

RAFAEL REIS DE REZENDE

**CARACTERIZAÇÃO MOLECULAR E FILOGENÉTICA DE DOIS  
GENOMOVÍRUS ASSOCIADOS A PLANTAS SILVESTRES NO  
BRASIL**

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

VIÇOSA  
MINAS GERAIS - BRASIL  
2017

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

T

R467c  
2017 Rezende, Rafael Reis de, 1991-  
Caracterização molecular e filogenética de dois  
genomovírus associados a plantas silvestres no Brasil / Rafael  
Reis de Rezende. – Viçosa, MG, 2017.  
viii, 34f. : il. (algumas color.) ; 29 cm.

Orientador: Poliane Alfenas Zerbini.  
Dissertação (mestrado) - Universidade Federal de Viçosa.  
Inclui bibliografia.

1. Vírus de planta. 2. *Genomoviridae*. 3.  
*Gemycircularvirus*. 4. Vírus de ssDNA circular. I. Universidade  
Federal de Viçosa. Departamento de Microbiologia. Programa de  
Pós-graduação em Microbiologia Agrícola. II. Título.

CDD 22. ed. 572.8

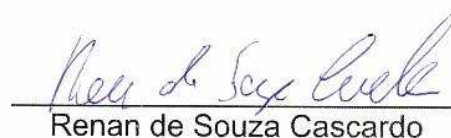
RAFAEL REIS DE REZENDE

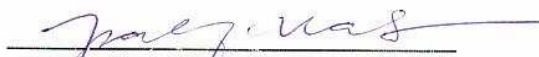
**CARACTERIZAÇÃO MOLECULAR E FILOGENÉTICA DE DOIS  
GENOMOVÍRUS ASSOCIADOS A PLANTAS SILVESTRES NO  
BRASIL**

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

APROVADA: 24 de fevereiro de 2017.

  
Marisa Vieira de Queiroz

  
Renan de Souza Cascardo

  
Poliane Alfenas Zerbini  
(Orientadora)

## **AGRADECIMENTOS**

Ao Triúno Criador, Autor da vida, por me inspirar no estudo da sua criação;

Aos meus pais por todo cuidado e conhecimento dado a mim;

A minha irmã pelo companheirismo e apoio;

A minha família por me ensinar os meus valores;

Ao meu Mestre, Peter, meu primeiro orientador da vida;

As minhas amigas Andréa, Claudinéia, Débora e Stefanni pelas orações, conselhos, encorajamento e principalmente pelas confusões;

A Debora por ser minha primeira cúmplice e aprendiz;

Ao meu marido Felipe, companheiro e amigo, pela paciência, apoio, cuidado e principalmente por me amar;

A Professora Juliana (UFSJ) que me acolheu e me ensinou a amar a microbiologia;

Ao MIND tumulto, esse laboratório alegre e extrovertido que torna nossos dias científicos de pura risadas e confusões;

A minha “Chefa” Poli, Professora Poliane, pelo valioso ensino em virologia e pelos conselhos para vida, pela fé em mim e paciência. Por me fazer apaixonar por vírus e em especial por toda cúmplice;

Aos Professores André, César, Fernandinha e Renan por toda disposição de sempre me ensinar as coisas que não estão nos artigos, mas sim na experiência;

Aos Professores Murilo e Denise pelas melhores aulas que tive durante o curso;

A todos nós do Mind pela amizade, parceria, malandragem e cumplicidade, mas em especial a:

Flávia (Coxinha) por todo apoio e cuidado e principalmente pelo vai...;

Fernanda pelas histórias de Mato Grosso e os lanchinhos fitness;

Paty por sempre se lembrar de tudo e sempre nos obrigar a fazer o que ela quer, e pelo laço que nos dois criamos;

Jú pelos momentos mais divertidos e das confusões que só aconteciam com nos dois juntos;

Aos estagiários Luan (nosso exemplo de paciência e seriedade), Bruna, Marcela e Anielly pela oportunidade de me ensinar a ensinar, a Thamylles (pelas atualizações das coisas que ocorrem fora do lab., é principalmente pelo companheirismo);

A toda equipe da pós-graduação que sobreviveu a mim;

A CAPES, CNPq e FAPEMIG pelo apoio financeiro.

À Universidade Federal de Viçosa e ao Departamento de Microbiologia pela oportunidade de crescimento pessoal e profissional.

A todos vocês sou grato e oro para que Deus os abençoe!

*Como o ferro afia o ferro, um amigo afia o outro.*

*Sábio Salomão*

***Combati o bom combate, completei a corrida, perseverai na fé!***

**Paulo, o apóstolo**

**Segunda carta a Timóteo, capítulo 4, verso 7.**

## **BIOGRAFIA**

*Sou biólogo, pesquisador, cristão adventista e reformador, desbravador, filho de Rosimar Reis e Onofre Mizael de Rezende. Em 2010, ingressei na Universidade Federal de São João del-Rei no curso de Ciências Biológicas. Durante a graduação me envolvi centro acadêmico, coordenação do curso e principalmente com a pesquisa. Junto a Professora Juliana Pereira Lyon defini minha trajetória como pesquisador em microbiologia.*

*No ano de 2015 fui aprovado neste programa para cursar o mestrado sob a orientação da Professora Poliane Alfenas Zerbini. Durante minha trajetória acadêmica nessa instituição fui eleito membro da Comissão Organizadora do Programa. Juntamente com os meus colegas participei da construção do Núcleo de Estudos em Microbiologia (NEMA), que tem como objetivo pesquisa, extensão e ampliação do ensino dos estudantes do programa. Quanto a minha pesquisa, junto à Professora Poliane, temos desenvolvido projetos de notável relevância científica. Meu crescimento nessa instituição tem sido de alto aproveitamento para minha vida acadêmica. Portanto sei que ainda tenho muito para crescer nessa instituição, o que me motivou a realizar meu doutorado na UFV.*

*Rafael Reis de Rezende*

## SUMÁRIO

RESUMO .....	vii
ABSTRACT.....	viii
INTRODUÇÃO.....	1
REVISÃO BIBLIOGRÁFICA .....	4
Biodiversidade viral .....	4
Genomovírus.....	8
REFERÊNCIAS BIBLIOGRÁFICAS.....	14
Identification and Characterization of Two <i>Gemycircularvirus</i> associated with wild plants in Brazil. ....	20
Identification and Characterization of Two <i>Gemycircularvirus</i> associated with wild plants in Brazil. ....	21
Abstract .....	21
Figure legends.....	25
References .....	26
Figure 1 .....	32
Figure 2 .....	33
CONCLUSÕES.....	34

## RESUMO

REZENDE, Rafael Reis de, M.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **Caracterização molecular e filogenética de dois genomovírus associados a plantas silvestres no Brasil.** Orientador: Poliane Alfenas Zerbini. Coorientador: Francisco Murilo Zerbini Júnior.

Vírus que possuem genoma constituído por DNA fita simples circular são amplamente distribuídos na natureza. As constantes descobertas aliadas à grande diversidade genética desses vírus têm ampliado a compreensão sobre a trajetória evolutiva desse grupo. A partir de 2010 um grupo divergente de vírus com genoma de DNA fita simples circular, denominado genomovírus, passou a ser identificado em diversos ambientes (fezes e esgoto) e associados a diferentes organismos (plantas, fungos, humanos, mamíferos, moluscos e insetos). Neste trabalho foram caracterizados dois genomovírus denominados *Momordica charantia associated gemycircularvirus* (MorGemy) e *Euphorbia heterophylla associated gemycircularvirus* (EupGemy) encontrados associados às plantas silvestres *Momordica charantia* e *Euphorbia heterophylla* respectivamente. Ambos os vírus possuem organização genômica típica dos genomovírus. Ambos os vírus apresentam *stem-loop* com nanonucleotídeo conservado em todos os genomovírus, sendo a sequência TAATRTTAT (N pode ser A ou G). Entretanto, no isolado MorGemy foi observado uma estrutura em *stem-loop* extra na possível origem de replicação não presente em outros genomovírus já descritos. Em ambos os vírus foi observada a presença de um íntron, no interior da ORF que codifica a proteína Rep sendo que o íntron observado em EupGemy é maior que o observado em MorGemy. Na proteína Rep foram observados os motivos I, II e III, e os domínios Walker A, B, C e GRS, elementos encontrados tipicamente na proteína Rep dos geminivírus. O SDT (Sequence Demarcation Tool) revelou que MorGemy tem 72% de identidade com *Pteropus associated gemycircularvirus 3* e *Hypericum associated gemycircularvirus 1*. Além disso, o EupGemy mostra 66% de identidade com *Odonata associated gemycircularvirus 1*. A taxa de identidade de nucleotídeos do genoma completo dos vírus isolados neste trabalho com outros genomovírus já descritos, permite classificar esses novos isolados como duas novas espécies.

## ABSTRACT

REZENDE, Rafael Reis de, M.Sc., Universidade Federal de Viçosa, February, 2017. **Molecular and phylogeny characterization of two genomoviruses associated with wild plants in Brazil.** Advisor: Poliane Alfenas Zerbini. Co-advisor: Francisco Murilo Zerbini Júnior.

