



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



**KARLA MARIA CANSANÇÃO GUIMARÃES**

**ESTRUTURA GENÉTICA DE UMA POPULAÇÃO DO BADNAVIRUS *DIOSCOREA*  
*BACILLIFORM ALATA VIRUS* (DBALV) QUE INFECTA INHAME (*Dioscorea spp.*) NO  
NORDESTE DO BRASIL**

**Rio Largo**

**2013**

**KARLA MARIA CANSANÇÃO GUIMARÃES**

**ESTRUTURA GENÉTICA DE UMA POPULAÇÃO DO BADNAVIRUS *DIOSCOREA*  
*BACILLIFORM ALATA VIRUS (DBALV)* QUE INFECTA INHAME (*Dioscorea*spp) NO  
NORDESTE DO BRASIL**

Dissertação de mestrado apresentada ao Programa de Pós Graduação em Proteção de Plantas do Centro de Ciências Agrárias da UFAL como requisito parcial para obtenção do grau de Mestre em Proteção de Plantas.

Orientador: Prof. Dr. Gaus Silvestre de Andrade Lima

**Rio Largo**

**2013**

**Catálogo na fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**  
**Bibliotecária Responsável: Fabiana Camargo dos Santos**

- G963e      Guimarães, Karla Maria Cansanção.  
             Estrutura genética de uma população do badnavirus *Dioscorea bacilliform alata virus* (BDALV) que infectam inhame (*Dioscorea spp.*) no nordeste do Brasil / Karla Maria Cansanção Guimarães. – 2013.  
             46 f. : il.
- Orientador: Gaus Silvestre de Andrade Lima.  
             Dissertação (Mestrado em Proteção de plantas) – Universidade Federal de Alagoas. Centro de Ciências Agrárias. Rio Largo, 2013.
- Bibliografia: f. 37- 46.
1. Inhame – Viroses. 2. *Dioscorea bacilliform alata virus*. 3. Vírus – Estrutura genética. I. Título.

CDU: 633.496:632.38

*A minha mãe Ana Maria, a minha  
filha Amanda Cansação, pelo  
carinho, amor, paciência e  
compreensão em todos os momentos  
da vida.*

***Dedico***

## AGRADECIMENTOS

*A Deus por me fazer forte e paciente em cada momento de dificuldade, pois a fé foi meu guia a cada passo;*

*A minha mãe Ana Maria, minha filha Amanda Cansação, pelo amor, apoio incondicional, compreensão e dedicação em todos os momentos da minha vida.*

*À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa, essencial para a completa realização do curso e deste trabalho;*

*À Universidade Federal de Alagoas (UFAL), através do Ao Centro de Ciências Agrárias (CECA), da Universidade Federal de Alagoas (UFAL), através do Programa de Pós Graduação em Proteção de Plantas pelo apoio institucional;*

*Ao meu orientador Prof. Dr. Gaus Andrade, pela oportunidade, confiança, amizade, e por todos os ensinamentos durante estes dois anos de estudo;*

*Á Dra. Sarah Cavalcanti pela co-orientação, confiança, amizade, conhecimentos transmitidos durante o desenvolvimento do trabalho e principalmente pela assistência e disposição em me ensinar e ajudar sempre que necessitei;*

*Á Prof.a Dra. Iraildes Assunção pela amizade, dedicação, ajuda durante as atividades laboratoriais e também pelas valiosas sugestões prestadas;*

*A Roberto Ramos Sobrinho pela assistência nas análises dos dados, e pela amizade e apoio (serei sempre grata);*

*A todos os Professores que compõem o curso de mestrado em Proteção de Plantas da Universidade Federal de Alagoas, pelos ensinamentos e contribuição na minha formação profissional;*

*Às amigas Deyse Rocha, Debora Tereza, Quitéria Cardoso, Joseane Castro, Sarah Cavalcanti pela amizade, companheirismo e convivência divertida e agradável;*

*Aos amigos do Laboratório de Fitopatologia Molecular CECA-UFAL, Jaqueline Figueiredo, Edilene Moraes, Mayra Ferro e em especial Joyce Lima, pela convivência, apoio, amizade e companhia e ajuda durante a realização deste trabalho;*

*Às amigas Deyse Rocha, Debora Tereza, Quitéria Cardoso, Joseane Castro, Sarah Cavalcanti pela amizade, companheirismo e convivência divertida e agradável;*

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

## RESUMO

O inhame (*Dioscorea* spp.) é uma cultura de grande importância socioeconômica para região Nordeste do Brasil devido ao valor nutritivo, energético e comercial de suas túberas, destacando-se pelo seu potencial para exportação, sobretudo para a Europa. Vírus são consideradas fatores limitantes do inhame por representar ameaça para produção, produtividade e movimento de germoplasma dessa cultura. O conhecimento da dinâmica da variabilidade genética de populações virais em plantas é necessário para entender como estas populações evoluem, bem como suas implicações para durabilidade de estratégias de manejo da doença baseadas na resistência do hospedeiro. Estudos para entender a estrutura genética de vírus de planta com genoma de dsDNA, a exemplo de membros de *Badnavirus* são escassos. Desta forma, o objetivo desse trabalho foi determinar a estrutura genética de populações de *Dioscorea AL bacilliformvirus* (DBALV) que infectam a cultura do inhame nos estados de Alagoas, Paraíba e Pernambuco, através da sequência parcial da região RT/RNaseH do genoma viral. Um total de 150 amostras foliares com sintomas típicos de infecção viral foi coletado em cinco áreas de cultivos comerciais de inhame no período de 2012/2013. As amostras foram indexadas para presença de *Badnavirus* por PCR com os oligonucleotídeos degenerados RP e FP. Os produtos de PCR foram purificados e diretamente sequenciados pela MacroGen Inc. (Seul, Coreia do Sul). As sequências parciais da região RT/RNaseH foram inicialmente submetidas ao BLASTn, e comparações pareadas utilizando a ferramenta SDT foram realizadas com outros *Caulimovirus* disponíveis no GenBank, para determinar o estado taxonômico dos isolados virais com base nos padrões estabelecidos pelo ICTV. A análise filogenética foi realizada utilizando-se o programa MEGA 5.0 através do método de Máxima Verossimilhança (ML). As análises de variabilidade genética e inferências sobre a estrutura de população foram realizadas utilizando-se o programa DnaSP versão 5.10. Os resultados das comparações pareadas de sequência revelaram que todos os isolados virais obtidos nesse trabalho pertencem à espécie DBALV. Uma população de 48 isolados virais foi utilizada para determinar a estrutura genética de DBALV. A análise da estrutura genética da população indicou que a variabilidade de DBALV é alta semelhante àquelas encontradas para vírus de ssDNA e que a mutação é um importante processo evolutivo que contribui para diversificação dessa espécie. O fato de DBALV estar sob possível influência de seleção purificadora sugere que a ocorrência de mutações não é suficiente para explicar completamente a sua variabilidade genética. Desta forma, estudos são necessários para demonstrar a influência de forças evolutivas adicionais atuando sobre essa população.

Palavras-chave: *Dioscorea* spp., *Badnavirus*, estrutura genética, *Dioscorea AL bacilliformvirus*, Nordeste do Brasil.

## ABSTRACT

The yam *Dioscorea spp.* is a crop of great economic and social importance to the Northeast Brazil due its nutritional value and trade of tubers, highlighting its potential for export, mainly to Europe. Viral diseases are considered limiting factors to the yam production and germplasm movement. The knowledge of the dynamics of genetic variability in plant viral populations is an important step to understand how these populations evolve, as well as its implications for durability of control strategies based on host resistance. Studies for understand the genetic structure of dsDNA plant viruses (*e.g.*, the members of *Badnavirus*) are scarce. Therefore, the aim of this work was determine the genetic structure of *Dioscorea AL bacilliform virus* (DBALV) populations that infect yam crop in Alagoas, Paraíba and Pernambuco States, by partial sequences of the RT/RNaseH region of the viral genome. A total of 150 leaf samples showing typical symptoms of viral infection were collected in five yam-growing areas in 2012 and 2013. The presence of badnavirus DNA was confirmed by PCR amplification using the degenerate oligonucleotides Badna RP and Badna FP. The PCR products were purified and directly sequenced by MacrogenInc (Seoul, South Korea). Partial sequences of the RT/RNaseH region were initially analyzed with BLASTn algorithm and the GenBank non-redundant nucleotide database to determine the viral species with which they shared greatest identity. These similar sequences from GenBank were used to determine the taxonomic state of the viral isolates based on criterion established by ICTV. The phylogenetic analysis was performed using the MEGA v.5.0. Genetic variability analysis and inference about the population structure were performed using the DnaSP program v.5.10. Pairwise sequence comparisons and phylogenetic tree showed that all viral isolates obtained in this work belong to the same badnavirus species: DBALV. A DBALV population comprising 48 isolates showed high genetic variability, similar to those found in ssDNA viruses. Mutation and purifying selection were important evolutionary mechanisms acting on evolution of this badnavirus population. However, additional studies are necessary to elucidate the influence of other evolutionary process shaping the genetic variability in DBALV infecting yam in Brazil.

**Keywords:** *Dioscorea* spp. *Badnavirus*, population structure, *Dioscorea AL bacilliform virus*, Northeast Brazil.

## LISTA DE TABELAS

**Tabela 1 - Localização, ano de coleta e número de amostras coletadas em três estados dos Nordeste no ano de 2012/2013.....23**

**Tabela 2 - Vírus do gênero *Badnavirus* e *Tungrovirus* utilizadas para as análises de Comparações pareadas e árvore de Máxima Verossimilhança..... 26**

**Tabela 3 - Estrutura genética da população de *DioscoreabacilliformAL virus* (DBALV) obtida de amostras de *Dioscorea* spp. coletadas em três estados da região nordeste do Brasil..... 32**

**Tabela 4 - Resultados dos diferentes testes de neutralidade realizados para o domínio RT da ORF III da população de *Dioscoreabacilliform AL virus* (DBALV) obtida de *Dioscorea* spp. no Nordeste do Brasil..... 34**

**Tabela 5 - Detecção de seleção utilizando os métodos ML disponíveis no DataMonkey..... 35**

## LISTA DE ILUSTRAÇÕES

**Figura 1 - A. Representação esquemática da organização genômica de *Commelinayellowmottle vírus* (CoYMV) membro tipo do gênero *Badnavirus*. O círculo completo representa o genoma de DNA de fita dupla. As setas largas indicam a posição das ORFs 1, 2 e 3. 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), Ribonuclease H (RH) (BRIDDON et al., 1999) ..... 16**

**Figura 2 - Mapa indicando os locais de coleta em áreas de cultivos de inhame (*Dioscoreaspp.*) nos Estados de Paraíba (em laranja), Pernambuco (em verde) e Alagoas (em azul).....23**

**Figura 3 -Plantas de inhame apresentando sintomas de infecção viral; A: cordão de sapato, B: Bolhosidades, C: Nanismo e D: Mosaico..... 24**

**Figura 4 -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), Ribonuclease H (RH) (BRIDDON et al., 1999)..... 25**

**Figura 5 -Amplificação do domínio RT/ RNaseH de badnavírus. Números 1-12 e 1-7 amostras provenientes de Bonito-PE; M, 1KB plusladder; C+, controle positivo e C-, controle negativo..... 28**

**Figura 6 - Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de sequências nucleotídicas da região RT/RNaseH dos *Badnavirus* desse trabalho e outras espécies do gênero disponíveis no GenBank..... 29**

**Figura 7 - Árvore de Máxima Verossimilhança baseada na sequência da região RT/RNaseH de isolados de desse trabalho e outras espécies de *Badnavirus* disponíveis no GenBank. O Rice tungrobacilliform vírus (RTBV) gênero *Tungrovirus* foi utilizado como “outgroup”..... 31**

## SUMÁRIO

|                                                                                                                 |           |
|-----------------------------------------------------------------------------------------------------------------|-----------|
| <b>1 INTRODUÇÃO.....</b>                                                                                        | <b>11</b> |
| <b>1.1 Aspectos gerais da cultura do inhame.....</b>                                                            | <b>11</b> |
| <b>1.2 Família <i>Caulimoviridae</i>.....</b>                                                                   | <b>13</b> |
| <b>1.3 Gênero <i>Badnavirus</i>.....</b>                                                                        | <b>15</b> |
| <b>1.4 Variabilidade e estrutura genética de populações de vírus com genoma de DNA.....</b>                     | <b>18</b> |
| 1.4.1 Mutação.....                                                                                              | 19        |
| 1.4.2 Recombinação.....                                                                                         | 21        |
| <b>2 MATERIAIS E MÉTODOS.....</b>                                                                               | <b>23</b> |
| <b>2.1 Coleta e Armazenamento das amostras.....</b>                                                             | <b>23</b> |
| <b>2.2 Amplificação da região RT e RNASE H de Badnavirus e sequenciamento.</b>                                  | <b>24</b> |
| <b>2.3 Comparação das sequencias e análise filogenética.....</b>                                                | <b>25</b> |
| <b>2.4 Descritores gerais da variabilidade da população e parametrização dos mecanismos evolutivos.....</b>     | <b>27</b> |
| <b>3 RESULTADOS E DISCUSSÃO.....</b>                                                                            | <b>28</b> |
| <b>3.1 Amplificação do domínio RT/ RNaseH de badnavírus e sequenciamento..</b>                                  | <b>28</b> |
| <b>3.2 Comparação das sequencias e análise filogenética.....</b>                                                | <b>28</b> |
| <b>3.3 Determinação da variabilidade genética da população e parametrização dos mecanismos evolutivos .....</b> | <b>32</b> |
| <b>REFERÊNCIAS.....</b>                                                                                         | <b>37</b> |

## 1. INTRODUÇÃO

### 1.1 Aspectos gerais da cultura do inhame

O inhame (*Dioscorea* spp.) é uma monotiledônea pertencente à família Dioscoreaceae, de hábito anual ou perene cujo órgão de reserva é conhecido como túbera, sendo considerada a segunda mais importante cultura de tubérculo (NWEKE et al., 1991).

