Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Fitopatologia
Programa de Pós-Graduação em Fitopatologia

Doctoral Thesis

Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics

LUCIANE DE NAZARÉ ALMEIDA DOS REIS

Brasília - DF
2020

LUCIANE DE NAZARÉ ALMEIDA DOS REIS

Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics

Thesis presented to the University of Brasília as a partial requirement for obtaining the title of Doctor in Phytopathology by the Post-Graduate Program in Phytopathology.

Advisor

Dra. Rita de Cássia Pereira Carvalho

Co-advisor

Dr. Leonardo Silva Boiteux

BRASÍLIA, DF– BRASIL
2020

FICHA CATALOGRÁFICA

Reis, A. N. L.

Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics

Luciane de Nazaré Almeida dos Reis.

Brasília, 2020.

Pages number p.:205

Doctoral Thesis - Programa de Pós-Graduação em Fitopatologia, Universidade de Brasília, Brasília, DF.

I- Tomato, NGS, *Geminiviridae*, *Begomovirus, Genomoviridae.*

II- Universidade de Brasília. PPG/FIT.

III- Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics

Aos meus pais Eliecê Almeida dos Reis e Lucival Nunes dos Reis. Ao meu irmão Luan Almeida dos Reis. Aos meus avós Deusarina Goes Almeida e Ubiratan Nascimento Almeida (In memorian). Ao meu Amor Gustavo Ribeiro

Dedico

Agradecimentos

A Deus, dono de toda a ciência, sabedoria e poder. Minha gratidão pelo dom da vida e por toda a força para que eu terminasse mais essa etapa da minha vida.

Aos meus pais Eliecê Almeida dos Reis e Lucival Nunes dos Reis por todo apoio e amor incondicional.

Ao meu irmão amado, Luan Almeida dos Reis.

Ao meu amor, meu namorado Gustavo Ribeiro por todo apoio, paciência e carinho durante todo esse tempo.

Aos pais do meu namorado Márcia Rodrigues e Carlos Rodrigues por toda a ajuda e por me receberem com todo o carinho na sua casa.

A minha orientadora professora Rita de Cássia Pereira Carvalho pela amizade, incentivo e orientação durante todos esses anos de mestrado e doutorado.

Ao meu co-orientador Leonardo Silva Boiteux por toda a ajuda, incentivo e orientação durante todos esses anos de doutorado.

Ao Dr. Fernando Lucas de Melo pela colaboração e incentivo.

A Dra. Maria Esther Noronha Fonseca Boiteux por todo incentivo e colaboração no trabalho.

Aos colegas do mestrado e doutorado Flávia, Josiane, Macária, Ikaro e Juliana.

Aos colegas de laboratório Felipe, Maria Luísa, Amanda, Jordânia e Vinícius.

Em especial agradeço a Josiane e Felipe pela amizade e carinho durante todos esses anos desde o meu mestrado, por terem tornado meus dias mais alegres no laboratório.

As minhas amigas a distância Ghaby Berberian e Kamille Vieira pelas conversas e apoio. A minha companheira e amiga de República Fernanda Kitano, pelo apoio e palavras de incentivo.

Aos amigos que fiz durante esses anos de UnB: Lucas, Catharine, Jamile, João Lucas, Kamila, Vitória, Elenice, Anna Sofya, Érica, Lincon, Sheila, Jefferson e Bianca agradeço por todo o carinho.

Aos professores: Juvenil Enrique Cares, Cleber Furlanetto, Adalberto Côrrea Café Filho, Carlos Hidemi Uesugi, Renato de Oliveira Resende, Fernando Lucas Melo, Maurício Rossato, Alice Kazuko Inoue-Nagata, Robert Neil Gerard Miller, Marisa Álvares da Silva Velloso Ferreira, Helson Mario Martins do Vale, José Carmine Dianese, Luís Eduardo Bassay Blum, Denise Vilela de Rezende e Danilo Batista Pinho.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).
Ao Programa de Pós-Graduação em Fitopatologia da Universidade de Brasília (PPG-FIT).
Ao Conselho Nacional de Pesquisa e Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de estudos.

Work carried out in the Department of Plant Pathology of the Institute of Biological Sciences of the University of Brasília (UnB), under the guidance of Dr. **Rita de Cássia Pereira Carvalho**. Financial Support by Conselho Nacional de Pesquisa e Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Embrapa.

Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics

LUCIANE DE NAZARÉ ALMEIDA DOS REIS
Thesis approved__/__/___by:

Dr. Cleber Furlanetto
Departament of Plant Pathology (UnB)
(Internal Examiner)

Érico de Campos Dianese
Universidade Federal de Goiás (UFG)
(External Examiner)

Dra. Mirtes Freitas Lima
Embrapa Vegetable Crops
(External Examiner)

Gabriel Sérgio Costa Alves
Departament of Cell Biology (UnB)
(External Examiner – Surrogate)

Dra. Rita de Cássia Pereira Carvalho
Departament of Plant Pathology (UnB)
(President)

BRASÍLIA, DF – BRASIL
2020

SUMMARY

LIST OF FIGURES

CHAPTER 1. LITERATURE REVIEW

countries where they were described: BR = Brazil; URU = Uruguay; ARG = Argentina; CHI = Chile. Two TGVV isolates (which complete sequences were obtained in the present study) are highlighted in red (MN928610 and MN928612). GenBank accession numbers of isolates classified/named as ToYVSV are the following: KC706641, KC706633, KC706631, KC706630, KC706653, KC706629, KC706638, KC706637, KC706634, KC706645, KC706651, KC7066, K7066, K7066, K7066, K7070, K7070 KC706644, KC706639, KC706640, KC706646, KC706636, KC706643, KC706642, EF459696, KJ413253, KR024026, KC136339, GQ387369, MN508216, KC136336, KC136337, and EF417915. GenBank accession numbers of isolates classified as TGVV are the following: JF803257, JF803255, JF803258, JF803256, JF803254, and JF803259. GenBank accession numbers of isolates classified as Tomato mottle wrinkle virus (ToMoWrV) are the following: KM243018, KM243019, KM243020, JQ714137, and KY555800. The DNA – A component of a tomato-infecting ToYVSV isolate from Bolivia was only partially characterized (GenBank JQ413300) and for this reason it was not included in the analyses.....119

Figure 2. Phylogenetic tree and Sequence Demarcation Tool (SDT) of a set of DNA–B component sequences showing the phylogenetic identity/distance among Tomato yellow vein streak virus (ToYVSV) and Tomato golden vein virus (TGVV) isolates. The isolates are identified by their accession number and by the acronym of the countries where they were described: BR = Brazil; URU = Uruguay; ARG = Argentina; CHI = Chile. TGVV isolates which complete sequences were obtained in the present study (MN928611 and MN928613) are highlighted in red. GenBank accession numbers of isolates classified as ToYVSV are the following: KC706655, KC706657, KC706665, KC706659, KC706656, KC706667, KC706662, KC706663, KC706660, KC706661 KC706666, KC706664, KC706658, KR024027, KC136340, MN508217, KC136338, and EF417916. GenBank accession number of the isolate classified as TGVV is JF803265. GenBank accession numbers of isolates classified as Tomato mottle wrinkle virus (ToMoWrV) are JQ714138 and KM243017. A tomato-infecting ToYVSV isolate from Bolivia was only partially characterized (GenBank JQ413300) and for this reason it was not included in the analyses.....121

Figure 3. Map of South America showing the geographical distribution of Tomato yellow vein streak virus (ToYVSV) and Tomato golden vein virus (TGVV) isolates. The red dots are representing the geographical areas of occurrence of tomato-infecting TGVV isolates in Brazil (the Federal District-DF and Goiás-GO, Minas Gerais-MG, and Rio de Janeiro-RJ States). The purple dots are indicating the geographical areas of ToYVSV occurrence in Brazil (a potato-infecting isolate in Rio Grande do Sul-RS State and tomato-infecting isolates in São Paulo-SP State) as well as ToYVSV isolates reported infecting tomato, bean, and *Capsicum annuum* crops in the South Cone of South America, including Argentina (AR), Chile (CH), Uruguay (UR). The DNA–A component of a tomato-infecting ToYVSV isolate from Bolivia (BO) was only partially characterized (GenBank JQ413300). 123

CHAPTER 4. A host-guided diversity and speciation of Bean golden mosaic virus isolates from Phaseolus species and from other legume and non-legume plants.

Figure 3. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the DNA–A component sequences of *Bean golden mosaic virus* (BGMV) isolates obtained from *Phaseolus lunatus,* unclassified *Phaseolus* species, and *Macroptilium lathyroides*, indicating their identities in relation to the reference BGMV (NC_004042) isolate (highlighted in red font color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: Isolates *P. lunatus*: [01 (KJ939748), 02 (KJ939739), 03 (KJ939749), 04 (KJ939738), 05 (KJ939746), 06 (KJ939743), 07 (KJ939750), 08 (KJ939741), 09 (KJ939751), 10 (KJ939737), 11 (KJ939744), 12 (KJ939747), 13 (KJ939740, 14 (KJ939745), 15 (KJ939752), 16 (KJ939753), 17 (KJ939742), 18 (KJ939730), 19 (KJ939728), 20 (KJ939727), 21 (KJ939726), 22 (KJ939729), 23 (KJ939736), 24 (KJ939762), 25 (KJ939760), 26 (KJ939754), 27 (KJ939763), 28 (KJ939759), 29 (KJ939761), 30 (KJ939758), 31 (KJ939757), 32 (KJ939755), 33 (KJ939765), 34 (KJ939756), 35 (KJ939712),36 (KJ939717),37 (KJ939715), 38 (KJ939714),39 (KJ939735), 40 (KJ939731), 41 (KJ939722), 42 (KJ939723), 43 (KJ939724), 44 (KJ939764), 45 (KJ939721), 46 (KJ939707), 47 (KJ939718), 48 (KJ939713), 49 (KJ939709), 50 (KJ939734), 51 (KJ939733), 52 (KJ939732), 53 (KJ939725), 54 (KJ939708), 55 (KJ939716), 56 (KJ939719), 57 (KJ939711), 58 (KJ939710), 59 (KJ939720)]; [Isolates from unclassified *Phaseolus* species : 01 (JF694453), 02 (JF694454), 03 (JF694450), 04 (F694451), and 05 JF694449, 06 (JF694452)]; [Isolates from *M. lathyroides*: 01 (JN419006), 02 (N419004), and 03 (JN419003)]……135

Figure 4. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the DNA–A component sequences of Bean golden mosaic virus (BGMV) isolates obtained from *Glycine max*, Macroptilium lathyroides, *Macroptilium erythroloma*, *Anadenanthera colubrina*, *Nicandra physalodes*, and *Solanum lycopersicum*, indicating their identities in relation to the reference BGMV (NC_004042) isolate (highlighted in red color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: [Isolates from *M. lathyroides*: 01 (KJ939725), 02 (KJ939714), 03 (KJ939707), 04 (KJ939756), 05 (KJ939708), 06 (KJ939732), 07 (KJ939764), 08 (KJ939733), 09 (KJ939709), 10 (KJ939717), 11 (KJ939715), 12 (KJ939734), 13 (JN419004), 14 (JN419003), and 15 (JN419006)]; [Isolate from *N. physalodes* (MN737555)]; [Isolates from *S. lycopersicum*: 01 (MN737552), 02 (MN737553), 03 (MN737554)]; [Isolate from G. max (FJ665283)]; [Isolate from *M. erythroloma* (MN822294)]; [Isolate from *A. colubrina* (MN734370)]…….....137

Figure 5. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the of DNA–B component sequences of Bean golden mosaic virus (BGMV) isolates obtained from unclassified Phaseolus species, Phaseolus vulgaris, *P. lunatus*, *Macroptilium lathyroides*, *M. erythroloma*, *Anadenanthera colubrina*, indicating their identities in relation to the reference BGMV (NC_004043) isolate (highlighted in red font color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: Isolates from

substitution model GTR + I + G selected by JModeltest v. 2.2. The analyzes were carried out by running 100 million generations and sampling every 2,000 generations after 2 million burn–in generation. Genome-wide pairwise matrix was generated by SDT v1.2. The isolates are identified by their name, by the GenBank accession number and by the acronym of the countries where they were described: BR = Brazil; TON = Tonga; USA = United States; ZAF = South Africa; NZ = New Zealand; CN = China; GH = Ghana; GE = Germany; LK = Sri Lanka; IR = Iran; BFA = Burkina Faso, and NL = Netherlands. GenBank accession numbers: 1. Pteropus associated gemycircularvirus 9 (KT732795); 2. Pacific flying fox faeces associated gemycircularvirus 4 (KT732796); 3. Thrips associated genomovirus 2 (KY308271); 4. Capybara Genomovirus 9 (MK483081); 5. Plant-associated genomovirus 12 (Tomato – MT214094); 6. Plant Genomovirus 12 (MH939425); 7. Capybara genomovirus 11 (MK483083); Capybara genomovirus 1 (MK483072); 9. Poaceae associated gemycircularvirus 1 (KT253577); 10. Poaceae associated gemycircularvirus 1 (KT253578);11. Poaceae associated gemycircularvirus 1 (KT253579);12. Plant Genomovirus 13 (MH939427); 13. Plant Genomovirus 13 (MH939434); 14. Faecal associated gemycircularvirus 1c (KF371641); 15. Blackbird associated gemycircularvirus 1 (KF371643); 16. Faecal associated gemycircularvirus 3 (KF371639); 17. Faeces associated gemycircularvirus 4 (KF371638); 18. Miniopterus associated gemycircularvirus 1 (KJ641719); 19. Soybean leaf associated gemycircularvirus 1 (KT598248); 20. Pteropus associated gemycircularvirus 3 (KT732797); 21. Bemisia associated genomovirus AdO (KY230614); 22. Hypericum japonicum associated circular DNA virus (KF413620); 23. Momordica charantia associated gemycircularvirus (MH047857); 24. Euphorbia heterophylla associated gemycircularvirus (MH047858); 25. Odonata associated gemycircularvirus 1 (KM598385); 26. Dragonfly associated circular virus 2 (JX185429); 27. Cassava associated gemycircularvirus 1 (JQ412057); 28. Faeces associated gemycircularvirus 22 (KT862253); 29. Sewage derived gemycircularvirus 1 (KJ547638); 30. Sewage associated gemycircularvirus 1 (KM821747); 31. Bromus associated gemycircularvirus 1 (KM510192); 32. Faeces associated gemycircularvirus 17 (KT862242); 33. Sclerotinia gemycircularvirus 1 (GQ365709); 34. Sewage associated gemycircularvirus 6 (KJ547636); 35. Pacific flying fox faeces associated gemycircularvirus 2 (KT732792); 36. Faeces associated gemycircularvirus 16 (KT862251); 37. Poecile atricapillus GI tract–associated gemycircularvirus (KT309029); 38. Pacific flying fox faeces associated gemycircularvirus 10 (KT732804); 39. Pteropus associated gemycircularvirus 8 (KT732806); 40. Pteropus associated gemycircularvirus 5 (KT732801); 41. Dragonfly associated circular virus 1 (JX185430); 42. Capybara genomovirus 2 (MK483074); 43. Plant associated genomovirus 2 (MH939414); 44. Black robin associated gemykibivirus 1 (KF371634); 45. Sewage associated gemykibivirus 3 (KJ547643); 46. HCBI8.215 virus Gemykibivirus (LK931483); 47. Rhinolophus associated gemykibivirus 1 (KJ641737); 48. Gemycircularvirus SL1 Gemykibivirus (KP133075); 49. Dragonfly associated gemyduguivirus 1 (JX185428); 50. Genomoviridae sp. (MK032706); 51. Gila monster associated gemykrogvirus (MH378453); 52. HCBI9.212 virus Gemykrogvirus (LK931484); 53. Caribou associated gemykrogvirus 1 (KJ938717); 54. Gila monster associated gemykrogvirus (MN954869); 55. Sewage associated gemycircularvirus 4

their accession number and by the acronym of the countries where they were described: BR = Brazil; URU = Uruguay; EC = Ecuador; JM = Jamaica; MEX = Mexico; CO = Colombia; USA = United States; VEN = Venezuela; GT = Guatemala. Species and GenBank accession numbers: Cabbage leaf curl virus – CaLCuV (MH359394); Rhynchosia golden mosaic Yucatan virus – RhGMYuV (KP641349); Rhynchosia golden mosaic Sinaloa virus – RhGMV (MK618662); Bean leaf crumple virus – BLCrV (KX857725); Bean calico mosaic virus – BcaMV (AF110189); Euphorbia mosaic virus – EuMV (DQ395342); Tomato twisted leaf virus – ToTLV (MK440292); Jacquemontia yellow vein virus – JacYVV (KY617094); Tomato severe leaf curl virus – ToSLCV (AF130415); Desmodium leaf distortion virus – DeLDV (DQ875870); Abutilon golden mosaic Yucatan virus – AbGMYV (KC430935); Sida golden mosaic Lara virus – SiGMLaV (JX857693); Chenopodium leaf curl virus – ChLCV (HM626515); Cotton leaf crumple virus – CLCrV (AY742220); Wissadula yellow mosaic virus – WYMV (KX691409); Tomato bright yellow mottle virus – ToBYMoV (KC791691); Tomato golden leaf spot virus – ToGLSV (KC626021); Tomato rugose yellow leaf curl virus – TRYLCV (JN381823); Tomato common mosaic virus – ToCmMV (KT203558); Sida yellow leaf curl virus – SiYLCV (EU710750); Tomato chlorotic leaf curl virus – ToCLCV (MK558058); Abutilon Brazil virus – AbBV (FN434438); Abutilon mosaic virus - AbMV (JF694482); Corchorus mottle virus – CoMoV (JQ805781); Sida mosaic Alagoas virus – SiMAV (JF694472); Sida yellow blotch virus – SiYBV (JX871380); Tomato golden leaf distortion virus – ToGLDV (HM357456); Tomato interveinal chlorosis virus2 – ToICV2 (MK087038); Tomato leaf curl purple vein virus – ToLCPVV (KY196216); Tomato yellow vein streak virus – ToYVSV (EF417915); Tomato golden mosaic virus – TGVV (JF80325); Tomato mottle leaf curl virus – ToMoLCV (JF803251); Macroptilium yellow net virus – MaYNV (JN418998); Tomato interveinal chlorosis virus – ToICV (JF803252); Tomato chlorotic mottle virus – ToCMoV (KC706542); Tomato bright yellow mosaic virus – ToBYMV (KC791690); Macroptilium yellow spot virus – MacYSV (JN419013); Bean golden mosaic virus – BGMV (M88686); Tomato rugose mosaic virus – ToRMV (AF2917050; Tomato severe rugose virus – ToSRV (KC004074); Tomato golden mosaic virus – TGMV (JF694490); Tomato mild mosaic virus – ToMMV (EU710752); Tomato yellow spot virus – ToYSV (DQ336350); Okra mottle virus – OMoV (EU914817); Tomato leaf distortion virus – ToLDV (KC706605) and Sida mottle Alagoas virus – SiMoAV (KX896415). 163

LIST OF TABLES

CHAPTER 1: LITERATURE REVIEW

CHAPTER 2. Metagenomics of Neotropical single-stranded DNA viruses in tomato cultivars with and without the *Ty*–1 gene

RESUMO GERAL

Reis, Luciane de Nazaré Almeida. Análise metagenômica da diversidade de begomovírus em tomateiro no Brasil Central e impacto do gene *Ty*-1 na dinâmica evolutiva viral. 2020. **Número de páginas (205). Tese (Doutorado em Fitopatologia) - Universidade de Brasília, Brasília, DF.**

O tomateiro (*Solanum lycopersicum* L.) é uma das principais hortaliças cultivadas no Brasil. Até o início década de 1990, a ocorrência de doenças causadas por espécies de *Begomovirus* (família *Geminiviridae*) era esporádica no país. Entretanto, a partir deste período, um complexo extremamente diverso de espécies de begomovírus emergiu no cultivo do tomateiro, coincidindo com a ampla dispersão geográfica do vetor *Bemisia tabaci* MEAM 1 (*Middle East-Asian Minor* 1= biótipo B). A maioria dos begomovírus apresenta genoma bipartido e apresentam níveis variados de eficiência de transmissão pelo vetor. A utilização mais intensa de híbridos resistentes/tolerantes (principalmente com o gene *Ty*–1) é um potentical fator no processo de evolução deste grupo de vírus no Brasil. A metagenômica aliada ao *Next-Generation Sequencing* – NGS é uma das ferramentas mais eficientes para analisar, em larga escala, a diversidade de populações virais em diferentes condições ambientais. Neste contexto, o objetivo geral do presente trabalho foi conduzir estudos de metagenômica sobre a diversidade de begomovírus em tomateiro no Brasil. Os objetivos específicos foram: **(a)** conduzir estudos comparativos da diversidade de begomovírus infectando tomateiros com e sem o gene *Ty*–1 e (**b)** catalogar as espécies virais predominantes e/ou novas espécies de *Begomovirus* ocorrendo em tomateiros com e sem o gene *Ty*–1. Para isto, 107 amostras foram coletadas em campos de produção em Goiás (n=56), Distrito Federal (n=27) e Minas Gerais (n=24) entre os anos de 2002 e 2016. O DNA total das amostras foi extraído e submetido a PCR usando *primers* para detecção de begomovírus e também com *primers* para região genômica ligada ao gene *Ty*–1. Posteriormente as amostras foram submetidas a um enriquecimento via RCA (*Rolling Circle Amplification*) e divididas em dois *pools*: tomateiros sem o gene *Ty–1* (n=64) e com o gene *Ty*–1 (n=43). Os dois *pools* foram sequenciados em uma plataforma *Illumina HiSeq* 2500. As sequências obtidas foram montadas no programa CLC *Genomics Workbench* 11.0 e analisadas no *Geneious* 10.1. As sequências então foram comparadas com sequências virais presentes no GenBank utilizando o algoritmo BLASTn. Pares de *primers* específicos foram desenhados visando recuperar o genoma completo e confirmar a presença dos vírus em amostras individuais dentro de cada *pool*. Os resultados destas análises estão descritos no **capítulo 2**. Foi observada uma maior diversidade de espécies virais (n=14) no *pool* de amostras sem o gene *Ty*–1 em comparação com aquelas obtidas de plantas com gene *Ty*–1 (n=6). Observou-se uma aparente filtragem entre as espécies detectadas nos dois *pools*. Foi também observada uma grande frequência de infecções mistas nas amostras, tendo casos da ocorrência simultânea de até cinco espécies em uma única amostra. Três potenciais novas espécies foram detectadas, duas em amostras sem o gene *Ty*–1 (MG-378 e GO-169) e uma em amostras contendo o gene *Ty*–1 (DF-640). Além disso, uma espécie do gênero *Gemycircularvirus* e um novo Alfasatélite foram detectados. *Tomato golden vein virus* (TGVV) foi uma das espécies amplamente detectadas nessas análises. Estudos

conduzidos no **capítulo 3**, mostraram que TGVV e *Tomato yellow vein streak virus* (ToYVSV) estão intimamente relacionados como indicado por análises empregando Sequence Demarcation Tool (SDT) e alinhamento MUSCLE. Dois grupos bem definidos foram identificados, consistentes com os critérios atuais para demarcação de espécies de *Begomovirus*, sendo também identificado um conjunto distinto características genômicas, biológicas e ecológicas específicas para cada espécie viral. Uma reavaliação dos isolados de TGVV e ToYVSV disponíveis no GenBank mostrou que uma grande fração está erroneamente classificada ao nível de espécie. A espécie *Bean golden mosaic virus* (BGMV) foi detectada em associação com tomateiro nas análises conduzidas no capítulo 1. No **capítulo 4** a diversidade de 161 isolados classificados como BGMV foi catalogada comparando suas sequências completas com o DNA–A e DNA–B do isolado de referência. Análises filogenéticas e com SDT indicaram que os isolados descritos coletivamente como BGMV compreendem, de fato, duas espécies distintas: uma que engloba isolados de BGMV de *Phaseolus vulgaris* e de uma ampla gama de hospedeiros (incluindo o tomateiro) e uma espécie estreitamente relacionada (com identidade variando de 89 a 91% em comparação com o isolado de referência de BGMV) principalmente associada ao feijão-lima (*P. lunatus*). O **capítulo 5** descreve as características moleculares de um *Gemycircularvirus* (2.189 nucleotídeos) identificado em associação com o tomateiro no Brasil Central. As análises mostraram que a espécie identificada compartilhou 99% de identidade com um vírus provisoriamente denominado como Plant-associated genomovirus 12 de *Larrea tridentata*. O **capítulo 6** descreve duas novas espécies de *Begomovirus* que foram identificadas em amostras de Minas Gerais e Goiás. Os genomas virais completos foram clonados, sequenciados via Sanger e provisoriamente denominados Tomato golden net virus – ToGNV (2.649 nucleotídeos) e Tomato yellow net virus – ToYNV (2.636 nucleotídeos). Ambos os vírus exibiram a organização do DNA–A com características típicas das espécies de begomovírus do Novo Mundo. No entanto, nenhum componente cognato do DNA–B foi encontrado, indicando que ToGNV e ToYNV provavelmente compreendem um grupo peculiar de begomovírus neotropicais monopartidos.

Palavras chaves: *Begomovirus*, diversidade, *Solanum lycopersicum* L., *Next-Generation Sequencing*, resistência genética

GENERAL ABSTRACT

Reis, Luciane de Nazaré Almeida. Metagenomic analysis of the begomovirus diversity in tomatoes in Central Brazil and impact of the *Ty*-1 tolerance gene on viral evolutionary dynamics. 2020. **Number of pages (205). Thesis (PhD in Phytopathology) - University of Brasília, Brasília, DF.**

Tomato (*Solanum lycopersicum* L.) is one of the main vegetable crops cultivated in Brazil. Until the early 1990s, the occurrence of diseases caused by *Begomovirus* species (family Geminiviridae) was sporadic in the country. However, from this period on, an extremely diverse complex of begomoviruses emerged in tomato fields, coinciding with the wide geographical dispersion of the *Bemisia tabaci* MEAM 1 (*Middle East-Asian Minor 1* = biotype B). Most begomoviruses have a bipartite genome and have varying levels of transmission efficiency by the vector. The more intense use of resistant/tolerant hybrids (mainly with the *Ty*–1 gene) is a potentical factor in the evolution process of this group of viruses in Brazil. Metagenomics combined with Next-Generation Sequencing (NGS) is one of the most efficient tools for large scale analysis of the diversity of viral populations in different environmental conditions. In this context, the general objective of the present work was to conduct metagenomics studies on the diversity of begomoviruses in tomatoes in Brazil. The specific objectives were: **(a)** to carry out comparative studies of the begomovirus diversity infecting tomato plants with and without the *Ty*–1 gene and **(b)** to catalog the predominant viral species and/or new begomoviruses occurring in tomato plants with and without the *Ty*–1 gene. For this, 107 samples were collected in production fields in Goiás (n = 56), Distrito Federal (n = 27) and Minas Gerais (n = 24) between the years 2002 and 2016. The total DNA of the samples was extracted and submitted to PCR using primers to detect begomovirus and also with primers for the genomic region linked to the *Ty*–1 gene. Subsequently, the samples were subjected to enrichment via RCA (Rolling Circle Amplification) and divided into two pools: tomatoes without the *Ty*–1 gene (n = 64) and with the *Ty*–1 gene (n = 43). The two pools were sequenced on an Illumina HiSeq 2500 platform. The obtained sequences were assembled using the CLC Genomics Workbench 11.0 program and analyzed in Geneious 10.1. The sequences were then compared to viral sequences present on GenBank using the BLASTn algorithm. Specific primer pairs were designed to recover the complete genome and confirm the presence of viruses in individual samples within each pool. The results of these analyzes are described in **Chapter 2**. A greater diversity of viral species (n = 14) was observed in the sample pool without the *Ty* – 1 gene compared to those obtained from plants with the *Ty* – 1 gene (n = 6). It was observed an apparent filtering effect among viral species detected in the two pools. A high frequency of mixed infections was also observed in the samples, with cases of the simultaneous occurrence of up to five species in a single sample. Three potential new species were detected, two in samples without the *Ty*–1 gene (MG-378 and GO-169) and one in samples containing the *Ty*–1 gene (DF-640). In addition, a species of the genus Gemycircularvirus and a new alpha-satellite were detected. Tomato golden vein virus (TGVV) was one of the species widely detected in these analyzes. Studies conducted in

Chapter 3 have shown that TGVV and Tomato yellow vein streak virus (ToYVSV) are closely related as indicated by analyzes using Sequence Demarcation Tool (SDT) and MUSCLE alignment. Two well-defined clusters were identified, consistent with the current criteria for demarcation of *Begomovirus* species. In addition, a distinct set of genomic, biological and ecological characteristics specific to each viral species was identified. A reassessment of the TGVV and ToYVSV isolates available on GenBank showed that a large fraction of them is erroneously classified at the species level. Bean golden mosaic virus (BGMV) was also detected in association with tomato in the analyzes carried out in Chapter 1. In **Chapter 4** the diversity of 161 isolates classified as BGMV was cataloged by comparing their complete sequences with the DNA–A and DNA–B components of reference isolate. Phylogenetic and SDT analyzes indicated that the isolates collectively described as BGMV actually comprise two distinct species: one that encompasses isolates of BGMV from *Phaseolus vulgaris* and from a wide range of hosts (including tomato) and a closely related species (with identity ranging from 89 to 91% compared to the reference BGMV isolate), which were mainly associated with lima beans (*P. lunatus*). **Chapter 5** describes the molecular characterization of a Gemycircularvirus (2,189 nucleotides) identified in association with tomato in Central Brazil. The analyzes showed that the gemycircularvirus shared 99% of identity with a virus tentatively named as Plant-associated genomovirus 12 of *Larrea tridentata*. **Chapter 6** describes two new *Begomovirus* species that were identified in samples from Minas Gerais and Goiás states. The complete viral genomes were cloned, sequenced via Sanger and tentatively named as Tomato golden net virus – ToGNV (2,649 nucleotides) and Tomato yellow net virus – ToYNV (2,636 nucleotides). Both viruses exhibited DNA–A organization with typical features of the New World begomovirus species. However, no cognate components of DNA–B were found, indicating that ToGNV and ToYNV might comprise a peculiar group of monopartite neotropical begomoviruses.

Keywords: *Begomovirus*, diversity, *Solanum lycopersicum* L., Next-Generation Sequencing, genetic resistance

GENERAL INTRODUCTION

The tomato (*Solanum lycopersicum* L.) is one of the main vegetable crops in the world, being cultivated across all continents. In Brazil, this crop has a high economic and social importance, due to its high demand for labor, production value and cultivation area (IBGE, 2020). According to FAOSTAT (2020), Brazil occupies the tenth global position in tomato production (58.168 hectares), reaching ≈ 4.1 million tons per year. The main tomato-producing states are Goiás (GO), São Paulo (SP) and Minas Gerais (MG) (IBGE, 2020). The almost uninterrupted cultivation of tomatoes throughout most of the year in most of the Brazilian regions favors the incidence of several diseases in crops for fresh-market and for processing. All groups of pathogens (*viz*. fungi, bacteria, nematodes, and viruses) have been reported infecting tomatoes under natural conditions on a global scale (Jones et al., 2014). However, viral pathogens are the ones that have the greatest difficulty in establishing effective control strategies. In Brazil, the main diseases of viral etiology are caused by species of the genera *Begomovirus*, *Crinivirus*, *Orthotospovirus*, *Tobamovirus* and *Potyvirus* (Lopes and Reis, 2011; Inoue-Nagata et al., 2016b).

Species classified in *Begomovirus* genus (Family: *Geminiviridae*) are characterized by single-stranded circular DNA (ssDNA) with either one genomic component (DNA–A for monopartite species) or two genomic components (DNA–A and DNA–B for bipartite species) separately encapsulated in twinned particles (Rojas et al., 2005a; Rojas et al., 2018; ICTV, 2020). More than one hundred species of begomoviruses have already been characterized infecting tomato in the world. Brazil is considered as one of the most important centers of diversity of bipartite begomoviruses (Fernandes et al., 2008). The transmission of viral species of this genus occurs naturally through a complex of cryptic species of whitefly (*Bemisia tabaci*) in a relationship with the vector that is characterized as persistent circulative (De Barro et al., 2011).

Currently, the highest incidence of viral diseases in tomato in Brazil are those caused by begomoviruses. The major relevance of the begomoviruses is due to a series of factors, including the type of dissemination and high population density of their vectors (whiteflies), the wide range of alternative hosts, and the genetic mechanisms to generate genetic diversity in this group of viruses, which favors the emergence of new species. The first report of a begomovirus in tomato in Brazil occurred in 1960 (Flores et al., 1960). However, until the 1990s, the occurrence of begomoviruses in tomato in the country was sporadic and without economic importance. In the early 1990s, with the introduction in the country of the polyphagous vector *B. tabaci* biotype B (= *B. tabaci Middle East-Asia*

Minor 1– MEAM1), a significant increase in the incidence and in the diversity of begomoviruses was observed (Ribeiro et al., 2003; Fernandes et al., 2008). Field surveys of begomoviruses associated with the tomato crop revealed an extremely diverse complex of viral species in Brazil (Ribeiro et al., 2003; Cotrim et al., 2007; Castillo-Urquiza et al., 2008; Fernandes et al., 2008). Currently, 21 begomoviruses have been reported infecting tomatoes in Brazil and all of them were accepted by the ICTV (Matyis et al., 1975; Ribeiro et al., 2003; Fernandes et al., 2006; Calegario et al., 2007; Ribeiro et al., 2007; Castillo-Urquiza et al., 2008; Fernandes et al., 2008; Albuquerque et al., 2012; Macedo et al., 2018; ICTV, 2020). The mechanisms of generating genetic variability in begomoviruses (mutation, recombination and pseudo-recombination) contribute to this current scenario (for review see Rossinck 1997; Seal et al 2006; Duffy and Holmes, 2008). Mutation and recombination are the most important mechanisms in begomoviruses, resulting in the emergence of new species and strains (Rocha et al., 2013). In fact, omparisons of sequences of begomovirus isolates reported in tomato in Brazil have indicated strong evidence of recombination events among viral species, resulting in a high degree of genetic diversity of these species in the country. An illustrative example is the case of *Tomato rugose mosaic virus* (ToRMV), which probably arose from a recombination event between *Tomato severe rugose virus* (ToSRV) and *Tomato chlorotic mottle virus* (ToCMoV) (Ribeiro et al., 2007).

The chemical control of the vectors has a low efficiency, making the use of resistant varieties the major strategy to minimize the losses caused by begomoviruses. In tomatoes, eight resistance genes/alleles have been reported: *Ty*–1 (Zamir et al., 1994), *Ty*–2 (Hanson et al., 2006), *Ty*–3 (Ji and Scott, 2006), *Ty*–4 (Ji et al., 2009), *ty*–5 (Anbinder et al., 2009), *Ty*–6 (Hutton et al., 2012), *tcm*–1 (Giordano et al., 2005b), and *tgr*–1 (Bian et al., 2007). Due to the extreme variability of the begomoviruses infecting tomato in the country, it is possible that new species and strains (not yet detected and/or characterized), may be occurring in the main producing regions. In fact, the increase of the areas with varieties and hybrids carrying the *Ty*–1 gene (Boiteux et al., 2007a) constitute a new selection factor towards more adapted isolates that may be even capable overcoming this factor.

Distinct strategies have been used to analyze the evolutionary processes capable of shaping the genetic-molecular structure of the begomovirus populations. The main methodological approach has been the sequencing of the complete viral genome (DNA–A and DNA–B components), which enables the characterization of the gene repertoire

and the elucidation of processes potentially involved in plant-virus interaction, providing crucial information for development of new control methods. Among the strategies available to assess viral diversity, metagenomics combined with Next-generation Sequencing – NGS has been providing great advances especially in the identification of new plant-infecting and plant-associated virus species (Barba et al., 2014; Pecman et al., 2017; Hadidi, 2019). In this context, the general objective of the present thesis was to carry out a study on the diversity of begomovirus species occurring in tomatoes in Central Brazil via metagenomic analysis using NGS. In addition, analyzes were also conducted to estimate the potential impacts of the introduction of resistant / tolerant tomato varieties (containing the *Ty*–1 gene) on viral evolutionary dynamics.

HYPOTHESES

- The use of tomato plants with the *Ty*–1 gene is restricting the genetic diversity of begomovirus populations in tomatoes.
- New species of begomovirus are occurring in the Central region of Brazil due to the selection pressure caused by presence of the *Ty*–1 gene.

GENERAL OBJECTIVE

- To conduct a study of metagenomic analysis of begomovirus diversity infecting tomatoes in Central Brazil in order to estimate the impact of the introduction of varieties containing the *Ty*–1 gene on the evolutionary dynamics of *Begomovirus* species.

SPECIFIC OBJECTIVES

- To elucidate the diversity of the begomoviruses infecting tomato varieties with and without the *Ty*–1 gene in Central Brazil.
- To catalog the predominant viral species and / or new species of begomovirus that are capable of overcoming *Ty*–1 resistance gene.

CHAPTER 1. LITERATURE REVIEW

1. The tomato

Tomato (*Solanum lycopersicum* L.) is classified in the class Magnoliopsida, order Tubiflorae, family Solanaceae and genus *Solanum* (Naturdata, 2020). The Solanaceae family contains 106 genera and ≈ 3,000 species. It has a cosmopolitan distribution, with South America being one of the main centers of diversity and endemism. In the Solanaceae family, in addition to tomatoes, other species of great economic importance are included, such as potatoes (*S. tuberosum* L.), eggplant (*S. melongena* L.), tobacco (*Nicotiana tabacum* L.), hot peppers (*Capsicum* spp.), sweet peppers (*C. annuum* L.), and scarlet eggplant (*S. aethiopicum* var. *gilo* L.). *Solanum* is the largest genus within the Solanaceae family contains around 1,500 species that are distributed throughout South America. In Brazil, about 350 species of the genus *Solanum* have been identified, many of which are endemic (Silva et al., 2006; Pereira et al., 2016).

The tomato domestication was carried out by indigenous tribes in Puebla and Vera Cruz in Mexico. The tomato was considered, for some time, as a poisonous plant, being employed only for ornamental purposes. In Brazil, commercial tomato cultivation was introduced by European immigrants at the end of the 19th century (Alvarenga and Coelho, 2013). The tomato crop is considered the most important vegetable in the world, being used for fresh consumption and for industrial processing (Vilela et al., 2012). China is the largest tomato producer followed by India, Turkey, and the United States. Brazil is currently the 10th world producer (FAO, 2020). The total area cultivated with tomatoes in the country is about 58,166 hectares (ha) with a production of 4.1 million tons and average yield of ≈ 58 tons per hectare.

In Brazil, the Southeast is the main tomato-producing region (≈ 45% of the total Brazilian production) followed by the Center-East region (≈ 30%) and the Northeast (≈ 13%). The State of Goiás (GO), located at Center-East region, is the main producer and it concentrates the largest area with tomato crops for industrial processing (1,290,134 tons), followed by São Paulo (SP) with 860,600 tons and Minas Gerais (MG) with 523,525 tons (IBGE, 2020). In Brazil, fresh-market tomato represents an important source of employment and income across its entire production chain (Vilela et al., 2012). It is estimated that tomato cultivation from soil preparation to up to harvesting, requires four

to five workers per hectare, generating an average of 106,000 direct jobs (Socoloski et al., 2017).

The type of conduction / management of the tomato crop is mainly defined by the growth habit of the plant (i.e. determined or indeterminate). The determined tomato is preferentially employed for industrial processing. This characteristic is conditioned by the recessive *self-prunning* gene gene (*sp*), which phenotype is a plant with reduced size and short internodes (Boiteux et al., 2012). In tomato cultivars with indeterminate growth, even after the appearance of flower buds, the plant continues to grow, with the simultaneous presence of ripe fruits and flower buds still opening (Silva et al., 2006; Alvarenga and Coelho, 2013). In Brazil, tomato planting is carried out almost all year round. This continuous cultivation represents a challenge for the growers, mainly due to disease and pest problems that can affect the crop in different degrees of severity. Under these growing conditions, tomato production can be affected by various pathogens, pests and virus vectors that may cause yield losses and/or a significant increase in production costs due to the use of pesticides (Lopes and Reis, 2011; Alvarenga and Coelho, 2013).

2. Main pathogens in tomato crop

2.1. Main fungal pathogens

The main fungal diseases of the tomato crop on a global scale are as follow: early blight caused by *Alternaria tomatophila* (= *A. linariae* and previously referred as *A. solani*), late blight (caused by the oomycete *Phytophthora infestans*), Septoria leaf spot (*Septoria lycopersici*), Fusarium wilts (caused by three races of *Fusarium oxysporum* f. sp. *lycopersici*), crown rot (caused by *F. oxysporum* f. sp. *radicis-lycopersici*), white mold (*Sclerotinia sclerotiorum*), gray leaf spot (*Stemphylium solani* e *S. lycopersici*), *Corynespora* spot (*Corynespora cassiicola*), adaxial powdery mildew (*Oidium neolycopersici*), abaxial powdery mildew (*Oidiopsis haplophylli*), *Cladosporium* spot (caused by different races of *Passalora fulva*), damping-offs (*Pythium* spp, *Phytophthora* spp. e *Rhizoctonia solani*) and Verticillium wilt caused by two races of *Verticillium dahliae* (Lopes and Reis, 2011; Jones et al., 2014).

2.2. Main bacterial pathogens

The main bacterial diseases causing significant damage to tomato production are: bacterial spot caused by a complex of *Xanthomonas* species; bacterial speck caused by *Pseudomonas syringae* pv. tomato, pith necrosis (*P. corrugata* and *P. mediterranea*);

bacterial wilt (caused by a complex of species and isolates of *Ralstonia solanacearum* and *R. pseudosolanacearum*); bacterial canker (*Clavibacter michiganensis* subsp. *michiganensis*) and soft rot (caused by a complex *Pectobacterium* and *Dickeya* species) (Lopes and Reis, 2011).

2.3. Main nematode pathogens

In Brazil, the main nematodes affecing the tomato crop are the causal agents of the root-knot disease, which are classified within the genus *Meloidogyne* (Pinheiro et al., 2014). Recently, populations of *Pratylenchus* sp. have also been reported inducing necrotic root lesions in tomatoes in Brazil.

2.4. Diseases of viral etiology

The economic importance of viral diseases in tomatoes is dependent upon the geographical region, the type of cultivation, and the vector dissemination and distribution. Isolates from about 286 viral species have been reported infecting tomatoes worldwide (Ong et al., 2020; Virus-HostDB, 2020) **(Figure 1).** In Brazil, the main viruses affecting tomato crops are classified in the genera *Begomovirus*, *Orthotospovirus*, *Crinivirus*, and *Tobamovirus* (Inoue-Nagata et al., 2016b). Isolates of *Cucumovirus*, *Potyvirus* and *Polerovirus* have also been reported in the crop as well as isolates of *Tobravirus* (Cupertino et al., 1991), *Amalgavirus* (Martins, 2017) and *Tymovirus* (Oliveira et al., 2013).

Some emerging tomato viruses have not yet been reported in Brazil. The viral pathogens present on the list of quarantine pests from the Ministério da Agricultura e Pecuária e Abastecimento (MAPA, 2020) are: Tomato black ring virus and Tomato ringspot virus (genus *Nepovirus*), Pepino mosaic virus (genus *Potexvirus*) and Perlagonium zonate spot virus (genus *Anulavirus*). Beside these species, there is the threat of introducing the Tomato brown rugose fruit virus (genus *Tobamovirus*) into the country, which is capable of 'breaking' the resistance controlled by the gene *Tm*–2^2 (Luria et al., 2017).

However, diseases caused by *Begomovirus* species (Family *Geminiviridae*) deserve special mention in Brazil because they induce severe symptoms and occur more frequently (due to high population densities of the whitefly vector and due to the wide range of alternative hosts of both vector and viral pathogens).

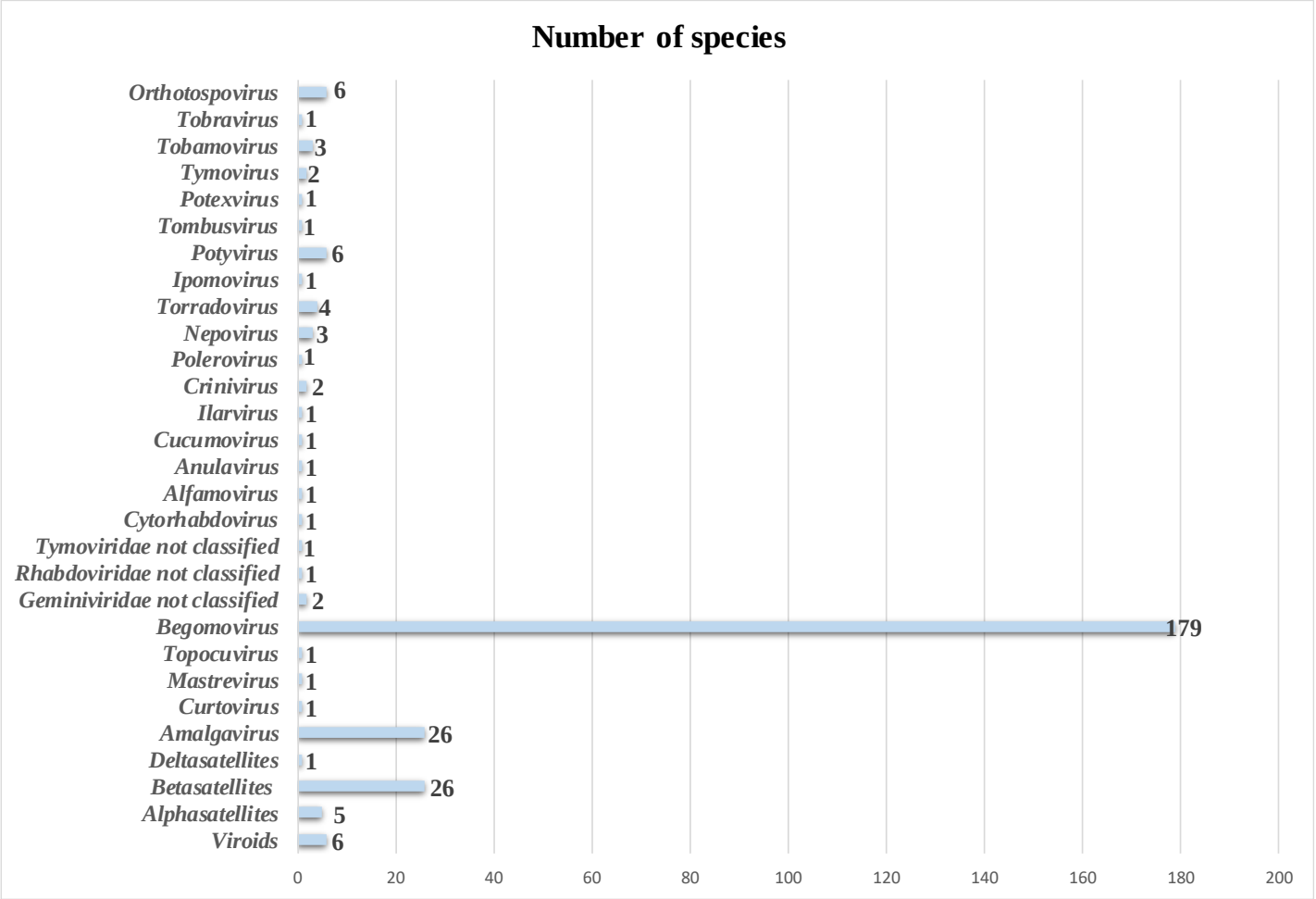

Figure 1. Total number of viruses classified by genus and/or family reported in association with tomato worldwide.

3. Family *Geminiviridae*

Species classified within the genera of the family *Geminiviridae* (Order: *Geplafuvirales*) are responsible for economic crop losses around the world mainly in tropical and subtropical regions. Viruses into *Geminiviridae* family are characterized by single-stranded circular DNA genomes, encapsulated in twinned icosahedral particles (18–20 x 30–32 nm), and may have only one (= monopartite species) or two (= bipartite species) DNA molecules (Varsani et al., 2014a; Brown et al., 2015; Rojas et al., 2018). Virus species within this family induce severe losses in a wide host range worldwide. Some of the major disease in terms of economic and social impacts are the ones caused African cassava mosaic virus (ACMV) in cassava in Africa, Bean golden mosaic virus (BGMV) and Bean golden yellow mosaic virus (BGYMV) on beans in the Americas; Beet curly top virus (BCTV) on eggplants in North America; Cotton leaf curl virus (CLCuV) inffecting cotton in Asia; Maize streak virus (MSV) on corn crops in Africa; Tomato yellow leaf curl virus (TYLCV) affecting tomato crops in Africa, the Americas, Asia, and Europe (Fondong, 2013) and a complex of bipartite and monopartite species also affecting tomato cultivation in South America (Inoue-Nagata et al., 2016a).

The family *Geminiviridae* is the largest family of plant viruses with 485 species described to date (ICTV, 2020). These species are distributed in nine genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* (**Table 1**). The classification in genera is based upon the host range, the type of insect vector(s), genomic organization and phylogenetic relationships (Brown et al., 2015; Varsani et al., 2017; ICTV, 2020). Beside the nine genera, two isolated species, accepted by ICTV, Citrus chlorotic dwarf associated virus (CCDaV) (Loconsole et al., 2012) and Mulberry mosaic dwarf associated virus (MMDaV) are classified in the *Geminiviridae* family (Lu et al., 2015; Ma et al., 2015). Other species (not yet accepted by ICTV) have been reported as potential new geminiviruses including: Mulberry crinckle leaf virus (Lu et al., 2015), Apple geminivirus – AGV (Liang et al., 2015), Grapevine geminivirus A– GGVA (Al Rwahnih et al., 2016), Tomato associated geminivirus 1 – TaGV1 (Fontenele et al., 2017), Tomato apical leaf curl virus – ToALCV (Vaghi Medina et al., 2018; Batista et al., 2019) and Passion fruit chlorotic mottle virus – PCMoV (Fontenele et al., 2018a). Recently, two species of geminiviruses were found in *Limeum africanum* L. and *Polygala garcinii* L. in South Africa and one in *Juncus maritimus* L. in France. The species were named Limeum africanum-associated virus – LaaV, Polygala garcinii-associated virus – PgaV and Juncus

maritimus-associated virus – JmaV (Claverie et al., 2018), respectively. Other recent report described a new geminivirus (Common bean curly stunt virus – CBCSV) in common beans (*Phaseolus vulgaris* L.), which recombination analyzes indicated that it may have a recombinant origin (Zhang et al., 2020).

Table 1. Genera classified in the family *Geminiviridae* (ICTV, 2020).

Genome type	Hosts	Genera	Type species	Vectors	Number of species classified in the genera
Bipartite	**Dicotyledoneous**	*Begomovirus*	*Bean golden yellow mosaic* (BGYMV)	Whitefly (*Bemisia tabaci*)	424
Monopartite	**Monocotyledoneous**	*Eragrovirus*	*Eragrostis curvula streak virus* (ECSV)	Unknown	1
		Mastrevirus	*Maize streak virus* (MSV)	leafhopper (*Cicadulina mbila*)	41
	Dicotyledoneous	*Becurtovirus*	*Beet curly top Iran virus* (BCTIV)	leafhopper (*Circulifer haematoceps*)	3
		Begomovirus	*Tomato yellow leaf curl virus* (TYLCV)	Whitefly (*Bemisia tabaci*)	424
		Capulavirus	*Euphorbia caput-medusae latent virus* (EcmLV)	Aphid (*Aphis craccivora*)	4
		Curtovirus	*Beet curly top virus* (BCTV)	leafhopper (*Circulifer tenellus*)	3
		Grablovirus	*Grapevine red blotch virus* (GRBV)	leafhopper (*Spissistilus festinus*)	3
		Mastrevirus	*Tobacco yellow dwarf virus* (TbYDV)	leafhopper (*Cicadulina mbila*)	41
		Topocuvirus	*Tomato pseudo-curly top virus* (TPCTV)	Membracídeo (*Micrutalis malleifera*)	1
		Turncurtovirus	*Turnip curly top virus* (TCTV)	leafhopper (*Circulifer haematoceps*)	3
Total number of species classified into genera of the family *Geminiviridae*					**483**

3.1. Genera of the *Geminiviridae* family

3.1.1. *Becurtovirus*

This genus is represented by isolates of three species: the type species Beet curly top Iran virus (BCITV) with the species Spinach curly top Arizona virus (SCTAV) and Exomis microphylla latente virus (EmLV) (ICTV, 2020). BCITV isolates have been reported only in Iran and they can infect more than three hundred species of dicotyledonous plants, such as tomato, beet (*Beta vulgaris* L.), *Beta vulgaris* subsp. *maritima*, cowpea (*Vigna unguiculata* L.), beans (*P. vulgaris* L.) and hot pepper (*Capsicum frutescens* L.) (Strausbaugh et al., 2017). The SCTAV species was reported infecting only spinach (*Spinacia oleracea* L.) in Arizona, USA (Hernández-Zepeda et al., 2013). An EmLV isolate was recently reported in *Exomis microphylla* L. (Claverie et al., 2018). The isolates of species classified in this genus are characterized by having the nonanucleotide ("TAA**G**ATT**C**C"), which is distinct from the other geminiviruses in the 4th and 8th positions, where T and A are typically found, respectively. The species of this genus have three ORFs (open reading frames) in the viral sense: V1 (capsid protein), V2 (movement protein) and V3 (movement protein) and two in the complementary sense: C1 (protein associated with replication) and C2 (transcription activation protein) (**Figure 2**) (Varsani and Krupovic, 2017; ICTV, 2020).

3.1.2. *Capulavirus*

This genus currently has four described species: Euphorbia caput-medusae latent virus (EcmLV), Alfalfa leaf curl virus (ALCV), French bean severe leaf curl virus (FbSLCV) and Plantago lanceolata latent virus (PlLV). Species of this genus were reported infecting Euphorbia caput-medusae in South Africa, beans (*P. vulgaris* L.) in India, alfalfa (*Medicago sativa* L.) in Spain and France, and *Plantago lanceolata* in Finland (Varsani et al., 2017). Capulaviruses have a genomic organization with two intergenic regions (similar to mastreviruses and becurtoviruses). The capulavirus isolates (in common with begomoviruses and curtoviruses) have a large ORF in the complementary sense (C3) that is incorporated into Rep. Another characteristic of the capulaviruses is the presence of potential ORFs (located in the 5' region of CP) that encode movement proteins (**Figure 2**). All capulaviruses are characterized by the nonanucleotide "TAA**T**ATT**A**C" (Bernardo Pauline et al., 2016; Varsani et al., 2017; ICTV, 2020).

3.1.3. *Curtovirus*

This genus is represented by three species: Beet curly top virus (BCTV), Horseradish curly top virus (HrCTV), and Spinach severe curly top virus (SSCTV). BCTV isolates infect a wide range of dicotyledoneous plants, including ≈ 300 species in 44 families (Strausbaugh et al., 2008; ICTV, 2020). The curtovirus genome (as seen in most of the geminivirus members) is composed by three ORFs in the viral sense (V1, V2 and V3) and four ORFs in the complementary sense (C1, C2, C3 e C4) (**Figure 2**) (Varsani et al., 2017; ICTV, 2020).

3.1.4. *Eragrovirus*

This genus is currently represented by isolates of the species Eragrostis curvula streak virus (ECSV). All isolates have been reported in *Eragrostis curvula* (Schrad.) Nees. in South Africa. Like becurtoviruses, ECSV isolates have a peculiar nonanucleotide with differences in the fourth and eighth position "TAA**G**ATT**C**C" (Varsani et al., 2014a; Varsani et al., 2017).

3.1.5. *Grablovirus*

This genus is represented by three species: the type-species Grapevine red blotch virus (GRBV), which was initially reported in cultivated grapevines (*Vitis vinifera* L.) (Krenz et al., 2012; Varsani et al., 2017). More recently, two new species have been accepted: one in wild grapevine (*Vitis* sp.), named Wild vitis latent virus – WvLV (Perry et al., 2018) and the other – Prunus latent virus (PrLV) – obtained from asymptomatic samples of plum (*Prunus salicina*) (Al Rwahnih Maher et al., 2018). Like the most members of the family *Geminiviridae*, the genomic organization of the viruses from the genus *Grablovirus* consists of three ORFs in the viral sense (V1, V2, and V3) and three in the complementary sense (C1, C2, and C3). ORF C3 is fully incorporated into C1 and its function remains unknown (Varsani et al., 2017) (**Figure 2**).

