



**UNIVERSIDADE FEDERAL DE ALAGOAS
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
PROTEÇÃO DE PLANTAS**



LUCAS JOBIM JORDÃO

**POSICIONAMENTO TAXONÔMICO E EFEITO DA
RECOMBINAÇÃO ENTRE ISOLADOS DO GÊNERO
*BADNAVIRUS***

Rio Largo - AL
2015

LUCAS JOBIM JORDÃO

**POSICIONAMENTO TAXONÔMICO E EFEITO DA
RECOMBINAÇÃO ENTRE ISOLADOS DO GÊNERO *BADNAVIRUS***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Proteção de Plantas do Centro de Ciências Agrárias da Universidade Federal de Alagoas como requisito parcial para obtenção do grau de Mestre em Proteção de Plantas.

Orientador: Prof. Dr. Gaus S. A. Lima
Coorientador: Dr. Roberto Ramos Sobrinho

Rio Largo - AL
2015

LUCAS JOBIM JORDÃO

**POSICIONAMENTO TAXONÔMICO E EFEITO DA
RECOMBINAÇÃO ENTRE ISOLADOS DO GÊNERO *BADNAVIRUS***

Dissertação apresentada ao Curso de Mestrado
em Proteção de Plantas, do Centro de Ciências
Agrárias da Universidade Federal de Alagoas.
Aprovada pela seguinte Banca Examinadora:

Prof. Dr. Gaus Silvestre de Andrade Lima

Banca Examinadora:

Prof. Dra. Iraildes Pereira Assunção – Universidade Federal de Alagoas (Examinador)

Prof. Dra. Sarah Jaqueline Cavalcanti da Silva – Universidade Federal de Alagoas
(Examinador)

Rio Largo - AL
2015

AGRADECIMENTOS

A Deus, acima de tudo, por sempre me guiar em todos os momentos da minha trajetória e por dar-me coragem para enfrentar todos os desafios;

Aos meus queridos pais, Luiz Carlos Jordão e Simone Jobim, pelo amor, incentivo e apoio em todas as horas;

A minhas irmãs Larissa, Livia e Marina por sempre estarem ao meu lado;

A todos familiares e amigos que contribuíram para meu crescimento;

Aos professores Drs. Gaus Silvestre de Andrade e Iraildes Pereira Assunção pela oportunidade, confiança, orientação, ensinamentos que foram fundamentais para minha formação. Serei sempre grato;

Aos Drs. Roberto Ramos e Leonardo Barbosa pela amizade, pelos ensinamentos, orientação e apoio durante minha trajetória no mestrado.

Aos meus amigos Mayra Machado, Aline Marques, Nayana Bruschi e Janaíne Rossane pelo companheirismo no dia-a-dia, pelos momentos de alegria, aprendizado e por tornarem tudo mais prazeroso;

A todos do Laboratório de Fitopatologia Molecular e Virologia Vegetal: Sarah, Jaqueline Figueredo, Joseane, Antônio, Maria Jussara, Fred, Renato Nascimento, Ana, Tamires, Laura, Jackeline Laurentino, Gilberlan, Jecylene Tomé, Daniel, pela amizade, ajuda mútua e ensinamentos durante todo esse período;

A todos os professores do Programa de Pós-graduação em Proteção de Plantas CECA-UFAL pelos conhecimentos passados, que contribuíram para a minha formação;

A todos os meus colegas do Programa de Pós-graduação em Proteção de Plantas do CECA-UFAL;

A CAPES pela bolsa de mestrado concedida, o que me permitiu adquirir inúmeros conhecimentos incluindo os fundamentais para a realização desse trabalho;

A todos aqueles que de alguma forma contribuíram para realização deste trabalho.

Obrigado!

RESUMO

Vírus pertencentes à família *Caulimoviridae* apresentam morfologia isométrica ou baciliforme e são divididos em oito gêneros, com o gênero *Badnavirus* sendo o mais importante devido ao grande número de espécies descritas infectando espécies cultivadas em todo o mundo. O presente trabalho teve como objetivo determinar o efeito da mutação e recombinação na evolução genômica de espécies de badnavírus. Inicialmente, um conjunto de dados de sequências completas de isolados de badnavírus foi obtido a partir do banco de dados não-redundantes GenBank. Alinhamentos múltiplos de sequências nucleotídicas foram preparados para os conjuntos de dados genoma completo, RT/RNaseH completa (1020pb), RT/RNaseH parcial (579pb), e ORFs I, II e III. Comparações pareadas, e análises filogenéticas e de recombinação foram realizadas para todos os conjuntos de dados individualmente. Um total de 71 isolados foi obtido, representando 32 espécies distintas. Análise de comparações pareadas de nucleotídeos para os conjuntos de dados RT/RNaseH e ORF III mostrou que alguns isolados considerados como espécies distintas compartilharam até 82% de identidade nucleotídica, ultrapassando o limiar de $\geq 80\%$ para demarcação de espécies de badnavírus. Árvores filogenéticas Bayesianas mostraram a formação de quatro grupos bem sustentados, com os *Badnavirus Groups* 1 e 3 formando caldos irmãos contendo predominantemente isolados infectando cana-de-açúcar e banana. Análise reticulada evidenciou um complexo padrão de recombinação, e pelo menos 13 eventos foram detectados com *breakpoints* de recombinação ocorrendo predominantemente na ORF III e região intergênica. Estes resultados mostram que mutação e recombinação são importantes mecanismos evolutivos atuando na diversificação de espécies de badnavírus, e que a região RT/RNaseH parcial é suficiente para determinar o posicionamento taxonômico da maioria das espécies descritas dentro do gênero.

Palavras-chave: *Caulimoviridae*. mecanismos evolutivos. Filogenia. variabilidade genética.

ABSTRACT

Viruses belonging to the family *Caulimoviridae* have isometric or bacilliform particles and are divided into eight genera, with the genus *Badnavirus* being the most important due the high number of species reported infecting cultivated plants worldwide. The present study aimed to determine the effect of mutation and recombination in the genomic evolution of badnavirus species. Firstly, a data set comprising badnavirus full-length genome sequences was obtained from the non-redundant database GenBank. Multiple nucleotide sequence alignments were prepared for the data sets complete genome, complete RT/RNaseH (1020 pb), partial RT/RNaseH (579 pb), and ORFs I, II and III. Pairwise comparisons, and phylogenetic and recombination analyses were performed for all data sets individually. A total of 71 isolates were obtained, representing 32 badnavirus species. Nucleotide pairwise comparisons for the data sets RT/RNaseH and ORF III showed that some distinct badnavirus species shared up to 82% of nucleotide identity, being higher than the $\geq 80\%$ threshold for badnavirus species demarcation. Bayesian phylogenetic trees showed four well supported clusters, with the *Badnavirus Groups* 1 and 3 being sister clades comprising predominantly isolates infecting sugarcane and banana. Network analysis evidenced a complex pattern of recombination, and at least 13 events were detected with recombination breakpoints occurring predominantly in the ORF III and Intergenic Region. These results showed that mutation and recombination are important evolutionary mechanisms acting on diversification of badnavirus species, and that the partial RT/RNaseH is sufficient to determine the taxonomic placement of most viral species described into this genus.

Key words: *Caulimoviridae*. evolutionary mechanisms. Phylogeny. genetic variability.

LISTA DE ILUSTRAÇÕES

Figura 1 - Representação esquemática do genoma do <i>Commelina yellow mottle virus</i> (CoYMV), espécie tipo do gênero <i>Badnavirus</i>	16
Figura 2 - Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de sequências nucleotídicas da região completa dos domínios RT/RNaseH (579 pb) de isolados de <i>Badnavirus</i> disponíveis no GenBank.....	36
Figura 3 - Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de sequências nucleotídicas da região parcial dos domínios RT/RNaseH (1020 pb) de isolados de <i>Badnavirus</i> disponíveis no GenBank.....	37
Figura 4 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados da região completa dos domínios RT/RNaseH de isolados de <i>Badnavirus</i> disponíveis no GenBank	38
Figura 5 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados da região parcial dos domínios de RT/RNaseH de isolados de <i>Badnavirus</i> disponíveis no GenBank	39
Figura 6 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados do genoma completo de isolados de <i>Badnavirus</i> disponíveis no GenBank	40
Figura 7 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF III de isolados de <i>Badnavirus</i> disponíveis no GenBank.....	41
Figura 8 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF I de isolados de <i>Badnavirus</i> disponíveis no GenBank	42
Figura 9 - Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF II de isolados de <i>Badnavirus</i> disponíveis no GenBank	43
Figura 10 - Análise reticulada baseada no conjunto de dados do genoma completo dos isolados de <i>Badnavirus</i>	44
Figura 11 - Análise reticulada baseada no conjunto de dados da ORF 3 dos isolados de <i>Badnavirus</i>	45
Figura 12 - Análise reticulada baseada no conjunto de dados da ORF 1 dos isolados de <i>Badnavirus</i>	46
Figura 13 - Análise reticulada baseada no conjunto de dados da ORF 2 dos isolados de <i>Badnavirus</i>	47

LISTA DE TABELAS

Tabela 1 - Sequências de badnavirus obtidas a partir do banco de dados não-redundantes GenBank.....	33
Tabela 2 - Eventos de recombinação detectados entre membros do gênero <i>Badnavirus</i>	35

SUMÁRIO

1 INTRODUÇÃO	11
2 REVISÃO DA LITERATURA	13
2.1 Família <i>Caulimoviridae</i>	13
2.2 Gênero <i>Badnavirus</i>	14
2.3 Sequências endógenas	17
2.4 Culturas de grande importância econômica infectadas por <i>Badnavirus</i>	18
2.5 Variabilidade e estrutura genética de populações de <i>Badnavirus</i>	20
2.5.1 Mutação	21
2.5.2 Recombinação	22
3 MATERIAL E MÉTODOS	23
3.1 Conjunto de dados	23
3.2 Alinhamento e comparação de sequências	23
3.3 Análises de recombinação	23
3.4 Análise filogenética	24
4 RESULTADOS	25
4.1 Conjunto de dados	25
4.2 Demarcação de espécies	25
4.3 Relacionamento Filogenético	26
4.4 Eventos de Recombinação	27
5 DISCUSSÃO	28
6 CONCLUSÕES	32
7 REFERÊNCIAS	48

1 INTRODUÇÃO

A família *Caulimoviridae* engloba vírus que apresentam morfologia isométrica ou baciliforme, possuindo tamanho de aproximadamente 35-50nm de diâmetro e 900nm de comprimento para os de formato baciliforme, e 45-50nm de diâmetro para os vírus com formato isométrico (GEERING; HULL, 2012). Vírus dentro dessa família são caracterizados por possuírem genoma composto por DNA semicircular de fita dupla medindo cerca de 7,2-9,2kb (FAUQUET et al., 2005; GEERING; HULL, 2012). A família *Caulimoviridae* é dividida em oito gêneros (*Badnavirus*, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Rosadnavirus*, *Solendovirus*, *Soymovirus* e *Tungrovirus*) de acordo com gama de hospedeiro, inseto vetor, filogenia e organização genômica (GEERING; HULL, 2012). O gênero *Badnavirus* é considerado o segundo em importância dentro das famílias de fitovírus com genoma de DNA, com diversas espécies sendo relatadas infectando culturas de grande importância econômica, como: cana-de-açúcar (*Saccharum* spp.), banana (*Musa* spp.), cacau (*Theobroma cacao* L.), pimenta (*Capsicum* spp.), inhame (*Dioscorea* spp.), citros (*citrus* spp.) (SILVA et. al, 2015; JAMES et. al, 2011a; KOUAKOU et. al, 2012; DEESHMA; BHAT, 2015; ENI et. al, 2008; JOHNSON et. al, 2012). Além dessa alta diversidade de espécies, populações de badnavírus também apresentam altos índices de variabilidade genética (MULLER et al. 2011; HARPER et al., 2005; GEERING et al., 2000; JOHNSON et. al, 2012; GUIMARÃES et al., 2015; SILVA et al., 2015).

O genoma dos badnavírus é composto por uma única molécula de dsDNA de 7,2-7,6kb, a qual forma um semicírculo interrompido por descontinuidades sítio-específicas que podem conter uma região intergênica poli A, e possuindo, em geral, três ORFs (*Open Reading Frames*) conhecidas como ORFs I, II e III (BOUHIDA; LOCKHART; OLSZEWSKI, 1993; HAGEN et al., 1993; HARPER; HULL, 1998). Até o presente momento, nenhuma função foi atribuída às ORFs I e II, porém a ORF3 codifica para uma poliproteína de 208 a 216 kDa que é proteoliticamente clivada, gerando as proteínas de movimento, do capsídeo viral, a aspartato protease (AP) (responsável pela clivagem da poliproteína), a transcriptase reversa (RT) e a ribonuclease H (domínios genômicos envolvidos no processo de replicação viral) (HARPER; HULL, 1998; LACO; BEACHY, 1994; MEDBERRY; LOCKHART; OLSZEWSKI, 1990).

A identificação de espécies de badnavírus é realizada através da comparação de sequências nucleotídicas dos domínios RT/RNaseH, com $\geq 80\%$ de identidade sendo o

critério estabelecido para demarcação de espécies (GEERING; HULL, 2012). A amplificação dessa região genômica viral é geralmente realizada por meio da técnica de PCR (*Polymerase Chain Reaction*) utilizando-se oligonucleotídeos específicos (YANG et al., 2003). Entretanto, esta técnica torna-se limitada durante processo diagnóstico, dado o fato de não ser capaz de distinguir entre sequências de badnavírus integradas ao genoma da planta hospedeira e cópias episomais dos genomas virais (BOUSALEM; DOUZERY; SEAL, 2009) e, também devido à grande diversidade e variabilidade genética entre membro desse grupo de vírus (MULLER et al., 2011; NDOWORA et al., 1999). Na tentativa de contornar tais dificuldades, métodos baseados na amplificação por círculo rolante (RCA) têm sido utilizados, sendo considerado um método confiável por amplificar apenas DNA circular e, portanto, evitando falso-positivos a partir de sequências virais endógenas (JAMES et. al, 2011a, 2011b).

O presente trabalho teve como objetivo determinar os efeitos da mutação e recombinação na evolução das espécies virais pertencentes ao gênero *Badnavirus*.

2 REVISÃO DE LITERATURA

2.1 Família *Caulimoviridae*

A família *Caulimoviridae* engloba vírus encapsidados em partículas não envelopadas, e podem ser de morfologia isométrica ou baciliforme. Possuem aproximadamente 35-50 nm de diâmetro e 900nm de comprimento (formato baciliforme) ou 45-50 nm de diâmetro e simetria icosaédrica (formato isométrico). O genoma dos *Caulimoviridae* é constituído por uma molécula de DNA semicircular de fita dupla com cerca de 7,2-9,2kb (FAUQUET et al., 2005; GEERING; HULL, 2012), uma região intergênica poli (A) presente em alguns membros da família, e descontinuidades de fita simples ou gaps em sítios específicos de ambas as fitas (HARPER et al., 2002).

