

UNIVERSIDADE FEDERAL DE ALAGOAS

CENTRO DE CIÊNCIAS AGRÁRIAS

PROGRAMA DE PÓS-GRADUAÇÃO EM PRODUÇÃO VEGETAL

Caio Henrique Loureiro de Hollanda Ferreira

Diversification into the genus *Badnavirus*: phylogeny and population genetic variability

Rio Largo

2018

CAIO HENRIQUE LOUREIRO DE HOLLANDA FERREIRA

Diversification into the genus *Badnavirus*: phylogeny and population genetic variability

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Produção Vegetal da Universidade Federal de Alagoas como requisito parcial para obtenção do grau de Mestre em Produção Vegetal.

Orientadora: Profa. Dra. Sarah Jacqueline Cavalcanti da Silva

Co-Orientador: Prof. Dr. Roberto Ramos Sobrinho

Rio Largo

2018

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Setorial do Centro de Ciências Agrárias
Bibliotecária Responsável: Myrtes Vieira do Nascimento

F383d Ferreira, Caio Henrique Loureiro de Hollanda
Diversification into the genus Badnavirus: phylogeny and
population genetic variability / Caio Henrique Loureiro de Hollanda
Ferreira – 2018.

71 f.; il.

Dissertação (Mestrado em Produção Vegetal) - Universidade
Federal de Alagoas, Centro de Ciências Agrárias. Rio Largo, 2018.
Orientação: Profa. Dra. Sarah Jacqueline Cavalcanti da Silva

Inclui bibliografia

1. Evolução 2. Network 3. Especiação I. Título

CDU: 632.38

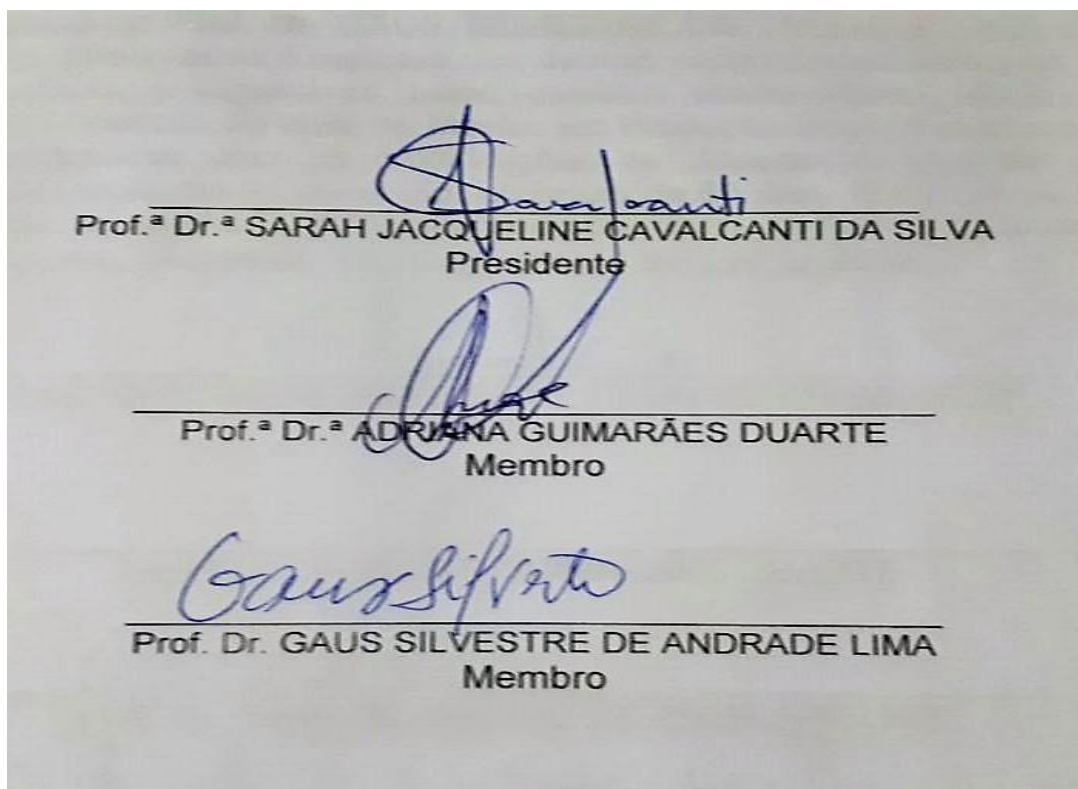
Folha de Aprovação

AUTOR: CAIO HENRIQUE LOUREIRO DE HOLLANDA FERREIRA

Diversification into the genus *Badnavirus*: phylogeny and population genetic variability

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Produção Vegetal da Universidade Federal de Alagoas e aprovada em 08 de fevereiro de 2018.

Banca Examinadora:



Ao meu avô Adelson Loureiro (*in memoriam*).

Agradecimentos

Sou grato primeiramente a Deus por ter me conduzido com força e saúde durante essa importante jornada.

Aos meus pais Alberto Hollanda e Adeane Loureiro que sempre me apoiaram e me ensinaram a lutar por meus objetivos de forma íntegra, justa e honesta. Hoje sou o que sou graças a eles.

A minha querida irmã Amanda Hollanda, por ser uma pessoa que me inspira, meu maior exemplo de superação e força de vontade.

Aos meus avós Alberto, Rosário, Amélia e Adelson (*in memoriam*) por cada palavra de sabedoria, cada gesto de carinho, torcida e cuidado.

A minha companheira Clarissa Karine que esteve sempre ao meu lado em todos os momentos, me incentivando, me entendendo e me apoiando incondicionalmente.

Aos professores Gaus Silvestre e Iraildes Assunção pelo acolhimento no laboratório de fitopatologia, por todo aprendizado durante esse tempo e pela confiança depositada em mim. Sou imensamente grato a vocês.

Aos meus orientadores Sarah Cavalcanti e Roberto Ramos por todo ensinamento e paciência. Foi um prazer ser orientado por vocês.

Aos pós-docs Fred Feijó e Mariote Netto por todo auxílio prestado e pelo acompanhamento diário nas atividades do laboratório.

Aos grandes amigos que fiz no laboratório, em especial Luan, Mayra, Tamires e Aninha. Todos vocês foram muito importantes nessa conquista.

Aos professores e colaboradores da Universidade Federal de Alagoas, especialmente os professores Renan, Vilma, Sônia e Marissônia (Embrapa) por toda dedicação e trabalho nessa bela profissão.

Deixo também meus agradecimentos a CAPES, CNPq e FAPEAL pelo incentivo educacional neste país.

Por fim, a todos que contribuíram direta ou indiretamente nessa conquista, meus sinceros agradecimentos.

RESUMO

A família *Caulimoviridae* compreende fitovírus com genomas semicirculares de DNA de fita dupla encapsulados em partículas isométricas ou baciliformes, sendo dividida em oito gêneros. O gênero *Badnavirus* é o mais importante devido ao seu alto número de espécies relatadas infectando plantas cultivadas em todo o mundo. O presente estudo teve como objetivo avaliar a variabilidade genética de populações e o posicionamento taxonômico/filogenético de espécies de badnavírus. Inicialmente, sequências completas para o genoma das espécies de badnavírus foram obtidas a partir do banco de dados não-redundante *GenBank*. Alinhamentos múltiplos das sequências nucleotídicas foram obtidos para os seguintes conjuntos de dados: genoma completo, ORFIII, RT/RNaseH completa (1020nt) e parcial (579nt). Comparações pareadas de sequências, análises filogenéticas, de recombinação e variabilidade genética foram realizadas para os conjuntos de dados. Um total de 127 isolados foram obtidos, representando 53 espécies de badnavírus. As comparações de sequências de nucleotídeos para o conjunto de dados RT/RnaseH (completa e parcial) mostraram que algumas espécies de badnavírus relatadas como distintas apresentam entre 57,1-82,5% de identidade, superior ao limite de 80,0% atualmente utilizado para a demarcação de espécies em *Badnavirus*. Resultados similares foram observados para os dados da ORFIII e genoma completo, reforçando que sequências parciais do domínio RT/RNaseH são suficiente para determinar o posicionamento taxonômico da maioria das espécies descritas neste gênero. Quatro grupos (*clusters 1-4*) bem suportados foram observados nas árvores filogenéticas para genoma completo e ORFIII, com os *cluster 1* e *3* formando grupos irmãos compreendendo espécies/isolados infectando predominantemente cana-de-açúcar (*Saccharum* spp.) e banana (*Musa* spp.). Análise de evolução em rede evidenciou alguns possíveis eventos de recombinação afetando a diversificação de espécies de badnavírus, com pelo menos 23 eventos independentes sendo detectados com pontos de recombinação ocorrendo predominantemente na ORFIII e região intergênica. Finalmente, altos índices de diversidade nucleotídica foram observados para a região RT/RnaseH parcial em populações de 12 espécies distintas de badnavírus. Estes resultados mostram que mutação e recombinação são mecanismos importantes atuando na evolução dos badnavírus.

Palavras-chave: Evolução; Network; Especiação.

ABSTRACT

The family *Caulimoviridae* comprises viruses with semicircular double-stranded DNA genomes encapsulated into isometric or bacilliform particles, being divided into eight genera. The genus *Badnavirus* is the most important due to its high number of species reported infecting cultivated plants worldwide. The present study aimed to evaluate the taxonomic/phylogenetic positioning and population genetic variability into the genus *Badnavirus*. Initially, a data set comprising badnavirus full-length genome sequences was obtained from the non-redundant GenBank database. Multiple nucleotide sequence alignments were obtained for the data sets: complete genome; ORF III; full (1020pb) and partial (579pb) RT/RNaseH. Pairwise sequence comparisons, phylogenetic and recombination analyses were performed for all data sets. A total of 127 isolates were obtained, representing 53 badnavirus species. Nucleotide pairwise comparisons for the data sets RT/RNaseH and ORF III showed that a number of distinct badnavirus species shared up to 82,5% of identity, higher than the 80% threshold currently used for species demarcation in the genus. Bayesian phylogenetic trees showed four well supported clusters, with clusters 1 and 3 being sister groups comprising predominantly isolates infecting sugarcane and banana. Non-tree-like evolution evidenced a complex pattern of recombination, and at least 23 independent events were detected with recombination breakpoints occurring predominantly in the ORF III and in the intergenic region. By the analysis of nucleotide diversity of the partial RT/RNaseH region in 12 badnavirus population, a high genetic variability was observed. These results showed that mutation and recombination are important evolutionary mechanisms acting on the diversification of badnaviruses, and that the partial RT/RNaseH sequence is sufficient to determine the taxonomic placement of most viral species described in this genus.

Keywords: Evolution; Network; Speciation.

Lista de Ilustrações

Figura 1	Representação esquemática do genoma de um Badnavirus típico.....	17
Figure 1	Midpoint-rooted Bayesian phylogenetic trees based on complete genome nucleotide sequences of badnaviruses.....	56
Figure 2	Midpoint-rooted Bayesian phylogenetic trees based on ORFIII nucleotide sequences of badnaviruses.....	57
Figure 3	Midpoint-rooted Bayesian phylogenetic trees based on complete RT/RNaseH nucleotide sequences of badnaviruses.....	58
Figure 4	Midpoint-rooted Bayesian phylogenetic trees based on partial RT/RNaseH nucleotide sequences of badnaviruses.....	59
Figure 5	Non-tree-like evolution based on the complete genome nucleotide sequences of badnaviruses.....	60
Figure 6	Non-tree-like evolution based on the ORFIII nucleotide sequences of badnaviruses.....	61
Figure 7	Non-tree-like evolution based on full RT/RNaseH nucleotide sequences of badnaviruses.	62
Figure 8	Non-tree-like evolution based on partial RT/RNaseH nucleotide sequences of badnaviruses.....	63

SUMÁRIO

1. INTRODUÇÃO	11
2. REVISÃO DE LITERATURA	13
2.1. Família <i>Caulimoviridae</i>	13
2.2. Gênero <i>Badnavirus</i>	15
2.3. Sequências endógenas	19
2.4. Culturas de importância econômica infectadas por badnavírus	20
2.5. Variabilidade e estrutura genética de populações de fitovírus	25
REFERÊNCIAS BIBLIOGRÁFICAS	28
CAPÍTULO I	38
Abstract	39
Introduction	40
Materials and Methods	42
<i>Badnavirus</i> data set	42
Sequence analysis	42
Phylogenetic inference	43
Recombination analysis	43
Population genetic variability	44
Results	44
<i>Badnavirus</i> isolates	44
Species demarcation	44
Phylogenetics analysis	45
Recombination events	46
Population genetic variability	47
Discussion	47
References	52
ANEXO I	70

1. INTRODUÇÃO

Os vírus pertencentes à família *Caulimoviridae* têm genomas de DNA fita dupla (dsDNA), semicirculares, com comprimento de 7,2-9,2 kbp, encapsulados em partículas isométricas ou baciliformes, que replicam via de um intermediário de RNA (pararetrovírus) (GEERING; HULL, 2012). A família é dividida em oito gêneros (*Badnavirus*, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Rosadnavirus*, *Solendovirus*, *Soymovirus* e *Tungrovirus*), de acordo com a gama de hospedeiro, inseto vetor, organização genômica e filogenia (GEERING; HULL, 2012; BHAT; HOHN; SELVARAJAN, 2016).

O gênero *Badnavirus* engloba vírus que possuem genoma de dsDNA semicircular de 7.2-9,2 kpb, encapsulados em partículas baciliformes não envelopadas (GEERING; HULL, 2012). O genoma viral apresenta descontinuidades sítio específicas em cada fita e abriga três ORFs, nomeadas I, II e III (BOUHIDA; LOCKHART; OLSZEWSKI, 1993; HAGEN et al. 1993; HARPER; HULL, 1998; GEERING; HULL, 2012; BHAT et al. 2016). As ORFs I e II codificam respectivamente proteínas associadas ao vírion (CHENG; LOCKHART; OLSZEWSKI, 1996) e de ligação de ácido nucleico (JACQUOT et al., 1996). A ORF III codifica para uma poliproteína de 208-216 kD, que é clivada proteoliticamente, gerando as proteínas de movimento (MP), da capa proteica (CP), a aspartate protease (AP- responsável pela clivagem da poliproteína), a transcriptase reversa (RT) e a ribonuclease H (RNaseH) (domínios genômicos que também estão envolvidos na replicação viral) (HARPER; HULL, 1998; MEDBERRY; LOCKHART; OLSZEWSKI, 1990).

Badnavírus são frequentemente transmitidos por cochonilha para diferentes espécies de culturas economicamente importantes: cana-de-açúcar (*Saccharum* spp.), banana (*Musa* spp.), cacau (*Theobroma cacao*), pimenta-peta (*Piper* spp.), inhame (*Dioscorea* spp.) e citros (*Citrus* spp.) (ENI et al., 2008; JAMES et al., 2011; JOHNSON et al., 2012; KOUAKOU et al., 2012; DEESHMA; BHAT, 2015; SILVA et al., 2015). Adicionalmente, estes podem existir como sequências endógenas pararetrovirais (EPRVs) (HOHN et al., 2008; STAGINNUS et al., 2009). Acredita-se que EPRVs integram-se por recombinação ilegítima no genoma da hospedeira, e sua presença não está necessariamente associada com a infecção (NDOWORA et al. 1999, GEERING et al., 2005). No entanto, alguns pararetrovírus são

capazes de "escapar" do genoma do hospedeiro e causar uma infecção sistêmica (DALLOT et al., 2001; LHEUREUX et al., 2003; CÔTE et al., 2010).

Atualmente 40 espécies de badnavírus são reconhecidas pelo ICTV (International Committee on Taxonomy of Viruses). Este órgão estabelece um limite padrão para demarcação de espécies de badnavírus de $\geq 80\%$ de identidade de nucleotídeos na região RT/RNase H (GEERING; HULL 2012). Contudo, identificar espécies de badnavírus tem se tornado um desafio. Primeiro, porque a principal forma de detecção e identificação destes vírus tem como base a amplificação da região RT-RNase H por PCR, técnica a qual não permite a distinção entre formas episomais e integradas do vírus. Segundo diferentes estudos demonstram que o critério de demarcação de espécies utilizado pelo ICTV é insuficiente para separar algumas espécies de badnavírus, especialmente aquelas que infectam banana e cana-de-açúcar (MULLER et al. 2011; KARUPPAIAH; VISWANATHAN; KUMAR, 2013; SILVA et al., 2015, JORDÃO, 2015). Por fim, além da grande diversidade de espécies, as populações de badnavírus também apresentam altos níveis de variabilidade molecular (GEERING et al., 2000; GEIJSKES et al., 2002; HARPER et al., 2002; YANG et al., 2003; MULLER, et al., 2005, ENI et al., 2008; KENYON, et al., 2008; MULLER, et al., 2011; GOVIND et al., 2014; GUIMARÃES et al., 2015; SILVA et al., 2015), gerada em função da ocorrência de frequentes eventos de recombinação e mutação (GOVIND et al., 2014; SILVA et al., 2015, JORDÃO, 2015).

A identificação taxonômica correta e determinação da estrutura (ou variabilidade) genética de populações impactam diretamente na durabilidade das medidas no controle de fitoviroses. Dessa forma, o objetivo desse trabalho foi avaliar a variabilidade genética de populações e o posicionamento taxonômico/filogenético de espécies de badnavírus.

2. REVISÃO DE LITERATURA

2.1. Família *Caulimoviridae*

A família *Caulimoviridae* compreende vírus de plantas que possuem genoma constituído por uma molécula de DNA semicircular de fita dupla (dsDNA), com tamanho entre 7,2 e 9,2 Kb, encapsulada em partículas não envelopadas de morfologia isométrica ou baciliforme, a depender do gênero (GEERING; HULL, 2012). Também são caracterizados pela presença da região intergênica poli (A) e descontinuidades de fita simples ou gaps em sítios específicos de ambas as fitas (HAPER et al., 2002).

Para divisão dos gêneros são adotados critérios que estão relacionados com o inseto vetor, gama de hospedeiro, organização genômica e relacionamento filogenético. De acordo com o Comitê Internacional de Taxonomia de Vírus (ICTV, 2016) a família *Caulimoviridae* é constituída por oito gêneros: *Caulimovirus* (11 espécies), *Soymovirus* (quatro espécies), *Solendovirus* (duas espécies), *Cavemovirus* (duas espécies), *Petuvirus* (uma espécie), *Rosadnavirus* (uma espécie), *Badnavirus* (40 espécies) (Anexo I) e *Tungrovirus* (uma espécie) (<https://talk.ictvonline.org/taxonomy/>). Membros dessa família são encontrados amplamente distribuídos geograficamente, sendo que a maioria das espécies dos gêneros *Tungrovirus* e *Badnavirus* se localizam nas regiões tropicais e subtropicais, os demais membros, são encontrados em regiões temperadas (GEERING; HULL, 2012).

Os hospedeiros naturais de espécies da família *Caulimoviridae* são angiospermas das classes Dicotyledonae e Monocotyledonae, a depender do gênero. Membros dos gêneros *Petuvirus*, *Soymovirus* e *Cavemovirus* apresentam infecção restrita apenas a plantas dicotiledôneas; *Tungrovirus* estão restritos a plantas monocotiledôneas, enquanto que os *Badnavirus*, por exemplo, infectam tanto plantas monocotiledôneas como dicotiledôneas (GEERING; HULL, 2012).