3.1.6. *Mastrevirus*

The genus *Mastrevirus* is currently represented by 41 species. Most isolates of these species have been reported infecting monocots. However, some isolates are capable of infecting dicots (with hosts within the Solanaceae and Fabaceae families) such as Tobacco yellow dwarf virus (TbYDV) (Trębicki et al., 2010) and Chickpea chlorotic

dwarf virus (CpCDV), which has been reported infecting chickpeas (*Cicer arietinum* L.) (Nahid et al., 2008). Initially, mastrevirus isolates were found only in the Old World (Asia, Africa and Europe) and Oceania (Australia). However, more recently, mastreviruses have also been reported in the Americas in *Panicum virgatum* L. (Agindotan et al., 2015) and sweet potato (*Ipomoea batatas* L.) (Kreuze et al., 2009; Cao et al., 2017). Recently, the first mastrevirus in the Americas was detected and identified through metagenomic analysis of leafhopper tissues (*Dalbulus maidis*) in Itumbiara–GO, Brazil (Fontenele et al., 2018b). Afterward, Sweet potato symptomless virus 1 (SPSMV 1) isolates have been reported in sweet potato clones collected across all Brazilian regions (Souza et al., 2018). The mastrevirus genome is composed of four ORFs, two in the viral sense – V1 (protein cover) and V2 (movement protein) – and two ORFs in the complementary sense (C1 and C2) that are related to replication. The mastrevirus isolates have the common nonanucleotide sequence of most geminiviruses, the only exception referring to SPSMV–1 isolates that have different nucleotides in the fourth and eighth position "TAA**G**ATT**C**C" (Cao et al., 2017; Souza et al., 2018).

3.1.7. *Topocuvirus*

The monotypic genus *Topocuvirus* is represented by the type species Tomato pseudo-curly top virus (TPCTV), which was reported infecting dicotyledonous species. Analyzes of the TPCTV genome revealed that this virus probably arose from a natural recombination between isolates of two distinct viral genera – *Mastrevirus* and *Begomovirus* (Briddon et al., 2010; King et al., 2011; ICTV, 2020).

3.1.8. *Turncurtovirus*

This genus is represented by isolates of three species: the type species Turnip curly top virus (TCTV), Turnip leaf roll virus – TuLRV and Sesame curly top virus – SeCTV (ICTV, 2020). TCTV and TuLRV isolates have been found in chinese cabbage (*Brassica rapa* L.), beet (*B. vulgaris* L.), lettuce (*Lactuca sativa* L.), basil (*Ocimum basilicum* L.) and radish (*Raphanus sativus* L.). Isolates of two new species of turncurtovirus have recently been reported infecting plants of *Sesamum indicum* L., which were called Sesame curly top virus – SeCTV and Sesame yellow mosaic virus – SeYMV (Hasanvand et al., 2018). All isolates have the same "TAA**T**ATT**A**C" sequence found at the origin of replication of begomovirus, curtovirus, topocuvirus, capulavirus, grablovirus and most mastreviruses.

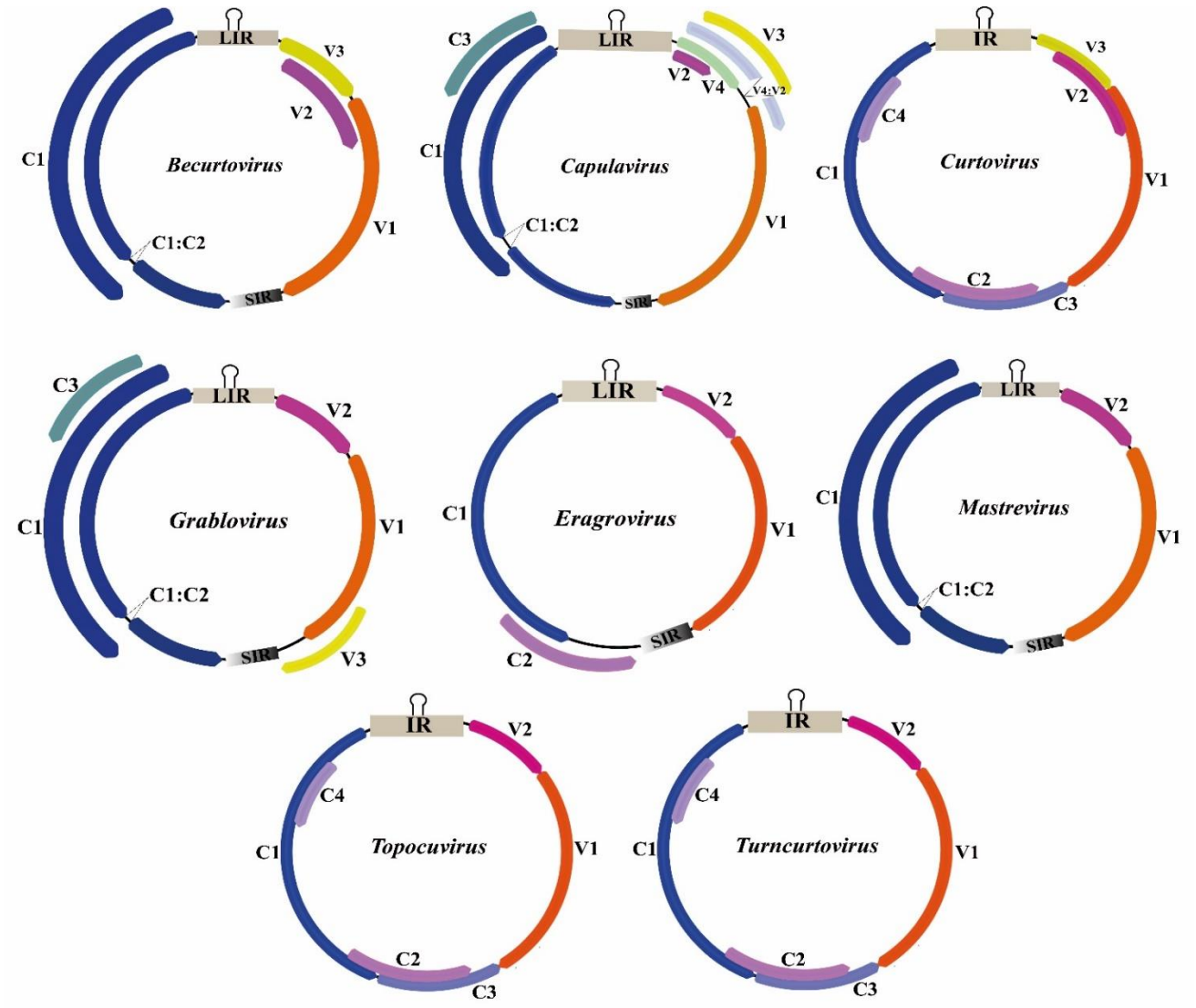

Figure 2. Genomic organization of *Becurtovirus*, *Capulavirus*, *Curtovirus*, *Grablovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus* and *Turncurtovirus* of species-isolates. The ORFs (*Open Reading Frames*) in the viral sense (V1, V2 & V3) and in the complementary sense (C1, C2, C3 & C4) are indicated above. LIR (*Long Intergenic Region*); SIR (*Short intergenic region*); V1 (*Coat Protein* – capsid protein); V2 (*Movement Protein*); V3 (*Regulatory gene*); C1 (*Replication associated protein*); C2 (*Trans-acting protein*); C3 (*Replication enhancer protein*) and C4 (symptom-determining protein).

3.1.9. *Begomovirus*

Currently, the genus *Begomovirus* (type species: *Bean golden yellow mosaic virus* – BGYMV) is the largest within the *Geminiviridae* (ICTV, 2020). These species include viruses that infect exclusively dicotyledoneous and are characterized by having either monopartite or bipartite genomes (Brown et al., 2015; Rojas et al., 2018). There is a

correlation between the type of begomovirus and its geographic distribution. In Australia and in Africa, Asia, and Europe (= Old World), most species have monopartite genomes. In the Americas (= New World), species with bipartite genomes are predominat (Melgarejo et al., 2013). As mentioned, the genomes of begomovirus species can be monopartite or bipartite, encoding from five to eight proteins distributed in one or two molecules of ssDNA (Brown et al., 2015; Varsani et al., 2017; Rojas et al., 2018). Monopartite species grouped the genes necessary for replication, encapsidation and viral movement in just one component, called DNA – A (Brown et al., 2015; Varsani et al., 2017). These species have an ambissense genomic organization, encoding two proteins in the viral sense (V1 and V2) and four in the complementary sense (C1, C2, C3 and C4). In the viral sense (**Figure 3**) it presents the V1 ORF that encodes the coat protein (CP), which is responsible for encapsidating the genome, transmission, and long-distance movement. The V2 ORF acts on the virus movement in the plant and the gene silencing suppression. In the complementary sense **(Figure 3)**, the ORF C1 encodes the protein associated with replication (Rep), C2 encodes the transcription-activating protein (TrAp), C3 encodes the protein that enhances viral replication (REn) and C4 is involved in the expression of symptoms, viral movement and post-transcriptional gene silencing(Gutierrez, 2002; Vanitharani et al., 2004; Gopal et al., 2007; Roy et al., 2019).

Bipartite begomoviruses have genomes comprising two genomic components DNA–A and DNA–B (≈ 2.6 Kb each). The DNA–A contains genes encoding proteins necessary for DNA replication, gene regulation and encapsidation, whereas the DNA–B is composed by genes encoding proteins involved in intracellular and intercellular movement. The two components do not show sequence similarity except for a common region (CR) of ≈ 200 nucleotides. The CR is conserved across the two components belonging to a same species and it is the starting point for the genomic replication process. The CR contains a sequence of nine nucleotides “TAATATTAC” (which is conserved in almost all geminiviruses), the Rep protein cleavage site and the TATA box (Argüello-Astorga and Ruiz-Medrano, 2001; Gutierrez, 2002; Zerbini et al., 2017). The DNA–A component of bipartite begomoviruses has an ambissense genomic organization and can encode four to six proteins (**Figure 3**). The DNA–A has an ORF in the viral direction (AV1) that encodes the coat protein (CP). CP is a multifunctional protein, because in addition to being responsible for capsid formation, it also acts on the accumulation of viral ssDNA (single-stranded DNA), transmission of the virus by the vector and in determining the vector specificity (Boulton, 2002). In addition, CP also plays a crucial

role in transporting viral DNA by interacting with host cell transporters. After viral infection, decapsidation occurs in the cytoplasm of the host plant and the entry of the viral ssDNA into the nucleus, which is subsequently facilitated by CP (Sharma and Ikegami, 2009; Kumar, 2019). In relation to the four ORFs in the complementary sense, ORF AC1 encodes the Rep protein. Rep is the most conserved protein and performs a wide range of functions within the host cell nucleus, such as: specific recognition of the origin of replication, guiding the DNA synthesis, dsDNA binding, helicase activity and interaction with various host proteins. Rep is also involved in the transcription process regulating the expression of certain viral genes (Argüello-Astorga and Ruiz-Medrano, 2001; Liang et al., 2015; Ruhel and Chakraborty, 2019). ORF AC2 encodes TrAp, which is a gene product needed to activate the expression of CP and BV1. The ORF AC3 encodes REn, which is a nuclear protein that interacts with Rep and increases the accumulation of viral DNA (Castillo et al., 2003; Kumar, 2019). AC4 is also present in bipartite begomoviruses, however, it has been demonstrated for some species that this ORF is not essential for viral infectivity, as observed for Tomato golden mosaic virus (ToGMV), Potato yellow mosaic virus (PYMV), Bean golden mosaic virus (BGMV), East African cassava mosaic Zanzibar virus (EACMZV) and Tomato chlorotic mottle virus (ToCMoV) (Sung and Coutts, 1995; Hoogstraten et al., 1996; Pooma and Petty, 1996; Bull et al., 2007; Fontenelle et al., 2007). However, in other bipartite begomoviruses, AC4 is a determinant factor for pathogenicity, being extremely necessary for viral infection as is the case of the East African cassava mosaic Cameroon virus (EACMCV) (Chen et al., 2019). AC4 is also related to the suppression of post-transcriptional gene silencing as demonstrated for the Sri Lanka cassava mosaic virus (SLCMV), ACMV (Gopal et al., 2007) and Tomato leaf curl Palampur virus (ToLCPaIV) (Kulshreshtha et al., 2019). ORF AC5 is present in isolates of some begomoviruses and can have a dual function as a pathogenicity factor and as a suppressor of post-transcriptional gene silencing (Liang et al., 2015).

The DNA–B component of bipartite begomoviruses comprises two ORFs, one in the viral sense (BV1 or NSP – nuclear shuttle protein) that encodes a nuclear transport protein and the other in the complementary sense (BC1 or MP) that encodes an intercellular movement protein. NSP assists viral DNA transport (replicated in the nucleus) to the cytoplasm. NSP interacts with MP (**Figure 3**) to transfer DNA to the MP complex, which is then systemically dispersed through plasmodesmata (Noueiry et al., 1994; Sanderfoot and Lazarowitz, 1996; Nawaz-ul-Rehman and Fauquet, 2009; Zerbini et al., 2017). In some bipartite begomoviruses of the “Old World”, only DNA–A is

sufficient for the systemic viral movement due to the presence of the AV2 protein (**Figure 3**). In the "New World" begomovirus DNA–A is dependent on DNA–B for systemic movement (Rojas et al., 2005a; Brown et al., 2015).

In order to classify all isolates of begomoviruses in a uniform manner and thereby obtain a standard classification, some general taxonomic guidelines have been proposed to define species and strains. In this way, a new monopartite or bipartite **species** must display the identity levels of its complete DNA–A genome less than 91% when compared to the complete genome of any other previously known begomovirus species. In turn, if the sequence of a given virus shares levels of identity greater than 91%, but less than 94% with the complete genome of all isolates described for that species, then it is classified as new **strain** (Brown et al., 2015).

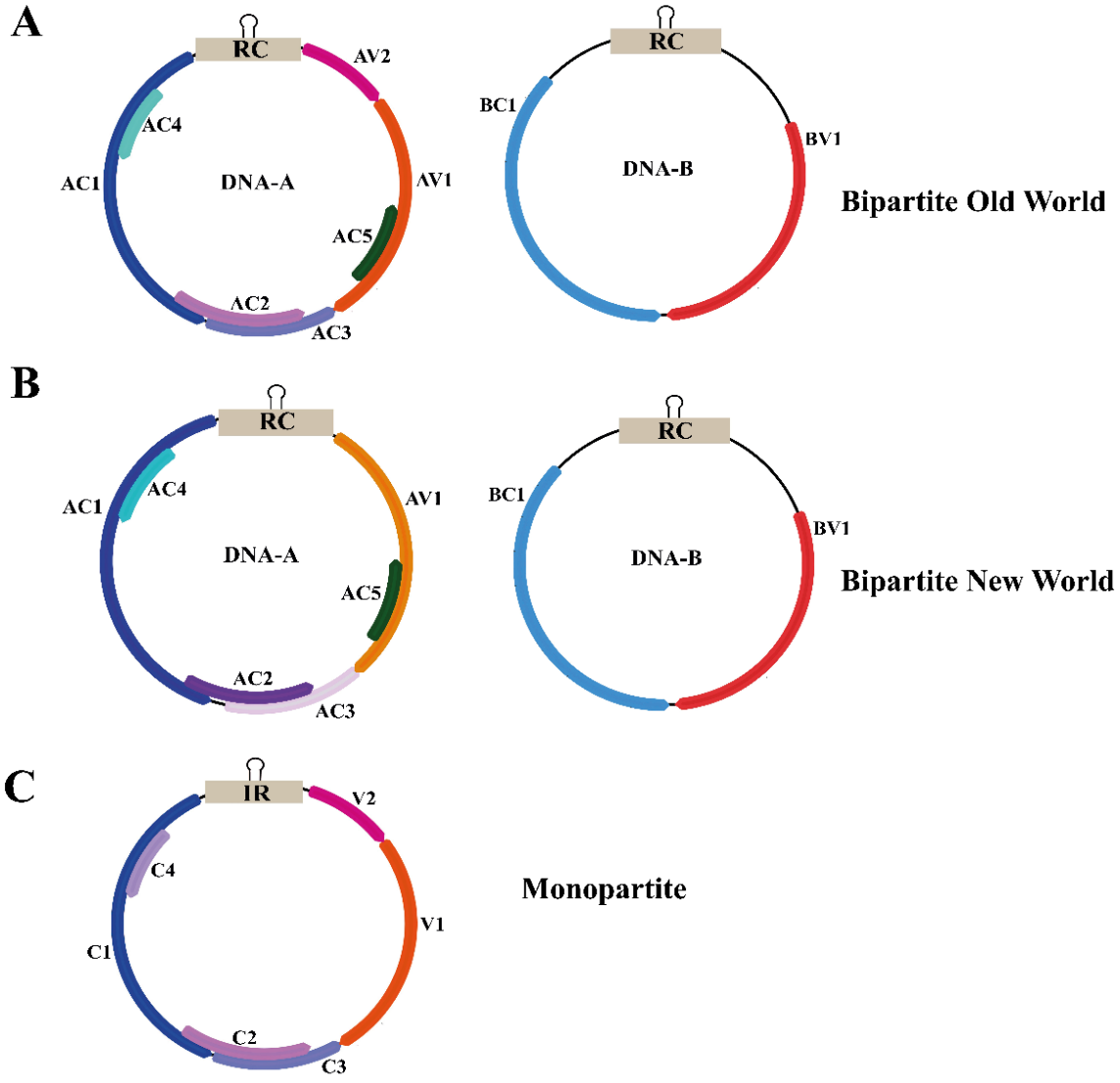

Figure 3. Typical genomic organization of monopartite and bipartite begomovirus. (A) Old World bipartite begomovirus; **(B)** New World bipartite begomovirus and **(C)** Monopartite begomovirus. The circles represent the viral genomes and the arrows indicate the position of the ORFs (*Open Reading Frames*) in the viral (V) and complementary (C) directions. CP (*Coat Protein* = capsid protein); AV2 (*Movement Protein*); Rep (*Replication associated protein*); TrAP (*Transciptional Activator Protein*); REn (*Replication enhancer*); AC4 (*Symptom-determining protein*); BC1 (*Movement protein* involved in cell-to-cell viral movement) and BV1 (*Nuclear shuttle protein*).

4. Transmission of begomovirus species

Under natural conditions, begomoviruses are transmitted by members of a complex of biotypes and cryptic species called the "*B. tabaci* complex". Transmission is done a persistent circulative (Rosen et al., 2015). This complex is considered as the most

destructive group of pests for world agriculture and is widely distributed in tropical and subtropical regions (De Barro et al., 2011). Currently, the *B. tabaci* complex is divided into 11 genetic groups encompassing 35 cryptic species (Dinsdale et al., 2010; De Barro et al., 2011; Boykin and De Barro, 2014). This classification is based upon sequence analysis of the mtCOI gene (mitochondrial cytochrome oxidase I) and its comparison with the consensus sequences described for the different species. Divergence levels above 3.5% has been the major criterion adopted to define a new member of the complex (Dinsdale et al., 2010; De Barro et al., 2011; Boykin and De Barro, 2014).

The species of the *B. tabaci* complex can also be differentiated by a number of biological properties, including host plant range, resistance to insecticides, ability to transmit different viral species and their ability to induce physiological disorders in a given group of host plants. The most prevalent (and harmful) species are those of the genetic group *B. tabaci* MEAM1 (*Middle East-Asia Minor* 1), formerly called *B. tabaci* biotype B and *B. tabaci* MED (*Mediterranean*), formerly known as *B. tabaci* biotype Q (Rosen et al., 2015). In Brazil, *B. tabaci* MEAM1 predominates, which is highly polyphagous, infecting more than 1,000 plant species and transmitting more than 300 virus species. The *B. tabaci* complex also involved in the transmission of *Crinivirus*, *Carlavirus*, *Torradovirus* and *Ipomovirus* species with different modes of transmission. Species related to these genera can cause diseases in ornamental and cultivated plants such as vegetables, cassava and cotton, resulting in decreased production (Barbosa et al., 2014; Gilbertson et al., 2015). Recently, the invasion of *B. tabaci* MED was detected in Brazil in the municipality of Barra do Quaraí in Rio Grande do Sul, in plants of *C. annuum* L. in a greenhouse and in *Ipomoea batatas* L. under field conditions (Barbosa et al., 2015).

The transmission of many plant viruses by *B. tabaci* and the ability of these insects to develop resistance to many pesticides make them one of the most devastating groups of pests for agriculture (Skaljac et al., 2017). Regarding the virus-vector interaction, studies with bipartite and monopartite begomovirus species indicate that the acquisition period may range from 10 to 60 minutes, inoculation period from 10 to 30 minutes, latency period from 17 to 20 hours and the whitefly remains viruliferous around 7 to 20 days (Santos et al., 2003; Rosen et al., 2015).

The genetic mechanisms involved in the circulation and transmission of whitefly viruses have not yet been fully elucidated. Gene expression studies of *B. tabaci* associated with virus transmission were conducted in order to obtain a broader understanding of how the vectors respond to food in agricultural crops infected by viruses, presenting different

modes of interaction with the vector. Assays were conducted to study the gene expression of *B. tabaci* colonizing tomato plants infected by Tomato yellow leaf curl virus (TYLCV – *Begomovirus* genus) and by Tomato chlorosis virus (ToCV – genus *Crinivirus*). The results showed that ≈100 genes were differentially expressed in whiteflies that fed on TYLCV-infected tomatoes, while more than 1,000 genes were identified in whiteflies that fed on ToCV-infected tomatoes. However, the results in three sampling times were very similar between the two viruses, with a greater number of genes between 24 and 72 hours and a smaller number of genes differentially expressed 48 hours after acquisition. In addition, a subgroup of genes was identified in the two treatments, suggesting that two viruses with different modes of transmission may have similar effects on *B. tabaci* (Kaur et al., 2017; Hasegawa et al., 2018).

The ability of begomoviruses to replicate in tissues of the whiteflies is still controversial. Initial studies with the TYLCV species concluded that the pathogen is able to replicate in the insect when it is under stress conditions. However, under normal conditions, the whitefly is able to prevent viral accumulation using its immune system (Pakkianathan et al., 2015). On the other hand, a different study demonstrated (via quantitative PCR) that the TYLCV DNA concentrations did not increase in the insects until 96 hours after the acquisition (Sánchez et al., 2016). Thus, these results are not in agreement with previous observations that TYLCV can replicate in *B. tabaci* (Sánchez et al., 2016). A later study indicated that most begomoviruses remain associated with the intestine of the insect and in its filter chamber. In these tissues, TYLCV genome could be transcribed and replicated. However, due to the activation of an immune-like response of the insect, there is inhibition of replication and a subsequent elimination of the virus (Czosnek et al., 2017).

The potential transovarian transmission of begomoviruses by their vectors is of great epidemiological relevance, since the vector could become a source of viral inoculum even in the absence of host plants and could, therefore, facilitate viral spread over long distances (Accotto and Sardo, 2009). Some studies were conducted with begomoviruses in recent years aiming to elucidate their potential *B. tabaci* transovarian transmission. A pioneer study was carried out using *B. tabaci* MEAM1 and *B. tabaci* MED and two begomoviruses TYLCV and Tomato yellow leaf curl Sardinia virus (TYLCSV). The results indicated that TYLCSV can be detected in eggs and nymphs, and to a lesser extent in adults of the first-generation progeny. On the other hand, TYLCV was not detected in any of the three stages (Bosco et al., 2004). More recent studies have confirmed the

transovarian transmission of TYLCV by seven species of the *B. tabaci* complex native to China. TYLCV was transmitted via *B. tabaci* eggs and nymphs, but this virus was not detected in adults, indicating that TYLCV is lost in the first-generation offspring and seems that it does not reach the adult stage. In addition, due to differences observed in the transovarian transmission efficiency of TYLCV, it can be concluded that this transmission capacity may vary according to the vector species (Guo et al., 2019).

5. Satellite DNAs associated with begomovirus species

The "Old World" monopartite begomoviruses are often accompanied by satellite DNA (alphasatellites and betasatellites). Until recently, it was considered that only monopartite viruses had this association with satellites. However, there is already a report of a new class of satellite DNA associated with bipartite begomoviruses from the "New World" (Rojas et al., 2005a; Brown et al., 2012). Currently, three types of satellite DNA associated with begomoviruses are described: alphasatellites, betasatellites and deltasatellites (Kumar et al., 2015; Kumar et al., 2017; ICTV, 2020).

Alphasatellites, also called DNA-1, belong to the family Alphasatellitidae, established in 2017 with two subfamilies, 11 genera and 71 species. The genera of alphasatellites associated with the geminivirus are found in the subfamily *Geminialphasatellitinae*, genera *Ageyesisatellite*, *Clecrusatellite*, *Colecusatellite* and *Gosmusatellite*. The demarcation of species within the subfamily *Geminialphasatellitinae*, indicates levels of identity less than 88% in comparison with complete sequences already known (Briddon et al., 2018; ICTV, 2020). Alphasatellite species are characterized by having a genome of ≈1,350 nts and a genomic organization composed of an ORF that encodes a Rep protein (similar to that of nanoviruses), a region rich in adenine and a structure in the form of "hairpin" containing the TAGTATTAC string (**Figure 4**). Unlike betasatellites, these molecules have the ability to self-replicate in the host plant, however they need their auxiliary virus for movement and transmission by the insect vector (Briddon et al., 2004; Zhou Xueping, 2013; Briddon et al., 2018). At least three types of alphasatellites (I, II, and III) are known (Rosario et al., 2013). Recently, type I alphasatellites have been shown to cause symptom attenuation when co-infecting plants with their helper viruses and its associated betasatellite, suggesting that they negatively regulate virulence to some degree, possibly reducing the accumulation of betasatellite molecules (Idris et al., 2011). Alphasatellites of types II and III were already

found in association with bipartite begomoviruses in the New World (Brazil, Cuba and Venezuela) (Paprotka et al., 2010; Romay et al., 2010).

Betasatellites (formerly known as DNA-β), belong to the *Tolecusatellitidae* family (established in 2016), which is composed by two genera (*Betasatellite* and *Deltasatellite*). The demarcation of species within the genera is based on nucleotide identity less than 91% when compared to other species already described (ICTV, 2020). Betasatellites are characterized by a small genome (≈1,360 nts) and they not share nucleotide identity with their auxiliary viruses. However, betasatellites depend on auxiliary viruses for their replication, encapsidation, cell-to-cell movement and long-distance movement (via phloem) as well as for transmission by the vector (Briddon et al., 2003; Zhou Xueping, 2013; Gnanasekaran et al., 2019). The genomic organization of these molecules displays highly conserved regions, composed of a region rich in adenine ("A-rich region"), a region called SCR (= satellite conserved region) that contains a structure in the form of "hairpin" encompassing the sequence "TAAGTATTAC" and a single ORF that encodes the βC1 protein (**Figure 4**). The βC1 gene product plays an important role in inducing symptoms in some hosts and in suppressing transcriptional and post-transcriptional gene silencing. Betasatellites can also affect the replication of their helper viruses (Briddon et al., 2004; Nawaz-ul-Rehman and Fauquet, 2009; Zhou Xueping, 2013; Kumar et al., 2014; Rosario et al., 2016; Gnanasekaran et al., 2019).

The **deltasatellites** are characterized by having a genome of ≈ 0.7 kb (**Figure 4**). They were initially found in association with monopartite begomoviruses from the Old World (Dry et al., 1997). Afterward, they were also found in association with bipartite begomoviruses from the New World (Fiallo-Olivé et al., 2012) and also with sweepovirus (Lozano et al., 2016). Unlike betasatellites and alphasatellites, deltasatellites do not encode any proteins. They are entirely dependent on the auxiliary begomovirus for replication and movement in plants as well as for transmission by *B. tabaci*. The presence of deltasatellites in some host-virus combinations results in reduced begomovirus accumulation and/or attenuated symptom expression (Fiallo-Olivé Elvira et al., 2016; Hassan et al., 2016).

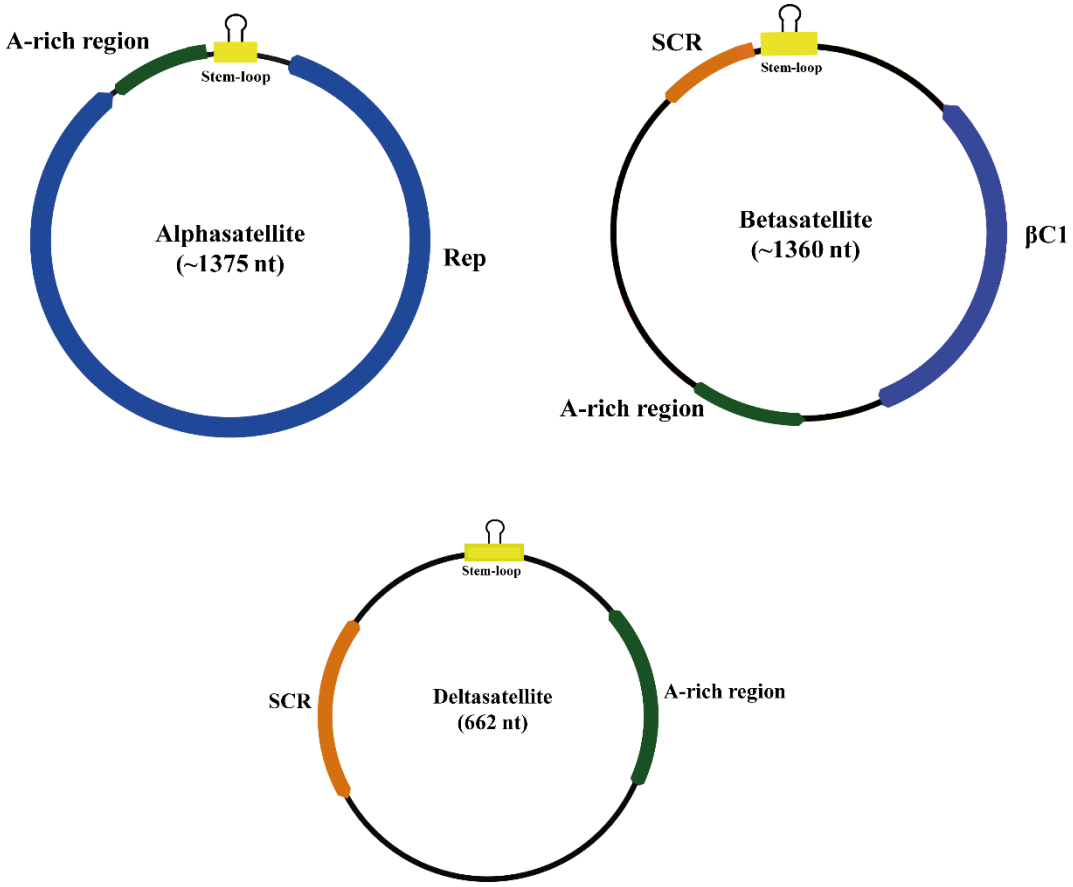

Figure 4. Genomic representation of satellite DNAs (Alfasatellites, Betasatellites, and Deltasatellites) that are found associated with isolates of *Begomovirus* species. Illustration of the main genomic characteristics: ORFs (*Open Reading Frames*) Rep (Replication-associated protein) in alphasatellites and βC1 and the Adenine (A)–rich region, which is present in all DNA satellites; SCR (= satellite conserved region) and stem-loop (conserved).

6. Replication of begomovirus in host cells

The replication of begomoviruses, as well as of all members of the *Geminiviridae* family, occurs in the nucleus of the host cell via a mechanism known as rolling circle replication. Initially, the viral particles are inoculated into the plant by the vector and upon entering the cell, the presence of the viral particles activates the host enzyme system that releases proteases, which degrade the protein layer exposing the viral genome, which is later transported to the nucleus (Stanley, 1995; Hanley-Bowdoin et al., 1999). Within the nucleus, replication can be divided into three functionally distinct stages. **In the first stage of replication**, single-stranded DNA (ssDNA) is converted to double-stranded DNA (dsDNA), which is known as the replicative form (RF). This intermediate form will

serve as a template for viral transcription and for the synthesis of new ssDNA strands by the rolling circle mechanism (Stanley, 1995; Monsalve-Fonnegra et al., 2002). This process is initiated by the connection of the Rep to a specific sequence in the CR composed by two repeated sequences called "iterons". Upon the binding of Rep, it cleaves the ssDNA chain initiating the viral replication cycle (Gutierrez, 2002; Monsalve-Fonnegra et al., 2002; Yadava et al., 2010; Pradhan et al., 2017). **The second stage of replication** is carried out by the rolling circle mechanism, which consists of using dsDNA as a template for the amplification of ssDNA. Rep is the protein responsible for the initiation of the reaction that involves a cleavage within the nonanucleotide "TAATATTAC" (which is conserved in most geminiviruses) located in the "stem loop" present in the intergenic region. **The third stage of replication** consists of the synthesis of ssDNA from dsDNA that takes place towards the end of the replication cycle, with the accumulation of viral genomes for encapsidation (Gutierrez, 2002; Monsalve- Fonnegra et al., 2002; Pradhan et al., 2017).

7. Genetic variability in begomovirus

Genetic variability in viral populations provides new opportunities for adaptation to new hosts as well as in relation to changes in environmental conditions. This genetic variability is generated by mechanisms such as mutation, recombination, and pseudo-recombination. These mechanisms occur frequently in *Begomovirus*, resulting in a high genetic variability that can lead to the emergence of new species and strains (Roossinck, 1997; Gutierrez et al., 2004; Seal et al., 2006).

7.1. Mutation – This type of genetic variation occurs due to the incorrect incorporation of nucleotides during viral replication. This low fidelity genetic mechanism can generate variability before the exchange of genomic fragments associated with recombination events (Roossinck, 1997; Seal et al., 2006; Duffy and Holmes, 2008). Mutations can be spontaneous or induced. Spontaneous mutations result from errors in base pairing during DNA replication. Mutations that alter only one base pair are called point mutations. These mutations are the most common and they can be caused by DNA base pair substitutions as well as by the loss or gain of a single base pair. When the mutation does not affect the sequence of the encoded polypeptide it is called a silent mutation. On the other hand, nonsense mutation is a type of mutation that occurs when a codon that corresponds to an amino acid becomes a stop codon, leading to the synthesis of an incomplete polypeptide.

Deletions and insertions cause major changes in DNA, due to the fact that they often affect the open reading frame, which normally results in loss of gene function (Simon-Loriere and Holmes, 2011; Madigan et al., 2016).

7.2. Recombination – Begomoviruses replicate their genomes using the host DNA polymerase, which has fidelity rates similar to those enzymes of RNA viruses (Duffy and Holmes, 2008). Mutational events are the main factors in the diversification of viral populations (García-Arenal et al., 2003). However, recombination can also play a significant role in generating virus diversity (Lima et al., 2013; Silva et al., 2014). The simultaneous presence of several begomoviruses under field conditions, all transmitted by the same vector, increases the frequency of mixed infections, where two or more viral species are present at the same time in a single plant. This situation increases the probability of occurrence of recombination and/or pseudo-recombination events among distinct viral genomic components, which can potentially accelerate the generation of recombinant genotypes with potentially novel biological features (García-Arenal et al., 2003; Seal et al., 2006; Silva et al., 2014). Recombination is the exchanging genomic segments between two strands of DNA or RNA during replication (Padidam et al., 1999; Lefeuvre and Moriones, 2015). Interspecific homologous recombination is considered an active source of begomovirus genetic diversity and contributes to viral evolution, including the emergence of new species (Lefeuvre et al., 2009; Kumar et al., 2010; Sahu et al., 2018). In fact, recombination events have been directly involved in the emergence of new diseases and epidemics in many cultivated plants. The first evidence of recombination among geminiviruses was obtained from studies of the "cassava mosaic" disease in Uganda. Sequence analyzes revealed that the causal agent of the "cassava mosaic", the East African cassava mosaic virus-Uganda (EACMV-UG) was most likely originated from interspecific recombination between East African cassava mosaic virus (EACMV) and ACMV (Zhou Xueping et al., 1997). Epidemics involving different members of the Tomato yellow leaf curl virus (TYLCV) complex in Spain, allowed the emergence of some recombinant pathogens such as Tomato yellow leaf curl Malaga virus (TYLCMaIV) and Tomato yellow leaf curl Auxarquia virus (TYLCAxV) (García-Andrés et al., 2007). In certain cases, recombinants can be more aggressive than their original parents. Still in Spain, it was observed that a recombinant virus between Tomato yellow leaf curl Sardinia virus (TYLCSV) and TYLCV had a broader host range, becoming the predominant begomovirus in that country (Monci et al., 2002; Mnari-Hattab et al., 2014).

A recent study showed that the begomoviruses (associated with the “cotton leaf roll”) disease are in a continuous process of evolution. Isolates originated from recombination events involving Cotton leaf curl Multan virus (CLCuMuV), Cotton leaf curl Kohran virus (CLCuKoV) and Cotton leaf curl Rajasthan virus (CLCuRaV) were detected in Asia. The results demonstrated interspecific and intraspecific recombination events, leading to significant structural changes in the DNA components of CLCuMuV isolates and the emergence of new variants. This genetic diversification can lead to the adaptation of the isolates to different hosts, thus increasing the threat to other crops. Differences may also be observed in viral transmission capacity as well as in the ability to overcome of resistance factors, being, therefore, a great challenge for the management of these cotton pathogens (Qadir et al., 2019). Recombination events between ToYMoV and Jacquemontia yellow vein virus (JacYVV) caused the emergence of a new monopartite begomovirus, which was able to induce major yield losses in tomato in Venezuela. This new species was tentatively named asTomato twisted leaf virus (Romay et al., 2019).

7.3. Pseudo-recombination – The presence of two genomic components in most of the “New World” begomoviruses promotes another mechanism, known as pseudo-recombination. This mechanism involves the exchange of genomic components between two different viruses (Andrade et al., 2006; Seal et al., 2006). The production of pseudo-recombinants requires a highly specific interaction of the Rep protein with the region around the origin of replication. For most begomoviruses, specific Rep binding sites include one inverted snf two direct repeats known as ‘iterons’. Species with DNA–A and DNA–B displaying identical iterons can eventually become pseudorecombinants. Other factors can also promote the formation of pseudorecombinants, the Rep protein has specificity determinants (SPDs) located in its motif I, which make iteron recognition specific. A mutation or deletion of the SPDs eliminates the ability of the Rep protein to specifically bind to the sequence (Orozco et al., 1997; Lima et al., 2013).

8. *Begomovirus* diversity in tomato in Brazil and the world

According to the criteria stablished by ICTV, 113 begomovirus species (**Table 2**) are currently reported as having tomatoes as their primary host (ICTV, 2020).

Table 2. Species classified in the genus *Begomovirus* already reported naturally infecting tomatoes (ICTV, 2020; Virus-HostDB, 2020)

Species of *Begomovirus* and acronyms	Local of initial detection	References
Ageratum yellow vein Hualian virus (AYVHuV)	Taiwan	(Tsai et al., 2011)
Ageratum yellow vein virus (AYVV)	Vietnan	Choi et al., 2019)
Chili leaf curl virus (ChiLCV)	India	(Venkataravanappa et al., 2016)
Chino del tomate virus (CdTV)	Mexico & USA	(Brown et al., 2000)
Chino del tomate Amazonas virus (CdTAV)	Brazil	(Fonseca et al., 2011)
Pepper golden mosaic virus (PepGMV)	United States	(Holguín-Peña et al., 2004)
Pepper huasteco yellow vein virus (PHYVV)	Mexico	(Moreno-Félix et al., 2016)
Tomato bright yellow mosaic virus (ToBYMV)	Brazil	(Fonseca et al., 2013)
Tomato bright yellow mottle virus (ToBYMoV)	Brazil	(Fonseca et al., 2013)
Tomato chino La Paz virus (ToChLPV)	Mexico	(Holguín-Peña et al., 2006)
Tomato chlorotic leaf curl virus (ToCLCV)	Brazil	(Quadros et al., 2019)
Tomato chlorotic leaf distortion virus (ToCILDV)	Venezuela	(Zambrano et al., 2011)
Tomato chlorotic mottle Guyane virus (ToCMoGFV)	French Guiana	(Lett et al., 2015)
Tomato chlorotic mottle virus (ToCMoV)	Brazil	(Ribeiro et al., 2003; Ribeiro et al., 2007)
Tomato common mosaic virus (ToCmMV)	Brazil	(Castillo-Urquiza et al., 2008)
Tobacco curly shoot virus (TbCSV)	China	(Li et al., 2004)
Tomato curly stunt virus (ToCSV)	South Africa	(Pietersen et al., 2000)
Tomato dwarf leaf virus (ToDfLV)	Argentina	(Medina and Lambertini, 2012)
Tomato enation leaf curl virus (ToELCV)	India	(Swarnalatha et al., 2014)
Tomato golden leaf distortion virus (ToGLDV)	Brazil	(Fonseca et al., 2010)
Tomato golden leaf spot virus (ToGLSV)	Brazil	(Fonseca and Boiteux, 2013)
Tomato golden mosaic virus (TGMV)	Brazil	(Matyis et al., 1975; Hamilton et al., 1984)
Tomato golden mottle virus (ToGMoV)	Mexico	(Mauricio-Castillo et al., 2007)
Tomato golden vein virus (ToGVV)	Brazil	(Fernandes et al., 2008)

Tomato interveinal chlorosis virus (ToICV)	Brazil	(Albuquerque et al., 2012)
Tomato latent virus (TLV)	Cuba	(Fuentes et al., 2016)
Tobacco leaf curl Thailand virus (TbLCTHV)	Thailand	(Knierim and Maiss, 2007)
Tomato leaf curl Anjouan virus (ToLCAnV)	Comoros	(Lefeuvre et al., 2007a)
Tomato leaf curl Arusha virus (ToLCArV)	Tanzania	(Shih et al., 2006a)
Tomato leaf curl Bangalore virus (ToLCBaV)	India	(Tiwari et al., 2010)
Tomato leaf curl Bangladesh virus (ToLCBV)	Bangladesh	(Shih et al., 1998)
Tomato leaf curl Burkina Faso virus (ToLCBFV)	Burkina Faso	(Ouattara et al., 2017)
Tomato leaf curl Cebu virus (ToLCCeV)	Philippines	(Tsai et al., 2011)
Tomato leaf curl China virus (ToLCCNV)	China	(Yang et al., 2011)
Tomato leaf curl Comoros virus (ToLCV)	Comoros	(Delatte et al., 2005)
Tomato leaf curl Diana virus (ToLCDiV)	Madagascar	(Lefeuvre et al., 2007b)
Tomato leaf curl Ghana virus (ToLCGV)	Ghana	(Osei et al., 2008)
Tomato leaf curl Guangdong virus (ToLCGdV)	China	(He et al., 2005)
Tomato leaf curl Guangxi virus (ToLCGxV)	China	(Xu Youping et al., 2007)
Tomato leaf curl Gujarat virus (ToLCGV)	India	(Chakraborty et al., 2003)
Tomato leaf curl Hainan virus (ToLHaiV)	China	(Zhang Hui et al., 2010)
Tomato leaf curl Hanoi virus (ToLCHaV)	Vietnam	(Cuong et al., 2011)
Tomato leaf curl Hsinchu virus (ToLCHsV)	China	(Tsai et al., 2006a)
Tomato leaf curl Iran virus (ToLCIV)	Iran	(Bahjatnia et al., 2004)
Tomato leaf curl Japan virus (ToLCJV)	Japan	(Ueda et al., 2005)
Tomato leaf curl Java virus (ToLCJaV)	Indonesia	(Kon Tatsuya et al., 2006)
Tomato leaf curl Joydebpur virus (ToLCJV)	India	(Tiwari et al., 2013)
Tomato leaf curl Karnataka virus (ToLCKV)	India	(Chatchawankanphanich and Maxwell, 2002)
Tomato leaf curl Karnataka virus 2 (ToLCKV2)	India	(Swarnalatha et al., 2019)
Tomato leaf curl Karnataka virus 3 (ToLCKV3)	India	(Swarnalatha et al., 2019)
Tomato leaf curl Kerala virus (ToLCKeV)	India	(Pasumarthy et al., 2010)

Tomato leaf curl Laos virus (ToLCLV)	Laos	(Tsai et al., 1999)
Tomato leaf curl Liwa virus (ToLCLwV)	Oman	(Khan et al., 2014)
Tomato leaf curl Madagascar virus (ToLCMGV)	Madagascar	(Delatte et al., 2005)
Tomato leaf curl Mahe virus (ToLCMahV)	Seychelles	(Scussel et al., 2018)
Tomato leaf curl Malaysia virus (ToLCMYV)	Malaysia	(Shih et al., 1998)
Tomato leaf curl Mali virus (ToLCMLV)	Mali	(Zhou et al., 2008)
Tomato leaf curl Mindanao virus (ToLCMiV)	Philippines	(Tsai et al., 2011)
Tomato leaf curl Moheli virus (ToLCMohV)	Comoros	(Lefeuvre et al., 2007b)
Tomato leaf curl Namakely virus (ToLCNamV)	Madagascar	(Lefeuvre et al., 2007b)
Tomato leaf curl New Delhi virus (ToLCNDV)	Bangladesh	(Varma and Malathi, 2003)
Tomato leaf curl New Delhi virus 2 (ToLCNDV2)	India	(Chaudhary et al., 2012)
Tomato leaf curl New Delhi virus 4 (ToLCNDV4)	India	(Swarnalatha P et al., 2013)
Tomato leaf curl Nigeria virus (TLCNV)	Nigeria	(Kon et al., 2009)
Tomato leaf curl Palampur virus (ToLPaIV)	India	(Heydarnejad et al., 2009)
Tomato leaf curl Patna virus (ToLCPaIV)	India	(Kumari et al., 2009)
Tomato leaf curl Philippines virus (ToLCPV)	Philippines	(Dolores and Bajet, 1995)
Tomato leaf curl Pune virus (ToLCPuV)	India	(Chowda et al., 2004)
Tomato leaf curl purple vein virus (ToLCPV)	Brazil	(Macedo et al., 2018)
Tomato leaf curl Rajasthan virus (ToLCRaV)	India	(Sivalingam et al., 2005)
Tomato leaf curl Seychelles virus (ToLCSV)	Seychelles	(Lefeuvre et al., 2007a)
Tomato leaf curl Sinaloa virus (ToLCSiV)	Nicaragua	(Rojas et al., 2005)
Tomato leaf curl Sri Lanka virus (ToLCLKV)	Sri Lanka	(Samarakoon et al., 2012)
Tomato leaf curl Sudan virus (ToLCSDV)	Sudan	(Idris et al., 2005)
Tomato leaf curl Sulawesi virus (ToLCSuV)	Indonesia	(Tsai et al., 2009)
Tomato leaf curl Taiwan virus (ToLCTWV)	China	(Zi-Fu et al., 2007)
Tomato leaf curl Tanzania virus (ToLCTV	Tanzania	(Chiang et al., 1997)
Tomato leaf curl Toliara virus (ToLCToV)	Madagascar	(Lefeuvre et al., 2007b)

Tomato leaf curl Uganda virus (ToLCUV)	Uganda	(Shih et al., 2006)
Tomato leaf curl Vietnam virus (ToLCVV)	Vietnam	(Ha et al., 2008)
Tomato leaf curl virus (ToLCV)	India	(Kumar et al., 2012)
Tomato leaf deformation virus (ToLDeV)	Peru	(Márquez-Martín et al., 2011)
Tomato leaf distortion virus (ToLDV)	Brazil	(Castillo-Urquiza et al., 2008)
Tomato mild mosaic virus (ToMMV)	Brazil	(Castillo-Urquiza et al., 2008)
Tomato mild yellow leaf curl Aragua virus (ToMYLCV)	Venezuela	(Romay et al., 2017)
Tomato mosaic Havana virus (ToMHaV)	Cuba & Nicaragua	(Zubiaur et al., 1998; Monger et al., 2008)
Tomato mottle leaf curl virus (ToMoLCV)	Brazil	(Albuquerque et al., 2012)
Tomato mottle Taino virus (ToMoTaV)	Cuba	(Ramos et al., 1997)
Tomato mottle virus (ToMoV)	United States	(Abouzid et al., 1992)
Tomato mottle wrinkle virus (ToMoWV)	Argentina	(Medina et al., 2015)
Tomato rugose mosaic virus (ToRMV)	Brazil	(Ribeiro et al., 2003; Fernandes et al., 2006)
Tomato rugose yellow leaf curl virus (TRYLCV)	South America	(Márquez-Martín et al., 2012; Guerrero et al., 2013; Fonseca, 2016)
Tomato severe leaf curl Kalakada virus (TSLCKV)	India	(Swarnalatha et al., 2014)
Tomato severe leaf curl virus (ToSLCV)	Mexico	(Mauricio-Castillo et al., 2006)
Tomato severe rugose virus (ToSRV)	Brazil	(Cotrim et al., 2007; Fernandes et al., 2008)
Tomato yellow leaf curl Axarquia virus (TYLCAxV)	Spain	(Anfoka et al., 2016)
Tomato yellow leaf curl China virus (TYLCCNV)	China	(Yin et al., 2001)
Tomato yellow leaf curl Guangdong virus (TYLCGV)	China	(He et al., 2005)
Tomato yellow leaf curl Indonesia virus (TYLCIDV)	Indonesia	(Tsai et al., 2006b)
Tomato yellow leaf curl Kanchanaburi virus (TYCKaV)	China	(Bagewadi and Naidu, 2016)
Tomato yellow leaf curl Malaga virus (TYLCMaV)	Spain	(Monci et al., 2002)
Tomato yellow leaf curl Mali virus (TYLCMLV)	Mali	(Sattar et al., 2013)
Tomato yellow leaf curl Sardinia virus (TYLCSV)	Italy & Spain	(Kheyr-Pour et al., 1991; Monci et al., 2002)
Tomato yellow leaf curl Shuangbai virus (TYLCShV)	China	(Zhao et al., 2015)
Tomato yellow leaf curl Thailand virus (TYLCTHV)	China	(Attathom et al., 1994)

Tomato yellow leaf curl Vietnam virus (TYLCVV)	Vietnam	(Ha et al., 2008)
Tomato yellow leaf curl virus (TYLCV)	Japan, Israel & Iran	(Bananej et al., 2004)
Tomato yellow leaf curl Yunnan virus (TYLCYnV)	China	(Ding et al., 2016)
Tomato yellow leaf distortion virus (ToYLDV)	Cuba	(Fiallo-Olivé et al., 2009)
Tomato yellow margin leaf curl virus (ToYMLCV)	Venezuela	(Nava et al., 2006)
Tomato yellow mottle virus (ToYMoV)	Costa Rica	(Maliano et al., 2012)
Tomato yellow spot virus (ToYSV)	Brazil	(Calegario et al., 2007)
Tomato yellow vein streak virus (ToYVSV)	Brazil	(Albuquerque et al., 2010)

In Brazil, common bean (Faria et al., 2016), tomato (Ribeiro et al., 2003; Andrade et al., 2006; Cotrim et al., 2007; Fernandes et al., 2008) and cowpea (Naito et al., 2019) are the most severely affected crops by diseases caused by begomoviruses. However, there are reports of occurrence of begomoviruses in other crops such as okra, potatoes, sweet potatoes, *Capsicum* species, and soybean (Zerbini et al., 2005; Inoue-Nagata et al., 2016a). The main symptoms observed in tomato production fields are: internerval yellowing, epinastia and dwarfism **(Figure 5).** Yield reduction of up to 60% can occur, mainly caused by significant lower number of fruits per plant (Giordano et al., 2005b; Lemos et al., 2010).

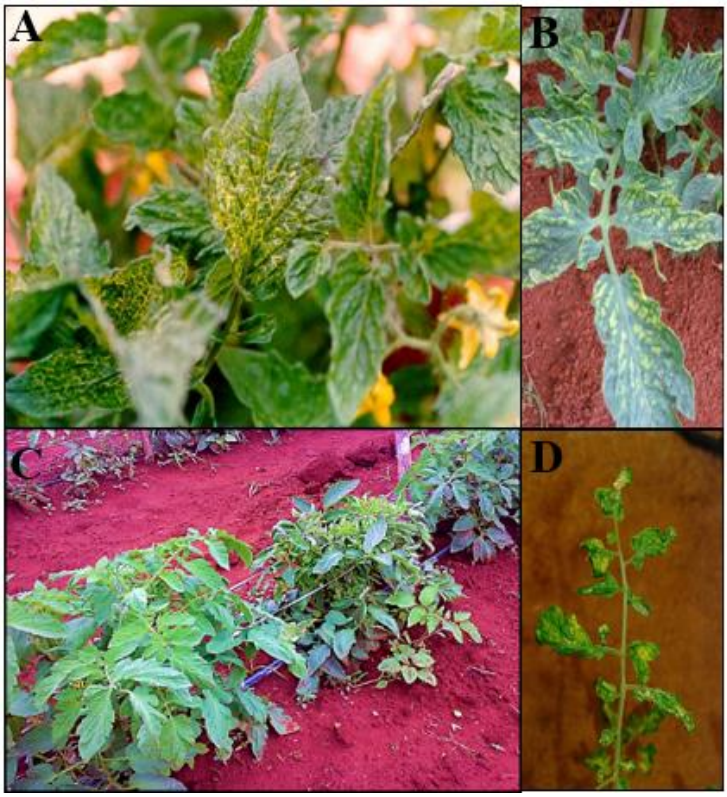

Figure 5. Typical symptoms of *Begomovirus* in tomato (*Solanum lycopersicum* L.). Internerval yellowing in A and B; C) Dwarfism; D) Leaf epinasty.

The first report of a begomovirus in Brazil was made in 1950 in *Euphorbia prunifolia* Jacq. (Costa and Bennett, 1950). The first report on tomato was done in the late 1950s, when plants with symptoms of golden mosaic and chlorosis were observed (Flores et al., 1960). Later, this virus was characterized and named Tomato golden mosaic virus (TGMV). Five other whitefly-transmitted viral species (in addition to TGMV) were identified, but without causing significant damage (Matyis et al., 1975). TGMV was one of the first begomoviruses to have the genome cloned and fully sequenced and became a

model for studies to elucidate molecular interactions of virus and host (Hamilton et al., 1984; Hanley-Bowdoin et al., 1999). TGMV was never considered an economically important viral species, probably due to the low transmission efficiency by the vector *B. tabaci* biotype A, which was the only species of whitefly present in the country at that time (Bedford et al., 1994; Ribeiro et al., 2003).

Since the early 1990s, after the introduction of *B. tabaci* biotype B in Brazil, a significant increase in reports of begomoviruses in tomato was observed (Ribeiro et al., 1994). Biotype B is extremely polyphagous and this vector was probably able to transmit endemic begomovirus species from the wild/native hosts to tomatoes (Ribeiro et al., 1994; Ribeiro et al., 1998; Ribeiro et al., 2003). Surveys conducted over the years (Ribeiro et al., 2003; Andrade et al., 2006; Cotrim et al., 2007; Fernandes et al., 2008) indicated a great diversity of *Begomovirus* species infecting tomatoes. Interestingly, some of these species have a restricted geographical distribution, while some were found to have a more widely distribution across the country **(Table 3).** Tomato severe rugose virus (ToSRV) is an example of a viral species with wide geographic distribution. ToSRV was reported by the first time in tomato in 1999 in Minas Gerais (MG). Afterwards, ToSRV was reported in almost all regions of Brazil (Rezende et al., 1997; Cotrim et al., 2008; Fernandes *et al.*, 2008). In contrast, Tomato yellow spot virus (ToYSV) was reported so far only in GO and MG and Tomato interveinal chlorosis virus (ToICV) in Pernambuco (PE) (Calegario et al., 2007).

The first nationwide survey of begomovirus species was conducted by Ribeiro et al. (2003), in which 23 isolates collected between 1994 and 1999 in fields located at the Midwest, Southeast and Northeast regions. At least seven potential new species were reported in this study, four of which were restricted to the Southeast and two identified exclusively in the Northeast. Subsequently, Cotrim et al. (2007) studied the diversity of begomoviruses in tomato-producing regions in São Paulo state. Tomato samples (n = 166) were collected between 2003 and 2004 and the presence of begomovirus was observed in ≈ 60% of the samples. Direct sequencing of the PCR products from 16 of these samples indicated the predominant presence of ToSRV as well as isolates of Sida mottle virus (SiMoV), Tomato yellow vein streak virus (ToYVSV) and one possible new viral species. Another important begomovirus diversity study was carried out with isolates collected between the years 2004 and 2005 (n = 138) in the states of Pernambuco, Bahia, Distrito Federal, Goiás and Minas Gerais. ToSRV was predominant virus (61% of samples) with report of TGVV,Tomato mottle leaf curl virus (ToMoLCV),ToYVSV isolates as well as

two potential new species (Fernandes et al., 2008). In another study, Castillo-Urquiza et al. (2008) analyzed 115 cloned viral genomes from tomato and weed samples collected in Rio de Janeiro in 2005 and Minas Gerais in 2007, which indicated the predominance of ToYVSV and ToCmMV species.

The large number of viral species infecting tomatoes in Brazilian conditions can be explained by the extreme susceptibility of this host to begomoviruses. Another explanation is the intensification of tomato crops in large areas of monoculture or contiguous areas (almost year-round planting in all producing regions) can also generate favorable conditions for the efficient propagation and survival of the virus and their vectors, thereby increasing the potential for emerging new viruses (Hanssen et al., 2010).

Currently, 21 begomoviruses have been reported naturally infecting tomatoes in Brazil (**Table 3**). In addition, a set of potential new species have been reported but are still in process of characterization (therefore, not yet accepted by ICTV), including: Tomato crinckle leaf yellow virus (ToCrLYV), Tomato mild leaf curl virus (ToMLCV), Tomato severe mosaic virus (ToSMV), Tomato infections yellows virus (ToIYV), Tomato crinckle virus (ToCrV), Tomato chlorotic vein virus (ToClVV), Tomato mosaic Barbados and Tomato yellow mosaic virus (ToYMV) (Andrade et al., 2006; Fernandes et al., 2008; Kitajima, 2020). Recently, two new bipartite species have been identified in Brazil, Tomato interveinal chlorosis virus-2 (ToICV2) in Luziânia in the state of Goiás (Rêgo-Machado et al., 2019) and Tomato chlorotic leaf curl virus (ToCLCV) in Igarapé-Mirim in Pará (Quadros et al., 2019).