Viruses that have genomes consisting of simple circular DNA are widely distributed in nature. The constant discoveries allied to the great genetic diversity of these viruses have broadened the understanding of the evolutionary trajectory of this group. In 2010 a divergent group of viruses with a circular ssDNA genome called Genomovirus was identified in several environments (feces and sewage) and associated with different organisms (plants, fungi, humans, mammals, mollusks and insects). In this work, two Genomoviruses found associated with wild plants were characterized. Both virus shows a structure Genomovirus-like. However two interesting facts were observed. The first, a presence a second structure on stem-loop in the MorGemy no related in other Genomovirus. Both virus presents stem-loop with nanonucleotide conserved in all Genomovirus. Being that the sequence TAATR TTAT (N could be A or G). We also observed the presence of a íntron in both genome, inside the ORF's Rep. The íntron of EufGmey is greater than MorGemy. In the Rep protein were observed the motif I, II and III. Interestingly, also were observed the presence of Motifs Walker A to C and GRS domain. These elementes are found on protein Rep's Geminivirus. The SDT releaved that MorGemy shows 72% of identity with *Pteropus associated gemycircularvirus 3* and *Hypericum associated gemycircularvirus 1*. In addition EupGemy shows 66% of identity with *Odonata associated gemycircularvirus 1*. The rate of identity of MorGemy and EupGemy with Genomovirus associate with plants combined with analysis phylogeny are an indicative that these virus can have a possible host-virus relation with plants.

## INTRODUÇÃO

Os vírus são micro-organismos amplamente distribuídos pelo planeta, encontrados infectando organismos de todos os três domínios da vida (*Archaea*, *Bacteria* e *Eukarya*) e associados a diversos ambientes. Além disso, os vírus são um reservatório genético inexplorado que possui grande importância ecológica como por exemplo, a capacidade de modular o crescimento de vários organismos e modulação dos ciclos bioenergéticos levando a liberação de matéria orgânica em grandes quantidades (Suttle, 2005; Ignacio-Espinoza et al., 2013; Suttle, 2007), demonstrando que os vírus possuem importantes papéis ecológicos, muitos ainda não explorados.

As descobertas de novos vírus têm demonstrado a existência de uma ampla diversidade de vírus de DNA fita simples (ssDNA) com genoma circular (Sierra et al., 2015; Kim et al., 2011). O estudo desses novos vírus depende da capacidade de obter amostras de genomas em quantidades e qualidades suficientes para poder realizar a caracterização molecular, evolutiva e biológica. Alcançar essa biodiversidade tem sido possível através de duas estratégias: metagenômica (Edwards and Rohwer, 2005; Mokili et al., 2012) e amplificação por círculo rolante (Johne et al., 2009; Haible et al., 2006; Inoue-Nagata et al., 2004). Portanto, uma ampla gama de novos vírus que não apresentam relação com doenças ou de interesse agrônômico têm sido descobertos em associação a diversos organismos e ambientes. Esses novos achados têm evidenciado que a biodiversidade viral apresenta um potencial ecológico a ser explorado (Simmonds et al., 2017).

Nesse contexto, genomovírus são vírus de DNA fita simples circular que formam uma nova família denominada de *Genomoviridae*. Esse grupo de vírus vem sendo descoberto associado a diversos ambientes e organismos por todo o mundo (Krupovic et al., 2016). O primeiro relato de um isolado dessa família foi realizado por Yu et al., (2010) infectando um fungo fitopatogênico e comopolita, o *Sclerotinia sclerotiorum*. O genomovírus *Sclerotinia sclerotiorum associated gemycircularvirus* (SsGemy-1) é capaz de induzir um efeito de hipovirulência sobre o fungo hospedeiro. As análises das sequências mostraram que o genoma deste vírus é composto por DNA circular de fita simples, com duas ORFs, uma codificadora da proteína capsial (na fita viral sentido senso) e outra codificadora da proteína associada a replicação (Rep) (na fita complementar sentido anti-senso). A proteína Rep do SsGemy-1 contém dois domínios conservados importantes para replicação por círculo rolante, presente também na Rep dos geminivírus (Rosario et al., 2012b). Além disso, apresentam uma estrutura de “*stem-loop*” na região de origem de replicação, onde há uma sequência de nonanucleotídeos conservada (5'-TAATRTTAT-3'). Esta estrutura também está presente nos geminivírus, nanovírus e nos circovírus (Comitê Internacional de Taxonomia Viral – ICTV. Disponível em: <https://talk.ictvonline.org/ictv-reports/>). Os genomovírus são um grupo bem caracterizado quanto à estrutura molecular e filogenia, no entanto, há pouca informação quanto a sua biologia, em especial quanto a sua gama de hospedeiro (Krupovic et al., 2016).

Devido à grande biodiversidade viral e seus efeitos ecológicos o estudo da composição dessa biodiversidade e suas relações ecológicas com outros organismos e ambientes são de extrema importância. Um exemplo é que os

vírus que infectam plantas silvestres podem se tornar infecções emergentes em sistemas agrícolas. Além disso, podem ser fontes de diversidade genética para os vírus já estabelecidos por meio da recombinação, dando origem a isolados ainda mais agressivos (Roossinck and García-Arenal, 2015). Neste trabalho, foram caracterizados molecularmente e filogeneticamente dois genomovírus encontrados associados a duas plantas silvestres, *Momordica charantia* e *Euphorbia heterofila*, no sudeste e sul do Brasil respectivamente.

## REVISÃO BIBLIOGRÁFICA

### Biodiversidade viral

Os vírus são micro-organismos parasitas celulares obrigatórios capazes de infectar diversos organismos distribuídos por todo planeta, apresentando uma alta abundância de espécies e tipos. Estudos têm estimado que existem em média  $10^{31}$  partículas vírus, com uma biomassa em torno de 200 milhões de toneladas em todo globo terrestre. Em oceanos essa abundância chega a  $\sim 10^7$  partículas virais por  $\text{mL}^{-1}$  (Suttle, 2013). Em consequência dessa alta abundância viral, os vírus estabelecem funções ecológicas de extrema importância, por exemplo, vírus marinhos são responsáveis por catalisar e modular a produção de nutrientes lisando micro-organismos e liberando nutrientes para o fitoplâncton, que sustentam toda cadeia alimentar nos oceanos. Além disso, o carbono liberado da lise celular na forma orgânica é usado pelo fitoplâncton, permitindo o crescimento populacional, consumindo carbono atmosférico e, conseqüentemente, favorecendo o ciclo do carbono (Carcer et al., 2015; Labonté and Suttle, 2013; Suttle, 2005). Esses eventos evidenciam que os vírus possuem papéis de grande importância biológica. Estudos de viromas estimam que existem milhares de genótipos de vírus de DNA a cada 200 L de água do mar, confirmando a abundância e a diversidade de vírus no nosso planeta. Em contraste, não sabemos muito sobre a composição, papel ecológico e evolutivo dessa importante biodiversidade. Esse fato associado a sua alta abundância, demonstra a importância do estudo da composição e papel biológico da biodiversidade viral (Breitbart and Rohwer, 2005; Suttle, 2005; Simmonds et al., 2017)

A exploração dessa biodiversidade viral depende diretamente de ferramentas que permitam a obtenção de genomas virais em quantidades e qualidade suficientes para análises moleculares, filogenéticas e, principalmente, biológicas. Atualmente duas ferramentas têm se destacado na exploração da biodiversidade viral. São elas a Metagenômica (Edwards and Rohwer, 2005) e a Amplificação por Círculo Rolante (*Rolling Circle Amplification* - RCA) (Haible et al., 2006; Inoue-Nagata et al., 2004). Exploração por metagenômica é uma estratégia que permite a obtenção de informação genômica em grandes quantidades. Essa técnica permite obter sequências a partir de amostras ambientais sem a necessidade de cultivo dos organismos. Inicialmente, a metagenômica foi desenvolvida com a finalidade de explorar as funções e a gama de sequências genômicas em uma comunidade microbiana. Nessa abordagem, os pesquisadores, analisando os dados metagenômicos, são capazes de identificar metabólitos microbianos possibilitando a descoberta de novos antibióticos e enzimas (Streit and Schmitz, 2004). O primeiro trabalho usando metagenômica no estudo de diversidade viral ocorreu em 2002 por Breitbart et al., (2002) que realizou um estudo de biodiversidade viral em amostras marinhas. Desde então, o número de trabalhos de biodiversidade viral a partir de amostras de diversos ambientes (água marinhas, solo, esgoto, entre outros) usando a metagenômica tem aumentando. Dessa forma, uma grande quantidade de novos vírus que não apresentam importância clínica e/ou econômica tem sido relatados (Mokili et al., 2012). Os estudos de metagenômica têm dado acesso a um grande número de sequências possibilitando análises da biodiversidade (riqueza e abundância de espécies). No entanto, os resultados de análises

metagenômicas não geram informações biológicas, portanto, uma grande quantidade de informações genômicas geradas por esses trabalhos não é associada a outros vírus já bem caracterizados. Sendo assim promovem uma maior exploração da composição da biodiversidade viral, porém não possibilitam a compreensão das associações e relações biológicas com os vírus já caracterizados. Sendo assim, o ICTV tem proposto que essas sequências sejam incluídas na taxonomia viral, baseada nas sequências, filogenia e nos fenótipos previstos a partir destas, sem a necessidade das informações biológicas. (Simmonds et al., 2017). Poder mensurar a composição da biodiversidade viral é um fator muito importante. Mas, além disso, são de fundamental importância a elucidação do papel biológico desses micro-organismos nos ecossistemas e suas relações com seus hospedeiros.