Os hospedeiros naturais de espécies da família *Caulimoviridae* são angiospermas das classes Dicotyledonae e Monocotyledonae. Dependendo do gênero, a transmissão natural do vírus pode ocorrer via inseto vetor (insetos das famílias Aleyrodidae, Aphididae, Cicadellidae e Pseudococcidae: ordem Hemiptera) ou por contato entre plantas hospedeiras, bem como por sementes ou grão de pólen. A transmissão pode ser também realizada via inoculação mecânica e enxertia (FAUQUET et al., 2005). 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).

Segundo o Comitê Internacional de Taxonomia de Vírus (ICTV), a família *Caulimoviridae* é dividida em oito gêneros (*Badnavirus*, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Rosadnavirus*, *Solendovirus*, *Soymovirus* e *Tungrovirus*) de acordo com gama de hospedeiro, inseto vetor, organização genômica e filogenia. O gênero *Badnavirus* é o que possui o maior número de espécies reconhecidas (FAUQUET et al., 2005). Recentemente, foi sugerido o reconhecimento de pelo menos mais duas novas espécies de badnavírus infectando cacau: Cacao swollen shoot Togo A virus (AJ781003) e Cacao swollen shoot Côte d'Ivoire D virus (JN606110) (KOUAKOU et. al, 2012).

A replicação dos membros da família *Caulimoviridae* é similar a dos retrovírus, com produção de um RNA intermediário de fita simples (ssRNA) através de transcrição reversa, formado para gerar dsDNA (DNA fita dupla) no citoplasma. 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. Tal DNA é transcrito pela RNA polymerase II do hospedeiro (DE KOCHKO et al.,1998). O RNA transcrito é 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 et al., 1990). Os dsDNA genômicos são encapsidados, produzindo assim novos vírions. O genoma desses vírus também 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). Outra característica importante dos vírus da família *Caulimoviridae* é sua capacidade de integração dentro do genoma do hospedeiro para replicação, apesar de não codificarem para uma proteína integrase. Dado essas características, Temin (1985) sugeriu a classificação dos membros desta família como "pararetrovírus", termo introduzido para definir vírus de plantas que diferem dos retrovírus por possuírem genoma de DNA e capacidade de integração ao genoma do hospedeiro.

2.2 Gênero *Badnavirus*

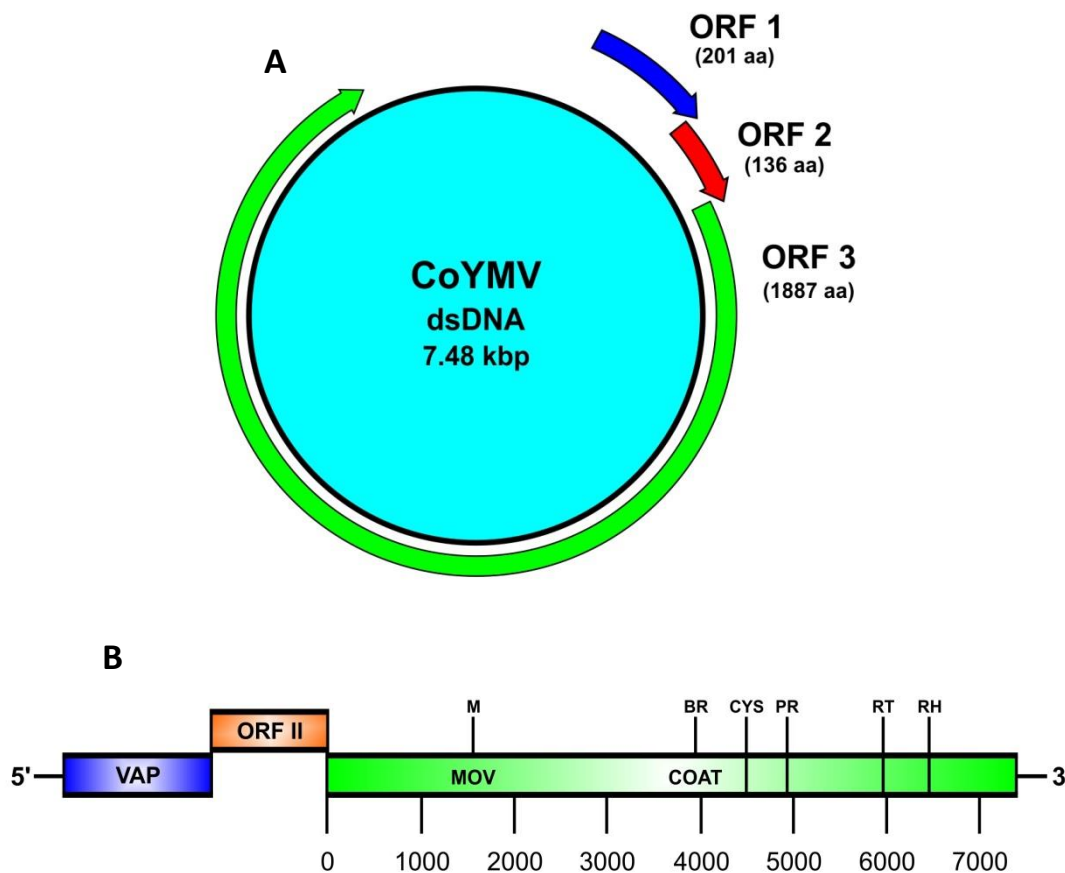
O gênero *Badnavirus* é o que apresenta maior diversidade de espécies dentro da família *Caulimoviridae*, com 32 espécies oficialmente reconhecidas pelo ICTV, é o segundo maior gênero de vírus de plantas com genoma de DNA (BOUHIDA et al., 1993; HAGEN et al., 1993). Badnavírus apresentam formato baciliforme com 95-130 nm de comprimento e 24-35 nm de largura (FAUQUET et al., 2005). O genoma dos badnavírus contém uma única molécula de dsDNA semicircular com cerca de 7,2-7,6 kb, apresentam descontinuidades sítio-específicas e podem conter uma região intergênica poli A (MEDBERRY et al., 1990).

O genoma da maioria dos badnavírus codifica para três ORFs (Figura 1A), porém algumas espécies podem possuir maior número de ORFs, como: *Taro baciliform virus* (TaBV) com quatro ORFs; *Cacao swollen shoot virus* (CSSV) com cinco; e *Citrus yellow mosaic virus* (CYMV) com seis ORFs presentes no seu genoma (HAGEN, 1993; YANG,

2003). Contudo, todos os badnavírus contêm as três ORFs principais (ORFs I, II e III) (BOUHIDA et al., 1993; HAGEN et al., 1993; HARPER; HULL, 1998).

As funções dos produtos das ORFs I e II permanecem desconhecidas. Entretanto, a proteína P1 de 20,8 a 23kDa codificada pela ORF I em *Comellina yellow mottle virus* (CoYMV) foi descrita em associação com vírions imaturos, enquanto a proteína P2 de 14,5 a 15kDa codificada pela ORF II foi encontrada associada com ambos os vírions imaturos e maduros (CHENG et al., 1996). A ORF III, codifica para uma poliproteína de 208 a 216kDa que é proteoliticamente clivada e gera, dentre outros produtos, as proteína de movimento e do capsídeo viral, a aspartato protease (AP; responsável pela clivagem da poliproteína), a transcriptase reversa (RT) e a ribonuclease H (conhecidos como domínios genômicos envolvidos no processo de replicação viral) (Figura 1B) (MEDBERRY et al., 1990; TZAFRIR et al., 1997; HARPER e HULL, 1998). Atualmente, a identificação de espécies de badnavírus é baseada na gama de hospedeiros, tipo de inseto vetor, e principalmente no critério de $\geq 80\%$ de identidade de sequências nucleotídicas da região RT/RNaseH (GEERING; HULL, 2012).

Figura 1. (A) Representação esquemática da organização genômica de *Comellina yellow mottle virus* (CoYMV) espécie-tipo do gênero *Badnavirus*. O círculo completo representa o genoma de DNA de fita dupla. As setas largas indicam a posição das ORFs I, II e III. (B) Representação linear do mapa genômico de badnavírus mostrando a identificação dos domínios: domínios de movimento (M), motivo de ligação ao RNA (BR), região conservada de cisteína (CYS), aspartato protease (PR), transcriptase reversa (RT) e ribonuclease H (RH) (BRIDDON et al., 1999).



A detecção de badnavírus tem sido realizada através de métodos como microscopia eletrônica de transmissão e de imuno-adsorção, testes sorológicos, e testes baseados na detecção de DNA (VEGA e SORDI, 1991; DELANOY et al., 2003; RIVAS et al., 2005; 2006; FIGUEIREDO et al., 2006; ANDRADE, 2007; VERZIGNASSI et al., 2009; ALEXANDRE et al., 2010; BRIOSO et al., 2011; JAMES et al., 2011; PENG 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 epissomais, 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 epissomais de espécies de Banana streak viruses (BSVs), dada a alta variabilidade genética presente neste complexo. (FIGUEIREDO et al., 2006; GEERING et al., 2000, 2005b; HARPER et al., 2005; HARPER; HULL, 1998).

Um método amplamente utilizado, e que apresenta alta eficiência na distinção entre sequências de badnavírus endógenas e epissomais, 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

Sequências genômicas de alguns pararetrovírus vegetais pertencentes ao gênero *Badnavirus* como *Banana streak OL virus* (BSOLV), *Banana streak GF virus* (BSGFV), *Petunia vein clearing virus* (PVCV), *Tobacco vein clearing virus* (TVCV), *Pineapple baciliform comosus virus* (PBCV) e *Dioscorea baciliform virus* (DBV) têm sido comumente observadas como sequências endógenas integradas ao genoma do hospedeiro (NDOWORA et al., 1999; GAYRAL; ISKRA-CARUANA, 2009; STAGINNUS; RICHERT-PÖGGELER, 2006; HERNANDEZ-RODRIGUEZ et al, 2013; SEAL, 2014). Estas formas virais endógenas muitas vezes exercem papel epidemiológico importante, já que podem dar origem a vírus epissomais e, conseqüentemente, iniciar processo de infecção viral. Condições de estresse como culturas de tecidos podem favorecer a liberação de sequências virais endógenas (DAHAL et al., 2000; DALLOT et al., 2001).

Dentre as principais culturas que apresentam alta frequência de integração de sequências de badnavírus pode-se citar a cultura da banana (*Musa* spp.), onde são comumente encontrado 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 consequente infecção do hospedeiro (GAYRAL et al., 2008; JAMES et al., 2011). Recentemente, pelo 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, 2013). Além da banana, culturas como abacaxi e inhame têm sido relatadas para conter sequências endógenas de badnavírus (SEAL, 2014; GAMBLEY et al., 2008).

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.), abacaxi (*Ananas comosus*), cana-de-açúcar (*Saccharum* spp.), banana (*Musa* spp.), citros (*citrus* spp.), cacau (*Theobroma cacao* L.) e pimenta (*Capsicum* spp.) (SILVA et al., 2015; JAMES et al., 2011a; KOUAKOU et al., 2012; DEESHMA; BHAT, 2015; ENI et al., 2008; JOHNSON et al., 2012). Em alguns casos, esta infecção pode interferir diretamente na produtividade da cultura, causando perdas significativas em qualidade e rendimento. Estas perdas podem variar 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 (DBALV) infectando *D. alata* na Nigéria (PHILLIPS, 1999). 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 recentes 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). A cultura do inhame desempenha grande importância na alimentação em regiões da África, onde são observados altos índices de produtividade (MAFRA, 1996). No Brasil, a região Nordeste destaca-se como principal

produtora, onde serve como fonte de alimento e auxilia na geração de emprego e renda (PEDRALLI, 2002). Dentre os vários problemas fitossanitários afetando esta cultura destacam-se as fitoviroses causadas por espécies de badnavírus (SEAL; MULLER, 2007; PHILLIPS et al., 1999). No Brasil, a espécie mais amplamente distribuída nos campos de cultivo de inhame é o *Dioscorea bacilliform AL virus* (DBALV) (GUIMARÃES et al., 2015). 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).

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* e, no caso do PBCoV, também pelas cochonilhas *Planococcus citri* e *D. neobrevipes* (GAMBLEY et al., 2008b; SETHER et al., 2012). Entretanto, uma associação clara entre sintomas e infecção por badnavírus em abacaxi ainda não foi demonstrada (GAMBLEY et al., 2008b).

Na cultura da cana-de-açúcar, são reconhecidas pelo ICTV duas espécies de badnavirus: *Sugarcane bacilliform IM virus* (SCBIMV) e *Sugarcane bacilliform MO virus* (SCBMOV), sendo comumente observados sintomas de manchas e estrias cloróticas e nanismo (VISWANATHAN et al., 1996). 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 et al., 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 et al., 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).

Diversas espécies de badnavirus 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). 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, além do

uso indiscriminado de agrotóxicos que possuem elevados custos econômicos e ambientais. 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., 2011; GEERING; HULL, 2012).

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 do 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).

Espécies do gênero *Badnavirus* têm causado grandes prejuízos para a cultura do cacau (*Theobroma cacao*) na África ocidental (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; HOFFMANN et al., 1997).

2.5 Variabilidade e estrutura genética de populações de fitovírus

A estrutura genética de populações de fitopatógenos refere-se à quantidade de variabilidade genética existente e de sua distribuição dentro e entre populações

(MILGROOM, 1995). O estudo da dinâmica da variabilidade genética de populações de fitopatógenos é necessário para conhecer como as populações evoluem (MCDONALD; ZHAN; BURDON, 1999). Uma vez compreendidas as maneiras como populações de fitopatógenos mantêm a variabilidade genética, é possível inferir sobre a taxa com a qual estas populações evoluem (MCDONALD; LINDE, 2002a). Programas de melhoramento visando resistência a patógenos podem ser orientados pelos conhecimentos da estrutura genética de populações. McDonald e Linde (2002b) propuseram um modelo para auxiliar na definição de estratégias de melhoramento visando aumentar a durabilidade da resistência, que se baseia no potencial evolutivo de populações de fitopatógenos. Por sua vez, o potencial evolutivo é determinado com base na estrutura genética das populações. Em síntese, o modelo considera a quantidade e a distribuição da variabilidade genética na população.

Em populações de fitovírus, o alto grau de diversidade e variabilidade genética é gerado, principalmente, por erros durante o processo de replicação do genoma viral (mutações), e rearranjos gênicos provocados por eventos de recombinação (GARCIA-ARENAL; FRAILE; MALPICA, 2003).