O gênero *Dioscorea* possui cerca de 600 espécies originárias da África, Ásia ou América do Sul (COURSEY, 1976; SANTOS, 1996) das quais somente dez são consideradas como comestíveis (O'HAIR, 1990). As espécies cultivadas incluem *Dioscorea alata* L., *D. cayennensis* Lam., *D. rotundata* Poir., *D. esculenta* (Lour) Burk, *D. bulbifera* L., *D. numulária* Lam., *D. pentaphylla* L., *D. hispida* Dennst., *D. trifida* L. e *D. dumetorum* (Kunth) Pax. Dentre as espécies cultivadas de inhame amarelo (*D. cayennensis* Lam.) e o inhame branco (*D. rotundata* Poir) são as mais importantes em termos de volume produzido e comercialização (ASIEDU & SARTIE, 2010). No Brasil têm destaque duas cultivares: Da Costa (*D. cayennensis* Lam), única destinada à exportação; e São Tomé (*D. alata* L.), menos cultivada nessa região (MOURA, 2005).

O inhame é cultivado em 4,59 milhões de hectares, com uma produção mundial de 51,4 milhões de toneladas (FAO, 2010). Na África Ocidental estão concentrados 93% da produção mundial, principalmente, em Benin, Costa do Marfim, Gana, Nigéria e Togo (ASIEDU; SARTIE, 2010). No cenário mundial, a Nigéria está em primeiro lugar, com a produção em 2012 de 37 milhões de toneladas. No cenário Sul americano, a Colômbia é a maior produtora, com 395 mil toneladas e 32 mil hectares. O Brasil é o segundo maior produtor da América do Sul, contribuindo com uma produção e área plantada em 2010 de aproximadamente 244 mil toneladas e 25 mil hectares, respectivamente (FAOSTAT, 2010). De acordo com o censo agropecuário, a região Nordeste foi responsável pela produção de 38.256 toneladas de inhame, tendo os estados de Alagoas, Paraíba e Pernambuco contribuído respectivamente com 2.872, 8.441 e 16.574 toneladas (IBGE, 2006).

A importância alimentar do inhame decorre do seu alto valor nutritivo e energético (SANTOS E MACÊDO, 2006). Túberas de inhame são uma importante fonte de carboidratos para milhões de pessoas nas regiões tropicais e subtropicais, principalmente na África, Caribe, Ásia, Américas Central e Sul e no Pacífico. Além disso, algumas espécies de inhame têm valor medicinal e ornamental (FAO, 2010; HOU et al., 2002).

O inhame desenvolve-se satisfatoriamente nos ecossistemas brasileiros, sobretudo na Região Nordeste, constituindo uma opção agrícola de grande potencial para ampliar o consumo no mercado interno e atender a demanda do mercado externo, movimentando a economia das zonas produtoras (SANTOS, 2002). Embora seja considerada, na maioria dos casos, uma cultura de subsistência, o inhame tem grande importância sócio-econômica no cenário da agricultura familiar no Nordeste do Brasil, com um significativo potencial de desenvolvimento, contribuindo para alimentação humana, beneficiando populações carentes, além de ser fonte de renda para pequenos e médios produtores (MENDES, 2005; SANTOS; MACÊDO, 2006).

Apesar da importância sócio-econômica que essa cultura representa para a região Nordeste do Brasil, sua produtividade ainda continua baixa, decorrente do manejo inadequado da cultura, do uso de túberas-semente de inferior qualidade, em função da indisponibilidade de material de melhor qualidade (O'HAIR, 1990; SANTOS, 2002; ANDRADE, 2007).

O plantio comercial do inhame é realizado utilizando-se túberas-sementes ou segmentos de túberas comerciais (SANTOS, 1996). O uso de sementes verdadeiras como propágulos é restrito a programas de melhoramento da cultura (ASIEDU; SARTIE, 2010). Esse método de propagação vegetativa, associado ao baixo nível tecnológico empregado, favorece a sobrevivência, acúmulo e disseminação de patógenos e pragas que podem ser introduzidas em áreas isentas via material de propagação (MENEZES, 2002; OLIVEIRA et al., 2006).

Doenças e pragas de inhame foram relatadas na maioria das áreas de plantio e constituem um dos principais entraves para a produção, armazenamento e exportação (MANTELL, 1993). Problemas fitossanitários, especialmente o ataque de nematóides, vírus e fungos, dificultam o cultivo do inhame (MOURA, 2005). Doenças causadas por vírus são a maior restrição na produção de inhame em vários países (OFFIE, 2001; ANDRADE, 2007).

Este trabalho teve como objetivo determinar estrutura genética de populações de DBALV que infectam cultura do inhame (*Dioscorea* spp.) no estado de Alagoas, Paraíba e Pernambuco através da sequência parcial da região RT/RNaseH da ORF3 do genoma viral, com a finalidade de fornecer informações que possam contribuir para o melhor entendimento da biologia dessa espécie e auxiliar os programas de melhoramento genético da cultura.

A propagação vegetativa, através de túberas-semente, facilita a sobrevivência, disseminação e o acúmulo de agentes virais em cultivos sucessivos (AMUSA et al., 2003). Doenças virais são de particular importância porque reduzem o vigor e subsequentemente causam a redução da produção e qualidade das túberas (AMUSA et al., 2003). Diversas

viroses têm sido descritas na cultura do inhame, por membros dos gêneros *Potyvirus*, *Badnavirus*, *Cucumovirus* e *Potexvirus*. Destacam-se pela importância e distribuição geográfica os vírus dos gêneros *Potyvirus* e *Badnavirus* (PHILLIPS et al., 1999; YANG et al., 2003; ODU et al., 2006, SEAL; MULLER, 2007). Duas espécies de potyvírus são bem caracterizadas, o *Yam mosaic virus* (YMV) e o *Yam mild mosaic virus* (YMMV), sendo o primeiro considerado o vírus mais importante e de distribuição geográfica mais ampla (MUNFORD; SEAL, 1997). *Dioscorea baciliform AL virus* (DBALV) é o badnavírus mais amplamente distribuído, sendo encontrado em infecções simples e/ou mistas com potyvírus (YANG et al., 2003; PIO-RIBEIRO et. al., 2005).

O conhecimento das viroses que ocorrem nas diferentes regiões produtoras é limitado (OFFIE, 2001) e assim medidas eficientes de manejo não podem ser adotadas. Considerando a importância da cultura do inhame para a região Nordeste e a escassez de estudos sobre as viroses que incidem sobre essa cultura, torna-se necessário a realização de pesquisas abrangentes e sistematizadas que abordem esse tema. A geração de informações sobre incidência, variabilidade genética e distribuição de vírus de inhame na região Nordeste, é de extrema importância para o desenvolvimento de testes diagnósticos confiáveis, adoção de medidas de controle mais duráveis, além de auxiliar programas de melhoramento no desenvolvimento de variedades resistentes e na troca de material nacional e internacional de germoplasma.

## 1.2 Família *Caulimoviridae*

A família *Caulimoviridae* inclui vírus de plantas com genoma de DNA de fita dupla (dsDNA) classificados como pararetrovírus (vírus de DNA que utilizam a transcriptase reversa no ciclo de replicação) e diferem dos retrovírus com base no seu genoma de DNA e na sua integração irregular dentro do genoma hospedeiro para a replicação. Os vírus dessa família apresentam partículas não envelopadas as quais com morfologia isométrica (50-52 nm de diâmetro) ou baciliforme (30 nm de diâmetro e 130 -150 nm de comprimento) (GEERING; HULL, 2012). A família é dividida em sete gêneros de acordo com o inseto vetor, gama de hospedeiros, organização do genoma e relacionamento filogenético: *Badnavirus*, *Caulimovirus*, *Cavemovirus*, *Petuvirus*, *Solendovirus*, *Soymovirus* *Tungrovirus*, são reconhecidos pelo International Committee on Taxonomy of Viruses (GEERING; HULL, 2012).

Os membros de *Caulimoviridae* possuem ampla distribuição geográfica, sendo que a maioria das espécies dos gêneros *Tungrovirus* e *Badnavirus* se localiza nas regiões tropicais e subtropicais, enquanto que *Petuvirus*, *Caulimovirus*, *Soymovirus* e *Cavemovirus* são encontrados em regiões temperadas (GEERING; HULL, 2012). A transmissão pode ocorrer de forma natural, via inseto vetor (Hemiptera das famílias Aleyrodidae, Aphididae, Cicadellidae e Pseudococcidae), por contato entre plantas hospedeiras, pelo pólen, pela semente e por propagação vegetativa. A transmissão pode ser também realizada de forma experimental por meio de técnicas como a inoculação mecânica e enxertia. Os sintomas causados por esses vírus são variáveis e dependem da espécie viral, hospedeira e condições climáticas. Mosaico, clareamento de nervuras, clorose entre nervuras e estrias são os sintomas mais frequentemente ocasionado pelos diferentes gêneros de *Caulimoviridae* (GEERING; HULL, 2012).

Os vírions contêm uma única molécula de DNA fita duplacircular medindo 7,2 a 9,2 Kb (FAUQUET et al., 2005; GEERING; HULL, 2012). O genoma apresenta descontinuidades ou gaps em sítios específicos de ambas as fitas (HARPER et al., 2002). O genoma dos caulimovírus contém de uma a oito ORFs e a organização genômica é uma das principais características utilizadas para distinguir os diferentes gêneros dessa família.

A replicação dos caulimovírus, em geral, envolve duas fases: transcrição de RNA a partir do DNA viral no núcleo e transcrição reversa deste RNA para gerar dsDNA no citoplasma. O ciclo de infecção em *Caulimoviridae* inicia com a entrada do vírus na célula hospedeira. Na célula, o vírus é direcionado para o núcleo, ancora no poro nuclear e é desencapsidado, expondo seu genoma. Em seguida, o dsDNA é importado para o interior do núcleo, onde as descontinuidades do genoma são seladas dando origem a um DNA superenovelado. Este se associa com proteínas histonas para formar minicromossomos, os quais são transcritos por uma RNA polimerase dependente de DNA do hospedeiro. A transcrição dos minicromossomos dá origem a um transcrito maior que o genoma, o mRNA35S ou mRNA34S, que possui uma redundância terminal de 35 a 270 nucleotídeos, dependendo da espécie. O mRNA35S serve como molde para expressão das ORFs virais e síntese da primeira fita do DNA (GEERING; HULL, 2012). 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 geralmente designada como nucleotídeo 1 (DE KOCHKO et al., 1998). No início da replicação, no citosol, o tRNA<sup>met</sup> do hospedeiro é

utilizado como primer para síntese da primeira fita, sendo ambas fitas replicadas pela Transcriptase Reversa (RT) viral.