Em contraste com a metagenômica a RCA permite a obtenção de genomas virais completos a partir de amostras ambientais, permitindo assim os estudos biológicos desses vírus. No entanto, essa técnica é restrita a genomas de ssDNA circular. Na RCA não há necessidade do uso de oligonucleotídeos específicos, assim, são usados oligonucleotídeos randômicos com poucos nucleotídeos permitindo à amplificação total do genoma sem a necessidade do conhecimento prévio da sequência ou de marcadores moleculares. Na RCA é usada a enzima DNA polimerase do bacteriófago phi29. Essa enzima tem a capacidade de sintetizar longas fitas de nucleotídeos com alta fidelidade. (Rector et al., 2004). A RCA está baseada na replicação por círculo rolante (RCR) usado por vírus de ssDNA circular (Comitê Internacional de Taxonomia Viral – ICTV. Disponível em: <https://talk.ictvonline.org/ictv-reports/>) e plasmídeos (Ruiz-Masó et al., 2015)

para replicar o genoma. Nesse sistema de replicação a DNA polimerase inicia o processo de polimerização a partir de um primer randômico, continuando a síntese até encontrar novamente o ponto de início da replicação. Quando isso acontece, ela provoca um deslocamento a jusante produzindo um concatamêro com alto peso molecular, com sequências repetidas de DNA fita simples (Rector et al., 2004; Ruiz-Masó et al., 2015).

As descobertas de novos vírus têm demonstrado à existência de uma ampla diversidade de vírus de ssDNA com genoma circular (Phan et al., 2015; Sierra et al., 2015; Kim et al., 2011). Os vírus de DNA fita simples compreendem um grupo cosmopolita, associados a hospedeiros em todos os domínios da vida (Ignacio-Espinoza et al., 2013). Além disso, são altamente diversos e de grande importância econômica, ambiental e clínica (Krupovic, 2013). Atualmente, são reconhecidas nove famílias de vírus com genomas de ssDNA. *Anelloviridae*, *Circoviridae* e *Parvoviridae* são os vírus associados a humanos, primatas, mamíferos e pássaros. As famílias *Inoviridae* e *Microviridae* estão associados a bactérias. Já as famílias *Geminiviridae* e *Nanoviridae* são vírus associados a plantas. Por fim, a família *Genomoviridae*. Essas oito famílias são bastante distintas entre si compartilhando apenas a composição do genoma de ssDNA. Entre essas diferenças estão a gama de hospedeiro, o tamanho e a estrutura do genoma, circular (Comitê Internacional de Taxonomia Viral – ICTV. Disponível em: <https://talk.ictvonline.org/ictv-reports/>).

O estudo da biodiversidade de vírus de ssDNA circular é um campo com alto potencial de exploração, por exemplo (Carcer et al., 2015) em um estudo de biodiversidade viral de água doce mostrou que 86% da comunidade viral é

composta por esses vírus de ssDNA, sendo que 65% desses vírus não apresentam similaridade com genomas virais conhecidos. Além disso, neste estudo os vírus de ssDNA não conhecidos, apresentaram similaridade de 79% a 89% com *Sclerotinia sclerotiorum associated gemycircularvirus-1* - o primeiro genomovírus descrito (Yu et al., 2010; Krupovic et al., 2016).

## **Genomovírus**

YU et al., (2010) realizaram o primeiro relato de um genomovírus, inicialmente denominado de *Sclerotinia sclerotiorum associated DNA vírus -1* e atualmente reconhecido como *Sclerotinia sclerotiorum associated gemycircularvirus-1* (SsGemy-1) (Krupovic et al., 2016). Um fato interessante é que o SsGemy-1 foi o primeiro relato de um vírus de genoma de DNA capaz de infectar um fungo. VAN DEN BRAND et al., 2012 relataram a descoberta de um genomovírus do gênero *Gemycircularvirus* associado a mamíferos (*Meles meles*). Por meio de análises filogenéticas os autores observaram que esses genomovírus formavam um clado fechado. Sendo assim, eles propuseram denominar provisoriamente esse grupo como uma nova família denominada de *Breviridae*, em referência ao seu pequeno genoma. No entanto, (Rosario et al., 2012a) em um estudo de diversidade viral de ssDNA circular em libélulas (*Odonata:Epiprocta*) encontraram vários genomovírus associados a esses insetos. Junto com outros achados de genomovírus associados a plantas (Dayaram et al., 2012), mamíferos (Van den Brand et al., 2012), mosquitos (Ng et al., 2011) e fungo (Yu et al., 2010). Por meio de análises filogenéticas propuseram nomear provisoriamente esse grupo em um novo gênero sem família denominado de *Gemycircularvirus* (Ge: Geminivírus-like, my:

mycovirus, circularvirus: genoma circular). O termo *Gemycircularvirus* foi o melhor aceito pela comunidade científica do que a nomenclatura provisória de *Breviridae*. Já em 2016, Krupovic et al. propuseram a criação de uma nova família denominada de *Genomoviridae* e a criação de oito gêneros dentro dessa família.

Genomovírus são vírus de ssDNA circular não envelopados com genoma em torno de 2200 nucleotídeos. O genoma desse grupo de vírus codifica duas putativas ORFs, em unidades transcricionais bidirecionais, sendo a CP (proteína capsial) codificada na fita viral e a Rep (proteína associada à replicação), codificada na fita complementar. Entre as extremidades 5' das duas ORFs, existe uma região intergência, denominada *large* e em alguns vírus, entre as extremidades 3' existe uma segunda região intergênica, denominada *small*. A presença de uma ou duas regiões intergênicas são usadas para diferenciar os dois tipos de genomas que os Genomovírus apresentam. O genoma tipo I é o que contém duas regiões intergênicas, e o genoma tipo II é o que contém apenas uma região intergênica (Figura 1). Quando não há presença da região intergênica *small* ocorre uma sobreposição das ORFs nos terminais 3' (genoma tipo II). A região intergênica *large* contém uma estrutura de *Stem-loop* com um nonanucleotídeo TAATR TTAT, importante para o início da replicação por círculo rolante (Rosario et al., 2012b). Vírus com genoma do tipo II possuem um íntron na região codificadora da Rep. Provavelmente esse íntron no interior da ORF da Rep passa por um processo de *splicing* gerando a proteína Rep funcional (Figura 1) (Sikorski et al., 2013; Male et al., 2016; Krupovic et al., 2016; Rosario et al., 2012a).

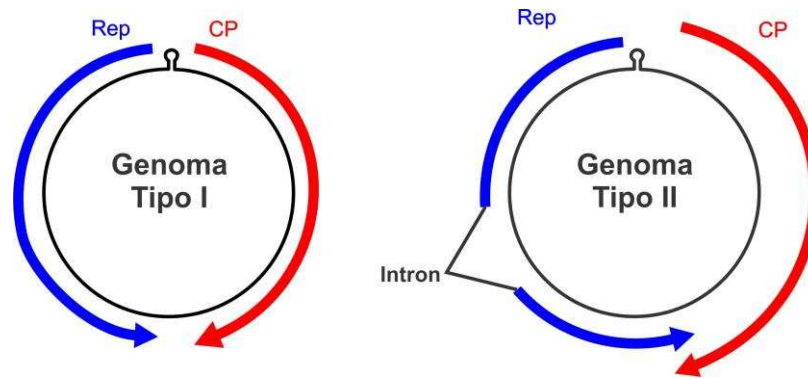


Figura 1: Estrutura genômica dos genomovírus com seus dois tipos de genomas (I e II) com as posições das ORFs REP (Proteínas associada a replicação) e CP (Proteína capsidial) em ambos tipos de genomas. O Genoma tipo 1 apresenta duas regiões intergênicas, *large* e *small*, sendo que a ORF Rep desse grupo não apresenta íntron. O genoma tipo II apresentam apenas a região intergênica *large*, devido a presença de um íntron na ORF da proteína REP gerando sobreposições da porção final das ORFs na região da região intergenicca *small*.

Estruturalmente o genoma dos genomovírus apresentam semelhanças com vários outros vírus de ssDNA circular como genimivirus e circovirus. As partículas dos genomovírus são isocaédricas com 20-22 diâmetros (Yu et al., 2010). O SsGemy-1 é o único genomovírus até o momento descrito sem a presença do íntron na Rep. Essa características assemelha os assemelha com os geminivírus do gênero *Begomovirus*. Já os demais genomovírus que apresentam um íntron no interior da ORF da proteína REP desta forma se assemelham aos geminivírus do gênero *Mastrevirus*.

