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

Tese de Doutorado

Viroma de plantas em áreas nativas e cultivadas do Distrito Federal

JOSIANE GOULART BATISTA

Brasília - DF
2020

JOSIANE GOULART BATISTA

Viroma de plantas em áreas nativas e cultivadas do Distrito Federal

Tese apresentada à Universidade de Brasília como requisito parcial para a obtenção do título de Doutor em Fitopatologia pelo Programa de Pós-Graduação em Fitopatologia.

Orientadora

Dra. Rita de Cássia Pereira Carvalho

BRASÍLIA
DISTRITO FEDERAL - BRASIL 2020

FICHA CATALOGRÁFICA

Batista, G.J.

Viroma de plantas em áreas nativas e cultivadas do Distrito Federal.

Josiane Goulart Batista.

Brasília, 2020.

Número de páginas p.: 154

Tese de Doutorado - Programa de Pós-Graduação em Fitopatologia, Universidade de Brasília, Brasília, DF.

Espécies florestais, cultivadas, *Geminiviridae*, *Genomoviridae.*

I.Universidade de Brasília. PPG/FIT.

II.Título. **Viroma de plantas em áreas nativas e cultivadas do Distrito Federal.**

A minha mãe Nelma e ao meu pai Josvaldo, as minhas filhas Giovanna e Raissa, a minha irmã Josefa e minha sobrinha Júlia.

Dedico

Agradecimentos

A Deus.

Aos meus pais Josvaldo e Nelma.

As minhas filhas Giovanna e Raissa, minha irmã Josefa e minha sobrinha Júlia.

A minha orientadora Professora Rita de Cássia Pereira Carvalho.

Aos colegas do mestrado e doutorado Caroline, Flávia, Luciane, Macária e Ikaro.

Aos colegas de laboratório Felipe Fochat, Mateus Malheiros e Vinícius.

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 Miller, Marisa Álvares da Silva Velloso Ferreira, Helson Mario Martins do Vale, José Carmine Dianese e Luís Eduardo Bassay Blum.

A Embrapa Recursos Genéticos e Biotecnologia e a pesquisadora Simone Graça Ribeiro.

A Embrapa Hortaliças e aos seus pesquisadores Mirtes Freitas Lima, Leonardo Silva Boiteux e Alexandre Furtado Silveira Mello.

A Companhia Urbanizadora da Nova Capital (NOVACAP) e seu funcionário Silomar.

A Estação Ecológica de Águas Emendadas (ESECAE).

A Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR).

A Fazenda Água Limpa (FAL-Universidade de Brasília-UnB).

A Estação Experimental de Biologia (EEB-UnB).

A Empresa de Assistência Técnica e Extensão Rural do Distrito Federal (EMATER-DF) e o Núcleo Rural de Taquara.

A Fundação de Apoio a Pesquisa do Distrito Federal (FAP-DF).

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.

Trabalho realizado junto ao Departamento de Fitopatologia do Instituto de Ciências Biológicas da Universidade de Brasília, sob orientação da **Dra. Profa. Rita de Cássia Pereira Carvalho**. Apoio Conselho Nacional de Pesquisa e Desenvolvimento Científico e Tecnológico (CNPq) e da Embrapa Recursos Genéticos e Biotecnologia.

Viroma de espécies arbóreas nativas, exóticas e cultivadas do Distrito Federal e entorno

JOSIANE GOULART BATISTA

TESE APROVADA em __/__/____por:

Dra. Mirtes Freitas Lima
Embrapa Hortaliças
(Examinador Externo)

Dr Alexandre Furtado Silveira Mello
Embrapa Hortaliças
(Examinador Externo)

Dra. Renato de Oliveira Resende
Departamento de Biologia Celular - UnB
(Examinador Interno)

Dr.Tatsuya Nagata
Departamento de Fitopatologia - UnB
(Examinador Interno-Suplente)

Dra. Rita de Cássia Pereira Carvalho
Departamento de Fitopatologia - UnB
(Orientadora Presidente)

BRASÍLIA - DISTRITO FEDERAL
BRASIL
2020

Sumário

Resumo Geral 1

General Abstract 4

Introdução Geral 7

Hipóteses do Trabalho 10

Objetivos Gerais 10

Objetivos Específicos 10

Capítulo 1 12

Revisão Bibliográfica 12

1. Importância florestal no contexto mundial 12

1.1. Importância das florestas para o Brasil 13

2. O Bioma Cerrado 14

3. Unidades de Conservação no Brasil 16

4. Patologia Florestal 19

4.1. Vírus em espécies florestais e arbóreas: estudos clássicos 21

4.2 Vírus em espécies arbóreas no Brasil 23

4.3 Vírus em espécies florestais e arbóreas: metagenômica 24

4.4. Descrição de novos vírus de DNA de fita simples 27

4.4.1 Família *Geminiviridae* 29

4.4.2 Família *Genomoviridae* 39

4.4.3. Família *Circoviridae* 43

4.5. Família *Caulimoviridae* 44

Referências Bibliográficas 47

Chapter 2 67

Single-stranded DNA viroma of vegetable and field crops and from native and exotic tree species growing in the Brazilian *Cerrado* biome 67

1. Introduction 70

2. Materials and Methods 72

2.1. Sampling sites 72

2.2. Details of sample collections 74

2.3. Plant material processing and viral circular DNA enrichment 74

2.4. Metagenomic analyses and bioinformatics (sequence trimming, de novo assembly and annotation) 75

2.5. Primers design and molecular validation of viruses identified by High-throughput sequencing - HTS 76
3. Results and Discussion 77
3.1 High performance sequencing: screening, assembly and analysis of contigs, and virus detection 77
4. Conclusion 99
References 114
Chapter 3 119
New genomovirus associated with trees and cultivated species in Central Brazil 119
1. Introduction 120
2. Materials and Methods 122
2.1. Sample collection locations - areas and plant species 122
2.2. List of plant species collected 122
2.3. Plant material processing and High-throughput sequencing - HTS 123
2.4. Bioinformatic and metagenomic analyzes 123
2.5. Specific primers design and virus detection 124
2.6. Sanger validation for High-throughput sequencing - HTS 126
2.7. Phylogenetic analysis, pairwise identity and recombination analysis 126
3. Results and Discussion 126
3.1. Molecular characterization and PCR detection 126
3.2. Phylogenetic analysis, pairwise identity and recombination analysis 134
4. Conclusion 140
References 140
Chapter 4 144
Broad spectrum of host range for *Geminiviridae* members 144
1. Introduction 145
2. Materials and Methods 147
2.1 Plant material collection and High-throughput sequencing - HTS 147
3. Results 149
4. APPENDIX – Attached and publications (accepted, submitted and to be submitted) from these Thesis. 151
Conclusões Gerais 152
Referências 152

LISTA DE FIGURAS

Chapter 2

Figure 1. Plant collection sites (tree and herbaceous species) growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. All samples displayed virus-like symptoms. Locations are represented by colors: **a. Anthropic Areas:** Arborização Urbana do Distrito Federal - AUDF (purple), Núcleo Rural Taquara - NRT (teal blue), Viveiro II da Companhia Urbanizadora da Nova Capital - VII-NOVACAP (red), Fazenda Água Limpa da Universidade de Brasília - AE-FAL (blue), **b. Conservation Areas:** Estação Ecológica de Águas Emendadas - ESECAE (yellow) and Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística - RECOR (orange) and **c. Transition Areas**: 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) and 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE - EEB) (green).....73

Figure 2. Distribution of viral species by collection sites. Microenvironments: Arborização Urbana do Distrito Federal (AUDF), Núcleo Rural Taquara (NRT), Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), Fazenda Água Limpa da Universidade de Brasília (AE-FAL), Estação Ecológica de Águas Emendadas (ESECAE), Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) and 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). Acronyms: *Cauliflower mosaic virus* - CaMV (*Caulimovirus/Caulimoviridae*); *Bean golden mosaic virus* - BGMV (DNA-A), Macroptilium bright yellow interveinal virus - MaBYIV, *Sida micrantha mosaic virus* - SiMMV (DNA-A), *Sweet potato leaf curl virus* - SPLCV, *Tomato mottle leaf curl virus* - ToMoLCV, *Tomato severe rugose virus* - ToSRV (DNA-A), (*Begomovirus/Genomoviridae*); Maize striate mosaic virus - MSMV (*Mastrevirus/Genomoviridae*); Tomato associated geminivirus 1 - TaGV1, Tomato apical leaf curl virus - ToALCV (unclassified in genus/*Geminiviridae*,); Vochysia rufa associated genomovirus - VoaGmV and Byrsonima crassiflora associated genomovirus - ByaGmV (unclassified in genus/*Genomoviridae*); Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus/Genomoviridae*); Ouratea duparquetiana associated gemykibivirus - OuaGmV (*Gemykibivirus/Genomoviridae*); Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus/Genomoviridae*);, Gila monster associated gemykrogvirus - GmaV1 (*Gemykrogvirus/Genomoviridae*). One unclassified ssDNA species, Caesalpinia pluviosa associated circular virus - CpaCV.....87

Figure 3. Single stranded DNA and double strand DNA viruses distribution detected in tree and herbaceous species from Single stranded DNA viruses distribution in tree and herbaceous species growing across distinct sites in three microenvironments of the Cerrado area of Brasília-DF (Central Brazil) (Federal District), Brazil. Acronyms: *Cauliflower mosaic virus* - CaMV (*Caulimovirus/Caulimoviridae*); *Bean golden mosaic virus* - BGMV (DNA-A), Macroptilium bright yellow interveinal virus - MaBYIV, *Sida micrantha mosaic virus* - SiMMV (DNA-A), *Sweet potato leaf curl virus* - SPLCV, *Tomato mottle leaf curl virus* - ToMoLCV, *Tomato severe rugose virus* - ToSRV (DNA-A), (*Begomovirus/Geminiviridae*); Maize striate mosaic virus - MSMV (*Begomovirus/Geminiviridae*), Tomato associated geminivirus 1 - TaGV1, Tomato apical leaf curl virus - ToALCV (unclassified in genus/*Geminiviridae*); Vochysia rufa associated genomovirus - VoaGmV and Byrsonima crassiflora associated genomovirus - ByaGmV (unclassified in genus/ *Genomoviridae*); Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus/Genomoviridae*); Ouratea duparquetiana associated

gemykibivirus - OuaGmV (*Gemykibivirus/Genomoviridae*); Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus/Genomoviridae*); Gila monster associated gemykrogvirus - GmaV1 (*Gemykrogvirus/Genomoviridae*). One unclassified ssDNA species, the Caesalpinia pluviosa associated circular virus - CpaCV. 95

Supplementary Figure 1 SF1. Symptomatic plants collected in Anthropic Areas - Arborização Urbana do Distrito Federal (AUDF): A. *Inga cylindrica*, B. *Caryocar brasiliense* and C. *Inga laurina*; Núcleo Rural Taquara (NRT): D. and E. *Solanum melongena* and F. *Sida* sp.; Viveiro II of Companhia Urbanizadora da Nova Capital (VII-NOVACAP): G. *Aspidosperma polyberon*, H. *Tabebuia chrysotricha* and I. *Cybistax antisyphilitica*; Fazenda Água Limpa da Universidade of Brasília (AE-FAL): J. *Gossypium hirsutum*, K. *Solanum lycopersicum* and L. *Beta vulgaris*. 101

Supplementary Figure 2 SF2. Symptomatic plants collected in Conservation Areas - Estação Ecológica de Águas Emendadas (ESECAE): A. *Bauhinia rufa*, B. *Miconia* sp. and C. *Anacardium humile*; Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR): D. *Byrsonima crassiflora*, E. *Baccharis tridentata* and F. *Calliandra dysantha*; Estação Experimental Experimental de Biologia da Universidade de Brasília (ARIE-AHE-EEB): G. *Manilkara zapota*, H. *Piper nigrum* and I. *Phaseolus5 vulgaris*. 102

Chapter 3

Figura 1. Genomic organization scheme and symptomatic plants which viruses were associated. In: A. *Solanum melongena* and B. Momordica charantia associated gemycircularvirus - MoaGmV; C. *Ouratea duparquetiana* and D. Ouratea duparquetiana associated gemykibivirus - OuaGmV; E. *Trembleya parviflora*; F. Gila monster-associated gemykrogvirus - GmaV1; G. Tecoma stans associated gemykolovirus; H. *Vochysia rufa* and I. Vochysia rufa associated genomovirus - VoaGmV; J. *Byrsonima crassiflora* and K. Byrsonima crassiflora associated genomovirus - ByaGmV and L. Eucalyptus urophylla associated genomovirus - EcaGmV. Scheme of genomic organization (replication associated protein gene (*rep*) in blue and coat protein gene (*cp*) in red) and plant photos: **A.** *Solanum melongena* and **B.** Momordica charantia associated with the gemicircular virus - MoaGmV; **C.** *Ouratea duparquetiana* and **D.** Ouratea duparquetiana associated gemykibivirus - OuaGmV; **E.** *Trembleya parviflora* and **F.** Gila monster associated gemykrogvirus - GmaV1; **G.** Tecoma stans gemykolovirus associated - TeaGmV; **H.** *Vochysia rufa* and **I.** Vochysia rufa-associated genomovirus - VoaGmV; **J.** *Byrsonima crassiflora* and **K.** Byrsonima crassiflora associated genomovirus; **L.** Eucalyptus urophylla associated genomovirus - EcaGmV. The nonanucleotide of each species is illustrated. The proposed taxonomic classification is next to the boxes containing the genome and plant. 129

Figura 2. Phylogenetic analysis and SDT using 60 representatives genomoviruses. **A.** Bayesian phylogenetic tree based on Rep aminoacids sequences. Numbers at nodes represent posterior probabilities. Geminiviruses sequences 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 in the Akaike Information Criterion (AIC). The analyzes were carried out by running 100 million generations and sampling every 2000 generations after 2 million burn-in generation. **B.** Genome-wide pair-wise matrix generated by SDT v1.2. Phylogenetic analysis and SDT using 60 representatives genomoviruses sequences added to sequences of EcaGmV,

VoaGmV, ByaGmV, GmaV1, TeaGmV, OuaGmV and MoaGmV (in black). The order, isolates names and accession number (GenBank) are: **1.** Spinach curly top Arizona virus (HQ443515); **2.** *French bean severe leaf curl virus* (JX094280); **3.** *Wheat dwarf virus* (EF536860); **4.** *African cassava mosaic virus* (FM877473); **5.** *Turnip leaf roll virus* (KT388088); **6.** *Beet curly top virus* (AF379637); **7.** *Faeces associated gemycircularvirus* 12 (KF371630); **8.** *Faeces associated gemycircularvirus 11* (KF371631); **9.** *Human genital-associated circular DNA virus-1* (KJ413144); **10.** Eucalyptus Urophilla associated genomovirus (EcaGmV); **11.** Vochisia rufa associated genomovirus (VoaGmV); **12.** Byrsonima crassiflora associated genomovirus (ByaGmV); **13.** *Sewage-associated gemycircularvirus 4* (KJ547634); **14.** *Gila monster - associated gemykrogvirus* (GmaV1); **15.** *HCBI9.212 virus* (LK931484); **16.** *Genomoviridae* sp. (MK032706); **17.** *Gila monster-associated gemycircularvirus* (MH378453); **18.** Car*ibou feces-associated gemycircularvirus* (KJ938717); **19.** *Pacific flying fox faeces associated gemycircularvirus-7* (KT732800); **20.** Tecoma stans associated gemykolovirus (TeaGmV); **21.** - Common bean-associated gemycircularvirus (KX434770); **22.** Common bean-associated gemycircularvirus (KX434768); **23.** *Pacific flying fox faeces associated gemycircularvirus-6* (KT732798); **24.** *Faeces associated gemycircularvirus 10* (KF371632); **25.** *Pacific flying fox faeces associated gemycircularvirus-1* (KT732790); **26.** *Sewage-associated gemycircularvirus 5* (KJ547635); **27.** *Human genital-associated circular DNA virus-1* (KJ413144); **28.** *Faeces associated gemycircularvirus 15* (KT862254); **29.** *Meles meles fecal virus Gemygorvirus* (JN704610); **30.** *Dragonfly-associated circular virus 3* (JX185428); **31.** Plant associated genomovirus 2 (MH939414); **32.** Ouratea duparquetiana associated gemikibivirus (OuaGmV); **33.** *Dragonfly associated Gemykibivirus 1* (JX185430); **34.** *Sewage associated gemycircularvirus 3* (KJ547643); **35.** *Black robin associated gemykibivirus 1* (KF371634); **36.** *HCBI8.215 virus* (LK931483); **37.** Gemycircularvirus SL1 Gemykibivirus (KP133075); **38.** *Rhinolophus associated gemykibivirus 1* (KJ641737); **39.** *Pacific flying fox faeces associated gemycircularvirus-10* (KT732804); **40.** Poecile atricapillus GI tract-associated gemycircularvirus (KT309029); **41.** *Pteropus associated gemycircularvirus 8* (KT732806); **42.** *Pteropus associated gemycircularvirus 5* (KT732801); **43.** *Faeces associated gemycircularvirus 17* (KT862242); **44.** *Miniopterus associated gemycircularvirus 1* (KJ641719); **45.** *Black robin associated gemykibivirus 1* (KF371634); **46.** *Faecal-associated gemycircularvirus 1* (KF371641); **47.** *Faeces associated gemycircularvirus 3* (KF371639); **48.** *Faeces associated gemycircularvirus 4* (KF371638); **49.** *Poaceae-associated gemycircularvirus 1* (KT253577); **50.** *Soybean leaf-associated gemycircularvirus 1* (KT598248); **51.** *Faeces associated gemycircularvirus 16* (KT862251); **52.** *Pacific flying fox faeces associated gemycircularvirus-2* (KT732792); **53.** *Sewage-associated gemycircularvirus 6* (KJ547636); **54.** *Sclerotinia sclerotiorum hypovirulence associated DNA virus 1* (GQ365709); **55.** *Bromus-associated circular DNA virus 3* (KM510192); **56.** *Sewage-associated gemycircularvirus 1* (KM821747); **57.** *Sewage derived gemycircularvirus 1* (KJ547638); **58.** *Faeces associated gemycircularvirus 22* (KT862253); **59.** *Cassava associated gemycircularvirus 1* (JQ412057); **60.** Momordica charantia associated gemycircularvirus (MoaGmV); **61.** Momordica charantia associated gemycircularvirus (MH047857); **62.** *Hypericum japonicum associated circular DNA virus* (KF413620); **63.** Bemisia-associated genomovirus AdO (KY230614); **64.** *Pteropus associated gemycircularvirus 3* (KT732797); **65.** Euphorbia heterophylla associated gemycircularvirus (MH047858); **66.** *Odonata associated gemycircularvirus-1*; **67.** *Dragonfly-associated circular virus 2* (JX185429).....135

LISTA DE TABELAS

Capítulo 1

Tabela 1. Relação dos gêneros da família *Geminiviridae* considerando o número de espécies por gênero, espécie-tipo, grupo de hospedeira, vetor e nonanucleotídeo. 32
Tabela 2. Relação dos gêneros da família *Genomoviridae* considerando o número de espécies, valor de identidade pareada em porcentagem (%) para demarcação de espécie para gênero, espécie-tipo e organismo associado. 41

Chapter 2

Table 1. DNA pools sending for High-Throughput Sequencing (HTS) from tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Samples were organized by microenvironment, collection sites, number of samples, and number of plant species. 74
Table 2. Numbers of reads, total number of contigs, number of contigs associated with ssDNA virus, total number of viruses, number of putative new viruses and the number of previously characterized viruses detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. 77
Table 3. Single stranded DNA and double stranded DNA virus organized by families and detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Genomic information of these viruses were obtained via High-Throughput Sequencing (HTS). 82
Table 4. Single stranded DNA and double stranded DNA virus detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Viruses were organizated considering collections sites, taxonomy and genomic organization and associated with the corresponding GenBank accession. Analyses were carried out via BLASTx in Geneious R11 program and NCBI website. 84
Supplementary Table 1 ST1. Distribution of plants by site, family, scientific name, number of samples for each symptom presented, and description of symptoms presented at the time of collection. 103
Supplementary Table 2 ST2. Oligonucleotides designed based on sequences obtained from High-Throughput Sequencing (HTS) and available in LVV-FITO UnB (Laboratório de Virologia Vegetal – Fitopatologia UnB) and PCR conditions.
.....109
Supplementary Table 3 ST3. Virus distribution by site and number of positive samples and total number of samples. 112

Chapter 3

Table 1. Oligonucleotides used to detect seven species of the *Genomoviridae*, designed using new sequences obtained from High-Throughput Sequencing - HTS, avaliable in LVV-FITO UnB (Laboratório de Virologia Vegetal – Fitopatologia UnB).. 125
Table 2. Single stranded DNA virus detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Viruses were organizated considering collections sites, taxonomy and genomic organization and associated with the corresponding GenBank accession. Analyses were carried out via BLASTx in Geneious R11 program and NCBI website 128
Table 3. Characteristics of the new *Genomoviridae* species organized by collection sites, genomic organization features, proposed new names, genome-full size, nonanucleotide, size genes: *cp*, *repA* and *rep* nucleotide (nt) and aminoacid (aa).....131
Table 4. Conserved motifs exhibited by *Genomoviridae* species. 133

Resumo Geral

Aluna: Josiane Goulart Batista. Título: Viroma de plantas em áreas nativas e cultivadas do Distrito Federal. Programa de Pós-Graduação em Fitopatologia. Data da defesa:30/01/2020. **Orientadora:** Rita de Cássia Pereira-Carvalho. Tese (Doutorado em Fitopatologia) – Universidade de Brasília, Brasília, DF.

O Brasil possui grande parte do seu território ocupado por vegetação nativa e áreas plantadas com espécies florestais. Nos últimos anos, a ação antrópica caracterizada pela expansão da agricultura e pecuária alteraram esse cenário de maneira significativa. A introdução de práticas que modificam o uso da terra pode influenciar na microbiota presente em ambientes naturais, de transição e agrícola. A região central do Brasil, predominantemente inserida no bioma Cerrado, é um polo de produção agrícola, onde as muitas áreas nativas foram substituídas por culturas extensivas. Pouco ainda se sabe sobre a diversidade viral presente nas espécies arbóreas nativas desse bioma. Eventos geradores de variabilidade em vírus de plantas podem propiciar a adaptação a novos ambientes e surgimento de novas espécies bem como facilitar de dispersão por artrópodes. Neste contexto, a expansão de fronteiras agrícolas pode favorecer o fluxo/trânsito de vírus entre os diferentes ambientes. Avanços nas tecnologias de sequenciamento e análise metagenômica têm permitido a detecção e caracterização molecular de um número crescente de espécies virais. Assim, o objetivo deste trabalho foi prospectar a distribuição de vírus associados a espécies arbóreas e cultivadas em diferentes microambientes da área de *Cerrado* de Brasília-DF (Distrito Federal), no Brasil Central. Ao todo, 296 amostras foliares (134 espécies e 49 famílias botânicas) foram obtidas em áreas de três microambientes: **(a) Áreas antrópicas** (n=176); **(b) Áreas de conservação** (n=60) e **(c) Áreas de transição** (n=60). As extrações de DNA foram realizadas individualmente e seguidas de enriquecimento via círculo rolante (RCA). As amostras foram agrupadas em cinco conjuntos e sequenciadas em uma plataforma Illumina Hiseq2500. Os resultados do *High-Throughput Sequencing* (HTS) geraram um total de 167.689,52 leituras formando 297.103 *contigs*, dos quais 503 mostraram identidade com vírus vegetais em análises com o algoritmo BLASTx. Foram recuperados 30 *contigs* (= 24 espécies virais), permitindo a identificação de 12 novas espécies das famílias: *Geminiviridae* (01), *Genomoviridae* (04), *Circoviridae* (04) bem como vírus de ssDNA sem classificação (03). Dezoito *contigs* (= 12 espécies virais já descritas) foram caracterizados nas famílias: *Geminiviridae* (08), *Genomoviridae* (02) e *Caulimoviridae* (02). O resultado de HTS foi validado empregando *primers* específicos que foram utilizados para detecção e/ou recuperação do genoma

completo (obtido mediante sequenciamento via *Sanger*). Foi observada uma prevalência de vírus em áreas de ambiente antrópico (com especial destaque para membros da família *Geminiviridae*). Poucos vírus foram detectados nas áreas de ambiente de conservação. No ambiente de conservação, as detecções foram de vírus classificados na família *Genomoviridae*. No ambiente de transição, observou-se que algumas espécies que ocorrem sob condições de cultivo também estão ocorrendo em espécies arbóreas. Espécies botânicas classificadas nas famílias Solanaceae e Fabaceae apresentaram o maior número de espécies de vírus. Seis *contigs* que apresentaram identidade com representantes da família *Genomoviridae* foram caracterizados molecularmente. Destes, quatro correspondiam a novas espécies e dois a espécies já conhecidas. Além disto, dois isolados de *Cauliflower mosaic virus* (*Caulimovirus*, *Caulimoviridae*) foram detectados em couve e em uma nova hospedeira – *Fevillea trilobata*. Foi também possível detectar uma nova espécie de ssDNA (até o momento sem classificação em gênero e família) em *Caesalpinia pluviosa*. Além disso, sequências do genoma completo de sete espécies virais foram recuperadas por HTS: *Sugarcane baciliforme virus* – SBCV (*Badnavirus*, *Caulimoviridae*), duas novas espécies de ssDNA (sem classificação) e quatro novas espécies de *Circoviridae*. **(Capítulo 2).** Foi acrescentada uma nova espécie da família *Genomoviridae* em *Eucalyptus urophylla*, totalizando sete espécies. Foram identificadas sequências de três espécies através de análises de identidade pareada e filogenia que ficaram próximas a uma espécie do gênero *Gemyvongvirus* detectadas em *Vochysia rufa*, *Byrsonima crassiflora* e *E. urophylla*, para as quais foram propostos os nomes Vochysia rufa associated genomovirus – VoaGmV, Byrsonima crassiflora associated genomovirus e Eucalyptus urophylla associated genomovirus, respectivamente. Sequências de duas novas espécies virais foram detectadas: Tecoma stans associated gemykolovirus (*Gemykolovirus*) em *Tecoma stans* e *Ouratea duparquetiana* associated gemykibivirus (*Gemykibivirus*) em *Ouratea duparquetiana*. Além disto, foram detectados dois isolados das espécies *Gila monster-associated gemykrogvirus* - (*Gemykrogvirus*) em *Trembleya parviflora* e Momordica charantia associated gemycircularvirus (*Gemycircularvirus*) em berinjela **(Capítulo 3).** Treze *contigs* apresentaram identidade com sequências de nove espécies da família *Geminiviridae*. Isolados das nove espécies virais foram considerados como: (**1) Nova espécie:** Macroptilium bright yellow interveinal virus – MaBYIV (*Begomovirus*) detectada em *Macroptilium erythroloma* em infecção mista com isolado de *Bean golden mosaic virus* – BGMV (*Begomovirus*). Ensaios de transmissão via mosca-branca (*Bemisia tabaci* MEAM 1) foram conduzidos usando como fonte de inóculo *M. erythroloma* para cultivares de feijão, caupi, amendoim, soja e tomate. A transmissão de MaBYIV ocorreu para plantas de feijão e soja. BGMV foi transmitido para todas as plantas. (**2) Isolados de duas espécies da família**

***Geminiviridae* sem relato no país e/ou com sequência divergente:** primeiro relato no país de um isolado de Tomato apical leaf curl virus – ToALCV proveniente do tomateiro e registro de dois isolados divergentes de Tomato associated geminivirus 1 – TaGV1 no tomateiro e em uma nova hospedeira (beterraba). Clones de ToALCV e TaGV1 foram inoculados via biobalística nas hospedeiras originais e em *Nicotiana benthamiana*. Somente o isolado de TaGV1 foi capaz de causar infecção e replicar-se em *N. benthamiana*. (**3) Isolados de quatro espécies virais já conhecidas detectados em novas hospedeiras:** dois isolados de Maize striate mosaic virus – MSMV (*Mastrevirus*) em cana de açúcar (*Saccharum officinarum*); novos isolados de BGMV em berinjela, angico (*Anadenanthera colubrina*) e macroptilium; um isolado de *Tomato severe rugose virus* (*Begomovirus*) em algodão (*Gossypium hirsutum*) e dois isolados de *Tomato mottle leaf curl virus* (*Begomovirus*) em leiteiro (*Euphorbia heterophylla*) e o outro em *Ouratea duparquetiana*. (**4) Isolados de cinco espécies (gênero *Begomovirus*) em hospedeiras já conhecidas:** vinte e quatro isolados de ToSRV, sendo em tomate (11), berinjela (11) *Nicandra physalodes* (1) e *Sida* sp. (1); três isolados de ToMoLCV em tomate; sete isolados de BGMV em feijão; quatro isolados de *Sida micrantha mosaic virus* (*Begomovirus*) em algodão (2), guanxuma (1) e *N. physalodes* (1) e quatro isolados de *Sweet potato leaf curl virus* (*Begomovirus*) em batata-doce e (**5) Isolados de *Mastrevirus* em hospedeira já descrita:** dois isolados de MSMV foram recuperados de milho (**Capítulo 4**). Os resultados do presente trabalho indicam uma considerável diversidade viral em espécies arbóreas, cultivadas e daninhas. HTS e metagenômica foram importantes ferramentas no estudo e caracterização da diversidade viral em diferentes ambientes naturais e antrópicos.

Palavras chaves: arbóreas, agrícolas, diversidade, ssDNA virus, dsDNA virus, Cerrado

General Abstract

Student: Josiane Goulart Batista. Title: Viroma of plants in native and cultivated areas of the Federal District. Graduate Program in Phytopathology. Defense date: 01/30/2020. **Advisor**: Rita de Cássia Pereira-Carvalho. Thesis (PhD in Phytopathology) - University of Brasília, Brasília, DF.

A large fraction of the Brazilian has territory is occupied by native vegetation and areas planted with forest species. In recent years, the anthropic action characterized by the expansion of agriculture and livestock production has significantly changed this scenario. The introduction of novel agricultural practices may affect the microbiota present in natural, transitional, and agricultural environments. The central region of Brazil, predominantly located within the *Cerrado* biome, is a major agricultural region, where native areas have been replaced by extensive crops. Little is known about the viral diversity present in native tree species in this biome. Events that generate variability in plant viruses can provide adaptation to new environments and the emergence of new species, as well as facilitate dispersion by arthropods. In this context, the expansion of agricultural frontiers may favor the flow of viruses among distinct environments. Advances in sequencing technologies and metagenomic analysis, have enabled the large-scale detection and molecular characterization of novel viral species. Thus, the objective of this work was to prospect the viruses associated with tree and cultivated species cultivated under different microenvironments in the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Altogether 296 leaf samples (from 134 species and 49 botanical families) were obtained in areas of three microenvironments: **(a) Anthropic areas** (n=176); **(b) Conservation areas** (n=60) and **(c) Transition areas** (n=60). DNA extractions were performed individually and were followed by enrichment by means of a rolling circle amplification reaction (RCA). The samples were grouped into five sets and sent for sequencing on an Illumina Hiseq2500 platform. The results of the High-Throughput Sequencing (HTS) generated a total of 167,689.52 readings forming 297,103 contigs, of which 503 showed identity with plant viruses in analyzes using the BLASTx algorithm. Thirty contigs (corresponding to 24 viral species) were recovered, allowing the identification of 12 new species of the families: *Geminiviridae* (01), *Genomoviridae* (04), *Circoviridae* (04) as well as unclassified ssDNA virus (03). Eighteen contigs (= 12 viral species already described) were characterized in the families: *Geminiviridae* (eight species), *Genomoviridae* (two species) and *Caulimoviridae* (two species). The HTS result was validated using specific primers that were used for detection and / or recovery of the complete genome. A prevalence of viruses was observed in areas of anthropic environment (with

special emphasis on members of the *Geminiviridae* family). Few viruses were detected in the conservation environment areas. In the conservation environment, the detections were of viruses classified in the *Genomoviridae* family. In the transition environment, it was observed that some species that occur under cultivation conditions are also occurring in trees. Botanical species classified in the families Solanaceae and Fabaceae presented the highest number of virus species. Six contigs that showed identity with representatives of the *Genomoviridae* family were characterized molecularly, detected by PCR and confirmed by Sanger. Of these, four corresponded to new species and two to species already known. In addition, two *Cauliflower mosaic virus* isolates (*Caulimovirus, Caulimoviridae*) were detected in cabbage (*Brassica oleraceae*) and in a new host – *Fevillea trilobata*. It was also possible to detect a new species of ssDNA in *Caesalpinia pluviosa* (so far unclassified by gender and family). In addition, sequences of the complete genome of seven viral species were recovered by HTS: *Sugarcane baciliforme virus* (*Badnavirus* / *Caulimoviridae*), two new species of ssDNA without classification and four new species of the family *Circoviridae*. **(Chapter 2)**. A new species of the *Genomoviridae* family was added to *Eucalyptus urophylla*, totaling seven species. Sequences of three species were identified through analysis of paired identity and phylogeny that were close to a species of the genus *Gemyvongvirus* detected in *Vochysia rufa*, *Byrsonima crassiflora* and *E. urophylla*, for which the names Vochysia rufa associated genomovirus, Byrsonima crassiflora associated genomovirus and Eucalyptus urophylla associated genomovirus, respectively. Sequences of two new viral species were detected: Tecoma stans associated gemykolovirus (*Gemykolovirus*) in Tecoma stans and Ouratea duparquetiana associated gemykibivirus (*Gemykibivirus*) in Ouratea duparquetiana. In addition, two isolates of the species *Gila monster-associated gemykrogvirus* (*Gemykrogvirus*) were detected in *Trembleya parviflora* and Momordica charantia associated gemycircularvirus (*Gemycircularvirus*) in eggplant **(Chapter 3)**. Thirteen contigs showed identity with sequences of nine species of the *Geminiviridae* family. Isolates from the nine viral species were considered as: **(1) New species:** Macroptilium bright yellow interveinal virus - MaBYIV (*Begomovirus*) detected in *Macroptilium erythroloma* in mixed infection with *Bean golden mosaic virus* isolate - BGMV *(Begomovirus).* Transmission tests via whitefly (*Bemisia tabaci* MEAM 1) were conducted using *M. erythroloma* as inoculum source for *Phaseolus vulgaris*, *Vigna unguiculata*, *Arachys hypogaea*, soy and tomato cultivars. The transmission of MaBYIV occurred to bean and soybean plants (with 40% efficiency). BGMV was transmitted to all plants. **(2) Isolates of two species of the Geminiviridae family with no report in the country and / or with divergent sequence:** one Tomato apical leaf curl virus isolate – ToALCV from tomato and corresponding to the first report of this virus in the country and two divergent

isolates of Tomato associated geminivirus 1 - TaGV1 in tomato and table beet. This is the first report of TaGV1 in beet. ToALCV and TaGV1 clones were inoculated via biobalistics in the original hosts and in *Nicotiana benthamiana*. Only the TaGV1 isolate was able to cause infection and replicate in *N. benthamiana*. **(3) Isolates of four known viral species detected in new hosts:** two isolates of Maize striate mosaic virus (*Mastrevirus*) in sugar-cane; new BGMV isolates in eggplant, angico (*Anadenanthera colubrina*) and macroptilium; one isolate of *Tomato severe rugose virus* (Begomovirus) in cotton (*Gossypium hirsutum*) and two isolates of *Tomato mottle leaf curl virus* (*Begomovirus*) in dairy (*Euphorbia heterophylla*) and the other in *Ouratea duparquetiana*. **(4) Isolates of five species (genus *Begomovirus*) in already known hosts:** twenty-four isolates of ToSRV, being in tomato (11), eggplant (11) *Nicandra physalodes* (01) and *Sida* sp. (01); three ToMoLCV isolates in tomatoes; seven BGMV isolates in beans; four isolates of *Sida micrantha mosaic virus* (*Begomovirus*) in cotton (02), *Sida* sp. (01) and *N. physalodes* (01) and four isolates of *Sweet potato leaf curl virus* (*Begomovirus*) in sweet-potatoes and **(5) *Mastrevirus* isolates in a host already described:** two MSMV isolates were recovered from maize **(Chapter 4)**. The results of the present work indicate a considerable viral diversity in tree species, field crops, and weeds. In conclusion, HTS and metagenomics were important tools in the study and characterization of viral diversity in different natural and anthropic environments.

Keywords: arboreal, agricultural, diversity, ssDNA virus, dsDNA virus, *Cerrado*.

Introdução Geral

As espécies arbóreas desempenham importante papel em ecossistemas florestais, agrícolas e urbanos em todo o mundo (Lindenmayer et al. 2012). A importância destas espécies arbóreas não se restringe apenas ao armazenamento de carbono (Harmon et al. 1986), mas por gerar também microambientes distintos, caracterizados por altos níveis de nutrientes do solo e riqueza de espécies de plantas (Harmon et al. 1986, Manning et al. 2006), além de fornecer habitat de nidificação e abrigo para inúmeras espécies de animais (> 350 espécies de mamíferos em todo o mundo) (Nowak and Walker 1999, Remm and Lõhmus 2011) incluindo até 30% da biota de vertebrados em um determinado tipo de vegetação (Rose et al. 2001, Gibbons and Lindenmayer 2002) e diversos microrganismos (Hardoim et al. 2015).

Espécies arbóreas abastecem o mercado através da oferta de produtos madeireiros e não madeireiros, movimentando em média 2% do PIB (Produto interno bruto) mundial (Food and Agriculture Organization of United Nations - FAO, 2018). Cerca de dois bilhões de pessoas dependem das florestas para subsistência (World Wide Fund for Nature - WWF, 2018). No Brasil o setor florestal representa 6% do PIB industrial e gera milhões de empregos diretos e indiretos (IBA - Indústria Brasileira de Árvores, 2018).

O Brasil é o país com a maior biodiversidade de espécies arbóreas do planeta, distribuída em seis biomas (Amazônia, Caatinga, Cerrado, Mata Atlântica, Pantanal e Pampa) (das Dores et al. 2017). O bioma Cerrado, também conhecido como savana brasileira é um *hotspot* de biodiversidade que está ameaçado devido à expansão agropecuária (WWF, 2018). Para preservar e conservar o seu ambiente natural foi criada a Lei n° 9985 que institui o Sistema Nacional de Unidades de Conservação da Natureza/ SNUC, estabelecendo critérios e normas para criação, implantação e gestão das unidades de conservação visando manter espaços territoriais protegidos para garantir um meio ambiente equilibrado (BRASIL, 2000).

Neste contexto, visando à conservação e preservação das espécies arbóreas principalmente do bioma Cerrado têm-se no Distrito Federal (DF) o Viveiro II da NOVACAP (Companhia Urbanizadora da Nova Capital) que fornece mudas para arborização. As unidades de conservação (UCs) já registradas no CNUC (Cadastro Nacional de Unidades de Conservação) são 2412 (somados os parques e florestas, estaduais e municipais) no Brasil e 107 no DF e entorno. Dentre as UCs do DF foram estudadas neste trabalho: **1)** Estação Ecológica de Águas Emendadas (ESECAE), **2)** Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR) e **3)** Área de Conservação: Área de Proteção Ambiental Cabeça de Veado (APACV-FAL) localizada na Fazenda Água Limpa - FAL (vinculada à Universidade de Brasília - UnB) e na RECOR. Estes locais fazem fronteira entre si. Essa APA é próxima a áreas de cultivos agrícolas, assim como uma área de conservação situada na Estação Experimental de Biologia (EEB - UnB), a qual provavelmente poderá ser considerada futuramente junto ao SNUC como Área de Interesse Ecológico (ARIE -EEB) devido às características por ela apresentada. Outra área de ocupação do Cerrado é o Núcleo Rural Taquara responsável pela produção de hortaliças que abastecem o DF (NRT).

As espécies arbóreas podem ser acometidas por pragas e doenças (incluindo aquelas de etiologia viral) que podem representar ameaça em potencial transitando do ambiente nativo para as lavouras e vice-versa (Rodríguez-Nevado et al. 2019). Importante salientar que a globalização faz com que haja intercâmbio de materiais vegetais entre países, inclusive material vegetal infectado com vírus, representando uma grande risco para espécies florestais em ecossistemas naturais e em áreas cultivadas em todo o mundo (Webster et al. 2007, Vincent et al. 2014, Fraile and García-Arenal 2016). No entanto, pouca atenção tem sido dedicada ao papel e ou ameaça que tais vírus podem representar para espécies arbóreas e seus potenciais impactos na redução da biodiversidade global.

Com a disponibilidade de novas tecnologias de sequenciamento ou tecnologias de alto rendimento (*Next Generation Sequencing*/NGS, também denominadas High-throughput sequencing/HTS) houve uma revolução nos estudos biológicos, e na virologia vegetal com o advento da metagenômica (Handelsman 2004). Para vírus de plantas, os trabalhos iniciaram em 2009 (Adams et al. 2009, Kreuze et al. 2009 Al Rwahnih et al, 2009) e vários estudos foram realizados depois disto (Wu et al. 2015, Hadidi et al. 2016, Posada-Cespedes et al. 2017, Xu et al. 2017, Glasa et al. 2019), entretanto para espécies arbóreas há poucos trabalhos (Roossinck et al. 2010, Jo et al. 2018, Bernardo et al. 2018, Fonseca et al. 2018). Os relatos de vírus em espécies arbóreas demonstram que a maioria corresponde a espécies novas (Muthukumar et al. 2009, Roossinck et al. 2010, Bernardo et al. 2018), embora exista também em trabalhos anteriores, relatos de vírus que ocorrem em espécies cultivadas infectando espécies arbóreas (Nienhaus and Castello 1989, Büttner et al. 2013, Vincent et al. 2014).

No Brasil, os estudos de vírus em espécies arbóreas são incipientes, principalmente para espécies nativas do país (Beserra Jr et al. 2011, Nicolini et al. 2012, Fonseca et al. 2018).

Neste contexto a hipótese do trabalho é de que existe diversidade de vírus, alguns deles, provavelmente desconhecidos para a Ciência ocorrendo em ambientes antrópicos, conservação e transição. Acredita-se também que espécies arbóreas possam atuar como reservatórios de vírus que ocorrem em culturas agronômicas quando próximas a áreas de ecossistemas naturais. Postula-se ainda que os diferentes ambientes possam favorecer eventos de recombinações que podem originar novos vírus que poderão causar perdas no futuro tanto para as espécies arbóreas como para as espécies cultivadas.

O termo “viroma” é definido como o genoma de todos os vírus que habitam um organismo ou ambiente específico (Virgin 2014). Sendo assim, o objetivo principal deste trabalho foi estudar o viroma de espécies arbóreas (nativas e exóticas), espécires daninhas e espécies cultivadas em áreas de diferentes microambientes: **A) Áreas Antrópicas**: 1- Arborização Urbana do Distrito

Federal (AUDF), 2- Núcleo Rural Taquara (NRT), 3- Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP) e 4- Área Experimental da Fazenda Água Limpa (AE-FAL); **B) Áreas de Conservação:** 1- Estação Ecológica de Águas Emendadas (ESECAE) e 2- Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR) e **C) Áreas de Transição**: 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE - EEB).

Hipóteses do Trabalho

Existe diversidade de vírus, alguns ainda desconhecidos para a Ciência, ocorrendo em espécies arbóreas em ecossistemas naturais.

Espécies arbóreas podem atuar como reservatório de vírus que infectam culturas agronômicas e situam-se próximas às áreas de sistemas nativos.

As regiões de transição entre sistemas naturais e de cultivo podem favorecer mudanças evolutivas e originar espécies virais novas.

Objetivos Gerais

- Prospectar vírus em espécies arbóreas, daninhas e cultivadas e;
- Identificar e caracterizar molecularmente vírus ocorrendo em espécies arbóreas, em áreas de conservação e de transição do Distrito Federal utilizando High-throughput sequencing/HTS.

Objetivos Específicos

- Identificar vírus em espécies arbóreas, cultivadas e daninhas coletadas em áreas de diferentes locais classificados em três tipos de ambientes:

A) Áreas Antrópicas:

1- Arborização Urbana do Distrito Federal (AUDF),

2- Núcleo Rural Taquara (NRT),

3- Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP) e

4- Área Experimental da Fazenda Água Limpa (AE-FAL).

B) Áreas de Conservação:

1- Estação Ecológica de Águas Emendadas (ESECAE) e

2- Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR).

C) Áreas deTransição:

1- "Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) e

2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB).

- Realizar a caracterização molecular de vírus encontrados via High-throughput sequencing/HTS,
- Realizar a caracterização biológica de pelo menos duas espécies virais encontradas em diferentes ambientes.
- Realizar análises filogenéticas dos novos vírus ocorrendo em espécies arbóreas, daninhas e cultivadas e
- Realizar análises de recombinação das espécies novas.

Capítulo 1

Revisão Bibliográfica

1. Importância florestal no contexto mundial

De acordo com o Banco Mundial (2018), o total de florestas no mundo cobre pouco menos de quatro bilhões de hectare (ha), correspondendo a cerca de 30% da superfície terrestre e inclui as áreas de florestas naturais e plantadas.

Os cinco países com maior área de florestas naturais são, nesta ordem: Rússia (20%), Brasil (12%), Canadá (9%), Estados Unidos - EUA (8%) e China (5%). Juntos, estes países, contam com mais de 50% da área de florestas em todo o mundo. As florestas nativas primárias somam 36% do total de área de florestas no planeta, mas houve redução de mais de 40 milhões de hectares desde 2000 (FAO, 2018), o que justifica uma maior atenção para preservar este ambiente natural.

Segundo relatórios do Banco Mundial, em 2018, 20% da população mundial (cerca de 1,6 bilhão de pessoas) dependiam das florestas naturais para sua subsistência, sendo que 350 milhões de pessoas que viviam dentro ou perto de florestas densas dependiam delas para sua subsistência e renda. Destes, cerca de 60 milhões de pessoas (especialmente comunidades indígenas) são totalmente dependentes de florestas.

As áreas de florestas plantadas representam 7% da área florestal mundial, o equivalente a 291 milhões de ha, e estão voltadas para abastecer o setor florestal. Os países com maior área de florestas plantadas são: a China representando 29% da área mundial, os EUA com 10%, a Rússia com 6%, Japão e Índia ambos com 4% e o Brasil apresentando 3% (FAO, 2018).

O setor florestal movimenta em média 2% do PIB mundial (FAO, 2018). No ano de 2018, o mercado mundial de produtos florestais, representado pelo valor total das exportações dos países, movimentou aproximadamente 230 bilhões de dólares. A estimativa é que o setor florestal gere 10 milhões de empregos diretos e entre 30 a 50 milhões de forma indireta (FAO, 2018).

O país de maior importância no comércio florestal é a China, em produção e consumo de produtos florestais, ultrapassando recentemente países como Canadá e EUA que são considerados potências no setor florestal (FAO, 2018).

As espécies florestais possuem grande importância econômica por abastecer o setor florestal fornecendo energia e/ou matéria-prima para a indústria da construção civil e de transformação em virtualmente todas as regiões do mundo (FAO, 2018). Além da importância econômica, as espécies florestais desempenham importante função ambiental e ecológica, contribuindo na conservação da água, na preservação do solo e da biodiversidade e na manutenção das condições climáticas (Crowther et al. 2015).

As florestas contribuem para a ciclagem de nutrientes (Power 2010), formação do solo (Pimentel and Kounang 1998), clima (Daily and Matson 2008) e regulação da água (De Groot et al. 2002). As florestas também são reconhecidas pelo seu potencial em recursos naturais e habitats para a fauna (Vitousek et al. 1986, Rojstaczer et al. 2001). Além disso, práticas de manejo integrado e agroflorestas podem contribuir na regulação do controle de pragas (Klein et al. 2006, Bale et al. 2007, Karp et al. 2013).

1.1. Importância das florestas para o Brasil

O Brasil representa a maior área de florestas tropicais do mundo. Aproximadamente 58% do território brasileiro é coberto por florestas naturais e plantadas (WWF - World Wild Fund, 2019).

O país é considerado o segundo maior em área florestal do planeta, superado pela Rússia que possui floresta temperada (SNIF - Setor Nacional de Informações Florestais, 2019). São estimados 485,8 milhões de hectares de florestas nativas e 7,8 milhões de hectares de florestas plantadas (SNIF, 2016). As condições edafoclimáticas, o espaço territorial e as áreas de florestas fizeram com que o Brasil se destacasse no mercado florestal.

No cenário mundial da economia florestal, o Brasil é o maior exportador de polpa para papel e o segundo maior exportador de celulose, perdendo apenas para o Canadá (FAO, 2018). Em 2018, o faturamento do setor florestal representou 1,1% de toda a riqueza gerada no país e 6% do PIB industrial, uma receita de mais de 60 bilhões de reais, gerando mais de 4 milhões de empregos diretos e indiretos (IBA, 2018).

De acordo com IBA (2018) existem 7,8 milhões de florestas plantadas, sendo as espécies predominantes: eucalipto (*Eucalyptus* spp. e *Corymbia* spp.: 72%), pinus (*Pinus* spp.: 20,7%), seringueira (*Hevea brasiliensis*: 2,27%), acácia (*Acacia mangium* e A. *mearnsii*: 1,93%), teca (*Tectona grandis*: 1,16%) e paricá (*Schizolobium amazonicum*: 1,15%). Outras espécies somam 0,79%. Os plantios florestais estão distribuídos por todo o país, entretanto, os principais estados com plantios de eucalipto são: Minas Gerais (MG), São Paulo (SP), Mato Grosso do Sul (MS), Bahia (BA), Rio Grande do Sul (RS) e Espírito Santo (ES). Os estados do Paraná (PR), Santa Catarina (SC) e RS se destacam na quantidade de área destinada às plantações de pinus. As plantações de seringueira, acácia, teca, paricá entre outras, estão localizadas principalmente nos estados do Mato Grosso (MT), RS, SP e Pará (PA).

O setor florestal brasileiro está representado em todos os estados federativos e os cinco estados brasileiros que mais exportaram, em 2018, em valor, foram SP, PR, BA, Espírito Santo e MS (SNIF, 2018).

2. O Bioma Cerrado

O Brasil é o país com maior biodiversidade do mundo e destaca-se no cenário mundial devido ao extenso território de florestas naturais que, de acordo com características peculiares, encontra-se dividido em seis biomas: Amazônia, Caatinga, Cerrado, Mata Atlântica, Pampa e Pantanal (MMA - Ministério do Meio Ambiente, 2017).

O Cerrado é uma das formações de savana mais ricas do mundo em termos de seres vivos e abrigam 5% de todas as espécies vivas na Terra e uma em cada dez espécies brasileiras (WWF, 2018).

De acordo com Rezende et al. (2008) a flora do bioma Cerrado está representada por mais de 12 mil plantas classificadas nas famílias Fabaceae, Melastomataceae, Vochysiaceae, Malpighiaceae, Clusiaceae, Erythroxylaceae e Myrtaceae. Deste total mais de cinco mil espécies foram consideradas endêmicas.

Embora não se conheça até o momento a diversidade viral nos diferentes Biomas brasileiros, para fungos, alguns trabalhos tem demonstrado a grande diversidade presente em Unidades de Conservação do DF, Estação Ecológica de Águas Emendadas - ESECAE [Sepúlveda-Chavera et al. (2008)] e Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística - RECOR [Pereira-Carvalho et al. (2004)]. Sepúlveda-Chavera et al. (2008) relataram mais de 21 espécies de fungos associados a uma única hospedeira, em amostras da ESECAE, enquanto Pereira-Carvalho et al. (2004) relataram mais de 44 espécies de fungos em uma única hospedeira em amostra da RECOR.

Além da sua biodiversidade única, o Cerrado é importante no abastecimento de água, produção de pastagem para gado e armazenamento de carbono. A maioria das nascentes dos rios brasileiros está em áreas do Cerrado, sendo oito das doze principais nascentes de bacias hidrográficas do Brasil (Overbeck et al. 2015).

O bioma Cerrado abrange aproximadamente dois milhões de Km² (em torno de 24% da superfície do território nacional) estando presente nos estados de Góias (GO), MT, MS, Tocantins (TO), SP, BA, Ceará (CE), Maranhão (MA) e Piauí (PI) e no DF (ICMBio, 2016).

A expansão agrícola no Cerrado teve início a partir da década 70, devido à modernização agrícola e as condições físicas e geográficas da região (Frederico 2012). Essa mudança no uso da terra fez com que apenas 51,54% de áreas remanescentes da vegetação original sejam

encontradas. A vegetação nativa deu lugar as *commodities*, monoculturas de grãos, como o milho (*Zea mays*), duas importantes fabáceas: soja (*Glycine max*) e feijão (*Phaseolus vulgaris*) (Fernandes and Pessôa 2011), hortaliças classificadas na família Solanaceae, como tomate (*Solanum lycopersicum*) e pimentão (*Capsicum annuum*), e ainda a *commodity* algodão (*Gossypium hirsutum*) (Rosolen et al. 2012b).

Desta forma, o Cerrado passou a ser observado como um novo cenário composto por áreas de sistemas nativos próximos a sistemas agrícolas e por vez, ambos os sistemas formando a agrofloresta. A pressão de seleção nos fitopatógenos, mais precisamente, nos vírus pode vir a favorecer o trânsito de espécies de um ambiente para o outro principalmente pela transmissão viral, via vetor e mecanicamente (Vincent et al. 2014).

3. Unidades de Conservação no Brasil

Segundo o World Database on Protected Areas (WDPA), órgão responsável pelo banco de dados de áreas de proteção, as áreas protegidas (APs) abrangem mais de 32 milhões de km² distribuídas por 193 países, podendo ser classificadas da seguinte forma: 15,4 % da superfície terrestre e das ilhas marinhas; 3,4% da área dos oceanos; 8,4% de todas as zonas marinhas, e 10,9% de todas as águas costeiras (Juffe-Bignoli et al. 2014, Franks 2016).

No Brasil, as áreas de proteção são regidas pela Lei 9.985/2000 que criou o Sistema SNUC. O SNUC é o conjunto de Unidades de Conservação (UCs) federais, estaduais e municipais com função de sistematizar e criar as UCs. As UCs são definidas como “espaços territoriais, que têm a função de assegurar a representatividade de amostras significativas e ecologicamente viáveis das diferentes populações, com características naturais relevantes, incluindo seus recursos ambientais, habitats e ecossistemas do território nacional e das águas jurisdicionais, preservando o patrimônio biológico existente” (Mercadante 2001, Drummond et al. 2006, Drummond et al. 2010).

As UCs se dividem em **Unidades de Proteção Integral:** que consistem em áreas que visam proteger a natureza de maneira mais restrita, permitindo apenas o uso indireto dos recursos naturais e **Unidades de Uso Sustentável:** áreas que buscam conciliar conservação da natureza com o uso sustentável dos recursos, é considerada menos restrita, permitindo a coleta e uso dos recursos naturais, entretanto de forma renovável (MMA, 2018).

As Unidades de Proteção Integral são: Estação Ecológica - ESEC, Reserva Biológica - REBIO, Parque Nacional - PARNA, Monumento Natural - MN e Refúgio da Vida Silvestre - REVIS, enquanto as Unidades de Uso Sustentável são: Área de Proteção Ambiental - APA, Área de Relevante Interesse Ecológico - ARIE, Floresta Nacional - FLONA, Reserva Extrativista - RESEX, Reserva de Fauna - REFAU, Reserva de Desenvolvimento Sustentável - RDS e Reserva Particular do Patrimônio Natural - RPPN (MMA, 2018).

O Brasil possui legalmente nove UCs que se dividem em 32 ESEC, 3 MN, 73 PARNA, 7 REVIS, 31 REBIO, 67 FLONA, 62 RESEX, 33 APA, 16 ARIE, 2 RDS, 1 RESEX, 1 REVIS e 634 RPPN (Não fazem parte desses números os parques e florestas, estaduais e municipais) (MMA, 2018).