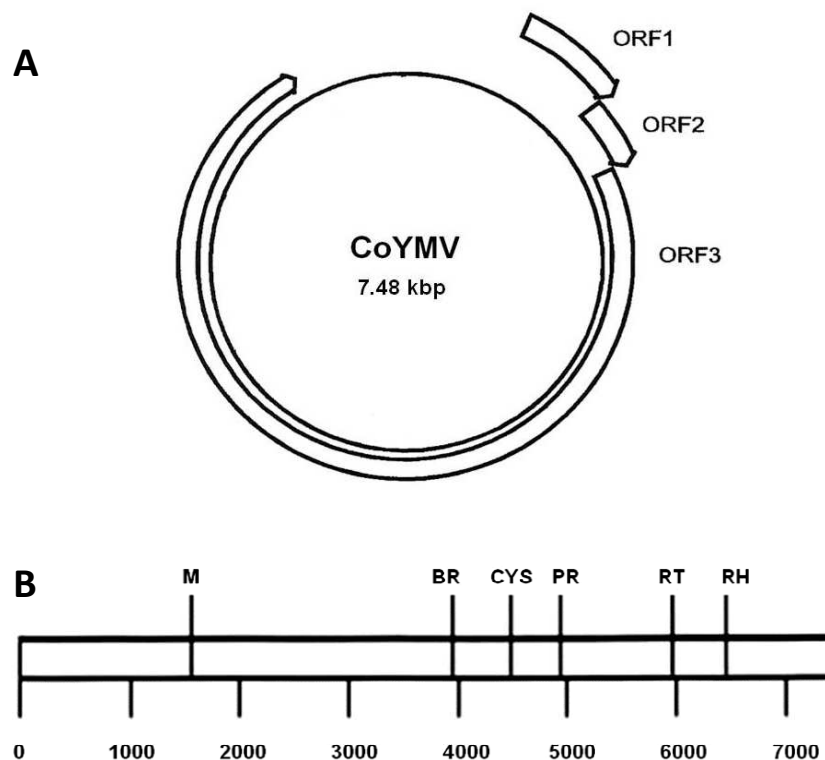
Em contraste com os retrovírus, o ciclo de replicação de *Caulimoviridae* é episomal e não envolve uma fase de integração no genoma da hospedeira (GEERING; HULL, 2012). Sequências virais podem ocorrer dispersas no genoma de plantas, originárias de uma infecção viral anterior que se fixou na linhagem germinativa. Tais sequências são conhecidas como sequências endógenas de pararetrovírus (Endogenous Pararetroviral Sequences - EPRVs) e são a classe mais abundante de sequências virais endógenas em diferentes espécies de plantas (GAYRAL; ISKRA-CARUANA, 2009). Um número cada vez maior de EPRVs em membros da família *Caulimoviridae* tem sido identificadas nos genomas de muitas espécies de plantas (GEERING; SCHARASCHKIN; TEYCHENEY, 2010). Todas as EPRVs descritas até o momento possuem um padrão de arranjo similar com repetições em *tandem*, duplicações internas, fragmentações e inversão do genoma viral. A maioria das EPRVs resulta em genomas virais parciais e não funcionais, porém várias integrações contêm toda extensão do genoma viral, com sequências de leitura aberta (ORFs) funcionais. Tais sequências podem então ser ativadas resultando na liberação do genoma viral funcional que infectará o hospedeiro (GAYRAL; ISKRA-CARUANA, 2009). Essas sequências são integradas no genoma hospedeiro, e podem dar origem a vírus episomais em determinadas condições. Os exemplos mais estudados foram *Banana streak OL virus* (BSOLV) (NDOWORA et al., 1999), *Banana streak GF virus* (BSGFV) (GAYRAL; ISKRA-CARUANA, 2009), *Petunia vein clearing virus* (PVCV) (Richert-Poeggeler, 2003) and *Tobacco vein clearing virus* (TVCV).

### 1.3 Gênero Badnavirus

*Badnavirus*, o mais numeroso gênero da família *Caulimoviridae*, são pararetrovírus com partícula viral baciliformes não envelopadas, medindo 30 nm largura x 130 –150 nm (GEERING; HULL, 2012). Seu genoma contém uma única molécula de DNA de fita dupla circular (dsDNA) de cerca de 7200-7600 pb que forma um círculo aberto interrompido por descontinuidades sítio-específicas e que pode conter uma região intergênica poli A (MEDBERRY et al., 1990; HULL et al., 2005). O membro tipo do gênero *Badnavirus* é o *Commelina yellow mottle virus* (CoYMV) (MIGLIORI, 1978). Todos os badnavirus codificam três principais ORFs (Figura 1A), e contém geralmente três descontinuidades (gaps) em sítios específicos (BOUHIDA et al., 1993; HARPE; HULL, 1998). As funções dos

produtos das ORFs I e II permanecem desconhecidas. Contudo, a proteína da ORF I de CoYMV está associada com vírions imaturos enquanto a ORF II é encontrada associada com ambos, imaturos e maduros (CHENG et al., 1996). A poliproteína codificada pela ORF3 contém domínios funcionais: rico em cisteína (CYS), motivo de ligação ao RNA (RB), aspartato protease (PR), e replicase viral (transcriptase reversa, RT e ribonuclease, RNase H-RH) (Figura 1B), os quais possuem domínios idênticos para os membros da família *Caulimoviridae* (MEDBERRY et al., 1990; LACO; BEACHY, 1994; HARPER; HULL, 1998).

**Figura 1A - Representação esquemática da organização genômica de *Commelina 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 1, 2 e 3. 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), Ribonuclease H (RH).**



Fonte: BRIDDON et al., 1999.

Esses vírus são relatados infectando uma ampla gama de culturas tropicais economicamente importantes como arroz (*Oryza sativa* L.) (OMURA et al., 1983), cana-de-açúcar (*Saccharum officinarum* L.) (LOCKHART et al., 1988), banana (*Musa* spp.) (LOCKHART; OLSZEWSKI, 1993), cacau (*Theobroma cacao* L.) (BRUNT, 1996), citros (*Citrus* spp.) (AHLAWAT et al., 1996), pimenta (*Capsicum* spp.) (LOCKHART et al., 1997), inhame (*Dioscorea* spp.) (PHILLIPS et al., 1999) e taro (*Colocasia esculenta* (L.) Schott.) (YANG et al., 2003). A transmissão dos membros do gênero *Badnavirus* ocorre principalmente por cochonilhas *Planococcus citri* (Hemiptera: Pseudococcidae). No entanto algumas espécies podem ser transmitidas por afídeos ou “lace bugs”. Todos os estágios móveis dos vetores são capazes de adquirir e transmitir o vírus. Transmissão por semente numa taxa de 30-63% tem sido relatada para *Kalanchoe top-spotting virus* (KTSV). O *Cacao swollen shoot virus*, CSSV têm sido transmitidos mecanicamente, no entanto este tipo de transmissão não foi observado para as outras espécies do gênero. Algumas espécies que infectam plantas lenhosas podem ser transmitidas por enxertia (GEERING; HULL, 2012).

Os *Badnavirus* foram relatados pela primeira vez em inhame em Barbados e em vários outros países do Caribe na década de 1970. Partículas baciliformes estavam presentes em infecção mista associada à doença da mancha marrom nos tubérculos de *D. alata* e *D. cayenensis* (MOHAMED; MANTELL, 1976). No Caribe, um potyvírus foi observado em combinação com o *Dioscorea bacilliforme virus* (DBV, gênero *Badnavirus*) em *D. alata* e *D. cayenensis* (MOHAMED; MANTELL, 1976). Posteriormente, partículas virais baciliformes foram confirmadas em amostras de *D. alata* no Caribe e na África Ocidental, bem como em *D. bulbifera* (Vanuatu), *D. dumetorum* (Papua Nova Guiné), *D. numularia* (Papua-Nova Guiné e Vanuatu), *D. cayenensis-rotundata* (África Ocidental, Brasil), *D. opposita* (Japão), *D. pentaphylla* (Vanuatu) e *D. trifida* (Barbados, Guiana) (HUGHES, 1997; PHILLIPS et al., 1999). O primeiro badnavírus caracterizado em inhame foi um isolado de *D. alata* obtido na Nigéria, designado como *Dioscorea alata bacilliform virus* (DaBV) (PHILLIPS et al., 1999). Posteriormente, SEAL; MULLER (2007) em estudos sobre a diversidade molecular de *Badnavirus* que infectam inhame nas ilhas sul do Pacífico, identificou a nova espécie *D. sansibarensis bacilliform virus*. O genoma de DSBV partilha 61.9 % de identidade de nucleotídeos com o DaBV (SEAL; MULLER, 2007; ENI, et al., 2008; KENYON et al., 2008).

Atualmente apenas duas espécies de badnavírus que infectam *Dioscorea* spp. são reconhecidas pelo ICTV, o *Dioscorea bacilliform AL virus* (DBALV) e *Dioscorea bacilliform SN virus* (DBSNV), provenientes de *D. alata* e *D. sansibarensis*, respectivamente

(GEERING;HULL, 2012). No entanto, estudos realizados na África e Sul do Pacífico sugerem alta diversidade de *Badnavirus* em inhame e a possível presença de até doze espécies (ENI, 2008).

A grande diversidade de sequências de *Badnavirus* dificulta um diagnóstico preciso e ameaça a circulação segura de germoplasma (SEAL; MULLER, 2007; ENI, et al., 2008; KENYON et al.,2008). Esses autores destacaram a necessidade de mais estudos para resolver a dificuldade na taxonomia e a distribuição mundial desse gênero. As badnavirose tem recebido maior atenção nas pesquisas devido à descoberta de que podem apresentar sequências virais integradas no genoma de plantas (HARPER et al., 2002). A identificação de badnavírus é baseada na amplificação parcial do genoma por PCR (*“Polymerase Chain Reaction”* = Reação em cadeia da Polimerase). Esta técnica não é capaz de distinguir entre sequências integradas ao genoma e cópias episomais dos vírus na planta hospedeira. Como consequência, o número de sequências de badnavírus tem crescido rapidamente, resultando na falta de clareza sobre como determinar o estado taxonômico dos membros desse gênero (BOUSALEM et al., 2008).

Informações relacionadas à diversidade genética de populações de badnavírus infectando inhame são escassas. ANDRADE, et al.(2007), realizou o primeiro registro detalhado com informações moleculares destes vírus no Brasil, relatando sua ocorrência nos estados de Alagoas, Bahia, Paraíba e Pernambuco. LIMA et al. (2013), em estudo sobre a variabilidade de sequências correspondentes a região RT/RNaseH e a distribuição de badnavírus infectando inhame nos estados de Alagoas, Paraíba e Pernambuco, demonstraram a ampla distribuição e prevalência de DBALV na região Nordeste. Sugere-se que a predominância dessa espécie nesta região ocorra devido à proximidade geográfica e a troca constante de material propagativo infectado entre as áreas de cultivo.

#### **1.4 Variabilidade e estrutura genética de populações de vírus com genoma de DNA**

Populações de vírus apresentam um alto grau de diversidade genética a qual é gerada por erros durante a replicação dos genomas. Para vírus, os dois mecanismos principais de geração de variabilidade são a mutação e recombinação (GARCIA-ARENAL, 2003).

### 1.4.1 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 (GARCIA-ARENAL, 2003).

A taxa de mutação mede o aparecimento de mutações espontâneas, em função do tempo. Taxas de mutação observadas diferem entre as espécies e também variam de acordo com o genoma de uma dada espécie. Este parâmetro é frequentemente mensurado como o número de substituições de nucleotídeos por sítio por ano. Como os vírus de RNA são replicados por polimerases de RNA dependentes de RNA, que não possuem a capacidade de correção de erro, eles apresentam as mais altas taxas de mutação entre  $10^{-4}$  e  $10^{-6}$  (DRAKE; HOLLAND, 1999). A taxa de evolução molecular para vírus de RNA também é alta, entre  $10^{-4}$  e  $10^{-2}$  subs/sítio/ano (JENKINS et al., 2002). Desta forma, vírus de RNA são considerados as entidades que evoluem mais rapidamente (HOLLAND et al., 1982).

