

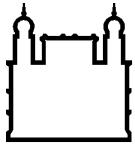
MINISTÉRIO DA SAÚDE  
FUNDAÇÃO OSWALDO CRUZ  
INSTITUTO OSWALDO CRUZ

Mestrado em Medicina Tropical

ANÁLISE DA DIVERSIDADE INTRA E INTER-HOSPEDEIRO DO  
DENV-2 EM AMOSTRAS DE PACIENTES COM DIFERENTES  
APRESENTAÇÕES CLÍNICAS

MARIA CELESTE TORRES

Rio de Janeiro  
Novembro de 2016



Ministério da Saúde

**FIOCRUZ**  
**Fundação Oswaldo Cruz**

**INSTITUTO OSWALDO CRUZ**  
**Programa de Pós-Graduação em Medicina Tropical**

***MARÍA CELESTE TORRES***

**ANÁLISE DA DIVERSIDADE INTRA E INTER-HOSPEDEIRO DO DENV-2 EM  
AMOSTRAS DE PACIENTES COM DIFERENTES APRESENTAÇÕES CLÍNICAS.**

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Medicina Tropical.

**Orientador (es):** Dra. Ana Maria Bispo de Filippis.  
Dr. Marcos César Lima de Mendonça.

**RIO DE JANEIRO**  
Novembro de 2016

Ficha catalográfica elaborada pela  
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

T693 Torres, Maria Celeste

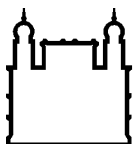
Análise da diversidade intra e inter-hospedeiro do DENV-2 em amostras de pacientes com diferentes apresentações clínicas / Maria Celeste Torres. – Rio de Janeiro, 2016.  
xviii, 128 f. : il. ; 30 cm.

Dissertação (Mestrado) – Instituto Oswaldo Cruz, Pós-Graduação em Medicina Tropical, 2016.

Bibliografia: f. 90-102

1. Dengue vírus sorotipo 2 (DENV-2). 2. Diversidade intra-hospedeiro. 3. Quasiespécies. 4. Quadro clínico. 5. Rio de Janeiro. I. Título.

CDD 616.91852



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

**INSTITUTO OSWALDO CRUZ**  
**Programa de Pós-Graduação em Medicina Tropical**

**AUTOR: MARIA CELESTE TORRES**

**ANÁLISE DA DIVERSIDADE INTRA E INTER-HOSPEDEIRO DO DENV-2 EM  
AMOSTRAS DE PACIENTES COM DIFERENTES APRESENTAÇÕES CLÍNICAS**

**ORIENTADOR (ES): Dra. Ana Maria Bispo de Filippis  
Dr. Marcos César Lima de Mendonça**

Aprovada em: \_\_\_\_/\_\_\_\_/\_\_\_\_

**EXAMINADORES:**

**Prof. Dr. Filipe Carvalho Costa - Presidente (IOC/Fiocruz)**

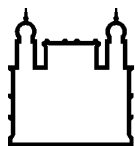
**Prof. Dra. Renata de Mendonça Campos (UFRJ)**

**Prof. Dr. Edson Oliveira Delatorre (IOC/Fiocruz)**

**Prof. Dr. Eduardo de Mello Volotão - Revisor (IOC/Fiocruz)**

**Prof. Dra. Gisela Freitas Trindade - Suplente (Biomanguinhos/Fiocruz)**

Rio de Janeiro, 10 de novembro de 2016.



Ministério da Saúde

**FIOCRUZ**

**Fundação Oswaldo Cruz**

**Anexar a cópia da Ata que será entregue pela SEAC já assinada.**

## **AGRADECIMENTOS**

À minha orientadora Dra Ana Bispo por ter me aberto as portas do Laboratório e ter me oferecido oportunidades impensadas no meu crescimento profissional. Pela paciência, pelas horas dedicadas e a atenção dispensada, pelo conhecimento compartilhado durante a realização deste trabalho, pela orientação em todos os momentos, pela compreensão, pelo seu apoio, e principalmente por acreditar em mim.

Ao meu co-orientador Dr Marcos Mendonça, pela sua imensa paciência, pelos seus milhões de desenhos na tentativa de que eu entendesse, por compartilhar o seu conhecimento, por me ensinar e ajudar a pensar sempre numa possível resposta aos meus questionamentos. Por me guiar sem se impor. Por ser ao meu critério, um excelente orientador.