O Distrito Federal (DF) e entorno possui um grande número de UCs, dentre elas: 5 APA, 12 ARIE, 2 ESES; 1 MN; 5 REBIO; 1 PARNA; 5 RPPN; 1 FLONA; 73 Parques e 2 Reserva Ecológica - RECO (uma nova categoria de unidade de conservação que possui finalidade científica). Por estar próximo à capital federal as UCs oferecem excelente oportunidade para a elaboração de estudos (IBRAM, 2018).

A Reserva Ecológica do IBGE (RECOR), o Parque Nacional de Brasília (PNB), a Estação Ecológica de Águas Emendadas (ESECAE) e a Estação Ecológica do Jardim Botânico de Brasília, formam o conjunto de unidades de conservação permanente do DF.

A ESECAE está localizada, no extremo nordeste do Distrito Federal, a uma distância de 50 km do centro Brasília (aqui considerado o marco zero de Brasília, cruzamento dos eixos

Rodoviário e Monumental da capital do Brasil) e ocupa uma área de mais de 10 mil ha, sendo destinada à educação ambiental, realização de pesquisas biológicas e ecológicas além de proteção do ambiente natural. Rocha et al. (2008) realizaram levantamento de flora na ESECAE e catalogaram 1.738 espécies, sendo 612 gêneros e 125 famílias.

A RECOR, criada em 1975 para ser um centro de pesquisas, encontra-se localizada a 35 km ao sul do centro de Brasília, fazendo limite a sudoeste com a Fazenda Água Limpa (FAL/UnB) e ocupando área de aproximadamente 1.398,9 ha (RECOR, 2018).

No Distrito Federal - DF, um importante reduto de produção de mudas do Cerrado é o viveiro II da NOVACAP (Companhia Urbanizadora da Nova Capital do Brasil) localizado próximo ao Parque Nacional (a dois km de distância), envolto por um ambiente nativo. A NOVACAP foi fundada em 1971, inicialmente com o Viveiro I produzindo espécies exóticas e ornamentais. Após algum tempo verificou-se que muitas espécies de plantas não se adaptavam ao ambiente do Cerrado e havia necessidade de manter espécies do ambiente original. Para isso criou-se (em 1971) o Viveiro II da NOVACAP para produção de mudas visando arborização do DF com espécies nativas e endêmicas do Cerrado tais como: ipês-amarelos (*Handroanthus serratifolius*), roxos (*Handroanthus impetiginosus*) e brancos (*Tabebuia roseo-alba*), quaresmeiras (*Miconia* spp.; *Tibouchina* spp.), sucupiras (*Pterodon emarginatus*) e copaíbas (*Copaifera langsdorffii*). A produção anual de mudas no viveiro II supera 300 mil mudas (NOVACAP, 2018).

Neste contexto de áreas de reserva, a Universidade de Brasília possui duas áreas destinadas à pesquisa: a Fazenda Água Limpa - FAL e a Estação Experimental de Biologia - EEB com infraestrutura básica voltada para o processo de ensino, pesquisa e extensão com ênfase às áreas de Agronomia, Biologia, Engenharia Florestal, Ecologia, Botânica, Zoologia, Fisiologia, Zootecnia e Fitopatologia (UnB, 2018). A seguir é feita uma breve descrição destas duas áreas.

A Fazenda Água Limpa (FAL) faz fronteira com a RECOR e ambos os locais possuem uma parte do seu território inserido na APA Cabeça de Veado. A FAL possui uma área de aproximadamente, 4.340 ha destinada à preservação (2.340 ha), na qual está inserida a APA Cabeça de Veado, conservação (800 ha) e produção agrícola (1.200 ha) (UnB, 2018).

A Estação Experimental de Biologia (EEB), localizada a cerca de 5 km do Campus Darcy Ribeiro da UnB, possui infraestrutura destinada à pesquisa de campo e casa de vegetação, contendo áreas destinadas a estudos fitopatológicos, de fruticultura e melhoramento genético. Na EEB existe uma área de preservação, onde são encontradas espécies endêmicas do Cerrado e exóticas (pinus e eucalipto). Esta área, ainda sem registro no SNUC, possivelmente será enquandrada como uma ARIE (UnB, 2018). Na EEB encontra-se também a Horta Experimental (HE-EEB) com várias espécies dentre fruteiras, hortaliças e grande culturas. Esta horta vem sendo mantida deste 2013 com finalidade de vitrine para estudantes de graduação e pós-graduação de disciplinas ministradas por professores de Fitopatologia da UnB. Não se usa nenhum produto químico na HE - EEB. A distância mais próxima da HE - EEB até a ARIE-EEB é de 30 m.

Áreas nativas de Cerrado do DF foram ocupadas e transformadas em fronteiras agrícolas. As denominadas “áreas rurais dirigidas” previstas no projeto urbanístico deram origem aos Núcleos Rurais, que são responsáveis por produção agrícola expressiva. Merece destaque o Núcleo Rural Taquara responsável pela maior produção de pimentão no DF, e de outras hortaliças como berinjela e tomate (Empresa de Assistência Técnica e Extensão Rural - EMATER, 2018).

4. Patologia Florestal

A patologia florestal é a área da ciência que se encarrega do estudo das doenças em espécies arbóreas. Há um baixo nível geral de doença em condições naturais, com exceções de doenças nativas devastadoras (Hansen and Goheen 2000, Jousimo et al. 2014, Islam et al. 2017).

A ação humana em ecossistemas naturais enfatizou a necessidade de abordagens globais sobre os impactos que o homem causa nos nichos ecológicos e na evolução das espécies, sendo que muitas destas mudanças evolutivas podem ser observadas em curto prazo (Palumbi 2001, Burdon and Thrall 2008, Jousimo et al. 2014, Islam et al. 2017). Este fato enfatizou a necessidade de compreender melhor os processos coevolutivos entre os fitopatógenos e as espécies arbóreas hospedeiras (Parker and Gilbert 2004, Schoettle and Sniezko 2007, Burdon and Thrall 2008, Hendry et al. 2011, Ennos 2014).

A identificação e a caracterização biológica dos agentes causais das doenças que afetam espécies arbóreas são os principais objetivos dos patologistas florestais que visam à produtividade (fungos que afetam a produção e conservação da madeira) e questões de conservação (agentes patogênicos que afetam as árvores em seu ambiente natural) (Peterson and Griffith 1999).

Exemplos bem documentados de epidemias florestais, causadas por fugos, podem ser citados: cancro da castanheira causada por *Cryphonectria parasitica* em castanheira (*Castanea dentata*), doença do olmo (*Ulmus* sp.) holandês causada por *Ophiostoma ulmi* e *Ophiostoma novo-ulmi*, morte súbita do carvalho (*Quecus robur*) por *Phytophthora ramorum*, ferrugem do pinho branco (*Pinus pinaster*) causada por *Cronartium ribicola* e mancha cinzenta em espécies de *Fraxinus* sp. causada por *Hymenoscyphus fraxineus* (Ghelardini et al. 2016).

A frequência da ocorrência de patógenos que levam ao declínio progressivo de árvores e a sua mortalidade se deve à expansão agropecuária, mudanças edafoclimáticas e a introdução de patógenos exóticos (Desprez-Loustau et al. 2011). Além disso, o monocultivo de espécies arbóreas, como o eucalipto que ocupa mais de 20 milhões de ha plantados distribuídos em 90 países (Booth 2013) tem favorecido a ocorrência de doenças, tais como: a ferrugem do eucalipto causada por *Puccinia psidii*, o mofo cinzento do eucalipto causado por *Botrytis cinerea* e a murcha bacteriana em eucalipto causada por *Ralstonia solanacearum* (Wingfield et al. 2012).

4.1. Vírus em espécies florestais e arbóreas: estudos clássicos

A riqueza de espécies arbóreas distribuídas pelo mundo leva os cientistas a levantarem hipóteses das árvores abrigarem uma diversidade de vírus. Os relatos de vírus ocorrendo em espécies arbóreas concentram-se na Europa e América do Norte (Nienhaus and Castello 1989, Büttner et al. 2013, Roossinck et al. 2015, Druciarek et al. 2019) enquanto relatos no Brasil são escassos (Lin et al. 1979, Lin et al. 1980, Lima and Gonçalves 1988, Seabra et al. 2001, Beserra Jr et al. 2011, Nicolini et al. 2012, Fonseca et al. 2018, Borges et al. 2019).

Dentre os estudos clássicos têm-se a revisão de Nienhaus e Castello (1989), no qual os autores compilaram dados de ocorrência de vírus em 46 espécies arbóreas dos EUA e Europa. Neste trabalho, foram relatados 33 vírus (identificação de espécie) e quatro gêneros. Na revisão, mais recente, realizada por Buttner et al. (2013) 32 espécies virais foram detectadas em 17 espécies arbóreas de importância econômica, dentre as quais: *Cherry leaf roll virus* - CLRV (*Nepovirus*), *Apple mosaic virus* - ApMV (*Ilarvirus*), *Arabis mosaic virus* - ArMV (*Nepovirus*), *Tomato mosaic virus* - ToMV (*Tobamovirus*) e *Plum pox virus* - PPV (*Potyvirus*).

Outros relatos de vírus ocorrendo em espécies arbóreas são: *European mountain ash ringspot-associated virus* - EMARaV (*Emaravirus*) disseminado na Europa e detectado em *Sorbus acuparia* (Mielke et al. 2008, Grimová et al. 2015, Druciarek et al. 2019) e a espécie CLRV acometendo *Betula* sp. na Finlândia, Suécia e Noruega (Jalkanen et al. 2007).

As espécies arbustivas também são acometidas por viroses. O ligustro (*Ligustrum japonicum*), um arbusto nativo da China, mas introduzido na Europa e América, pode ser infectado fora do seu ambiente nativo pelas espécies *Prunus necrotic ringspot virus* - PNRSV (*Ilarvirus*), CLRV (*Nepovirus*), *Tomato bushy stunt virus* - TBSV (*Tombusvirus*) e *Tomato black ring virus* - TBRV (*Nepovirus*) (Cooper 1993). Ainda para ligustro, Scott e Zimmerman (2008) relataram *Ligustrum necrotic ringspot virus* - LigRSV (*Carlavirus*).

Borondyko et al. (2007) identificou em *Robinia pseudoacacia*, espécie arbórea nativa dos EUA, três isolados de *Strawberry latent ringspot virus* - SLRSV (gênero *Nepovirus*) usando gama de hospedeiras, sorologia e Reação em Cadeia da Polimerase (PCR).

Informações adicionais sobre ocorrência de vírus em espécies arbóreas referem-se às espécies de importância econômica como citros (*Citrus* sp.) (Matsumura 2016), framboesa (*Rubus idaeus*) (Martin et al. 2013), macieira (*Malus domestica*) (Liang et al. 2015b), pessegueiro (*Prunus persica*) (Marais et al. 2014), oliveira (*Olea europaea*) (Grieco et al. 2000, Choueiri et al. 2015) e cacaueiro (*Theobroma cacao*) (Ploetz 2016).

Vicent et al. (2014) realizaram, na Austrália, um estudo inoculando espécies de plantas com 17 vírus, incluindo *Tomato spotted wilt virus* - TSWV (*Orthotospovirus*), *Bean yellow mosaic virus* - BYMV (gênero *Potyvirus*), *Alfafa mosaic virus* - AMV (*Alfamovirus*) e *Clitoria chlorosis virus* - CLCV (*Potyvirus*). O resultado obtido foi a infecção de 15 espécies botânicas classificadas em oito famílias Araliaceae, Asteraceae, Dilleniaceae, Haemodoraceae, Fabaceae, Malvaceae, Poaceae e Solanaceae, demonstrando a ampla gama de hospedeiras destas espécies virais.

Características intrínsecas aos vírus de plantas, como tamanho diminuto variando na ordem de dezenas a centenas de nanômetros, pequena variação morfológica e parasitismo obrigatório são fatores que dificultam a identificação de vírus. Neste contexto, as técnicas clássicas de detecção viral, como os métodos biológicos, sorológicos e moleculares demandam maior tempo para identificação de novas espécies virais (Lima 2015) e a necessidade de suspeita de ocorrência de infecção viral, segundo a sintomatologia observada.

Com o HTS aliado as análises bioinformáticas, a obtenção de sequências de genomas virais, identificação e caracterização de vírus conhecidos e desconhecidos, em plantas infectadas, estão atualmente entre as aplicações mais bem-sucedidas dessas tecnologias (Barba et al. 2014).

4.2 Vírus em espécies arbóreas no Brasil

Poucos estudos foram realizados de detecção de vírus em espécies arbóreas no Brasil. Até o momento, com base na morfologia das partículas e propriedades biológicas e sorológicas, algumas espécies de vírus foram relatadas infectando espécies de *Senna* (sinonímia *Cassia*), sendo duas delas relatadas nas décadas de 1970 e 1980. Lin et al. (1979) relataram em *Cassia sylvestris* uma possível nova espécie do gênero *Carlavirus*, enquanto Lin et al. (1980) associaram *Cassia mild mosaic virus* - CMMV (*Carlavirus*) com os sintomas de morte súbita em *Cassia macranthera*. Almeida et al. (2002) relataram uma estirpe de *Soybean mosaic virus* - SMV (*Potyvirus)* em *Senna occidentalis*. Anos mais tarde, Nicolini (2012) ao realizar estudos sorológicos e moleculares em *Cassia hoffmannseggii* apresentando sintomas de mosaico, coletada no município de Lava Pratos, situado no estado de Pernambuco, detectou uma nova espécie do gênero *Tymovirus* a qual propôs o nome Cassia yellow mosaic-associated virus (CAYMaV) e *Cowpea aphid borne mosaic virus* - CABMV (*Potyvirus*).

Beserra et al. (2011) detectaram dois novos vírus (uma nova espécie do gênero *Potyvirus* denominada Senna virus Y, e a outra na ordem *Tymovirales*, gênero não definido, denominada Senna virus X) em amostras sintomáticas de *Senna macranthera* em área urbana do município de Viçosa.

Farias (2012) detectou uma espécie de *Begomovirus* em *Mimosa caesalpinifolia* coletada em Cuiabá-MT apresentando sintomas de manchas cloróticas e Batista (2014) ao realizar estudos sorológicos, de vírus importantes na agricultura, em mudas de espécies arbóreas coletadas em dois viveiros do DF e em área de preservação florestal, detectou a presença de espécies do gênero *Orthotospovirus* em cerca de 60% de cinquenta e oito espécies botânicas avaliadas. Outras espécies virais classificadas nos gêneros *Potyvirus*, *Tobamovirus* e *Cucumovirus* foram detectadas em menor número percentual.

Basso et al. (2015) coletaram 74 amostras de árvores frutíferas, nos estados de MG e RS, e encontraram um novo vírus ssDNA circular monopartido. Ao realizarem testes de infectividade os autores confirmaram a capacidade do clone em infectar mudas de macieira e pereira, mas não *Nicotiana benthamiana*. O nome Temperate fruit decay-associated virus (TFDaV) foi proposto por Basso et al. (2015).

Santos (2016) coletou mudas de espécies arbóreas no DF e realizou estudo metagenômico. Após análises bioinformáticas a maior sequência genômica obtida foi de 9.529 nucleotídeos (nt) e apresentou identidade com os vírus da ordem *Picornavirales*. Foi proposto que o genoma encontrado seria membro dessa ordem, denominado Hovenia dulcis associated virus (HDAV).

Fonseca et al. (2018) estudou o viroma de plantas de seringueira (*Hevea brasiliensis*) coletadas na Amazônia. Nas análises bioinformáticas, estes autores obtiveram *contigs* relacionados a família *Tymoviridae* e fizeram a proposta de uma nova espécie relacionada ao gênero *Maculavirus* denominada Hevea brasiliensis virus (HBrV).

Borges et al. (2019) detectaram, por RT-PCR, *Tomato chlorosis virus* - ToCV (*Crinivirus*) em mudas de teca (*Tectona grandis*) coletadas em viveiro do Distrito Federal com sintomas de clorose internerval e com alta infestação de mosca-branca (*Bemisia tabaci*).

Apesar disso, pouco ainda se conhece sobre vírus em espécies arbóreas no Brasil.

4.3 Vírus em espécies florestais e arbóreas: metagenômica

A metagenômica (também referida como genômica ambiental) é a análise genômica da estrutura e função das sequências nucleotídicas isoladas diretamente de uma amostra ambiental, especialmente de uma comunidade de microrganismos. O desenvolvimento da metagenômica decorreu da evidência de que microrganismos não cultivados representam a grande maioria dos organismos na maioria dos ambientes na Terra (Handelsman 2004).

As tecnologias de sequenciamento de nova geração (High-Throughput Sequencing - HTS) vem sendo usadas para descrever várias tecnologias de sequenciamento modernas, dentre as quais: *Illumina*, 454, Pacific Biosciences, IonTorrent, Nanopore. O advento do HTS, aliado a bioinformática e análises metagenômicas revolucionaram os estudos biológicos, inclusive no campo da virologia vegetal (Adams and Fox 2016).

O HTS fornece sequenciamento altamente eficiente e rápido de DNA, ou RNA, de genomas de vírus e viróides de plantas e de pequenos RNAs específicos gerados durante o processo de infecção (miRNAs ou siRNAs). Embora Kleeff et al. (2016) tenham demonstrado que sRNA (small RNA) do vetor *Bemisia tabaci* possa ser encontrado em floema de tomate, não há estudos indicando a influência destes resultados usando HTS para detecção viral a partir de RNAs específicos. HTS tem sido utilizado em vários estudos em virologia de plantas, bem como na detecção e identificação dos patógenos já conhecidos, análise da diversidade e evolução do genoma e estudos da epidemiologia (Hadidi et al. 2016).

As análises bioinformáticas geradas pelo HTS fornecem dados virtuais de sequenciamento do genoma e identificação de novos vírus ou de espécies virais já conhecidas. Entretanto, para detecção viral e inclusão de um novo vírus na taxonomia são necessários ainda detecção por PCR e sequenciamento *Sanger*, além de estudos biológicos e sorológicos (Hadidi et al. 2016).

A aplicação do HTS na área da Virologia Vegetal iteve início em 2009 com os trabalhos de Adams et al. (2009), Al Rwahnih et al. (2009) e Kreuze et al. (2009). Desde então houve aumento significativo no número de novos vírus de plantas descobertos e caracterizados, tanto em plantas hospedeiras como insetos vetores (Kaur et al. 2016). Mais de 100 novos vírus de plantas de DNA e RNA pertencentes a diferentes gêneros e famílias foram relatados nos últimos anos (Hadidi and Barba 2012, Barba et al. 2014, Ho and Tzanetakis 2014, Barba et al. 2015, Roossinck et al. 2015, Wu et al. 2015, Varsani et al. 2017b). HTS vem sendo amplamente

utilizado por muitos grupos de pesquisa e *papers* ilustrando esta abordagem para detecção viral vem sendo publicados (Maree et al. 2018; Massart et al. 2019).

Estimativas de Mokili et al. (2012) apontavam para o conhecimento de menos de 1% da diversidade viral. Desde 2009 a descoberta de novos vírus tem sido realizada em plantas cultivadas, como batata-doce (Kreuze et al. 2009), mandioca (Monger et al. 2010), tomate (Li et al. 2012), citros (Ruiz-Ruiz et al. 2011), milho (Adams et al. 2013), cana-de-açúcar (Candresse et al. 2014), videira (Molenaar et al. 2015), nectarinas (Bag et al. 2015, Villamor et al. 2016), e também em espécies de plantas nativas com a utilização de HTS em abordagens ecogenômicas (Stobbe and Roossinck 2014, Roossinck et al. 2015, Bernardo et al. 2018).

Loconsole et al. (2012), na Itália, identificaram cinco vírus infectando plantas do gênero *Prunus*, *Citrus*, *Vitis*, *Ficus*, *Corylus*, *Diospyros* e *Morus*.

Uma revisão de estudos de vírus empregando HTS e metagenômica em plantas perenes e frutíferas compilaram dados dos últimos cinco anos. Em geral, foram descobertos novos vírus com genomas tão diferenciados que levaram à proposta de criação de novos gêneros ao ICTV (Maliogka et al. 2018).

Recentemente, foram descobertas novas espécies da família *Geminiviridae* infectando amoreira, macieira, videira e citrus (Varsani et al. 2017b).

Roosinck et al. (2010) realizaram análises metagenômicas em plantas provenientes de área de conservação da Costa Rica e identificaram 344 espécies virais, sendo que 70% correspondiam a espécies virais desconhecidas pertencentes às famílias *Bromoviridae*, *Chrysoviridae*, *Totiviridae*, *Endornaviridae*, *Tymoviridae*, *Partitiviridae*, *Potyviridae*, *Closteroviridae*, *Luteoviridae*, *Narnaviridae* e *Caulimoviridae*.

MacDiarmid et al. (2013) detectaram aproximadamente 300 espécies de vírus em plantas nativas de uma pradaria nos EUA (*Tallgrass Prairie Preserve*). Deste total, apenas 18 são conhecidas e classificados na família *Potyviridae*.

Estudos ecogenômicos, utilizando a metodologia de enriquecimento de partículas, foram conduzidos em 2010 e 2012, nas cidades de Camargue e Cabo Ocidental, na França e África do Sul, respectivamente. Foram coletadas plantas nativas sintomáticas e assintomáticas, em agroecossistema (ambiente natural e cultivado próximos - 4,5 Km). Coletivamente, havia presença de vírus em 25% a 58% das amostras analisadas, sendo que a maioria das amostras era assintomática. Das sequências obtidas de amostragem da África do Sul apenas 10% foram identificadas com espécies virais conhecidas (Bernardo et al. 2018).

Nos estudos da relação vírus-hospedeira, para espécies arbóreas, nem sempre há presença de sintomas e alguns vírus podem estar latentes na hospedeira (Hadidi et al. 1998, Hadidi et al. 2003, Hadidi et al. 2011). Isto favorece o livre trânsito destes vírus que podem ser patogênicos a outras plantas e sua infecção resulta na redução do rendimento e no declínio das plantas (Hadidi and Barba 2012, Barba et al. 2015).

Metagenômica de vírus em espécies arbóreas nativas evidenciam a necessidade de explorar esse universo que tem se mostrado vasto, devido à maioria das espécies de vírus de plantas conhecidas terem sido caracterizados a partir de plantas domesticadas (Scheets et al. 2011, Thapa et al. 2012). Além disso, estudos crescentes de vírus em ecossistemas naturais indicam o potencial dos vírus de plantas de transitarem de um ambiente natural para as lavouras (Rodríguez-Nevado et al. 2017, Bernardo et al. 2018, Rodríguez-Nevado et al. 2019).

4.4. Descrição de novos vírus de DNA de fita simples

Os vírus com genomas de ssDNA (single strand DNA) representam um grupo altamente diverso e de importância econômica e ecológica por infectar e serem encontrados em associação com organismos dos três domínios de vida (Krupovic 2013). Muitos vírus novos de ssDNA foram descritos recentemente, incluindo aqueles associados a algas (Nagasaki et al. 2005, Tomaru et al.

2011), fungos (Yu et al. 2010, Yu et al. 2013), insetos (Wang et al. 2007) e arqueobactérias (Pietilä et al. 2009, Mochizuki et al. 2012).

Pesquisadores de grupos do ICTV (2019) classificaram os vírus de ssDNA em treze famílias usando como critérios, o organismo no qual o vírus foi encontrado e a organização genômica. A seguir esta classificação é apresentada e entre parênteses encontram-se informações sobre o organismo ao qual o vírus foi encontrado. Dois grandes grupos foram propostos: **a)** vírus que possuem a Rep (replication-associated protein gene). Encontram-se membros das famílias *Inoviridae* e *Microviridae* (bactérias), *Bacilladnaviridae* (microrganismos eucarióticos), *Circoviridae* e *Parvoviridae* (humanos, outros vertebrados e invertebrados), *Smacoviridae* (humanos, outros vertebrados, invertebrados e bactérias), *Genomoviridae* (humanos, outros vertebrados, invertebrados e fungos), *Geminiviridae* e *Nanoviridae* (plantas) e **b)** vírus que não possuem a Rep: *Anelloviridae* (humanos, outros vertebrados e invertebrados), *Pleolipoviridae* (arqueobactérias), *Spiraviridae* (arqueobactérias) e *Bidnaviridae* (invertebrados), sendo este último, o único que não possui genoma circular. Outros vírus descobertos por metagenômica permanecem sem classificação e são denominados CRESS (proteínas associadas à replicação circular que codificam vírus de DNA de cadeia simples) (Simmonds et al. 2017).

As abordagens em estudos metagenômicos e estudos tradicionais utilizando clonagem e sequenciamentos convencionais permitiram o conhecimento da enorme diversidade genética de vírus, fornecendo informações valiosas sobre sua evolução (Krupovic 2013). Além disso, a utilização da DNA polimerase (fago phi29) num método simples denominado RCA (*Rolling circle amplification*) amplifica genomas circulares extraídos de amostras ambientais, animais e plantas (Rosario et al. 2012b).

O uso da RCA na detecção de vírus de plantas acarretou a descoberta de novas espécies, com destaque para as famílias *Geminiviridae* (Varsani et al. 2017b, Zerbini et al. 2017) e *Genomoviridae* (Krupovic et al. 2016, Varsani and Krupovic 2017).

A grande maioria dos vírus ssDNA replicam seus genomas usando o mecanismo do círculo rolante, que envolve a remoção de uma das cadeias em um intermediário replicativo de dsDNA por uma endonuclease codificada por vírus (Gutierrez 1999).

Os vírus com genomas de ssDNA estão entre os menores vírus conhecidos. Estudos tem indicado que o pequeno tamanho do genoma resulta em altas taxas de substituição de nucleotídeos e de recombinação (Duffy et al. 2008, Sanjuán et al. 2010, Martin et al. 2011), que é amplamente semelhante ao dos vírus de RNA (Holmes 2011).

A seguir serão apresentadas características de vírus ssDNA classificados nas famílias *Geminiviridae, Genomoviridae* e *Circoviridae.* Serão apresentadas também características de *Caulimoviridae* (dsDNA).

4.4.1 Família *Geminiviridae*

A família *Geminiviridae* engloba vírus que infectam plantas, possuem partículas icosaédricas geminadas e genoma composto por um ou dois componentes de ssDNA (2.5-5.2 kb). Vírus dessa família causam doenças em importantes culturas agrícolas e tem distribuição mundial. As plantas infectadas apresentam ampla gama de sintomas, incluindo nanismo; crescimento distorcido; riscas e estrias foliares em plantas monocotiledôneas, enrolamento foliar, ondulação, distorção, mosaico/mosqueado, amarelecimento internerval e clorose em plantas dicotiledôneas (Inoue-Nagata et al. 2016).

Estes vírus costumam existir em complexos de doenças e apresentam altas taxas de recombinação e mutação, permitindo-lhes adaptar-se rapidamente a novos hospedeiros e ambientes (Reyes et al. 2013).

Membros da família codificam uma proteína iniciadora de replicação, a Rep, que se liga de maneira específica a sequências de DNA, os iterons, funcionando como elementos essenciais para a replicação específica de vírus (Argüello-Astorga and Ruiz-Medrano 2001).

A região N-terminal de Rep possui atividades específicas de corte, ligação e ligação a DNA (Fontes et al. 1992; Orozco et al. 1997; Chatterji et al. 2000), enquanto o terminal C (aa 120-361) funciona autonomamente como helicase de 3 'a 5' (Choudhury et al. 2006; Clérot & Bernardi, 2006). As helicases são classificadas em cinco superfamílias (SF1-SF5) onde todas as proteínas possuem função de ATPases associada a várias atividades celulares. A Rep de pequenos vírus de DNA são classificados na superfamília helicase SF3 de acordo com a identidade de sequência (Walker et al. 1982, Gorbalenya et al. 1990).

O N terminal da Rep contém três sequências conservadas chamados motivos I, II e III que são característicos de muitos mecanismos de replicação de círculo rolante (Ilyina and Koonin 1992, Koonin and Ilyina 1992). O motivo I (FLTYP) é necessário para ligação específica de dsDNA (double strand DNA) enquanto o motivo II (HLH) é um local de ligação de metal que pode estar envolvido na conformação de proteínas e clivagem de DNA (Orozco and Hanley-Bowdoin 1998, Gutierrez 1999, Argüello-Astorga and Ruiz-Medrano 2001). O motivo III (YxxKD/E) é um sítio catalítico de clivagem do DNA, com o grupo hidroxila do resíduo Y formando uma ligação covalente com o grupamento 5 fosforil da fita de DNA clivada (Laufs et al. 1995b, Orozco and Hanley-Bowdoin 1996). O domínio GRS contribui para integridade da proteína Rep (Nash et al. 2011). As helicases da família SF3 possuem três motivos de assinatura conservados: Walker A [envolvido na ligação do ATP; GxxxxGK (T/S)], Walker B (envolvido na hidrólise de ATP; DxxD ou xxxxDD) e motivo C (um resíduo de asparagina conservado que interage com o gama Pi do ATP e uma molécula de água 'apical')(George et al. 2014).

Atualmente, a família é representada por nove gêneros *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus* e *Turncurtovirus* e duas novas espécies não classificadas em nenhum gênero *Citrus chlorotic dwarf associated virus* - CCDaV e *Mulberry mosaic dwarf associated virus* - MMDaV (Varsani et al. 2017a). As

relações dos nove gêneros, quantidade de espécies aceitas, tipo de hospedeira, vetor e algumas observações estão presentes na **Tabela 1**.

Os critérios de demarcação de um novo gênero na família *Geminiviridae* incluem: hospedeira (monocotiledôneas ou dicotiledôneas), tipo de vetor (cigarrinhas, soldadinho, moscas brancas, pulgões), organização e identidade pareada do genoma completo (mono ou bipartido) e relações filogenéticas (Zerbini et al. 2017).

Tabela 1. Relação dos gêneros da família *Geminiviridae* considerando o número de espécies por gênero, espécie-tipo, grupo de hospedeira, vetor e nonanucleotídeo.

Gênero (número de espécies)	**Espécie-tipo**	**Hospedeira**	**Vetor**	**Nonanucleotídeo***
Becurtovirus (3)	*Beet curly top Iran virus* - BCTIV	Dicotiledônea	Cigarrinha (*Circulifer haematoceps*)	TAA**G**ATT**C**C
Begomovirus[1] (409)	*Bean golden yellow mosaic virus* - BGYMV	Dicotiledônea	Mosca branca (*Bemisia tabaci*)	TAATATTAC
Capulavirus (4)	*Euphorbia caput-medusae latent virus* - EcmLV	Dicotiledônea	Afídeo (*Aphis craccivora*)	TAATATTAC
Curtovirus (3)	*Beet curly top virus* – BCTV	Dicotiledônea	Cigarrinha (*Circulifer tenellus*)	TAATATTAC
Eragrovirus (1)	*Eragrostis curvula streak virus* – ECSV	Monocotiledônea	Desconhecido	TAA**G**ATT**C**C
Grablovirus (3)	*Grapevine red blotch virus* – GRBD	Dicotiledônea	Desconhecido	TAATATTAC
Mastrevirus[2] (40)	*Maize streak virus* - MSV	Monocotiledônea e Dicotiledônea	Cigarrinha (*Cicadulina* sp.)	TAATATTAC e TAA**G**ATT**C**C
Topocuvirus [3] (1)	*Tomato pseudo-curly top virus* - TPCTV	Dicotiledônea	Membracídeo (*Micrutalis malleifera*)	TAATATTAC
Turncurtovirus (2)	*Turnip curly top virus* - TCTV	Dicotiledônea	Cigarrinha (*Circulifer haematoceps*)	TAATATTAC
Citrus chlorotic dwarf associated virus - CCDaV	-	Dicotiledônea	Desconhecido	TAATATTAC
Mulberry mosaic dwarf associated virus - MMDaV	-	Dicotiledônea	Desconhecido	TAATATTAC

[1] Três classes de satélites de DNA circulares foram descritas associadas a begomovírus: betasatélites, alphasatélites e deltasatélites (Zhou 2013, Lozano et al. 2016). [2]*Wheat dwarf India virus* demonstrou associar-se com os alfasatélites e betasatélites (Kumar et al. 2014). [3]Genoma similar aos curtovírus. Comitê Internacional de Taxonomia de Vírus ICTV (https://talk.ictvonline.org/). Acesso em 21.12.2019. *Diferenças no nonanucleotídeo encontram-se em negrito.

O gênero *Becurtovirus* agrupa três espécies aceitas. Estes vírus possuem genoma monopartido, infectam dicotiledôneas e são transmitidos por cigarrinhas (ICTV, 2019). O nonanucleotídeo é diferenciado das espécie de outros gêneros pois possui a sequência TAA**G**ATT**C**C ao invés de TAATATTAC. O critério para classificação de uma nova espécie nesse gênero é a identidade pareada de 80% com algum membro já estabelecido, e de 94% para uma nova estirpe (Varsani et al. 2014b).

As quatro espécies classificadas em *Capulavirus* possuem genoma monopartido (ICTV, 2019). Isolados de *Alfalfa leaf curl virus* - ALCV são transmitidos por pulgão (Roumagnac et al. 2015). Todos os capulavírus conhecidos têm o nonanucleotídeo TAATATTAC. Atualmente é adotado um limiar de demarcação de espécies de 78% (Varsani et al. 2017b). Assim, quaisquer dois capulavírus cujas sequências genômicas tenham um valor de identidade de nucleotídeos (nt) em pares acima de 78% devem ser considerados isolados da mesma espécie.

O gênero *Curtovirus* inclui três espécies reconhecidas, cujos membros são patógenos economicamente importantes na América do Norte e no Irã (Chen et al. 2011). Os isolados da espécie tipo, *Beet curly top virus* - BCTV apresentam uma ampla gama de hospedeiras incluindo mais de 300 espécies de plantas dicotiledôneas em 44 famílias de plantas (Strausbaugh et al. 2008). Estes vírus são monopartidos e transmitidos por cigarrinhas. A identidade pareada de sequência nucleotídica do genoma completo menor que 77% é o limiar para demarcação de espécie no gênero *Curtovirus* (Varsani et al. 2014a).

A espécie *Eragrostis curvula streak virus* - ECSV é a única representante do gênero *Eragrovirus* (ICTV, 2019). Todos os isolados conhecidos foram encontrados infectando *Eragrostis curvula* na região de KwaZulu-Natal na África do Sul e são monopartidos (Varsani et al. 2009). Como os becurtovírus, os eragrovírus têm a sequência nonanucleotídica TAAGATTCC. O limiar de demarcação de espécie é a identidade pareada de sequência do

genoma completo menor que 94% para ser considerado membros/variantes da mesma estirpe (Varsani et al. 2014b).

O gênero *Grablovirus* agrupa três espécies que infectam videira (*Vitis vinifera*) e pessegueiro (*Prunus persica*) (ICTV, 2019). Isolados de *Grapevine red blotch virus* - GRBV são transmitidos pelo membracídeo *Spissistilus festinus* (Hemiptera: Membracidae) (Bahder et al. 2016). Os isolados encontrados que compartilham menos que 80% de identidade pareada do genoma completo com membros de espécies de grablovírus devem ser classificados como membros de uma espécie distinta (Varsani et al. 2017b).

A espécie *Tomato pseudo-curly top virus* - TPCTV é a única representante do gênero *Topocuvirus*. Isolados dessa espécie são transmitidos por um membracídeo (*Micrutallis malleifera*) (Briddon et al. 1996). Não existe ainda um limiar para demarcação de espécie (ICTV, 2019).

O gênero *Turncurtovirus* possui duas espécies. Estes vírus são monopartidos, transmitidos por cigarrinha e encontrados em dicotiledôneas (Razavinejad and Heydarnejad 2013, Kamali et al. 2016). O critério de demarcação de espécie é a identidade pareada da sequência genômica completa apresentando valor abaixo de 80% com outros membros e abaixo de 95% serão considerados estirpes (Varsani et al. 2014b).

Os membros do gênero *Mastrevirus* são transmitidos por cigarrinhas (Webb 1987, Gutierrez 1999) e possuem genomas monopartidos (2.5-2.7 kb) que codificam quatro proteínas separadas por duas regiões intergênicas [grande região intergênica (LIR) e pequena região intergênica (SIR)]. As duas proteínas codificadas no sentido viral são a proteína do movimento (MP) e a proteína capsidial (CP) e no sentido complementar são codificadas as proteínas Rep e Rep A associadas à replicação (Wright et al. 1997). A Rep é expressa através de um mecanismo de *splicing* de transcrição dos genes para C1:C2 [3], enquanto a Rep A é expressa através de uma transcrição que abrange a ORF C1(Wright et al. 1997).

O gênero *Mastrevirus* é o segundo maior gênero da família *Geminiviridae,* com 40 espécies aceitas (Zerbini et al. 2017) conhecidas por infectarem espécies monocotiledôneas e algumas poucas espécies infectam dicotiledôneas em associação com moléculas de satélite ssDNA (Kumar et al. 2014, Hamza et al. 2018), com aproximadamente metade do tamanho do genoma do vírus.

Na América do Sul há relatos de isolados de mastrevírus em batata (Peru e Uruguai) (Kreuze et al. 2009, Cao et al. 2017). Recentemente, no Brasil foi detectada em cigarrinha (*Dalbulus maidis*) e milho (*Zea mays*) uma nova espécie do gênero apresentando identidade de 63% com outros isolados de mastrevírus, denominada “Maize striate mosaic virus - MSMV” (Fontenele et al. 2018).

O gênero *Begomovirus* inclui 409 espécies aceitas, cujos isolados infectam plantas dicotiledôneas e são transmitidas por moscas-brancas (*Bemisia tabaci*). O genoma pode ser bipartido (duas moléculas de ssDNA circulares, cada uma com cerca de 2,6 kb) ou monopartido (uma única molécula de ssDNA circular de aproximadamente 2,7 kb) (Brown et al. 2015). Algumas espécies ocorrem em associação com ssDNA circulares conhecidos como satélites.

Três classes de satélites circulares de DNA foram descritas associadas aos begomovírus: betasatélites, alfasatélites e deltasatélites (Zhou 2013, Lozano et al. 2016). Os begomovírus monopartidos são frequentemente associados a betasatélites ou alfasatélites que possuem, geralmente, metade do tamanho do vírus. Os betasatélites estão associados à indução de sintomas, na supressão do silenciamento gênico transcricional e pós-transcricional e podem afetar os genes associados ao ácido jasmônico. Os alfasatélites foram identificados principalmente em begomovírus monopartidos que associam-se a betasatélites e, não possuem contribuições conhecidas para a patogênese em complexos da doença begomovírus-betasatélites. Os deltasatélites foram propostos por Lozano et al. (2016) agrupando satélites associados a begomovírus encontrados em malváceas em Cuba, moscas brancas na Flórida, plantas do gênero

Ipomoea na Espanha e associados aos vírus *Tomato leaf curl virus* - ToLCV e *Sweet potato leaf curl virus* - SPLCV. Alguns estudos sobre o papel dos deltassatélites vem sendo conduzidos (Fiallo-Olivé et al. 2012, Hassan et al. 2016).

Isolados de mais de 60 espécies do gênero *Begomovirus* são responsáveis por causar, somente no Brasil, doenças em tomate (*Solanum lycopersicum*), pimentão (*Capsicum annum*), caupi (*Vigna unguiculata*), feijão (*Phaseolus vulgaris*), algodão (*Gossypium hirsutum*), mandioca (*Manihot esculenta*), fava (*Phaseolus lunatus*) e batata doce (*Ipomoea batatas*) (Faria et al. 2000, Silva 2006, Inoue-Nagata et al. 2016). Além disto estes vírus infectam espécies de *Macroptilium*, *Datura*, *Calopogonium*, *Chenopodium* e *Desmodium* que desempenham um papel epidemiológico importante, servindo como reservatórios de vários vírus. Estas plantas podem apresentar infecções mistas, portanto, novos vírus recombinantes podem surgir. Vários estudos demonstraram que a recombinação é um dos principais mecanismos geradores de variabilidade genética em begomovírus no mundo (García-Andrés et al. 2007, García-Andrés et al. 2006, García-Andrés et al. 2007, Monci et al. 2002, Pita et al. 2001) e no Brasil (Galvão et al. 2003, Inoue-Nagata et al. 2006, Ribeiro et al. 2007, Silva et al. 2012).

Estudos de metagenômica evidenciam que plantas não cultivadas servem como fonte de inóculo, abrigando alta diversidade de begomovírus, e contribuem para a evolução viral. Notavelmente, na variação genética do vírus é mantido um grau de heterogeneidade na sequência que seja conveniente para adaptação na natureza, entretanto, sem alterar as sequências essenciais (Rodríguez-Negrete et al. 2019).

A ampla disseminação de begomovírus pelo mundo e a vasta gama de hospedeiras estão associados ao seu vetor, mosca branca, responsável pela colonização de mais de mil espécies de plantas distribuídas em cerca de 74 famílias botânicas (Abd-Rabou et al. 2010, Navas-Castillo et al. 2011). A mosca branca facilita o trânsito entre hospedeiros cultivados e não cultivados, contribuindo para a permanência, evolução e epidemiologia do vírus. No Brasil, o aumento da

incidência e severidade de doenças causadas por espécies do gênero *Begomovirus* se deve a introdução do biótipo B da mosca branca no país (Faria, 2000). Uma nova classificação foi proposta para os biótipos de moscas brancas. De acordo com a subdivisão proposta por Dinsdale et al. (2010), usando as sequências de nucleotídeos do gene que codifica a citocromo oxidase mitocondrial, o biótipo B (predominante no Brasil) está classificado como: *Middle East-Asia Minor* 1 (MEAM1) (Barbosa et al. 2014). Além disto o *Mediterranean* (MED= biótipo Q) já está presente no Brasil (Barbosa et al. 2015). De acordo com Moraes et al. (2018) MEAM 1 continua prevalecendo no país e MED, já foi relatado em cinco estados das regiões Sul e Sudeste, entretanto sempre associado a plantas ornamentais.

A mosca branca já foi relatada no Brasil infestando várias culturas como repolho (*Brassica oleracea* var. capitata), algodão, soja, feijão, tomate (Villas Bôas et al. 2001; Lourenção & Nagai, 1994) e ainda é capaz de transmitir vírus dos gêneros: *Begomovirus*, *Carlavirus*, *Ipomovirus*, *Crinivirus* e *Torradovirus* (Barbosa, 2012). O tipo de relação vírus-vetor para mosca branca-begomovírus é do tipo circulativa não propagativa (Brown, 1997; Rubinstein & Czosnek, 1997; Ghanim et al. 1998; Morin et al. 1999).

No Brasil, 25 espécies de *Begomovirus* já foram relatados infectando tomateiro em todo o país (Ambrozevicius et al. 2002, Galvao et al. 2003, Ribeiro et al. 2003, Zerbini et al. 2005, Andrade et al. 2006, Ribeiro 2006, Calegario et al. 2007, Fernandes et al. 2008, Chaves et al. 2017). Dentre estas, cita-se as espécies: *Tomato severe rugose virus* - ToSRV, *Sida micrantha mosaic virus* - SiMMV e *Tomato mottle leaf curl virus* (ToMoLCV).

O componente DNA A pode replicar autonomamente e produzir partículas virais, mas requer o componente DNA B para movimento intra e intercelular na planta (Brown et al. 2015). De acordo com Galvao et al. (2003), entretanto, o DNA A sozinho de *Tomato chlorotic mottle virus* - ToCMoV é capaz de estabelecer infecção sistêmica em plantas de *Nicotiana benthamiana*. Os dois componentes genômicos de *Begomovirus* apresentam identidade de sequência na região

comum (RC) de aproximadamente 200 - 250 pares de bases. Esta região é altamente conservada para os dois componentes de uma espécie viral e contém a sequência necessária para o início da replicação viral (Timmermans et al. 1994, Orozco and Hanley-Bowdoin 1996, Faria et al. 2000, Argüello-Astorga and Ruiz-Medrano 2001).

Em espécies de genoma monopartido, toda a informação genética para replicação e movimento está presente em apenas um componente (Lazarowitz and Shepherd 1992). A replicação do DNA viral ocorre no núcleo da célula pelo mecanismo de amplificação em círculo rolante (RCA) através da formação de um DNA de fita dupla chamado de forma replicativa (RF) que em seguida, será o molde para a transcrição viral e para a síntese de novas fitas de ssDNA (Gutierrez 1999, Alberter et al. 2005). O DNA A, pode codificar de quatro a sete proteínas. As ORFs (open read frame - fase aberta de leitura) codificadas no sentido viral são: AV1 (capa protéica - CP), AV2 (proteína de movimento - MP) (Höfer et al. 1997). No sentido complementar são encontradas: AC1 (proteína associada à replicação-Rep), AC2 (proteína transativadora-TrAp), AC3 (proteína potencializadora da replicação viral- Ren) (Settlage et al. 1996), AC4 importante na manifestação de sintomas e, pode ainda suprimir a resposta do hospedeiro à expressão da Rep (Vanitharani et al. 2004) e AC5 (proteína determinante de patogenicidade). A ORF AC5 foi anotada em algumas espécies, incluindo Mungbean yellow mosaic India virus. A proteína codificada por AC5 está envolvida na supressão do silenciamento gênico (Li et al. 2015). Nas espécies bipartidas classificadas no gênero *Begomovirus* do Novo Mundo a ORF AV2 não está presente (Höfer et al. 1997). A AV1 codifica para proteína capsidial (CP), com funções de proteção da informação genética viral, direcionamento nuclear, exportação nuclear de DNA transmissão e especificidade do vírus pelo vetor *Bemisia tabaci* (Briddon et al. 1990, Azzam et al. 1994, Höfer et al. 1997, Harrison et al. 2002). A ORF AC2 codifica a proteína transativadora da transcrição (TrAP) influenciando a transcrição e subsequente expressão dos genes CP e NSP no sentido viral dos componentes A e B respectivamente (Pratap et al. 2011). A ORF AC3

codifica a proteína Ren de aproximadamente 16kDa, que é requerida para replicação eficiente do DNA viral, sendo considerada como um fator potencializador da replicação possivelmente por interagir com a Rep no reconhecimento da origem de replicação (Settlage et al. 1996). O produto da ORF AC4 estimula a proliferação celular (Rojas et al. 2005). No componente DNA B as ORFs BV1 (sentido viral) e BC1 (sentido complementar) codificam proteínas relacionadas ao movimento viral na planta. A BV1 codifica a proteína NSP "nuclear shuttle protein" encarregada de transportar moléculas de ssDNA e dsDNA (DNA fita dupla) viral do núcleo até o citoplasma e A BC1 codifica a proteína de movimento MP encarregada do transportar de ssDNA ou dsDNA célula a célula via plasmodesma (Noueiry et al. 1994, Sanderfoot and Lazarowitz 1996, Ward and Lazarowitz 1999, Frischmuth et al. 2007). O movimento de espécies bipartidas no interior de uma planta é mediado pelas proteínas codificadas pelos genes presentes no DNA-B. Por outro lado, nas espécies monopartidas a CP atua na dispersão viral também através de interações com os plasmodesmas (Gafni and Epel 2002).

4.4.2 Família *Genomoviridae*

A família *Genomoviridae* engloba vírus de genoma ssDNA circulares pequenos (~ 2-2,4 kb) e que possuem dois genes que codificam duas proteínas bidirecionalmente: a de iniciação à replicação de círculo rolante (Rep) no sentido complemental e a proteína do capsidial (CP) no sentido viral (Krupovic et al. 2016). Esses vírus se caracterizam também por não apresentarem uma proteína de movimento (MP) (Krupovic et al. 2016).

Os genomovírus estão distribuídos em diversos ambientes e encontram-se associados a uma ampla gama de organismos. Atualmente, existem nove gêneros estabelecidos na família *Genomoviridae*: *Gemycircularvirus*, *Gemyduguivirus*, *Gemygorvirus*, *Gemykibivirus*, *Gemykolovirus*, *Gemykrogvirus*, *Gemykroznavirus*, *Gemytondvirus* e *Gemyvongvirus* (Varsani and Krupovic 2017).

Os critérios taxonômicos para classificação de novas espécies dentro da família *Genomoviridae* é que o novo vírus apresente 78% de identidade pareada com alguma espécie já estabelecida (Varsani and Krupovic 2017). Para cada um dos noves gêneros têm-se um critério para demarcação de espécie (% de identidade pareada) e a espécie tipo presentes na **Tabela 2**, incluindo na tabela a quantidade de espécies já aceitas pelo ICTV (2019) e organismos associados.

Tabela 2. Relação dos gêneros da família *Genomoviridae* considerando o número de espécies, valor de identidade pareada em porcentagem (%) para demarcação de espécie para gênero, espécie-tipo e organismo associado.

Gênero	Valor de identidade pareada[1] (%)*	Espécie tipo	N° de espécies aceitas ICTV*	Organismos associados
Gemycircularvirus	44	*Sclerotinia gemycircularvirus 1*	43	Fungos, plantas, animais
Gemyduguivirus	57-62	*Dragonfly associated gemyduguivirus 1*	1	Insetos
Gemygorvirus	49	*Starling associated gemygorvirus 1*	5	Aves
Gemykibivirus	43	*Dragonfly associated gemykibivirus 1*	16	Insetos
Gemykolovirus	37	*Pteropus associated gemykolovirus 1*	2	Mamíferos
Gemykrogvirus	33	*Bovine associated gemykrogvirus 1*	3	Mamíferos
Gemykroznavirus	56-61	*Rabbit associated gemykroznavirus 1*	1	Mamíferos
Gemytondvirus	53-61	*Ostrich associated gemytondvirus 1*	1	Aves
Gemyvongvirus	56-62	*Human associated gemyvongvirus 1*	1	Mamíferos

[1] Valor de identidade pareada refere-se ao critério de demarcação de espécie dentro do gênero (Varsani & Krupovic, 2017). Comitê Internacional de Taxonomia de Vírus - ICTV (https://talk.ictvonline.org/). *Número de espécies aceitas. Acesso em 21.12.2019.

Representantes da família *Genomoviridae* apresentam sequências altamente divergentes. Nas análises das proteínas a CP é mais divergente que a Rep (Varsani and Krupovic 2017).

Os nonanucleotídeos são variáveis ('TAWWDWRN'). Na região N terminal contém motivos importantes para iniciar a replicação por círculo rolante (RCR). Alguns desses motivos são conservados em ssDNA, fagos e plasmídeos que se replicam usando o mecanismo RCR (Ilyina and Koonin 1992, Vega-Rocha et al. 2007, Rosario et al. 2012a, Krupovic 2013). A presença de um único resíduo catalítico de tirosina no motivo III classifica os representantes de genomovírus, geminivírus, bacilladnavírus, circovírus e nanovírus como membros da superfamília II (Ilyina and Koonin 1992, Krupovic 2013).

Membros da família *Genomoviridae* apresentam a sequência conservada do motivo I, que pode estar envolvida no reconhecimento de sequências de iterons associadas à origem da replicação. Esta sequência é predominantemente (uuTYxQ) (u indica resíduos hidrofóbicos e x qualquer resíduo) com exceção dos representantes dos gemikolovírus atualmente conhecidos e gemykrogvírus (Varsani and Krupovic 2017). O motivo II dos membros de *Genomoviridae* (xHxHx), semelhante ao encontrado para os representantes de *Geminiviridae*, e em trabalhos iniciais mostraram que as histidinas nesse motivo coordenam íons metálicos divalentes, Mg2þ ou Mn2þ, que são co-fatores importantes para a atividade da endonuclease na origem da replicação (Koonin 1993, Laufs et al. 1995a). O motivo III (YxxK) baseado em outros estudos da Rep, pode estar envolvido na clivagem do dsDNA e subsequente ligação covalente da Rep através do resíduo de tirosina catalítica na

extremidadedo do produto clivado (Laufs et al. 1995a, Orozco and Hanley-Bowdoin 1998, Timchenko et al. 1999, Steinfeldt et al. 2006, Rosario et al. 2012a). O resíduo lisina no motivo III é proposto para mediar a ligação e o posicionamento durante a catálise (Vega-Rocha et al. 2007). Um quarto motivo conservado (GRS), é encontrado apenas em geminivírus e genomovírus (Nash et al. 2011).

A Rep é uma proteína multifuncional, com atividade de endonucleases e de helicase. A atividade helicase da Rep é mediada por motivos conservados conhecidos como Walker A (GxxxxGKT), Walker B (uuDDu) e motivo C (uxxN) localizados em um domínio de ligação a NTP do terminal C (Gorbalenya et al. 1990, Koonin 1993, Choudhury et al. 2006, Clérot and Bernardi 2006).

No Brasil já foram descritas espécies da família *Genomoviridae* associadas à feijão (Lamas et al. 2016), *Momordica charantia* e *Euphorbia heterophylla* (de Rezende et al. 2018).

4.4.3. Família *Circoviridae*

A família *Circoviridae* engloba vírus que infectam animais e causam doenças que geram perdas como: *Porcine circovirus 2* (PCV2) (Ciacci-Zanella et al. 2006). São vírus com menor tamanho que infectam animais de 1.8 a 2.1 kb, monopartidos. Apesar destes vírus terem sido descritos infectam animais, já se tem relatos de vírus de animal replicando em plantas, no caso o Providence virus um patógeno de inseto que causou infecção em feijão caupi (*Vigna unguiculata*) (Jiwaji et al. 2019).

Os membros desta família são classificados em dois gêneros, *Circovirus* e *Cyclovirus* e ambos possuem apenas duas ORFs, a Rep e CP. As diferenças entre os gêneros são a posição da origem da replicação em relação às regiões codificantes nos circovírus é a Rep e nos cyclovírus a CP e o comprimento das regiões intergênicas (Rosario et al. 2012b). Os circovírus possuem duas regiões intergênicas entre as ORFs, e os cyclovírus não possuem ou é menor que a dos circovírus. Além disso, já foram observados *splicings* em genomas de cyclovírus enquanto não há relatos em circovírus. Os membros dos gêneros possuem uma estrutura stem-loop com nonanucleotídeo conservado [(T/ n)A(/ t)TATTAC] (Mankertz et al. 1997, Rosario et al. 2017).

Os membros do gênero *Circovirus* foram identificados em vertebrados, enquanto os membros do gênero *Cyclovirus* foram identificados em vertebrados e invertebrados (Rosario et al. 2012b). A espécie-tipo para o gênero *Circovirus* é o *Porcine circovirus 1* e a espécie-tipo para o gênero *Cyclovirus* é *Human-associated cyclovirus 8* (Breitbart et al. 2017).

Para cada gênero, o limiar de demarcação de espécies é de 80% de identidade de sequência de nucleotídeos considerando todo genoma (Breitbart et al. 2017). Os circovírus possuem partículas de simetria icosaédrica enquanto que não há estudos sobre as partículas dos cyclovírus (Breitbart et al. 2017), o primeiro possui 39 espécies enquanto o segundo 48 espécies (ICTV, 2019).

4.5. Família *Caulimoviridae*

A família *Caulimoviridae* é a única família contendo vírus de DNA de fita dupla que infectam plantas e replicam por transcrição reversa (RT) (Harper et al. 2002). Os membros desta família têm genomas circulares de 7.000 a 9.200 pb com descontinuidades em ambas as cadeias que codificam de 1 a 8 ORFs (Lefkowitz et al. 2017). O ciclo de replicação é episomal e não requerem integração no genoma da hospedeira, no entanto, a integração pode ocorrer durante o reparo não-homólogo da junção final de dsDNA nos genomas do hospedeiro, deixando um impressão digital de infecções passadas (Geering et al. 2014).

A família *Caulimoviridae* está representada por oito gêneros que se distinguem principalmente com base na morfologia das partículas e organização do genoma (King et al. 2011, Geering et al. 2012). Seis destes gêneros, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Rosadnavirus*, *Soymovirus* e *Solendovirus* possuem partículas isométricas com 52 nm de diâmetro, enquanto dois gêneros, *Badnavirus* e *Tungrovirus*, possuem vírus com formato de partícula alongado e rígidos (130 a 150 nm x 30 nm) (Hull 1996, King et al. 2011, Geering et al. 2012). Isolados de espécies classificadas nos gêneros *Badnavirus* e *Caulimovirus* são os mais importantes. Estes dois gêneros da família *Caulimoviridae* possuem o maior número de espécies aceitas pelo ICTV (2019), 57 (*Badnavirus*) e 13 (*Caulimovirus*), respectivamente. Os critérios para demarcação de espécies no gênero *Caulimovirus* e *Badnavirus* são: hospedeiras, sequência nucleotídica da polimerase (RT + RNAse H) superiores a 20%, sequências de produtos gênicos e vetor (somente para *Badnavirus*).