A taxa universal de mutação para microrganismos com genoma de DNA é entre  $10^{-10}$  a  $10^{-6}$  mutações por base por geração, enquanto essa taxa varia em torno de um padrão médio de 0,003 mutações por genoma por geração, um resultado conhecido como regra de Drake (DRAKE et al., 1991). Embora essa taxa reflita estimativas obtidas para alguns vírus de DNA, a sua universalidade é uma questão em aberto. Por exemplo, a taxa de mutação do ssDNA do fago  $\Phi$ X174 é mais alta do que a predita por esse modelo, enquanto os papilomavírus evoluem mais lentamente que o previsto (RANEY; DELONGCHAMP; VALENTINE, 2004). Estudos tem demonstrado que vírus de ssDNA como os parvovírus (SHACKELTON et al., 2006) e geminivírus (DUFFY; HOLMES, 2009), podem apresentar taxas evolutivas semelhantes as dos vírus de RNA.

As taxas de mutação correspondem à fidelidade das polimerases utilizadas na replicação: vírus de RNA (os quais utilizam RNA polimerase dependentes de RNA, RdRps) mutam mais rápido que os retrovírus (DNA polimerase dependente de RNA; RdDps ou transcriptase reversa, RTs), as quais mutam mais rápido que os vírus de DNA (DNA polimerase) (DUFFY, SHACKELTON; HOLMES, 2008). DNA polimerases podem possuir domínios de correção de erros, os quais reduzem a taxa de mutação durante a replicação do DNA por ao menos uma ordem de magnitude (GARCIA-DIAZ; BEBENEK, 2007). Ainda não se conhece polimerases de RNA com capacidade de correção de erro. No caso de RTs, a fidelidade é mais alta que a de RdRps (embora permaneça mais baixa que a de DNA polimerases), resultando em menores taxas de mutação em retrovírus que em vírus de RNA.

Por exemplo, quando não integrados no genoma do hospedeiro, taxas de 0.1-0.2 mutações por genoma por geração têm sido estimadas para Human immunodeficiency virus 1 (HIV-1) (DRAKE et al., 1998), uma taxa aproximadamente cinco vezes menor do que a observada em vírus que replicam usando RdRps (DRAKE, 1993).

Estudos de bactérias e sistemas animais indicaram que a taxa de mutação dos vírus de dsDNA e ssDNA diferem significativamente (DUFFY; SHAKELTON; HOLMES, 2008).

O impacto das mutações pontuais tem sido estudado para begomovírus, vírus ssDNA que infectam plantas. Sob diferentes condições de seleção, como presença de um efeito gargalo (população inicial pequena do vírus, período curto de aquisição pelo vetor), transferências sucessivas entre hospedeiros sem emprego do vetor, e inoculação em plantas resistentes, isolados de MSV apresentaram alta frequência de mutação, da ordem de  $10^{-4}$  e  $10^{-5}$  (ISNARD et al., 1998).

Resultados similares foram obtidos num experimento controlado de análise da taxa de variabilidade genética do begomovírus *Tomato yellow leaf curl China virus* (TYLCCNV) onde foi encontrada uma frequência média de mutação de  $3,5 \times 10^{-4}$  e  $5,3 \times 10^{-4}$  após 60 dias de infecção em *N. benthamiana* e tomateiro (*Solanum lycopersicon*), respectivamente (GE et al., 2007).

Uma série de experimentos de evolução em longo prazo (de 6 a 32 anos) também revelaram alta frequência de mutação, entre 2 e  $3 \times 10^{-4}$ , para MSV e *Sugarcane streak Réunion virus* (SSRV), sugerindo que mastrevírus provavelmente não co-divergem com seus hospedeiros (HARKINS et al., 2009). Estes resultados discordam com a hipótese de aparente co-divergência entre alguns mastrevírus e seus hospedeiros, o que implicaria em taxas de substituições de apenas  $10^{-8}$  subs/sítio/ano na natureza (WU et al., 2008).

Estudos sobre a evolução de populações de begomovírus que infectam tomateiro (ToCmMV, ToCMoV, ToRSV e ToYVSV) e a planta daninha *Blainvillea* sp. (BIYSV) revelaram que a frequência de mutações dessas espécies ( $10^{-2}$  e  $10^{-4}$ ) são maiores ou iguais àquelas observadas em vírus de RNA. Altas taxas de mutação ( $10^{-4}$ ) também foram relatadas para parvovírus e circovírus, vírus de ssDNA que infectam vertebrados (GALLIAN et al., 2002; BIAGINI et al., 2004). Semelhante aos geminivírus, os parvovírus e circovírus replicam seu genoma via mecanismo de círculo rolante, sugerindo que os altos níveis de heterogeneidade relatados podem refletir erros de replicação (ARGUELLO-ASTORGA et al., 2004). Foi sugerido que os mecanismos de correção de erro associados à replicação de DNA em eucariotos não sejam eficientes na replicação por círculo rolante e, ou na replicação de ssDNA (VAN DER WALT et al., 2008).

As informações sobre taxas de mutação para vírus de dsDNA que infectam plantas são escassas, sendo relatada apenas para o *Cauliflower mosaic virus* (CaMV), a qual estima-se que é em torno de 10<sup>-5</sup> subs/sítio/ano (FROISSART et al., 2005).

#### 1.4.2 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, 1999a). Eventos de recombinação têm sido demonstrados estar associados com o aumento da gama de hospedeira (GIBBS; WEILLER, 1999) virulência (ZHOU et al., 1997; PITA et al., 2001) potencial evolutivo e adaptação local (MONCI et al., 2002).

A recombinação é um evento bastante comum em vírus de plantas com genoma de ssDNA, a exemplo dos geminivírus (PADIDAM; BEACHY; FAUQUET, 1999a; LEFEUVRE et al., 2007b). A elevada frequência de recombinação nesse grupo de vírus pode ser em parte explicada pela existência de uma possível estratégia de replicação dependente de recombinação (RDR) (JESKE; LUTGEMEIER; PREISS, 2001; PREISS; JESKE, 2003) em adição à replicação por círculo rolante (RCR) (SAUNDERS; BEDFORD; STANLEY, 2001), e pela ocorrência frequente de infecções mistas (TORRES-PACHECO et al., 1996; SANZ et al., 2000; PITA et al., 2001; RIBEIRO et al., 2003; GARCÍA-ANDRÉS et al., 2006; DAVINO et al., 2009), com a evidência de infecção do mesmo núcleo da célula por mais de um begomovírus (MORILLA et al., 2004).

A análise comparativa da distribuição de sítios de recombinação dentro do genoma de diversas famílias de vírus de ssDNA sugeriu a distribuição não aleatória dos sítios e também uma tendência significativa para estes caírem tanto fora como na periferia dos genes. Além disso, foi particularmente observado que poucos sítios de recombinação foram encontrados dentro de genes que codificam proteínas estruturais, a exemplo da proteína capsidial (LEFEUVRE et al., 2007a). Esses resultados sugerem que a seleção natural agindo contra vírus que expressam proteínas recombinantes é o principal determinante na distribuição não aleatória dos sítios de recombinação na maioria das famílias de vírus de ssDNA.

Os vírus que pertencem à família *Caulimoviridae* são considerados objetos ideais para estudos de recombinação. Esta família de plantas tem recebido atenção adicional desde a descoberta que capazes de se integrar no genoma da planta hospedeira (SQUIRES et al., 2011).

Gal et al. (1991), criaram um sistema para monitorar recombinação homóloga em CaMV. Estes integraram um multímero de sequências do vírus no DNA nuclear de nabo (*Brassica napus*) arranjadas de maneira que a forma epissomal do CaMV só poderia ser obtida por eventos de recombinação. Os eventos de recombinação foram detectados pelo surgimento dos sintomas de infecções virais. Os vírus recombinantes apresentaram regiões homólogas derivadas de duas diferentes estirpes de CaMV. Na maioria das plantas transgênicas, uma única estirpe de vírus foi detectada. Cerca de metade das plantas transgênicas continham vírus do mesmo tipo, sugerindo um “hotspot” para recombinação. Eventos de recombinação homóloga também foram relatados para ativação do genoma epissomal de BSV e TVCV, integrados nos genomas de seus hospedeiros (NDOWORA et al., 1999; STAGINNUS; RICHERT-PÖGGELER, 2006).

Embora tenha sido sugerido que eventos de recombinação contribuem para o nível de variação genética entre os membros dos gêneros *Caulimovirus* (FROISSART et al., 2005) e *Tungrovirus* (ARBOLEDA; AZZAM et al., 2000), a recombinação entre badnavírus de inhame em condições naturais parece ser rara. A investigação de eventos de recombinação em espécies de badnavírus que infectam inhame identificou apenas um único evento ocorrendo na região que codifica para a transcriptase reversa entre as espécies tentativas DeBV A e DeBV B. No entanto como as análises foram realizadas para um pequeno fragmento do genoma, recombinações em outras partes do genoma não podem ser excluídas (BOUSALEM et al., 2009).

## 2. MATERIAIS E MÉTODOS

### 2.1 Coleta e Armazenamento das amostras

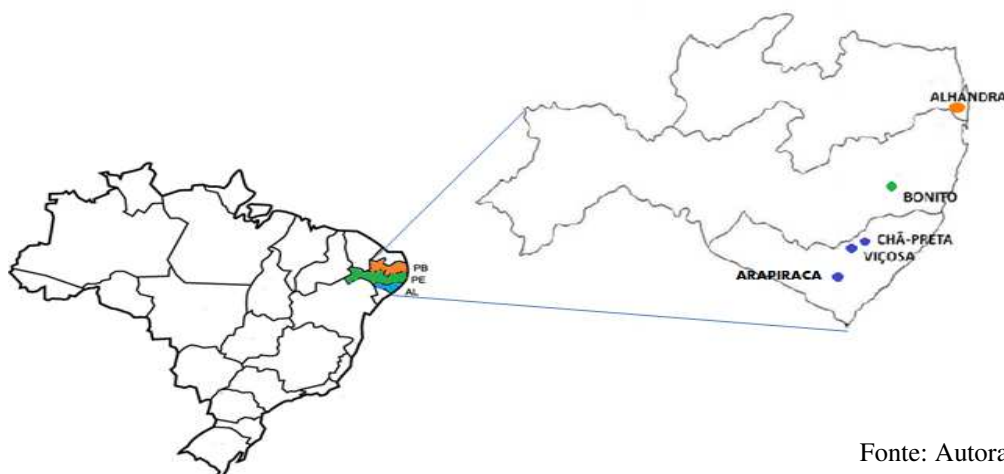
Um total de 150 amostras foliares de inhame foram coletadas entre julho/novembro de 2012 e julho 2013 em cinco áreas de cultivo nas principais regiões produtoras da Zona da Mata dos estados de Alagoas, Pernambuco e Paraíba (Figura 2, Tabela 1). Foram coletadas apenas plantas que apresentavam sintomas típicos de infecção viral, como mosaico, deformação foliar, cordão de sapato e nanismo (Figura 3). As amostras foram armazenadas em sacos plásticos e transportadas para o laboratório, onde foram mantidas a  $-80^{\circ}\text{C}$  até a extração de DNA.

**Tabela 1** -Localização, ano de coleta e número de amostras coletadas em três estados do Nordeste no ano de 2012/2013.

| Local de coleta<br>(Município) | Data da coleta | Número de<br>amostras coletadas | Código das<br>amostras |
|--------------------------------|----------------|---------------------------------|------------------------|
| Arapiraca-AL                   | Agosto 2012    | 30                              | 1A a 30A               |
| Alhandra-PB                    | Julho 2012     | 30                              | 1P a 30P               |
| Bonito-PE                      | Julho 2013     | 30                              | 1B a 30B               |
| Chã Preta-AL                   | Setembro 2012  | 30                              | 1CP a 30CP             |
| Viçosa-AL                      | Setembro 2012  | 30                              | 1V a 30V               |

Fonte: Autora, 2013.

**Figura 2** - Mapa indicando os locais de coleta em áreas de cultivos de inhame (*Dioscorea* spp.) nos Estados de Paraíba (em laranja), Pernambuco (em verde) e Alagoas (em azul).



Fonte: Autora, 2013.

**Figura 3 - Plantas de inhame apresentando sintomas de infecção viral; A: cordão de sapato, B: Bolhosidades, C: Nanismo e D: Mosaico.**



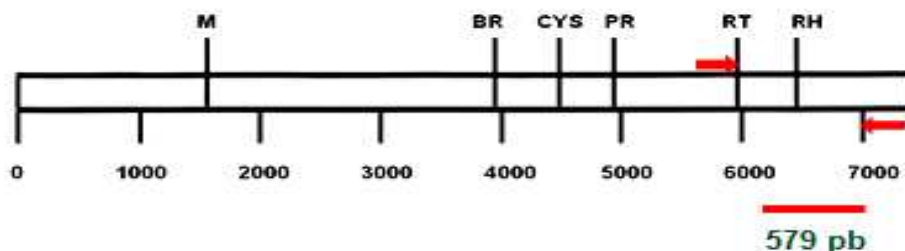
Fonte: Autora, 2013.