A Rep é uma proteína associada a replicação por círculo rolante de plasmídeos e vírus de ssDNA circular. Esta proteína apresenta três regiões motivos conservados, denominados de motivo I, II e III. A presença de um ou dois resíduos de tirosina no motivo III separa as proteínas Rep em dois grupos, denominados de superfamílias. A superfamília família I apresenta dois resíduos de tirosina, e incluem as proteínas Rep dos bacteriófagos e os de

plasmídeos. Já a família II apresenta apenas um resíduo de tirosina que inclui os vírus de ssDNA circular de eucariotos. As Reps dos genomovírus apresentam apenas um resíduo de tirosina, ou seja, pertence da família II juntos com os outros vírus que infectam eucariotos (Rosario et al., 2012b). Um fato interessante é que a Rep dos genomovírus são relacionadas a REP dos geminivírus devido a presença do domínio GRS presente apenas em geminivírus (Heyraud-Nitschke et al., 1995; Laufs et al., 1995). Isso demonstra uma aproximação evolutiva entre esses dois grupos, que fica evidente através de análises filogenéticas (Krupovic et al., 2016). Sequência de nucleotídeos da proteína Rep revelaram que genomovírus e geminivírus são uns grupos filogeneticamente irmãos (Figura 2). A Rep dos mastrevírus e dos genomovírus são muito similares, porém a CP é distinta. Portanto, essa informação sugere que genomovírus e geminivírus possuem um ancestral em comum, sendo que os genomovírus adquiriram uma CP distinta. Além disso, por ser um vírion estruturalmente tão básico quando comparado aos geminivírus, em relação ao tamanho e número de genes, é um indicativo que possivelmente o ancestral dos genomovírus possa ter uma relação direta com o pré-mastrevírus, ancestral dos mastrevírus (Ng et al., 2014; Nawaz-ul-Rehman and Fauquet, 2009; Krupovic et al., 2009)

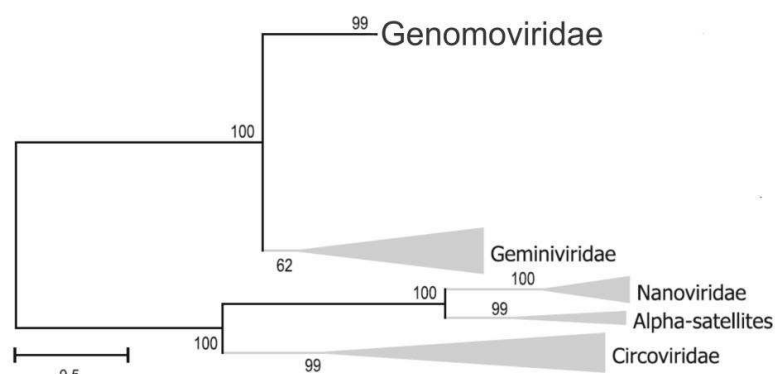


Figura 2. Filogenia dos vírus ssDNA circular (Adptado de KRUPOVIC et al., 2016)

Os genomovírus têm sido descobertos associados a diversos ambientes e organismos, como fungos (Yu et al., 2010), carnívoros de pequeno porte da América do Norte (Conceição-Neto et al., 2015; Van den Brand et al., 2012), morcegos (Male et al., 2016), crianças com encefalite (Zhou et al., 2015), sangue humano portador do HIV (Uch et al., 2015), sangue de ratos de laboratório (Li et al., 2015), plantas (Male et al., 2015; Du et al., 2014; Kraberger et al., 2015b; Dayaram et al., 2012), pássaros (Hanna et al., 2015), mosquitos (Ng et al., 2011), fezes de rena (Ng et al., 2014), esgoto (Kraberger et al., 2015a; Phan et al., 2015) e fezes humanas (Phan et al., 2015). Até o momento existem quatro relatos da presença de gemycircularvírus associados a plantas. O *Cassava associated circular DNA vírus 1* (Dayaram et al., 2012), *Hypericum japonicum associated circular DNA vírus 1* (Du et al., 2014), *Poaceae-associated gemycircularvirus 1* (Male et al., 2015) e o *Bromus associated circular DNA vírus 1* (Kraberger et al., 2015b). No entanto há pouca informação sobre a biologia dos genomovírus. Até o momento, além de estarem associados a diversos organismos e ambientes apenas o SsGemy-1 já foi demonstrado ser capaz de infectar o fungo *Sclerotinia sclerotiorum* (Yu et al., 2010). Além disso também foi demonstrado por (Ng et al., 2014) que um genomovírus foi capaz de replicar em *Nicotiana benthamiana*, porém sem apresentar sintomas. Mais recentemente, Liu et al., (2016) demonstraram que o SsGemy-1 foi capaz de infectar um inseto micofágico quando a forma larval do inseto se alimenta do fungo infectado pelo vírus.

Atualmente, existem 120 organismos virais classificados através de análises *in silico* como genomovírus, porém há pouca informação quanto a

suas propriedades biológicas (Simmonds et al., 2017; Krupovic et al., 2016). A alta distribuição desses vírus com o aumento das descobertas de genomovírus a cada dia comparado com esses três relatos ainda não fornecem informações suficientes para elucidar sobre a biologia dos genomovírus.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Van den Brand, J.M. a., Van Leeuwen, M., Schapendonk, C.M., Simon, J.H., Haagmans, B.L., Osterhaus, A.D.M.E., and Smits, S.L. (2012). Metagenomic Analysis of the Viral Flora of Pine Marten and European Badger Feces. *J. Virol.* 86: 2360–2365.
- Breitbart, M. and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13: 278–284.
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., and Rohwer, F. (2002). Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci.* 99: 14250–14255.
- Carcer, A. de D., Lopez-Bueno, A., Pearce, D.A., and Alcami, A. (2015). Biodiversity and distribution of polar freshwater DNA viruses. *Sci. Adv.* 1: e1400127–e1400127.
- Conceição-Neto, N., Zeller, M., Heylen, E., Lefrère, H., Mesquita, J.R., and Matthijnsens, J. (2015). Fecal virome analysis of three carnivores reveals a novel nodavirus and multiple gemycircularviruses. *Virol. J.* 12: 79.
- Dayaram, A., Opong, A., Jäschke, A., Hadfield, J., Baschiera, M., Dobson, R.C.J., Offei, S.K., Shepherd, D.N., Martin, D.P., and Varsani, A. (2012). Molecular characterisation of a novel cassava associated circular ssDNA virus. *Virus Res.* 166: 130–135.
- Dayaram, A., Potter, K.A., Pailles, R., Marinov, M., Rosenstein, D.D., and Varsani, A. (2015). Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. *Infect. Genet. Evol.* 30: 278–287.
- Doyle, J.J. and Doyle, J.L. (1987). A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochem. Bull.* 19: 11–15.
- Du, Z., Tang, Y., Zhang, S., She, X., Lan, G., Varsani, A., and He, Z. (2014). Identification and molecular characterization of a single-stranded circular DNA virus with similarities to *Sclerotinia sclerotiorum* hypovirulence-

- associated DNA virus 1. *Arch. Virol.* 159: 1527–1531.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Edwards, R.A. and Rohwer, F. (2005). Opinion: Viral metagenomics. *Nat. Rev. Microbiol.* 3: 504–510.
- Haible, D., Kober, S., and Jeske, H. (2006). Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J. Virol. Methods* 135: 9–16.
- Hanna, Z.R., Runckel, C., Fuchs, J., DeRisi, J.L., Mindell, D.P., Van Hemert, C., Handel, C.M., and Dumbacher, J.P. (2015). Isolation of a Complete Circular Virus Genome Sequence from an Alaskan Black-Capped Chickadee ( *Poecile atricapillus* ) Gastrointestinal Tract Sample. *Genome Announc.* 3: e01081-15.
- Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J., and Gronenborn, B. (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Res.* 23: 910–916.
- Ignacio-Espinoza, C.J., Solonenko, S.A., and Sullivan, M.B. (2013). The global virome: not as big as we thought? *Curr. Opin. Virol.* 3: 566–571.
- Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B., and Nagata, T. (2004). A simple method for cloning the complete begomovirus genome using the bacteriophage phi 29 DNA polymerase. *J. Virol. Methods* 116: 209–211.
- Johne, R., Mu, H., Rector, A., Ranst, M. Van, and Stevens, H. (2009). Rolling-circle amplification of viral DNA genomes using phi29 polymerase.: 205–211.
- Kim, M.-S., Park, E.-J., Roh, S.W., and Bae, J.-W. (2011). Diversity and Abundance of Single-Stranded DNA Viruses in Human Feces. *Appl. Environ. Microbiol.* 77: 8062–8070.
- Kraberger, S., Argüello-Astorga, G.R., Greenfield, L.G., Galilee, C., Law, D., Martin, D.P., and Varsani, A. (2015a). Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered

- from a sewage treatment oxidation pond. *Infect. Genet. Evol.* 31: 73–86.
- Kraberger, S., Farkas, K., Bernardo, P., Booker, C., Argüello-Astorga, G.R., Mesléard, F., Martin, D.P., Roumagnac, P., and Varsani, A. (2015b). Identification of novel Bromus- and Trifolium-associated circular DNA viruses. *Arch. Virol.* 160: 1303–1311.
- Krupovic, M. (2013). Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Curr. Opin. Virol.* 3: 578–586.
- Krupovic, M., Ghabrial, S.A., Jiang, D., and Varsani, A. (2016). Genomoviridae: a new family of widespread single-stranded DNA viruses. *Arch. Virol.* 161: 2633–2643.
- Krupovic, M., Ravantti, J.J., and Bamford, D.H. (2009). Geminiviruses : a tale of a plasmid becoming a virus. 11: 1–11.
- Labonté, J.M. and Suttle, C.A. (2013). Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* 7: 2169–2177.
- Laufs, J., Jupin, I., David, C., Schumacher, S., Heyraud-Nitschke, F., and Gronenborn, B. (1995). Geminivirus replication: Genetic and biochemical characterization of rep protein function, a review. *Biochimie* 77: 765–773.
- Li, W., Gu, Y., Shen, Q., Yang, S., Wang, X., Wan, Y., and Zhang, W. (2015). A novel gemycircularvirus from experimental rats. *Virus Genes* 51: 302–305.
- Liu, S., Xie, J., Cheng, J., Li, B., Chen, T., Fu, Y., Li, G., Wang, M., Jin, H., Wan, H., and Jiang, D. (2016). Fungal DNA virus infects a mycophagous insect and utilizes it as a transmission vector. *Proc. Natl. Acad. Sci.* 113: 12803–12808.
- Male, M.F., Kami, V., Kraberger, S., and Varsani, A. (2015). Genome Sequences of Poaceae-Associated Gemycircularviruses from the Pacific Ocean Island of Tonga. *Genome Announc.* 3: e01144-15.
- Male, M.F., Kraberger, S., Stainton, D., Kami, V., and Varsani, A. (2016). Cycloviruses, gemycircularviruses and other novel replication-associated protein encoding circular viruses in Pacific flying fox (*Pteropus tonganus*)