Membros do gênero *Caulimovirus* podem apresentar de seis a sete ORFs que codificam proteínas: de movimento, fator de transmissão pelo pulgão, proteína de ligação ao DNA, capa

protéica, poliproteína da polimerase (protease, RT e ribonuclease H) e proteína transativadora/ viroplasmina (Guilley et al. 1982, Richins et al. 1987).

A espécie-tipo deste gênero é o *Cauliflower mosaic virus* - CaMV. Segundo Scholthof et al. (2011) o CaMV ocupou o sexto lugar entre os dez vírus de plantas mais importantes em termos científicos e econômicos. Essa notoriedade se deve exclusivamente aos principais avanços conceituais que foram feitos na virologia de plantas usando o CaMV como modelo. O genoma do CaMV possui sete ORFs, que codificam seis proteínas virais bem caracterizadas (P1-P6): P1 (40 kDa), proteína de movimento de célula a célula; P2 (18 kDa), proteína associada à transmissão por pulgões; P3 (15 kDa), proteína com duplo papel no movimento de célula a célula e na transmissão por pulgões; P4 (56 kDa), proteína do capsídeo; P5 (78 kDa), precursor poliproteico de proteinase, transcriptase reversa e ribonuclease H; e P6 (62 kDa), proteína multifuncional, o principal componente dos corpos citoplasmáticos de inclusão (viroplasma ou fábricas virais) associados ao movimento intracelular do vírus, desenvolvimento de sintomas e defesas do hospedeiro (Franck et al. 1980). Como o CaMV se replica pela transcrição reversa de um intermediário de RNA, também é classificado no supergrupo pararetrovírus. Dois principais RNAs transcritos (RNA pré-genômico 35S e subgenômico 19S RNA) são transcritos a partir do genoma do DNA (Fütterer et al. 1989).

No Brasil, o CaMV já foi descrito infectando vegetais como brócolis (*Brassica oleraceae* var. *italica*), couve-flor (*Brassica oleracea* var. *botrytis*), couve (*Brassica oleracea* var. *acephala*), canola (*B. napus*), repolho chinês (*B. rapa* ssp. *Pekinensis*), goivo

(*Matthiola incana*), agrião (*Nasturtium officinale*), mostarda branca (*Sinapsis alba*) (compiladas por Kitajima 2015) e rabanete *Raphanus raphanistrum* (Rodrigues et al. 2019).

A maioria dos membros do gênero *Badnavirus* codificam em seu genoma três ORFs conservadas (ORF I, ORF II e ORF III) (Bhat et al. 2016). A função da proteína P1 e P2 são desconhecidas, enquanto P3 é uma poliproteína que consiste na: proteína de movimento capa protéica, aspártica- protease e domínios RT / RNase H1 nessa ordem.

Espécies classificadas em *Badnavirus* já foram descritas infectando abacaxi (*Ananas comosus*), banana (*Musa* spp.), cacau (*Theobroma cacao*), citrus (*Citrus* spp.), inhame (*Dioscorea* spp.), taro (*Colocasia esculenta*), pimenta do reino (*Piper nigrum*), espécies do gênero *Capsicum* e a espécie arbórea betula (*Betula* spp) (Bhat et al. 2003; Yang et al. 2003; Eni et al. , 2008; James et al. 2011; Johnson et al. 2012; Kouakou et al. 2012; Deeshma & Bhat, 2015; Silva et al. 2015; Deeshma & Bhat, 2017; Rumbou et al. 2018; Xu et al. 2019).

Ao todo, 57 espécies já foram aceitas pelo ICTV, dentre as quais *Sugarcane baciliforme virus* (SBCV) que já foi usada como promotor em ensaios de expressão transiente em monocodiledôneas e dicotiledôneas (Schenk et al. 1999).

Referências Bibliográficas

Abd-Rabou S, Simmons AM (2010) Survey of reproductive host plants of Bemisia tabaci (Hemiptera: Aleyrodidae) in Egypt, including new host records. Entomological News 121:456–465.

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, Miano DW, Kinyua ZM, et al (2013) Use of next-generation sequencing for the identification and characterization of Maize chlorotic mottle virus and Sugarcane mosaic virus causing maize lethal necrosis in Kenya. Plant Pathology 62:741–749.

Adams MJ, Lefkowitz EJ, King AMQ, et al (2017) Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). Archives of Virology 162:2505–2538

Adorno T, Horkheimer M (2007) Scholar (10). Mass communication. Soc. 349–83.

Alberter B, Rezaian MA, Jeske H (2005) Replicative intermediates of *Tomato leaf curl virus* and its satellite DNAs. Virology 331:441–448.

Ambrozevicius LP, Calegario RF, Fontes EPB, et al (2002) Genetic diversity of begomovirus infecting tomato and associated weeds in Southeastern Brazil. Fitopatologia brasileira 27:372–377.

Andrade EC, Manhani GG, Alfenas PF, et al (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.

Argüello-Astorga GR, 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.

Bag S, Al Rwahnih M, Li A, et al (2015) Detection of a new luteovirus in imported nectarine trees: a case study to propose adoption of metagenomics in post-entry quarantine. Phytopathology 105:840–846.

Bahder BW, Zalom FG, Jayanth M, Sudarshana MR (2016) Phylogeny of geminivirus coat protein sequences and digital PCR aid in identifying Spissistilus festinus as a vector of grapevine red blotch-associated virus. Phytopathology 106:1223–1230.

Bale JS, Van Lenteren JC, Bigler F (2007) Biological control and sustainable food production. Philosophical Transactions of the Royal Society Biological 363:761–776.

Barba M, Czosnek H, Hadidi A (2014) Historical perspective, development and applications of next-generation sequencing in plant virology. Viruses 6:106–36. doi: 10.3390/v6010106.

Barba M, Ilardi V, Pasquini G (2015) Control of pome and stone fruit virus diseases. In: Advances in Virus Researchearch. Elsevier, pp 47–83

Barbosa L da F, Marubayashi JM, De Marchi BR, et al (2014) Indigenous American species of the Bemisia tabaci complex are still widespread in the Americas. Pest Management Science 70:1440–1445.

Barbosa L da F, Yuki VA, Marubayashi JM, et al (2015) First report of Bemisia tabaci Mediterranean (Q biotype) species in Brazil. Pest Management Science 71:501–504.

Basso MF, da Silva JCF, Fajardo TVM, et al (2015) A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. Virus Researchearch 210:27–33.

Bernardo P, Charles-Dominique T, Barakat M, et al (2018a) Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. EMBO J 12:173.

Beserra Jr. JE a., Carvalho MG De, Barguil BM, Zerbini FM (2011) Partial genome sequence of a *Potyvirus* and of a virus in the order *Tymovirales* found in *Senna macranthera* in Brazil. Tropical Plant Pathology 36:116–120. doi: 10.1590/S1982-56762011000200008.

Bhat AI, Devasahayam S, Sarma YR, Pant RP (2003) Association of a badnavirus in black pepper (*Piper nigrum* L.) transmitted by mealybug (*Ferrisia virgata*) in India. Current Science 1547–1550.

Bôas GLV, França FH, Macedo N, Moita AW (2001) Avaliação da preferência de *Bemisia argentifolii* por diferentes espécies de plantas1. Horticultura Brasileira.

Booth TH (2013) Eucalypt plantations and climate change. Forest Ecology and Management 301:28–34.

Borges RCF, Fontes MG, Macedo MA de, et al (2019) First Report of Tomato Chlorosis Virus Infecting *Tectona grandis* Associated With Infestation of *Bemisia tabaci* Mediterranean in Central Brazil. Plant Disease 103:2704.

Borodynko N, Hasiow B, Figlerowicz M, Pospieszny H (2007) Identification of the new strain of *Strawberry latent ringspot virus* isolated from black locust (*Robinia pseudoacacia* L.). Journal of Phytopathology 155:738–742.

Breitbart M, Delwart E, Rosario K, et al (2017) ICTV virus taxonomy profile: Circoviridae. Journal of General Virology 98:1997.

Briddon RW, Bedford I, Tsai JH, Markham PG (1996) Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. Virology 219:387–394.

Briddon RW, Pinner MS, Stanley J, Markham PG (1990) Geminivirus coat protein gene replacement alters insect specificity. Virology 177:85–94.

Brown JK (1997) The biology and molecular epidemiology of the *Geminiviridae* subgroup III. In: Plant-microbe Interactions 2. Springer, pp 125–195.

Brown JK, Zerbini FM, Navas-Castillo J, et al (2015) Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. Archives of Virology 160:1593–1619.

Burdon JJ, Thrall PH (2008) Pathogen evolution across the agro-ecological interface: implications for disease management. Evolutionary Applications 1:57–65.

Büttner C, von Bargen S, Bandte M, Mühlbach H-P (2013) Forest diseases caused by viruses. In: Infectious Forest Diseases. CABI: Wallingford, Oxfordshire, UK, pp 50–75.

Calegario RF, Ferreira S de S, Andrade EC de, Zerbini FM (2007) Characterization of *Tomato yellow spot virus*, a novel tomato-infecting begomovirus in Brazil. Pesquisa Agropecuária Brasileira 42:1335–1343.

Candresse T, Filloux D, Muhire B, et al (2014) Appearances can be deceptive: revealing a hidden viral infection with deep sequencing in a plant quarantine context. PLoS One 9:e102945.

Cao M, Lan P, Li F, et al (2017) Genome characterization of sweet potato symptomless virus 1: a mastrevirus with an unusual nonanucleotide sequence. Archives of Virology 162:2881–2884.

Chaves ALR, Eiras M, Rodrigues LK, et al (2017) First report of *Tomato mottle leaf curl virus* infecting tomato in São Paulo State, Brazil. Summa Phytopathol 43:353.

Chen L-F, Vivoda E, Gilbertson RL (2011) Genetic diversity in curtoviruses: a highly divergent strain of *Beet mild curly top virus* associated with an outbreak of curly top disease in pepper in Mexico. Archives of Virology 156:547–555.

Choudhury NR, Malik PS, Singh DK, et al (2006) The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. Nucleic Acids Research 34:6362–6377.

Choueiri E, Freiji A, Kubaa RA, Elbeaino T (2015) First report of *Cucumber mosaic virus* and *Olive latent ringspot virus* on olive trees in Lebanon. Journal of Plant Pathology 97.

Ciacci-Zanella J, Mores N, Simon N, et al (2006) Identificação do circovírus suíno tipo 2 por reação em cadeia da polimerase e por imunoistoquímica em tecidos suínos arquivados desde

1988 no Brasil. Ciência Rural 36:1480–1485.

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.

Cooper JI (1993) Virus diseases of trees and shrubs. Springer Science & Business Media
Crowther TW, Glick HB, Covey KR, et al (2015) Mapping tree density at a global scale. Nature 525:201.

Daily GC, Matson PA (2008) Ecosystem services: From theory to implementation. Proceedings of the National Academy of Sciences 105:9455–9456.

das Dores, V. M., Estevão, C. I. M., & Rocha, S. A. (2017). Análise da percepção ambiental dos alunos da unifimes sobre o conceito de unidades de conservação da biodiversidade um estudo de caso. In Anais Colóquio Estadual de Pesquisa Multidisciplinar (ISSN-2527-2500) & Congresso Nacional de Pesquisa Multidisciplinar.

de Faria JC, Zerbini FM (2000) Família *Geminiviridae*-taxonomia, replicação e movimento. Embrapa Arroz e Feijão-Artigo em periódico indexado.

De Groot RS, Wilson MA, Boumans RMJ (2002) A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecological Economics 41:393–408.

de Rezende RR, Mar TB, Páez LMC, et al (2018) Complete genome sequences of two gemycircularviruses associated with non-cultivated plants in Brazil. Archives of Virology 1–4.

Deeshma KP, Bhat AI (2017) Occurrence of endogenous *Piper yellow mottle virus* in black pepper. Virus Disease 28:213–217.

Desprez-Loustau M-L, Feau N, Mougou-Hamdane A, Dutech C (2011) Interspecific and intraspecific diversity in oak powdery mildews in Europe: coevolution history and adaptation to their hosts. Mycoscience 52:165–173.

Dinsdale A, Cook L, Riginos C, et al (2010) Refined global analysis of Bemisia tabaci (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. Annals of the Entomological Society of America 103:196–208.

Druciarek T, Lewandowski M, Tzanetakis IE (2019) First Report of European Mountain Ash Ringspot-Associated Emaravirus in Sorbus aucuparia in Poland. Plant Disease 103:166.

Drummond JA, FRANCO JL de A, Ninis AB (2006) O Estado das Áreas Protegidas do Brasil.

Drummond JA, FRANCO JL de A, OLIVEIRA D de (2010) Uma análise sobre a história e a situação das unidades de conservação no Brasil. Conservação da Biodiversidade Legis e Políticas Públicas Brasília Ed Câmara.

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.

Eni AO, Hughes J d'A, Asiedu R, Rey MEC (2008) Sequence diversity among badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin and Nigeria. Archives of Virology 153:2263–2272.

Ennos RA (2014) Resilience of forests to pathogens: an evolutionary ecology perspective. Forestry: An International Journal of Forest Research 88:41–52.

Faria JC, Bezerra IC, Zerbini FM, et al (2000) Situação atual das geminiviroses no Brasil. Embrapa Arroz e Feijão-Artigo em periódico indexado.

Fernandes FR, de Albuquerque LC, de Britto Giordano L, et al (2008) Diversity and prevalence of Brazilian bipartite begomovirus species associated to tomatoes. Virus Genes 36:251–258.

Fernandes PA, Pessôa VLS (2011) O cerrado e suas atividades impactantes: uma leitura sobre o garimpo, a mineração e a agricultura mecanizada. Observatorium: Revista Eletrônica de Geografia: 3.

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

Fonseca PLC, Badotti F, de Oliveira TFP, et al (2018) Virome analyses of Hevea brasiliensis using small RNA deep sequencing and PCR techniques reveal the presence of a potential new virus. Journal of Virology 15:184.

Fontenele RS, Alves-Freitas DMT, Silva PIT, et al (2018) Discovery of the first maize-infecting mastrevirus in the Americas using a vector-enabled metagenomics approach. Archives of Virology 163:263–267.

Franck A, Guilley H, Jonard G, et al (1980) Nucleotide sequence of cauliflower mosaic virus DNA. Cell 21:285–294.

Franks P (2016) Advancing equity in protected area conservation. International Institute for Environment and Development.

Frederico S (2012) Expansão da fronteira agrícola moderna e consolidação da cafeicultura científica globalizada no Oeste da Bahia. Boletim Campineiro de Geografia 2:279–301.
Frischmuth S, Wege C, Hülser D, Jeske H (2007) The movement protein BC1 promotes redirection of the nuclear shuttle protein BV1 of Abutilon mosaic geminivirus to the plasma membrane in fission yeast. Protoplasma 230:117–123.

Fütterer J, Gordon K, Pfeiffer P, et al (1989) Differential inhibition of downstream gene expression by the cauliflower mosaic virus 35S RNA leader. Virus Genes 3:45–55.

Gafni Y, Epel BL (2002) The role of host and viral proteins in intra-and inter-cellular trafficking of geminiviruses. Physiological and Molecular Plant Pathology 60:231–241.

Galvao RM, Mariano AC, Luz DF, et al (2003) A naturally occurring recombinant DNA-A of a typical bipartite begomovirus does not require the cognate DNA-B to infect Nicotiana benthamiana systemically. Journal of General Virology 84:715–726.

García-Andrés S, Accotto GP, Navas-Castillo J, Moriones E (2007) Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. Virology 359:302–312.

García-Andrés S, Monci F, Navas-Castillo J, Moriones E (2006) Begomovirus genetic diversity in the native plant reservoir Solanum nigrum: evidence for the presence of a new virus species of recombinant nature. Virology 350:433–442.

Geering ADW, Hull R, King AMQ, et al (2012) Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses.

Geering ADW, Maumus F, Copetti D, et al (2014) Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. Nature Communications 5:5269.

Ghelardini L, Pepori AL, Luchi N, et al (2016) Drivers of emerging fungal diseases of forest trees. Forest Ecology and Management 381:235–246.

Gorbalenya AE, Koonin E V, Wolf YI (1990) A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Lett 262:145–148.

Grieco F, Alkowni R, Saponari M, et al (2000) Molecular detection of olive-infecting viruses. In: IV International Symposium on Olive Growing 586. pp 737–740.

Grimová L, Marek M, Konrady M, Ryšánek P (2015) Newly identified host range of *European mountain ash ringspot-associated* virus (EMAR aV) and its distribution in the C zech Republic. Forest Pathology 45:177–189.

Guilley H, Dudley RK, Jonard G, et al (1982) Transcription of *Cauliflower mosaic virus* DNA: detection of promoter sequences, and characterization of transcripts. Cell 30:763–773 Gutierrez C (1999) Geminivirus DNA replication. Cellular and Molecular Life Sciences 56:313–329.

Hadidi A, Barba M (2012) Next-generation sequencing: Historical perspective and Current applications in Plant Virology. Petria 22:262–277.

Hadidi A, Barba M, Candresse T, Jelkmann W (2011) Virus and virus-like diseases of pome and stone fruits. APS Press/American Phytopathological Society.

Hadidi A, Flores R, Candresse T, Barba M (2016) Next-generation sequencing and genome editing in plant virology. Frontiers in Microbiology 7:. doi:10.3389/fmicb.2016.01325.

Hadidi A, Flores R, Randles J, Semancik J (2003) Viroids: properties, detection, diseases and their control. Csiro Publishing.

Hadidi A, Khetarpal RK, Koganezawa H (1998) Plant virus disease control. APS press.

Hamza M, Tahir MN, Mustafa R, et al (2018) Identification of a dicot infecting mastrevirus along with alpha-and betasatellite associated with leaf curl disease of spinach (*Spinacia oleracea*) in Pakistan. Virus Researchearch 256:174–182.

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews 68:669–685.

Hansen EM, Goheen EM (2000) Phellinus weirii and other native root pathogens as determinants of forest structure and process in western North America. Annual Review of Phytopathology 38:515–539.

Harper G, Hull R, Lockhart BEN, Olszewski N (2002) Viral sequences integrated into plant genomes. Annual Review of Phytopathology 40:119–136.

Hendry AP, Kinnison MT, Heino M, et al (2011) Evolutionary principles and their practical application. Evolutionary Applications 4:159–183.

Ho T, Tzanetakis IE (2014) Development of a virus detection and discovery pipeline using next generation sequencing. Virology 471–473:54–60. doi: 10.1016/j.virol.2014.09.019.

Höfer P, Bedford ID, Markham PG, et al (1997) Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. Virology 236:288–295.

Holmes EC (2011) What does virus evolution tell us about virus origins? Journal of Virology 85:5247–5251.

Hull R (1996) Molecular biology of rice tungro viruses. Annual Review of Phytopathology 34:275–297.

IBÁ (2017) Industria Brasileira de Arvores. pp 1–80.

IBÁ (2018) INSTITUTO BRASILEIRO DE ÁRVORES (IBA). https://www.iba.org/dados-estatisticos. Accessed 10 Dec 2019.

IBRAM (2018) Relatório Anual - INSTITUTO BRASÍLIA AMBIENTAL - IBRAM. http://iba.org/images/shared/Biblioteca/IBA_RelatorioAnual2016. Accessed 10 Dec 2019.

ICTV (2019) ICTV online. International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/. Accessed 21 Dec 2019.

Ilyina T V, Koonin E V (1992) Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Research 20:3279–3285.

Inoue-Nagata AK, Lima MF, Gilbertson RL (2016) 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, Martin DP, Boiteux LS, et al (2006) New species emergence via recombination among isolates of the Brazilian tomato infecting Begomovirus complex. Pesquisa Agropecuária Brasileira 41:1329–1332.

Islam W, Zhang J, Adnan M, et al (2017) Plant virus ecology: a glimpse of recent accomplishments. Applied Ecology and Environmental Research 15:691–705.

James AP, Geijskes RJ, Dale JL, Harding RM (2011) Development of a novel rolling-circle amplification technique to detect Banana streak virus that also discriminates between integrated and episomal virus sequences. Plant Disease 95:57–62.

Jiwaji M, Matcher GF, de Bruyn M-M, et al (2019) Providence virus: An animal virus that replicates in plants or a plant virus that infects and replicates in animal cells? PLoS One 14:e0217494.

Johnson AMA, Borah BK, Gopal DVRS, Dasgupta I (2012) Analysis of full-length sequences of two Citrus yellow mosaic badnavirus isolates infecting Citrus jambhiri (Rough Lemon) and *Citrus sinensis* L. Osbeck (*Sweet Orange*) from a nursery in India. Virus Genes 45:600–605.

Jousimo J, Tack AJM, Ovaskainen O, et al (2014) Ecological and evolutionary effects of fragmentation on infectious disease dynamics. Science (80-) 344:1289–1293.

Juffe-Bignoli D, Burgess ND, Bingham H, et al (2014) Protected Planet Report 2014 Cambridge. UK UNEP-WCMC.

Kamali M, Heydarnejad J, Massumi H, et al (2016) Molecular diversity of turncurtoviruses in Iran. Archives of Virology 161:551–561.

Karp DS, Mendenhall CD, Sandí RF, et al (2013) Forest bolsters bird abundance, pest control and coffee yield. Ecology Letters 16:1339–1347.

Kaur N, Hasegawa DK, Ling K-S, Wintermantel WM (2016) Application of genomics for understanding plant virus-insect vector interactions and insect vector control. Phytopathology 106:1213–1222.

King AMQ, Lefkowitz E, Adams MJ, Carstens EB (2011) Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses. Elsevier.

Klein A, STEFFAN-DEWENTER I, Tscharntke T (2006) Rain forest promotes trophic interactions and diversity of trap-nesting Hymenoptera in adjacent agroforestry. Journal of Animal Ecology 75:315–323.

Koonin E V (1993) A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Research 21:2541–2547.

Kouakou K, Kébé BI, Kouassi N, et al (2012) Geographical distribution of cacao swollen shoot virus molecular variability in Côte d'Ivoire. Plant Disease 96:1445–1450.

Kreuze JF, Perez A, Untiveros M, et al (2009a) 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. doi: 10.1016/j.virol.2009.03.024.

Kreuze JF, Perez A, Untiveros M, et al (2009b) 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.

Kumar J, Kumar J, Singh SP, 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.

Lamas NS, Fontenele RS, Melo FL, et al (2016) Complete genome sequence of a genomovirus associated with common bean plant leaves in Brazil. Genome Announcements 4:e01247-16.

Laufs J, Schumacher S, Geisler N, et al (1995) Identification of the nicking tyrosine of geminivirus Rep protein. FEBS Lett 377:258–262.

Lazarowitz SG, Shepherd RJ (1992) Geminiviruses: genome structure and gene function. Critical Reviews in Plant Sciences 11:327–349.

Li H, Deng Z, Wu T, et al (2012) Microwave-assisted extraction of phenolics with maximal antioxidant activities in tomatoes. Food Chem 130:928–936.

Liang P, Navarro B, Zhang Z, et al (2015) Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. Journal of General Virology 96:2411–2420.

Lima JAA (2015) Virologia essencial & viroses em culturas tropicais. Fortaleza Ed UFC.

Lima JAA, Gonçalves MFB (1988) *Cassia occidentalis*, potencial reservatório de potyvirus que infetam caupi. Fitopatologia Brasileira 13:365–368.

Lin MT, Kitajima EW, Costa CL, others (1980) Association of cassia mild mosaic virus with dieback of Cassia macranthera in central Brazil. Plant Disease 64:587–589.

Lin MT, Kitajima EW, Cupertino FP, Costa CL (1979) Properties of a possible carlavirus isolated from a cerrado native plant, Cassia sylvestris. Plant Disease Report.

Loconsole G, Saldarelli P, Doddapaneni H, et al (2012) Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family *Geminiviridae*. Virology 432:162–172.

Lourenção AL, Nagai H (1994) Surtos populacionais de *Bemisia tabaci* no Estado de São Paulo. Bragantia 53:53–59.

Lozano G, Trenado HP, Fiallo-Olivé E, et al (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.

MacDiarmid R, Rodoni B, Melcher U, et al (2013) Biosecurity implications of new technology and discovery in plant Virus Researchearch. PLOS Pathogens 9:e1003337.

Maliogka VI, Minafra A, Saldarelli P, et al (2018) Recent advances on detection and characterization of fruit tree viruses using high-throughput sequencing technologies. Viruses 10:436.

Mankertz A, Persson F, Mankertz J, et al (1997) Mapping and characterization of the origin of DNA replication of porcine circovirus. Journal of Virology 71:2562–2566.

Marais A, Faure C, Couture C, et al (2014) Characterization by deep sequencing of divergent Plum bark necrosis stem pitting-associated virus (PBNSPaV) isolates and development of a broad-spectrum PBNSPaV detection assay. Phytopathology 104:660–666.

Martin DP, Lefeuvre P, Varsani A, et al (2011) Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. PLOS Pathogens 7:e1002203.

Martin RR, MacFarlane S, Sabanadzovic S, et al (2013) Viruses and virus diseases of Rubus. Plant Disease 97:168–182.

Matsumura EE (2016) Estudo das populações de vírus presentes em plantas de citros cultivadas em uma região afetada pela morte súbita dos citros.

Mercadante M (2001) Uma década de debate e negociação: a história da elaboração da Lei do SNUC. Direito Ambient das áreas protegidas o regime jurídico das unidades Conserv Rio Janeiro Forense Univ 190–231.

Mielke N, Weber M, Khan S, Muehlbach H-P (2008) Detection of European mountain ash ringspot-associated virus (EMARAV) in *Sorbus aucuparia* L. by a specific antiserum and reverse transcription-PCR. Forest Pathology 38:371–380.

MMA M do MA (2017) http://www.mma.gov.br/biomas. Accessed 21 Dec 2019.

MMA M do MA (2018). https://www.mma.gov.br/areas-protegidas/unidades-de-conservação. Accessed 21 Dec 2019.

Mochizuki T, Krupovic M, Pehau-Arnaudet G, et al (2012) Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. Proceedings of the National Academy of Sciences 109:13386–13391.
Mokili JL, Rohwer F, Dutilh BE (2012) Metagenomics and future perspectives in virus discovery. Current Opinion in Virology2:63–77. doi: 10.1016/j.coviro.2011.12.004.

Molenaar N, Burger JT, Maree HJ (2015) Detection of a divergent variant of grapevine virus F by next-generation sequencing. Archives of Virology 160:2125–2127.

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 WA, Alicai T, Ndunguru J, et al (2010) The complete genome sequence of the Tanzanian strain of *Cassava brown streak virus* and comparison with the Ugandan strain sequence. Archives of Virology 155:429–433.

Morin S, Ghanim M, Sobol I, Czosnek H (2000) The GroEL protein of the whitefly Bemisia tabaci interacts with the coat protein of transmissible and nontransmissible begomoviruses in the yeast two-hybrid system. Virology 276:404–416.

Muthukumar V, Melcher U, Pierce M, et al (2009) Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. Virus Research 141:169–173.

Nagasaki K, Tomaru Y, Takao Y, et al (2005) Previously unknown virus infects marine diatom. Applied and Environmental Microbiology71:3528–3535.

Nash TE, Dallas MB, Reyes MI, et al (2011) Functional analysis of a novel motif conserved across geminivirus Rep proteins. Journal of Virology 85:1182–1192.

Navas-Castillo J, Fiallo-Olivé E, Sánchez-Campos S (2011) Emerging virus diseases transmitted by whiteflies. Annual Review of Phytopathology 49:219–248.

Nicolini C, Pio-Ribeiro G, Andrade GP, et al (2012) A distinct tymovirus infecting Cassia hoffmannseggii in Brazil. Virus Genes 45:190–194.

Nienhaus F, Castello JD (1989) Viruses in Forest Trees. Annual Review of Phytopathology27:165–186. doi: 10.1146/annurev.py.27.090189.001121.

Noueiry AO, Lucas WJ, Gilbertson RL (1994) Two proteins of a plant DNA virus coordinate

nuclear and plasmodesmal transport. Cell 76:925–932.

NOVACAP (2018) COMPANHIA URBANIZADORA DA NOVA CAPITAL DO NOVACAP - BRASIL. http://www.novacap.df.gov.br/ 2018. Accessed 21 Dec 2019.

Orozco BM, Hanley-Bowdoin L (1996) A DNA structure is required for geminivirus replication origin function. Journal of Virology 70:148–158.

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

Overbeck GE, Vélez-Martin E, Scarano FR, et al (2015) Conservation in Brazil needs to include non-forest ecosystems. Diversity and Distributions 21:1455–1460.

Palumbi SR (2001) Humans as the world's greatest evolutionary force. Science (80-) 293:1786–1790.

Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant-pathogen interactions. Annual Review of Ecology, Evolution, and Systematics Syst 35:675–700.

Pereira-Carvalho RC., Sanchez M., Bezerra ACC., et al (2004) A micobiota da Reserva Ecológica do IBGE. In: IBGE Reserva Ecológica do Instituto Brasileiro de Pesquisa Geográfica e Estatística. Brasília.

Peterson PD, Griffith CS (1999) Hermann von Schrenk and the beginning of forest pathology in the US. For Host Today 29–34.

Pietilä MK, Roine E, Paulin L, et al (2009) An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Molecular Microbiology 72:307–319.

Pimentel D, Kounang N (1998) Ecology of soil erosion in ecosystems. Ecosystems 1:416–426.

Pita JS, Fondong VN, Sangare A, et al (2001) Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. Journal of General Virology 82:655–665.

Ploetz R (2016) The impact of diseases on cacao production: a global overview. In: Cacao Diseases. Springer, pp 33–59.

Power AG (2010) Ecosystem services and agriculture: tradeoffs and synergies. Philosophical Transactions of the Royal Society B 365:2959–2971.

Pratap D, Kashikar AR, Mukherjee SK (2011) Molecular characterization and infectivity of a Tomato leaf curl New Delhi virus variant associated with newly emerging yellow mosaic disease of eggplant in India. 8:305.

Razavinejad S, Heydarnejad J (2013) Transmission and natural hosts of Turnip curly top virus. Iranian Journal of Pathology 49:27–28.

RECOR (2018) RECOR – Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística. https://recor.org.br/. Accessed 21 Dec 2019.

Reyes MI, Nash TE, Dallas MM, et al (2013) Peptide aptamers that bind to geminivirus replication proteins confer a resistance phenotype to tomato yellow leaf curl virus and tomato mottle virus infection in tomato. Journal of Virology 87:9691–9706.

Rezende A V, Walter BMT, Fagg CW, et al (2008) Flora vascular do bioma cerrado. Cerrado Ecologia e flora Embrapa, Brasília 1028–1059.

RIBEIRO S da G (2006) Diversity and host interactions of emerging tomato Begomoviruses in Brazil. 2006.

Ribeiro SG, Ambrozevicius LP, Avila AC, et al (2003) Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. Archives of Virology 148:281–295.

Ribeiro SG, Martin DP, Lacorte C, et al (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.

Richins RD, Scholthof HB, Shepherd RJ (1987) Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Research15:8451–8466.

Rodrigues LK, de Oliveira AM, Chaves ALR, et al (2019) *Cauliflower mosaic virus* naturally infects wild radish (*Raphanus raphanistrum*) in Brazil. Australas Plant Disease Notes 14:26.

Rodríguez-Negrete EA, Morales-Aguilar JJ, Domínguez-Duran G, et al (2019) High-Throughput Sequencing Reveals Differential Begomovirus Species Diversity in Non-Cultivated Plants in Northern-Pacific Mexico. Viruses 11:594.

Rodríguez-Nevado C, Gavilán R, Pagán I (2019) Host abundance and identity determine the epidemiology and evolution of a generalist plant virus in a wild ecosystem. Phytopathology.

Rodríguez-Nevado C, Montes N, Pagán I (2017) Ecological factors affecting infection risk and population genetic diversity of a novel potyvirus in its native wild ecosystem. Frontiers

in Plant Science8:1958.

Rojas MR, Hagen C, Lucas WJ, Gilbertson RL (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annual Review of Phytopathology43:361–394.

Rojstaczer S, Sterling SM, Moore NJ (2001) Human appropriation of photosynthesis products. Science (80-) 294:2549–2552.

Roossinck MJ, García-Arenal F (2015) Ecosystem simplification, biodiversity loss and plant virus emergence. Current Opinion in Virology 10:56–62. doi: 10.1016/j.coviro.2015.01.005.

Roossinck MJ, Martin DP, Roumagnac P (2015) Plant virus metagenomics: advances in virus discovery. Phytopathology 105:716–727.

Roossinck MJ, Saha P, Wiley GB, et al (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. Molecular Ecology 19:81–88.

Rosario K, Breitbart M, Harrach B, et al (2017) Revisiting the taxonomy of the family *Circoviridae*: establishment of the genus *Cyclovirus* and removal of the genus *Gyrovirus*. Archives of Virology 162:1447–1463.

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

Rosolen V, Resende TM, Borges EN, et al (2012) Impactos da substituição da vegetação original do Cerrado brasileiro em sistemas agrícolas: alteração do carbono orgânico do solo e δ13C. Investig Geográficas 39–47.

Roumagnac P, Granier M, Bernardo P, et al (2015) *Alfalfa leaf curl virus*: An aphid-transmitted geminivirus. Journal of Virology 89:9683–9688.

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

Ruiz-Ruiz S, Navarro B, Gisel A, et al (2011) Citrus tristeza virus infection induces the accumulation of viral small RNAs (21–24-nt) mapping preferentially at the 3′-terminal region of the genomic RNA and affects the host small RNA profile. Plant Molecular Biology 75:607–619.

Rumbou A, Candresse T, Marais A, et al (2018) A novel badnavirus discovered from Betula sp. affected by birch leaf-roll disease. PLoS One 13:e0193888.

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.

Sanjuán R, Nebot MR, Chirico N, et al (2010) Viral mutation rates. Journal of Virology 84:9733–9748.

Scheets K, Blinkova O, Melcher U, et al (2011) Detection of members of the Tombusviridae in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. Virus Research 160:256–263.

Schoettle AW, Sniezko RA (2007) Proactive intervention to sustain high-elevation pine ecosystems threatened by white pine blister rust. Journal of Forest Research 12:327–336.

Scholthof KBG, Adkins S, Czosnek H, et al (2011) Top 10 plant viruses in molecular plant pathology. Molecular Plant Pathology 12:938–954. doi: 10.1111/j.1364-3703.2011.00752.x.

Scott SW, Zimmerman MT (2008) The complete sequence of ligustrum necrotic ringspot virus, a novel carlavirus. Archives of Virology 153:393.

Seabra P V, Rivas EB, Duarte LML, et al (2001) Detecção de Carlavirus em *Senna macranthera*. Ornamental Horticulture 7.

Sepúlveda-Chavera G., Pereira-Carvalho RC., Sanchez M;, Dianese J. (2008) Micobiota. In: Seduma.

Settlage SB, Miller AB, Hanley-Bowdoin L (1996) Interactions between geminivirus replication proteins. Journal of Virology 70:6790–6795.

Silva SJC da (2006) Detecção, caracterização molecular e diversidade genética de begomovirus que infectam fava (*Phaseolus lunatus* L.).

Silva SJC, Castillo-Urquiza GP, Hora-Júnior BT, et al (2012) Species diversity, phylogeny and genetic variability of begomovirus populations infecting leguminous weeds in northeastern Brazil. Plant Pathology 61:457–467.

Simmonds P, Adams MJ, Benk\Ho M, et al (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nature Reviews Microbiology 15:161.

SNIF - 2019 Sistema Nacional de Informações Florestais http://snif.florestal.gov.br/pt-br/especies-florestais. Acessed 21 Dec 2019.

Steinfeldt T, Finsterbusch T, Mankertz A (2006) Demonstration of nicking/joining activity

at the origin of DNA replication associated with the rep and rep′ proteins of porcine circovirus type 1. Journal of Virology 80:6225–6234.

Stobbe AH, Roossinck MJ (2014) Plant virus metagenomics: what we know and why we need to know more. Frontiers in Plant Science 5:150. doi: 10.3389/fpls.2014.00150.

Strausbaugh CA, Rearick E, Camp S (2008) Influence of curly top and Poncho Beta on storability of sugarbeet. Journal of Sugar Beet Research 45:31.

Thapa V, Melcher U, Wiley GB, et al (2012) Detection of members of the secoviridae in the tallgrass prairie preserve, osage county, Oklahoma, USA. Virus Research 167:34–42.

Timchenko T, De Kouchkovsky F, Katul L, et al (1999) A single rep protein initiates replication of multiple genome components of faba bean necrotic yellows virus, a single-stranded DNA virus of plants. Journal of Virology 73:10173–10182.

Timmermans MCP, Das OP, Messing J (1994) Geminiviruses and their uses as extrachromosomal replicons. Annual Review of Plant Biology 45:79–112.

Tomaru Y, Takao Y, Suzuki H, et al (2011) Isolation and characterization of a single-stranded DNA virus infecting Chaetoceros lorenzianus Grunow. Applied and Environmental Microbiology 77:5285–5293.

UnB (2018) UnB - Universidade de Brasília. http://www.fal.unb.br/. Accessed 21 Dec 2019.

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

Varsani A, Krupovic M (2017) Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family *Genomoviridae*. Virus Evolution 3:vew037.

Varsani A, Navas-Castillo J, Moriones E, et al (2014) 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, et al (2017a) *Capulavirus* and *Grablovirus*: two new genera in the family Geminiviridae. Archives of Virology 162:1819–1831. doi: 10.1007/s00705-017-3268-6.

Varsani A, Shepherd DN, Dent K, et al (2009) A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virology Journal 6:36.

Vega-Rocha S, Byeon I-JL, Gronenborn B, et al (2007) Solution structure, divalent metal and DNA binding of the endonuclease domain from the replication initiation protein from porcine circovirus 2. Journal of Molecular Biology 367:473–487.

Villamor DE V, Mekuria TA, Pillai SS, Eastwell KC (2016) High-throughput sequencing identifies novel viruses in nectarine: insights to the etiology of stem-pitting disease. Phytopathology 106:519–527.

Vincent SJ, Coutts B a., Jones R a C (2014) Effects of introduced and indigenous viruses on native plants: Exploring their disease causing potential at the agro-ecological interface. PLoS One 9:6–9. doi: 10.1371/journal.pone.0091224.

Vitousek PM, Ehrlich PR, Ehrlich AH, Matson PA (1986) Human appropriation of the products of photosynthesis. Bioscience 36:368–373.

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. EMBO J 1:945–951.

Wang YJ, Yao Q, Chen KP, et al (2007) Characterization of the genome structure of Bombyx mori densovirus (China isolate). Virus Genes 35:103–108.

Ward BM, Lazarowitz SG (1999) Nuclear export in plants: use of geminivirus movement proteins for a cell-based export assay. Plant Cell 11:1267–1276.

Webb MD (1987) Species recognition in Cicadulina leafhoppers (Hemiptera: Cicadellidae), vectors of pathogens of Gramineae. Bulletin of Entomological Research 77:683–712.

Wingfield MJ, De Beer ZW, Slippers B, et al (2012) One fungus, one name promotes progressive plant pathology. Molecular Plant Pathology 13:604–613.

Word Bank (2019) WORLD BANK. https://www.worldbank.org/en/topic/forests. Accessed 21 Dec 2019.

Wright EA, Heckel T, Groenendijk J, et al (1997) Splicing features in maize streak virus virion-and complementary-sense gene expression. The Plant Journal 12:1285–1297.

Wu Q, Ding S-W, 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.

WWF (2018) - WORLD WILDLIFE FUND. http://www.wwf.org.br. Accessed 21 Dec 2019.

Yang IC, Hafner GJ, Dale JL, Harding RM (2003) Genomic characterisation of taro bacilliform virus. Archives of Virology 148:937–949.

Yu X, Li B, Fu Y, et al (2010) A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. Proceedings of the National Academy of Sciences 107:8387–8392.

Yu X, Li B, Fu Y, et al (2013) Extracellular transmission of a DNA mycovirus and its use as a natural fungicide. Proceedings of the National Academy of Sciences 110:1452–1457.

Zerbini FM, Andrade EC, Barros DR, et al (2005) Traditional and novel strategies for geminivirus management in Brazil. Australas Plant Pathology34:475–480.

Zerbini FM, Briddon RW, Idris A, et al (2017) ICTV Virus Taxonomy Profile: *Geminiviridae*. Journal of General Virology 98:131–133. doi: 10.1099/jgv.0.000738.

Zhou X (2013) Advances in understanding begomovirus satellites. Annual Review of Phytopathology 51:357–381

Chapter 2

Single-stranded DNA viroma of vegetable and field crops and from native and exotic tree species growing in the Brazilian *Cerrado* biome

Resumo

O Cerrado é um bioma neotropical com ocorrência endêmica de mais de 4000 plantas superiores, incluindo as principais espécies arbóreas e herbáceas. É também a principal região de produção agrícola no Brasil. No entanto, pouco se sabe ainda sobre vírus naturalmente associados a árvores (nativas e exóticas) e espécies agrícolas neste bioma. Os avanços nas análises metagenômicas intensificaram a taxa de descoberta de vírus associados a plantas. Neste trabalho nós avaliamos a diversidade de vírus de DNA fita simples (ssDNA) em plantas que crescem no bioma Cerrado (Brasil Central). Ao todo, 296 amostras de folhas (correspondendo a 131 espécies em 49 famílias) foram obtidas em três microambientes: **(a) Áreas Antrópicas** - Arborização Urbana do Distrito Federal (AUDF), Núcleo Rural Taquara (NRT), Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP) e Área Experimental da Fazenda Água Limpa (AE-FAL); **(b) Áreas de Conservação** - Estação Ecológica de Águas Emendadas (ESECAE) e Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR) e **(c) Área de Transição - "**Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) e Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). Extrações de DNA foram enriquecidas em moléculas circulares via Rolling Circle Amplification (RCA) e submetidas ao High-Throughput Sequencing (HTS). Foram gerados 297.103 *contigs* sendo que 503 deles apresentaram identidade do BLASTx para vírus de plantas e 24 espécies virais foram anotadas. *Primers* específicos foram desenhados e utilizados para detecção e/ou recuperação do genoma completo mediante *Sanger*, validando-se assim o resultado de HTS. A distribuição de vírus nas áreas dos diferentes microambientes mostrou maior prevalência de vírus em ambiente antrópico pertencente à família *Geminiviridae* enquanto em ambiente de conservação poucos vírus foram detectados. Neste ambiente as detecções foram de vírus classificados na família *Genomoviridae*. Do total de 49 famílias, 13 tiveram ao menos uma espécie positiva para algum vírus. Detecção viral em espécies das famílias Fabaceae e Solanaceae foram as mais frequentes. Os 30 genomas recuperados (a partir de 39 *contigs*) foram equivalentes a 24 espécies virais a saber: **a)** 12 novas espécies classificados nas famílias: *Geminiviridae* (uma), *Genomoviridae* (4), *Circoviridae* (4) e ssDNA sem classificação (3). A provável espécie nova classificada na família *Geminiviridae* foi detectada em *Macroptilium erythroloma* (AUDF) em infecção mista com *Bean golden mosaic virus* - BGMV. As quatro novas espécies de genomovírus foram detectados em *Vochysia rufa* e *Byrsonima crassiflora* (RECOR); em *Tecoma stans* (AUDF) e *Ouratea duparquetiana* (ARIE-EEB). Foram propostos os nomes (seguindo o nome da planta que está associado):

Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV (sem classificação em gênero), Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) e *Ouratea duparquetiana* associated gemykibivirus - OuaGmV (*Gemykibivirus*) e**b)** Dezoito *contigs* (12 espécies) referiam-se à espécies já descritas e foram classificados nas famílias *Caulimoviridae* (2), *Genomoviridae* (2) e *Geminiviridae* (14). O CaMV (*Caulimovirus* - *Caulimoviridae*), possui ampla gama de hospedeiras e foi detectado em couve (*Brassica oleraceae*) e primeiro relato ocorrendo naturalmente em fevilha (*Fevillea trilobata*), na AHE-EEB. Dois isolados das espécies Gila monster-associated gemykrogvirus - (*Gemykrogvirus*) - GmaV1 e Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) foram detectadas em *Trembleya parviflora* e berinjela (*Solanum melongena*), respectivamente, a primeira na RECOR e a segunda no NRT. Quatorze *contigs* apresentaram homologia com sequências de oito espécies da família *Geminiviridae* já conhecidas para Ciência, sendo duas espécies, sem classificação em gênero, encontradas na AE-FAL: Tomato apical leaf curl virus - ToALCV, em tomate (*Solanum lycopersicum*), e o Tomato associated geminivirus 1 - TaGV1 em tomate e beterraba (*Beta vulgaris*). Isolado de uma espécie do gênero *Mastrevirus*, Maize striate mosaic virus - MSMV foi detectado em milho (*Zea mays*) e cana de açúcar (*Saccharum officinarum*) na AHE-EEB. Isolados de duas espécies do gênero *Begomovirus* (*Bean golden mosaic virus* - BGMV e *Tomato severe rugose virus* - ToSRV) foram detectados em cinco áreas de dois microambientes analisados. Isolados de BGMV foram detectados em feijão (*Phaseolus vulgaris*) na AHE-EEB e AE-FAL, em *M. erythroloma* na AUDF, em *Anadenanthera colubrina* na ARIE-EEB e em berinjela, na NRT. Isolados de ToSRV foram detectados em tomate da AHE-EEB e AE-FAL, algodão (*Gossypium hirsutum*) (AE-FAL) e berinjela, guanxuma (*Sida* sp.) e joá de capote (*Nicandra physalodes*) do NRT. Este é o primeiro relato de ToSRV ocorrendo naturalmente em algodão. Isolados de *Sida micrantha mosaic virus* - SiMMV foram encontrados em áreas do microambiente antrópico, em plantas de algodão na AE-FAL, em guanxuma e joá de capote no NRT. Isolados de *Tomato mottle leaf curl virus* - ToMoLCV foram encontrados somente na AHE-EEB em amostras de tomate, na planta daninha *Euphorbia heterophylla* e arbórea *Ouratea duparquetiana*. Isolados de SPLCV foram detectados nos ambientes da AHE-EEB, AE-FAL e NRT em amostras de sua hospedeira natural, a batata doce (*Ipomoea batatas*). As análises metagenômicas indicaram uma vasta diversidade enigmática de vírus ssDNA em árvores que crescem neste bioma altamente ameaçado.

Palavras chave: espécies arbóreas; antrópico; diversidade de vírus, áreas de conservaçao, áreas de transição

Abstract

The *Cerrado* is a Neotropical biome with an endemic occurrence of over 4000 higher plants, including major tree and herbaceous species. It is also the major region for field crop production in Brazil. However, little is yet known on viruses naturally associated with trees

(native and exotic) and agricultural species in this biome. Advances in metagenomic analyses have intensified the discovery rate of plant-associated viruses. In this work were evaluated DNA single stranded virus diversity in plants growing in *Cerrado* biome (Central Brazil). In all, 296 leaf samples (corresponding to 131 species in 49 families) were obtained from three microenvironments **(a) Anthropic Areas** - Arborização Urbana do Distrito Federal (AUDF), Núcleo Rural Taquara (NRT), Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP) e Área Experimental da Fazenda Água Limpa (AE-FAL); **(b) Conservation Areas** - Estação Ecológica de Águas Emendadas (ESECAE) e Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR) and **(c) Transition Areas -"**Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) e Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). DNA extractions were enriched in circular molecules via Rolling Circle Amplification (RCA) and subjected to High-Throughput Sequencing (HTS), Illumina sequencing. A total of 297,103 contigs were generated and 503 showed BLASTx identity for plant viruses and 24 viral species were annotated. Specific primers were designed and used for detection and/or recovery the complete genome by Sanger, validating HTS result. Virus distribution in different areas of microenvironments showed a higher prevalence of viruses in an anthropic environment belonging to family *Geminiviridae* while in a conservation environment few viruses were detected. In this environment the detections were of viruses classified in family *Genomoviridae*. Thirteen of the 49 families had least one positive virus detection. Virus detection in Fabaceae and Solanaceae families were the most frequent. Thirty genomes were recovered (from 39 contigs) corresponding to 24 viral species: **a)** 12 new species classified into families: *Geminiviridae* (1), *Genomoviridae* (4), *Circoviridae* (4) and unclassified ssDNA (3). The probable new species classified in family *Geminiviridae* was detected in *Macroptilium erythroloma* (AUDF) in mixed infection with *Bean golden mosaic virus* - BGMV isolates. The four new genomovirus species were detected in *Vochysia rufa* and *Byrsonima crassiflora* (RECOR); in *Tecoma stans* (AUDF) and *Ouratea duparquetiana* (ARIE-EEB). The following names were proposed (following the name of the associated plant): Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV (not classified by genus), Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) and Ouratea duparibivirus associated gemyvirus OuaGmV (*Gemykibivirus*) and **b)** Eighteen contigs (12 species) referred to species already described were classified in families *Caulimoviridae* (2), *Genomoviridae* (2) and *Geminiviridae* (14). The CaMV (*Caulimovirus* - *Caulimoviridae*) has a wide host range and was detected in cabbage (*Brassica oleraceae*) and first report occurring naturally in antidote cannoon (*Fevillea trilobata*) in AHE-EEB. Two isolates of Gila monster-associated gemycircularvirus - (*Gemykrogvirus*) - GmaV1 and Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) were detected in *Trembleya parviflora* and eggplant (*Solanum melongena*), respectively, the first in RECOR and the second in NRT areas. Fourteen contigs showed homology with eight *Geminiviridae* species already known to science, but two of them, not classified in genus, were found from samples of AE-FAL: Tomato apical leaf curl virus - ToALCV in tomato (*Solanum lycopersicum*) and Tomato associated geminivirus 1 - TaGV1 in tomato and beet (*Beta vulgaris*). Isolates of *Mastrevirus*,

Maize striate mosaic virus - MSMV, were detected in maize (*Zea mays*) and sugarcane (*Saccharum officinarum*) in AHE-EEB. Isolates of two *Begomovirus* species (*Bean golden mosaic virus* - BGMV e *Tomato severe rugose virus* - ToSRV) were detected in five areas from two microenvironments analised. BGMV isolates were detected in beans (*Phaseolus vulgaris*) in AHE-EEB and AE-FAL, in *M. erythroloma* in AUDF, in *Anadenanthera colubrina* in ARIE-EEB and in eggplant from NRT samples. ToSRV isolates were detected in tomato from AHE-EEB and AE-FAL, cotton (*Gossypium hirsutum*) (AE-FAL), eggplant, guanxuma (*Sida* sp.) and "shoo-fly plant" (*Nicandra physalodes*) from NRT. This is the first report of ToSRV occurring naturally occurring in cotton. *Sida micrantha mosaic virus* - SiMMV isolates were found at anthropic microenvironments in cotton plants from AE-FAL samples, in guanxuma and joá de capote, from NRT. Isolates of *Tomato mottle leaf curl virus* - ToMoLCV were found only in AHE-EEB in tomato samples, in wild poinsettia *Euphorbia heterophylla* and *Ouratea duparquetiana*. Isolates of SPLCV were detected in AHE-EEB, AE-FAL and NRT environments in samples of its natural host, sweet potato (*Ipomoea batatas*). Metagenomic analysis indicated a vast enigmatic diversity of ssDNA viruses in trees growing in this highly threatened biome.

Keywords: trees species; anthropic; virus diversity, conservation areas, transition areas.

1. Introduction

Tree species play an important role in the economy, society, and environment. Brazil is internationally recognized for its biodiversity and for preserving and conserving its natural resources via protected areas distributed into distinct biomes in order to maintain adequate levels of biodiversity, to protect water reservoirs and fountains, and to preserve ecological corridors among biomes (Drummond et al. 2010). However, over the years native areas have been replaced by urban areas and by monocultures with an intense expansion of the agricultural boundaries and areas with pastures (Fernandes and Pessôa 2011, Rosolen et al. 2012a). One of the biotic problems that affect tree orchards, native and planted forests as well as field and vegetable crops refers to occurrence of pests and diseases. Among the major

diseases are those caused by viruses (Gonthier and Nicolotti 2013). The proximity between native ecosystems and agricultural fields may contributed to transition of novel plant pathogens from one environment to another, especially in the case of vector-borne viruses (Vincent et al. 2014, Hadidi et al. 2016).

Studies on viruses occurring in tree species are thus far mainly reported in the Northern Hemisphere (Roossinck et al. 2010, MacDiarmid et al. 2013, Vincent et al. 2014, Navarro et al. 2018). In Brazil, despite the exuberant richness of tree species, virus studies occurring in these species are incipient and there are few reports such as Nicolini et al. (2012), Batista (2014) and Fonseca et al. (2018). The submicroscopic nature of the viruses coupled with traditional detection techniques require prior information of at least which viral genus may be involved with a given disease outbreak. However, this scenario is now changing. The recent development of High-Throughput Sequencing - HTS technologies coupled with metagenomic studies have allowed the identification of new viral species and the detection of known species. This approach is currently consolidated as a fast and accurate strategy of obtaining a large number of complete genomes of plant-associated viruses (Quan et al. 2008).

The *Cerrado* biome is a savannah-like Neotropical ecosystem, encompassing ≈ 2 million km^2 with a very diversified flora and fauna (Marquis, 2002). The *Cerrado* flora is composed by over 4000 higher plant species, including major tree and herbaceous species. The areas encompassing the *Cerrado* biome is also the major region for field crop production in Brazil, including important crops such as soybean, bean, maize, cotton, and processing tomato. However, little is known on viruses naturally associated with trees (native and

exotic). Due the lack of information an virus species diversity, geographical distribution and host range in forest crops of economic and/or environmental importance, hypotheses were raised that there are virus species, some still unknown to science occurring in tree species and proximity between cultivated and agricultural environments may favor the transit of viruses from one environment to another. As mentioned, advances in metagenomic analyses have intensified the discovery rate of plant-associated viruses. In this context, in the present work we assessed single stranded DNA (ssDNA) virus diversity in plants (native trees, exotic trees, field crops and weeds) growing across distinct microenvironments of Brazilian Cerrado in Brazil Central: areas with anthropic action; areas of conservation, and transitional areas (with close proximity between conservation and anthropic environments).

