprllhccfgctwqsqmggviglgalvgkvtvetlcafqvrivrglvfvlkrlpvyatkeqeetsqvltqvlrvvnnvd eansearrqsfqgvveyfalelfnpnvsinvrrivqsclallasrtgsevselleplyqpllqplvgrplrsktveqqvgtvtalnfclalrpplkl tqelisflqealqiaeadetvwvikfmnpkvanslnklrtaciellctamawadfktqnqselrskiismffkskltsrtseivavakeglrqviq qqrmpkellqsslrpilvnlahtknlnmpllqglarllellanwfnvtlggkllehlrkwlepeklaqcqkswkageepkiaaaiielfhllps aagkflddlvtltielesalppgqfyseinspyrlpvtkflnryptaavdyflarlcqpkyfrrfmyiirsdagqplreelakspekiiasafpefi aksdasagqeslsrpststgdeglgtpqveasipsastnvapqdayfqglslvktlvklmpnwlqnnrcifdtlvlmwksparisrlqneqe lnlvqvkeskwlvkcflnylrhdkteinvlfdilsiflfrtridftflkefyiievaegyppnmkrtlllhflnlfqsrqlghdhlvvvmqmlilp mlahafqngqtwdvvdsaiiktivdkllddppeevsadydeplriellqlatlllkylqtdlvhhrkelikfgwnhlkredsaskqwafvnvc hfleayqapekiilqvfvallrtcqpenkmlvkqaldilmpalprrlplgdsrmpiwirytkkilveeghsipnlihifqlivrhsdlfyscraq fvpqmvnslsrlglpynttaenrrlaielaglvvnwerqrqsemkivpandgtgqnadglshasagsvdpkhpdgssfsedpskrvkve pglqsicvmspggassipnietpgsgggqpdeefkpnaameemiinflirvwffpemvalviepkdkeaslmykqaldllsqalevwpn anvkfnylekllnnlppsqskdpstalaqgldvmnkvlekqphlfirnninhisqilepcfkfkvldagksmccllkmvyvafppegsntt qdvkmlyqkveeliqkhlaavatpqtsgednsgsmvsfvlyviktlaevhknfiepvnlvrllqrlardmgssigshvrqgqrsdpdsavt ssrqgadvgvvianlksvlgliservmaipdckrpvtqilnslllsekgtdssvllsildvikgwieedmtkpgvsiasstflspkdvvsflqrls qvdkqnftpsaaeewdkkyiellyglcadsnkyahslrhevfqkverqyllgirakdpemrmkfftlyheslgrmlftrlqyiiqiqdwea lsdvfwlkqgldlllailvedksitlapnsakvpplvvagtigdsigpqpmvldvpegseeapltvdsfiakhaqflnemsklqvadlviplr elahtdanvayhlwvlvfpivwvtlhkeeqvalakpmitllskdyhkkqaahrpnvvqalleglqlshpqprmpselikyigktynawh ialalleshvmlflndtkcseslaelyrllneedmrcglwkkrsitaetraglslvqhgywqraqslfyqamvkatqgtynntvpkaemcl weeqwlccasqlsqwdvlvdfgkmvenyeilldslwkqpdwaylkdhvipkaqvedspklriiqsyfslhekstngvaeaentvgkg vdlaleqwwqlpemsihakisllqqfqqlvevqesariivdiangnklsgnsavgvhgglyadlkdiletwrlripnewdsssvwydllq wrnemynavidafkdfgstnsqlhhlgyrdkawnvnklahiarkqglyevcvsvlekmyghstmevqeafvkireqakaylemkgel tsglnlinstnleyfsvkhkaeifrlkgdfllklndceganlaysnaislfknlpkgwiswgnycdmayketheeiwleysvscflqgikfgi pnsrghlarvlyllsfdtpnepvgrsfdkyleqipnwvwlswipqlllslqrteaphcklvlmkvatvfpqalyywlrtyllerrdvasksey grmamaqqrmqqnvsganaaapmgladgnarmtgqsggssagenhtpqgaqsgggvgsqdgnssqiqeperpdgnmpsgndqs lhqgssgndgggqaalrrnsalslvasaasafdaakdimealrskhsnlageleillteigsrfvtlpeerllavvnallhrcykyptattaevpqs lkkelsgvcracfsadavnkhvdfvreykqdferdldpdsaatfpatlselterlkhwknvlqsnvedrfpavlkledesrvlrdfhvvdvei pgqyftdhevapdhtvkldrvaadipivrrhgssfrrltligsdgsqrhfivqtsltpnarsderilqlfrvmnrmfdkhkesrrrhicihtpiii pvwsqvrmveddlmystflevyenhcarndreadlpitffkeqlnqaisgqispdavvdlrlqayneitksfvtesifsqymyktlvsgnh mwafkkqfaiqlalssfmsfmlqiggrspnkilfakntgkifqtdfhpaydangmiefnepvpfrltrnlqaffshfgveglvvsamcaaa qavvspkqsqllwyhlamffrdellswswrrplgmplatvvgagnlnpvdfkqkvttnvenvigritgiapqyiseeeengmdppqsv qrgvaelveaaltprnlcmmdptwhpwf*
Ferrochelatase	N. benthamiana	55	3%	1	94%	Bom	SOL Genomics	metgtisrllkptkvpsssfssftsnssiccsrtrlmppfhsshggremnsdereitilsseltlkkssvveglslhrpvhkrdvvgktfcsvgls tfpgsivesssqtteekigvlllnlggpdtlhdvqpflfnlfadpdiirlprlfrflqrplaqlisvlrapksk**EGYAAIGGGSPLR**kitd eqasvlkialetkevpanvyvamrywhpfteeavhqikrdgitklvvlplypqysisttgssvralqnifkgdsylsrlpvavieswyqrqg yiksmanlieqelhnftkpeevmiffsahgvpisyvenagdpyrdqmeeciflimkelkareinndhtlayqsrvgpvqwlkpytdevl velgqkgvksllavpvsfvsehietleeidmeykelalesgienwgrvpalnctssfitdladavievalpstmamststgteeevdnnpmq yfmkmlfgsvlafvflflpkmvsafnknil*

Tabela 1 (cont.)

Urease accessory protein UREG putative urease accessory protein G	A. thaliana N. benthamiana	30	7%	1	94%	Bom	SOL Genomics NCBI	massdhhmndhhhhhhhdhdhnhedtkatswvgadgkvyhshdglaphshepiyspgyfsrrapplndrnfseraftigiggpvgtg ktalmlalckilrekyslaavtndiftkedgeflikhgalpeeriravetggcphaairedisinlgpleelsnlykadillcesggdnlaanfsr **ELADYIIYIIDVSAGDKIPR**kggpgitqadllvinktdlapavgadlsvmerdalqmrdggpfvfaqvkhgvgvedivnh ilqswevatgnkkr*
Universal stress protein (UspA)	A. thaliana	26	5%	1	94%	Excelente	SOL Genomics NCBI	mtspkkpppesdrpltaimhqpasprfhlgtpnsganrkvaiavdlsdesayavkwavqnylrpgdavillhvrptsvlygadwgaidv svdtaneesqrkleddfdnftttksndlaqplveanipfkihivkdhdmkerlclever**LGLSAVIMGSR**gfgasrrspkggmlgs vsdycvhhcvcpvvvvrypddkdattgndsvtdvlhpvpeeepiyhdasdktadsekas*
Lipoxygenase homology domain-containing protein 1	A. thaliana	20 24	5% 4%	1 1	94% 94%	Excelente Excelente	SOL Genomics	mgvaaqvnqmwfnlmivlffvsissisaedcvytayirtgsiikagtdsnisltlydangyglriknieawgglmgpgynyfer**GNLD IFSGK**gpcvngpickmnltsdgtgphhgwycnyvevtvtgakkqcnqqlftvnqwlgtdvspykltairnncknkyesgelkplyd sesfsivdvi* mgvstqlnyfwfnlfilffcfsissiygsevdcvytvyvrtsskikagtdsiisltlynangygirinnleawgglmgpgynyfer**GNLDI FSGR**gpclngpickmnltsdgtgsghgwycnyvevtvtgvhrecrqqhfeveqwiatdhspyqltfiknlcrktksgeylsvsaedlp asvenlsfsgedlpasvenlpvsgenlvvfdaavm*
Unknown Protein	A. thaliana	33	3%	1	94%	Excelente	SOL Genomics	milyraeadvcdafisfledswtfesshikqkqkqvndeldkcedylvnlaihvlsmykdelgpsianlrqlgenlkrpeiinnekl**ETV EVKR**rleeeymdaeaklvtvfsvvdgikrqfygqdgtisrkgdekmkelldtlqkleddfesinrpmleletlgktdaksketlegspss sptqptnddsasfvdalpriqtprpdhtsnpetvvltprrwmrenvrrslslainqttgtlessfkspkhrkgksmdpaaelsrlklelelends rihpseeindwcd*
ATP binding/serine-threonine kinase	Tomateiro	53	4%	1	94%	Excelente	SOL Genomics	msllikvgelnsscqaltsldhamlvaylisslflflsiinsfsillqsfsplkmkcfqfsnddhnktrkstsgqssasvftdsdlkgseyssqnv stdistvsfsclsnkkachlkvftveelksatknfsrslmlgeggfggvykgvlkdtniavkqlsqrglqghkewvtevnvlgvvehqnlv kligycaeddergiqrllvyeflpnrsvqdhlisrymsplpwetrlniar**DAARGLAYLHEGMEFQIIFR**dfkssnilldek wnaklsdfglarlgpsdglshvstavvgtvgyaapeyiqtgrltsksdvwsygvflyelitgrrpldrnkpkneqkllewvrphlsdlkkfe qildprldgkysiksaqklaaianrclvkhprnrpkmsevlemvnqvvkateakipqtpidesstpnvdddmfmvrcltasrrlvehtsre nkllvwklwkpklvssn*
PRLI-interacting factor M (cysteine proteinase inhibitor 6)	A. thaliana	23	8%	1	75%	Excelente	NCBI	vlgfcneemalvggvgdvpanqnsgeveslarfavdehnkkenallefarvvkakeqvvagtlhhltleileagqkklyeakvwvkpwl nfkelqeftpasdapaitssdlgckqgehesgwrevpgddpevkhvaeqavktiqqr**SNSLFPYELLEVVHAK**aevtgeaa kynmllklkrgekeekfkvevhknhegalhvnhaeqhhd*
Receptor-like protein kinase At3g21340	Tomateiro	84	3%	1	93%	Excelente	SOL Genomics	mllnmallhlmiliyslfsvsslqfteksianattitkpgcdskcgsltvpfpfgignqtgcsidpsfditcnvsfnppkaflsgkdievvdild dhilvknnvgsrcydqagalinddsinfslgstsfsfsdlnslmvvgcddlalilgyegrnftsgcislcskkediidgycsgigccqtsiptgl ksfvsltrslnnhtnvssfnpcgyaflgepdkfifkssdlsnstfrnkvieevpvvidwiignsdsctvakksadyacgensvcvdsktglgg yrcsckpgyqgnpyispgcidvnecenenpcdgicnnfpggysctcphgqigdgkkdghgcipknskspilqlslglcfgflalvisatw iylgikrrrlirlretffqknggllmltqklrsneggmkyaakiftaaelekstnnyaedrilgrggygtvykgvlpdkrvvaikksrtmdvsqi elfineviilsevnhrnvvkllgccleseypllvyeyiskgtlyyhihdggdqtrwfswenrlriaseaagalaylhsaastpvihrdvkstnill denytakisdfgasrlvsldqtqvttlvqgtlgylpeyfhtsqlteksdvysfgvvlaelltgrmpldttasekernlaaffvrsikenrlfqvle trvlregsfeqcqgvaelakrclr**LTSEERPTMKEVAMELEGLR**kftkhpwsktqqcqdeesigliteqtldlyainmntn fmsngefsgqqsldsrmmlqihspr*
Potassium channel Cyclic nucleotide-binding	Tomateiro	33	8%	1	93%	-	SOL Genomics	mktedlliyachsptrkriarkifefikkasknvicniiltfkstytmiddpkcialsktwlslgnlhdrsiwspyvmsmdwsvttlttgygd llhavateemmvymlfdlgltgcyssfaqknnlpvrleeqmldhlsmmhrtdtvglqqqeiletlprvdkaylfhgvshdvlfqlvsemk aeyfppkediilqneastdlyilvngavdlishrngmdqvvegrgclwrswycllnsiqeildigaek**FHISLTKVLTEDGALI EDIAVIR**dgdhlvlstsen*
Plant-specific domain TIGR01615 family protein	Tomateiro	48	3%	1	93%	Bom	SOL Genomics	mpfpmkiqpvdfstveessr**CDSFKPVPKSR**fkrlferqfsgllrssapeklvtgedlignkkdaseelepssvclakmvqtfiee gedkhrcnrsrcscfncngtesseeendsvncfgesnqncssdaceilksvvscpsvlernvlaeitkiieknrmvkrkdnfirkmvvdgll ameydasicesrwektpstpagayeyidvvtegerliididfrsefeiarstrsykcllqvlpnifvgkadrlqkivhllteaaklslkkkgmp cppwrkveyvkakwlstytrmtpvlmpiasnstsepgtacktkeaikeetsevfrgemnlvfggkssplad nnpksagtsplfacddek nlavklgdifegkssplaenntkrastsslfacdieknamaaqlgepleikpkdssncarkmtglt sliedht*