Table 3. Geographic distribution of 21 species of begomovirus reported naturally infecting tomatoes in Brazil (ICTV, 2020; Kitajima, 2020).

Species (acronym)	Geographic distribution	References
Chino del tomate Amazonas virus (CdTAV)	AM	(Fonseca et al., 2011)
Tomato bright yellow mosaic virus (ToBYMV)	BA	(Fonseca et al., 2013)
Tomato bright yellow mottle virus (ToBYMoV)	TO	(Fonseca et al., 2013)
Tomato chlorotic mottle virus (ToCMoV)	BA, MG, DF, ES, PE & RJ	(Ribeiro et al., 2003; Ribeiro et al., 2007)
Tomato common mosaic virus (ToCmMV)	RJ, MG & ES	(Castillo-Urquiza et al., 2008; Barbosa et al., 2016)
Tomato golden leaf distortion virus (ToGLDV)	TO	(Fonseca et al., 2010)
Tomato golden leaf spot virus (ToGLSV)	TO	(Fonseca et al., 2010)
Tomato golden mosaic virus (TGMV)	BA, DF, MG, PR, RN, RJ & SP	(Matyis et al., 1975; Hamilton et al., 1984)
Tomato golden vein virus (ToGVV)	GO, DF & MG	(Fernandes et al., 2008)
Tomato interveinal chlorosis virus (ToICV)	PE	(Albuquerque et al., 2012)
Tomato leaf distortion virus (ToLDV)	MG & RJ	(Castillo-Urquiza et al., 2008)
Tomato mild mosaic virus (ToMMV)	MG & RJ	(Castillo-Urquiza et al., 2008)
Tomato mottle leaf curl virus (ToMoLCV)	MG, GO, DF & PE	(Albuquerque et al., 2012)
Tomato rugose mosaic virus (ToRMV)	MG, GO, DF, SP, PR & BA	(Ribeiro et al., 2003; Fernandes et al., 2006)
Tomato rugose yellow leaf curl virus (TRYLCV)	RS	(Fonseca et al., 2016)

Tomato severe rugose virus (ToSRV)	DF, GO, MG, RJ, SP, PE & RS	(Rezende et al., 1997; Cotrim et al., 2007; Fernandes et al., 2008)
Tomato yellow spot virus (ToYSV)	GO & MG	(Calegario et al., 2007)
Tomato yellow vein streak virus (ToYVSV)	DF, GO, MG, RS, RJ & SP	(Faria et al., 1997; Albuquerque et al., 2010)
Tomato leaf curl purple vein virus (ToLCPVV)	PI	(Macedo et al., 2018)
Tomato interveinal chlorosis virus-2 (ToICV2)	GO	(Rêgo-Machado et al., 2019)
Tomato chlorotic leaf curl virus (ToCLCV)	PA	(Quadros et al., 2019)

Amazonas (AM); Bahia (BA); Tocantins (TO); Minas Gerais (MG); Distrito Federal (DF); Espírito (ES); Pernambuco (PE); Paraná (PR); Rio Grande do Norte (RN); Rio Grande do Sul (RS); Rio de Janeiro (RJ); São Paulo (SP) and Pará (PA).

In addition, there are other four begomoviruses reported on tomato under Brazilian conditions. However, these begomoviruses were initially reported affecting alternative weed hosts such as *Euphorbia heterophylla* L., *Sida rhombifolia* L., *Sida santaremnensis* L. and *Sida micranta* L. (**Table 4**).

Table 4. *Begomovirus* species reported infecting tomatoes in Brazil that were previously reported infecting alternative hosts (ICTV, 2020; Kitajima, 2020).

***Begomovirus* species and acronyms**	**Distribution**	**References**
Euphorbia yellow mosaic virus (EuYMV)	DF & GO	(Barreto et al., 2013)
Sida micranta mosaic virus (SimMV)	MG	(Calegario et al., 2004)
Sida mottle virus (SiMoV)	SP	(Cotrim et al., 2007)
Sida yellow net virus (SiYNV)	AM & RJ	(Fernandes, 2015)

Amazonas (AM); Minas Gerais (MG); Distrito Federal (DF); Rio de Janeiro (RJ); São Paulo (SP) and Goiás (GO).

In Brazil, we can overall consider ToRSV, ToCMoV, ToRMV, ToYVSV, and ToMoLCV as the predominant species affecting tomatoes (Ribeiro et al., 2003; Calegario et al., 2007; Fernandes et al., 2008; Albuquerque et al., 2010) **(Figure 6)**.

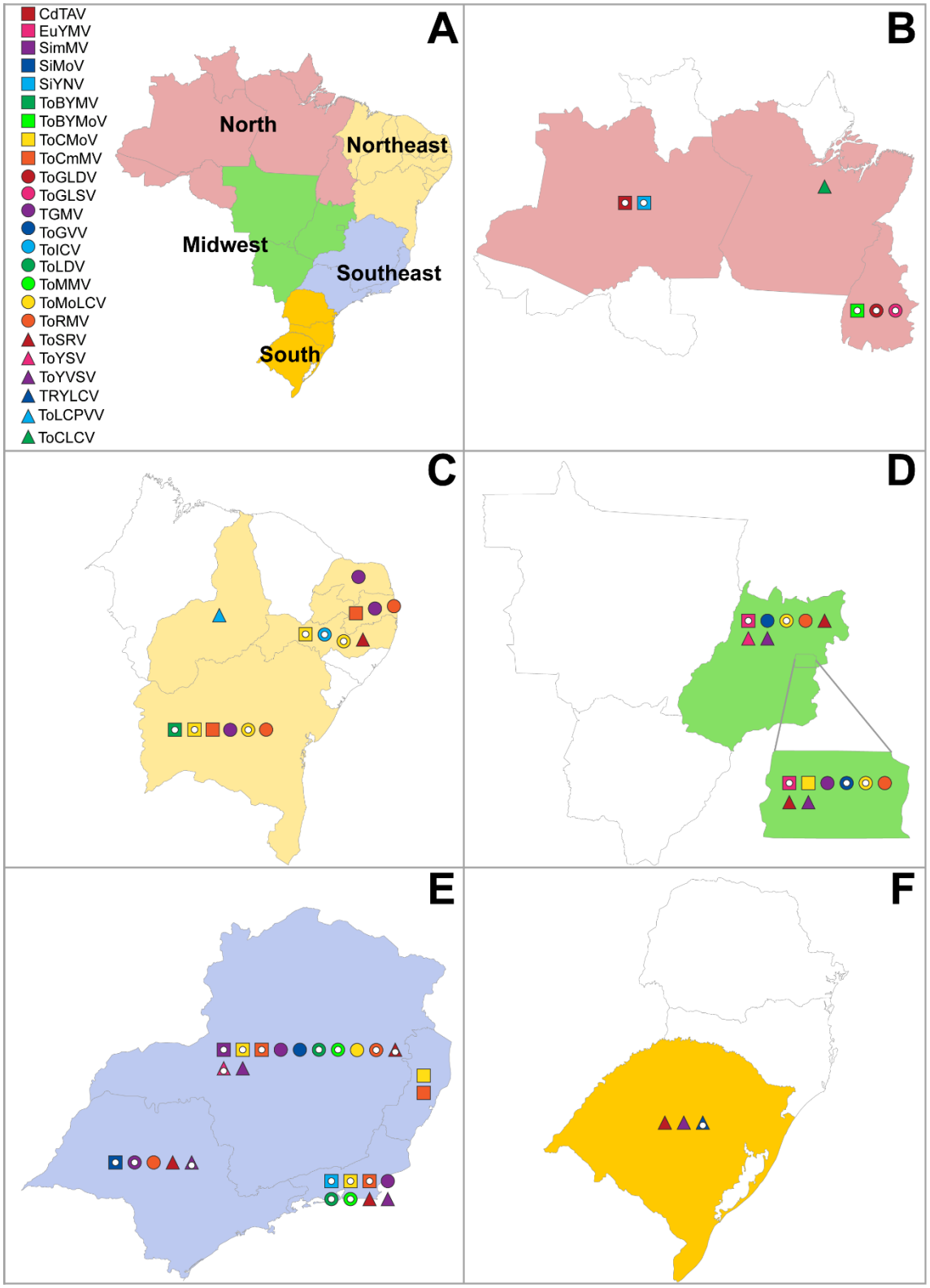

Figure 6. Map of the geographic distribution of begomovirus species reported infecting tomato in Brazil. (A) Acronyms of species with their respective colors and the map of Brazil divided by regions; (B) North Region; (C) Northeast Region; (D) Midwest Region; (E) Southeast Region and (F) South Region. Colors with white dots in the middle, refer to the first report of the virus species. Chino del tomate Amazonas virus (CdTAV); Euphorbia yellow mosaic virus

(EuYMV); Sida micranta mosaic virus (SiMMV); Sida mottle virus (SiMoV); Tomato bright yellow mosaic virus (ToBYMV); Tomato bright yellow mottle virus (ToBYMoV); Tomato chlorotic mottle virus (ToCMoV); Tomato common mosaic virus (ToCmMV); Tomato golden leaf distortion virus (ToGLDV); Tomato golden mosaic virus (TGMV); Tomato golden vein virus (TGVV); Tomato interveinal chlorosis virus (ToICV); Tomato leaf distortion virus (ToLDV); Tomato mild mosaic virus (ToMMV); Tomato mottle leaf curl virus (ToMoLCV); Tomato rugose mosaic virus (ToRMV); Tomato severe rugose virus (ToSRV); Tomato yellow spot virus (ToYSV); Tomato yellow vein streak virus (ToYVSV); Sida yellow net virus (SiYNV); Tomato rugose yellow leaf curl virus (TRYLCV); Tomato leaf curl purple vein virus (ToLCPVV) and Tomato chlorotic leaf curl virus (ToCLCV).

9. Resistance genes to begomovirus characterized in tomato

The use of cultivars with genetic resistance is the most effective control strategy to minimize losses caused by diseases of viral etiology (Boiteux et al., 2012a). In the case of diseases caused by begomoviruses, the use of host genetic resistance is the most economically and environmentally sustainable strategy. Chemical control of the vector is almost impossible in situations involving migration of large viruliferous whitefly populations of from older crops to early established fields in association with the presence of vector subpopulations with resistance to different groups of insecticides (Silva et al., 2009; Yao et al., 2017).

The first breeding programs for begomovirus resistance in cultivated tomatoes (*S. lycopersicum*) were based on the search for resistance genes/alleles in wild *Solanum* species (Boiteux et al., 2012a; Pereira-Carvalho et al., 2014; Dhaliwal et al., 2019). Several accessions of different species of wild tomatoes were identified as potential sources of resistance within the germplasm of *S. pimpinellifolium*, *S. peruvianum*, *S. chilense*, *S. habrochaites* and *S. cheesmaniae*. Several resistance genes/loci have been identified and they were introgressed in commercial cultivars (Pereira-Carvalho et al., 2014; Dhaliwal et al., 2019).

Eight resistance genes /loci were characterized and mapped in the tomato genome: *Ty*–1 (Zamir et al., 1994), *Ty*–2 (Hanson et al., 2006), *Ty*–3 (Ji and Scott, 2006), *Ty*–4 (Ji et al., 2009), *ty*–5 (Anbinder et al., 2009), *Ty*–6 (Hutton and Scott, 2014), *tcm*–1 (Giordano et al., 2005a) and *tgr*–1 (Bian et al., 2007). It is important to emphasize that, until now, all characterized genes in tomato do not confer immunity-like responses, but

they provide, in general, high levels of partial resistance and/or tolerance (*sensu* Cooper and Jones, 1983).

The semi-dominant *Ty*–1 gene (Table 5): This gene is one of the most used in breeding programs in the Americas, France, Italy, and Israel. *Ty*–1 is located (in repulsion phase) in the chromosomal region containing resistance genes for other pathogens, including the *Mi*1.2 gene, which confers resistance to *Meloidogyne* species (Pereira-Carvalho et al., 2010; Verlaan et al., 2011). The phenotypic expression of the *Ty*–1 gene against a wide range of begomoviruses is best described as a tolerance response, since the plant containing this factor allows a mild manifestation of symptoms in the apical meristematic regions (Boiteux et al., 2007a). This tolerance manifests against a wide spectrum of monopartite and bipartite begomoviruses and its action is related to the inhibition of viral movement, being more efficient in conditions of low inoculum pressure (Zamir et al., 1994; Boiteux et al., 2007a). Genetic mapping conducted by Verlaan et al. (2011) allowed the cloning and identification of the *Ty*–1 gene. It encodes an RNA-dependent RNA polymerase (RDR) belonging to the RDRy type, representing an entirely new class of genes that confers resistance by intensifying the levels of transcriptional gene silencing (Butterbach et al., 2014). Subsequent studies have shown that the *Ty*–1 gene also confers resistance to *Beet curly top virus* (BCTV) (genus *Curtovirus*) in transformed *Nicotiana benthamiana* plants. Interestingly, plants of *N. benthamiana* transformed with the *Ty*–1 gene showed, under conditions of TYLCV and *Ageratum yellow vein betasatellite* (AYVB) co-infection, a higher intensity of symptoms and a higher concentration of TYLCV, when compared with plants inoculated only with TYLCV (Voorburg et al., 2020). The *Ty*–1 gene does not confer resistance to RNA viruses as already demonstrated for Tomato spotted wilt orthotospovirus (TSWV) and Cucumber mosaic virus (CMV) (Butterbach et al., 2014). However, simultaneous infections with RNA viruses can compromise resistance against begomoviruses as shown previously during TYLCV and CMV co-infection, where an increase in TYLCV concentration was observed due to the inhibition of the transcriptional gene silencing response by CMV 2b RNAi suppressor protein (González et al., 2010; Hamera et al., 2012; Butterbach et al., 2014).

The *Ty*–2 resistance gene (Table 5): This gene confers resistance to different TYLCV isolates present in different countries, such as Taiwan, Vietnam, India, and Israel. However, this resistance is not effective against isolates from northern India, Thailand, the Philippines, and Central America. (Hanson et al., 2006). In Brazil, *Ty–2* gene was

found to be an efficient source of resistance to ToRMV using inoculation with *B. tabaci* biotype B (Boiteux et al., 2007b). This gene has recently been determined to belong to the family of resistance genes containing nucleotide binding domain and leucine-rich repeats (NB-LRR) (Yamaguchi et al., 2018).

The *Ty*–3 resistance gene/allele (Table 5): This gene was initially described as a new gene/allele introgressed from *S. chilense*. However, subsequent molecular studies have shown that the *Ty*–1 and *Ty*–3 genes are more likely allelic variants of the same gene (Verlaan et al., 2013).

The *Ty*–4 resistance gene (Table 5): It is a gene that is less efficient than *Ty*–1 and it was found to be not effective for all TYLCV isolates (Ji et al., 2009; Kadirvel et al., 2013).

The *ty*–5 resistance gene (Table 5): Segregation studies have suggested that resistance in the accession 'TY172' is controlled by three genes, two of which have partially dominant effects and the other recessive. However, additional studies have indicated that the resistance observed in 'TY172' is conferred by a dominant QTL (Quantitative Trait Loci = locus of quantitative characteristic) with great effect called *Ty*–5 and by other smaller QTLs, mapped on chromosome 4 (Anbinder et al., 2009; Wang et al., 2018). Functional mapping and validation studies have shown that a *Pelo* – Protein pelota homolog (involved in the ribosome recycling phase in protein synthesis) is, in fact, the gene product of a recessive factor (*ty*–5). Silencing this gene in a susceptible inbred line made transgenic plants highly resistant to TYLCV. Therefore, the *Pelo* gene offers an alternative route to promote resistance to TYLCV and other related viruses (Lapidot et al., 2015; Wang et al., 2018).

The *Ty*–6 resistance gene (Table 5): The action of this gene is incomplete dominance, with an intermediate response when *Ty*–6 is heterozygous. Another important information is that *Ty*–6 complements the resistance conferred by the *Ty*–1 and *ty*–5 genes in pyramidized inbred lines (Gill et al., 2019).

The *tcm*–1 gene (Table 5): It was the first recessive gene described for resistance to begomovirus in tomatoes (Giordano et al., 2005b). Studies conducted to elucidate the efficiency spectrum of this source, indicated that *tcm*–1 is effective against bipartite species from Brazil and monopartite species from Europe (Pereira-Carvalho et al., 2010; Pereira-Carvalho et al., 2015). Inheritance studies conducted in Brazil and Spain indicated that the expression of resistance to both ToCMoV and TYLCV is better explained by a monogenic recessive model (Giordano et al., 2005b; García-Cano et al., 2008; Pereira-

Carvalho, 2009), although the participation of a second minor gene (also recessive) cannot be ruled out in some virus-specific responses (García-Cano et al., 2008). The original source of the *tcm*-1 gene (the inbred line 'TX-468-RG') also showed recessive resistance to a set of species of the Tomato yellow leaf curl disease (TYLCD) complex. The resistance response is associated with limitation of systemic virus accumulation and absence of symptom expression (García-Cano et al., 2008). Studies carried out with the TYLCV isolate from Israel, indicated that the resistance of 'TX-468-RG' interferes with the viral translocation over long distances within the infected plant and causes a reduction in the viral accumulation in the apical leaves (Pereira-Carvalho, 2009).

The *tgr*–1 gene (Table 5): Another recessive gene was named as *tgr*-1 and it was characterized in the inbred line 'FLA-653'. Plants with this gene showed high levels of resistance to the begomovirus TYLCV (Bian et al., 2007).

Although the use of resistant cultivars constitutes the most efficient strategy for reducing losses caused by begomoviruses, a considerable increase in the use of these genetic materials could result in the selection of specific viral isolates, accelerating the change in the population composition which may culminate in the emergence of isolates capable of overcoming these resistance/tolerance factors. In fact, some studies have been published reporting the breakdown of resistance of cultivars containing the *Ty*–1 gene by TYLCV strains. In Morocco, Italy and Spain tomato plants containing *Ty*–1 have been identified exhibiting severe symptoms caused by TYLCV. The analyzes revealed the presence of a new viral variant which was found to be derived from a recombination event between the TYLCV and TYLCSV, encompassing a non-coding region of the viral genome between the origin of replication and the beginning of the V2 gene (Belabess et al., 2015; Panno et al., 2018; Granier et al., 2019; Torre et al., 2019). Studies have also reported one strain of TYLCV able to overcome the resistance mediated by the *Ty*–2 gene (Ohnishi et al., 2016) and also a strain of Tomato leaf curl Bangalore virus (ToLCBV) leading to great losses in production in India (Tiwari et al., 2010). This breakdown of resistance can be explained by multiple changes in replication efficiency, and viral gene expression. Another hypothesis is that the region involved in the recombination event may be less prone to transcriptional gene silencing (Voorburg et al., 2020).

Table 5: Genes of resistance against *Begomovirus* characterized in tomato.

Gene	Chromosome / Reference	Resistance mechanism	Viral species that resistance has been demonstrated	Original source of resistance
***Ty*-1**	6 (Zamir et al., 1994)	RNA-dependent RNA polymerase (Verlaan et al., 2013); Movement of the long-distance virus (Tobar, 2013).	TYLCV (Zamir et al., 1994); ToCMoV (Boiteux et al., 2007; Tobar, 2013); ToRMV (Pereira-Carvalho et al., 2009); ToSRV & ToYVSV (Tobar, 2013).	*S. chilense*
***Ty*-2**	11 (Hanson et al., 2006)	Domain containing leucine-rich repeats - NB-LRR (Yamaguchi et al., 2018)	TYLCV (Hanson et al., 2006). ToSRV (Boiteux et al., 2007)	*S. habrochaites*
***Ty*-3**	6 (Ji & Scott, 2006)	Not yet characterized	TYLCV & ToMoV (Ji & Scott, 2006)	*S. chilense*
***Ty*-4**	3 (Ji et al., 2008)	Not yet characterized	TYLCV & ToMoV (Nakhla et al.,2005)	*S. chilense*
***ty*-5**	4 (Anbinder et al., 2009)	Protein translation (Lapidot et al. 2015).	TYLCV & ToMoV (Anbinder et al., 2009)	*S. peruvianum*
***Ty*-6**	10 (Hutton & Scott, 2013)	Not yet characterized	TYLCV & ToMoV (Hutton, 2012)	*S. chilense*
***tcm*-1**	6 (Machado, 2013)	Less viral accumulation & translocation, restriction in viral replication, less efficiency in viral movement (Pereira-Carvalho et al., 2009)	ToYLCV; ToSRV; ToRMV; ToYVSV & ToCMoV (Pereira-Carvalho et al., 2009; Machado, 2013)	TX-468
***tgr*-1**	6 (Bian et al., 2007)	Not yet characterized	TYLCV (Bian et al., 2007)	FLA 653’

10. *Next Generation Sequencing* (NGS) applied to Plant Virology

The use of next-generation sequencing (NGS) technology is a powerful tool for large-scale detection and identification of new virus species in different crops. Several modern sequencing technologies are being employed including: Illumina, 454, Pacific Biosciences, IonTorrent and Nanopore (Barba and Hadidi, 2015; Adams and Fox, 2016; Hadidi, 2019; Villamor et al., 2019). NGS allows massive sequencing of biological and/or environmental samples (including complex genomes) and generating an enormous number of reads (short sequences) in small time intervals. The demand for NGS technologies was generated by the need for large-scale sequencing in a more economical, effective, and fast way. These demands were not being met with the automated sequencing obtained with the use of "first generation" Sanger sequencer machines.

Analyzes using metagenomics combined with NGS for characterization of viruses or viroids are collectively called "virome" (Barba et al., 2014; Villamor et al., 2019). To carry out studies on the diversity of plant viruses, samples are subjected to previous procedures for enriching viral particles via semi-purification protocols followed by extraction of nucleic acid (DNA or RNA) as well as extraction of dsRNA and sRNAs (= small RNAs produced as a result of plant defense mechanisms against viruses, such as gene silencing). For enrichment of DNA viruses, a protocol is performed employing Rolling Circle Amplification – RCA (Idris et al., 2014; Roossinck et al., 2015; Kathurima et al., 2016; Massart et al., 2019).

Studies with plant viruses using metagenomics combined with NGS started in 2009 (Adams et al., 2009; Al Rwahnih et al., 2009; Kreuze et al., 2009). Those studies allowed the characterization of host-specific viromas and their insect vectors, which has considerably increased the pace of discovery of new viruses as well as the expansion of databases of complete sequences of many previously known virus species. It is estimated that more than 100 new plant viruses associated with different genera and families have been identified in recent years (Hadidi and Barba, 2012; Barba et al., 2014; Ho and Tzanetakis, 2014; Barba and Hadidi, 2015; Roossinck et al., 2015; Wu et al., 2015; Bernardo et al., 2018). Three new viroids were also discovered using this type of technology: Persimmon viroid 2 (PVd2), Grapevine latent viroid (GLVd) and Apple chlorotic fruit spot viroid (ACFSVd) (Ito et al., 2013; Zhang et al., 2014; Leichtfried et al., 2019).

A new species of the genus *Foveavirus* (named as Grapevine virus T – GVT) was identified in a recent work involving a grape transcriptome analysis (Jo et al., 2017). A

new member of the genus *Badnavirus* was detected in *Betula pendula* and *B. pubescens* and named as Birch leaf roll-associated virus (BLRaV) (Rumbou et al., 2018). The new species Cherry virus A (CVA) and Little cherry virus 1 (LChV– 1) were detected in apricot samples in Hungary (Baráth et al., 2018). The complete genome of a new rhabdovirus infecting corn and wheat in Argentina was obtained via NGS as named as Maize yellow striate virus (MYSV) (Maurino et al., 2018). In the genus *Dioscorea*, the complete genome of two Badnavirus species (Dioscorea nucleou bacilliform RT virus 1 – DBRTV1 and Dioscorea nucleous bacilliform RT virus 2 – DBRTV2) and one of potyvirus (Yam mosaic virus – YMV) were also obtained via NGS. This work also highlighted the NGS usefulness in detecting virus species in in tissue culture of *Dioscorea* species (Bömer et al., 2018). Recently, the new species *Prunus virus* F (PrVF) was detected via NGS on sweet cherry (*Prunus avium* L.) in Belgium (Tahzima et al., 2019). The NGS strategy also allow the discovery a great viral diversity in different peach cultivars. A total of eight viruses and viroids (belonging to five families) were identified: Peach latent mosaic viroid (PLMV) of the genus *Pelamoviroid* (Family *Avsunviroidae*); two new Nectarine species stem – pitting associated virus (NSPaV) and Peach-associated luteovirus (PaLV) (Family *Luteoviridae*); Plum bark necrosis stem pitting-associated virus (PBNSPaV) of the genus *Ampelovirus* (Family *Closteroviridae*); Apple chlorotic leaf spot virus (ACLSV) from the genus *Trichovirus*, Asian prunus virus 1 and 3 (APV1 and APV3) from the genus *Foveavirus*; Cherry green ring mottle virus (CGRMV) and Cherry necrotic rusty mottle virus (CNRMV) both from the genus *Robigovirus* (Family *Betaflexiviridae*); and Peach virus D (PeVD) belonging to the genus *Marafivirus* (Family *Tymoviridae*) (Xu et al., 2019).

Recently, in Mexico, 132 species of uncultivated plants belonging to 34 families were submitted to metagenomic analysis combined with NGS. These analyzes indicated a great diversity of *Begomovirus* in uncultivated plants of the families Brassicaceae, Convolvulaceae, Curcubitaceae, Euphorbiaceae, Fabaceae, Malvaceae, and Solanaceae. Fourteen begomoviruses with monopartite genomes and five with bipartite genomes were detected, in addition to the curtoviruses BCTV and TPCTV (Rodríguez-Negrete et al., 2019). In grapevine, NGS has been shown to be a very efficient method in detecting viruses and satellites (Baráth et al., 2018). Recently a new species classified in *Vitivirus* (Family *Betaflexiviridae*) was identified in Argentina. The proposed name for the new species was Grapevine virus L (GVL) (Debat et al., 2019). In Hungary, the NGS proved to be very efficient in identifying commonly found virus species as well as others that had

not yet been detected in that country. In this study, the following species were detected: Grapevine chrome mosaic virus – GCMV (*Nepovirus*), Grapevine leafroll-associated virus 1 – GLRaV1 (*Ampelovirus*), Grapevine virus A – GVA and Grapevine virus B – GVB (*Vitivirus*); seven species of the order Tymovirales: Grapevine fleck virus – GFKV (*Maculavirus*), Grapevine asteroid mosaic associated virus – GaMaV, Grapevine syrah virus 1 – GSyV1 (*Marafivirus*), Grapevine rupestres stem pitting-associated virus – GRSPaV (*Foveavirus*), Grapevine pinot gris virus – GPGV (*Trichovirus*), Grapevine red globe virus – GRGV and Grapevine rupestris vein feathering – GRVFV and a species of the genus *Idaeovirus*, Raspberry bushy dwarf virus – RBDV. It was also possible to detect a satellite DNA Grapevine satellite virus – GSV and three viroids Hop stunt viroid – HSVd, Grapevine yellow speckled viroid 1 – GYSVd1 and Grapevine yellow speckled viroid 2 – GYSVd2 (Czotter et al., 2018).

In tomato, studies involving analyzes of the viroma through NGS contributed significantly to the detection and discovery of new viral species present in this crop in Brazil and in other countries. Studies with tomato samples collected in California, Mexico and Arizona, were able to detect by Illumina sequencing, Potato spindle tuber viroid – PSTVd, Pepino mosaic virus – PepMV and a new species member of the *Potyvirus* genus named Tomato necrotic stunt virus – ToNStV (Li et al., 2012). In Slovakia, it was possible to obtain the complete genome of the *Carlavirus* Potato virus M – PVM (Glasa et al., 2019).

Recently, a survey aiming to study the diversity, distribution, and evolution of viruses from tomato in China. Leaf samples were collected in 2013 and a total of 22 virus species were identified belonging to the genera *Polerovirus*, *Crinivirus*, *Ilarvirus*, *Cucumovirus*, *Tobamovirus*, *Carlavirus*, *Potyvirus*, *Orthotospovirus*, *Begomovirus* and *Amalgavirus*. In addition, it was also possible to detect a mycovirus of the genus *Mitovirus* and a probable new species, named as Tomato yellow mottle-associated virus – TYMaV. Analysis of the genomic and phylogenetic organization indicated that TYMaV is a possible new member of the genus *Cytorhabdovirus* (Xu et al., 2017a).

In Brazil, tomato samples with typical symptoms of viral infections were collected in Brazlândia (DF), Campinas (SP) and Araguari (MG). The results showed a great viral diversity of natural occurrence in the Brazilian production fields, such as the species Groundnut ringspot orthotospovirus – GRSV and Tomato spotted wilt orthotospovirus – TSWV (*Orthotospovirus*); Pepper ringspot virus – PepRSV (*Tobravirus*); Tomato chlorosis virus – ToCV (*Crinivirus*); Sida micranta mosaic virus – SiMMV and Tomato

severe rugose virus – ToSRV (*Begomovirus*); Tomato blistering mosaic virus – ToBMV (*Tymovirus*); Pepper yellow mosaic virus – PepYMV and Potato virus Y – PVY (*Potyvirus*); and Pepper mild mottle virus – PMMoV (*Tobamovirus*). This study was the first report of Southern tomato virus – STV belonging to the genus *Amalgavirus*, in the country and report of a probable new species of *Ilarvirus* due to its low identity with the species *Ageratum latent virus* – ALV and Parietaria mottle virus – PMoV (Martins, 2017).

In recent years, new members of the *Geminiviridae* family have been detected by NGS. The geminiviruses Citrus chlorotic dwarf-associated virus – CCDaV and Mulberry mosaic dwarf-associated virus – MMDaV were reported in citrus in Turkey and apple tree in China, respectively (Loconsole et al., 2012; Ma et al., 2015). A new geminivirus, Apple geminivirus – AGV was reported on apple in China. The AGV shows genomic organization different from the other members of the *Geminiviridae* family, and has the ability to infect hosts such as *N. benthamiana*, *N. glutinosa* and tomatoes (Liang et al., 2015). Another geminivirus was detected in vines (*V. vinifera* L.), cultivars Black Beet and Nagano Purple, from South Korea. The sequence of the monopartite circular genome was confirmed in the samples by RCA, Sanger cloning and sequencing. Phylogenetic analyzes allowed the classification of this isolate in the *Geminiviridae* family. The proposed name for this new species was Grapevine geminivirus A – GGVA (Al Rwahnih et al., 2016).

Fontenelle et al. (2017) identified a new geminivirus in tomato and cleome plants collected in the central region of Brazil. The proposed name for the new species was Tomato associated geminivirus 1 – TaGV1. The protein associated with TaGV1 replication is similar to that of species of *Capulavirus* genus (62–70% identity) while the CP is closer to the species Tomato pseudo-curly top virus (genus *Topocuvirus*), indicating that this species is a member of a probable new genus within the *Geminiviridae* family (Fontenele et al., 2017).

All the examples mentioned above reinforce the importance of NGS as a universal, fast and accurate method for the discovery and detection of virus species, enabling a greater understanding of viral diversity in different species of cultivated and non-cultivated plants. This contributes in a very significant way to plant virology, especially regarding control measures, since information about the diversity of virus species is crucial in establishing effective management strategies.

11. Family *Genomoviridae*

Viruses with single-stranded DNA genomes are associated with organisms in all domains of life (*Archaea*, *Bacteria* and *Eukarya*). The discovery of new viruses with circular ssDNA has revealed a wide diversity within this group (Rosario et al., 2012b; Varsani and Krupovic, 2017). These viruses share low, but significant degrees of similarity with viruses with circular ssDNA in the *Geminiviridae*, *Circoviridae* and *Nanoviridae* families. Because of this, it was proposed that these new ssDNA viruses (despite having similarities to members of these three families) would be more appropriately allocated within a new family (Rosario et al., 2012a; Krupovic et al., 2016).

In 2016 the International Committee on Taxonomy of Viruses created two new virus families with ssDNA genomes (*Pleolipoviridae* and *Genomoviridae*). The family *Genomoviridae* is represented by nine genera: *Gemycircularvirus*, *Gemyduguivirus*, *Gemygorvirus*, *Gemykibivirus*, *Gemykolovirus*, *Gemykrogvirus*, *Gemykroznavirus*, *Gemytondvirus* and *Gemyvongvirus* (Krupovic et al., 2016; ICTV, 2020). Genomoviruses are circular non-enveloped ssDNA viruses with a genome size of around 2.0 to 2.4 kb. Viruses belonging to this family generally have only two ORFs in bidirectional transcriptional units, CP (protein encoded in the viral sense) and Rep (protein associated with replication) in the complementary sense. All the genera share a typical intergenic region that varies in size, and in some cases the presence of an intron in the Rep coding region (**Figure 7**). The intron inside the Rep undergoes a splicing process generating the functional Rep (Male et al., 2016; Varsani and Krupovic, 2017). Nonanucleotide sequences are variable for each genus among genomoviruses (Varsani and Krupovic, 2017).

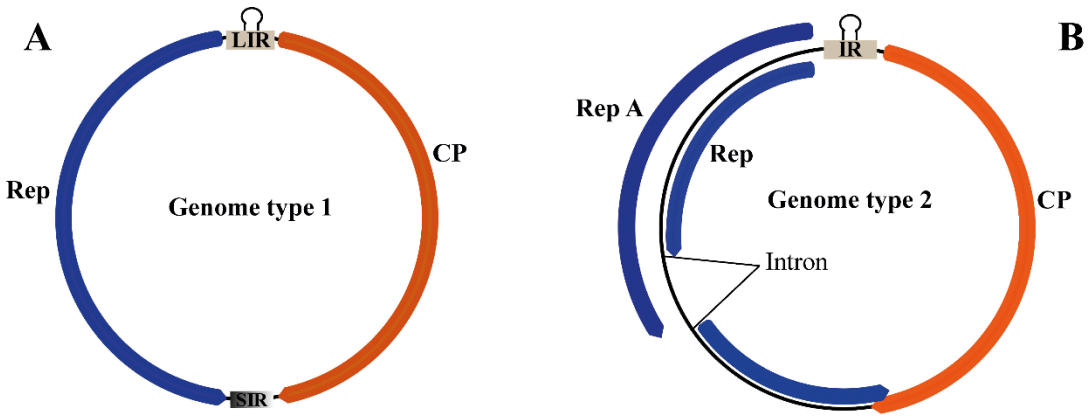

Figure 7. Genomic organization of *Genomoviridae* family (Genome type 1 and type 2). ORFs (*Open Reading Frames*) Rep (*Replication associated protein* - protein associated with replication) and CP (*Coat protein* – capsid protein). **(A**) Genome type 1 presenting two intergenic regions LIR (*Long intergenic region*), SIR (*Short intergenic region*) and the Rep ORF of this group does not present an Intron; **(B)** Type 2 genome shows only one intergenic region, Rep A and the presence of an Intron in the Rep ORF.

The terminal N region contains important motifs for initiating rolling circle replication (RCR). Some of these motifs are conserved in ssDNA, phages and plasmids that replicate using the RCR mechanism (Krupovic, 2013; Varsani and Krupovic, 2017). Rep is a multifunctional protein that has essential functions in replication, with activity of endonuclease and helicase. Rep helicase activity is mediated by conserved motifs known as Walker A (**GxxxxGKT**), Walker B (**uuDDu**) and C motif (**uxxN**) located in a C-terminal NTP-binding domain (Choudhury et al., 2006; Clérot and Bernardi, 2006). The taxonomic criteria for the classification of new species within the *Genomoviridae* family is that the new virus must present 78% of paired identity when compared to the already established species. A differential criterion is adopted for the demarcation of species within each genus (Krupovic et al., 2016; Varsani and Krupovic, 2017).

The ***Gemyduguivirus*** **genus**: has only one species Dragonfly associated gemyduguivirus 1. The identity for the classification of a new species in the genus is between 57% to 62% with other available sequences. The Dragonfly associated gemyduguivirus 1 was found in association with insects, and shows the conserved nonanucleotide sequence "TAATATTAT". (Rosario et al., 2012a; Varsani and Krupovic, 2017; ICTV, 2020).

The ***Gemygorvirus*** **genus** aggregates five species: Canine associated gemygorvirus 1, Mallard associated gemygorvirus 1, Pteropus associated gemygorvirus

1, Sewage derived gemygorvirus 1 and the Starling associated gemygorvirus 1. Members of this genus have been associated to birds (Van den Brand et al., 2012; Sikorski et al., 2013; Kraberger et al., 2015; Male et al., 2016; Steel et al., 2016). The conserved nonanucleotide sequences of this genus are variable in the fourth, fifth and seventh position "TAT**WW**A**W**AS". The criterion for the classification of a new species within the genus is to have a nucleotide identity less than 49%, when compared to other species (Varsani and Krupovic, 2017).

At the ***Gemykibivirus*** **genus** 16 species have been accepted so far. The species type is Dragonfly associated gemykibivirus 1 (DaGmV – 1), which was characterized in association to insects (Rosario et al., 2012a; ICTV, 2020). The conserved nonanucleotide sequence of this genus is variable in four different positions ("TATA**WWM**A**V**"). The demarcation criterion for the classification of new species within the genus is to have a nucleotide identity of less than 43% (Varsani and Krupovic, 2017).

The **genus *Gemykolovirus*** has 2 accepted species Pteropus associated gemykolovirus 1 (type-species) and Pteropus associated gemykolovirus 2, that have been reported in association to mammals (Male et al., 2016; ICTV, 2020). The conserved nonanucleotide sequence of this genus is variable at fifth and sixth position("TAAT**RY**TAT") The demarcation criterion for the classification of new species in *Gemykolovirus* is identity of less than 37% with other species (Varsani and Krupovic, 2017).

The **genus *Gemykrogvirus*** has three accepted species: Bovine associated gemykrogvirus 1 (type-species), Caribou associated gemykrogvirus 1, and Sewage derived gemykrogvirus 1 reported in association with mammals (Lamberto et al., 2014; Ng et al., 2014; Kraberger et al., 2015). The nonanucleotide sequence is the same found in the genus *Gemyduguivirus*.The demarcation criterion for the classification of new species within the genus is a nucleotide identity of less than 33% with other sequence available (Varsani and Krupovic, 2017).

The **genus *Gemykroznavirus***, is monotypic with the species Rabbit associated Gemykroznavirus 1, also reported in association with mammals. The nonanucleotide sequence is the same found in the genera *Gemyduguivirus* and *Gemykrogvirus.* The identity of nucleotide to considere and to classify as a newspecies in the genus is between 56% to 61% (Sikorski et al., 2013; ICTV, 2020).

The **genus *Gemytondvirus***, monotypic, is represented by Ostrich associated Gemytondvirus 1 that was found in association with birds. The sequence has a conserved

nonanucleotide sequence "TAA**C**ATT**G**A". The identity of nucleotide for the classification of a new species in this genus is between 53% to 61% (Sikorski et al., 2013; Varsani and Krupovic, 2017; ICTV, 2020).

In the monotypic genus ***Gemyvongvirus***, the species Human associated gemyvongvirus 1 was found in association with mammals. The identity of nucleotide for the classification of a new species in the genus is between 56% to 62%. This genus is characterized by the conserved nonanucleotide sequence "TAA**A**TAA**G**A" (Varsani and Krupovic, 2017; ICTV, 2020).

In Brazil a new genomovirus has been identified associated with common bean plants (*Phaseolus vulgaris* L.) collected in Pernambuco state. The proposed name for the species was common bean-associated gemycircularvirus – CbaGmV (Lamas et al., 2016). Analyzes with two samples of capybara feces (*Hydrochoerus hydrochaeris*) in Brazil, revealed a great diversity of ssDNA viruses, in which 148 complete sequences of viruses that belonged to the family were identified *Microviridae*, 14 *Genomoviridae*, a new type of *Cyclovirus* (Family *Circoviridae*) it is a *Smacovirus* (Family *Smacoviridae*) (Fontenele et al., 2019).

11.1 *Gemycircularvirus*

The genus Gemycircularvirus has the largest number of described species, 43 in total. The genus is characterized by a circular ssDNA genome with 2.1 to 2.3 kb and two ORFs, one encoding the capsid protein (in the viral sense) and another encoding the Rep protein (in the complementary sense). The Rep protein contains two conserved domains important for RCR, also present in the Rep proteins of the geminiviruses. In addition, they have a stem-loop structure in the region of origin of replication, where there is a conserved nonanucleotide sequence – TAAT**RY**TAT (Rosario et al., 2012a; Krupovic et al., 2016; Varsani and Krupovic, 2017).

The first described gemycircularvirus (the type species of the genus) was Sclerotinia sclerotiorum hypovirulence associated circular DNA virus 1 (SsHADV–1), isolated from the fungus *Sclerotinia sclerotiorum*, causing hypovirulence phenotype (Yu et al., 2010). Currently, two techniques have been used to exploration and discovery of these new ssDNA viruses: the amplification via rolling circle (Rolling Circle Amplification – RCA) (Inoue-Nagata et al., 2004) and metagenomics (Edwards and Rohwer, 2005). As a result, many species of gemycircularvirus have been reported in association with plants. The species Cassava associated circular DNA virus (CasCV) was

identified in Ghana (Africa) in association with cassava leaves (*M. esculenta*) infected with *Collectotrichum* sp. and *Plectosphaerella* sp. (Dayaram et al., 2012). In Vietnam, the Hypericum japonicum-associated circular DNA virus (HJasCV) was isolated from plants of *H. japonicum* with yellow mosaic symptoms (Du et al., 2014). On the island of Tonga, three gemycirculaviruses were identified in 43 grass samples, using Illumina sequencing. The species Brachiaria deflexa-associated circular DNA molecule 1 (BdaCM–1) and Brachiaria deflexa-associated circular DNA molecule 2 (BdaCM–2) were detected in *Brachiaria deflexa* (Schumach.) C. E. Hubb.) and Poaceae-associated gemycircularvirus 1 (PaGmV–1) in sugar cane (Male et al., 2015). Analyzes of fungal viromas from soybean leaf samples in the United States identified 22 mycovirus sequences, including a new gemycircularvirus called Soybean leaf-associated gemycircularvirus 1 (SlaGemV–1) (Marzano and Domier, 2016). Recently, two gemycircularvirus Momordica charantia associated gemycircularvirus (MoaGmV) and Euphorbia heterophylla associated gemycircularvirus (EuaGmV) were found in Brazil associated with the weeds *Momordica charantia* L. and *Euphorbia heterophylla* L. collected in Minas Gerais and Rio Grande do Sul (De Rezende et al., 2018).

Different strategies have been used to analyze the evolutionary processes capable of shaping the genetic-molecular structure of begomovirus populations. In this context, the general objective of the present work was to carry out a study on *Begomovirus* species diversity occurring on tomato in Central Brazil through metagenomic analysis via NGS. In addition, analyzes will also be conducted to estimate the potential impact of the introduction of varieties containing the *Ty*–1 gene on viral evolutionary dynamics.

CHAPTER 2

Metagenomics of Neotropical single-stranded DNA (ssDNA) viruses in tomato cultivars with and without the *Ty*–1 gene

Luciane de Nazaré Almeida dos Reis[1], Maria Esther de Noronha Fonseca[2], Simone Graça Ribeiro[3], Fernanda Yuri Borges Naito[1], Leonardo Silva Boiteux[1,2]*, and Rita de Cássia Pereira-Carvalho[1]*

[1]Universidade de Brasília (UnB), Dept. Fitopatologia, Área de Virologia Vegetal, Brasília-DF, Brazil.

[2]National Center for Vegetable Crops Research (CNPH), Embrapa Hortaliças, Brasília – DF, Brazil.

[3]Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, Brazil.

Work submitted to Viruses

Resumo

Um complexo de begomovírus (*Geminiviridae*) pode causar graves perdas de produção de tomate em regiões neotropicais. Nesse trabalho, a estratégia de *next-generation sequencing* (NGS) foi empregada para avaliação em larga escala da diversidade do vírus de DNA de fita simples (ssDNA) em tomates com e sem o gene de amplo espectro resistência *Ty*–1 no Brasil Central. As amostras foliares (n = 107) exibindo sintomas típicos de begomovírus foram coletadas em condições de campo. As amostras individuais foram enriquecidas com DNA circulares e subdivididas em dois conjuntos (com e sem *Ty*–1) e sequenciadas em uma plataforma Illumina. As validações por PCR com primers específicos para os vírus e sequenciamento Sanger confirmaram um total 15 vírus de ssDNA e/ou agentes subvirais (ocorrendo principalmente em infecções mistas). Esta multiplicidade viral destaca a desvantagem potencial de se empregar resistência do tipo vírus-específica em plantações de tomate. Uma maior diversidade viral (14 versus 6 espécies) foi observada em tomates sem o gene *Ty*–1. Um gemycirculavírus (*Genomoviridae*), um novo alfa-satélite e duas novas espécies de *Begomovirus* foram identificados exclusivamente em amostras sem o gene *Ty*–1, enquanto um novo begomovírus foi encontrado apenas em amostras com o gene *Ty*–1. Esse último vírus foi o único encontrado induzindo sintomas severos em plantas com o gene *Ty*–1 nesse levantamento. Os resultados indicaram a necessidade de mais estudos sobre a potencial adaptação viral ao *Ty*–1 e os efeitos desse gene na filtragem do tipo espécie-específica em um subconjunto de agentes virais/subvirais de ssDNA.

Palavras chaves: *Begomovirus*, tomate, gene de resistência, NGS, viroma.

Abstract

A complex of begomoviruses (*Geminiviridae*) can cause severe tomato yield losses in the Neotropics. Here, next-generation sequencing was employed for large–scale assessment of single-stranded (ss) DNA virus diversity in tomatoes either harboring or lacking the large–spectrum *Ty*–1 tolerance gene in Central Brazil. Leaves exhibiting begomovirus–like symptoms (n=107) were field–collected; circular DNA–enriched individual samples were subdivided into pools (with and without *Ty*–1) and Illumina–sequenced. Virus–specific PCR and Sanger dideoxy sequencing validations confirmed 15 ssDNA virus/subviral agents in total (occurring mainly in mixed infections), which highlight the potential drawbacks of employing virus-specific resistance in tomato breeding. More

viruses (14 versus six species) were observed in tomatoes without the *Ty*–1 gene. A gemycircularvirus (*Genomoviridae*), a new alpha–satellite, and two novel *Begomovirus* species were identified exclusively in samples without the *Ty*–1 gene, whereas a novel begomovirus was found only in the *Ty*–1 pool. This last virus was the only found inducing severe symptoms in plants carrying the *Ty*–1 gene in our survey. Our results indicated the need for further studies on the potential viral adaptation to *Ty*–1 and its virus-specific filtering effects on a subset of ssDNA viral/subviral agents.
Keywords: begomoviruses, tomato, resistance gene, NGS, virome.

1. Introduction

The *Geminiviridae* is the largest family of plant–infecting viruses with ≈ 468 species described to date, which are currently allocated in nine genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus* (ICTV, 2020). The classification at the genus level is based upon host range, the associated insect vector(s), genomic organization, and phylogenetic relationships (Brown et al., 2015; Varsani et al., 2017; ICTV, 2020). In 2016, two novel viral families with non–enveloped, circular, single-stranded DNA (ssDNA) genomes with size ranging from 2.0 to 2.4 kb were created and named as *Pleolipoviridae* and *Genomoviridae*. The *Genomoviridae* family also comprises nine genera: *Gemycircularvirus*, *Gemyduguivirus*, *Gemygorvirus*, *Gemykibivirus*, *Gemykolovirus*, *Gemykrogvirus*, *Gemykroznavirus*, *Gemytondvirus*, and *Gemyvongvirus* (ICTV, 2020).

The genus *Begomovirus* is composed by whitefly-transmitted species with one (= monopartite) or two (= bipartite) circular, ssDNA genomic component(s) with ≈ 2.6 kb that are encapsulated separately into twinned particles formed by two incomplete icosahedrons (Brown et al., 2015; Rojas et al., 2018). The begomovirus transmission is characterized as being non-propagative, circulative and is carried out by insects members of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex (De Barro et al., 2011). The begomoviruses display a set of mechanisms for generating genetic variability such as mutation, recombination, and pseudo-recombination, which have direct influence in the continuous emergence of new species that are often reported in this genus (Ribeiro et al., 2003; Seal et al., 2006; Silva et al., 2014).

The tomato (*Solanum lycopersicum* L.) crop is grown year-round under distinct cultivation systems across major tropical and subtropical regions (FAO, 2020). In Brazil, outbreaks of *Begomovirus* species in tomatoes become more intensively reported after the invasion of *B. tabaci Middle East-Asia Minor 1* (MEAM 1 = biotype B) in the early 1990s (Ribeiro et al., 1994). The well-known biological attributes of *B. tabaci* MEAM 1 (*viz*. large host range, ability to transmit a wide range of viral species, and adaptation to distinct environmental conditions) facilitated the rapid dispersal of tomato-infecting begomoviruses across all major producing areas of the country (Ribeiro et al., 2003). Field surveys conducted afterward have revealed an extremely diverse complex of *Begomovirus* species (composed mainly by bipartite viruses), occurring in all Brazilian biomes. Currently, over 21 tomato-infecting *Begomovirus* species have been characterized and most of them are already accepted by the International Virus Taxonomy Committee (ICTV). Some of these viruses are listed below in alphabetical order: *Chino del tomate Amazonas virus* – CdTAV, *Tomato bright yellow mosaic virus* – ToBYMV, *Tomato bright yellow mottle virus* – ToBYMoV, *Tomato chlorotic mottle virus* – ToCMoV, *Tomato common mosaic virus* – ToCmMV, *Tomato golden leaf distortion virus* – ToGLDV, *Tomato golden leaf spot virus* – ToGLSV, *Tomato golden mosaic virus* – TGMV, *Tomato golden vein virus* – TGVV, *Tomato interveinal chlorosis virus* – ToICV, *Tomato leaf distortion virus* – ToLDV, *Tomato mild mosaic virus* – ToMMV, *Tomato mottle leaf curl virus* – ToMoLCV, *Tomato rugose mosaic virus* – ToRMV, *Tomato severe rugose virus* – ToSRV, *Tomato yellow spot virus* – ToYSV, *Tomato yellow vein streak virus* – ToYVSV, *Tomato rugose yellow leaf curl virus* – ToRYLCV, *Tomato leaf curl purple vein virus* – ToLCPVV and, more recently, Tomato interveinal chlorosis virus 2 – ToICV2, and *Tomato chlorotic leaf curl virus* – ToCLCV (Quadros et al., 2019; Rêgo-Machado et al., 2019). In addition, begomoviruses initially reported in alternative weed hosts are also occasionally reported infecting tomatoes such as *Sida mottle virus* – SiMoV and *Sida micrantha mosaic virus* – SiMMV (Calegario et al., 2004; Cotrim et al., 2007). Currently, ToSRV (a bipartite species) and ToMoLCV (a monopartite species) are the most widespread and economically important begomoviruses with occurrence reported across all major tomato-producing regions, including Central Brazil. The remaining viral species have an overall more restricted (sometimes endemic) geographic distribution (Inoue-Nagata et al., 2016a).

The preferential strategy for begomovirus management in tomatoes is the employment of cultivars with genetic resistance/tolerance, since the use of insecticides for controlling viruliferous vector populations is neither efficient nor economically and environmentally sustainable (Silva et al., 2009; Yao et al., 2017). Currently, eight resistance/tolerance genes/alleles to begomovirus have been characterized in *Solanum* (section *Lycopersicon*) germplasm: *Ty*–1 (Zamir et al., 1994), *Ty*–2 (Hanson et al., 2006), *Ty*–3 (Ji and Scott, 2006), *Ty*–4 (Ji et al., 2009), *ty*–5 (Anbinder et al., 2009), *Ty*–6 (Hutton and Scott, 2014), *tcm*–1 (Giordano et al., 2005b), and *tgr*–1 (Bian et al., 2007). The *Ty*–1 gene/locus introgressed from *S. chilense* LA 1969 (Zamir et al. 1994) is by far the most employed genetic factor in tomato breeding programs across the globe. In Brazil, cultivars carrying the *Ty*–1 gene have been widely used, mainly across producing regions in Central Brazil (Boiteux et al., 2012a; Boiteux et al., 2012b). The *Ty*–1 gene is located on chromosome 6 in a genomic region in repulsion phase linkage with resistance genes against other pathogens, including the *Mi*–1.2 gene that confers resistance in tomato to the three most important root–knot nematode species: *M. incognita*, *M. javanica*, and *M. arenaria* (Pereira-Carvalho et al., 2010; Verlaan et al., 2011). Molecular markers capable of monitoring the presence of the *Ty*–1 gene/locus in tomato cultivars are now available (Maxwell et al., 2006; Caro et al., 2015; Jung et al., 2015).

The phenotypic expression of the *Ty*–1 gene is best described as a tolerance response (Cooper and Jones, 1983), since plants harboring this factor allow for a mild manifestation of symptoms mainly in the apical meristematic regions, which is followed by a progressive recovery as the plant growth/development advances (Boiteux et al., 2007a). This tolerant reaction is expressed against a relatively large number of monopartite and bipartite begomoviruses and it is related to the inhibition of viral movement, being more efficient under low inoculum conditions (Zamir et al., 1994; Boiteux et al., 2007a). Genetic studies conducted by Verlaan et al. (2011) showed that the *Ty*–1 gene encodes an RNA–dependent RNA polymerase. Therefore, the *Ty*–1 gene is representing an entirely new class of disease resistance/tolerance genes that operates by intensifying the levels of transcriptional silencing of viral genes. More recent studies have shown that the *Ty*–1 gene can also confer resistance to Beet curly top virus (a viral species of the genus *Curtovirus*) in genetically-transformed *Nicotiana benthamiana* plants (Voorburg et al., 2020). However, no information is yet available about the effects of the

Ty–1 gene on ssDNA viruses and subviral agents described in association with tomatoes in the Neotropical areas.

Next-generation sequencing (NGS) technologies have intensified the advances in elucidating many aspects of plant-microbe interactions by enabling the generation of a huge amount of low-cost sequence data of both hosts and pathogens (Knief, 2014). Currently, metagenomic analyses with NGS are the best tools available for large–scale assessment of viral diversity under distinct environmental conditions. NGS has contributed significantly in the sequencing of complete genomes as well as in detecting novel Plant-associated viral species (Adams and Fox, 2016; Hadidi, 2019; Villamor et al., 2019). In this regard, NGS technologies have also contributed to reveal the viral diversity associated with the tomato crop. Illumina sequencing of tomato samples collected in California, Arizona, and Mexico allowed for the detection of *Potato spindle tuber viroid*, *Pepino mosaic virus* as well as a new *Potyvirus* species (Li et al., 2012). NGS analyses of tomato samples in China allowed for the detection of *Polerovirus*, *Orthotospovirus*, *Crinivirus*, *Ilarvirus*, *Cucumovirus*, *Tobamovirus*, *Carlavirus*, *Potyvirus*, *Begomovirus*, and *Amalgavirus* species as well as a new *Cytorhabdovirus* (Xu Chenxi et al., 2017b). A metagenomic analysis with NGS also allowed the identification of a new geminivirus (Tomato associated geminivirus 1 – TaGV1) in association of tomato and *Cleome affinis* plants collected in Central Brazil (Fontenele et al., 2017). More recently, NGS analysis with samples from 132 plant species belonging to 34 botanic families in Mexico detected *Becurtovirus* and *Topocuvirus* species as well as a large diversity of monopartite and bipartite begomoviruses in members of the families Brassicaceae, Convolvulaceae, Curcubitaceae, Euphorbiaceae, Fabaceae, Malvaceae, and Solanaceae (Rodríguez-Negrete et al., 2019).

Due to the extreme variability of tomato-infecting begomoviruses in the Neotropics, it is possible that not yet identified species and strains can be emerging in these areas. The increase in the crop acreage with tomato varieties and hybrids harboring the *Ty*–1 gene may represent a relevant selection factor on viral populations that could make them either more adapted or even capable of entirely overcoming this tolerance factor (Boiteux et al., 2012a; Boiteux et al., 2012b). However, the viral diversity associated with the *Ty*–1 gene and other tomato resistance/tolerance factors were not yet extensively studied. The complete sequence information of the DNA–A and DNA–B genomic segments generated by NGS provides large–scale assessment tools to study viral

population diversity in virtually all tomato–virus pathosystems. In this context, the objective of the present work was to carry out metagenomic analyses aiming to reveal the diversity of *Begomovirus* species as well as other ssDNA viruses and subviral agents in tomato cultivars either lacking or harboring the *Ty*–1 gene in Central Brazil.

2. Materials and Methods

2.1. Tomato leaf samples and confirmation of the presence/absence of the *Ty*–1 gene/locus in the genome of the tomato samples by employing a cleaved amplified polymorphic sequence (CAPS) marker system

Foliar samples of field-grown tomato cultivars/hybrids (with and without the *Ty*–1 tolerance gene) were collected from 2001 to 2016 across three geographic regions (Goiás State – GO, the Federal District – DF, and Minas Gerais State – MG). In order to confirm the presence of the *Ty*–1 gene/locus, we performed PCR assays with the DNA of all the 107 tomato leaf samples. We employed the primer pair UWTyF / UWTyR, which is capable of generating a CAPS marker linked to this tomato genomic region (Maxwell et al., 2006). This codominant marker system is able to discriminate the dominant resistance allele (*Ty*–1) from the susceptible recessive allele (*ty*–1) after cleavage with the restriction enzyme *Taq* I (Maxwell et al., 2006). In order to reveal these alternative alleles for the *Ty*–1 gene/locus, PCR products (amplicons) were cleaved with the enzyme *Taq* I for two hours at a constant temperature of 65°C. The products obtained after cleavage were analyzed in 1% agarose gels, stained in ethidium bromide, and visualized under ultraviolet light.