2. Materials and Methods

2.1. Sampling sites

For study of viral community naturally associated with trees (native and exotic) as well as agricultural species cultivated in *Cerrado* biome from Central Brazil, the following microenvironmental areas were selected: **(1)** Areas with strong anthropic action (afforestation, nursery and agricultural crops); **(2)** Areas of conservation and **(3)** Areas of transition with production of several agricultural crops, but close to natural conservation areas. The following collection sites were classified as: **anthropic areas:** Área de Urbanização do Distrito Federal - AUDF, Núcleo Rural Taquara - NRT, Viveiro II da Companhia Urbanizadora da Nova Capital - VII-NOVACAP and Área Experimental da Fazenda Água Limpa da Universidade de Brasília - AE-FAL. The following areas were

classified as **conservation areas:** Estação Ecológica de Águas Emendadas - ESECAE and Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística - RECOR. The Estação Experimental de Biologia da Universidade de Brasília - ARIE-AHE-EEB (also belonging to UnB) was considered as a **transitional area** due to its proximity among conservation and anthropic. The anthropic areas of AE-FAL and NRT are located near the conservation areas, RECOR (10 km distance) and ESECAE (20 km distance), respectively. The locations are shown in **Figure 1.**

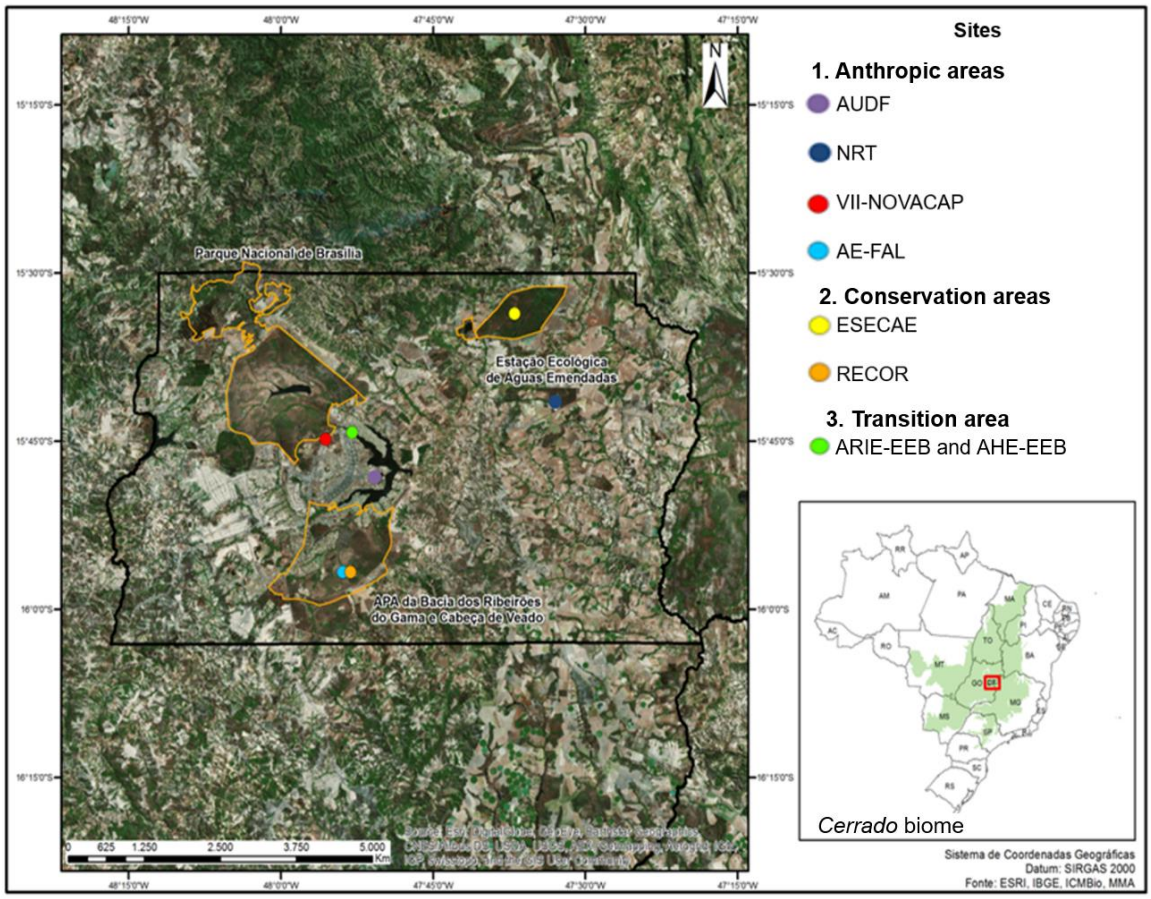


Figure 1. Plant collection sites (tree and herbaceous species) growing across distinct sites in three microenvironments of the *Cerrado* biome of Brasília-DF (Federal District), Central Brazil. All samples displayed virus-like symptoms. Locations are represented by colors: **a. Anthropic Areas:** Arborização Urbana do Distrito Federal - AUDF (purple), Núcleo Rural Taquara - NRT (teal blue), Viveiro II da Companhia Urbanizadora da Nova Capital - VII-NOVACAP (red), Fazenda Água Limpa da Universidade de Brasília - AE-FAL (blue), **b. Conservation Areas:** Estação Ecológica de Águas Emendadas - ESECAE (yellow) and

Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística - RECOR (orange) and **c. Transition Areas**: 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE - EEB) and 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE - EEB) (green).

2.2. Details of sample collections

All collections were performed along the year 2016 and in first quarter of the year 2017. Three surveys were performed at ESECAE and two surveys at RECOR. The criterion for sample selections was done in according with the presence of virus-like symptoms (**Supplementary Figure SF1 and SF2**). Two-hundred ninety-six (296) samples were collected being 78 from native tree species, 35from exotic tree species, 18 from agricultural species and, comprising the total of 49 botanical families in different areas. Samples collected per microenvironment and per areas were as follow: **(a) Areas with an environment with anthropic action:** AUDF (26 samples, including thirteen tree species and one weed), NRT (21 samples corresponding, two agricultural species and two weeds), Nursery of VII-NOVACAP (79 samples, corresponding to 29 tree species) and AE-FAL (50 samples, being 20 agricultural species); **(b) Conservation areas:** ESECAE (30 samples, including 18 tree species) and RECOR (30 samples being 20 tree species) and **(c) Transitional areas:** ARIE-EEB (29 samples with 20 of them being tree species) and AHE-EEB (31 samples from 22 plant species). We also considered ARIE-EEB as a transitional area because its close to AHE-EEB and subject to human action. Sampling criteria are described in **Supplementary Table (ST1)** (attached at the end of this paper).

For some plant samples were necessary identify the plant species by botanists and using *rbcl* primers in according to Fazekas et al. (2012).

2.3. Plant material processing and viral circular DNA enrichment

The symptoms were photographed, and leaves were cleaned with the aid of a brush. Total DNA was extracted individually from samples collected using the adapted CTAB protocol (Boiteux et al. 1999). Circular DNA enrichment was performed via Rolling Circle Amplification (RCA) individually (Inoue-Nagata et al. 2004) before to prepare the pools for HTS sequencing. DNA pools sent for HTS are present in **Table 1**.

For RCA reactions were performed using 10 μL final volume reaction containing: 0.1 μL Phi-29 DNA polymerase (10,000 U/mL), 1 μL of Phi-29 DNA polymerase enzyme buffer (NEB), 1 μL BSA (10X), 1 μL dNTPs (2.5 mM), 1 μL Thioprotected Primer (50 μM), 4.9 μL Milli-Q water and 1 μL extracted DNA (20 ng / μL).

The reaction was incubated at 30 ° C for 18 hours and at 65 ° C for 10 minutes to inactivate the enzyme. The five DNA pools were sent by collection site: ESECAE and RECOR; VII-NOVACAP, ARIE-EEB and AHE-EEB, AE-FAL and AUDF.

Table 1. DNA pools sending for High-Throughput Sequencing (HTS) from tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Samples were organized by microenvironment, site of collection, number of samples, and number of plant species.

Microenvironment	**Collection site (pool)**	**#samples***	**#plant species****
Anthropic areas	AUDF[1] and NRT[2]	47	20
	VII-NOVACAP[3]	79	29
	AE-FAL[4]	50	20
Conservation areas	ESECAE[5] and RECOR[6]	60	38
Transition areas	ARIE-EEB and AHE-EEB[7]	60	42

[1]Arborização Urbana do Distrito Federal (AUDF) and [2]Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB).*# of samples = Number of samples. **# of species = Number of species.

2.4. Metagenomic analyses and bioinformatics (sequence trimming, de novo assembly and annotation)

RCA products were sequenced by Illumina HiSeq 2500 platform (2 × 125 paired ends) at Macrogen Inc. (South Korea). Sequence analyzes were performed essentially as previously described (Kreuze et al. 2009, Adams and Fox 2016). The following steps were performed: (i) elimination of low quality sequences and adapters; (ii) reassembly of sequences and organization into contigs, using the CLC Genomics Workbench 10 program; (iii) validation of contigs by local alignments using BLASTx algorithm (Altschul et al. 1990) against a virus database (Viral RefSeq-GenBank); (iv) Afterward, the BLASTx results were compared to ViralRefSeq and another round of BLASTx analyses were performed with the general database. Contigs with higher coverages were selected; (v) the contigs showing identity to plant viruses after alignments were then analyzed with the assistance of the Geneious R11 program (Kearse et al. 2012). Contigs were mapped in order to obtain the complete genome sequences. Afterward, a BLASTn analysis was performed with the entire genome and the most similar sequence was used to transfer annotation. ORFs were confirmed using the Conserved Domain Architecture Retrieval Tool (CDART) and the ORF finder programs.

2.5. Primers design and molecular validation of viruses identified by High-throughput sequencing - HTS

Specific rimers for contigs recovered by HTS, were designed and used to detection virus in individual plants. Primers available in LVV-FITO-UnB (Laboratório de Virologia Vegetal – Fitopatologia UnB) were also used (**Supplementary Table ST2** - attached at the end of this paper).

PCR reactions were performed for a total volume of 12.5 µL containing 1.25 µL Taq Polymerase 10X Buffer (100 mM Tris-HCl, pH 8.3 and 500 mM KCl, Invitrogen); 0.38 µL MgCl2 (50 mM, Invitrogen); 0.25 µL dNTPs (2.5 mM, Invitrogen); 0.25 µL of each Primer

Foward and Reverse (10 μM); 0.1 μL of the enzyme Taq DNA Polymerase (5U / μL, Invitrogen); 9.02 μL MiliQ water and 1 μL DNA (which was diluted in MilliQ water 1:10).

PCR amplification reactions (from the RCA) in this step are described below: 1. denaturation - 94 ° C for 2 minutes and followed by 35 cycles each of 94 ° C for 30 seconds; 2. annealing (according to the primer used in the reaction, **Supplementary Table ST2**) for 55 seconds and 3. Extension - 72 ° C for 1 minute. PCR product visualized on 1% agarose gel stained with ethidium bromide.

The PCR product was cloned or sent for sequencing. Purification of DNA fragments obtained from PCR was performed according to the availability of two GFX (GE Healthcare) and Pure Link (Invitrogen) purification kits following manufacturers' guidelines. The generated amplicon was observed in 0.8% agarose gel electrophoresis, purified and cloned into pJET 1.2 plasmid (Thermo Fisher Scientific, USA) or pGEM T Easy vector (Promega). Sanger sequencing was performed at Myleus (Minas Gerais, Brazil) and Embrapa Hortaliças (Brasília, Brazil).

3. Results and Discussion

3.1 High performance sequencing: screening, assembly and analysis of contigs, and virus detection

The Illumina HiSeq 2500 sequencing platform was employed for the viroma study. The sequence reads were selected by quality in the CLC Genomics program and the adapters were removed. Subsequently, reads were used to assemble the contigs. Data from five libraries were described by collection site, number of samples, number of plants, number of reads, number of contigs, total number of complete viral genomes, number of putative new viruses,

and the number of previously characterized viruses. This complete set of information is provided in **Table 2**.

Table 2. Collection sites, numbers of reads, total number of contigs, number of contigs associated with DNA virus, total number of contigs viruses, number of putative new viruses and the number of previously characterized viruses detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil.

Collection site (pool)	**# reads**	**# contigs**	**# contigs associated with viruses**	**# total of contigs viruses**	**# new virus / viruses**	**# viruses previously Detected**
AUDF[1] and NRT[2]	32,257.21	100,205	149	8	4	4
VII-NOVACAP[3]	27,792.32	29,430	43	1	1	0
AE-FAL[4]	21,118.14	44,697	51	7	0	7
ESECAE[5] and RECOR[6]	31,164.53	63,842	92	4	3	1
ARIE-EEB and AHE-EEB[7]	29,849.94	54,923	168	10	4	6
Total	167,689.52	293,097	503	30	12	18

[1]Arborização Urbana do Distrito Federal (AUDF) and [2]Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB).*# = Number.

Viruses cause substantial losses in many fruit and vegetable crops. They may remain on wild and native plants, also semi-domesticated and weed species occurring nearby environments. In the present work we used HTS and metagenomic analyzes to study collections of symptomatic plants across microenvironments of the *Cerrado* biome. A total were analyzed 296 samples (131 species in 49 botanical families) distributed in anthropic areas (afforestation - AUDF, nursery - VII-NOVACAP and crops - AE-FAL and NRT) conservation areas (ESECAE and RECOR) and transition areas (ARIE-EEB and AHE-EEB). To study this *Cerrado* viroma, the Illumina HiSeq 2500 sequencing was performed. Based on BLASTx coverage, a total of 503 virus-associated contigs were detected. The anthropic areas with tree and crops species (AUDF and NRT) and the transitional (ARIE-EEB and AHE-EEB) areas displayed higher numbers of virus-associated contigs (149 and 168, respectively) when compared with anthropic areas with only crops (AE-FAL) areas (51). After contig mapping, 30 complete genomes - 24 species viral (distributed in the five libraries) were obtained with 12 of them representing new ssDNA virus species. Eighteen contigs were found to be isolates of previously reported ssDNA virus species (16) and dsDNA virus species (2).

The anthropic areas (AUDF and NRT) and the transitional (ARIE-EEB and AHE-EEB) areas showed higher amounts contigs of viruses, 10 and 8, respectively. In these microenvironments newer viruses were found, while the conservation areas only known species were detected. In terms of ssDNA and dsDNA virus diversity, members of the

following families were found: *Geminiviridae* (50%), *Genomoviridae* (20%), *Circoviridae* (13%), ssDNAs (10%) and *Caulimoviridae* (7%) (**Table 3**).

Table 3. Single stranded DNA and double stranded DNA viruses, organized by families and detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Genomic information an these viruses were obtained via High-Throughput Sequencing (HTS).

Collection site (pool)	***Geminiviridae***	***Genomoviridae***	***Circoviridae***	***Caulimoviridae***	**ssDNA unclassified**
AUDF[1] and NRT[2]	4	2	2	0	0
VII-NOVACAP[3]	0	0	0	0	1
AE-FAL[4]	7	0	0	0	0
ESECAE[5] and RECOR[6]	0	3	0	0	1
ARIE-AHE-EEB [7]	4	1	2	2	1
Total	15	6	4	2	3

[1]Arborização Urbana do Distrito Federal (AUDF) and [2]Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB)..

The results analyses of 39 contigs (30 genomes and 24 species viral) via BLASTx in Geneious R11 program and on considerering NCBI website are shown in **Table 4**, separated by collection sites, and taxonomy informations besides size full-genome, e-value and access number of GenBank.

Table 4. Single stranded DNA and double stranded DNA viruses detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* biome of Brasília-DF (Federal District), Central Brazil. Viruses were organizated considering collections sites, taxonomy and genomic organization and associated with the corresponding GenBank accessions. Analyses were carried out via BLASTx in Geneious R11 program and NCBI website.

Site	Viral Families	Virus Genus	24 species (39 contigs)	length (size - nts*)	Accession number in GenBank
AUDF[1] and NRT[2]	*Geminiviridae*	*Begomovirus*	*Bean golden mosaic virus* DNA-A	2617	AIN35848
			Bean golden mosaic virus DNA-B	2592	AYA73381
			Macroptilium bright yellow interveinal virus DNA-A	2611	YP005352908
			Macroptilium bright yellow interveinal virus DNA-B	2578	YP005352911
			Sida micrantha mosaic virus DNA-A	2676	AGH29615
			Sida micrantha mosaic virus DNA-B	2654	AGH29596
			Tomato severe rugose virus DNA-A	2593	APZ74282
			Tomato severe rugose virus DNA-B	2568	AGH29959
	Genomoviridae	*Gemykolovirus*	Tecoma stans associated gemykolovirus	2221	QCX29420
		Gemycircularvirus	Momordica charantia associated gemycirvcularvirus	2193	AXI69777
	Circoviridae	*Circovirus*	Bat circovirus	2230	YP009110680
		Cyclovirus	Mouse associated cyclovirus 1	2324	YP009315918

VII-NOVACAP[3]	unclassified	ssDNA unclassified	Caesalpinia pluviosa associated circular virus	2302	YP009551352
AE-FAL[4]	*Geminiviridae*	*Begomovirus*	*Bean golden mosaic virus* DNA-A	2617	AIN35848
			Bean golden mosaic virus DNA-B	2594	AYA73381
			Sida micrantha mosaic virus DNA-A	2675	AWX66243
			Sida micrantha mosaic virus DNA-B	2586	AGH29596
			Tomato severe rugose virus DNA-A	2593	APZ74282
			Tomato severe rugose virus DNA-B	2569	AGH29959
			Sweet potato leaf curl virus	2829	ABG90917
		Mastrevirus	Maize striate mosaic virus	2746	AST11833
		unclassified	Tomato apical leaf curl virus	2875	AZP54640
			Tomato associated geminivirus 1	2573	QFR15868
ESECAE[5] RECOR[6]	*Genomoviridae*	*Gemykrogvirus*	Gila monster-associated gemykrogvirus	2196	QCQ85257
		unclassified	Vochysia rufa associated genomovirus	1895	QCW23624
			Byrsonima crassiflora associated virus	1922	YP009021860
	unclassified	ssDNA unclassified	Lake Sarah-associated circular virus-36	2143	ALE29743
ARIE-EEB and AHE-EEB[7]	*Geminiviridae*	*Begomovirus*	*Bean golden mosaic virus* DNA-A	2618	AIN35848
			Bean golden mosaic virus DNA-B	2584	AYA73381
			Tomato severe rugose virus DNA-A	2596	ACO55139
			Tomato severe rugose virus DNA-B	2569	AGH29959
			Tomato mottle leaf curl virus	2631	AGH29930

		Mastrevirus	Maize striate mosaic virus	2746	AST11833
	Genomoviridae	*Gemykibivirus*	Ouratea duparquetiana associated gemykibivirus	2201	QCX29358
	Caulimoviridae **	*Badnavirus*	*Sugarcane bacilliform virus*	7467	AOV63232
		Caulimovirus	*Cauliflower mosaic virus*	8039	APG80689
	Circoviridae	*Circovirus*	Bat circovirus	1899	YP009110680
		Cyclovirus	Bat circovirus POA/2012/II	1712	AUM61764
	unclassified	ssDNA unclassified	Lake Sarah-associated circular virus-9	1168	AXH75115

[1]Arborização Urbana do Distrito Federal (AUDF) and [2] Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). *Nts = Nucleotides. ** Species classified in family *Caulimoviridae* show double strand DNA - dsDNA. In gray: six new species not associated with plants and showing low nucleotide identity between. Four of them were classified in *Circoviridae* family, showing o low identity with these viral species and two remain unclassified (probably CRESS DNA - more information in Duffy et al. 2019).

Considering all the microenvironments, it was possible to identify the diversity in ssDNA viruses found via HTS, metagenomic and bioinformatic analysis.

Considering the **anthropic areas**, 51 positive samples out of 15 plant species were obtained for at least one virus. These informations are presented below.

In AUDF and NRT pool were detected isolates of *Bean golden mosaic virus* - BGMV (DNA-A and DNA-B), *Tomato severe rugose virus* - ToSRV (DNA-A and DNA-B), *Sida micrantha mosaic virus* - SiMMV (DNA-A and DNA-B) and Macroptilium bright yellow interveinal virus - MBYIV (DNA-A and DNA-B) (*Begomovirus*/*Geminiviridae*) in different plants as illustrated in **Figure 2**. Two isolates of family *Genomoviridae* classified as Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) and as a new species Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) were detected in eggplant and *Tecoma stans* respectively. Two isolates of family *Circoviridae* were classified as two new species and showed identity of (39% and 56%) with Bat circovirus (*Circovirus*) and Mouse associated cyclovirus 1 (*Cyclovirus*). Primers have been synthesized and future studies will be performed to detection these circovirus in individually plants.

In VII-NOVACAP pool, Caesalpinia pluviosa associated circular virus - CpaCV, a new species of ssDNA was detected in *Caesalpinia pluviosa*. This virus is unclassified in family and may be considered CRESS DNA [more information in Duffy et al. 2019)].

In AE-FAL pool were detected isolates of BGMV (DNA A and B), ToSRV (DNA A and B), SiMMV (DNA-A and DNA-B) and the monopartite SPLCV (*Begomovirus/Geminiviridae*). Other species from *Geminiviridae* family were also detected including MSMV (*Mastrevirus*) and two capula-like species ToALCV and TaGV1 (not classified in genus yet).

For **conservation areas** ESECAE and RECOR, samples from three plant species were confirmed positive for different virus. In these areas genomovirus were detected: Gila

monster-associated gemycircularvirus - GmaV1 (*Gemykrogvirus/Genomoviridae*) and two new species proposed as Vochysia rufa associated genomovirus - VoaGmV and Byrsonima crassiflora associated genomovirus - ByaGmV (unclassified genus/*Genomoviridae*). Besides this, one isolate (CRESS DNA - More information in Duffy et al. 2019) was considered a new virus species by show low identity (12%) with Lake Sarah-associated circular virus-36 (unclassified ssDNA). Primers have been synthesized and future studies will be performed to virus-host detection.

Considering the **transition areas**, 15 samples from10 plant species were positive. Species of family *Caulimoviridae* with double strand DNA were found in this environment. In ARIE-EEB and AHE-EEB pool were detected isolates to: BGMV (DNA-A and DNA-B), ToSRV (DNA-A and DNA-B), ToMoLCV (*Begomovirus*/*Geminiviridae*), MSMV (*Mastrevirus*/*Geminiviridae*), *Sugarcane bacilliform virus* - SCBV (*Badnavirus/Caulimoviridae*), *Cauliflower mosaic virus* - CaMV (*Caulimovirus*/*Caulimoviridae*). Four isolates of news species were detected in this area: Ouratea duparquetiana associated gemykibivirus - OuaGmV (*Gemykibivirus/Genomoviridae*), two isolates showing low aminoacids identity with Bat circovirus *(Circovirus/Circoviridae/*) and Bat circovirus POA/2012/I (*Cyclovirus/Circoviridae*), and one isolate exhibiting low aminoacids identity with Lake Sarah-associated circular virus (unclassified ssDNA). Primers have been synthesized and future studies will be performed for circovirus-host detection.

Other ssDNA viruse sequences and also new alpha-satellite were recovered by HTS and are under analysis (data not shown).

Evaluating the scenario constituted in the anthropic environment by cultivation sites (NRT, AE-FAL), urban afforestation (AUDF) and nursery (VII-NOVACAP), notably, in cultivated areas there was a higher amount of detection of viral species (**Figure 2**).

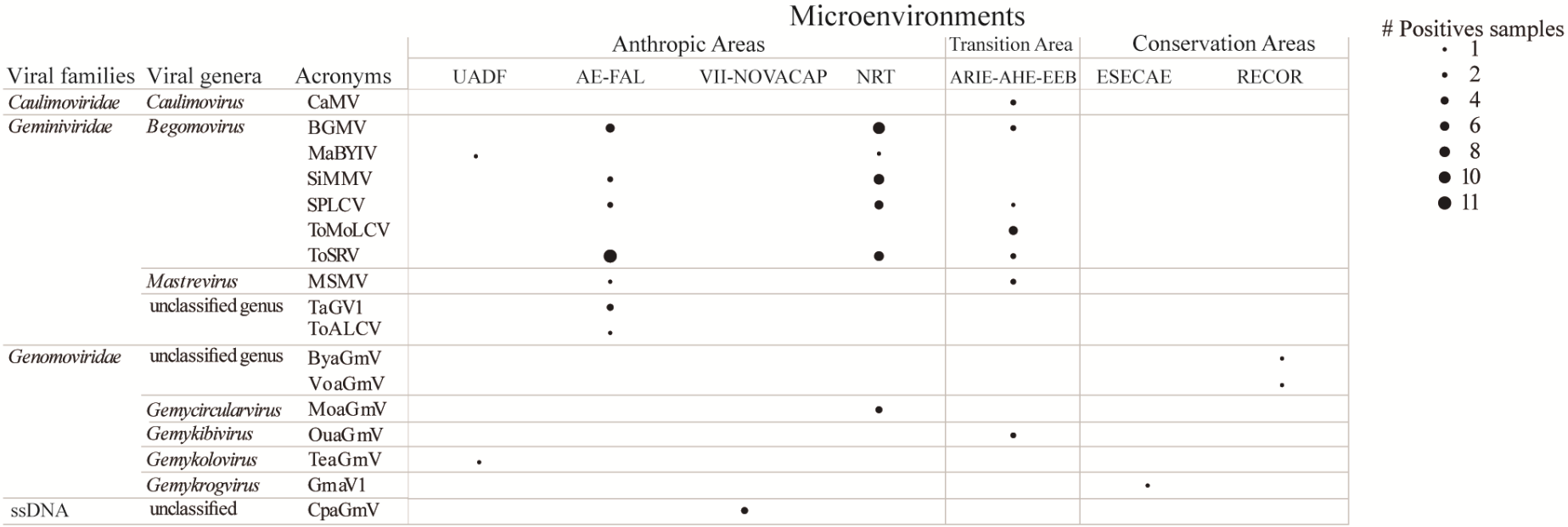


Figure 2. Distribution of viral species by collection sites. Microenvironments: Arborização Urbana do Distrito Federal (AUDF), Núcleo Rural Taquara (NRT), Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), Fazenda Água Limpa da Universidade de Brasília (AE-FAL), Estação Ecológica de Águas Emendadas (ESECAE), Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) and 2-Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). Acronyms: *Cauliflower mosaic virus* - CaMV (*Caulimovirus/Caulimoviridae*); *Bean golden mosaic virus* - BGMV (DNA-A), Macroptilium bright yellow interveinal virus - MaBYIV, *Sida micrantha mosaic virus* - SiMMV (DNA-A), *Sweet potato leaf curl virus* - SPLCV, *Tomato mottle leaf curl virus* - ToMoLCV, *Tomato severe rugose virus* - ToSRV (DNA-A), (*Begomovirus/Genomoviridae*) ; Maize striate mosaic virus - MSMV (*Mastrevirus/Genomoviridae*), Tomato associated geminivirus 1 - TaGV1, Tomato apical leaf curl virus - ToALCV (unclassified in genus/*Geminiviridae*); Vochysia rufa associated genomovirus - VoaGmV and Byrsonima crassiflora associated genomovirus - ByaGmV (unclassified in genus/*Genomoviridae*); Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus/Genomoviridae*); Ouratea duparquetiana associated gemykibivirus - OuaGmV (*Gemykibivirus/Genomoviridae*); Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus/Genomoviridae*); Gila monster associated gemykrogvirus - GmaV1 (*Gemykrogvirus/Genomoviridae*). One unclassified ssDNA species, Caesalpinia pluviosa associated circular virus - CpaCV..

In AUDF, two isolates of *Begomovirus* (BGMV and MaBYIV) were detected in *Macroptilium erythroloma* (Fabaceae). The first report of BGMV infecting *M. erythroloma* and biological transmission assays are described in Batista et al. (2020a). (**Attached I/paper entitled: Molecular Confirmation of *Macroptilium erythroloma* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil**).

The MaBYIV sequences for A and B component were recovered by HTS and showed the typical genomic organization of the New World begomoviruses. The complete DNA-A sequence exhibited paired identity less than 91% (considered threshold for novel species in *Begomovirus* genus). The biological transmission assays were performed confirming whitefly transmission. These results are described in Batista et al. (2020b). **(Attached II/paper entitled: Characterization in *Macroptilium erythroloma* of a Novel Neotropical Legume-Infecting Bipartite Begomovirus species.).**

In NRT, was possible to detected isolates of BGMV, ToSRV and MoaGmV in eggplant (*Solanum melongena* - Solanaceae). This is the first report of BGMV and MoaGmV in eggplant. Samples of weeds *Sida* sp. (Malvaceae) and *Nicandra physalodes* (Solanaceae) were positive for SiMMV and ToSRV. The sequence for *rbcl* (RuBisCO gene) will allowed to identify the species of *Sida*. Only one sweet potato plant was found to be positive for the SPLCV infection. These virus species already been described in these hosts.

Isolate of TeaGMV (*Gemykolovirus/Genomoviridae*) was detected only in *Tecoma stans* (Bignoniaceae) samples. Roossinck et al. (2010) described a virus-associated contig in a single sample from the Bignoniaceae family. However, without detection assays in individual plant species.

In VII-NOVACAP, isolates of CpaCV (ssDNA unclassified) were detected in two samples of *Handroanthus serratifolius* (Bignoniaceae) and one *Caesalpinia pluviosa* (Fabaceae) sample. No virus had been described on these hosts until the moment.

In AE-FAL were detected different species of *Geminiviridae* family. Isolates of: BGMV in bean (*Phaseolus vulgaris* - Fabaceae), ToSRV in tomato (*Solanum lycopersicum* - Solanaceae) and cotton (*Gossypium hirsutum* - Malvaceae), SiMMV in cotton, SPLCV in sweet potato (*Ipomoea batatas* - Convolvulaceae) and the recently described MSMV in maize (*Zea mays* - Poaceae), ToALCV in tomato and TaGV1 in beet (*Beta vulgaris* - Amaranthaceae) and tomato.

Here, we report by the first time a natural infection of cotton samples by ToSRV. In Brazil, the first and only report of bipartite begomovirus species in cotton, *Cotton chlorotic spot virus* (CCSV), was done in Paraíba, considering an isolate sharing 77.8% of nucleotide identity with ToCMV (DNA-A) and DNA-B sharing 67.8% of nucleotide identity with ToYVSV (de Almeida et al. 2013). In biological transmission assays using *Bemisia tabaci* MEAM 1, ToSRV and *Tomato golden vein virus* (TGVV) isolates failed to be transmitte for cotton and other hosts (Macedo et al. 2015). In AE-FAL cotton sample was infected with both virus ToSRV and SiMMV.

The newly described ToALCV species in Argentina (Vaghi Medina et al. 2017) was found for the first time in Brazil [Batista et al. (2018). **(Attached III/Paper entitled: First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil)].** Two isolates of TaGV1, exhibited a divergent sequence from those TaGV1 previously described by Fontenele et al (2017) were also detected. Besides, isolates of TAGV1 was found in a new host, the beet and biological tests were performed [Batista et al. (2020c). **(Attached IV/Paper entitled: Characterization of genetically divergent Tomato-associated geminivirus 1 isolates from table beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*)].**

The EEB is considered in this work a transitional environment due to the proximity between the cultivated areas (AHE-EEB) and conservation areas (ARIE-EEB). In this

microenvironment it is emphasized the presence of whitefly. The EEB has a preservation area near cultivated areas of vegetables and fruit that are showcases for students (Agronomy, Biologist and Forest Engineering) called AHE-EEB. Isolates of BGMV were detected in beans and "angico" tree species (*Anadenanthera colubrina* - Fabaceae). More details about first report of BGMV in *Anadenanthera colubrina* are described in Batista et al (2020d). **(Attached V/Paper entitled: First report of *Bean golden mosaic virus* in *Anadenanthera colubrina*).** In EEB we also detected ToSRV in tomato, SPLCV in sweet potato, ToMoLCV in tomato, wild poinsettia (*Euphorbia heterophylla* - Euphorbiaceae) and *Ouratea duparquetiana* (Ochnaceae). Mixed infections were observed in samples from this microenvironment. One sample of tomato was infected with ToSRV and ToMoLCV isolates.

On the other hand, Euphorbiaceae species are known begomovirus hosts (Fernandes et al. 2011) while no virus records were done for Ochnaceae species (Virus Host Database). Isolate of ToMoLCV is widely distributed in Brazil, mainly northeast region in tomatoes in (Vu et al. 2015). It has been described occurring in the states of BA, DF, ES, MG, RJ, PR, PE, PI and SP (Vu et al. 2015, Chaves et al. 2017). However, this is the first report of ToMoLCV (monopartite begomovirus) in *Euphorbia heterophylla* and *Ouratea duparquetiana* [(= *Campylospermum duparquetianum* (Baill.) Tiegh].

In general, *Begomovirus* species have a more limited number of hosts when in comparison with another important virus genera as *Orthotospovirus* and *Potyvirus*. However, the evolutionary mechanisms (high mutation and recombination rate) (Padidam et al. 1999) and the transit through its vector that has a large host spectrum more than 60 botanical families (De Barro et al. 2011) could be allow adaptation to new plant species. Among weeds the main hosts are species classified in three families: Malvaceae, Euphorbiaceae and

Fabaceae (Morales and Anderson 2001). However, Ochnaceae species can also be considered host of Begomovirus species as demonstrated here.

Recently MSMV (*Mastrevirus*) was described, infecting maize and leafhoppers. In this work, besides affecting maize, two isolates of MSMV were found in sugarcane (*Sugarcane officinarum*). Information on the first report of MSMV in sugarcane in Brazil can be found Batista et al. (2020e). **(Attached VI/Paper entitled: Natural Infection of Sugarcane (*Saccharum officinarum*) by Maize striate mosaic virus, Plant Disease Submitted).**

One isolate of a new species OuaGmV (*Gemykibivirus/Genomoviridae*) was detected in *O. duparquetiana.* This is the first report of a species in *Gemykibivirus* genus associated plant. In Brazil a species associated to humans has already been described (Phan et al. 2015).

Classified into family *Caulimoviridae*, were possible to detect two isolates of CaMV, in cabbage (*Brassica oleracea* Brassicaceae) and *Fevillea trilobata* (Cucurbitaceae). This is the first report of CaMV in *Fevillea*.

In conservation areas (RECOR) were detected the novel species of *Genomoviridae* family, possibly forming a new genus, ByaGMV and VoaGmV, in *Byrsonima crassiflora* (Malpighiaceae) and *Vochysia rufa* (Vochysiaceae), respectively. One isolate of GmaV1 (genus *Gemykrogvirus*) were detected in *Trembleya parviflora* (Melastomataceae) at ESECAE. Further details of the *Genomoviridae* viruses: ByaGmV, VoaGmV, MoaGmV, OuaGmV, TeaGmV and GmaV1 are described in **Chapter 3**.

A total of 64 samples (22%) (13 botanical families = 24.5%) showed at least one positive sample in our surveys (**Figure 3**). The individual results of the detections are in **Supplementary Table ST3**.

Botanical families

Viral families	Viral genera	Acronyms	Amaranthaceae	Anacardiaceae	Annonaceae	Apiaceae	Apocynaceae	Araliaceae	Asteraceae	Bignoniaceae	Bombacaeae	Brassicaceae	Calophyllaceae	Caricaceae	Caryocaraceae	Clusiaceae	Convolvulaceae	Curcubitaceae	Dioscoreaceae	Erythroxylaceae	Euphorbiaceae	Fabaceae	Hypericaceae	Lamiaceae	Lauraceae	Liliaceae	Magnoliaceae	Malpighiaceae	Malvaceae	Melastomataceae	Meliaceae	Moraceae	Moringaceae	Myrtaceae	Ochnaceae	Passifloraceae	Peraceae	Piperarceae	Poaceae	Proteaceae	Rosaceae	Rubiaceae	Rutaceae	Sapindaceae	Sapotaceae	Simaroubaceae	Solanaceae	Styracaceae	Urticaceae	Vebernaceae	Vochysiaceae
Caulimoviridae	*Caulimovirus*	CaMV										1						1																																	
Geminiviridae	*Begomovirus*	BGMV																				8																								3				
		MaBYIV																				1																												
		SiMMV																											4																	1				
		SPLCV															4																																	
		ToMoLCV																			1														1												3				
		ToSRV																											3																	16				
	Mastrevirus	MSMV																																					4												
	unclassified genus	TaGV1	1																																											2				
		ToALCV																																												1				
Genomoviridae	unclassified genus	ByaGmV																										1																						
		VoaGmV																																																1
	Gemycircularvirus	MoaGmV																																												1				
	Gemykibivirus	OuaGmV																																	1																
	Gemykolovirus	TeaGmV								1																																									
	Gemykrogvirus	GmaV1																												1																					
ssDNA	unclassified	CpaGmV								2												1																													

Figure 2 Single stranded DNA and double strand DNA viruses distribution detected in tree and herbaceous species from d Single stranded DNA viruses distribution in tree and herbaceous species growing across distinct sites in three microenvironments of the Cerrado area of Brasília-DF (Central Brazil) (Federal District), Brazil. Acronyms: *Cauliflower mosaic virus* - CaMV (*Caulimovirus/Caulimoviridae*); *Bean golden mosaic virus* - BGMV (DNA-A), Macroptilium bright yellow interveinal virus - MaBYIV, *Sida micrantha mosaic virus* - SiMMV (DNA-A), *Sweet potato leaf curl virus* - SPLCV, *Tomato mottle leaf curl virus* - ToMoLCV, *Tomato severe rugose virus* - ToSRV (DNA-A), (*Begomovirus/Geminiviridae*); Maize striate mosaic virus - MSMV (*Begomovirus/Geminiviridae*), Tomato associated geminivirus 1 - TaGV1, Tomato apical leaf curl virus - ToALCV (unclassified in genus/*Geminiviridae*); Vochysia rufa associated genomovirus - VoaGmV and Byrsonima crassiflora associated genomovirus - ByaGmV (unclassified in genus/ *Genomoviridae*); Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus/Genomoviridae*); Ouratea duparquetiana associated gemykibivirus - OuaGmV (*Gemykibivirus/Genomoviridae*); Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus/Genomoviridae*); Gila monster associated gemykrogvirus - GmaV1 (*Gemykrogvirus/Genomoviridae*). One unclassified ssDNA species, the Caesalpinia pluviosa associated circular virus - CpaCV.

The Solanaceae family which includes a large number of important vegetable crops [e.g. tomato, pepper (*Capsicum* sp) and eggplant] and the Fabaceae family with important vegetable and legumes [beans and soybeans (*Glycine max*)] (Dzoyem et al. 2014) are well-known hosts of virus species from *Geminiviridae* family. Here, plants of these families were the ones with the largest number samples as well as largest number of samples with positive detection to *Geminiviridae* members. In comparison, the virus incidence was higher in the Solanaceae family (60%).

Plant species of thirty-six families did not showed positive results for any virus of ssDNA recovery by HTS and Sanger. The families were: Amaranthaceae, Anacardiaceae, Annonaceae, Apiaceae, Apocynaceae, Araliaceae, Asteraceae, Bombacaceae, Caricaceae, Caryocaraceae, Clusiaceae, Calophyllaceae, Clusiaceae, Dioscoreaceae, Erythroxylaceae, Hypericaceae, Lamiaceae, Lauraceae, Liliaceae, Magnoliaceae, Meliaceae, Moraceae, Moringaceae, Myrtaceae, Passifloraceae, Peraceae, Piperaceae, Proteaceae, Rosaceae, Rubiaceae, Rutaceae, Sapindaceae, Sapotaceae, Styracaceae, Simaroubaceae, Urticaceae, Verbenaceae e. Some of them, encompassing important plant species for RNA viruses. The lack of positive results samples is, in fact, not a rare outcome. For example, in a geometagenomic study no virus was detected in families Ebenaceae, Proteaceae and Rhamnaceae collected in South Africa (Bernardo et al. 2018).

RNA of these samples (analysed here for DNA virus) were sending to sequencing (results not showed here).

Metagenomic approaches in samples from natural environments are showing thar a majority of viruses detected are novel reports as observed in studies carried out in native forests of Costa Rica (Roossinck et al. 2010) prairie plants in the United States (Muthukumar et al. 2009), in samples of agricultural and native species from France and South Africa (Bernardo et al. 2018). Here we obtained about 46% new virus species, a result close to that

found in France and South Africa where they also sampled native and agricultural species as well. Evaluating the prevalence of viruses in genomic libraries of agricultural species the number of novel viruses was higher than in genomic libraries composed by samples of native species. This result corroborates the one reported by Bernardo et al. (2017). All together, these results are reinforcing the hypothesis that relates host abundance to pathogen prevalence. In this hypothesis, increasing the number of hosts also will increase the pathogen prevalence (Agrawal et al. 2006, Keesing et al. 2010), similar to what occurs in monocultures.

Notably, sequences like those already available at GenBank database were the easiest to detect and annotate. GenBank data show a higher number of virus sequences related to agricultural species, which may be related to the amount of virus found in native species lower than those of cultivated species. This fact is an indirect evidence of how divergent the ssDNA viruses in native species can be. The large amount of the obtained contigs with no hit with previous ssDNA viruses might be related to species that were not yet discovered. In fact, so far studies have long focused on commercially important agricultural and ornamental species (Lefeuvre et al. 2019). Difficulties in finding virus-related sequences in other native environments and wild species have been reported elsewhere (Rosario and Breitbart 2011, Brum et al. 2016, Bernardo et al. 2018). In general, about 70% of sequences generated in viral metagenonic studies do not have identity sequence to the ones currently available in public databases (Rosario and Breitbart 2011). Bernardo et al. (2017) reports that 30.9% of contigs had no homology with any GenBank sequence. According to the methodologies used in the present work, only about 0.17% of the contigs were detected as having identity to plant and plant-associated viruses. Another bias in metagenomic studies is the structural integrity of the nucleic acids of all samples. Despite anatomical and physiological differences, every effort was made to ensure equal conditions for isolation of nucleic acids from both cultivated and native species.

The ocurrence of *Genomoviridae* and *Geminiviridae* species in native and cultivated plants suggests that these viruses are transiting across different environments. Although the *Genomoviridae* family has a higher number of viruses associated fungi and animals, new findings indicate their widespread presence in plants. For instance, the accepted species of *Genomoviridae*:*Cassava associated gemycircularvirus* 1 (Dayaram et al. 2012); *Hypericum japonicum associated gemycircularvirus* 1 (Du et al. 2014); *Poaceae associated gemycircularvirus* 1 (Male et al. 2015); *Bromus associated gemycircularvirus 1* (Kraberger et al. , 2015) and *Soybean associated gemycircularvirus 1* (Marzano and Domier, 2016). In Brazil, novel genomovirus species have been characterized in beans (Lamas et al. 2016), *Momordica charantia* (Cucurbitaceae) and *Euphorbia heterophylla* (Euphorbiaceae) (de Rezende et al. 2018).

Metagenomic studies are revealing a large diversity of geminiviruses in woody plants such as citrus (*Citrus* sp.) (Lu et al. 2015), apples (*Malus domestica*) (Liang et al. 2015a), mulberry (*Morus* sp.) (Loconsole et al. 2012), grapevine (*Vitis* sp.) (Al Rwahnih et al. 2017) and also in herbaceous weeds (Rodríguez-Negrete et al. 2019).

Considering dsDNA, CaMV is one of the ten most important viruses (Scholthof et al. 2011), affecting plants of Brassicaceae and Solanaceae family (Qiu and Schoelz 1992, Piqué et al. 1995, Yasaka et al. , 2014).

Found in a smaller number (four), *Circoviridae* viruses, are known to affect animals. However, recent work describes that an animal virus (Providence virus - PrV) was tested on bean plants and was able to replicate (Jiwaji et al. 2019).

New unclassified ssDNA sequences have been discovered and are being considered as CRESS DNA (Duffy et al 2019). Here we detect two sequences of unclassified ssDNA. In Brazil a new virus of ssDNA has been identified in apple, pear and vines and Temperate fruit decay-associated virus (TFDaV) species was proposed (Basso et al. 2015).

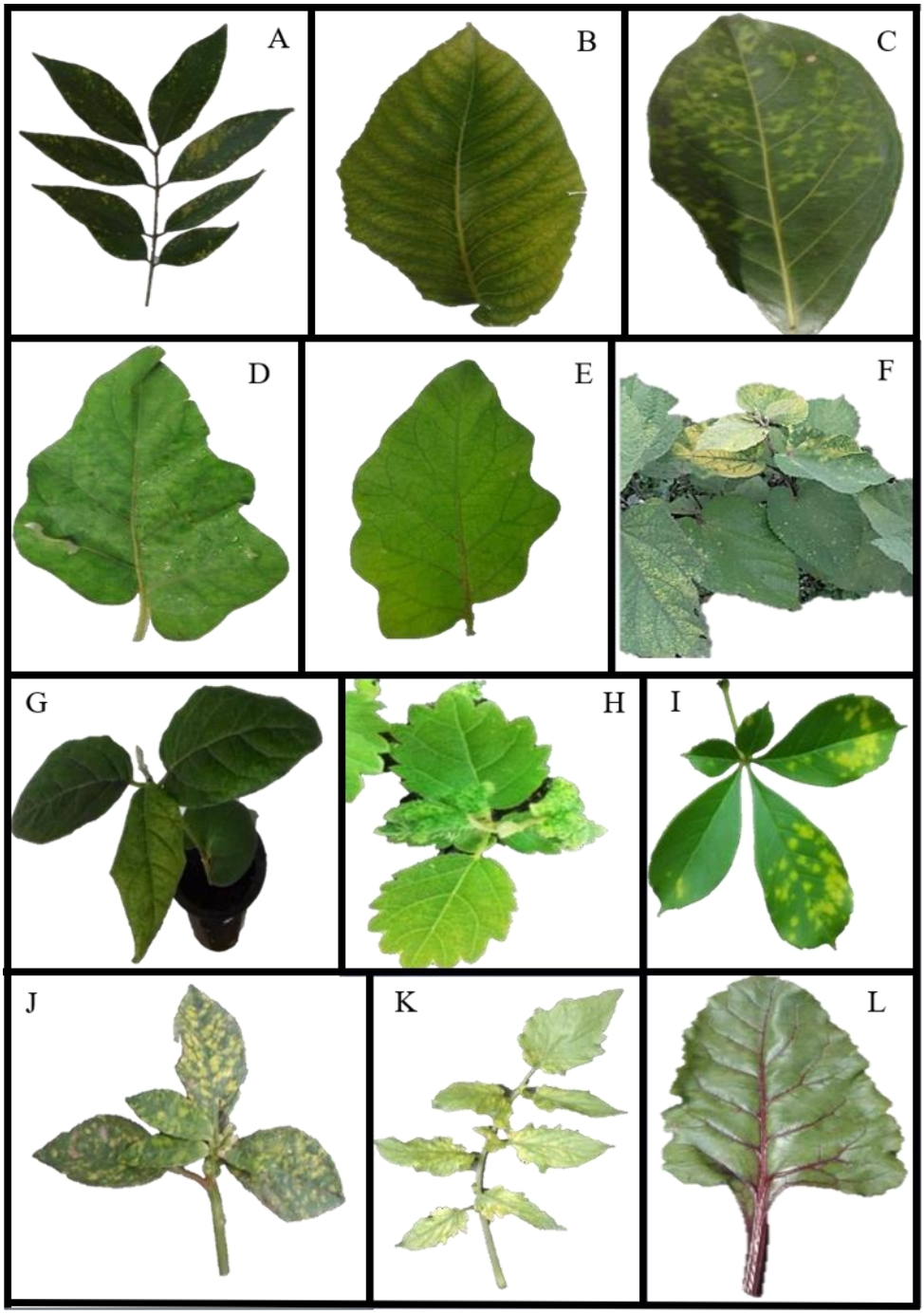
Native plants may act as a virus researchervoir (Cooper and Jones 2006) or alternatively, infections occurring in cultivated species may cause damage to native species (Jones 2009, Alexander et al. 2014, Jones and Coutts 2015). In addition to viral accumulation over time, it causes declines in density and distribution of native species (Flory and Clay 2013, Vincent et al. 2014). Increasing evidence indicates that there is a wide diversity of plant virus species in natural ecosystems. Results from a study in Iberian Peninsula (Portugal and Spain) showed isolate of *Potyvirus* (family *Potyviridae*) has always been present and prevalent in oak forests (Rodríguez-Negrete et al. 2019). Virus ecology seeks to elucidate the virus-plant interactions of the possible benefits or harms that new viruses cause. Viruses are known to play a role in preventing the overgrowth of homogeneous plant populations, as monocultures, promoting the adaptation of their hosts to environmental changes (Lefeuvre et al. 2019). On the other hand, the emergence of new diseases occurs after changes in the host ranges of the virus, which is driven by adaptive viral evolution in response to new ecological conditions (Jones 2009). These new ecological conditions include the introduction of viruses and vectors into new areas, intensification of agriculture, urbanization, ecological changes and changes in climatic conditions (Pagán et al. 2012, Roossinck and García-Arenal 2015).

4. Conclusion

The *Cerrado* is a Neotropical biome with an endemic occurrence of over 4000 higher plants and it is also the major region for field crop production in Brazil. In total, 30 entire viral genome (24 species) sequences were recovered via HTS, corresponding to 12 new species and 12 species that already have been described. Seven new first report of ssDNA virus on cultivated (4), tree species (2) and weeds (1) hosts were done. For dsDNA a first report on cultivated. The positive ssDNA viruses were higher in the anthropic and transition microenvironments, with a lower prevalence in the conservation microenvironments. In the

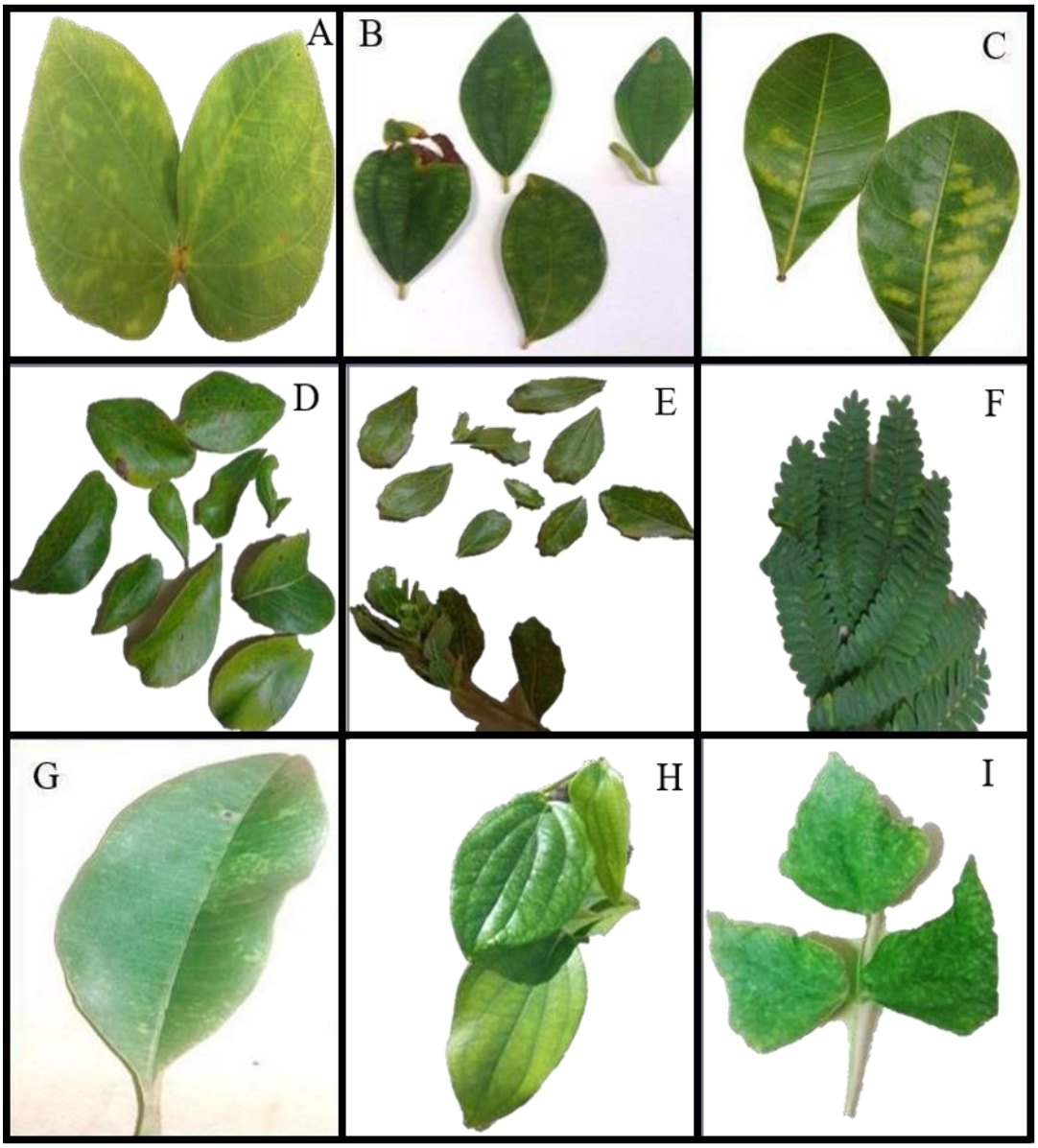
transitional environment, the proximity between cultivated and native species may favor the viruses traffic. Therefore, a metagenomic approach coupled with HTS was able to provide fresh insights into the diversity of plant-associated viruses in different environments of the *Cerrado* biome. Our metagenomic analyses indicated a vast cryptic diversity of ssDNA viruses in native and exotic trees as well as field and vegetable crops growing in this highly threatened biome. This large-scale documentation of viral species associated with plants present in this biome will be important in assisting in the development of surveillance systems of novel viral species and their potential transit across environments.

Supplementary Figure 1 - SF1.



Supplementary Figure 1 SF1. Symptomatic plants collected in Anthropic Areas - Arborização Urbana do Distrito Federal (AUDF): **A.** *Inga cylindrica*, **B.** *Caryocar brasiliense* and **C.** *Inga laurina*; Núcleo Rural Taquara (NRT): **D.** and **E.** *Solanum melongena* and **F.** *Sida* sp.; Viveiro II of Companhia Urbanizadora da Nova Capital (VII-NOVACAP): **G.** *Aspidosperma polyberon*, **H.** *Tabebuia chrysotricha* and **I.** *Cybistax antisyphilitica*; Fazenda Água Limpa da Universidade de Brasília (AE-FAL): **J.** *Gossypium hirsutum*, **K.** *Solanum lycopersicum* and **L.** *Beta vulgaris*.

Supplementary Figure 2 - SF2.



Supplementary Figure 2 SF2. Symptomatic plants collected in Conservation Areas - Estação Ecológica de Águas Emendadas (ESECAE): **A.** *Bauhinia rufa*, **B.** *Miconia* sp. and **C.** *Anacardium humile*; Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR): **D.** *Byrsonima crassiflora*, **E.** *Baccharis tridentata* and **F.** *Calliandra dysantha*; 1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB). **G.** *Manilkara zapota*, **H. Piper** *nigrum* and **I.** *Phaseolus vulgaris*.

Supplementary Table 1 - ST1

Table 1 ST1. Distribution of plants by site, family, scientific name, number of samples for each symptom presented, and description of symptoms presented at the time of collection.