À Dra. Rita Maria Ribeiro Nogueira, pelo apoio e confiança que tornaram possível a realização dessa dissertação.

À coordenadora da Pós-graduação em Medicina Tropical, Dra. Martha Suarez Muttis pelo apoio constante durante o desenvolvimento deste trabalho.

À unidade de genética das doenças infecciosas do Instituto Pasteur da França, ao seu diretor o Dr Anavaj Sakuntabhai e a sua equipe, por me receber no seu laboratório e me permitir uma oportunidade maravilhosa, por me ajudar no desenvolvimento dos meus experimentos e pelo apoio financeiro ao. Ao Dr Etienne Simon-Lorriere por me orientar e me ensinar sobre cada metodologia empregada nos experimentos, pela sua paciência, pelo o seu conhecimento transmitido e por me salvar em cada momento de desespero “bioinformático”.

Ao Dr. Eduardo de Mello Volotão por aceitar ser o revisor deste trabalho.

Ao Dr. Pedro Cabello, Dr. Leonardo Soares Bastos e Dra Claudia Codeço pela disponibilidade e apoio na aplicação da estatística neste trabalho.

Aos membros da banca examinadora, Dr. Filipe Carvalho Costa, Dra. Renata de Mendonça Campos, Dr. Edson Oliveira Delatorre, Dr. Eduardo de Melo Volotão e Dra. Gisela Freitas Trindade por aceitarem participar da avaliação deste trabalho.

Às minhas colegas que viraram amigas e equipe de todos os dias no laboratório, Aline Silva, Cintia Damasceno e Flávia Levy, por me lembrar todo dia o bom que é trabalhar em equipe, por sermos cúmplices, pelo apoio incondicional, pela ajuda e pela força em cada momento, pelo empenho em me ensinar o português e não deixar que o meu “portunhol” seja um impedimento para nada, por cada risada que aliviou os momentos de estresse, e por cada momento compartilhado, dentro e fora da bancada. Obrigada meninas por me fazer sentir como em casa e por essa amizade que ultrapassou o mestrado e que sei que levarei para o resto da vida. Tenho a certeza de que esta instância não teria sido o mesmo sem vocês!

À toda equipe do Laboratório de Flavivírus, pela colaboração, paciência e pronta resposta em todos os momentos que precisei de ajuda.

À minha família e amigos, por seu apoio incondicional em cada decisão que tomei na minha vida, por estarem sempre presentes apesar da distância, por me darem força para perseguir os meus sonhos, pelo carinho e compreensão e por sempre torcerem por mim.

Ao meu namorado e amigo Juan, por ter me acompanhado nesta viagem, por ter me feito rir até nos momentos de máximo desespero, por lutar comigo, por toda paciência, por resistir a cada dor de cabeça e cada uma das minhas loucuras, pelo amor, por ser a minha equipe!

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro.

E a todas as pessoas que contribuíram direta e indiretamente para a realização deste trabalho.

**“Nada em biologia faz sentido exceto à luz da evolução” Theodosius Dobzhansky 1973.**



## RESUMO

### ANÁLISE DA DIVERSIDADE INTRA E INTER-HOSPEDEIRO DO DENV-2 EM AMOSTRAS DE PACIENTES COM DIFERENTES APRESENTAÇÕES CLÍNICAS