Tabela 1 (cont.)

Cellulose synthase-like C6 glycosyltransferase family 2	Tomateiro	61	4%	1	93%	Baixa	SOL Genomics	mdlniyfdhgitsqiglfwqqvrapllvpflkimvilclamslmlfvdkvymgiviifikilrkkpekkykwepikddlelanssypmvli qipmynerevyqlsigaacnlswpanriliqvlddstdpiikvmveqeckrwaskgvnikyeirdnrkgykagalkegmkhnyvk**M CDFVAIFDADFQPEFDFLLR**tvpflvhnpeiglvqarwkfvnsdeclltrmqemsldyhfkveqevgssnhaffgfngta gvwrisalneaggwkdrttvedmdlavraglkgwkfvyigdlkvknelpstfkayryqqhrwscgpanlfkkmvgeiasnkkvsiwk rfyliysfffvrkiiahmvtfifycviipttvlipevqvpiwgavyipstitllnavgtprsfyllifwilfenvmslhrtratfiglfevgrvnewv vteklgdgikaktstkgsrrsrfrigewynkvhfvevfvglylmlcgwydysfgknrfyiylflqgvaflvagfgyigvfvpns*
Peptide chain release factor 2	N. benthamiana	47	2%	1	92%	Excelente	SOL Genomics	makmaadpfsvrieatcsscplnlrssktatfriqahnnqhtddknrfykelgtfalkrkiedlvlraemlaptalefeearrlkqeeiireydl wddvaksnedlvqlaesakavdalkdlrykaeeaklitelagmdsinydflkqaytacvsvnktldkyemskllrepydmegacvtiesg segiyskiwaekltrmyikwaekqghkarivekqdsesggikyvmielefksaygylsgergihcmngssegkfdlskdgaaaidvipl flesspdlqidendleittsyeeergrtspsliihhiptglqvqstgersrfanklkalnrlk**AKLLIVMR**eqgvsnlasirssdmssswn qvirryvfhpnklvqdmktgvqlsdltavlngniepfigahinsrrh*
Ferredoxin-NADP reductase	Tomateiro	42	6%	1	91%	Baixa	SOL Genomics	**MAHSAISQVSVAVPLQTDSSFR**rstfkatsltfsdkscismpsidlkatrsrsqyivcmsvqhaskakvsvsplslenate pplniykpkepytativsverivgpkapgetchividhdgnlpywegqsygvvppgenpkkpgsphnvrlysiastrygdsfdgktaslc vrravyydpetgkedpsknggvcsnflcnakpgdkvkitgpsgkimllpennpnathimigtgtgvapfrgylrrmfmesvpmkfngla wlflgvantdslyddeftkylndypgnfrydralsreqknnkggkmyvqdkieeysdeifklldegahiyfcglkgmmpgiqdtlkrva eergesweqklsqlkknkqwhvevy*
Unknown Protein	Tomateiro	127	2%	1	91%	Baixa	SOL Genomics	**MQDKDAMNSAFFEGVVCEDR**ekisprisgwnreessremnsphldlpqvkqlanpivpiaqtreanleeiatkspesri dirptsssirrdsttgvcspkdedpisqkwkqmdgriagsensgdltvpyagingtstkeafsqaqntqvttlinrpnrgnlntndpasstnr nphehlgrhertgetsscgrgeerhssatdphdleaktmakdnnnhqkenfneqqdknkmnmdiqpaasaagtsnfsfamtstsshltp qlnagnspgkynnqdkggreekanehsheqakekqvqdhkqqhqpeeekrrleginldqsqvnaekqtdqtkqsytqggrksddqsn kareeyhnnfpkisnnysrydpklhndntrkqagqadqantrnpnnhlqgqqlnnnnkneqnsepapytviqsfaarlrynqaqketpit lnepvhttrqgfpavlidendyyvklaeickytlvgkftntmprmelvrksfilqtqlmggvkithfnsrhvyidldneldyqtvwtklkm niegqamriqawtpdftpeeetpvvpiwisipglpwhyynkvflttilesigkvlfldsptsqrtrgsttrvkatgmrsvqskeemrrlkkke leveknsmerssgklneqrkdqkedttktpnqpnkkqkgsneknedqtqprthmtsqqqkkiiqnqeeqgqqedrqeeqwqiqrrkq pkiqeqviskavwrpvtppmqgnnsshqheqgntgisilptqniftnlevqekqerqqagqdgikeaasqinqpnnknlndqpvknnp nnhkpphqqaagsnkskrtgidlslpipkapnflnvdagltdevvggtdgecqeiatnmqegdtkggnlphamheglvtdprtdlrasa mnsqqqreqtqqhvqnsgnrkqtakekieeqvekmkdkdqrkksvsmgtestpksknkpskqkrdaakrrqnkqqdkdfeqvqev reelcnkfvmvddnqglnilplqvqymtppsseppdkmqqkcrvntepimdeyavinsedeivgdnqsledsddndetseapirafsp hndqtleneiqqvtnsqclsprglhydrfqlrkqdantvtagrpntrlfssrssq*
Beta-D-glucosidase, family 3, N-terminal	Tomateiro	69	4%	1	89%	Baixa	SOL Genomics	mgrfsipmkgfvlfclwaviaeaeylkykdpkqpmvtrikdlmkrmsleekigqmtqierkvalpdvmkqyfigsvlsgggsvpapk asaedwinmvneiqkgalstrlgipmiygidavhghnnvynatifphniglgvtrdpdlvkrigaatalevratgipyafapciavcrdpr wgrcyesysedhnivrtmteiipglqgdlpansrkgvpfvegklkvaacakhfvgdggtvk**GIDENNTVINSNGLFGIH MPAYYNSIIK**gvstvmvsysswngekmhanrdlvtgflkdrlkfrgfvisdwqgidritspphanytysvqagvsagidmimv penykefidaltllvkdniipmsriddavkrilrvkftmglfenpladlslvnqlgsqehrelareavrkslvllkngkstsqpllplpkkapki lvagthadnlgyqcggwtiewqgvagndltvgttilsaikktvdpytqvvyqqnpdanfvesnkfdyaivvvgevpyaemmgdssnlt itepgpstinnvcgavkcvvvvvsgrpvvlepyvekidglvaawlpgtegqgvadvlfgdygftgklartwfksvdqlpmnvgdrnyn plfpfgfgltqavkln*

Tabela 1 (cont.)