Site	Family	Scientific name	N°* samples	Symptoms**
AUDF[1]	Bignoniaceae	*Tecoma stans*	1	CS
		Jacaranda decurrens	1	PC
		Tabebuia chrysotricha	1	B
	Caryocaraceae	*Caryocar brasiliense*	3	Y; Mo
	Fabaceae	*Inga laurina*	5	Mo
		Inga marginata	2	Y
		Cassia podalyriifolia	1	PC
		Caesalpinia echinata	1	Y
		Clitoria fairchildiana	1	M
		Bowdichia virgilioides	1	PC
		Macroptilium erythroloma	1	Y
	Malpighiaceae	*Malpighia glabra*	2	PC
	Moraceae	*Morus nigra*	2	M
	Passifloracea	*Passiflora* sp.	4	M
NRT[2]	Solanaceae	*Solanum melongena*	17	M
		Nicandra physalodes	2	M
	Malvaceae	*Sida* sp.	1	M
	Convolvulaceae	*Ipomoea batatas*	1	M
VII-NOVACAP[3]	Anacardiaceae	*Schinus terebinthifolius*	1	Mo
			1	LC
	Apocynaceae	*Aspidosperma polyberon*	6	Mo; LC
	Bignoniaceae	*Tecoma stans*	2	Mo
		Tabebuia chrysotricha	4	M
			4	LC
			1	PC
		Handroanthus serratifolius	3	EP

			2	M
			1	S
			2	CS
			2	LC; PC
		Cybistax antisyphilitica	3	CS
		Handroanthus heptaphyllus	2	M
		Jacaranda puberula	2	LC; PC
	Bombacaceae	*Pseudobombax grandiflorum*	6	Y
	Clusiaceae	*Calophyllum brasiliense*	1	M
	Fabaceae	*Copaifera langsdorfii*	2	PC
		Amburana cearensis	1	M
		Anadenanthera macrocarpa	7	M
		Caesalpinia pluviosa	1	M
		Myroxylon balsamum	1	PC
		Copaifera sp.	1	PC
		Hymenaeae stignocarpa	2	CS
		Platymiscium floribundum	2	M
		Parquia leguminosae	1	M
		Hymenaeae courbaril	2	CS
	Melastomataceae	*Tibouchina granulosa*	1	M
	Meliaceae	*Cedrela fissilis*	2	Y
	Moringaceae	*Moringa oleífera*	6	Y
	Myrtaceae	*Eugenia tomentosa*	1	PC
	Rutaceae	*Esenbeckia leiocarpa*	1	M
	Solanaceae	*Solanum lycocarpum*	1	CS
		Solanum americanum	2	M
	Urticaceae	*Cecropia pachystachya*	1	PC
	Verbenaceae	*Cytharexyllum myrianthum*	1	Y
AE-FAL[4]	Amaranthaceae	*Chenopodium quinoa*	1	M
		Beta vulgaris	5	M

		Spinacia oleracea	1	M
	Apiaceae	*Daucus carota*	1	M
	Asteraceae	*Lactuca sativa*	1	Mo
	Brassicaceae	*Raphanus sativus*	1	M
	Convolvulaceae	*Ipomoea batatas*	1	M; PN
			1	M; B
			2	M
			1	M; CV
	Curcubitaceae	*Curcubita pepo var. melopepo*	1	M
		Curcubita pepo	1	M; B
	Dioscoreaceae	*Dioscorea cayenensis*	1	M
	Fabaceae	*Phaseolus vulgaris*	2	M; B
			1	M
			1	B; Y
			1	B
			1	Y; B; D
		Crotalaria juncea	1	M
	Liliaceae	*Allium sativum*	1	Y
	Malvacea	*Gossipium hirsutum*	3	PC
	Poaceae	*Pennisetum americanum*	1	Y
		Saccharum officinarum	1	St
		Zea mays	1	Y; St
	Rubiaceae	*Coffea arabica*	1	NP
	Solanaceae	*Solanum lycopersicum*	11	LC; Y
			4	LC; Y; NL
		Capsicum annuum	2	M
ESECAE[5]	Anacardiaceae	*Anacardium humile*	1	NP
	Araliaceae	*Sheflera vinosa*	1	M
	Asteraceae	*Baccharis tridentata*	2	PC
		Piptocarpha rotundifolia	1	PC

	Caryocaraceae	*Caryocar brasiliense*	2	NL
			2	Y
			1	B
			1	PC
	Fabaceae	*Bauhinia* sp.	1	PC
		Dalbergia miscolobium	1	Y
	Lamiaceae	*Aegiphila verticillata*	1	M
	Malpighiaceae	*Banisteria argyrophylla*	1	PC
		Byrsonima sp.	1	PC
	Melastomataceae	*Trembeya parviflora*	2	PC
			1	NP
		Miconia albicans	1	M
		Miconia sp.	2	M
	Proteaceae	*Roupala montana*	1	Y
	Rubiaceae	*Palicorea rigida*	2	Y
	Simaroubaceae	*Simarouba versicolor*	2	CV
	Styracaceae	*Styrax ferrugineus*	2	Mo
		Lipia sp.	1	M
RECOR[6]	Anacardiaceae	*Tapirira guianensis*	1	PC
	Asteraceae	*Baccharis tridentata*	2	PC; B; Y
	Caryocaraceae	*Caryocar brasiliense*	1	CS
	Erythroxylaceae	*Erytroxilum suberosum*	2	CS
	Fabaceae	*Stryphnodendron adstringens*	2	PC; Mo
		Mimosa sp.	1	Y
		Calliandra dysantha	1	Y
		Bauhnia rufa	2	PC
	Hypericaceae	*Vismia sp.*	1	M
	Lamiaceae	*Aegiphilla sellowiana*	1	CS
	Malpighiaceae	*Byrsonima crassiflora*	1	EP; NL
	Melastomataceae	*Leandra melastomoides*	2	PC

	Myrtaceae	*Myrcia tomentosa*	1	CS
	Passifloraceae	*Passiflora cincinatta*	2	PC
	Peraceae	*Pera glabrata*	2	PC; NL
	Poaceae	*Bambusa vulgaris*	2	Y
	Rosaceae	*Rubus sp.*	1	M
	Solanaceae	*Solanum lycocarpum*	2	M
	Vochysiaceae	*Vochysia rufa*	2	CV; Y
		Qualea multiflora	1	Y
AHE-EEB[7]	Asteraceae	*Lactuca sativa*	1	Mo
	Brassicaceae	*Brassica oleracea*	1	M
	Caricaceae	*Carica papaya*	1	Y
	Convolvulaceae	*Ipomoea batatas*	1	Y
	Cucurbitaceae	*Curcubita moschata*	1	CS
		Luffa cyllindrica	2	PC
		Fevillea trilobata	1	CS
	Euphorbiaceae	*Manihot esculenta*	2	PC; M
		Euphorbia heterophylla	1	LC; M; D
	Fabaceae	*Glycine max*	1	M
		Arachis hypogaea	1	PC
		Phaseolus vulgaris	2	B; M
		Cajanus cajan	2	Y
	Liliaceae	*Allium sativum*	1	Y
	Passifloraceae	*Passiflora edulis*	1	B
	Piperaceae	*Piper nigrum*	3	M; PC
	Poaceae	*Zea mays*	1	Y
		Saccharum officinarum	1	Y
	Rutaceae	*Citrus limon*	1	M
	Solanaceae	*Capsicum annuum*	2	M
		Capsicum frutescens	2	M
		Solanum lycopersicum	2	LC; M

ARIE-EEB[7]	Anacardiaceae	*Anacardium humile*	1	Y
		Haplopappus laricifolius	3	Y
	Annonaceae	*Trigynaea duckei*	1	M
	Calophyllaceae	*Kielmeyera coriacea*	2	M
	Fabaceae	*Dalbergia sp.*	2	CS
		Piptadenia sp.	1	M
		Enterolobium cyclocarpum	2	PC
		Plathymenia foliolosa	1	Mo
		Pterodon sp	1	Y
		Peltophorum dubium	2	PC
		Dalbergia nigra	1	M
	Lamiaceae	*Aegiphylla sp.*	1	M
	Lauraceae	*Endlicheria pyriformis*	2	PC
	Magnoliaceae	*Magnolia grandiflora*	1	M
	Melastomataceae	*Miconia albicans*	1	M
		Miconia sp.	1	M
	Ochnaceae	*Ouratea duparquetiana****	2	PC
	Rutaceae	*Pilocarpus jaborandi*	2	C
	Sapindaceae	*Litchi chinesis*	1	M
	Sapotaceae	*Manilkara zapota*	1	M

[1]Arborização Urbana do Distrito Federal (AUDF) and [2]Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB).*N°= number. **Symptoms: Y = yellowing; B = blistering C = chlorosis; CV = chlorosis vein; LC = leaf curl; EP = epinasty; NL = necrotic lesions; CS = chlorotic spots; M = mosaic; Mo = mottle; S = stunting, N = necrosis; PC = chlorotic; NP = necrotic points ans St=striate. *** *Ouratea duparquetiana* (= Synonymia *Campylospermum duparquetianum* (Baill.) Tiegh.)

Supplementary Table 2 - ST2

Table 2 ST2. Oligonucleotides designed based on sequences obtained from High-Throughput Sequencing (HTS) and available in LVV-FITO UnB (Laboratório de Virologia Vegetal – Fitopatologia UnB) and PCR conditions.

Family	Genus	Species (acronym) - component	Primer Name	Sense (Foward) 5'-3' / Antisense (Reverse) 5'-3'	Ta[1]	Expected Amplicon	References
Geminiviridae	*Begomovirus*	*Bean golden mosaic virus* (BGMV) - DNA A	1BGMV_A_For/ 1BGMV_A_Rev	GTGCGTGAATCCATGACCGT/ ATTCACGCACAGGGGAACG	55	2600	HTS
		Bean golden mosaic virus (BGMV) - DNA B	Primer_BF/ Primer_BR	GGGTCYTTRAGSGAGAAR/ TGKKTYTCATGWAKMCGMTG	55	2600	HTS
		Tomato severe rugose virus (ToSRV) - DNA A	ToSRV-A_S_For/ ToSRV-A_S_ReV	GATTCTCGTATTTCCCTGCCTCC/ AGAATCATACTGAGAACGCCTTGC	60	2600	HTS
		Tomato severe rugose virus (ToSRV) - DNA B	ToSRV-B_S_For/ ToSRV-B_S_Rev	GACAGAAGAATGACGGACAATGAATC/ CATGAATTTCGACTATAACTGAACCTACG	60	2600	HTS
		Sida micrantha mosaic virus (SiMMV) - DNA A	SiMMV_BamHIF/ SiMMV_BamHIR	GGATCCCTCATGGCGCCAGATG/ CGAAATGCCCAAGCGGGATCC	59	2600	HTS
		Tomato mottle leaf curl virus (ToMoLCV)	ToMoLCV-A_S_For/ ToMoLCV-A_S_Rev	GTTGCCCATCTTCGTGTAGTTCT/ CAACCTCATCTCCACGTGCTC	60	2600	HTS
		Sweet potato leaf curl virus (SPLCV)	MA292/ MA293	CCYTAGGGTTCGAGCTVTGTTCGG/ TTTATTAATTDTTRTGCGAATC	48	823	Lozano et al. , 2009

	Mastrevirus	Maize mosaic streak virus (MSMV)	J455-F/ J455-R	ACCCTTCTTAACTTCCACCACGGCAGAA/ GGTAATTGTCTGATGGTTACCTCCTACA	60	2700	Fontenele et al. , 2018
	Unclassified	Tomato associated geminivirus 1 (TaGV1)	Cap2KpnI-F/ Cap2KpnI-R	GGTACCCCCCTTGGAAATGTAGTCTGCAAC/ GGTACCTTTGAGGAGAGAGGTATACTTCG	66	2600	HTS
		Tomato apical leaf curl virus (ToALCV)	Cap1PstI-F/ Cap1PstI-R	CTGCAGAYTTGCGCGGATCGATTAAT/ CTGCAGAAATGCGTTGTAACTTCTCGGATAT	68	2900	HTS
Genomoviridae	Unclassified	Eucalyptus urophylla associated genomovirus (EcaGmV)	F1_ EUaVPINT/ EUaVPINT _R1	GTGTCCCACTGATAACG/ CCAGAGAGCGAGACAGA	55	1886	HTS
		Vochysia rufa associated genomovirus (VoaGmV)	15_6033F/ 15_6033R	GAATAGGCCTCGAACGAC/ GCATTTTACCAGCGCCTAC	60	1895	HTS
		Byrsonima crassiflora associated genomovirus (ByaGmV)	15_5347F/ 15_5347R	CGTTTATCTTTGCGCTTCG/ ACGAGTCAACCAGTTCCGTC GGA	60	1922	HTS
	Gemykrogvirus	Gila monster-associated gemykrogvirus (GmaV1)	CaribouEcoR1-F/ CaribouEcoR1-R	GAATTCGTCCATGTTGAATTG/ GACGAATTCTCCGACACTTGT	59	2196	HTS

	Gemykibivirus	Ouratea duparquetiana associated gemykibivirus (OuaGmV)	7_6560F/ 7_6560R	CTACGATAAGACCCGCCACC/ AGTGGACCCTAATACGCTTG	60	2201	HTS
	Gemykolovirus	Tecoma stans associated gemykolovirus (TeaGmV)	TickPst1-F/ TickPst1-R	CTGCAGGTAATGAGCAGGGT/ CTGCAGCTAATGTAGTAGTC	52	2200	HTS
	Gemycircularvirus	Momordica charantia associated gemycircularvirus (MoaGmV)	11_205F/ 11_205R	GACCACCTCTGATGTCGT/ GGTCTTGGATTCTTTCATGG	60	2193	HTS
ssDNA	Unclassified	Caesalpinia pluviosa associated gemycircularvirus (CpaCV)	13_3705F/ 13_3705R	AGAACGGTTGGGCGTCAA/ CTCGCGTAGGTTGACTGG	60	2302	HTS
Caulimoviridae	*Caulimovirus*	*Cauliflower mosaic virus* (CaMV)	CaMV-F/ CaMV-R	GAAGGCCATTACGCCAACGAATGTCCT/ ATGGATTTTCTGAACATACT	48.5	840	Pappu & Druffel, 2009

[1]Ta = annealing temperature.

Supplementary Table 3- ST3.

Tabela ST3. Virus distribution by site and number of positive samples and total number of samples.

Site	Species – acronym	Species NPS*/NTS**
AUDF[1]	*Bean golden mosaic virus* – BGMV	*Macroptilium erythroloma* 1/1
	Tecoma stans associated gemycircularvirus -TeaGmV	*Tecoma stans* 1/1
	Macroptilium bright yellow interveinal virus - MaBYIV	*Macroptilium erythroloma* 1/1
NRT[2]	*Bean golden mosaic virus* – BGMV	*Solanum melongena* 3/17
	Tomato severe rugose virus - ToSRV	*Sida* sp. 1/1; *Nicandra physalodes* 1/2; *Solanum melongena* 11/17
	Sweet potato leaf curl virus - SPLCV	*Ipomoea batatas* 1/1
	Sida micrantha mosaic virus - SiMMV	*Sida* sp. 1/1; *Nicandra physalodes*1/2
	Momordica charantia associated gemycircularvirus – MoaGmV	*Solanum melongena* 1/17
FAL[3]	*Bean golden mosaic virus* – BGMV	*Phaseolus vulgaris* 5/5
	Tomato severe rugose virus - ToSRV	*Solanum lycopersicum* 9/15; *Gossypium hirsutum* 2/3
	Sweet potato leaf curl virus - SPLCV	*Ipomoea batatas* 2/5
	Sida micrantha mosaic virus - SiMMV	*Gossypium hirsutum* 2/3
	Maize striate mosaic virus - MSMV	*Zea mays* 1/1
	Tomato apical leaf curl virus - ToALCV	*Solanum lycopersicum* 1/15
	Tomato associated geminivirus - TaGV1	*Solanum lycopersicum* 2/15; *Beta vulgaris* 1/6
VII-NOVACAP[4]	Caesalpinia pluviosa associated circular virus - CpaCV	*Handroanthus serratifolius* 2/10; *Caesalpinia pluviosa* 1/1
ESECAE[5]	Gila monster associated gemykrogvirus - GmaV1	*Trembleya parviflora* 1/1
RECOR[6]	Vochysia rufa associated genomovirus - VoaGmV	*Vochysia rufa* 1/1
	Byrsonima crassiflora associated genomovirus - ByaGmV	*Byrsonima crassiflora* 1/1
EEB[7]	*Bean golden mosaic virus* – BGMV	*Phaseolus vulgaris* 2/2; *Anadenanthera macrocarpa* 1/1
	Tomato severe rugose virus - ToSRV	*Solanum lycopersicum* 2/4
	Sweet potato leaf curl virus - SPLCV	*Ipomoea batatas* 1/1

	Maize striate mosaic virus - MSMV	*Zea mays* 2/4; *Saccharum officinarum*2/6
	Tomato mottle leaf curl virus - ToMoLCV	*Solanum lycopersicum* 3/4; *Euphorbia heterophylla* 1/1; *Ouratea duparquetiana* 1/2
	Ouratea duparquetiana associated gemycircularvirus – OuaGmV	*Ouratea duparquetiana* 1/2
	Cauliflower mosaic virus – CaMV	*Brassica oleracea* 1/1; *Fevillea trilobata* 1/1

[1]Arborização Urbana do Distrito Federal (AUDF) and [2]Núcleo Rural Taquara (NRT), [3]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [4]Fazenda Água Limpa da Universidade de Brasília (AE-FAL), [5]Estação Ecológica de Águas Emendadas (ESECAE), [6]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR), [7]1-"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) e 2- Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB).* NPS= number of positives samples, NTS= number of total samples.

References

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.

Agrawal AA, Lau JA, Hambäck PA (2006) Community heterogeneity and the evolution of interactions between plants and insect herbivores. The Quarterly Review of Biology 81:349–376.

Al Rwahnih M, Alabi OJ, Westrick NM, et al (2017) Description of a Novel Monopartite Geminivirus and Its Defective Subviral Genome in Grapevine. Phytopathology 107:240–251. doi: 10.1094/PHYTO-07-16-0282-R.

Alexander HM, Mauck KE, Whitfield AE, et al (2014) Plant-virus interactions and the agro-ecological interface. The European Journal of Plant Pathology 138:529–547. doi: 10.1007/s10658-013-0317-1.

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403–410.

Basso MF, da Silva JCF, Fajardo TVM, et al (2015) A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. Virus Research 210:27–33.

Batista JG (2014) Ocorrência natural de espécies virais e avaliação de potenciais hospedeiras experime ntais de begomovírus, potyvírus e tospovírus em plantas arbóreas e berinjela (Solanum melongena). Universidade de Brasilia.

Batista JG, Melo FL, Pereira-Carvalho RC, et al (2018) First Report of *Tomato apical leaf curl virus* Infecting Tomato in Brazil. Plant Disease.

Bernardo P, Charles-Dominique T, Barakat M, et al (2018) Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. EMBO J 12:173.

Boiteux LS, Fonseca MEN, Simon PW (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.

Brown JK, Zerbini FM, Navas-Castillo J, et al (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Archives of Virology160:1593–1619.

Brum JR, Ignacio-Espinoza JC, Kim E-H, et al (2016) Illuminating structural proteins in viral "dark matter" with metaproteomics. Proceedings of the National Academy of Sciences 113:2436–2441.

Chaves ALR, Eiras M, Rodrigues LK, et al (2017) First report of *Tomato mottle leaf curl virus* infecting tomato in São Paulo State, Brazil. Summa Phytopathol 43:353.

Cooper I, Jones RAC (2006) Wild plants and viruses: under-investigated ecosystems. Advances in Virus Research 67:1–47.

Dayaram A, Opong A, Jäschke A, et al (2012) Molecular characterisation of a novel cassava associated circular ssDNA virus. Virus Research166:130–135. doi: 10.1016/j. Virusres. 2012.03.009.

de Almeida MMS, Jain S, Barroso PAV, et al (2013) Complete sequence of a new bipartite begomovirus infecting cotton plants in Brazil. Genome Announcements 1:e00661-13.

De Barro PJ, Liu S-S, Boykin LM, Dinsdale AB (2011) Bemisia tabaci: a statement of species status. Annual Review of Entomology 56:1–19.

de Rezende RR, Mar TB, Páez LMC, et al (2018) Complete genome sequences of two gemycircularviruses associated with non-cultivated plants in Brazil. Archives of Virology1–4.

Drummond JA, FRANCO JL de A, OLIVEIRA D de (2010) Uma análise sobre a história e a situação das unidades de conservação no Brasil. Conserv da Biodiversidade Legis e Políticas Públicas Brasília Ed Câmara.

Du Z, Tang Y, Zhang S, et al (2014) Identification and molecular characterization of a single-stranded circular DNA virus with similarities to S*clerotinia sclerotiorum hypovirulence-associated DNA virus 1*. Archives of Virology 159:1527–1531.

Dzoyem JP, Kuete V, Eloff JN (2014) Biochemical parameters in toxicological studies in Africa: significance, principle of methods, data interpretation, and use in plant screenings. In: Toxicological Survey of African Medicinal Plants. Elsevier, pp 659–715.

Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM (2012) DNA barcoding methods for land plants. In: DNA Barcodes. Springer, pp 223–252.

Fernandes FR, Albuquerque LC, de Oliveira CL, et al (2011) Molecular and biological characterization of a new Brazilian begomovirus, euphorbia yellow mosaic virus (EuYMV), infecting Euphorbia heterophylla plants. Archives of Virology156:2063.

Fernandes PA, Pessôa VLS (2011) O cerrado e suas atividades impactantes: uma leitura sobre o garimpo, a mineração e a agricultura mecanizada. Observatorium Revista Eletrônica de Geografia 3.

Flory SL, Clay K (2013) Pathogen accumulation and long-term dynamics of plant invasions. Journal of Ecology 101:607–613.

Fonseca PLC, Badotti F, de Oliveira TFP, et al (2018) Virome analyses of Hevea brasiliensis using small RNA deep sequencing and PCR techniques reveal the presence of a potential new virus. Virology Journal 15:184.

Fontenele RS, Lamas NS, Lacorte C, et al (2017) A novel geminivirus identified in tomato and cleome plants sampled in Brazil. 240:175–179. doi: 10.1016/j. Virusres. 2017.08.007.

Gonthier P, Nicolotti G (2013) Infectious Forest Diseases. Cabi.

Hadidi A, Flores R, Candresse T, Barba M (2016) Next-generation sequencing and genome

editing in plant Virology. Frontiers in Microbiology 7:. doi: 10.3389/fmicb.2016.01325.

Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. The Journal of Virological Methods 116:209–211.

Jiwaji M, Matcher GF, de Bruyn M-M, et al (2019) Providence virus: An animal virus that replicates in plants or a plant virus that infects and replicates in animal cells? PLoS One 14:e0217494.

Jones RAC (2009) Plant virus emergence and evolution: Origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. Virus Research141:113–130. doi: 10.1016/j. Virusres. 2008.07.028.

Jones RAC, Coutts BA (2015) Spread of introduced viruses to new plants in natural ecosystems and the threat this poses to plant biodiversity. Molecular Plant Pathology 16:541–545.

Kearse M, Moir R, Wilson A, et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Keesing F, Belden LK, Daszak P, et al (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. Nature 468:647.

Kraberger S, Farkas K, Bernardo P, et al (2015) Identification of novel Bromus-and Trifolium-associated circular DNA viruses. Archives of Virology160:1303–1311.

Kreuze JF, Perez A, Untiveros M, et al (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. doi: 10.1016/j. Virology. 2009.03.024.

Lamas NS, Fontenele RS, Melo FL, et al (2016) Complete genome sequence of a genomovirus associated with common bean plant leaves in Brazil. Genome Announcements 4:e01247-16.

Lefeuvre P, Martin DP, Elena SF, et al (2019) Evolution and ecology of plant viruses. Nature Reviews Microbiology 17:632–644.

Liang P, Navarro B, Zhang Z, et al (2015) Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. Journal of General Virology 96:2411–2420.

Loconsole G, Giampetruzzi A, Roberto R, et al (2012) Next generation sequencing and metagenomic analysis advances plant virus diagnosis and discovery. Journal of Plant Pathology 94.

Lu Q-Y, Wu Z-J, Xia Z-S, Xie L-H (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (Morus alba L.). Archives of Virology 160:2135–2138.

MacDiarmid R, Rodoni B, Melcher U, et al (2013) Biosecurity implications of new technology and discovery in plant virus research. PLOS Pathogens 9:e1003337.

Macedo MA, Michereff Filho M, Navas-Castillo J, Inoue-Nagata AK (2015) Host range and whitefly transmission efficiency of Tomato severe rugose virus and Tomato golden vein virus in tomato plants. Tropical Plant Pathology40:405–409.

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--15.

Marquis RJ (2002) The cerrados of Brazil: ecology and natural history of a neotropical savanna. Columbia University Press.

Marzano S-YL, Domier LL (2016) Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. Virus Research 213:332–342.

Morales FJ, Anderson PK (2001) The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Archives of Virology146:415–441.

Muthukumar V, Melcher U, Pierce M, et al (2009) Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. Virus Research 141:169–173.

Navarro B, Minutolo M, De Stradis A, et al (2018) The first phlebo-like virus infecting plants: a case study on the adaptation of negative-stranded RNA viruses to new hosts. Molecular Plant Pathology 19:1075–1089.

Nicolini C, Pio-Ribeiro G, Andrade GP, et al (2012) A distinct tymovirus infecting Cassia hoffmannseggii in Brazil. Virus Genes 45:190–194.

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

Pagán I, González-Jara P, Moreno-Letelier A, et al (2012) Effect of biodiversity changes in disease risk: exploring disease emergence in a plant-virus system. PLOS Pathogens 8:e1002796.

Piqué M, Mougeot J-L, Geldreich A, et al (1995) Sequence of a cauliflower mosaic virus strain infecting solanaceous plants. Gene 155:305–306.

Qiu SG, Schoelz JE (1992) Three regions of cauliflower mosaic virus strain W260 are involved in systemic infection of solanaceous hosts. Virology 190:773–782.

Quan P-L, Briese T, Palacios G, Lipkin WI (2008) Rapid sequence-based diagnosis of viral infection. Antiviral Research 79:1–5.

Rodríguez-Negrete EA, Morales-Aguilar JJ, Domínguez-Duran G, et al (2019) High-Throughput Sequencing Reveals Differential Begomovirus Species Diversity in Non-Cultivated Plants in Northern-Pacific Mexico. Viruses 11:594.

Roossinck MJ (2010) Lifestyles of plant viruses. Philosophical Transactions of the Royal Society B 365:1899–1905. doi: 10.1098/rstb.2010.0057.

Roossinck MJ, García-Arenal F (2015) Ecosystem simplification, biodiversity loss and plant virus emergence. Current Opinion in Virology 10:56–62. doi: 10.1016/j.coviro.2015.01.005.

Roossinck MJ, Saha P, Wiley GB, et al (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. Molecular Ecology 19:81–88.

Rosario K, Breitbart M (2011) Exploring the viral world through metagenomics. Current Opinion in Virology. Virology 1:289–297.

Rosolen V, Resende TM, Borges EN, et al (2012) Impactos da substituição da vegetação original do Cerrado brasileiro em sistemas agrícolas: alteração do carbono orgânico do solo e δ13C. Investigaciones Geográficas 39–47.

Scholthof KBG, Adkins S, Czosnek H, et al (2011) Top 10 plant viruses in molecular plant pathology. Molecular Plant Pathology 12:938–954. doi: 10.1111/j.1364-3703.2011.00752.x.

Vaghi Medina CG, Teppa E, Bornancini VA, et al (2017) Tomato Apical Leaf Curl Virus: A Novel, Monopartite Geminivirus Detected in Tomatoes in Argentina. Frontiers in Microbiology 8:2665. doi: 10.3389/fmicb.2017.02665.

Vincent SJ, Coutts B a., Jones R a C (2014) Effects of introduced and indigenous viruses on native plants: Exploring their disease causing potential at the agro-ecological interface. PLoS One 9:6–9. doi: 10.1371/journal.pone.0091224.

Vu S, Melgarejo TA, Chen L, et al (2015) Evidence that tomato mottle leaf curl virus from Northeastern Brazil is an indigenous New World monopartite begomovirus. Phytopathology 105:S4.

Yazdi HRB, Heydarnejad J, Massumi H (2008) Genome characterization and genetic diversity of beet curly top Iran virus: a geminivirus with a novel nonanucleotide. Virus Genes 36:539–545.

Zhao L, Rosario K, Breitbart M, Duffy S (2019) Eukaryotic circular rep-encoding single-stranded DNA (CRESS DNA) viruses: Ubiquitous viruses with small genomes and a diverse host range. Advances in Virus Research 103:71–133.

Chapter 3

New genomovirus associated with trees and cultivated species in Central Brazil

Resumo

A família *Genomoviridae*, proposta em 2016, agrupa atualmente nove gêneros de vírus de ssDNA que foram encontrados em associação com diversos organismos, incluindo plantas. Desde 2016 um aumento crescente em espécies da familia *Genomoviridae*, associadas a diferentes ambientes, vem sendo observado. O objetivo deste trabalho, foi caracterizar sequências virais pertencendo à família *Genomoviridae* recuperadas via High-throughput sequencing (HTS) provenientes de amostras foliares de espécies arbóreas (nativas e exóticas) e cultivadas na região Central do Brasil. Ao todo 246 amostras foliares (correspondendo 41 famílias) foram obtidas. Inicialmente realizou-se extrações individuais de DNA, seguida de RCA (*Rolling Circle Amplification*) e envio em pools para HTS. Dentre os *contigs* gerados, sete *contigs* foram equivalentes a sete espécies classificadas em *Genomoviridae*. Primers específicos foram desenhados e procedeu-se a detecção viral em amostras individuais de plantas. Resultados positivos foram observados em plantas de sete famílias botânicas. O genoma completo foi recuperado via *Sanger* e análises de identidade e filogenia foram realizadas. Sequências de três espécies novas se posicionaram próximas ao gênero *Gemyvongvirus*, entretanto, possuem organização genômica diferente deste gênero. Isolados destas espécies foram detectados em *Vochysia rufa* (Vochysiaceae), *Byrsonima crassiflora* (Malpighiaceae) e *Eucalyptus urophylla* (Myrtaceae). Os nomes propostos para estas espécies foram: Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV e Eucalyptus urophylla associated genomovirus - EcaGmV. Sequências de duas novas espécies nomeadas como Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) e *Ouratea duparquetiana* associated gemykibivirus - OuaGmV (*Gemykibivirus*) foram detectadas em *Tecoma stans* (Bignoniacaeae) e *Ouratea duparquetiana* (Ochnaceae) respectivamente. Além disto, dois isolados das espécies *Gila monster-associated gemykrogvirus* - (*Gemykrogvirus*) - GmaV1 e Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) foram detectados em *Trembleya parviflora* (Melastomataceae) e *Solanum melongena* (Solanaceae). respectivamente. Clones infecciosos serão produzidos e ensaios futuros serão realizados para determinar se algum destes genomovírus é capaz de se replicar em plantas. Esses resultados constituem exemplos raros de genomovírus associados a plantas.

Palavras-chave: *Genomoviridae*, *Gemyvongvirus*, *Gemykolovirus*, *Gemykibivirus*, *Gemykrogvirus*, *Gemycircularvirus*, espécies arbóreas, cultivadas

Abstract

The *Genomoviridae* family, proposed in 2016, currently accommodates nine genera of ssDNA viruses in association with a diversity of organisms, including plants. Since 2016, a crescent increase in species of family *Genomoviridae* family, associated with different environments, has been observed. The objective of this work was to characterize seven viral sequences belonging to family *Genomoviridae* recovered by High-Throughput Sequencing

(HTS) from leaf samples of native and cultivated trees in Central Brazil. All 246 samples (41 families) foram obtained. Firstly, DNA was individually extracted, followed by RCA (Rolling Circle Amplification) and sent in pools for HTS. Seven contigs were equivalent to seven species of family *Genomoviridae*. Specific primers were designed to detection in individuall plant species. The full genome of seven species were recovered by Sanger and paired identity and phylogeny analyzes were performed. Sequences of three new species were close to species of *Gemyvongvirus*. However, they have different genomic organization of this genus. Isolates from this species were detected in *Vochysia rufa* (Vochysiaceae), *Byrsonima crassiflora* (Malpighiaceae) and *Eucalyptus urophylla* (Myrtaceae). The proposed names were: Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV and Eucalyptus urophylla associated genomovirus - EcaGmV. Sequences of two new species: Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) and Ouratea duparquetiana associated gemykibivirus - OuaGmV (*Gemykibivirus*) were detected in Tecoma stans (Bignoniacaeae) and Ouratea duparquetiana (Ochnaceae) respectively. In addition, two isolates of Gila monster-associated gemykrogvirus - (*Gemykrogvirus*) - GmaV1 and Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) species were detected in *Trembleya parviflora* (Melastomataceae) and *Solanum melongena* (Solanaceae), respectively. Infecctious clones will be produced and future assays will be performed to determine if any of these genomoviruses can replicate in plants. These findings constitute rare examples of plant-associated genomoviruses.

Keywords: *Genomoviridae*, *Gemyvongvirus*, *Gemykolovirus*, *Gemykibivirus*, *Gemykrogvirus*, *Gemycircularvirus*, tree species, crops

1. Introduction

The family *Genomoviridae*, proposed in 2016 by the International Committee on Virus (ICTV), currently encompasses nine genera: *Gemycircularvirus*, *Gemykibivirus*, *Gemygorvirus*, *Gemykolvirus*, *Gemyvongvirus*, *Gemytondvirus*, *Gemykroznavirus* and *Gemyduguivirus* (Varsani and Krupovic 2017).

Members of *Genomoviridae* family showing icosahedral particles with a diameter of 20 to 22 nm (Yu et al. 2010, Sikorski et al. 2013) encapsiding a small circular single stranded DNA (2.1 to 2.3 kb) encodes two proteins, the coat protein - CP (in viral sense) and the replication iniciation protein (Rep). Some genomes may have a spliced Rep protein in addition to a conserved stem-loop structure containing the origin of DNA replication. Phyllogenetic analyses Rep genomoviruses form a sister clade to geminiviruses. Genomoviruses have a high sequence diversity of 47% among genus species. The accepted identity criterion for species is 78% paired

identity (Krupovic, et al. 2016) when comparing the complete genome. Species with values larger than 78% are considered as variants.

With the advent of metagenomic new viruses, including the members of *Genomoviridae* family, are being discovered in a wide diversity of organisms. Initially isolated from fungi (Yu et al. 2010), genomoviruses are also associated with animals (Dayaram et al. 2015, Fontenele et al. 2019, Kraberger et al. 2018, Li et al. 2015, Nakasu et al. 2017, Rosario et al. 2012a), humans (Halary et al. 2016, Lamberto et al. 2014, Phan et al. 2015), and plants (Dayaram et al. 2012, Kraberger et al. 2015, Marzano and Domier 2016, Richet et al. 2019). They are also present in environmental samples such as sewage, faeces and water (Conceição-Neto et al. 2015, da Silva Assis et al. 2016, Sikorski et al. 2013) and in plants. To date, viruses belonging to *Genomoviridae* family have not been reported as animal or plant pathogens. However, Ng et al. (2014) got positive infection inoculating a virus associated with ancient caribou feces associated virus, the - ACFV genomovirus in *Nicotiana benthamiana*. The virus multiplied in the plant but without induce symptoms.

Other viruses, accepted by ICTV, were found associated with plants as:*Cassava associated gemycircularvirus 1* (Dayaram et al. 2012); *Hypericum japonicum associated gemycircularvirus 1* (Du et al. 2014); *Poaceae associated gemycircularvirus 1* (Male et al. 2015); *Bromus associated gemycircularvirus 1* (Kraberger et al. 2015); *Soybean associated gemycircularvirus 1* (Marzano and Domier, 2016). In tree species of the only record refers to a gemycircularvirus associated with olive tree in Italy (Chiumenti et al. 2019).

In Brazil, studies of genomovirus are scarce, and their occurrence has been reported in beans (*Phaseolus vulgaris*) (Lamas et al. 2016) and in weeds (*Momordica charantia* and *Euphorbia heterophylla*) (Rezende et al. 2018).

Considering the prodigious diversity of the Brazilian native flora, especially in the *Cerrado* biome, and the gigantic area covered by its main monocultures [beans, soybeans (*Glycine max*),

cotton (*Gossypium hirsutum*), tomatoes (*Solanum lycopersicum*), sweet pepper (*Capsicum annuum*), economic and ecological impact impact is expected to occur, as shown here with the detection of seven *Genomoviridae* species collected in conservation, afforestation, nursery and agricultural cultivation environments, in the Central region of Brazil.

2. Materials and Methods

2.1. Sample collection locations - areas and plant species

The collection sites were areas containing trees (native and exotic). The microenvironments were classified as: (1) area with strong anthropic action (afforestation, nursery and agricultural crops); (2) conservation areas and (3) transition areas showing agricultural crops close to conservation area. The following collection sites were classified as anthropic microenvironments: Arborização Urbana do Distrito Federal - AUDF, Núcleo Rural Taquara - NRT, Viveiro II da Companhia Urbanizadora da Nova Capital - VII-NOVACAP. Conservation areas are two ecological reserves in Brasília (Estação Ecológica de Águas Emendadas - ESECAE and Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística - RECOR); and the "Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) and Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB) was chosen as a transitional environment due to the close proximity of a conservation area and area used for agricultural experiments. The locations are showed on **Figure 1** of this Thesis.

2.2. List of plant species collected

All collections were performed in 2016 and the first quarter of 2017. A total of 246 samples (41 botanical families) of different host species showing virus-like symptoms, including mosaic, mottle, blistering, yellowing and leaves deformation were collected. Here were aadded to three

samples *Eucalyptus urophylla* samples from VII-NOVACAP. Further information can be found in as described in Supplementary **Table ST1**.

2.3. Plant material processing and High-throughput sequencing - HTS

The symptoms were photographed, and leaves were cleaned with the aid of a brush. Total DNA was extracted individually from each sample using a CTAB protocol (Boiteux et al. 1999). Circular DNA enrichment was performed via RCA - Rolling Circle Amplification (Inoue-Nagata et al. 2004) before preparing the pools for sequencing, later submitted to HTS.

The RCA reaction was performed for a final volume of 10 μL final volume reaction containing: 0.1 μL Phi-29 DNA polymerase (10,000 U/mL), 1 μL of Phi-29 DNA polymerase enzyme buffer (NEB), 1 μL BSA (10X), 1 μL dNTPs (2.5 mM), 1 μL Thioprotected Primer (50 μM), 4.9 μL Milli-Q water and 1 μL extracted DNA (20 ng / μL).

The reaction was incubated at 30 ° C for 18 hours, and then the reaction was submitted for 10 min at 65 ° C for inactivation of the enzyme. Four DNA pools were separated by the collection sites ESECAE and RECOR (60 samples, 38 plant species); VII-NOVACAP (79 samples, 29 plant species); ARIE-EEB and AHE-EEB (60 samples, 39 plant species), and AUDF and NRT in Brasília-DF (47 samples, 20 plant species).

2.4. Bioinformatic and metagenomic analyzes

The RCA products were sequenced by Illumina HiSeq 2500 platform (2 × 125 paired ends) at Macrogen Inc. (South Korea). Sequence analyzes were performed essentially as previously described (Kreuze et al. 2009, Adams and Fox 2016). The following steps were performed: (i) elimination of low quality sequences and adapters; (ii) reassembly of sequences and organization into contigs, using the CLC Genomics Workbench 10 program; (iii) validation of contigs by local alignments using BLASTx algorithm (Altschul et al. 1990) against a virus database (Viral

RefSeq-GenBank); (iv) Afterward, the BLASTx results were compared to ViralRefSeq and another round of BLASTx analyses were performed with the general database. Contigs with higher coverages were selected; (v) the contigs showing identity to plant viruses after alignments were then analyzed with the assistance of the Geneious R11 program (Kearse et al. 2012). Contigs were mapped in order to obtain the complete genome sequences. Afterward, a BLASTn analysis was performed with the entire genome and the most similar sequence was used to transfer annotation. ORFs were confirmed using the Conserved Domain Architecture Retrieval Tool (CDART) and the ORF finder programs.

2.5. Specific primers design and virus detection

For detection of seven viruses found via High-Throughput Sequencing (HTS), specific primers were designed to contigs recovered by HTS using Geneious R11 program based on the sequence obtained from the Illumina sequencing (**Table 1**).

Amplification reactions were performed in a total volume of 12.5 µL containing 1.25 Taq Polymerase 10X Buffer (100 mM Tris-HCl, pH 8.3 and 500 mM KCl, Invitrogen), 0.38 µL MgCl2 (50 mM, Invitrogen), 0.25 µL dNTPs (2.5 mM, Invitrogen), 0.25 µL of each primer listed in **Table 1** (10uM), 0.1 µL of Taq Polymerase enzyme (5U / µL, Invitrogen), 9.02 µL of MiliQ water and 1 µL of RCA (Diluted RCA 1:10 MiliQ water). The reactions were amplified in thermal cycler programmed under the following conditions: initial temperature 94 ° C for 3 minutes and then 35 cycles each with 94 ° C for 30 seconds (denaturation); Ta ° C (see **Table 1**) for 45 seconds (annealing) and 72 ° C for 3 minutes (extension). The product of PCR was visualized on 1% agarose gel and stained with ethidium bromide.

Table 1. Oligonucleotides used to detect seven species of the *Genomoviridae*, designed using new sequences obtained from High-Throughput Sequencing - HTS, or avaliable in LVV-FITO UnB (Laboratório de Virologia Vegetal – Fitopatologia UnB).

Species	Name of primer	Foward 5'-3' / Reverse 5'-3'	Ta[1]	Amplicon
Vochysia associated genomovirus - VoaGmV	**15_6033F/ 15_6033R**	GAATAGGCCTCGAACGAC/ GCATTTTACCAGCGCCTAC	60	1895
Byrsonima associated genomovirus - ByaGmV	**15_5347F/ 15_5347R**	CGTTTATCTTTGCGCTTCG/ ACGAGTCAACCAGTTCCGTC GGA	60	1922
Eucalyptus urophylla associated genomovirus - EcaGmV	**F1_ EUaVPINT/ EUaVPINT _R1**	GTGTCCCACTGATAACG/ CCAGAGAGCGAGACAGA	55	1886
Gila monster-associated gemycircularvirus – GmaV1	**CaribouEcoR1-F/ CaribouEcoR1-R**	GAATTCGTCCATGTTGAATTG/ GACGAATTCTCCGACACTTGT	59	2196
Ouratea associated gemykibivirus - OuaGmV	**7_6560F/ 7_6560R**	CTACGATAAGACCCGCCACC/ AGTGGACCCTAATACGCTTG	60	2201
Tecoma associated gemykolovirus - TeaGmV	**TickPst1-F/ TickPst1-R**	CTGCAGGTAATGAGCAGGGT/ CTGCAGCTAATGTAGTAGTC	52	2200
Momordica charantia associated gemycircularvirus - MoaGmV	**11_205F/ 11_205R**	GACCACCTCTGATGTCGT/ GGTCTTGGATTCTTTCATGG	60	2193

[1]Ta= annealing temperature.

2.6. Sanger validation for High-throughput sequencing - HTS

The PCR product or clone was sent for sequencing. Purification of DNA fragments obtained from PCR was performed in according to the availability of GFX (GE Healthcare) purification kit following manufacturers' guidelines. The generated amplicon was observed in 0.8% agarose gel electrophoresis, purified and cloned into pGEM T Easy vector (Promega) and send to sequencing at Myleus (Minas Gerais, Brazil) and Embrapa Hortaliças (Brasília, Brazil). Internal primers were designed to recovery the complete genome sequence.

2.7. Phylogenetic analysis, pairwise identity and recombination analysis

Sixty-seven sequences representing all genera of the*Genomoviridae* family were included in the phylogenetic analysis. Among them are species aligned using the BLASTn algorithm. The nucleotide sequence of Rep was used in the analyses. The alignment was built using MUSCLE (Edgar, 2004), and the jModelTest program (Posada, 2008) was used to provide the best fit nucleotide substitution model for each data set. The phylogenetic tree was built by Bayesian inference.

The trees will be edited in the FigTree program (Rambaut ,2012) and by the Evolview (He, et al. 2016) online server http://www.evolgenius.info/evolview/.

For comparisons of the paired complete sequences, was used the software SDT v.1.2 (Muhire et al. 2014), and RDP4 program (Martin et al. 2015) for the recombination analysis.

3. Results and Discussion

3.1. Molecular characterization and PCR detection

To study the diversity of ssDNA virus, samples of tree species (native and exotic) growing across distinct areas from three microenvironments (anthropic, conservation and transition) of the *Cerrado* biome of Central Brazil were collected showed symptoms similar to those caused by viruses.

The use of HTS and metagenomic analysis detected seven sequences that when performing BLASTx (**Table 2**) aligned with species of family *Genomoviridae*.

Specific primers were designed, and viruses were detected by PCR and the sequences were recovered by Sanger in *Tecoma stans* (AUDF), eggplant (*Solanum melongena*) (NRT), *Eucalyptus urophylla* (VII-NOVACAP), *Trembleya parviflora* (ESECAE), *Vochysia rufa* and *Byrsonima crassiflora* (RECOR) and *Ouratea duparquetiana* (ARIE-EEB). **Figure 1** shows the genomic organization and photos of the plants in which the viruses were found (except: *Tecoma stans* and *Eucalyptus urophylla*).

Table 2. Single stranded DNA viruses (*Genomoviridae* family) detected in tree and herbaceous species growing across distinct sites in three microenvironments of the *Cerrado* area of Brasília-DF (Federal District), Central Brazil. Viruses were organizated considering collection sites, taxonomy and genomic organization in association with the corresponding GenBank accession number and identity (percentage). Analyses were carried out via BLASTx in Geneious R11 program, and NCBI website.

Sites	Genus	Species	GenBank Accession number	Identity (%)
RECOR[4]	unclassified	Vochysia rufa associated genomovirus	QCW23624	20
		Byrsonima crassiflora associated genomovirus	YP009021860	20
NOVACAP[5]		Eucalyptus urophylla associated genomovirus	QCS35893	40
ESECAE[3]	*Gemykrogvirus*	Gila monster-associated gemykrogvirus	QCQ85257	89
ARIE-EEB and AHE-EEB*[6]	*Gemykibivirus*	Ouratea duparquetiana associated gemykibivirus	QCX29358	71
AUDF[1]	*Gemykolovirus*	Tecoma stans associated gemykolovirus	QCX29420	74
NRT[2]	*Gemycircularvirus*	Momordica charantia associated gemycirvcularvirus	AXI69777	90

[1]Arborização Urbana do Distrito Federal (AUDF), [2]Núcleo Rural Taquara (NRT), [3]Estação Ecológica de Águas Emendadas (ESECAE), [4]Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR),[5]Viveiro II da Companhia Urbanizadora da Nova Capital (VII-NOVACAP), [6]"Área de Relevante Interesse Ecológico" - Estação Experimental de Biologia (ARIE-EEB) and Área da Horta Experimental da Estação Experimental de Biologia (AHE-EEB)* No genomovirus was found in the AHE-EEB.

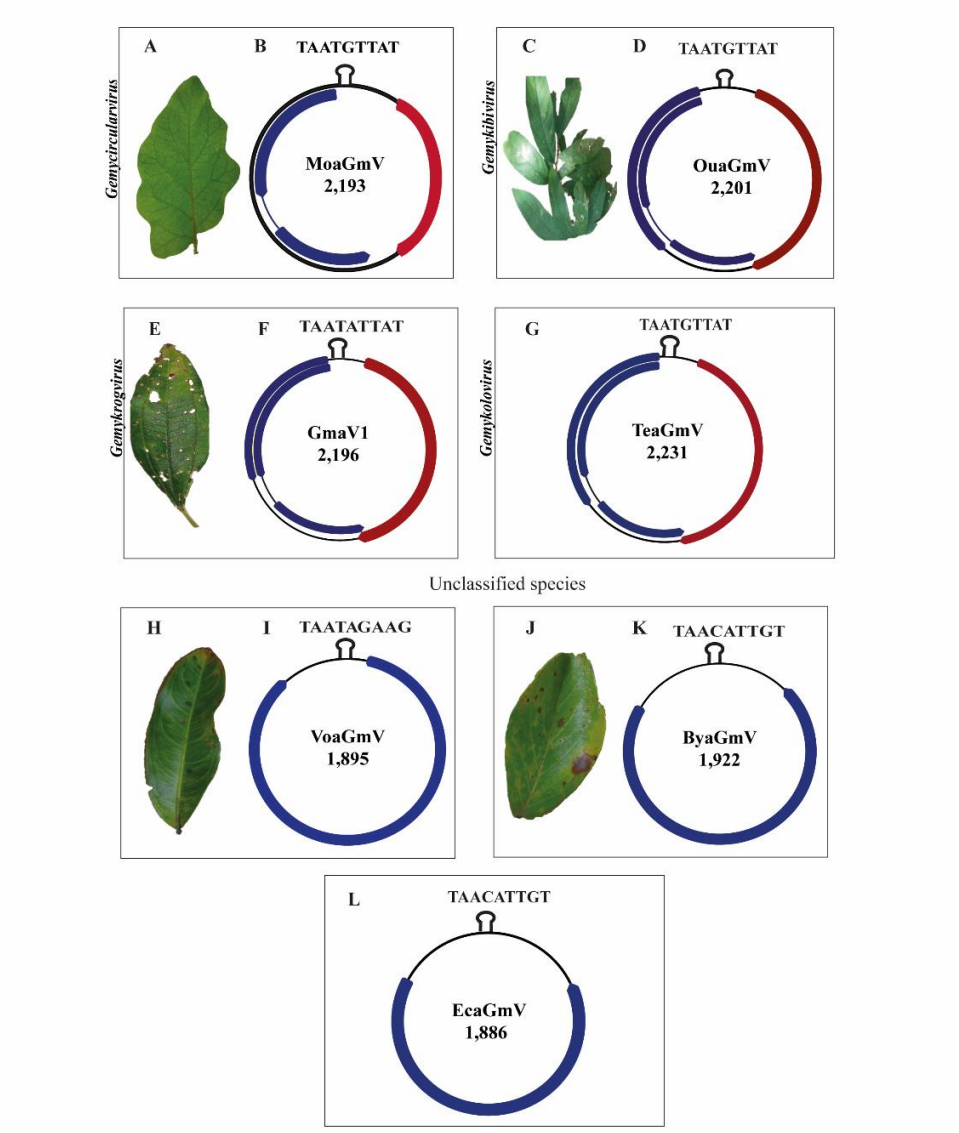


Figure 1. Genomic organization scheme and symptomatic plants which viruses were associated. In: A. *Solanum melongena* and B. Momordica charantia associated gemycircularvirus - MoaGmV; C. *Ouratea duparquetiana* and D. Ouratea duparquetiana associated gemykibivirus - OuaGmV; E. *Trembleya parviflora*; F. Gila monster-associated gemykrogvirus - GmaV1; G. Tecoma stans associated gemykolovirus; H. *Vochysia rufa* and I. Vochysia rufa associated genomovirus - VoaGmV; J. *Byrsonima crassiflora* and K. Byrsonima crassiflora associated genomovirus - ByaGmV and L. Eucalyptus urophylla associated genomovirus - EcaGmV. Scheme of genomic organization (replication associated protein gene (*rep*) in blue and coat protein gene (*cp*) in red) and plant photos: **A.** *Solanum melongena* and **B.** Momordica charantia associated with the gemicircular virus - MoaGmV; **C.** *Ouratea duparquetiana* and **D.** Ouratea duparquetiana associated gemykibivirus - OuaGmV; **E.** *Trembleya parviflora* and **F.** Gila monster associated gemykrogvirus - GmaV1; **G.** Tecoma stans gemykolovirus associated - TeaGmV; **H.** *Vochysia rufa* and **I.** Vochysia rufa-associated genomovirus - VoaGmV; **J.** *Byrsonima crassiflora* and **K.** Byrsonima crassiflora associated genomovirus; **L.** Eucalyptus urophylla associated genomovirus - EcaGmV. The nonanucleotide of each species is illustrated. The proposed taxonomic classification is next to the boxes containing the genome and plant.

The low identity among five sequences when compared to sequences from other species [aminoacids (aa) and nucleotides (nt)] available in GenBank showed that they are likely new species to be classified in family*Genomoviridae*, since the species demarcation criterion in the genera of this family. equals the identity of complete genome nucleotides of less than 78% with other species already deposited (Varsani and Krupovic 2017).

For the five new species the names (acronyms) were proposed: Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV, Eucalyptus urophylla associated genomovirus - EcaGmV, Ouratea duparquetiana associated gemykibivirus - OuaGmV and Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*). The other two species Momordica charantia associated gemycircularvirus - MoaGmV (de Rezende et al. 2018) and Gila monster associated gemykrogvirus - GmaV1 (Somayaji et al. 2018) are species already described previously

In analyzing the seven sequences, they all presented stem-loops containing a nonanucleotide sequence that is probably important for the initiation of rolling circle replication. Sizes range from 1.8 kb to 2.3 kb.

The sequences detected vary belongs to genes that encode protein coat (*cp*) and replication-associated protein (*rep*) proteins, and some have splicing in *rep* and / or replication-associated protein (*repA*). The sequences of VoaGmV, ByaGmV and EcaGmV species have only *rep*, while GmaV1, OuaGmV, TeaGmV have *cp*, *repA* and *rep* form splicing. The MoaGmV species has only *cp* and *repA*. **Table 3** shows collection site information, proposed name, size, nonanucleotide sequence, size of ORFs in nucleotides and amino acids of genes: *cp*, *rep*, *repA*.

Table 3. Characteristics of the new *Genomoviridae* species organized by collection sites, genomic organization features, proposed new names, genome-full size, nonanucleotide, size genes: *cp*, *repA* and *rep* nucleotide (nt) and aminoacid (aa).

Genus	Species	Size (nt[1])	Nonanucleotide	CP[3](nt[1]/aa[2])	RepA[4] (nt[1]/aa[2])	Rep splicing[5](nt[1])
Unclassified	Vochysia rufa associated genomovirus	1895	TAATAGAAG	Unknown	Unknown	1599/533
	Byrsonima crassiflora associated genomovirus	1922	TAACATTGT	Unknown	Unknown	1347/449
	Eucalyptus urophylla associated genomovirus	1886	TATACTTTC	Unknown	Unknown	549/183
Gemykrogvirus	Gila monster-associated gemykrogvirus	2196	TAATATTAT	912/304	621/207	990/330
Gemykibivirus	Ouratea duparquetiana associated gemykibivirus	2201	TAATGTTAT	870/290	756/252	960/320
Gemykolovirus	Tecoma stans associated gemykolovirus	2221	TAATGTTAT	927/309	762/254	1047/349
Gemycircularvirus	Momordica charantia associated gemycircularvirus	2193	TAATGTTAT	921/307	Unknown	1014/338

[1]nt = Nucleotide; [2]aa = amino acid; [3]*cp* = coat protein; [4]*repA* = replication-associated protein, [5]*rep* splicing = replicon protein with splicing intron region.

Some species contains a potential stem-loop structure with a nonanucleotide ('TAWWDWRN') motif at its apex, which is likely to be important for rolling-circle replication initiation conserved and motifs described by Krupovic et al. 2016, the motifs conserved are in **Table 4.**

The isolate GmaV1 showed motifs II, III, C, Walker A, B and GRS domain similar to *Gemykrogvirus* genus, only motif I had different amino acids. Vochysia associated genomovirus showed only motif II similar to the species of the genus *Gemycircularvirus*, motifs I, Walker A, B and GRS domain have conserved amino acids, however different sequence from those described, motif C and III are unknown. Byrsonima associated genomovirus has motifs II, III, Walker A, B and GRS domain different from those described, but with conserved amino acids, motif C is similar to species of genus *Gemykibivirus*, motif I is unknown. In Ouratea associated genomovirus motifs I and III were found, which are different from those already described, while motifs II, Walker B and GRS domain are similar to the species of the genus *Gemykibivirus*. Tecoma associated genomovirus presents motifs III, Walker B and GRS domain similar to species of the genus *Gemykolovirus*, motifs I, II, Walker A and motif C are unknown. Momordica charantia associated gemycircularvirus has all similar motifs to the species of the genus *Gemycircularvirus*, except Motive C which was not found in this isolate. Eucalyptus urophylla associated genomovirus had no known reason.

Table 4. Conserved motifs exhibited by *Genomoviridae* genome sequence.

Species	**Motif I**	**MotifII**	**Motif III**	**Domain GRS**	**WALKER A**	**WALKER B**	**Motif C**
Gila monster-associated gemykrogvirus	IITFPQ	VHYHV	YVGK	TAFDYFGAHCNIKSIR	GPTRTGKT	IFDDI	MCMV
Vochysia rufa associated genomovirus	SLTLSQ	FLHL	unknown	DGVTYHPNFKKVKNY	SQHTTGKT	ILDDL	Unknown
Byrsonima crassiflora associated genomovirus	unknown	ALHLS	YAIP	KFMDIAGYHPNIKPIK	LGSLLGKT	ILDDL	YLSN
Ouratea duparquetiana associated gemykibivirus	LLTYPQ	IHLHA	YAFS	RVFDVDGRHPNVVRG	unknown	VFDDM	Unknown
Tecoma stans associated gemykolovirus	unknown	unknown	YVGK	RTFKVGTRVPNIRVRR	unknown	IFDDM	Unknown
Momordica charantia associated gemycircularvirus	LLTLPP	LHLHV	YAIK	DVFDVDGRHPNVEPSK	GKSRTGKT	VFDDI	Unknown

The study of conserved motifs allows the understanding of the rolling circle replication mechanism and evolution.

The N-terminal region has important rolling circle motifs (I, II, and III) that are well conserved in many ssDNA viruses, phages, and plasmids that replicate using the RCR mechanism (Ilyina and Koonin 1992, Vega-Rocha et al. 2007a, Duffy and Breitbart 2012, Krupovic, 2013). The GRS motif in the Rep sequence is found only in geminivirus and genomovirus. In geminiviruses, it allows spatial arrangements of motifs II and III demonstrating that GRS is essential for geminivirus replication (Nash et al. 2011) and is likely to be the case for genomoviruses as well.