A dengue atualmente é a arbovirose de maior impacto no mundo em termos de morbidade e mortalidade. Casos de dengue com confirmação laboratorial são reportados no Brasil desde a década de 80. Atualmente a doença é hiperendêmica com quatro sorotipos de vírus dengue (DENV1-4) circulando no país e associados a quadros que variam de dengue clássico a dengue grave, representando um importante problema de saúde pública. Cada sorotipo é subdividido em genótipos e linhagens, com virulência variável. Os DENV são vírus RNA e como tal, propensos a erros durante a replicação, comportando-se como quasiespecies dentro de cada hospedeiro. O genótipo asiático/americano do DENV-2 tem circulado no estado do Rio de Janeiro desde 1990, e no ano de 2008 um surto de DENV-2 foi associado com o aumento da gravidade, alta taxa de letalidade e concomitante mudança na faixa etária mais afetada. Os fatores virais envolvidos na patogênese grave permanecem ainda pouco caracterizados. Dentro desse contexto, analisamos a diversidade genética intra-hospedeiro do DENV-2 a fim de determinar a possível associação com a gravidade da doença. Foram estudadas 29 amostras de soros DENV-2 positivos do período 1999-2011, oriundos de pacientes do estado do Rio de Janeiro, clinicamente classificados como dengue clássico (48%), dengue com sinais de alarme (33%) e dengue grave (19%). O tipo de infecção também foi classificada nos pacientes em primária (52%) e secundária (48%). Genomas virais completos foram sequenciados empregando uma abordagem livre de amplicon, na plataforma NextSeq500 da Illumina. Variações de um único nucleotídeo (SNV) e as suas frequências foram detectadas entre todos os “reads” sequenciados para cada amostra. Todas as amostras analisadas corresponderam ao genótipo asiático/americano que tem circulado no Brasil desde 1990. No entanto, confirmamos a existência de duas linhagens para este genótipo no estado do Rio de Janeiro. A linhagem I, do período 1999-2001, apresentou neste estudo uma frequência menor de SNV não sinônimas e um perfil distinto de variação gênica, em comparação com a linhagem II, identificada a partir de 2007. Por outro lado, os casos das infecções secundárias apresentaram uma maior carga viral (CV) (Mediana=5,43x10<sup>3</sup>), número de SNV ( $\bar{x}$ =129) e índices de diversidade nucleotídica, quando comparados com os das infecções primárias (Mediana CV=2,33x10<sup>2</sup>; média do número de SNV=34). Os casos de dengue grave se apresentaram nesta amostragem com uma menor diversidade viral intra-hospedeiro e um padrão de SNV não sinônimas diferente do padrão achado para as amostras de dengue clássico com e sem sinais de alarme. Isto poderia sugerir que as SNV não sinônimas achadas nos casos clássicos com ou sem sinais de alarme não estariam envolvidas na patogênese grave. A função que cumpre cada uma destas mutações minoritárias na evolução da doença requer estudos mais detalhados. Os resultados gerados neste trabalho contribuem ao entendimento da dinâmica viral do DENV-2, e reforçam a importância de estudos complementares que permitam elucidar definitivamente a relação entre as subpopulações virais com a patogenia e apresentação clínica da infecção.

## ABSTRACT

### INTRA AND INTER-HOST DIVERSITY ANALYSIS OF DENV-2 IN SAMPLES OF PATIENTS WITH DIFFERENT CLINICAL OUTCOMES

Dengue fever is currently the arboviral disease of major impact in the world in terms of morbidity and mortality. Dengue cases with laboratory confirmation are reported in Brazil since the 80s, and nowadays the disease is hyperendemic with four serotypes of dengue virus (DENV1-4) circulating in the country. They are associated with clinical outcomes ranging from dengue fever to severe dengue, representing a major public health problem for the country. Each serotype is subdivided into genotypes and lineages of variable virulence. DENV are RNA viruses that behave as quasispecies within each host. DENV-2 Asian/American genotype has been circulating in the state of Rio de Janeiro since 1990. In 2008, a DENV-2 outbreak was associated with increased severity, high rate of mortality and concomitant change in the affected age group. Viral factors involved in severe pathogenesis remain poorly characterized. In this context, we analyzed the genetic intra-host diversity of DENV-2, in order to determine the possible association with disease severity. Thus, we studied 29 DENV-2 positive serum samples from the 1999-2011 period, obtained from patients of the state of Rio de Janeiro, clinically classified as dengue fever (48%), dengue with warning signs (33%) and severe dengue (19 %). The type of infection has also been identified in patients with primary (52%) and secondary (48%). Full-length viral genomes were deep sequenced using an amplicon-free approach, on the Illumina NextSeq500 sequencing system. Single nucleotide variants (SNV) and their frequencies among all sequencing reads were determined for each sample. All analyzed samples corresponded to the Asian/American genotype that has been circulating in Brazil since its introduction in 1990. We confirmed the circulation of two lineages for this genotype in the state of Rio de Janeiro. Linage I, of the period 1999-2001, showed a lower frequency of non-synonymous SNV and distinct profile of genetic variation, in comparison to lineage II identified since 2007. On the other hand, higher viral load (VL) (Median=5,43x10<sup>3</sup>), number of SNV ( $\bar{x}$ =129), and diversity indexes values were detected in cases of secondary infections compared to primary infection cases (VL Median=2,33x10<sup>2</sup>, SNV  $\bar{x}$ =34). In this cohort, cases with severe dengue presented lower diversity indexes than classic dengue cases with or without warning signs, as well as a different non-synonymous SNV pattern comparing to those last ones. This may suggest that non-synonymous SNV found in severe cases could be involved in severe pathogenesis. The function that each of these minority mutations may accomplish in the evolution of the disease requires more detailed studies. Results generated in this study contribute to the understanding of the viral dynamics of DENV-2, and emphasize the importance of additional studies to definitely elucidate the relationship between the subpopulations with viral pathogenesis and clinical presentation of the infection.