## **2.2 Amplificação da região RT e RNASE H de Badnavirus e sequenciamento**

Para a extração do DNA total foram utilizados 100-200 mg de tecido foliar proveniente de folhas jovens de plantas de inhame, de acordo com Doyle & Doyle (1987).

A presença de badnavírus foi verificada por meio de reações de amplificação por PCR de fragmentos de 579 pb utilizando o par de oligonucleotídeos degenerados BadnaFP (5'-ATGCCITTYGGIITIAARAAYGCICC – 3') e BadnaRP (5'-CCAYTTRCAIACISCICCCCAICC – 3') (Yang, 2003), baseado no domínio RT/RNaseH da ORF 3 do genoma viral de badnavírus (Figura 4).

**Figura 4 -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), Ribonuclease H (RH).**



Fonte: BRIDDON et al., 1999.

As reações foram realizadas em volume total de 60  $\mu$ L, contendo 6  $\mu$ L de tampão 10X (KCl 100 mM; Tris-HCl 100 mM pH 9,0; Triton-X 1%), 4,8  $\mu$ L de dNTPs 2,5 mM, 1,8  $\mu$ L de  $MgCl_2$  50 mM, 3  $\mu$ L de cada oligonucleotídeo 10  $\mu$ M, 1U de Taq DNA Polimerase, 1  $\mu$ L de DNA (10 a 100 ng) e 40  $\mu$ L de água ultrapura. As condições de amplificação foram: desnaturação inicial a 94°C por 4 min, 94°C por 30 s, 50°C por 30 s, 72°C por 30 s. Os ciclos foram repetidos 35 vezes exceto o passo inicial e a extensão final a 72°C por 10 min. Os produtos amplificados foram analisados em gel de agarose a 1,5% (p/v), após eletroforese em tampão TAE (Tris-Ácido Acético, EDTA 0,5M pH 8,0) a 80V por 1 h, corados com brometo de etídeo e visualizados sob luz ultravioleta.

Os produtos de amplificação, com fragmento de ~580 bp, foram purificados utilizando o kit GFX<sup>TM</sup> PCR DNA and Gel Band Purification (GE Healthcare) conforme instruções do fabricante. O sequenciamento foi realizado pela Macrogen Inc. (Seul, Coréia do Sul) a partir dos produtos de PCR purificados.

### 2.3 Comparação das sequências e análise filogenética

As sequências da região RT/RNaseH foram inicialmente submetidas ao algoritmo BLASTn (Altschul et al., 1990) e comparadas com banco de dados não-redundante de nucleotídeos no GenBank (<http://www.ncbi.nlm.nih.gov>) para determinar a porcentagem de identidade das sequências de nucleotídeos entre os isolados. Análises de comparações pareadas com a ferramenta Species Demarcation Tool v. 1.0 (Muhire et al., 2013) entre sequências obtidas nesse trabalho e outras espécies de *Caulimoviridae* disponíveis no GenBank (Tabela 2), foram utilizadas para determinar o verdadeiro estado taxonômico dos

isolados virais. De acordo com os critérios de ICTV (GEERING; HULL, 2012).para demarcação de espécies dentro do gênero *Badnavirus*, diferenças de sequência dentro da RT/RNaseH de mais do que 20% são consideradas suficiente para identificar os vírus como espécies distintas.

Alinhamentos múltiplos de sequências parciais de nucleotídeos da RT/RNaseH preparados com o programa MUSCLE (EDGAR, 2004) disponível no pacote computacional MEGA 5 (Molecular Evolutionary Genetics Analysis)(TAMURA et al., 2007) foram utilizados para análises filogenéticas

Uma árvore de Máxima Verossimilhança (ML) foi inferida usando RAxML v. 7.0.3 (STAMATAKIS, 2006) assumindo o modelo de substituição de nucleotídeos General Time Reversible (GTR), com o modelo de distribuição gama da taxa de heterogeneidade. A robustez de cada ramo individual foi estimada a partir de bootstrap de 1000 replicações. As árvores foram visualizadas e editadas usando FigTree (tree.bio.ed.ac.uk/software/figtree) e CorelDraw.

**Tabela 2 - Vírus do gênero *Badnavirus* e *Tungrovirus* utilizadas para as comparações de sequências e análises filogenéticas.**

| Espécies                                          | Acrônimo   | Nº de acesso GenBank  |
|---------------------------------------------------|------------|-----------------------|
| <i>Banana streak GF virus</i>                     | BSGFV      | AY493509              |
| <i>Banana streak Ol virus</i>                     | BSOLV      | AJ002234              |
| <i>Banana streak mysore virus</i>                 | BSMYV      | AY805074              |
| <i>Banana streak VN virus</i>                     | BSVNV      | AY750155              |
| <i>Bougainvillea chlorotic vein banding virus</i> | BCVBV      | EU034539 (=NC011592)  |
| <i>Cacao swollen shoot virus</i>                  | CSSV       | L14546 (= NC001574)   |
| <i>Citrus yellow mosaic virus</i>                 | CIYMV      | AF347695 (=NC003382)  |
| <i>Commelia yellow mottle virus</i>               | ComyMY     | X52938 (=NC001343)    |
| <i>Dioscorea bacilliform SN virus</i>             | DBSNV      | DQ822073 (=NC009010)  |
| <i>Dioscorea bacilliform Al virus</i>             | DBALV      | X 94576               |
| <i>Pineapple bacilliform comosus virus</i>        | PBV        | NC14648               |
| <i>Rice tungro bacilliform virus</i>              | RTBV       | NC001914              |
| <i>Sugarcane bacilliform IM virus</i>             | SCBIMV-QLD | AJ277091 (=NC003031)  |
| <i>Sugarcane bacilliform Mor virus</i>            | SCBMV-MOR  | M89923 (=NC008017)    |
| <i>Taro bacilliform virus</i>                     | TABV       | AF357836 (=NC_004450) |

Fonte: Autora, 2013.

## 2.4 Descritores gerais da variabilidade da população e parametrização dos mecanismos evolutivos

Descritores da variabilidade molecular foram estimados para a população viral, incluindo número total de sítios segregantes ( $s$ ), número total de mutações ( $\Theta$ ), número médio de diferença de nucleotídeos entre sequências ( $k$ ), número de haplótipos ( $h$ ), diversidade de nucleotídeos ( $\pi$ ), diversidade haplotípica ( $H_d$ ), taxa de mutação da população estimada por Watterson baseada no número total de sítios segregantes ( $\theta_w$ ), taxa de mutação da população estimada por Watterson baseada no número total de mutações ( $\theta$ - $\Theta$ ) usando DnaSP v. 5.10 (ROZAS et al., 2009). Adicionalmente, o índice de fixação  $F$  de Wright (WRIGHT, 1951) calculado usando DnaSP v. 5.10 (ROZAS et al., 2009) foi utilizado para detecção de subdivisão da população de DBALV por região geográfica.

Quatro tipos de testes de neutralidade foram empregados para testar a hipótese de ocorrência de seleção em populações utilizando o programa DnaSP (ROZAS et al., 2009): Tajima's  $D$ , Fu and Li's  $D^*$  e  $F^*$  e o teste baseado em número de substituições sinônimas ( $D_s$ ) e não-sinônimas ( $D_n$ ) com o modelo de Pamilo-Bianchi-Li (PBL). Para detecção de sítios sob seleção positiva e negativa as sequências parciais foram analisadas utilizando quatro diferentes métodos baseados em Máxima Verossimilhança (ML) disponível no servidor DataMonkey ([www.datamonkey.org](http://www.datamonkey.org)): Single-Likelihood Ancestor Counting (SLAC), Fixed-Effect Likelihood (FEL), Random-Effect Likelihood (REL) and Partitioning for Robust Inference of Selection (PARRIS) (DELPORTE et al., 2010).

Para evitar falsos resultados de seleção causados por recombinação, o conjunto de dados inicialmente foi analisado usando Genetic Algorithm Recombination Detection (GARD), a fim de detectar supostos eventos de recombinação. SLAC também foi usada para estimar as taxas sinônimas e não-sinônimas ( $d_N/d_S$ ) baseada na árvore filogenética inferida no GRAD.

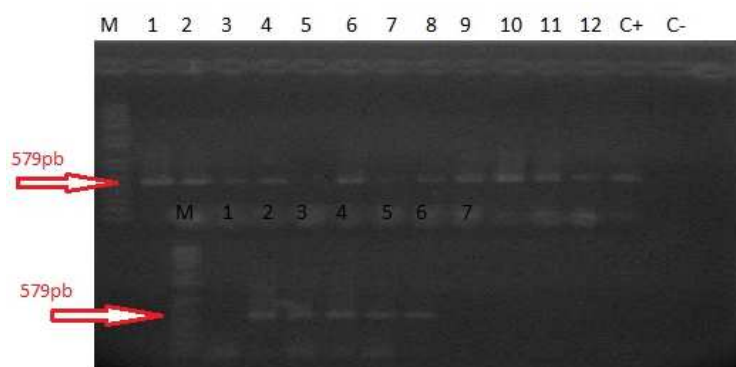
O modelo de substituição de nucleotídeos general time reversible nucleotide (GTR) foi assumido e fatores Bayes maiores que 50 e  $P$ -value menores que 0,1 foram usados com limiares de significância para os métodos REL e FEL.

### 3. RESULTADOS e DISCUSSÃO

#### 3.1 Amplificação do domínio RT/ RNaseH de badnavírus e sequenciamento

Um total de 150 amostras foram analisadas, das quais 108 foram positivas para a presença de badnavírus com base na detecção de um fragmento de 579 pb após amplificação da região RT/RNaseH da ORF3 do genoma viral (Figura 5). Todas as amostras positivas para badnavírus foram sequenciadas, no entanto foi possível montar contigs de boa qualidade a partir de somente 48 das sequências.

**Figura 5 - Amplificação do domínio RT/ RNaseH de badnavírus. Números 1-12 e 1-7 amostras provenientes de Bonito-PE; M, 1KB plus ladder; C+, controle positivo e C-, controle negativo.**

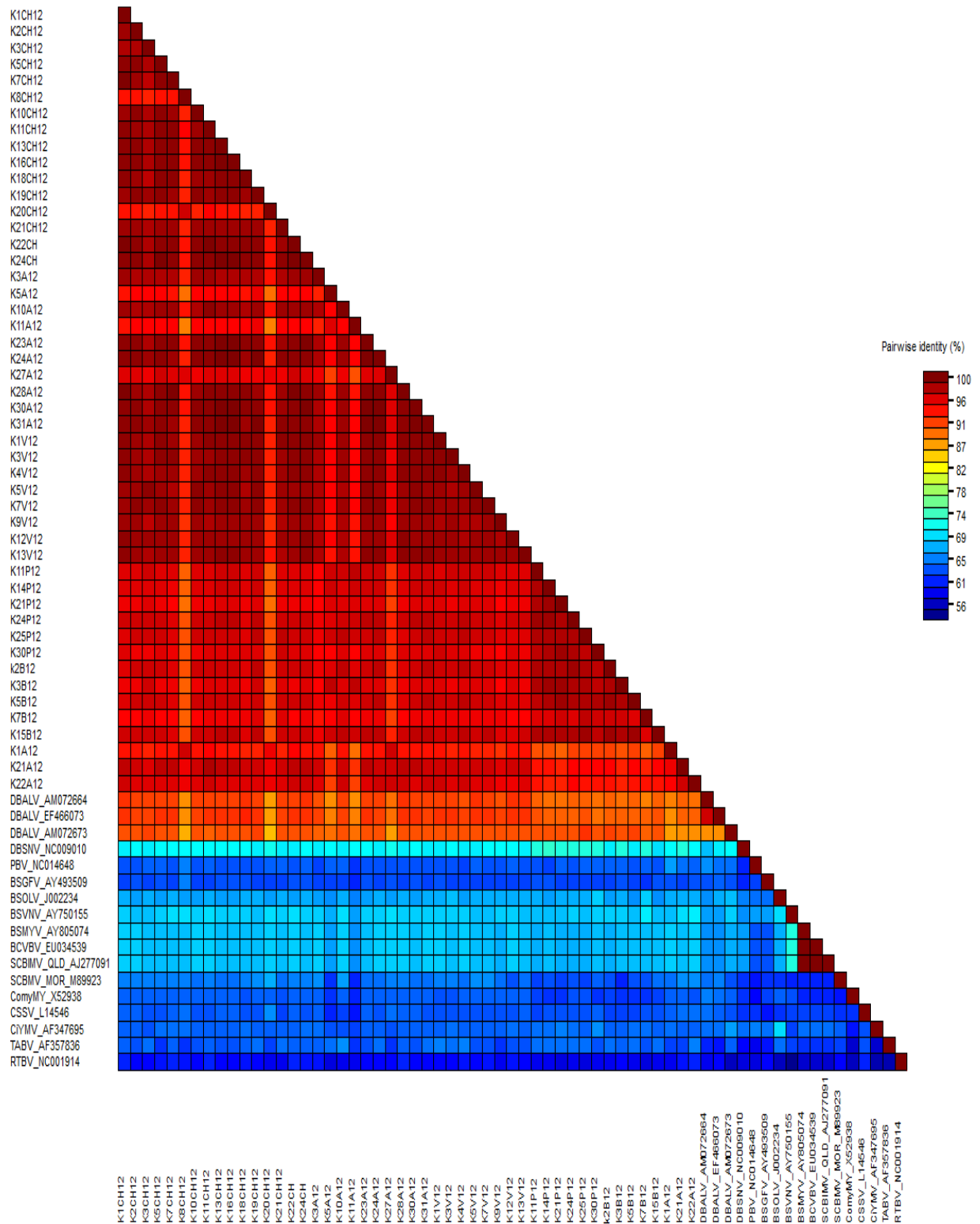


Fonte: Autora, 2013.