2.5.1 Mutação

A mutação é o processo pelo qual nucleotídeos que não estão presentes na fita molde são incorporados na fita filha durante a replicação dos ácidos nucléicos, com a taxa de mutação medindo o aparecimento de mutações espontâneas em função do tempo (GARCIA-ARENAL; FRAILE; MALPICA, 2003). É conhecido que taxas de mutações para vírus com genomas de RNA e DNA de fita simples (ssDNA) são extremamente elevadas (DUFFY; SHACKELTON; HOLMES, 2008). As altas taxas de mutação observadas para vírus de RNA (10^{-4} a 10^{-2} subs/sítio/ano) são explicadas pelo fato de esses vírus serem replicados por polimerases de RNA dependentes de RNA que não possuem capacidade de correção de erro (DRAKE; HOLLAND, 1999; JENKINS et al., 2002). Entretanto, apesar de serem replicados por polimerases de DNA dependentes de DNA que possuem capacidade de correção de erro, vírus de ssDNA evoluem tão rapidamente quanto os vírus de RNA (SHACKELTON et al., 2006; DUFFY; SHACKELTON; HOLMES, 2008; DUFFY; HOLMES, 2009). Isso pode ser explicado por duas hipóteses principais: primeiro, o tempo de duração do intermediário de fita dupla formado durante processo de

replicação de vírus de ssDNA não é suficientemente longo para que haja correção de erros; e segundo, que mutações espontâneas poderiam ocorrer na fita de DNA durante a fase de ssDNA viral (DUFFY; SHACKELTON; HOLMES, 2008; DUFFY; HOLMES, 2009).

Estudos com bactérias e sistemas animais indicaram que a taxa de mutação dos vírus de dsDNA e ssDNA diferem significativamente, com vírus de ssDNA evoluindo muito mais rapidamente (DUFFY; SHACKELTON; HOLMES, 2008). Infelizmente, informações sobre taxas de mutação para vírus de dsDNA que infectam plantas são escassas, sendo relatada apenas para o *Cauliflower mosaic virus* (CaMV), o qual apresentou taxas em torno de 10^{-5} subs/sítio/ano (FROISSART et al., 2005). Segundo Firth et al. (2010), os vírus dsDNA são frequentemente descritos evoluindo através de associações com seus hospedeiros, onde se espera um padrão de baixas taxas de substituição de nucleotídeos.

2.5.2 Recombinação

Recombinação é o processo pelo qual segmentos de uma fita de DNA ou RNA tornam-se incorporados na fita de um indivíduo diferente durante o mecanismo de replicação (PADIDAM; BEACHY; FAUQUET; 1999). Eventos de recombinação têm sido demonstrados estar associados com vírus aumentando sua gama de hospedeiros (GIBBS; WEILLER, 1999), virulência (PITA et al., 2001; ZHOU et al., 1997) e potencial evolutivo e adaptação local (MONCI et al., 2002).

Recombinação é um evento bastante comum em fitovírus com genoma de ssDNA como os geminivírus (LEFEUVRE et al., 2007; PADIDAM; BEACHY; FAUQUET, 1999), estando diretamente relacionada com os altos índices de variabilidade genética observados em populações de begomovírus (LIMA et al., 2013; ROCHA et al., 2013; RAMOS-SOBRINHO et al., 2014). A elevada frequência de recombinação nesse grupo de vírus pode ser, em parte, explicada pela ocorrência frequente de infecções mistas (DAVINO et al., 2009; GARCÍA-ANDRES et al., 2006; PITA et al., 2001; RIBEIRO et al., 2003; SANZ et al., 2000; TORRES-PACHECO et al., 1996), com evidência de infecção do mesmo núcleo da célula por mais de um begomovírus (MORILLA et al., 2004). Outro mecanismo associados às altas taxas de recombinação é a replicação dependente de recombinação (JESKE; LUTGEMEIER; PREISS, 2001).

Os vírus que pertencem à família *Caulimoviridae* (dsDNA) são considerados objetos ideais para estudos de recombinação. Esta família de vírus de plantas tem recebido atenção adicional desde a descoberta de que é capaz de se integrar no genoma da planta hospedeira (SQUIRES et al., 2011). Eventos de recombinação homóloga foram relatados para ativação dos genomas epissomais de BSV e TVCV integrados nos genomas de seus hospedeiros (NDOWORA et al., 1999; STAGINNUS; RICHERT-PÖGGELER, 2006). Estudos sugerem ainda que eventos de recombinação contribuem para a variação genética observada entre membros dos gêneros *Caulimovirus* (FROISSART et al., 2005) e *Tungrovirus* (ARBOLEDA; AZZAM, 2000).

3 MATERIAL E MÉTODOS

3.1 Conjunto de dados

Sequência do genoma completo de espécies do gênero *Badnavirus* foram obtidas a partir do banco de dados não-redundante GenBank (www.ncbi.nlm.nih.gov/genbank; acessado em maio de 2015). A partir dos genomas completos, foram obtidos conjuntos de dados das sequências nucleotídicas para as ORFs I, II, III, e domínios RT/RNaseH completo (1020 pb) e parcial (579 pb). O conjunto de dados RT/RNaseH parcial refere-se à região genômica amplificada pelos oligonucleotídeos mais amplamente utilizados para detecção e identificação de espécies de badnavírus (YANG et al., 2003).

3.2 Alinhamento e comparação de sequências

Sequências completas de aminoácidos das ORFs I, II, e III foram individualmente alinhadas utilizando o algoritmo MUSCLE (EDGAR, 2004), manualmente editadas no pacote MEGA 6 (TAMURA et al., 2013), e convertidas em sequências nucleotídicas para análises posteriores. Para estes três conjuntos de dados foram então utilizados para obtenção de alinhamento dos genomas completos. A partir da sequência obtida da ORF3, foram obtidos os conjunto de dados dos domínios RT/RNaseH (completo e parcial).

Para confirmar a taxonomia atribuída aos isolados de badnavírus obtidos a partir do GenBank, os conjuntos de dados genoma completo, ORFs I, II, III e domínios RT/RNaseH (completo e parcial) foram analisadas utilizando o programa *Sequence Demarcation Tool* v. 1.2 (MUHIRE et al., 2013).

3.3 Análises de recombinação

Evidência de evolução em rede para os conjuntos de dados genoma completo e ORFs I, II e III foi avaliada utilizando-se o método Neighbor-Net implementado no programa SplitsTree v. 4.10 (HUSON; BRYANT, 2006). Possíveis parentais e locais (*breakpoints*) de recombinação, para o conjunto de dados genoma completo, foram determinados utilizando-se os métodos RDP, Geneconv, Boot-scan, Maximum Chi Square, Chimaera, SisterScan e 3Seq implementados no pacote RDP v. 4.0 (MARTIN et al., 2015). Os alinhamentos foram analisados com definições padrão para os diferentes métodos e a significância estatística foi inferida por um *P*-valor menor que o Bonferroni corrigido de 0,05. Somente eventos de recombinação detectados no mínimo por quatro diferentes métodos foram considerados confiáveis.

3.4 Análise filogenética

Inferência filogenética Bayesiana para os conjuntos de dados genoma completo, ORFs I, II e III e domínios RT/RNaseH (completo e parcial) dos isolados de badnavírus foi realizada no *web portal* CIPRES (MILLER et al., 2010) usando MrBayes v. 3.2.3 (RONQUIST et al., 2012). O melhor modelo de substituição de nucleotídeos foi determinado para cada conjunto de dados usando MrModeltest 2.3 (POSADA; BUCKLEY, 2004) de acordo com o *Akaike Information Criterion* (AIC). O modelo evolutivo GTR+I+G foi utilizado para todos os conjuntos de dados. Análises foram rodadas para cada conjunto de dados por 20 milhões de gerações usando quatro cadeias e amostrando a cada 2.000 gerações, para um total de 10.000 árvores. As primeiras 2.500 árvores foram descartadas como uma fase de *burn-in*. Probabilidades posteriores (RANNALA; YANG, 1996) foram determinadas a partir de uma árvore consenso *majority-rule* gerada com as 7.500 árvores remanescentes. As árvores foram visualizadas e editadas nos programas FigTree v. 1.4 (tree.bio.ed.ac.uk/software/figtree) e Inkscape (<https://inkscape.org/pt/>).

4 RESULTADOS

4.1 Conjunto de dados

Um total de 32 espécies distintas de badnavírus com sequências genômicas completas foram obtidas a partir do GenBank (Tabela 1). Como algumas espécies foram representadas por mais de um isolado/sequência, o conjunto de dados compreendeu 71 isolados, com *Citrus yellow mottle virus* (CYMV) tendo o maior número de representantes (sete isolados), e 15 espécies sendo representadas por apenas um isolado/sequência (Tabela 1).

4.2 Demarcação de espécies

Na análise de comparações pareadas para o conjunto de dados RT/RNaseH parcial (579pb), mais amplamente utilizado na identificação de espécies de badnavírus, foi observado que as porcentagens de identidade de nucleotídeos entre as diferentes espécies variou entre 57,0 a 82,2%, portanto, algumas comparações excederam o limiar de $\geq 80\%$ de identidade para demarcação de espécies dentro do gênero *Badnavirus* (GEERING; HULL, 2012). Isolados de SCBGAV compartilharam 80-82% e 81-82% de identidade de nucleotídeos com isolados de BSOLV e BSCAV, respectivamente. Entretanto, isolado de BSOLV e BSCAV exibiram porcentagens máximas de identidade nucleotídica de 76-77%. O isolado de SCBBBV compartilhou 78-80% e 82% de identidade de nucleotídeos com isolados de SCBIMV e SCBMOV, respectivamente. Já os isolados de SCBIMV e SCBMOV exibiram porcentagens máximas de identidade nucleotídica de 79-80%. Para todas as espécies representadas por mais de um isolado/sequência, porcentagens de identidade foram superiores a 80% dentro de uma mesma espécie (Figura 2). Resultados similares foram observados para o conjunto de dados RT/RNaseH completa (1020pb) (Figura 3).

Quando analisado o conjunto de dados para ORF III (onde estão presentes os domínios RT/RNaseH), pode-se observar as mesmas incongruências taxonômicas para os isolados de SCBGAV, BSOLV e BSCAV. Entretanto, isolados de SCBBBV, SCBIMV e SCBMOV apresentaram porcentagens de identidade nucleotídicas abaixo de 80% (dados não mostrados). Para o conjunto de dados de genomas completos, todas as comparações entre espécies distintas de badnavírus foram abaixo de 80% de identidade de nucleotídeos,

com identidade máxima de 79,5% entre isolados de SCBGAV comparados com isolado de BSOLV e BSCAV (dados não mostrados). Finalmente, para o conjunto de dados ORF I, porcentagens de identidade $\geq 80\%$ foram observadas entre isolados de SCBGAV, BSOLV e SCBGAV e entre SCBIMV e SCBMOV (dados não mostrados). Entretanto, para a ORF II, todas as comparações foram abaixo de 78% de identidade de nucleotídeos (dados não mostrados).

4.3 Relacionamento filogenético

Como a região RT/RNaseH é amplamente utilizada para determinação da posição taxonômica de membros do gênero *Badnavirus* (GEERING; HULL, 2012), árvores filogenéticas de inferência Bayesiana foram obtidas para os conjuntos de dados genoma completo, ORFs I, II e III, e domínios RT/RNaseH (parcial e completo), no intuito de demonstrar se o relacionamento filogenético observado para os conjunto de dados RT/RNaseH parcial (579pb) e completo (1020pb) refletem o agrupamento inferido para o genoma completo dos isolados de badnavírus,

Para o conjunto de dados RT/RNaseH completo, foram observados os três principais grupos de badnavírus (*Badnavirus Groups* 1, 2 e 3) descritos por Muller et al. (2011). Adicionalmente, houve a formação de um grupo bem definido e relacionado com o *Badnavirus Group* 3, aqui denominado *Badnavirus Group* 4 (Figura 4). Quando analisada a árvore filogenética para RT/RNaseH parcial (como utilizado por MULLER et al., 2011), os quatro principais grupos de badnavírus são ainda observados, com algumas incongruências topológicas (Figura 5). Entretanto, a definição e valores de sustentação dos ramos foram bem menores, indicando que o conjunto de dados RT/RNaseH parcial não tem sinal filogenético suficiente para o correto agrupamento de todos os isolados de badnavírus. Alguns isolados pertencentes ao *Badnavirus Group* 1, para a filogenia da RT/RNaseH completa, foram posicionados fora deste grupo na filogenia para a RT/RNaseH parcial (Figura 5).

Na árvore filogenética para o genoma completo, foi observada ainda a formação dos quatro principais grupos presentes nas filogenias para RT/RNaseH (completa e parcial), porém com algumas incongruências quanto ao relacionamento filogenético desses grupos (Figura 6). Para o conjunto de dados genoma completo, o *Badnavirus Group* 3 formou um grupo irmão com o *Badnavirus Group* 1, sendo formados predominantemente

por badnavírus infectando cana-de-açúcar e banana (Figura 6). Resultados similares foram observados para a filogenia da ORF III (Figura 7), a qual representa aproximadamente 80% do genoma completo dos badnavírus. Estes resultados reforçam a ideia de que badnavírus infectando cana-de-açúcar e banana são proximamente relacionados, como indicado pelas análises de comparação de sequências.

As filogenias para as ORFs I e II apresentaram topologia bastante distinta quando comparadas com as árvores para os outros conjuntos de dados, além de baixos valores de sustentação dos ramos formados (Figuras 8-9). Estes resultados sugerem que os conjuntos de dados ORFs I e II possuem sinal filogenético insuficiente para inferir sobre o relacionamento das diversas espécies de badnavírus analisadas aqui.

4.4 Eventos de recombinação

Análise reticulada para o genoma completo e ORFs I, II e III revelou forte evidência de eventos de recombinação contribuindo para o processo evolutivo dos isolados de badnavírus, sendo observados possíveis eventos de recombinação intra e interespecífico (Figura 10-13). Estes eventos foram mais pronunciados entre os badnavírus infectando cana-de-açúcar e banana em todos os conjuntos de dados analisados (Figura 10-13). As redes formadas para os dados de genoma completo e ORF III foram similares (Figuras 10-11), e bastante divergentes das redes para as ORFs I e II (Figuras 12-13), como observado na análise filogenética. Além do efeito de recombinação, foi evidenciado o efeito de mutação na diversificação dos membros do gênero *Badnavirus*, indicado pelos ramos longos associados a diversos isolados (Figura 10-13).

Para adicionalmente investigar possíveis parentais e locais (*breakpoints*) de recombinação, o conjunto de dados para o genoma completo foi analisado utilizando o pacote RDP4. Como evidenciado pela análise reticulada, um complexo padrão de recombinação foi observado (Tabela 2). Pelo menos 13 eventos independentes de recombinação foram detectados entre os isolados de badnavírus, com sete desses eventos ocorrendo em espécies infectando cana-de-açúcar e banana (Tabela 2). A maioria dos *breakpoints* de recombinação foi localizada dentro da ORF III e Região Intergênica, com apenas quatro eventos com *breakpoints* na ORF I (Tabela 2).