## ÍNDICE

	<b>Agradecimentos</b> .....	<b>v</b>
	<b>RESUMO</b> .....	<b>viii</b>
	<b>ABSTRACT</b> .....	<b>ix</b>
	<b>Índice de figuras</b> .....	<b>xiii</b>
	<b>Lista de Tabelas</b> .....	<b>xv</b>
	<b>Lista de Siglas e Abreviaturas</b> .....	<b>xvi</b>
<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>1</b>
	<b>1.1 Breve descrição da doença</b> .....	<b>1</b>
	<b>1.2 Breve histórico</b> .....	<b>1</b>
	<b>1.3 Dengue vírus</b> .....	<b>2</b>
	1.3.1 Estrutura do vírus .....	<b>3</b>
	1.3.2 Estrutura gênica.....	<b>3</b>
	1.3.3 Replicação dos DENV .....	<b>5</b>
	<b>1.4 Diversidade dos DENV</b> .....	<b>7</b>
	1.4.1 Subpopulações de DENV .....	<b>9</b>
	<b>1.5 Transmissão</b> .....	<b>12</b>
	<b>1.6 Manifestações clínicas e patogênese da doença</b> .....	<b>12</b>
	1.6.1 Patogênese.....	<b>14</b>
	<b>1.7 Epidemiologia da doença</b> .....	<b>15</b>
	1.7.1 No mundo .....	<b>15</b>
	1.7.2 No Brasil .....	<b>16</b>
	1.7.3 No estado do Rio de Janeiro .....	<b>17</b>
	<b>1.8 Diagnóstico laboratorial das infecções por DENV</b> .....	<b>18</b>
	<b>1.9 Prevenção e controle</b> .....	<b>20</b>
	<b>1.10 Justificativa</b> .....	<b>22</b>
<b>2</b>	<b>OBJETIVOS</b> .....	<b>24</b>
	<b>2.1 Objetivo Geral</b> .....	<b>24</b>
	<b>2.2 Objetivos Específicos</b> .....	<b>24</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	<b>25</b>
	<b>3.1 Desenho do estudo</b> .....	<b>25</b>
	<b>3.2 Considerações éticas</b> .....	<b>25</b>

<b>3.3</b>	<b>Amostragem.....</b>	<b>25</b>
<b>3.4</b>	<b>Classificação das amostras quanto à gravidade do quadro clínico ..</b>	<b>26</b>
<b>3.5</b>	<b>Classificação das amostras segundo o tipo de resposta imune (primária ou secundária).....</b>	<b>26</b>
3.5.1	Elisa para a determinação do título de anticorpos de tipo IgG anti-DENV.....	27
<b>3.6</b>	<b>Extração do RNA viral.....</b>	<b>28</b>
<b>3.7</b>	<b>Determinação da carga viral.....</b>	<b>28</b>
3.7.1	Construção da curva de calibração .....	28
3.7.2	RT-PCR em tempo real .....	28
<b>3.8</b>	<b>Sequenciamento profundo .....</b>	<b>29</b>
3.8.1	Extração do RNA .....	30
3.8.2	Tratamento com DNase.....	30
3.8.3	Tratamento com RNase H .....	30
3.8.4	Síntese do cDNA .....	32
3.8.5	PCR quantitativa em tempo real (qPCR) .....	34
3.8.6	Preparo das bibliotecas .....	35
3.8.7	Análises das bibliotecas geradas: Bioanalyzer 2100 .....	36
3.8.8	Construção dos pools das amostras selecionadas.....	37
3.8.9	Acondicionamento das bibliotecas.....	37
3.8.10	Sequenciamento profundo.....	37
<b>3.9</b>	<b>Análise bioinformática .....</b>	<b>38</b>
<b>3.10</b>	<b>Análise da diversidade genética intra- e inter-hospedeiro .....</b>	<b>38</b>
<b>3.11</b>	<b>Análise filogenética.....</b>	<b>39</b>
<b>3.12</b>	<b>Análise estatística .....</b>	<b>40</b>
<b>4</b>	<b>RESULTADOS .....</b>	<b>41</b>
<b>4.1</b>	<b>Amostragem.....</b>	<b>41</b>
<b>4.2</b>	<b>Aspectos técnicos.....</b>	<b>45</b>
4.2.1	Tratamento com DNase e RNase H, e posterior síntese do cDNA .....	45
4.2.2	Construção dos pools das bibliotecas de sequenciamento profundo.....	47
4.2.3	Sequenciamento profundo.....	47
<b>4.3</b>	<b>Análise filogenética.....</b>	<b>49</b>