2.2. Viral isolates and preliminary confirmation of the presence of begomoviruses in the tomato leaf samples

All the tomato samples/isolates (n=107) were collected from plants showing distinct degrees of begomovirus-like symptoms (*viz.* apical and interveinal chlorosis, yellow spots**,** golden mosaic, severe rugose mosaic, apical leaf deformation, and stunting). Each individual sample was subjected to total DNA extraction using a modified (high pH buffer) 2X CTAB + organic solvents protocol (Boiteux et al., 1999). These

samples/isolates were stored at -20 °C and they currently comprise a section of the begomovirus collection of the Plant Breeding Laboratory at CNPH (Brasília-DF, Brazil). The purified total DNA was subjected to polymerase chain reaction (PCR) assays aiming to confirm the presence of begomovirus (es) in these tomato leaf samples. Amplicons derived from a segment of the DNA – A component were obtained using the 'universal' primer pairs 'PAL1v1978/ PAR1c496' (Rojas et al., 1993) and 'BegomoAFor1' / 'BegomoARev1' (Ha et al., 2006), which produce two large and non-overlapping segments (≈ 1120 bp and ≈ 1205 bp, respectively). Amplicons derived from a segment of the DNA – B component (≈ 690 bp) were obtained using the 'universal' primer pair 'PBL1v2040'/ 'PCRc1'(Rojas et al., 1993). The obtained amplicons were analyzed in 1% agarose gels, stained in ethidium bromide, and visualized under ultraviolet light. Only samples displaying begomovirus-derived amplicons were selected for a subsequent enrichment of circular DNAs via rolling circle amplification and for next-generation sequencing-NGS (see sections below).

2.3. Enrichment via rolling circle amplification of circular DNA molecules on each individual sample

The virus-derived circular DNA molecules in the samples were selectively enriched by rolling circle amplification (RCA) assays (Inoue-Nagata et al., 2004). After DNA analysis on agarose gel and via NanoVue Plus®, the concentrations were adjusted to 1 microgram per sample and then used to make up the two pools. The CAPS-characterized samples were then subdivided into two pools: one composed by DNAs of tomato plants without the *Ty*–1 **(Table 1)** gene (n=64) and one composed by DNAs of tomato samples with the *Ty*–1 **(Table 2)** gene (n=43).

Table 1. Identification of 64 samples (= isolates) exhibiting begomovirus–like symptoms that were obtained from tomato plants without the *Ty*–1 gene/locus in Central Brazil. Information is provided about the region where the isolate was collected, year of collection, and the respective isolate code.

Geographic region	Year of collection	Isolate code
Goiás State-GO	2003	GO–023, GO–046, GO–109, GO–111, GO–118, GO–120, GO–130, GO–134, GO–136, GO–137, GO–142, GO–143, GO–144, GO–168, GO–169, GO–191, GO–192, GO–221, GO–244, GO–245, GO–248, GO–249, GO–250, GO–251.
	2004	GO–298, GO–299, GO–301, GO–322, GO–336.
	2006	GO–384, GO–390.
	2011	GO–493.
	2012	GO–505, GO–511.
	2015	GO–594.
Federal District-DF	2003	DF–018, DF–023, DF–028, DF–043, DF–045, DF–046, DF–050, DF–062.
	2005	DF–166, DF–167, DF–211.
	2010	DF–330.
	2011	DF–447, DF–453.
	2013	DF–544.

	2014	DF–566.
	2016	DF–667.
Minas Gerais State-MG	2001	MG–046.
	2002	MG–012, MG–015, MG–016, MG–018, MG–029.
	2010	MG–073, MG–113, MG–150.
	2012	MG–325.
	2015	MG–378, MG–388.

Table 2. Identification of 43 samples (= isolates) exhibiting begomovirus–like symptoms that were obtained from tomato plants harboring the *Ty*–1 gene/locus in Central Brazil. Information is provided about the region where the isolate was collected, year of collection, and the respective isolate code.

Geographic region	Year of collection	Isolate code
Goiás State-GO	2003	GO–145, GO–148, GO–149, GO–151, GO–157, GO–161, GO–164.
	2004	GO–247, GO–305, GO–307, GO–308, GO–320, GO–326, GO–330.
	2007	GO–371.
	2010	GO–479, GO–487, GO–490.
	2013	GO–550, GO–582, GO–583.
Federal District-DF	2007	DF–227, DF–236, DF–238.
	2008	DF–252.
	2010	DF–339.
	2011	DF–438.
	2013	DF–529, DF–550, DF–556.
	2016	DF–640.
Minas Gerais State-MG	2010	MG–092, MG–122, MG–169, MG–282, MG–283, MG–284, MG–285, MG–286, MG–287.
	2012	MG–326.
	2015	MG–383, MG–387.

2.4. Next-generation sequencing (NGS) of the two tomato DNA pools and analysis of the NGS-derived sequences

The sample pools (with and without *Ty*–1 gene) were subjected to high performance sequencing in an Illumina platform with the HiSeq 2500 system (Macrogen Inc., South Korea). The NGS–derived sequences were analyzed according to the following workflow: **(1)** elimination of low–quality reads; **(2)** re-assembly of the sequences using the program CLC Genomics Workbench 10; and **(3)** validation of the contigs via BLASTx and BLASTn algorithms by comparing with the ssDNA virus database of the GenBank (https://www.ncbi.nlm.nih.gov/). The viral contigs were annotated and the trimmed reads were mapped back to the annotated genome using the tool 'Map to reference' available in the Geneious 11.0 program (Kearse et al., 2012). The conserved regions/motifs present in the begomovirus genomes such as: nonanucleotide, TATA box, stem loop, and iterons were also selectively analyzed (Argüello-Astorga and Ruiz-Medrano, 2001). Additionally, individual identification of the viruses was obtained in the NGS–derived dataset by using the SeqMan NGene Metagenomic sequence analysis software (DNAStar, Madison, WI). Viral contigs were analyzed against the RefSeq viral database (NCBI) at a very high stringency conditions (minimum match percentage = 99%).

2.5. Design of a collection of viral species-specific PCR primers for detection in individual samples

For the confirmation of the viral species detected in each individual sample, specific PCR primers (for both DNA–A and DNA–B genomic segments) were designed in opposite and overlapping directions. Primer design was carried out based upon the consensus contigs obtained with the Geneious 11.0 program **(Table 3)**. Virus-specificity of the primers was double–checked *in silico* by using the Primer-Blast tool and in preliminary PCR assays using as template DNA samples from a reference collection of the NGS-identified viral isolates.

2.6. Validation of NGS-derived information via PCR assays with virus-specific primers

PCR assays with the previously selected virus-specific primers **(Table 3)** were carried with in all 107 individual DNA samples. These assays were performed in order to validate the NGS results. PCR were carried with a total volume of 12.5 μL, containing 1.25 μL of *Taq* polymerase buffer (100 mM Tris–HCl, pH 8.3 and 500 mM KCl), 0.40 μL $MgCl_2$ (50 mM), 0.25 μL, dNTPs (2.5 mM), 0.25 μL of each primer (forward and reverse) (10 μM), 2 μL of DNA, 8.0 μL of Milli-Q® water (Millipore, Bedford, MA, USA) and and 0,1 μL *Taq* DNA polymerase (5 U/μL). The reactions were amplified in thermal cycler (Bio-Rad Laboratories, Hercules, CA) programmed for 35 cycles with the following conditions: initial denaturation: 94 °C for 3 minutes, denaturation: 94 °C for 30 seconds, annealing (ranging from 46 to 60 °C, according to the primer pair employed; **Table 3**) for 45 seconds, extension 72 °C for 3 minutes and final extension 72 °C for 10 minutes. The begomovirus-derived amplicons were observed to 1.5% agarose–gel electrophoresis stained with ethidium bromide and visualized under UV–light.

2.7. Sanger dideoxy sequencing validation of virus-specific PCR amplicons

Direct Sanger dideoxy sequencing reactions of positive virus–derived amplicons were carried out to double check the viral diversity observed in a subset of individual samples. Sequencing reactions were performed at the Genomic Analysis Laboratory (at CNPH), employing the same virus-specific primer pairs (**Table 3**) in one ABI PRISM 3130 sequencer using the BigDye® Terminator Cycle Sequencing Ready Reaction Kit version 3.1 protocol (Applied Biosystems, São Paulo-SP, Brazil). After contig assembling and quality evaluation, the obtained sequences were analyzed using the BLASTn algorithm. This tool was used to compare our sequences with the ones retrieved from the GenBank - NCBI public database (https://www.ncbi.nlm.nih.gov/), aiming to verify the sample–associated viral species. We adopted the current pairwise identities of 91% and 94% as the demarcation thresholds to identify *Begomovirus* species and strains, respectively (Brown et al., 2015).

Table 3. PCR primer pairs designed based upon Next-generation Sequencing (NGS)-derived viral consensus sequences for validation of the *Begomovirus* species as well as single-stranded DNA viruses and subviral agents identified in the tomato DNA sample pools (with the *Ty*–1 gene versus without the *Ty*–1 gene). For = forward direction and Rev = reverse direction.

Viral Species	**Primer Name**	**Sequence 5'–3'**	**Annealing temperature (T°C)**
Bean golden mosaic virus (BGMV) DNA–A	BGMV–For	GTGCGTGAATCCATGACCGT	55
	BGMV–Rev	ATTCACGCACAGGGGAACG	
Cleome leaf crumple virus (CILCrV) DNA–A	CILCrV–A–For	GACTCGACGTTCTGTGGT	51
	CILCrV–A–Rev	TCCTAGTCGGGGCTCACT	
Cleome leaf crumple virus (CILCrV) DNA–B	CILCrV–B–For	TAGGAAAGCAAAACGAGAATGGAA	58
	CILCrV–B–Rev	GCTTTCCTAAATCGCAATTGATC	
Tomato severe rugose virus (ToSRV) DNA–A	ToSRV–For5.1	AGCGTCGTTAGCTGTCTGGCA	58
	ToSRV–Rev5	TGCCGCAGAAGCCTTGAACGCACCT	
Tomato severe rugose virus (ToSRV) DNA–B	ToSRV–B–For	AAACCCACACGAAAGCAGAGTTT	55
	ToSRV–B–Rev	CACCACGTCTATACATATTGTCCAGG	

Euphorbia yellow mosaic virus (EuYMV) DNA–A	EuYMV–A–R–For	GGGGTTCCAAGTCCAATAAAGATGA	52
	EuYMV–A–R–Rev	CAGACACCTTATATTTGCCGGATTC	
Euphorbia yellow mosaic virus (EuYMV) DNA–B	EuYMV–B–R–For	GCCGAGGATAGAGGACACCAA	60
	EuYMV–B–R–Rev	CCAGGCCCAAACGCATTATATTTTATC	
Tomato chlorotic mottle virus (ToCMoV) DNA–A	ToCMoV–A–For	TTTGGGCCGCTCTTTTGGG	47
	ToCMoV–A–Rev	CAAACTGAATGGGCCTTAAA	
Tomato chlorotic mottle virus (ToCMoV) DNA–B	ToCMoV–B–For	GTATTTGTTCTGGGTGCAATCATAAAAC	55
	ToCMoV–B–Rev	TTGTACTAATGACACATTATTCAATCACGA	
Tomato golden vein virus (TGVV) DNA–A	TGVV–A–For1	AAAGGAAGATAATTCAAATATAGGGA	51
	TGVV–A–Rev1	ATCTTCCTTTACTCACGTTCCTGAT	
Tomato golden vein virus (TGVV) DNA–B	TGVV–B–S–For	CCCACTTTCCATAACCTACATGAGA	55
	TGVV–B–S–For	GGAGAGAAAATTGATAAGATCGGCATC	
Tomato mottle leaf curl virus (ToMLCV) DNA–A	ToMoLCV–For	TGTGGTCCAGTCAATAAATG	47

	ToMoLCV–Rev	TGACTGGACCACATAGTAAA	
Tomato common mosaic virus (ToCmMV) DNA–A	ToCmMV–For1	ATTGCTCTCAACTTCTGTGC	54
	ToCmMV–Rev2	GCAATCCCTGGTGTCCTCAC	
Tomato rugose mosaic virus (ToRMV) DNA–A	ToRMV–A–For	TGAAAGTAATTTTGACCCAATC	52
	ToRMV–A–Rev	CAATTCATATGAGTTTTAGAGCAGC	
Sida micrantha mosaic virus (SiMMV) DNA–A	SiMMV–For	GATCTCGCTCCCCCCTCT	58
	SiMMV–Rev	AGATCGCACGACAACCAG	
Plant-associated genomovirus 2	Gemy–For	GCTCTGAATCAAATCTCGCTTACTTG	54
	Gemy–Rev	CGATGTTGATTGGTTGGAAGCAAA	
New *Begomovirus* Species #1 DNA–A	DF–640–A–For	GTTGACTGACATTTGCCTT	47
	DF–640–A–Rev	TGTCAGTCAACAATCTATACACA	
New *Begomovirus* Species #1 DNA–B	DF–640–B–For	GTTGTTTCAAGGGCGTCGAC	55
	DF–640–B–Rev	CAACATCAGACATCCAGCAATAATAAACT	
New *Begomovirus* Species #2 DNA–A	1ToBYMV–A–For	ATCCATGTCCTCGGCAGTCT	55

	1ToBYMV–A–Rev	TCACGCACAGAGGAACGC	
New *Begomovirus* Species #3 DNA–A	Abuti–A–For	GGACTCCAGGGGGCAAAA	55
	Abuti–A–Rev	AGTCCCGTCCGTACCACTTG	
Alpha–satellite	Alfa–For	TGGTGTCCTGGCTTATAT	46
	Alfa–Rev	GGCGGAGTCCTTTTTTTT	

3. Results

3.1. NGS detection of previously reported *Begomovirus* species in the two pools of samples (with and without the *Ty*–1 gene)

The total number of reads per viral species/genomic component obtained in the pool without the *Ty*–1 gene is presented in **Table 4.** The total number of reads per viral species/genomic component obtained in pool from plants with the *Ty*–1 gene is presented in **Table 5**. After assembly, 19,487 contigs were obtained in the pool without the *Ty*–1 gene and 7,045 contigs in the sample pool with the *Ty*–1 gene. Even though with a slightly different number of evaluated samples in the pools with (n=43) and without (n=64) the *Ty*–1 gene, BLASTn analyses of the Illumina HiSeq 2500 sequencing against a reference GenBank collection of ssDNA viruses revealed a greater diversity of viral species in the pool of tomato samples lacking the *Ty*–1 gene. Ten begomoviruses were found in the pool without the *Ty*–1 gene *viz*. Bean golden mosaic virus – BGMV (only DNA–A was recovered; GenBank MT214083**),** Cleome leaf crumple virus – ClLCrV (both DNA–A and DNA–B were recovered; MN337873 and MN337872, respectively), Tomato severe rugose virus – ToSRV (DNA–A and DNA–B; MT214084 and MT214085, respectively), Euphorbia yellow mosaic virus – EuYMV (DNA–A and DNA–B; MN746971 and MN746970, respectively), Sida micrantha mosaic virus – SiMMV (only DNA–A; MT214092), Tomato chlorotic mottle virus – ToCMoV (DNA–A and DNA–B; MT214086 and MT214087, respectively), Tomato golden vein virus – TGVV (DNA–A and DNA–B; MN928610 and MN928611, respectively), Tomato mottle leaf curl virus – ToMoLCV (only DNA–A was recovered, confirming it as a monopartite species; MT214088), Tomato common mosaic virus – ToCmMV (only the DNA–A of this bipartite species was recovered; MT214089) and Tomato rugose mosaic virus – ToRMV (DNA–A and DNA–B; MT214090 and MT214091, respectively) **(Table 4).** In the pool harboring the *Ty*–1 gene, four previously described bipartite *Begomovirus* species were recovered with both DNA–A and DNA–B components *viz*. ToSRV (DNA–A: MT215001; DNA–B: MT215002), ToCMoV (DNA–A: MT215003; DNA–B: MT215004), TGVV (DNA–A: MN928612; DNA–B: MN928613), and ToRMV (DNA–A: MT215006; DNA–B: MT215007). The DNA–A component (MT215005) of the monopartite species ToMoLCV was also recovered. ToSRV and ToRMV displayed the two highest numbers of reads, indicating their relative predominance in the tomato samples with the *Ty*–1 gene. Some of the Neotropical tomato-infecting *Begomovirus*

species (included on the RefSeq database) displayed overall high identity levels (e.g. > 97% identity in the case of the DNA–B component that is shared by the species ToSRV and ToRMV). This implies that some of our reads **(Tables 4 and 5)** were most likely counted more than once. For this reason, virus identification was double-checked using SeqMan NGene with high stringency parameters (99%). The validation of the NGS results via PCR assays with virus-specific primers coupled with Sanger dideoxy sequencing was also a very important tool to verify the presence of each individual virus species described here.

Table 4. Viral circular, single-stranded DNA species detected after Illumina Hiseq sequencing in the pool of tomato DNA samples lacking the *Ty*–1 gene.

Viral Species	**N° of *reads****	**Size (nts)**
Bean golden mosaic virus (BGMV) DNA–A	63,525	2.626
Cleome leaf crumple virus (CILCrV) DNA–A	566	2.560
Cleome leaf crumple virus (CILCrV) DNA–B	702	2.664
Tomato severe rugose virus (ToSRV) DNA–A	3,225,120	2.593
Tomato severe rugose virus (ToSRV) DNA–B	4,018,351	2.572
Euphorbia yellow mosaic virus (EuYMV) DNA–A	1,122	2.609
Euphorbia yellow mosaic virus (EuYMV) DNA–B	1,822	2.579
Tomato chlorotic mottle virus (ToCMoV) DNA–A	5,971,019	2.620
Tomato chlorotic mottle virus (ToCMoV) DNA–B	1,111,227	2.600
Tomato golden vein virus (TGVV) DNA–A	2,639,961	2.562
Tomato golden vein virus (TGVV) DNA–B	977,027	2.512
Tomato mottle leaf curl virus (ToMLCV) DNA–A	1,784,881	2.632
Tomato common mosaic virus (ToCmMV) DNA–A	1,070,674	2.560
Tomato rugose mosaic virus (ToRMV) DNA–A	3,267,808	2.619
Tomato rugose mosaic virus (ToRMV) DNA–B	4,742,730	2.571

Sida micrantha mosaic virus (SiMMV) DNA–A	1,221,062	2.688
Plant-associated genomovirus 2	119	2.189
New *Begomovirus* Species #2 DNA–A	427,646	2.649
New *Begomovirus* Species #3 DNA–A	2,839	2.636
New Alpha–satellite	155,793	1.321

Table 5. Viral circular, single-stranded DNA species detected after Illumina Hiseq sequencing in the pool of tomato DNA samples harboring the *Ty*–1 gene.

Viral Species	**N° of *reads***	**Size (nts)**
Tomato severe rugose virus (ToSRV) DNA–A	7,181,771	2.592
Tomato severe rugose virus (ToSRV) DNA–B	5,782,296	2.570
Tomato golden vein virus (TGVV) DNA–A	2,358,838	2.561
Tomato golden vein virus (TGVV) DNA–B	1,401,684	2.590
Tomato chlorotic mottle virus (ToCMoV) DNA–A	4,519,040	2.623
Tomato chlorotic mottle virus (ToCMoV) DNA–B	811,733	2.565
Tomato mottle leaf curl virus (ToMLCV) DNA–A	2,644,606	2.631
Tomato rugose mosaic virus (ToRMV) DNA–A	7,964,942	2.618
Tomato rugose mosaic virus (ToRMV) DNA–B	5,780,864	2.649
New *Begomovirus* Species #1 DNA–A	1,270,494	2.605
New *Begomovirus* Species #1 DNA–B	84,022	2.603

3.2. NGS detection of putative three novel *Begomovirus* species as well as a new alpha–satellite species and a *Gemycircularvirus* (*Genomoviridae*) in the tomato samples

We were also able to identify three putative new *Begomovirus* species (one in the pool with the *Ty*–1 gene and two species in the pool without the *Ty*–1 gene). In the pool of samples with the *Ty*–1 gene, a putative new virus (named here as species #1 = isolate DF–640) displayed a bipartite genome organization having a DNA–A component with 2,605 nts and a DNA–B component with 2,603 nts (GenBank DNA–A: MT215017 and DNA–B: MT215018). The putative new species #1 displayed the highest level of identity (85%) with Tomato rugose yellow leaf curl virus (TRYLCV) isolates. The isolate DF–640 was recovered from a field-grown tomato plant in vicinities of Gama city (in the Federal District) with severe symptoms, indicating a putative increase in virulence in relation to the *Ty*–1 gene. Two putative new species were detected in the pool lacking the *Ty*–1 gene. The first one was tentatively named here as new species #2 (= isolate MG–378) and displayed only the DNA–A component with 2,649 nts (GenBank MT214095), Tomato bright yellow mottle virus (ToBYMoV) displayed the highest identity level (84%) to the new species #2. Additional PCR assays were carried out using the isolate MG–378 as template, but no amplicon for the putative cognate DNA–B component was recovered (data not shown), indicating that it is more likely a monopartite virus. The new species # 3 is more likely also a monopartite begomovirus with a DNA–A genome with 2,636 nts (GenBank MT214096). The new species # 3 displayed the highest identity level (84%) to *Abutilon mosaic Brazil virus* (AbMV). In addition, a novel alpha–satellite species (MT214093) and a gemycircularvirus (Family: *Genomoviridae*) (MT214094) species were also detected exclusively in the pool of tomato samples without the *Ty*–1 gene **(Table 4).**

3.3. Confirmation via PCR assays with virus-specific primers and Sanger dideoxy sequencing of the viral and subviral ssDNA species present in each individual tomato sample and quantification of mixed infections

After carrying out PCR assays with virus-specific primers (**Table 3**) and Sanger sequencing, it was possible to catalog all the viral and subviral ssDNA species present in each individual tomato sample comprising the two pools. In the samples of the pool without the *Ty*–1 gene, it was possible to confirm the presence of all *Begomovirus* species

reported initially by the analyses of the NGS-derived results **(Table 6)**. ToSRV was the most prevalent begomovirus, mainly in samples from Goiás-GO State. In the samples of the pool with the *Ty*–1 gene, all species identified after NGS analyses were also confirmed via PCR assays with virus-specific primers **(Table 7)**. In addition, it is important to highlight that the majority of the samples displayed mixed infections with two to five viral species being simultaneously detected in a single tomato plant **(Figures 1 and 2)**.

Table 6. Relative frequency of begomovirus and other circular single-stranded DNA viruses detected after Illumina Hiseq sequencing of 63 tomato DNA samples lacking the *Ty*–1 gene.

Viral species* followed by the respective number of occurrences in each region	**Goiás State-GO**	**Federal District-DF**	**Minas Gerais State-MG**
ToSRV (32+9+5) = 46	GO–046, GO–109, GO–118, GO–120, GO–130, GO–134, GO–136, GO–137, GO–142, GO–143, GO–144, GO–168, GO–169, GO–191, GO–192, GO–221, GO–244, GO–245, GO–248, GO–249, GO–250, GO–251, GO–298, GO–299, GO–301, GO–322, GO–336, GO–390, GO–493, GO–505, GO–511, GO–594.	DF–043, DF–166, DF–167, DF–211, DF–447, DF–453, DF–544, DF–566, DF–667.	MG–012, MG–018, MG–029, MG–150, MG–388.
TGVV (23+8+5) = 36	GO–046, GO–109, GO–130, GO–134, GO–137, GO–142, GO–143, GO–168, GO–169, GO–191, GO–192, GO–221, GO–244, GO–245, GO–248, GO–249, GO–250, GO–	DF–023, DF–028, DF–045, DF–046, DF–050, DF–062, DF–167, DF–211.	MG–015, MG–016, MG–018, MG–029, MG–046.

	298, GO–299, GO–301, GO–322, GO–336, GO–493.		
ToCMoV **(21+8+5) = 34**	GO–023, GO–046, GO–109, GO–111, GO–120, GO–130, GO–134, GO–136, GO–137, GO–143, GO–144, GO–191, GO–245, GO–249, GO–250, GO–251, GO–298, GO–299, GO–301, GO–322, GO–390.	DF–018, DF–028, DF–043, DF–045, DF–046, DF–050, DF–167, DF–566.	MG–015, MG–018, MG–046, MG–073, MG–150.
ToCmMV **(1+0+1) = 2**	GO–023.	---	MG–388.
BGMV **(1+2+0) = 3**	GO–142.	DF–045, DF–046.	---
CILCrV **(0+0+1) = 1**	---	---	MG–150
EuYMV **(0+0+2) = 2**	---	---	MG–012, MG–016.
ToMLCV **(4+8+1) = 13**	GO–299, GO–384, GO–505, GO–594.	DF–018, DF–023, DF–028, DF–050, DF–062, DF–330, DF–453, DF–566.	MG–325.
ToRMV **(19+1+1) = 21**	GO–109, GO–118, GO–130, GO–134, GO–136, GO–137, GO–143, GO–144, GO–168, GO–191, GO–192, GO–244, GO–248, GO–250, GO–251, GO–298, GO–322, GO–336, GO–505.	DF–043.	MG–150.

SiMMV **(8+3+0) = 11**	GO–118, GO–120, GO–134, GO–168, GO–245, GO–248, GO–301, GO–511.	DF–045, DF–050, DF–166.	---
***Plant-associated genomovirus* 2** **(2+0+0) = 2**	GO–298, GO–301.	---	---
Alpha–satellite **(0+4+0) = 4**	---	DF–023, DF–028, DF–050, DF–062.	---
New *Begomovirus* species #2 **(0+0+1) = 1**	---	---	MG–378.
New *Begomovirus* species #3 **(1+0+0) = 1**	GO–169.	---	---

*ToSRV = *Tomato severe rugose virus*, TGVV = *Tomato golden vein virus*, ToCMoV = *Tomato chlorotic mottle virus*, ToCmMV = *Tomato common mosaic virus*, BGMV = *Bean golden mosaic virus*, CILCrV = *Cleome leaf crumple virus*, EuYMV = *Euphorbia yellow mosaic virus*, ToMLCV = *Tomato mottle leaf curl virus*, ToRMV = Tomato rugose mosaic virus, and SiMMV = *Sida micrantha mosaic virus*.

Table 7. Relative frequency of begomovirus and other circular single-stranded DNA viruses in association with 43 tomato DNA samples harboring the *Ty*–1 gene detected after Illumina Hiseq sequencing.

Viral species* followed by the respective number of occurrences in each region	**Goiás State-GO**	**Federal District-DF**	**Minas Gerais State-MG**
ToSRV (14+5+7) = 26	GO–145, GO–148, GO–151, GO–157, GO–161, GO–164, GO–247, GO–330, GO–371, GO–479, GO–487, GO–490, GO–550, GO–582.	DF–236, DF–339, DF–438, DF–550, DF–556.	MG–169, MG–285, MG–286, MG–287, MG–326, MG–383, MG–387.
TGVV (12+3+0) = 15	GO–145, GO–148, GO–149, GO–151, GO–305, GO–320, GO–326, GO–371, GO–479, GO–490, GO–582, GO–583.	DF–236, DF–238, DF–438.	---
ToCMoV (12+1+9) = 22	GO–145, GO–148, GO–149, GO–305, GO–320, GO–326, GO–330, GO–371, GO–479, GO–490, GO–582, GO–583.	DF–550.	MG–092, MG–122, MG–169, MG–282, MG–283, MG–284, MG–285, MG–286, MG–287.
ToMLCV (4+8+1) = 13	GO–307, GO–320, GO–326, GO–582.	DF–227, DF–236, DF–252, DF–339, DF–438, DF–529, DF–550, DF–556.	MG–326.
ToRMV (13+1+8) = 22	GO–145, GO–148, GO–149, GO–151, GO–157, GO–161, GO–164, GO–247, GO–307, GO–308,	DF–227.	MG–092, MG–122, MG–169, MG–282, MG–283, MG–284, MG–285, MG–286.

	GO–320, GO–330, GO–479.		
New *Begomovirus* species #1 **(0+1+0) = 1**	---	DF–640.	---

*ToSRV = *Tomato severe rugose virus*, TGVV = *Tomato golden vein virus*, ToCMoV = *Tomato chlorotic mottle virus*, and ToRMV = *Tomato rugose mosaic virus*.

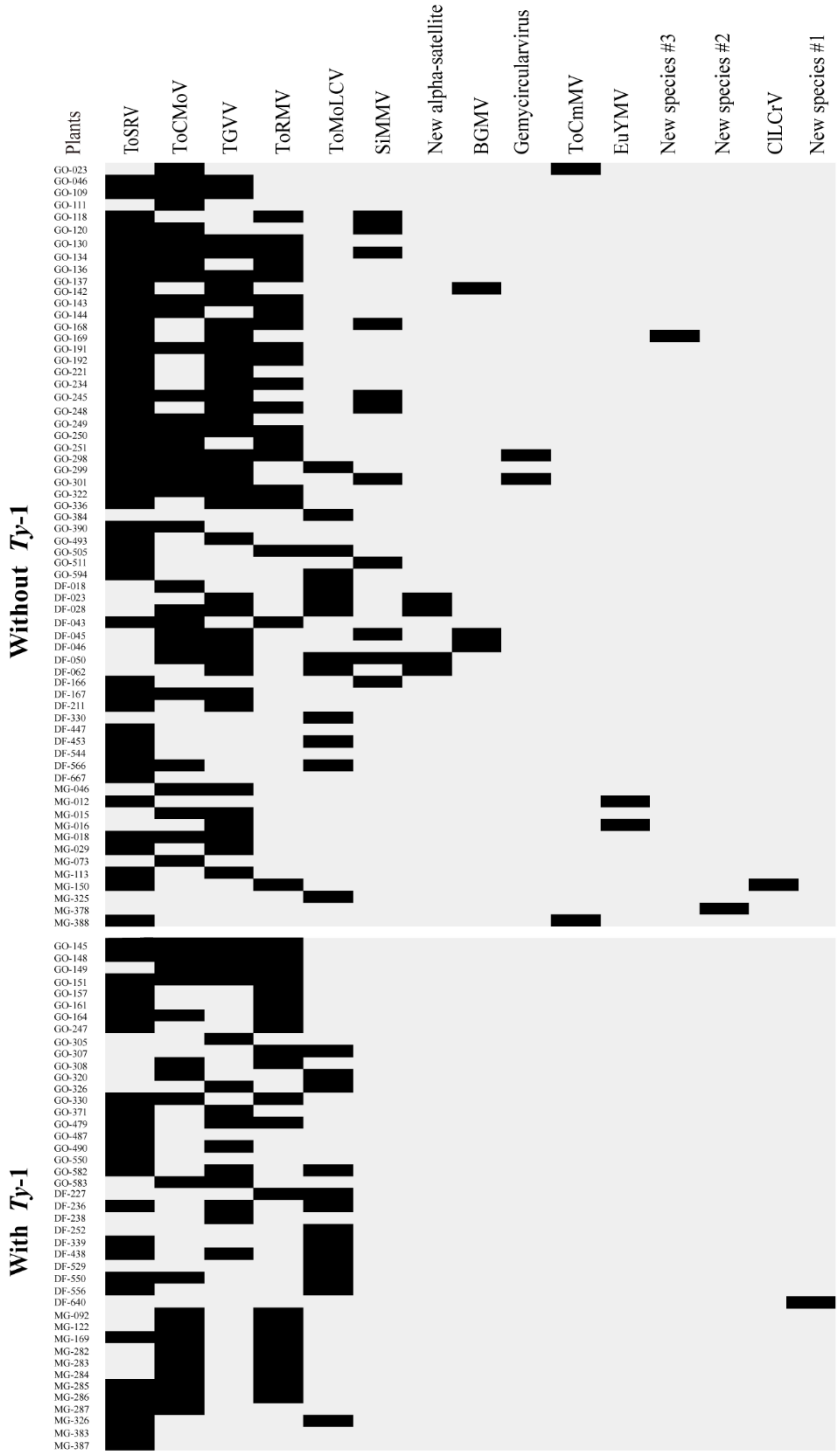

Figure 1. Frequency and relative predominance of *Begomovirus* species and single-stranded DNA (ssDNA) viruses detected with via Illumina Hiseq sequencing of tomato samples with (n=43) and without (n=64) the *Ty*–1 gene. Results were validated by PCR assays with virus-specific primers and by Sanger dideoxy sequencing. Viruses detected: Tomato severe rugose virus (ToSRV); Tomato golden vein virus (TGVV); Tomato chlorotic mottle virus (ToCMoV); Tomato rugose mosaic virus (ToRMV); Tomato mottle leaf curl virus (ToMoLCV); Sida micrantha

mosaic virus (SiMMV); Bean golden mosaic virus (BGMV); Tomato common mosaic virus (ToCmMV); Euphorbia yellow mosaic virus (EuYMV) and Cleome leaf crumple virus (CILCrV). A new alpha-satellite species and three putative novel *Beomovirus* species (= New species #1, New species #2, and New species #3) were also detected. Black bars in each line are indicating the presence of a given virus in a given individual sample = isolates (left column). Isolate with **GO** abbreviation = isolates collected in Goiás State; **DF** abbreviation = isolates collected in the Federal District and **MG** abbreviation = isolates collected in Minas Gerais State, in Central Brazil.

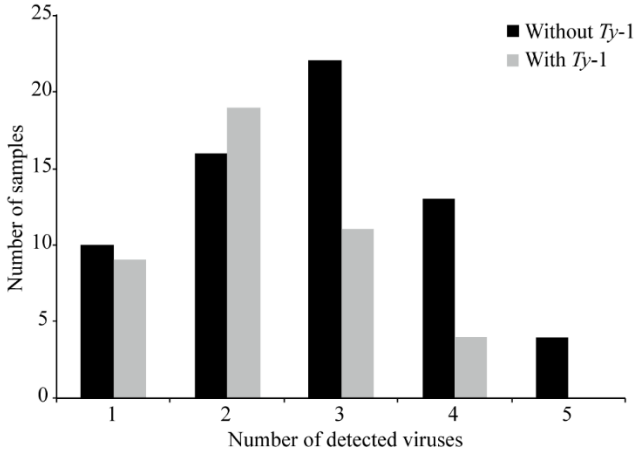

Figure 2. Number of samples displaying single and mixed (ranging from two to five viruses per sample) infections with *Begomovirus* species and single-stranded DNA (ssDNA) viruses detected with Illumina Hiseq sequencing of tomato samples with (n=43) and without (n=64) the *Ty*–1 gene. Results were validated by PCR assays with virus-specific primers and by Sanger dideoxy sequencing.

4. Discussion

Over 286 viral species have been reported infecting tomatoes worldwide (Virus-HostDB, 2020). In Brazil, the tomato crop is also affected by several virus–induced diseases of great economic importance (Inoue-Nagata et al., 2016b). Diseases caused by begomoviruses are among the most important ones for the tomato crop in the country, mainly due to the widespread presence of their very efficient vector: *B. tabaci* MEAM 1 (De Barro et al., 2011; Rosen et al., 2015). A large number of surveys have been carried out in tomato fields across many Brazilian regions after the introduction of *B. tabaci* MEAM 1 in the early 1990s and they are revealing the presence of an extremely diverse

complex of *Begomovirus* species. Currently, over 20 begomoviruses have been described infecting tomatoes under natural conditions (Inoue-Nagata et al., 2016a).

The emergence *per se* of a large number of novel viral species is somewhat expected since the begomoviruses display a well–known set of mechanisms for generating genetic variability such as mutation, recombination, and pseudo-recombination (Ribeiro et al., 2003; Seal et al., 2006; Sahu et al., 2018). The scenario of immense begomovirus variability in the Neotropics favors the emergence of new species, which can be intensified by the frequent occurrence of mixed infections. However, there is a surprisingly scarce amount of information quantifying the frequency of mixed infections of tomato plants by members of the Neotropical *Begomovirus* species complex under natural conditions. Our NGS-derived results displayed a substantial number of the tomato samples with events of co–infection in both pools (with and without the *Ty*–1 gene). The simultaneous presence of distinct virus species detected in single plants ranged from two to up to five (**Figures 1 and 2**). However, it is interesting to highlight that the *Ty*–1 gene did not have a significant impact on reducing the overall number of multiple viral infections, since samples with this genetic factor displayed non–significant differences when compared to samples without this gene (chi–square test = 6.5193; *p*–value = 0.1635, which was found to be not significant at $p < 0.05$).

In our study, in addition to the detection of *Begomovirus* species already reported in the Neotropics, it was possible to detect two putative new *Begomovirus* species in the samples without the *Ty*–1 and one novel *Begomovirus* species in a sample with the *Ty*–1 gene. The putative new species # 1 (DF–640) displays all typical features of the New World bipartite begomoviruses, having a DNA–A with a size of 2,605 nts and a DNA–B component with 2,603 nts. The new species # 1 displayed the highest identity level (85%) with *Tomato rugose yellow leaf curl virus* (TRYLCV). Only the DNA–A components were found in the putative new species # 2 (= isolate MG–378) and in the new species # 3 (= isolate GO–169), suggesting that both might be novel monopartite viruses. The new species # 2 (2,649 nts) displayed the highest identity (84%) with the *Tomato bright yellow mottle virus* (ToBYMoV) and the new species # 3 (2,636 nts) displayed the highest identity level (84%) with *Abutilon mosaic Brazil virus* (AbMBV). According to the current criteria for species demarcation in the genus *Begomovirus*, nucleotide identities of the DNA–A component that are less than 91% with the complete DNA–A genome of any other known begomovirus sequence will correspond to a new species (Brown et al.,

2015). The overall low number of samples detected with these putative new begomoviruses indicates that they may represent extremely rare emergence events of novel viral variants. Therefore, it is most likely that we were able to identify these emerging viruses here due to the enhanced analytical power provided by the NGS technology.

A putative novel alpha–satellite (with 1,321 nts) was also detected in four isolates (DF–023, DF–028, DF–050, and DF–062) collected in distinct areas of the Federal District in plants lacking the *Ty*–1 gene. Alpha–satellite DNA molecules are subviral agents classified in the family Alphasatellitidae that have been found in association with *Begomovirus* (Briddon et al., 2018; ICTV, 2020). Here, alpha–satellite isolates were found in samples with mixed infections with distinct viral species *viz*. isolate DF–023 (mixed with TGVV and ToMLCV), DF–028 (mixed with TGVV, ToCMoV, and ToMLCV), isolate DF–050 (mixed infection with TGVV, ToCMoV, ToMLCV, and SiMMV), and isolate DF–062 (mixed infection with TGVV and ToMLCV). The genera of alpha–satellites associated with the geminiviruses are found in the subfamily *Geminialphasatellitinae*, genus *Ageyesisatellite*, *Clecrusatellite*, *Colecusatellite* and *Gosmusatellite*. Nucleotide identity less than 88% (in comparison with complete sequences of the known alpha–satellites) is the criterion currently used for the classification of a new species within the family *Geminialphasatellitinae* (Briddon et al., 2018; ICTV, 2020). The alpha–satellite isolates found in the present study showed the highest level of identity (81%) with other New World species that were found in association with bipartite begomoviruses in Brazil, Cuba, and Venezuela (Paprotka et al., 2010; Romay et al., 2010). Thus, according to the demarcation within the subfamily, the alpha–satellite is more likely a new species, probably of the genus *Clecrusatellite*, which is composed by species found in association with bipartite *Begomovirus* from the New World (Paprotka et al., 2010; Romay et al., 2010; ICTV, 2020). All four alpha–satellite isolates reported here were found in constant association with two begomoviruses (TGVV and ToMLCV). Therefore, additional bioassays will be necessary to identify which associated *Begomovirus* species is able to transreplicate this novel alpha–satellite.

A Plant-associated genomovirus 12, classified into the genus *Gemycircularvirus* (family *Genomoviridae*), was also detected in two tomato samples from the pool without the *Ty*–1 gene (isolates GO–298 and GO–301). Both isolates were collected in Leopoldo de Bulhões, Goiás-GO State in 2004. These isolates displayed 98% identity to Capybara

genomovirus 9 isolate cap1_561 (MK483081.1) from Brazil. The gemycircularviruses have ssDNA and some species of this genus have been reported in association with plants (Male et al., 2015; Marzano and Domier, 2016). Studies with these viral species are recent, since the *Genomovoridae* family was only established in 2016 by the ICTV (Krupovic et al., 2016). In Brazil, two gemycircularviruses (Momordica charantia–associated gemycircularvirus – MoaGmV and Euphorbia heterophylla–associated gemycircularvirus – EuaGmV) were found in samples obtained from weeds *Momorcadia charantia* and *Euphorbia heterophylla*, respectively (De Rezende et al., 2018). However, according to our knowledge, this is the first report of a gemycircularvirus associated with tomatoes in Brazil and worldwide.

In the present study, the complete DNA–A sequences of the begomoviruses BGMV, SiMMV, and ToCmMV were detected and subsequently confirmed in the individual samples via PCR assays and Sanger sequencing. BGMV was found in two samples collected in the Gama–DF region in 2003 (isolates DF–045 and DF–046) and in one sample collected in Leopoldo de Bulhões–GO (isolate GO–142). In fact, BGMV has been previously found in association with tomatoes in the Submédio São Francisco River valley in Northeast Brazil (Lima et al., 2001). However, this initial detection was carried out by using only DNA–A specific probes without additional molecular characterization of the putative BGMV isolates (Lima et al., 2001). Therefore, our work is the first to characterize tomato-infecting BGMV isolates. Interestingly, the DNA–B components of these BGMV isolates were not recovered from the samples that were positive for DNA–A component of this virus; indicating that these isolates might be using one alternative DNA–B component from another co-infecting species. Additional bioassays will be necessary to confirm this hypothesis.

SiMMV was detected in association with tomatoes in Goiás State (nine samples) and in the Federal District (three samples). SiMMV was already reported infecting tomatoes in 2004 in São Joaquim de Bicas–MG (Calegario et al., 2004). Interestingly, all SiMMV isolates were found only in the pool without the *Ty*–1 gene (**Table 6**), suggesting virus-specific filtering effects by this genetic factor. It will be of interest to challenge plants harboring the *Ty*–1 gene with infectious SiMMV clones to confirm this potential high level of resistance to this pathogen. This work is now underway.

Somewhat surprising, only the DNA–A component of the bipartite species ToCmMV was detected in two samples of the pool without the *Ty*–1 gene (GO–023 and MG–388) collected in Luziânia-GO and Viçosa-MG, respectively. The isolate GO–023 is mixed infection with ToCMoV and the isolate MG–388 is mixed infection with ToSRV. The absence of the DNA–B component of ToCmMV is also suggesting that these isolates might be using this component of these co-infecting species. This hypothesis remains to be investigated. ToCmMV was initially reported infecting tomato plants collected in 2005 in Paty do Alferes in Rio de Janeiro and Coimbra-MG (Castillo-Urquiza et al., 2008). In field surveys carried out in Espírito Santo, ToCmMV was identified as the only bipartite *Begomovirus* species infecting tomatoes between the years 2007 and 2011 (Barbosa et al., 2016). However, ToCmMV was not yet reported in Goiás State (GO–023), having a predominant occurrence in regions comprising the Atlantic Rain Forest biome and vicinities. Even though both ToCmMV isolates were found in the pool without the *Ty*–1 gene, there are reports indicating that this virus can replicate and cause mild symptoms in tomato plants carrying this tolerance factor (manuscript in preparation).

The DNA–A and DNA–B genome sequences of EuYMV and CILCrV were also recovered in our NGS analyses only from samples without the *Ty*–1 gene. EuYMV was detected in two samples in Minas Gerais State (isolates MG–012 and MG–016) collected in 2002 and CILCrV was found in one sample collected in Minas Gerais State (MG–150) in 2010. EuYMV was first characterized infecting the weed *E. heterophylla* (Fernandes et al., 2011) and CILCrV was first reported infecting the weed *Cleome affinis* (Paprotka et al., 2010). However, according to our knowledge, this is the first report of these two viral species naturally associated with tomatoes. The detection of these two species reinforces the hypothesis that weeds can serve as a natural reservoir for begomoviruses that may able to move and be able to infect cultivated plants such as tomatoes.

We found that the NGS analyses in combination with PCR assays with virus-specific primers and Sanger sequencing to be powerful tools that allowed us to assess the relative prevalence of the predominant *Begomovirus* species in distinct geographic areas of Central Brazil. In the present work, we were also able to recover the complete genomes of the monopartite species ToMoLCV as well as the sequences of the DNA–A and DNA–B components of the bipartite species ToCMoV, TGVV, ToRMV, and ToSRV, which were detected in both pools (with and without the *Ty*–1 gene) of samples **(Figure 1)**. ToSRV has been described as the predominant begomovirus species as indicated by

independent surveys carried out across all tomato-producing regions in Brazil (Rezende et al., 1997; Cotrim et al., 2007; Fernandes et al., 2008). ToSRV was also the predominant species in our study, being found in 46 samples of pool without the *Ty*–1 gene and in 26 samples in the pool harboring the *Ty*–1 gene **(Figure 1)**. This ability of ToSRV isolates replicate in plants with the *Ty*–1 gene could also be considered as an additional factor explaining the overall predominance of the virus under Brazilian conditions. ToSRV is predominant in the central and meridional regions of Brazil (Fernandes et al., 2008; Rocha et al., 2013), whereas ToMoLCV is predominant in the Northeast region (Fernandes et al., 2008). However, ToMoLCV is also often found in the Central Brazil (Albuquerque et al., 2012), which was confirmed by our results. ToCMoV has already been reported across the Northeast, Southeast, and Central Brazil (Ribeiro et al., 2003; Ribeiro et al., 2007). However, our results indicated that besides the Federal District, a large number of ToCMoV-infected tomato samples were also identified in Goiás State. TGVV is commonly found in Central Brazil (Albuquerque et al., 2012) and our results are in agreement with this observation. ToRMV is a recombinant viral species with genomic contributions of ToSRV and ToCMoV (Ribeiro et al., 2007). In accordance with our results, ToRMV was found to be one of the predominant viral species in the central region of Brazil, especially in the Goiás State (Ribeiro et al., 2003; Fernandes et al., 2006).

The present work is the first exploratory study on the potential impact of the *Ty*–1 gene on the diversity of Neotropical *Begomovirus* species. It was possible to observe putative filtering effects as well as gene-specific viral selection in samples with the *Ty*–1 gene, indicating a potential evolution of viral populations more adapted to this genetic factor. It would be interesting to know if the viruses detected in the apical mild symptoms in plants carrying the *Ty*–1 gene are indeed able to escape its effects or if the occurrence of multiple infections on these plants makes a more permissive cellular environment. However, an illustrative example is the isolate DF–640 that was recovered from a field–grown tomato plant carrying the *Ty*–1 gene in the Federal District, which was displaying severe disease symptoms. This strong susceptible–like reaction associated with the isolate DF–640 may indicate its potential ability to overcome the *Ty*–1 gene. Another possibility is that the isolate DF–640 may represent a singular "host switch" event that is not necessary associated with viral adaptation to the *Ty*–1 gene. The production of infectious DF–640 clones is now underway and they will be used to verify this hypothesis. Nevertheless, it is well documented in the literature that the increase in the acreage of

cultivars harboring resistance genes such as the *Ty*–1 can result in strong selection forces towards more aggressive viral isolates, accelerating the change in the composition of the viral population and potentially culminating with the loss of effectivity of the source of resistance/tolerance. Recently, it was reported the "breakdown" of *Ty*–1 mediated resistance/tolerance by TYLCV strains in Morocco, Italy, and Spain. Tomato plants with the *Ty*–1 gene showing severe symptoms caused by TYLCV were collected and the analyzes revealed the presence of a novel virus derived from a recombination event between TYLCV and *Tomato yellow leaf curl Sardinia virus* (TYLCSV) in which a non–coding region between the origin of replication and the start of the V2 gene were switched (Belabess et al., 2015; Panno et al., 2018; Granier et al., 2019; Torre et al., 2019). In Brazil, a study was carried out, evaluating begomovirus diversity in samples of a susceptible processing tomato cultivar ('Heinz 9553') and a *Ty*–1 harboring cultivar ('BRS Sena'). ToSRV and ToMoLCV were detected in both cultivars, being ToSRV the most prevalent. Mutations were detected in the isolates of both viral species with a greater number of substitution mutations occurring in the ToSRV and ToMoLCV DNA – A sequences obtained from 'BRS Sena', indicating that these viral isolates are suffering stronger selection pressure which was most likely imposed by the presence of the *Ty*–1 gene (Rêgo, 2016). Studies have also reported the "breakdown" of the resistance mediated by the *Ty*–2 gene caused by a strain of TYLCV (Ohnishi et al., 2016) and by a strain of the Tomato leaf curl Bangalore virus (ToLCBV) in India (Tiwari et al., 2010). The effectivity loss of the *Ty*–2 mediated resistance to this virus was explained by a combination of changes in replication efficiency, viral gene expression and by the recombination events in viral genomic regions that may be less prone to transcriptional gene silencing (Voorburg et al., 2020).

Our preliminary set of analyses showed no unique (i.e. pool–specific) polymorphisms among a subset of *Begomovirus* species found in the two pools (data not shown). Several point mutations were found, but none of them was specific to the viruses present in pool with or without the *Ty*–1 gene. Thus, another plausible explanation for some of the reported field events of *Ty*–1 mediated resistance/tolerance "breakdown" under Brazilian conditions could be related to some natural synergistic interactions with distinct group of viruses. In fact, it has been demonstrated that the *Ty*–1 gene does not confer resistance to major tomato-infecting RNA viruses such as *Tomato spotted wilt virus* – TSWV and *Cucumber mosaic virus* – CMV (Butterbach et al., 2014). However,

it has been demonstrated that RNA viruses can compromise resistance against begomoviruses as previously shown during TYLCV and CMV co–infection, where there was a significant increase in TYLCV concentration that was due to the inhibition of the transcriptional gene silencing response by CMV 2b RNAi suppressor protein (González et al., 2010; Hamera et al., 2012; Butterbach et al., 2014). In the present work, it was not possible to assess the diversity of RNA viruses associated with the samples because the employed methodological approach did not allow us to analyze this group of viruses.

5. Conclusion

The results reported here provide useful information about the population dynamics of begomoviruses associated with tomato crops across three major tomato-producing regions of Central Brazil in the last decade. However, in order to carry out a more precise study on the potential selective impact of the *Ty*–1 locus on begomovirus diversity and evolution, a distinct experimental strategy would be probably more appropriate, since our analysis was conducted on samples collected in different regions of a large country, in different years and from tomato plants grown in different microenvironmental situations. Therefore, it is possible that these variables (geographic area, climate, and year) can generate some biases that may not allow us to estimate the actual effect of *Ty*–1 gene. For this purpose, the analysis could be more appropriately conducted on samples collected from experimental plots cultivated with tomato isolines with and without the *Ty*–1 gene. On the other side, our ecologically–oriented approach allowed us to carry out a more ample exploration of an array of environments which may enhance the opportunity to detect a larger number of yet undescribed viral species associated with the tomato crop. Even though with a slightly different number of evaluated samples in the pools with (n=43) and without (n=64) the *Ty*–1 gene, virus-specific PCR assays and Sanger sequencing validations of NGS–derived data indicated greater diversity (14 versus six species) in samples lacking this gene. Moreover, two novel *Begomovirus* species, one gemycircularvirus (*Genomoviridae*) and one alpha–satellite were identified exclusively in samples without the *Ty*–1, whereas a novel begomovirus was found exclusively in the *Ty*–1 gene pool. These results indicated a potential viral adaptation to this tolerance factor as well as virus-specific filtering effects of the *Ty*–1 on a subset of single-stranded DNA viruses and subviral agents. However, these hypotheses

will be better tested with tomato isolines (with and without the *Ty*–1 gene) after controlled experiments employing infectious clones.

CHAPTER 3

Tomato yellow vein streak virus and *Tomato golden vein virus*: A reappraisal of the species status of two South American begomoviruses based upon genome-wide pairwise identity of multiple isolates

[1]Luciane de Nazaré Almeida dos Reis, [2]Maria Esther N. Fonseca, [2]Leonardo S. Boiteux, [1]Rita de Cássia Pereira–Carvalho.

[1]Departamento de Fitopatologia, Universidade de Brasília (UnB), Brasília – DF, Brazil.
[2]National Center for Vegetable Crops Research (CNPH), Embrapa Vegetable Crops (Hortaliças), Brasília – DF, Brazil.

Work submitted to Virus Genes

Resumo

Tomato yellow vein streak virus (ToYVSV) e o Tomato golden vein virus (TGVV) são begomovírus bipartidos da América do Sul que apresentam estreita relação genética. As identidades de DNA–A entre os isolados ToYVSV e TGVV exibem uma variação contínua (de 89 a 100%), o que tem gerado incertezas quanto ao real status taxonômico desses vírus. Um estudo abrangente com todos os isolados virais disponíveis foi realizado utilizando o Sequence Demarcation Tool (SDT) e alinhamentos via MUSCLE. Dois grupos bem definidos foram identificados, consistentes com os critérios atuais para demarcação de espécies de *Begomovirus*. Além disso, nossa reavaliação reconheceu uma variedade de isolados com nomes errôneos e um conjunto distinto de características genômicas, biológicas e ecológicas específicas para cada isolado.

Abstract

Tomato yellow vein streak virus (ToYVSV) and Tomato golden vein virus (TGVV) are closely-related bipartite begomoviruses from South America. The DNA–A identities among ToYVSV and TGVV isolates display a continuum (from 89 to 100%), that has generated uncertainty concerning their actual taxonomic status. A comprehensive study with all available viral isolates was conducted employing Sequence Demarcation Tool (SDT) and multiple MUSCLE alignments. Two clear-cut clusters were identified, consistent with the current criteria for *Begomovirus* species demarcation. Moreover, our reappraisal recognized an array of misnamed isolates and a distinctive set of species/isolate–specific genomic, biological, and ecological features.

Viruses of the genus *Begomovirus* (family *Geminiviridae*) are efficiently transmitted by members of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex (ICTV, 2020). Their single-stranded DNA (ssDNA) genomes consist of either one (DNA–A only) or two/bipartite (DNA–A and DNA–B) components that are replicated in the nuclei of their host cells (Rojas et al., 2018). The genus *Begomovirus* aggregates the largest number of species within the family *Geminiviridae* (ICTV, 2020). Due to the increasing number of isolates that have been recently characterized within this genus, a more robust set of taxonomic rules was established for novel species demarcation (Brown et al., 2015). In the first proposed set of criteria, a new species was only defined when the nucleotide identity levels the of complete DNA–A component was less than 89% in comparison with all the available virus

sequences (Fauquet et al., 2008). In 2015, a new set of criteria was established that determined standardized comparative analyses employing the MUSCLE alignment in combination with Sequence Demarcation Tool (SDT). In the current classification system, a novel species can only be defined when the nucleotide identity of the entire DNA–A displays less than 91% in comparison with the complete genome of any other known begomovirus sequence. If the DNA–A sequence of a given virus shares less than 94% identity with the complete DNA–A genome of all the previously described isolates for that species it is then classified as a new strain (Brown et al., 2015).

Begomoviruses are reported infecting tomatoes (*Solanum lycopersicum* L.), potatoes (*S. tuberosum* L.), beans (*Phaseolus vulgaris* L.), cowpea [*Vigna unguiculata* (L.) Walp.], cotton (*Gossypium hirsutum* L.) as well as more than 100 dicotyledonous species around the world (Inoue-Nagata et al., 2016a; Naito et al., 2019). In Brazil, the invasion of *B. tabaci Middle East-Asia Minor 1* (MEAM 1 = biotype B) in the early 1990s favored the rapid spread of begomoviruses across the main tomato-producing areas of the country (Ribeiro et al., 2003). In addition, the mechanisms of generating genetic variability in begomovirus (mutation, recombination, and pseudo-recombination) can lead to a more intense natural emergence of novel species (Ribeiro et al., 2003; Seal et al., 2006). In fact, tomato field surveys conducted after begomovirus outbreaks revealed a wide array of viral species (mainly with bipartite genomes) affecting this crop under Brazilian conditions. Thus far, 21 tomato-infecting species have been characterized in the country with the most prevalent ones being: Tomato severe rugose virus – ToSRV; Tomato mottle leaf curl virus – ToMoLCV (monopartite); Tomato chlorotic mottle virus – ToCMoV; Tomato common mosaic virus – ToCmMV; Tomato golden vein virus – TGVV, and Tomato yellow vein streak virus – ToYVSV (Faria et al., 1997; Ribeiro et al., 2003; Calegario et al., 2007; Castillo-Urquiza et al., 2008; Albuquerque et al., 2012; Macedo et al., 2018; Quadros et al., 2019; Rêgo-Machado et al., 2019). In addition, some begomoviruses that were reported infecting weed hosts were also described in tomatoes including: Sida mottle virus – SiMoV and Sida micrantha mosaic *virus* – SimMV (Calegario et al., 2004; Cotrim et al., 2007).