5 DISCUSSÃO

Os badnavírus vêm sendo relatados em culturas economicamente importantes em todo o mundo (SILVA et al., 2015; JAMES et al., 2011a; KOUAKOU et al., 2012; DEESHMA; BHAT, 2015; ENI et al., 2008; JOHNSON et al., 2012), podendo induzir grandes perdas na qualidade e produtividade destas culturas, como descrito em cacau na Costa do Marfim (KOUAKOU et al., 2012) e banana em Uganda (DAHAL et al., 2000; DANIELLS et al., 2001). Além disso, isolados de badnavírus apresentam alta diversidade e variabilidade genética (MULLER et al., 2011; KARUPPAIAH et al., 2013; GUIMARÃES et al., 2015; SILVA et al., 2015).

Sequências nucleotídicas da região parcial dos domínios RT/RNaseH (579pb), presentes na ORF III, vem sendo amplamente utilizadas para identificação de espécies pertencentes ao gênero *Badnavirus* (GUIMARÃES et al., 2015; SILVA et al., 2015; MULLER et al., 2011). Oligonucleotídeos universais descritos por Yang et al. (2003) são bastante eficientes na amplificação dessa região genômica nos badnavírus. Entretanto, tais oligonucleotídeos são incapazes de distinguir entre sequências virais endógenas integradas ao genoma do hospedeiro e sequências episomais responsáveis pelo processo infeccioso (JAMES et al., 2011a). Além disso, a existência de diferentes espécies de badnavírus proximamente relacionadas tem dificultado a classificação taxonômica dentro desse grupo de vírus (FIGUEIREDO et al., 2006; GEERING et al., 2000, 2005b; HARPER et al., 2005; HARPER; HULL, 1998; LHEREUX et al., 2007).

No presente trabalho, análise de comparações pareadas de sequências nucleotídicas de amplo conjunto de dados de diferentes badnavírus mostrou que o limiar de $\geq 80\%$ de identidade para a região RT/RNaseH [parcial (579pb) ou completa (1020pb)] (GEERING; HULL, 2012) permitiu o posicionamento taxonômico da maioria das espécies analisadas. Entretanto, não foi possível distinguir entre algumas espécies de badnavírus infectando banana e cana-de-açúcar, como observado em trabalhos publicados anteriormente (LOCKHART; OLSZEWSKI, 1993; AUTREY, 1991; MULLER et al., 2011; KARUPPAIAH et al., 2013; SILVA et al., 2015). Foi anteriormente observado que as regiões intergênicas e RNaseH de isolados de Sugarcane bacilliform viruses (SCBV) e Banana streak viruses (BSV) compartilham até 90% de identidade nucleotídica, reforçando a hipótese de que esses isolados pertenceriam a uma única espécie de badnavírus (LOCKHART; AUTREY, 1991; BRAITHWAITE et al., 1997), e explicaria a aparição

frequente de BSV em cana-de-açúcar. Interessantemente, mesmo analisando uma região genômica mais ampla como a ORF III (representa cerca de 80% do genoma completo dos badnavírus), foram obtidos resultados similares aos observados para a RT/RNaseH.

Na tentativa de solucionar tais problemas taxonômicos observados para alguns badnavírus, principalmente infectando cana-de-açúcar e banana, são sugeridas algumas alternativas. Primeiro, considerado o critério de $\geq 80\%$ de identidade para a região conservada RT/RNaseH na demarcação de espécies de badnavírus, isolados de BSV e SCBV muito proximamente relacionados (identidade nucleotídica $\geq 80\%$) seriam considerados como uma única espécie viral. Segundo, um novo limiar de identidade nucleotídica para RT/RNaseH poderia ser estabelecido. Neste caso, foi observado que os isolados analisados compartilharam máximo de identidade de 82%, portanto, critérios acima desse valor poderiam ser adotados. Entretanto, seria interessante analisar um maior número de sequências/isolados para a região RT/RNaseH antes da definição de um novo critério. Terceiro, dada à ineficiência para diferenciação de algumas espécies de badnavírus utilizando critério de identidade de sequências da RT/RNaseH, o posicionamento taxonômico poderia ser realizado através de sequências do genoma completo, como observado para outros grupos de fitovírus (BRWON et al., 2015).

Análise filogenética baseada na região RT/RNaseH parcial (529pb) mostrou que não há uma separação clara entre isolados de SCBV e BSV em grupos distintos de badnavírus (MULLER et al., 2011). A estreita relação entre isolados de SCBV e BSV, e a estrutura polifilética de BSV (GAYRAL; ISKRA-CARUANA, 2009) são fortes argumentos suportando a hipótese de que banana e cana-de-açúcar teriam um ancestral comum, ou ainda que banana pode ter sido o hospedeiro original para SCBV, bem como para outros badnavírus que agrupam com BSV (GAYRAL; ISKRA-CARUANA, 2009).

O relacionamento filogenético observado para todos os conjuntos de dados analisados corroboram com as análises de comparações pareadas de sequências nucleotídicas, onde se pode observar a não separação de isolados de SCBV e BSV em filogrupos distintos. É importante notar que, apesar de algumas incongruências topológicas, as filogenias para RT/RNaseH relembram o agrupamento inferido para o genoma completo dos membros do gênero badnavírus. Além da formação bem definida dos três grupos de badnavírus descritos por Muller et al. (2011), um quarto grupo filogenético é proposto aqui, o *Badnavirus Group 4*, o qual representa um clado separado dos outros badnavírus. Assim como observado por Muller et al. (2011), quando analisadas

apenas as filogenias para RT/RNaseH, os clados *Badnavirus Group 1* e *Badnavirus Group 3*, compostos predominantemente por espécies que infectam cana-de-açúcar e banana, são mais distantemente relacionados, com o *Badnavirus Group 3* sendo mais próximo do *Badnavirus Group 4* na árvore filogenética para RT/RNaseH completa. Entretanto, quando analisadas as filogenias para ORF III e genomas completos, fica evidente a proximidade genética entre os *Badnavirus Groups 1* e *3*, onde formaram grupos irmãos. Estes resultados reforçam a ideia de que os badnavírus infectando cana-de-açúcar e banana podem ter um histórico evolutivo comum.

A alta variabilidade genética observada em badnavírus tem sido atribuída à replicação inacurada por transcrição reversa presente em *Caulimoviridae* (BOUSALEM et al., 2009). Transcriptases reversas são conhecidas por serem altamente tendenciosas a produzirem erros em retrovírus e retroelementos para os quais a taxa de fidelidade foi estimada (DRAKE et al., 1998; SVAROVSKAIA et al., 2003). Embora não seja conhecida a medida precisa da fidelidade da transcriptase reversa (RT) codificada pelos vírus pertencentes às famílias *Caulimoviridae* e *Hepadnaviridae*, acredita-se que as altas taxas de mutação encontradas para esses vírus se deve à perda da atividade de correção de erros de suas RTs (DRAKE et al., 1998; SVAROVSKAIA et al., 2003). Outro mecanismo evolutivo que influencia a diversidade e variabilidade genética dos badnavírus é a recombinação (GOVIND et al., 2014).

Análise de evolução em rede mostrou forte evidência de eventos de recombinação contribuindo para o processo evolutivo de algumas espécies de badnavírus, principalmente aqueles infectando cana-de-açúcar e banana. Estes resultados foram reforçados pela análise de recombinação para o conjunto de dados de genomas completos, onde um complexo padrão de eventos de recombinação foi detectado. Novamente, a maioria dos eventos envolveu isolados de SCBV e BSV, com *breakpoints* de recombinação ocorrendo predominantemente na ORF III e região intergênica.

Um das premissas básicas para que haja recombinação é a ocorrência de infecções mistas, inclusive com presença dos vírus no núcleo de uma mesma célula vegetal (ZHOU et al., 1997; FONDONG et al., 2000; GRAHAM et al., 2010). Infecção cruzada entre badnavírus em cana-de-açúcar e banana são comumente relatadas, porém parece ser um evento incomum para badnavírus infectando outros hospedeiros (KOUAKOU et al., 2012; KARUPPAIAH et al., 2013; SILVA et al., 2015). Portanto, isto poderia explicar, em partes, a origem recombinante comum entre isolados de SCBV e BSV, como evidenciado

pela análise reticulada. Adicionalmente, eventos frequentes de recombinação poderiam explicar a alta diversidade de espécies de badnavírus comumente encontrada nesses hospedeiros (SILVA et al., 2015; SMITH et al., 1996; BRAITHWAITE et al., 1997). Além da recombinação contribuindo para diversificação dos badnavírus, mutação parece realmente ser o mecanismo evolutivo mais importante atuando na evolução desse grupo de vírus.

6 CONCLUSÕES

- Mutação é o principal mecanismo evolutivo atuando na evolução dos membros do gênero *Badnavirus*;
- Isolados de SCBV e BSV possuem uma estreita relação filogenética;
- A região RT/RNaseH parcial (579pb) é suficiente para demarcação da maioria das espécies de badnavírus atualmente descritas, porém o limiar de $\geq 80\%$ de identidade nucleotídica sozinho é ineficiente para diferenciar todas as espécies de SCBV e BSV.

Tabela 1. Sequências de badnavirus obtidas a partir do banco de dados não-redundantes GenBank.

Espécie/Acrônimo	#acesso no GenBank
<i>Banana streak VN virus</i> /BSVNV	DQ092436
BSVNV	AY750155
BSVNV	KJ013510
<i>Banana streak IM virus</i> /BSIMV	HQ659760
BSIMV	KJ013508
BSIMV	HQ593112
<i>Banana streak OL virus</i> /BSOLV	DQ859899
BSOLV	KJ013506
BSOLV	DQ451009
BSOLV	JQ409539
BSOLV	JQ409540
<i>Sugarcane bacilliform GA virus</i> /SCBGAV	FJ824813
SCBGAV	FJ824814
<i>Banana streak CA virus</i> /BSCAV	HQ593111
BSCAV	KJ013511
<i>Banana streak UA virus</i> /BSUAV	HQ593107
<i>Kalanchoë top-spotting virus</i> /KTSV	AY180137
<i>Banana streak GF virus</i> /BSGFV	AY493509
BSGFV	KJ013507
<i>Banana streak mysore virus</i> /BSMyV	KJ013509
BSMyV	KF724854
BSMyV	AY805074
BSMyV	KF724856
BSMyV	KF724855
<i>Sugarcane Bacilliform GD virus</i> /SCBGDV	FJ439817
<i>Cycad leaf necrosis virus</i> /CLNV	EU853709
<i>Pineapple bacilliform comosus virus</i> /PBCV	GQ398110
PBCV	GU121676
<i>Cacao swollen shoot virus</i> /CSSV	L14546
CSSV	AJ534983
CSSV	AJ609019
CSSV	AJ609020
<i>Cacao swollen shoot Togo A virus</i> /CSSTAV	AJ781003
<i>Cacao swollen shoot Côte d'Ivoire D virus</i> /CSSCIDV	JN606110
<i>Hybiscus bacilliform virus</i> /HBV	KF875586
<i>Citrus yellow mosaic virus</i> /CYMV	JN006805
CYMV	AF347695
CYMV	EU708317
CYMV	JN006806
CYMV	EU489744

CYMV	EU489745
CYMV	EU708316
<i>Dioscorea bacilliforma</i> SN virus/DBSNV	DQ822073
<i>Piper yellow mottle virus</i> /PYMV	KC808712
PYMV	KJ873041
PYMV	KJ873042
PYMV	KJ873043
<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>Sugarcane bacilliform IM virus</i> /SCBIMV	JN377533
SCBIMV	JN377536
SCBIMV	JN377537
SCBIMV	AJ277091
<i>Sugarcane Bacilliform BB Virus</i> /SCBBBV	JN377535
<i>Sugarcane bacilliform MO virus</i> /SCBMOV	JN377534
SCBMOV	M89923
<i>Commelina yellow mottle virus</i> /CoYMV	X52938
<i>Dracaena mottle virus</i> /DMV	DQ473478
<i>Pelargonium vein banding virus</i> /PVBV	GQ428155
<i>Gooseberry vein banding associated virus</i> /GVBV	JQ316114
GVBV	HQ852249
GVBV	HQ852250
GVBV	HQ852251
GVBV	HQ852248
<i>Rubus yellow net virus</i> /RYNV	KM078034
RYNV	KM078034
<i>Grapevine vein clearing virus</i> /GVCV	KJ725346
GVCV	JF301669
GVCV	JF301669
<i>Yacon necrotic mottle virus</i> /YNMV	KM229702

Tabela 2. Eventos de recombinação detectados entre isolados de badnavírus obtidos a partir do Genbank.

Evento	Recombinante	Breakpoints*		Parentais		Métodos†	P value‡
		Inicial	Final	Minor	Major		
1	JN006805_CYMV	1277	274	AF347695_CYMV	JN006806_CYMV	RGBMCS ³	5.27E-97
2	L14546_CSSV AJ534983_CSSV	3	377	AJ781003_CSSV	AJ609020_CSSV	RGBMCS	2.28E-26
3	EU489744_CYMV EU489745_CYMV	5546	6674	Unknown	JN006806_CYMV	RMC ³	2.82E-21
4	JN377533_SCBIMV	3148	3242	Unknown	JN377537_SCBIMV	RGBMC	6.02E-20
5	HQ593107_BSUAUV	450	681	HQ593109_BSULV	KJ013511_BSCAV	RGBMCS	3.04E-18
6	KJ873043_PYMV	2111 (?)	3539	Unknown	KJ873042_PYMV	RGBMCS	3.34E-23
7	KJ873042_PYMV	5825	6499 (?)	KJ873043_PYMV	KC808712_PYMV	RGBMCS	4.96E-18
8	JN377534_SCBMOV	400	515	AJ277091_SCBIMV	M89923_SCBMOV	RGBMC	7.68E-12
9	AY493509_BSGFV KJ013507_BSGFV	403	1868	KF724855_BSM _y V	AY180137_KTSV	RBMCS	1.87E-09
10	HQ593110_BSUMV HQ593108_BSUIV HQ593109_BSULV	285	469	HQ593111_BSCAV	Unknown	RGBMCS	3.63E-08
11	JN377533_SCBIMV JN377537_SCBIMV	6385	396 (?)	JN377534_SCBMOV	JN377535_SCBBBV	RBMS	5.41E-09
12	JQ409539_BSOLV DQ859899_BSOLV KJ013506_BSOLV DQ451009_BSOLV JQ409540_BSOLV	890	1731	Unknown	HQ593111_BSCAV	RBMCS	7.06E-05
13	AY180137_KTSV	1652	1891	Unknown	HQ593107_BSUAUV	RBMCS	1.23E-03

Figura 2. Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de seqüências nucleotídicas da região completa dos domínios RT/RNaseH (579 pb) de isolados de Badnavirus disponíveis no GenBank.