- faeces. *Infect. Genet. Evol.* 39: 279–292.
- Mokili, J.L., Rohwer, F., and Dutilh, B.E. (2012). Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* 2: 63–77.
- Muhire, B.M., Varsani, A., and Martin, D.P. (2014). SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9: e108277.
- Nawaz-ul-Rehman, M.S. and Fauquet, C.M. (2009). Evolution of geminiviruses and their satellites. *FEBS Lett.* 583: 1825–1832.
- Ng, T.F.F. et al. (2014). Preservation of viral genomes in 700-y-old caribou feces from a subarctic ice patch. *Proc. Natl. Acad. Sci.* 111: 16842–16847.
- Ng, T.F.F., Willner, D.L., Lim, Y.W., Schmieder, R., Chau, B., Nilsson, C., Anthony, S., Ruan, Y., Rohwer, F., and Breitbart, M. (2011). Broad Surveys of DNA Viral Diversity Obtained through Viral Metagenomics of Mosquitoes. *PLoS One* 6: e20579.
- Phan, T.G., Mori, D., Deng, X., Rajindrajith, S., Ranawaka, U., Fan Ng, T.F., Bucardo-Rivera, F., Orlandi, P., Ahmed, K., and Delwart, E. (2015). Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. *Virology* 482: 98–104.
- Rector, A., Tachezy, R., and Van Ranst, M. (2004). A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *J. Virol.* 78: 4993–8.
- Roossinck, M.J. and García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr. Opin. Virol.* 10: 56–62.
- Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., Breitbart, M., and Varsani, A. (2012a). Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J. Gen. Virol.* 93: 2668–2681.
- Rosario, K., Duffy, S., and Breitbart, M. (2012b). A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Arch. Virol.* 157: 1851–1871.

- Ruiz-Masó, J., Machón, C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., and Del Solar, G. (2015). Plasmid Rolling-Circle Replication. *Microbiol. Spectr.* 3: PLAS-0035-2014.
- Sierra, H., Cordova, M., Chen, C.-S.J., and Rajadhyaksha, M. (2015). Confocal Imaging–Guided Laser Ablation of Basal Cell Carcinomas: An Ex Vivo Study. *J. Invest. Dermatol.* 135: 612–615.
- Sikorski, A., Massaro, M., Kraberger, S., Young, L.M., Smalley, D., Martin, D.P., and Varsani, A. (2013). Novel myco-like DNA viruses discovered in the faecal matter of various animals. *Virus Res.* 177: 209–216.
- Simmonds, P. et al. (2017). Consensus statement: Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* 15: 161–168.
- Streit, W.R. and Schmitz, R.A. (2004). Metagenomics – the key to the uncultured microbes. *Curr. Opin. Microbiol.* 7: 492–498.
- Suttle, C.A. (2013). Viruses: unlocking the greatest biodiversity on Earth 1. *Genome* 56: 542–544.
- Suttle, C.A. (2005). Viruses in the sea. *Nature* 437: 356–361.
- Suttle, C. a (2007). Marine viruses — major players in the global ecosystem. *Nat. Rev. Microbiol.* 5: 801–812.
- Uch, R., Fournier, P.-E., Robert, C., Blanc-Tailleux, C., Galicher, V., Barre, R., Jordier, F., de Micco, P., Raoult, D., and Biagini, P. (2015). Divergent Gemycircularvirus in HIV-Positive Blood, France. *Emerg. Infect. Dis.* 21: 2096–2098.
- Varsani, A. and Krupovic, M. (2017). Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. *Virus Evol.* 3: 1–14.
- Yu, X., Li, B.B., Fu, Y., Jiang, D., Ghabrial, S.A., Li, G., Peng, Y., Xie, J., Cheng, J., Huang, J., and Yi, X. (2010). A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc. Natl. Acad. Sci.* 107: 8387–8392.

Zhou, C., Zhang, S., Gong, Q., and Hao, A. (2015). A novel gemycircularvirus in an unexplained case of child encephalitis. *Virology Journal*. 12: 197.

## Identification and Characterization of Two *Gemycircularvirus* associated with wild plants in Brazil.

Rafael Reis de Rezende<sup>1ad</sup>, Talita Bernardon Mar<sup>1bd</sup>, Lina Marcela Cortés Páez<sup>1ad</sup>, André da Silva Xavier<sup>bd</sup>, César Augusto Diniz Xavier<sup>bd</sup>, Jesus Navas Castillo<sup>c</sup>, Francisco Murilo Zerbini<sup>bd</sup>, Poliane Alfenas-Zerbini<sup>bd#</sup>

<sup>a</sup>Departamento de Microbiologia, Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), Universidade Federal de Viçosa, Viçosa, MG 36570-000, Brazil

<sup>b</sup>Departamento de Fitopatologia, Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), Universidade Federal de Viçosa, Viçosa, MG 36570-000, Brazil

<sup>c</sup>Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga-Consejo Superior de Investigaciones Científicas, 29750 Algarrobo-Costa, Málaga, Spain

<sup>d</sup>National Research Institute on Plant-Pest Interactions, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

<sup>1</sup>These authors contributed equally to this work

#Corresponding author: Poliane Alfenas-Zerbini

Phone: (+55-31) 3899-2953; Fax: (+55-31) 3899-2240; Email:

palfenas@ufv.br

Short title: Characterization of Two Gemonoviruses: *Gemycircularvirus* plant-associated

Key words: Genomovirus, *Gemycircularvirus*.

Number of words in the main text (including figure legends): 4,333

Number of figures: 2

## Identification and Characterization of Two *Gemycircularvirus* associated with wild plants in Brazil.

Rafael Reis de Rezende, Talita Bernardon Mar, Lina Marcela Cortés Páez, André da Silva Xavier, César Augusto Diniz Xavier, Jesus Navas Castillo, Francisco Murilo Zerbini, Poliane Alfenas-Zerbini

### Abstract

Genomovirus is a group of virus spread around the world in association with several organisms and environments. In this work, we identified two genomoviruses of *Gemycircularvirus* genus associated with two non-cultivated plants in Brazil, *Momordica charantia* and *Euphorbia heretofila*. Both viruses present the general structure of genomoviruses and revealed some unique features. The ability of these viruses to infect organisms in the three life domains and their spread in several diverse environments suggest that the genomoviruses are an important virus group in nature.

Since 2010, when it was first discovered, genomoviruses were described in association with different environments and organisms, such as animals, fungi and plants (Krupovic et al., 2016). *Sclerotinia sclerotiorum hypovirulence-associated DNA 1* (SsHADV-1), the first genomovirus to be characterized, causes hypovirulence in the phytopathogenic fungi *Sclerotinia sclerotiorum*. SsHADV-1 was also reported as the first ssDNA virus infecting fungi (Yu et al., 2010). Its genome consists of a circular single stranded DNA with 2100

nucleotides and encodes two ORFs corresponding to a replication associated protein (Rep) and a capsid protein (Cap) (Krupovic et al., 2016). Moreover, these viruses contain a nonanucleotide sequence within a replicative associated stem-loop (Rosario et al., 2012b). Phylogenetics analyses using sequences of Rep protein from several ssDNA circular viruses showed that the genomoviruses cluster on single clade and that these viruses are similar to mastrevirus. Besides that, it was suggested that this group is a new genus temporarily named *Gemycircularvirus* (Rosario et al., 2012a). Recently, a new family, named *Genomoviridae*, was created by the International Committee on Taxonomy of Viruses to group all gemycircularviruses (Krupovic et al., 2016). Eight genera belonging to the family were defined based on phylogenetic relationship and named *Gemycircularvirus*, *Gemykibivirus*, *Gemygorvirus*, *Gemykolovirus*, *Gemyvongvirus*, *Gemykrogvirus*, *Gemytondvirus* and *Gemykronzavirus*. For gemycircularviruses, the ICTV adopted the optimal species demarcation criteria of 78% pairwise identity as a conservative value (Varsani and Krupovic, 2017).

Two non-cultivated plants, *Momordica charantia* and *Euphorbia heretofila*, with yellow mosaic symptoms, were collected in the Southeastern and Southern Brazil, respectively. The total DNA was extracted (Doyle and Doyle, 1987) and the circular DNA was amplified by Rolling Circle Amplification (Johne et al., 2009; Inoue-Nagata et al., 2004). The digestion of RCA products with the restriction enzymes *Apal* or *HindIII* produced a fragment of about 2100 nucleotides, which suggests the presence of genomoviruses. The 2100 nucleotide fragments from both samples were cloned on pBlueScrip II SK (+) and completely sequenced by primer walking (Macrogen, Korea). The Blastn

analysis of the assembled contigs showed similarity with genomovirus, as expected. Pairwise sequence comparisons were performed using Sequence Demarcation Tool (SDT) (Muhire et al., 2014) and the results indicate that the viruses should be classified as two new genomovirus species. The virus isolated from *M. charantia* showed the highest nucleotide sequence identity (72%) with *Hypericum associated gemycircularvirus 1* (HypGemy, #KF413620) and *Pteropus associated gemycircularvirus 3* (PteGemy, # KT732797). The isolate from *E. heterophylla* showed the highest nucleotide sequence identity (67%) with *Odonata associated gemycircularvirus 1* (OdoGemy-1, #KM598385). The isolated viruses showed 66% of nucleotide sequence identity between them. Since the genomovirus species demarcation criterion is that a new sequence must have less than 78% pairwise identity to any other known genomovirus, the isolate should be classified as a new species, and we propose the names *Momordica charantia associated gemycircularvirus* (MomGemy) and *Euphorbia heterophylla associated gemycircularvirus* (EupGemy) to the viruses isolated from *M. charanti* and *E. heterophylla*, respectively.