Rep is a multifunctional protein with endonuclease and helicase activities. In the C-terminal region, Rep helicase activity is mediated by conserved motifs known as Walker A, Walker B, and motif C (Gorbalenya et al. 1990; Koonin, 1993; Choudhury et al. 2006; Clerot and Bernardi 2006). The domain found in the Rep proteins of eukaryotic ssDNA viruses belongs to the superfamily helicase 3 (Gorbalenya et al. 1990; Koonin 1993).

3.2. Phylogenetic analysis, pairwise identity and recombination analysis

Sixty-seven amino acid sequences from Rep representing the family *Genomoviridae* were aligned using MUSCLE, with aphylogenetic tree constructed using Bayesian inference (**Figure 2**).

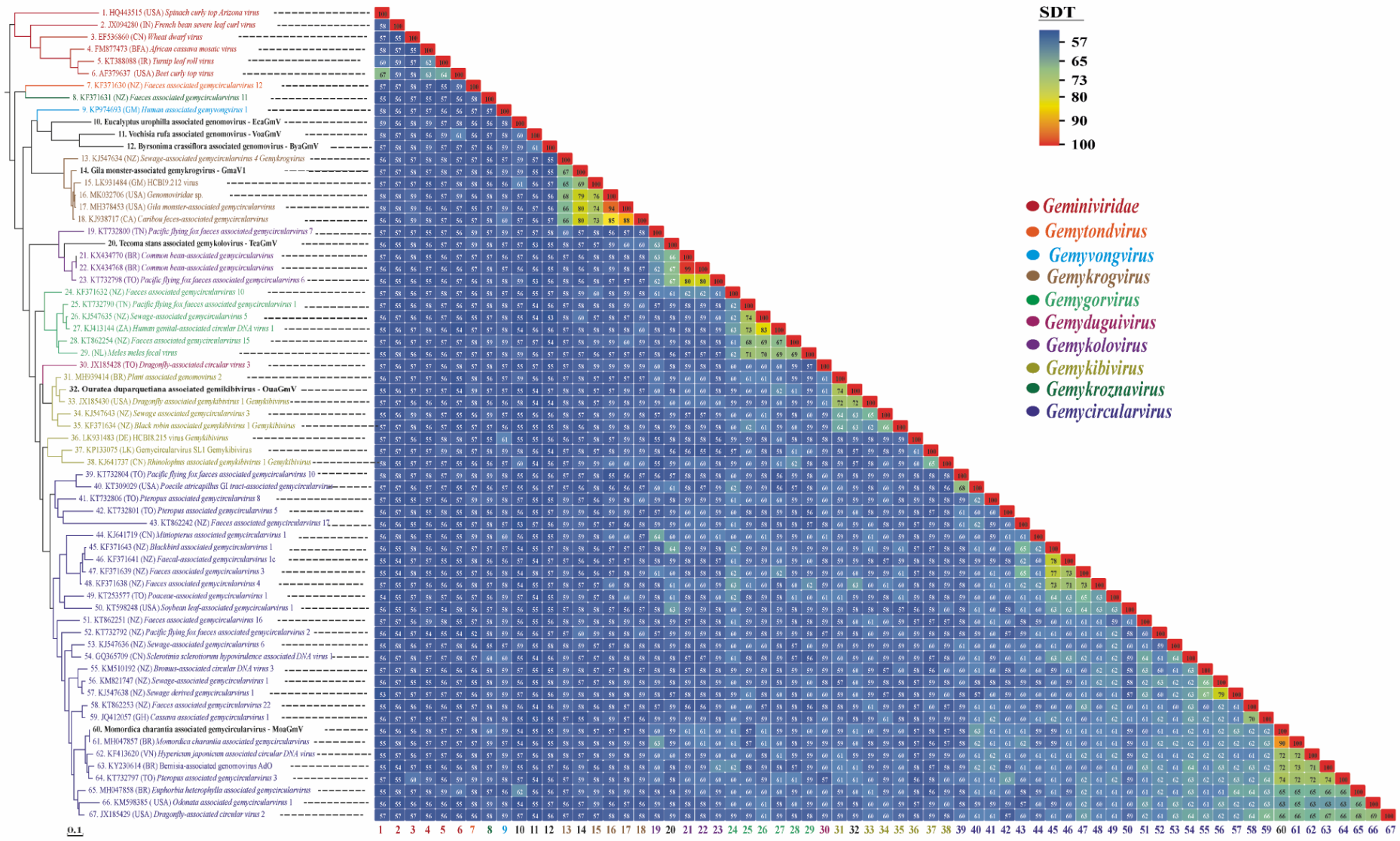

SDT
57
65
73
80
90
100
Geminiviridae
Gemytondvirus
Gemyvongvirus
Gemykrogvirus
Gemygorvirus
Gemyduguivirus
Gemykolovirus
Gemykibivirus
Gemykroznavirus
Gemycircularvirus
1. HQ443515 (USA) Spinach curly top Arizona virus
2. JX094280 (IN) French bean severe leaf curl virus
3. EF536860 (CN) Wheat dwarf virus
4. FM877473 (BFA) African cassava mosaic virus
5. KT388088 (IR) Turnip leaf roll virus
6. AF379637 (USA) Beet curly top virus
7. KF371630 (NZ) Faeces associated gemycircularvirus 12
8. KF371631 (NZ) Faeces associated gemycircularvirus 11
9. KP974693 (GM) Human associated gemyvongvirus 1
10. Eucalyptus urophilla associated genomovirus - EcaGmV
11. Vochisia rufa associated genomovirus - VoaGmV
12. Byrsonima crassiflora associated genomovirus - ByaGmV
13. KJ547634 (NZ) Sewage-associated gemycircularvirus 4 Gemykrogvirus
14. Gila monster-associated gemykrogvirus - GmaV1
15. LK931484 (GM) HCBI9.212 virus
16. MK032706 (USA) Genomoviridae sp.
17. MH378453 (USA) Gila monster-associated gemycircularvirus
18. KJ938717 (CA) Caribou feces-associated gemycircularvirus
19. KT732800 (TN) Pacific flying fox faeces associated gemycircularvirus 7
20. Tecoma stans associated gemykolovirus - TeaGmV
21. KX434770 (BR) Common bean-associated gemycircularvirus
22. KX434768 (BR) Common bean-associated gemycircularvirus
23. KT732798 (TO) Pacific flying fox faeces associated gemycircularvirus 6
24. KF371632 (NZ) Faeces associated gemycircularvirus 10
25. KT732790 (TN) Pacific flying fox faeces associated gemycircularvirus 1
26. KJ547635 (NZ) Sewage-associated gemycircularvirus 5
27. KJ413144 (ZA) Human genital-associated circular DNA virus 1
28. KT862254 (NZ) Faeces associated gemycircularvirus 15
29. (NL) Meles meles fecal virus
30. JX185428 (TO) Dragonfly-associated circular virus 3
31. MH939414 (BR) Plant associated genomovirus 2
32. Ouratea duparquetiana associated gemikibivirus - OuaGmV
33. JX185430 (USA) Dragonfly associated gemykibivirus 1 Gemykibivirus
34. KJ547643 (NZ) Sewage associated gemycircularvirus 3
35. KF371634 (NZ) Black robin associated gemykibivirus 1 Gemykibivirus
36. LK931483 (DE) HCBI8.215 virus Gemykibivirus
37. KP133075 (LK) Gemycircularvirus SL1 Gemykibivirus
38. KJ641737 (CN) Rhinolophus associated gemykibivirus 1 Gemykibivirus
39. KT732804 (TO) Pacific flying fox faeces associated gemycircularvirus 10
40. KT309029 (USA) Poecile atricapillus GI tract-associated gemycircularvirus
41. KT732806 (TO) Pteropus associated gemycircularvirus 8
42. KT732801 (TO) Pteropus associated gemycircularvirus 5
43. KT862242 (NZ) Faeces associated gemycircularvirus 17
44. KJ641719 (CN) Miniopterus associated gemycircularvirus 1
45. KF371643 (NZ) Blackbird associated gemycircularvirus 1
46. KF371641 (NZ) Faecal-associated gemycircularvirus 1c
47. KF371639 (NZ) Faeces associated gemycircularvirus 3
48. KF371638 (NZ) Faeces associated gemycircularvirus 4
49. KT253577 (TO) Poaceae-associated gemycircularvirus 1
50. KT598248 (USA) Soybean leaf-associated gemycircularvirus 1
51. KT862251 (NZ) Faeces associated gemycircularvirus 16
52. KT732792 (NZ) Pacific flying fox faeces associated gemycircularvirus 2
53. KJ547636 (NZ) Sewage-associated gemycircularvirus 6
54. GQ365709 (CN) Sclerotinia sclerotiorum hypovirulence associated DNA virus 1
55. KM510192 (NZ) Bromus-associated circular DNA virus 3
56. KM821747 (NZ) Sewage-associated gemycircularvirus 1
57. KJ547638 (NZ) Sewage derived gemycircularvirus 1
58. KT862253 (NZ) Faeces associated gemycircularvirus 22
59. JQ412057 (GH) Cassava associated gemycircularvirus 1
60. Momordica charantia associated gemycircularvirus - MoaGmV
61. MH047857 (BR) Momordica charantia associated gemycircularvirus
62. KF413620 (VN) Hypericum japonicum associated circular DNA virus
63. KY230614 (BR) Bemisia-associated genomovirus AdO
64. KT732797 (TO) Pteropus associated gemycircularvirus 3
65. MH047858 (BR) Euphorbia heterophylla associated gemycircularvirus
66. KM598385 (USA) Odonata associated gemycircularvirus 1
67. JX185429 (USA) Dragonfly-associated circular virus 2
0.1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67

Figure 2. Phylogenetic analysis and SDT using 60 representatives genomoviruses. **A.** Bayesian phylogenetic tree based on Rep aminoacids sequences. Numbers at nodes represent posterior probabilities. Geminiviruses sequences 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 in the Akaike Information Criterion (AIC). The analyzes were carried out by running 100 million generations and sampling every 2000 generations after 2 million burn-in generation. **B.** Genome-wide pair-wise matrix generated by SDT v1.2. Phylogenetic analysis and SDT using 60 representatives genomoviruses sequences added to sequences of EcaGmV, VoaGmV, ByaGmV, GmaV1, TeaGmV, OuaGmV and MoaGmV (in black). The order, isolates names and accession number (GenBank) are: **1.** Spinach curly top Arizona virus (HQ443515); **2.** *French bean severe leaf curl virus* (JX094280); **3.** *Wheat dwarf virus* (EF536860); **4.** *African cassava mosaic virus* (FM877473); **5.** *Turnip leaf roll virus* (KT388088); **6.** *Beet curly top virus* (AF379637); **7.** *Faeces associated gemycircularvirus* 12 (KF371630); **8.** *Faeces associated gemycircularvirus 11* (KF371631); **9.** *Human genital-associated circular DNA virus-1* (KJ413144); **10.** Eucalyptus Urophilla associated genomovirus (EcaGmV); **11.** Vochisia rufa associated genomovirus (VoaGmV); **12.** Byrsonima crassiflora associated genomovirus (ByaGmV); **13.** *Sewage-associated gemycircularvirus 4* (KJ547634); **14.** *Gila monster - associated gemykrogvirus* (GmaV1); **15.** *HCBI9.212 virus* (LK931484); **16.** *Genomoviridae* sp. (MK032706); **17.** *Gila monster-associated gemycircularvirus* (MH378453); **18.** C*aribou feces-associated gemycircularvirus* (KJ938717); **19.** *Pacific flying fox faeces associated gemycircularvirus-7* (KT732800); **20.** Tecoma stans associated gemykolovirus (TeaGmV); **21.** -Common bean-associated gemycircularvirus (KX434770); **22.** Common bean-associated gemycircularvirus (KX434768); **23.** *Pacific flying fox faeces associated gemycircularvirus-6* (KT732798); **24.** *Faeces associated gemycircularvirus 10* (KF371632); **25.** *Pacific flying fox faeces associated gemycircularvirus-1* (KT732790); **26.** *Sewage-associated gemycircularvirus 5* (KJ547635); **27.** *Human genital-associated circular DNA virus-1* (KJ413144); **28.** *Faeces associated gemycircularvirus 15* (KT862254); **29.** *Meles meles fecal virus Gemygorvirus* (JN704610); **30.** *Dragonfly-associated circular virus 3* (JX185428); **31.** Plant associated genomovirus 2 (MH939414); **32.** Ouratea duparquetiana associated gemikibivirus (OuaGmV); **33.** *Dragonfly associated Gemykibivirus 1* (JX185430); **34.** *Sewage associated gemycircularvirus 3* (KJ547643); **35.** *Black robin associated gemykibivirus 1* (KF371634); **36.** *HCBI8.215 virus* (LK931483); **37.** Gemycircularvirus SL1 Gemykibivirus (KP133075); **38.** *Rhinolophus associated gemykibivirus 1* (KJ641737); **39.** *Pacific flying fox faeces associated gemycircularvirus-10* (KT732804); **40.** Poecile atricapillus GI tract-associated gemycircularvirus (KT309029); **41.** *Pteropus associated gemycircularvirus 8* (KT732806); **42.** *Pteropus associated gemycircularvirus 5* (KT732801); **43.** *Faeces associated gemycircularvirus 17* (KT862242); **44.** *Miniopterus associated gemycircularvirus 1* (KJ641719); **45.** *Black robin associated gemykibivirus 1* (KF371634); **46.** *Faecal-associated gemycircularvirus 1* (KF371641); **47.** *Faeces associated gemycircularvirus 3* (KF371639); **48.** *Faeces associated gemycircularvirus 4* (KF371638); **49.** *Poaceae-associated gemycircularvirus 1* (KT253577); **50.** *Soybean leaf-associated gemycircularvirus 1* (KT598248); **51.** *Faeces associated gemycircularvirus 16* (KT862251); **52.** *Pacific flying fox faeces associated gemycircularvirus-2* (KT732792); **53.** *Sewage-associated gemycircularvirus 6* (KJ547636); **54.** *Sclerotinia sclerotiorum hypovirulence associated DNA virus 1* (GQ365709); **55.** *Bromus-associated circular DNA virus 3* (KM510192); **56.** *Sewage-associated gemycircularvirus 1* (KM821747); **57.** *Sewage derived gemycircularvirus 1* (KJ547638); **58.** *Faeces associated gemycircularvirus 22* (KT862253); **59.** *Cassava associated gemycircularvirus 1* (JQ412057); **60.** Momordica charantia associated gemycircularvirus (MoaGmV); **61.** Momordica charantia associated gemycircularvirus (MH047857); **62.** *Hypericum japonicum associated circular DNA virus* (KF413620); **63.** Bemisia-associated genomovirus AdO (KY230614); **64.** *Pteropus associated gemycircularvirus 3* (KT732797); **65.** Euphorbia heterophylla associated gemycircularvirus (MH047858); **66.** *Odonata associated gemycircularvirus-1*; **67.** *Dragonfly-associated circular virus 2* (JX185429).

Phylogenetic analysis, shows **Figure 2A**, grouped the VoaGmV, ByaGmV and EcaGmV species into the same clade as the *Humam associated gemyvongvirus 1* (KP974693) species classified in the genus *Gemyvongvirus*. The species GmaV1, TeaGmV, OuaGmV and MoaGmV grouped into clades with representatives of the genera *Gemykrogvirus*, *Gemykolovirus*, *Gemykibivirus* and *Gemycircularvirus*, respectively.

Pairwise identity comparisons using the genome-full (nucleotides-nt) of the seven viruses with the sequences of representatives of all genera of the *Genomoviridae* family is show in **Figure 2B.** The isolate of EcaGmV was closer to the *Geminiviridae* species: *African cassava mosaic virus* - ACMV (FM877473) (59%), VoaGmV closest to *Beet curly top virus* - BCTV (AF379637) (61%).

The fact that these species share high identity with species within *Geminiviridae* is due to the relationship of the unique evolutionary origin of the Rep containing intron shared by some members of the *Geminiviridae* and shown virtually by all *Genomoviridae* (Zhao et al. 2019).

The isolates of EcaGmV, VoaGmV and ByaGmV showed 58-59% pairwise identity with *Humam associated gemyvongvirus 1* (KP974693). The pairwise identity between isolates EcaGmV, VoaGmV and ByaGmV of 59 to 61%.

Although the value of paired identity and phylogeny indicate that the isolates EcaGmV, VoaGmV and ByaGmV could be classified into the genus *Gemyvongvirus*. The characteristics of genomic organization show that they are distinct because they only have the rep gene.

The GmaV1 isolate shared 80% pairwise identity with the Gila monster-associated gemycircularvirus (MH378453) and *Caribou faeces-associated gemycircularvirus* (KJ938717) isolates belonging to the genus *Gemykrogvirus*.

The TeaGmV isolate shared 67% pairwise identity with Common bean-associated gemycircularvirus (KX434768) and Pacific flying fox faeces associated gemycircularvirus (KT732798) isolates belonging to the genus *Gemykolovirus*.

The OuaGmV isolate shared 74% and 72% pairwise identity with Plant associated genomovirus 2 (MH939414) and *Dragonfly associated gemykibivirus 1* (JX185430) isolates, respectively, both belonging to the genus *Gemykibivirus*.

The MoaGmV isolate showed 90% identity pairwise with another MoaGmV isolate (MH047857) belonging to the genus *Gemycircularvirus*.

Our results show species classified into four of the nine genera of the family (*Gemykrogvirus*, *Gemykolovirus*, *Gemykibivirus* and *Gemycircularvirus*), and three species of the family *Genomoviridae* that can be accommodated in a new genus.

The *Genomoviridae* family is composed of many viruses associated with various organisms, including several recovered from plant leaf samples (Varsani and Krupovic 2017).

Only representatives of species classified in the genus *Gemycircularvirus* were found in association with plants: *Cassava associated gemycircularvirus 1* (Dayaram et al. 2012); *Hypericum japonicum associated gemycircularvirus 1* (Du et al. 2014); *Poaceae associated gemycircularvirus 1* (Male et al. 2015); *Bromus associated gemycircularvirus 1* (Kraberger et al., 2015); *Soybean associated gemycircularvirus 1* (Marzano and Domier, 2016). However, recently in Brazil three Gemycircularvirus species were discovered on common bean-associated gemycircularvirus (Lamas et al. 2017), Euphorbia heterophylla associated gemycircularvirus and Momordica charantia associated gemycircularvirus (De Rezende et al. 2018).

Caribou faeces associated virus, classified in the genus *Gemykrogvirus*, was found in reindeer feces, based on the hypothesis that the virus was replicating in plants ingested by the animal. A biological assay was carried out on *Nicotiana benthamiana* plants, demonstrated its ability to replicate in this plant, however, without symptoms (Ng et al. 2014). Oother species of the genus were found associated with cattle (*Bovine associated gemykrogvirus 1*) and in sewage (*Sewage derived gemykrogvirus 1*) (Varsani and Krupovic, 2016).

The genera *Gemyvongvirus*, *Gemyduguivirus*, *Gemykroznavirus* and *Gemytondvirus* present only one species, and described associated with humans (Zhang et al. 2016), insects (Rosario et al. 2012), mammals (Sikorski et al. 2013), and bats. (Sikorski et al. 2013), respectively.

In the genus *Gemykolovirus*, viruses associated with bats are present while in the genus *Gemygorvirus* there are species found in birds, canids and sewer. The genus *Gemykibivirus* have

more species and a larger number of possible hosts, namely: mammals (including humans), birds and sewer (Varsani and Krupovic, 2016).

Recombination analyzes for nucleotide sequence indicated that only the TeaGmV species showed evidence of being recombinant in seven statistical methods RDP (p-value = 6.62E-18), GENECONV (p-value = 1.657E-11), BootScan (p -value = 2,468E-23), MaxChi (p-value = 3,362E-21), Chimera (p-value = 1,951E-11), SiScan (p-value = 6,654E-34) and 3Seq (p-value = 6.535E-25). Theanalyzes showed that the TeaGmV sequences in the analyzed dataset most closely resemble the parent sequence; the largest parent is an isolate of Soybean leaf-associated gemycircularvirus 1 (KT598248), and the smallest parent is the common bean-associated gemycircularvirus (KX434770). The initial breakpoint is at nucleotide 1144 and the final breakpoint is at nucleotide 2184 involving sequences. from Rep.

Viruses are constantly changing to adapt to new hosts. Recombination is an important evolutionary ally. In addition, studies have shown that some viruses function as 'helper viruses' allowing other viruses to access that cellular machinery (Imai et al. 2019). Genomoviruses are not known to have the movement protein, however, they have been reported in plants worldwide. The fact that the first species of the *Genomoviridae* family was found on a fungal host (*Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1*) (Yu et al. 2010, 2013), raised hypotheses that genomoviruses would be associated with fungi present in the plants. So far, the genomovirus-plant interaction remains unknown. Another relevant fact in the classification of these new genomes in the *Genomoviridae* family is due to the recognition of viruses identified through metagenomic studies (Varsani and Krupovic, 2016). Despite the importance of traditional criteria (including biological properties: such as host range, pathology, vectors, particle shape, and sequence data). A new approach has been taken for this family based only on sequence data (Simmonds et al. 2017).

4. Conclusion

The *Genomoviridae* family comprises viruses spread around the world and associated with different organisms, including plants. In this work we characterize, via HTS, and detected three novel genomoviruses three: Vochysia rufa associated genomovirus - VoaGmV, Byrsonima crassiflora associated genomovirus - ByaGmV, Eucalyptus urophylla associated genomovirus - EcaGmV. Two new species Tecoma stans associated gemykolovirus - TeaGmV (*Gemykolovirus*) and *Ouratea duparquetiana* associated gemykibivirus - OuaGmV (*Gemykibivirus*). In addition, two isolates of the species: *Gila monster-associated gemycircularvirus* - (*Gemykrogvirus*) - GmaV1 e Momordica charantia associated gemycircularvirus - MoaGmV (*Gemycircularvirus*) all associated with native and cultivated plants present in the *Cerrado* biome. Our studies area a first step of further research to better understand the nature of the interaction of these rare viruses with plants.

References

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.

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403-410.

Boiteux L (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.

Chiumenti M, Greco C, Antelmi I, et al (2019) Molecular characterisation of a novel gemycircularvirus associated with olive trees in Italy. Virus Research 263:169-172.

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, et al (2015) Fecal virome analysis of three carnivores reveals a novel nodavirus and multiple gemycircularviruses. Virology Journal 12:79.

da Silva Assis MR, Vieira CB, Fioretti JM, et al (2016) Detection and molecular characterization of gemycircularvirus from environmental samples in Brazil. Food and Environmental Virology 8:305-309.

Dayaram A, Opong A, Jäschke A, et al (2012) Molecular characterisation of a novel cassava associated circular ssDNA virus. Virus Research 166:130-135.

Dayaram A, Potter KA, Pailes R, et al (2015) Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. Infection, Genetics and Evolution 30:278-287.

de Rezende RR, Mar TB, Páez LMC, et al (2018) Complete genome sequences of two gemycircularviruses associated with non-cultivated plants in Brazil. Archives of Virology 1-4.

Du Z, Tang Y, Zhang S, et al (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.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792-1797.

Fontenele RS, Lacorte C, Lamas NS, et al (2019) Single stranded DNA viruses associated with capybara faeces sampled in Brazil. Viruses 11:710.

Gorbalenya AE, Koonin E V, Wolf YI (1990) A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Letter 262:145-148.

Halary S, Duraisamy R, Fancello L, et al (2016) Novel single-stranded DNA circular viruses in pericardial fluid of patient with recurrent pericarditis. Emerging Infectious Diseases journal 22:1839.

He Z, Zhang H, Gao S, et al (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Research 44:W236-W241.

Ilyina T V, Koonin E V (1992) Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Research 20:3279-3285.

Imai R, Nagai M, Oba M, et al (2019) A novel defective recombinant porcine enterovirus G virus carrying a porcine torovirus papain-like cysteine protease gene and a putative anti-apoptosis gene in place of viral structural protein genes. Infection, Genetics and Evolution 75:103975.

Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. Journal of Virology Methods 116:209-211.

Koonin EV (1993) A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Research 21:2541-2547.

Kraberger S, Farkas K, Bernardo P, et al (2015) Identification of novel Bromus-and Trifolium-associated circular DNA viruses. Archives of Virology 160:1303-1311.

Kraberger S, Hofstetter RW, Potter KA, et al (2018) Genomoviruses associated with mountain and western pine beetles. Virus Research 256:17-20.

Kreuze JF, Perez A, Untiveros M, et al (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.

Lamas NS, Fontenele RS, Melo FL, et al (2016) Complete genome sequence of a genomovirus associated with common bean plant leaves in Brazil. Genome Announcements 4:e01247-16.

Lamberto I, Gunst K, Müller H, et al (2014) Mycovirus-like DNA virus sequences from cattle serum and human brain and serum samples from multiple sclerosis patients. Genome Announcements 2:e00848-14.

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:e01144-15.

Martin DP, Murrell B, Golden M, et al (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution 1.

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

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

Nakasu EYT, Melo FL, Michereff-Filho M, et al (2017) Erratum to: Discovery of two small circular ssDNA viruses associated with the whitefly Bemisia tabaci. Archives of Virology 162:3563-3566.

Nash TE, Dallas MB, Reyes MI, et al (2011) Functional analysis of a novel motif conserved across geminivirus Rep proteins. Journal of Virology 85:1182-1192.

Ng TFF, Chen L-F, Zhou Y, et al (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.

Phan TG, Mori D, Deng X, et al (2015) Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. Virology 482:98-104.

Posada D (2008) jModelTest: phylogenetic model averaging. Molecular Biology and

Evolution 25:1253-1256.

Rambaut A (2012) FigTree v1. 4. Molecular evolution, phylogenetics and epidemiology. Edinburgh Univ Edinburgh, Institute of Evolutionary Biology and Ecology.

Richet C, Kraberger S, Filloux D, et al (2019) Novel circular DNA viruses associated with Apiaceae and Poaceae from South Africa and New Zealand. Archives of Virology 164:237-242.

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

Sikorski A, Massaro M, Kraberger S, et al (2013) Novel myco-like DNA viruses discovered in the faecal matter of various animals. Virus Research 177:209-216.

Simmonds P, Adams MJ, Benk\Ho M, et al (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nature Reviews Microbiology 15:161.

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:vew037.

Vega-Rocha S, Byeon I-JL, Gronenborn B, et al (2007) Solution structure, divalent metal and DNA binding of the endonuclease domain from the replication initiation protein from porcine circovirus 2. Journal of Molecular Biology 367:473-487.

Yu X, Li B, Fu Y, et al (2010) A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. Proceedings of the National Academy of Sciences 107:8387-8392.

Yu X, Li B, Fu Y, et al (2013) Extracellular transmission of a DNA mycovirus and its use as a natural fungicide. Proceedings of the National Academy of Sciences 110:1452-1457.

Zhang W, Li L, Deng X, et al (2016) Viral nucleic acids in human plasma pools. Transfusion 56:2248-2255.

Zhao L, Lavington E, Duffy S (2019) A comprehensive genealogy of the replication associated protein of CRESS DNA viruses reveals a single origin of intron-containing Rep. BioRxiv 687855.

Chapter 4

Broad spectrum of host range for *Geminiviridae* members

Resumo

A expansão das fronteiras em áreas nativas do Cerrado tem possibilitado o trânsito de pragas e doenças entre os ecossistemas nativos e agrícolas. As doenças causadas por vírus da família *Geminiviridae* geram grandes perdas na agricultura. Dessa forma, o objetivo deste trabalho foi avaliar o potencial de espécies arbóreas, daninhas e agrícolas como potenciais hospedeiras de geminiviruses. Para isto, foram coletadas amostras de folhas (296 de 131 espécies e 49 famílias) em três microambientes do bioma Cerrado: antrópico; área de conservação e área de transição. Extrações de DNA total, purificado foram enriquecidas em moléculas circulares via Rolling Circle Amplification - RCA e submetidas ao sequenciamento de Illumina HiSeq2500. As análises metagenômicas permitiram recuperarnove espécies da família *Geminiviridae* em diferentes ambientes. Uma espécie nova denominada Macroptilium bright yellow interveinal virus - MaBYIV (*Begomovirus*) foi recuperada em macroptilium (*Macroptilium erythroloma)*. Esta mesma amostra estava infectada com *Bean golden mosaic virus* - BGMV. Ensaios biológicos foram conduzidos utilizando *Bemisia tabaci* MEAM 1 (*Middle East Asian Minor 1*) e confirmaram a transmissão tanto de MaMYIV quanto de BGMV. Isolados das espécies TaGV1 (Tomato associated geminivirus 1) e ToALCV (Tomato apical leaf curl virus) (próximos ao *Capulavirus*) foram detectados em tomate (*Solanum lycopersicum*). Dois isolados de TaGV1, um de tomate e outro de beterraba (*Beta vulgaris*) apresentaram sequência divergente das sequências já descritas. Em ensaios biológicos, via biobalística, um dos isolados de TaGV1 foi capaz de induzir sintomas em *Nicotiana benthamiana* mas não foi capaz de replicar em tomate e beterraba. A espécie ToALCV foi detectada pela primeira vez no Brasil, apresentando 96% de identidade com isolados da Argentina. Foram realizados ensaios de biobalística em plantas de tomate e *N. benthamiana*, entretanto não houve infecção. A espécie Maize striate mosaic virus - MSMV (*Mastrevirus*) foi detectada em milho (*Zea mays*) e na nova hospedeira, cana de açúcar (*Saccharum officinarum*). Isolados de espécies do gênero *Begomovirus*, como BGMV foram detectados em berinjela (*Solanum melongena*), angico (*Anadenanthera colubrina*) e macroptilium; um isolado de *Tomato severe rugose virus* - ToSRV em algodão (*Gossypium hirsutum*) e dois isolados de *Tomato mottle leaf curl virus* - ToMoLCV, sendo um em leiteiro (*Euphorbia heterophylla*) e o outro em *Ouratea duparquetiana*. Além disto foi possível detectar também vinte e seis isolados de ToSRV em tomate (11), berinjela (11) joá de capote (*Nicandra physalodes*) (1) e sida (*Sida* spp); sete isolados de BGMV em feijão; quatro isolados de *Sida micrantha mosaic virus* - SiMMV em algodão (2), guanxuma (1) e joá de capote (1) e *Sweet potato leaf curl virus* - SPLCV em batata-doce (*Ipomoea batatas*).

Abstract

The expansion of borders in native areas of the Cerrado has enabled the transit of pests and diseases between native and agricultural ecosystems. Diseases caused by viruses of the Geminiviridae family generate large losses in agriculture. Thus, the objective of this work was

to evaluate the potential of tree, weed and agricultural species as potential hosts of geminiviruses. For this, leaf samples (296 of 131 species and 49 families) were collected from three microenvironments of the Cerrado biome: anthropic; conservation area and transition area. Purified DNA preparations were enriched in circular molecules via Rolling Circle Amplification - RCAand subjected to Illumina HiSeq2500 sequencing. Metagenomic analyzes identified nine species of the *Geminiviridae* in different areas. A novel species Macroptilium bright yellow interveinal virus - MaBYIV (*Begomovirus*) was recovered from macroptilium (*Macroptilium erythroloma*). The same sample showed mixed infection with Bean golden mosaic virus - BGMV. Biological assays were performed using *Bemisia tabaci* MEAM 1 (Middle East Asian Minor 1) and confirmed the transmission of MaMYIV and BGMV. Isolates of TaGV1 (Tomato associated geminivirus 1) and ToALCV (Tomato apical leaf curl virus) species (closed to Capulavirus genus) were detected in tomato (*Solanum lycopersicum*). Two isolates of TaGV1, being one of tomato and other beet (*Beta vulgaris*) showed divergent sequence from the sequences already described. In bioassay biological assays, one of the TaGV1 isolates was able to induce symptoms in *Nicotiana benthamiana* but was unable to replicate in tomato and beet. The ToALCV species was first detected in Brazil, showing 96% identity with isolates from Argentina. Biobalistics assays were performed in tomato and *N. benthamiana* plants, however there was no infection. Maize striate mosaic virus - MSMV (*Mastrevirus*) was detected in maize (*Zea mays*) and by the first time in sugarcane (*Saccharum officinarum*). Isolates of *Begomovirus* species such as BGMV were detected in eggplant (*Solanum melongena*), angico (*Anadenanthera colubrina*) and macroptilium; one isolate of *Tomato severe rugose virus* - ToSRV on cotton (*Gossypium hirsutum*) and two isolates of *Tomato mottle leaf curl virus* - ToMoLCV in *Euphorbia heterophylla* and the other in *Ouratea duparquetiana*. In addition it was possible to detect twenty-six isolates of ToSRV in tomato (11), eggplant (11) “shoo-fly plant” (*Nicandra physalodes*) (1) and guanxuma (Sida spp); seven BGMV isolates in beans; four isolates of *Sida micrantha mosaic virus* - SiMMV in cotton (2), guanxuma (1) and “shoo-fly plant” (1) and *Sweet potato leaf curl virus* - SPLCV in sweet potato (*Ipomoea batatas*).

1. Introduction

The *Cerrado* Biome is considered a biodiversity *hotspot* however, the expansion of agricultural borders has threatened this biome and favored the transit of pests and diseases. Among the diseases caused by viruses of the family *Geminiviridae* generate large losses in agriculture.

Geminiviridae is a family that groups important plant viruses that cause losses in agriculture around the world. Actually, the family consists of the following genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Glabrovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus* (Varsani et al. 2017). Additionally, the species *Citrus chlorotic dwarf associated virus* (CCDaV) (Loconsole et al. 2012) and *Mulberry mosaic dwarf associated virus* (MMDaV (Lu et al. 2015; Ma et al. 2015). To be considered a family member, the following

criteria are observed: host ranges, insect vector, genomic organization and phylogenetic relationships (ICTV, 2019). Geminiviruses have icosahedral particle and the genome may have one or two components, circular DNAs of 2.5-5.2 kb transmitted by different vectors including whitefly, leafhoppers, aphids and treehoppers (ICTV, 2019). Geminiviruses genomes can have up to 8 genes and two are highly conserved, which encode the rep (replication associated protein) and cp (protein coat) proteins. In addition to all also possess the movement protein (mp). The genomes of becurtoviruses, capulaviruses, grabloviruses, eragroviruses, mastreviruses, and CCDaVs all have long intergenic region (LIR) and short intergenic region (SIR) (Argüello-Astorga and Ruiz-Medrano 2001; Londoño Riego-Ruiz et al. 2010; Varsani et al. 2017). Introns have already been described in becurtoviruses, capulaviruses, grabloviruses, mastreviruses, CCDaV and MMDaV in Rep and RepA (Heckel et al. 1997, Golden et al. 2013, Filloux et al. 2014, Heydarnejad et al. 2017). Recombination has played a significant role in the evolution of geminiviruses (Padidam et al. 1999, Varsani et al. 2017).The use of RCA (Inoue-Nagata et al. 2004) and the high-throughput sequencing (HTS) allied to metagenomic analyzis (Roossinck et al. 2015, Simmonds et al. 2017; Bernardo et al. 2018), have been contributed to discovery a new plant virus, including geminivirus. In South America, new monopartite divergent geminiviruses were recently identified infecting tomato (*Solanum lycopersicum*) Tomato associated geminivirus 1 (TaGV1) in Brazil (Fontenele et al. 2018) and tomato apical leaf curl virus (ToALCV) in Argentina. Additionally, new species of the genus *Mastrevirus* were found infecting leafhopper (*Dalbulus maidis*) and maize (*Zea mays*) in Brazil (Fontenele et al. 2018).

Members of the genus *Mastrevirus* are transmitted by leafhoppers (Gutierrez 1999; Webb 1987). The genome is monopartite (size 2.5-2.7 kb) consists of four genes that encode proteins: movement protein (MP/AV2) and coat protein (CP/AV1) in the viral sense; Rep (AC1:AC2) and RepA (AC2) proteins associated with replication in the complementary sense. Rep is expressed

by a gene transcription splicing mechanism for C1:C2. And two intergenic regions (large intergenic region [LIR] and small intergenic region [SIR]) (Wright et al. 1997).

Members of the genus *Begomovirus* are transmitted by whiteflies (*Bemisia tabaci*). The genome is monopartite (DNA-A) and bipartite (DNA-A and DNA-B). The genome bipartite consists four open reading frames (ORFs): one encoding the coat protein (AV1) (but the AV2 ORF is only found in monopartite begomoviruses and 'Old World') in the virion sense, and four ORFs/proteins (AC1/Rep; AC2/TrAP; AC3/Ren; AC4 and AC5) in the complementary sense strand. Two genes encoding viral DNA movement-associated proteins (BV1/NSP and BC1/MP) were annotated in the DNA-B genome (Rojas et al. 2005, Stanley et al. 2005, Li et al. 2015).

The genus *Begomovirus* is a group of whitefly-transmitted plant virus species (*Bemisia tabaci*). Begomoviruses are divided into monopartite (DNA-A) and bipartite (DNA A and B), however, only DNA-A is used for taxonomic classification (Brown et al. 2012). Advances in sequencing technologies and use of bioinformatics have favored the discovery of new begomovirus species. Species of the genus *Begomovirus* have been reported to infect potato plants (*Ipomoea batatas*), cassava (*Manihot esculenta*), beans (*Phaseolus vulgaris*), tomatoes (*Solanum lycopersicum*), soybean (*Glycine max*), and weeds such as *Macroptilium* sp., *Calopogonium* sp. and *Desmodium* sp. Weed infections are known to serve as a source of inoculum for crops of agronomic interest. In Brazil, is *Bean golden mosaic virus* - BGMV is an important species widely distributed in the country (Sobrinho et al. 2012).

In order to know more about the role of tree species (native and exotic), weeds and cultivated as hosts and/or reservoirs of species of the *Geminiviridae* family, this study was conducted in samples collected from the Cerrado biome.

2. Materials and Methods

2.1 Plant material collection and High-throughput sequencing - HTS

In 2016, a total of 296 symptomatic samples were collected, 81, 18 and 35 of native tree species, exotic agricultural, respectively, comprising 44 botanical families, from the urban forestation of DF - AUDF (26 samples – thirteen tree species and one weed), Viveiro II of the Urbanization Company of Nova Capital - VII-Novacap (79 samples - 29 tree species), Experimental Area of Clean Water Farm belonging to the University of Brasília (UnB) - AE-FAL (50 samples - 20 agricultural species) and NRT (21 samples - 2 agricultural species, 2 weeds), anthropic environments; and from the Ecológical Station of Amended Water - ESECAE (30 samples - 18 tree species) and Ecológical Reserve of the Brazilian Institute of Geography and Statistics - RECOR (30 samples - 20 tree species) - conservation environments; of the Experimental Biology Station (UnB) in conservation environment a possible "Relevant Ecológical Interest Area" (ARIE-EEB) (29 samples - 20 tree species) and agricultural cultivars Experimental Vegetable Garden Area of the Experimental Biology Station (AHE-EEB) 31 samples - 22 plant species, being considered here a transitional environment due to the proximity. The locations of the environments are present in **Figure 1** of chapter 2 of this work and the collected plant species are listed in **Table ST1** of chapter 2 of this work.

The samples were submitted to total DNA extraction (protocol adapted from Boiteux et al. 1999) and after RCA enrichment (Inoue-Nagata et al. 2004). The samples were pooled according to **Table 1** of Chapter 2 of this paper and sent to the Hight-throughput sequence (HTS) using the Illumina Hiseq2500 platform.

Metagenomic analyzes were performed at LVV-Fito / UnB and at the Cell Biology Laboratory / UnB according to analyzes performed by other research groups (Kreuze et al. 2009, Adams and Fox 2016). From the sequences obtained by HTS, overlapping specific primers (in the Geneious R11 program for obtaining the complete genome in the individual samples) were designed and obtained in the literature and are listed in **Table ST2** chapter 2 of this work.

The complete methodologies used for each virus are described in the related articles.

3. Results

Metagenomics analyzes allowed nine species of the *Geminiviridae* family identification. A novel species MaBYIV (Macroptilium bright yellow interveinal virus), TaGV1 (Tomato associated geminivirus 1), ToALCV (Tomato apical leaf curl virus) (both not classified by genus), Maize striate mosaic virus - MSMV (*Mastrevirus*), *Bean golden mosaic virus* - BGMV and a new species called - MaBYIV. Beside this, isolates of *Begomovirus* species such as BGMV were detected in eggplant (*Solanum melongena*), angico (*Anadenanthera colubrina*) and macroptilium; one isolate of *Tomato severe rugose virus* - ToSRV on cotton (*Gossypium hirsutum*) and two isolates of *Tomato mottle leaf curl virus* - ToMoLCV in *Euphorbia heterophylla* and the other in *Ouratea duparquetiana*. In addition it was possible to detect twenty-six isolates of ToSRV in tomato (11), eggplant (11) shoo-fly plant (*Nicandra physalodes*) (1) and guanxuma (*Sida* spp); seven BGMV isolates in beans; four isolates of *Sida micrantha mosaic virus* - SiMMV in cotton (2), guanxuma (1) and "shoo-fly" (1) and *Sweet potato leaf curl virus* - SPLCV in sweet potato (*Ipomoea batatas*).

In NRT, was detected isolates of BGMV in eggplant by the first time. Samples of weeds *Sida* sp. (Malvaceae) and *Nicandra physalodes* (Solanaceae) were positive for SiMMV and ToSRV. Eleven plants eggplant were detection ToSRV. Only one sweet potato plant was found to be positive for the SPLCV infection.

In AE-FAL were detected different species of *Geminiviridae* family. Isolates of: BGMV in beans, ToSRV in tomatoes and cotton, SiMMV in cotton, SPLCV in sweet potato.

Here, we report by the first time a natural infection of cotton samples by ToSRV. In Brazil, the first and only report of bipartite begomovirus species in cotton (*Cotton chlorotic spot virus* - CCSV), was done in Paraíba (de Almeida et al. 2013).

The EEB is considered in this work a transitional environment due to the proximity between the cultivated areas (AHE-EEB) and conservation areas (ARIE-EEB). Isolates of BGMV were detected in beans. ToSRV detected in tomato, SPLCV in sweet potato, ToMoLCV in tomato, wild poinsettia (*Euphorbia heterophylla* - Euphorbiaceae) and *Ouratea duparquetiana* (Ochnaceae). Mixed infections were observed in samples from this microenvironment. One sample of tomato was infected with ToSRV and ToMoLCV isolates.

On the other hand, Euphorbiaceae species are known begomovirus hosts (Fernandes et al. 2011) while no virus records were done for Ochnaceae species (Virus Host Database).

In AUDF, BGMV and MaBYIV were detected in *Macroptilium erythroloma* (Fabaceae). The role of the *M. erythroloma* plant as a source of inoculum of these viruses (mixed infection) in a whitefly transmission assay for plants important for agriculture was evaluated. The first report of BGMV infecting *M. erythroloma* and the result of the assay for this virus are described in Batista et al 2020a **(ATTACHED I).** The MaBYIV sequence showed 84% identity and 100% coverage, according to the criterion for new *Begomovirus* species the threshold is 91% paired identity (Brown et al. 2015), therefore a new species. The report of this new species of begomovirus and the result of the transmission assays are in Batista et al 2020b **(ATTACHED II).**

In AE-FAL ToALCV and TaGV1 were detected in tomato and beet and tomato, respectively. The newly discovered ToALCV species in Argentina (Vaghi Medina et al. 2017) was found for the first time in Brazil and after performing molecular and biological assays were described in Batista et al. 2018 **(ATTACHED III).** The TaGV1 species (Fontenele et al. 2017) after molecular characterization proved to be a divergent sequence from those described for TaGV1. Besides being found a new host, the beet) and performed biological tests being described in Batista et al. 2020c **(ATTACHED IV).**

In AHE-EEB were detected BGMV in tree species *Anadenanthera colubrina*. These results are showed in Batista et al. 2020d **(ATTACHED V)**.

In AHE-EEB were detected MSMV in maize and by the firs time in sugarcane (*Sugarcane officinarum*). These results are described in Batista et al. 2020e **(ATTACHED VI)**.

BGMV isolate detected and recovered in eggplant was included in the disease note: Natural Infection of Tomato, Eggplant and *Nicandra physalodes* (Solanaceae) by *Bean golden mosaic virus* isolates in Brazil. This disease note will be part of the work of doctoral student Luciane Reis and therefore will not enter the version of this Thesis.

4. APPENDIX – Attached and publications (accepted, submitted and to be submitted) from these Thesis.

Batista; JG., Pereira-Carvalho RC., Malheiros MF., et al (2020a) Molecular Confirmation of *Macroptilium erythroloma* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil. (Plant Disease/Disease Note - **Submitted**, **Attached I**).

Batista; JG., Fochat; F., Nery FMB., et al (2020b) Characterization in *Macroptilium erythroloma* of a Novel Neotropical Legume-Infecting Bipartite Begomovirus species. (Viruses/ Full paper - **To be submitted**, **Attached II**).

Batista; JG., Melo; FL., Pereira-Carvalho RC., et al (2018) First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil. (Plant Disease/Disease Note **- Published, Attached III**).

Batista; JG., Fochat; F., Nery; FMB, et al (2020c) Characterization of genetically divergent Tomato associated geminivirus 1 isolates from table beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*). (Tropical Plant Pathology/Full paper - **Submitted, Attached IV**).

Batista; JG., Fochat; F., Nery; FMB., et al (2020d) First report of *Bean golden mosaic virus* in *Anadenanthera colubrina*. (Plant Disease/Disease Note - **To submitted**, **Attached V**).

Batista; JG., Fochat; F., Miranda; BEC., et al (2020e) Natural Infection of Sugarcane (*Saccharum officinarum*) by Maize striate mosaic virus (*Geminiviridae*) in Brazil. (Plant Disease/ Disease Note **- Accepted, Attached VI**).

Batista; JG., Fochat; F., Nery; FMB, et al (2020f) New Genomovirus associated with trees and cultivated species in Central Brazil. (Viruses/Full paper - **Chapter 3**).

Batista; JG., Fochat; F., Nery; FMB., et al (2020g) Single-stranded DNA viroma of vegetable and field crops and from native and exotic tree species growing in the Brazilian Cerrado biome. (Viruses/Full paper - **Chapter 2**).

Conclusões Gerais

Conclui-se que as espécies árboreas podem ser reservatórios de vírus importantes para agricultura e de vírus ainda desconhecidos.

Os vírus estão transitando entre os diferentes ambientes, entretanto em ambiente de conservação poucos vírus foram encontrados.

As espécies BGMV e ToSRV são predominantes na região do Cerrado. As plantas das famílias Fabaceae e Solanaceae são as principais hospedeiras de vírus.

Estudos usando o HTS aliado as análises metagenômicas e bioinformática possibilitam identificar vírus em diferentes ambientes e famílias botânicas.

Referências

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.

Argüello-Astorga GR, 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.

Batista JG, Melo FL, Pereira-Carvalho RC, et al (2018) First Report of Tomato apical leaf curl virus Infecting Tomato in Brazil. Plant Disease Journal.

Bernardo P, Charles-Dominique T, Barakat M, et al (2018) Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. EMBO J 12:173.

Bernardo P, Golden M, Akram M, et al (2013) Identification and characterisation of a highly divergent geminivirus: evolutionary and taxonomic implications. Virus Research 177:35–45. doi: 10.1016/j.virusres.2013.07.006.

Bozorgi N, Heydarnejad J, Kamali M, Massumi H (2017) Splicing features in the expression of the complementary-sense genes of Beet curly top Iran virus. Virus Genes 53:323–327.

Brown JK, Zerbini FM, Navas-Castillo J, et al (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Archives of Virology 160:1593–1619.

Candresse T, Filloux D, Muhire B, et al (2014) Appearances can be deceptive: revealing a hidden viral infection with deep sequencing in a plant quarantine context. PLoS One 9:e102945.

de Almeida MMS, Jain S, Barroso PAV, et al (2013) Complete sequence of a new bipartite begomovirus infecting cotton plants in Brazil. Genome Announcements 1:e00661-13.

de Viana ML, Giamminola E, Russo R, Ciaccio M (2014) Morphology and genetics of Anadenanthera colubrina var. cebil (Fabaceae) tree from Salta (Northwestern Argentina). Revista de biologia tropical 62:757–767.

Fernandes FR, Albuquerque LC, de Oliveira CL, et al (2011) Molecular and biological characterization of a new Brazilian begomovirus, euphorbia yellow mosaic virus (EuYMV), infecting Euphorbia heterophylla plants. Archives of Virology156:2063.

Fontenele RS, Alves-Freitas DMT, Silva PIT, et al (2018) Discovery of the first maize-infecting mastrevirus in the Americas using a vector-enabled metagenomics approach. Archives of Virology 163:263–267.

Fontenele RS, Lamas NS, Lacorte C, et al (2017) A novel geminivirus identified in tomato and cleome plants sampled in Brazil. Virus Research 240:175–179. doi: 10.1016/j.virusres.2017.08.007.

ICTV (2019) ICTV online. International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/. Accessed 21 Dec 2019.

Inoue-Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. Journal of Virological Methods 116:209–211.

Kreuze JF, Perez A, Untiveros M, et al (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.

Li F, Xu X, Huang C, et al (2015) The AC5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. New Phytol 208:555–569. doi: 10.1111/nph.13473.

Loconsole G, Saldarelli P, Doddapaneni H, et al (2012) Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. Virology 432:162–172. doi: 10.1016/j. Virology .2012.06.005.

Londoño A, Riego-Ruiz L, Argüello-Astorga GR (2010) DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. Archives of Virology 155:1033–1046.

Lu Q-Y, Wu Z-J, Xia Z-S, Xie L-H (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (Morus alba L.). Archives of Virology 160:2135–2138.

Ma Y, Navarro B, Zhang Z, et al (2015) Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease. Journal of General Virology 96:2421–2434. doi: 10.1099/vir.0.000175.

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

Roossinck MJ, Martin DP, Roumagnac P (2015) Plant virus metagenomics: advances in virus discovery. Phytopathology 105:716–727.

Simmonds P, Adams MJ, Benk\Ho M, et al (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nature Reviews Microbiology 15:161.

Varsani A, Roumagnac P, Fuchs M, et al (2017a) Capulavirus and Grablovirus: two new genera in the family Geminiviridae. Archives of Virology 162:1819–1831. doi: 10.1007/s00705-017-3268-6.

Varsani A, Roumagnac P, Fuchs M, et al (2017b) *Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. Archives of Virology 162:1819–1831.

Von Reis A (1964) Taxonomic Study of Anadenanthera Contributions from the Gray Herbarium of Harvard University 193:3–65.

Wright EA, Heckel T, Groenendijk J, et al (1997) Splicing features in maize streak virus virion- and complementary-sense gene expression. The Plant Journal 12:1285–1297.

APPENDIX – Attached and publications (published, accepted, submitted and to be submitted) from these Thesis.

Batista; JG., Pereira-Carvalho RC., Malheiros MF., et al (2020a) Molecular Confirmation of *Macroptilium erythroloma* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil. (Plant Disease/Disease Note - **Submitted**, **Attached I**).

Batista; JG., Fochat; F., Nery FMB., et al (2020b) Characterization in *Macroptilium erythroloma* of a Novel Neotropical Legume-Infecting Bipartite Begomovirus species. (Viruses/ Full paper - **To be submitted**, **Attached II**).

Batista; JG., Melo; FL., Pereira-Carvalho RC., et al (2018) First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil. (Plant Disease/Disease Note - **Published, Attached III**).

Batista; JG., Fochat; F., Nery; FMB, et al (2020c) Characterization of genetically divergent Tomato associated geminivirus 1 isolates from table beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*). (Tropical Plant Pathology/Full paper - **Submitted, Attached IV**).

Batista; JG., Fochat; F., Nery; FMB., et al (2020d) First report of *Bean golden mosaic virus* in *Anadenanthera colubrina*. (Plant Disease/Disease Note - **To submitted**, **Attached V**).

Batista; JG., Fochat; F., Miranda; BEC., et al (2020e) Natural Infection of Sugarcane (*Saccharum officinarum*) by Maize striate mosaic virus (*Geminiviridae*) in Brazil. (Plant Disease/ Disease Note - **Accepted, Attached VI**).

Batista; JG., Fochat; F., Nery; FMB, et al (2020f) New Genomovirus associated with trees and cultivated species in Central Brazil. (Viruses/Full paper - **Chapter 3**).

Batista; JG., Fochat; F., Nery; FMB., et al (2020g) Single-stranded DNA viroma of vegetable and field crops and from native and exotic tree species growing in the Brazilian Cerrado biome. (Viruses/Full paper - **Chapter 2**).

Attached I.

Molecular Confirmation of *Macroptilium erythroloma* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil (2020a)

Batista; JG., Pereira-Carvalho RC., Malheiros MF., et al - Plant Disease/Disease Note

Submitted

plant disease

Molecular Confirmation of Macroptilium erythroloma (Fabaceae) as a Natural Host of Bean golden mosaic virusin Brazil.

Journal:	*Plant Disease*
Manuscript ID	PDIS-12-19-2669-PDN
Manuscript Type:	Plant Disease Note
Date Submitted by the Author:	20-Dec-2019
Complete List of Authors:	Batista, Josiane; Universidade de Brasília, Fitopatologia Pereira Carvalho, Rita; Universidade de Brasília, Phytopathology Malheiros, Mateus; Universidade de Brasília, Phytopathology Rezende, Denise; Universidade de Brasilia, Plant Pathology Melo, Fernando; Universidade de Brasília, Biologia Celular dos Reis, Luciane de Nazaré; Universidade de Brasilia, Fitopatologia; Fonseca, Maria Esther; Embrapa Vegetable Crops, Plant Breeding Lab Boiteux, Leonardo; Embrapa Hortaliças, Plant Breeding
Keywords:	Virus, BGMV, Macroptilium erythroloma, Bemisia tabaci

Molecular Confirmation of *Macroptilium erythroloma* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil.

J. G. Batista; R. C. Pereira-Carvalho; M. F. Malheiros; D. V. Rezende; F. L. Melo and **L. N. A. Reis.** Departament of Plant Pathology, Universidade de Brasília (UnB), Brasília–DF, Brazil; **M. E. N. Fonseca** and **L. S. Boiteux**, National Center for Vegetable Crops Research (CNPH), Embrapa Hortaliças, Brasília–DF, Brazil.

Bean golden mosaic virus (BGMV) is a widely distributed *Begomovirus* species (family *Geminiviridae*), being responsible for severe yield losses in the bean (*Phaseolus vulgaris* L.) crop in Brazil. Nationwide surveys have detected BGMV isolates infecting cultivated as well as weed hosts (Bracero et al. 2003; Sobrinho et al. 2014). Here, plants of *Macroptilium erythroloma* (Mart. ex Benth.) Urban displaying golden mosaic-like symptoms were collected Brasilia–DF area in 2016. Total DNA extraction was performed from infected leaf samples using a modified CTAB protocol (Boiteux et al., 1999). Sanger dideoxy sequencing of the barcoding gene *rbcL* confirmed the taxonomic status of the sample. Purified DNA was also submitted to Rolling Circle Amplification (Inoue-Nagata et al. 2004) and sequenced on the Illumina Hiseq 2500 platform in Macrogen (South Korea). Sequence analysis was performed in the CLC Genomics Workbench program after manually refining and assembling the reads in higher quality contigs. Contigs were imported into the Geneious software and compared via BLASTn algorithm to a GenBank viral database. The *M. erythroloma* BGMV–derived contigs for DNA–A (MN822294) displayed identities of ≈ 99% with various BGMV isolates infecting bean (e.g. KJ939798) and *M. lathyroides* (e.g. KJ939766). The 2617 nts of the DNA–A component displayed all five ORFs: AV1 (CP) in the viral sense, and AC1 (Rep), AC2 (TrAP), AC3 (Ren), AC4 and AC5 in the complementary sense. The sequence was confirmed by Sanger. The contigs for DNA–B (MN822293) displayed 98.2% identity to one bean BGMV isolate from Minas Gerais State, Brazil (MG334553). The 2594 nts of the DNA–B component displayed two ORFs: BV1 (NSP) in the viral sense and BC1 (MP) in the complementary-sense. Plants of bean (*P. vulgaris* cv. Carioca), soybean (*Glycine max* cv. Monsoy 9144), peanut (*Arachis hypogaea* cv. BRS Havana), cowpea (*Vigna unguiculata* cv. Seridó), and tomato (*Solanum lycopersicum* cv. Santa Clara) were employed in infectivity assays. One week before the assay, BGMV–infected *M. erythroloma* plants were caged with viruliferous adults of *Bemisia tabaci* Middle East Asian Minor 1 (MEAM 1) whiteflies for virus

acquisition and then transferred to cages with five plants of each plant species. Onenon-inoculated plant of each species was employed as negative control. One apicalleaf sample per test plant was collected 15 days after inoculation (DAI). Detection of BGMV was performed 21 DAI, employing PCR with specific primers. PCR results indicated that BGMV isolate from *M. erythroloma* was transmitted to all fiveplant species. In the genus *Macroptilium*, BGMV has been reported thus far infecting only the species *M. lathyroides* (L.) Urban (Bracero et al. 2003; Sobrinho et al. 2014). However, *M. erythroloma* was assigned in the early 1980s (i.e. before the invasion of *B. tabaci* MEAM 1) as a putative host of BGMV based only upon symptoms and biological tests (Chagas et al. 1981). Therefore, we provide here the first formal biological and molecular confirmation that BGMV can naturally infect *M. erythroloma*. In addition, the BGMV isolate from *M. erythroloma* was able to infect accessions of bean, soybean, peanut, cowpea, and tomato under our experimental conditions. Therefore, the presence of *Macroptilium* weeds in association with commercial bean fields might be a major epidemiological concern, since they may serve as year-round sources ofBGMV inoculum for this crop.