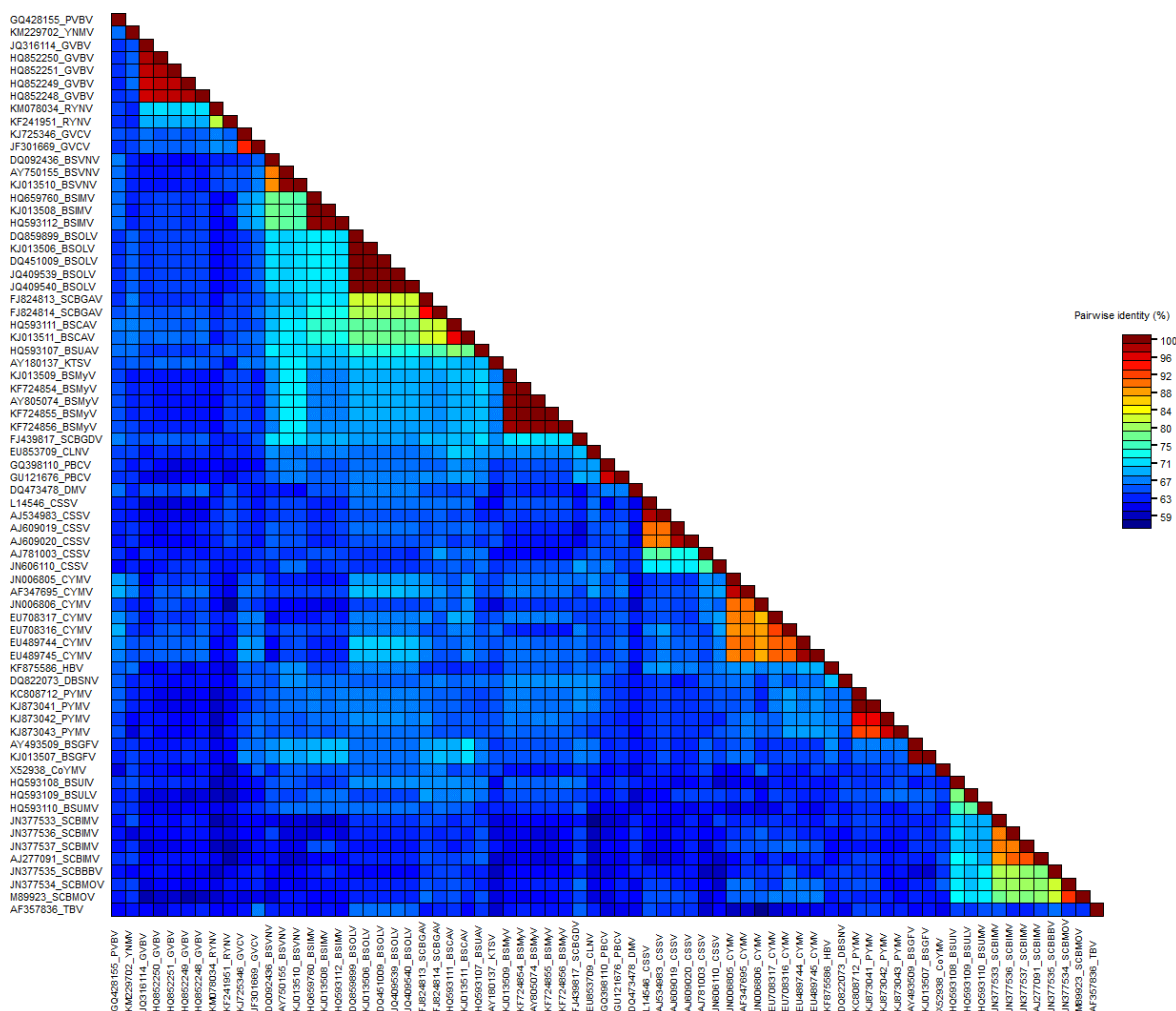


Figura 3. Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de seqüências nucleotídicas da região parcial dos domínios RT/RNaseH (1020 pb) de isolados de Badnavirus disponíveis no GenBank.

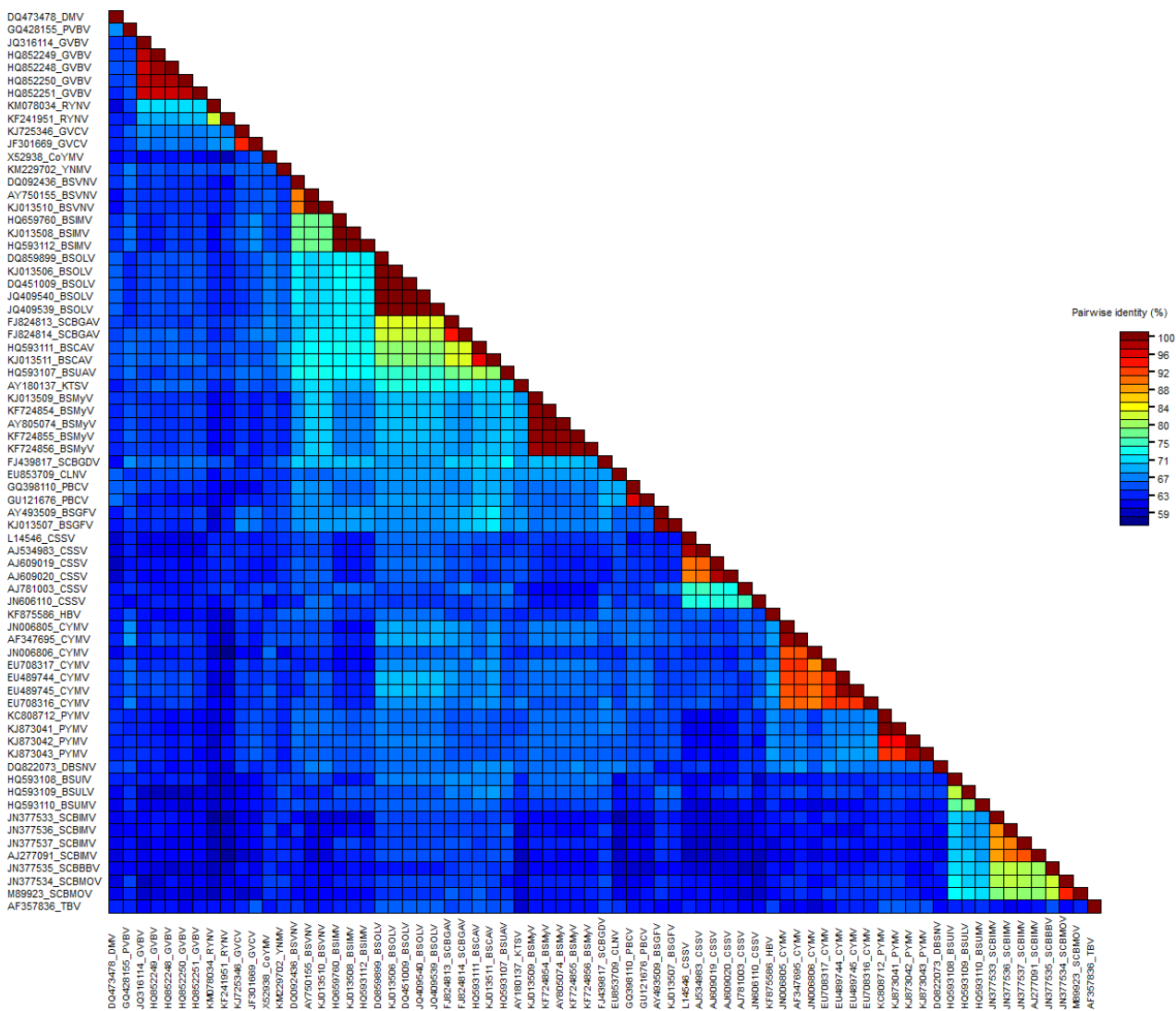


Figura 4. Análise filogenética de inferência Bayesiana baseada no conjunto de dados da RT/RNaseH completa de espécies de badnavírus. *Badnavirus Groups 1* (cinza), 2 (verde), 3 (azul) e 4 (laranja).

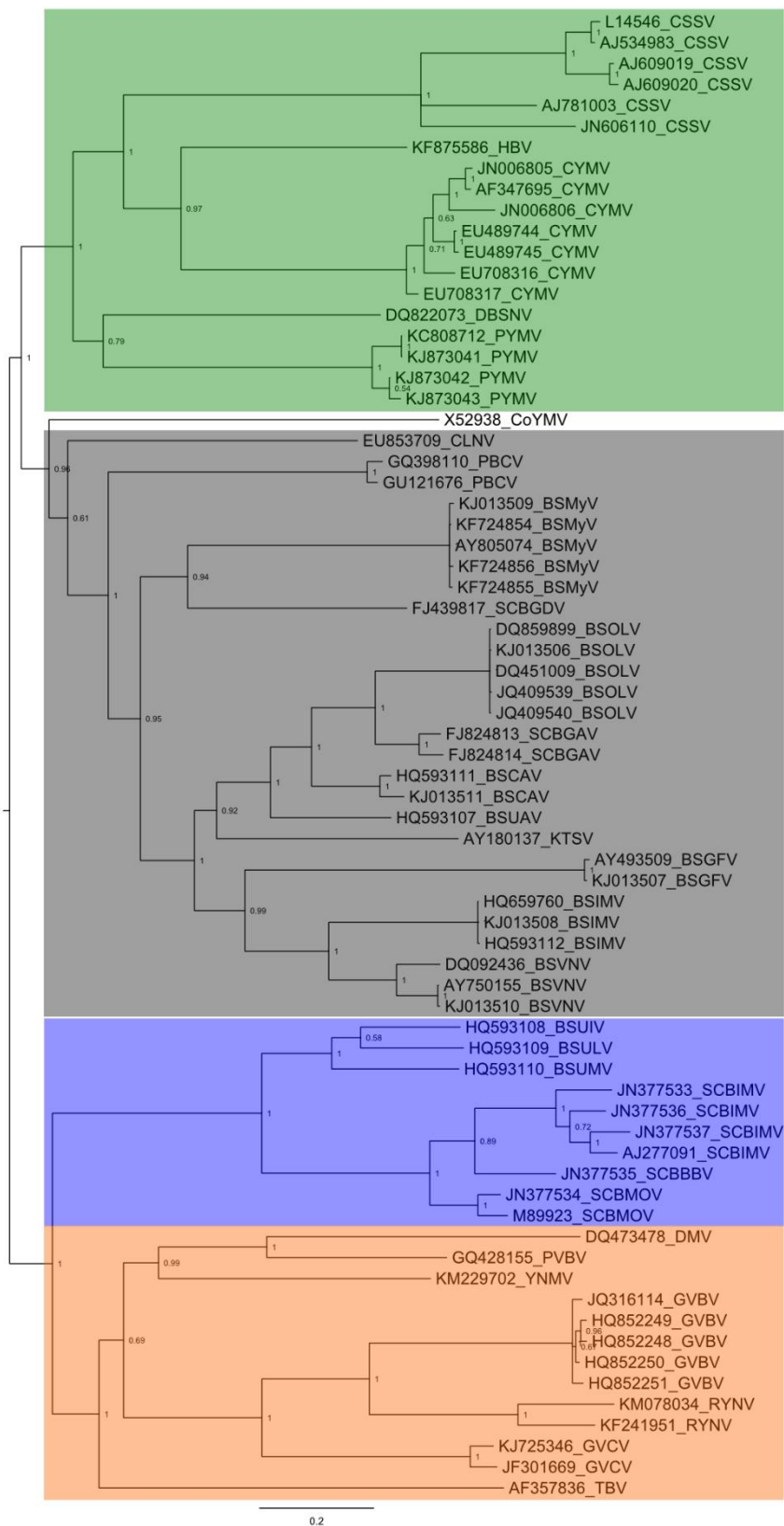


Figura 5. Análise filogenética de inferência Bayesiana baseada no conjunto de dados da RT/RNaseH parcial de espécies de badnavírus. *Badnavirus Groups 1* (cinza), 2 (verde), 3 (azul) e 4 (laranja).

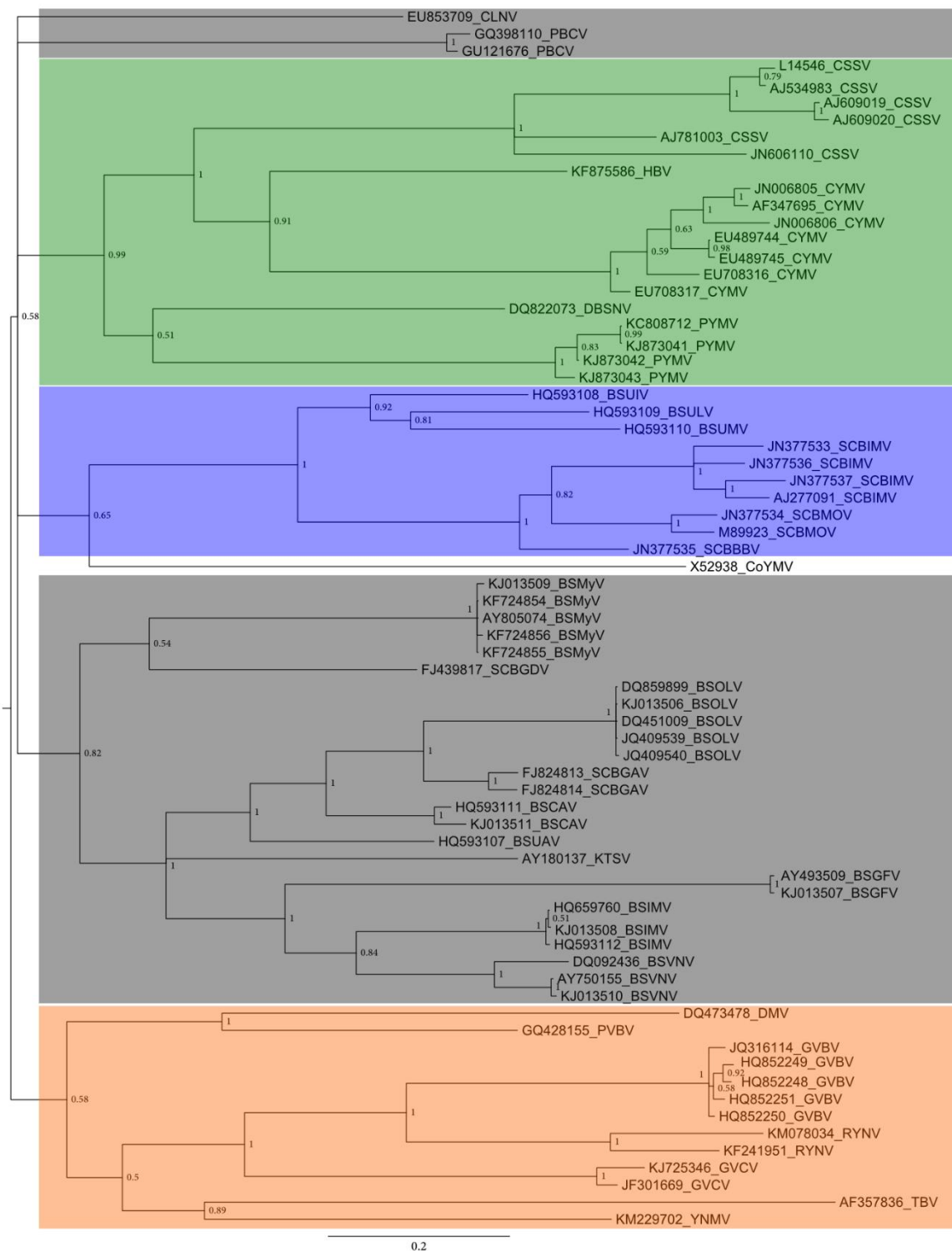


Figura 6. Análise filogenética de inferência Bayesiana baseada no conjunto de dados do genoma completo de espécies de badnavírus. *Badnavirus* Groups 1 (cinza), 2 (verde), 3 (azul) e 4 (laranja).