A transmissão natural do vírus pode ocorrer via inseto vetor (insetos das famílias Aleyrodidae, Aphididae, Cicadellidae e Pseudococcidae: Hemiptera) ou por contato entre plantas hospedeiras, bem como por sementes ou grão de pólen, dependendo do gênero. A transmissão também pode ser realizada via inoculação mecânica e enxertia (FAUQUET et al., 2005). Além disso, muitas espécies de plantas que são infectadas por membros da família

Caulimoviridae são propagadas vegetativamente o que contribui para o processo de disseminação desses vírus.

Os sintomas característicos na planta associados aos *Caulimoviridae* podem variar de acordo com o hospedeiro, a espécie viral envolvida, e condições climáticas, sendo os mais comumente observados: clareamento de nervuras, estrias, mosaico, manchas necróticas, linhas e manchas cloróticas, nanismo, enrolamento e má formação foliar e enrolamento de folíolos. (GEERING; HULL, 2012).

Os genomas contêm entre um e oito ORFs (*open reading frames*), dependendo do gênero. As proteínas codificadas comuns a todos os gêneros são: proteína de movimento (MP), proteína do capsídeo (CP), aspartato protease (AP), transcriptase reversa (TR) e ribonuclease H (GEERING; HULL, 2012).

Os membros da família *Caulimoviridae* são classificados como Pararetrovírus, termo sugerido por Temin (1985) para definir vírus de plantas que diferem dos retrovírus por possuírem genoma de DNA e capacidade de integração ao genoma do hospedeiro. Assim como os Retrovírus, polimerizam dsDNA (DNA de fita dupla) através de um intermediário de RNA de fita simples (ssRNA) pela ação da enzima transcriptase reversa no citoplasma do hospedeiro. Entretanto, os pararetrovírus não codificam para uma proteína integrase, embora apresentem capacidade de integração dentro do genoma do hospedeiro para replicação. Essa integração genômica é conhecida como sequência endógena de pararetrovírus (Endogenous Pararetroviral Sequences - EPRVs) correspondendo a classe mais abundante de sequências virais endógenas em diferentes espécies de plantas (GAYRAL; ISKRA-CARUANA, 2009).

O tipo de replicação é citoplasmática/nuclear, ou seja, o vírion chega à célula, sendo desencapsidado pela ação de proteinases celulares e libera o dsDNA genômico que é transportado para o núcleo onde as descontinuidades do DNA são reparadas e “mini cromossomos” são formados. O DNA é transcrito pela RNA polimerase II do hospedeiro (DE KOCHKO et al.,1998), sendo então traduzido, resultando nos produtos proteicos relacionados aos genes virais existentes (FÜTTERER et al., 1997; POOGGIN et al., 1999). O RNA pré-genômico é retrotranscrito pela transcriptase reversa no citoplasma produzindo

novos dsDNA genômicos (MEDBERRY; LOCKHART; OLSZEWSKI, 1990). Os dsDNA genômicos são encapsidados, produzindo assim novos vírions. O genoma desses vírus contém uma sequência complementar ao tRNAMet da planta que corresponde ao sítio de iniciação da replicação do DNA. Geralmente, este sítio está localizado dentro ou adjacente à região intergênica (região não-codificante) e é designado como nucleotídeo um (DE KOCHKO et al.,1998).

2.2. Gênero *Badnavirus*

Dentre os gêneros da família *Caulimoviridae*, os *Badnavirus* apresentam a maior diversidade, com 40 espécies oficialmente reconhecidas pelo ICTV (<https://talk.ictvonline.org/taxonomy/>). Segundo Fauquet et al. (2005) corresponde ao segundo gênero mais numeroso entre os fitovírus com genoma de DNA, sendo *Commelina yellow mottle virus* (ComYMV) a espécie tipo do gênero, descrita pela primeira vez em *Commelina diffusa* L. na Ilha de Guadalupe por Migliori e Lastra (1978).

Apresentam partícula viral baciliforme não envelopada com aproximadamente 30 nm de largura x 130 nm de comprimento, porém, partículas de 60 – 900 nm de comprimento são comumente observadas (GEERING; HULL, 2012; BHAT; HOHN; SELVARAJAN, 2016). A origem do nome Badnavírus (*bacilliform DNA virus*) está relacionada a forma de sua partícula viral associada a constituição do seu genoma (DNA).

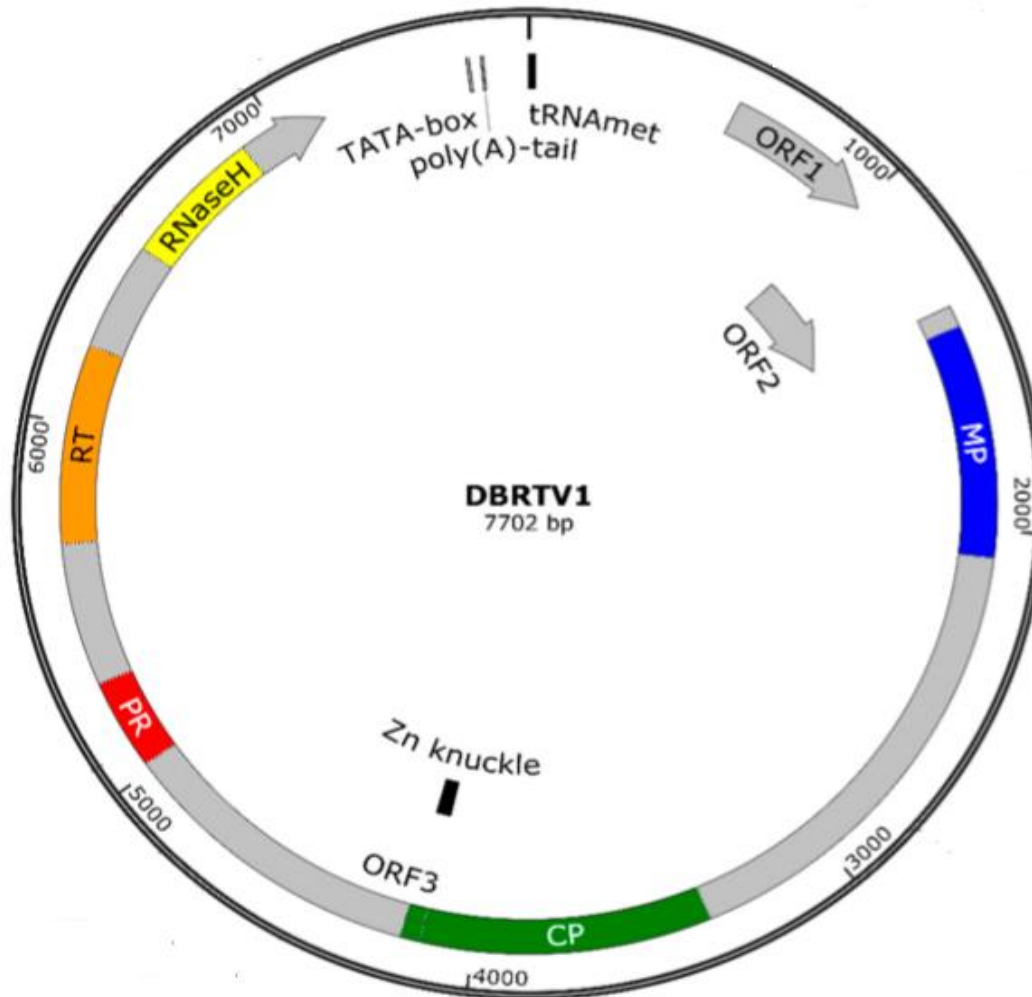
O genoma é constituído por uma única molécula de DNA de fita dupla, semi-circular, com tamanho variando de 7,2 a 9,2 kb (GEERING; HULL, 2012). Apresentam descontinuidades, que representam domínios relacionados a síntese de DNA durante a replicação, sítio-específicas e podem conter uma região intergênica poli A (MEDBERRY; LOCKHART; OLSZEWSKI, 1990).

Os *Badnavirus* codificam seu genoma para três ORFs conservadas (ORF I, II e III), entretanto, algumas espécies podem apresentar ORFs adicionais, como: *Piper yellow mottle virus* (PYMoV), *Bougainvillea spectabilis chlorotic vein banding virus* (BsCVBV), *Dioscorea bacilliform SN virus* (DBSNV), *Grapevine Roditis leaf discoloration-associated virus* (GRLDaV), *Taro bacilliform virus* (TaBV) e *Yacon necrotic mottle virus* (YNMoV) que

possuem uma ORF adicional; *Cacao swollen shoot virus* (CSSV) e *Pagoda yellow mosaic associated virus* (PYMAV), possuem duas; *Citrus yellow mosaic virus* (CYMV) e *Taro bacilliform CH virus* (TaBCHV), três e *Dracaena mottle virus* (DMV) e *Rubus yellow net virus* (RYNV) possuem quatro ORF's adicionais (BHAT; HOHN; SELVARAJAN, 2016).

As funções dos produtos das ORFs I e II ainda são pobremente compreendidas. Entretanto, a proteína P1 de aproximadamente 23kDa, codificada pela ORF I em *Comellina yellow mottle virus* (CoYMV), foi descrita em associação com vírions imaturos, enquanto que a proteína P2 de 14,5 a 15 kDa, codificada pela ORF II, foi encontrada em associação com vírions maduros e imaturos com função relacionada à encapsidação do vírion (CHENG; LOCKHART; OLSZEWSKI, 1996; JACQUOT et al., 1996; HOHN; FÜTTERER, 1997; BHAT; HOHN; SELVARAJAN, 2016). A ORF III é maior ORF, variando seu comprimento entre 5100 a 6000 pb, codificando uma poliproteína de 208 a 216 kDa que é proteoliticamente clivada, gerando, dentre outros produtos, as proteínas de movimento (MP) e da capa proteica (CP), aspartil protease (AP; responsável pela clivagem da poliproteína) a transcriptase reversa (RT) e a ribonuclease H (conhecidos como domínios genômicos conservados envolvidos no processo de replicação viral) (Figura 1) (HARPER; HULL, 1998; MEDBERRY; LOCKHART; OLSZEWSKI, 1990; BHAT; HOHN; SELVARAJAN, 2016).

Figura 1. Representação esquemática da organização genômica de *Dioscorea bacilliform* RT virus 1, apresentando as ORFs I, II e III e as proteínas comuns aos *Badnavirus* *Capa proteica* (CP), *movimento* (MP), *aspartil protease* (PR), *transcriptase reversa* (RT) e *ribonuclease H* (RNaseH) (BÖMER et al, 2016).



Para demarcação de espécies do gênero *Badnavirus* são considerados os seguintes critérios: gama de hospedeiro, especificidades do inseto vetor, e principalmente, a determinação da sequência de nucleotídeos da região genômica que inclui os domínios Transcriptase reversa e RNaseH (RT/RNaseH) localizados na ORF III do genoma viral. O ICTV estabelece um limite padrão de $\geq 80\%$ de identidade de nucleotídeos na região RT/RNaseH para demarcação de espécies deste gênero (GEERING; HULL, 2012).

A detecção de badnavírus tem sido realizada através de métodos como microscopia eletrônica de transmissão (visualização da partícula viral) e de imuno-adsorção, testes sorológicos, e testes baseados na detecção de DNA (VEGA; SORDI, 1991; DELANOY et al., 2003; RIVAS et al., 2005; FIGUEIREDO et al., 2006; ANDRADE, 2007; VERZIGNASSI et al., 2009; ALEXANDRE et al., 2010; BRIOSO; PEREIRA; GASPAROTO, 2011; JAMES et al., 2011a; PENG; GUO; WANG et al., 2012).

Testes sorológicos como ELISA (Enzyme Linked Immuno Sorbent Assay) podem ser vantajosos devido aos baixos custos, relativa eficiência e rapidez. Porém, apresentam desvantagens na detecção de espécies apresentando alta variabilidade genética, sendo comum obtenção de resultados falso-negativos (FIGUEIREDO; BRIOSO, 2007).

A técnica de PCR (Polymerase Chain Reaction) é também bastante utilizada na detecção e identificação de espécies do gênero *Badnavirus*, possuindo também limitações por não ser capaz de distinguir entre sequências virais integradas ao genoma do hospedeiro e cópias episomais dos vírus na planta (BOUSALEM et al., 2009). Dada a grande diversidade e variabilidade genética dos badnavírus, testes de PCR são utilizados para detecção de regiões genômicas conservadas presentes na ORF III (MEDBERRY; LOCKHART; OLSZEWSKI, 1990), mais especificamente os domínios tRNAMet, e RT/RNaseH (LOCKHART; OLSZEWSKI, 1993; GEERING et al., 2000).

Na tentativa de distinguir entre sequências virais endógenas e episomais, foi desenvolvida técnica de PCR-Imunocaptura, que inicialmente se utiliza um antissoro específico para depois obter a amplificação da partícula viral via PCR. (HARPER et al., 1999). Entretanto, foi observado que, apesar da alta eficiência na detecção de diversas espécies, esta técnica não era capaz de distinguir entre sequências integradas e episomais de espécies de Banana streak viruses (BSVs), dada a alta variabilidade genética presente neste complexo. (HARPER; HULL, 1998; GEERING et al., 2000, 2005b; HARPER et al., 2005; FIGUEIREDO et al., 2006).

Um método amplamente utilizado, e que apresenta alta eficiência na distinção entre sequências de badnavírus endógenas e episomais, é a amplificação do genoma completo por círculo rolante (RCA), evitando assim a detecção de vírus endógenos ou Endogenous

Pararetroviruses (EPRVs) integrados ao genoma do hospedeiro, mesmo para aquelas espécies apresentando alta variabilidade genética como no caso de BSVs (JAMES et al., 2011a, 2011b).

2.3. Sequências endógenas

Badnavirus e outros pararetrovirus, como *Cavemovirus*, *Caulimovirus*, *Petuvirus*, *Solendovirus* e *Tungrovirus*, também estão presentes como sequências fragmentadas integradas aos genomas de plantas hospedeiras (EPRVs). *Tobacco vein clearing virus* (gênero *Cavemovirus*), *Petunia vein clearing virus* (gênero *Petuvirus*) (RICHERT-PÖGGELER et al., 2003) e *Banana streak GF virus* (gênero *Badnavirus*) (GAYRAL; ISKRA-CARUANA, 2009) são exemplos de EPRVs comumente observadas na família *Caulimoviridae*.

A integração dos pararetrovirus ocorre por recombinação ilegítima em genomas hospedeiros, e sua presença não está necessariamente associada à infecção. As sequências endógenas representam geralmente formas fragmentadas e lineares do DNA viral circular original (BHAT; HOHN; SELVARAJAN, 2016).

Sequências virais endógenas muitas vezes exercem papel epidemiológico importante, já que podem originar vírus epissomais e, conseqüentemente, iniciar o processo de infecção viral. Condições de estresse dos hospedeiros tendem a favorecer este processo (DAHAL et al., 2000; DALLOT et al., 2001). Algumas espécies de badnavírus já foram descritas ocorrendo na forma de EPRVs, como: *Banana streak OL virus* (BSOLV), *Pineapple baciliform comosus virus* (PBCV), *Dioscorea baciliform virus* (DBV) *Banana streak GF virus* (BSGFV) (NDOWORA et al., 1999; STAGINNUS; RICHERT-PÖGGELER, 2006; GAYRAL; ISKRA-CARUANA, 2009; SEAL, 2014).

A cultura da banana (*Musa* spp.) destaca-se por apresentar uma alta frequência de integração de sequências de badnavírus, e comumente são encontrados dois tipos de sequências integradas: formas virais incompletas, onde não é observada infecção viral; e formas virais completas, onde, sob condições de estresse, sequências virais endógenas são liberadas com conseqüente infecção do hospedeiro (GAYRAL; ISKRA-CARUANA, 2009;

JAMES et al., 2011a). Ao menos três sequências de badnavírus integradas ao genoma da banana foram caracterizadas molecularmente: endogenous *Banana streak* OL virus (eBSOLV), endogenous *Banana streak* GF virus (eBSGFV) e endogenous *Banana streak* IM virus (eBSIMV), confirmando assim a alta diversidade destas sequências presentes no genoma do hospedeiro, o que dificulta diagnóstico preciso para espécies de BSVs (CHABANNES et al., 2013).

A presença de sequências endógenas no genoma da planta hospedeira tem se tornado um desafio à diagnose, taxonomia, movimento seguro de germoplasma e o manejo das doenças causadas pelos badnavírus. Alguns vírus endógenos podem apenas servir como componentes de genomas de plantas, enquanto outros poderiam fornecer imunidade viral (METTE et al., 2002; CHABANNES et al., 2013; BHAT; HOHN; SELVARAJAN, 2016).

Algumas técnicas são empregadas para diferenciar as formas epissomais e integradas do vírus, tais como: hibridação *in situ* de sondas virais marcadas para o cromossomo, hibridação Southern da sonda viral para o DNA digerido do hospedeiro, imunocaptura (IC)-PCR, amplificação por círculo rolante (RCA) e sequenciamento genômico da planta (BHAT; HOHN; SELVARAJAN, 2016).

2.4. Culturas de importância econômica infectadas por badnavírus

Badnavírus têm sido descritos infectando culturas de grande importância econômica em todo o mundo, como: inhame (*Dioscorea* spp.), taro (*Colocasia esculenta* L.) abacaxi (*Ananas comosus*), cana-de-açúcar (*Saccharum* spp.), banana (*Musa* spp.), citros (*Citrus* spp.), cacau (*Theobroma cacao* L.) e pimenta (*Capsicum* spp.) (YANG et al., 2003; ENI et al., 2008; JAMES et al., 2011a; JOHNSON et al., 2012; KOUAKOU et al., 2012; DEESHMA; BHAT, 2015; SILVA et al., 2015).

Em alguns casos, esta infecção pode interferir diretamente na produtividade da cultura, causando perdas significativas em qualidade e rendimento. Relatos de perdas econômicas associadas a diferentes espécies de badnavírus em diversas culturas variam entre 10% a 90% (BHAT; HOHN; SELVARAJAN, 2016). Essa variação ocorre em função da espécie botânica, cultivar/variedade, idade e época de infecção na planta, percentual de

plantas infectadas, virulência da estirpe/espécie viral, e da interação com outros grupos de vírus em uma mesma planta (AHLAWAT et al., 1996; COMSTOCK & LOCKHART, 1996; BHAT et al., 2003; GEIJSKES, 2003; BARANWAL et al., 2005; GEERING, 2009; BESSE et al., 2010).

Em 1999, foi caracterizado o primeiro badnavírus na cultura do inhame, *Dioscorea bacilliform* AL virus (DBALV) infectando *D. alata* na Nigéria (PHILLIPS, 1999) Este vírus pode induzir distorção foliar severa, sendo comumente encontrado em infecções mistas com vírus do gênero *Potyvirus* (PIO-RIBEIRO et al., 2005; YANG et al., 2003). Em seguida, foi descoberto um novo badnavírus infectando *D. sansibarensis* em ilhas do Sul do Pacífico, o *Dioscorea bacilliform* SN virus (DBSNV) (GEERING; HULL, 2012). Entretanto estudos realizados na África e Sul do Pacífico sugerem alta diversidade de espécies de badnavírus em inhame, com possível presença de até doze espécies distintas (ENI et al., 2008). Os sintomas causados por badnavirose na cultura do inhame são: mosaico, clorose nas folhas e deformação foliar, levando a uma redução da formação de açúcar e do armazenamento de amido, reduzindo significativamente o rendimento e qualidade dos tubérculos. O vírus é transmitido vegetativamente e através de cochonilhas *Planococcus citri* (Risso)(Hemiptera: Pseudococcidae) (BHAT; HOHN; SELVARAJAN, 2016).