#### 3.2 Comparação de sequências e análise filogenética

As análises preliminares utilizando o algoritmo BLAST e comparações de sequências pareadas com SDT revelaram que as 48 sequências obtidas neste trabalho apresentaram maior identidade 85 a 97% com três diferentes isolados (AM072664, AM072673 e EF46607) da espécie já descrita *Dioscorea bacilliform AL virus* (DBALV) (Figura 6). Esse resultado corrobora estudos anteriores sobre a diversidade de vírus que incidem sobre inhame na região Nordeste do Brasil, os quais também demonstraram que DBALV é a única espécie de badnavírus ocorrendo nos estados de Alagoas, Bahia, Paraíba e Pernambuco (ANDRADE, 2007; LIMA et al., 2013). Uma situação oposta é observada em países da África Ocidental e Sul do Pacífico, onde estudos de diversidade de badnavírus em inhame revelaram a ocorrência de alta diversidade de espécies, sendo encontradas ao menos as duas espécies aceitas pelo ICTV, DBALV e *Dioscorea bacilliform SN virus* (DBSNV) (GEERING; HULL, 2012) além de outras espécies tentativas (ENI et al., 2008; KENYON et al., 2008; BOUSALEM et al, 2009).

**Figura 6 - Matriz bidimensional representando a porcentagem de identidade de comparações pareadas de seqüências nucleotídicas da região RT/RNaseH dos *Badnavirus* desse trabalho e outras espécies do gênero disponíveis no GenBank.**



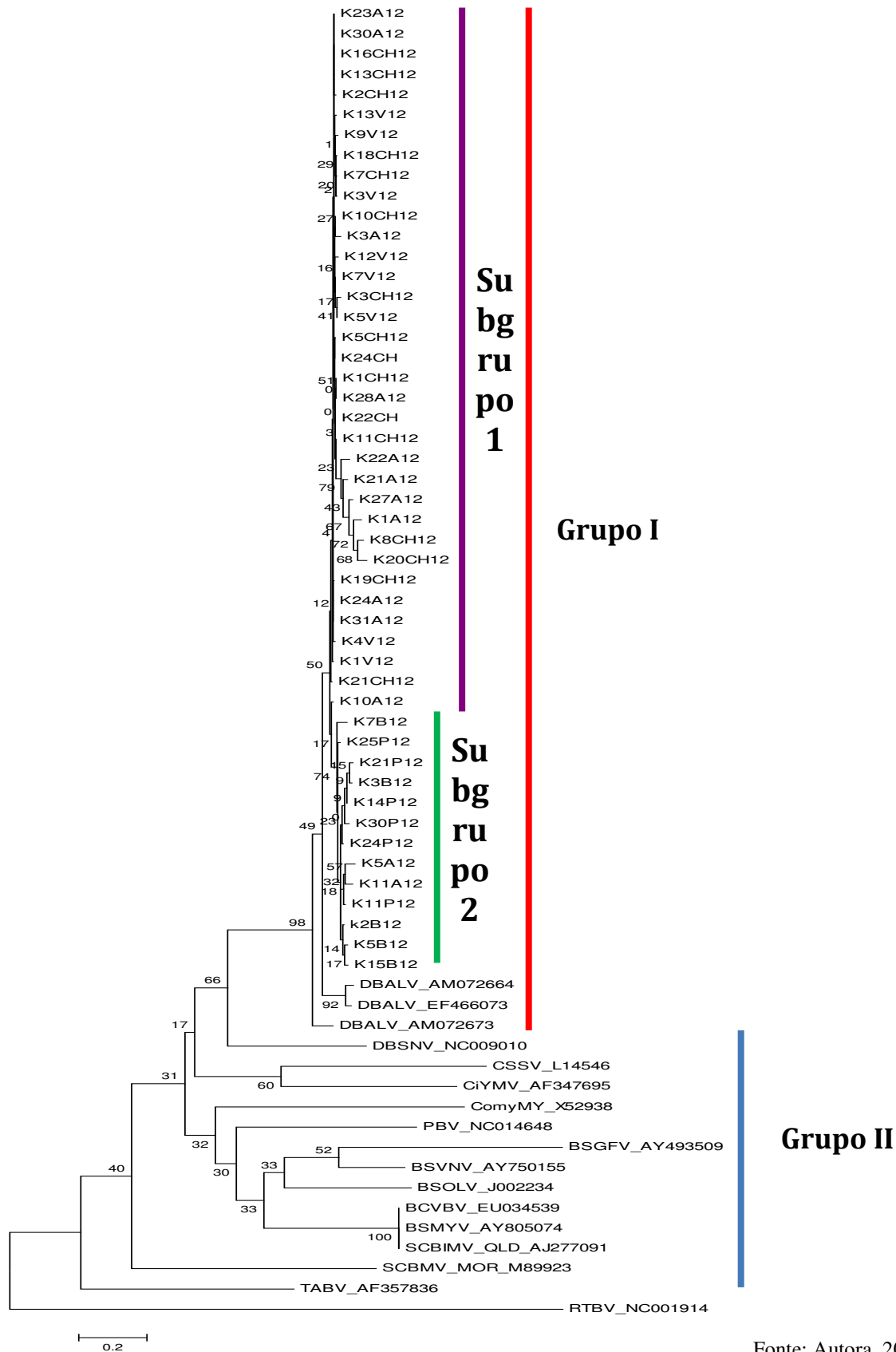
Fonte: Autora, 2013.

Os relacionamentos filogenéticos foram analisados com base nas sequências de nucleotídeos do domínio RT/RNaseH da ORF3 dos badnavírus. Uma árvore de Máxima Verossimilhança (ML) foi construída incluindo sequências dos isolados obtidos nesse estudo e sequências referências do gênero *Badnavirus* disponíveis no GenBank (Figura 7). Essa análise indicou que 48 isolados formaram um ramo monofilético (Grupo I) com a espécie DBALV, as quais possuem um ancestral comum com a espécie *Dioscorea bacilliform SN virus* (DBSNV). Os isolados mostraram uma tendência a agrupar por região geográfica, formando dois subgrupos dentro do grupo I. O subgrupo 1 contém isolados provenientes dos municípios de Arapiraca, Chã Preta e Viçosa, todos pertencentes ao estado de Alagoas, enquanto o subgrupo 2 inclui todos os isolados provenientes dos municípios de Alhandra e Bonito, oriundos da Paraíba e Pernambuco, respectivamente. O grupo II inclui as outras espécies pertencentes ao gênero *Badnavirus* (Figura 7).

Análises filogenéticas baseadas no domínio RT/RNaseH de badnavírus que infectam *Dioscorea* spp. nas Ilhas do Sul do Pacífico revelaram alta diversidade genética, sendo encontradas 10 diferentes espécies, incluindo o DBALV. Estas agruparam mais intimamente com relação às suas hospedeiras do que por origem geográfica (KENYONet al., 2008). No entanto, em estudos sobre a diversidade de badnavírus que incidem sobre o inhame no Caribe, América do Sul e África, houve uma correlação significativa entre os isolados virais e sua distribuição geográfica, sendo os grupos separados principalmente por sua origem. Grupos formados com base na hospedeira também foram observados (BOUSALEM et al., 2009).

A espécie DBALV é eficientemente transmitida por inoculação mecânica (ODU et al., 2006) e pela cochonilha *Planococcus citri* (PHILLIPS et al., 1999). Contudo a taxa de transmissão de badnavírus por cochonilha é geralmente considerada baixa, sendo o maior risco de disseminação do vírus causado pelo de material propagativo infectado (LOCKHART; OLSZEWSKI, 1993). Desta forma, acredita-se que a propagação vegetativa seja um dos principais fatores que contribuem para a prevalência de DBALV na região nordeste do Brasil, devido à proximidade geográfica das áreas de cultivo avaliadas e ausência de barreiras fitossanitárias, permitindo assim a prática comum de troca de túberas-semente infectada entre as mesmas.

**Figura 7 -** Árvore de Máxima Verossimilhança baseada na sequência da região RT/RNaseH de isolados de DBALV obtido neste trabalho e outras espécies de Badnavirus disponíveis no GenBank. O Rice tungro bacilliform vírus (RTBV), gênero *Tungrovirus*, foi utilizado como “outgroup”.



### 3.3 Determinação da variabilidade genética da população e parametrização dos mecanismos evolutivos

A análise dos descritores demonstrou que a população de DBALV apresenta uma alta variabilidade genética quando comparada a begomovírus (ssDNA) que infectam tomateiro (Tomato common mosaic virus, ToCmMV; Tomato chlorotic mottle virus, ToCMoV, *Tomato severe rugose virus*, ToSRV e *Tomato yellow vein streak virus*, ToYVSV), fava (Bean golden mosaic virus, BGMV) e plantas daninhas, *Blainvillea rhomboidea* (Blainvillea yellow spot virus, BIYSV) e *Cleome affinis* (Cleome leaf crumple virus, CILCrV), com exceção de *Macroptilium* spp. (Macroptilium yellow spot virus, MaYSV). Por exemplo, valores para diversidade nucleotídica ( $\pi$ ) são 0,0333 para DBALV, 0,0154 para BGMV, 0,02466 para BIYSV, 0,0191 para CILCrV, 0,0572 para MaYSV, 0,0143 para ToCmMV, 0,0071 para ToCMoV, 0,0102 para ToSRV e 0,021 para ToYVSV (Tabela 3).

**Tabela 3 -Variabilidade genética da população de *Dioscorea bacilliformAL* virus (DBALV) obtida de amostras de *Dioscorea* spp. coletadas em três estados da região nordeste do Brasil.**

| População           | Número de sequências | Tamanho do genoma | $s^*$ | Eta <sup>†</sup> | $k^{\ddagger}$ | $\pi^{\S}$ | $h^I$ | Hd <sup>¶</sup> | $\theta-w^{\#}$ | $\theta-Eta^*$ |
|---------------------|----------------------|-------------------|-------|------------------|----------------|------------|-------|-----------------|-----------------|----------------|
| DBALV               | 48                   | 435               | 86    | 98               | 14.484         | 0.0333     | 40    | 0.992           | 0.0452          | 0.0515         |
| BGMV <sup>a</sup>   | 20                   | 2620              | 251   | 265              | 40.321         | 0.0154     | 18    | 0.989           | 0.0271          | 0.0295         |
| BIYSV <sup>b</sup>  | 7                    | 2661              | 200   | 211              | 65.619         | 0.0246     | 7     | 1.000           | 0.0307          | 0.0323         |
| CILCrV <sup>c</sup> | 14                   | 2756              | 253   | 267              | 51.758         | 0.0191     | 14    | 1.000           | 0.0294          | 0.0311         |
| MaYSV <sup>d</sup>  | 10                   | 2658              | 402   | 419              | 150.17         | 0,0572     | 10    | 1.000           | 0,0537          | 0,0542         |
| ToCmMV <sup>b</sup> | 22                   | 2560              | 103   | 104              | 36.645         | 0.0143     | 20    | 0.987           | 0.0110          | 0.0111         |
| ToCMoV <sup>b</sup> | 22                   | 2619              | 135   | 138              | 18.351         | 0.0070     | 22    | 1.000           | 0.0141          | 0.0144         |
| ToSRV <sup>b</sup>  | 27                   | 2588              | 148   | 159              | 26.530         | 0.0102     | 26    | 0.997           | 0.0148          | 0.0159         |
| ToYVSV <sup>b</sup> | 26                   | 2562              | 49    | 49               | 5.3810         | 0.0021     | 25    | 0.997           | 0.0050          | 0.0050         |

<sup>a</sup>Ramos-Sobrinho (2010)

<sup>b</sup>Rocha et al. (2013)

<sup>c</sup>Silva et al. (2011)

<sup>d</sup>Silva et al. (2012)

\* Número total de sítios segregantes.