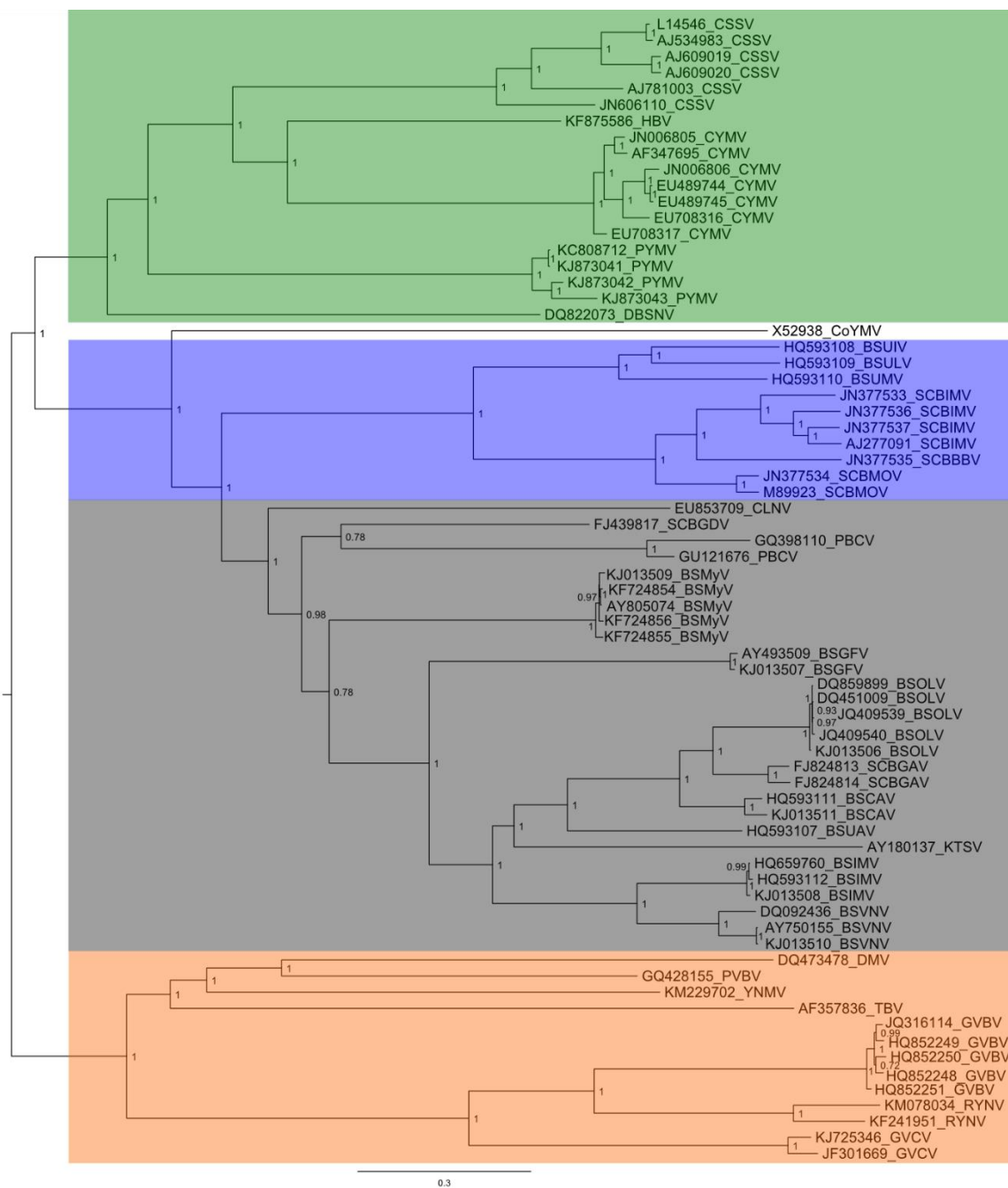


Figura 7. Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF III de espécies de badnavírus. *Badnavirus Groups 1* (cinza), 2 (verde), 3 (azul) e 4 (laranja).

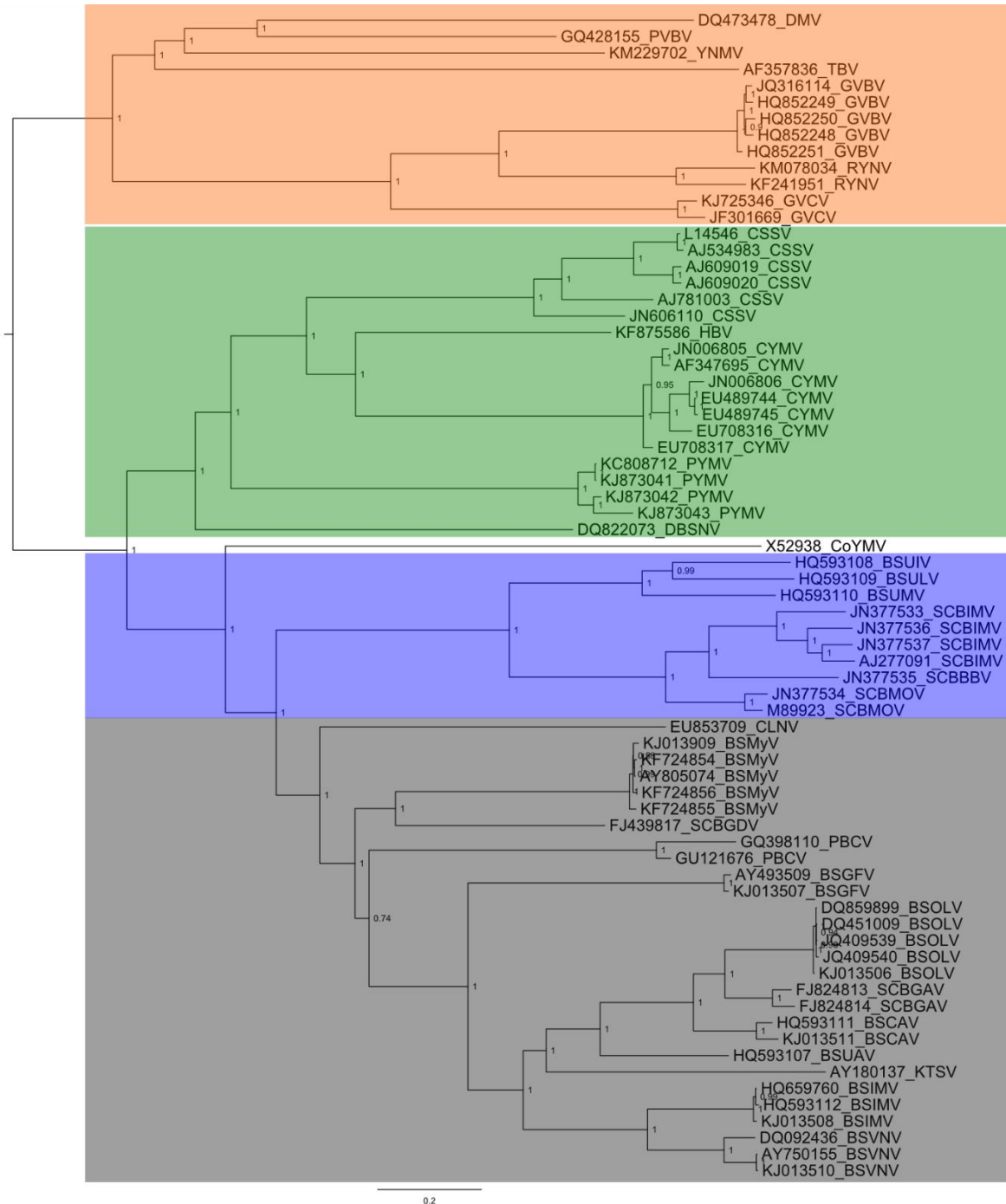


Figura 8. Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF I de espécies de badnavírus. *Badnavirus Groups 1* (cinza), 2 (verde), 3 (azul) e 4 (laranja).

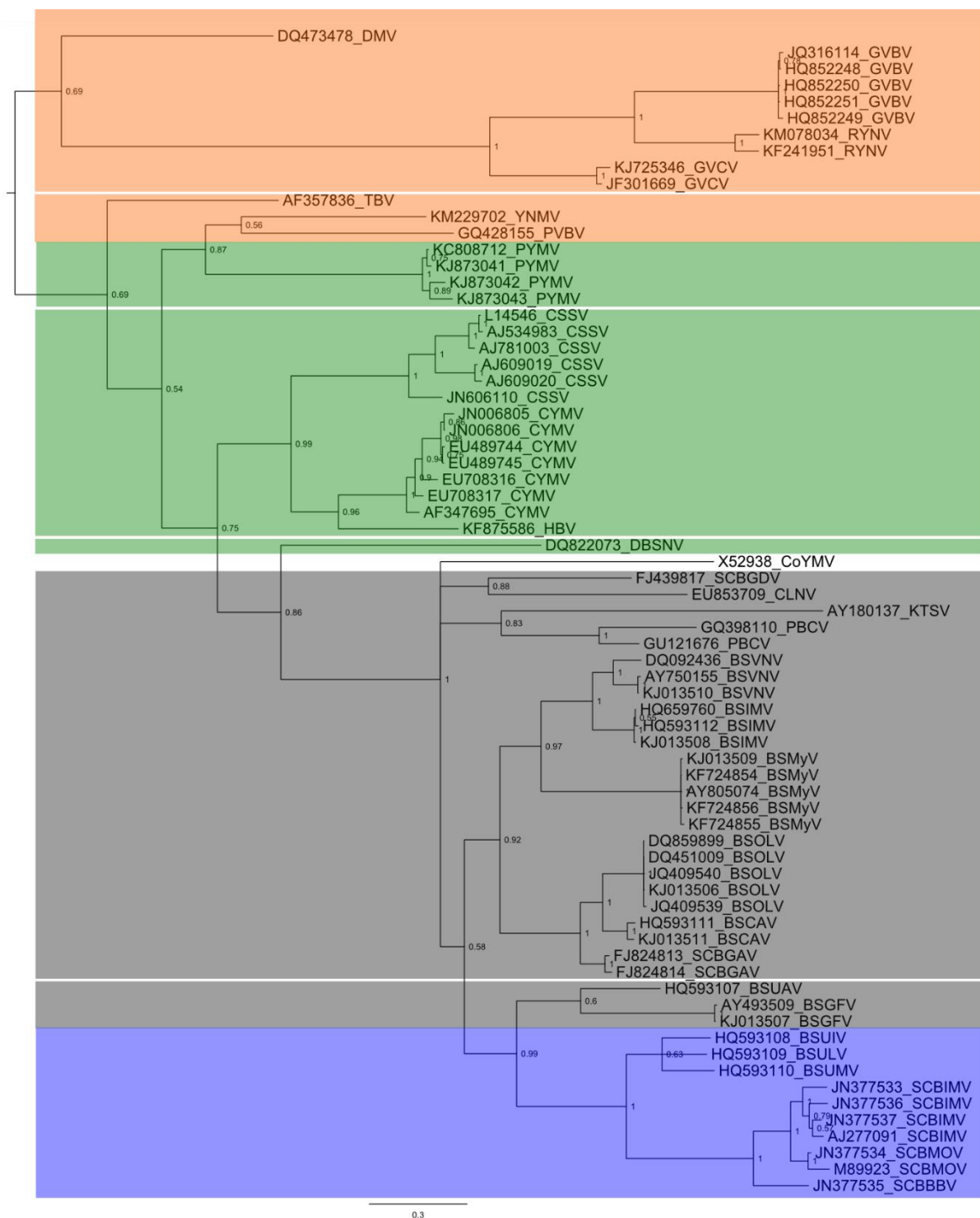


Figura 9. Análise filogenética de inferência Bayesiana baseada no conjunto de dados da ORF II de espécies de badnavírus. *Badnavirus Groups 1* (cinza), 2 (verde), 3 (azul) e 4 (laranja).

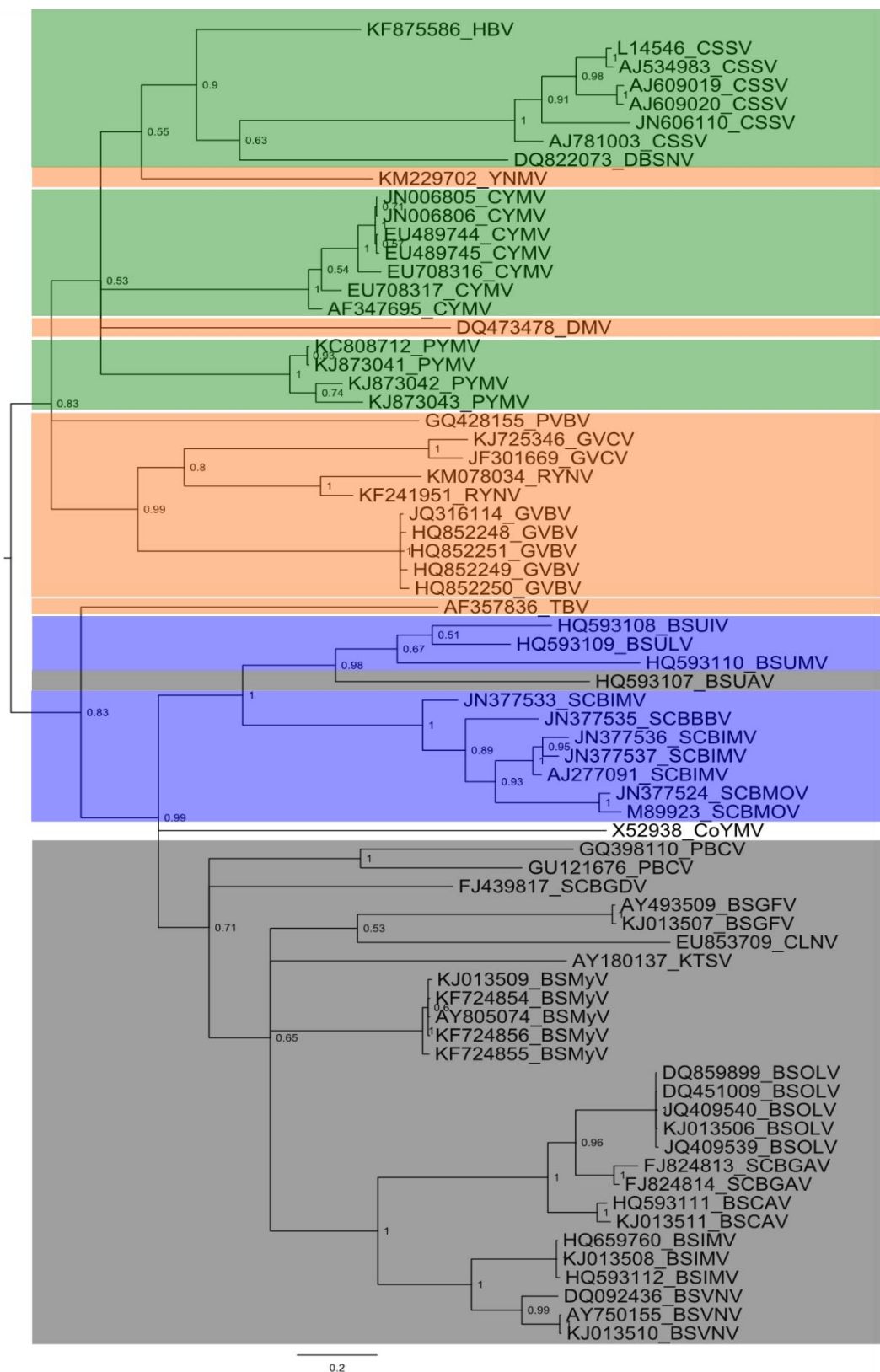


Figura 10. Análise reticulada baseada no genoma completo dos isolados de badnavírus. *Badnavirus Groups 1* (cinza), *2* (verde), *3* (azul) e *4* (laranja).