Estudos com ELISA e Southern hybridization envolvendo oito espécies de inhame suportam a presença de DBV endógeno (eDBV) em *D. rotundata* (SEAL et al., 2014). Umber et al. (2014) relataram a caracterização molecular de sequências de eDBV no genoma de inhame na África do complexo *D. cayensis-rotundata* mostrando uma sequência proveniente de várias partes de genomas de badnavirus, resultando em uma estrutura em mosaico. Relataram também que os eDBVs pertencem a pelo menos quatro badnavirus distintos o que indica vários eventos de incorporação genômica independentes. Até o momento não há nenhum relatório sobre a possibilidade de eDBV gerar partículas de vírus infecciosas sob condições de estresse abiótico (BHAT; HOHN; SELVARAJAN, 2016).

No ano de 1995, Wakman et al. detectaram *Pineapple bacilliform virus* pela primeira vez em híbridos de abacaxi na Austrália. Pelo menos duas espécies de badnavírus são reconhecidas infectando a cultura do abacaxi: *Pineapple bacilliform* CO virus (PBCoV) e

Pineapple bacilliform ER virus (PBERV). Ambos são transmitidos pela cochonilha *Dysmicoccus brevipes* (Cockerell) (Hemiptera: Pseudococcidae) e, no caso do PBCoV, também pelas cochonilhas *Planococcus citri* e *D. neobrevipes* (GAMBLEY et al., 2008; SETHER et al., 2012). Entretanto, uma associação clara entre sintomas e infecção por badnavírus em abacaxi ainda não foi demonstrada (GAMBLEY et al., 2008). Sequências endógenas também foram identificadas na cultura do abacaxi (SETHETTER et al., 2012).

Sugarcane bacilliform virus foi descrito pela primeira vez em Cuba em 1985 e posteriormente em muitos países pelo mundo (AUTREY; BOOLELL; JONES, 1995). Na cultura da cana-de-açúcar, são reconhecidas pelo ICTV quatro espécies de badnavírus: *Sugarcane bacilliform IM virus* (SCBIMV), *Sugarcane bacilliform MO virus* (SCBMOV), *Sugarcane bacilliform Guadeloupe A virus* (SCBGAV) e *Sugarcane bacilliform Guadeloupe D virus* (SCBGDV), sendo comumente observados sintomas de manchas e estrias cloróticas e nanismo, embora, muitas plantas encontrem-se assintomáticas (VISWANATHAN; ALEXANDER; GARG, 1996). Trabalho desenvolvido por Li et al. (2010) mostrou dentre outros fatores, uma diminuição do teor de sacarose e peso do caule em amostras de cana-de-açúcar infectadas com SCBV na China. Diversas espécies e híbridos comerciais de cana-de-açúcar (*S. barberi*, *S. officinarum*, *S. robustum*, *S. spontaneum* e *S. sinense*) já foram relatados sendo infectados por diferentes espécies de badnavírus (VISWANATHAN; ALEXANDER; GARG, 1996; SILVA et al., 2015).

Trabalho realizado na Índia mostrou que pelo menos três novas espécies de badnavírus estão associadas à cultura da cana-de-açúcar: *Sugarcane bacilliform BB virus* (SCBBBV), *Sugarcane bacilliform BO virus* (SCBBOV) e *Sugarcane bacilliform black reunion virus* (SCBBRV) (KARUPPAIAH; VISWANATHAN; KUMAR, 2013). Recentemente, foi demonstrada uma alta diversidade de espécies de badnavírus infectando diferentes espécies/híbridos de cana-de-açúcar em um banco de germoplasma no Brasil, com relato de uma possível espécie nova: *Sugarcane bacilliform AL virus* (SCBALV) (SILVA et al., 2015).

Desde o seu primeiro relato em 1974 no Marrocos, o *Banana streak virus* (BSV), foi reportado em vários países da África, da América do Sul e do Pacífico e, provavelmente,

ocorre mundialmente onde há cultivo de banana (LOCKHART, 1995; LOCKHART; JONES, 2000; KUMAR et al., 2015). Diversas espécies de badnavírus já foram relatadas na cultura da bananeira (*Musa spp.*) causando uma doença conhecida como estria clorótica, as quais são consideradas um complexo de espécies denominado BSVs (LHEUREUX et al., 2007; JAMES et al., 2011b). Os sintomas da doença variam e são influenciados pela cultivar, espécie do vírus e condições ambientais. Além disso, plantas infectadas com BSV podem apresentar-se assintomáticas, sendo a temperatura considerada um fator importante na expressão da doença. O BSV é transmitido horizontalmente às bananas de forma semi-persistente por cochonilhas onde a espécie *Planococcus citri* é o mais prevalente, e verticalmente por micropropagação em massa – principal método de propagação de plantas de banana (BHAT; HOHN; SELVARAJAN, 2016).

As perdas causadas por este complexo de badnavírus podem variar de 6 a 90 % (DAHAL et al., 2000; DANIELLS et al., 2001), levando a grandes prejuízos aos produtores. Pelo menos dez diferentes badnavírus já foram relatadas compondo esse complexo de espécies em banana em todo o mundo: *Banana streak OL virus* (BSOLV), *Banana streak GF virus* (BSGFV), *Banana streak MY virus* (BSMYV), *Banana streak VN virus* (BSVNV), *Banana streak IM virus* (BSIMV), *Banana streak UA virus* (BSUAV), *Banana streak UI virus* (BSUIV), *Banana streak UL virus* (BSULV), *Banana streak UM virus* (BSUMV) e *Banana streak CA virus* (BSCAV) (HARPER; HULL, 1998; LHEUREUX et al., 2007; JAMES et al., 2011a; GEERING; HULL, 2012).

Staginnus et al., (2009) relataram quatro espécies de BSV apresentando sequências integradas ao genoma de plantas, são elas: BSOLV, BSGFV, BSIMV e BSMYV, denominadas de eBSV.

Citrus yellow mosaic virus (CYMV) foi inicialmente relatado em laranja doce e, posteriormente, em outras espécies cítricas (DAKSHINAMURTI; REDDY, 1975). Este badnavírus emergiu no sul da Índia causando a doença do mosaico dos citros, podendo ser transmitida através de enxertia (AHLAWAT et al., 1996). Recentemente, CYMV foi relatado em diferentes cultivares de citros: Sweet Orange, Pummelo, Acid Lime, Sweet Orange SOP, Sweet Orange Nagri, Acid Lime AP e Rangpur Lime, o qual se apresentou filogeneticamente

próximo do *Cacao swollen shoot virus* (CSSV) (JOHNSON et al., 2012). É especialmente comum e severo em Sweet Orange, apresentando uma variação de incidência de 10 a 70% e a redução do rendimento da fruta foi de 77%, os frutos das árvores afetadas apresentaram 10% menos de suco e ácido ascórbico (AHLAWAT et al., 1996).

O badnavirus *Cacao swollen shoot virus* (CSSV), infectando cacau, foi descoberto pela primeira vez na região leste de Gana em 1936. Provoca uma doença devastadora que afeta o cultivo de cacau na maioria dos países da África Ocidental. A doença é caracterizada por inchaço de raízes e caule e as folhas tornam-se vermelhas, podendo matar a planta em dois anos (POSNETTE, 1947; POSNETTE, 1950; THRESH, 1980; KEBE et al., 2006; KOUAKOU et. al, 2012). No Brasil, é considerada praga quarentenária ausente. A doença é causada por pelo menos três espécies distintas de badnavírus: *Cacao swollen shoot virus* (CSSV), *Cacao swollen shoot Togo virus* (CSSToV) e *Cacao swollen shoot Côte d'Ivoire D virus* (CSSCIV), indicando a existência de uma alta variabilidade entre isolados de CSSV ocorrendo nos principais países produtores de cacau (KEBE et al., 2006; KOUAKOU et. al, 2012). Na Costa de Marfim, maior produtor de cacau do mundo, a sustentabilidade da produção foi comprometida desde o aparecimento desta doença. Os sintomas da infecção podem ser observados em diversos órgãos da árvore (folhas, galhos, brotos), podendo afetar a produtividade e até mesmo causar morte da planta (CASTEL et al., 1980).

A disseminação de badnavírus ocorre principalmente por propagação vegetativa em algumas espécies como *Commelina yellow mottle virus* (ComYMV), *Kalanchoe top-spotting virus* (KTSV), *Piper yellow mottle virus* (PYMoV), *Cocoa swollen shoot virus* (CSSV) e *Taro bacilliform virus* (TaBV) e também podem ser transmitidas por sementes (HEARON; LOCKE, 1984; QUAINOO; WETTEN; ALLAINGUILLAUME, 2008; HAREESH; BHAT, 2010; DEESHMA; BHAT, 2014; BHAT; HOHN; SELVARAJAN, 2016.) A transmissão horizontal de badnavírus ocorre através de várias espécies de cochonilha, enquanto as espécies *Gooseberry vein banding virus* (GVBaV), *Rubus yellow net virus* (RYNV) e *Spiraea yellow leaf spot virus* (SYLSV) são transmitidas por afídeos (BHAT; HOHN; SELVARAJAN, 2016).

2.5. Variabilidade genética e estrutura de populações de fitovírus

De acordo com Milgroom (1995) a estrutura genética de populações pode ser definida em função da quantidade de variações genéticas existentes e a sua distribuição nas populações. O estudo da dinâmica da variabilidade genética nas populações de fitopatógenos é importante, pois permite a compreensão da evolução das populações (MCDONALD; ZHAN; BURDON, 1999). Dessa forma, permite-se deduzir a taxa de evolução de uma população (MCDONALD; LINDE, 2002).

O conhecimento sobre as estruturas genéticas é muito empregado em programas de melhoramento genético visando resistência a patógenos. McDonald e Linde (2002) propuseram um modelo para auxiliar na definição de estratégias de melhoramento, que permitem aumentar a duração da resistência baseado no potencial evolutivo populacional de fitopatógenos.

Em populações de fitovírus, o alto grau de diversidade e variabilidade genética é conferido, sobretudo, por erros ocorridos durante o processo de replicação do genoma viral, as mutações e por rearranjos genômicos provocados por eventos de recombinação e pseudorecombinação (GARCIAARENAL; FRAILE; MALPICA, 2003).

A mutação pode ser definida como um processo pelo qual os nucleotídeos externos à fita-molde do ácido nucléico são incorporados à fita-filha durante sua replicação e a taxa de mutação avalia o surgimento de mutações espontâneas em função do tempo (GARCIAARENAL; FRAILE; MALPICA, 2003). No geral, as taxas de mutação em vírus de RNA e ssDNA são bastante elevadas (DUFFY; SHACKELTON; HOLMES, 2008).

Nos vírus de RNA, essas taxas elevadas são explicadas pela replicação utilizando polimerases de RNA dependentes e que não possuem capacidade de correção do erro (DRAKE; HOLLAND, 1999; JENKINS et al., 2002). Para os vírus de ssDNA apesar do sistema de correção de erro existir, esses vírus tendem a evoluir tão rapidamente quanto os de RNA (SHACKELTON et al., 2006; DUFFY; SHACKELTON; HOLMES, 2008).

Existem duas hipóteses que explicam esse fenômeno: a primeira, o tempo de duração do intermediário de fita dupla formado durante a replicação do ssDNA não é suficiente para que os erros sejam corrigidos, e a segunda: as mutações espontâneas podem ocorrer no DNA durante a fase de ssDNA (DUFFY; SHACKELTON; HOLMES, 2008; DUFFY; HOLMES, 2009).

Apesar do consenso em que, os vírus de ssDNA evoluem mais rapidamente que os de dsDNA, as informações sobre as taxas de mutação para vírus de dsDNA que infectam plantas ainda são escassas, com poucos relatos na literatura e são descritos evoluindo através de associações com seus hospedeiros (FIRTH et al., 2010).

A recombinação ocorre quando segmentos de uma fita de ácido nucleico são incorporados na fita de um indivíduo diferente durante a replicação (PADIDAM; BEACHY; FAUQUET; 1999). Esses eventos podem estar associados ao aumento da gama de hospedeiros, da virulência, do potencial evolutivo e da adaptação (GIBBS; WEILLER, 1999; PITA et al., 2001; MONCI et al., 2002).

Em fitovírus com DNA de fita simples, a recombinação é um processo bastante comum, sendo relacionada com os altos índices de variabilidade observados em populações como Begomovirus (PADIDAM; BEACHY; FAUQUET, 1999; LEFEUVRE et al., 2007; LIMA et al., 2013; ROCHA et al., 2013; RAMOS-SOBRINHO et al., 2014), nesse caso, infecções mistas e a replicação dependente de recombinação explicam elevadas taxas de recombinação.

Além da importância econômica, badnavírus são considerados objetos ideais para estudos de recombinação e receberam atenção especial após a descoberta de sua capacidade de interação com o genoma do hospedeiro (SQUIRES et al., 2011). Recombinação homóloga foi relatada durante a ativação dos genomas episomais em BSV e TVCV que estavam integrados ao genoma dos hospedeiros (NDOWORA et al., 1999; STAGINNUS; RICHERT-PÖGGELER, 2006). Algumas pesquisas apontam que os eventos de recombinação tendem a contribuir para variação genética observada em alguns *Caulimoviridae* (ARBOLEDA; AZZAM, 2000; FROISSART et al., 2005).

Eventos de recombinação têm sido comumente descritos para espécies de badnavirus que infectam banana, cana-de-açúcar e inhame (SHARMA et al., 2015; BÖOMER et al., 2017). Em banana e cana-de-açúcar, já foram descritos eventos de recombinação interespecífica entre as espécies SCBV e BSV, além de recombinação intraspecífica respectivamente, em SCBV e BSV (SHARMA et al., 2015). Já em inhame, um provável evento de recombinação interespecífica entre DBALV e DBSNV deu origem a espécie DBRTV3 (BÖOMER et al., 2017).

REFERÊNCIAS BIBLIOGRÁFICAS

- AHLAWAT, Y.S.; PANT, R.P.; LOCKHART, B.E.L.; SRIVASTAVA, M.; CHAKRABORTY, N.K.; VARMA, A. Association of a badnavirus with citrus mosaic disease. **Plant Disease**, v. 80, p. 590-592, 1996.
- ALEXANDRE, M.A.V.; DUARTE, L. M. L.; RIVAS, E. B.; GALLETI, S. R. Vírus detectados em plantas ornamentais no período 2004 a 2008. **Revista Brasileira de Horticultura Ornamental**, v. 16, p. 95- 100, 2010.
- ANDRADE, G.P. **Diagnóstico fitossanitário da cultura do inhame (*Dioscorea spp.*) em áreas produtoras do nordeste do Brasil**. 2007. Tese (Doutorado em Fitopatologia) - Universidade Federal Rural de Pernambuco (UFRPE), Recife, 2007.
- ARBOLEDA, M.; AZZAM, O. Inter and intra site genetic diversity of natural field populations of Rice Tungro Bacilliform Virus in the Philippines. **Archives of Virology**, v. 145, n. 2, p. 275-89, 2000.
- AUTREY, L.J.C.; BOOLELL, S.; JONES, P. Distribution of sugarcane bacilliform virus in various geographical regions. In: Congress of the International Society of Sugarcane Technologists, XXI, Bangkok, 1995. **Proceedings**. Bangkok: Kasetsart University, p. 657.
- BARANWAL, V.K.; MAJUMDER, S.; AHLAWAT, Y. S.; SINGH, R.P. A novel approach for simultaneous detection of Citrus yellow mosaic virus and Citrus greening bacterium by multiplex polymerase chain reaction. **Indian Journal of Biotechnology**, v. 4, p. 528-33, 2005.
- BESSE, S.; GURGELI, P.; RAMEL, M. E.; BALMELLI, C. Characterisation of mixed virus infections in Ribes species in Switzerland. 1st International Conference on Virus and other Graft Transmissible Diseases of Fruit Crops. **Julius Kühn Archives**, n. 427, p. 214-9, 2010.
- BHAT, A.I.; DEVASAHAYAM, S.; SARMA, Y. R.; PANT, R. P. Association of a badnavirus in black pepper (*Piper nigrum* L.) transmitted by mealybug (*Ferrisia virgata*) in India. **Current Science**, v. 84, p. 1547-50, 2003.
- BHAT, A. I.; HOHN, T; SELVARAJAN, R. Badnaviruses: the current global scenario. **Viruses**, v. 8, n. 6, p. 177, 2016.
- BÖMER, M.; TURAKI, A. A.; SILVA, G.; KUMAR, P.; SEAL, S. E. A sequence-independent strategy for amplification and characterisation of episomal badnavirus sequences reveals three previously uncharacterised yam badnaviruses. **Viruses**, v. 8, n. 7, p. 188, 2016.
- BÖMER, M.; RATHNAYAKE, A. I.; VISENDI, P.; SILVA, G.; SEAL, S. E. Complete genome sequence of a new member of the genus Badnavirus, *Dioscorea bacilliform* RT virus 3, reveals the first evidence of recombination in yam badnaviruses. **Archives of virology**, p. 1-6, 2017.

BOUHIDA, M. L.; LOCKHART, B. E.; OLSZEWSKI, N. E. An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. **Journal of General Virology**, v. 74, p. 15-22, 1993.

BOUSALEM, M.; DURAND, O.; SCARCELLI, N.; LEBAS, B. S. M.; KENYON, J. L.; MARCHAND, L.; LEFORT, F.; SEAL, S. E. Dilemmas caused by endogenous pararetroviruses regarding the taxonomy and diagnosis of yam (*Dioscorea* spp.) badnaviruses: analyses to support safe germplasm movement. **Archives of Virology**, v. 154, p.297–314, 2009.

BRIOSO, P. S. T.; PEREIRA, J. C. R.; GASPAROTO, A. Ocorrência de badnavirus em frutos de bananeira no estado do Amazonas, Brasil. **Revista Brasileira de Fruticultura**, v. 33, n. 4, p. 1353-1355, 2011.

CASTEL, C. A.; DJIEKPOR, Y. K.; PARTIOT, E. K.; SEGBOR, A. M. Le swollen shoot du cacaoyer au Togo. Les différentes formes de viroses et leurs conséquences économiques. **Café, Cacao, Thé**, v. 24, n. 2, p. 131–146, 1980.

CHABANNES, M.; BAURENS, F. C.; DUROY, P. O.; BOCS, S.; VERNEREY, M.S.; RODIER-GOUD, M.; BARBE, V.; GAYRAL, P.; ISKRA-CARUANA, M. L. Three infectious viral species lying in wait in the banana genome. **Journal of Virology**, v. 87, p. 8624-8637, 2013.

CHENG, C.P.; LOCKHART, B.E.; OLSZEWSKI, N.E. The ORF I and II Proteins of Commelina yellow mottle virus are virion-associated. **Virology**, v. 223, p. 263-271, 1996.

COMSTOCK, J. C.; LOCKHART, B. E. L. Effect of sugarcane bacilliform virus on biomass production of three sugarcane cultivars. **Sugar Cane**, v. 4, p. 12-5, 1996.

CÔTE, F. X.; GALZI, S.; FOLLIOU, M.; LAMAGNÈRE, Y.; TEYCHENEY, P. Y.; ISKRA-CARUANA, M. L. Micropropagation by tissue culture triggers differential expression of infectious endogenous Banana streak virus sequences (eBSV) present in the B genome of natural and synthetic interspecific banana plantains. **Molecular Plant Pathology**, v. 11, p. 137-144, 2010.

DAHAL, G. O. R.; ORTIZ, R.; TENKOUANO, A.; HUGHES, J. D. A.; THOTTAPPILLY, G.; VUYLSTEKE, D.; LOCKHART, B. E. L. Relationship between natural occurrence of banana streak badnavirus and symptom expression, relative concentration of viral antigen, and yield characteristics of some micropropagated *Musa* spp. **Plant Pathology**, v. 49, p. 68-79, 2000.