<b>4.4</b>	<b>Análise da diversidade viral .....</b>	<b>54</b>
4.4.1	Índices de complexidade da população viral e da diversidade nucleotídica .....	54
4.4.2	Espectro de mutantes .....	58
4.4.3	Mutações fixas no genoma viral .....	69
<b>5</b>	<b>DISCUSSÃO .....</b>	<b>73</b>
5.1	Amostragem.....	73
5.2	Aspectos técnicos.....	75
5.3	Análise filogenética.....	76
5.4	Análise da diversidade viral .....	76
<b>6</b>	<b>CONCLUSÕES .....</b>	<b>83</b>
<b>7</b>	<b>PERSPECTIVAS .....</b>	<b>85</b>
<b>8</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>87</b>
<b>9</b>	<b>APÊNDICES E/OU ANEXOS .....</b>	<b>100</b>

**APÊNDICE I** - Fluxograma das análises bioinformáticas.

**APÊNDICE II** - SNV reportadas para cada amostra.

**APÊNDICE III** - Mutações fixas no genoma consenso de cada amostra.

## ÍNDICE DE FIGURAS

<b>Figura 1.1</b> Estrutura do dengue vírus.....	3
<b>Figura 1.2</b> Estrutura genética do DENV.....	4
<b>Figura 1.3</b> Ciclo de replicação do DENV.....	6
<b>Figura 1.4</b> Representação esquemática das quasiespécies virais em um hospedeiro infectado.....	10
<b>Figura 1.5</b> Classificação clínica da dengue segundo a OMS.....	13
<b>Figura 1.6</b> Modelo da teoria das infecções sequenciais por DENV.....	14
<b>Figura 1.7</b> Distribuição da dengue no mundo.....	16
<b>Figura 1.8</b> Métodos empregados no diagnóstico das infecções por DENV.....	19
<b>Figura 3.1</b> Processo de fragmentação do cDNA e marcação com sequências índices.....	36
<b>Figura 4.1</b> Localização geográfica e temporal dos casos incluídos neste estudo.....	43
<b>Figura 4.2</b> Variação da carga viral (cópias/ml) com o acondicionamento das amostras para o sequenciamento profundo.....	46
<b>Figura 4.3</b> Relação entre a porcentagem dos <i>reads</i> específicos para DENV-2 e a profundidade média atingida no sequenciamento profundo.....	49
<b>Figura 4.4(a-d)</b> Árvore filogenética de máxima verossimilhança obtida para o genoma completo de DENV-2.....	51
<b>Figura 4.5(a-f)</b> Número de variações nucleotídicas e medidas de diversidade viral segundo o tipo de resposta imune e a gravidade do quadro clínico.....	57
<b>Figura 4.6</b> Frequência das SNV achadas para os casos pertencentes à Linhagem I ou II.....	58
<b>Figura 4.7</b> Frequência do tipo de SNV com respeito à região do genoma viral.....	59
<b>Figura 4.8</b> Frequência percentual das SNV localizadas no marco de leitura da poliproteína para cada amostra.....	60
<b>Figura 4.9</b> Frequência absoluta das mutações não sinônimas segundo a sua posição no genoma.....	61
<b>Figura 4.10</b> Frequência relativa das mutações não sinônimas ao comprimento dos genes.....	62
<b>Figura 4.11(a-b)</b> Distribuição das substituições não sinônimas ao longo do genoma segundo a classificação do quadro clínico.....	62