Isolates described as either ToYVSV or TGVV have been reported infecting tomatoes as well as other hosts across South America (Arruabarrena et al., 2016; Vaghi Medina et al., 2018; Varela et al., 2018; Bornancini et al., 2020). The close phylogenetic relationship as well as the multiple and independent descriptions of novel viral isolates of these two putatively distinct species have generated some uncertainty in relation to their taxonomic status. The first ToYVSV isolate was described infecting tomato in Campinas, São Paulo State – SP, Brazil in

1995 (Faria et al., 1997). This initial description was done with partial sequences of the DNA–A (1,303 nts; U79998) and DNA–B components (1,077 nts; U80042) obtained after PCR assays with the degenerate primer pairs 'PAC1v1978'/'PAV1c715' and 'PBC1v2039'/ 'PBV1c800' (Rojas et al., 1993) respectively. Subsequently, the partial (1,320 nts) DNA–A segment (encompassing the replication-associated protein – AC1 and the coat protein – AV1 genes) of a novel tomato-infecting ToYVSV isolate was characterized in Campinas–SP (AY829113) in 2004. The first complete DNA–A genome sequences of tomato-infecting isolates designated as ToYVSV were reported in 2007 (EF417915 = NC_010949 and EF459696). Additional surveys in Paty do Alferes (in Rio de Janeiro-RJ State) provided the complete DNA–A genome characterization of 23 novel tomato-infecting isolates that were named as ToYVSV (Rocha et al., 2013). Meanwhile, independent analyses indicated that ToYVSV isolates were also associated with a leaf deformation disease of potatoes known to occur in Southern Brazil since the 1980s (Daniels and Castro, 1985). The complete DNA–A and DNA–B sequences of the potato-infecting ToYVSV-Ba3 isolate (collected in 1983 in Rio Grande do Sul) were deposited at GenBank. The reference ToYVSV DNA–A component is EF417915 (= NC_010949)] and the reference DNA–B component is EF417916 (= NC_010950) (Albuquerque et al., 2010). Isolates referred to as ToYVSV have been also reported in association with tomato, bean, and *Capsicum* crops in the South Cone of South America, including Argentina, Uruguay, and Chile (Arruabarrena et al., 2016; Varela et al., 2018; Bornancini et al., 2020). In contrast, the first TGVV isolates were reported in 2004–2005 as putative novel species closely related to ToYVSV. Partial DNA–A genome characterization was carried with three tomato-infecting TGVV isolates (AY751742, DQ346649, and DQ346650) collected in Central Brazil. In 2011, the complete DNA–A sequence of five tomato-infecting TGVV isolates from inland areas of Central Brazil (including the Federal District-DF as well as Goiás-GO and Minas Gerais-MG States) were obtained (JF803254, JF803255, JF803256, JF803257, JF803258, and JF803259). The DNA–A component of the isolate DF [BR:Ita1220:03] (= NC_038807) was established as the TGVV reference sequence.

Even though the complete DNA–A of ToYVSV reference isolate (EF417916 = NC_010949) displays 89.22% nucleotide identity with the TGVV reference isolate (NC_038807), our preliminary analysis employing complete DNA–A segments of a subgroup of the isolates from the GenBank identified as either ToYVSV or TGVV showed a continuum with their identity levels ranging from 89 to 100%. This observation caused some uncertainty in relation to their actual taxonomic status. Moreover, a subgroup of GenBank isolates with

distinct names (either ToYVSV or TGVV) displayed 98 to 100% identities, indicating that they may represent dubious/erroneous descriptions of the very same viral species. On the other hand, a subgroup of isolates also designated as either ToYVSV or TGVV shared DNA–A nucleotide identity >87.5%, indicating potential inaccuracy of their nomenclature. In view of these apparently conflicting and complex aspects on the taxonomic status and nomenclature of these isolates, we carried out a comprehensive set of analyses aiming to clarify the genetic relationships among isolates previously characterized as either ToYVSV or TGVV.

Complete DNA–A sequences of 42 isolates named as either ToYVSV (n=36) or TGVV (n=6) were retrieved from the GenBank (www.ncbi.nlm.nih.gov). In addition, efforts to characterize additional ToYVSV and TGVV isolates were also carried out in the present work. Foliar samples of tomato cultivars showing typical begomovirus–induced symptoms (interveinal chlorosis, apical leaf deformation, yellow mosaic, rugosity, and dwarfism) were collected in across producing regions in GO, DF, and MG regions in Central Brazil. Foliar samples of tomato cultivars harboring the *Ty*–1 gene, but expressing conspicuous symptoms were also collected. These samples were subsequently subjected to total DNA extraction using a modified CTAB protocol (Boiteux et al., 1999) and stored at -20 °C. To confirm the presence of the *Ty*–1 gene, Polymerase Chain Reaction (PCR) assays were performed with the pair of *UWTyF* / *UWTyR* primers which allow for the detection of polymorphic codominant Cleavage Amplified Polymorphic Sequence (CAPS) markers associated with the resistant dominant allele (*Ty*–1) and with the susceptible recessive allele (*ty*–1) after cleavage with *Taq* I restriction enzyme. Total genomic DNA purifications were further enriched for circular ssDNA molecules via rolling circle amplification – RCA (Inoue-Nagata et al., 2004). Afterward, the samples were grouped into two contrasting pools: one harboring the *Ty*–1 gene and the other lacking the *Ty*–1 gene. The contrasting pools were submitted to the high–performance sequencing at an Illumina HiSeq 2500 platform at the Macrogen Inc. (South Korea). The sequences were assembled in the CLC Genomics Workbench program 10. The generated contigs were validated by BLASTn and compared to a ssDNA virus database. The genomes of the Illumina-derived viral species were analyzed and amplified using the Geneious 11.0 program. Both pools of samples showed the presence of TGVV–related sequences, but not ToYVSV– related sequences. Therefore, PCR assays were performed for the detection of individual TGVV isolates using a pair of species–specific primers (TGVV For1: 5'–AAA GGA AGA TAA TTC AAA TAT AGG GA–3'/ TGVV Rev1: 5'–ATC TTC CTT TAC TCA CGT TC CTG AT–3') designed in Geneious 11.0. Multiple MUSCLE alignments were performed in SDT v1.2 (Muhire et al., 2014) and the phylogenetic

constructions were performed using the Geneious 11.0 program by the PhyML method, model F81 with 1,000 bootstrap replications. The figures were elaborated with Adobe Illustrator CC and EvolView.

We were able to recover the complete DNA–A and DNA–B components of two novel TGVV–related isolates: one from the pool without the *Ty*–1 gene [(the DNA–A component with 2,562 nts (= GenBank MN928610) and the DNA–B component with 2,534 nts (= MN928611)] and one from the pool of plants carrying *Ty*–1 gene [the DNA–A component with 2,561 nts (= MN928612) and the DNA–B component with 2,575 nts (= MN928613)]. A set of analyses using phylogenetic analysis and SDT was then carried out with these two novel isolates plus all 43 isolates available at GenBank that were named as either ToYVSV or TGVV. Our results showed two clear-cut clusters of isolates, which is consistent with the current criteria for *Begomovirus* species classification (Brown et al., 2015). The first subgroup was composed by tomato-infecting isolates named in the GenBank as either ToYVSV or TGVV which displayed two common features: **(1)** tomato as a natural host and **(2)** geographical occurrence in inland areas of Central and South-East Brazil. These isolates displayed identity levels ranging from 95 to 100% among them, indicating that they represent descriptions of a single viral species. Therefore, a large fraction of this cluster of viruses is composed by misnamed isolates. These viruses should, according to our analysis, be collectively referred to as TGVV isolates, since they have nucleotide identity levels above 91% with the corresponding reference isolate of this viral species (**Figure 1**). The second phylogenetic subgroup was composed by the reference ToYVSV–Ba3 (the potato-infecting isolate from South Brazil = EF417915) as well as by an array of isolates also named as ToYVSV from the South Cone of South America (*viz*. KC136336; KC136337; KC136339; GQ387369, KY555801, and MN508216). The overall nucleotide identity range of these isolates in comparison with the reference ToYVSV isolate ranged from 96% to 100% (**Figure 1**).

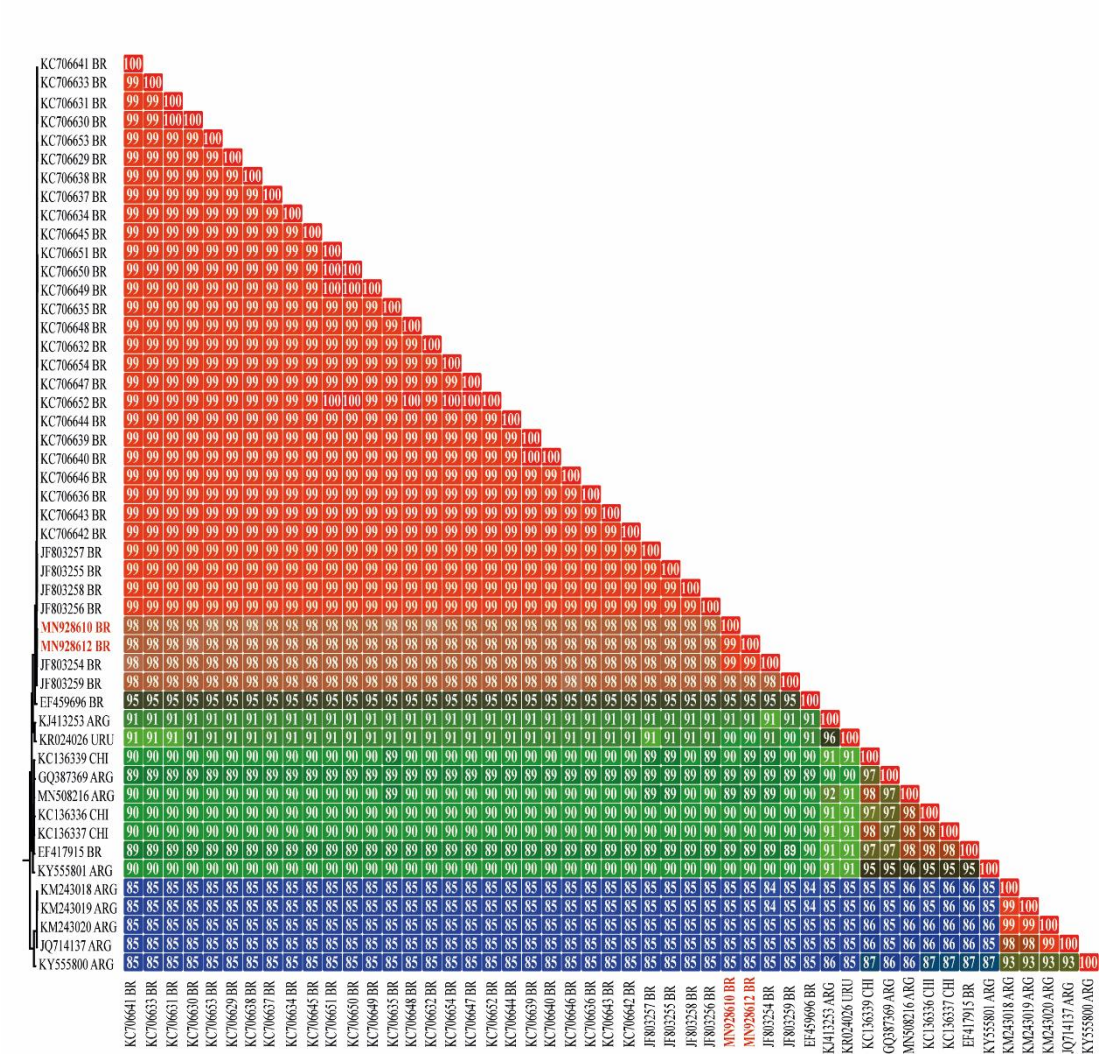

Figure 1. Phylogenetic tree and Sequence Demarcation Tool (SDT) of a set of DNA–A component sequences showing the phylogenetic identities/distances among *Tomato yellow vein streak virus* (ToYVSV) and *Tomato golden vein virus* (TGVV) isolates. These isolates are identified by their accession number and by the acronym of the countries where they were described: BR = Brazil; URU = Uruguay; ARG = Argentina; CHI = Chile. Two TGVV isolates (which complete sequences were obtained in the present study) are highlighted in red (MN928610 and MN928612). GenBank accession numbers of isolates classified/named as ToYVSV are the following: KC706641, KC706633, KC706631, KC706630, KC706653, KC706629, KC706638, KC706637, KC706634, KC706645, KC706651, KC7066, K7066, K7066, K7066, K7070, K7070 KC706644, KC706639, KC706640, KC706646, KC706636, KC706643, KC706642, EF459696, KJ413253, KR024026, KC136339, GQ387369, MN508216, KC136336, KC136337, and EF417915. GenBank accession numbers of isolates classified as TGVV are the following: JF803257, JF803255, JF803258, JF803256, JF803254, and JF803259. GenBank accession numbers of isolates classified as *Tomato mottle wrinkle virus* (ToMoWrV) are the following: KM243018, KM243019, KM243020, JQ714137, and KY555800. The DNA–A component of a

tomato-infecting ToYVSV isolate from Bolivia was only partially characterized (GenBank JQ413300) and for this reason it was not included in the analyses.

These ToYVSV isolates from the South Cone of South America have been reported infecting not only tomatoes, but distinct hosts (such as potato, *Capsicum*, and beans) in Chile, Argentina, and Uruguay. In addition, a partially characterized tomato-infecting ToYVSV isolate was reported in Bolivia (JQ413300). Is important to highlight the identification of two tomato-infecting isolates that displayed distinct genetic features: KJ413253 from Argentina and KR024026 from Salto (Uruguay). These isolates were the two most divergent ones within this subgroup, displaying identity levels of 91% with isolates of both species (TGVV and ToYVSV). However, according to our analyses these isolates should be referred to as ToYVSV, since they have closer relationship to the reference isolate U79998. *Tomato mottle wrinkle virus* (ToMoWrV), which was reported infecting tomatoes in Argentina, was found to be the begomovirus with the closest genetic relationship to ToYVSV and TGVV (**Figure 1**). The results observed with the available DNA–B sequences were similar to those described for the DNA – A component (**Figure 2**).

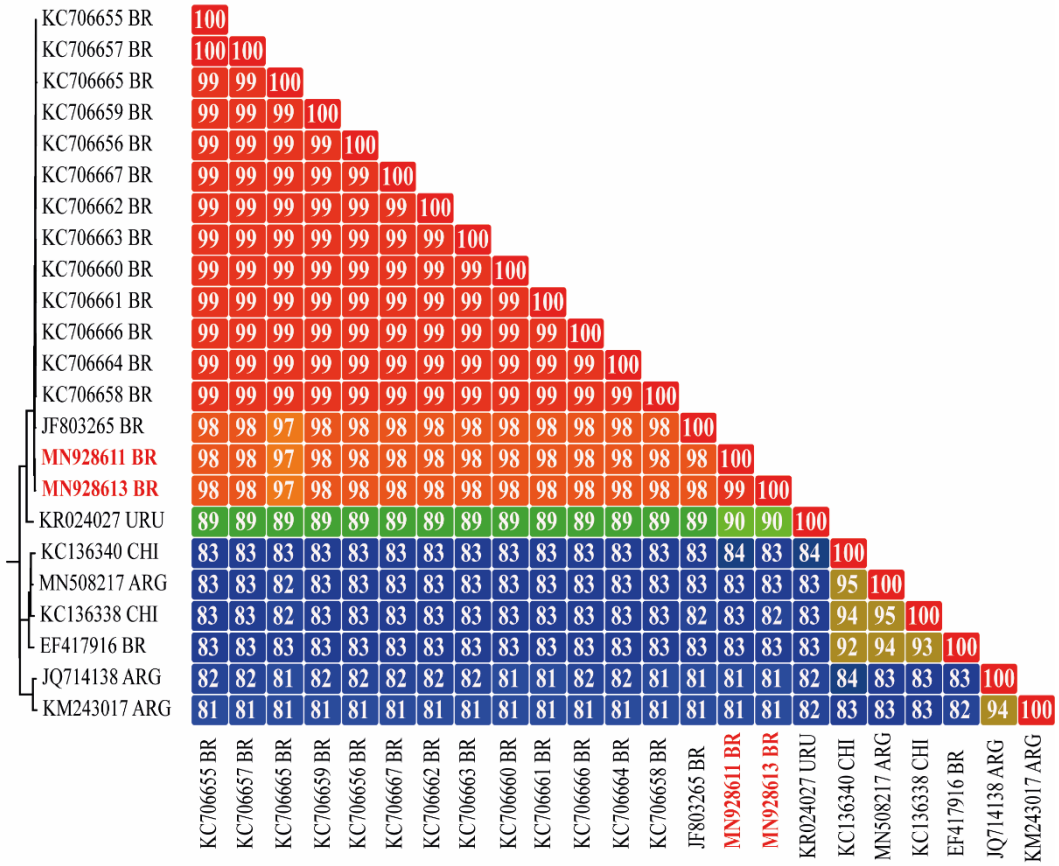

Figure 2. Phylogenetic tree and Sequence Demarcation Tool (SDT) of a set of DNA–B component sequences showing the phylogenetic identity/distance among *Tomato yellow vein streak virus* (ToYVSV) and *Tomato golden vein virus* (TGVV) isolates. The isolates are identified by their accession number and by the acronym of the countries where they were described: BR = Brazil; URU = Uruguay; ARG = Argentina; CHI = Chile. TGVV isolates which complete sequences were obtained in the present study (MN928611 and MN928613) are highlighted in red. GenBank accession numbers of isolates classified as ToYVSV are the following: KC706655, KC706657, KC706665, KC706659, KC706656, KC706667, KC706662, KC706663, KC706660, KC706661 KC706666, KC706664, KC706658, KR024027, KC136340, MN508217, KC136338, and EF417916. GenBank accession number of the isolate classified as TGVV is JF803265. GenBank accession numbers of isolates classified as *Tomato mottle wrinkle virus* (ToMoWrV) are JQ714138 and KM243017. A tomato-infecting ToYVSV isolate from Bolivia was only partially characterized (GenBank JQ413300) and for this reason it was not included in the analyses.

As previously discussed, ToYVSV and TGVV are closely-related viruses that were independently described in different years, hosts as well as geographic areas (**Figure**

3). The original viral descriptions were also carried out with distinct amount of genomic information (i.e. partial versus complete DNA–A sequences), which generated dubious information about the taxonomic status and nomenclature of these pathogens. In this context, the present study is the first comprehensive attempt to clarify the taxonomic status and isolate nomenclature of these two economically important New World bipartite begomoviruses. Our genome-wide pairwise identity analyses of multiple isolates indicated that a substantial fraction of the 43 GenBank isolates identified as either ToYVSV or TGVV were deposited with an erroneous virus name. For example, the same research group responsible for the characterization of the original ToYVSV isolate in São Paulo State, Brazil (U79998) deposited another putative tomato-infecting ToYVSV isolate in 2004 (AY829113). However, our analyses indicated that AY829113 was, in fact, one of the first partial sequences available for TGVV. Likewise, the complete DNA–A sequence of one of the first available TGVV isolates (named as isolate G–22 = EF459696) was also misnamed as ToYVSV. Our analyses indicated that isolate G–22 displays 94.68% identity with the TGVV reference isolate. Similar inaccuracy in relation to virus nomenclature was observed in a distinct study with tomato-infecting isolates collected in Rio de Janeiro State, South-East Brazil (Rocha et al., 2013). Our analyses indicated that 26 isolates from this survey deposited as ToYVSV in the GenBank (KC706629–KC706640; KC706642; KC706643; KC706645–KC706650; and KC 706652–KC706654) are misnamed and they should be reclassified as TGVV isolates.

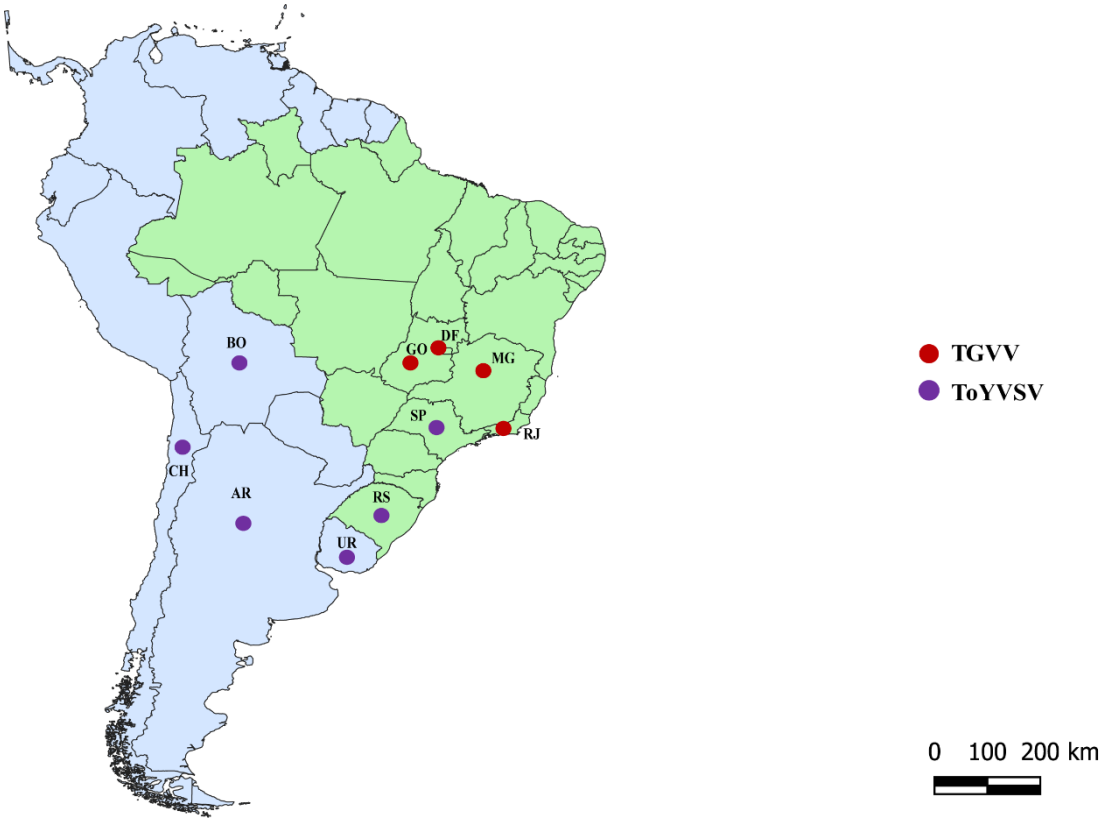

Figure 3. Map of South America showing the geographical distribution of *Tomato yellow vein streak virus* (ToYVSV) and *Tomato golden vein virus* (TGVV) isolates. The red dots are representing the geographical areas of occurrence of tomato-infecting TGVV isolates in Brazil (the Federal District-DF, Goiás-GO, Minas Gerais-MG, and Rio de Janeiro-RJ States). The purple dots are indicating the geographical areas of ToYVSV occurrence in Brazil (a potato-infecting isolate in Rio Grande do Sul-RS State and tomato-infecting isolates in São Paulo-SP State) as well as ToYVSV isolates reported infecting tomato, bean, and *Capsicum annuum* crops in the South Cone of South America, including Argentina (AR), Chile (CH), Uruguay (UR). The DNA–A component of a tomato-infecting ToYVSV isolate from Bolivia (BO) was only partially characterized (GenBank JQ413300).

We also carry out analyses of the genomic region encompassing the common region (CR) of the DNA–A and DNA–B components of all available TGVV and ToYVSV isolates. Our results showed that the TGVV and ToYVSV isolates are harboring distinct cognate iterons as well as distinct Rep iteron–related domains (Rep IRDs) (Argüello-Astorga and Ruiz-Medrano, 2001). The analyzed ToYVSV isolates displayed the **GGGGA** iteron (Rep IRD = MP**L**PKRF**L**VN), whereas the TGVV isolates displayed a **GGGTC** iteron (Rep IRD = MP**P**PKRF**T**VN). Positions with amino acid polymorphisms are represented in bold/underlined. In the 3th position, apolar amino acids were observed, being

a leucine (L) residue detected in the ToYVSV isolates and a proline (P) residue in the TGVV isolates. For the polymorphism of the 8th amino acid residue, it was observed an apolar amino acid (Leucine – L) in ToYVSV and a polar amino acid (Treonine – T) in the TGVV isolates. These observations reinforce the notion that TGVV and ToYVSV are, in fact, distinct species (Argüello-Astorga and Ruiz-Medrano, 2001). The divergent tomato-infecting ToYVSV isolate from Uruguay displayed the iteron **GGGGA** (Rep IRD = MP**L**PKRF**Q**VN), whereas the other divergent isolate from Argentina (from which the DNA–B component is not available) displayed Rep IRD = MP**P**PKRF**Q**VN. These genetically divergent ToYVSV isolates showed a distinct amino acid residue at the 8th position in comparison with other isolates from the same species. However, the isolate of Uruguay displayed Rep IRD more similar to that of the other ToYVSV isolates.

We also examined potential differences across isolates for the structural helix 4 motif (which amino acid sequence is strongly conserved across geminiviruses) (Arguello-Astorga et al., 2004). The predicted sequence in the TGVV isolates was: LSK**A**LNILKEEQPRDYVL**HL**D**KIQS**HVQ**KIF**AKAPAPWVPI**F**ELSSFT**HVP**DEMQ QWA, whereas for the ToYVSV isolates the predicted amino acid sequence was: PST**A**LNILKEEQPRDYVL**HL**D**KIRT**HVQ**RIF**AKAPTPWVSPFQLSS**F**T**NVP**DEMQ EW. The highlighted amino acid residues are the ones that are predicted to interact with the plant retinoblastoma–related protein (pRBR) in order to modulate the overall host gene expression (Arguello-Astorga et al., 2004). The TGVV reference isolate showed the identical helix 4 motif when compared with a set of isolates previously classified as ToYVSV from Rio de Janeiro (Rocha et al., 2013). Helix 4 motifs observed in the two isolates divergent isolates from Uruguay and Argentina were more similar to the others ToYVSV isolates.

Moreover, we analyzed the quasi-palindromic DNA–A segment [ACTT– (N7) – AAGT] which is an structural element conserved across the CP gene promoters of several members of the *Geminiviridae* family (Cantú-Iris et al., 2019). Most of the ToYVSV isolates displayed the sequence ACTT–A**G**GCGCT–AAGT. However, stable differences were observed in the 6th nucleotide [ACTT–A**A**GCGCT–AAGT] across the tomato-infecting TGVV isolates from MG, DF, GO (Central Brazil) as well as in isolates from Rio de Janeiro (Rocha et al., 2013). Interestingly, ToYVSV isolates from Argentina, Uruguay, and Chile also displayed an adenine (**A**) in this site. We also performed another set of analyses aiming to verify the presence of the ORF AC5 across the GenBank collection of TGVV and ToYVSV isolates. The ORF AC5 has been identified in a subgroup of

begomoviruses, and it gene product is supposed to act as a pathogenicity factor by suppressing RNA silencing-based antiviral host defenses (Li et al., 2015a). Our analysis indicated that the ORF AC5 is present in most of the isolates of both species. However, the ORF AC5 displayed a variable range in size (from 228 to 252 nts) in the TGVV isolates, whereas all ToYVSV isolates displayed a standard size of 291 nts. The divergent ToYVSV isolate KJ413253 from Argentina was the only one devoid of the ORF AC5.

Hence, in agreement with the new set of criteria for taxonomic demarcation of *Begomovirus* species (Brown et al., 2015), our work gives support to the notion that ToYVSV and TGVV are closely-related but distinct and valid *Begomovirus* species with EF417915 and JF803254 being their reference DNA–A genomic sequences, respectively. Moreover, our reappraisal recognized an array of misnamed isolates as well as a peculiar/distinctive set of species–specific genomic, biological, and ecological features. According to our analyses, the current collection of 45 ToYVSV and/or TGVV isolates at the GenBank is, in fact, composed by 35 instead of seven TGVV isolates and by a group of nine instead of 35 ToYVSV isolates. Therefore, a significant fraction of the ToYVSV and TGVV isolates currently available at GenBank should be renamed in order to avoid further misunderstandings. ToYVSV isolates are so far reported infecting a wider range of natural hosts (e.g. tomato, potato, *Capsicum*, and beans) and it can be also mechanically transmitted to *Nicotiana benthamiana* (Albuquerque et al., 2010). It was also observed a prevalence of ToYVSV isolates in meridional (high latitude) areas with mild climates across the South Cone of South America. On the other hand, the TGVV isolates were found to be restricted to subtropical inland areas of Central and South-East Brazil (**Figure 3**) and they were reported infecting only tomatoes thus far. From the plant breeding standpoint, this information will be useful in guiding germplasm screening strategies in order to develop resistant/tolerant cultivars to each specific virus as well as to each specific host and region in South America.

CHAPTER 4

A host-guided diversity and speciation of *Bean golden mosaic virus* isolates from *Phaseolus* species and from other legume and non-legume plants

L. N. A. Reis[1], J. G. Batista[1], M. L. F. Oliveira[1], L. S. Boiteux[1,2], M. E. N. Fonseca[2], J. C. Faria[3], R. C. Pereira–Carvalho[1].

[1]Universidade de Brasília (UnB), Departamento de Fitopatologia, Área de Virologia Vegetal, Brasília–DF.

[2]National Center for Vegetable Crops Research (CNPH), Embrapa Hortaliças, Brasília–DF.

[3]Embrapa Arroz e Feijão, Goiânia–GO, Brazil.

Work submitted to Virus Genes

Resumo

O feijão-comum (*Phaseolus vulgaris*) e o feijão-lima (*Phaseolus lunatus*) são os principais hospedeiros de Bean golden mosaic virus (BGMV). Este begomovírus também foi descrito infectando outras leguminosas e culturas solanáceas. Regras taxonômicas foram recentemente estabelecidas para a demarcação de espécies de *Begomovirus*. No entanto, as identidades de DNA–A entre isolados designados como BGMV exibem uma variação contínua (89–100%), em claro conflito com os critérios taxonômicos para uma única espécie. Aqui, avaliamos a diversidade de 161 isolados classificados como BGMV comparando suas sequências completas de DNA–A e DNA–B com o isolado de referência. Os isolados de BGMV foram agrupados em quatro grupos discriminados principalmente pelas hospedeiras leguminosas originais. As análises empregando Sequence Demarcation Tool (SDT) indicaram que os isolados descritos coletivamente como BGMV compreendem, de fato, duas espécies distintas: uma que engloba isolados de BGMV (de *P. vulgaris* e de uma ampla gama de hospedeiros) e uma espécie intimamente relacionada com BGMV (com identidade variando de 89 a 91% com o isolado de referência) que foi encontrada associada principalmente com o feijão-lima. Além disso, reconhecemos um conjunto de características genômicas específicas das espécies, como iterons e seus motivos proteicos associados a Rep. Para esta nova espécie de *P. lunatus*, sugerimos um nome previamente proposto – Lima bean golden mosaic virus (LBGMV).

Palavras chaves: *Begomovirus*, Feijão-comum, Feijão-lima, Sequence Demarcation Tool

Abstract

Common beans (*Phaseolus vulgaris*) and Lima beans (*Phaseolus lunatus*) are the major hosts of Bean golden mosaic virus (BGMV). This begomovirus was also described infecting other legumes species and solanaceous crops. Robust taxonomic rules were established for *Begomovirus* species demarcation. However, DNA–A identities among isolates designated as BGMV display a continuous variation (89–100%), in a clear conflict with the taxonomic criteria for a single viral species. Here, we assessed the diversity of 161 isolates classified as BGMV by comparing their complete DNA–A and DNA–B sequences with the reference isolate. BGMV isolates were clustered into four groups mainly discriminated by the original legume hosts. Sequence Demarcation Tool (SDT) analyses indicated that isolates collectively described as BGMV comprise, in fact, two distinct species: one encompassing *bona fide* BGMV isolates (from *P. vulgaris* and from a wide range of hosts) and one closely-related species (identities ranging from 89–91% to the reference BGMV) mainly associated with Lima beans. Moreover,

we recognized a set of species–specific genomic features of the iterons and their Rep–associated protein motifs. For this novel species from *P. lunatus* we suggest the proposed name – *Lima bean golden mosaic virus* (LBGMV).

Keywords: *Begomovirus*, common bean, Lima bean, Sequence Demarcation Tool

---

Viral species of the family *Geminiviridae* are responsible for significant yield losses in many economically important crops across tropical and subtropical regions (Rojas et al. 2018). *Begomovirus* is the largest genus within the family with over 400 species described (ICTV, 2020). These viruses are reported infecting exclusively dicotyledons and they are quite efficiently transmitted by a cryptic species complex of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) (Rojas et al. 2018). Begomoviruses are characterized by circular, single-stranded (ss) DNA genomes, encapsulated in twinned icosahedral particles (18–20 *x* 30–32 nm). These viruses can have either only one (= monopartite) or two (= bipartite) DNA molecules (Brown et al., 2015; Rojas et al., 2018). Due to the increasing number of viruses that have been characterized within the *Begomovirus* genus, robust taxonomic rules have been established in recent years for novel species demarcation. In the first proposed criteria set, a new species was defined only when nucleotide identity levels of the complete DNA–A component was less than 89%, in comparison with all available viral sequences (Fauquet et al., 2008). Subsequently, Brown et al. (2015) elaborated a novel standardized set of criteria that requires comparative analyses employing multiple MUSCLE alignments in combination with Sequence Demarcation Tool (SDT). In the current classification system, a novel species can only be defined when the entire DNA–A nucleotide sequence identity displays less than 91% in comparison with the complete reference genome of a given *Begomovirus* species. If a given DNA–A sequence shares 94% identity with the complete DNA–A genome that virus is then classified as a new strain (Brown et al., 2015).

Bean golden mosaic virus (BGMV) and Bean golden yellow mosaic virus (BGYMV) are the most important bean–infecting *Begomovirus* species in the Americas and Caribbean region, being the causal agents of "bean golden mosaic disease" (Gilbertson et al., 1991; Gilbertson et al., 1993; Faria et al., 2016). In Brazil, only BGMV has been reported and thus far it is the most important begomovirus of common beans (*Phaseolus vulgaris* L.) and Lima beans (*P. lunatus* L.) (Gilbertson et al., 1991; Faria and Maxwell, 1999; Faria et al., 2016). More recently, the weed–associated Macroptilium yellow spot virus (MaYSV) has been

reported as an emergent pathogen of beans in Northeast Brazil (Silva et al., 2012; Sobrinho et al., 2014). However, BGMV is still very important in that geographic region (Sobrinho et al., 2014). In many traditional bean-producing regions the cultivation of this crop has become almost unfeasible due to high levels of BGMV incidence (Faria et al., 2016). Yield losses caused by BGMV in common beans may vary from 90,000 to 280,000 tons, and isolates of this virus have been described infecting other legume hosts [*viz*. *Glycine max* (L.) Merr. and *Macroptilium lathyroides* (L.) Urb.] after the invasion of *B. tabaci Middle East-Asia Minor 1* (MEAM 1 = biotype B) in the early 1990s (Fernandes et al., 2009; Silva et al., 2012; Sobrinho et al., 2014). More recently, BGMV isolates have been described infecting the legume weed *Macroptilium erythroloma* (Mart. ex Benth.), a Fabaceae tree [*Anadenanthera colubrina* (Vell.) Brenan] as well as non-legume (solanaceous) hosts such as tomatoes (*Solanum lycopersicum* L.), eggplant (*S. melongena* L.), and the weed *Nicandra physalodes* (L.) Gaertn.

Studies dealing with BGMV diversity in *Phaseolus* species as well as in other legume hosts have been conducted under Brazilian conditions (Faria and Maxwell, 1999; Wyant et al., 2012; Sobrinho et al., 2014). Overall, the results indicated relatively low genetic variability among BGMV populations, but distinct host-guided genetic diversity was observed (Sobrinho et al. 2014). These viral variants were initially classified as novel BGMV strains which were mainly associated with *P. lnatus* samples (Wyant et al., 2012; Sobrinho et al., 2014). However, DNA–A identities among isolates designated as either BGMV or BGMV strains that are available in public databases display a continuous variation (89–100%), which is in a clear conflict with the established taxonomic criteria for a single viral species. In fact, the observation that a putative novel viral species (distinct from BGMV) might be associated with Lima beans was done previously by Faria and Maxwell (1999). They suggested the name Lima bean golden mosaic virus (LBGMV) for one of these isolates. However, this nomenclature was not adopted more likely because this initial description was done only with a partial genomic sequence of 1185 bp (= GenBank U92531) encompassing a segment of the Rep protein (rep) and the coat protein (cp). In addition, the standard taxonomic rules for novel *Begomovirus* species demarcation were not well-established at that time.

Due to the economic and biological importance of the BGMV–bean pathosystem, we decided to carry out an extensive analysis in order to catalog the genetic variability of all available isolates classified as either BGMV or BGMV strains from *Phaseolus* species and other legume hosts as well as novel isolates identified in non-legume (solanaceous) hosts. This work was carried out by analyzing the complete DNA–A and DNA–B sequences of these

isolates and by comparing them with both components of the reference BGMV isolate (NC_004042 for DNA–A and NC_004043 for DNA–B). For this purpose, all 161 available complete DNA–A genomic sequences of BGMV were retrieved from the GenBank database (https://www.ncbi.nlm.nih.gov). The original hosts of these BGMV isolates were the following: six unclassified *Phaseolus* species, 77 *P. vulgaris*, 56 *P. lunatus* as well as one isolate from soybean, three from tomatoes, 15 from *M. lathyroides*, one from *M. erythroloma*, one from *N. physalodes*, and one from *A. colubrina*. In addition, 12 complete DNA–B genomes were also retrieved from the NCBI database corresponding to five isolates from unclassified *Phaseolus* species, two from *P. vulgaris*, one from *P. lunatus*, two from *M. lathyroides*, one from *M. erythroloma* and one from *A. colubrina*. Phylogenetic analyzes were carried out employing genomic information of these 161 BGMV isolates with the complete sequence of the DNA–A component. The phylogenetic tree was generated from the alignment of the complete DNA–A component of each isolate, using the MUSCLE program implemented by Geneious 11.0 (PhyML method, model F81 with 1,000 bootstrap replications). Multiple MUSCLE alignments were performed in SDT v1.2 (Muhire et al., 2014) and the figures were elaborated with Adobe Illustrator CC and EvolView (He et al., 2016). Comparative analyses were also carried out with Geneious 11.0 program (Kearse et al., 2012) using genomic information from the quasi-palindromic DNA–A segment [ACTT– (N7)–AAGT] which is an structural element conserved across the Coat protein (CP) gene sequences of several *Geminiviridae* genera (Cantú-Iris et al., 2019). We also analyzed the nucleotide sequences of the common region (CR) of the cognate DNA–A and DNA–B components as well as the replication–associated protein (Rep) motifs (Argüello-Astorga et al., 1994; Argüello-Astorga and Ruiz-Medrano, 2001).

Phylogenetic analysis of a set of the full–genomes of the DNA–A components indicated a clear-cut discrimination of the BGMV isolates in four clusters according mainly to their original legume hosts (**Figure 1**). The **Group #1** was composed by BGMV isolates reported infecting mainly *P. vulgaris*, but also soybean, *P. lunatus*, tomato, *N. physalodes*, *M. erythroloma*, and *A. colubrina*. The **Group #2** was composed by BGMV isolates obtained from *M. lathyroides*, whereas the **Group #3** encompassed BGMV isolates mainly obtained from *P. lunatus*, but also from unclassified *Phaseolus* species and *M. lathyroides*. Finally, the **Group #4** was composed by only two divergent BGMV isolates reported infecting *M. lathyroides*. Analyzes using SDT and MUSCLE alignments, including the isolates of the Group #1 and Group #2 as well as the DNA–A genome of the BGMV reference isolate (NC_004042), showed identity levels ranging from 96–97% among them (**Figure 2**). These results indicated that all

these viruses are isolates with close genetic relationship to the reference BGMV species. SDT analyses employing the isolates belonging to **Groups #3 and #4** displayed identity levels ranging from 89 to 91% in relation to the reference BGMV isolate (NC_004042). In fact, the majority of the isolates from *P. lunatus* and from unclassified *Phaseolus* species displayed identity levels of 91% when compared to NC_004042. Exceptions were observed in three isolates (KJ939711, KJ939710, and KJ939720), which nucleotide identities ranged from 94–95% to the reference BGMV isolate **(Figure 3).** Moreover, some isolates also classified as BGMV (*viz*. KJ939735, KJ939731, KJ939719, JF694451, JF694449, and JF694452) displayed identities of 90% **(Figure 3)**, indicating that they are more likely isolates of a novel species according to the current criteria for the classification in the *Begomovirus* genus (Brown et al., 2015). On the other hand, SDT analyses among the **Groups #3 and #4** isolates showed identity levels ranging from 95–99%, indicating that they belong to the same species.

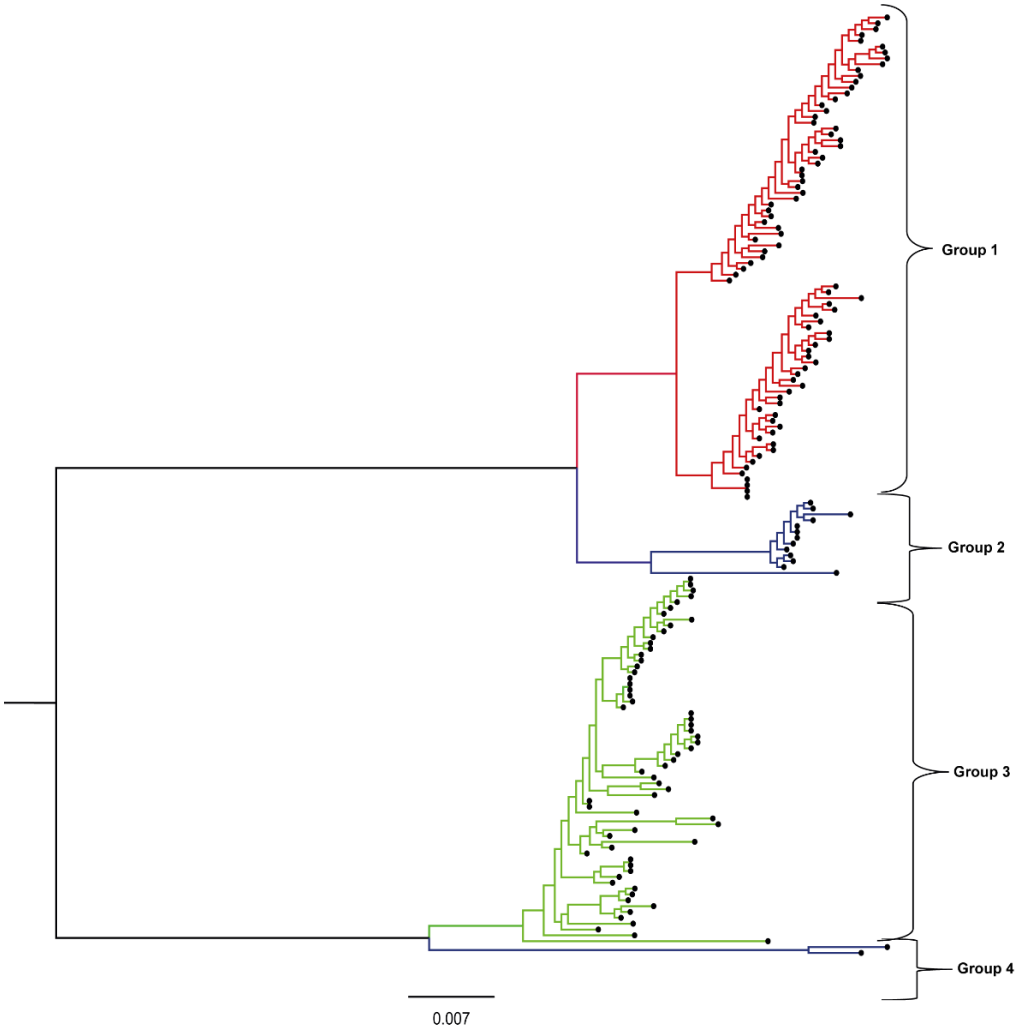

Figure 1. Phylogenetic tree of a set of full-genome DNA–A components showing the phylogenetic identities/distances of 161 *Bean golden mosaic virus* (BGMV) isolates available at the GenBank. Midpoint-rooted ML with 1,000 bootstrap replications. **Group #1** was composed by BGMV isolates reported infecting *Phaseolus vulgaris*, soybean (*Glycine max*), tomato (*Solanum lycopersicum*), *Nicandra physalodes*, *Macroptilium erythroloma*, and *Anadenanthera colubrina* (with branches in red), **Group #2** was composed by BGMV isolates obtained from *Macroptilium lathyroides* (with branches in blue) and **Group #3** was composed by BGMV isolates obtained from *P. lunatus* (with branches in green), and **Group #4** was composed by two highly divergent BGMV isolates reported infecting *M. lathyroides* (with branches also in blue).

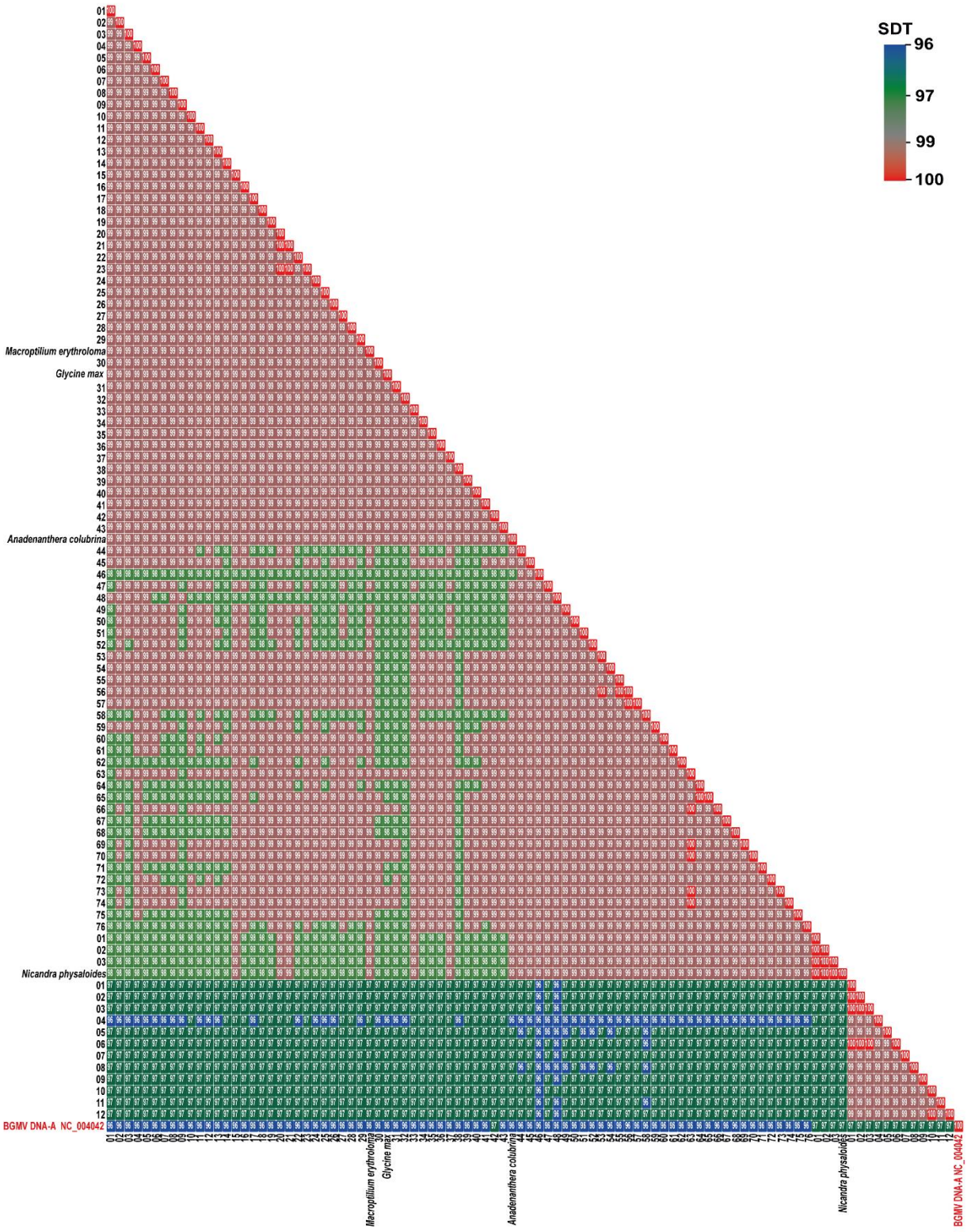

Figure 2. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the DNA–A component sequences of isolates obtained from *Phaseolus vulgaris*, *Macroptilium lathyroides, Macroptilium erythroloma*, *Anadenanthera colubrina*, *Nicandra physalodes*, *Glycine max* and *Solanum lycopersicum* indicating their identities in relation to the reference

(NC_004042) *Bean golden mosaic virus* (BGMV) sequence (indicated in red font color). BGMV isolates from *P. vulgaris* are identified by a numerical order and they correspond to the following GenBank accessions: [Isolates *P. vulgaris:* 01 (KJ939839), 02 (KJ939838), 03 (KJ939810), 04 (KJ939848), 05 (KJ939829), 06 (KJ939836), 07 (KJ939786), 08 (KJ939815), 09 (KJ939845), 10 (KJ939837), 11 (KJ939822), 12 (KJ939824), 13 (KJ939832), 14 (KJ939823), 15 (KJ939811), 16 (KJ939798), 17 (KJ939841), 18 (KJ939809), 19 (KJ939816), 20 (KJ939801), 21 (KJ939805), 22 (KJ939795), 23 (KJ939813), 24 (KJ939849), 25 (KJ939852), 26 (KJ939818), 27 (KJ939781), 28 (KJ939840), 29 (KJ939783), 30 (KJ939782), 31 (KJ939803), 32 (KJ939842), 33 (KJ939853), 34 (KJ939793), 35 (KJ939812), 36 (MG334552), 37 (KJ939843), 38 (KJ939851), 39 (KJ939792), 40 (KJ939802), 41 (KJ939850), 42 (KJ939799), 43 (KJ939806), 44 (KJ939844), 45 (KJ939826), 46 (KJ939847), 47 (KJ939835), 48 (KJ939830), 49 (KJ939821), 50 (KJ939831), 51 (KJ939819), 52 (KJ939825), 53 (KJ939827), 54 (KJ939788), 55 (KJ939787), 56 (KJ939785), 57 (KJ939820), 58 (KJ939833), 59 (KJ939828), 60 (KJ939780), 61 (KJ939784), 62 (KJ939790), 63 (KJ939779), 64 (KJ939817), 65 (KJ939800), 66 (KJ939789), 67 (KJ939794), 68 (KJ939807), 69 (KJ939808), 70 (KJ939791), 71 (KJ939796), 72 (KJ939797), 73 (KJ939814), 74 (KJ939804), 75 (KJ939834), and 76 (KJ939846)]; [Isolate from *M. erythroloma* (MN822294)]; [Isolate from *Glycine max* (FJ665283)]; [Isolate from *A. colubrina* (MN734370)]; [Isolate from *N. physalodes* (MN737555)]; [Isolates from *S. lycopersicum*: 01 (MN737552), 02 (MN737553), 03 (MN737554)]; [Isolates from *Macroptilium lathyroides*: 01 (KJ939725), 02 (KJ939714), 03 (KJ939707), 04 (KJ939756), 05 (KJ939708), 06 (KJ939732), 07 (KJ939764), 08 (KJ939733), 09 (KJ939709), 10 (KJ939717), 11 (KJ939715), 12 (KJ939734)].

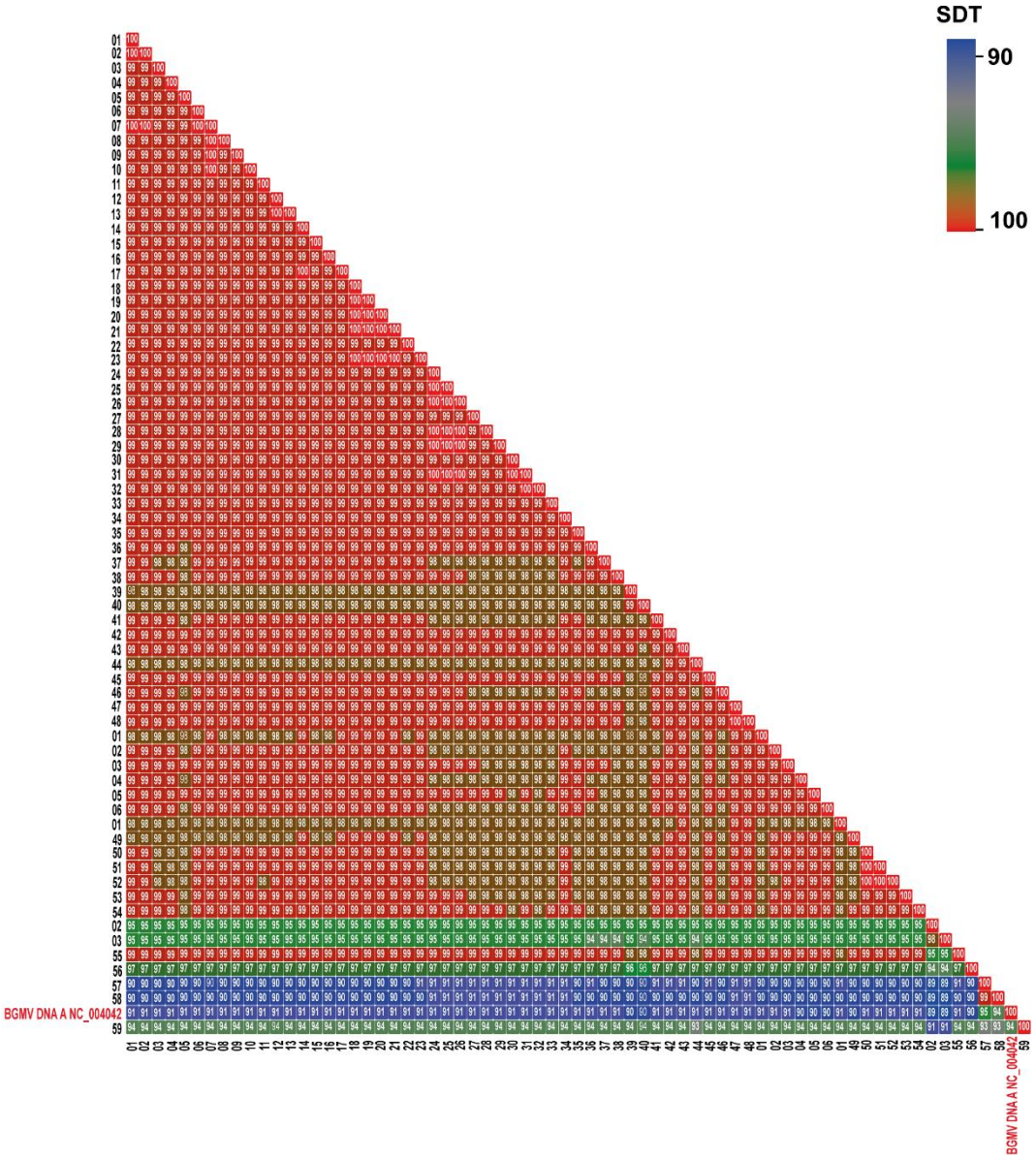

Figure 3. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the DNA–A component sequences of *Bean golden mosaic virus* (BGMV) isolates obtained from *Phaseolus lunatus*, unclassified *Phaseolus* species, and *Macroptilium lathyroides*, indicating their identities in relation to the reference BGMV (NC_004042) isolate (highlighted in red font color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: Isolates *P. lunatus*: [01 (KJ939748), 02 (KJ939739), 03 (KJ939749), 04 (KJ939738), 05 (KJ939746), 06 (KJ939743), 07 (KJ939750), 08 (KJ939741), 09 (KJ939751), 10 (KJ939737), 11 (KJ939744), 12 (KJ939747), 13 (KJ939740, 14 (KJ939745), 15 (KJ939752), 16 (KJ939753), 17 (KJ939742), 18 (KJ939730), 19 (KJ939728), 20 (KJ939727), 21 (KJ939726), 22 (KJ939729), 23 (KJ939736), 24 (KJ939762), 25 (KJ939760), 26 (KJ939754), 27 (KJ939763), 28 (KJ939759), 29 (KJ939761), 30 (KJ939758), 31 (KJ939757), 32 (KJ939755), 33 (KJ939765), 34 (KJ939756), 35 (KJ939712),36 (KJ939717),37 (KJ939715), 38 (KJ939714),39 (KJ939735), 40

(KJ939731), 41 (KJ939722), 42 (KJ939723), 43 (KJ939724), 44 (KJ939764), 45 (KJ939721), 46 (KJ939707), 47 (KJ939718), 48 (KJ939713), 49 (KJ939709), 50 (KJ939734), 51 (KJ939733), 52 (KJ939732), 53 (KJ939725), 54 (KJ939708), 55 (KJ939716), 56 (KJ939719), 57 (KJ939711), 58 (KJ939710), 59 (KJ939720)]; [Isolates from unclassified *Phaseolus* species : 01 (JF694453), 02 (JF694454), 03 (JF694450), 04 (F694451), and 05 JF694449, 06 (JF694452)]; [Isolates from *M. lathyroides*: 01 (JN419006), 02 (N419004), and 03 (JN419003)].