Catalase isozyme 2	Tomateiro	57	1%	1	85%	Bom	SOL Genomics	mdpykyrpssafnspfcttnsgapvfnnnssltvgargpvlledyhlveklanfdreriaervvhargasakgffevthdiahltcadflrap gvqtpvivrfstvihergspetlrdprgfavkfytregnfdlvgnnfpvffirdgmkfpdmvhalkpnpkshiqenwrvldffshhpeslh mftflfddigipqdyrhmdgsgvhtftlinragkstyvkfhwkptcgvkslleeeairvgganhshatqdlydsiaagnypewklfiqim dpehedkfdfdpldvtktwpedilplqpvgrlvlnknidnffneneqlafcpsivvpgvyysddkmlqtrifsysdtqryrlgpnylqlpa napkcahhnnhydgsmnfmhrdeeidyfpsrydqvrhaevypipstvcsgkrekciiqkennfkqpgeryrsftpdrqerfirrwveal sdpr**ITYEIR**siwitywsqadkslgqklasrlnvrpsi*
Catalase isozyme B (Salicilic acid binding catalase)	Tomateiro	57	2%	1	89%	Excelente	NCBI	mdpykhrpssgsnstfwttnsgapvwnnnsaltvgergpilledyhlieklaqfdreripervvhargasakgffevthdishltcadflrap gvqtpvivr**FSTVVHER**gspetlrdprgfavkfytregnfdlvgnnmpvffirdgmkfpdmvhafkpspktnmqenwrivdffs hhpeslhmfsflfddvgiplnyrhmegfgvntytylinkdgkphlvkfhwkptcgvkclldeavtvggtchshatkdltdsiaagnype wklyiqtidpdhedrfdfdpldvtktwpediiplqpvgrmvlnknidnffaeneqlafcpaiivpgihysddkllqtrifsyadtqrhrlgpn ylmlpvnapkcayhnnhhdgsmnfmhrdeevnyfpsrfdaarhaekvpipprvltgcrekcvidkennfqqageryrsfdparqdrfl qrwvdalsdprithelrgiwisywsqcdaslgqklasrlnlkpnm*
Catalase related subgroup	Tomateiro	57	7%	3	100%	Excelente	SOL Genomics	mivtqyrpssaydtpflttnaggpvynnvssltvgprgpvlledyylieklatfdrekipervvhargasakgffevthdishltcadflr**AP GAQTPVICRFSTVVHER**gspesirdirgfavkfytregnfdlvgnnvpvffnrdak**SFPDTIR**alkpnpkshiqenwri ldffsflpeslhtfaffyddvclptdyrhmegfgvhayqlinkegkahyvkfhwkptcgvkcmseeeairvggtnhshatkdlydsiaag nypewklfiqtmdpedvdkfdfdpldvtktwpedllplipvgrlvlnrnidnffaeneqlafnpghivpgiyysedkllqtrifayadtqrh rigpnymqlpvnapkcghhnnhrdgamnmthrdeevdylpsrfdpcrpaeqypipscvlngrr**TNCVIPK**ennfkqageryrs wepdrqdryinkwveslsdprvtheirsiwisylsqadkscgqkvasrltvkptm*
Major latex-like protein (putative PR-10)	N. benthamiana	17	5%	1	81%	Bom	SOL Genomics	mgvkgk**LIASVEVK**cgghlvhdifhtkthhisnicpskiqnfeihegdsvkvgsvvswkynddgkdkiskqmieaidhetktit wkliggdllelynsftiitscdhewttwtflyekktedipeplvllgfvlhvtkdieghllk*
Kinase interacting family protein (helicase-like protein)	N. benthamiana	107	1%	1	80%	Bom	SOL Genomics	mlqraatnayswwaashirtkqskwleqslqdmeekvedviklieedgdsfakraemyykkrpelinfveesyrayralaerydklsrel qsanntiatlfpeqiqlamdeeddygtprmpknfpqvpttganipnvppkapvkdmkgllktgtlqflgkktankardanketpksgltk dealeeidklqrdilslqtvkefvkssyqsgiaktmeiehqivekhqkicsledefgearvieddeartlmaeaalkscqetlaqlqekqeqst neakeefkkieearkklksfrqkylgdpadetepdekddesagvtdvssqctsleeigekmeslhgkmneqfdassmsslтvtqlaekid elvnkvvsleaavstqtvlierlrteacglhthirtleddkaatltgdthnlnvr**VTMLEEKLK**giqdlnkdvetqnsslkthfaeahns lgqlsfkltsvkpdeevdetdssqdeeedladirlqnepekkenhvsaseakkeqdpvsaseikkeqdpvsvvsdkevqedtkslkkhve pleptvaekgeervssksessvhheqkpleddkdddltwqqmlltgledkekilleyttilrnykelkrkliemekkerdtefevtlqirel kstiskkdeeieslrlklslvqgnasesseskeekhqdlnpsddrslkpedtpkneekdeqdakiilidqcsslspveeklrmgidalldenl nfwlrfssafhqiqkfkttaqdlqgeittlketkegssktdmkseirpiykhlreiqneltvwleqslslkdelkrrfsslcsiqeeitkalkdg veedeirfsshqaakfqgevlnmkqennkvreeleagvehvttlqvdvektlrkldhqfdvggnqpqltnsasrsriplrafifgtkvkksk rsffhhnrkyqvlkggvpl*
Galactosylgalactosyl xylosylprotein 3-beta-glucuronosyltransferase 1	N. benthamiana	36	7%	1	80%	Baixa	SOL Genomics	mapfggfddaknsdfsfeikppvvnvkeemkdvviprpdnvvvnsvklpglgeeevhgkfdyvsrkllivvtptynralqayyllrlse vlklvkspllwvvvemnvasaetadilrktgvmyrhlvcsknmtdikdrgvhqrnvalehiehhrlngivyfadddniyslelfesirsinr fgtwpvamlaqsksk**AILEGPVCNGSQVIGWHTNEKSK**qlrrfhvdmsgfafnstilwdpkkwhrptsdpirqldn vkegfqettfieqivedesqmeavppgcsrvlnwhlhleahgavypggwllqknldaiistt*
Solute carrier family 2 facilitated glucose transporter member 8	N. benthamiana	57	6%	1	82%	Baixa	SOL Genomics	meeqllssegnkfpistscsasvvlstcvvmcgslaygfavgysspvqsgimadlglsiaeystfgailtlggtigalvsgtiadmagrkitm wimdlcfilgwlsiifaksvwwldigrflmgigagllllyvapiyiaeiapksir**GGCTAVFAFMLYFGFSLMFLIGNI LTWR**tlavvgvipslvqlpgiffipespprwlakigldkeveaslqylrgknvdisletaeikdnvesftklsgsryldmfdrryahslivgl gimillqsggtdaissfassifekagcsasfatttmgfiqlpfaamgifllldsagrrpvlmvasagtclgnflvgvgflckdydqmsqltatfvl vgilvfaiffsmgvggaattivseifpmnikgsagslaivcnwftswivtyafnfllfewspsgvffmftffsglmilfvakivpetkgrtleei qasltlln*
Katanin p60 ATPase-containing subunit A	N. benthamiana	50	4%	1	82%	Baixa	SOL Genomics	m**ADEPALTRWTFPFGR**rkvpardatdngdsvqngnsstmngngsvrntaelsifeqynqangvqsagvdgkpeksllpp fesaemrtlaesltrdilrgnpdvkwesikgletakrllkeavvmpikypkyfkglltpwkgillfgppgtgktmlakavatecnttffnisas svvskwrgdseklikvlfelarhhapstifldeidaiisqrgearsehessrrlktelliqmdglnrtdelvfvlaatnlpweldaamlrrlekril vplpeaearcgmfedllpslpeeaplpydllvektegfsgsdirllckeaamqplrrlmaeldkreelvpedelpnvgpitvtdievalkntr psahlhaprydkfnsdygshalq*

Tabela 1 (cont.)

Mutator-like transposase	Tomateiro	58	2%	1	81%	Bom	SOL Genomics	mcppikihndvgvkvyldqmrvnldfftkyplcitlkdcghynefhrvivndiinsdvrlsqnnidlysnnsirligmdldgvvndnsev dndiicdhsnlfvaenqiynnketlkevmrhvglvekfsfrvarcnaskervyarrqgitdvvavlimdnyidpskvytpkdvaddmlkl hgvsltyiqawrakekavklvrgdpaesyarlpghilplayaivdsendaswtwffeqfreaygvrqnmcfmsdrnesiwkgttnvype sehyaciwhlsvnvlknfnrnted lkmlffslakaytkqqfetimgkidqidtrirpylfdigyskwsrsysnckrtwtmtsniaeslnnvn ilarrlpvisllefmrvtiqrwihkhneeadkttsnltkkydvylqk**IITLSCNMR**vipstvdlhavaegakkyivnlntrmcscgrfq hyeipcghaiavlryrkfheadfcsafyslknfkdvvttknraatdthtyppm*
LINE-type retrotransposon Lib DNA	N. benthamiana	58	2%	1	82%	Baixa	SOL Genomics	meiekiasqqssqlaeksiqtpghlagsrqdsdigdispvhdsniqqigagidgqnvdhrhegetgiaqqeqrkcdvqqqqshrpvqnp qsvsnlnkekeitqqfqrptrvqdsreavhlaknshpamhsyelairsdngqdldneldynmvwtkqrmtiagqvmriqawtptfkpe eetplvpiwislpelpwhcynkefittllspigrvlyldsasinktrgsqarvkvqvdltkerpphiwmgyigeditdgrwqkieydsipdy cyyyk**HQGHKEMNCMVK**qrdednkkrkemekkprkdtaqivnaelqinkekdngrsdmanskkgvvinkldlamigll finnnykqeynttgegmethtqqrqeqsqtdaegiqvhrdhrnkqkrqevqdsrtenavyqqetrknsgidsmlpspaslyidtvgvae ggevgcgqmdtrfhhdrndkgknkiveqgscsnvdkeppdkiqlpnqqinfsyknknpsntgkapidhqrdnld*
Cytosolic 6-phosphogluconate dehydrogenase	N. benthamiana	53	2%	1	55%	Bom	NCBI	mavtriglaglavmgqnlalniaekgfpisvynrttskvdetvqrakvegnlpvygfhdpasfvnsiqkprvvimlvkagapvdqtiatla ahleqgdciidggnewyenterrekameergllylgmgvsggeegarngpslmpggsfeaykyiedillkvaaqvpdsgpcvtyigkg gsgnfvkmvhngieygdmqliseaydvlksvgkltnselqqvfsewnkgellsflieitadifsikddqgsghlvdkvldktgmkgtgk wtvqqaaelsvaaptieasldsrflsglkderveaakvfqgdfssnlpvdk**AQLIEDVR**qalyaskicsyaqgmniikaksmekg wslnlgaelariwkggciiraifldrikkaydrnsdlanllvdpefaqeimdrqaawrrvvclainngvstpgmsaslayfdsyrrdrlpanl vqaqrdyfgahtyervdmpgsfhtewfkiaraakm*
Histidinol dehydrogenase	A. thaliana	45	2%	1	54%	Excelente	NCBI	mpirlsrasadfdtqfrqflaakrevsadveaaaraivddvarrgdaalleatakfdr**LELSAATLR**vtaseieaavkacdtatvdaltf ardrieafhlkqlpqdqrftdvagvelgwrwsaiesaglyvpggtaaypssvlmnaipakvagvsrlvmvvpapdgklnplvlaaahlg gvseiyrvggaqavaalahgtatiapvakivgpgnayvaaakrlvfgkvgidmiagpsevlviaddsgnadwiaadllaqaehdasaqsi litdsrrladdvaravesqlktlpraaiagaswadfgaiievdhldeavpladaiaaehleimtadpdafaakvrnagaiflgahtpeaigdyv ggsnhvlptarsarfssglgvldfmkrtsllkcgpeqlralgpaamtlgkaegldahsrsvgirlnlp*
Glicina Dehidrogenase P (AHRD V1)	Tomateiro infectado com ToYSV	113	7%	4	100%	Excelente	SOL Genomics	merarklanrailkrlvsqskqsrsneipspsslyrpsryvsslspytfqarnsvksfntqqvrsisvealkpsdtfprrhnsatpeeqtkmaef cgfqsldalidatvpqsirsesmklpkfdggltesqmiehmqklasknkvfksyigmgyyntyvppvilrnllenpawytqytpyqaeis qgrlesllnyqtmitdltglpmsnaslldegtaaaeamamcnnilkgkkktfliannchpqtieicktradgfdlkvvtvdlkdidyksgdv cgvlvqypgtegeildygefiknahahgvkvvmasdllaltmlkppgelgadivvgsaqr**FGVPMGYGGPHAAFLATS QEYKR**mmpgr**IIGLSVDSTGKPALR**mamqtreqhirrdkatsnictaqallanmaamyavyhgpeglktigqrvhgl agtfsaglkk**LGTVEVQDLPFFDTVK**vkcsdak**AIADVATKNDINVR**ivdnntitvsfdetttledvddlfkvfal gkpvpftaqsiaqevenlipsgltretpflthqifnsyhtehellrylhklqskdlslchsmiplgsctmklnattemmpvtwpsfanihpfa pteqaagyqemfddlgallctitgfdsfslqpnagaageyaglmvirayhmsrgdhhrnvciipvsahgtnpasaamcgmkivavgtda kgninieelrkaaeahkdnlsalmvtypsthgvyeegideickiihdnggqvymdganmnaqvgltspgfigadvchlnlhktfciphg gggpgmgpigvkkhlapylpshpvvstggipspdqskplgaisaapwgsalilpisytyiammgskgltdaskiailsanymakrlekh ypvlfrgvngtcahefiidlrgfkntagiepedvakrlidygfhgptmswpvpgtlmieppteseskaeldrfcdalisireeiaqiekgnvdi nnnvlkgaphppsmlmadawtkpysreyaaypapwlrsakfwpttgrvdnvygdrnlictllpvsemaeekaata*
Transketolase	N. benthamiana infectada com ToYSV Tomateiro infectado com ToYSV	80,1	14 a 17%	8	100%	Excelente	NCBI SOL Genomics	massslplsqailsrsiprhgsssstnsqfspsslslptfsglksttsttfrrrtlpspvavrspeirasaavetlektdnalveksvntir**FLAID AVEK**ansghpglpmgcapighilydeimrynpknpywfnrdrfvlsaghgcmlqyallhlagydavreedlksfrqwgsktpghp enfetpgveyttgplgqgianavglalsekhlaarfnkpgaeivdhytycilgdgcqmegisqeaaslaghwglgkliafyddnhisidgdt eiaftedvgarfealgwhviwvkngntgydeiraaikeakavtdkptmik**VTTTIGFGSPNKANSYSVHGSALGA K**veatrqnlgwpyepfhvpedvkshwsr**HVPEGAALEAGWNTK**faeyekkypeeaadlksiitgelpagwek**ALP TYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLMK**mfgdfqkntpeernlrfgv rehgmgaicngialhspglipycatffvftdymraairisalsesrviyvmthdsiglgedgpthqpiehlasframprnvlmlrpadgneta gaykvailksk**TPSILALSR**qklpqlagtsiegtakggyvtsdnssgnnpdviligtgseleiavkaadelrkegkavrvvslvswelf deqsaeykesvlpaavtarvsieagstfgwekivgskgkaigidrfgasapagkiykeygitaeaviaaakqvs*
Anidrase carbônica	Tomateiro infectado com ToYSV	34,4	18%	4	100%	Excelente	NCBI SOL Genomics	mstasinncltlspaqaslkkptrpvafarlgnsssssssipslirnepvfaaptpiinpivreemakesydqaiaalekllsekaelgpvaaarv dqitaelksadggk**AFDPVEHMK**agfihfktekydtnpalygelakgqspkfmvfacsdsr**VCPSHVLNFQPGEAF MVRNIANMVPAYDKVR**ysgvgaaieyavlhlkvenivvighsacggikglmslpedgsestafiedwvkiclpakak**VL ADHGGKEFAHQCTACEK**eavnvslgnlltypfvreglvkktlalkggyydfvkggfelwglefglspplsv*

Tabela 1 (cont.)