DAKSHINAMURTI, V.; REDDY, G.S. Mosaic - a transmissible disorder of sweet oranges. **Indian Phytopathology**, v. 28, p. 398-399, 1975.

DALLOT, S.; ACCUNA, P.; RIVERA, C.; RAMIREZ, P.; COTE, F.; LOCKHART, B. E. L.; CARUANA, M. L. Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of Banana streak virus integrated into the genome of the FHIA21 hybrid (*Musa* AAAB). **Archives of Virology**, v. 146, p. 2179-2190, 2001.

DANIELLS, J. B.; GEERING, A. D. W.; BRYDE, N. J.; THOMAS, J. E. The effect of banana streak virus on the growth and yield of dessert bananas in tropical Australia. **Annals of Applied Biology**, v. 139, p. 51-60, 2001.

DE KOCHKO, A.; VERDAGUER, B.; TAYLOR, N.; CARCAMO, R.; BEACHY, R. N.; FAUQUET, C. Cassava Vein Mosaic Virus (CsVMV), type species for a new genus of plant double stranded DNA viruses. **Archives of Virology**, v. 143, n. 5, p. 945-962, 1998.

DEESHMA, K.P.; BHAT, A.I. Further evidence of true seed transmission of Piper yellow mottle virus in black pepper (*Piper nigrum* L.). **Journal of Plantation Crops**, v. 42, p. 289-293, 2014.

DEESHMA, K. P.; BHAT, A. I. Complete genome sequencing of Piper yellow mottle virus infecting black pepper, betelvine, and Indian long pepper. **Virus Genes**, v. 50, p. 172–175, 2015.

DELANOY, M.; SALMON, M.; KUMMERT, J.; FRISON, E.; LEPOIVRE, P. Development of realtime PCR for the rapid detection of episomal Banana streak virus (BSV). **Plant Disease**, v.87, p.33-8, 2003.

DRAKE, J. W.; HOLLAND, J. J. Mutation rates among RNA viruses. **Proceedings of the National Academy of Sciences**, v. 96, n. 24, p. 13910-13913, 1999.

DUFFY, S.; SHACKELTON, L. A.; HOLMES, E. C. Rates of evolutionary change in viruses: patterns and determinants. **Nature Reviews Genetics**, v. 9, p. 267–276, 2008.

ENI, A. O.; HUGHES, J. D.; ASIEDU, R.; REY, M. E. Sequence diversity among *Badnavirus* isolates infecting yam (*Dioscorea spp.*) in Ghana, Togo, Benin and Nigeria. **Archives of Virology**, v. 153, p. 2263-2272, 2008.

FAUQUET, C. M.; MAYO, M. A.; MANILOFF, J.; DESSELBERGER, U.; BALL, L. A. **Virus taxonomy: VIII report of the International Committee on Taxonomy of Viruses**. Amsterdam: Academic Press. 2005.

FIGUEIREDO, D. V.; MEISSNER FILHO, P. E.; SILVA NETO, S. P.; BRIOSO, P. S. T. Detecção e análise da variabilidade de sequências do Banana streak virus (BSV) em bananeiras no Brasil. **Summa Phytopathologica**, v. 32, p. 118-23, 2006.

FIGUEIREDO, D. V.; BRIOSO, P. S. T. PCR multiplex para a detecção de BSV e CMV em bananeiras micropropagadas. **Summa Phytopathology**, v. 33, n. 2, p. 118-123, 2007.

FIRTH, C.; KITCHEN, A.; SHAPIRO, B.; SUCHARD, M. A.; HOLMES, E. C.; RAMBAUT, A. Using time-structured data to estimate evolutionary rates of doublestranded DNA viruses. **Molecular Biology and Evolution**, v. 27, p. 2038-2051, 2010.

FROISSART, R.; ROZE, D.; UZEST, M.; GALIBERT, L.; BLANC, S.; MICHALAKIS, Y. Recombination every day: abundant recombination in a virus during a single multicellular host infection. **PLoS Biology**, v. 3, n. 3, p. 89, 2005.

FÜTTERER, J.; ROTHNIE, H. M.; HOHN, T.; POTRYKUS, I. Rice tungro bacilliform virus open read frame II and III are translated from polycistronic pre genomic RNA by leaky scanning. **Journal of General Virology**, v. 71, n. 10, p. 7984-9, 1997.

GAMBLEY, C. F.; GEERING, A. D. W.; STEELE, V.; THOMAS, J. E. Identification of viral and non-viral reverse transcribing elements in pineapple (*Ananas comosus*), including members of two new badnavirus species. **Archives of Virology**, v. 153, p. 1599–1604, 2008.

GÁRCIA-ARENAL, F.; FRAILE, A.; MALPICA, J. M.; Variation and evolution of plant virus populations. **International Microbiology**, v. 6, p. 225-232, 2003.

GAYRAL, P.; ISKRA-CARUANA, M. Phylogeny of Banana streak virus reveals recent and repetitive endogenization in the genome of its banana host (*Musa* sp.). **Journal of Molecular Evolution**, v. 69, p. 65-80, 2009.

GEERING, A. D. W.; MCMICHAEL, L. A.; DIETZGEN, R. G.; THOMAS, J. E. Genetic diversity among Banana streak virus isolates from Australia. **Phytopathology**, v. 90, p. 921-927, 2000.

GEERING, A. D. W.; POOGGIN, M. M.; OLSZEWSKI, N. E.; LOCKHART, B. E. L.; THOMAS, J. E. Characterisation of banana streak mysore virus and evidence that its DNA is integrated in the B genome of cultivated *Musa*. **Archives of Virology**, v. 150, n. 4, p. 787-796, 2005.

GEERING, A.D.W. Viral pathogens of banana: outstanding questions and options for control. **Acta Horticulturae**, v. 828, p. 39-50, 2009.

GEERING, A. D. W.; HULL, R. Family Caulimoviridae. In: KING A.M.Q.; ADAMS M.J.; CARSTENS E.B.; LEFKOWITZ, E.J. **Virus Taxonomy: 9th Report of the International Committee on Taxonomy of Viruses**. London: Elsevier Academic Press, p. 429-443, 2012.

GEIJSKES, R. J.; BRAITHWAITE, K. S.; DALE, J. L.; HARDING, R. M.; SMITH, G. R. Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. **Archives of Virology**, v. 147, p. 2393-2404, 2002.

GEIJSKES, R.J. **Characterisation of an australian isolate of Sugarcane bacilliform virus**. PhD Thesis. Queensland University of Technology, 2003.

GIBBS, M. J; WEILLER, G. F. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. **Proceedings of the National Academy of Sciences**, v. 96, n. 14, p. 8022-8027, 1999.

GOVIND, P. R.; SUSHEEL, K. S.; DEEPTI, S.; MEENAKSHI, A.; PRIYANKA, S.; VIRENDRA, K. B. Genetically diverse variants of *Sugarcane bacilliform virus* infecting sugarcane in india and evidence of a novel recombinant *Badnavirus* variant. **Journal of Phytopathology**, v. 162, p. 779–787, 2014.

GUIMARÃES, K. M.; SILVA, S. J. C.; MELO, A. M.; RAMOS-SOBRINHO, R.; LIMA, J. S.; ZERBINI, F. M.; ASSUNÇÃO, I. P.; LIMA, G. S. A. Genetic variability of badnaviruses infecting yam (*Dioscorea* spp.) in northeastern Brazil. **Tropical Plant Pathology**, v. 40, p. 111-118, 2015.

HAGEN, L. S.; JACQUEMOND, M.; LEPINGLE, A.; LOT, H.; TEPFER, M. Nucleotide sequence and genomic organisation of cacao swollen shoot virus. **Virology**, v. 196, p. 619-628, 1993.

HAREESH, P.S.; BHAT, A.I. Seed transmission of Piper yellow mottle virus in black pepper (*Piper nigrum* L.). **Journal of Plantation Crops**, v. 38, n. 1, p. 62-65, 2010.

HARPER, G.; HULL, R. Cloning and sequence analysis of Banana Streak Virus DNA. **Virus Genes**, v. 17, p. 271-278, 1998.

HARPER, G.; DAHAL, G.; THOTTAPPILLY, G.; HULL, R. Detection of episomal Banana Streak Badnavirus by IC-PCR. **Journal of Virological Methods**, v.79, n. 1, p.1-8, 1999.

HARPER, G.; HARPER, G.; HULL, R.; LOCKHART, B.; OLSZEWSKI, N. Viral sequences integrated into plant genomes. **Annual Review of Phytopathology**, v. 40, p. 119-136, 2002.

HEARON, S. S.; LOCKE, J. C. Graft, pollen and seed transmission of an agent associated with top spotting in *Kalanchoe blossfeldiana*. **Phytopathology**, v. 74, p. 347-50, 1984.

HOHN, T.; FÜTTERER, J. The proteins and functions of plant pararetroviruses: knowns and unknowns. **Critical Reviews in Plant Sciences**, v. 16, p. 133-161, 1997.

HOHN, T.; RICHERT-POGGELER, K. R.; HARPER, G.; SCHAWARZACHER, T.; TEO, C.; TEYCHENEY, P. Y.; ISKRA-CARUANA, M. L.; HULL, R. Evolution of integrated plant viruses. In: ROOSINCK, M. (Eds.). **Plant Virus Evolution**. Heidelberg: Academic Springer. pp. 54–76, 2008.

JACQUOT, E.; HAGEN, L. S.; JACQUEMOND, M.; YOT, P. The open reading frame 2 product of cacao swollen shoot badnavirus is a nucleic acid-binding protein. **Virology**, v. 225, p. 191-195, 1996.

JAMES, A. P.; GEIJSKES, R. J.; DALE, J. L.; HARDING, R. M. Development of a novel rolling-circle amplification technique to detect Banana streak virus that also discriminates between integrated and episomal virus sequences. **Plant Disease**, v. 95, p. 57-62, 2011a.

JAMES, A. P.; GEIJSKES, R. J.; DALE, J. L.; HARDING, R. M. Molecular characterization of six badnavirus species associated with leaf streak disease of banana in East Africa. **Annals of Applied Biology**, v. 158, p. 346-353, 2011b.

JENKINS, G. M.; RAMBAUT, A.; PYBUS, O. G.; HOLMES, E. C. Rates of molecular evolution in RNA viruses: a quantitative phylogenetics analysis. **Journal of Molecular Evolution**, v. 54, p. 152–161, 2002.

JOHNSON, A. M.; BORAH, B. K.; SAI GOPAL, D. V.; DASGUPTA, I. 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**, v. 45, p. 600–605, 2012.

KARUPPAIAH, R.; VISWANATHAN, R.; KUMAR, V. G. Genetic diversity of Sugarcane bacilliform virus isolates infecting Saccharum spp. in India. **Virus Genes**, v. 46, p. 505–516, 2013.

KEBE, B. I.; KOFFIE, K.; N'GUESSAN, K. F.; ASSIRI, A. A.; ADIKO, A.; AKÉ, S.; ANNO, P. A. Le swollen shoot en Côte d'Ivoire: situation actuelle et perspectives. In: INTERNATIONAL COCOA RESEARCH CONFERENCE, 15, 2006, San José. **Anais**. San José, p. 907-922, 2006.

KOUAKOU K.; KÉBÉ, B. I.; KOUASSI, N.; MULLER, E. Geographical distribution of *Cacao swollen shoot virus* molecular variability in Côte d'Ivoire. **Plant Disease**, v. 96, p. 1445-1450, 2012.

KUMAR, P. L.; SELVARAJAN, R.; ISKRA-CARUANA, M. L.; CHABANNES, M.; HANNA, R. Biology, etiology, and control of virus diseases of banana and plantain. *Advances Virus Research*, v. 91, p. 229–269, 2015.

LEFEUVRE, P.; MARTIN, D. P.; HOAREAU, M.; NAZE, F.; DELATTE, H.; THIERRY, M.; VARSANI, A.; BECKER, N.; REYNAUD, B.; LETT, J. M. Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. **Journal of General Virology**, v. 12, p. 3458-68, 2007.

LHEUREUX, F.; CARREEL, F.; JENNY, C.; LOCKHART, B. E. L.; ISKRA-CARUANA, M. L. Identification of genetic markers linked to banana streak disease expression in inter-specific *Musa* hybrids. **Theoretical and Applied Genetics**, v. 106, p. 594-598, 2003.

LHEUREUX, F.; LABOUREAU, N.; MULLER, E.; LOCKHART, B.E.; ISKRA-CARUANA, M. L. Molecular characterization of banana streak acuminata Vietnam virus isolated from *Musa acuminata siamea* (banana cultivar). **Archives of Virology**, v. 152, p. 1409-1416, 2007.

LI, W. F.; HUANG, Y. K.; JIANG, D. M.; ZHANG, Z. X.; ZHANG, B. L.; LI, S. F. Detection of Sugarcane bacilliform virus isolate and its influence on yield and quality of cane in Yunnan. **Acta Phytopathologica Sinica**, v. 6, p. 651-654, 2010.

LIMA, J. S.; LIMA, A. T. M.; CASTILLO-URQUIZA, G. P.; SILVA, S. J. C.; ASSUNÇÃO, I. P.; MICHEREFF, S. J.; ZERBINI, F. M.; LIMA, G. S. A. Variabilidade genética de isolados de badnavírus infectando inhame (*Dioscorea* spp.) no Nordeste do Brasil. **Tropical Plant Pathology**, v.38, p.349-353, 2013.

LOCKHART, B. E. L.; OLSZEWSKI, N.E. Serological and genomic heterogeneity of banana streak badnavirus: implications for virus detection in *Musa* germplasm, *Breeding*

Banana and Plantain for Resistance to Diseases and Pests. **CIRAD/INIBAP**, p. 105-113, 1993.

LOCKHART, B. E. L. **Banana streak Badnavirus infection in Musa: epidemiology, diagnosis, and control**. Taipei: ASPAC Food & Fertilizer Technology Center, 1995.

LOCKHART, B. E. L.; JONES, D. R. Banana streak mosaic. In: JONES, D.R. (Ed.). **Diseases of Banana, Abaca and Enset**. Wallingford: CAB International. 2000

MCDONALD, B. A.; ZHAN, J.; BURDON, J. J. Genetic structure of *Rhynchosporium secalis* in Australia. **Phytopathology**, v .89, p .639-45, 1999.

MCDONALD, B. A.; LINDE, C. The population genetics of plant pathogens and breeding strategies for durable resistance. **Euphytica**, v. 124, p. 163–180, 2002.

MEDBERRY, S. L.; LOCKHART, B. E. L.; OLSZEWSKI, N. E. Properties of Commelina yellow mottle virus complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. **Nucleic Acids Research**, v. 18, p. 5505-5513, 1990.

METTE, M. F.; KANNO, T.; AUFSATZ, W.; JAKOWITSCH, J.; Van Der WINDEN, J.; MATZKE, M. A.; MATZKE, A. J. M. Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. **EMBOJ**, v. 21, p. 461–469, 2002.

MIGLIORI, A.; LASTRA, R. Study of viruses on *Commelina diffusa* Burm. in Guadeloupe. **Annales de Phytopathologie**, v. 4, p. 467-77, 1978.

MILGROOM, M.G. Analysis of population structure in fungal plant pathogens. In: LESLIE, J. F.; FREDERIKSEN, R. A. (Eds.) **Disease analysis through genetics and biotechnology**. Ames: Iowa State University. 1995.

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**, v. 303, n. 2, p. 317-26, 2002.

MULLER, E.; DUPUY, V.; BLONDIN, L.; BAUFFE, F.; DAUGROIS, J. H.; LABOUREAU, N.; ISKRA-CARUANA, M. L. High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. **Virus Research**, v. 160, p. 414-419, 2011.

NDOWORA, T. D. G.; DAHAL, G.; LAFLEUR, D.; HARPER, G.; HULL, R.; OLSZEWSKI, N. E.; LOCKHART, B. Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. **Virology**, v. 255, p. 214-220, 1999.

PADIDAM, M.; BEACHY, R.N.; FAUQUET, C.M. A phage single-stranded DNA (ssDNA) binding protein complements ssDNA accumulation of a geminivirus and interferes with viral movement. **Journal of Virology**, v. 73, p. 1609-1616, 1999.

PENG, J.; GUO, L.; WANG, G. Development of a LAMP method for rapid detection of Banana streak virus. **Acta Phytopathologica Sinica**, v. 42, p. 2, 2012.

PHILLIPS, S.; BRIDDON, R. W.; BRUNT, A. A.; HULL, R. The partial characterization of a Badnavirus infecting the greater asiatic or water yam (*Dioscorea alata*). **Journal of Phytopathology**, v. 147, p. 265-269, 1999.

PIO-RIBEIRO, G.; MELO-FILHO, P. A.; KITAJIMA, E. M.; LIMA, S. K.; ANDRADE, G. P.; DOMINGOS, C. A. Detecção de potyvírus em inhame em áreas produtoras de Pernambuco e obtenção de matrizes in vitro. **Fitopatologia Brasileira**, v. 30, p. 183, 2005.

PITA, J. S.; FONDONG, V. N.; SANGARÉ, A.; OTIM-NAPE, G. W.; OGWAL, S.; FAUQUET, C. M. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. **Journal of General Virology**, v. 82, p. 655-665, 2001.

POOGGIN, M. M.; FÜTTERER, J.; SKRYABIN, K. G.; HOHN, T. A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. **Journal of General Virology**, v. 80, p. 2217-28, 1999.

POSNETTE, A.F. Virus diseases of cacao in West Africa. 1. Cacao viruses 1A, 1B, 1C and 1D. **Annals of Applied Biology**, v. 34, p. 388-402, 1947.

POSNETTE, A. F. Virus diseases of cacao in West Africa VII: Virus transmission by different vector species. **Annals of Applied Biology**, v. 37, p. 378-384, 1950.

QUAINOO, A.K.; WETTEN, A.C.; ALLAINGUILLAUME, J. Transmission of cocoa swollen shoot virus by seeds. **Journal of Virological Methods**, v. 150, n. 1, p. 45-49, 2008.

RAMOS-SOBRINHO, R.; XAVIER, C. A. D.; PEREIRA, H. M. B.; LIMA, G. S. A.; 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**, v. 95, p.2540-2552, 2014.

RICHERT- PÖGGELER, K. R.; NOREEN, F.; SCHWARZACHER, T.; HARPER, G.; HOHN, T. Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petúnia. **EMBO Journal**, v. 22, p. 4836-4845, 2003.

RIVAS, E.B.; DUARTE, L.M.L.; ALEXANDRE, M.A.V.; FERNANDES, F.M.C.; HAKAKAVA, R.; CHAGAS, C.M. A new Badnavirus species detected in *Bougainvillea* in Brazil. **Journal of General Plant Pathology**, v. 71, p. 438-40, 2005.

ROCHA, C. S.; CASTILLO-URQUIZA, G. P.; LIMA, A. T. M.; SILVA, F. N.; XAVIER, C. A. D.; HORA-JÚNIOR, B. T.; BESERRA-JÚNIOR, J. E. A.; MALTA, A. W. O.; MARTIN, D. P.; VARSANI, A.; ALFENAS-ZERBINI, P.; MIZUBUTI, E. S. G.; ZERBINI, F. M. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location., **Journal of General Virology** v. 87, p. 5784-5799, 2013.

SHACKELTON, L. A.; RAMBAUT, A.; PYBUS, O. G.; HOLMES, E. C. JC virus evolution and its association with human populations. **Journal of Virology**, v. 80, p. 9928-9933, 2006.