The putative ORFs of MomGemy and EupGemy were detected by ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>). MomGemy and EupGemy have two ORFs, one in the virion-sense strand, which encodes the putative CP protein, and the other in the complementary-sense strand, which encodes a putative Rep protein. The two ORFs overlap at the end and are very similar to those observed in the genomovirus classified in the genus *Gemycircularvirus*, which demonstrates high similarity with *Cassava associated gemycircularvirus 1* (Dayaram et al., 2012), several *Dragonfly associated gemycircularviruses*

(Rosario et al., 2012a), *Odonata associated gemycircularvirus 2* (Dayaram et al., 2015) and several *Pacific flying fox faeces associated gemycircularviruses* (Male et al., 2016). Unlike SsHADV-1, which has two intergenic regions between the ORFs, only one intergenic region was observed between the two ORFs, in both viruses (Figure 1A).

Within the intergenic region, we observed the stem-loop with the nonanucleotide sequence TAATnTTAT, which is conserved in all genomoviruses; TAATGTTAT for MomGemy, and TAATATTAT for EupGemy. Moreover, an atypical structure consisting in a hairpin embedded in the major stem-loop was observed in the intergenic region of MomGemy. (Figure 1A). Aiming to ensure that this atypical structure was not an artifact, we obtained and sequenced distinct clones. The sequences of the origin of replication were identical in all of them.

In both MomGemy and EupGemy viruses, the Rep has the typical catalytic domains important for the Rolling Circle Replication (Rosario et al., 2012b). In both genomoviruses, we first identified three conserved motifs in the Rep protein: Motif I, Motif II and Motif III. Motif I recognizes and binds to the nonanucleotide within the stem-loop, and motifs II and III are involved in DNA cleavage. We also identified the NTP-binding helicase domain, formed by Walker A, Walker B and C motifs. Besides that domain, the GRS domain, which is required for the replication in geminiviruses, was located between the motifs II and III (Laufs et al., 1995; Rosario et al., 2012b) (Figure 1C). Inside the ORF of the Rep protein from both viruses, we identified an intron that may be processed for Rep expression (Figure 1A) through alignment and by manual inspection. Interestingly, introns of same size and located at the same genome

position are observed in all genomoviruses, with the exception of SsDAHV-1.

The Rep coding sequence of 61 genomovirus-like viruses was aligned using MUSCLE alignment (Edgar, 2004) and implemented in the MEGA 6 software. Then, a phylogenetic tree was constructed using Bayesian inference performed with MrBayes v3.2, implemented in the CIPRES Science Gateway, with model substitution GTR+I+G, and selected by MrModeltest v. 2.2 in the Akaike Information Criterion (AIC). The analyses were carried out running 100.000.000, sampling at every 5000 generations to produce the distribution of tree. The first 4000 trees were excluded as burn-in. Trees were visualized using FigTree program. The phylogenetic analysis indicates that MomGemy and EupGemy shared at least 56% of identify with other Genomoviruses belonging to *Gemycircularvirus* genus. Besides, both viruses cluster in a single clade, which suggests a relationship between them (Figure 2).

The identification of new genomoviruses on wild plants in Brazil reaffirm that the family *Genomoviridae* is a group of viruses spread around on world. Also, the presence of MomGemy and EupGemy in plants indicates that the *Gemycircularvirus* genus is more predominant and diverse than other genera of the *Genomoviridae* family. Further studies should be conducted for deeper understanding about the ecological relevance of genomoviruses.

## Figure legends

Figure 1. Molecular Characterization of MomGemy and EupGemy. (A) Genomic organization of MomGemy and EupGemy, indicating ORF regions and the predicted intron within Rep sequence. (B) Stem-loop structure

predicted for MomGemy and EupGemy. (C) Aminoacid sequence alignment of typical catalytic domains conserved within Rep from genomoviruses.

Figure 2. Bayesian phylogenetic tree based in full-length nucleotide sequence of sixty-one representative species of the eight genera from *Genomoviridae* family. Number at the nodes represent the number of nucleotides substitutions per site. The geminivirus representants was use as root.

## References

- Van den Brand, J.M. a., Van Leeuwen, M., Schapendonk, C.M., Simon, J.H., Haagmans, B.L., Osterhaus, A.D.M.E., and Smits, S.L. (2012). Metagenomic Analysis of the Viral Flora of Pine Marten and European Badger Feces. *J. Virol.* 86: 2360–2365.
- Breitbart, M. and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13: 278–284.
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., and Rohwer, F. (2002). Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci.* 99: 14250–14255.
- Carcer, A. de D., Lopez-Bueno, A., Pearce, D.A., and Alcamí, A. (2015). Biodiversity and distribution of polar freshwater DNA viruses. *Sci. Adv.* 1: e1400127–e1400127.
- Conceição-Neto, N., Zeller, M., Heylen, E., Lefrère, H., Mesquita, J.R., and Matthijnssens, J. (2015). Fecal virome analysis of three carnivores reveals a novel nodavirus and multiple gemycircularviruses. *Virol. J.* 12: 79.
- Dayaram, A., Opong, A., Jäschke, A., Hadfield, J., Baschiera, M., Dobson, R.C.J., Offei, S.K., Shepherd, D.N., Martin, D.P., and Varsani, A. (2012). Molecular characterisation of a novel cassava associated circular ssDNA

- virus. *Virus Res.* 166: 130–135.
- Dayaram, A., Potter, K.A., Pailles, R., Marinov, M., Rosenstein, D.D., and Varsani, A. (2015). Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. *Infect. Genet. Evol.* 30: 278–287.
- Doyle, J.J. and Doyle, J.L. (1987). A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochem. Bull.* 19: 11–15.
- Du, Z., Tang, Y., Zhang, S., She, X., Lan, G., Varsani, A., and He, Z. (2014). Identification and molecular characterization of a single-stranded circular DNA virus with similarities to *Sclerotinia sclerotiorum* hypovirulence-associated DNA virus 1. *Arch. Virol.* 159: 1527–1531.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Edwards, R.A. and Rohwer, F. (2005). Opinion: Viral metagenomics. *Nat. Rev. Microbiol.* 3: 504–510.
- Haible, D., Kober, S., and Jeske, H. (2006). Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J. Virol. Methods* 135: 9–16.
- Hanna, Z.R., Runckel, C., Fuchs, J., DeRisi, J.L., Mindell, D.P., Van Hemert, C., Handel, C.M., and Dumbacher, J.P. (2015). Isolation of a Complete Circular Virus Genome Sequence from an Alaskan Black-Capped Chickadee ( *Poecile atricapillus* ) Gastrointestinal Tract Sample. *Genome Announc.* 3: e01081-15.
- Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J., and Gronenborn, B. (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Res.* 23: 910–916.
- Ignacio-Espinoza, C.J., Solonenko, S.A., and Sullivan, M.B. (2013). The global virome: not as big as we thought? *Curr. Opin. Virol.* 3: 566–571.
- Inoue-Nagata, A.K., Albuquerque, L.C., Rocha, W.B., and Nagata, T. (2004). A simple method for cloning the complete begomovirus genome using the

- bacteriophage phi 29 DNA polymerase. *J. Virol. Methods* 116: 209–211.
- Johne, R., Mu, H., Rector, A., Ranst, M. Van, and Stevens, H. (2009). Rolling-circle amplification of viral DNA genomes using phi29 polymerase.: 205–211.
- Kim, M.-S., Park, E.-J., Roh, S.W., and Bae, J.-W. (2011). Diversity and Abundance of Single-Stranded DNA Viruses in Human Feces. *Appl. Environ. Microbiol.* 77: 8062–8070.
- Kraberger, S., Argüello-Astorga, G.R., Greenfield, L.G., Galilee, C., Law, D., Martin, D.P., and Varsani, A. (2015a). Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. *Infect. Genet. Evol.* 31: 73–86.
- Kraberger, S., Farkas, K., Bernardo, P., Booker, C., Argüello-Astorga, G.R., Mesléard, F., Martin, D.P., Roumagnac, P., and Varsani, A. (2015b). Identification of novel Bromus- and Trifolium-associated circular DNA viruses. *Arch. Virol.* 160: 1303–1311.
- Krupovic, M. (2013). Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Curr. Opin. Virol.* 3: 578–586.
- Krupovic, M., Ghabrial, S.A., Jiang, D., and Varsani, A. (2016). Genomoviridae: a new family of widespread single-stranded DNA viruses. *Arch. Virol.* 161: 2633–2643.
- Krupovic, M., Ravantti, J.J., and Bamford, D.H. (2009). Geminiviruses : a tale of a plasmid becoming a virus. 11: 1–11.
- Labonté, J.M. and Suttle, C.A. (2013). Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* 7: 2169–2177.
- Laufs, J., Jupin, I., David, C., Schumacher, S., Heyraud-Nitschke, F., and Gronenborn, B. (1995). Geminivirus replication: Genetic and biochemical characterization of rep protein function, a review. *Biochimie* 77: 765–773.
- Li, W., Gu, Y., Shen, Q., Yang, S., Wang, X., Wan, Y., and Zhang, W. (2015). A novel gemycircularvirus from experimental rats. *Virus Genes* 51: 302–305.