Serina hidroximetil-transferase	N. benthamiana infectada com ToYSV Tomateiro Infectado com ToYSV	57,3	2%	1	98%	Excelente	NCBI SOL Genomics	mamamalrklsssvnkssrplfsasslyyksslpdeavydkenprvtwpkqlnsplevidpeiadiielekarqwkglelipsenftslsv mqavgsvmtnkysegypgaryyggneyidmaetlcqkraleafrldpakwgvnvqplsgspsnfqvytallkphdrimaldlphggh lshgyqtdtkkisavsiffetmpyrldestgyidydqleksatlfrpk**LIVAGASAYAR**lydyarirkvcdkqkavlladmahisgl vaagvipspfdyadvvttttthkslrgprgamiffrkglkevnkqgkevfydyedkinqavfpglqggphnhtitglavalkqattpeyray qeqvlsnsskfakalsekgydlvsggtenhlvlvnlknkgidgsrvekvlelvhiaanktvpgdvsamvpggirmgtpaltsrgfveed fvkvaeyfdaavslalkvkaeskgtklkdfvealqtssyvqseisklkhdveefakqfptigfekatmkynk*
Catalase 4 OS	N. benthamiana infectada com ToYSV Tomateiro infectado com ToYSV	56,7	3%	1	98%	Excelente	NCBI SOL Genomics	mdpykhrpssafnspfwttnsgapiwnnnssltvgargpilledyhlveklanfdreripervvhargasakgffevthdishltcadflrap gvqtpvivrfstvihergspetlrdprgfavkfytregnfdlvgnnlpvffvrdgmkfpdmvhalkpnpknhiqenwrildffshfpeslh mftflfddlgvpqdyrhmdgfgvntytlinkagkavyvkfhwkttsgikclleeeaikvgganhshatqdlhdsiaagnypewklfvqti dpehedkfdfdpldvtktwpediiplqpvgrlvlnknidnffaeneqlafcpaivvpgvyysddkmlqtrifsyadsqrhr**LGPNYL QLPANAPK**cahhnnhhegfmnfihrdeevnyfpsrydpvrhaerfpippaicsgrrekcgiekennfkqpgeryrswapdrqdrf arrwvdalsdprvtheirsvwisywsqadrslgqkiashlstrpni*
S-adenosil metionina sintase 1	N. benthamiana infectada com ToYSV Tomateiro infectado com ToYSV	43,2	4%	1	84%	Excelente	NCBI SOL Genomics	maaldtflftsesvneghpdklcdqvsdavldaclaedpdskvacetctktnmvmvfgeittkanvdyekivretcrnigfvsadvgldad hckvlvnieqqspdiaqgvhghftkrpeeigagdqghmfgyatdetpelmplshvlatklgarltevrkngtcawlrpdgktqvtveyrn esgarvpvrvhtvlistqhdetvtndeiaadlkehvikpvipeqyldektifhlnpsgr**FVIGGPHGDAGLTGR**kiiidtyggw gahgggafsgkdptkvdrsgayvarqaaksivasglarrcivqvsyaigvpeplsvfvdtygtgripdkeilkivkenfdfrpgmiiinldl kkggngrylktaayghfgrddpdftwevvkplkwekpsa*
Thiamine thiazole synthase 3	N. benthamiana infectada com ToYSV Tomateiro infectado com ToYSV	37,8	4%	1	84%	Excelente	NCBI SOL Genomics	msisaagvatglganvelksnvgsssssvagvrlftsrkaqlrrcaapatsaslysdanydlnnykfapikesivaremtrrymtdmithadt dvvvvgagsaglscayelsknpnvkvaiveqsvspgggawlggqlfsamivrkpahrfldeievpyeemenyvvikhaalftstimskl larpnvklfnavaaedliirgdrvsgvvtnwalvaqnhntqscmdpnvmeakvvvvsscghdgpfgatgvkrlrsigmiesvpgmkcld mnaaedaivkhtrevvpgmivtgmevaeidgspr**MGPTFGAMMISGQK**aahlalralglpnevdgnykpnvhpelvlast ddmtasa*
Probable helicase	Tomateiro infectado com ToYSV	91,4	2%	1	84%	Excelente	SOL Genomics	mavdknkleeealslrfykivlswdylrlikesdrkkgkgdddnalvlkkaknsykdvqdylatfepllfeevkaqiiqgkkddeeetlw mkavtvgcseidgfhfpmiscsdaesiqqndllllsnkefgdgkrlptayafalvedrrpdkirlrmhlsgevkqlntqeieacsrllsmrpl vtenakllhvlkicslstiareyvalrsvsslpfkdlilsaadsnrstedqawkisrplkeflesnhnksqldainaglsrktfvliqgppgtgktq tilgilsailhatpsrvhsnrvklssvkrgpelsmsdkykhwgkaspwlggtnpldqempidgddgffptsgndlkpevvnssrkyr**VR VLVCAPSNSALDEIVLR**ilntgirdendrayspkivriglkahhsvqavsmdylveqrlsgmdsqigdrqkqggpvkdkds irasildeavivfstlsfsaspvftklnrgfdvviideaaqavepstllplsngckqvflvgdpvqlpatvispiagkfgyctslferlqragypv qmlktqyrmhpeirnfpsrefyeealedgpdvevqtkrswheyrcfgpfcffdihdgkesqpsgsgswqnvdevefvlamyhklvsg ypelksssrlaiispyryqvkllrqkfretfgvesdkvvdintvdgfqgrekdvaifscvraskdkgigfvadyrrmnvgitrarssvlvvgs astlrkdarwqnlvesaekrnalhkvskpyaeffseenlkllkvevahdkheappedmdidvpiaaetdhapqedwgyageeggyded *
Ran-binding protein 9	Tomateiro infectado com ToYSV	54	1%	1	81%	Bom	SOL Genomics	mdssltlslshtnsdthdligtyfleqwrlkskaavrksesvmgvsenqylmeidseeceeekeelpteldtlhssggfsivgsdklsvrypn ahlhghdvgvvqanrpapckrlvyyfeifvenagtkgqiaigfttpgfklrrqpgwesnsygyhgddgllyrgqgkgeafgpaystgdtv ggginyasqelfftkngtvvgtvfkdvkgplfptvavhsqhevvtvnfgkrpfvfdlkayeaqerak**QNSTIEK**lsipqnasygivra ylqhygyedtlktfdvesrstlppisfvqengfaedvsvyslnqrkvlrqlirsgqigdafgklrelypqilqdgtsaicfllhcqnfielvrvgk leeavlygrtefekfyklgdyddlvkdcaallayeqpqkssvgyllgdsqrdivadavnamvlstnpsvkdsreclhsrldrllrqlsacflak rtfngdqgegfhlhrilnsgrkg*

Tabela 2. Avaliação da interação entre a MP com ou sem domínio transmembrana (MPΔtmd) com proteínas de plantas hospedeiras do ToYSV por meio do sistema duplo-híbrido de levedura, e confirmação da interação pela quantificação da atividade da enzima β-galactosidase.

Proteínas/genes de plantas hospedeiras	Gene de referência TAIR[1] - Sol Genomics	pDEST22 +MP		pDEST22 +MPΔtmd		pDEST32 +MP		pDEST32 +MPΔtmd		Atividade β-gal	Atividade β-gal	Resultado final
		cl. 1	cl. 2	cl. 1	cl. 2	cl. 1	cl. 2	cl. 1	cl. 2	pDEST22 +MP/ MPΔtmd	pDEST22 vazio	
Probable helicase	Solyc01g104640.2.1	-	-	+	+	-	-	-	-	Negativo	-	Negativo
Ran-bind protein 9	Solyc03g033380.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
CBL (cystathionine beta-lyase)	AT3G57050	-	-	-	+	-	-	-	-	Positivo	Positivo	Negativo
RXT3-like protein C1259.07 Histone deacetylation protein Rxt3	Solyc05g056340.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
AGT (serine-glyoxylate aminotransferase)	AT2G13360	-	-	-	-	-	-	-	-	-	-	Negativo
Lipid-associated family protein (PLAT-plant-stress domain-containing protein)*	AT2G22170	-	-	-	-	-	-	-	-	-	-	Negativo
Gamma-aminobutyrate transaminase-like protein	AT3G22200	-	-	-	-	-	-	-	-	-	-	Negativo
Pistil extensin like protein (Fragment) Pollen Ole e 1 allergen and extension*	AT2G34700	-	-	-	-	-	-	-	-	-	-	Negativo
Cobalamin synthesis protein P (CobW-like)	AT1G15730	-	+	-	-	-	-	-	-	Positivo	Positivo	Negativo
Universal stress protein (UspA)	AT1G11360	-	-	-	-	-	-	-	-	-	-	Negativo
Plant-specific domain TIGR01615 family protein	Solyc09g065760.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
Lipoxygenase homology domain-containing protein 1*	AT4G39730	-	-	-	-	-	-	-	-	-	-	Negativo
Transketolase	AT3G60750	-	-	-	-	-	-	-	-	-	-	Negativo
THI1; protein homodimerization	AT5G54770	-	-	+	-	-	-	-	-	Negativo	-	Negativo
Afadin- and alpha-actinin-binding protein*	Solyc07g019530.2.1	-	-	-	+	-	-	-	-	Positivo	Positivo	Negativo
Urease accessory protein UREG	AT2G34470	-	-	-	-	-	-	-	-	-	-	Negativo
Os07g0175100 protein (Fragment) Homeodomain-like containing protein	Solyc07g040960.1.1	-	-	-	-	-	-	-	-	-	-	Negativo
Peptide chain release factor 2	AT3G57190	-	-	-	-	-	-	-	-	-	-	Negativo
Major latex-like protein (putative PR-10)	AT1G70890	-	-	-	-	-	-	-	-	-	-	Negativo
RNA-binding glycine-rich protein-1b	Solyc10g051390.1.1	+	+	-	-	-	-	-	-	Positivo	Positivo	Negativo
Nodulin-related	AT2G03440	-	-	-	-	-	-	-	-	-	-	Negativo
Anidrase carbônica	Solyc05g005490.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
Glicina Dehidrogenase P (AHRD V1)	Solyc08g065220.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
ASR4 (Fragment) ABA/WDS induced protein	Solyc04g071620.2.1	-	-	-	-	-	-	-	-	-	-	Negativo
Kinase interacting family protein (helicase-like protein)	AT5G10500	-	-	-	-	-	-	-	-	-	-	Negativo
Controle negativo pDEST32 vazio + pDEST22-MP	-	-	-	-	-	-	-	-	-	-	-	Negativo
Controle negativo pDEST22 vazio + pDEST32-MP	-	-	-	-	-	-	-	-	-	-	-	Negativo
Controle negativo pDEST32 vazio + pDEST22-MPΔ	-	-	-	-	-	-	-	-	-	-	-	Negativo
Controle negativo pDEST22 vazio + pDEST32-MPΔ	-	-	-	-	-	-	-	-	-	-	-	Negativo
Controle positivo pDEST22+pDEST32 NIK:NSP	-	+	+	+	+	+	+	+	+	+	+	Positivo

[1]TAIR: The Arabidopsis Information Resource.