<b>Figura 4.12</b> Frequência relativa das SNV não sinônimas para cada gene com respeito à linhagem do genótipo asiático/americano.....	64
<b>Figura 4.13</b> Frequência relativa das variantes compartilhadas segundo a sua localização no genoma e à classificação clínica dos casos.....	65
<b>Figura 4.14</b> Percentagens de variantes compartilhadas nas amostras segundo a sua classificação clínica.....	67
<b>Figura 4.15</b> Frequência relativa das mutações fixas achadas nos genomas virais com respeito à classificação clínica e à resposta imune dos casos.....	71

## LISTA DE TABELAS

<b>Tabela 1.1</b> Função de cada um dos componentes do genoma viral.....	4
<b>Tabela 1.2</b> Genótipos disponíveis para cada sorotipo de DENV.....	7
<b>Tabela 1.3</b> Vacinas candidatas contra a dengue.....	21
<b>Tabela 3.1</b> Critérios para a classificação da resposta imune.....	26
<b>Tabela 3.2</b> Condições de desnaturação do RNA e hibridização de sondas específicas para o RNA ribossomal e RNA carrier.....	31
<b>Tabela 3.3</b> Condições de ação para a RNase H.....	31
<b>Tabela 3.4</b> Condições do tratamento com DNase I para remoção das sondas de DNA.....	32
<b>Tabela 3.5</b> Condições de síntese da primeira fita simples de cDNA.....	33
<b>Tabela 3.6</b> Condições de síntese da fita complementar do cDNA fita simples.....	33
<b>Tabela 3.7</b> Condições do ensaio <i>in-house</i> PCR em tempo real.....	34
<b>Tabela 4.1</b> Características das amostras selecionadas.....	42
<b>Tabela 4.2</b> Resultado do PCR em tempo real das amostras processadas, pré e post-tratamento com DNase, RNase H e síntese do cDNA.....	45
<b>Tabela 4.3</b> Número de <i>reads</i> e cobertura do genoma das amostras sequenciadas no sistema NextSeq 500.....	48
<b>Tabela 4.4</b> Número de variantes e índices de complexidade e diversidade nucleotídica calculados para cada amostra.....	55
<b>Tabela 4.5</b> Mutações de elevada frequência inter-hospedeiro.....	69
<b>Tabela 4.6</b> Mutações fixas nos genomas virais achadas como SNV.....	72



## LISTA DE SIGLAS E ABREVIATURAS

- ABTS – 2,2'-azino-bis (ácido 3-etilbenzotiazolina-6-sulfônico).
- AL – acrilamida linear.
- AL – dengue clássico com sinais de alarme.
- ANOVA – análise da variância.
- BR – Brasil.
- C – proteína do capsídeo.
- Cc M – concentração molar.
- cDNA – ácido desoxirribonucleico complementar, do inglês *complementar deoxyribonucleic acid*.
- CEP – comitê de ética em pesquisa.
- CL – dengue clássico.
- D.C – depois de Cristo.
- DENV – vírus dengue.
- DENV-1, DENV-2, DENV-3, DENV-4 – vírus dengue sorotipo 1, 2, 3 e 4.
- DF – febre da dengue.
- DH – dengue hemorrágico.
- DNA – ácido desoxirribonucleico, do inglês *deoxyribonucleic acid*.
- dNTP – desoxirribonucleotideo trifosfato, do inglês *deoxynucleotide triphosphate*.
- DTT – ditioneitol.
- EDTA – ácido etilenodiaminotetracético, do inglês *Ethylenediamine tetraacetic acid*.
- ELISA – ensaio de imunoabsorção enzimática, do inglês *Enzyme-Linked ImmunoSorbent Assay*.
- ENV – proteína do envelope.
- ES – Estado do Espírito Santo.
- FHD – febre hemorrágica por dengue.
- GR – dengue grave.
- HCl – ácido clorídrico.
- HCV – vírus da hepatite C.
- HIV – vírus da imunodeficiência humana.
- H<sub>2</sub>O – água.
- IgG – imunoglobulina G.
- IgM – imunoglobulina M.
- M – glicoproteína de membrana.

MgCl<sub>2</sub> – cloreto de magnésio.

ML – máxima verossimilhança.

mM – concentração milimolar.

MS – Ministério de Saúde.

NaCl – cloreto de sódio.

NGS – sequenciamento de nova geração, do inglês *next generation sequencing*.

nM – concentração nanomolar.

NS – não estrutural.

PBS – tampão fosfato salino, do inglês *phosphate buffered saline*.

PCR – reação em cadeia da polimerase, do inglês *polymerase chain reaction*.