References

Boiteux, L. S., et al. 1999. J. Am. Soc. Hort. Sci. 124: 32.

Bracero, V., et al. 2003. Plant Disease, 87: 1022.

Chagas, C. M. et al. 1981. Arq. Inst. Biol. 48: 113.

Inoue-Nagata, A. K., et al. 2004. J. Virol. Meth. 116: 209.

Sobrinho, R. R., et al. 2014. J. Gen. Virol. 95: 2540.

Funding: This research had financial support from grants from Embrapa, FAPDF, CAPES, and CNPq. J. G. Batista; M. F. Malheiros and L. N. A. Reis were supported by a scholarship from CNPq.

Attached II

Characterization in *Macroptilium erythroloma* of a Novel Neotropical Legume-Infecting Bipartite Begomovirus species (2020b)

Batista; JG., Fochat; F., Nery FMB., et al - Viruses/Full paper

To be submitted

viruses

MDPI

Article

Characterization in *Macroptilium erythroloma* of a Novel Neotropical Legume-Infecting Bipartite *Begomovirus* species.

Josiane Goulart Batista[1], Felipe Fochat Silva[1], Flávia Milene Barros Nery[1], Fernando Lucas de Melo[1], Leonardo Silva Boiteux[2], Mateus Fonseca Malheiros[1], Simone da Graça Ribeiro[3], Denise Vilela de Rezende[1], Rita de Cássia Pereira-Carvalho[1]

[1]Departament of Plant Pathology, Universidade de Brasília, 70910-900, Brasília, Brazil.

[2]Embrapa Hortaliças, Brasília, 70275-970, Brasília, Brazil.

[3]Embrapa Recursos Genéticos e Biotecnologia, Brasília, 70770-917, Brasília, Brazil.

Abstract: Begomoviruses are whitefly-transmitted, circular single–stranded DNA viruses, encapsidated in either one (monopartite) or two (bipartite) particles. The continuous presence of weed hosts around and within fields of major susceptible crops has intensified the flow of viruses as well as the exchange of genetic information among viral genotypes in mixed infections, creating a favorable scenario to the emergence of novel species in Neotropical areas. Here, high throughput sequencing was employed along with Sanger dideoxy sequencing to characterize a novel begomovirus species occurring in mixed infection with *Bean golden mosaic virus* – BGMV in the legume weed *Macroptilium erythroloma*. This new species (tentatively named as Macroptilium bright yellow interveinal virus – MaBYIV), showed a typical genomic organization of the New World bipartite begomoviruses. MaBYIV displayed recombination signals of the begomoviruses BGMV and *Tomato mottle leaf curl virus*. The original *M. erythroloma* plants (co-infected with MaBYIV and BGMV) were used in transmission assays with *Bemisia tabaci Middle East Asian Minor* 1. BGMV was transmitted to bean, cowpea, peanut, soybean, and tomato cultivars, whereas MaBYIV was transmitted only to bean and soybean cultivars. Our biological and molecular characterization indicated that *M. erythroloma* may play a role as a year–round reservoir of both MaBYIV as well as BGMV under Brazilian conditions.

Keywords: begomovirus; weed; whitefly; host range, epidemiology.

1. Introduction

Geminiviridae family is composed by viruses with circular single–stranded DNA (ssDNA) genome and twinned particles, which can cause serious production losses in a wide range of crops across the world [1-3]. Classification criteria in *Geminiviridae* include vector species, host range, genome organization, and pairwise sequence identities of the complete DNA–A genome [4-10]. Currently, *Geminiviridae* family comprises nine genera (*Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Grablovirus*, *Eragrovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus*) as well as two species (*Citrus chlorotic dwarf associated virus* and *Mulberry mosaic dwarf associated virus*), which are yet unclassified [8]. The high number of characterized *Begomovirus*

species is explained by their small circular DNA genomes associated with an array of mechanisms of generating genetic variability such as mutation, recombination and pseudo-recombination. These genetic/genomic features lead to a natural emergence of new species. Currently, several species of *Begomovirus* have been added to the genera considering the species demarcation criterion of 91% of pairwise identity of the complete genome of the DNA-A component [5]. This group of viruses is efficiently transmitted in a non-propagative circulative relation by whitefly [*Bemisia tabaci Middle East Asian Minor* 1 (MEAM1)] [11]. An increase of incidence and severity of diseases caused by species of this genus is due the introduction of *B. tabaci* MEAM1 in Brazil [12]. Whitefly is able to colonize about 600 species of plants distributed in 74 botanical families [13]. In Brazil, whiteflies can colonize various crops such as cabbage, cotton, soybeans, beans, tomatoes [14, 15]. Besides this, *B. tabaci* transmits other viral species of the genera *Begomovirus as Carlavirus*, *Ipomovirus*, *Crinivirus* and *Torradovirus* [16].

Begomovirus species exhibit twined particles encapsidating the ssDNA, monopartite (DNA–A) or bipartite (DNA–A and DNA–B) genomes, with 2.6 and 2.8 kb in size. The DNA–A presents to six ORFs (open reading frames). In the viral sense, AV1 gene codifies coat protein (CP) associated with vector specificity determination and viral movement. Begomovirus species from Old World Word presents AV2 gene codifying a protein involved in symptoms expression and viral movement. Four ORFs in complementary-sense are found AC1, AC2, AC3 and AC4 codifying a Replication protein (Rep), Trans-activating protein (TrAP), Replication enhancer protein (REN) and AC4 protein, respectively [17]. These proteins involved in a variety of functions are essential for the infectious cycle of the virus: replication, plant cell cycle interference, temporal regulation of viral gene expression, and suppression of host antiviral responses [17, 18]. The ORF AC5 has been annotated in some species, including Mungbean yellow mosaic India virus. AC5-encoded protein is involved in suppression of gene silencing [19]. The DNA–B codifies two 2 ORFs, BV1 gene (viral-sense) and BC1 gene (complementary–sense), codifying a nuclear shuttle protein (NSP) and movement protein (MP), respectively, facilitating the intracellular and intercellular transport [17, 20]. Cognates DNA-A and DNA-B components have no sequence identity except for a common region (CR) of approximately 200-250 bp that is highly conserved between both components of the same viral species [17, 18]. The CR includes a stem-loop structurecontaining the nonanucleotide TAATATTAC, as the TATA box sites and iterons associated with regulation of replication [21-23]. Thus, CR harbor vital regulatory signals that influences the replication and the coupling of the bipartite genomes [24]. The identification of novel begomoviruses is currently facilitated by the large number of methodological tools, including PCR assays with universal primers, rolling circle amplification (RCA) technique, high-throughput sequencing (HTS) and platforms of bioinformatics.

In the Neotropical areas, a wide array of *Begomovirus* species have been reported infecting legume crops such as bean (*Phaseolus vulgaris*), lima bean (*P. lanatus*), soybean (*Glycine max*), and cowpea (*Vigna unguiculata*) [25] as well as a wide array of weeds including species of *Calopogonium*, *Desmodium* and *Macroptilium* [16, 26]. *Bean golden mosaic virus* (BGMV) is one of the most important *Begomovirus* species in Brazil, being widely distributed and is responsible for yield losses of up to 100% in beans and lima beans [26]. Studies show a great diversity among BGMV isolates infecting cultivated and uncultivated hosts [27]. Weeds as may play an important epidemiological role acting as reservoirs for various viruses, favoring conditions for mixed infections and recombination events. Several studies have shown that recombination plays an important role in generating genetic variability in begomovirus in Brazil [28-31] and worldwide [32-36]. Recent epidemics of begomoviruses reflect favorable combinations of plant, vector, and viral (e.g. emergence

of a novel recombinant virus) factors. Such epidemics typically result in co-infection of plants with different begomovirus, leading to the appearance of further variants, especially recombinants. Whitefly is the main agent for the spread of begomovirus, making transit between cultivated and uncultivated hosts, contributing to virus permanence, evolution and epidemiology.

The genus *Macroptillum* (Benth.) Urban (*Leguminosae*: *Papilionoideae*: *Phaseolinae*) is widely disseminated in the Americas, occurring from the southern United States to Argentina and Uruguay. In Brazil, it is represented by at least ten species distributed across all geographic regions [37, 38]. Eight *Begomovirus* species have already been reported infecting species of the genus *Macroptilium* and all were accepted by the ICTV, four in Brazil [31], two in Jamaica [39, 40], one in Puerto Rico [41] and one in the United States [41]. In the present work, we performed HTS and metagenomic analysis of the viruses found in the weed *M. erythroloma*. A molecular and biological characterization new species of begomoviruses previously named Macroptilium bright yellow interveinal virus –– MaBYIV (DNA–A and DNA–B) occurring in mixed infection with BGMV (DNA–A and DNA–B) in *M. erythroloma* (Benth.) Urban.

2. Materials and Methods

2.1. Sample Collected and processing

As part of a project involving disease surveillance studies of native, agricultural, and wild species in Central Brazil, several foliar samples displaying virus-like symptoms were collected in 2016. Among these, samples of *Macroptilium* plants showing symptoms of bright yellow interveinal and foliar deformation were collected in Brasilia–DF. This plant species was identified as *M. erythroloma* by botanists and also by the Sanger dideoxy sequencing of the barcoding *rbcL* gene [42] at CNPH. The samples were photographed and processed for HTS. Total DNA extraction was performed [43] and used as template for RCA (Rolling Circle Amplification) assays [44]. The sample was sent for sequencing on the Illumina Hiseq 2500 platform in Macrogen (South Korea).

2.2. Molecular Characterization

The HTS was initially performed in the CLC Genomics Workbench program by cleaning and assembling the reads forming quality contigs. Contigs were imported into the Geneious program and were confronted with a GenBank viral database through BLASTn and BLASTx. Four contigs were related to isolates: Macroptilium yellow net virus isolate BR:Mur1:09 segment DNA–A (JN418998), Macroptilium yellow net virus isolate BR:Mur1:09 segment DNA–B (JN418999), Bean golden mosaic virus isolate BR:Par4:12 segment DNA–A (KJ939798) and Bean golden mosaic virus isolate Ponte Nova–MG segment DNA–B (MG334553) showing identities of 84.4%, 78.4 %, 99.9% and 98.2%, respectively. Abutting primers were designed based on the sequence obtained by HTS for a putative new species related to MaYNV (MaBYIV_F: 5'–CGT CTC CTT ATC GCA CGT TGA C–3' and MaBYIV_R: 5'–TCG CAT TGA CTT AGA GTG C–3'). In order to recover the full viral genome, the PCR was performed under the following conditions: 94 ºC for 3 minutes, 35 cycles of 94 ºC for 30 seconds, 60 ºC for 45 seconds, 72 ºC for 3 minutes, and a final extension of 72 ºC for 10 minutes. Specific primers [45] were used for detection of BGMV. The PCR products (amplicons) were purified and sequenced aiming to identify the associated ssDNA viruses.

2.3. Phylogenetic, Pairwise Identity and Recombination Analyses

To establish phylogenetic relationships, data were selected that include DNA–A sequences (n=30) and DNA–B sequences (n=24) from representative begomovirus species obtained from the NCBI database **(**www.ncbi.nlm.nih.gov), including both components (DNA–A and DNA–B) of our putative new species and of BGMV isolates. The components were evaluated in different trees. All sequences in the datasets were aligned with MUSCLE algorithm [46]. For tree, the nucleotide substitution model was chosen using jModeltest v2.1.6 [47] the best nucleotide substitution model was GTR + I with 1,000 bootstrap replications. Phylogenetic reconstructions using MrBayes sequence software and the mid–rooted tree. The pairwise sequence identities of the complete genomes were determined using the Sequence Demarcation Tool (SDT) v1.2 software [48]. Whole genome recombination analyzes were performed using the RDP4 program [49] using six statistical methods: GENECONV, MaxChi, Bootscan, 3Seq, Chimera, and SiScan.

2.4. Infectivity Assay in alternative hosts by whitefly

Six plants of each *P. vulgaris* cv. Carioca, soybean (*G. max*), peanut (*Arachis hypogaea*), cowpea (*V. unguiculata*), and tomato (*Solanum lycopersicum* cv. Santa Clara) were used in the inoculation assays. Five plants of each species were inoculated and one uninoculated (negative control). *Macroptillum erythroloma* plants (with mixed infection of BGMV and the putative new species) were used as inoculum source. Infected plants were kept isolated in a greenhouse. One week before the assay, the source plants were caged in the presence of aviruliferous adult whiteflies for virus acquisition. The inoculation assay was performed in boxes containing the viruliferous adult whiteflies, one plant of *M. erythroloma* (BGMV and novel species) served as source of inoculum. The test plants were bean, soybean, peanut, cowpea, and tomato cultivars. After 15 days of exposure to the viruliferous insects, one sample per plant was collected (third leaf counting from the apex to downwards). Inoculated plants were evaluated for symptoms at 7, 14, 21, 28, 35, and 60 days after inoculation (DAI). Detection of BGMV and new species was performed at 21 DAI by total DNA extraction and PCR using specific primers.

3. Results

3.1 Molecular Characterization of a novel recombinant begomovirus

Macroptilium erythroloma samples were collected in urban settings in Brasilia–DF area and they displayed symptoms of bright internerval yellowing and mosaic in the leaves and pods (**Figure 1**A and **1**B). *Macroptillum* plants were previously reported as alternative hosts of begomoviruses that are able to infect cultivars of important legume crops.

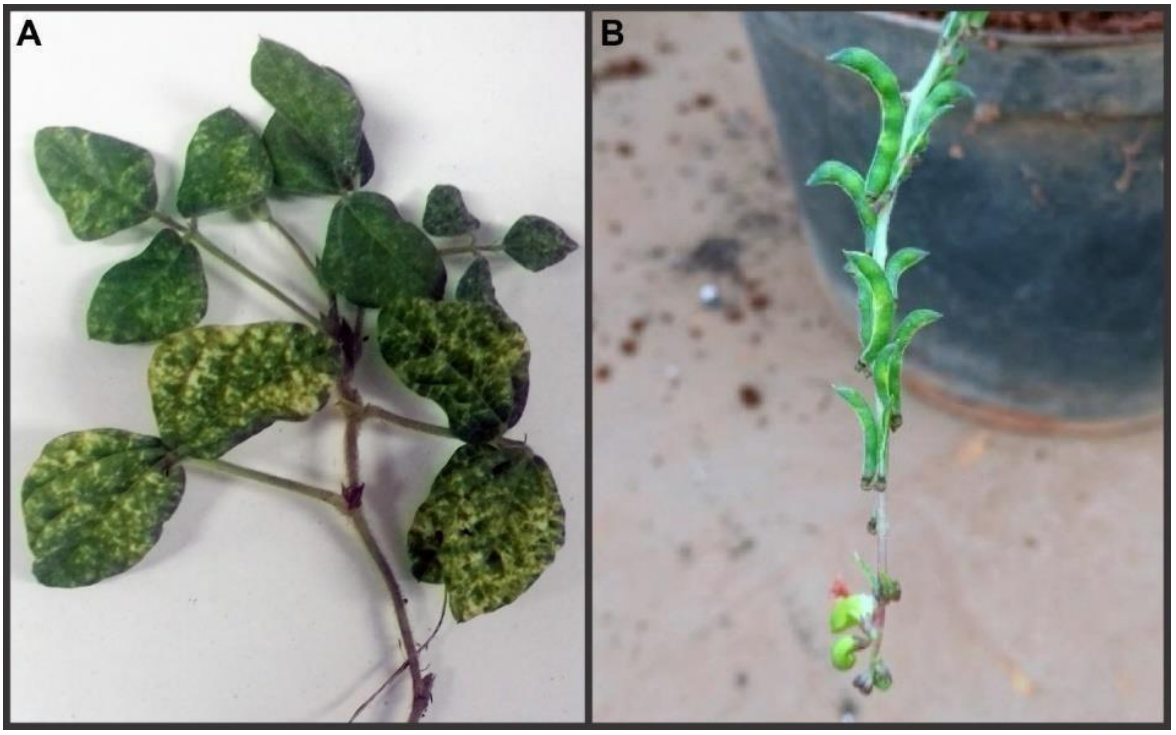


Figure 1.A-B. *Macroptilium erythroloma* plant showing begomovirus-like symptoms. A. Foliar bright internerval yellowing and mosaic caused by mixed infection with *Bean golden mosaic virus* (BGMV) and with a novel species – Macroptilium bright yellow interveinal virus (MaBYIV). B. Pod symptoms in *M. erythroloma* plant with mixed infection (BGMV and MaBYIV).

The obtained HTS-derived nucleotide sequences of *M. erythroloma* infected by BGMV and MaBYIV were analyzed on the CLC Genomics program and compared with the NCBI RefSeq viral database. Four contigs displayed identity to the following begomoviruses: Macroptilium yellow net virus – MaYNV (DNA–A and DNA–B components) and BGMV (DNA–A and DNA–B components). This observation was verified by performing PCR with MaYNV–specific primers by amplifying the complete genome using specific primers for BGMV detection. The PCR product was sent for Sanger sequencing and confirmed for both viral species presence. The closest contigs related to new species MaBYIV showed 2611 nt for DNA–A and 2578 nt for DNA–B and identities of 84.4%, 78.4% with GenBank accessions JN418998 (DNA–A component of MaYNV) and JN418999 (DNA–B component of MaYNV), respectively. The identity of 84.4% of DNA A indicates a putative new *M. erythroloma*–associated begomovirus species in according to the established criteria for genera species demarcation criteria (91% DNA - A nucleotide sequence pairwise identity). The genomic organization observed was typical of bipartite begomoviruses from New World. The DNA–A component (**Figura 2**A) displayed the five ORFs (codify proteins / amount of amino acids): AV1 (CP / 252 AAs) in the viral sense, and AC1 (Rep / 362 AAs), AC2 (TrAP / 132 AAs) AC3 (Ren / 133 AAs) and AC4 (AC4 protein / 98 AAs) in the complementary-sense. The DNA–B component (**Figura 2**B) have two ORFs (codify proteins / amount of amino acids): BV1 (NSP / 255 AAs) in the viral sense and BC1 (MP / 294 AAs) in the complementary-sense. DNA–A and DNA–B components shared the conserved nonanucleotide 5'-TAATATTAC-3' as part of a stem-loop structure at the origin of replication. The cognate components DNA–A and DNA–B showed identical iterons, GGGA, repeated and inverted twice immediately prior to TATA-box (**Figure 1C**). The domains found were Motif I: FLTYPR, Motif II: PHLHI, Motif III: DVKTYVEK, Walker A: GDSRTGTGKTMW, Walker B:

IIDDV, Motif C: VLCN and do not have Arg Finger. Comparison of SDT paired identity by selecting representative and species-related begomovirus species indicated that it was a new species, the isolate with the highest paired identity was Tomato interveinal chlorosis virus isolate PE [BR: BSF2729: 04] (JF803253) 84.5% and Macroptilium yellow net virus isolate BR: Mur1: 09 segment DNA–A (JN418998) 84.4% (**Figure 2B**).

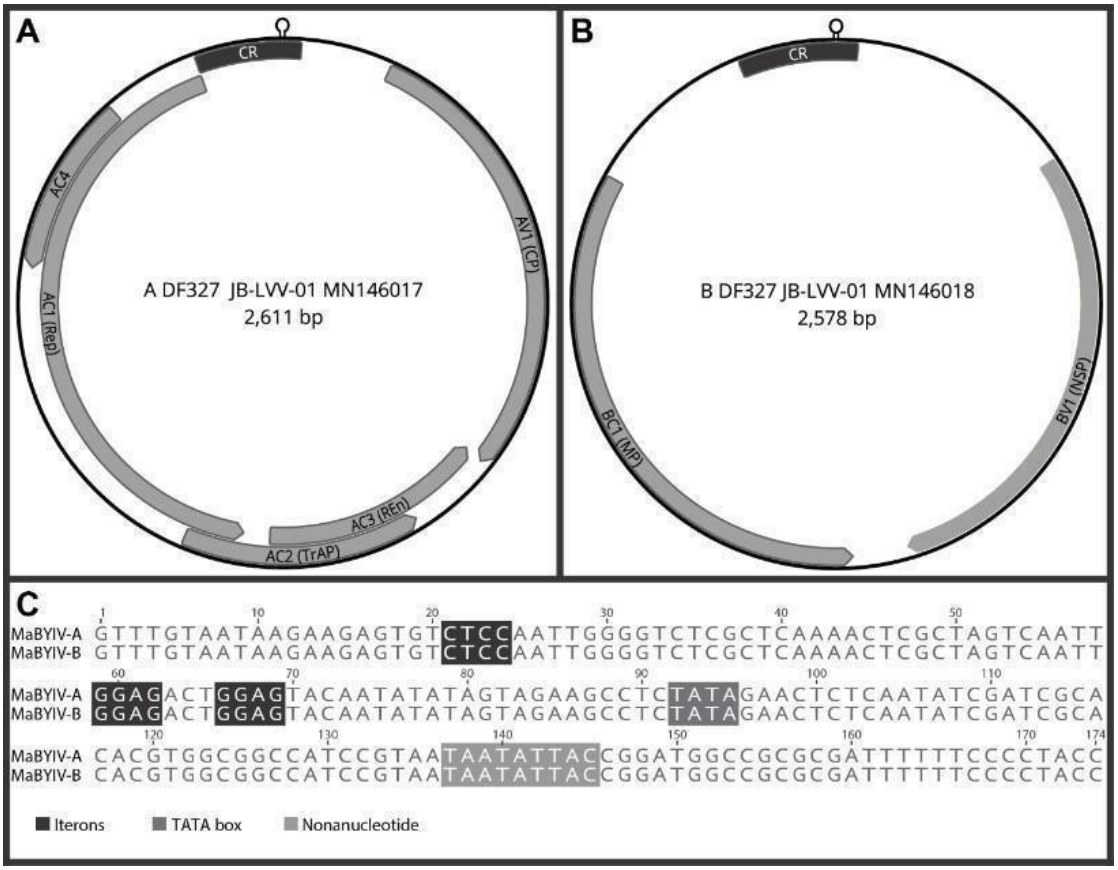


Figure 2. Genomic organization of Macroptilium bright yellow interveinal virus – MaBYIV and common region of the cognate components (DNA–A and DNA–B). **(A)** Five genes were annotated in the DNA–A component: AV1 (coat protein - CP), AC1 (replication protein – Rep), AC2 (transcriptional activating protein – TrAP), AC3 (replication promoting protein – REn), AC4, and CR (common region); **(B)** Two genes were annotated in the DNA–B component: BV1 (nuclear carrier protein - NSP), BC1 (movement protein – MP and CR (common region). Both components displayed stem-loop. **(C)** The common region (CR) and the DNA–A and DNA–B components of MaBYIV with iterons, TATA box and nonanucleotide between stem-loop.

The BGMV related contigs had sizes of 2617 nts for DNA–A (MN822293) and 2592 nts for DNA–B (MN822294) with identities of 99.9% and 98.2% with GenBank accessions: KJ939798 (BGMV DNA–A) and MG334553 (BGMV DNA–B), respectively. The DNA–A from BGMV KJ939798 isolate is related samples of bean and *Macroptilium lathyroides* plants in the northeastern and central regions of Brazil, including the Federal District [27]. There are no published data for DNA–B of a BGMV isolate, however in its denomination refers to the state of Minas Gerais (Brazil). The DNA–A component displayed the five ORFs: AV1 (CP / 252 AAs) in the viral sense, and AC1 (Rep / 362 AAs), AC2 (TrAP / 130 AAs), AC3 (Ren / 133 AAs) and AC4(AC4)

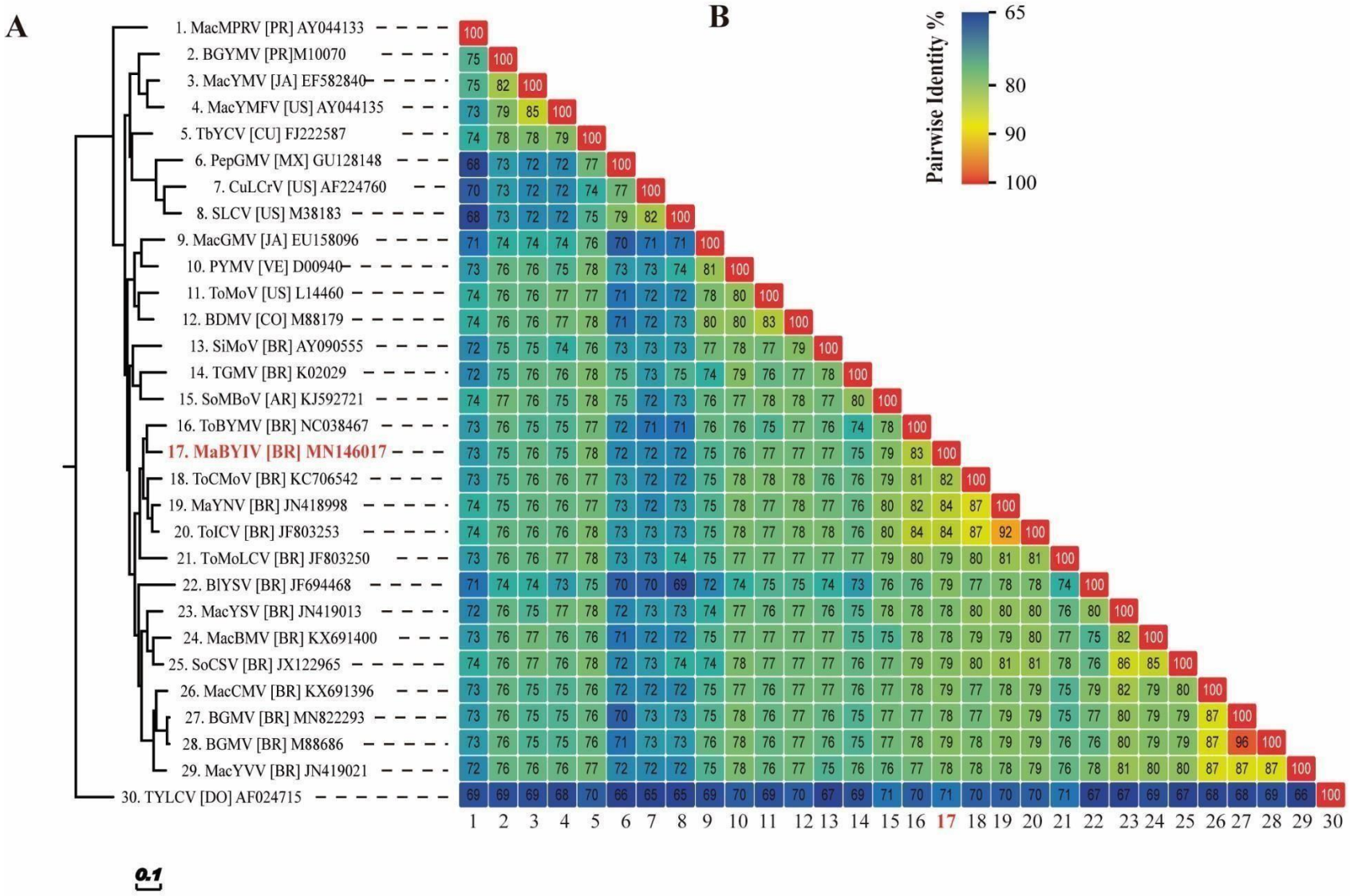
protein / 86 AAs) in the complementary sense. The DNA–B component displayed two ORFs: BV1 (NSP / 254 AAs), in the viral sense, and BC1 (MP / 294 AAs) in the complementary-sense.

3.2 Phylogenetic analysis and recombination

Phylogenetic analysis of the complete DNA–A genome of representative begomoviruses indicates for that the new species (MaBYIV) is clustered in the same clade with tomato bright yellow mosaic virus isolate – ToBYMV (NC038467) and distantly related to BGMV isolates (**Figure 3**A). For DNA–B component, the sequence of the new species (MaBYIV) was grouped in a separate branch (**Figure 3**C).

viruses

MDPI



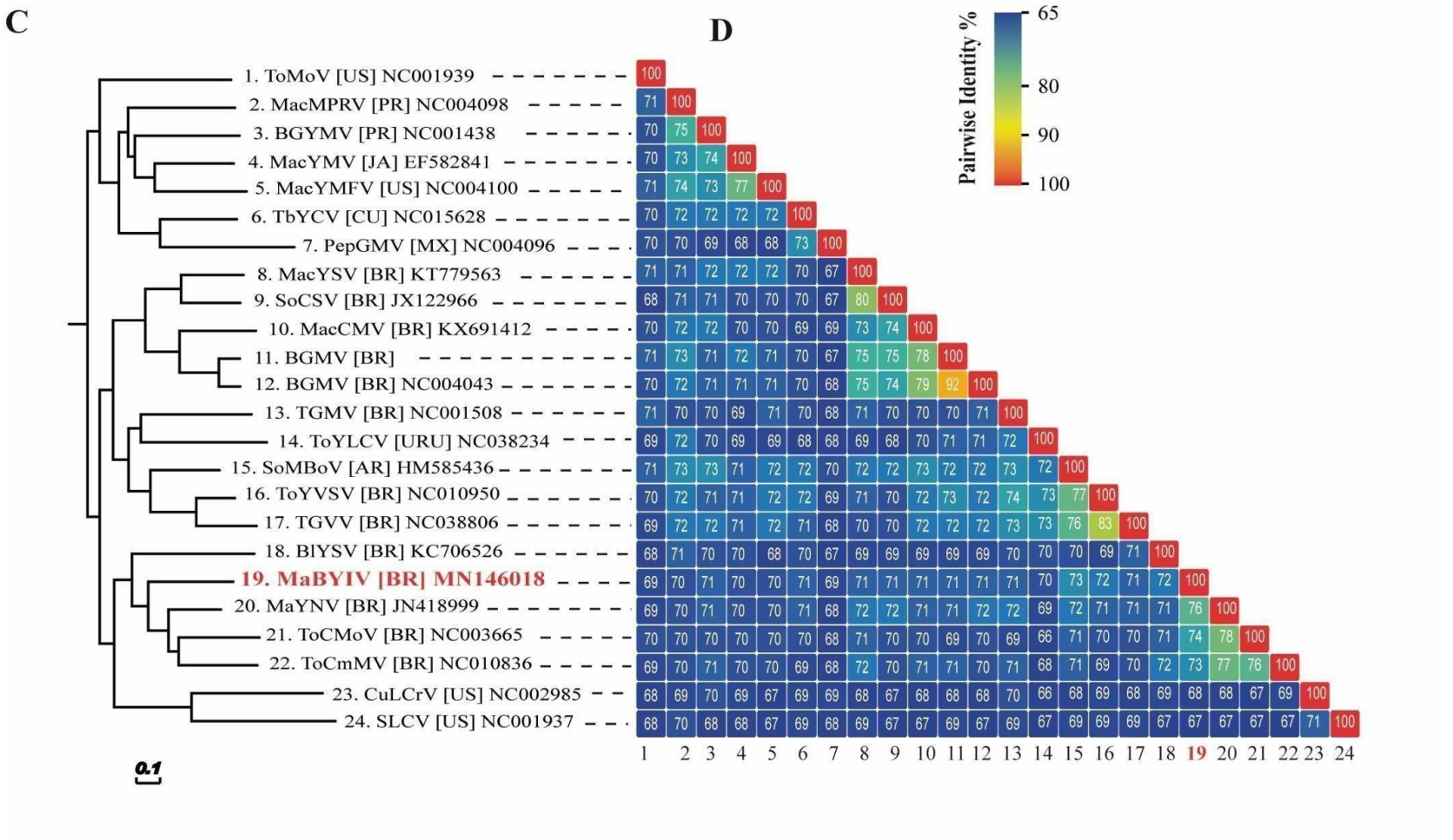

C
D
Pairwise Identity %
65
80
90
100
1. ToMoV [US] NC001939
2. MacMPRV [PR] NC004098
3. BGYMV [PR] NC001438
4. MacYMV [JA] EF582841
5. MacYMFV [US] NC004100
6. TbYCV [CU] NC015628
7. PepGMV [MX] NC004096
8. MacYSV [BR] KT779563
9. SoCSV [BR] JX122966
10. MacCMV [BR] KX691412
11. BGMV [BR]
12. BGMV [BR] NC004043
13. TGMV [BR] NC001508
14. ToYLCV [URU] NC038234
15. SoMBoV [AR] HM585436
16. ToYVSV [BR] NC010950
17. TGVV [BR] NC038806
18. BlYSV [BR] KC706526
19. MaBYIV [BR] MN146018
20. MaYNV [BR] JN418999
21. ToCMoV [BR] NC003665
22. ToCmMV [BR] NC010836
23. CuLCrV [US] NC002985
24. SLCV [US] NC001937
0.1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Figure 3. Phylogenetic tree representation for DNA–A and DNA–B components of Macroptilium bright yellow interveinal virus – MaBYIV and *Bean golden mosaic virus* – BGMV (**A and C**). Pairwise identity analysis in SDT was carried out only using the information of the DNA–A component of MaBYIV and BGMV. Bayesian phylogenetic tree sequence and the mid-rooted tree. The best nucleotide replacement model was GTR + I with 1,000 bootstrap replications. (**B and D**). Comparison of the pairwise identity of DNA–A components of begomovirus. The order, isolates names and accession number (GenBank) are**: DNA-A**: **1**. MacMPRV [PR] AY044133 *Macroptilium mosaic Puerto Rico virus*; **2**. BGYMV [PR] M10070 *Bean golden yellow mosaic virus*; **3**. MacYMV [JA] EF582840 *Macroptilium yellow mosaic virus*; **4**. MacYMFV [US] AY044135 *Macroptilium yellow mosaic* Florida *virus*; **5**. TbYCV [CU] FJ222587 *Tobacco yellow crinkle virus*; **6**. PepGMV [MX] GU128148 *Pepper golden mosaic virus*; **7**. CuLCrV [US] AF224760 *Cucurbit leaf curl virus*; **8**. SLCV [US] M38183 *Squash leaf curl virus*; **9**. MacGMV [JA] EU158096 *Macroptilium golden mosaic virus*; **10**. PYMV [VE] D00940 *Potato yellow mosaic virus*;**11**. ToMoV [US] L14460 *Tomato mottle virus* **12**. BDMV [CO] M88179 *Bean dwarf mosaic virus*; **13**. SiMoV [BR] AY090555 *Sida mottle virus*; **14**. TGMV [BR] K02029 *Tomato golden mosaic virus*; **15**. SoMBoV [AR] KJ592721 *Solanum mosaic Bolivia virus*; **16**. ToBYMV [BR] NC038467 *Tomato bright yellow mosaic virus*; **17**. MaBYIV – MN146017 Macroptilium bright yellow interveinal virus; **18**. ToCMoV [BR] KC706542 *Tomato chlorotic mottle virus*; **19**. MaYNV [BR] JN418998 Macroptilium yellow net virus; **20**. ToICV [BR] JF803253 *Tomato interveinal chlorosis virus*; **21**. ToMoLCV [BR] JF803250 *Tomato mottle leaf curl virus*; **22**. BlYSV [BR] JF694468 *Blainvillea yellow spot virus*; **23**. MacYSV [BR] JN419013 *Macroptilium yellow spot virus*; **24**. MacBMV [BR] KX691400 *Macroptilium bright mosaic virus*; **25**. SoCSV [BR] JX122965 *Soybean chlorotic spot virus*; **26**. MacCMV [BR] KX691396 *Macroptilium common mosaic virus*; **27**. BGMV [BR] MN822294 Bean golden mosaic virus; **28**. BGMV [BR] M88686 *Bean golden mosaic virus*; **29**. MacYVV [BR] JN419021 *Macroptilium yellow vein virus*; **30**. TYLCV [DO] AF024715 *Tomato yellow leaf curl virus*. **DNA - B**: **1**. ToMoV [US] NC001939 *Tomato mottlevirus*; **2**. MacMPRV [PR] NC004098 *Macroptilium mosaic Puerto Rico virus*; **3**. BGYMV [PR] NC001438 *Bean golden yellow mosaic virus*; **4**. MacYMV [JA] EF582841 *Macroptilium yellow mosaic virus*; **5**. MacYMFV [US] NC004100 *Macroptilium yellow mosaic Florida virus*; **6**. TbYCV [CU] NC015628 *Tobacco yellow crinkle virus*; **7**. PepGMV [MX] NC004096 *Pepper golden mosaic virus*; **8**. MacYSV [BR] KT779563 *Macroptilium yellow spot virus*; **9**. SoCSV [BR] JX122966 *Soybean chlorotic spot virus*; **10**. MacCMV [BR] KX691412 *Macroptilium common mosaic virus*; **11**. BGMV [BR] MN822294 *Bean golden mosaic virus*; **12**. BGMV [BR] NC004043 *Bean golden mosaic virus*; **13**. TGMV [BR] NC001508 *Tomato golden mosaic virus*; **14**. ToYLCV [URU] NC038234 *Tomato rugose yellow leaf curl virus*; **15**. SoMBoV [AR] HM585436 *Solanum mosaic Bolivia virus*; **16**. ToYVSV [BR] NC010950 *Tomato yellow vein streak virus*; **17**. TGVV [BR] NC038806 *Tomato golden vein virus*; **18**. BlYSV [BR] KC706526 *Blainvillea yellow spot virus*; **19**. MaBYIV - MN146018 Macroptilium bright yellow interveinal virus; **20**. MaYNV [BR] JN418999 Macroptilium yellow net virus; **21**. ToCMoV [BR] NC003665 *Tomato chlorotic mottle virus*; **22**. ToCmMV [BR] NC010836 *Tomato common mosaic virus*; **23**. CuLCrV [US] NC002985 *Cucurbit leaf crumple virus*; **24**. SLCV[US] NC001937 *Squash leaf curl virus*.

All six statistical methods indicated evidence of recombination: RDP (3,463 x 10^{-9}), GENECONV (1,050 x 10^{-2}), MaxChi (1,695 x 10^{-10}), Bootscan (4,094 x 10^{-7}), 3Seq (5,556 x 10^{-5}), Chimera (2,529 x 10^{-11}), and SiScan (7,379 x 10^{-14}). The recombination breakpoint is located in the region starting at nucleotide 1918 and ending at 2394 (i.e. 476 nts). The major parental genotype is BGMV and the minor parental genotype is Tomato mottle leaf curl virus – ToMoLCV. Weeds may serve as year-round reservoirs of numerous begomoviruses, favoring the emergence of new species through pseudo-recombination as well as recombination events. One well-illustrated case is the Sida micrantha mosaic-associated viruses (SimMV) [50]. A survey of begomoviruses infecting weeds (within the Fabaceae family) in Brazil found six viral species, four of them new. In the MaYSV populations, recombination events and high degree of variability were found [31]. A new species of begomovirus infecting *M. lathyroides* was described and named as Macroptilium common mosaic virus (MacCMV). This novel virus clustered with BGMV in phylogenetic analyses but no recombination events were identified [51].

3.3 Infectivity assays with BGMV and with a novel recombinant begomovirus

After 15days, observations were made regarding vector preference, symptom evaluations and collection of the third apex leaf to perform the detection. Observations regarding the vector were that cowpea, peanut and soybean plants were less visited than bean and tomato plants, because they presented low amounts of whitefly life stages, while tomato and bean plants presented large amounts. The fact that the whitefly prefers bean and tomato plants corroborates a study that demonstrates this preference [2]. Symptoms were observed from seven days after inoculation, such as (plant / symptom / number of plants): carioca bean / blister / 3 and mosaic / 3; tomato / mottled / 6 and soybean / mottled / 5 and chlorosis / 2, peanuts / ribbed chlorosis / 4 and asymptomatic / 2, mottled seridó / 1 and asymptomatic / 5. The uninoculated plants had chlorotic spots related to the presence of eggs and whitefly nymphs, since they were placed in the presence of aviruliferous whiteflies. By performing PCR detection using MaBYIV-specific primers, bean and cowpea plants were positive, two samples from each co-infected with BGMV. For BGMV, all inoculated plants were positive except two cowpea plants. The amplicons were purified and sequenced confirming to be BGMV and MaBYIV. Assay results are listed in **Table 1**.

Fabaceae plants such as *Calopogonium mucunoides*, *Canavalia* sp., *Centrosema brasilianum*, *Macroptilium atropurpureum* and *M. lathyroides*, as well as *Desmodium* sp. are known as invasive plants and alternative hosts of begomovirus in the field [7]. Plants of *M. erythroloma* was assigned in the early 1980s (i.e. before the invasion of B. tabaci MEAM 1) as a putative host of BGMV based only upon symptoms and biological tests [52]. Plants of *M. lathyroides* were identified as a common hosts of a wide array begomoviruses and they could act as a mixing milieu from which recombinant viruses could emerge [31, 53, 54]. species. Wild and/or invasive legume species are known to be good hosts for the insect vectors and begomoviruses [31]. Weeds have high adaptability and distribution and might serve as alternative hosts for begomovirus survival and spread [55, 56] in the absence of major crops. Epidemiological studies focused on the occurrence of genetic recombination among begomoviruses in distinct hosts or vectors [57] as well as the survival of begomoviruses in plant debris and during unfavorable conditions should be of interest.

Table 1. Results of the transmission assay via viruliferous adult whiteflies [*Bemisia tabaci Middle East Asian Minor* 1]. Plant species employed in this assay are listed by common and scientific name. Plants were scored for the observed symptoms. Viral detection assays for Bean golden mosaic virus – BGMV and Macroptilium bright yellow interveinal virus – MaBYIV were also carried out with the inoculated plants.

Common and scientific name	**Observed symptoms**	**BGMV detection (NPP/NTP*)**	**MaBYIV detection (NPP/NTP *)**
Peanut (*Arachis hypogea*)	Mosaic	**Positive** (5/5)	Negative
Bean (*Phaseolus vulgaris* cv. Carioca)	Blistering and mosaic	**Positive** (5/5)	**Positive** (2/5)
Soybean (*Glycine max*)	Blistering	**Positive** (5/5)	**Positive** (2/5)
Cowpea (*Vigna unguiculata*)	Asymptomatic	**Positive** (2/5)	Negative
Tomato (*Solanum lycopersicum* cv. Santa Clara)	Mosaic	**Positive** (5/5)	Negative

* **NPP/NTP** = number of positive plants / total number of plants.

4. Conclusion

Our results suggest the presence a new bipartite begomovirus species associated to the weed legume *M. erythroloma*, reinforcing the high levels of genetic diversity of the genus *Begomovirus* in New World area. The present information can be useful in the establishment of effective preemptive management tactics for the newly identified susceptible crops (beans and soybeans).

Acknowlegments

The authors are grateful to the colleagues **Maurício Rossato a**nd **Bruno Eduardo Cardozo de Miranda** from the Departament of Plant Pathology, Universidade de Brasília (UnB) and **Maria Esther de Noronha Fonseca (CNPH)** for their help in some set of analyses as well as in reviewing the manuscript.

References

1. Stanley, J.; Bisaro, D.; Briddon, R.; Brown, J.; Fauquet, C.; Harrison, B.; Rybicki, E.; Stenger, D., Family geminiviridae. *Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses* **2005**, 301-326.
2. Zerbini, F. M.; Briddon, R. W.; Idris, A.; Martin, D. P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Roumagnac, P.; Varsani, A.; Consortium, I. R., ICTV virus taxonomy profile: Geminiviridae. *The Journal of general virology* **2017,** 98, (2), 131.

3. Rojas, M. R.; Macedo, M. A.; Maliano, M. R.; Soto-Aguilar, M.; Souza, J. O.; Briddon, R. W.; Kenyon, L.; Rivera Bustamante, R. F.; Zerbini, F. M.; Adkins, S., World management of geminiviruses. *Annual review of phytopathology* **2018,** 56, 637-677.
4. Adams, M.; King, A.; Carstens, E., Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Archives of virology* **2013,** 158, (9), 2023-2030.
5. Brown, J. K.; Zerbini, F. M.; Navas-Castillo, J.; Moriones, E.; Ramos-Sobrinho, R.; Silva, J. C.; Fiallo-Olivé, E.; Briddon, R. W.; Hernández-Zepeda, C.; Idris, A., Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Archives of virology* **2015,** 160, (6), 1593-1619.
6. Muhire, B.; Martin, D. P.; Brown, J. K.; Navas-Castillo, J.; Moriones, E.; Zerbini, F. M.; Rivera-Bustamante, R.; Malathi, V.; Briddon, R. W.; Varsani, A., A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of virology* **2013,** 158, (6), 1411-1424.
7. Varsani, A.; Martin, D. P.; Navas-Castillo, J.; Moriones, E.; Hernández-Zepeda, C.; Idris, A.; Zerbini, F. M.; Brown, J. K., Revisiting the classification of curtoviruses based on genome-wide pairwise identity. *Archives of virology* **2014,** 159, (7), 1873-1882.
8. Varsani, A.; Roumagnac, P.; Fuchs, M.; Navas-Castillo, J.; Moriones, E.; Idris, A.;Briddon, R. W.; Rivera-Bustamante, R.; Murilo Zerbini, F.; Martin, D. P., Capulavirus and Grablovirus: two new genera in the family Geminiviridae. *Archives of Virology* **2017,** 162, (6), 1819-1831.
9. Zerbini, F. M.; Briddon, R. W.; Idris, A.; Martin, D. P.; Moriones, E.; Navas-Castillo, J.; Rivera-Bustamante, R.; Roumagnac, P.; Varsani, A.; Ictv Report, C., ICTV VirusTaxonomy Profile: Geminiviridae. *The Journal of general virology* **2017,** 98, (2), 131-133.
10. Varsani, A.; Navas-Castillo, J.; Moriones, E.; Hernández-Zepeda, C.; Idris, A.; Brown, J. K.; Zerbini, F. M.; Martin, D. P., Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. *Archives of virology* **2014,** 159, (8), 2193-2203.
11. Hogenhout, S. A.; Ammar, E.-D.; Whitfield, A. E.; Redinbaugh, M. G., Insect vector interactions with persistently transmitted viruses. *Annual. Review. Phytopathology.* **2008,** 46, 327-359.
12. de Faria, J. C.; Zerbini, F., Família Geminiviridae-taxonomia, replicação e movimento. *Embrapa Arroz e Feijão-Artigo em periódico indexado (ALICE)* **2000**.
13. Brown, J. K.; Frohlich, D.; Rosell, R. C., The sweetpotato or silverleaf whiteflies: biotypes of Bemisia tabaci or a species complex? *Annual review of entomology* **1995,** 40, (1), 511-534.
14. Bôas, G. L. V.; França, F. H.; Macedo, N.; Moita, A. W., Avaliação da preferência de Bemisia argentifolii por diferentes espécies de plantas1. *Horticultura. brasileira* **2001,** 19, (2).
15. Lourenção, A.; Nagai, H., Surtos populacionais de Bemisia tabaci no Estado de SãoPaulo. *Bragantia* **1994,** 53, (1), 53-59.
16. Navas-Castillo, J.; Fiallo-Olivé, E.; Sánchez-Campos, S., Emerging virus diseases transmitted by whiteflies. *Annual review of phytopathology* **2011,** 49, 219-248.
17. Fondong, V. N., Geminivirus protein structure and function. *Molecular plant pathology***2013,** 14, (6), 635-649.

18. Hanley-Bowdoin, L.; Bejarano, E. R.; Robertson, D.; Mansoor, S., Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology* **2013,** 11, (11), 777.
19. Li, F.; Xu, X.; Huang, C.; Gu, Z.; Cao, L.; Hu, T.; Ding, M.; Li, Z.; Zhou, X., The AC5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. *New Phytologist* **2015,** 208, (2), 555-569.
20. Zhang, S. C.; Wege, C.; Jeske, H., Movement proteins (BC1 and BV1) of Abutilon mosaic geminivirus are cotransported in and between cells of sink but not of source leaves as detected by green fluorescent protein tagging. *Virology* **2001,** 290, (2), 249-260.
21. Argüello-Astorga, G.; Ruiz-Medrano, R., An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. *Archives of virology* **2001,** 146, (8), 1465-1485.
22. Sanz-Burgos, A. P.; Gutiérrez, C., Organization of thecis-Acting Element Required for Wheat Dwarf Geminivirus DNA Replication and Visualization of a Rep Protein–DNA Complex. *Virology* **1998,** 243, (1), 119-129.
23. Argüello-Astorga, G.; Guevara-Gonzalez, R.; Herrera-Estrella, L.; Rivera-Bustamante, R., Geminivirus replication origins have a group-specific organization of iterative elements: a model for replication. *Virology* **1994,** 203, (1), 90-100.
24. Orozco, B.; Hanley-Bowdoin, L., A DNA structure is required for geminivirus replication origin function. *Journal of virology* **1996,** 70, (1), 148-158.
25. Naito, F. Y.; Melo, F. L.; Fonseca, M. E. N.; Santos, C. A.; Chanes, C. R.; Ribeiro, B. M.; Gilbertson, R. L.; Boiteux, L. S.; de Cássia Pereira-Carvalho, R., Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. *Archives of virology* **2019,** 164, (7), 1907-1910.
26. Morales, F. J., History and current distribution of begomoviruses in Latin America. *Advances in virus research* **2006,** 67, 127-162.
27. Sobrinho, R. R.; Xavier, C. A. D.; de Barros Pereira, H. M.; de Andrade Lima, G. S.; Assunção, I. P.; Mizubuti, E. S. G.; Duffy, S.; Zerbini, F. M., Contrasting genetic structure between two begomoviruses infecting the same leguminous hosts. *Journal of General Virology* **2014,** 95, (11), 2540-2552.
28. Galvao, R. M.; Mariano, A. C.; Luz, D. F.; Alfenas, P. F.; Andrade, E. C.; Zerbini, F. M.; Almeida, M. R.; Fontes, E. P., A naturally occurring recombinant DNA-A of a typical bipartite begomovirus does not require the cognate DNA-B to infect Nicotiana benthamiana systemically. *Journal of General Virology* **2003,** 84, (3), 715-726.
29. Inoue-Nagata, A. K.; Martin, D. P.; Boiteux, L. S.; Giordano, L. d. B.; Bezerra, I. C.; Ávila, A. C. d., New species emergence via recombination among isolates of the Brazilian tomato infecting Begomovirus complex. *Pesquisa Agropecuária Brasileira* **2006,** 41, (8), 1329-1332.
30. Ribeiro, S. G.; Martin, D. P.; Lacorte, C.; Simões, I. C.; Orlandini, D. R.; Inoue-Nagata, A. K., Molecular and biological characterization of Tomato chlorotic mottle virus suggests that recombination underlies the evolution and diversity of Brazilian tomato begomoviruses. *Phytopathology* **2007,** 97, (6), 702-711.
31. Silva, S.; Castillo-Urquiza, G.; Hora-Júnior, B.; Assunção, I.; Lima, G.; Pio-Ribeiro, G.; Mizubuti, E.; Zerbini, F., Species diversity, phylogeny and genetic variability of

begomovirus populations infecting leguminous weeds in northeastern Brazil. *Plant Pathology* **2012,** 61, (3), 457-467.

32. Pita, J.; Fondong, V.; Sangare, A.; Otim-Nape, G.; Ogwal, S.; Fauquet, C., Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **2001,** 82, (3), 655-665.
33. Monci, F.; Sánchez-Campos, S.; Navas-Castillo, J.; Moriones, E., A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Virology* **2002,** 303, (2), 317-326.
34. García-Andrés, S.; Monci, F.; Navas-Castillo, J.; Moriones, E., Begomovirus genetic diversity in the native plant reservoir Solanum nigrum: evidence for the presence of a new virus species of recombinant nature. *Virology* **2006,** 350, (2), 433-442.
35. García-Andrés, S.; Accotto, G. P.; Navas-Castillo, J.; Moriones, E., Founder effect, plant host, and recombination shape the emergent population of begomoviruses that cause the tomato yellow leaf curl disease in the Mediterranean basin. *Virology* **2007,** 359, (2), 302-312.
36. García-Andrés, S.; Tomás, D. M.; Sánchez-Campos, S.; Navas-Castillo, J.; Moriones, E., Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. *Virology* **2007,** 365, (1), 210-219.
37. Delgado-Salinas, A.; Lewis, G. P., A new species of *Macroptilium* (Benth.) Urb.(Leguminosae: Papilionoideae: Phaseolinae) from North-Eastern Brazil. *Kew Bulletin* **2008,** 63, (1), 151-154.
38. Fevereiro, V. B., Macroptilium (Bentham) Urban do Brasil (Leguminosae-Faboideae-Phaseoleae-Phaseolinae). *Arquivos do Jardim Botânico do Rio de Janeiro* **1987,** 28,109-180.
39. Collins, A. M.; Brown, J. K.; Rehman, M. M.; Roye, M. E., Complete nucleotide sequence of an isolate of Euphorbia mosaic virus that infects Euphorbia heterophylla and Wissadula amplissima in Jamaica. *Archives of virology* **2009,** 154, (11), 1859.
40. Amarakoon, I.; Roye, M.; Briddon, R.; Bedford, I.; Stanley, J., Molecular and biological characterization of Macroptilium yellow mosaic virus from Jamaica. *Plant pathology* **2008,** 57, (3), 417-426.
41. Idris, A.; Hiebert, E.; Bird, J.; Brown, J. K., Two newly described begomoviruses of *Macroptilium lathyroides* and common bean. *Phytopathology* **2003,** 93, (7), 774-783.
42. Fazekas, A. J.; Kuzmina, M. L.; Newmaster, S. G.; Hollingsworth, P. M., DNAbarcoding methods for land plants. In *DNA barcodes*, Springer: 2012; pp 223-252.
43. Boiteux, L.; Fonseca, M.; Simon, P., 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* **1999,** 124, (1), 32-38.
44. Inoue-Nagata, A. K.; Albuquerque, L. C.; Rocha, W. B.; Nagata, T., A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. *Journal Virology Methods* **2004,** 116, (2), 209-11.
45. Bonfim, K.; Faria, J. C.; Nogueira, E. O.; Mendes, É. A.; Aragão, F. J., RNAi-mediated resistance to *Bean golden mosaic virus* in genetically engineered common bean (*Phaseolus vulgaris*). *Molecular Plant-microbe interactions* **2007,** 20, (6), 717-726.