SHARMA, S. K.; KUMAR, P. V.; GEETANJALI, A. S.; PUN, K. B.; BARANWAL, V. K. Subpopulation level variation of banana streak viruses in India and common evolution of banana and sugarcane badnaviruses. **Virus genes**, v. 50, n. 3, p. 450-465, 2015.

SEAL, S.; TURAKI, A.; MULLER, E.; KUMAR, P.L.; KENYON, L.; FILLOUX, D.; GALZI, S.; LOPEZ-MONTES, A.; ISKRA-CARUANA, M. L. The prevalence of badnaviruses in West African yams (*Dioscoreacayenensis-rotundata*) and evidence of endogenous pararetrovirus sequences in their genomes. **Virus Research**, v. 186, p. 144-154, 2014.

SETHER, D. M.; MELZER, M. J.; BORTH, W. B.; HU, J. S. Pineapple bacilliform CO virus: Diversity, Detection, Distribution, and Transmission. **Plant Disease**, v. 96, n. 12, p. 1798-1884, 2012.

SILVA, J. M.; JOBIM, L. J.; RAMOS-SOBRINHO, R.; LIMA, J. S.; ASSUNÇÃO, I. P.; CRUZ, M. M.; LIMA, G. S. A. Incidence and species diversity of badnaviruses infecting sugarcane from a germplasm collection in Brazil. **Tropical Plant Pathology**, v. 40, p. 212-217, 2015.

SQUIRES, J.; GILLESPIE, T.; SCHOELZ, J.E; PALUKAITIS, P. Excision and episomal replication of cauliflower mosaic virus integrated into a plant genome. **Plant Physiology**, v. 155, p. 1908-1919, 2011.

STAGINNUS, C.; RICHERT-PÖGGELER, K.R. Endogenous pararetroviruses: two faced travelers in the plant genome. **Trends in Plant Sciences**, v. 11, p. 485-491, 2006.

STAGINNUS, C.; ISKRA-CARUANA, M.; LOCKHART, B.; HOHN, T.; RICHERT-PÖGGELER, K. R. Suggestions for a nomenclature of endogenous pararetroviral sequences in plants. **Archives of Virology**, v. 154, p. 1189-1193, 2009.

SVAROVSKAIA, E. S.; CHESLOCK, S. R.; ZHANG, W. H.; HU, W. S.; PATHAK, V. K. Retroviral mutation rates and reverse transcriptase fidelity. **Frontiers in Bioscience**, v. 8, p.117–134, 2003.

TEMIN, H. M. Reverse transcription in the eukaryotic genome: Retroviruses, pararetroviruses, retrotransposons, and retrotranscripts. **Molecular Biology and Evolution**, v. 2, p. 455-468, 1985.

THRESH, M. J. **The origin and epidemiology of some important plant virus diseases**. London: Academic Press. 1980.

UMBER, M.; FILLOUX, D.; MULLER, E.; LABOUREAU, N.; GALZI, S.; ROUMAGNAC, P.; ISKRA-CARUANA, M. L.; PAVIS, C.; TEYCHENEY, P. Y.; SEAL, S. The genome of African yam (*Dioscorea cayenensis-rotundata* complex) hosts endogenous

sequences from four distinct badnavirus species. **Molecular Plant Pathology**, v. 15, p. 790-801, 2014.

VEGA, J.; SORDI, R. Virus baciliforme em cana-de-açúcar detectado no Brasil através de Microscopia eletrônica de imunoabsorção. **Fitopatologia Brasileira**, v. 16, p. 26, 1991.

VERZIGNASSI, J. R.; MATOS, A. P. D.; SANTOS, M. D. F.; POLTRONIERI, L. S.; BENCHIMOL, R. L.; SANCHES, N. F. Mancha negra do abacaxi no Pará. **Summa Phytopathologica**, v. 35, n. 1, p. 76, 2009.

VISWANATHAN, R., ALEXANDER, K. C., GARG, I. D. Detection of sugarcane bacilliform virus in sugarcane germplasm. **Acta Virologica**, v. 40, n. 1, p. 5–8, 1996.

Virus Taxonomy: 2016 Release. **ICTV**. International Committee on Taxonomy of Viruses, 2016. Disponível em: <<https://talk.ictvonline.org/taxonomy>>. Acesso em: 10 jan. 2017.

WAKMAN, W.D.; TEAKLE, D.S.; THOMAS, J.E.; DIETZGEN, R. G. Presence of clostero-like virus and a bacilliform virus in pineapple plants in Queensland. **Australian Journal of Agricultural Research**, v. 46, p. 947-958, 1995.

YANG, I.C.; HAFNER, G.J.; DALE, J. L.; HARDING, R. M. Genomic characterization of taro bacilliform virus. **Archives of Virology**, v. 148, p. 937-949, 2003.

CAPÍTULO I

Diversification into the genus *Badnavirus*: phylogeny and population genetic variability

Caio H. L. de H. Ferreira¹, Lucas J. Jordão¹, Roberto Ramos-Sobrinho¹, Francisco M.

Zerbini², Sarah J. C. da Silva¹, Iraildes P. Assunção¹, Gaus S. A. Lima^{1*}

¹Setor de Fitossanidade/Centro de Ciências Agrárias, Universidade Federal de Alagoas, Rio Largo, AL, Brazil, 57100-000

²Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, Brazil, 36570-900

* To whom correspondence should be addressed.

Email: gaus@ceca.ufal.br

Email: robertorsobrinho@gmail.com

Phone: (+55-82) 3261-3470; Fax: (+55-82) 3261-1351

Abstract

The family *Caulimoviridae* comprises viruses with semicircular double-stranded DNA genomes encapsulated into isometric or bacilliform particles, being divided into eight genera. The genus *Badnavirus* is the most important due to its high number of species reported infecting cultivated plants worldwide. The present study aimed to evaluate the phylogenetic positioning and population genetic variability into the genus *Badnavirus*. Data sets comprising badnavirus full-length genome and partial RT/RNaseH sequences were retrieved from the non-redundant GenBank database. Multiple nucleotide sequence alignments were obtained for the data sets: complete genome, ORFIII, full (1020pb) and partial (579pb) RT/RNaseH domains. A total of 127 genome sequences were obtained, representing 53 badnavirus species. Nucleotide pairwise comparisons for RT/RNaseH showed that only a few isolates reported as distinct *Badnavirus* species shared $\geq 80\%$ identity, the current threshold used for species demarcation into this genus. Bayesian phylogenetic trees for the complete genome and ORFIII showed four well supported clusters (badnavirus groups 1-4), with clusters 1 and 3 being sister groups comprising predominantly species/isolates infecting sugarcane and banana. Non-tree-like evolution evidenced putative recombination events affecting the *Badnavirus* diversification, and at least 23 independent events were detected with recombination breakpoints occurring mostly in the ORFIII and intergenic region. High levels of nucleotide diversity were observed for the partial RT/RNaseH region in populations of 11 distinct badnaviruses. These results showed that mutation and recombination are important mechanisms acting on the evolution of badnaviruses, and that the partial RT/RNaseH sequences are sufficient for species demarcation of most viruses reported in this genus.

Key words: recombination, evolution, species demarcation.

Introduction

Viruses belonging to the family *Caulimoviridae* have semicircular, double-stranded (ds)DNA genomes, 7.2-9.2 kbp in length, encapsidated into isometric or bacilliform particles, and which replicate through an RNA intermediate (plant pararetroviruses; Geering and Hull 2012). This family is divided into eight genera (*Badnavirus*, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Rosadnavirus*, *Solendovirus*, *Soymovirus* and *Tungrovirus*) according to host range, insect vector, genomic organization and phylogeny (Bath et al. 2016; Geering and Hull 2012). Badnaviruses are transmitted mostly by mealybugs (a few species by aphids) in a semi-persistent manner (Bath et al. 2016; Geering and Hull 2012) and are among the most important plant viruses with a DNA genome.

The semicircular dsDNA of badnaviruses has site-specific discontinuities and at least three ORFs, named I, II and III (Bouhida et al. 1993; Geering and Hull 2012; Hagen et al. 1993; Harper and Hull 1998). Proteins codified by ORFs I and II have been reported to be virion-associated (Cheng et al. 1996) and nucleic acid-binding (ORF II; Jacquot et al. 1996). ORFIII encodes a polyprotein of 208-216 kDa that is proteolytically cleaved generating the movement and coat proteins, the aspartate protease responsible for the polyprotein cleavage, the reverse transcriptase (RT) and ribonuclease H (RNaseH) (both genomic domains involved in the viral replication) (Harper and Hull 1998; Medberry et al. 1990). The criterion of 80% nucleotide sequence identity for the RT/RNaseH domains was established for species demarcation in the genus *Badnavirus* (Geering and Hull 2012), and specific primer pairs are largely used to amplify this viral genomic region (Yang et al. 2003). However, different studies have shown this criterion to be insufficient to separate some badnavirus species, mostly those infecting banana and sugarcane (Karuppaiah et al. 2013; Muller et al. 2011;

Silva et al. 2015). Furthermore, the existence of endogenous badnavirus sequences represents a great challenge for taxonomy and diagnosis of members into this genus.

Different badnavirus species have been reported infecting economically important crops such as sugarcane (*Saccharum* spp.), banana (*Musa* spp.), cacao (*Theobroma cacao* L.), black pepper (*Piper* spp.), yam (*Dioscorea* spp.) and citrus (*Citrus* spp.) (Deeshma and Bhat 2015; Eni et al. 2008; James et al. 2011; Johnson et al. 2012; Kouakou et al. 2012; Silva et al. 2015), with 40 badnavirus species being currently recognized by the International Committee on Taxonomy of Viruses (<https://talk.ictvonline.org/taxonomy/>). Badnaviruses have also been reported as integrated sequences into their host genomes, commonly referred to as endogenous pararetroviruses (EPRVs; Hohn et al. 2008; Staginnus et al. 2009), and it is known that some EPRVs are able to ‘scape’ from the host genome and establish a systemic infection (episomal forms; Côte et al. 2010; Dallot et al. 2001; Lheureux et al. 2003). Besides this high species diversity, badnavirus populations also show high levels of molecular variability (Geering et al. 2000; Geijskes et al. 2002; Govind et al. 2014; Guimarães et al. 2015; Harper et al. 2005; Silva et al. 2015). The lack of a proofreading activity of the badnavirus reverse transcriptase (Svarovskaia et al. 2003) and the occurrence of recombination events (Govind et al. 2014) may directly affect the high genetic variability observed in this viral group.

A *in silico* large-scale study was carried out to obtain more information about the genetic relationship and variability in *Badnavirus*. The threshold for nucleotide sequence comparisons of the RT/RNaseH genomic region currently used for species demarcation identifies most reported badnaviruses. However, this criterion alone is unable to differentiate all sugarcane- and banana-infecting badnaviruses. A new badnavirus phylogenetic clade was proposed, here named badnavirus group 4. Additionally, it was observed this is a highly

diverse viral group, with recombination and mutation being important factors contributing to high levels of nucleotide diversity observed in some badnavirus populations.

Material and methods

Badnavirus data set

Full-length genome sequences of badnaviruses were retrieved from the non-redundant GenBank database (www.ncbi.nlm.nih.gov/genbank; accessed on Nov 2017) (Table 1). Data sets of nucleotide sequences of the ORFIII, and full (1020pb) and partial (579pb) RT/RNaseH domains were obtained from the complete genomes. The partial RT/RNaseH data set comprises the genomic region amplified by the primer pairs largely used for detection and identification of badnaviruses (Yang et al. 2003). A previous analysis using sequences from ORF I and II showed that these regions are inconclusive, presenting low support and insufficient phylogenetic signal, for this reason were not included in this study (data not shown).

Sequence analysis

Multiple amino acid sequence alignments were prepared for the ORFIII using the MUSCLE algorithm (Edgar 2004), manually edited in the MEGA7 package (Kumar, Stecher, Tamura 2015) and returned to nucleotide sequences for posterior analyses. The RT/RNaseH sequences (full and partial) were obtained from the ORFIII data set. Additionally, multiple nucleotide alignments were obtained for the complete genome. To confirm taxonomy attributed to the badnavirus isolates retrieved from GenBank, pairwise nucleotide sequence comparisons were performed to all data sets using Sequence Demarcation Tool (SDT) v.1.2 (Muhire et al. 2013).

Phylogenetic inference

In order, to demonstrate if the phylogenetic relationship observed for the RT/RNaseH reflects the clustering inferred for the complete genomes of badnaviruses, Bayesian phylogenetic trees were obtained for the complete genome, ORFIII and RT/RNaseH (full and partial) data sets. Analyses were run using MrBayes v. 3.2 (Ronquist et al. 2012) through the CIPRES web portal (Miller et al. 2010), assuming GTR+I+G as the evolutionary model. Two replicates with four chains each for 20 million generations and sampling every 2,000 generations were used. The first 2,500 trees were discarded as a burn-in phase in each run. Posterior probabilities (Rannala and Yang 1996) were determined from a majority-rule consensus tree generated with the 15,000 remaining trees. The trees were edited in FigTree v.1.4 (ztree.bio.ed.ac.uk/software/figtree) and Inkscape (<https://inkscape.org/pt/>).

Recombination analysis

Evidence of non-tree-like evolution was assessed for the complete genome, ORFIII, and RT/RNaseH (full and partial) data sets using the Neighbor-Net method implemented in SplitsTree v.4.10 (Huson and Bryant 2006). Putative parental sequences and recombination breakpoints for the complete genome data set were determined using the methods RDP, Geneconv, Boot-scan, Maximum Chi Square, Chimaera, SisterScan and 3Seq implemented in the RDP v.4.0 package (Martin et al. 2015). Alignments were analyzed with default settings for the different methods and statistical significance was inferred by a *P*-value lower than a Bonferroni-corrected cut-off of 0.05. Only events detected by at least five different methods were considered to be reliable.

Population genetic variability

Partial nucleotide sequences of the RT/RNaseH region of badnaviruses infecting different hosts were retrieved from the non-redundant GenBank database (www.ncbi.nlm.nih.gov/genbank; accessed on Dez 2017). The mean pairwise number of nucleotide differences per site (nucleotide diversity, π) was estimated for each population using DnaSP v. 6.10 (Rozas et al. 2017).

Results

Badnavirus isolates

A total of 127 full-length genomes were obtained from GenBank (Table 1), comprising 53 different badnavirus species. *Cacao swollen shoot virus* (CSSV) was the badnavirus represented by the higher number of sequences/isolates (12), while 25 species were represented by one only sequence (Table 1).

Species demarcation

Pairwise sequence comparisons for the partial (579pb) and full (1020pb) RT/RNaseH, largely used for badnavirus species identification, showed percent nucleotide identities ranging from 57.2 to 82.4% and 57.1 to 83.8%, respectively, among species (Supplementary Material). Therefore, some comparisons exceeded the currently used 80% nucleotide identity criterion for species demarcation into the genus *Badnavirus* (Geering and Hull 2012).

According to the partial RT/RNaseH region, *Sugarcane bacilliform GA virus* (SCBGAV) isolates shared 80.5-82.0% and 81.3-82.4% identity with *Banana streak OL virus* (BSOLV) and *Banana streak CA virus* (BSCAV) isolates, respectively, with BSOLV

and BSCAV showing up to 77.7% identity. *Sugarcane bacilliform BB virus* (SCBBBV) showed 77.4-80.4% and 78.3-82.0% nucleotide identity with *Sugarcane bacilliform IM virus* (SCBIMV) and *Sugarcane bacilliform MO virus* (SCBMOV), respectively, while SCBIMV and SCBMOV isolates shared 77,9-79,8% identity. To the full RT/RNaseH, the percentages of nucleotide identity were slightly increased: SCBGAV shared 81.0-82.4% and 81.9-82.5% identity with BSOLV and BSCAV, respectively, with BSOLV and BSCAV showing up to 79.0% identity. SCBBBV showed 78.7-81.2% and 78.8-81.9% identity with SCBIMV and SCBMOV, respectively, while SCBIMV and SCBMOV isolates shared 78.4-80.9% identity.

The *Sweet potato badnavirus B* (SPBV) and *Sweet potato pakakuy virus* (SPPV) isolates shared 81.9% and 83.8% of nucleotide identity for the partial and full RT/RNaseH data sets, respectively. For all badnavirus represented by more than one sequence/isolate, percent nucleotide identities were higher than 80.0% within species.

When analyzed the ORFIII data set, which comprises the RT/RNaseH domains, SCBGAV showed highest nucleotide identity of 80.3% and 79.8% with BSOLV and BSCAV, respectively. However, SCBBBV, SCBIMV and SCBMOV shared up to 75,7% identity. For the complete genome, all pairwise comparisons between distinct badnaviruses were lower than 80% identity, with SCBGAV isolates showing the highest level of nucleotide identity (79.5%) with BSOLV and BSCAV isolates (Supplementary Material).

Phylogenetic analysis

In the complete genome Bayesian phylogenetic tree, the three badnavirus clusters (badnavirus groups 1, 2 and 3) described by Muller et al. (2011) were observed (Figure 1). Additionally, a fourth clade can be observed, here named badnavirus group 4 (Figure 1). The clusters 1 and 3 formed sister groups, being predominantly comprised by badnaviruses

infecting sugarcane and banana (Figure 1). Similar results were observed for the ORFIII data set (Figure 2), which represents ~80% of the complete genome. These results reinforce the idea that badnaviruses infecting sugarcane and banana are closely related, as indicated by the pairwise comparisons and recombination analyses (see below).

When analyzed the phylogenetic trees for the RT/RnaseH data sets (partial and full), badnavirus groups are still evidenced, but with many topological incongruences and very low resolution (Figures 3 and 4). Some species (mainly the cacao-infecting isolates) in the badnavirus group 2 clustered with isolates in groups 3 and 4, while the other group 2 isolates clustered with badnaviruses in group 1 (Figure 3). However, statistical support was considerably smaller for trees based on RT/RnaseH sequences than ORFIII and complete genome, indicating the RT/RNaseH sequences have insufficient phylogenetic signal.

Recombination events

Non-tree-like evolution for the complete genome, ORF III and RT/RNaseH revealed evidence of putative recombination events affecting the evolution of badnaviruses, with both intra- and interspecies recombination being observed (Figures 5-8). These events were more pronounced among badnaviruses infecting sugarcane and banana (phylogenetic groups 1 and 3), and cacao (phylogenetic group 2). Besides recombination, a strong mutation effect was evidenced in the diversification of members into this genus, indicated by the long branches associated with many isolates (Figures 5-8).

To investigate putative parental sequences and recombination breakpoints, the complete genome data set was analyzed using the RDP4 package. Based on a stringent set of criteria, at least 23 independent recombination events were detected among badnavirus isolates, with ten of them involving species associated with sugarcane and banana (Table 2).

Additional recombination events involved viruses infecting citrus, cacao, black pepper, pineapple and grapevine. Most recombination breakpoints were located in ORF III and in the intergenic region (Table 2).

Population genetic variability

A total of 299 RT/RNaseH partial sequences were obtained for 11 distinct badnavirus species, and the nucleotide diversity of each dataset/population was calculated (Table 3). Additionally, 96 *Dioscorea bacilliform AL virus* (DBALV) partial sequences from Guimarães et al. (2015) were added.

High values of per-site nucleotide diversity were observed for *Banana streak UA virus* (BSUAV), *Banana streak UL virus* (BSULV), *Banana streak UM virus* (BSUMV), *Piper yellow mottle virus* (PYMoV), SPBV and SCBIMV, suggesting high genetic variability for the partial RT/RNaseH region in different badnavirus species/populations (Table 3). The most variable sequence data sets were BSULV ($\pi = 0.1467$), PYMoV ($\pi = 0.0944$) and SCBIMV ($\pi = 0.0876$), being higher than the previously reported DBALV ($\pi = 0.0534$).