- Liu, S., Xie, J., Cheng, J., Li, B., Chen, T., Fu, Y., Li, G., Wang, M., Jin, H., Wan, H., and Jiang, D. (2016). Fungal DNA virus infects a mycophagous insect and utilizes it as a transmission vector. *Proc. Natl. Acad. Sci.* 113: 12803–12808.
- Male, M.F., Kami, V., Kraberger, S., and Varsani, A. (2015). Genome Sequences of Poaceae-Associated Gemycircularviruses from the Pacific Ocean Island of Tonga. *Genome Announc.* 3: e01144-15.
- Male, M.F., Kraberger, S., Stainton, D., Kami, V., and Varsani, A. (2016). Cycloviruses, gemycircularviruses and other novel replication-associated protein encoding circular viruses in Pacific flying fox (*Pteropus tonganus*) faeces. *Infect. Genet. Evol.* 39: 279–292.
- Mokili, J.L., Rohwer, F., and Dutilh, B.E. (2012). Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* 2: 63–77.
- Muhire, B.M., Varsani, A., and Martin, D.P. (2014). SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9: e108277.
- Nawaz-ul-Rehman, M.S. and Fauquet, C.M. (2009). Evolution of geminiviruses and their satellites. *FEBS Lett.* 583: 1825–1832.
- Ng, T.F.F. et al. (2014). Preservation of viral genomes in 700-y-old caribou feces from a subarctic ice patch. *Proc. Natl. Acad. Sci.* 111: 16842–16847.
- Ng, T.F.F., Willner, D.L., Lim, Y.W., Schmieder, R., Chau, B., Nilsson, C., Anthony, S., Ruan, Y., Rohwer, F., and Breitbart, M. (2011). Broad Surveys of DNA Viral Diversity Obtained through Viral Metagenomics of Mosquitoes. *PLoS One* 6: e20579.
- Phan, T.G., Mori, D., Deng, X., Rajindrajith, S., Ranawaka, U., Fan Ng, T.F., Bucardo-Rivera, F., Orlandi, P., Ahmed, K., and Delwart, E. (2015). Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. *Virology* 482: 98–104.

- Rector, A., Tachezy, R., and Van Ranst, M. (2004). A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *J. Virol.* 78: 4993–8.
- Roossinck, M.J. and García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr. Opin. Virol.* 10: 56–62.
- Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., Breitbart, M., and Varsani, A. (2012a). Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Eiprocta). *J. Gen. Virol.* 93: 2668–2681.
- Rosario, K., Duffy, S., and Breitbart, M. (2012b). A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Arch. Virol.* 157: 1851–1871.
- Ruiz-Masó, J., Machón, C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., and Del Solar, G. (2015). Plasmid Rolling-Circle Replication. *Microbiol. Spectr.* 3: PLAS-0035-2014.
- Sierra, H., Cordova, M., Chen, C.-S.J., and Rajadhyaksha, M. (2015). Confocal Imaging–Guided Laser Ablation of Basal Cell Carcinomas: An Ex Vivo Study. *J. Invest. Dermatol.* 135: 612–615.
- Sikorski, A., Massaro, M., Kraberger, S., Young, L.M., Smalley, D., Martin, D.P., and Varsani, A. (2013). Novel myco-like DNA viruses discovered in the faecal matter of various animals. *Virus Res.* 177: 209–216.
- Simmonds, P. et al. (2017). Consensus statement: Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* 15: 161–168.
- Streit, W.R. and Schmitz, R.A. (2004). Metagenomics – the key to the uncultured microbes. *Curr. Opin. Microbiol.* 7: 492–498.
- Suttle, C.A. (2013). Viruses: unlocking the greatest biodiversity on Earth 1. *Genome* 56: 542–544.
- Suttle, C.A. (2005). Viruses in the sea. *Nature* 437: 356–361.
- Suttle, C. a (2007). Marine viruses — major players in the global ecosystem.

Nat. Rev. Microbiol. 5: 801–812.

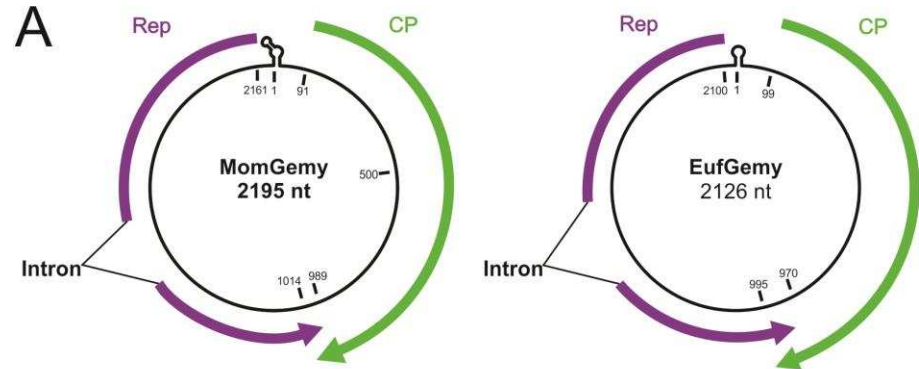
Uch, R., Fournier, P.-E., Robert, C., Blanc-Tailleux, C., Galicher, V., Barre, R., Jordier, F., de Micco, P., Raoult, D., and Biagini, P. (2015). Divergent Gemycircularvirus in HIV-Positive Blood, France. *Emerg. Infect. Dis.* 21: 2096–2098.

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

Yu, X., Li, B.B., Fu, Y., Jiang, D., Ghabrial, S.A., Li, G., Peng, Y., Xie, J., Cheng, J., Huang, J., and Yi, X. (2010). A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc. Natl. Acad. Sci.* 107: 8387–8392.

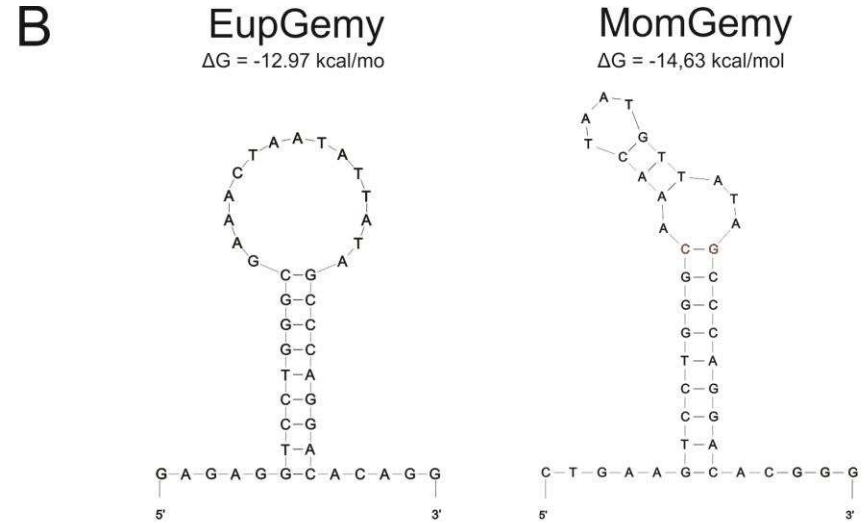
Zhou, C., Zhang, S., Gong, Q., and Hao, A. (2015). A novel gemycircularvirus in an unexplained case of child encephalitis. *Virol. J.* 12: 197.

**Figure 1**



*Euphorbia gemycircularvirus* - EufGemy  
 GTTGACGGCAGAGATGACTGGGTACGAGAATCTGGTCTTGGCGTGGGACAATCATTAGTGGGTTATCTTGC GCGTCGCACTAGAACGCTCGG  
 GGGGCCCCCCCGCGAAGGGGGCACCTCCCCCCTCCCTCGCTTCTTTGTGCTGGGCGCGCTGGTTAATAGATATACTGGGGTCTTTTGGT  
 AACTCAGGTTTATCCAGAG