46. Edgar, R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* **2004,** 32, (5), 1792-1797.
47. Darriba, D.; Taboada, G. L.; Doallo, R.; Posada, D., jModelTest 2: more models, new heuristics and parallel computing. *Nature methods* **2012,** 9, (8), 772.
48. Muhire, B. M.; Varsani, A.; Martin, D. P., SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PloS one* **2014,** 9, (9), e108277.
49. Martin, D. P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B., RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus evolution* **2015,** 1,(1).
50. Jovel, J.; Reski, G.; Rothenstein, D.; Ringel, M.; Frischmuth, T.; Jeske, H., Sida micrantha mosaic is associated with a complex infection of begomoviruses different from Abutilon mosaic virus. *Archives of Virology* **2004,** 149, (4), 829-841.
51. Passos, L. S.; Rodrigues, J. S.; Soares, É. C. S.; Silva, J. P.; Zerbini, F. M.; Araújo, A. S. F.; Beserra, J. E. A., Complete genome sequence of a new bipartite begomovirus infecting *Macroptilium lathyroides* in Brazil. *Archives of virology* **2017,** 162, (11), 3551-3554.
52. Chagas, C.; Vicente, M.; Barradas, M., *Macroptilium erythroloma* (Mart. ex Benth.) Urb.-Leguminosae-a possible reservoir of bean golden mosaic virus (BGMV). *Arquivos do Instituto Biológico, São Paulo* **1981,** 48, (1/4), 113-116.
53. Idris, A.; Bird, J.; Brown, J., First report of a bean-infecting begomovirus from *Macroptilium lathyroides* in Puerto Rico that is distinct from bean golden mosaic virus. *Plant disease* **1999,** 83, (11), 1071-1071.
54. Ramos, P.; Fernández, A.; Castrillo, G.; Diaz, L.; Echemendía, A.; Fuentes, A.; Peral, R.; Pujol, M.; Ascencio-Ibanez, J.; Rivera-Bustamante, R., Macroptilium yellow mosaic virus, a new begomovirus infecting *Macroptilium lathyroides* in Cuba. *Plant disease* **2002,** 86, (9), 1049-1049.
55. Gilbertson, R. L.; Rojas, M. R.; Russell, D. R.; Maxwell, D. P., Use of the asymmetric polymerase chain reaction and DNA sequencing to determine genetic variability of bean golden mosaic geminivirus in the Dominican Republic. *Journal of General Virology* **1991,** 72, (11), 2843-2848.
56. Bedford, I.; Kelly, A.; Banks, G.; Briddon, R.; Cenis, J.; Markham, P., Solanum nigrum: an indigenous weed reservoir for a tomato yellow leaf curl geminivirus in southern Spain. *European Journal of Plant Pathology* **1998,** 104, (2), 221-222.
57. Tan, P. H.; Wong, S. M.; Wu, M.; Bedford, I. D.; Saunders, K.; Stanley, J., Genome organization of ageratum yellow vein virus, a monopartite whitefly-transmitted geminivirus isolated from a common weed. *Journal of General Virology* **1995,** 76, (12), 2915-2922.

Attached III

First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil (2018)

Batista; JG., Melo; FL., Pereira-Carvalho RC., et al - Plant Disease/Disease Note

Published

APS Publications

Previous

DISEASE NOTES

OPEN

First Report of Tomato Apical Leaf Curl Virus Infecting Tomato in Brazil

J. G. Batista, F. L. Melo, R. C. Pereira-Carvalho, D. M. T. Alves-Freitas, and S. G. Ribeiro

Affiliations

Authors and Affiliations
J. G. Batista[1]
F. L. Melo[1]
R. C. Pereira-Carvalho[1]
D. M. T. Alves-Freitas[2]
S. G. Ribeiro[2] †
[1]Departamento de Fitopatologia, Universidade de Brasília, Brasília, DF, Brazil; and
[2]Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brazil

Published Online: 7 May 2019 https://doi.org/10.1094/PDIS-09-18-1636-PDN

Tomato (*Solanum lycopersicum*) is one of the most important vegetable crops in Brazil. The occurrence of begomovirus (genus *Begomovirus*, family *Geminiviridae*) infections is a predominant biotic constraint affecting the tomato production in the country and can cause up to 100% yield losses (Inoue-Nagata et al. 2016). Besides the plethora of tomato-infecting begomoviruses present in South America, new monopartite divergent geminiviruses were recently identified, tomato-associated geminivirus 1 (TaGV1) in Brazil (Fontenele et al. 2017) and tomato apical leaf curl virus (ToALCV) in Argentina (Vaghi Medina et al. 2018). In 2016, 15 tomato samples showing symptoms of mosaic, leaf curling, and necrosis were collected from an experimental field in Brasília, DF, Brazil. Total DNA was extracted from each sample using the cetyltrimethylammonium bromide method and was used in rolling circle amplification (RCA) reactions with phi-29 DNA polymerase (Inoue-Nagata et al. 2004). RCA-amplified samples were pooled and sequenced in an Illumina Hi-Seq2500 platform at Macrogen (South Korea) using a Nextera DNA Library Prep kit. The sequencing generated 21,128,652 (2 × 100 paired-end) reads, which were processed and assembled with CLC Genomic Workbench

version 9.0 (Qiagen Bioinformatics). The resulting contigs (44,967 contigs) were compared with a local viral RefSeq database using BlastN and BlastX with Geneious R10.2. One circular contig (2,875 nucleotides) was identified exhibiting 96% identity to the ToALCV genome. Based in the contig sequence, primers Cap1PstIF (5′-CTGCAGAYTTGCGCGGATCGATTAAT-3′) and Cap1PstIR (5′-CTGCAGAAATGCGTTGTAACTTCTCGGATAT-3′) overlapping in a *Pst*I site were designed and used in polymerase chain reactions (PCRs) to detect individual infected plants and recover the complete viral genome. A fragment of ~2.9 kb was amplified only from one of the 15 initially collected samples (FAL-18). This fragment was cloned and was sequenced using Sanger methodology. The sequence from clone ToALCV:BR:Brasilia:Tom18 (MH539677) displayed 99% identity with the contig identified in the Illumina sequencing data and 96% identity with the Argentinian isolates of ToALCV (MG491195 to MG491197). To confirm the PCR diagnosis, Southern hybridization analysis was conducted with the total DNA from all plant samples using ^{32}P-labeled ToALCV:BR:Brasilia:Tom18 complete genome as a probe, and the typical geminivirus ssDNA and dsDNA replicative intermediate forms were observed only in the sample FAL-18. The complete genome of ToALCV:BR:Brasilia:Tom18 is 2,875 nt long with six open reading frames (ORFs). The nucleotide sequences of the ORFs and the deduced amino acid sequences of the proteins of ToALCV:BR:Brasilia:Tom18 share high identities with those of the Argentinian isolates. For V1, V3, C1, C1:C2, and C3 the identities are 96 to 97% and 94 to 98% for nucleotide and amino acid sequences, respectively. V2 shares 97% nucleotide and 92% amino acid identities with ToALCV-AR isolates. This is the first report of the occurrence of ToALCV in Brazil. Additional studies are needed to better understand the biology of ToALCV, its host range, and vector transmission and to evaluate the risk of ToALCV for the tomato crop in Brazil.

Funding: Financial support was provided by grants from Empresa Brasileira de Pesquisa Agropecuária, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Apoio à Pesquisa do Distrito Federal. J. G. Batista was supported by a scholarship from CAPES and D. M. T. Alves-Freitas by a fellowship from CNPq.

APS Publications

The American Phytopathological Society

(APS)

3340 Pilot Knob Road, St. Paul, MN 55121 USA

+1.651.454.7250

FAX +1.651.454.0766

Attached IV

Characterization of genetically divergent Tomato associated geminivirus 1 isolates from table beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*) (2020c)

Batista; JG., Fochat; F., Nery; FMB, et al - Tropical Plant Pathology/Full paper

Submitted

Tropical Plant Pathology

Characterization of genetically divergent Tomato–associated geminivirus 1 isolates from table beet (Beta vulgaris) and tomato (Solanum lycopersicum)

--Manuscript Draft--

Manuscript Number:	
Full Title:	Characterization of genetically divergent Tomato–associated geminivirus 1 isolates from table beet (Beta vulgaris) and tomato (Solanum lycopersicum)
Article Type:	Original Article
Funding Information:	
Abstract:	Metagenomic approaches in conjunction with high–throughput sequencing (HTS) have been efficiently employed for the discovery of novel plant–associated viral species, including members of the Geminiviridae family. In the present work, the HTS strategy allowed for the identification of genetically divergent Tomato–associated geminivirus 1 (TaGV1) isolates infecting under natural field conditions in Brazil a novel host – table beet – and tomato. Viral contigs with 2,572 nucleotides in length were obtained with identity levels of 81% to previously reported TaGV1 isolates (from tomato and Cleome affinis) and ≈ 71% identity with a range of capulaviruses. Phylogenetic analyses of the TaGV1 isolates using either their complete genomes or Rep gene sequences indicated the genus Capulavirus as the closest group, whereas the CP gene information indicated a closest relationship with members of the genus Topocuvirus . Leaf curling and overall reduction in plant size were the prevalent TaGV1–associated symptoms observed in Nicotiana benthamiana plants after biolistic inoculation assays. No conspicuous symptoms were observed in inoculated tomato and table beet plants. However, systemic virus infection was confirmed in all three hosts by Southern hybridization and PCR assays. Further infectivity assays with a new set of table beet and tomato accessions as well as a large number of indicator plants are needed to establish the host range of this new geminivirus.
Corresponding Author:	Rita de Cássia Pereira de Carvalho Universidade de Brasilia Instituto de Ciencias Biologicas Brasília, DF BRAZIL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universidade de Brasilia Instituto de Ciencias Biologicas
Corresponding Author's Secondary Institution:	
First Author:	Rita de Cássia Pereira de Carvalho, Professor
First Author Secondary Information:	
Order of Authors:	Rita de Cássia Pereira de Carvalho, Professor
	Josiane Goulart Batista, Doctorate
	Felipe Fochat Silva, Undergraduate
	Flávia Milene Barros Nery, Doctorate
	Fernando Lucas de Melo, Professor
	Leonardo Silva Boiteux, PhD
	Maria Esther de Noronha Fonseca, Researcher
	Cristiano Lacorte, Researcher
	Dione Mendes Teixeira Alves-Freitas, Researcher
	Simone da Graça Ribeiro
Order of Authors Secondary Information:	

Characterization of genetically divergent Tomato–associated geminivirus 1 isolates from table beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*)

Josiane Goulart Batista[1], Felipe Fochat Silva[1], Flávia Milene Barros Nery[1], Fernando Lucas de Melo[1], Leonardo Silva Boiteux[2], Maria Esther de Noronha Fonseca[2], Cristiano Lacorte[3], Dione Mendes Teixeira Alves-Freitas[3], Simone da Graça Ribeiro[3], Rita de Cássia Pereira–Carvalho[*].

Corresponding authors:

Rita de Cássia Pereira–Carvalho, ORCID: 0000-0002-9905-189X, rcpcarvalho@unb.br

[1]Universidade de Brasília (UnB), Departamento de Fitopatologia, 70910–900, Brasília–DF, Brazil.

[2]Embrapa Hortaliças (CNPH), Brasília–DF, Brazil.

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

Abstract

Metagenomic approaches in conjunction with high–throughput sequencing (HTS) have been efficiently employed for the discovery of novel plant–associated viralspecies, including members of the *Geminiviridae* family. In the present work, the HTS strategy allowed for the identification of genetically divergent Tomato–associated geminivirus 1 (TaGV1) isolates infecting under natural field conditions in Brazil a novel host – table beet – and tomato. Viral contigs with 2,572 nucleotides in length were obtained with identity levels of 81% to previously reported TaGV1 isolates (from tomato and *Cleome affinis*) and ≈ 71% identity with a range of capulaviruses. Phylogenetic analyses of the TaGV1 isolates using either their complete genomes or *Rep* gene sequences indicated the genus *Capulavirus* as the closest group, whereas the *CP* gene information indicated a closest relationship with members of the genus *Topocuvirus*. Leaf curling and overall reduction in plant size were the prevalent TaGV1–associated symptoms observed in *Nicotiana benthamiana* plants after biolistic inoculation assays. No conspicuous symptoms were observed in inoculated tomato and table beet plants. However, systemic virus infection was confirmed in all three hosts by Southern hybridization and PCR assays. Further infectivity assays with a new set of table beet and tomato accessions as well as a large number of indicator plants are needed to establish the host range of this new geminivirus.

Keywords: beet, tomato, *Geminiviridae*, NGS

Introduction

Geminiviridae is a large family of plant–associated viruses, encompassing so far nine genera viz. *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Glabrovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus* (Varsani et al. 2017). *Citrus chlorotic dwarf associated virus* – CCDaV (Loconsole et al. 2012) and *Mulberrymosaic dwarf associated virus* – MMDaV (Lu et al. 2015, Ma et al. 2015) were also considered as members of the family. However, CCDaV and MMDaV were not yet classified at genus level (Varsani et al. 2017). Members of the *Geminiviridae* family exhibit small (2600–5200 nucleotides) single–stranded DNA (ssDNA) genomes organized in one or two circular molecules. Viral genomes are packed into particles formed by two incomplete icosahedra and the virus species are designated as either monopartite (only with the DNA–A component) or bipartite (with both DNA–A and DNA–B components) (Hesketh et al. 2018, Xu et al. 2019). A characteristic set of ORFs (open reading frames) is codified by members of the *Geminiviridae* family. In bipartite viruses, both components operate in a coordinated manner, ensuring replication and viral movement through the host plant tissues. In addition, distinctive genomic features are also present, including the common region – CR (with 180–200 nucleotides), which includes the virus replication origin (Ori), introns, and the conserved nonanucleotide "TAATATTAC" sequence (Argüello-Astorga et al. 1994, Laufs et al. 1995, Argüello–Astorga and Ruiz–Medrano 2001). Insects such as whiteflies, leafhoppers, aphids, and treehoppers are the natural vectors reported for many *Geminiviridae* viruses (Varsani et al. 2017, Zerbini et al. 2017). The associated vector species together with the information about host range, genomic organization and phylogenetic relationships are the criteria currently used for species classification in the *Geminiviridae* family (Varsani et al. 2017, Zerbini et al. 2017). The criteria for novel species demarcation in the *Geminiviridae* family are genomic pairwise identities ranging from 78% to 91% (Zerbini et al. 2017).

Begomovirus is considered as the most economically important genus within the *Geminiviridae* family in Neotropical areas (Rojas et al. 2018). However, the scenario in relation *Geminiviridae* diversity is sharply changing in this geographic region. Novel monopartite geminiviruses were recently identified, including Tomato associated geminivirus 1 (TaGV1) in Brazil (Fontenele, et al. 2017) and tomato apical leaf curl virus (ToALCV) in Argentina and Brazil (Vaghi Medina et al. 2018; Batista et al. 2018). TaGV1 has a monopartite genome organization with three genes in the viral sense [the coat protein gene (*CP*) and two overlapping genes of yet unknown function (one ofthem

is related to a putative movement protein *MP* gene and the other related to *Reg* gene)] and with two genes in the complementary sense, coding for replication–associatedproteins (*Rep* and *Rep*A). The presence of a long intergenic region (LIR) is also a genomic feature of TaGV1. Isolates of this virus were first described infecting plants of tomato and *Cleome affinis* DC in Brazil (Fontenele et al. 2017). On the other hand, ToALCV has been considered a type member of a new putative genus of the *Geminiviridae* family. The RepA, Rep and C3 proteins of the ToALCV isolates are similar to that of the capulaviruses, while the CP, MP and V3 proteins are related to topocuvirus, curtovirus and becurtovirus (Vaghi Medina et al. 2018), indicating a putative recombinant origin of this virus.

Metagenomic approaches in conjunction with high-throughput sequencing (HTS) have been efficiently employed for the discovery of new plant–associated viral species (Adams et al. 2009, Al Rwahnih et al. 2009, Kreuze et al. 2009), including members of the *Geminiviridae* family (Varsani et al., 2017). In the present study, we describe the complete genomes of genetically divergent TaGV1 isolates naturally infecting a novel host – table beet – and a previously reported host – tomato. In addition, we provide comparative analyses of these isolates with sequences of the previously reported TaGV1 isolates and also with ToALCV isolates from Argentina and Brazil (Vaghi Medina et al. 2017, Batista et al. 2018).

Materials & Methods

Sample collection, DNA purification, and rolling cycle amplification (RCA) assays

Five table beet and 15 tomato leaf samples were obtained from plants displaying virus–like symptoms during field surveys carried out in 2016 at Fazenda Água Limpa (UnB) in Brasília–DF. Table beet plants displayed leaf yellowing (Fig. 1A) and the tomato samples showed mosaic and leaf curling (Fig. 1B). Total DNA was purified with a modified CTAB protocol (Boiteux et al. 1999) and employed as template in RCA assays (Inoue–Nagata et al. 2004).

Sequencing, contig assembly, and virus identification

The RCA products were sequenced in an Illumina HiSeq 2500 platform at Macrogen Inc. (South Korea). A total 21,128,652 (2 × 100 paired–end) reads was obtained by the high throughput sequencing (HTS). Sequences were processed and assembled using CLC Genomic Workbench 9.0 version (Qiagen Bioinformatics). This

procedure allowed for the generation of 44,967 contigs, which were compared to a viral sequence database using BLASTx algorithm. The sequence of one contig (named contig–15 with ≈ 2,570 nts) displayed identity levels of 81% with the TaGV1 isolates from Cleome (MF072689) and tomato (MF072688) (Fontenele et al. 2017). This contig also showed ≈ 71% identity with a range of capulaviruses. Overlapping primers (Cap2KpnI-F 5'–GGT ACC CCC CTT GGA AAT GTA GTC TGC AAC–3' and Cap2KpnI-R 5'–GGT ACC TTT GAG GAG AGA GGT ATA CTT CG–3') containing a *Kpn*I restriction site were designed for detection and recovery of the complete genome of the contig–15. Amplicons with the expected size were observed after analysis of PCR products in 0.8% agarose gel electrophoresis. Specific amplicons were purified and cloned into pJET 1.2 plasmid (Thermo Fisher Scientific, San Jose, CA, USA) essentially as described (Fontenele et al. 2019). Three clones derived from two TaGV1 isolates from table beet and one from tomato were sequenced at Macrogen Inc. (South Korea) using a primer walking strategy. Sequences were assembled and analyzed in the Geneious R11 software (Kearse et al. 2012). The complete genome sequences were deposited at GenBank/NCBI database (www.ncbi.nlm.nih.gov) under the following accessions numbers: MN527305 (table beet isolate) and MN334783 (tomato isolate). Protein domains were identified in the translated ORFs employing the NCBI Conserved Domain tool (Marchler–Bauer et al. 2015) as well as with the InterProScan tool in the Geneious R11 package (Kearse et al., 2012).

Pairwise identity and phylogenetic relationships among isolates and viral species

For the phylogenetic analyses of TaGV1 isolates, a total of 39 representative sequences of the nine genera from *Geminiviridae* family were selected, including the previously reported TaGV1 isolates from *C. affinis* and tomato (Fontenele et al. 2017). Sequences of ToALCV isolates, a new monopartite geminivirus detected in Argentina (Vaghi Medina et al. 2018) and in Brazil (Batista et al. 2018) were also included in the analyses. The alignments were performed using the computer program MUSCLE (Edgar, 2004). For the phylogenetic analyses with the complete viral genome information, the following parameters were used: Neighbor–joining (NJ) inference and the tree with Jukes–Cantor nucleotide substitution model and 1,000 bootstrap replications using Geneious R11 software. Transcript amino acid analyses of the *CP* and *Rep* sequences were performed and a maximum likelihood phylogenetic tree was generated using the LG + G substitution model with the approximate likelihood ratio test (aLRT) branch support

in Geneious R11 software. The pairwise identity of the complete genomes as well as the amino acid sequences of the CP and Rep proteins of the 39 selected *Geminiviridae* family sequences were analyzed using SDT v1.2 (Muhire et al. 2014).

Biolistic inoculation/infectivity assays

The complete TaGV1genome of one table beet isolate was cleaved using the restriction enzyme *Kpn*. Afterward, it was self–ligated using T4 DNA ligase enzyme. The ligation product was precipitated (5 μg) in tungsten microparticles and bombarded into foliar tissues essentially as described by Blawid et al. (2013). Seven plants of each table beet (cultivar 'Boro'), *Nicotiana benthamiana*, and tomato cultivar 'Santa Clara' were employed in these assays. As negative controls, five plants of each species were bombarded with tungsten particles without DNA. The plants were kept in greenhouse benches. At 21 days after inoculation (DAI) leaf curling, individual table beet, *N. benthamiana*, and tomato plants were evaluated for symptom expression and total DNA extracted of each plant using a modified CTAB protocol (Boiteux et al. 1999). PCR assays were performed with the specific primers used for TaGV1 detection. The PCR products were purified, cloned and sequenced. The Sanger dideoxy sequencing of these amplicons was also carried out to confirm TaGV1 infection. Virus replication was assessed by Southern hybridization. The membrane was hybridized with a αP32 dCTP labeled probe specific for the full genome of the tomato TaGV1 isolates. The replicative forms of TaGV1 were analyzed in the inoculated plants by Southern hybridization assays.

Results

Identification of TaGV1–related contigs in HTS and virus detection in individual table beet and tomato samples

Virus replication in the field samples was assessed by Southern hybridization analyses. It was possible to detect replicative TaGV1forms in one tomato and one table beet plant. The complete sequence of the contig–15 (that displayed identity levels to TaGV1 isolates) was obtained using the previously described overlapping primers. These amplicons were cloned and the complete sequence was obtained via primer walking. The PCR product was purified, cloned and sequenced. The genomic information generated by Sanger dideoxy sequencing was employed to verify the obtained sequences of both table beet and tomato TaGV1 isolates. The complete genomes of table beet (MN527305) and

tomato (MN334783) isolates displayed 2,573 nts in size (Fig. 1C). They had 99% identity to each other, displaying only one single nucleotide polymorphism in the LIR and another in the *Rep* gene. Both isolates displayed in their genomes the conserved "TAATATTAC" nonanucleotide motif. Three ORFs were annotated in the viral sense viz. the *CP* gene (with 711 nts), codifying a protein with 237 amino acids (AAs) and two other superimposed ORFs representing a putative *Reg* gene (with 348 nts), codifying a protein of 116 AAs, and putative movement protein – *MP* gene (with 255 nts), codifying a protein of with 85 AAs. Two ORFs were annotated in the complementary sense viz. the *Rep* (with 999 nts), encoding a protein of 333 AAs (which is probably transcribed from a spliced transcript), and the overlapping *Rep*A gene (786 nts), codifying for a protein of 262 AAs. The N-terminal endonuclease domain displayed the conserved motif I (FLTYPK), motif II (PHIHC) and motif III unknown. The C-terminal superfamily 3 helicase domain contains conserved walker A (GTRCGKTAWAR), walker B (different from what has been described for geminiviruses – FDDIP) and motif C (VLCN) (Kazlauskas et al. 2018). The putative CP, Rep, and RepA proteins were found to be highly related to the corresponding geminivirus proteins. MP protein displayed a transmembrane domain, which was detected by the InterProScan. Reg protein did not share any conserved domains with the ones available in protein database.

Pairwise identity and phylogenetic relationships among isolates and viral species

The analyses of the complete genome and individual genes (*Rep* and *CP*) were carried out using SDT v1.2 (Muhire et al. 2014). Phylogenetic analyses were also done with our new TaGV1isolates and a total of 39 *Geminiviridae* sequences representativeof nine genera. Six outgroup species were selected as well as four TaGV1 isolates and two ToALCV isolates from Argentina and Brazil. In the phylogenetic analyses using the complete genome information, our tomato and table beet isolates displayed high divergence levels (with 83% pairwise identities) with the two previously described TaGV1 isolates. Our tomato and table beet isolates also showed 60–63% identity to capulaviruses, which are below the 78% threshold for novel species demarcation in the genus *Capulavirus* (Varsani et al. 2017). Moreover, comparing these novel tomato and table beet isolates with ToALCV isolates found in Argentina (Vaghi Medina et al. 2017) and in Brazil (Batista et al. 2018), the pairwise identities were 65% and 64%, respectively. With other geminiviruses the identity levels were below 56% (Fig. 1D and Fig. 1E). Analysis employing the *CP* gene information showed 92% identity with TaGV1 isolates,

46% with ToALCV isolates, 36% with ToPCTV (*Tomato pseudo-curly top virus* – genus *Topocuvirus*) and 19 to 27% with other geminiviruses. For the *Rep* gene, higher identities with sequences of TaGV1 isolates were also observed, showing 91% with previously described TaGV1 isolates and 72% with ToALCV and other capulaviruses. The *Rep* gene identity with other geminiviruses was around 28 to 34%. The phylogenetic analyses using the complete genome (Fig. 1D) and *Rep* genes clustered our tomato and table beet isolates with the TaGV1 isolates and ToALCV isolates as well as with the capulaviruses. In the *CP*-derived phylogenetic tree a relationship of TaGV1 and the ToPCTV isolates was observed. These genetic relationships indicate that TaGV1 isolates might be potential recombinants between capulaviruses and topocuviruses.

Biobalistic infectivity assays

At 21 days after inoculation, leaf curling and dwarfism were observed in some *N. benthamiana* plants. Plant death was also observed in one *N. benthamiana* plant. The PCR assays indicated that two symptomatic plants of *N. benthamiana* were infected by TaGV1 infection. The replicative forms of TaGV1 were detected in a single *N. benthamiana* plant. No conspicuous symptoms were observed in inoculated tomato and table beet plants. Mock-inoculated table beet and tomatoes plants were free of conspicuous symptoms during the entire evaluation period.

Discussion

Begomovirus is considered thus far as the most important genus within the *Geminiviridae* family affecting the tomato crop in South America. The table beet crop has been reported as a host of viral species of the *Geminiviridae* family in many countries around the world (Rojas et al. 2018). To our knowledge, no viruses of this group were detected infecting this vegetable crop in Brazil. In the present work we identified genetically divergent TaGV1 isolates in foliar tissues of table beet and tomato plants. TaGV1 isolates were previously described infecting tomato and plants of the weed *C. affinis* in Brazil (Fontenele et al. 2017). Our novel tomato and table isolates displayed 83% identity with the previously described TaGV1 isolates, 65% with ToALCV isolates and 63% with members of the genus *Capulavirus*. In addition, these novel isolates showed differences in the overall genome size when compared to previously described TaGV1 isolates from *C. affinis* and tomato, with a deletion of 38 nts being detected in the LIR region. The genomes of becurtoviruses, capulaviruses, grabloviruses, eragroviruses,

mastreviruses, and CCDaVs all have long intergenic region (LIR) and short intergenic region (SIR) (Argüello-Astorga and Ruiz-Medrano 2001, Londoño et al. 2010, Varsani et al. 2017). TaGV1 displayed two genes (*RepA* and *Rep*) that are potentially involvedin virus replication. *Rep* gene encodes a fusion protein in which a splicing processing step has been identified. Introns in either *Rep* or *Rep*A genes have already been described in becurtoviruses, capulaviruses, grabloviruses, mastreviruses, CCDaV, MMDaV and begomovírus (e.g. Mungbean yellow mosaic virus) (Wright et al. 1997, Shivaprasad et al. 2005, Bernardo et al. 2013, Candresse et al. 2014, Bozorgi et al. 2017). The sizes of *CP*, *Rep* and *Rep*A genes of the novel isolates were identical, corroborating previous observations that these proteins are highly conserved across geminiviruses (Varsani et al. 2017). The nucleotide identities of these novel TaGV1 isolates with the original ones were 68% for *RepA*, 92% for *CP* and 76% for *Rep*. The *CP* gene shares similarities with ToPCTV. For the putative *MP* gene, all the other isolates have 348 nts encoding a protein of 116 AAs with 81% identity. The putative Reg proteins displayed different sizes: 255 nts (85 AAs) in our novel isolates, whereas the original tomato and *C. affinis* isolates has 264 nts (with 88 AAs and 83% identity). Putative MP and Reg proteins share similarities with curtoviruses and becurtoviruses. Coincidentally, the genera *Becurtovirus* and *Curtovirus* have important species able to infect *B. vulgaris*, including the *Beet curlytop Iran virus* – BCTIV (Yazdi et al. 2008) and the *Beet curly top virus* – BCTV (Murphy et al. 1995), respectively. It is also important to highlight that BCTIV and BCTV isolates have been also reported infecting tomatoes in Iran (Heydarnejad et al. 2007) and Mexico (Guzmán et al. 1996), respectively.

The phylogenetic tree using complete genome information placed the novel table beet and tomato isolates into a clade composed by other isolates of TaGV1 and ToALCV as well as with *Capulavirus* species, indicating that they may share a common ancestor (Figure 1D). Similar relationships were observed in phylogenetic analyzes using the Rep protein information. Interestingly, using CP protein information, the TaGV1 isolates, instead of grouping with species of the genus *Capulavirus*, grouped with ToPCTV (genus *Topocuvirus*) isolates. ToPCTV is a virus which has not been found thus far in South America (Briddon et al. 1996). Therefore, there is a significant amount of evidences reinforcing the hypothesis of the recombinant nature of TaGV1 isolates, which have capulavirus and topocuvirus as the most likely ancestral strains (Fontenele et al. 2017). In fact, recombination has played a significant role in the evolution of geminiviruses (Padidam et al. 1999, Varsani et al. 2017). Topocuviruses and curtoviruses are examples

of a modular genomic arrangements that arisen by recombining fragments of distinct genera within the *Geminiviridae* family (Stanley et al. 1986, Briddon et al. 1996, Klute et al. 1996, Varsani et al. 2009). The identity of the nucleotide sequence of becurtoviruses virion sense genes such as the BCTIV is mainly related to that of short–lived viruses, while complementary sense genes are distally related to mastreviruses (Yazdi et al. 2008, Bozorgi et al. 2017).

Biolistic inoculation induced leaf curl symptoms in *N. benthamiana*, which is in contrast with previously generated TaGV1 clone that was also able to infect *N. benthamiana*, but without inducing symptoms (Fontenele et al. 2017). Therefore, these TaGV1isolates are also biologically distinct from the previous reported. Theinformation available so far about TaGV1 isolates indicates that they have relatively wide range of taxonomically distinct hosts such as tomato (Solanaceae), table beet (Amaranthaceae), and *Cleome affinis* (Cleomaceae). In addition, the complete genome of the TaGV1 isolates as well as their *Rep* gene sequences indicated a close relationship to ToALCV and capulaviruses. For *CP* gene-derived information, TaGV1 displayed a closer relationship to ToALCV and topocuviruses. Moreover, it was also found that TaGV1 isolates can infect and induce symptoms in *N. benthamiana*, but tomato and table beet are more likely asymptomatic hosts. Therefore, further studies will be conducted to investigate the geographical distribution and potential effects of TaGV1 across distinct cultivars/accessions of their host plants. Additional infectivity assays with a new set of table beet and tomato varieties as well as a wide range of indicator plants are needed to identify the host range of this new geminivirus, including plants of the genus *Cleome* from which one of the original TaGV1 isolates was also found (Fontenele et al. 2017).

Acknowledgements

This research was supported by grants from Embrapa, Capes, CNPq, and FAPDF.

References

Adams IP, Glover RH, Monger WA, 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.

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.

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, 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.

Batista JG, Melo FL, Pereira–Carvalho RC, Alves–Freitas DMT, Ribeiro SG (2018) First report of Tomato apical leaf curl virus infecting tomato in Brazil. Plant Disease doi.org/10.1094/PDIS-09-18-1636-PDN

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

Blawid R, Fontenele RS, Lacorte C, Ribeiro SG (2013) Molecular and biological characterization of corchorus mottle virus, a new begomovirus from Brazil. Archives of virology 158: 2603–2609.

Boiteux LS, Fonseca MEN, Simon PW (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.

Bozorgi N, Heydarnejad J, Kamali M, Massumi H (2017) Splicing features in the expression of the complementary–sense genes of Beet curly top Iran virus. Virus Genes 53: 323–327.

Briddon RW, Bedford ID, Tsai JH, Markham PG (1996) Analysis of the nucleotide sequence of the treehopper–transmitted geminivirus, tomato pseudo–curly top virus, suggests a recombinant origin. Virology 219: 387–394.

Candresse T, Filloux D, Muhire B, Julian C, Galzi S, Fort G, Bernardo P, Daugrois J–H, Fernandez E, Martin DP (2014) Appearances can be deceptive: Revealing a hidden viral infection with deep sequencing in a plant quarantine context. PloS one 9: e102945.

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.

Guzmán P, Goldberg N, Liddell C, Gilbertson R (1996) Detection and partial characterization of Beet curly top geminivirus (BCTV) isolates infecting beans, peppers and tomatoes in California and New Mexico. Annual Report of the bean improvement cooperative 39: 75–76.

Hesketh EL, Saunders K, Fisher C, Potze J, Stanley J, Lomonossoff GP, Ranson NA (2018) The 3.3 Å structure of a plant geminivirus using cryo-EM. Nature communications 9: 2369.

Heydarnejad J, Hosseini Abhari E, Bolok Yazdi H, Massumi H (2007) Curly top of cultivated plants and weeds and report of a unique curtovirus from Iran. Journal of phytopathology 155: 321–325.

Inoue–Nagata AK, Albuquerque LC, Rocha WB, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage phi29 DNA polymerase. Journal of virological methods 116: 209–211.

Kazlauskas D, Varsani A, Krupovic M (2018) Pervasive chimerism in the replication-associated proteins of uncultured single-stranded DNA viruses. Viruses 10: 187.

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

Klute KA, Nadler SA, Stenger DC (1996) Horseradish curly top virus is adistinct subgroup II geminivirus species with rep and C4 genes derived from a subgroup III ancestor. Journal of general virology 77: 1369-1378.

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.

Laufs J, Traut W, Heyraud F, Matzeit V, Rogers SG, Schell J, Gronenborn B (1995) In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. Proceedings of the national academy of sciences (PNAS) 92: 3879–3883.

Loconsole G, Saldarelli P, Doddapaneni H, Savino V, Martelli GP, Saponari M (2012) Identification of a single–stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. Virology 432: 162–172.

Londoño A, Riego–Ruiz L, Argüello–Astorga GR (2010) DNA–binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. Archives of virology 155: 1033–1046.

Lu Q–Y, Wu Z–J, Xia Z–S, Xie L–H (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (*Morus alba* L.). Archives of virology 160: 2135–2138.

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.

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

Murphy F, Fauquet C, Bishop D, Ghabrial S, Jarvis A, Martelli G, Mayo M, Summers M (1995) Virus taxonomy: Sixth report of the International Committee on Taxonomy of Viruses. Archives of virology 10: 350–354.

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

Rojas MR, Macedo MA, Maliano MR, Soto–Aguilar M, Souza JO, Briddon RW, Kenyon L, Rivera–Bustamante RF, Zerbini FM, Adkins S (2018) World management of geminiviruses. Annual review of phytopathology 56: 637–677.

Shivaprasad P, Akbergenov R, Trinks D, Rajeswaran R, Veluthambi K, Hohn T, Pooggin MM (2005) Promoters, transcripts, and regulatory proteins of Mungbean yellow mosaic geminivirus. Journal of virology 79: 8149–8163.

Stanley J, Markham P, Callis R, Pinner M (1986) The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus. The EMBO journal 5:1761–1767.

Vaghi Medina CG, Teppa E, Bornancini VA, Flores CR, Marino–Buslje C, Lopez Lambertini PM (2017) Tomato apical leaf curl virus: A novel, monopartite geminivirus detected in tomatoes in Argentina. Frontiers in microbiology 8: 2665.

Varsani A, Roumagnac P, Fuchs M, Navas–Castillo J, Moriones E, Idris A, Briddon RW, Rivera–Bustamante R, Murilo Zerbini F, Martin DP (2017) Capulavirus and Grablovirus: two new genera in the family Geminiviridae. Archives of virology 162: 1819–1831.

Varsani A, Shepherd DN, Dent K, Monjane AL, Rybicki EP, Martin DP (2009) A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virology journal 6: 36.

Wright E, Heckel T, Groenendijk J, Davies J, Boulton M (1997) Splicing features in maize streak virus virion and complementary–sense gene expression. The plantjournal 12: 1285–1297.

Xu X, Zhang Q, Hong J, Li Z, Zhang X, Zhou X (2019) Cryo–EM structure of a begomovirus geminate particle. International journal of molecular sciences 20: 1738.

Yazdi HB, Heydarnejad J, Massumi H (2008) Genome characterization and genetic diversity of beet curly top Iran virus: A geminivirus with a novel nonanucleotide. Virus genes 36: 539–545.

Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas–Castillo J, Rivera–Bustamante R, Roumagnac P, Varsani A, Consortium IR (2017) ICTV virus taxonomy profile: Geminiviridae. The Journal of general virology 98: 131.

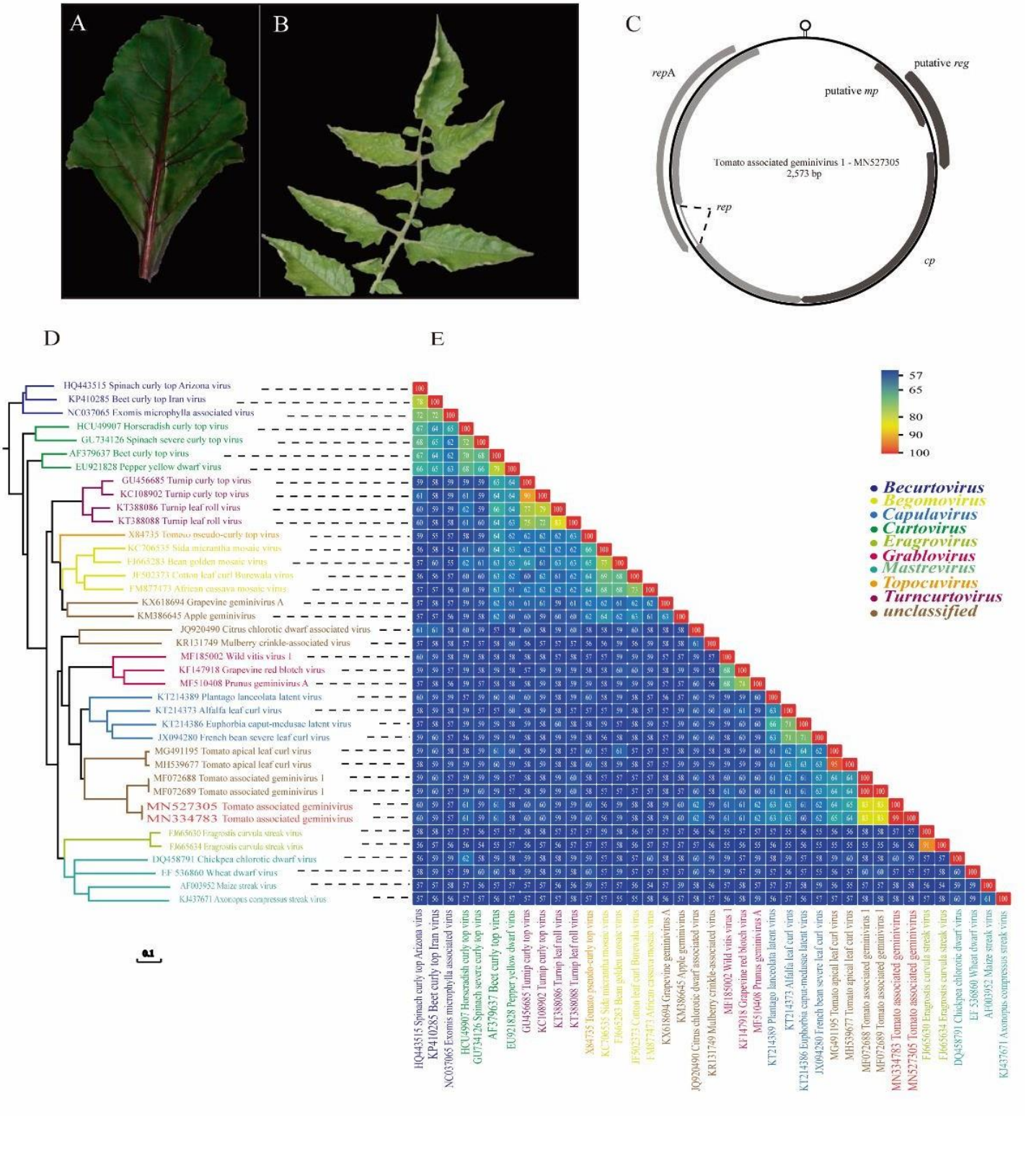


Fig. 1. (A) Field–collected table beet plant with leaf yellowing symptoms; **(B)** tomato plant with mosaic and leaf curling symptoms; **(C)** Genomic organization of Tomato associated geminivirus 1 – TaGV1 (GenBank MN527305) displaying five ORFs (open reading frames). In the viral sense were annotated the coat protein gene (*CP*) and two overlapping genes of yet unknown function (one of them is related to a putative movement protein – *MP* gene). In the complementary sense were annotated the genes coding for the replication-associated proteins (*Rep* and *Rep*A). **(D)** Neighbor–joining phylogenetic tree with 1,000 bootstrap replications. **(E)** Genome–wide pairwise matrix generated by SDT v1.2. Representative geminiviruses were employed for phylogenetic analyses and SDT, including all TaGV1 isolates as well as isolates of the recently discovered Tomato apical leaf curl virus (ToALCV) from Argentina and Brazil.

Attached V

First report of *Bean golden mosaic virus* in *Anadenanthera colubrina*.

Batista; JG., Fochat; F., Nery; FMB., et al (2020d) - Plant Disease/Disease Note

To submitted, **Attached V**

Molecular Confirmation of *Anadenanthera colubrina* (Fabaceae) as a Natural Host of *Bean golden mosaic virus* in Brazil.

J. G. Batista, F. F. Silva, Flávia M. B. Nery, F. L. Melo, L. N. A. Reis, and **R. C. Pereira-Carvalho**, Departament of Plant Pathology, Universidade de Brasília (UnB), Brasília–DF, Brazil; **M. E. N. Fonseca** and **L. S. Boiteux**, National Center for Vegetable Crops Research (CNPH), Embrapa Hortaliças, Brasília–DF, Brazil.

Bean golden mosaic virus (BGMV) is one of the most important viruses that occur in Brazil causing serious losses in bean production (*Phaseolus vulgaris* L. family Fabaceae family). In addition to beans, it has been identified in soybean (*Glycine max*) and weeds of the genus *Euphorbia* sp. and *Macroptilium* sp. As studies of tree species as host viruses are part of a surveillance work of the UnB Virology Laboratory, angico plants (*Anadenanthera colubrina* family Fabaceae) were collected at the Experimental Biology Station of the University of Biology (EEB-UnB).) and in the urban afforestation of the Federal District, in the second semester of 2016 with symptoms of chlorotic points. A. colubrina is an important tree species in South America (Viana et al., 2013) and widely distributed in Bolivia, Paraguay and Brazil (Von, 1964). In EEB-UnB there is a transition area where tree species are very close to field experiments with vegetables and whiteflies (*Bemisia tabaci* Middle East Asian Minor 1 - MEAM 1) Total DNA extraction was performed following a protocol adapted from (Boiteux et al., 1999) and enriched by RCA - Rolling Circle Amplification (Inoue-Nagata et al. 2004). The purified DNA was also subjected to Rolling Circle Amplification (Inoue-Nagata et al. 2004) and sequenced on the Illumina Hiseq 2500 platform in Macrogen (South Korea). Sequence analysis was performed on the CLC Genomics Workbench program after manually refining and mounting readings on higher quality contigs. Contigs were imported into Geneious software and compared via BLASTn algorithm to a GenBank viral database. Contigs

derived from the angico isolate - DF-126_JB-LVV-07 - BGMV for DNA-A (MN734371) exhibited ≈ 99% identities with various bean-infecting BGMV isolates (e.g., KJ939798) and *Macroptilium lathyroides* (e.g., JN419003) and soybean (e.g., FJ665283). The 2617 nts (nucleotides) of the DNA-A component and presented all five ORFs - nts / aa (aminoacids) AV1 (CP) - 756 nts / in the viral sense and AC1 (Rep) - 1086 nts /, AC2 (TrAP) - 390 nts /, AC3 (Ren) - 399 nts /, AC4 - 258nts / and AC5 - 756 nts / in the complementary sense. DNA-B contigs (MN734370) exhibited 97% identity for a bean BGMV isolate (MG334553). The 2584 nts of DNA - B component exhibited two ORFs (nts / yy): BV1 (NSP) - 762 nts / in the viral sense and BC1 (MP) - 882 nts / in the complementary sense. Phylogenetic analyzes, with all isolates of full BGMV - DNA-A sequences adding Angico isolate - DF-126_JB-LVV-07, indicated the formation of a distinct clade with *M. lathyroides* isolates.Thus, *A. colubrina* plants near commercial bean and soybean fields can serve as BGMV reservoirs in the inter-crop periods of these crops.

References

Boiteux, L. S., et al. 1999. J. Am. Soc. Hort. Sci. 124: 32.

de Viana, M. L., et al. 2014 Rev de Biol Trop 62:757–767.

Inoue-Nagata, A. K., et al. 2004. J. Virol. Meth. 116: 209.

Von Reis A, 1964. Taxonomic Study of Anadenanthera. Contributions from the Gray Herbarium of Harvard University 193:3–65.

Attached VI

Natural Infection of Sugarcane (*Saccharum officinarum*) by Maize striate mosaic virus (*Geminiviridae*) in Brazil (2020e)

Batista; JG., Fochat; F., Miranda; BEC., et al - Plant Disease/ Disease Note

Accepted.

plant disease

Natural Infection of Sugarcane (Saccharum officinarum) by Maize striate mosaic virus (Geminiviridae) in Brazil.

Journal:	*Plant Disease*
Manuscript ID	PDIS-12-19-2670-PDN
Manuscript Type:	Plant Disease Note
Date Submitted by the Author:	20-Dec-2019
Complete List of Authors:	Batista, Josiane; Universidade de Brasília, Fitopatologia Melo, Felipe ; Universidade de Brasília, Plant Pathology Miranda, Bruno ; Universidade de Brasilia, Plant Pathology Rossato, Mauricio; Universidade de Brasilia Instituto de Ciencias Biologicas, Departamento de Fitopatologia Melo, Fernando; Universidade de Brasília, Biologia Celular Ribeiro, Simone; Embrapa Recursos Genéticos e Biotecnologia, PBI Fonseca, Maria Esther; Embrapa Vegetable Crops, Plant Breeding Lab Boiteux, Leonardo; Embrapa Hortaliças, Plant Breeding Pereira Carvalho, Rita; Universidade de Brasília, Phytopathology
Keywords:	sugarcane, MSMV, mastrevirus

Natural Infection of Sugarcane (*Saccharum officinarum*) by Maize striate mosaic virus (*Geminiviridae*) in Brazil.

J. G. Batista, **F. F. S. Melo**, **B. E. C. Miranda**, **M. Rossato** and **F. L. Melo**, Dept. Fitopatologia, Universidade de Brasília (UnB), Brasília–DF, Brazil; **S. G. Ribeiro**, Embrapa Recursos Genéticos & Biotecnologia, Brasília–DF, Brazil; **M. E. N. Fonseca** and **L. S. Boiteux**, Embrapa Hortaliças (CNPH), Brasília–DF, Brazil; **R. C. Pereira–Carvalho**, Dept. Fitopatologia, Universidade de Brasília (UnB), Brasília–DF.

Brazil is currently the leading country in sugar and ethanol production from sugarcane (*Saccharum officinarum* L.). This crop is affected worldwide by many virus-induced diseases, which are responsible for severe yield losses (Gonçalves et al., 2012). Members of the *Mastrevirus* genus (family *Geminiviridae*) are efficiently transmitted by leafhoppers and are widespread across sugarcane-producing regions in Africa and Asia (Boukari et al., 2017). Recently, a new mastrevirus – maize striate mosaic virus (MSMV) – was described in association with maize (*Zea mays* L.) and leafhoppers in Brazil (Fontenele et al., 2018). During field surveys conducted in 2016 and 2018 in Brasília–DF, Central Brazil, we observed sugarcane and maize plants with mastrevirus-like symptoms (≈ 50% incidence). Five sugarcane leaf samples with mild chlorotic streak and 14 maize leaf samples with striate mosaic were collected. Total genomic DNA was extracted from each sample of both species using a modified CTAB method (Boiteux et al., 1999). Purified DNA was used as template in rolling-circle amplification (RCA) assays (Inoue-Nagata et al. 2004). Afterward, the RCA products were pooled and sequenced in an Illumina Hi-Seq2500 platform at Macrogen Inc. (South Korea), using a Nextera DNA Library Prep kit. The sequencing generated 29.882,284 paired-end reads, which were assembled with CLC Genomic Workbench 9.0 version (Qiagen Bioinformatics). The resulting 54,923 contigs were compared with a local viral RefSeq database using BLASTn and BLASTx algorithms in the Geneious® R11 package. One circular contig (with 2,746 nucleotides) was recovered, displaying identities higher than 99% with MSMV isolates from maize (MF167297–MF167307). For detection of the full-length genome by Sanger dideoxy sequencing, abutting primers were used (Fontenele et al. 2018) in individual sugarcane and maize samples. MSMV–specific amplicons were obtained from two maize and two sugarcane isolates. These amplicons were purified and Sanger-sequenced at Myleus (Minas Gerais, Brazil). The complete

genome of the sugarcane MSMV isolate (MN535988) displayed four ORFs (open reading frames) with structural features typical of mastreviruses. In the viral sense were annotated the V1 gene (codifying the coat protein – CP) and the V2 gene (= the movement protein – MP). In the complementary sense were annotated the C1 gene (= RepA – replication associated protein) and the C1:C2 (Rep). All gene products of the MSMV sugarcane isolate displayed identities ≥ 99% to the maize isolates. To our knowledge, this is the first global report of the natural infection of sugarcane by MSMV isolates. The potential negative impacts of MSMV infection in sugarcane are yet to be determined. However, sugarcane crops are maintained by vegetative propagation of elite clones, favoring long-term virus establishment under field conditions. Thus, the detection of this new mastrevirus may represent an alert to the sugarcane production sector in order to monitor its potential damage in this economically important field crop.

References

Boiteux, L. S., et al. 1999. J. Am. Soc. Hort. Sci. 124: 32.

Boukari, W., et al., 2017. Virol J. 14:116.

Fontenele, R. S., et al. 2018. Arch. Virol. 163: 263.

Gonçalves, M. C., et al. 2012. Funct. Plant Sci. Biotech. 108: 166.

Inoue-Nagata, A., et al. 2004. J. Virol. Meth. 116: 209.

Funding: Financial support was provided by grants from Embrapa, CAPES, CNPq and FAPDF.

13-Jan-2020

Dear Dr. Pereira Carvalho:

Note ID PDIS-12-19-2670-PDN entitled "Natural Infection of Sugarcane (Saccharum officinarum) by Maize striate mosaic virus (Geminiviridae) in Brazil.", which you submitted to Plant Disease, has been reviewed by two experts and by me. Both reviewers recommend the note with major revision and I agree with their assessment. The comments of the reviewers are included at the bottom of this letter or attached.

Therefore, I invite you to respond to the reviewers' comments and revise your note.

To revise your note, log in to https://mc.manuscriptcentral.com/plantdisease and enter your Author Center, where you will find your note listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your note number has been appended to denote a revision.

O isolado de BGMV detectado e recuperado em berinjela proveniente do Núcleo Rural Taquara (NRT) foi incluído na disease note: **Natural Infection of Tomato, Eggplant and *Nicandra physaloides* (Solanaceae) by *Bean golden mosaic virus* isolates in Brazil**. Esta disease note fará parte do trabalho da aluna de doutorado Luciane Reis e portanto não entrará na versão desta Tese.

Natural Infection of Tomato, Eggplant and *Nicandra physalodes* (Solanaceae) by *Bean golden mosaic virus* isolates in Brazil.

L. N. A. Reis, **J. G. Batista, F. F. S. Melo, F. L. Melo, R. C. Pereira-Carvalho,** and Dept. Fitopatologia, Universidade de Brasília (UnB), Brasília–DF, Brazil; **M. E. N. Fonseca** and **L.S. Boiteux,** Embrapa Vegetable Crops (CNPH), Brasília–DF, Brazil.

Bean golden mosaic virus (BGMV) is one of the most important pathogens of beans (*Phaseolus vulgaris* L.) and lima beans (*P. lunatus* L.) in Brazil. More recently, after the invasion of *Bemisia tabaci Middle East Asian Minor* 1, BGMV isolates have been also described infecting other Fabaceae species such as *Glycine max*, *Macroptilium lathyroides*, and *Mucuna pruriens* (Fernandes et al., 2009; Silva et al., 2012; Sobrinho et al., 2014). During field surveys leaf samples of tomato (*Solanum lycopersicum* L.), eggplant (*Solanum melongena* L.) and shoo-fly plant (*Nicandra physalodes* (L.) Pers.) exhibiting golden leaf mosaic, overall dwarfism and apical chlorotic spots were collected across Central Brazil from 2001–2016. Total DNA was extracted with a CTAB method and used as template in rolling circle amplification (RCA) reactions (Inoue-Nagata et al., 2004). RCA-amplified samples were pooled and sequenced in an Illumina Hi-Seq2500 platform at Macrogen Inc. (South Korea) using a Nextera DNA Library Prep kit. The sequencing generated reads, which were processed and assembled with CLC Genomic Workbench 9.0 version (Qiagen Bioinformatics) resulting in six contigs with identity with BGMV. The contigs were compared with a local viral RefSeq database using BlastN and BlastX using Geneious R10.2. Three BGMV-related DNA–A (with 2,617 - 2,626 nucleotides - nts) and DNA–B (with 2,583 – 2592 nts) contigs were identified. Universal begomovirus primers for A (PAL1978/PAR496) and B (PBL1v2040/PCRc1) were used for detection (Rojas et al., 1993). Besides, BGMV-specific PCR primers were designed for the DNA–A segment BGMV_A_For (5’-GTG CGT GAA TCC ATG ACC GT -3 '), BGMV_A_Rev (5'-ATT CAC GCA CAG GGG AAC G -3’). These primers were used in PCR reactions to detect into the pools the BGMV-infected individual plants. The PCR products were directly Sanger sequenced (ABI Prism 3100) at CNPH. It was possible to

detect BGMV in three tomato samples collected in 2003 in Leopoldo de Bulhões, Goiás State (isolate GO–142), and collected in 2003 in Gama, the Federal District–DF (DF–045 and DF–046). BGMV was also found in one *N. physalodes* sample collected in Brazlândia–DF in 2016 (DF–644) and in three eggplant sample **(**DF–308R, DF–312R and DF–318R) collected in Brasília–DF in 2016. All isolates had a complete genome recovered by HTS. These isolates displayed DNA–A genome identities between 99.2–99.7% to a wide range of BGMV isolates. The entire DNA–A components from the isolates DF–045 (MN737552), DF–046 (MN737553), GO–142 (MN737554), DF–644 (MN737555) and eggplant (DF–308R, DF–312R and DF–318R), displayed six open reading frames: coat protein gene (AV1), replication associated protein gene (AC1), trans-acting protein gene (AC2), replication enhancer gene (AC3), symptom determinant gene (AC4 and AC5). The DNA–B component was characterized only in the *N. physalodes* sample and displayed two ORFs: nuclear shuttle protein (BV1) and movement protein (BC1). To our knowledge, this is the first report of BGMV infecting members of the Solanaceae family, expanding the host range of this viral species. The identification of BGMV isolates able to infect tomatoes, eggplant and *N. physalodes* may intensify the flow and exchange of genetic information among viral genotypes in mixed infections in these new hosts, favoring the emergence of novel *Begomovirus* species with the potential ability to infect across Fabaceae and Solanaceae members.

References:

Fernandes, F. R. et al. 2009. Arch. Virol. 154:1567.

Inoue-Nagata, A. et al. 2004. J. Virol. Meth. 116:209.

Rojas, M. R., et al. 1993. Plant Dis. 77:340.

Silva, S. J. C. et al. 2012. Plant. Pathol. 61:457.

Sobrinho, R. R. et al. 2014. J. Gen. Virol. 95:2540.

Funding: This research had financial support from grants from Embrapa, CAPES, and CNPq. L. N. A. Reis and J.G. Batista were supported by a scholarship from CAPES and CNPq.