* Proteínas selecionadas para teste de interação proteína-proteína in vivo através de BiFC.

Análise in vitro da interação entre MP-ToYSV e proteínas de plantas hospedeiras. Ensaios de duplo-híbrido empregando a MP-ToYSV íntegra e MP sem domínio transmembrana (MP-Δtmd) foram conduzidos para avaliar a interação com as 25 proteínas de plantas hospedeiras do ToYSV selecionadas por pull-down (Tabela 2). Os resultados indicaram que pelo menos seis proteínas eram fortes candidatas a interagirem com a MP-ToYSV (Figura 4). Leveduras transformadas com os 19 clones restantes não apresentaram crescimento em meio seletivo, indicando ausência de interação.

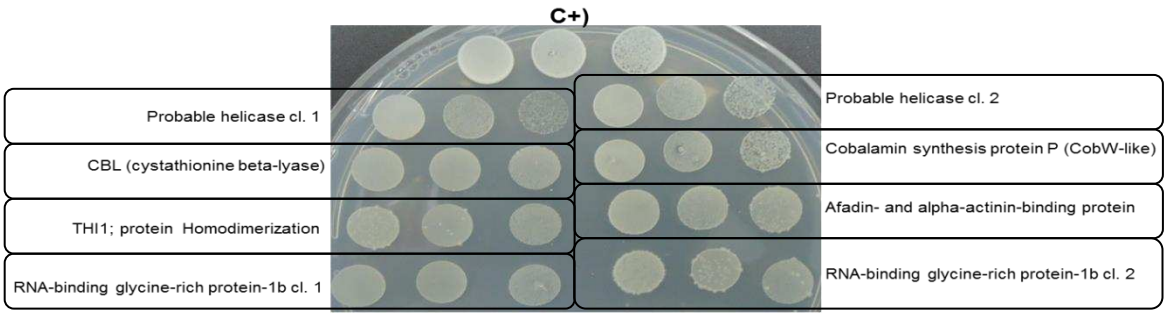


Figura 4. Avaliação in vitro da interação da MP-ToYSV com proteínas de plantas hospedeiras identificadas em ensaio de pull-down. Proteínas candidatas a interagirem com a MP-ToYSV pelo sistema duplo-híbrido de levedura. Como controle positivo (C+) utilizaram-se as proteínas NIK e NSP.

A quantificação da atividade da enzima β-galactosidase foi realizada para confirmar a interação da MP-ToYSV com as seis proteínas selecionadas. As proteínas "probable helicase" e THI1 apresentaram resultado negativo no teste, descartando-se a possibilidade de interagirem com a MP. As outras quatro proteínas apresentaram resultado positivo (embora fraco em comparação ao controle positivo) e continuaram a ser consideradas candidatas a

interagirem com a MP. Para certificar-se de que as interações eram específicas, estas quatro proteínas candidatas remanescentes foram também avaliadas em combinação com os respectivos vetores pDEST22 ou pDEST32 vazios. Verificou-se que estas combinações com vetores vazios eram capazes de ativar o gene marcador responsável por codificar a enzima β-galactosidase, fornecendo resultados falso-positivos de interação proteína-proteína pelo sistema duplo-híbrido e nos testes de quantificação da atividade da enzima β-galactosidase. A quantificação numérica via medição em espectrofotômetro foi impossibilitada devido a sua baixa atividade e, consequentemente, baixa densidade óptica.

Das 25 proteínas testadas pelo sistema duplo-híbrido, quatro serão também avaliadas para interação proteína-proteína in vivo pelo método de complementação de fluorescência bimolecular (BiFC) (Tabela 2). Os quatro genes correspondentes estão sendo subclonados em vetores SPYCE e SPYNE (C- e N-terminal YFP) para os testes de BiFC utilizando a MP-ToYSV clonada nestes mesmos vetores. Posteriormente as culturas serão co-agroinfiltradas em folhas de N. benthamiana e a interação será avaliada conforme descrito por Mendes et al. (2013). Ensaios de co-imunoprecipitação também serão realizados, conforme descrito por Carvalho et al. (2008a).

DISCUSSÃO

A MP dos geminivírus é responsável pelo movimento célula-a-célula, pode atuar no movimento sistêmico e em pelo menos alguns casos (ToYSV e Bean dwarf mosaic virus, BDMV) também pode influenciar na sintomatologia e agressividade viral (Hou et al., 2000; Rojas et al., 2005, Zhou et al., 2011; Antunes, 2011). O sucesso da infecção sistêmica do hospedeiro é dependente de interações entre proteínas virais e inúmeros fatores do hospedeiro (Florentino et al., 2006; Carvalho et al., 2008b; revisado por Lozano-Durán et al., 2011). A identificação destes fatores que interagem com a MP é o primeiro passo para entender o mecanismo de transporte intercelular do genoma viral coordenado pela MP e, em uma última

análise, auxiliar na busca de fontes de resistência do hospedeiro contra o vírus. No contexo do movimento viral, já são conhecidos alguns fatores do hospedeiro que interagem com a MP dos begomovírus (Krenz et al., 2010; Lewis & Lazarowitz, 2010; Balmant, 2011; Zhou et al., 2011). Balmant (2011), por meio de uma biblioteca de cDNA de A. thaliana infectada com o Cabbage leaf curl virus (CaLCuV) e utilizando a MP deste mesmo vírus como isca, identificou a interação com a proteína ribossomal S5A (RPS5A), confirmou sua interação in vitro e demonstrou que a infecção pelo CaLCuV regula negativamente a sua expressão, sugerindo que esta interação seja favorável para o ciclo infeccioso viral. Krenz et al. (2010), também utilizando uma biblioteca de cDNA de A. thaliana, identificaram e confirmaram a interação in vitro e in vivo da proteína MP do AbMV com a proteína cpHSC70. A interação facilita o movimento viral e está ligada à indução de sintomas. Lewis & Lazarowitz (2010) demonstraram que a MP do CaLCuV também interage in vitro e in vivo com a proteína sinaptotagmina A (SYTA) de A. thaliana, e que essa interação favorece o movimento viral. Zhou et al. (2011), utilizando a proteína MP do BDMV como isca, identificaram interação com a histona H3 em tomateiro e N. benthamiana.

A expressão constitutiva ou transiente da etiqueta NTAPi-MP em plantas hospedeiras do ToYSV seguida da purificação dos heterocomplexos protéicos pela metodologia TAP original ou com modificações não permitiu a identificação de proteínas que interagem com a MP do ToYSV, nem mesmo aquelas proteínas de plantas que já tiveram sua interação confirmada com a MP de begomovírus. Em plantas, o método TAP tem sido utilizado com sucesso em A. thaliana (Rubio et al., 2005; Van Leene et al., 2007; Dufresne et al., 2008; Thivierge et al., 2008), N. benthamiana (Rohila et al., 2004) e O. sativa (Rohila et al., 2006). Também tem sido utilizado com eficiência no estudo da interação entre proteínas de potyvírus com proteínas de A. thaliana (Dufresne et al., 2008; Thivierge et al., 2008). O método TAP é considerado mais eficiente quando as proteínas são abundantes e citoplasmáticas (Jensen et al., 2008) e tem tido maior destaque em estudos com células de mamíferos e leveduras (Xu et

al., 2010). O tamanho da etiqueta TAP (aprox. 20 kDa) em comparação às etiquetas FLAG ou 6His (aprox. 1 kDa), tem sido apontado como um possível problema devido à possibilidade de alterações estruturais ou funcionais que causem instabilidade na proteína etiquetada (Gloeckner et al., 2007; Abu-Farha et al., 2008). Outro ponto que tem sido abordado como um possível fator de insucesso é o fato do processo de purificação ser longo e envolver inúmeras etapas, de forma que apenas complexos protéicos muito estáveis e abundantes permaneceriam complexados (Rohila et al., 2009).

Burckstummer et al. (2006) e Van Leene et al. (2007 e 2008), baseando-se no método TAP original, desenvolveram uma nova etiqueta denominada protein G/streptavidin binding peptide (GS)-TAP, que uma vez fusionada à proteína de interesse, é expressa em maior quantidade em células vegetais e propicia a recuperação de uma alta quantidade do heterocomplexos em comparação à etiqueta TAP. Gloeckner et al. (2007), na tentativa de reduzir o tamanho da etiqueta TAP original, desenvolveram a etiqueta streptavidin binding peptide-tag II and FLAG tag (SF)-TAP reduzindo a etiqueta para 4.6 kDa e demonstraram maior estabilidade da proteína alvo etiquetada.

A tentativa de identificação dos complexos protéicos utilizando duas abordagens de espectrometria de massa, MALDI-TOF/TOF e ION TRAP, reforça a hipótese de que a ineficiência observada neste trabalho tenha sido realmente no método de purificação TAP. A espectrometria de massa utilizando MALDI-TOF/TOF tem como principal vantagem a maior tolerância a contaminantes presentes nas amostras peptídicas e, em contrapartida, menor sensibilidade em termos de concentração peptídica. Por outro lado, a espectrometria de massa utilizando ION-TRAP-LC apresenta maior sensibilidade em termos de concentração peptídica e, em contrapartida, baixa tolerância a contaminantes (Abu-Farha et al., 2008; Collins & Choudhary, 2008).

Foi demonstrado que uma protease de *A. thaliana* tem afinidade pelas esferas de IgG-agarose utilizadas na primeira etapa de purificação TAP, e consequentemente interfere na

formação de heterocomplexos e/ou na ligação de proteínas alvo sensíveis a proteases (Rohila et al., 2004). Desta forma, o uso do inibidor de proteases do tipo cisteína E-64 nesta etapa foi fundamental. Neste mesmo sentido, a etapa inicial de cross-linking também foi fundamental para a manutenção e recuperação dos heterocomplexos, bem como evitar a formação de complexos não específicos. Mesmo assim, a composição dos complexos proteícos quando se utiliza estes sistemas pode não refletir diretamente a composição dos complexos em células em condições naturais.

Ao contrário do método TAP, o método pull-down utilizando a proteína 6His-MP-ToYSV em fase estacionária permitiu detectar diversas proteínas de hospedeiros candidatas a interagirem com a MP do ToYSV, sugerindo ser uma metodologia relativamente eficiente para a identificação de interações proteína-proteína. A principal desvantagem deste método é a possibilidade de ligações inespecíficas proteína-proteína ou presença de proteínas contaminantes (interações falso-positivas). Ajustes ao protocolo e uso em paralelo de outras proteínas virais como controle (ex. 6His-CP-ToYSV ou 6His-GFP) podem auxiliar na seleção das proteínas candidatas a interagir com a proteína alvo para posterior análise pelos sistemas duplo-híbrido ou BiFC. Bruckner (2012), utilizando este mesmo tipo de ensaio, na tentativa de encontrar proteínas parceiras da TCTP (translationally controlled tumour protein) de tomateiro a partir de proteínas totais de N. benthamiana, identificou algumas proteínas candidatas, incluindo proteínas relacionadas com a síntese de ABA, sorbitol, urease G, e cobalina, catalases e uma proteína semelhante a extensinas específicas de pistilo, ambas também detectadas por este sistema em nossas buscas, sugerindo assim que estas podem apresentar ligação com a resina Ni-NTA e/ou interação inespecífica com a proteína utilizada como isca, uma vez que a interação destas foi negativa pelo sistema duplo-híbrido, tanto com TCTP quanto com a MP do ToYSV.