Discussion

Badnaviruses have been reported infecting economically important crops worldwide (Deeshma and Bhat 2015; Eni et al. 2008; James et al. 2011; Johnson et al. 2012; Kouakou et al. 2012; Silva et al. 2015), causing great losses in cacao (Kouakou et al. 2012) and banana (Dahal et al. 2000; Daniells et al. 2001). They comprise a highly diversified group, and recent studies have shown that badnavirus populations exhibit high levels of genetic variability (Guimarães et al. 2015; Karuppaiah et al. 2013; Muller et al. 2011; Silva et al. 2015). Nucleotide sequence comparisons of the partial RT/RNaseH domains (579pb) and a

threshold of $\geq 80\%$ identity have largely been used for species demarcation in the genus *Badnavirus* (Geering and Hull 2012; Guimarães et al. 2015; Muller et al. 2011; Silva et al. 2015), with universal primers described by Yang et al. (2003) being very efficient to amplify this genomic region. However, the existence of closely related badnavirus species which share more than 80% nucleotide identity in that region suggests this criterion needs to be revised (Harper et al. 2005; Karuppaiah et al. 2013; Muller et al. 2011; Silva et al. 2015).

In the present study, nucleotide sequence comparisons of a large sequence data set showed that the threshold of $\geq 80\%$ identity for the RT/RNaseH genomic region [partial (579pb) or full (1020pb)] allowed the species demarcation of most badnavirus isolates with full-length genome sequences available in GenBank. However, as previously reported, it was not possible to distinguish a few badnaviruses from banana and sugarcane (Karuppaiah et al. 2013; Muller et al. 2011; Silva et al. 2015), even when analyzed the entire ORF III (which corresponds to $\sim 80\%$ of the badnavirus genome).

We suggest a few alternates to tentatively solve the identification problems observed for badnaviruses infecting banana and sugarcane. First, maintaining the criterion of $\geq 80\%$ nucleotide sequence identity for the RT/RNaseH genomic region, closely related banana streak viruses (BSV) and sugarcane bacilliform viruses (SCBV) species sharing $\geq 80\%$ nucleotide sequence identity should be considered as different strains of a same species. Taking account biological aspects as host range (Geering and Hull 2012), cross infection involving sugarcane and banana-infecting badnaviruses has been observed (Lockhart and Autrey, 1988; Bouhida et al., 1993; Jones and Lockhart, 1993), which reinforces the idea SCBV and BSV sharing more than 80% identity may belong to a same viral species. Additionally, the isolates SPPV (#access FJ560943) and SPBVB (#access FJ560944), previously reported as distinct species, must belong to a same badnavirus species as they

share more than 80% identity for RT/RNaseH and the same host range, besides their narrow phylogenetic relationship. Second, a new threshold for species demarcation based on nucleotide identity of the RT/RNaseH domain could be established. In our analysis, it was observed the badnavirus species showed no more than 82,5% nucleotide identity for the RT/RnaseH domain, and therefore higher values could be adopted as the new threshold for species demarcation. Third, considering the inefficiency to differentiate some badnavirus species using the currently established criteria, full-length genome sequences could be used for taxonomy.

Phylogenetic analysis showed clear admixture of SCBVs and BSVs isolates into at least two distinct badnavirus clusters (*i.e.*, badnavirus groups 1 and 3 *sensu* Muller et al., 2011). The close genetic relationship among sugarcane and banana-infecting badnaviruses, and the polyphyletic structure of these viruses (Gayral and Iskra-Caruana 2009; Muller et al. 2011), strongly support the hypothesis of a host shift, although it is not possible to determine whether sugarcane or banana was the original host (Gayral and Iskra-Caruana 2009; Muller et al. 2011).

The phylogenetic relationships observed for all data sets analyzed, in which SCBVs and BSVs isolates are grouped in sister clusters, agree with results of pairwise sequence comparisons. Besides the three well defined badnavirus groups reported by Muller et al. (2011), a fourth phylogenetic cluster is proposed here, named badnavirus group 4. In the phylogeny based on full RT/RNaseH, the groups 1 and 3 (both composed predominantly by sugarcane and banana-infecting viruses) are more distantly related, with group 3 being closer to group 4. However, in phylogenies based either on ORFIII or complete genomes, a close relationship between groups 1 and 3 was evident, so they can be considered as sister groups.

These results reinforce the hypothesis that sugarcane and banana-infecting badnaviruses have a common evolutionary history.

The high levels of nucleotide diversity for the partial RT/RNaseH region in some badnavirus species (BSUAL, BSULV, BSUMV, PYMoV, SPBVB, SCBIMV and DBALV) are comparable to values observed for RNA viruses (Guimarães et al 2015). However, this variability is lower when compared to less conserved regions [ORFs I, II and intergenic region (IGR)], with IGR showing the highest values of genetic diversity (0.468) followed by ORFII 0.367; Sharma et al., 2015). These results suggest although the partial RT/RNaseH region be variable in different badnavirus populations, non-conserved regions tend to have higher nucleotide diversity.

The high genetic variability observed for badnaviruses has been attributed to error-prone replication by their reverse transcriptase (Bousalem et al. 2008). Reverse transcriptases (RT) are known to produce errors in retroviruses and retroelements for which the fidelity rates have been estimated (Svarovskaia et al. 2003). Although fidelity rates of caulimovirid RTs have not been estimated, it is believed that the lack of proofreading activity contributed to the high mutation rates observed for these viruses (Svarovskaia et al. 2003). Nevertheless, the contribution of recombination to the genetic diversity of badnaviruses must also be considered (Govind et al. 2014).

One of the basic assumptions for successful recombination is the occurrence of mixed infections, with the presence of the viruses in the same host cell (García-Andrés et al. 2006; Graham et al. 2010; Zhou et al. 1997). Cross infection of badnaviruses seems to be an uncommon event, however it has been reported for viruses infecting sugarcane and banana, which could explain the putative recombinant origin of SCBV and BSV isolates (Bath et al. 2016; Bouhida et al. 1993; Lockhart and Autrey 1988; Sharma et al., 2015). Here, besides

the putative recombination events affecting the evolution of SCBV and BSV isolates, it also seems an important evolutionary mechanism for diversification of cacao-infecting badnaviruses.

Nucleotide sequence comparisons of the partial RT/RNaseH region are sufficient for species demarcation of most badnaviruses currently known. However, the $\geq 80\%$ nucleotide identity threshold alone is unable to differentiate all SCBV and BSV species and should be reviewed. Finally, mutation and putative recombination events may be involved with the high levels of genetic variability and diversity observed in *Badnavirus*.

REFERENCES

- Bhat AI, Hohn T, Selvarajan R (2016) Badnaviruses: The Current Global Scenario. *Viruses* 177:1-29.
- Bouhida ML, Lockhart BE, Olszewski NE (1993) An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *Journal of General Virology* 74:15-22.
- Bousalem M, Douzery EJP, Seal SE (2008) Molecular taxonomy, phylogeny, and evolution of plant reverse transcribing viruses (Caulimoviridae) inferred from the reverse transcriptase sequences. *Archives of Virology* 153:1085-1102.
- Cheng CP, Lockhart BE, Olszewski NE (1996) The ORF I and II Proteins of Commelina yellow mottle virus are virion-associated. *Virology* 223:263-271
- Côte FX, Galzi S, Folliot M, Lamagnère Y, Teycheney PY, Iskra-Caruana ML (2010) Micropropagation by tissue culture triggers differential expression of infectious endogenous Banana streak virus sequences (eBSV) present in the B genome of natural and synthetic interspecific banana plantains. *Molecular Plant Pathology* 11:137-144.
- Dahal G, Ortiz R, Tenkouano A, Hughes Jd'A, Thottappilly G, Vuylsteke D, Lockhart BEL (2000) Relationship between natural occurrence of banana streak badnavirus and symptom expression, relative concentration of viral antigen, and yield characteristics of some micropropagated Musa spp. *Plant Pathology* 49:68-79.
- Dallot S, Accuna P, Rivera C, Ramirez P, Cote F, Lockhart BEL, Caruana ML (2001) Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of Banana streak virus integrated into the genome of the FHIA21 hybrid (Musa AAAB). *Archives of Virology* 146:2179-2190.
- Daniells JW, Geering ADW, Bride NJ, Thomas JE (2001) The effect of banana streak virus on the growth and yield of dessert bananas in tropical Australia. *Annals of Applied Biology* 139:51-60.
- Deeshma KP, Bhat AI (2015) Complete genome sequencing of Piper yellow mottle virus infecting black pepper, betelvine, and Indian long pepper. *Virus Genes* 50:172–175.
- Edgar RC (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:1–19.
- Eni AO, Hughes Jd, Asiedu R, Rey ME (2008) Sequence diversity among Badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin and Nigeria. *Archives of Virology* 153:2263-2272.

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.

Gayral P, Iskra-Caruana M (2009) Phylogeny of Banana Streak Virus Reveals Recent and Repetitive Endogenization in the Genome of Its Banana Host (*Musa* sp.). *Journal of Molecular Evolution* 69:65-80.

Graham AP, Martin DP, Roye ME (2010) Molecular characterization and phylogeny of two begomoviruses infecting *Malvastrum americanum* in Jamaica: Evidence of the contribution of inter-species recombination to the evolution of malvaceous weed-associated begomoviruses from the Northern Carribean. *Virus Genes* 40:256-266.

Geering ADW, Hull R (2012) Family Caulimoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (Eds.). *Virus Taxonomy*. 9th Report of the International Committee on Taxonomy of Viruses. London, UK. Elsevier Academic Press. pp. 429-443.

Geering ADW, Mcmichael LA, Dietzgen RG, Thomas JE (2000) Genetic diversity among Banana streak virus isolates from Australia. *Phytopathology* 90:921-927.

Geijskes RJ, Braithwaite KS, Dale JL, Harding RM, Smith GR (2002) Sequence analysis of an Australian isolate of sugarcane bacilliform badnavirus. *Archives of Virology* 147:2393-2404.

Govind PR, Susheel KS, Deepti S, Meenakshi A, Priyanka S, Virendra KB (2014) Genetically Diverse Variants of Sugarcane bacilliform virus Infecting Sugarcane in India and Evidence of a Novel Recombinant Badnavirus Variant. *Journal of Phytopathology* 162:779–787.

Guimarães KM, Silva SJC, Melo AM, Ramos-Sobrinho R, Lima JS, Zerbini FM, Assunção IP, Lima GSA (2015) Genetic variability of badnaviruses infecting yam (*Dioscorea* spp.) in northeastern Brazil. *Tropical Plant Pathology* 40:111-118.

Hagen LS, Jacquemond M, Lepingle A, Lot H, Tepfer M (1993) Nucleotide sequence and genomic organisation of cacao swollen shoot virus. *Virology* 196:619-628.

Harper G, Hart D, Moulton S, Hull R, Geering A, Thomas J (2005) Diversity of banana streak virus isolates in Uganda. *Archives of Virology* 150:2407-2420.

Harper G, Hull R (1998) Cloning and sequence analysis of Banana Streak Virus DNA. *Virus Genes* 17:271-278.

Hohn T, Richert-Poggeler KR, Harper G, Schwarzacher T, Teo C, Teycheney PY, Iskra-Caruana ML, Hull R (2008) Evolution of integrated plant viruses. In: Roossinck M (Eds.) *Plant Virus Evolution*. Heidelberg, Germany. Academic Springer. pp. 54–76.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23:254–267.

Jacquot E, Hagen LS, Jacquemond M, Yot P (1996) The open reading frame 2 product of cacao swollen shoot badnavirus is a nucleic acid-binding protein. *Virology* 225:191-195.

James AP, Geijskes RJ, Dale JL, Harding RM (2011) Molecular characterization of six badnavirus species associated with leaf streak disease of banana in East Africa. *Annals of Applied Biology* 158:346-353.

Johnson AM, Borah BK, Sai Gopal DV, 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.

Jones DR, Lockhart BEL (1993) Banana streak disease. Musa fact sheet No. 1, International network for improvement of banana and plantain, Montpellier, France.

Karuppaiah R, Viswanathan R, Kumar VG (2013) Genetic diversity of Sugarcane bacilliform virus isolates infecting Saccharum spp. in India. *Virus Genes* 46:505–516.

Kouakou K, Kébé BI, Kouassi N, Muller E (2012) Geographical distribution of Cacao swollen shoot virus molecular variability in Côte d’Ivoire. *Plant Disease* 96:1445-1450.

Kumar S, Stecher G, Tamura K (2015) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*.

Lheureux F, Carreel F, Jenny C, Lockhart BEL, Iskra-Caruana ML (2003) Identification of genetic markers linked to banana streak disease expression in inter-specific Musa hybrids. *Theoretical and Applied Genetics* 106:594-598.

Lockhart BEL, Autrey LJC (1988) Occurrence in sugarcane of a bacilliform virus related serologically to banana streak virus. *Plant Disease* 72: 230-233.

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* 1:1-5.

Medberry SL, Lockhart BEL, Olszewski NE (1990) Properties of Commelina yellow mottle virus complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. *Nucleic Acids Research* 18:5505-5513.

Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T (2010) The CIPRES portals. CIPRES. <http://www.phylo.org/sub-sections/portal> (Accessed 15 Aug 2016).

Muller E, Dupuy V, Blondin L, Bauffe F, Daugrois JH, Laboureau N, Iskra Caruana ML (2011) High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. *Virus Research* 160:414-419.

Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A (2013) A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). *Archives of Virology* 158:1411-1424.

Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution* 43:304-311.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Systematic Biology Advance Access* 61:539-42.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Orsins SE, Sánchez-Gracia A. (2017) DnaSP v6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution*, 34(12): 3299-3302.

Sharma, S. K., Kumar, P. V., Geetanjali, A. S., Pun, K. B., & Baranwal, V. K. (2015). Subpopulation level variation of banana streak viruses in India and common evolution of banana and sugarcane badnaviruses. *Virus genes*, 50(3), 450-465.

Silva JM, Jobim LJ, Ramos-Sobrinho R, Lima JS, Assunção IP, Cruz MM, Lima GSA (2015) Incidence and species diversity of badnaviruses infecting sugarcane from a germplasm collection in Brazil. *Tropical Plant Pathology* 40:212-217.

Staginnus C, Iskra-Caruana M, Lockhart B, Hohn T, Richert-Pöggeler KR (2009) Suggestions for a nomenclature of endogenous pararetroviral sequences in plants. *Archives of Virology* 154:1189-1193.

Svarovskaia ES, Cheslock SR, Zhang WH, Hu WS, Pathak VK (2003) Retroviral mutation rates and reverse transcriptase fidelity. *Frontiers in Bioscience* 8:117-134.

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

Zhou X, Liu Y, Calvert L, Munoz C, Otim-Nape GW, Robinson DJ, Harrison BD (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology* 78:2101-2111.

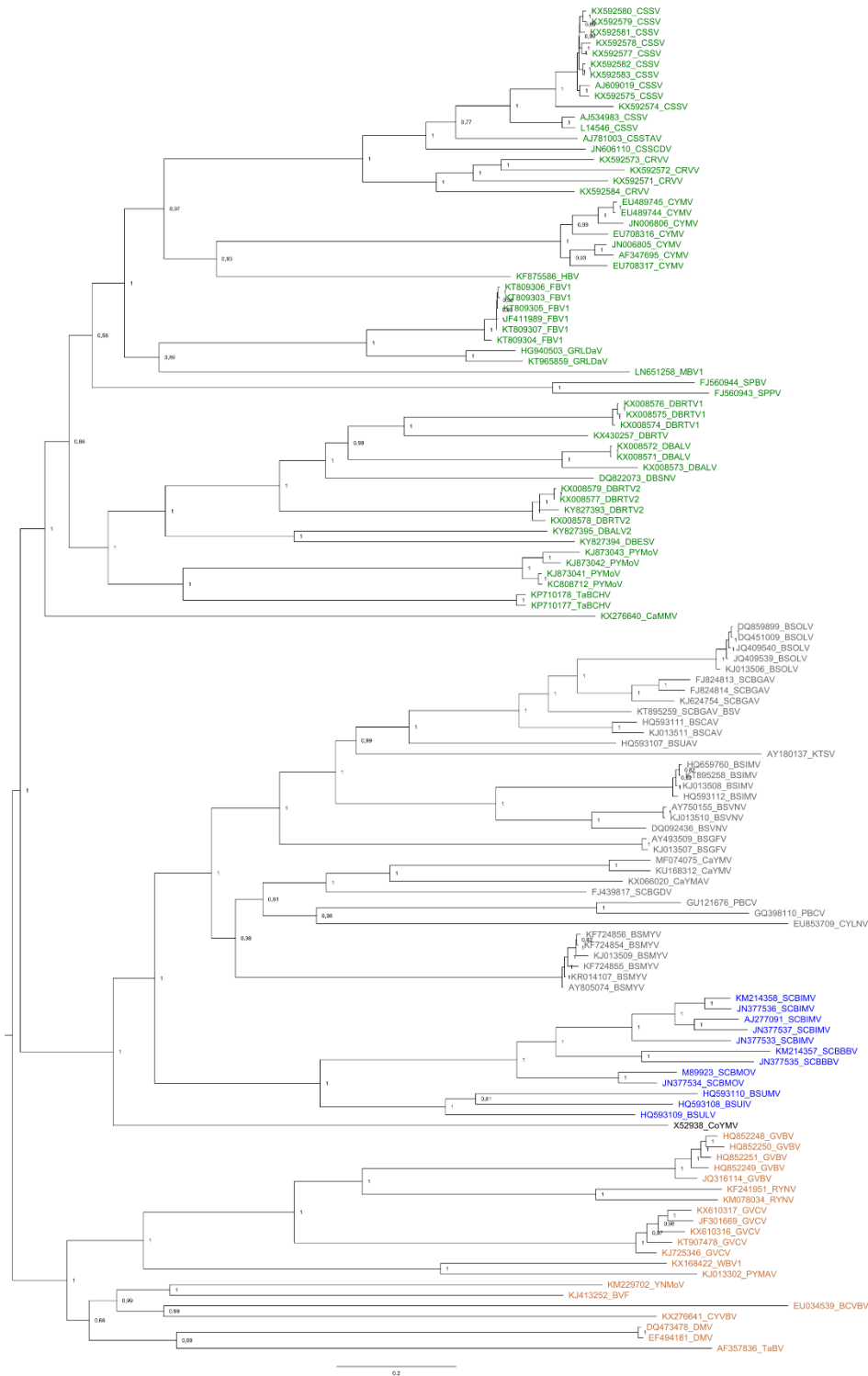


Figure 1. Midpoint-rooted Bayesian phylogenetic trees based on complete genome nucleotide sequences of badnaviruses, indicating badnavirus groups 1 (in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV) (not highlighted) was placed apart from the four main clusters.