Comparisons of a subgroup of BGMV isolates obtained from the weed legume host *M. lathyroides* (clustered in the **Group #2)** showed identities of around 97% with the reference isolate (NC_004042) **(Figure 4)**. However, a distinct subgroup of isolates also from *M. lathyroides* (classified as BGMV) clustered in the **Groups #3 and #4** showed identity levels of 89–90% (e.g. JN419004 and JN419003) and 91% (e.g. JN419006). Therefore, as previously observed by Silva et al. (2012), *M. lathyroides* seems to be an "universal" host of distinct BGMV–related isolates **(Figure 4)** as well as from other legume-infecting viral species. When compared to *P. lunatus* isolates, a subgroup of *M. lathyroides* isolates from the **Groups #3 and #4** displayed identity levels ranging from 95 to 99%, indicating they are isolates of the same species **(Figure 3)**. On the other hand, the BGMV isolates from tomato (MN737552, MN737553, and MN737554), soybean (FJ665283), *M. erythroloma* (MN822294), *N. physalodes* (MN737555), and *A. colubrina* (MN734370) displayed identities ranging from 96–97% with the reference BGMV sequence **(Figure 4)**. Comparative analyses with the DNA–B sequences available at GenBank showed that isolates from unclassified *Phaseolus* species, *P. vulgaris*, *P. lunatus*, *A. colubrina*, and *M. erythroloma* displayed identities ranging from 89–100%. However, the sequences of isolates obtained from *M. lathyroides* displayed the lowest identity levels (79% and 82%) when compared to all available sequences **(Figure 5)**.

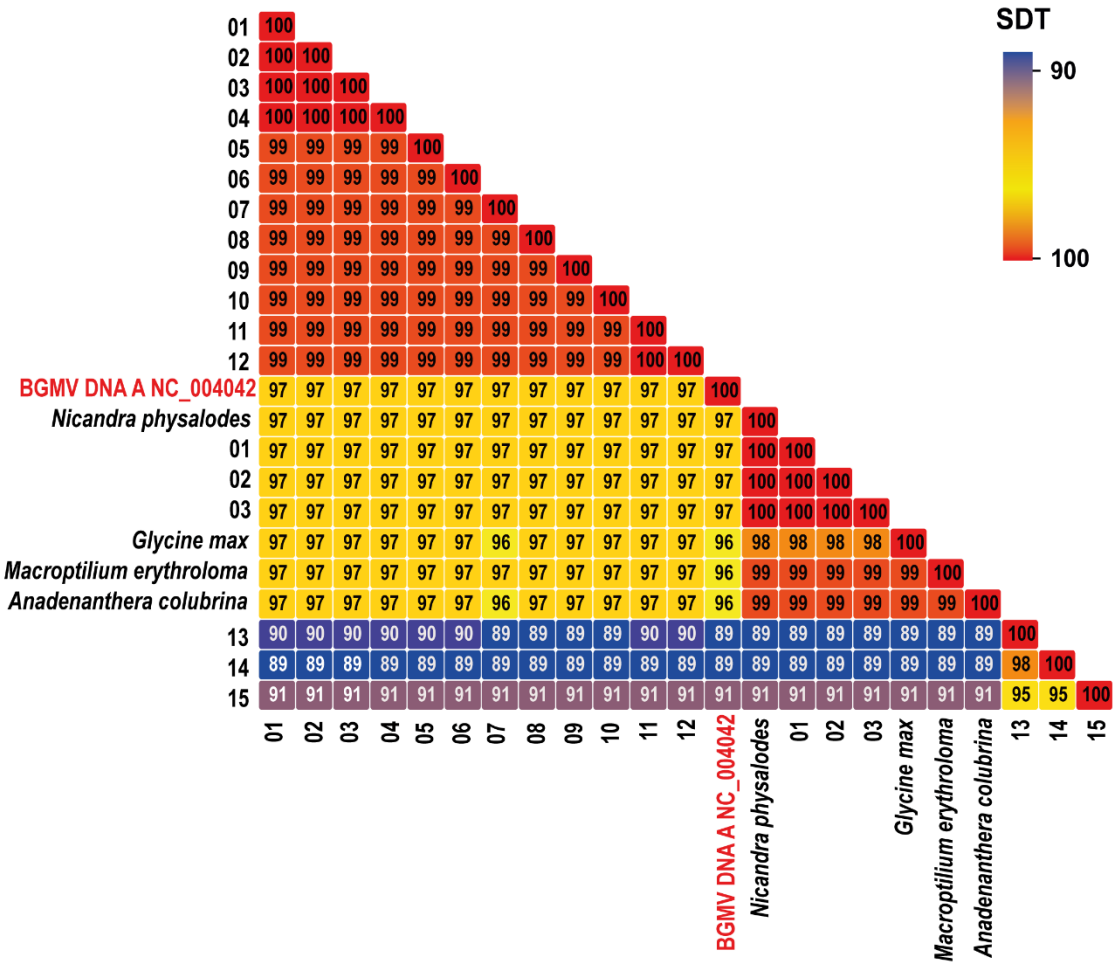

Figure 4. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the DNA–A component sequences of *Bean golden mosaic virus* (BGMV) isolates obtained from *Glycine max*, *Macroptilium lathyroides*, *Macroptilium erythroloma*, *Anadenanthera colubrina*, *Nicandra physalodes*, and *Solanum lycopersicum*, indicating their identities in relation to the reference BGMV (NC_004042) isolate (highlighted in red color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: [Isolates from *M. lathyroides*: 01 (KJ939725), 02 (KJ939714), 03 (KJ939707), 04 (KJ939756), 05 (KJ939708), 06 (KJ939732), 07 (KJ939764), 08 (KJ939733), 09 (KJ939709), 10 (KJ939717), 11 (KJ939715), 12 (KJ939734), 13 (JN419004), 14 (JN419003), and 15 (JN419006)]; [Isolate from *N. physalodes* (MN737555)]; [Isolates from *S. lycopersicum*: 01 (MN737552), 02 (MN737553), 03 (MN737554)]; [Isolate from *G. max* (FJ665283)]; [Isolate from *M. erythroloma* (MN822294)]; [Isolate from *A. colubrina* (MN734370)].

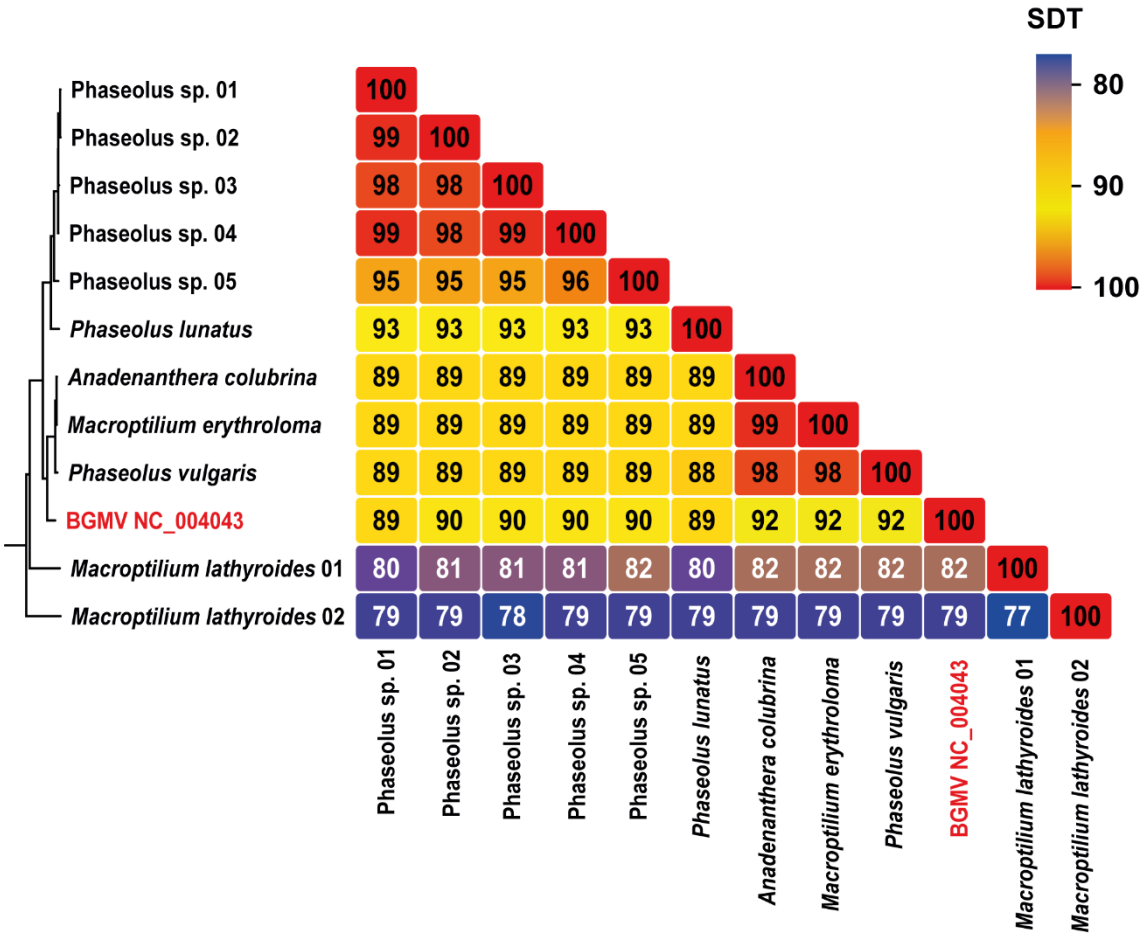

Figure 5. Pairwise identity analysis in Sequence Demarcation Tool (SDT) was carried out using the information of the of DNA–B component sequences of *Bean golden mosaic virus* (BGMV) isolates obtained from unclassified *Phaseolus* species, *Phaseolus vulgaris*, *P. lunatus*, *Macroptilium lathyroides*, *M. erythroloma*, *Anadenanthera colubrina*, indicating their identities in relation to the reference BGMV (NC_004043) isolate (highlighted in red font color). BGMV isolates from these hosts are identified by a numerical order and they correspond to the following GenBank accessions: Isolates from *Phaseolus* sp. 01 (JF694457), *Phaseolus* sp. 02 (JF694456), *Phaseolus* sp. 03 (JF694458), *Phaseolus* sp. 04 (JF694459), *Phaseolus* sp. 05 (JF694455); isolate from *P. lunatus* (MH925107); isolate from *A. colubrina* (MN734371); isolate from *P. vulgaris* (MG334553); isolates from *M. lathyroides* 01 (JN419008), and 02 (JN419017).

In order to reinforce the hypothesis that isolates named as BGMV may represent at least two distinct viruses, we also carried out analyses of the genomic region encompassing the common region (CR) of the DNA–A and DNA–B components of all available isolates. The iterons of the reference isolate as well as across all BGMV isolates with identity levels greater than 91%, displayed the sequence **GGTGT** (Rep iteron–

related domain – Rep IRD = **MPPPKRFKIN) (Figure 6)** (Argüello-Astorga et al., 1994; Argüello-Astorga and Ruiz-Medrano, 2001). The CR of the DNA–A and DNA–B components of isolates corresponding to putative new species also showed distinct iterons. The iteron found in the sequences of *P. lunatus*, unclassified *Phaseolus* species, and in a subgroup of the *M. lathyroides* isolates (JN419003 and JN419006) was **GGGGT** and the inverted sequence **ACCCC** (Rep IRD = **MPPPKRFKIS)** differing from the reference BGMV isolate in the last amino acid residue that was replaced by a serine. An exception was observed in the isolate KJ939719 that showed a distinct Rep IRD = **MPPPKRFRIS**. In the DNA–A genome of the *M. lathyroides* isolate (JN419004) and in the corresponding DNA–B genome (JN419017), was found the iteron **GGTAC** and its inverted **GTACC** sequence (Rep IRD = MPPPKRFKI**S) (Fig.7).**

We also examined potential differences across isolates classified as BGMV for the structural helix 4 motif, which the amino acid sequence is strongly conserved across geminiviruses (Arguello-Astorga et al., 2004). The sequence of the BGMV isolates compared to the reference isolate showed minor amino acid differences. All isolates differed with the reference isolate at position 184th. The polar amino acid tyrosine (Y) is present in the reference isolate, whereas in other isolates this residue was replaced by the non–polar amino acid phenylalanine (F) **(Figure 6)**. Other BGMV isolates showed additional, but not biologically relevant differences at positions 167th (e.g. isolates of *M. lathyroides*) as well as 175th and 178th (e.g. the isolate KJ939720 from *P. lunatus*) **(Figure 6)**. Isolates previously classified as BGMV when compared with reference BGMV isolate also displayed differences in position 184th. The reference isolate displayed the polar amino acid tyrosine (Y), whereas in the other isolates this residue was replaced by the non–polar amino acid phenylalanine (F) **(Figure 7)**. Another difference in relation to the BGMV reference sequence was at position 175th in which the basic amino acid lysine (K) is present. However, most of the other sequences displayed the non–polar amino acid proline (P), except for the sequences of *M. lathyroides* (JN419004 and JN419003), which have a polar amino acid glutamine (G) **(Figure 7, panel A)**. A subgroup of isolates showed distinct but not significant differences at positions 171, 197 and 198, including *M. lathyroides* (JN419004; JN419003) and 181th (the isolates KJ939735 and KJ939731 from *P. lunatus*) **(Fig. 7)**. The highlighted amino acid residues **(Figure 7, panel B)** are the ones predicted to interact with the plant retinoblastoma–related protein (pRBR) in order to modulate the overall host gene expression (Arguello-Astorga et al., 2004).

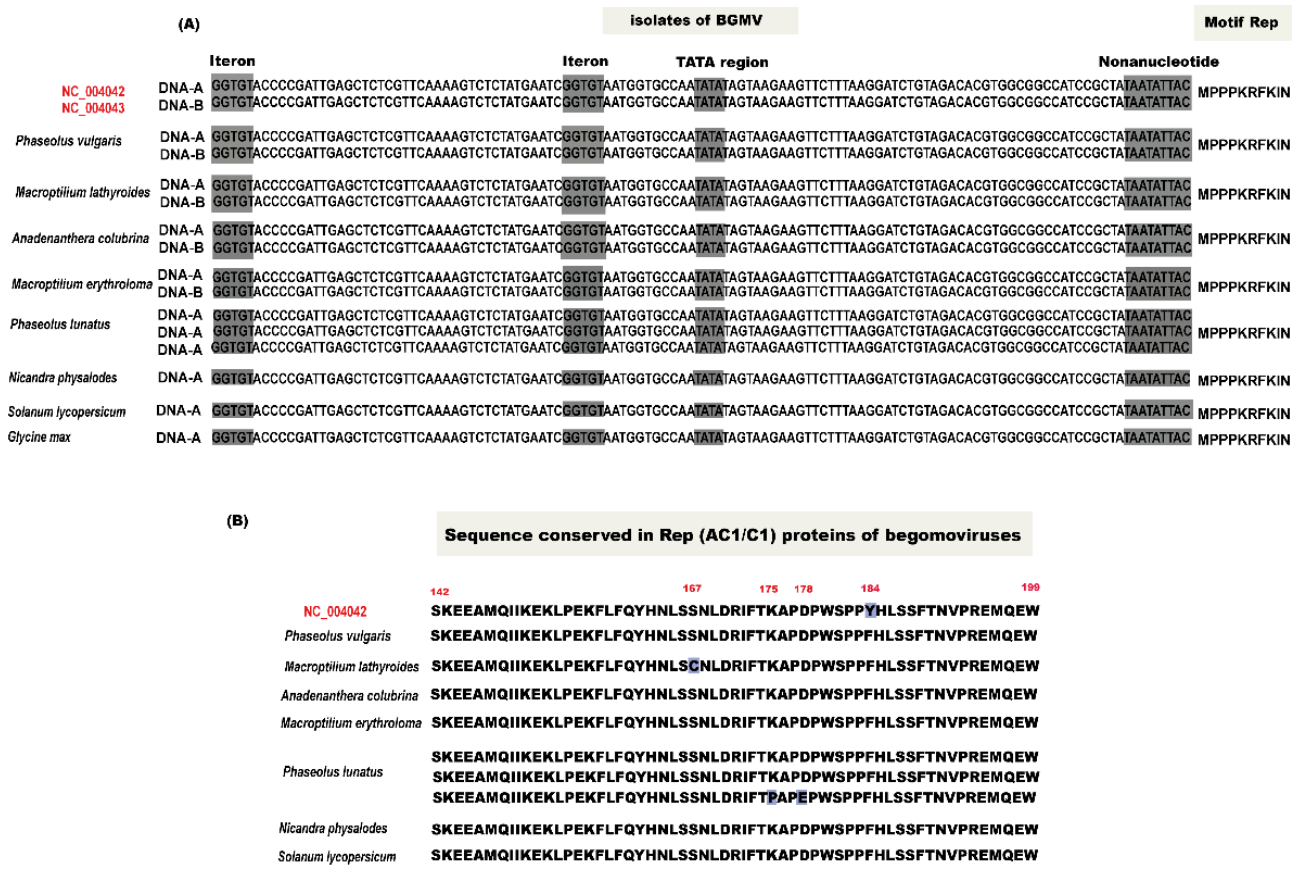

Figure 6. Common region, iterons and motifs of the Replication–associated protein (Rep) with the reference DNA–A and DNA–B sequences of *Bean golden mosaic virus* – BGMV (highlighted in red font color) compared with other isolates with identity levels greater than or equal to 96%. **Panel (A):** Iterons, TATA region, nonanucleotide and Rep motif; **Panel (B):** Conserved Rep protein sequence (ranging from 142 to 199 nucleotides). GenBank accessions: *Phaseolus vulgaris* DNA–A (NC_004042), DNA–B (NC_004043); *Phaseolus vulgaris*: DNA–A (KJ939833), DNA–B (MG334553); *Macroptilium lathyroides*: DNA–A (KJ939776), DNA–B (JN419008); *Anadenanthera colubrina*: DNA–A (MN734370), DNA–B (MN734371); *Macroptilium erythroloma*: DNA–A (MN822294), DNA–B (MN822293); *Phaseolus lunatus*: DNA–A (KJ939711), (KJ939710) and (KJ939710); *Nicandra physalodes* (MN737555); and *Solanum lycopersicum* (MN737552).

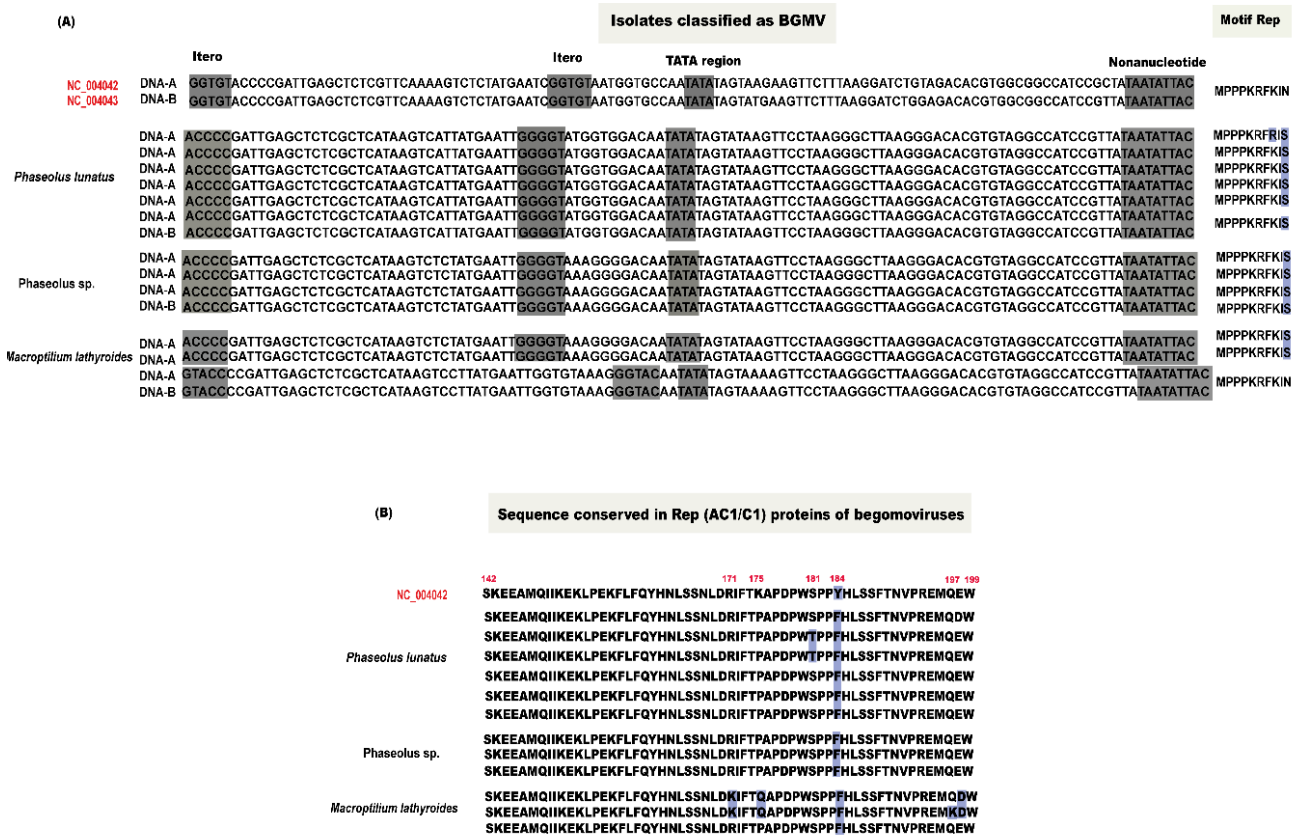

Figure 7. Common region, iterons and motifs of the Replication–associated protein (Rep) with the reference DNA–A and DNA–B sequences of *Bean golden mosaic virus* (BGMV) in red, compared with other isolates with identity levels between 89% and 91%. **Panel (A):** Iterons, TATA region, nonanucleotide and Rep motif; **Panel (B):** Sequence conserved of the Rep protein (from 142 to 199 nucleotides). GenBank accessions: *Phaseolus vulgaris* NC: DNA–A (NC_004042), DNA–B (NC_004043); *Phaseolus lunatus*: DNA–A (KJ939719), DNA–A (KJ939735), DNA–A (KJ939731), DNA–A (KJ939764), DNA–A (KJ939709), DNA–A (KJ939725), DNA – B (MH925107); unclassified *Phaseolus* species: DNA–A (JF694452), DNA–A (JF694451), DNA–A (JF694449), DNA–B (JF694454); *Macroptilium lathyroides*: DNA–A (JN419003), DNA–A (JN419006), DNA–A (JN419004), DNA–B (JN419017).

We also carried out analyzes of the quasi-palindromic DNA – A segment [ACTT–(N7)–AAGT] which is an structural element conserved across the coat protein (CP) gene

promoter of several members of the *Geminiviridae* family (Cantú-Iris et al., 2019). The sequences of the BGMV isolates as well as of the reference isolate are reported in **Fig. 8**. Results showed a motif diversity among BGMV isolates when comparing a specific region (ACTT–GTCGCCC–AAGT) with the reference isolate (NC_004042). Overall, the sequences found were: G**GC**G**A**CC, GG**T**G**A**CC, G**GC**G**A**CC, G**G**C**AA**CC, G**GT**G**T**CC and G**GCC**CCC. The differences in these sequences were detected mainly from the 2nd the 5th position **(Fig. 8).** Seventy–three isolates from *P. vulgaris* displayed the sequence **GG**C**GACC**, with the exception of the sequences of *P. vulgaris* isolates KJ939851 and KJ939795 that displayed the sequence **GGTGACC** with differences in the 3th position. Isolates from *P. lunatus* (KJ939710; KJ939720, and KJ939719) that shared identity greater than 96% to the reference, presented the sequence **GG**T**G**T**CC**, with the exception of the KJ939719 isolate, which displayed the sequence G**GCC**CCC. Twelve isolates from *M. lathyroides* (with identity levels of 97%) displayed the sequence G**G**C**AA**CC. On the other hand, isolates obtained from tomato, *M. erythroloma*, *N. physalodes*, *A. colubrina*, and soybean displayed the sequence G**G**CG**A**CC **(Fig. 8)**. The sequence found in isolates classified as BGMV were ACTT–G**G**C**C**CCC–AAGT, with the exception of a subgroup of isolates that displayed the alternative sequences (*viz*. G**A**C**C**C**T**C, G**G**C**C**CC**G**, and G**G**C**C**CC**T**C), with differences in the 2nd, 6th and 7th positions **(Fig. 8)**. Thirty isolates from *P. lunatus* and all isolates from unclassified *Phaseolus* species (with identity levels ranging from 90% to 91% with BGMV) displayed the G**G**C**C**CCC sequence. Exceptions were found, however, in six *P. lunatus* isolates (KJ939731; KJ939764; KJ939707; KJ939721; KJ939722, and KJ939723) in which the G**A**C**C**C**T**C sequence was annotated. Exceptions were also found in a group of 17 isolates (KJ939746; KJ939752; KJ939738; KJ939749; KJ939744; KJ939751; KJ939748; KJ939739; KJ939743; KJ939741; KJ939737; KJ939750; KJ939747; KJ939740; KJ939742; KJ939745, and KJ939753) that displayed the alternative G**G**C**C**CC**G** sequence. Three isolates from *M. lathyroides* classified as BGMV (with identity levels of 89–91%) displayed the alternative sequences G**G**C**C**CC**T**C (JN419006) and G**G**C**C**CCC (JN419004 and JN419003). The AC5 ORF has been identified in a subgroup of begomoviruses and its gene product is supposed to act as a pathogenicity factor by suppressing RNA silencing-based antiviral host defenses (Li et al., 2015a). We could detect the AC5 ORF (with a size of 252 nts) in a wide range of isolates obtained from *P. vulgaris*, *P. lunatus*, *M. lathyroides*, tomato, *N. physalodes*, soybean, *A. colubrina*, and *M. erythroloma*. Interestingly, two isolates reported infecting

M. lathyroides (JN419004 and JN419003) displayed the AC5 ORF with a distinct size of 276 nts.

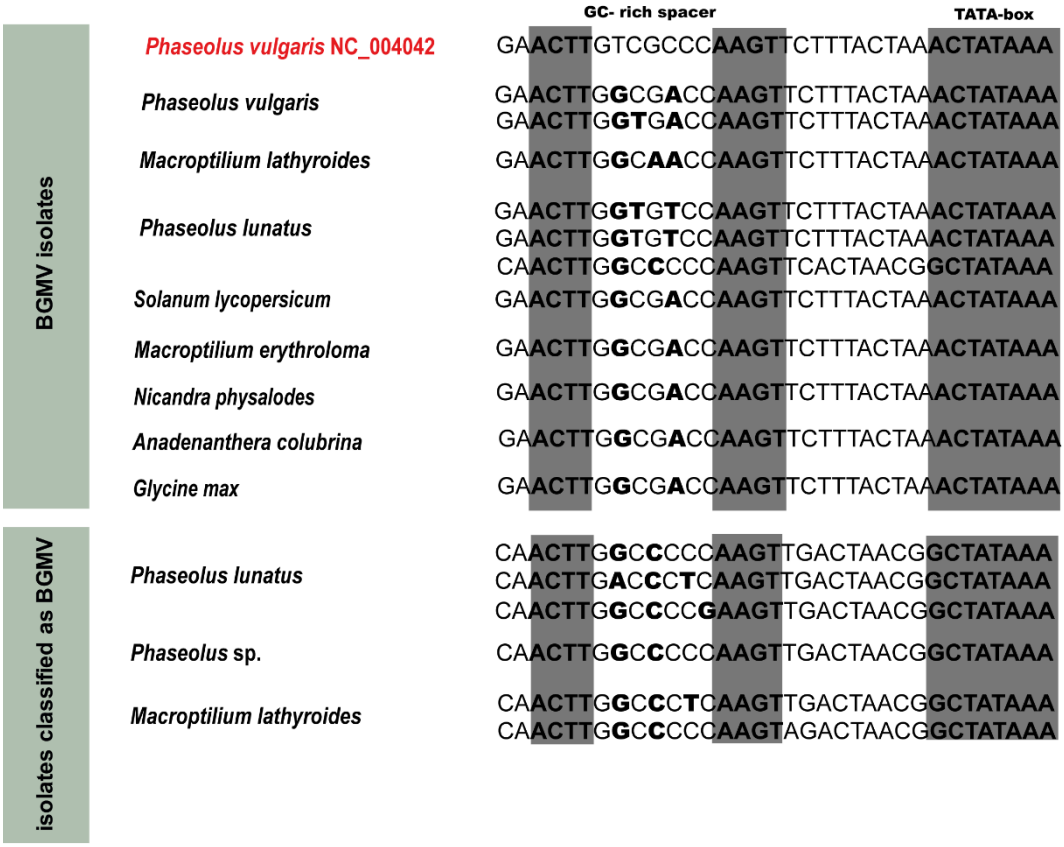

Figure 8. Symmetric region ACTT– (N7) – AAGT of isolates described as *Bean golden mosaic virus* (BGMV). Intergenic region sequences of the DNA–A component of the BGMV reference isolate (highlighted in red font color) was compared with other BGMV isolates with identity levels greater than or equal to 96%. Comparisons were also carried out with other isolates displaying identity level ranging from 89% to 91% (all these isolates were formerly classified as BGMV).

In conclusion, the 161 isolates previously classified as either BGMV or BGMV strains were clustered into four phylogenetic groups that were discriminated mainly by their original legume host. SDT analyses indicated (according to the current criteria for *Begomovirus* species demarcation) that isolates collectively named as BGMV comprise, in fact,of two viral species: one encompassing *bona fide* BGMV isolates mainly from *P. vulgaris* (but also including other legume and solanaceous hosts) plus one closely-related viral species (with identity levels below

91%) that was mainly associated with Lima beans and also *M. lathyroides*. Differences were also annotated among isolates of these viruses for a set of genomic features. Due to all these evidences, we concluded that this latter group of isolates correspond to a novel *Begomovirus* species for which is given the tentative name Lima bean golden mosaic virus (LBGMV), as previously suggested by Faria and Maxwell (1999). The reason why these genetically divergent BGMV isolates were not previously identified as novel species can be explained by the fact that the majority of these isolates was characterized before the novel *Begomovirus* species demarcation rules were established (Brown et al., 2015). Thus, with the current taxonomic criteria, these viral variants can now be unambiguously recognized as being isolates of a new *Begomovirus* species. In this context, the reclassification of a subset of BGMV isolates available at GenBank with less than 91% identity to the reference genome (NC_004042) seems to be necessary in order to avoid additional misunderstandings and, especially, for improving the classification system of these closely-related legume-infecting *Begomovirus* species.

It is interesting to highlight that variability in symptom the expression caused by BGMV isolates in beans has been detected as early as the 1960s in Brazil (Costa, 1965). Three distinct whitefly-transmitted diseases were previously identified as 'golden mosaic'; 'mottled dwarf', and 'crumpling'. These symptoms were so distinctive that Costa (1965) speculated that they were more likely induced by a complex of closely related viruses. Therefore, a plausible hypothesis is that one of these distinct diseases described by Costa (1965) could be caused by LBGMV isolates yet undetected at that time. Thus, it will be interesting to carry out comparative assays inoculating *P. lunatus* and *P. vulgaris* cultivars with BGMV and LBGMV isolates to identify if there is a peculiar set of symptoms associated with each virus. So far, isolates of BGMV have been reported on *P. lunatus*, but no natural infection of *P. vulgaris* by LBGMV. Practical implications of this pathogen diversity on classical and biotech *Phaseolus* resistance breeding programs are expected since virus-specific genetic factors may not be simultaneously effective against LBGMV and BGMV isolates (Faria et al., 2016).

As previously mentioned, the original description of the putative LBGMV species was done by Faria and Maxwell (1999), but only with a genomic segment (1185 bp) encompassing the partial Rep protein (rep) and coat protein (cp) genes of a single isolate (U92531). Our BLASTn analyses employing U92531 indicated identity levels ranging from 94.36% to 97.91% with a collection of 56 accessions composed mainly by isolates obtained from *P. lunatus* (Wyant et al., 2012; Sobrinho et al., 2014), from a group of unclassified *Phaseolus* species (Wyant et al., 2012) as well as two isolates from *M. lathyroides* (Silva et al. 2012). The highest

identity levels of U92531 to *bona fide* BGMV isolates was 89.74% (e.g. KJ939776), indicating that the isolate described by Faria and Maxwell (1999) was, in fact, the first report of a novel legume-infecting *Begomovirus* species distinct from BGMV.

CHAPTER 5

Complete genomic sequence of a *Gemycircularvirus* species detected in natural association with open-field tomatoes in Brazil

Luciane de Nazaré Almeida dos Reis[1·] Maria Esther de Noronha Fonseca[2·] Leonardo Silva Boiteux[2·] Josiane Goulart Batista[1·] Flávia Milene Barros dos Santos[1·] Rita de Cássia Pereira–Carvalho[1].

[1]Universidade de Brasília, Departamento de Fitopatologia, Área de Virologia Vegetal, Brasília – DF.

[2]National Center for Vegetable Crops Research (CNPH), Embrapa Hortaliças, Brasília – DF.

Work submitted to Journal of Plant Pathology

Resumo

A estratégia de *Next-Generation Sequencing* (NGS) foi utilizada para a identificação de um *Gemycircularvirus* (Família: *Genomoviridae*) associado com amostras foliares de tomateiro em campos de produção no Brasil Central. O genoma viral (ssDNA com 2.189 nucleotídeos) foi amplificado por meio de primers específicos e sequenciados por Sanger. As análises do BLASTn e da ferramenta Sequence Demarcation Tool (SDT) mostraram que a espécie identificada no tomateiro compartilha 99% de identidade com um vírus provisoriamente denominado como Plant-associated genomovirus 12 de *Larrea tridentata* (Zygophyllaceae). Análises filogenéticas usando as sequências de aminoácidos da proteína associada à replicação (Rep) dos genomas de *Genomoviridade*, confirmaram esse vírus como um membro do gênero *Gemycircularvirus*, representando o primeiro relato mundial de sua associação natural ao tomateiro.

Abstract

Using a next-generation sequencing (NGS) approach, we identified a *Gemycircularvirus* (*Genomoviridae*) in association with foliar samples of open-field tomatoes in Brazil. The viral ssDNA genome (2,189 nucleotides) was amplified via specific–primers and Sanger–sequenced. BLASTn and sequence demarcation tool (SDT) analyses showed 99% identity with a virus tentatively named as Plant-associated genomovirus 12 from *Larrea tridentata* (Zygophyllaceae). Phylogenetic analyzes using the amino acid sequences of the replication–associated (Rep) protein from *Genomoviridade* genomes confirmed this virus as a member of the genus *Gemycircularvirus*, representing the first worldwide report of its natural association with tomatoes.

In Brazil, tomato (*Solanum lycopersicum* L.; family Solanaceae) cultivation is of great economic importance, with the Southeast region being the main producer (45%), followed by the Central region with 30% of the total production (IBGE 2020). A large number of diseases of viral etiology have been reported in tomatoes under Brazilian conditions, varying according to the region, the type of cultivation and the dissemination and distribution of the vectors. Worldwide, isolates of ≈ 286 viral species have been reported either infecting or in association with tomatoes, including many circular single-stranded DNA (ssDNA) viruses (Virus-HostDB, 2020). Novel approaches for large–scale sequencing are allowing the discovery of a

wide diversity of circular ssDNA viruses (Rosario et al., 2012b; Krupovic et al., 2016). Viruses with ssDNA genomes were found in hosts of all three domains of life (Archaea, Bacteria, and Eukarya) across a large array of environments. The initial genomic analyses of a subgroup of these viruses indicated significant, but with overall low levels of similarity with circular ssDNA viruses classified within the *Geminiviridae*, *Circoviridae*, and *Nanoviridae* families. Because of this low genomic identities, it was proposed that these divergent ssDNA viruses should be placed into distinct families (Rosario et al., 2012a; Krupovic et al., 2016). In 2016, the International Committee on Taxonomy of Viruses created two new families of ssDNA: *Pleolipoviridae* and *Genomoviridae* (ICTV, 2020). The *Genomoviridae* family is currently composed by nine genera: *Gemycircularvirus*, *Gemyduguivirus*, *Gemygorvirus*, *Gemykibivirus*, *Gemykolovirus*, *Gemykrogvirus*, *Gemykroznavirus*, *Gemytondvirus*, and *Gemyvongvirus* (Krupovic et al., 2016; ICTV, 2020). The genus *Gemycircularvirus* has the largest number of described species thus far (43 in total), being characterized by circular ssDNA genomes (from 2.1 to 2.3 kb) with two open–reading frames (ORFs), one encoding the capsid protein (in the viral sense) and another encoding (in the complementary sense) the replication–associated protein (Rep). The Rep protein contains conserved domains that are crucial for rolling circle replication (RCR), which are also present in the Rep proteins from geminiviruses. In addition, they have a stem–loop structure in the region of origin of replication, where a conserved nonanucleotide sequence (**TAATATTAT**) is located (Rosario et al., 2012b; Krupovic et al., 2016; Varsani and Krupovic, 2017).

Currently, the rolling circle amplification – RCA (Inoue-Nagata et al., 2004) and the metagenomics allied to Next-generation Sequencing – NGS (Edwards and Rohwer, 2005; Pantaleo and Chiumenti, 2018; Liu et al., 2020) are the major techniques employed for the discovery of new circular ssDNA viruses in various organisms and ecosystems. The first gemycircularvirus described was the fungal *Sclerotinia sclerotiorum* hypovirulence–associated circular DNA virus 1 (SsHADV–1) (Yu et al., 2010). Afterwards, many genomoviruses have been reported in association with a large array of organisms such as insects (Dayaram et al., 2012; Rosario et al., 2012a; Li et al., 2015b), rats (Li et al. 2015), humans (Lamberto et al., 2014; Phan et al., 2015; Halary et al., 2016) as well as in environmental samples such as sewage, faeces, and water (Sikorski et al., 2013; Conceição-Neto et al., 2015; Da Silva Assis et al., 2016). In addition, these viruses have been found across different kingdoms, being reported also in association with plant species such as citrus (Chabi-Jesus et al. 2020); cassava (Dayaram et al. 2012); olive tress (Chiumenti et al. 2019);

Hypericum japonicum Thunb. (Hypericaceae) (Du et al. 2014); Poaceae species (Male et al. 2015), and soybeans(Dayaram et al., 2012; Du et al., 2014; Kraberger et al., 2015; Male et al., 2015; Marzano and Domier, 2016; Chiumenti et al., 2019; Chabi-Jesus et al., 2020). Here, we described an experimental approach using NGS in combination with Sanger dideoxy sequencing for the identification and complete genome characterization of a *Gemycircularvirus* (*Genomoviridae*) found in natural association with open-field tomatoes in Central Brazil.

Seventy–two tomato leaf samples displaying virus–like symptoms were collected across commercial fields in Goiás State (GO). Total genomic DNA was extracted from these samples, and circular DNA was enriched via RCA (Inoue-Nagata et al., 2004). The RCA products from all tomato samples were pooled and sequenced using a *Illumina* Hiseq platform at Macrogen Inc. (South Korea). The sequences were assembled and analyzed in the *CLC Genomics Workbench* 10 program. The contigs were validated via BLASTn and compared to the GenBank database of ssDNA viruses (https://www.ncbi.nlm.nih.gov/). A specific open–primer pair ('Gemy For': 5'–GCT CTG AAT CAA ATC TCG CTT ACT TG–3' / 'Gemy Rev': 5'–CGA TGT TGA TTG GTT GGA AGC A–3') was designed (Geneious 11.0 program) to anneal to a conserved Rep region of a putative *Gemycircularvirus* species found in the initial NGS analysis. This primer pair had the double purpose of detecting this putative *Gemycircularvirus* species in individual tomato samples and to amplify its complete genome. The obtained PCR amplicons were purified with Gel DNA purification (Ludwig Biotecnologia, Alvorada–RS, Brazil) and then Sanger–sequenced, employing the same *Gemycircularvirus-specific* primer pair. Sanger dideoxy sequencing was carried out at Genomic Analysis Laboratory (at CNPH), using the BigDye® Terminator Cycle Sequencing Ready Reaction Kit version 3.1 protocol (Applied Biosystems, São Paulo–SP, Brazil). After contig assembling and quality evaluation, the obtained sequences were analyzed using the BLASTn algorithm. Seventy–four sequences representing all genera of the *Genomoviridae* family were retrieved from NCBI and included in the phylogenetic analyses, employing the Rep amino acid sequences. The phylogenetic tree was built by Bayesian inference (MrBayes v. 3.2.) with amino acid substitution model GTR + I + G selected by JModeltest v. 2.2 (Posada, 2008). The trees were edited in the FigTree program (Rambaut, 2012) and in the Evolview (He et al., 2016) at the online server http://www.evolgenius.info/evolview/. For comparisons of the complete sequences, it was used the software SDT v.1.2 (Muhire et al., 2014). The

RDP4 program (Martin et al., 2015) was employed for analysis of potential recombination events.

After Illumina sequencing, the assembly of the contigs and analysis via BLASTn revealed a sequence with ≈ 98% identity to the viral species Plant-associated genomovirus 12 (genus *Gemycircularvirus*) that was reported in association with *Larrea tridentata* (Sesse & Mocino ex DC.) Coville (MH939425) from the family Zygophyllaceae. After PCR with the virus-specific primers, the gemycircularvirus was detected in two tomato samples (codified as GO–298 and GO–301) collected in Leopoldo de Bulhões, Goiás-GO in 2004. The amplicons corresponding to the full viral genome were validated via Sanger dideoxy sequencing. Phylogenetic and SDT analyzes confirmed the results obtained initially by BLASTn, where the tomato–associated isolates showed a 99% identity with Plant-associated genomovirus 12, belonging to the genus *Gemycircularvirus* (**Figure 1**). SDT analyses also showed that Plant-associated genomovirus 12 is closely-related to other species of this genus because it shared identity levels ranging from 78 and 84% with *Pteropus–associated gemycircularvirus* 9, Pacific flying faeces–associated gemycircularvirus 4 (PffaGmV–1), Capybara genomovirus 9 (CapGV1–9), Capybara genomovirus 11 (CapGV1–11) and Thrips–associated genomovirus 2 (Male et al., 2016; Kraberger et al., 2017; Fontenele et al., 2019). The current taxonomic rule for the classification of a new *Genomoviridae* species is the overall identity of 78% when compared to the already established species (Varsani and Krupovic, 2017; ICTV, 2020). All previously mentioned viruses (expect for *Pteropus–associated gemycircularvirus* 9) have not yet been accepted as formal species by the ICTV.

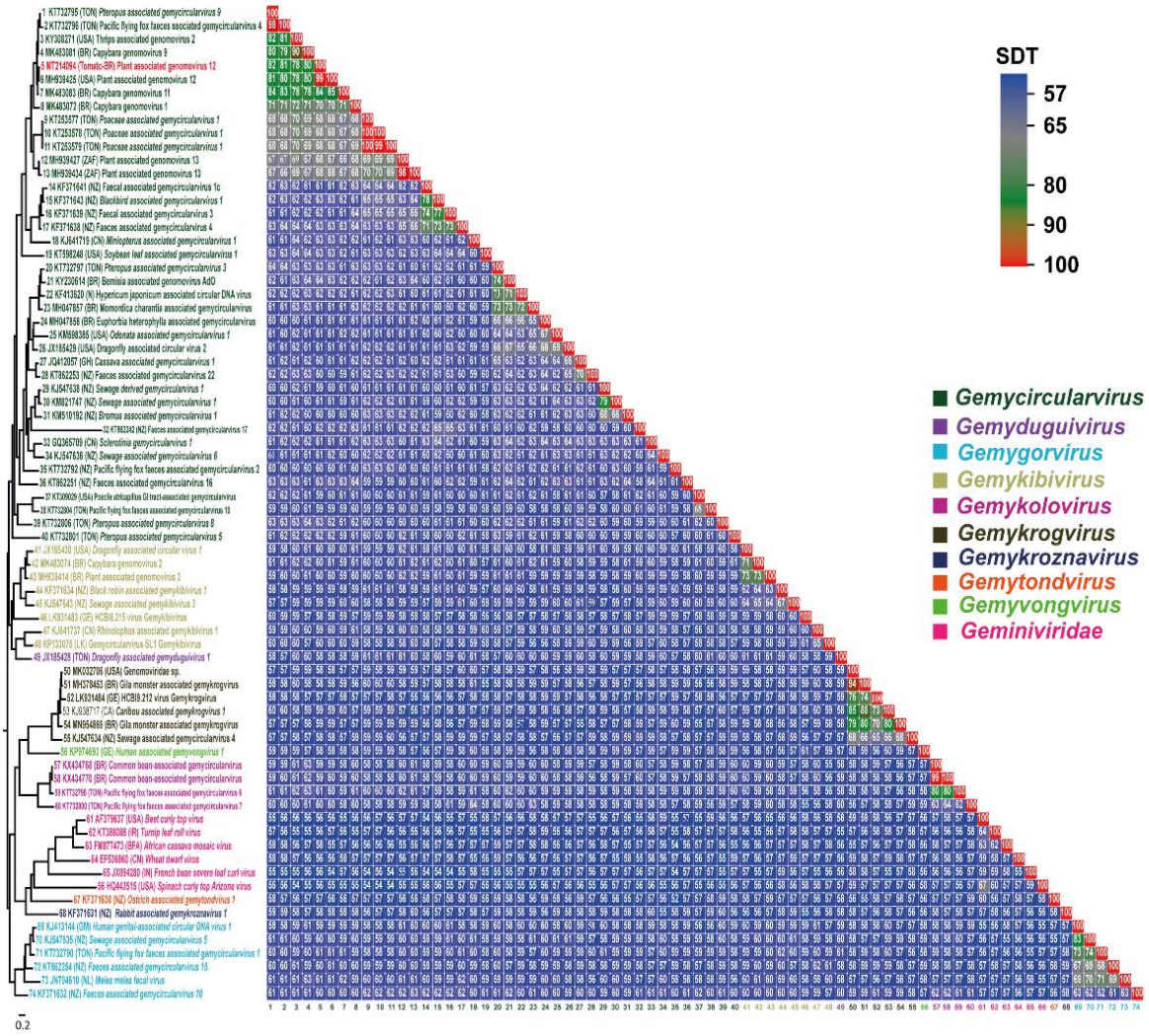

Figure 1. Phylogenetic and sequence demarcation tool (SDT) analyses using 74 representatives genomovirus sequences, including the Plant-associated genomovirus 12 isolate that was described in association with tomato leaf samples in the present work (highlighted in red font color). Bayesian phylogenetic tree was based upon the replication–associated (Rep) protein sequences. Sequences of geminiviruses were used as outgroups. The Rep coding sequences were aligned using MUSCLE, and phylogenic tree was constructed using Bayesian inference performed with MrBayes v3.2, with amino acid substitution model GTR + I + G selected by JModeltest v. 2.2. The analyzes were carried out by running 100 million generations and sampling every 2,000 generations after 2 million burn–in generation. Genome-wide pairwise matrix was generated by SDT v1.2. The isolates are identified by their name, by the GenBank accession number and by the acronym of the countries where they were described: BR = Brazil; TON = Tonga; USA = United States; ZAF = South Africa; NZ = New Zealand; CN = China; GH = Ghana; GE = Germany; LK = Sri Lanka; IR = Iran; BFA = Burkina Faso, and NL = Netherlands. GenBank accession numbers: 1. Pteropus associated gemycircularvirus 9 (KT732795);

2. Pacific flying fox faeces associated gemycircularvirus 4 (KT732796); 3. Thrips associated genomovirus 2 (KY308271); 4. Capybara Genomovirus 9 (MK483081); 5. Plant-associated genomovirus 12 (Tomato – MT214094); 6. Plant Genomovirus 12 (MH939425); 7. Capybara genomovirus 11 (MK483083); Capybara genomovirus 1 (MK483072); 9. Poaceae associated gemycircularvirus 1 (KT253577); 10. Poaceae associated gemycircularvirus 1 (KT253578);11. Poaceae associated gemycircularvirus 1 (KT253579);12. Plant Genomovirus 13 (MH939427); 13. Plant Genomovirus 13 (MH939434); 14. Faecal associated gemycircularvirus 1c (KF371641); 15. Blackbird associated gemycircularvirus 1 (KF371643); 16. Faecal associated gemycircularvirus 3 (KF371639); 17. Faeces associated gemycircularvirus 4 (KF371638); 18. Miniopterus associated gemycircularvirus 1 (KJ641719); 19. Soybean leaf associated gemycircularvirus 1 (KT598248); 20. Pteropus associated gemycircularvirus 3 (KT732797); 21. Bemisia associated genomovirus AdO (KY230614); 22. Hypericum japonicum associated circular DNA virus (KF413620); 23. Momordica charantia associated gemycircularvirus (MH047857); 24. Euphorbia heterophylla associated gemycircularvirus (MH047858); 25. Odonata associated gemycircularvirus 1 (KM598385); 26. Dragonfly associated circular virus 2 (JX185429); 27. Cassava associated gemycircularvirus 1 (JQ412057); 28. Faeces associated gemycircularvirus 22 (KT862253); 29. Sewage derived gemycircularvirus 1 (KJ547638); 30. Sewage associated gemycircularvirus 1 (KM821747); 31. Bromus associated gemycircularvirus 1 (KM510192); 32. Faeces associated gemycircularvirus 17 (KT862242); 33. Sclerotinia gemycircularvirus 1 (GQ365709); 34. Sewage associated gemycircularvirus 6 (KJ547636); 35. Pacific flying fox faeces associated gemycircularvirus 2 (KT732792); 36. Faeces associated gemycircularvirus 16 (KT862251); 37. Poecile atricapillus GI tract–associated gemycircularvirus (KT309029); 38. Pacific flying fox faeces associated gemycircularvirus 10 (KT732804); 39. Pteropus associated gemycircularvirus 8 (KT732806); 40. Pteropus associated gemycircularvirus 5 (KT732801); 41. Dragonfly associated circular virus 1 (JX185430); 42. Capybara genomovirus 2 (MK483074); 43. Plant associated genomovirus 2 (MH939414); 44. Black robin associated gemykibivirus 1 (KF371634); 45. Sewage associated gemykibivirus 3 (KJ547643); 46. HCBI8.215 virus Gemykibivirus (LK931483); 47. Rhinolophus associated gemykibivirus 1 (KJ641737); 48. Gemycircularvirus SL1 Gemykibivirus (KP133075); 49. Dragonfly associated gemyduguivirus 1 (JX185428); 50. Genomoviridae sp. (MK032706); 51. Gila monster associated gemykrogvirus (MH378453); 52. HCBI9.212 virus Gemykrogvirus (LK931484); 53. Caribou associated gemykrogvirus 1 (KJ938717); 54. Gila monster associated gemykrogvirus (MN954869); 55. Sewage associated gemycircularvirus 4 (KJ547634); 56. Human associated gemyvongvirus 1 (KP974693); 57. Common bean-associated gemycircularvirus (KX434768); 58. Common bean-associated gemycircularvirus (KX434770); 59. Pacific flying fox faeces associated gemycircularvirus 6 (KT732798); 60. Pacific flying fox faeces associated gemycircularvirus 7 (KT732800); 61. Beet curly top virus (AF379637); 62. Turnip leaf roll virus (KT388088); 63. African cassava mosaic virus (FM877473); 64. Wheat dwarf virus (EF536860); 65. French bean severe leaf

curl virus (JX094280); 66. Spinach curly top Arizona virus (HQ443515); 67. Ostrich associated gemytondvirus 1 (KF371630); 68. Rabbit associated gemykroznavirus 1 (KF371631); 69. Human genital-associated circular DNA virus 1 (KJ413144); 70. Sewage associated gemycircularvirus 5 (KJ547635); 71. Pacific flying fox faeces associated gemycircularvirus 1 (KT732790); 72. Faeces associated gemycircularvirus 15 (KT862254); 73. Meles meles fecal virus (JN704610); 74. Faeces associated gemycircularvirus 10 (KF371632).

The tomato–associated gemycircularvirus isolates displayed a genome of 2,189 nucleotides (nts) in size (**Figure 2**). The genome contains three ORFs: one capsid protein (CP) in the viral sense (with 906 nts) and two ORFs in the complementary sense (RepA with 735 nts and Rep with 1008 nts). These two last ORFs overlap at the end as observed across species of the genus *Gemycircularvirus*. In the intergenic region (with 128 nts), we annotated a "stem–loop" with the **TAATGTTAT** nonanucleotide sequence, typical of the genus *Gemycircularvirus*. Another peculiar feature is the presence of a "splicing" (intron) within the Rep, with a size of 124 nts (**Figure 2**), which is typical of the gemycircularviruses and very similar to that present in the genera *Becurtovirus*, *Capulavirus*, *Grablovirus* and *Mastrevirus* (*Geminiviridae* family) (Varsani & Krupovic, 2017; ICTV, 2018). Conserved motifs and domains present in the Rep protein were also identified. Motifs I, II, III, and GRS (*Geminivirus* Rep Sequence) and helicase motifs are conserved in all ssDNAs of the *Genomoviridae* family (Varsani and Krupovic, 2017). Here, the motif I (**LLTYAQ**) was found in the N–terminal Rep domain. This motif is very important for dsDNA binding during replication. We also found the motif II (**THLHV**), which is believed to be involved in the coordination of DNA cleavage, and the motif III (**YAVK**) which contains a tyrosine residue essential for dsDNA cleavage and assigns the covalent form at the 5' terminus of the cleaved DNA. Motifs I, II, and III are also present in the genomes of geminiviruses, nanoviruses, and circoviruses (Rosario et al., 2012b; Varsani and Krupovic, 2017). The GRS domain (**GHHPNITP**), located between motifs II and III, was also identified in the tomato–associated viral isolates, it has a peculiar function when compared to geminiviruses, since it is involved in the ssDNA cleavage reaction that occurs during the initiation of RCR. This domain is related to replication initiation across geminiviruses, and it more likely have the same function in genomoviruses (Nash et al., 2011; Varsani and Krupovic, 2017). The helicase domain (C–terminal terminus of Rep protein) could also be identified, which has the well–characterized motifs Walk A, Walk B, and C. The GRS domain is related to the denaturation of the double–

strand DNA during viral replication. In the tomato–associated gemycircularvirus the amino acid sequences of the motifs were as follow: Walk A (**LYGPSRLGKTVW**), Walk B (**YAVFDDMRG**) and C (**PDRNALWIC**). Walk A and Walk B motifs were first identified in enzymes that require the use of ATP. Walk A is part of the P–loop motif that is employed for ATP recognition (Walker et al., 1982; Rosario et al., 2012b). Walk B is responsible for ATP binding and hydrolysis, and the motif C interacts with ATP phosphate and water molecules through a conserved asparagine residue (Choudhury et al., 2006; George et al., 2014; Varsani and Krupovic, 2017).

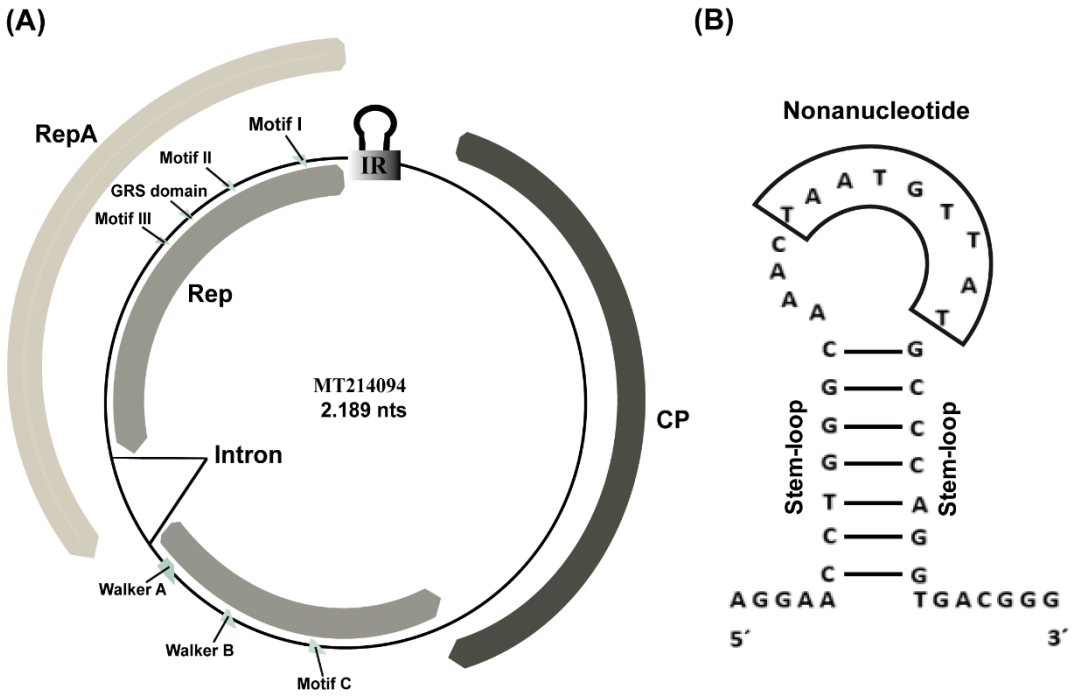

Figure 2. Diagrammatic representation of the genomic organization of an isolate of Plant-associated genomovirus 12 detected in natural association with open-field tomatoes in Central Brazil. **Panel (A):** The tomato–associated circular genome (GenBank MT214094) displayed 2,189 nucleotides (nts) in size. The genome contains three open reading frames (ORFs): one capsid protein (CP) in the viral sense (with 906 nts) and two ORFs in the complementary sense (RepA with 735 nts and Rep with 1008 nts). An intron is located within the ORF Rep. Arrows are indicating the location of the motifs I, II, III, and C as well as the GRS domain and the Walker A and B motifs. **Panel (B):** Intergenic region (with 128 nts) showing a conserved "stem–loop" which contains the nonanucleotide sequence **TAATGTTAT** (highlighted).