É razoável supor que os mecanismos da interação entre fatores virais e do hospedeiro não envolvam somente interações diretas, mas também (e talvez predominantemente)

interações indiretas envolvendo mais de uma proteína viral ou do hospedeiro, que juntamente com outras proteínas possibilitam a formação de heterocomplexos estáveis. Contudo, esta não pode ser uma justificativa que permita explicar o insucesso na busca de proteínas parceiras, uma vez que ambos os métodos utilizados (TAP e pull-down), em teoria permitem a recuperação destes heterocomplexos. Porém, na prática a grande limitação tem sido a recuperação dos heterocomplexos em quantidade suficiente para serem detectados e identificados por MS (Silva, 2011; Cascardo, 2011; Bruckner, 2012). Para contornar este problema seguimos diretamente para a digestão tríptica em solução, ou seja, sem submeter às proteínas recuperadas a eletroforese. Entretanto, os heterocomplexos recuperadados dessa forma não apresentaram a pureza necessária para análise em MS.

A identificação ou comprovação de interações proteína-proteína pelo sistema duplo-híbrido tem sido amplamente utilizada por constituir um método simples, rápido e de baixo custo (Navaratnam, 2009). Diversas interações proteína-proteína envolvendo geminívirus e seus hospedeiros já foram identificadas por este sistema, a exemplo de NSP com uma acetilase nuclear de A. thaliana (AtNSI) (McGarry et al., 2003), uma GTPase citosólica (NSP-interacting GTPase, NIG) (Carvalho et al., 2008a), uma proteína receptora do tipo PERK (proline-rich extensin-like receptor kinase) localizada na membrana plasmática (Florentino et al., 2006) e membros da família LRR-RLK, também localizados na membrana plasmática e designados NSP-interacting kinase (NIK) 1, 2 e 3 (Mariano et al., 2004; Fontes et al., 2004). As principais limitações deste sistema são o grande número de falso-positivos e o fato de proteínas com domínio transmembrana poderem não atingir o núcleo, e com isso não ocorrer à ativação transcricional dos genes repórteres, obtendo-se um resultado falso-negativo. Resultados falsos negativos também podem ocorrer devido a diferenças nas modificações pós-traducionais de proteínas de eucariotos superiores quando estas são analisadas em leveduras (Jensen et al., 2008). Aliado aos testes de duplo híbrido, outros métodos de identificação de interações poderiam ter sido utilizados neste caso, como por exemplo, a co-

imunoprecipitação (Zhou et al., 2011), far-Western blotting (Wu et al., 2007; Dombrovsky et al., 2007; Machida & Mayer, 2009; Fernández-Calvino et al., 2010) ou ensaios de BiFC (Sánchez-Durán et al., 2011).

A proteína MP dos begomovírus apresenta um domínio central associado à ligação com a membrana plasmática (Aberle et al., 2002; Zhang et al., 2002; Frischmuth et al., 2004). Mesmo se a MP for fusionada a um sinal nuclear (presente nos vetores pDEST22 ou 32), esta vai se localizar, ao menos parcialmente, na membrana plasmática (Frischmuth et al., 2004). Estas características dificultam ou podem mesmo impossibilitar a utilização do sistema duplo-híbrido da forma tradicional. A clonagem da MP sem o domínio transmembrana foi importante para contornar este problema, contudo, a retirada deste domínio pode ter afetado ou até abolido a capacidade da MP de interagir com proteínas do hospedeiro, novamente fornecendo resultados falso-negativos. Contudo, Krenz et al. (2010) utilizaram apenas a região N-terminal (posição 1-116 de amonoácidos) da MP de AbMV para detectar a interação pelo sistema duplo-híbrido desta proteína com a proteína HSC70 de *A. thaliana*. De fato, é esperado que os domínios que possibilitam a interação entre a MP e fatores do hospedeiro não tenham sobreposição com o domínio transmembrana.

Outros sistemas de busca ou identificação de interações entre proteínas poderão fornecer informações adicionais sobre fatores do hospedeiro e a MP dos begomovírus. A exemplo, a construção de bibliotecas de cDNA a partir de mRNA de plantas hospedeiras (sadias e infectadas) do vírus e o uso destes genes/proteínas como presa e a proteína viral de interesse como isca no sistema duplo-híbrido poderia viabilizar a identificação de proteínas candidatas a interagirem com a proteína viral (McGarry et al., 2003; Castillo et al., 2004; Lozano-Durán et al., 2011; Balmant, 2011). Outros métodos de identificação de complexos proteícos incluem ProCoDeS (Proteomic Complex Detection using Sedimentation) (Hartman et al., 2007; Segura et al., 2010) e far-Western blotting (Wu et al., 2007; Dombrovsky et al., 2007; Fernández-Calvino et al., 2010).

Alguns softwares têm sido recentemente disponibilizados para a busca de interactomas ou predição in silico de interações proteína-proteína de plantas, a exemplo do STRING (string-db.org; Franceschini et al., 2013), DIP (dip.doe-mbi.ucla.edu/dip/Search.cgi?SM=2; Salwinski et al., 2004), IntAct (www.ebi.ac.uk/intact/; Hermjakob et al., 2004), MINT (mint.bio.uniroma2.it/mint/Welcome.do; Chatr-aryamontri et al., 2007), BIND Database (www.bindingdb.org/; Bader et al., 2003), IMEx (www.imexconsortium.org/; Orchard et al., 2007) e BioGRID (thebiogrid.org; Stark et al., 2006). Outra abordagem que tem uma tendência promissora são algoritimos para predição de interação direta proteína-proteína que levam em consideração características como as estruturas terciárias das proteínas estudadas, denominada análise em docking (Ritchie, 2008). Contudo, há pouca ou nenhuma informação disponível sobre proteínas virais ou seu interactoma. Neste sentido, a confiabilidade da meta-análise é considerada limitada e pode ser controversa, principalmente devido a falhas na curagem ou atualização dos dados (Zhang et al., 2010).

A avaliação da interação in vivo por BiFC-YFP e/ou co-imunoprecipitação das quatro proteínas selecionadas (Tabela 2) poderá trazer mais informações, assim como o uso de outras metodologias de busca e indentificação de interações da MP com proteínas de hospedeiros do ToYSV.

AGRADECIMENTOS

Ao Instituto Nacional de Ciência e Tecnologia em Interações Planta-Praga (INCTIPP), CAPES e CNPq (bolsa de doutorado para MFB), CNPq (concessão 483607/2013-4 para FMZ) e FAPEMIG (concessão APQ-00949-09 para FMZ). Ao Núcleo de Análise de biomoléculas da Universidade Federal de Viçosa por disponibilizar os equipamentos para a realização das análises de MS. Os autores também agradecem o apoio financeiro das agências: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos

(FINEP) e Sistema Nacional de Laboratórios em Nanotecnologias (SisNANO)/Ministério da Ciência, Tecnologia e Informação (MCTI).

REFERÊNCIAS

Aberle HJ, Rutz ML, Karayavuz M, Frischmuth S, Wege C, Hulser D & Jeske H (2002) Localizing the movement proteins of Abutilon mosaic geminivirus in yeast by subcellular fractionation and freeze-fracture immuno-labelling. Archives of Virology 147:103-117.

Abu-Farha M, Elisma F & Figeys D (2008) Identification of protein-protein interactions by mass spectrometry coupled techniques. Advances in Biochemical Engineering/Biotechnology 110:67-80.

Almeida AMR & Lima JAA (2001) Princípios e técnicas de diagnose aplicados em fitovirologia. Londrina: Embrapa Soja, 186p.

Alves-Junior M, Alfenas-Zerbini P, Andrade EC, Esposito DA, Silva FN, Cruz ACF, Ventrella MC, Otoni WC & Zerbini FM (2009) Synergism and negative interference during co-infection of tomato and Nicotiana benthamiana with two bipartite begomoviruses. Virology 387:257-266.

Amberg DC, Burke DJ & Strathernt JN (2005) Assay of β-galactosidase in yeast: Assay of crude extracts. Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, EUA.

Ambrozevicius LP, Calegario RF, Fontes EPB, Carvalho MG & Zerbini FM (2002) Genetic diversity of begomoviruses infecting tomato and associated weeds in Southeastern Brazil. Fitopatologia Brasileira 27:372-377.

Andrade EC, Ambrozevicius LP, Calegario RF, Fontes EPB & Zerbini FM (2002) Molecular cloning and characterization of Tomato chlorotic mottle virus (TCMV), a new tomato-infecting begomovirus. Virus Reviews and Research 7:153-155.

Andrade EC, Manhani GG, Alfenas PF, Calegario RF, Fontes EPB & Zerbini FM (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. Journal of General Virology 87:3687-3696.

Antunes TFS (2011) Determinantes virais envolvidos na adaptabilidade diferencial de dois begomovírus em tomateiro e Nicotiana benthamiana. 73 pág. Dissertação MS (Bioquímica Agrícola). Universidade Federal de Viçosa.

Arguello-Astorga G, Herrera-Estrella L & Rivera-Bustamante R (1994) Experimental and theoretical definition of geminivirus origin of replication. Plant Molecular Biology 26:553-556.

Arguello-Astorga G, Lopez-Ochoa L, Kong LJ, Orozco BM, Settlage SB & Hanley-Bowdoin L (2004) A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. Journal of Virology 78:4817-4826.

Ascencio-Ibanez JT, Sozzani R, Lee TJ, Chu TM, Wolfinger RD, Cella R & Hanley-Bowdoin L (2008) Global analysis of arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiology 148:436-454.

Bader GD, Betel D & Hogue CW (2003) BIND: the Biomolecular Interaction Network Database. Nucleic Acids Research 31:248-250.

Baliji S, Lacatus G & Sunter G (2010) The interaction between geminivirus pathogenicity proteins and adenosine kinase leads to increased expression of primary cytokinin responsive genes. Virology 402:238-247.

Balmant KM (2011) Identificação de um fator do hospedeiro, RPS5A, envolvido na interação com a proteína de movimento (MP) de geminivírus. 64 pág. Dissertação MS (Genética e Melhoramento). Universidade Federal de Viçosa.

Barbosa LF, Marubayashi JM, De Marchi BR, Yuki VA, Pavan MA, Moriones E, Navas-Castillo J & Krause-Sakate R (2014) Indigenous American species of the Bemisia tabaci complex are still widespread in the Americas. Pest Management Science 70:1440-1445.

Basso MF (2010) Desenvolvimento de ferramentas sorológicas e moleculares para identificação de vírus em videiras e cochonilhas, alterações fisiológicas e na qualidade enológica da uva de videiras infectadas. 114f. Dissertação (Mestrado em Agronomia) - Curso de Pós-graduação em Agronomia, Universidade Estadual de Ponta Grossa, PR.

Bernardo P, Golden M, Akram M, Naimuddin M, Nadarajan N, Fernandez E, Granier M, Rebelo AG, Peterschmitt M, Martin DP & Roumagnac P (2013) Identification and characterization of a highly divergent geminivirus: evolutionary and taxonomic implications. Virus Research 177:35-45.

Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini FM, Zhou X & Fauquet CM (2008) Recommendations for the classification and nomenclature of the DNA-beta satellites of begomoviruses. Archives of Virology 153:763-781.

Briddon RW, Patil BL, Bagewadi B, Nawaz-Ul-Rehman MS & Fauquet CM (2010) Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evolutionary Biology 10:1-17.