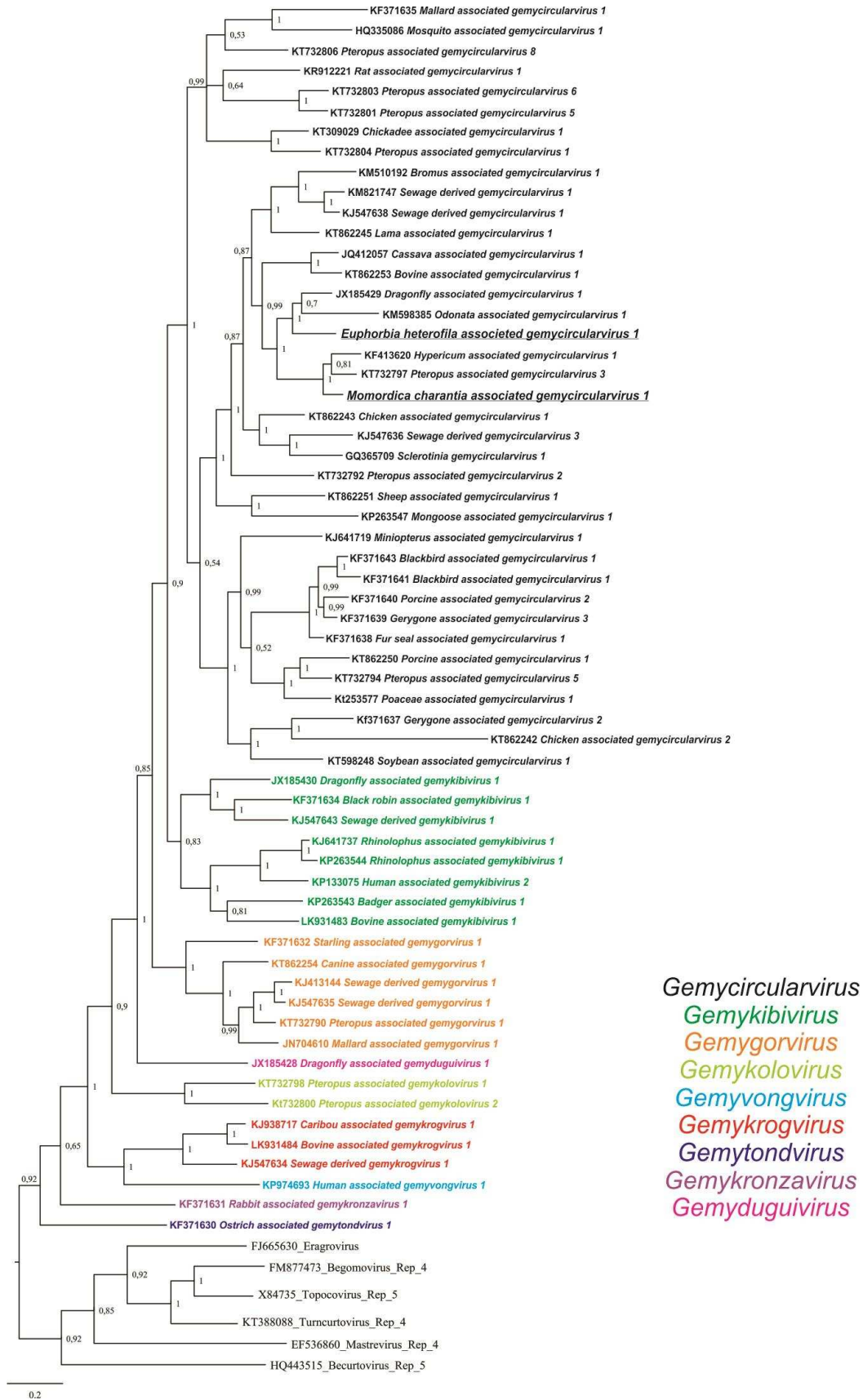
*Mormodica charantia gemycircularis* - MomGemy  
 GTAAGTTTTCCGATATCAAGAGCGGAAACAAAAAGGGTAAACCCGCTCGCCGCTGGCAACGGTCAATTGGCTGTTGCGGCTCAATACCGT  
 ACCACCCAGAGGGGACCCCGGTACTCCTTCGCCTTGTTTGGGTATGTGACTCATGCTGACTAAAG



	GRS						
	Motif I	Motif II	domain	Motif III	Walker A	Walker B	Motif C
MomGemy	GLITYPQDGD	GGHLHFVFA	GRHPNVEP	YDYAIKDG	LWGKSRITGKTTW	YAVFDDIRG	SIWVCSNDP
EufGemy	FLITYAQDGD	GGHLHFVFA	GRHPNIES	YNYAIKDG	LYGESRTGKTLW	YAVFDDIRG	SIWVSNDDP
SsHADV-1	VLLITYAQDGE	GGHLHFVFA	GRHPNIEK	YDYAIKDG	LYGPSQTGKTSW	YAVFDDIRG	SIWCSNADP
CasCV	ALITYAQDGA	GGVHLHFVFI	GRHPNIEP	YDYAIKDG	LYGDSRSRGKTLW	YAVFDDIRG	SIWISNDDP
BasCV-3	ALITYAQDGE	GGVHLHFVFI	GRHPNIEP	YDYAIKDG	VYGESRTGKTLW	YAVFDDIRG	SIWLSNDDP
PaGmV-1	VLLITYAQDGD	GGHLHFVFA	GRHPNIEP	YDYAIKDG	YFGASRLGKTLW	YAVFDDMRG	CIWLSNADP
SlaGmV-1	FLITYSQDAG	RGTHLHFVFA	GRHPNIEP	YDYATKHG	IFGPSQTGKTSW	YVVLDDMEG	CIWLSNDDP
HJassCV	VLLITYAQDGD	GGHLHFVFA	GRHPNIEP	YDYAIKDG	LYGDSRSRGKTLW	YAVFDDIRG	SIWLSNDDP
PfffaGmV-1	VLLITYPQDGN	GGHLHFVFA	GRHPNIEP	YDYAIKDG	LYGDSRSRGKTLW	YAVFDDIRG	TIWVCSNDP
DfasC-2	FLITYPQDGN	GGHLHFVFA	GCHPNIEP	YDYAIKDG	LYGESRTGKTLW	YAVFDDIRG	SIWISNDDP
OdaGmV-1	FLITYAQDGD	GGHLHFVFA	GRHPNIEP	YDYAIKDG	LEGGNGSGQGPY	VCFVDDIRG	SIWISNDDP
FaGmV-16	VLLITYAQDGD	GGVHLHFVFI	GRHPNIEP	YDYAIKDG	LYGPSRTGKTSW	YAVFDDMRG	SIWCSNADP
PfffaGmV-11	ALITYSQDGN	GGHLHFVFA	GCHPNIEP	YDYAIKDG	LYGDSRSRGKTLW	YAVFDDIQG	SIWLSNDDP
PfffaGmV-14	ALITYSQDGD	GGHLHFVFA	GRHPNIEP	YDYAIKDG	LYGDSRSRGKTLW	YAVFDDIQG	SIWLSNDDP
FaGmV-4	VLLITYAQDGD	GGHLHFVFA	GRHPNIEP	YDYAIKDG	IYGDTRIGKTLW	YAVFDDMRG	VIWLSNDDP
SaGmV-7a	VLLITYAQDGD	GGVHLHFVFI	GCHPNIEP	YDYAIKDG	VYGESRTGKTLW	YAVFDDIRG	SIWLSNDDP

	Stem 5'	Origin	Stem 3'
MomGemy	TCTGAAGTCCCTGGGC	--AAACTAATGTTAT	----AGCCC--AGGACACGGGCACAG
EufGemy	CGAGAGGTCTGGGC	-GAACTAATATTAT	----AGCCC--AGGACACAGGACACA
SsHADV-1	TTGAGTCTGGGCT	--AAACTAATATTAT	----AGCC--CAGGACGGGGACT
CasCV	CAAGTGTCTGCGGTGA	AAACTAATATTAT	----ACGG--CAGGACACAGGGCAGC
BasCV-3	AAAAAGTCCCTGGGC	--AAACTAATGTTAT	----AGCCC--AGGACGGGGACACC
PaGmV-1	CGTAGGAACCTGGGC	--AAACTAATGTTAT	----AGCCC--AGGTGACGGGTGACA
HJassCV	TAATGGGTCTGGGGG	--AAAGTAATGTTAT	----CCC--CCAGACACAGGACACG
PfffaGmV-5	TCTGTGTCTGGGC	-AAACTAATATTAT	----AGCCC--AGGACACAGGGCACA
DfasCV-2	TGAGAAGTCCCTGGGC	--AAACTAATATTAT	----AGCCC--AGGACACAGGACACA
OdaGmV-12	CCGAAAGTCCCTGGGC	-GAACTAATATTAT	----AGCC--AGGTACAGGGCACA
FaGmV-16	GAAGTCGTGTTCTGC	--AAACTAATATTAT	----AGCAG--AACAGACGGACACAG
PfffaGmV-11	AAGTGGACTGGGGG	-GAACTAATATTAT	----AGCC--GCAGGACAGGGACGG
PfffaGmV-14	ACTGGACTGCGGGCA	--AAACTAATATTAT	----GCC--CGCAGGACAGGGGACT
SaGmV-7a	AAAAAGTCCCTGGGC	-GAACTAATGTTAT	----AGCCC--AGGGACAGGGGACA
FaGmV-4	TGAAAGTGTCTGGC	--AAACTAATGTTAT	----AGCCA--GAACACGGACTCACG

Figure 2



## CONCLUSÕES

- Os MorGemy e EupGemy são novos genomovírus pertencentes ao gênero *Gemycircularvirus* isolados de duas plantas silvestres no Brasil.
- A análise de SDT mostrou que MorGemy tem 72% de identidade com *Pteropus associated gemycircularvirus 3*, *Hypericum associated gemycircularvirus 1* e que ambos compartilham 66% de identidade.
- Ambos vírus se agruparam em um mesmo clado, o que sugere uma relação evolutiva entre os dois isolados.
- Ambos os vírus apresentam nonanucleotídeo conservado em todos os genomovírus, sendo sequência TAATRTTAT.
- MorGemy e EupGemy compartilham todas as características encontradas nos genomovírus. No entanto, em MorGemy é descrito pela primeira vez a presença de uma estrutura de grampo no *stem-loop*.
- Foi observado um íntron em ambos os genomas, no interior da ORF da proteína Rep. O íntron de EupGemy é maior do que o íntron de MorGemy.
- Na proteína Rep foram observados os motivos I, II e III, além dos domínios Walker A, B, C e GRS, encontrados na proteína dos geminivírus.
- O agrupamento de MorGemy e EupGemy com outros genomovírus associados com plantas em um mesmo gênero são um indicativo de que esses vírus podem ter uma possível relação de hospedeiro-vírus com plantas.
- A presença de MorGemy e EupGemy em plantas silvestres é um indicativo de que de gemycircularvírus é o gênero mais predominante e diversos da família *Genomoviridae*.