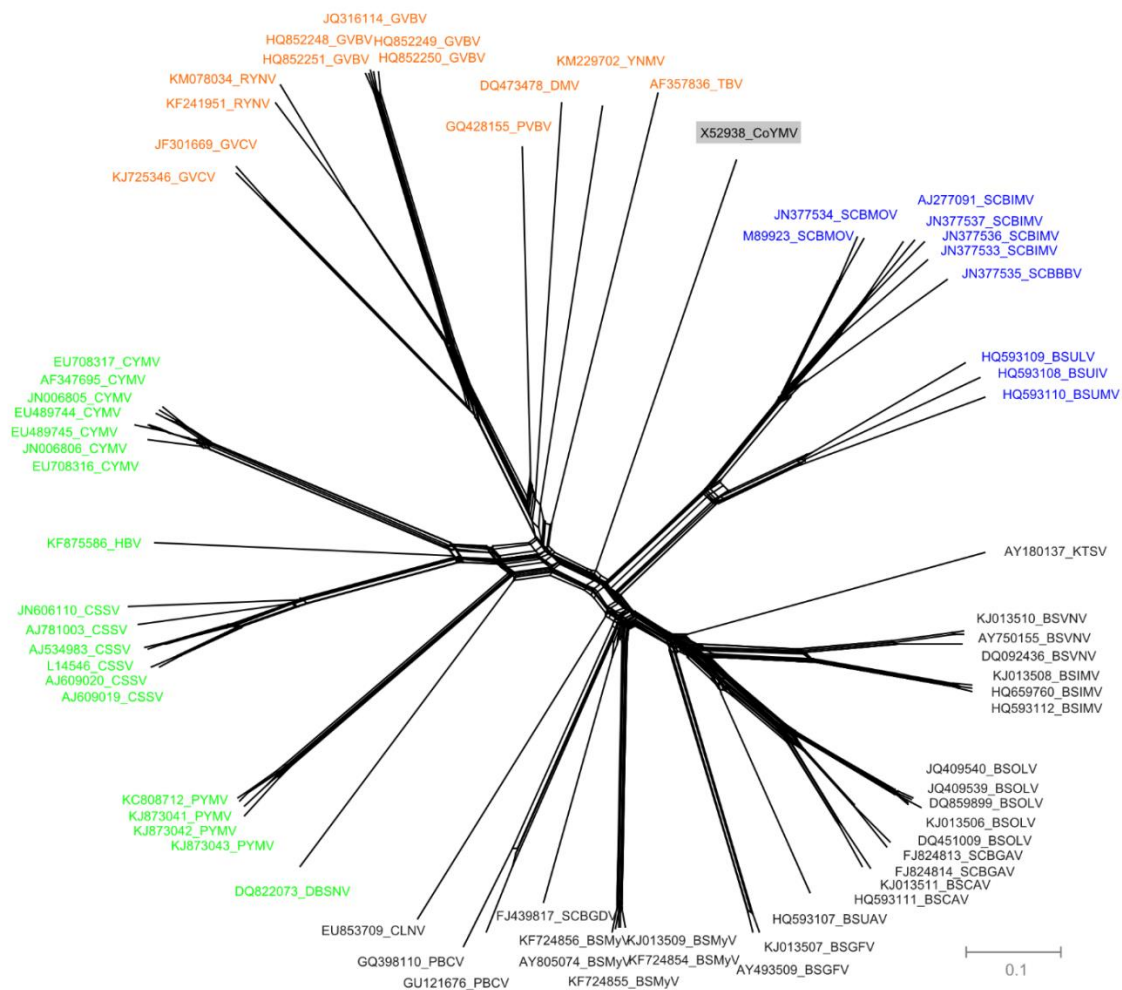


Figura 11. Análise reticulada baseada na ORF 1 dos isolados de badnavírus. *Badnavirus* Groups 1 (cinza), 2 (verde), 3 (azul) e 4 (laranja).

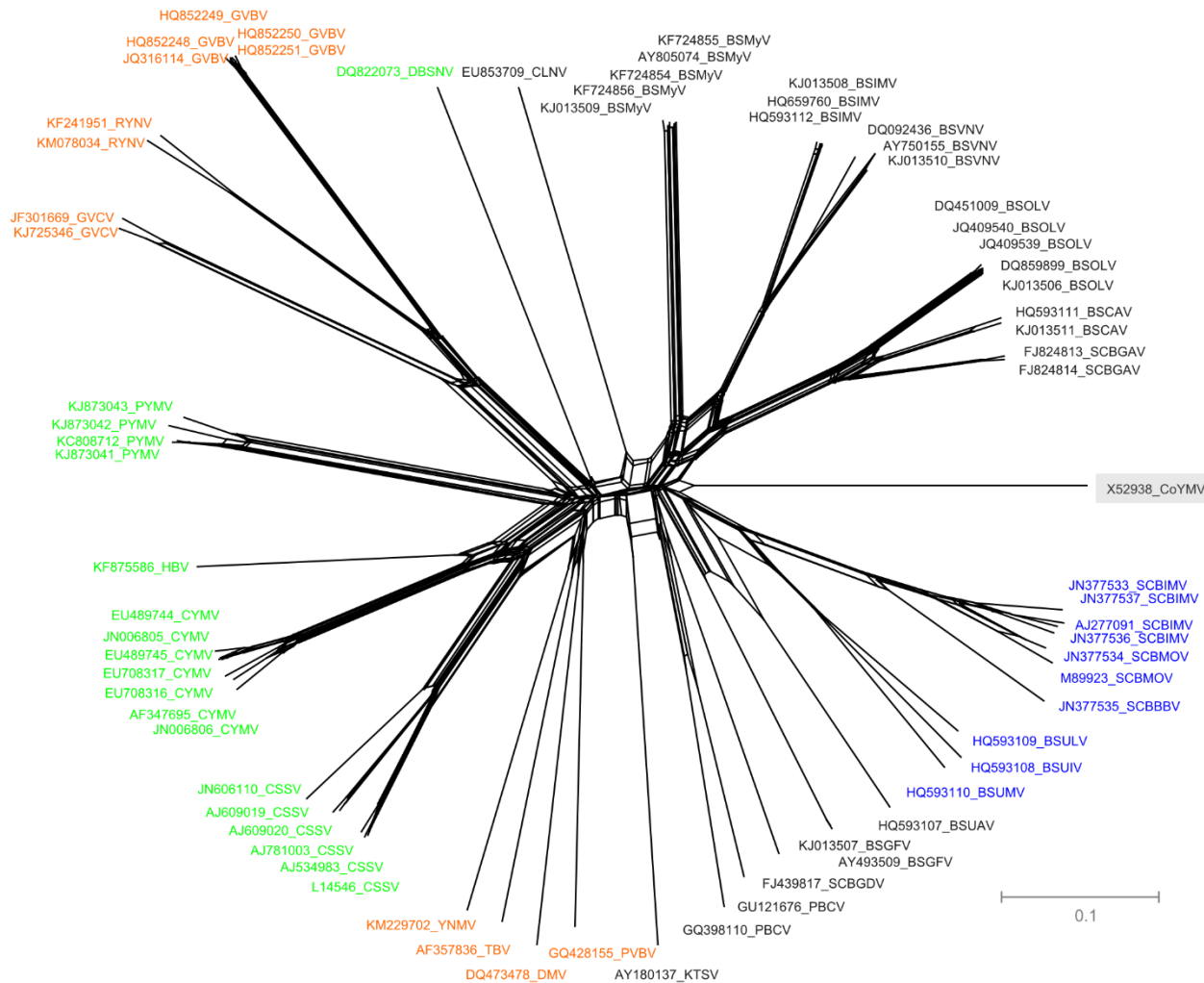


Figura 12. Análise reticulada baseada na ORF 2 dos isolados de badnavírus. *Badnavirus* Groups 1 (cinza), 2 (verde), 3 (azul) e 4 (laranja).

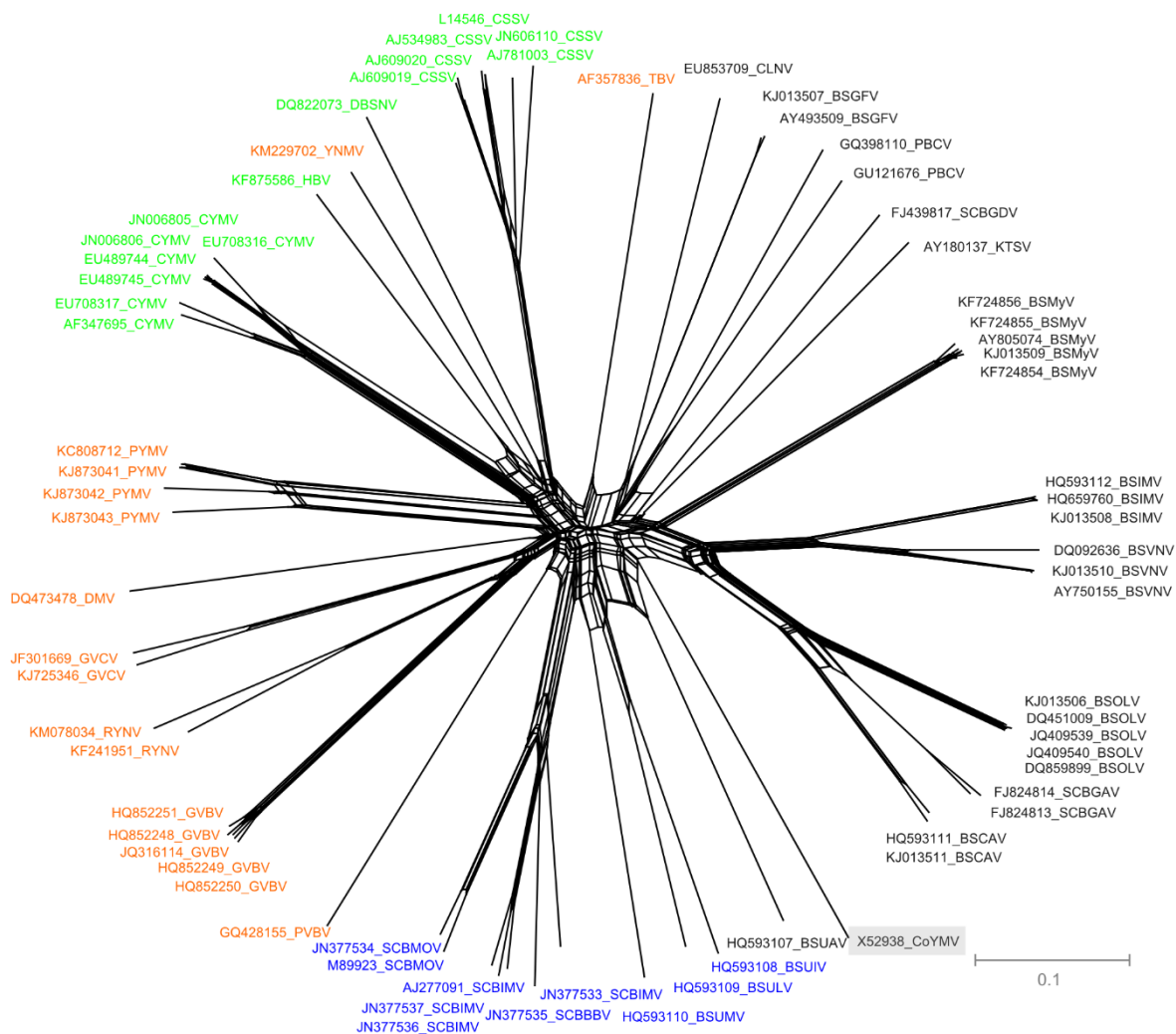
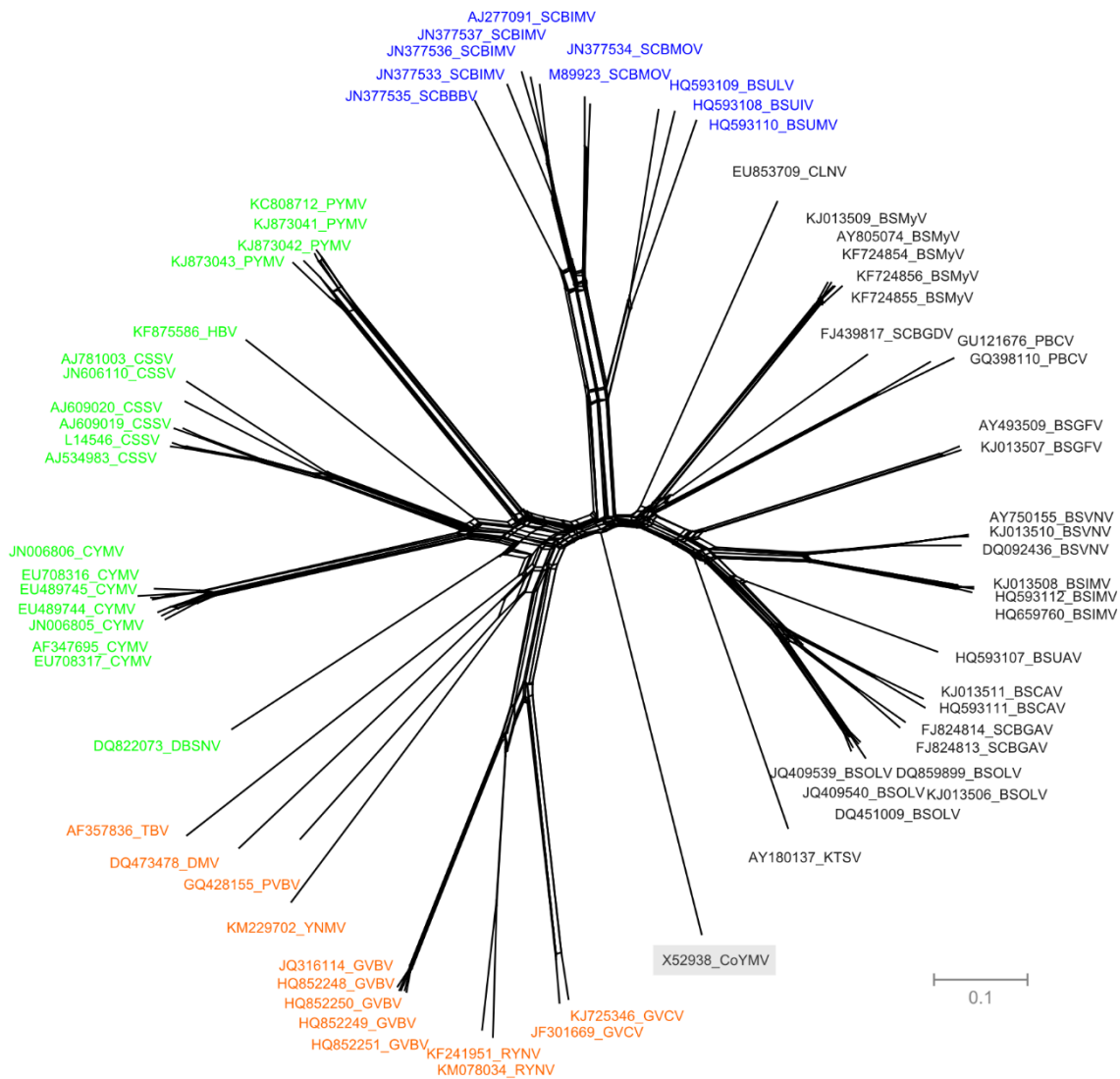