Briddon RW (2003) Cotton leaf curl disease, a multicomponent begomovirus complex. Molecular Plant Pathology 4:427-434.

Brown JK, Fauquet CM, Briddon RW, Zerbini FM, Moriones E & Navas-Castillo J (2012) Family Geminiviridae. In: AMQ King, MJ Adams, EB Carstens & EJ Lefkowitz (Eds), Virus Taxonomy. 9th Report of the International Committee on Taxonomy of Viruses, pp. 351-373. Elsevier Academic Press, London, UK.

Bruckner FP (2012) Aspectos da interação entre a proteína TCTP e o potyvírus PepYMV na infecção de tomateiro e Nicotiana benthamiana. 81 pág. Dissertação MS (Microbiologia Agrícola). Universidade Federal de Viçosa.

Burckstummer T, Bennett KL, Preradovic A, Schutze G, Hantschel O, Superti-Furga G & Bauch A (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. Nature Methods 3:1013-1019.

Calegario RF, Ferreira SS, Andrade EC & Zerbini FM (2007) Characterization of Tomato yellow spot virus (ToYSV), a novel tomato-infecting begomovirus from Brazil. Pesquisa Agropecuária Brasileira 42:1335-1343.

Carmo LST, Resendo RO, Silva LP, Ribeiro SG & Mehta A (2013) Identification of host proteins modulated by the virulence factor AC2 of Tomato chlorotic mottle virus in Nicotiana benthamiana. Proteomics 13:1947-1960.

Carvalho CM, Fontenelle MR, Florentino LH, Santos AA, Zerbini FM & Fontes EPB (2008a) A novel nucleocytoplasmatic traffic GTPase identified as a functional targert of the bipartite geminivirus nuclear shuttle protein. The Plant Journal 55:869-880.

Carvalho CM, Santos AA, Pires SR, Rocha CS, Saraiva DI, Machado JPB, Mattos EC, Fietto LG & Fontes EPB (2008b) Regulated nuclear trafficking of rpL10A mediated by NIK1 represents a defense strategy of plant cells against virus. PLoS Pathogens 4:e1000247.

Cascardo RS (2011) Estudo do envolvimento das proteínas Sl-Gal83 e TCTP na infecção de hospedeiros suscetíveis pelo potyvírus Pepper yellow mosaic virus (PepYMV). 76 pág. Dissertação MS (Genética e Melhoramento). Universidade Federal de Viçosa.

Castillo AG, Kong LJ, Hanley-Bowdoin & Bejarano ER (2004) Interaction between a geminivirus replication protein and the plant sumoylation system. Journal of Virology 78:2758-2769.

Chatr-aryamontri A, Ceol A, Palazzi LM, Nardelli G, Schneider MV, Castagnoli L & Cesareni G (2007) MINT: the Molecular INTeraction database. Nucleic Acids Research 35:D572-574.

Chung HY & Sunter G (2014) Interaction between the transcription factor AtTIFY4B and begomovirus AL2 protein impacts pathogenicity. Plant Molecular Biology 86:185-200.

Clough SJ & Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant Journal 16:735-743.

Collins MO & Choudhary JS (2008) Mapping multiprotein complexes by affinity purification and mass spectrometry. Current Opinion in Biotechnology 19:324-330.

Dufresne PJ, Thivierge K, Cotton S, Beauchemin C, Ide C, Ubalijoro E, Laliberte JF & Fortin MG (2008) Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles. Virology 374:217-227.

Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini FM & Zhou X (2008) Geminivirus strain demarcation and nomenclature. Archives of Virology 153:783-821.

Fernandes AV, Andrade EC, Alfenas PF, Galvão RM, Zerbini FM & Fontes EPB (1998) Clonagem e caracterização de novos geminivírus que infectam Sida rhombifolia. Fitopatologia Brasileira 23:317.

Fernández-Calvino L, Goytia E, López-Abella D, Giner A, Urizarna M, Vilaplana L & López-Moya JJ (2010) The helper-component protease transmission factor of Tobacco etch potyvirus binds specifically to an aphid ribosomal protein homologous to the laminin receptor precursor. Journal of General Virology 91:2862-2873.

Florentino LH, Santos AA, Fontenelle MR, Pinheiro GL, Zerbini FM, Baracat-Pereira MC & Fontes EPB (2006) A PERK-like receptor kinase interacts with the geminivirus nuclear shuttle protein and potentiates viral infection. Journal of Virology 80:6648-6656.

Fondong VN (2013) Geminivirus protein structure and function. Molecular Plant Pathology 14:635-649.

Fontes EPB, Eagle PA, Sipe PS, Luckow VA & Hanley-Bowdoin L (1994) Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. Journal of Biological Chemistry 269:8459-8465.

Fontes EPB, Santos AA, Luz DF, Waclawovsky AJ & Chory J (2004) The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. Genes & Development 18:2545-2556.

Forler D, Kocher T, Rode M, Gentzel M, Izaurralde E & Wilm M (2003) An efficient protein complex purification method for functional proteomics in higher eukaryotes. Nature Biotechnology 21:89-92.

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C & Jensen LJ (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Research 41:D808-D815.

Frischmuth S, Kleinow T, Aberle HJ, Wege C, Hulser D & Jeske H (2004) Yeast two-hybrid systems confirm the membrane association and oligomerization of BC1 but do not detect an interaction of the movement proteins BC1 and BV1 of Abutilon mosaic geminivirus. Archives of Virology 149:2349-2364.

Frischmuth S, Wege C, Hulser D & Jeske H (2007) The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast. Protoplasma 230:117-123.

Fu Q, Li S & Yu D (2010) Identification of an Arabidopsis Nodulin-related protein in heat stress. Molecules and Cells 29:77-84.

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G & Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141-147.

Gilbertson RL, Sudarshana M, Jiang H, Rojas MR & Lucas WJ (2003) Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: Insights into DNA trafficking. Plant Cell 15:2578-2591.

Gloeckner CJ, Boldt K, Schumacher A, Roepman R & Ueffing M (2007) A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. Proteomics 7:4228-4234.

Gully D, Moinier D, Loiseau L & Bouveret E (2003) New partners of acyl carrier protein detected in Escherichia coli by tandem affinity purification. FEBS Letters 548:90-96.

Gutierrez C (1999) Geminivirus DNA replication. Cellular and Molecular Life Sciences 56:313-329.

Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S & Robertson D (1999) Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. Critical Reviews in Plant Sciences 18:71-106.

Hartman NT, Sicilia F, Lilley KS & Dupree P (2007) Proteomic Complex Detection Using Sedimentation. Analytical Chemistry 79:2078-2083.

Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, Margalit H, Armstrong J, Bairoch A, Cesareni G, Sherman D & Apweiler R (2004) IntAct: an open source molecular interaction database. Nucleic Acids Research 32: D452-455.

Hou YM, Sandres R, Ursin VM & Gilbertson RL (2000) Transgenic plants expressing geminivirus movement proteins: abnormal phenotypes and delayed infection by Tomato mottle virus in transgenic tomatoes expressing the Bean dwarf mosaic virus BV1 or BC1 proteins. Molecular Plant-Microbe Interactions 13:297-308.

Jensen BP, Sharp CF, Gardiner SJ & Begg EJ (2008) Development and validation of a stereoselective liquid chromatography-tandem mass spectrometry assay for quantification of S- and R-metoprolol in human plasma. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life 865:48-54.

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

Kinter M & Sherman NE (2000) Protein sequencing and identification using tandem mass spectrometry. New York, John Wiley Interscience, 301p.

Kjemtrup S, Sampson KS, Peele C, Nguyen L, Conkling M, Thompson W & Roberson D (1998) Gene silencing from DNA carried by a geminivirus. Plant Journal 14:91-100.

Kleinow T, Holeiter G, Nischang M, Stein M, Karayavuz M, Wege C & Jeske H (2008) Post-translational modifications of Abutilon mosaic virus movement protein (BC1) in fission yeast. Virus Research 131:86-94.

Knuesel M, Wan Y, Xiao Z, Holinger E, Lowe N, Wang W & Liu X (2003) Identification of novel protein-protein interactions using a versatile mammalian tandem affinity purification expression system. Mol Cell Proteomics 2:1225-1233.

Krenz B, Windeisen V, Wege C, Jeske H & Kleinow T (2010) A plastid-targeted heat shock cognate 70 kDa protein interacts with the Abutilon mosaic virus movement protein. Virology 401:6-17.

Kunik T, Salomon R, Zamir D, Navot N, Zeidan M, Michelson I, Gafni Y & Czosnek H (1994) Transgenic tomato plants expressing the Tomato yellow leaf curl virus capsid protein are resistant to the virus. Biotechnology 12:500-504.

Latham JR, Saunders K, Pinner MS & Stanley J (1997) Induction of plant cell division by Beet curly top virus gene C4. Plant Journal 11:1273-1283.

Laufs J, Schumacher S, Geisler N, Jupin I & Gronenborn B (1995) Identification of the nicking tyrosine of geminivirus Rep protein. FEBS Letters 377:258-262.

Lazarowitz SG (1992) Geminiviruses: Genome structure and gene function. Critical Reviews in Plant Sciences 11:327-349.

Lazarowitz SG & Beachy RN (1999) Viral movement proteins as probe or intracellular and intercellular trafficking in plants. Plant Cell 11:535-548.

Lewis JD & Lazarowitz SG (2010) Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. Proceedings of the National Academy of Sciences, USA 107:2491-2496.

Liu L, Chung HY, Lacatus G, Baliji S, Ruan J & Sunter G (2014) Altered expression of Arabidopsis genes in response to a multifunctional geminivirus pathogenicity protein. BMC Plant Biology 14:1-16.

Lozano-Durán R, Rosas-Díaz T, Gusmaroldi G, Luna AP, Taconnat L, Deng XW & Bejarano ER (2011) Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in Arabidopsis thaliana. The Plant Cell 23:1014-1032.

Lozano-Durán R, Rosas-Díaz T, Luna AP & Bejarano ER (2011) Identification of host genes involved in geminivirus infection using a reverse genetics approach. PLoS ONE 6:e22383:1-13.

Luque A, Sanz-Burgos AP, Ramirez-Parra E, Castellano MM & Gutierrez C (2002) Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. Virology 302:83-94.

Machida K & Mayer BJ (2009) Detection of protein-protein interactions by far-western blotting. Methods in Molecular Biology, Protein Blotting and Detection 536:313-329.

Mansoor S, Briddon RW, Zafar Y & Stanley J (2003) Geminivirus disease complexes: An emerging threat. Trends in Plant Sciences 8:128-134.

Mariano AC, Andrade MO, Santos AA, Carolino SM, Oliveira ML, Baracat-Pereira MC, Brommonshenkel SH & Fontes EPB (2004) Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. Virology 318:24-31.

Martinez-Trujillo M, Limones-Briones V, Cabrera-Ponce JL & Herrera-Estrella L (2004) Improving transformation efficiency of Arabidopsis thaliana by modifying the floral dip method. Plant Molecular Biology Reporter 22:63-70.

McGarry RC, Barron YD, Carvalho MF, Hill JE, Gold D, Cheung E, Kraus WL & Lazarowitz SG (2003) A novel Arabidopsis acetyltransferase interacts with the geminivirus movement protein NSP. The Plant Cell 15:1605-1618.

Melcher U (2000) The '30K' superfamily of viral movement proteins. Journal of General Virology 81:257-266.

Mendes GC, Reis PAB, Calil IP, Carvalho HH, Aragão FJL & Fontes EPB (2013) GmNAC30 and GmNAC81 integrate the endoplasmic reticulum stress- and osmotic stress-induced cell death responses through a vacuolar processing enzyme. PNAS 110:19627-19632.