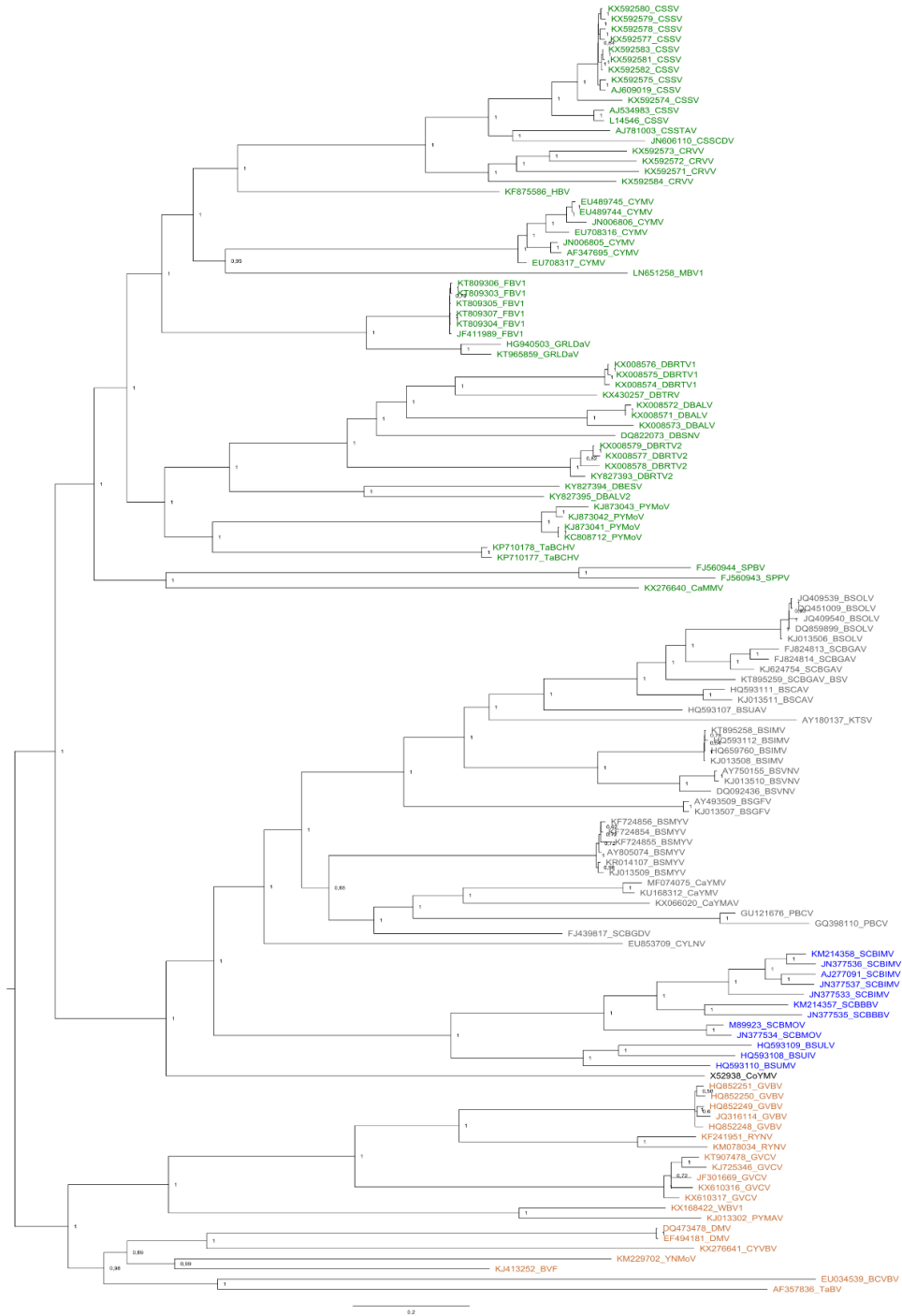


Figure 2. Midpoint-rooted Bayesian phylogenetic trees based on ORF 3 nucleotide sequences of badnaviruses, indicating badnavirus groups 1 (in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV) (not highlighted) was placed apart from the four main clusters.

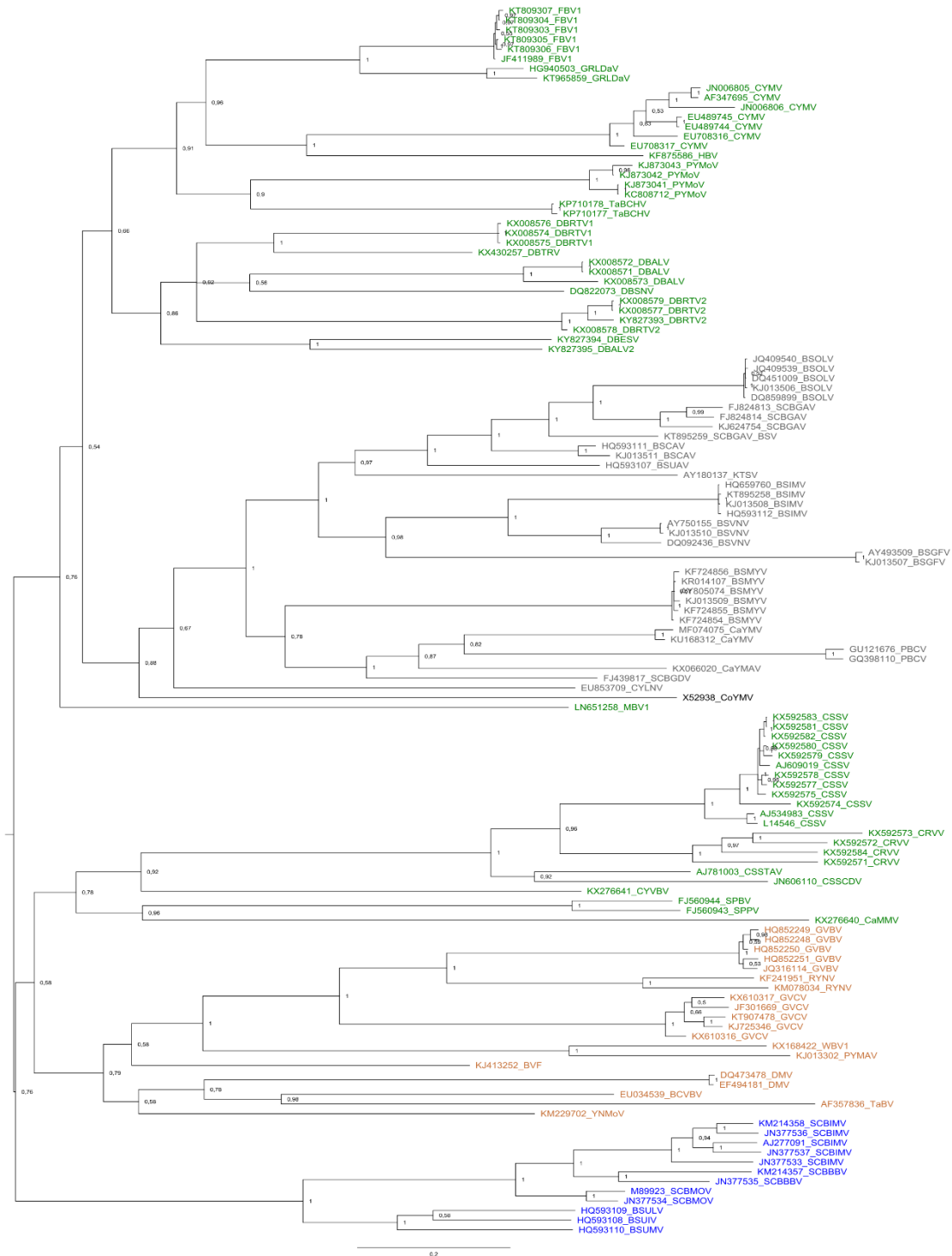


Figure 3. Midpoint-rooted Bayesian phylogenetic trees based on complete RT/RNaseH nucleotide sequences of badnaviruses, indicating badnavirus groups 1 (in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV) (not highlighted) was placed apart from the four main clusters.

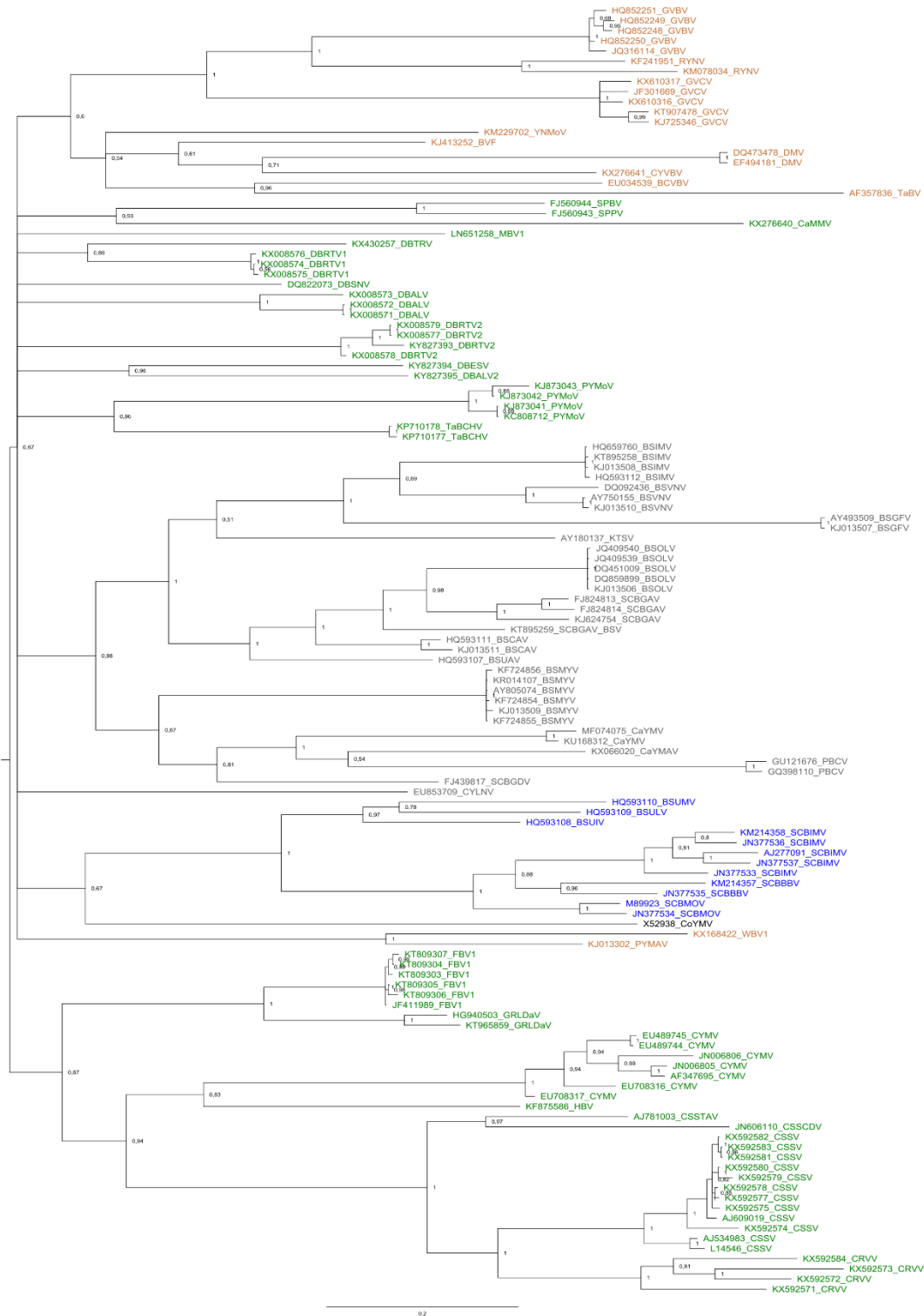


Figure 4. Midpoint-rooted Bayesian phylogenetic trees based on partial RT/RNaseH nucleotide sequences of badnaviruses, indicating badnavirus groups 1 (in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV) (not highlighted) was placed apart from the four main clusters.

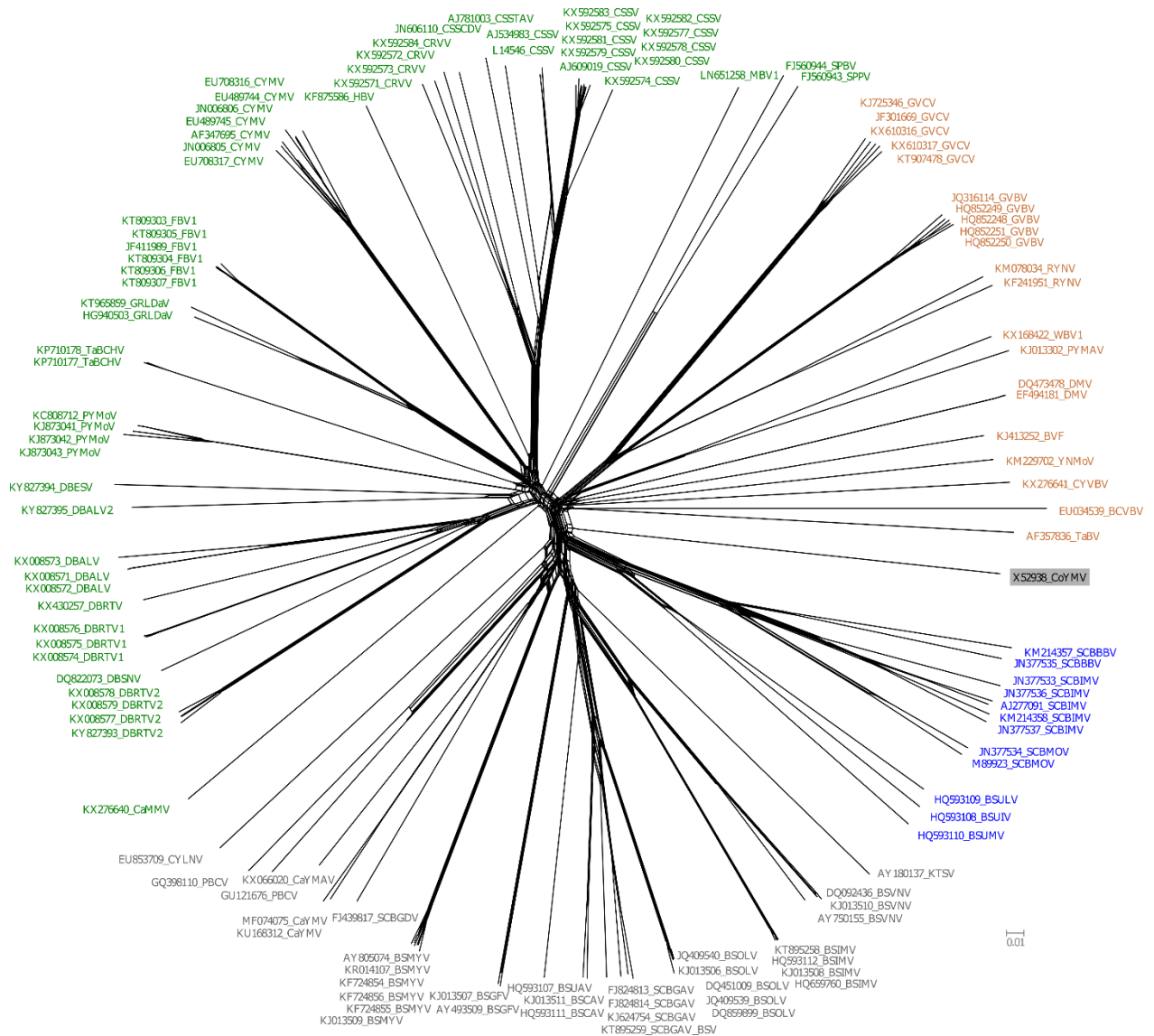


Figure 5. Non-tree-like evolution based on the complete genome nucleotide sequences of badnaviruses. Badnavirus groups 1 (indicated in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV, highlighted in grey, was placed apart from the four main groups.

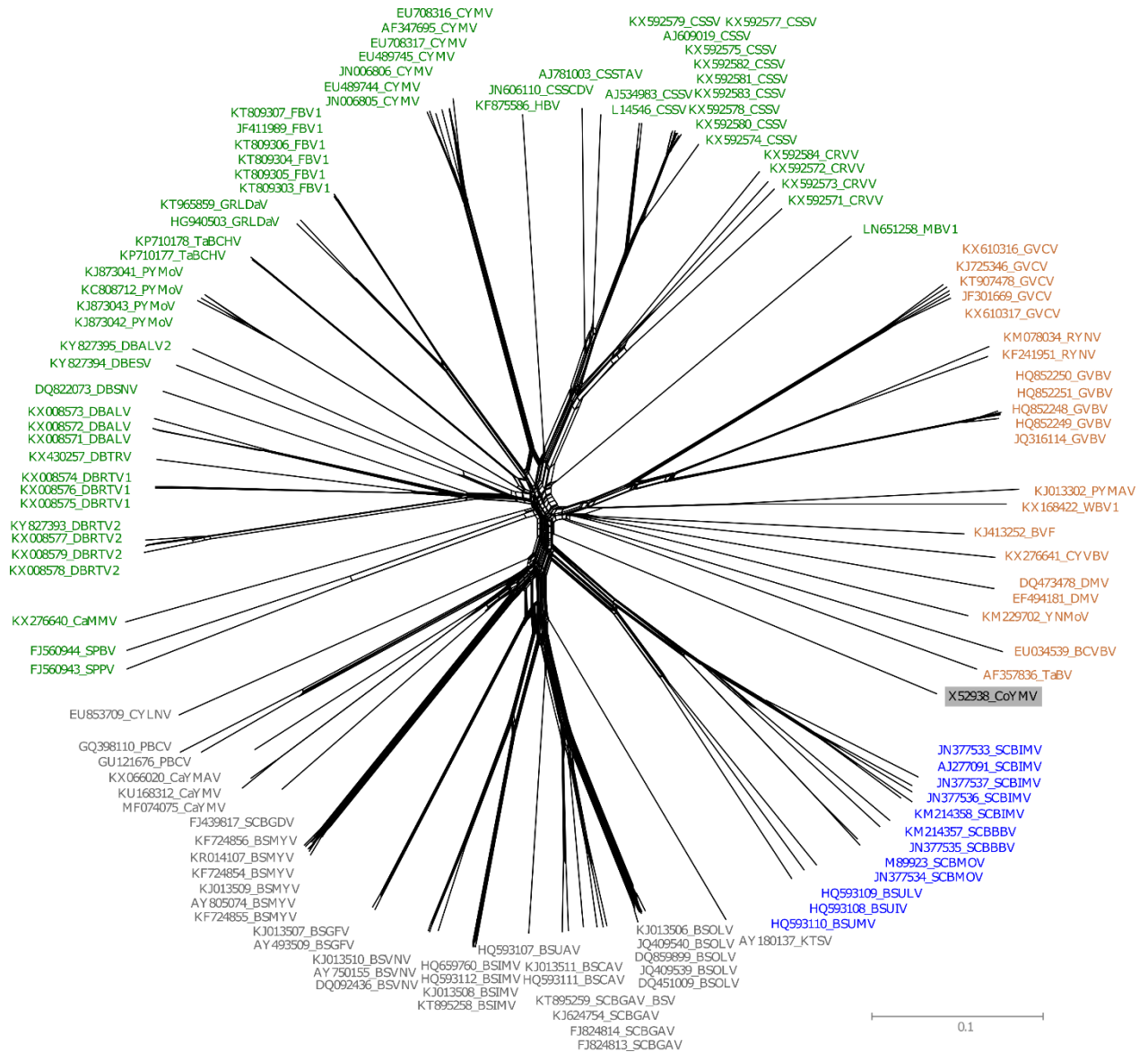


Figure 6. Non-tree-like evolution based on the ORFIII nucleotide sequences of badnaviruses. Badnavirus groups 1 (indicated in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV, highlighted in grey, was placed apart from the four main groups.