Recombination analyses for nucleotide sequence of the tomato–associated gemycircularvirus showed evidence of recombination events in four statistical methods: MaxChi (*p-value* = 1.890 *x* 10^{-9}), Chimera (*p-value* =1.374 *x* 10^{-5}), SiScan (*p-value* = 2.883 *x* 10^{-3}) and 3Seq (*p-value* = 3.883 *x* 10^{-7}). The analyzes showed that this gemycircularvirus closely resemble its major parent, which is an isolate of the Plant-associated genomovirus 14 (MH939452), whereas the minor parent is the Soybean leaf–associated gemycircularvirus 1 – SlaGemV1 (KT598248) (Marzano and Domier, 2016). The initial breakpoint is at nucleotide 458 and the final breakpoint is at nucleotide 1807 (involving sequences from the CP and Rep genes).

In conclusion, phylogenetic analyzes using the amino acid sequences of the replication–associated (Rep) protein from available *Genomoviridae* genomes confirmed that the tomato–associated virus (GenBank MT214094) as a member of the genus *Gemycircularvirus*. SDT analyzes also corroborated this conclusion. To our knowledge, this the first worldwide report of a *Gemycircularvirus* in natural association with either tomatoes or other Solanaceae species. This virus displayed a close genetic relationship (99% identity) to Plant-associated genomovirus 12, which was reported in samples of the creosote bush (*L. tridentata*; family Zygophyllaceae) that is a plant native from desertic areas of North America. Therefore, an intriguing aspect (which will require additional studies) is how viral isolates sharing high levels of genetic identity can interact with two taxonomically divergent plant species (from distinct botanic families) in such contrasting environments. In this context, further studies will necessary to elucidate the type of interaction of this virus with tomatoes as well as the potential effects (beneficial/detrimental) it may cause in this vegetable crop.

CHAPTER 6

Tomato golden net virus (ToGNV) and Tomato yellow net virus (ToYNV): Two novel begomoviruses from the Neotropics with monopartite genomes.

[1]Luciane de Nazaré Almeida dos Reis, [1,2]Leonardo S. Boiteux, [2]Maria Esther N. Fonseca, [1]Rita de Cássia Pereira–Carvalho.

[1]Departamento de Fitopatologia, Universidade de Brasília (UnB), Brasília – DF, Brazil.
[2]National Center for Vegetable Crops Research (CNPH), Embrapa Vegetable Crops (Hortaliças), Brasília – DF, Brazil.

Work submitted to Archives of Virology

Resumo

Dois novos begomovírus que infectam o tomateiro foram descobertos por sequenciamento Illumina no Brasil. Os genomas virais completos foram clonados e sequenciados por Sanger. Esses vírus foram detectados em campos comerciais no Brasil e exibiram DNA–A com a organização típica de espécies de begomovírus do Novo Mundo. As espécies foram provisoriamente denominadas Tomato golden net virus (ToGNV) e Tomato yellow net virus (ToYNV). A maioria dos begomovírus de áreas neotropicais possui genomas bipartidos. No entanto, nenhum componente cognato de DNA–B foi encontrado para ToGNV e ToYNV. Portanto, eles compreendem um grupo peculiar de begomovírus neotropicais com genomas monopartidos, que apresentam um enorme interesse biológico, molecular e genético.

Abstract

Two novel tomato-infecting begomoviruses were originally discovered via Illumina sequencing assays, subsequently they were cloned and Sanger–sequenced. These viruses were detected in commercial fields in Brazil and displayed DNA–A components with typical organization of New–World *Begomovirus* species. They were tentatively named as Tomato golden net virus (ToGNV) and Tomato yellow net virus (ToYNV). The majority of the begomoviruses from Neotropical areas has bipartite genomes. However, no cognate DNA–B components were found for ToGNV and ToYNV. Hence, they comprise a peculiar group of Neotropical begomoviruses with monopartite genomes, which have enormous biological, molecular, and genetic interest.

The *Geminiviridae* family (Order: *Geplafuvirales*) is composed by viruses with single-stranded DNA (ssDNA) genomes, which is currently composed of nine genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus*, and *Turnucurtovirus*. The classification at the genus level is based upon the host range, the type of insect(s) vector(s), genomic organization, and phylogenetic relationships (Loconsole et al., 2012; Ma et al., 2015; Varsani et al., 2017; Rojas et al., 2018; ICTV, 2020). The genus *Begomovirus* aggregates the largest number of species within the *Geminiviridae* family (ICTV, 2020). The *Begomovirus* genus is composed by whitefly-transmitted species with one (= monopartite) or two (= bipartite) circular, ssDNA genomic component(s) with ≈ 2.6 kb that are separately encapsulated into twinned particles formed by two incomplete icosahedrons

(Rojas et al., 2018). The begomovirus transmission is characterized as being non–propagative, circulative and is done by insects members of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex (Rojas et al., 2018).

The tomato (*Solanum lycopersicum* L.) crop is grown year-round under distinct cultivation systems across major tropical and subtropical regions. The first reports of tomato-infecting *Begomovirus* species in Brazil were done between the years 1950 and 1970 (Flores et al., 1960). However, begomovirus outbreaks in tomatoes increased after *B. tabaci Middle East-Asia Minor 1* (MEAM 1 = biotype B) invasion in the early 1990s. Field surveys conducted subsequently revealed a complex of *Begomovirus* species (composed mainly by bipartite viruses), occurring across all Brazilian biomes. So far, over 21 *Begomovirus* species have been reported naturally infecting tomatoes (Matyis et al., 1975; Ribeiro et al., 2003; Calegario et al., 2007; Cotrim et al., 2007; Ribeiro et al., 2007; Castillo-Urquiza et al., 2008; Fernandes et al., 2008; Albuquerque et al., 2010; Albuquerque et al., 2012; Macedo et al., 2018; Quadros et al., 2019; Rêgo-Machado et al., 2019). In addition, viruses associated with Malvaceae weeds were also reported infecting tomatoes (Calegario et al., 2004; Cotrim et al., 2007). Currently, Tomato severe rugose virus – ToSRV and Tomato mottle leaf curl virus – ToMoLCV are the most widespread and economically important begomoviruses in this crop. ToMoLCV was first erroneously described as being a bipartite species by Albuquerque et al. (2012) (Flores et al., 1960), but, is was subsequently characterized as a monopartite virus (Loconsole et al., 2012; Ma et al., 2015; Varsani et al., 2017; Rojas et al., 2018; ICTV, 2020). In fact, ToMoLCV is thus far the only monopartite species reported infecting tomatoes under Brazilian conditions. In the present study, Next-generation sequencing (NGS) was employed in combination with Sanger dideoxy sequencing for the identification and complete genome characterization of two new monopartite *Begomovirus* species.

As part of systematic field surveys for tomato-infecting viruses, leaf tissues were collected from tomato plants growing in commercial fields between 2003 and 2015 across three regions of Central Brazil [the Federal District (DF), Goiás State (GO), and Minas Gerais State (MG)]. The samples were collected from tomato plants with typical begomovirus symptoms such as foliar chlorosis (generating typical yellow and golden net patterns), leaf deformation, yellow mosaic, and dwarfism. These samples were subsequently subjected to total DNA extraction using the modified CTAB protocol with organic solvents (Boiteux et al., 1999) and stored at -20 °C. After confirming the presence of a begomovirus through PCR (polymerase chain reaction) using the degenerate primers ‘PAR1c496’/ ‘PAL1v1978’ (Rojas

et al., 1993), viral circular DNAs were amplified by rolling–circle amplification (RCA) (Inoue-Nagata et al., 2004). High–throughput sequencing (HTS) was then performed on an Illumina HiSeq2500 platform (Macrogen Korea). The HTS–derived information was analyzed according to the following workflow: (1) elimination of low–quality reads; (2) re–assembly of the sequences using the program CLC Genomics Workbench 10; and (3) validation of the contigs with the BLASTn algorithm by comparing with the ssDNA Genbank virus database (https://www.ncbi.nlm.nih.gov/). The viral contigs were annotated and the trimmed reads were mapped back to the annotated genome using the tool 'Map to reference' available in the Geneious 11.0 program.

As a result of the Illumina HiSeq 2500 sequencing, 32,691,974 million reads were obtained from the sample pool. After assembly, in the CLC Genomics Workbench 11 program, 19,487 contigs were obtained. Most of these contigs were composed by previously characterized *Begomovirus* species such as: Bean golden mosaic virus (BGMV), Cleome leaf crumple virus (CILCrV), Tomato severe rugose virus (ToSRV), Euphorbia yellow mosaic virus (EuYMV), Tomato chlorotic mottle virus (ToCMoV), Tomato golden vein virus (TGVV), Tomato mottle leaf curl virus (ToMoLCV), Tomato common mosaic virus (ToCmMV), Tomato rugose mosaic virus (ToRMV) and Sida micranta mosaic virus (SimMV). However, two full–length DNA–A contigs displayed identity levels of less than 91%. According to the current taxonomic criterion of the genus *Begomovirus* (Brown et al., 2015), both isolates were two putative new species.

The DNA–A genome of the initially named new species #1 (427,646 reads) displayed 84% identity to Tomato bright yellow mottle virus (ToBYMV), whereas the species #2 (2,839 reads) displayed 85% identity to Tomato chlorotic leaf curl virus (ToCLCV). The new species #1 was detected in a single sample with code MG–378, collected in 2015 in Itacarambi–MG and the new species #2 in two samples (GO–169 and GO–189) collected in 2003 in Leopoldo de Bulhões–GO and Goianópolis–GO, respectively. Open primers specific for the DNA–A component of each putative new viral species were designed. For the DNA–B detection, the universal primer pair 'PBL1v2040'/'PCRc1' (Rojas et al., 1993) were used. After exhaustive PCR assays the DNA–A of both putative new species was detected. However, no DNA–B component or DNA–B segment was amplified, indicating that they might be monopartite viruses.

The complete DNA – A sequence of both viruses was confirmed via Sanger dideoxy sequencing employing the specific primer pairs, Sequencing reactions were carried out at CNPH (Plant Breeding Lab) using the BigDye® Terminator Cycle Sequencing Ready Reaction Kit version 3.1 protocol (Applied Biosystems, São Paulo–SP, Brazil). The DNA – A components of the viruses were cloned into pSL1180 plasmid vector (Addgene). For the new species #1 (MG–378 sample), the RCA products were individually cleaved with *Apa*I. For putative new species #2 (GO–169), RCA products were cleaved with *Kpn*I. The vector pSL1180, previously cleaved with the corresponding specific enzyme, was introduced into *Escherichia coli* DH5α by transformation. Viral inserts were Sanger–sequenced at CNPH. Full–length genomes were assembled using Geneious 11.0. Sequences were initially analyzed using the BLASTn algorithm and sequence identity to the closest begomoviruses was calculated with Species Demarcation Tool v.1.2 (SDT) (Martin et al., 2015). Full–length genomes were aligned with MUSCLE. Phylogenetic trees based on DNA–A alignments were generated by Maximum likelihood phylogenies – PhyML (Guindon et al., 2009) with HKY + I+ G nucleotide substitution model selected by jModelTest (Posada, 2008) with 1,000 bootstrap replications. Figures were elaborated with Adobe Illustrator CC and EvolView (He et al., 2016). To detect potential recombination events, the software RDP 4.77 program (Martin et al., 2015) was used. Recombination events were considered reliable only if they were detected by at least four of the seven methods implemented by the program.

The DNA – A of species #1 was detected in sample MG–378, collected in 2015 in Itacarambi–MG. The name Tomato golden net virus–ToGNV (GenBank MT214095) was proposed and displayed a genome with 2,649 nucleotides (nts). Species #2 was detected in two samples (GO–169 and GO–189) and the proposed name Tomato yellow net virus–ToYNV (GenBank MT214096) presenting a genome with of 2,636 nts. ToGNV and ToYNV showed genomic organization typical of New World begomoviruses with the DNA–A encoding one Coat protein (CP) in the virion sense strand, and four in the complementary–sense strand Replication associated protein (Rep), Trans–acting protein (TrAP), Replication enhancer (REn) and possible symptom determinant, and the silencing suppression gene (C4) (**Figure 1A**). All components have the conserved nonanucleotide (5'–**TAATATTAC**–3') located at the origin of replication. In the intergenic region (IR) in addition to the nonanucleotide, the analyzes identified in Rep the iterons: **TGGTGTC** for TGNV (Rep IRD = **MPRNQNSFRLA**) and **GGAGT** for ToYNV (Rep IRD = **MPRNPKSFRLQ**) **(Figure 1B)** (Argüello-Astorga and Ruiz-Medrano, 2001).

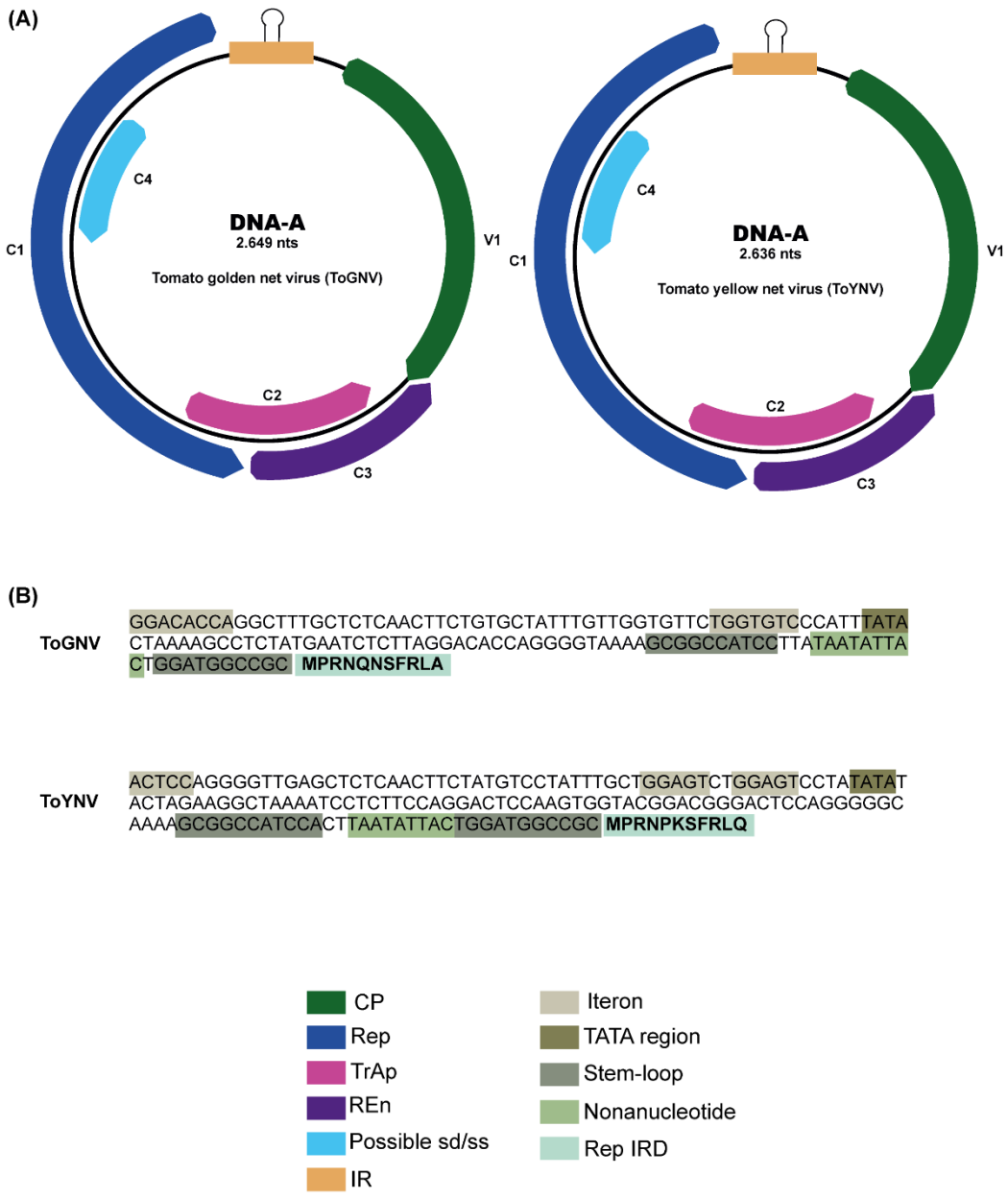

Figure 1. Genomic organization of the two new tomato-infecting monopartite *Begomovirus* species. **Panel A:** Diagrammatic representation of the circular genomes of Tomato golden net virus (ToGNV) and Tomato yellow net virus (ToYNV) and their respective open reading frames (ORFs). The ORFs AV1, AC1, AC2, AC3 and AC4 are color-coded according to the putative function of their protein products. CP = capsid protein; Rep = replication-associated protein; TrAp = transactivator protein; Ren = replication enhancer; sd = possible symptom determinant; ss = possible silencing suppressor; IR = intergenic region, encompassing the hairpin; **Panel B:** A segment of the intergenic region showing iterons, TATA region, nonanucleotide, stem-loop and at the end Rep = IRD (Rep Iteron-Related Domain).

The pairwise nucleotide sequence identities of DNA–A of two new species and other begomoviruses were calculated using the SDT. The analysis showed that the ToGNV shared in general 71–85% with other begomoviruses, whereas the closest identity levels of ToYNV ranged from 74% to 85% (**Figure 2**). Phylogenetic analyzes showed that Tomato bright yellow mottle virus (ToBYMoV – KC791691) and Tomato golden leaf spot virus (ToGLSV – KC626021) were the viruses with closest genetic relationship (85% identity) to ToGNV. On the other hand, Abutilon Brazil virus (AbBV – FN434438), Abutilon mosaic Brazil virus (AbMBV – JF694482) and ToCLCV (MK558058) were the viral species with the closest genetic relationship to ToYNV (**Figure 2**).

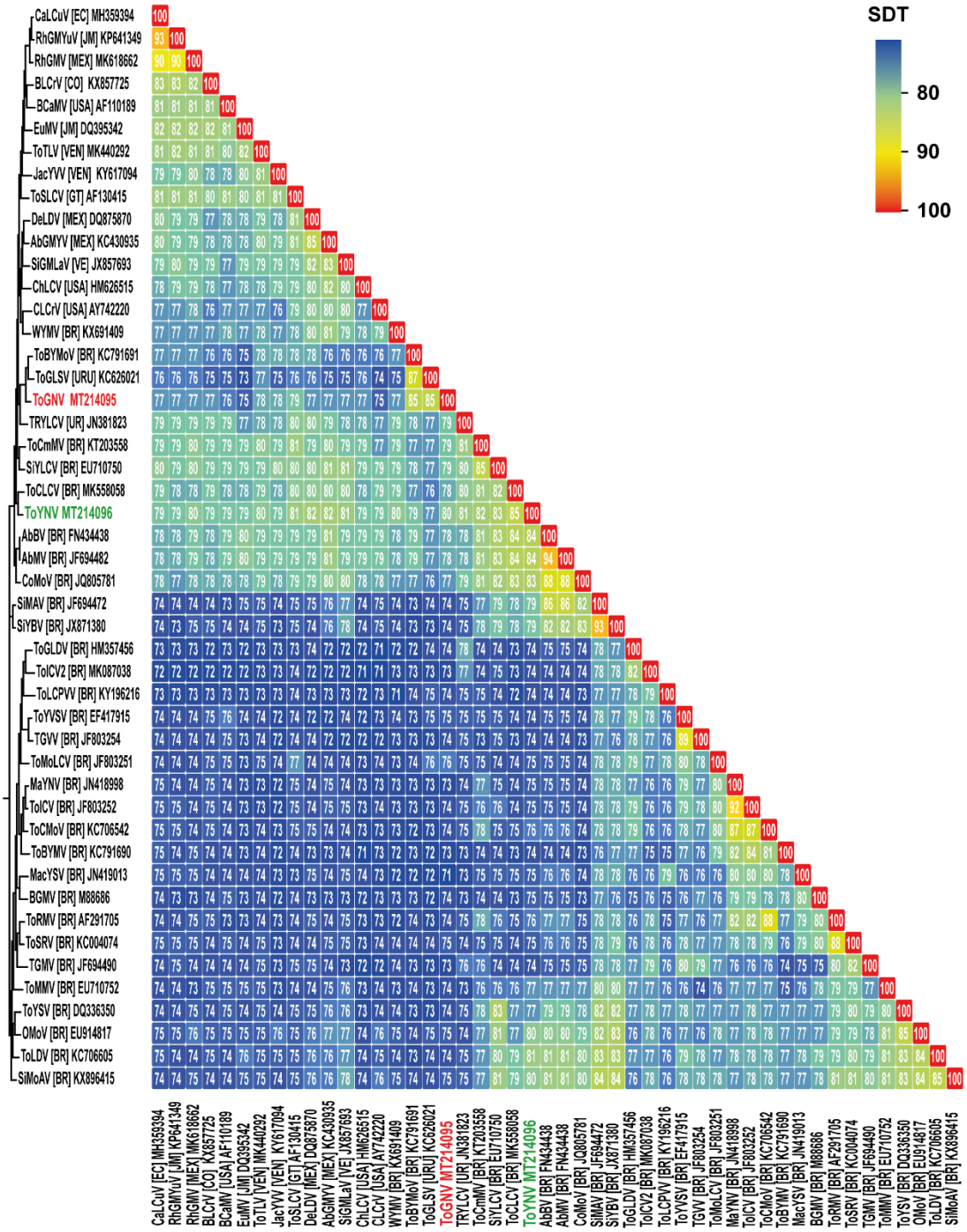

Figure 2. Pairwise identity in Sequence Demarcation Tool (SDT) analysis carried out using the information of the DNA–A sequences of selected New World *Begomovirus* species showing their phylogenetic identities/distances with two new tomato-infecting species: Tomato golden net virus – ToGNV = MT214095 (in red) and Tomato yellow net virus – ToYNV = MT214096 (in green). These *Begomovirus* species were identified by their accession number and by the acronym of the countries where they were described: BR = Brazil; URU = Uruguay; EC = Ecuador; JM = Jamaica; MEX = Mexico; CO = Colombia; USA = United States; VEN = Venezuela; GT = Guatemala. Species and

GenBank accession numbers: Cabbage leaf curl virus – CaLCuV (MH359394); Rhynchosia golden mosaic Yucatan virus – RhGMYuV (KP641349); Rhynchosia golden mosaic Sinaloa virus – RhGMV (MK618662); Bean leaf crumple virus – BLCrV (KX857725); Bean calico mosaic virus – BcaMV (AF110189); Euphorbia mosaic virus – EuMV (DQ395342); Tomato twisted leaf virus – ToTLV (MK440292); Jacquemontia yellow vein virus – JacYVV (KY617094); Tomato severe leaf curl virus – ToSLCV (AF130415); Desmodium leaf distortion virus – DeLDV (DQ875870); Abutilon golden mosaic Yucatan virus – AbGMYV (KC430935); Sida golden mosaic Lara virus – SiGMLaV (JX857693); Chenopodium leaf curl virus – ChLCV (HM626515); Cotton leaf crumple virus – CLCrV (AY742220); Wissadula yellow mosaic virus – WYMV (KX691409); Tomato bright yellow mottle virus – ToBYMoV (KC791691); Tomato golden leaf spot virus – ToGLSV (KC626021); Tomato rugose yellow leaf curl virus – TRYLCV (JN381823); Tomato common mosaic virus – ToCmMV (KT203558); Sida yellow leaf curl virus – SiYLCV (EU710750); Tomato chlorotic leaf curl virus – ToCLCV (MK558058); Abutilon Brazil virus – AbBV (FN434438); Abutilon mosaic virus - AbMV (JF694482); Corchorus mottle virus – CoMoV (JQ805781); Sida mosaic Alagoas virus – SiMAV (JF694472); Sida yellow blotch virus – SiYBV (JX871380); Tomato golden leaf distortion virus – ToGLDV (HM357456); Tomato interveinal chlorosis virus2 – ToICV2 (MK087038); Tomato leaf curl purple vein virus – ToLCPVV (KY196216); Tomato yellow vein streak virus – ToYVSV (EF417915); Tomato golden mosaic virus – TGVV (JF80325); Tomato mottle leaf curl virus – ToMoLCV (JF803251); Macroptilium yellow net virus – MaYNV (JN418998); Tomato interveinal chlorosis virus – ToICV (JF803252); Tomato chlorotic mottle virus – ToCMoV (KC706542); Tomato bright yellow mosaic virus – ToBYMV (KC791690); Macroptilium yellow spot virus – MacYSV (JN419013); Bean golden mosaic virus – BGMV (M88686); Tomato rugose mosaic virus – ToRMV (AF2917050; Tomato severe rugose virus – ToSRV (KC004074); Tomato golden mosaic virus – TGMV (JF694490); Tomato mild mosaic virus – ToMMV (EU710752); Tomato yellow spot virus – ToYSV (DQ336350); Okra mottle virus – OMoV (EU914817); Tomato leaf distortion virus – ToLDV (KC706605); Sida mottle Alagoas virus – SiMoAV (KX896415).

Recombination analyzes performed with the RDP 4 program, showed no evidence of recombination events for ToYNV. However, for ToGNV, there was evidence of recombination events in seven RDP statistical methods (*p-value* = 1.747×10^{-13}), GENECONV (*p-value* = 5.362×10^{-17}), BootScan (*p-value* = 1.943×10^{-11}), MaxChi (p-value = 9.102×10^{-8}), Chimaera (*p-value* = 2.360×10^{-7}), SiScan (*p-value* = 7.445×10^{-8}), 3Seq (p-value = 4.509×10^{-9}). The analyzes showed that this species closely resemble the major parent which is an isolate of ToBYMoV, and the minor parent is Tomato severe rugose virus (ToSRV). The initial

breakpoint is at nucleotide 1270 and the final breakpoint is at nucleotide 1476 involving sequences from Rep, TrAP and REn genes.

The majority of tomato-infecting begomoviruses from Neotropical areas has bipartite genomes. In our exhaustive set of analyzes, we were unable to detect cognate DNA–B components in association with both ToYNV and ToGNV, indicating that they more likely represent new monopartite begomoviruses. Currently ToMoLCV and ToLCPVV are the only two tomato-infecting monopartite species reported infecting tomatoes in Brazil, which have the typical genomic organization of New World begomoviruses (Ferro et al., 2017; Macedo et al., 2018). Our overall analyses of the genomes of all four tomato-infecting monopartite species (ToMoLCV, and ToLCPVV plus ToYNV and ToGNV) indicated that they are devoid of the V2 movement protein which is present across the Old World monopartite begomoviruses (Rojas et al., 2018). These observations suggest that the New–World monopartite species may employ a distinct mechanism to perform this function. A study with the Old–World monopartite species Tomato yellow leaf curl virus (TYLCV), indicated that the proteins V1 (CP) and C4 may also be involved in the systemic viral movement within the tomato cells (Rojas et al., 2001). Therefore, it is possible that V1 and C4 proteins display analogous function to the BC1 protein responsible for cell–to–cell movement in bipartite begomoviruses (Rojas et al., 2001). ToYNV was detected in species–specific PCR assays as occurring in multiple infections with the bipartite species ToSRV and TGVV (in both GO–169 and GO–189 samples), indicating that this novel virus may share the DNA–B component with either one or both of these tomato-infecting viruses. Conversely, no co-infection was detected in association with ToGNV (sample MG–378), suggesting that it might be a *bona fide* monopartite virus. In conclusion, ToYNV and ToGNV comprise a peculiar group of Neotropical monopartite begomoviruses, which convey enormous biological and molecular interest. Additional characterization studies with infectious clones of these two new viruses are necessary to elucidate their interaction with susceptible as well as resistant tomato cultivars and in relation to their host range.

GENERAL CONCLUSIONS

- Using NGS combined with metagenomics, it was possible to catalog the begomovirus diversity in samples of susceptible (without *Ty*–1) and resistant (with *Ty*–1) tomatoes, collected in the Southeast (Minas Gerais) and Midwest regions (Goiás and Distrito Federal) of Brazil.

- Tomato severe rugose virus (ToSRV), Tomato golden vein virus (TGVV), Tomato chlorotic mottle virus (ToCMoV), Tomato mottle leaf curl virus (ToMoLCV) and Tomato rugose mosaic virus (ToRMV) were the predominant species in both pools of RCA samples.

- Mixed infections were predominant in most of the tomato samples from both resistant and susceptible pools.

- Three new *Begomovirus* species were found, one with a bipartite genome in a resistant tomato (DF-640) with the proposed name of Tomato mosaic severe dwarf virus (ToMSDV). And two others in susceptible samples MG-378, GO-169 and GO-189, possibly containing monopartite genomes, which were provisionally named Tomato golden net virus (ToGNV) and Tomato yellow net virus (ToYNV) respectively.

- A new Alfasatellite was identified in susceptible samples from the Federal District, always in mixed infection with ToCMoV, TGVV and ToMoLCV. The tentatively name was New world bipartite associated alphasatellite.

- Sequence analyzes of Tomato yellow vein streak virus (ToYVSV) and Tomato golden vein virus (TGVV) species indicated that both are closely related. Our reassessment showed the need for a reclassification of misnamed isolates available at GenBank.

- The diversity of 161 isolates classified as BGMV indicated that they comprise two distinct species: BGMV isolates from *Phaseolus vulgaris* and from a wide range of hosts, and a closely related species (with identity ranging from 89 to 91% wotj the reference BGMV isolate) which are occurring mainly in *Phaseolus lunatus*. For this new viral species from *P. lunatus*, we suggest a previously proposed name – Lima bean golden mosaic virus (LBGMV).

- In the present work, a *Gemycircularvirus* was identified in association with two tomato samples collected in production fields in Goiás. The species was identified and tentatively named as Plant-associated genomovirus 12, also identified in *Larrea tridentata*. This is the first worldwide report of a gemycircularvirus in natural association with tomato.

REFERENCES

Abouzid AM, Polston J, Hiebert E. 1992. The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from tomatoes in Florida. Journal of General Virology, 73: 3225-3229.

Accotto GP, Sardo L. 2009. Transovarial transmission of begomoviruses in *Bemisia tabaci*. In: Bemisia: Bionomics and Management of a Global Pest. Springer, pp. 339-345.

Adams I, Fox A. 2016. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis. In: Current research topics in plant virology. Springer, pp. 323-335.

Adams IP, Glover R, Monger W, Mumford R, Jackeviciene E, Navalinskiene M, Samuitiene M, Boonham N. 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Molecular Plant Pathology, 10: 537-545.

Agindotan BO, Domier LL, Bradley CA. 2015. Detection and characterization of the first North American mastrevirus in switchgrass. Archives of Virology, 160: 1313-1317.

Al Rwahnih M, Alabi O, Westrick N, Golino D, Rowhani A. 2016. Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. Phytopathology, 107: 240-251.

Al Rwahnih M, Daubert S, Golino D, 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, Alabi OJ, Westrick NM, Golino D. 2018. Prunus geminivirus A: a novel Grablovirus infecting Prunus spp. Plant Disease: PDIS-09-17-1486-RE.

Albuquerque LC, Martin D, Avila A, Inoue-Nagata A. 2010. Characterization of tomato yellow vein streak virus, a begomovirus from Brazil. Virus Genes, 40: 140-147.

Albuquerque LC, Varsani A, Fernandes F, Pinheiro B, Martin D, Ferreira P, Lemos T, Inoue-Nagata A. 2012. Further characterization of tomato-infecting begomoviruses in Brazil. Archives of Virology, 157: 747-752.

Alvarenga M, Coelho F. 2013. Valor nutricional, p. p. 23-29. In: Alvarenga MAR (ed.), Tomate: produção em campo, casa de vegetação e hidroponia, 2 ed, Universidade de Lavras.

Anbinder I, Reuveni M, Azari R, Paran I, Nahon S, Shlomo H, Chen L, Lapidot M, Levin I. 2009. Molecular dissection of Tomato leaf curl virus resistance in tomato line TY172 derived from Solanum peruvianum. Theoretical and Applied Genetics, 119: 519-530.

Andrade E, Manhani G, Alfenas P, Calegario R, Fontes E, Zerbini F. 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.

Anfoka G, Al-Talb M, Haj-Ahmad F. 2016. A new isolate of tomato yellow leaf curl Axarquia virus associated with tomato yellow leaf curl disease in Jordan. Journal of Plant Pathology: 145-149.

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

Argüello-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.

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

Arruabarrena A, Rubio L, González-Arcos M, Maeso D, Fiallo-Olivé E, Moriones E. 2016. First Report of the *Begomovirus* Tomato yellow vein streak virus Infecting Tomato in Uruguay. Plant Disease, 100: 231-231.

Attathom S, Chiemsombat P, Kositratana W, Sae-Ung N. 1994. Complete nucleotide sequence and genome analysis of bipartite tomato yellow leaf curl virus in Thailand. Kasetsart Journal (Natural Science), 28: 632-639.

Bagewadi B, Naidu RA. 2016. First report of Tomato yellow leaf curl Kanchanaburi virus in eggplant and tomato in Cambodia. Plant Disease, 100: 233-233.

Bahjatnia S, Izadpanah K, Barry DI, Rezaian M. 2004. Molecular characterization and taxonomic position of the Iranian isolate of Tomato leaf curl virus. Iranian Journal of Plant Pathology 40.

Bananej K, Kheyr-Pour A, Salekdeh GH, Ahoonmanesh A. 2004. Complete nucleotide sequence of Iranian tomato yellow leaf curl virus isolate: further evidence for natural recombination amongst begomoviruses. Archives of Virology, 149: 1435-1443.

Baráth D, Jaksa-Czotter N, Molnár J, Varga T, Balássy J, Szabó LK, Kirilla Z, Tusnády GE, Preininger É, Várallyay É. 2018. Small RNA NGS revealed the presence of Cherry virus A and Little cherry virus 1 on apricots in Hungary. Viruses, 10: 318.

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

Barba M, Hadidi A. 2015. An overview of plant pathology and application of next-generation sequencing technologies. CAB Reviews, 10: 1-21.

Barbosa JC, Albuquerque L, Rezende J, Inoue-Nagata A, Bergamin Filho A, Costa H. 2016. Occurrence and molecular characterization of Tomato common mosaic virus (ToCmMV) in tomato fields in Espírito Santo state, Brazil. Tropical Plant Pathology, 41: 62-66.

Barbosa LF, Marubayashi J, De Marchi B, Yuki V, Pavan M, 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.

Barbosa LF, Yuki V, Marubayashi J, De Marchi B, Perini F, Pavan M, de Barros D, Ghanim M, Moriones E, Navas-Castillo J. 2015. First report of *Bemisia tabaci* Mediterranean (Q biotype) species in Brazil. Pest Management Science, 71: 501-504.

Barreto S, Hallwass M, Aquino O, Inoue-Nagata A. 2013. A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. Phytopathology, 103: 436-444.

Batista J, Melo F, Pereira-Carvalho R, Alves-Freitas D, Ribeiro S. 2019. First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil. Plant Disease, 103: 1443.

Bedford ID, Briddon RW, Brown JK, Rosell R, Markham PG. 1994. Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. Annals of Applied Biology, 125: 311-325.

Belabess Z, Dallot S, El-Montaser S, Granier M, Majde M, Tahiri A, Blenzar A, Urbino C, Peterschmitt M. 2015. Monitoring the dynamics of emergence of a non-canonical recombinant of Tomato yellow leaf curl virus and displacement of its parental viruses in tomato. Virology, 486: 291-306.

Bernardo P, Charles-Dominique T, Barakat M, Orlet P, Fernadez E, Filloux D, Hatnady P, Rebelo TA, Cousin. S. R, Mesleard F, Cohez D, Yavercovski N, Roumagnac P. 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. The ISME Journal 12: 173-184.

Bernardo P, Muhire B, François S, Deshoux M, Hartnady P, Farkas K, Kraberger S, Filloux D, Fernandez E, Galzi S. 2016. Molecular characterization and prevalence of two capulaviruses: Alfalfa leaf curl virus from France and Euphorbia caput-medusae latent virus from South Africa. Virology, 493: 142-153.

Bian XY, Thomas M, Rasheed M, Saeed M, Hanson P, De Barro P, Rezaian M. 2007. A recessive allele (*tgr*-1) conditioning tomato resistance to geminivirus infection is associated with impaired viral movement. Phytopathology, 97: 930-937.

Boiteux L, Fonseca M, Giordano L, Melo P. 2012a. Melhoramento genético. Produção de tomate para processamento industrial. Brasília: Embrapa: 31-50.

Boiteux L, Fonseca M, Simon P. 1999. Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. Journal of the American Society for Horticultural Science, 124: 32-38.

Boiteux L, Fonseca M, Viera J, Pereira-Carvalho R. 2012b. Melhoramento para resistência a doenças virais. Melhoramento de Plantas para Condições de Estresses Bióticos. Visconde de Rio Branco, MG: Suprema: 89-119.

Boiteux L, Oliveira V, Silva C, Makishima N, Inoue-Nagata A, Fonseca M, Giordano L. 2007a. Reaction of tomato hybrids carrying the *Ty*-1 locus to Brazilian bipartite *Begomovirus* species. Horticultura Brasileira, 25: 20-23.

Boiteux L, Pereira-Carvalho R, Inoue-Nagata A, Fonseca M, Resende R, Fernández-Muñoz R. 2007b. Reação de acessos de tomateiro portando o gene *Ty*-2 (introgredido de *Solanum habrochaites* f. glabratum) a um isolado de begomovírus de genoma bipartido. Horticultura Brasileira, 25: S97.

Bömer M, Rathnayake AI, Visendi P, Silva G, Seal SE. 2018. Complete genome sequence of a new member of the genus *Badnavirus*, Dioscorea bacilliform RT virus 3,

reveals the first evidence of recombination in yam badnaviruses. Archives of Virology, 163: 533-538.

Bornancini VA, Irazoqui JM, Flores CR, Vaghi Medina CG, Amadio AF, López Lambertini PM. 2020. Reconstruction and Characterization of Full-Length *Begomovirus* and Alphasatellite Genomes Infecting Pepper through Metagenomics. Viruses, 12: 202.

Bosco D, Mason G, Accotto G. 2004. TYLCSV DNA, but not infectivity, can be transovarially inherited by the progeny of the whitefly vector *Bemisia tabaci* (Gennadius). Virology, 323: 276-283.

Boykin LM, De Barro PJ. 2014. A practical guide to identifying members of the *Bemisia tabaci* species complex: and other morphologically identical species. Frontiers in Ecology and Evolution, 2: 45.

Briddon RW, Bull SE, Amin I, Idris AM, Mansoor S, Bedford ID, Dhawan P, Rishi N, Siwatch SS, Abdel-Salam AM. 2003. Diversity of DNA β, a satellite molecule associated with some monopartite begomoviruses. Virology, 312: 106-121.

Briddon RW, Bull SE, Amin I, Mansoor S, Bedford ID, Rishi N, Siwatch SS, Zafar Y, Abdel-Salam AM, Markham PG. 2004. Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus–DNA β complexes. Virology, 324: 462-474.

Briddon RW, Heydarnejad J, Khosrowfar F, Massumi H, Martin DP, Varsani A. 2010. Turnip curly top virus, a highly divergent geminivirus infecting turnip in Iran. Virus Research, 152: 169-175.

Briddon RW, Martin DP, Roumagnac P, Navas-Castillo J, Fiallo-Olivé E, Moriones E, Lett J-M, Zerbini FM, Varsani A. 2018. *Alphasatellitidae*: A new family with two subfamilies for the classification of geminivirus-and nanovirus-associated alphasatellites. Archives of Virology, 163: 2587-2600.

Brown JK, Fauquet C, Briddon R, Zerbini F, Moriones E, Navas-Castillo J. 2012. *Geminiviridae*. In: Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses, King AMQA, M.J.; Carstens, E.B. & Lefkowitz, E.J (ed.) Associated Press, Elsevier Inc., London, Waltham, San Diego, pp. 351-373.

Brown JK, Ostrow KM, Idris AM, Stenger DC. 2000. Chino del tomate virus: relationships to other begomoviruses and identification of A-component variants that affect symptom expression. Phytopathology, 90: 546-552.

Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JC, Fiallo-Olivé E, Briddon RW, Hernández-Zepeda C, Idris A. 2015. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Archives of Virology, 160: 1593-1619.

Bull SE, Briddon RW, Sserubombwe WS, Ngugi K, Markham PG, Stanley J. 2007. Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. Journal of General Virology, 88: 1624-1633.

Butterbach P, Verlaan MG, Dullemans A, Lohuis D, Visser RG, Bai Y, Kormelink R. 2014. Tomato yellow leaf curl virus resistance by *Ty*-1 involves increased cytosine methylation of viral genomes and is compromised by Cucumber mosaic virus infection. Proceedings of the National Academy of Sciences, 111: 12942-12947.

Calegario R, Ferreira S, Andrade C, Zerbini F. 2004. Caracterização de um isolado do begomovírus Sida micrantha mosaic virus (SiMMV) obtido de tomateiro. Fitopatologia Brasileira, 29: 150.

Calegario R, Ferreira S, Andrade E, Zerbini F. 2007. Characterization of Tomato yellow spot virus, a novel tomato-infecting begomovirus in Brazil. Pesquisa Agropecuária Brasileira, 42: 1335-1343.

Cantú-Iris M, Pastor-Palacios G, Mauricio-Castillo J, Bañuelos-Hernández B, Avalos-Calleros J, Juárez-Reyes A, Rivera-Bustamante R, Argüello-Astorga G. 2019. Analysis of a new begomovirus unveils a composite element conserved in the CP gene promoters of several *Geminiviridae* genera: Clues to comprehend the complex regulation of late genes. PLoS One, 14: e0210485.

Cao M, Lan P, Li F, Abad J, Zhou C, Li R. 2017. Genome characterization of sweet potato symptomless virus 1: a mastrevirus with an unusual nonanucleotide sequence. Archives of Virology, 162: 2881-2884.

Caro M, Verlaan MG, Julián O, Finkers R, Wolters A-MA, Hutton SF, Scott JW, Kormelink R, Visser RG, Díez MJ. 2015. Assessing the genetic variation of *Ty*-1 and *Ty*-3 alleles conferring resistance to Tomato yellow leaf curl virus in a broad tomato germplasm. Molecular Breeding, 35: 132.

Castillo-Urquiza G, Beserra J, Bruckner F, Lima A, Varsani A, Alfenas-Zerbini P, Zerbini F. 2008. Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. Archives of Virology, 153: 1985-1989.

Castillo AG, Collinet D, Deret S, Kashoggi A, Bejarano ER. 2003. Dual interaction of plant PCNA with geminivirus replication accessory protein (Ren) and viral replication protein (Rep). Virology, 312: 381-394.

Chabi-Jesus C, Najar A, Fontenele RS, Kumari SG, Ramos-González PL, Freitas-Astúa J, Kraberger S, Varsani A. 2020. Viruses representing two new genomovirus species identified in citrus from Tunisia. Archives of virology: 1-5.

Chakraborty S, Pandey P, Banerjee M, Kalloo G, Fauquet C. 2003. Tomato leaf curl Gujarat virus, a new begomovirus species causing a severe leaf curl disease of tomato in Varanasi, India. Phytopathology, 93: 1485-1495.

Chatchawankanphanich O, Maxwell DP. 2002. Tomato leaf curl Karnataka virus from Bangalore, India, appears to be a recombinant begomovirus. Phytopathology, 92: 637-645.

Chaudhary A, Kumar R, Mukherjee S. 2012. Tomato leaf curl new Delhi Virus 2. [Genbank]

Chen K, Khatabi B, Fondong VN. 2019. The AC4 protein of a cassava geminivirus is required for virus infection. Molecular Plant-Microbe Interactions, 32: 865-875.

Chiang BT, Nakhla M, Maxwell D, Schoenfelder M, Green S. 1997. A new geminivirus associated with a leaf curl disease of tomato in Tanzania. Plant Disease, 81: 111-111.

Chiumenti M, Greco C, Antelmi I, Sion V, Altamura G, Nigro F, Saldarelli P. 2019. Molecular characterisation of a novel gemycircularvirus associated with olive trees in Italy. Virus Research, 263: 169-172.

Choudhury NR, Malik PS, Singh DK, Islam MN, Kaliappan K, Mukherjee SK. 2006. The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. Nucleic Acids Research, 34: 6362-6377.

Chowda R, Colvin J, Muniyappa V .2004. Molecular characterization of a begomovirus from tomato at Pune, Maharastra, India. [Genbank]

Claverie S, Bernardo P, Kraberger S, Hartnady P, Lefeuvre P, Lett J-M, Galzi S, Filloux D, Harkins GW, Varsani A. 2018. From spatial metagenomics to molecular characterization of plant viruses: A geminivirus case study. In: Advances in virus research, vol. 101. Elsevier, pp. 55-83.

Clérot D, Bernardi F. 2006. DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus. Journal of Virology, 80: 11322-11330.

Conceição-Neto N, Zeller M, Heylen E, Lefrère H, Mesquita JR, Matthijnssens J. 2015. Fecal virome analysis of three carnivores reveals a novel nodavirus and multiple gemycircularviruses. Virology Journal, 12: 79.

Cooper J, Jones A. 1983. Responses of plants to viruses: proposals for the use of terms. Phytopathology, 73: 127-128.

Costa A. 1965. Three whitefly-transmitted virus diseases of beans in São Paulo, Brazil. FAO Plant Protection Bull, 13: 1-12.

Costa A, Bennett C. 1950. Whitefly transmitted mosaic of *Euphorbia prunifolia*. Phytopathology, 40.

Cotrim MAA, Krause-Sakate R, Narita N, Zerbini FM, Pavan MA. 2007. Diversidade genética de begomovírus em cultivos de tomateiro no Centro-Oeste Paulista. Summa Phytopathologica, 33: 300-303.

Cuong HV, Le Van Hai TNT, Hao NB. 2011. Molecular characterization of Tomato leaf curl Hainan virus and Tomato leaf curl Hanoi virus in Vietnam. J. ISSAAS, 17: 70-82.

Cupertino FP, De Sá PB, Kitajima EW. 1991. Propriedades biológicas de um isolado do vírus do anel do pimentão causando faixa amarela em tomateiros no Distrito Federal. Fitopatologia Brasileira, 10: 251-256.

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

Czotter N, Molnar J, Szabó E, Demian E, Kontra L, Baksa I, Szittya G, Kocsis L, Deak T, Bisztray G. 2018. NGS of virus-derived small RNAs as a diagnostic method used to determine viromes of Hungarian vineyards. Frontiers in Microbiology, 9: 122.

Da Silva Assis MR, Vieira CB, Fioretti JM, Rocha MS, de Almeida PIN, Miagostovich MP, Fumian TM. 2016. Detection and molecular characterization of gemycircularvirus from environmental samples in Brazil. Food and environmental virology, 8: 305-309.

Daniels J, Castro L. 1985. Ocorrencia do virus do mosaico deformante da batata no Rio Grande do Sul. Fitopatologia Brasileira, 10: 306.

Dayaram A, Opong A, Jäschke A, Hadfield J, Baschiera M, Dobson RC, Offei SK, Shepherd DN, Martin DP, Varsani A. 2012. Molecular characterisation of a novel cassava associated circular ssDNA virus. Virus Research, 166: 130-135.

De Barro P, Liu S-S, Boykin L, Dinsdale A. 2011. Bemisia tabaci: a statement of species status. Annual Review of Entomology, 56: 1-19.

De Rezende RR, Mar TB, Páez LMC, da Silva Xavier A, Xavier CAD, Navas-Castillo J, Zerbini FM, Alfenas-Zerbini P. 2018. Complete genome sequences of two gemycircularviruses associated with non-cultivated plants in Brazil. Archives of Virology, 163: 3163-3166.

Debat H, Zavallo D, Brisbane RS, Vončina D, Almeida RP, Blouin AG, Al Rwahnih M, Gomez-Talquenca S, Asurmendi S. 2019. Grapevine virus L: a novel vitivirus in grapevine. European Journal of Plant Pathology, 155: 319-328.

Delatte H, Martin DP, Naze F, Goldbach R, Reynaud B, Peterschmitt M, Lett J-M. 2005. South West Indian Ocean islands tomato begomovirus populations represent a new major monopartite begomovirus group. Journal of General Virology, 86: 1533-1542.

Dhaliwal M, Jindal S, Sharma A, Prasanna H. 2019. Tomato yellow leaf curl virus disease of tomato and its management through resistance breeding: a review. The Journal of Horticultural Science and Biotechnology: 1-20.

Ding M, Li T, Fang Q, Zhang Z, Zhou X. 2016. Tomato yellow leaf curl Yunnan virus, a new begomovirus species associated with tomato yellow leaf curl disease in China. Journal of Plant Pathology: 337-340.

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

Dolores L, Bajet N 1995, posting date. Occurrence of tomato yellow leaf curl virus in the Philippines. [Online.]

Dry IB, Krake LR, Rigden JE, Rezaian MA. 1997. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. Proceedings of the National Academy of Sciences, 94: 7088-7093.

Du Z, Tang Y, Zhang S, She X, Lan G, Varsani A, He Z. 2014. Identification and molecular characterization of a single-stranded circular DNA virus with similarities to Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1. Archives of Virology, 159: 1527-1531.

Duffy S, Holmes EC. 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. Journal of Virology, 82: 957-965.

Edwards RA, Rohwer F. 2005. Viral metagenomics. Nature Reviews Microbiology, 3: 504-510.

FAO.2020. Food and agriculture organization of the United Nations (FAO). [http://faostat3.fao.org]. Accessed on 6 February, 2020.

Faria JC, Aragão F, Souza T, Quintela E, Kitajima E, Ribeiro SG. 2016. Golden mosaic of common beans in Brazil: management with a transgenic approach. Embrapa Arroz e Feijão-Artigo em periódico indexado (ALICE).

Faria JC, Souza-Dias J, Slack S, Maxwell D. 1997. A New Geminivirus Associated with Tomato in the State of São Paulo, Brazil. Plant Disease 81: 423-423.

Faria JC, Maxwell DP. 1999. Variability in geminivirus isolates associated with Phaseolus spp. in Brazil. Phytopathology, 89: 262-268.

Fauquet C, Briddon R, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X. 2008. Geminivirus strain demarcation and nomenclature. Archives of Virology, 153: 783-821.

Fernandes NAN. 2015. Begomoviroses no cultivo do tomateiro no brasil: variabilidade e caracterização de novas espécies virais e diversidade do vetor bemisia tabaci. Thesis.

Fernandes FR, Albuquerque L, Giordano L, Boiteux L, Avila A, Inoue-Nagata A. 2008. Diversity and prevalence of Brazilian bipartite *Begomovirus* species associated to tomatoes.Virus Genes 36: 251–258. .

Fernandes JJ, Carvalho M, Andrade E, Brommonschenkel S, Fontes E, Zerbini F. 2006. Biological and molecular properties of Tomato rugose mosaic virus (ToRMV), a new tomato-infecting begomovirus from Brazil. Plant Pathology, 55: 513-522.

Fernandes FR, Albuquerque LC, de Oliveira CL, Cruz AR, da Rocha WB, Pereira TG, Naito FY, Dias NdM, Nagata T, Faria JC. 2011. Molecular and biological characterization of a new Brazilian begomovirus, Euphorbia yellow mosaic virus (EuYMV), infecting Euphorbia heterophylla plants. Archives of Virology, 156: 2063.

Fernandes FR, Cruz A, Faria J, Zerbini F, Aragao FJ. 2009. Three distinct begomoviruses associated with soybean in central Brazil. Archives of virology, 154: 1567-1570.

Ferro MM, Ramos-Sobrinho R, Silva JT, Assunção IP, Lima GS. 2017. Genetic structure of populations of the begomoviruses Tomato mottle leaf curl virus and Sida mottle Alagoas virus infecting tomato (*Solanum lycopersicum*) and Sida spp., respectively. Tropical Plant Pathology, 42: 39-45.

Fiallo-Olivé E, Martínez-Zubiaur Y, Moriones E, Navas-Castillo J. 2012. A novel class of DNA satellites associated with New World begomoviruses. Virology, 426: 1-6.

Fiallo-Olivé E, Martínez-Zubiaur Y, Rivera-Bustamante R. 2009. Tomato yellow leaf distortion virus, a new bipartite begomovirus infecting tomato in Cuba. Plant Pathology, 58: 785-785.

Fiallo-Olivé E, Tovar R, Navas-Castillo J. 2016. Deciphering the biology of deltasatellites from the New World: maintenance by New World begomoviruses and whitefly transmission. New Phytologist, 212: 680-692.

Flores E, Silberschmidt K, Kramer M. 1960. Observações de "clorose infecciosa" das malváceas em tomateiros do campo. O Biológico, 26: 65-69.

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

Fonseca M, Boiteux L, Costa A .2016. Tomato rugose yellow leaf curl virus (*Geminiviridae*) in tomatoes associated with biological invasion of *Bemisia tabaci* mediterranean (biotype Q) into subtropical Brazil. [Genbank]

Fonseca M, Boiteux LS .2013. Tomato golden leaf spot virus infecting tomato plants. [Genbank]

Fonseca M, Boiteux LS, Fernandes N. 2010. Tomato golden leaf distortion virus, a new begomovirus infecting tomato in Tocantins state, Brazil. [Genbank]

Fonseca M, Boiteux LS, Fernandes N. 2011. Genomic organization of Chino del tomate Amazonas virus a bipartite begomovirus endemic to the Amazon region of Brazil. [Genbank]

Fonseca M, Boiteux LS, Fernandes N. 2013. Tomato bright yellow mosaic virus and Tomato bright yellow mottle virus: two new tomato-infecting begomoviruses in Brazil. [Genbank]

Fontenele RS, Abreu RA, Lamas NS, Alves-Freitas DM, Vidal AH, Poppiel RR, Melo FL, Lacorte C, Martin DP, Campos MA. 2018a. Passion fruit chlorotic mottle virus: molecular characterization of a new divergent geminivirus in Brazil. Viruses, 10: 169.

Fontenele RS, Alves-Freitas DM, Silva PI, Foresti J, Silva PR, Godinho MT, Varsani A, Ribeiro SG. 2018b. Discovery of the first maize-infecting mastrevirus in the Americas using a vector-enabled metagenomics approach. Archives of Virology, 163: 263-267.

Fontenele RS, Lacorte C, Lamas NS, Schmidlin K, Varsani A, Ribeiro SG. 2019. Single stranded DNA viruses associated with capybara faeces sampled in Brazil. Viruses, 11: 710.

Fontenele RS, Lamas NS, Lacorte C, Lacerda ALM, Varsani A, Ribeiro SG. 2017. A novel geminivirus identified in tomato and cleome plants sampled in Brazil. Virus Research, 240: 175-179.

Fontenelle MR, Luz DF, Gomes APS, Florentino LH, Zerbini FM, Fontes EP. 2007. Functional analysis of the naturally recombinant DNA-A of the bipartite begomovirus Tomato chlorotic mottle virus. Virus Research, 126: 262-267.

Fuentes A, Carlos N, Ruiz Y, Callard D, Sánchez Y, Ochagavía ME, Seguin J, Malpica-López N, Hohn T, Lecca MR. 2016. Field trial and molecular characterization of RNAi-transgenic tomato plants that exhibit resistance to tomato yellow leaf curl geminivirus. Molecular Plant-Microbe Interactions, 29: 197-209.

García-Andrés S, Tomás DM, Sánchez-Campos S, Navas-Castillo J, Moriones E. 2007. Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. Virology, 365: 210-219.

García-Arenal F, Fraile A, Malpica JM. 2003. Variation and evolution of plant virus populations. International Microbiology, 6: 225-232.

García-Cano E, Resende R, Boiteux L, Giordano L, Fernández-Muñoz R, Moriones E. 2008. Phenotypic expression, stability, and inheritance of a recessive resistance to monopartite begomoviruses associated with tomato yellow leaf curl disease in tomato. Phytopathology, 98: 618-627.

George B, Ruhel R, Mazumder M, Sharma VK, Jain SK, Gourinath S, Chakraborty S. 2014. Mutational analysis of the helicase domain of a replication initiator protein reveals critical roles of Lys 272 of the B′ motif and Lys 289 of the β-hairpin loop in geminivirus replication. Journal of General Virology, 95: 1591-1602.

Gilbertson RL, Batuman O, Webster CG, Adkins S. 2015. Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. Annual Review of Virology, 2: 67-93.

Gilbertson RL, Faria JC, Ahlquist P, Maxwell DP. 1993. Genetic diversity in geminiviruses causing bean golden mosaic disease: the nucleotide sequence of the infectious cloned DNA components of a Brazilian isolate of bean golden mosaic geminivirus. Phytopathology, 83: 709-709.