Miozzo L, Napoli C, Sardo L & Accotto GP (2014) Transcriptomics of the interaction between the monopartite phloem-limited geminivirus Tomato yellow leaf curl sardinia virus and Solanum lycopersicum highlights a role for plant hormones, autophagy and plant immune system fine tuning during infection. PLoS ONE 9:e89951.

Monci F, Sanchez-Campos S, Navas-Castillo J & Moriones E (2002) A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. Virology 303:317-326.

Morales FJ & Anderson PK (2001) The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Archives of Virology 146:415-441.

Murashige T & Skoog F (1962) A revised medium for rapid growth and biossays with tobacco tissue cultures. Plant Physiology 15:473-497.

Navaratnam DS (2009) Yeast two-hybrid screening to test for protein-protein interactions in the auditory system. Methods in Molecular Biology 493:257-268.

Nitsch JP & Nitsch C (1969) Haploid plants from pollen grains. Science 163:85-87.

Noueiry AO, Lucas WJ & Gilbertson RL (1994) Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. Cell 76:925-932.

Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bidwell S, Bridge A, Briganti L, Brinkman FS, Cesareni G, Chatr-aryamontri A, Chautard E, Chen C, Dumousseau M, Goll J, Hancock RE, Hannick LI, Jurisica I, Khadake J, Lynn DJ, Mahadevan U, Perfetto L, Raghunath A, Ricard-Blum S, Roechert B, Salwinski L, Stümpflen V, Tyers M, Uetz P, Xenarios I & Hermjakob H (2012) Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. Nature Methods 9:345-350.

Orozco BM & Hanley-Bowdoin L (1998) Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. Journal of Biological Chemistry 273:24448-24456.

Palmer KE & Rybicki EP (1998) The molecular biology of mastreviruses. Advances in Virus Research 50:183-234.

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

Petty IT, Carter SC, Morra MR, Jeffrey JL & Olivey HE (2000) Bipartite geminivirus host adaptation determined cooperatively by coding and noncoding sequences of the genome. Virology 277:429-438.

Petty ITD, Miller CG, Meade-Hash TJ & Schaffer RL (1995) Complementable and noncomplementable host adaptation defects in bipartite geminiviruses. Virology 212:263-267.

Poojari S, Alabi OJ, Fofanov VY & Naidu RA (2013) A leafhopper-transmissible DNA virus with novel evolutionary lineage in the family Geminiviridae implicated in Grapevine Redleaf Disease by next-generation sequencing. PLoS ONE 8:6:e64194.

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M & Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology 17:1030-1032.

Rohila JS, Chen M, Cerny R & Fromm ME (2004) Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. Plant Journal 38:172-181.

Rohila JS, Chen M, Chen S, Chen J, Cerny R, Dardick C, Canlas P, Xu X, Gribskov M, Kanrar S, Zhu JK, Ronald P & Fromm ME (2006) Protein-protein interactions of tandem affinity purification-tagged protein kinases in rice. Plant Journal 46:1-13.

Rojas MR, Hagen C, Lucas WJ & Gilbertson RL (2005) Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. Annual Review of Phytopathology 43:361-394.

Rojas MR, Noueiry AO, Lucas WJ & Gilbertson RL (1998) Bean dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. Cell 95:105-113.

Romay G, Chirinos D, Geraud-Pouey F & Desbiez C (2010) Association of an atypical alphasatellite with a bipartite New World begomovirus. Archives of Virology 55:1843-1847.

Rubio V, Shen YP, Saijo Y, Liu YL, Gusmaroli G, Dinesh-Kumar SP & Deng XW (2005) An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. Plant Journal 41:767-778.

Sakamoto T, Deguchi M, Brustolini OJB, Santos AA, Silva FF & Fontes EPB (2012) The tomato RLK superfamily: phylogeny and functional predictions about the role of the LRRII-RLK subfamily in antiviral defense. BMC Plant Biology 12:1-18.

Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU & Eisenberg D (2004) The Database of Interacting Proteins: 2004 update. Nucleic Acids Research 32:D449-451.

Sambrook J & Russel D (2001) Molecular Cloning - A Laboratory Manual (3a ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. 2344p.

Sánchez-Durán MA, Dallas MB, Ascencio-Ibañez JT, Reyes MI, Arroyo-Mateos M, Ruiz-Albert J, Hanley-Bowdoin L & Bejarano ER (2011) Interaction between geminivirus replication protein and the SUMO-conjugating enzyme is required for viral infection. Journal of Virology 85:9789-9800.

Sanderfoot AA & Lazarowitz SG (1995) Cooperation in viral movement: The geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. Plant Cell 7:1185-1194.

Sanderfoot AA & Lazarowitz SG (1996) Getting it together in plant virus movement: Cooperative interactions between bipartite geminivirus movement proteins. Trends in Cell Biology 6:353-358.

Santos AA, Lopes KVG, Apfata JAC & Fontes EPB (2010) NSP-interacting kinase, NIK: a transducer of plant defence signalling. Journal of Experimental Botany 61:3839-3845.

Schaffer RL, Miller CG & Petty ITD (1995) Virus and host-specific adaptations in the BL1 and BR1 genes of bipartite geminiviruses. Virology 214:330-338.

Segura MP, Lilley KS & Dupree P (2010) Proteomic Complex Detection using Sedimentation (ProCoDeS): screening for proteins in stable complexes and their candidate interaction partners. Biochemical Society Transactions 38:923-927.

Settlage SB, Miller AB, Gruissem W & Hanley-Bowdoin L (2001) Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. Virology 279:570-576.

Shevchenko A, Tomas H, Havlis J, Olsen JV & Mann M (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols 1:2856-2860.

Silva FN (2011) Interações intra- e intermoleculares nos begomovírus e seus efeitos na adaptação viral e na maquinaria celular do hospedeiro. 126 pág. Tese DS (Fitopatologia). Universidade Federal de Viçosa.

Stanley J, Bisaro DM, Briddon RW, Brown JK, Fauquet CM, Harrison BD, Rybicki EP & Stenger DC (2005) Family Geminiviridae. pp. 301-326 In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U & Ball LA. (Eds.) Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. San Diego: Elsevier Academic Press.

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

Stanley J (1995) Analysis of African cassava mosaic virus recombinants suggest strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. Virology 206:707-712.

Stenger DC, Revington GN, Stevenson MC & Bisaro DM (1991) Replicational release of geminivirus genomes from tandemly repeated copies: Evidence for rolling-circle replication of a plant viral DNA. Proceedings of the National Academy of Sciences 88:8029-8033.

Thivierge K, Cotton S, Dufresne PJ, Mathieu I, Beauchemin C, Ide C, Fortin MG & Laliberte JF (2008) Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. Virology 377:216-225.

Trinks D, Rajeswaran R, Shivaprasad PV, Akbergenov R, Oakeley EJ, Veluthambi K, Hohn T & Pooggin MA (2005) Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. Journal of Virology 79:2517-2527.

Van Leene J, Stals H, Eeckhout D, Persiau G, De Slijke EV, Van Isterdael G, De Clercq A, Bonnet E, Laukens K, Remmerie N, Henderickx K, De Vijlder T, Abdelkrim A, Pharazyn A, Van Onckelen H, Inze D, Witters E & De Jaeger G (2007) A tandem affinity purification-based technology platform to study the cell cycle interactome in Arabidopsis thaliana. Molecular & Cellular Proteomics 6:1226-1238.

Van Leene J, Witters E, Inzé D & De Jaeger G (2008) Boosting tandem affinity purification of plant protein complexes. Trends in Plant Science 13:517-520.

Vanitharani R, Chellappan P, Pita JS & Fauquet CM (2004). Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. Journal of Virology 78:9487-9498.

Varsani A, Navas-Castillo J, Moriones E, Hernández-Zepeda C, Idris A, Brown JK, Zerbini FM & Martin DP (2014). Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. Archives of Virology 159:2193-2203.

Were HK, Winter S & Maiss E (2004) Viruses infecting cassava in Kenya. Plant Disease 88:17-22.

Xie Q, Sanz-Burgos AP, Guo H, Garcia JA & Gutierrez C (1999) Grab proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. Plant Molecular Biology 39:647-656.

Xu X, Song Y, Li Y, Chang J, Zhang H & An L (2010) The tandem affinity purification method: An efficient system for protein complex purification and protein interaction identification. Protein Expression and Purification 72:149-156.

Wang H, Hao L, Shung C-Y, Sunter G & Bisaro DM (2003) Adenosine Kinase is inactivated by geminivirus AL2 and L2 proteins. The Plant Cell 15:3020-3032.

Ward BM, Medville R, Lazarowitz SG & Turgeon R (1997) The geminivirus BL1 movement protein is associated with endoplasmic reticulum-derived tubules in developing phloem cells. Journal of Virology 71:3726-3733.

Wu Y, Li Q & Chen X-Z (2007) Detecting protein-protein interactions by far western blotting. Nature Protocols 2:3278-3284.

Zhang Y, Gao P &Yuan JS (2010) Plant protein-protein interaction network and interactome. Current Genomics 11:40-46.

Zhang SC, Ghosh R & Jeske H (2002) Subcellular targeting domains of Abutilion mosaic geminivirus movement protein BC1. Archives of Virology 147:2349-2363.

Zhang B, Pan X, Cobb GP & Anderson TA (2006) Plant microRNA: A small regulatory molecule with big impact. Developmental Biology 289:3-16.

Zhou Y, Rojas MR, Park MR, Seo YS, Lucas WJ & Gilbertson RL (2011) Histone H3 interacts and colocalizes with the Nuclear Shuttle Protein and the Movement Protein of a geminivirus. Journal of Virology 85:11821-11832.

CONCLUSÕES FINAIS

Capítulo 1:

1. Plantas lenhosas perenes também podem suportar a infecção por vírus de DNA.

2. O TFDaV compartilha características de circo-, nano- e geminivírus.

3. O TFDaV provavelmente será classificado em uma nova família e gênero.

4. O TFDaV pode ser um ancestral comum dos vírus de DNA fita simples circular.

5. RCA, PCR e ELISA indireto foram efetivos na detecção do TFDaV.

Capítulo 2:

1. No total, foram diferencialmente expressos genes pertencentes a 70 famílias diferentes.

2. Diversos genes MIR foram regulados pela infecção viral, possivelmente para aumentar a suscetibilidade do hospedeiro.

3. Diversos genes MIR foram regulados durante a infecção pelo ToYSV, possivelmente atuando na regulação de respostas de defesa do hospedeiro.

4. Diversos genes MIR tiveram sua expressão aumentada ou diminuída de 5 dpi para 10 dpi em plantas infectadas pelo ToYSV.

5. Diversos fatores de transcrição e outras proteínas diferencialmente expressas durante a infecção viral tiveram seus genes MIR reguladores identificados.

6. vsiRNAs podem atuar na regulação de mRNAs do hospedeiro de forma a torná-lo mais suscetível a infecção viral.

7. Validação da expressão dos genes MIR por qPCR e da degradação de seus alvos por degradoma poderá fornecer evidências adicionais da regulação das respostas de defesa do hospedeiro pelos begomovírus.

Capítulo 3:

1. O método TAP original ou com modificações não permitiu obter complexo protéicos em quantidade e pureza.

2. Pull-down foi um método promissor, porém necessita de ajustes.

3. MALDI-TOF/TOF ou ION-TRAP foram relativamente eficientes.

4. Recomenda-se o uso de outras etiquetas: GS-TAP ou SF-TAP.

5. Uso de outros sistemas de busca ou identificação de interações: bibliotecas de cDNA.

6. Uso de outros métodos de avaliação da interação: CoIP e BiFC.

7. Associação com ferramentas de bioinformática: banco de dados e predição baseada na análise em docking.