Figura 13. Análise reticulada baseada na ORF 3 dos isolados de badnavírus. *Badnavirus* Groups 1 (cinza), 2 (verde), 3 (azul) e 4 (laranja).



7 REFERÊNCIAS

- AHLAWAT, Y. S. et al. Association of a badnavirus with citrus mosaic disease in India. **Plant Disease**, Saint Paul, v. 80, p. 590-592, 1996.
- ALEXANDRE, M.A.V. et al. 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**. Tese (Doutorado em Fitopatologia) - Universidade Federal Rural de Pernambuco (UFRPE). Recife, PE, 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**, Austria, v. 145, n. 2, p. 275-89, 2000.
- BARANWAL, V.K. et. al. A novel approach for simultaneous detection of Citrus yellow mosaic virus and Citrus greening bacterium by multiplex polymerase chain reaction. **Indian J. Biotech**, v. 4, p. 528-33, 2005.
- BESSE, S. et al. 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**, p. 214-9, 2010.
- BHAT, A.I et al. Association of a badnavirus in black pepper (*Piper nigrum* L.) transmitted by mealybug (*Ferrisia virgata*) in **IndiaCurr**, v. 84, p. 1547-50, 2003.
- 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. et al. 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**, Jaboticabal, v. 33, n. 4, p. 1353-1355, 2011.
- CHABANNES, M. et al. Three infectious viral species lying in wait in the banana genome. **Journal of Virology**, v. 87, p. 8624–8637, 2013.
- CASTEL, C. et al. Le swollen shoot du cacaoyer au Togo. Les différentes formes de viroses et leurs conséquences économiques. **Café, Cacao The**, v. 24, p. 131–146, 1980.
- CHENG, Y. H.; YANG, J. S.; YEH, S. D. Efficient transformation of papaya by coat protein gene of papaya ringspot virus mediated by agrobacterium following liquid-phase wounding of embryogenic tissues with caborundum. **Plant Cell Reports**, v. 16, p. 127-132, 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.

DAHAL, G. O. R. et al. 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. A. P. et al. Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of Banana streak virus integrated in the genome of FHIA 21 hybrid (*Musa* AAAB). **Archives of Virology**, v.146, p. 2179-2190, 2001.

DANIELLS, J. B. et al. The effect of banana streak virus on the growth and yield of dessert bananas in tropical Australia. **Annals of Applied Biology**, Warwick, v. 139, p. 51-60, Mar. 2001.

DAVINO, S.; NAPOLI, C.; DELLACROCE, C.; MIOZZI, L.; NORIS, E.; DAVINO, M.; ACCOTTO, G. P. Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. **Virus Research**, v. 143, n.1, p.15-23, 2009.

DE KOCHKO, A. et al. *Cassava Vein Mosaic Virus* (CsVMV), type species for a new genus of plant double stranded DNA viruses. **Archives of Virology**, v. 143, p.945-962, 1998.

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. et al. Development of realtime PCR for the rapid detection of episomal Banana streak virus (BSV). **Plant Disease**, v.87, p.33-8, 2003.

DRAKE, R. R. et al. Identification of a nucleotide binding site in HIV-1 integrase. **Proceedings of the National Academy of Sciences**, v. 95, n. 8, p. 4170-5, 1998.

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.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v 32, p. 1792-1797, 2004.

- ENI, A. O. et al. 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. et al. Eight Report of the International committee on Taxonomy of Viruses. Amsterdam. Elsevier. **Virus taxonomy**, 2005.
- FIGUEIREDO, D. et al. 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**, Jaguariúna, v. 33, n. 2, p. 118-123, set. 2007.
- FIRTH, C.; KITCHEN, A.; SHAPIRO, B.; SUCHARD, M. A.; HOLMES, E. C.; RAMBAUT, A. Using time-structured data to estimate evolutionary rates of double-stranded 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 multi-cellular host infection. **PLOS Biol**, v. 3, n. 3, p. 89, 2005.
- FÜTTERER, J. et al. 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, 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**, Austria, v. 153, p. 1599–1604, 2008b.
- GARCÍA-ANDRÉS, S.; MONCI, F.; NAVAS-CASTILLO, J.; MORIONES, E. Begomovirus genetic diversity in the native plant reservoir *Solanum nigrum*: Evidence for the presence of a new virus species of recombinant nature. **Virology**, v. 350, n. 2, p. 433-42, 2006.
- GARCIA-ARENAL, F. et al. 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.
- 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. et al. Characterisation of banana streak mysore virus and evidence that its DNA is integrated in the B genome of cultivated *Musa*. **Archives of Virology**, New York, v. 150, n. 4, p. 787- 796, Apr. 2005b.
- 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 UK. Elsevier Academic Press, p. 429-443, 2012.
- 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.
- 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. et al. 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.
- GRAHAM, A.P.; MARTIN, D.P.; ROYE, M.E. 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 Caribbean. *In: VIRUS GENES*, v. 40, p. 256-266, 2010.
- GUIMARÃES, K.M. et al. Genetic variability of badnaviruses infecting yam (*Dioscorea* spp.) in northeastern Brazil. **Tropical Plant Pathology**, v. 40, p. 111-118, 2015.
- HAGEN, L. S. et al. Nucleotide sequence and genomic organisation of cacao swollen shoot virus. **Virology**, v. 196, p.619-628, 1993.
- HARPER, G. et al. Diversity of banana streak virus isolates in Uganda. **Archives of Virology**, New York, v. 150, n. 12, p. 2407-2420, 2005.
- HARPER, G. et al. Viral sequences integrated into plant genomes. **Annual Review of Phytopathology**, v. 40, p.119-136, 2002.
- HARPER, G.; HULL, R. Cloning and sequence analysis of Banana Streak Virus DNA. **Virus Genes**, v. 17, p. 271-278, 1998.
- HUSON, D. H.; BRYANT, D. Application of phylogenetic networks in evolutionary studies. **Molecular Biology and Evolution**, v. 23, p. 254–267, 2006.

- JAMES, A. P. et al. Molecular characterization of six badnavirus species associated with leaf streak disease of banana in East Africa. **Annals of Applied Biology**, Warwick, v. 158, p. 346-353, 2011b.
- 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, 2011
- 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.
- JESKE, H., LUTGEMEIER, M.; PREISS, W. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. **EMBO Journal**, v.20, p.6158-6167, 2001.
- JOHNSON, A. M. et al. 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, M. et al. Genetic diversity of Sugarcane bacilliform virus isolates infecting Saccharum spp. in India. **Virus Genes**, v. 46, p. 505–516, 2013.
- KEBE, B. I. et al. Le swollen shoot en Côte d’Ivoire: situation actuelle et perspectives. **In: 15th international cocoa research conference San José, Costa Rica**, p. 907-922, 2006.
- KOUAKOU, K. et al. Geographical distribution of *Cacao swollen shoot virus* molecular variability in Côte d’Ivoire. **Plant Disease**, v. 96, p. 1445-1450, 2012.
- LACO, G. S.; BEACHY, R. N. Rice tungro bacilliform virus encodes reverse transcriptase, DNA polymerase and ribonuclease H activities. **Proceedings of the National Academy of Sciences USA**, v. 91, p. 2654-2658, 1994.
- 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**, London, v. 12, p. 3458-68, 2007.
- LIMA, J.S. et al. 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.; AUTREY, L.J.C. Mealy bug transmission of sugar cane bacilliform and sugar cane clostero-like viroses. Paper presented at the third Sugarcane Pathology Workshop of the International Society Technologists, 1991.
- MAFRA, R. C. **Recomendações técnicas para o cultivo do cará**. Brasília: Centro Nacional de Pesquisa de Hortaliças – EMBRAPA, 1986.

- MARTIN D. P. et al. RDP4: Detection and analysis of recombination patterns in virus genomes. **Virus Evolution**, p. 1-5, 2015.
- MCDONALD, B. A. et al. 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**, New York, v. 124, p. 163–180, 2002b.
- 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.
- MILGROOM, M.G. Analysis of population structure in fungal plant pathogens. In: Leslie, J. F. e Frederiksen, R. A. (Eds.) **Disease analysis through genetics and biotechnology**. Iowa State University, Ames. p.213-229, 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**, New York, v. 303, n. 2, p. 317-26, 2002.
- MORILLA, G.; KRENZ, B.; JESKE, H.; BEJARANO, E. R.; WEGE, C. Tête à tête of tomato yellow leaf curl virus and tomato yellow leaf curl sardinia virus in single nuclei. **Journal of Virology**, Washington, v. 78, p. 10715-10723, 2004.
- MUHIRE B., et al. A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus Mastrevirus (family Geminiviridae). **Archives of Virology**, v. 158, p. 1411–1424, 2013.
- MULLER, E. et al. 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. et al. 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.
- PEDRALLI, G. Dioscoreaceae e Araceae: Aspectos taxonômicos, Etnobotânicos e Espécies nativas com Potencial para Melhoramento Genético. In: II SIMPÓSIO NACIONAL SOBRE AS CULTURAS DO INHAME E TARO, 1. 2002. João Pessoa. **Anais**. João Pessoa: EMEPA-PB, 2002.
- PENG, J.; GUO, L.; WANG, G. Development of a LAMP method for rapid detection of Banana streak virus. **Acta Phytopathol. Sin**, v. 42, p. 2, 2012.

PHILLIPS, S. et al. 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. et al. 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**, London, v. 82, p. 655-665, 2001.

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**, London, v. 82, p. 655-665, 2001.

POOGGIN, M. et al. 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.

POSADA, D.; BUCKLEY, T. Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. **Systematic Biology**, v. 53, p. 793–808, 2004.

RAMOS-SOBRINHO, R., et al. Contrasting genetic structure between two begomoviruses infecting the same leguminous Hosts. **Journal of General Virology**, v. 95, p.2540–2552, 2014.

RANNALA, B.; YANG, Z. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. **Journal of Molecular Evolution**, v. 43, p. 304-311, 1996.

RIBEIRO, S. G.; AMBROZEVIĆIUS, L. P.; AVILA, A. C.; BEZERRA, I. C.; CALEGARIO, R. F.; FERNANDES, J. J.; LIMA, M. F.; MELLO, R. N.; ROCHA, H.; ZERBINI, F. M. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. **Archives of Virology**, Austria, v. 148, p. 281-295, 2003.

ROCHA, C. S. et al. Brazilian begomovirus populations are highly recombinant, rapidly evolving, and segregated based on geographical location., **Journal of General Virology** v. 87, p. 5784–5799, 2013.

RODRIGUEZ, E. et al. First report of Pineapple bacilliform comosus virus (PBCoV) and endogenous Pineapple pararetrovirus-1 (ePPRV-1) in pineapple plants in Cuba. **New Disease Reports**, v. 28, p. 2, 2013.

RONQUIST, F. et al. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. **Systematic Biology Advance Access**, 2012.

SANZ, A. I.; FRAILE, A.; GARCÍA-ARENAL, F.; ZHOU, X.; ROBINSON, D. J.; KHALID; S.; BUTT, T.; HARRISON, B. D. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. **Journal of General Virology**, London, v. 81, p. 1839-1849, 2000.

SEAL, S. et al. The prevalence of badnaviruses in West African yams (*Dioscorea cayenensis-rotundata*) and evidence of endogenous pararetrovirus sequences in their genomes. **Virus Research**, v. 186, p. 144-54, 2014.

SEAL, S.; MULLER, E. Molecular analysis of a full-length sequence of a new yam badnavirus from *Dioscorea sansibarensis*. **Archives of Virology**, v. 152, p. 819-825, 2007.

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

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.

SILVA, J.M. et al. Incidence and species diversity of badnaviruses infecting sugarcane from a germplasm collection in Brazil. **Tropical Plant Pathology**, 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.

SVAROVSKAIA, E.S. et al. Retroviral mutation rates and reverse transcriptase fidelity. **Frontiers in Bioscience**, v.8, p.117–134, 2003.

TAMURA, K. et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Molecular Biology and Evolution**, v. 30, p. 2725-2729, 2013.

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

TORRES-PACHECO, I.; GARZÓN-TIZNADO, J. A.; BROWN, J. K., BECERRA-FLORA, A.; RIVERA-BUSTAMANTE, R. F. Detection and distribution of geminiviruses in Mexico and the Southern United States. **Phytopathology**, Saint Paul, v. 86, p. 1186-1192, 1996.

TZAFRIR, I. et al. The N-terminal portion of the 216-kDa polyprotein of Commelina yellow mottle badnavirus is required for virus movement but not for replication. **Virology**, v. 232, p. 359-68, 1997.

VEGA, J.; SORDI, R. Virus baciliforme em cana-de-açúcar detectado no Brasil através de Microscopia eletrônica de imunoabsorção. **Fitopatol. Bras**, v. 16: XXVI. 1991.

VERZIGNASSI, J. R. et al. Mancha negra do abacaxi no Pará. **Summa Phytopathologica**, Botucatu, 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(1), p. 5–8, 1996

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.