† Número total de mutações.

‡ Número médio de diferenças de nucleotídeos entre sequências (Taxa de mutação da população estimada por Tajima, $\theta$ ).

§ Diversidade de nucleotídeos.

I Número de haplótipos.

¶ Diversidade haplotípica.

# Taxa de mutação da população estimada por Watterson baseada no número total de sítios segregantes.

\* Taxa de mutação da população estimada por Watterson baseada no número total de mutações.

Fonte: Autora, 2013.

A alta diversidade de sequências de RT/RNaseH observada para badnavírus que infectam inhame é consistente com aquela encontrada para badnavírus de banana (JAMES et al., 2011), cacau (MULLER; SACKKEY 2005), cana-de-açúcar (MULLER et al., 2011) e taro (YANG et al., 2003). Contudo, a variabilidade genética baseada na região RT/RNaseH entre isolados de uma mesma espécie de badnavírus de inhame se mostrou muito baixa (BOUSALEM et al., 2009). A alta variabilidade genética de badnavírus tem sido atribuída à replicação inacurada por transcrição reversa presente em *Caulimoviridae* (BOUSALEM et al., 2009).

A população de DBALV foi obtida de três estados da região nordeste e para verificar a existência de subdivisão dessa população foi calculado o índice de Fixação F de Wrigth (Fst). De acordo com Weir (1996), valores de Fst de 0 a 0,05 indicam pouca diferenciação genética; de 0,05 a 0,15 moderada diferenciação; de 0,15 a 0,25, grande diferenciação; e  $> 0,25$ , alta diferenciação. A análise da diferenciação da população revelou um valor Fst total de 0,35085 para o gene da transcriptase reversa (RT) indicando a existência de uma diferenciação da população de DBALV por região geográfica. Isto pode ser evidenciado na árvore de Máxima Verossimilhança (ML), onde os isolados virais provenientes de Bonito-PE e Alhandra-PB, compartilham o mesmo clado (Figura 7). Esses resultados sugerem que a subpopulação de DBALV proveniente de Pernambuco e Paraíba é distinta daquela oriunda de Alagoas.

Testes de neutralidade foram usados para acessar o tipo de seleção que está atuando sobre a sequência codificadora do domínio RT da ORF III de DBALV. Valores negativos foram obtidos, mas não estatisticamente suportados a  $p < 0.10$ , para Tajima D, Fu & Li D e Fu & Li F, não sendo possível confirmar que seleção purificadora está atuando na população. Contudo, a pressão de seleção pode ser mais acuradamente identificada pela proporção entre a taxa de substituições não sinônimas (dN) e a taxa de substituições sinônimas (dS). A razão dN/dS é esperada exceder uma unidade quando a seleção natural promove mudanças na sequência de proteínas (seleção diversificadora), enquanto numa razão mais baixa que uma unidade é esperada se a seleção natural suprime mudanças na proteína (seleção purificadora ou negativa) (YANG; BIELAWSKI, 2000). A razão  $dN/dS < 1$  estimada para o domínio RT indica que seleção purificadora (negativa) está atuando sobre esta população (Tabela 4).

**Tabela 4 - Resultados dos diferentes testes de neutralidade realizados para o domínio RT da ORF III da população de *Dioscorea bacilliform* AL virus (DBALV) obtida de *Dioscorea* spp. no Nordeste do Brasil.**

| População | ORF* | Tajima D          | Fu & Li D     | Fu & Li F    | dN/dS        |
|-----------|------|-------------------|---------------|--------------|--------------|
| DBALV     | RT   | -1,27385(ns)<br>† | - 1,63074(ns) | -1,79041(ns) | 0.096759(ns) |

\*RT, Transcriptase reversa;

†ns, valores não significantes a  $p < 0.10$

Fonte: Autora, 2013.

Os testes adicionais para detecção de seleção ocorrendo na população de DBALV, utilizando SLAC, REL, FEL e PARRIS, revelaram vários sítios sob seleção purificadora (negativa), corroborando os dados da razão dN/dS (Tabela 5).

Seleção purificadora também foi encontrada como uma das principais forças evolutivas atuando sobre vírus de genoma de ssDNA como os begomovírus BGMV em fava (RAMOS-SOBRINHO, 2010), CILCrV em *Cleome affinis* (SILVA et al., 2011), MaYSV em *Macroptilium* spp. (SILVA et al., 2012), ToCmMV, ToCMoV, ToSRV e ToYVSV em tomateiro (ROCHA et al., 2013) e vírus de genoma de dsDNA, como o *hepdnavírus* Avian hepatitis E virus (aHEV) em aves (BÁNYAI et al., 2010).

Na ausência de seleção positiva, mecanismos como amostragem de hospedeiras e vetor induzindo efeitos gargalo e fundador (GARCIA-ARENAL, 2003), aliado a um novo ambiente biogeográfico tem provavelmente contribuído para formação de linhagens distintas de badnavírus de inhame (BOUSALEM et al., 2009).

**Tabela 5 - Detecção de seleção utilizando os métodos ML disponíveis no DataMonkey.**

|                     | SELEÇÃO<br>POSITIVA | SELEÇÃO NEGATIVA                                                                                                                                                              |
|---------------------|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SLAC <sup>#</sup>   | <b>NÃO</b>          | 1,4,7,13,29,30,40,41,43,54,55,56,58,59,60,67,69,80,<br>82,90,102,111,118,124,135,139                                                                                          |
| FEL <sup>*</sup>    | <b>NÃO</b>          | 1,4,7,13,23,28,29,30,32,40,41,43,45,54,55,56,58,59,<br>60,67,69,79,80,82,89,90,102,103,105,109,111,112,<br>117,118,119,124,135,139,140,145                                    |
| REL <sup>†</sup>    | <b>NÃO</b>          | 1,4,7,13,23,27,28,29,30,40,41,43,44,45,48,54,55,56,<br>58,59,60,67,69,79,80,82,89,90,95,98,99,102,103,1<br>05,107,109,111,112,117,118,119,124,129,130,135,<br>139,140,142,145 |
| PARRIS <sup>§</sup> | <b>NÃO</b>          |                                                                                                                                                                               |
| GARD <sup>‡</sup>   |                     | Não foi detectado recombinação                                                                                                                                                |

<sup>#</sup> Single-Likelihood Ancestor Counting (SLAC),<sup>\*</sup> Fixed-Effect Likelihood (FEL)

Fonte: Autora, 2013.

<sup>†</sup> Random-Effect Likelihood (REL)<sup>§</sup> Partitioning for Robust Inference of Selection (PARRIS)<sup>‡</sup> Genetic Algorithm Recombination Detection (GARD)

#### **4. CONCLUSÕES**

A análise da estrutura genética da população indicou que a variabilidade de DBALV é alta, semelhante àquela encontrada para vírus de ssDNA e que a mutação é um importante processo evolutivo contribuindo para a diversificação dessa espécie. O fato de DBALV estar sobre possível influência de seleção purificadora sugere que a ocorrência de mutações não é suficiente para explicar completamente a sua variabilidade genética. Desta forma, estudos são necessários para demonstrar a influência de forças evolutivas adicionais atuando sobre esta população.

## REFERÊNCIAS BIBLIOGRÁFICAS

AHLAWAT Y. S. P. et al. A. Association of a badnavirus with citrus mosaic disease in India. **Plant Disease**, Saint Paul, v.80, n.5, p.590-592, 1996.

ALTSCHUL, S.F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, U.S.A, v. 215, p. 403 - 410, 1990.

AMUSA, N. A. *et al.* Yam diseases and its management in Nigeria. **Journal of Biotechnology**, African, v. 2, p. 497 - 502, 2003.

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, PE, 2007.

ANDRADE, G. P. et al. Planta de *Dioscorea alata* infectada com YMMV no estado de Pernambuco. **Fitopatologia Brasileira**, Brasília, DF, v. 31, p. S 344, 2006.

ARBOLEDA M.; AZZAM O. Inter- and intra-site genetic diversity of natural field populations of rice tungro bacilliform virus in the Philippines. *Arch Virol*, v. 145(2), p 275-89, 2000.

ASIEDU, R.; SARTIE, A. Crops that feed the World 1. Yams: Yams for income and food security. **Food Security**, v. 2, p. 305-315, 2010.

BÁNYAI K. et al. Putative novel genotype of avian hepatitis E virus, Hungary. *Emerg Infect Dis.* V.8, p. 1365-1368, 2010.

BIAGINI, P. et al. Anelloviruses. In FAUQUET, C. M. et al. **Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses**. New York: Academic Press, 2004.

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**, Great Britain, v. 74, p.15-22, 1993.

BOUSALEM, M. et al. Taxonomy, molecular phylogeny and evolution of plant reverse transcribing viruses (family *Caulimoviridae*) inferred from full-length genome and reverse transcriptase sequences. **Archives of Virology**, v. 153, p.1085-1102, 2008.

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.

BRIDDON R. W. P. S.; BRUNT, A.; HULL, R. Analysis of the sequence of *Dioscorea Alata bacilliform virus*: comparison to others members of the badnavirus group. **Virus Genes**, v. 18, p. 277-283, 1999.

BRUNT A. A. et al. **Plant Viruses Online**: Descriptions and Lists from the VIDE Database, 1996. Disponível em: <<http://biology.anu.edu.au/Groups/MES/vide/>> Acesso em: 15 maio. 2013.

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.

COURSEY, D. G. Yams - *Dioscorea* spp. (Dioscoreaceas). In: SIMMONDS, N.W. **Evolution of crop plants**, New York. Longman, p. 70-74, 1976.

DAVINO S. et al. Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. **Virus Res**, 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.

DELPORT W. et al. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. **Bioinformatics**, V. 19, p. 2455-7, 2010

DOYLE, J.J.; DOYLE, J.L. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. **Phytochemistry Bulletin**, v. 19, p. 11-15, 1987.

DRAKE J.W. A constant rate of spontaneous mutations in DNA- based microbes. **Proc Natl Acad Sci. USA**, V. 88, n.16, p. 7160 –7164, 1991.

DRAKE R.R. et al. Identification of a nucleotide binding site in HIV-1 integrase. **Proc Natl Acad Sci, USA**, V. 95, n.8, p. 4170-5, 1998.

DRAKE J.W.; HOLLAND J.J. Mutation rates among RNA viruses. **Proc Natl Acad Sci. USA**, V. 96, n. 24, p. 13910-13913, 1999.

DRAKE J.W. RATES OF SPONTANEOUS mutations among RNA viruses. **Proc Natl Acad Sci. USA**, V. 90, n. 9, p. 4171–4175, 1993.

DUFFY S.; HOLMES E.C. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. **J Gen Virol.** V. 90, p. 1539–1547, 2009.

DUFFY S.; SHACKELTON L.A.; HOLMES E.C. Rates of evolutionary change in viruses: patterns and determinants. **Nat Rev Genet.** V.9, p. 267–276, 2008

EDGAR RC. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. **Bmc Bioinformatics**, v. 5, p. 1-19, 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.

FAO.FAOSTAT.Agricultural statistics database. Rome: **World Agricultural Information Centre**, 2010. Disponível em: <<http://faostat.fao.org/default.aspx>>. Acesso em: 14 jun. 2013.