Gilbertson RL, Faria JC, Hanson SF, Morales FJ, Ahlquist P, Maxwell DP, Russell DR. 1991. Cloning of the complete DNA genomes of four bean-infecting geminiviruses and determining their infectivity by electric discharge particle acceleration. Phytopathology, 81: 980-985.

Gill U, Scott JW, Shekasteband R, Ogundiwin E, Schuit C, Francis DM, Sim S-C, Smith H, Hutton SF. 2019. *Ty*-6, a major begomovirus resistance gene on chromosome 10, is effective against Tomato yellow leaf curl virus and Tomato mottle virus. Theoretical and Applied Genetics, 132: 1543-1554.

Giordano L, Fonseca MS, JBC, Inoue-Nagata AB, LS 2005a. Efeito da infecção precoce por Begomovirus com genoma bipartido em características de frutos de tomate industrial. Horticultura Brasileira 23: 815-818.

Giordano L, Silva-Lobo V, Santana F, Fonseca M, Boiteux L. 2005b. Inheritance of resistance to the bipartite Tomato chlorotic mottle begomovirus derived from Lycopersicon esculentum cv.'Tyking'. Euphytica, 143: 27-33.

Glasa M, Šoltys K, Predajňa L, Sihelská N, Budiš J, Mrkvová M, Kraic J, Mihálik D, Ruiz-García AB. 2019. High-throughput sequencing of Potato virus M from tomato in Slovakia reveals a divergent variant of the virus. Plant Protection Science, 55: 159-166.

Gnanasekaran P, KishoreKumar R, Bhattacharyya D, Vinoth Kumar R, Chakraborty S. 2019. Multifaceted role of geminivirus associated betasatellite in pathogenesis. Molecular Plant Pathology, 20: 1019-1033.

González I, Martínez L, Rakitina DV, Lewsey MG, Atencio FA, Llave C, Kalinina NO, Carr JP, Palukaitis P, Canto T. 2010. Cucumber mosaic virus 2b protein subcellular targets and interactions: their significance to RNA silencing suppressor activity. Molecular Plant-Microbe Interactions, 23: 294-303.

Gopal P, Kumar PP, Sinilal B, Jose J, Yadunandam AK, Usha R. 2007. Differential roles of C4 and βC1 in mediating suppression of post-transcriptional gene silencing: evidence for transactivation by the C2 of Bhendi yellow vein mosaic virus, a monopartite begomovirus. Virus Research, 123: 9-18.

Granier M, Tomassoli L, Manglli A, Nannini M, Peterschmitt M, Urbino C. 2019. First report of TYLCV-IS141, a Tomato yellow leaf curl virus recombinant infecting tomato plants carrying the Ty-1 resistance gene in Sardinia (Italy). Plant Disease, 103: 1437-1437.

Guerrero EB, De Francesco A, García M, Balatti P, Bó ED. 2013. First Report of Tomato rugose yellow leaf curl virus Infecting Tomato in Argentina. Plant Disease, 97: 1662-1662.

Guindon S, Delsuc F, Dufayard J-F, Gascuel O. 2009. Estimating maximum likelihood phylogenies with PhyML. In: Bioinformatics for DNA sequence analysis. Springer, pp. 113-137.

Guo Q, Shu YN, Liu C, Chi Y, Liu YQ, Wang XW. 2019. Transovarial transmission of tomato yellow leaf curl virus by seven species of the *Bemisia tabaci* complex indigenous to China: Not all whiteflies are the same. Virology, 531: 240-247.

Gutierrez C. 2002. Strategies for geminivirus DNA replication and cell cycle interference. Physiological and Molecular Plant Pathology, 60: 219-230.

Gutierrez C, Ramirez-Parra E, Castellano MM, Sanz-Burgos AP, Luque A, Missich R. 2004. Geminivirus DNA replication and cell cycle interactions. Veterinary Microbiology, 98: 111-119.

Ha C, Coombs S, Revill P, Harding R, Vu M, Dale J. 2008. Molecular characterization of begomoviruses and DNA Satellites from Vietnam: additional evidence that the new world geminiviruses were present in the old world prior to continental separation. Journal of General Virology, 89: 312-326.

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

Hadidi A. 2019. Next-Generation Sequencing and CRISPR/Cas13 editing in viroid research and molecular diagnostics. Viruses, 11: 120.

Hadidi A, Barba M. 2012. Next-generation sequencing: Historical perspective and current applications in plant virology. Petria, 22: 262-277.

Halary S, Duraisamy R, Fancello L, Monteil-Bouchard S, Jardot P, Biagini P, Gouriet F, Raoult D, Desnues C. 2016. Novel single-stranded DNA circular viruses in pericardial fluid of patient with recurrent pericarditis. Emerging infectious diseases, 22: 1839.

Hamera S, Song X, Su L, Chen X, Fang R. 2012. Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. The Plant Journal, 69: 104-115.

Hamilton W, Stein V, Coutts R, Buck K. 1984. Complete nucleotide sequence of the infectious cloned DNA components of tomato golden mosaic virus: potential coding regions and regulatory sequences. The EMBO Journal, 3: 2197-2205.

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.

Hanson P, Green S, Kuo G. 2006. Ty-2, a gene on chromosome 11 conditioning geminivirus resistance in tomato. Report of the Tomato Genetics Cooperative 56: 17-18.

Hanssen IM, Lapidot M, Thomma BP. 2010. Emerging viral diseases of tomato crops. Molecular Plant-Microbe Interactions, 23: 539-548.

Hasanvand V, Kamali M, Heydarnejad J, Massumi H, Kvarnheden A, Varsani A. 2018. Identification of a new turncurtovirus in the leafhopper Circulifer haematoceps and the host plant species Sesamum indicum. Virus Genes, 54: 840-845.

Hasegawa DK, Chen W, Zheng Y, Kaur N, Wintermantel WM, Simmons AM, Fei Z, Ling K-S. 2018. Comparative transcriptome analysis reveals networks of genes activated in the whitefly, *Bemisia tabaci* when fed on tomato plants infected with Tomato yellow leaf curl virus. Virology, 513: 52-64.

Hassan I, Orílio AF, Fiallo-Olivé E, Briddon RW, Navas-Castillo J. 2016. Infectivity, effects on helper viruses and whitefly transmission of the deltasatellites associated with sweepoviruses (genus *Begomovirus*, family *Geminiviridae*). Scientific Reports, 6: 1-12.

He ZF, Yu H, Luo F. 2005. Molecular characteristics of DNA-A of tomato leaf curl Guangdong virus isolate G2. Wei sheng wu xue bao= Acta Microbiologica Sinica, 45: 48-52.

He Z, Zhang H, Gao S, Lercher MJ, Chen W-H, Hu S. 2016. Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Research, 44: W236-W241.

Hernández-Zepeda C, Varsani A, Brown JK. 2013. Intergeneric recombination between a new, spinach-infecting curtovirus and a new geminivirus belonging to the genus *Becurtovirus*: first New World exemplar. Archives of Virology, 158: 2245-2254.

Heydarnejad J, Mozaffari A, Massumi H, Fazeli R, Gray AJ, Meredith S, Lakay F, Shepherd DN, Martin DP, Varsani A. 2009. Complete sequences of tomato leaf curl Palampur virus isolates infecting cucurbits in Iran. Archives of Virology, 154: 1015-1018.

Ho T, Tzanetakis IE. 2014. Development of a virus detection and discovery pipeline using next generation sequencing. Virology, 471: 54-60.

Holguín-Peña R, Arguello-Astorga G, Brown J, Rivera-Bustamante R. 2006. A new strain of Tomato chino La Paz virus associated with a leaf curl disease of tomato in Baja California Sur, Mexico. Plant Disease, 90: 973-973.

Holguín-Peña R, Juárez RV, Rivera-Bustamante R. 2004. Pepper golden mosaic virus affecting tomato crops in the Baja California Peninsula, Mexico. Plant disease, 88: 221-221.

Hoogstraten R, Hanson S, Maxwell D. 1996. Mutational analysis of the putative nicking motif in the replication-associated protein (AC1) of bean golden mosaic geminivirus. Molecular Plant-Microbe Interactions: MPMI, 9: 594-599.

Hutton S, Scott J. 2014. *Ty*-6, a major begomovirus resistance gene located on chromossome 10. Research Reports. , 64: 14-16.

Hutton S, Scott J, Schuster D. 2012. Recessive resistance to Tomato yellow leaf curl virus from the tomato cultivar Tyking is located in the same region as Ty-5 on chromosome 4. HortScience, 47: 324-327.

IBGE. 2020. Levantamento sistemático da produção agrícola - Pesquisa mensal de previsão e acompanhamento das safras agrícolas (Abril/2020). Instituto Brasileiro de Geografia e Estatística. [https://sidra.ibge.gov.br/]. Accessed on 10 April, 2020.

ICTV. 2020. International Committee on Taxonomy of Viruses. [https://talk.ictvonline.org/]. Accessed on 10 May, 2020.

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

Idris AM, Brown J. 2005. Evidence for interspecific-recombination for three monopartite begomoviral genomes associated with the tomato leaf curl disease from central Sudan. Archives of Virology, 150: 1003-1012.

Idris MA, Shahid MS, Briddon RW, Khan A, Zhu J-K, Brown JK. 2011. An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. Journal of General Virology, 92: 706-717.

Inoue-Nagata AK, Albuquerque L, Rocha W, Nagata T. 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage φ29 DNA polymerase. Journal of Virological Methods, 116: 209-211.

Inoue-Nagata AK, Lima MF, Gilbertson RL. 2016a. A review of geminivirus diseases in vegetables and other crops in Brazil: current status and approaches for management. Horticultura Brasileira, 34: 8-18.

Inoue-Nagata AK, Lopes C, Reis A, Pereira R, Quezao-Durval A, Pinheiro J, LIMA M. 2016b. Doenças do Tomateiro. In: Manual de Fitopatologia Amorim LBF, A.; Rezende, J. A. M.; Camargo, L. E. A. 5ª ed. Ceres. p. 697-729.

Ito T, Suzaki K, Nakano M, Sato A. 2013. Characterization of a new apscaviroid from American persimmon. Archives of Virology, 158: 2629-2631.

Ji Y, Scott J. 2006. *Ty*-3, a begomovirus resistance locus linked to *Ty*-1 on chromosome 6 of tomato. Tomato Genetics Cooperative 56: 22-25.

Ji Y, Scott J, Schuster D. 2009. Molecular mapping of *Ty*-4, a new Tomato yellow leaf curl virus resistance locus on chromosome 3 of tomato. Journal of the American Society for Horticultural Science 134: 281–288.

Jo Y, Song MK, Choi H, Park J-S, Lee J-W, Lian S, Lee BC, Cho WK. 2017. Genome Sequence of Grapevine Virus T, a Novel Foveavirus Infecting Grapevine. Genome announcements, 5: e00995-00917.

Jones JB, Zitter TA, Momol TM, Miller SA. 2014. Compendium of tomato diseases and pests. 1:15-119.

Jung J, Kim HJ, Lee JM, Oh CS, Lee H-J, Yeam I. 2015. Gene-based molecular marker system for multiple disease resistances in tomato against Tomato yellow leaf curl virus, late blight, and verticillium wilt. Euphytica, 205: 599-613.

Kadirvel P, De la Peña R, Schafleitner R, Huang S, Geethanjali S, Kenyon L, Tsai W, Hanson P. 2013. Mapping of QTLs in tomato line FLA456 associated with resistance to a virus causing tomato yellow leaf curl disease. Euphytica, 190: 297-308.

Kathurima T, Ateka E, Nyende A, Holton T. 2016. The rolling circle amplification and next generation sequencing approaches reveal genome wide diversity of Kenyan cassava mosaic geminivirus. African Journal of Biotechnology, 15: 2045-2052.

Kaur N, Chen W, Zheng Y, Hasegawa DK, Ling K-S, Fei Z, Wintermantel WM. 2017. Transcriptome analysis of the whitefly, Bemisia tabaci MEAM1 during feeding on tomato infected with the crinivirus, Tomato chlorosis virus, identifies a temporal shift in gene expression and differential regulation of novel orphan genes. BMC Genomics, 18: 370.

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

Khan AJ, Akhtar S, Singh AK, Al-Shehi AA, Al-Matrushi AM, Ammara U, Briddon RW. 2014. Recent evolution of a novel begomovirus causing tomato leaf curl disease in the Al-Batinah region of Oman. Archives of Virology, 159: 445-455.

Kheyr-Pour A, Bendahmane M, Matzeit V, Accotto GP, Crespi S, Gronenborn B. 1991. Tomato yellow leaf curl virus from sardinia is a whitefly-transmitted monoparatite geminivirus. Nucleic Acids Research, 19: 6763-6769.

King AM, Lefkowitz E, Adams MJ, Carstens EB. 2011. Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses, vol. 9. Elsevier.

Kitajima EW. 2020. An annotated list of plant viruses and viroids described in Brazil (1926-2018). Biota Neotropica, 20.

Knief C. 2014. Analysis of plant microbe interactions in the era of next generation sequencing technologies. Frontiers in Plant Science, 5: 216.

Knierim D, Maiss E. 2007. Application of Phi29 DNA polymerase in identification and full-length clone inoculation of tomato yellow leaf curl Thailand virus and tobacco leaf curl Thailand virus. Archives of Virology, 152: 941-954.

Kon T, Rojas M, Gilbertson R. 2009. Molecular characterization of tomato-infecting Begomovirus from Nigeria. [Genbank]

Kon T, Hidayat SH, Hase S, Takahashi H, Ikegami M. 2006. The natural occurrence of two distinct begomoviruses associated with DNAβ and a recombinant DNA in a tomato plant from Indonesia. Phytopathology, 96: 517-525.

Kraberger S, Argüello-Astorga GR, Greenfield LG, Galilee C, Law D, Martin DP, Varsani A. 2015. Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. Infection, Genetics and Evolution, 31: 73-86.

Kraberger S, Polston JE, Capobianco HM, Alcalá-Briseño RI, Fontenele RS, Varsani A. 2017. Genomovirus genomes recovered from Echinothrips americanus sampled in Florida, USA. Genome Announc., 5: e00445-00417.

Krenz B, Thompson JR, Fuchs M, Perry KL. 2012. Complete genome sequence of a new circular DNA virus from grapevine. Journal of Virology, 86: 7715-7715.

Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, Simon R. 2009. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology, 388: 1-7.

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

Krupovic M, Ghabrial SA, Jiang D, Varsani A. 2016. Genomoviridae: a new family of widespread single-stranded DNA viruses. Archives of Virology, 161: 2633-2643.

Kulshreshtha A, Kumar Y, Roshan P, Bhattacharjee B, Mukherjee SK, Hallan V. 2019. AC4 protein of tomato leaf curl Palampur virus is an RNA silencing suppressor and a pathogenicity determinant. Microbial Pathogenesis, 135: 103636.

Kumar RV. 2019. Plant antiviral immunity against geminiviruses and viral counter-defense for survival. Frontiers in Microbiology, 10.

Kumar A, Kumar J, Khan J. 2010. Sequence characterization of cotton leaf curl virus from Rajasthan: phylogenetic relationship with other members of geminiviruses and detection of recombination. Virus Genes, 40: 282-289.

Kumar J, Kumar J, Singh S, Tuli R. 2014. Association of satellites with a mastrevirus in natural infection: complexity of wheat dwarf India virus disease. Journal of Virology, 88: 7093-7104.

Kumar SP, Patel SK, Kapopara RG, Jasrai YT, Pandya HA. 2012. Evolutionary and molecular aspects of Indian tomato leaf curl virus coat protein. International Journal of Plant Genomics, 2012.

Kumar RV, Singh A, Singh A, Yadav T, Basu S, Kushwaha N, Chattopadhyay B, Chakraborty S. 2015. Complexity of begomovirus and betasatellite populations associated with chilli leaf curl disease in India. Journal of General Virology, 96: 3143-3158.

Kumar RV, Singh D, Singh A, Chakraborty S. 2017. Molecular diversity, recombination and population structure of alphasatellites associated with begomovirus disease complexes. Infection, Genetics and Evolution, 49: 39-47.

Kumari P, Chattopadhyay B, Singh A, Chakraborty S. 2009. A new begomovirus species causing tomato leaf curl disease in Patna, India. Plant Disease, 93: 545-545.

Lamas NS, Fontenele RS, Melo FL, Costa AF, Varsani A, Ribeiro SG. 2016. Complete genome sequence of a genomovirus associated with common bean plant leaves in Brazil. Genome Announcements, 4: e01247-01216.

Lamberto I, Gunst K, Müller H, zur Hausen H, de Villiers E-M. 2014. Mycovirus-like DNA virus sequences from cattle serum and human brain and serum samples from multiple sclerosis patients. Genome Announcements, 2: e00848-00814.

Lapidot M, Karniel U, Gelbart D, Fogel D, Evenor D, Kutsher Y, Makhbash Z, Nahon S, Shlomo H, Chen L. 2015. A novel route controlling begomovirus resistance by the messenger RNA surveillance factor pelota. PLoS Genetics, 11: e1005538.

Lefeuvre P, Delatte H, Naze F, Dogley W, Reynaud B, Lett JM. 2007a. A new tomato leaf curl virus from the Seychelles archipelago. Plant Pathology, 56: 342-342.

Lefeuvre P, Lett JM, Varsani A, Martin DP. 2009. Widely conserved recombination patterns among single-stranded DNA viruses. Journal of Virology, 83: 2697-2707.

Lefeuvre P, Martin D, Hoareau M, Naze F, Becker N, Dellate H, Reynaud B, Lett J. 2007b. *Begomovirus* 'melting pot' in the South West Indian Ocean Islands: Molecular diversity and evolution through recombination. [Genbank]

Lefeuvre P, Moriones E. 2015. Recombination as a motor of host switches and virus emergence: geminiviruses as case studies. Current Opinion in Virology, 10: 14-19.

Leichtfried T, Dobrovolny S, Reisenzein H, Steinkellner S, Gottsberger RA. 2019. Apple chlorotic fruit spot viroid: a putative new pathogenic viroid on apple characterized by next-generation sequencing. Archives of Virology, 164: 3137-3140.

Lemos P, Almeida M, Bastos C, Inoue-Nagata A. 2010. Avaliação do efeito da begomovirose na qualidade do fruto de tomate para processamento industrial. Horticultura Brasileira, 28: 1142-1147.

Lett JM, De Bruyn A, Hoareau M, Ouattara A, Claverie S, Dalmon A, Laplace D, Lefeuvre P, Hostachy B. 2015. Tomato chlorotic mottle Guyane virus: a novel tomato-infecting bipartite begomovirus from French Guiana. Archives of Virology, 160: 2887-2890.

Li F, Xu X, Huang C, Gu Z, Cao L, Hu T, Ding M, Li Z, Zhou X. 2015a. The AC5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant

that suppresses RNA silencing-based antiviral defenses. New Phytologist, 208: 555-569.

Li R, Gao S, Hernandez A, Wechter W, Fei Z, Ling K-S. 2012. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. PloS One, 7: e37127.

Li W, Gu Y, Shen Q, Yang S, Wang X, Wan Y, Zhang W. 2015b. A novel gemycircularvirus from experimental rats. Virus Genes, 51: 302-305.

Li ZH, Zhou X, Zhang X, Xie Y. 2004. Molecular characterization of tomato-infecting begomoviruses in Yunnan, China. Archives of Virology, 149: 1721-1732.

Liang P, Navarro B, Zhang Z, Wang H, Lu M, Xiao H, Wu Q, Zhou X, Di Serio F, Li S. 2015. Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. Journal of General Virology, 96: 2411-2420.

Lima MF, Bezerra I, Ribeiro S, Ávila A. 2001. Distribuição de geminivírus nas culturas do tomate e pimentão em doze municípios do Submédio do Vale São Francisco. Fitopatologia Brasileira, 26: 81-85.

Lima TMA, Sobrinho RR, Gonzalez-Aguilera J, Rocha CS, Silva SJ, Xavier CA, Silva FN, Duffy S, Zerbini FM. 2013. Synonymous site variation due to recombination explains higher genetic variability in begomovirus populations infecting non-cultivated hosts. Journal of General Virology, 94: 418-431.

Liu Q, Wang H, Ling Y, Yang SX, Wang XC, Zhou R, Xiao YQ, Chen X, Yang J, Fu WG. 2020. Viral metagenomics revealed diverse CRESS-DNA virus genomes in faeces of forest musk deer. Virology Journal, 17: 1-9.

Loconsole G, Önelge N, Potere O, Giampetruzzi A, Bozan O, Satar S, De Stradis A, Savino V, Yokomi R, Saponari M. 2012. Identification and characterization of Citrus yellow vein clearing virus, a putative new member of the genus Mandarivirus. Phytopathology, 102: 1168-1175.

Lopes CA, Reis A. 2011. Doenças do tomateiro cultivado em ambiente protegido. In. Corcórdia: Embrapa Hortaliças, Brasília – DF, pp. Circular técnica, 100.

Lozano G, Trenado HP, Fiallo-Olivé E, Chirinos D, Geraud-Pouey F, Briddon RW, Navas-Castillo J. 2016. Characterization of non-coding DNA satellites associated with sweepoviruses (genus *Begomovirus*, Geminiviridae)–definition of a distinct class of begomovirus-associated satellites. Frontiers in Microbiology, 7: 162.

Lu QY, Wu ZJ, Xia ZS, Xie LH. 2015. Complete genome sequence of a novel monopartite geminivirus identified in mulberry (*Morus alba* L.). Archives of Virology, 160: 2135-2138.

Luria N, Smith E, Reingold V, Bekelman I, Lapidot M, Levin I, Elad N, Tam Y, Sela N, Abu-Ras A. 2017. A new Israeli *Tobamovirus* isolate infects tomato plants harboring *Tm*-22 resistance genes. PloS One, 12: e0170429.

Ma Y, Navarro B, Zhang Z, Lu M, Zhou X, Chi S, Di Serio F, Li S. 2015. Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease. Journal of General Virology, 96: 2421-2434.

Macedo M, Albuquerque L, Maliano M, Souza J, Rojas M, Inoue-Nagata A, Gilbertson R. 2018. Characterization of tomato leaf curl purple vein virus, a new monopartite New World begomovirus infecting tomato in Northeast Brazil. Archives of Virology, 163: 737-743.

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA. 2016. Microbiologia de Brock-14ª Edição. Artmed Editora.

Male MF, Kami V, Kraberger S, Varsani A. 2015. Genome sequences of Poaceae-associated gemycircularviruses from the Pacific Ocean island of Tonga. Genome Announcements, 3: e01144-01115.

Male MF, Kraberger S, Stainton D, Kami V, Varsani A. 2016. Cycloviruses, gemycircularviruses and other novel replication-associated protein encoding circular viruses in Pacific flying fox (*Pteropus tonganus*) faeces. Infection, Genetics and Evolution, 39: 279-292.

Maliano M, Melgarejo T, Roja M, Barboza N, Gilbertson R. 2012. Tomato yellow mottle virus, a new begomovirus in Costa Rica. [Genbank]

MAPA. 2020. Instrução Normativa no 06, de 16 de maio de 2005. [Online.] Accessed on 10 April, 2020.

Márquez-Martín B, Aragón-Caballero L, Fiallo-Olivé E, Navas-Castillo J, Moriones E. 2011. Tomato leaf deformation virus, a novel begomovirus associated with a severe disease of tomato in Peru. European Journal of Plant Pathology, 129: 1-7.

Márquez-Martín B, Maeso D, Martínez-Ayala A, Bernal R, Federici MT, Vincelli P, Navas-Castillo J, Moriones E. 2012. Diverse population of a new bipartite begomovirus infecting tomato crops in Uruguay. Archives of Virology, 157: 1137-1142.

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus evolution, 1:1.

Martins TP. 2017. Identificação de vírus em tomateiro através de análise por sequenciamento de alto desempenho.Thesis.

Marzano SYL, Domier LL. 2016. Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. Virus Research, 213: 332-342.

Massart S, Chiumenti M, De Jonghe K, Glover R, Haegeman A, Koloniuk I, Komínek P, Kreuze J, Kutnjak D, Lotos L. 2019. Virus detection by high-throughput sequencing of small RNAs: Large-scale performance testing of sequence analysis strategies. Phytopathology, 109: 488-497.

Matyis J, Silva D, Oliveira A, Costa A. 1975. Purificação e morfologia do vírus do mosaico dourado do tomateiro. Summa Phytopathologica, 1: 267-275.

Mauricio-Castillo J, Argüello-Astorga G, Alpuche-Solís A, Monreal-Vargas C, Díaz-Gómez O, De La Torre-Almaraz R. 2006. First Report of Tomato severe leaf curl virus in México. Plant Disease, 90: 1116-1116.

Mauricio-Castillo J, Argüello-Astorga G, Ambriz-Granados S, Alpuche-Solís A, Monreal-Vargas C. 2007. First Report of Tomato golden mottle virus on *Lycopersicon esculentum* and *Solanum rostratum* in Mexico. Plant Disease, 91: 1513-1513.

Maurino F, Dumón AD, Llauger G, Alemandri V, de Haro LA, Mattio MF, Del Vas M, Laguna IG, Pecci MdlPG. 2018. Complete genome sequence of maize yellow striate virus, a new cytorhabdovirus infecting maize and wheat crops in Argentina. Archives of Virology, 163: 291-295.

Maxwell DM, C, Salus M, Montes L, Mejía L. 2006. Tagging begomovirus resistance gene. [Online] Accessed on 10 April, 2020.

Medina CV, Lambertini PL. 2012. Tomato dwarf leaf virus, a New World begomovirus infecting tomato in Argentina. Archives of Virology, 157: 1975-1980.

Medina CV, Martin D, Lambertini PL. 2015. Tomato mottle wrinkle virus, a recombinant begomovirus infecting tomato in Argentina. Archives of Virology, 160: 581-585.

Melgarejo TA, Kon T, Rojas MR, Paz-Carrasco L, Zerbini FM, Gilbertson RL. 2013. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. Journal of Virology: JVI. 00234-00213.

Mnari-Hattab M, Zammouri S, Pellegrin F, Gauthier N. 2014. Natural occurrence of begomovirus recombinants associated with tomato yellow leaf curl disease co-existing with parental viruses in tomato crops and weeds in Tunisia. Journal of Plant Pathology, 96: 195-200.

Monci F, Sánchez-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.

Monger W, Mumford R, Antonio García E, Boa E. 2008. Occurrence of Tomato mosaic Havana virus in Nicaragua. Plant Pathology, 57: 387-387.

Monsalve- Fonnegra Z, Arguello-Astorga GR, Rivera-Bustamante RF. 2002. Geminivirus replication and gene expression. n: Plant Viruses as Molecular Pathogens, Khan JAD, J. Plant Viruses as Molecular Pathogens.2ª ed. Dijskara. Nova York. p. 257-277.

Moreno-Félix M, Rodríguez-Negrete E, Meléndrez-Bojórquez N, Camacho-Beltrán E, Leyva-López N, Méndez-Lozano J. 2016. Presented at the V International Symposium on Tomato Diseases: Perspectives and Future Directions in Tomato Protection 1207.

Muhire BM, Varsani A, Martin DP. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PloS One, 9: e108277.

Nahid N, Amin I, Mansoor S, Rybicki E, Van Der Walt E, Briddon R. 2008. Two dicot-infecting mastreviruses (family *Geminiviridae*) occur in Pakistan. Archives of Virology, 153: 1441-1451.

Naito FY, Melo FL, Fonseca MEN, Santos CA, Chanes CR, Ribeiro BM, Gilbertson RL, Boiteux LS, de Cássia Pereira-Carvalho R. 2019. Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. Archives of Virology, 164: 1907-1910.

Nash TE, Dallas MB, Reyes MI, Buhrman GK, Ascencio-Ibañez JT, Hanley-Bowdoin L. 2011. Functional analysis of a novel motif conserved across geminivirus Rep proteins. Journal of virology, 85: 1182-1192.

Naturdata. 2020. [Online.] Accessed on 5 February, 2020.

Nava A, Patte C, Hiebert E, Polston J. 2006. Detection and variability of begomoviruses in tomato from the Andean states of Venezuela. Plant Disease, 90: 61-66.

Nawaz-ul-Rehman MS, Fauquet CM. 2009. Evolution of geminiviruses and their satellites. FEBS letters, 583: 1825-1832.

Ng TFF, Chen LF, Zhou Y, Shapiro B, Stiller M, Heintzman PD, Varsani A, Kondov NO, Wong W, Deng X. 2014. Preservation of viral genomes in 700-y-old caribou feces from a subarctic ice patch. Proceedings of the National Academy of Sciences, 111: 16842-16847.

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

Ohnishi J, Yamaguchi H, Saito A. 2016. Analysis of the mild strain of Tomato yellow leaf curl virus, which overcomes *Ty*-2 gene–mediated resistance in tomato line H24. Archives of Virology, 161: 2207-2217.

Oliveira VC, Nagata T, Guimarães FC, Ferreira FA, Kitajima EW, Nicolini C, de Oliveira Resende R, Inoue-Nagata AK. 2013. Characterization of a novel tymovirus on tomato plants in Brazil. Virus Genes, 46: 190-194.

Ong SN, Taheri S, Othman RY, Teo CH. 2020. Viral disease of tomato crops (*Solanum lycopesicum* L.): an overview. Journal of Plant Diseases and Protection, 1:1.

Orozco BM, Miller AB, Settlage SB, Hanley-Bowdoin L. 1997. Functional domains of a geminivirus replication protein. Journal of Biological Chemistry, 272: 9840-9846.

Osei M, Akromah R, Shih S, Lee L, Green S. 2008. First report and molecular characterization of DNA A of three distinct begomoviruses associated with tomato leaf curl disease in Ghana. Plant Disease, 92: 1585-1585.

Ouattara A, Tiendrébéogo F, Lefeuvre P, Claverie S, Hoareau M, Traoré EV, Barro N, Traoré O, Lett J-M. 2017. Tomato leaf curl Burkina Faso virus: a novel tomato-infecting monopartite begomovirus from Burkina Faso. Archives of Virology, 162: 1427-1429.

Padidam M, Sawyer S, Fauquet CM. 1999. Possible emergence of new geminiviruses by frequent recombination. Virology, 265: 218-225.

Pakkianathan BC, Kontsedalov S, Lebedev G, Mahadav A, Zeidan M, Czosnek H, Ghanim M. 2015. Replication of Tomato yellow leaf curl virus in its whitefly vector, *Bemisia tabaci*. Journal of Virology, 89: 9791-9803.

Panno S, Caruso AG, Davino S. 2018. The nucleotide sequence of a recombinant tomato yellow leaf curl virus strain frequently detected in Sicily isolated from tomato plants carrying the *Ty*-1 resistance gene. Archives of Virology, 163: 795-797.

Pantaleo V, Chiumenti M. 2018. iral metagenomics: Methods and Protocols. Methods in Molecular Biology, 1746:63-75.

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

Pasumarthy KK, Choudhury NR, Mukherjee SK. 2010. Tomato leaf curl Kerala virus (ToLCKeV) AC3 protein forms a higher order oligomer and enhances ATPase activity of replication initiator protein (Rep/AC1). Virology Journal, 7: 128.

Pecman A, Kutnjak D, Gutiérrez-Aguirre I, Adams I, Fox A, Boonham N, Ravnikar M. 2017. Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches. Frontiers in Microbiology, 8: 1998.

Pereira-Carvalho R. 2009. Expressão fenotípica e mecanismos de ação de genes envolvidos na resistência ampla e begomovírus monopartidos e bipartidos em tomate.Thesis.

Pereira-Carvalho R, Boiteux L, Fonseca M, Díaz-Pendón J, Moriones E, Fernández-Muñoz R, Charchar J, Resende R. 2010. Multiple resistance to Meloidogyne spp. and to bipartite and monopartite *Begomovirus* spp. in wild Solanum (Lycopersicon) accessions. Plant Disease, 94: 179-185.

Pereira-Carvalho R, Díaz-Pendón J, Fonseca M, Boiteux L, Fernández-Muñoz R, Moriones E, Resende R. 2015. Recessive resistance derived from Tomato cv. Tyking-limits drastically the spread of Tomato yellow leaf curl virus. Viruses, 7: 2518-2533.

Pereira-Carvalho R, Tobar L, Dianese É, Fonseca M, Boiteux L. 2014. Melhoramento genético do tomateiro para resistência a doenças de etiologia viral: avanços e perspectivas.

Pereira IS, Rodrigues VF, Vega MRG. 2016. Flavonoides do Gênero Solanum. Revista Virtual de Química, 8: 4-26.

Perry KL, McLane H, Thompson JR, Fuchs M. 2018. A novel grablovirus from non-cultivated grapevine (Vitis sp.) in North America. Archives of Virology, 163: 259-262.

Phan TG, Mori D, Deng X, Rajindrajith S, Ranawaka U, Ng TFF, Bucardo-Rivera F, Orlandi P, Ahmed K, Delwart E. 2015. Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. Virology, 482: 98-104.

Pietersen G, Idris A, Krüger K, Brown J. 2000. Tomato curly stunt virus, a new begomovirus of tomato within the Tomato yellow leaf curl virus-IS cluster in South Africa. Plant Disease, 84: 810-810.

Pinheiro JB, Boiteux LS, Pereira RB, Almeida MRA, Carneiro RMG. 2014. Identificação de espécies de meloidogyne em tomateiro no Brasil. Embrapa Hortaliças-Boletim de Pesquisa e Desenvolvimento (INFOTECA-E).

Pooma W, Petty IT. 1996. Tomato golden mosaic virus open reading frame AL4 is genetically distinct from its C4 analogue in monopartite geminiviruses. Journal of General Virology, 77: 1947-1951.

Posada D. 2008. jModelTest: phylogenetic model averaging. Molecular biology and evolution, 25: 1253-1256.

Pradhan B, Van Tien V, Dey N, Mukherjee SK. 2017. Molecular Biology of Geminivirus DNA Replication. Avid Science, 1: 1-33.

Qadir R, Khan ZA, Monga D, Khan JA. 2019. Diversity and recombination analysis of Cotton leaf curl Multan virus: a highly emerging begomovirus in northern India. BMC Genomics, 20: 274.

Quadros AF, Silva JP, Xavier CA, Zerbini FM, Boari AJ. 2019. Two new begomoviruses infecting tomato and Hibiscus sp. in the Amazon region of Brazil. Archives of Virology, 164: 1897-1901.

Rambaut A. 2012. FigTree v1. 4. Molecular evolution, phylogenetics and epidemiology. Edinburgh, UK: Retrieved from http://tree.bio.ed.ac.uk/software/figtree [Google Scholar].

Ramos P, Guerra O, Peral R, Oramas P, Guevara R, Rivera-Bustamante R. 1997. Taino tomato mottle virus, a new bipartite geminivirus from Cuba. Plant Disease, 81: 1095-1095.

Rêgo-Machado CM, Nakasu EY, Blawid R, Nagata T, Inoue-Nagata AK. 2019. Complete genome sequence of a new bipartite begomovirus infecting tomato in Brazil. Archives of Virology, 164: 2873-2875.

Rêgo CM. 2016. Diversidade genômica de begomovírus em tomateiros resistente (BRS SENA) e susceptível (H-9553).Doctor Degree Thesis.

Rezende W, Militão Neto V, Goulart L, GiovaniniI M, Juliatti F, Fernandes J. 1997. Mixed infection by geminiviruses in tomato plants detected by LIS-SSCP-PCR. Fitopatologia Brasileira 22: 338-339.

Ribeiro S, Ambrozevicius L, Avila A, Bezerra I, Calegario R, Fernandes J, Lima M, De Mello R, Rocha H, Zerbini F. 2003. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. Archives of Virology, 148: 281-295.

Ribeiro S, De Ávila A, Bezerra I, Fernandes J, Faria J, Lima M, Gilbertson R, Maciel-Zambolim E, Zerbini F. 1998. Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. Plant Disease, 82: 830-830.

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

Ribeiro S, Mello L, Boiteux L, Kitajima E, Faria J. 1994. Tomato infection by a geminivirus in the Federal District, Brazil. Fitopatologia Brasileira, 19: 330.

Rocha CS, Castillo-Urquiza GP, Lima AT, Silva FN, Xavier CA, Hora-Júnior BT, Beserra-Júnior JE, Malta AW, Martin DP, Varsani A, Alfenas-Zerbini P, Mizubuti ESG, Zerbini FM. 2013. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location. Journal of Virology: JVI. 00155-00113.

Rodríguez-Negrete EA, Morales-Aguilar JJ, Domínguez-Duran G, Torres-Devora G, Camacho-Beltrán E, Leyva-López NE, Voloudakis AE, Bejarano ER, Méndez-Lozano J. 2019. High-throughput sequencing reveals differential *Begomovirus* species diversity in non-cultivated plants in Northern-Pacific Mexico. Viruses, 11: 594.

Rojas A, Kvarnheden A, Marcenaro D, Valkonen J. 2005. Sequence characterization of tomato leaf curl Sinaloa virus and tomato severe leaf curl virus: phylogeny of New World begomoviruses and detection of recombination. Archives of Virology, 150: 1281-1299.

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

Rojas MR, Hagen C, Lucas W, Gilbertson R. 2005a. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annual Review of Phytopathology, 43: 361-394.

Rojas MR, Macedo M, Maliano M, Soto-Aguilar M, Souza J, Briddon R, Kenyon L, Rivera Bustamante R, Zerbini F, Adkins S. 2018. World management of geminiviruses. Annual Review of Phytopathology, 56: 637-677.

Rojas RM, Jiang H, Salati R, Xoconostle-Cázares B, Sudarshana M, Lucas WJ, Gilbertson RL. 2001. Functional analysis of proteins involved in movement of the monopartite begomovirus, Tomato yellow leaf curl virus. Virology, 291: 110-125.

Romay G, Chirinos D, Geraud-Pouey F, Desbiez C. 2010. Association of an atypical alphasatellite with a bipartite New World begomovirus. Archives of Virology, 155: 1843-1847.

Romay G, Chirinos DT, Geraud-Pouey F, Gillis A, Mahillon J, Desbiez C, Bragard C. 2017. Molecular and biological characterization of a new Tomato mild yellow leaf curl Aragua virus strain producing severe symptoms in tomato. Virus Genes, 53: 939-942.

Romay G, Geraud-Pouey F, Chirinos DT, Mahillon M, Gillis A, Mahillon J, Bragard C. 2019. Tomato twisted leaf virus: A novel indigenous new world monopartite begomovirus infecting tomato in Venezuela. Viruses, 11: 327.

Roossinck MJ. 1997. Mechanisms of plant virus evolution. Annual Review of Phytopathology, 35: 191-209.

Roossinck MJ, Martin D, Roumagnac P. 2015. Plant virus metagenomics: Advances in virus discovery. Phytopathology, 105: 716-727.

Rosario K, Dayaram A, Marinov M, Ware J, Kraberger S, Stainton D, Breitbart M, Varsani A. 2012a. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). Journal of General Virology, 93: 2668-2681.

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

Rosario K, Marr C, Varsani A, Kraberger S, Stainton D, Moriones E, Polston JE, Breitbart M. 2016. Begomovirus-associated satellite DNA diversity captured through vector-enabled metagenomic (VEM) surveys using whiteflies (Aleyrodidae). Viruses, 8: 36.

Rosario K, Padilla-Rodriguez M, Kraberger S, Stainton D, Martin DP, Breitbart M, Varsani A. 2013. Discovery of a novel mastrevirus and alphasatellite-like circular DNA in dragonflies (Epiprocta) from Puerto Rico. Virus Research, 171: 231-237.

Rosen R, Kanakala S, Kliot A, Pakkianathan BC, Farich BA, Santana-Magal N, Elimelech M, Kontsedalov S, Lebedev G, Cilia M. 2015. Persistent, circulative transmission of begomoviruses by whitefly vectors. Current Opinion in Virology, 15: 1-8.

Roy A, Zhai Y, Ortiz J, Neff M, Mandal B, Mukherjee SK, Pappu HR. 2019. Multiplexed editing of a begomovirus genome restricts escape mutant formation and disease development. PloS one, 14.

Ruhel R, Chakraborty S. 2019. Multifunctional roles of geminivirus encoded replication initiator protein. Virus Disease, 30: 66-73.

Rumbou A, Candresse T, Marais A, Theil S, Langer J, Jalkanen R, Büttner C. 2018. A novel badnavirus discovered from Betula sp. affected by birch leaf-roll disease. PloS One, 13: e0193888.

Sahu AK, Verma RK, Gaur R, Sanan-Mishra N. 2018. Complexity and recombination analysis of novel begomovirus associated with spinach yellow vein disease in India. Plant Gene, 13: 42-49.

Samarakoon S, Balasuriya A, Rajapaksha R, Wickramarachchi W. 2012. Molecular detection and partial characterization of tomato yellow leaf curl virus in Sri Lanka. Pakistan Journal of Biological Sciences, 15: 863-870.

Sánchez PAG, Mesa HJ, Montoya MM. 2016. Next generation sequence analysis of the forage peanut (*Arachis pintoi*) virome. Revista Facultad Nacional de Agronomía, Medellín, 69: 7881-7891.

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 CD, Ávila AC, Resende RO. 2003. Estudo da interação de um begomovírus isolado de tomateiro com a mosca branca. Fitopatologia Brasileira, 28: 664-673.

Sattar MN, Kvarnheden A, Saeed M, Briddon RW. 2013. Cotton leaf curl disease–an emerging threat to cotton production worldwide. Journal of General Virology, 94: 695-710.

Scussel S, Claverie S, Hoareau M, Moustache R, Delatte H, Lefeuvre P, Lett J-M. 2018. Tomato leaf curl Mahé virus: a novel tomato-infecting monopartite begomovirus from the Seychelles. Archives of virology, 163: 3451-3453.

Seal S, VandenBosch F, Jeger M. 2006. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. Critical Reviews in Plant Sciences, 25: 23-46.

Sharma P, Ikegami M. 2009. Characterization of signals that dictate nuclear/nucleolar and cytoplasmic shuttling of the capsid protein of Tomato leaf curl Java virus associated with DNAβ satellite. Virus Research, 144: 145-153.

Shih S, Green S, Tsai W, Ssekyewa C. 2006. Molecular characterization of a begomovirus associated with tomato leaf curl disease in Uganda. Plant Disease, 90: 246-246.

Shih S, Roff M, Nakhla M, Maxwell D, Green S. 1998. A new geminivirus associated with a leaf curl disease of tomato in Malaysia. J Zhiwu Baohuxue Hui Huikan, 40: 435-435.

Shih S, Tsai W, Green S, Lee L. 2006a. Molecular characterization of a distinct begomovirus associated with Tomato Leaf Curl Disease in Arusha of Tanzania. Plant Disease, 90: 1550-1550.

Sikorski A, Massaro M, Kraberger S, Young LM, Smalley D, Martin DP, Varsani A. 2013. Novel myco-like DNA viruses discovered in the faecal matter of various animals. Virus Research, 177: 209-216.

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

Silva JBC, Giordano L, Furumoto O, Boiteux L, França F, Villas-Bôas G, Castelo-Branco M, Medeiros M, Marquelli W, Silva W, Lopes C, Ávila A, Nascimento W, Pereira W. 2006. Cultivo de tomate para industrialização. [Genbank]

Silva LD, Omoto C, Bleicher E, Dourado P. 2009. Monitoring the susceptibility to insecticides in *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) populations from Brazil. Neotropical Entomology, 38: 116-125.

Silva SJC, Castillo-Urquiza G, Hora-Júnior B, Assunção I, Lima G, Pio-Ribeiro G, Mizubuti E, Zerbini F. 2012. Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. Plant Pathology, 61: 457-467.

Simon-Loriere E, Holmes EC. 2011. Why do RNA viruses recombine? Nature Reviews Microbiology, 9: 617-626.

Sivalingam P, Malathi V, Varma A. 2005. A new begomovirus infecting tomato in Rajasthan, India. [Genbank]

Skaljac M, Kanakala S, Zanic K, Puizina J, Pleic IL, Ghanim M. 2017. Diversity and phylogenetic analyses of bacterial symbionts in three whitefly species from Southeast Europe. Insects, 8: 113.

Sobrinho RR, Xavier CAD, de Barros Pereira HM, de Andrade Lima GS, Assunção IP, Mizubuti ESG, Duffy S, Zerbini FM. 2014. Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. Journal of General Virology, 95: 2540-2552.

Socoloski A, Grzebieluckas C, dos Santos JSC, Stieler MC, de Lima AdFA. 2017. Economic analysis of vegetable crop production: a study with family farmers. Custos e Agronegócio Online, 13: 389-407.

Souza C, Rossato M, Melo F, Boiteux L, Pereira-Carvalho R. 2018. First Report of Sweet Potato Symptomless Virus 1 Infecting Ipomoea batatas in Brazil. Plant Disease, 102: 2052-2052.

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

Steel O, Kraberger S, Sikorski A, Young LM, Catchpole RJ, Stevens AJ, Ladley JJ, Coray DS, Stainton D, Dayaram A. 2016. Circular replication-associated protein encoding DNA viruses identified in the faecal matter of various animals in New Zealand. Infection, Genetics and Evolution, 43: 151-164.

Strausbaugh CA, Eujayl I, Wintermantel W. 2017. Beet curly top virus strains associated with sugar beet in Idaho, Oregon, and a Western US Collection. Plant Disease, 101: 1373-1382.

Strausbaugh CA, Wintermantel W, Gillen A, Eujayl IA. 2008. Curly top survey in the western United States. Phytopathology, 98: 1212-1217.

Sung Y, Coutts RH. 1995. Mutational analysis of potato yellow mosaic geminivirus. Journal of General Virology, 76: 1773-1780.

Swarnalatha P, Jalali S, Reddy K. 2014. Molecular characterization of Begomovirus associated with Tomato leaf curl disease in India. [Genbank]

Swarnalatha S; Venkataravanappa V, Jalali SK, M. 2019. Tomato leaf curl Karnataka virus 2, [Genbank]

Swarnalatha P, Manasa M, Venkataravanappa V, Jalali S, Reddy K. 2013. Tomato leaf curl new Delhi Virus 4. [Genbank]

Tahzima R, Foucart Y, Massart S, De Jonghe K. 2019. First report of Prunus virus F infecting sweet cherry cultivars using high-throughput sequencing in Belgium. New Disease Reports, 40: 7-7.

Tiwari N, Padmalatha K, Singh V, Haq Q, Malathi V. 2010. Tomato leaf curl Bangalore virus (ToLCBV): infectivity and enhanced pathogenicity with diverse betasatellites. Archives of Virology, 155: 1343-1347.

Tiwari N, Singh VB, Sharma P, Malathi V. 2013. Tomato leaf curl Joydebpur virus: a monopartite begomovirus causing severe leaf curl in tomato in West Bengal. Archives of Virology, 158: 1-10.

Torre C, Agüero J, Aranda M. 2019. First evidence of Tomato yellow leaf curl virus-Israel IS76 recombinant isolates associated with severe yellow leaf curl epidemics in resistant tomatoes in Spain. Plant Disease, 103: 780.

Trębicki P, Harding R, Rodoni B, Baxter G, Powell K. 2010. Vectors and alternative hosts of Tobacco yellow dwarf virus in southeastern Australia. Annals of Applied Biology, 157: 13-24.

Tsai W, Green S, Shi S 2006a, posting date. Molecular identification of Taiwan Ageratum- Infecting begomoviruses. [Genbank]

Tsai W, Nakhla M, Maxwell D, Green S, Black L. 1999. Tomato leaf curl geminivirus from Laos. [Genbank]

Tsai W, Shih S, Green S, Akkermans D, Jan F-J. 2006b. Molecular characterization of a distinct tomato-infecting begomovirus associated with yellow leaf curl diseased tomato in Lembang, Java Island of Indonesia. Plant Disease, 90: 831-831.

Tsai W, Shih S, Green S, Lee L, Luther G, Ratulangi M, Sembel D, Jan F-J. 2009. Identification of a new begomovirus associated with yellow leaf curl diseases of tomato and pepper in Sulawesi, Indonesia. Plant Disease, 93: 321-321.

Tsai W, Shih S, Kenyon L, Green S, Jan FJ. 2011. Temporal distribution and pathogenicity of the predominant tomato-infecting begomoviruses in Taiwan. Plant Pathology, 60: 787-799.

Ueda S, Takeuchi S, Okabayashi M, Hanada K, Tomimura K, Iwanami T. 2005. Evidence of a new Tomato yellow leaf curl virus in Japan and its detection using PCR. Journal of General Plant Pathology, 71: 319-325.

Vaghi Medina CG, Teppa E, Bornancini VA, Flores CR, Marino-Buslje C, López Lambertini PM. 2018. Tomato apical leaf curl virus: A novel, monopartite geminivirus detected in tomatoes in Argentina. Frontiers in Microbiology, 8: 2665.

Van den Brand JM, van Leeuwen M, Schapendonk CM, Simon JH, Haagmans BL, Osterhaus AD, Smits SL. 2012. Metagenomic analysis of the viral flora of pine marten and European badger feces. Journal of Virology, 86: 2360-2365.

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.

Varela G, Avalos V, Reyna P, Laguna IG, Pardina PR. 2018. Identification, molecular characterization and relative incidence of begomoviruses infecting bean crops in northwestern Argentina: an update. Australasian Plant Pathology, 47: 343-350.

Varma A, Malathi V. 2003. Emerging geminivirus problems: a serious threat to crop production. Annals of Applied Biology, 142: 145-164.

Varsani A, Krupovic M. 2017. Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. Virus Evolution, 3: 1-14.

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

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

Venkataravanappa V, Swarnalatha P, Reddy CL, Chauhan N, Reddy MK. 2016. Association of recombinant Chilli leaf curl virus with enation leaf curl disease of tomato: a new host for chilli begomovirus in India. Phytoparasitica, 44: 213-223.

Verlaan MG, Hutton SF, Ibrahem RM, Kormelink R, Visser RG, Scott JW, Edwards JD, Bai Y. 2013. The tomato yellow leaf curl virus resistance genes Ty-1 and Ty-3 are allelic and code for DFDGD-class RNA–dependent RNA polymerases. PLoS Genetics, 9: e1003399.

Verlaan MG, Szinay D, Hutton SF, de Jong H, Kormelink R, Visser RG, Scott JW, Bai Y. 2011. Chromosomal rearrangements between tomato and *Solanum chilense* hamper mapping and breeding of the TYLCV resistance gene *Ty*-1. The Plant Journal, 68: 1093-1103.

Vilela NJ, Melo PCT, Boiteux LS. 2012. Perfil socioeconômico da cadeia agroindustrial no Brasil. In: Produção de tomate para processamento industrial, Clemente FMVTB, L. S. Produção de tomate para processamento industrial. Brasília: Embrapa, 2012. p.17-27.

Villamor D, Ho T, Al Rwahnih M, Martin R, Tzanetakis I. 2019. High throughput sequencing for plant virus detection and discovery. Phytopathology, 109: 716-725.

Virus-HostDB. 2020. Host Index. [https://www.genome.jp/virushostdb/]. Accessed on 15 March, 2020.

Voorburg CM, Yan Z, Bergua-Vidal M, Wolters AMA, Bai Y, Kormelink R. 2020. *Ty*-1, a universal resistance gene against geminiviruses that is compromised by co-replication of a betasatellite. Molecular Plant Pathology, 21: 160-172.

Walker JE, Saraste M, Runswick MJ, Gay NJ. 1982. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO Journal, 1: 945-951.

Wang Y, Jiang J, Zhao L, Zhou R, Yu W, Zhao T. 2018. Application of Whole Genome Resequencing in Mapping of a Tomato Yellow Leaf Curl Virus Resistance Gene. Scientific Reports, 8: 1-11.

Wu Q, Ding SW, Zhang Y, Zhu S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. Annual Review of Phytopathology, 53: 425-444.

Wyant PS, Strohmeier S, Schäfer B, Krenz B, Assunção IP, de Andrade Lima GS, Jeske H. 2012. Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. Virology, 427: 151-157.

Xu C, Sun X, Taylor A, Jiao C, Xu Y, Cai X, Wang X, Ge C, Pan G, Wang Q. 2017a. Diversity, distribution, and evolution of tomato viruses in China uncovered by small RNA sequencing. Journal of virology: JVI. 00173-00117.

Xu, Y, Li S, Na C, Yang L, Lu M. 2019. Analyses of virus/viroid communities in nectarine trees by next-generation sequencing and insight into viral synergisms implication in host disease symptoms. Scientific Reports, 9: 1-12.

Xu C, Sun X, Taylor A, Jiao C, Xu Y, Cai X, Wang X, Ge C, Pan G, Wang Q. 2017b. Diversity, distribution, and evolution of tomato viruses in China uncovered by small RNA sequencing. Journal of Virology: JVI. 00173-00117.

Xu Y, Cai X, Zhou X. 2007. Tomato leaf curl Guangxi virus is a distinct monopartite begomovirus species. European Journal of Plant Pathology, 118: 287-294.

Yadava P, Suyal G, Mukherjee SK. 2010. *Begomovirus* DNA replication and pathogenicity. Current Science (00113891), 98.

Yamaguchi H, Ohnishi J, Saito A, Ohyama A, Nunome T, Miyatake K, Fukuoka H. 2018. An NB-LRR gene, TYNBS1, is responsible for resistance mediated by the *Ty*-2 *Begomovirus* resistance locus of tomato. Theoretical and Applied Genetics, 131: 1345-1362.

Yang X, Guo W, Ma X, An Q, Zhou X. 2011. Molecular characterization of Tomato leaf curl China virus infecting tomato in China and functional analyses of its associated betasatellite. Applied and Environmental Microbiology, 7: 3092-3101.

.

Yao FL, Zheng Y, Huang XY, Ding XL, Zhao JW, Desneux N, He YX, Weng QY. 2017. Dynamics of *Bemisia tabaci* biotypes and insecticide resistance in Fujian province in China during 2005–2014. Scientific Reports, 7: 1-12.

Yin Q, Yang H, Gong Q, Wang H, Liu Y, Hong Y, Tien P. 2001. Tomato yellow leaf curl China virus: monopartite genome organization and agroinfection of plants. Virus Research, 81: 69-76.

Yu X, Li B, Fu Y, Jiang D, Ghabrial SA, Li G, Peng Y, Xie J, Cheng J, Huang J. 2010. A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. Proceedings of the National Academy of Sciences, 107: 8387-8392.

Zambrano K, Geraud-Pouey F, Chirinos D, Romay G, Marys E. 2011. Tomato chlorotic leaf distortion virus, a new bipartite begomovirus infecting *Solanum lycopersicum* and Capsicum chinense in Venezuela. Archives of Virology, 156: 2263-2266.

Zamir D, Ekstein-Michelson I, Zakay Y, Navot N, Zeidan M, Sarfatti M, Eshed Y, Harel E, Pleban T, Van-Oss H. 1994. Mapping and introgression of a Tomato yellow

leaf curl virus tolerance gene, *Ty*-1. Theoretical and Applied Genetics, 88: 141-146.

Zerbini FM, Andrade E, Barros D, Ferreira S, Lima A, Alfenas P, Mello R. 2005. Traditional and novel strategies for geminivirus management in Brazil. Australasian Plant Pathology, 34: 475-480.

Zerbini FM, Briddon R, Idris A, Martin D, Moriones E, Navas-Castillo J, Bustamante R, Roumagnac P, Varsani A, Consortium IR. 2017. ICTV virus taxonomy profile: Geminiviridae. The Journal of General Virology, 98: 131.

Zhang R, Wu X, Jiang X, Wu X, Luan X, Cheng X. 2020. Molecular characterization of common bean curly stunt virus: a novel recombinant geminivirus in China. Archives of Virology, 165: 257-260.

Zhang Z, Qi S, Tang N, Zhang X, Chen S, Zhu P, Ma L, Cheng J, Xu Y, Lu M. 2014. Discovery of replicating circular RNAs by RNA-seq and computational algorithms. PLoS Pathogens, 10: 1-12.

Zhang H, Hu G, Zhou X. 2010. Molecular characterization of Tomato leaf curl Hainan virus, a new begomovirus, and evidence for recombination. Journal of Phytopathology, 158: 829-832.

Zhao L, Ding M, Yon Y, Zhang X, Zhang Z .2015. Identification of a novel begomovirus associated with betasatellites infecting tomato in china. [Genbank]

Zhou YC, Noussourou M, Kon T, Rojas M, Jiang H, Chen L-F, Gamby K, Foster R, Gilbertson R. 2008. Evidence of local evolution of tomato-infecting begomovirus species in West Africa: characterization of tomato leaf curl Mali virus and tomato yellow leaf crumple virus from Mali. Archives of Virology, 153: 693-706.

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

Zhou X, Liu Y, Calvert L, Munoz C, Otim-Nape GW, Robinson DJ, Harrison BD. 1997. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. Journal of General Virology, 78: 2101-2111.

Zi-Fu H, Hao Y, Ming-Jie M, Fang-Fang L, Yi-Han L, Sui-Tao W. 2007. Tomato yellow leaf curl disease in Guangdong is caused by Tomato leaf curl Taiwan virus. Chinese Journal of Agricultural Biotechnology, 4: 127-131.

Zubiaur YM, De Blas C, Quiñones M, Castellanos C, Peralta EL, Romero J. 1998. Havana tomato virus, a new bipartite geminivirus infecting tomatoes in Cuba. Archives of Virology, 143: 1757-1772.