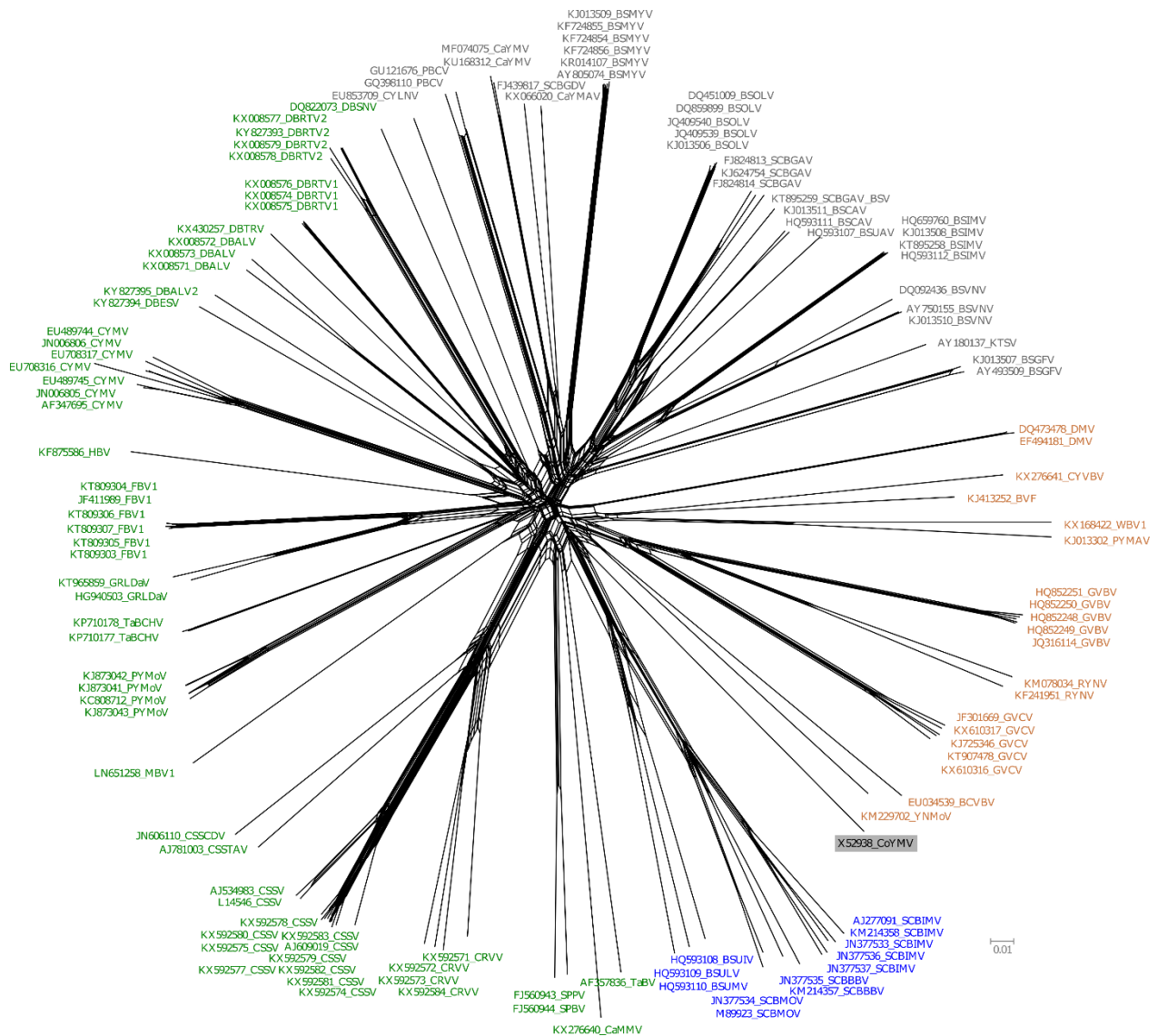


Figure 7. Non-tree-like evolution based on full RT/RNaseH nucleotide sequences of badnaviruses. Badnavirus Groups 1 (indicated in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV, highlighted in grey), was placed apart from the four main groups.

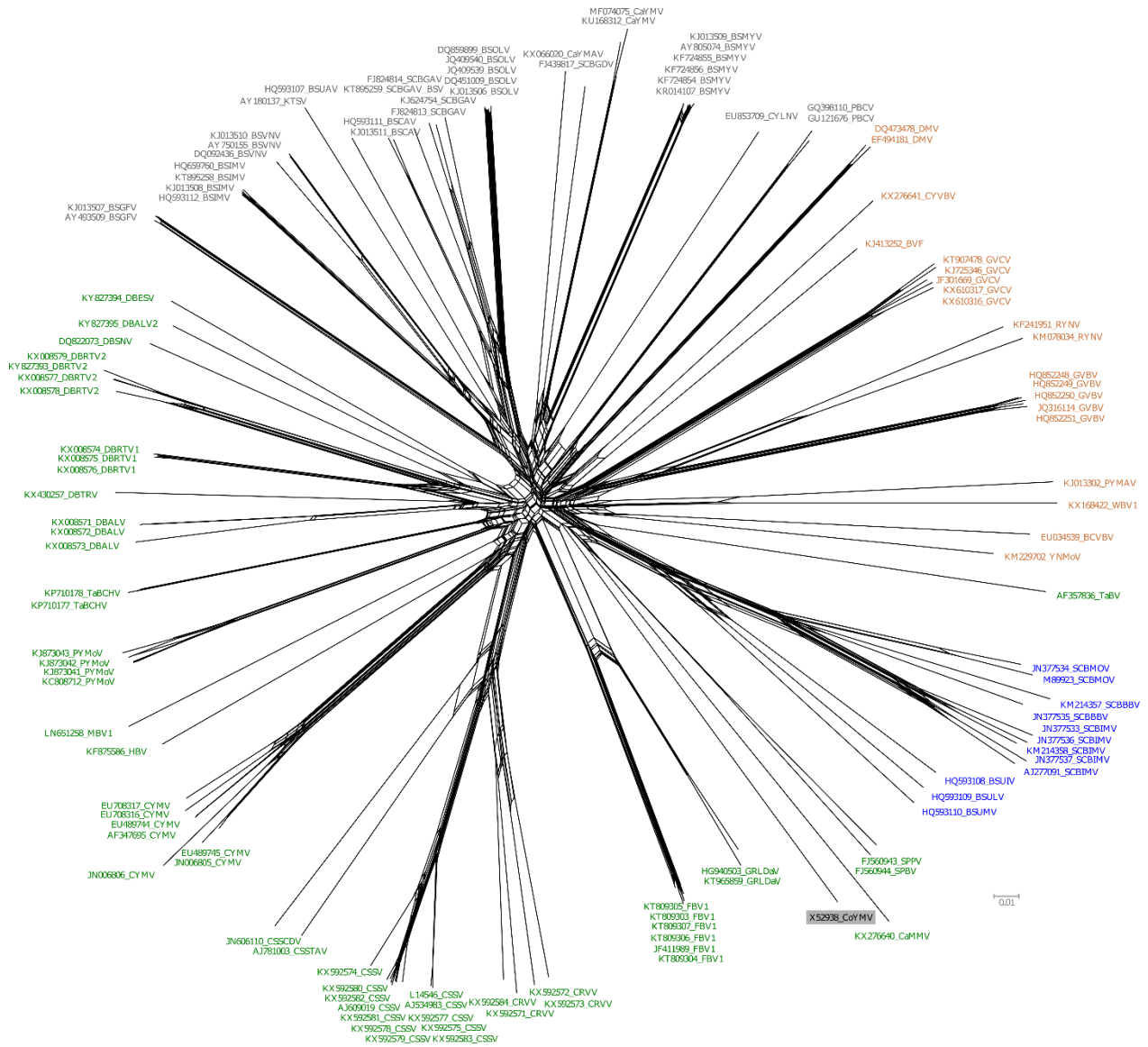


Figure 8. Non-tree-like evolution based on partial RT/RNaseH nucleotide sequences of badnaviruses. Badnavirus Groups 1 (indicated in grey), 2 (green), 3 (blue) and 4 (orange). *Commelina yellow mottle virus* (CoYMV, highlighted in grey), was placed apart from the four main groups.

Table 1. Full-length badnavirus sequences retrieved from the non-redundant GenBank database on Nov 2017.

Species/Acronym	GenBank access#
<i>Banana streak CA virus</i> /BSCAV	HQ593111, KJ013511
<i>Banana streak GF virus</i> /BSGFV	AY493509, KJ013507
<i>Banana streak IM virus</i> /BSIMV	HQ593112, KJ013508, HQ659760, KT895258
<i>Banana streak MY virus</i> /BSMYV	KR014107, KJ013509, KF724855, KF724856, AY805074, KF724854
<i>Banana streak OL virus</i> /BSOLV	JQ409540, JQ409539, KJ013506, DQ451009, DQ859899
<i>Banana streak UA virus</i> /BSUAV	HQ593107
<i>Banana streak UI virus</i> /BSUIV	HQ593108
<i>Banana streak UL virus</i> /BSULV	HQ593109
<i>Banana streak UM virus</i> /BSUMV	HQ593110
<i>Banana streak virus Acuminata Yunnan</i> /BSVNV	DQ092436, AY750155, KJ013510
<i>Blackberry virus F</i> /BVF	KJ413252
<i>Bougainvillea spectabilis chlorotic vein-banding virus</i> /BCVBV	EU034539
<i>Cacao mild mosaic virus</i> /CaMMV	KX276640
<i>Cacao swollen shoot CD virus</i> /CSSCDV	JN606110
<i>Cacao swollen shoot Togo A virus</i> /CSSTAV	AJ781003
<i>Cacao red vein virus</i> /CRVV	KX592584, KX592573, KX592572, KX592571, KX592582, KX592583, KX592581, KX592580, KX592579, KX592578, KX592577, KX592575, KX592574, AJ609019, AJ534983, L14546
<i>Cacao swollen shoot virus</i> /CSSV	
<i>Cacao yellow vein-banding virus</i> /CYVBV	KX276641
<i>Canna yellow mottle virus</i> /CaYMV	KU168312, MF074075
<i>Canna yellow mottle-associated virus</i> /CaYMAV	KX066020
<i>Citrus yellow mosaic virus</i> /CYMV	EU489745, EU489744, EU708316, JN006806, JN006805, EU708317, AF347695
<i>Commelina yellow mottle virus</i> / CoYMV	X52938
<i>Cycad leaf necrosis virus</i> /CYLNV	EU853709
<i>Dioscorea bacilliform AL virus</i> /DBALV	KX008573, KX008572, KX008571
<i>Dioscorea bacilliform RT virus</i> /DBRTV	KX430257
<i>Dioscorea bacilliform RT virus 1</i> /DBRTV1	KX008576, KX008575, KX008574
<i>Dioscorea bacilliform RT virus 2</i> /DBRTV2	KX008579, KX008578, KX008577, KY827393
<i>Dioscorea bacilliform SN virus</i> /DBSNV	DQ822073
<i>Dioscorea bacilliform AL virus 2</i> /DBALV2	KY827395
<i>Dioscorea baciliform ES virus</i> /DBESV	KY827394
<i>Dracaena mottle virus</i> /DMV	DQ473478, EF494181
<i>Fig badnavirus 1</i> /FBV-1	KT809307, KT809306, KT809305, KT809304, KT809303, JF411989
<i>Gooseberry vein banding virus</i> /GVBV	HQ852251, HQ852250, HQ852249, HQ852248, JQ316114
<i>Grapevine roditis leaf discoloration-associated virus</i> /GRLDaV	HG940503, KT965859
<i>Grapevine vein-clearing virus</i> /GVCV	KX610317, KX610316, KT907478, JF301669, KJ725346
<i>Hibiscus bacilliform virus</i> /HBV	KF875586
<i>Kalanchoe top-spotting virus</i> /KTSV	AY180137
<i>Mulberry badnavirus 1</i> /MBV-1	LN651258
<i>Pagoda yellow mosaic associated virus</i> /PYMAV	KJ013302
<i>Pineapple bacilliform comosus virus</i> /PBCV	GU121676, GQ398110
<i>Piper yellow mottle virus</i> /PYMoV	KJ873043, KJ873041, KJ873042, KC808712

<i>Rubus yellow net virus</i> /RYNV	KF241951, KM078034
<i>Sugarcane bacilliform GA virus</i> /SCBGAV SCBGAV/BSV	FJ824813, FJ824814, KJ624754 KT895259
<i>Sugarcane bacilliform GD virus</i> /SCBGDV	FJ439817
<i>Sugarcane bacilliform IM virus</i> /SCBIMV	AJ277091, JN377537, JN377536, JN377533, KM214358
<i>Sugarcane bacilliform MO virus</i> /SCBMOV	M89923, JN377534
<i>Sugarcane bacilliform BB virus</i> /SCBBBV	KM214357, JN377535
<i>Sweet potato pakakuy virus</i> /SPPV	FJ560943
<i>Sweet potato badnavirus B</i> /SPBV	FJ560944
<i>Taro bacilliform CH virus</i> /TaBCHV	KP710178, KP710177
<i>Taro bacilliform virus</i> /TaBV	AF357836
<i>Wisteria badnavirus 1</i> /WBV-1	KX168422
<i>Yacon necrotic mottle virus</i> /YNMoV	KM229702

Table 2. Putative recombination events detected within badnavirus isolates, based on complete genome sequences.

Event	Breakpoints*		Recombinant	Parents		Methods†	P value‡
	Begin	End		Minor	Major		
1	202	1499	JN006805_CYMV	JN006806_CYMV	AF347695_CYMV	RGBMCS3	1.7 E-82
2	833	1183	AY805074_BSMYV	Unknown	KR014107_BSMYV	RGBMCS3	8.6 E-20
3	3493	7233	^KJ873042_PYMoV	KJ873043_PYMoV	KC808712_PYMoV	RGBMCS3	6.4 E-28
4	7283(?)	202(?)	^EU489745_CYMV EU489744_CYMV	JN006805_CYMV	JN006806_CYMV	RGBMCS3	7.3 E-20
5	202(?)	4651	^JN006806_CYMV	EU489745_CYMV	AF347695_CYMV	RGBMCS3	1.7 E-44
6	7510	80	HQ852251_GVBV	Unknown	HQ852249_GVBV	RGBMCS3	5.3 E-15
7	4342	4833(?)	^HQ852250_GVBV HQ852248_GVBV	Unknown	HQ852249_GVBV	RGBMCS3	3.1 E-14
8	348	499	^KJ013508_BSIMV	Unknown	KT895258_BSIMV	RGBMCS	3.1 E-13
9	1007	1186	^JN377534_SCBMOV	KM214358_SCBIMV	M89923_SCBMOV	RGBMCS3	4.6 E-12
10	3733	3848	^JN377533_SCBIMV	Unknown	KM214357_SCBBBV	RGBMCS	7.1 E-11
11	1066(?)	1309	^GU121676_PBCV	Unknown	GQ398110_PBCV	RBMC3	5.0 E-06
12	6306	647(?)	^KF724854_BSMYV KF724856_BSMYV	Unknown	AY805074_BSMYV	RGMC3	1.5 E-10

KJ013509_BSMYV							
13	7050	7146(?)	^JN377534_SCBMOV AJ277091_SCBIMV M89923_SCBMOV KM214357_SCBBBV	FJ824813_SCBGAV	Unknown	RGBS3	1.8 E-10
14	4356	4863	^DQ092436_BSVNV AY750155_BSVNV KJ013510_BSVNV	Unknown	KJ013508_BSIMV	RBMC3	2.1 E-10
15	865	1234	HQ593107_BSUAV	HQ593109_BSULV	HQ593111_BSCAV	RBMC3	2.6 E-10
16	7463(?)	202(?)	^AF347695_CYMV JN006805_CYMV	Unknown	EU708317_CYMV	RGBMC3	2.1 E-09
17	7290(?)	7401	^DQ092436_BSVNV AY750155_BSVNV	FJ439817_SCBGDV	HQ659760_BSIMV	RGCS3	2.0 E-08
18	7078	210	^AJ534983_CSSV L14546_CSSV	Unknown	KX592583_CSSV KX592582_CSSV	RBMC3	8.8 E-08
19	1400(?)	2363	AY493509_BSGFV KJ013507_BSGFV	KU168312_CaYMV	AY180137_KTSV	RBMC3	2.2 E-09
20	5022	6244(?)	^KT965859_GRLDaV HG940503_GRLDaV	FJ824813_SCBGAV	KX592572_CRVV	BMCS3	4.9 E-07
21	601	1206	^JN377536_SCBIMV KM214358_SCBIMV	M89923_SCBMOV	KM214357_SCBBBV	RGBMCS	2.6 E-10

AJ277091_SCBIMV
 JN377537_SCBIMV

22	393	702	[^] HQ593107_BSUA	Unknown	DQ092436_BSVNV	RBMCS3	4.7 E-05
23	1055	1388(?)	[^] KX610317_GVCV	KJ725346_GVCV	JF301669_GVCV	RBMCS	2.4 E-06

* Numbering starts at the 5' end of the minus-strand primer-binding site and increases clockwise. (?), breakpoint could not be precisely pinpointed.

† R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, Chimera; S, SisScan; 3, 3SEQ.

‡ The reported *P* values are for the methods indicated in red, and they are the lowest *P* values calculated for the region in question.

[^] The recombinant sequence may have been misidentified (one of the identified parents might be the recombinant).

Table 3. Genetic variability of the RT/RNaseH region of badnavirus populations infecting distinct hosts.

Population	No. of sequences	Length (nt)	π^* (SD)[†]
BSGFV	13	668	0.0209 (± 0.0069)
BSMYV	10	659	0.0054 (± 0.0013)
BSOLV	35	668	0.0164 (± 0.0038)
BSUAV	8	662	0.0456 (± 0.0039)
BSULV	12	668	0.1467 (± 0.0133)
BSUMV	8	668	0.0579 (± 0.0061)
PYMoV	27	579	0.0944 (± 0.0037)
SPBVB	35	668	0.0418 (± 0.0088)
CSSV	26	575	0.0334 (± 0.0055)
CSSCDV	7	575	0.0162 (± 0.0019)
SCBIMV	118	725	0.0876 (± 0.0034)
DBALV[‡]	96	435	0.0534 (± 0.0054)

* Pairwise, per-site nucleotide diversity.

[†] Standard deviation.

[‡] Guimarães et al., 2015.

ANEXO I

Lista das espécies de *Badnavirus* reconhecidos pelo ICTV até novembro de 2017:

Aglaonema bacilliform virus
Banana streak GF virus
Banana streak IM virus
Banana streak MY virus
Banana streak OL virus
Banana streak UA virus
Banana streak UI virus
Banana streak UL virus
Banana streak UM virus
Banana streak VN virus
Bougainvillea chlorotic vein banding virus
Cacao swollen shoot CD virus
Cacao swollen shoot Togo A virus
Cacao swollen shoot virus
Canna yellow mottle virus
Citrus yellow mosaic virus
Commelina yellow mottle virus
Dioscorea bacilliform AL virus
Dioscorea bacilliform SN virus
Fig badnavirus 1
Gooseberry vein banding associated virus
Grapevine Roditis leaf discoloration-associated virus
Grapevine vein clearing virus
Kalanchoë top-spotting virus
Mulberry badnavirus 1

Pagoda yellow mosaic associated virus

Pineapple bacilliform CO virus

Pineapple bacilliform ER virus

Piper yellow mottle virus

Rubus yellow net virus

Schefflera ringspot virus

Spiraea yellow leafspot virus

Sugarcane bacilliform Guadeloupe A virus

Sugarcane bacilliform Guadeloupe D virus

Sugarcane bacilliform IM virus

Sugarcane bacilliform MO virus

Sweet potato pakakuy virus

Taro bacilliform CH virus

Taro bacilliform virus

Yacon necrotic mottle virus