FAUQUET, C. et al. **Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses**. Amsterdam, Elsevier Academic Press. 2005.

FROISSART R. et al. Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. **PLoS Biol.** v. 3, n. 3, p.89, 2005.

GALLIAN P. et al. High genetic diversity revealed by the study of TLMV infection in French hemodialysis patients. **J. Med. Virol.** v. 67, n. 4, p. 630-635, 2002.

GARCÍA-ANDRÉS S. et al. 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.; FRAILE, A.; MALPICA, J.M. Variation and evolution of plant virus populations. **International Microbiology**, v. 6, p. 225-232, 2003.

GARCIA-DIAZ M.; BEBENEK K. Multiple functions of DNA polymerases. **CRC Crit Rev Plant Sci.** v. 26, n. 2, p. 105-122, mar. 2007.

GAUHL F.; PASBERG-GAUHL, C.; HUGHES, J. D'A. First report of banana streak badnavirus in plantain landraces in southern Cameroon, Central Africa. **Plant Disease**, v. 81, p. 1335-1335, 1997.

GAL S. et al. Genomic homologous recombination in planta. **EMBO J.** v. 10, n. 6, p. 1571-1578, jun. 1991.

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

GEERING A.D.; SCHARASCHKIN T.; TEYCHENEY P.Y. The classification and nomenclature of endogenous viruses of the family Caulimoviridae. **Arch Virol.** v. 155, n.1, p. 123-31, 2010

GE, L. M. et al. Genetic structure and population variability of tomato yellow leaf curl China virus. **J Virol.** v. 81, p. 5902-5907, 2007

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.

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. **Proc Natl Acad Sci.** v. 96, n. 14, p. 8022-7, 1999

GOUDOU-URBINO, C. K. G.; QUIOT, J. B.; DUBERN, J.A etiology and ecology of a yam mosaic disease in Burkina Faso. **Tropical Sciences**, v. 36, p.34-40, 1996.

HARPER, G. et al. Viral sequences integrated into plant genomes. **Annual Review of Phytopathology**, v. 40, p.119-136, 2002.

HARPER, G. et al. The diversity of *Banana streak virus* isolates in Uganda. **Archives of Virology**, v. 150, p. 2407-2420, 2005.

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

HOLLAND, J. J. et al. Rapid evolution of RNA genomes. **Science**. v. 215, p. 1577- 1585, 1982.

HOU, W. C. et al. Yam (*Dioscorea alata*) tuber mucilage antioxidant activities in-vitro. **Planta Medica**, v.68, n. 12, p. 1072-1076, 2002.

HUGHES, J. D. A.; DONGO, L; ATIRI, G. I. Viruses infecting cultivated yams (*Dioscorea alata* and *D. rotundata*) in Nigeria. **Phytopathology**, v. 87, p. 45, 1997.

HULL R. et al. Family *Caulimoviridae*. In: FAUQUET, C.M.; MAYO, M.A.; MANILOFF, J.; DESSELBERGER, U.; BALL, A.L. (Ed.). **Virus taxonomy**: classification and nomenclature of viruses. San Diego: Elsevier, 2005. p. 385-396. (Eighth Report of the International Committee on Taxonomy of Viruses).

IBGE - Instituto Brasileiro de Geografia e Estatística. **Censo Agropecuário**. Rio de Janeiro, Brasil. 2006.

ISNARD M. et al. Uasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars. *J Gen Virol*.v. 79, p. 3091-9, 1998.

JAMES A.P. et al. 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. et al. Rates of molecular evolution in RNA viruses: a quantitative phylogenetics analysis. **J Mol Evol**. v. 54, p. 152–161, 2002.

KENYON, L.; LEBAS, B.S.M.; SEAL, S.E. Yams (*Dioscorea* spp.) from the South Pacific Islands contain many novel badnaviruses: implications for international movement of yam germplasm. **Archives of Virology**, v. 3, p. 877–889, 2008.

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. et al. Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. *J Gen Virol.* v. 12, p. 3458-68, 2007.

LIMA, S.L. et al. Variabilidade genética de isolados de badnavírus infectando inhame (*Dioscorea* spp.) no Nordeste do Brasil. **Tropical Plant Pathology**, v.39, n. 4, p. 349-353, 2013.

LOCKART, B. E. L. et al. Identification of *Piper yellow mottle virus*, a mealybug transmitted badnavirus infecting *Piper* spp. in Southeast Asia. **European of Plant Pathology**, v. 103, p. 303-311, 1997.

LOCKHART, B. E. L.; AUTREY, L. J. C. Occurrence in sugarcane of a bacilliform virus related serologically to banana streak virus. **Plant Disease**, v. 72, p. 230-233, 1988.

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.

LLORENS, C. et al. Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. **Biology Direct**, v. 4, p. 41, 2009.

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.

MIGLIORI A.; LASTRA, R. Etude de virus présent chez *Commelina diffusa* Burm. Guadeloupe. **Annals Phytopathology**, v. 10, p. 467-477, 1978.

MOHAMED, K. A.; MANTELL, S. H. Incidence of virus symptoms in yam (*Dioscorea* sp.) foliage in the commonwealth Caribbean. **Tropical Agriculture**, v. 53, p. 255-261, 1976.

MONCI F. et al. 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.

MORILLA, G. et al. 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.

MOURA, R. M. Doenças do inhame-da-costa (*Dioscorea cayennensis*). In: KIMATI, H.; MORIM, L.; REZENDE, J. A. M; BERGAMIN FILHO, A.; CAMARGO, L. E. A. **Manual de Fitopatologia: doenças das plantas cultivadas**. 4 ed. São Paulo: Agronômica Ceres, p. 415-419, 2005.

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**, 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23340592>>. Acesso em: jul. 2013.

MULLER E. et al. High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. **Virus Res.** v. 160, n. 2, p. 414 - 419, 2011.

MULLER E.; SACKKEY S. Molecular variability analysis of five new complete cocoa swollen shoot virus genomic sequences. **Archives of Virology**, v. 150, p. 53-66, 2005.

MUNFORD R.A.; SEAL, S.E. Rapid single-tube immunocapture RT-PCR for the detection of two yam potyviruses. **Journal of Virological Methods**, v. 69, p. 73–79, 1997.

NDOWORA T. D. et al. Evidence that *Badnavirus* infection in *Musa* can originate from integrated pararetroviral sequences. **Virology**, v. 255, p. 214-220, 1999.

NEWKE F. I. et al. Production costs in the yam-based cropping systems of south-western Nigeria. **Resource and Crop Management Division**. Research Monograph No 6, IITA Ibadan. 29 pp, 1991.

ODU, B.O. et al. Responses of white yam (*Dioscorea rotundata*) cultivars to inoculation with three viruses. **Plant Pathology**, v. 53, p. 141–147, 2004.

ODU, B.O. et al. Reaction of White Guinea yam (*Dioscorea rotundata* Poir.) genotypes to virus diseases in four agroecological zones in Nigeria. **Journal of Phytopathology**, v. 154, p. 11-12, 2006.

- OFFIE, S.K. Virus and viral diseases of sub-Saharan Africa: analysis of responses to questionnaires by scientists in sub-Saharan Africa. **Plant Virology in Sub-Saharan Africa**, p. 128-136, 2001.
- O'HAIR, S. K. Tropical root and tuber crops. **Horticultural Reviews**, v. 12, p. 157-196, 1990.
- OMURA T. et al. Improved purification procedure and RNA segments of rice ragged stunt virus. **Annals of the Phytopathological Society of Japan**, v. 49, p. 670-675, 1983.
- 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, 1999a.
- PADIDAM, M.; SAWYER, S.; FAUQUET, C.M. Possible emergence of new geminiviruses by frequent recombination. **Virology**, v. 265, p. 218-224, 1999b.
- 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. S183, 2005.
- PITA, J.S. et al. 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.
- PREISS, W.; JESKE, H. Multitasking in replication is common among geminiviruses. **Journal of Virology**, v. 77, p. 2972-2980, 2003.
- RAMOS, N. F. et al. Presença dos vírus da mancha anelar e do amarelo letal em frutos de mamoeiro comercializados. **Tropical Plant Pathology**, v. 33, p. 449-452, 2008.
- RAMOS-SOBRINHO R. et al. Genetic structure of a population of the begomovirus Bean golden mosaic virus (BGMV) that infects lima bean (*Phaseolus lunatus* L.) in the state of Alagoas, Brazil. Program and Abstracts, 6th International Geminivirus Symposium and 4th International ssDNA Comparative Virology Workshop. México, 2010.
- RANEY, J. L.; DELONGCHAMP R. R.; VALENTINE C. R. Spontaneous mutant frequency and mutation spectrum for gene A of \_X174 grown in *E.coli*. **Environ. Mol. Mutagen.** v.44, p. 119-127, 2004.

RIBEIRO, S.G. et al. Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. **Archives of Virology**, v. 148, p. 281-295, 2003.

RITZINGER, C. H. S. P. et al. Aspectos fitossanitários da cultura do inhame. Cruz das Almas: **Documentos EMBRAPA/SPI**, 2003, 39p.

ROZAS, et al. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, p. 1451-1452, 2009.

SANTOS, E. S. Inhame (*Dioscorea* spp.): aspectos básicos da cultura. João Pessoa: **EMEPA-PB/SEBRAE**, 1996. 158p.

SANTOS, E. S. Manejo sustentável da cultura do inhame (*Dioscorea* sp.) no Nordeste do Brasil. 2002. In: II Simpósio Nacional sobre as Culturas do Inhame e do Taro. **Anais eletrônicos**. João Pessoa, PB, EMEPA-PB, 2002. Disponível em: <[http://www.emepa.org.br/anais/an\\_sincit2\\_v1.php](http://www.emepa.org.br/anais/an_sincit2_v1.php)>. Acesso em: 25 mai. de 2013.

SANTOS, E. S.; MACÊDO, L. S. Tendências e potencialidades da cultura do Inhame (*Dioscorea* sp.) no Nordeste do Brasil. **Anais eletrônicos**. João Pessoa, PB, EMEPA-PB, 2006. Disponível em: <<http://www.emepa.org.br/php>>. Acesso em: 15 mai. 2013.

SANZ, A.I. et al. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. **Journal of General Virology**, v. 81, p. 1839-1849, 2000.

SAUNDERS, K.; BEDFORD, I.D.; STANLEY, J. Pathogenicity of a natural recombinant associated with ageratum yellow vein disease: implications for geminivirus evolution and disease aetiology. **Virology**, v. 282, p. 38-47, 2001.

SHACKELTON, L.A. et al. JC vírus evolution and its association with human populations. **J Virol**.v. 80, p. 9928–9933, 2006.

STAGINNUS, C.; RICHERT-PÖGGELER, K.R. Endogenous pararetroviruses: two-faced travelers in the plant genome. **Trends Plant Sci.**v. 11, p. 485-91, 2006.

STAMATAKIS, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatic**, v. 21, p. 2688-90, 2006.

SEAL, S.; MULLER, E. Molecular analysis of a full-length sequence of a new yam *badnavirus* from *Dioscorea salsibarensis*. **Archives of Virology**, v.152, p. 829-832, 2007.

TAMURA, K. et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v. 24, p. 1596-1599, 2007.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res.** v. 22, p. 4673–4680, 1994.

THOUVENEL, J. C.; FAUQUET, C. Yam mosaic, a potyvirus infecting *Dioscorea cayenensis* in the Ivory Coast. **Annals of Applied Biology**, v.93, p. 4, 1979.

TORRES-PACHECO, I. et al. Detection and distribution of geminiviruses in Mexico and the Southern United States. **Phytopathology**, v. 86, p. 1186-1192, 1996.

YANG, Z.; BIELAWSKI, J.P. Statistical methods for detecting molecular adaptation. **Trends in Ecology and Evolution**, v.15, p. 496-503, 2000.

YANG, I.C. et al. Sequence diversity of South Pacific isolates of Taro bacilloform virus and the development of a PCR-based diagnostic test. **Archives of Virology**, v. 148, p. 1957-1968, 2003.

ZHOU, X. et al. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. **Journal of General Virology**, v. 78, p. 2101-2111, 1997.

WU, B. et al. Assessment of codivergence of Mastreviruses with their plant hosts. **BMC Evol Biol.** v. 8, p. 335, 2008

VAN DER WALT, E. et al. Experimental observations of rapid Maize streak virus evolution reveal a strand-specific nucleotide substitution bias. **Virology J.** v.5. p. 104, 2008.