PDK – células de rim de cão, do inglês *primary dog kidney*.

PFU – unidade formadora de placa, do inglês *plaque forming unit*.

pH – coeficiente que indica o grau de acidez ou basicidade de uma solução aquosa.

Pi (π) – índice Pi da diversidade nucleotídica.

pM – concentração picomolar.

PRI – infecção primária.

prM – glicoproteína de membrana imatura.

PRNT – teste de neutralização de redução de placas, do inglês *Plaque reduction neutralization test*.

Q – score de qualidade.

qPCR – reação em cadeia da polimerase quantitativa, do inglês *quantitative polymerase chain reaction*.

RJ – Estado do Rio de Janeiro.

RNA – ácido ribonucleico, do inglês *ribonucleic acid*.

RT-PCR – transcrição reversa seguida da reação em cadeia da polimerase, do inglês *reverse transcription followed by polymerase chain reaction*.

SCD – síndrome de choque por dengue.

SD – desvio padrão, do inglês *standard deviation*.

SEC – infecção secundária.

Sn – índice de Shannon da complexidade da população viral.

SNS – substituição não sinônima.

SP – Estado de São Paulo.

SNV – substituição de nucleotídeo único, do inglês *single nucleotide variation*.

SS – substituição sinônima.

SVS – Secretaria de vigilância em saúde.

UTR – região não codificante de proteínas, do inglês *untranslated region*.

WHO – Organização Mundial da Saúde, do inglês *World Health Organization*.

# 1 INTRODUÇÃO

## 1.1 Breve descrição da doença

A Dengue é hoje a arbovirose de maior importância epidemiológica no mundo devido ao seu alto índice de morbimortalidade. A doença causada por um dos quatro sorotipos do vírus dengue (DENV), DENV-1, DENV-2, DENV-3 e DENV-4, é endêmica em mais de 100 países e nos últimos 50 anos, vem progressivamente alcançando o perfil de pandemia global, sendo estimado o número de infecções anuais em torno de 390 milhões, dos quais 96 milhões manifestam ao menos algum sinal de gravidade (Bhatt et al., 2013). O homem é o único hospedeiro capaz de desenvolver a doença, a qual é transmitida por mosquitos do gênero *Aedes*, tendo o *Aedes aegypti* como a principal espécie vetorial na região das Américas. (WHO, 2009).

## 1.2 Breve histórico

O termo “dengue” provavelmente originou-se na Espanha no início do século XIX, sendo um homônimo para a expressão de origem africana “Ki Denga Pepo”, ou “Denga”, cujo significado é “pancada ou golpe causado por um espírito mau que provocava um ataque doloroso”. Em 1828, a literatura médica inglesa adotou o termo durante uma epidemia, ocorrida no Caribe, de doença exantemática com artralgia (Halstead, 1980; Schatzmayr, 2008). Existem registros históricos da ocorrência de doença clinicamente compatível com a dengue na China publicados durante a Dinastia Chin, 265 – 420 D.C., sendo relatada formalmente nas Dinastias Tang, 610 D.C, e Sung do Norte, 992 D.C. A enfermidade foi chamada naquela época de “veneno da água”, já havendo a associação de insetos voadores com a água (Gubler, 1998; Weaver & Vasilakis, 2009). Surtos ocorridos nas Índias Francesas Ocidentais e no Panamá em 1635 e 1699, respectivamente, também foram relacionados à Dengue, porém os primeiros registros de grandes epidemias ocorreram nos anos de 1779 e 1780, atingindo os continentes da Ásia, África e América do Norte. Estes episódios importantes da dengue ocorreram, quase que simultaneamente em 1779, em Jakarta (Indonésia), e Egito, e em 1780 na Filadélfia (Gubler, 1998; Halstead, 1980). Em 1907, enquanto Ashburn e Craig desenvolviam

uma série de experimentos para investigar a febre da dengue (DF) nas Filipinas, encontraram um agente filtrável e infeccioso no sangue humano (Ashburn & Craig, 1907). A transmissão do DENV pelo mosquito *Aedes aegypti* foi primeiramente sugerida por Bancroft, em 1906, (Bancroft, 1906) e mais tarde confirmada por Cleland e colaboradores, em 1918, ao testarem, em voluntários humanos, a capacidade de transmissão do DENV pelos mosquitos *Culex fatigans* e *Ae. aegypti*, obtendo sucesso apenas com o último (Cleland et al., 1919).

O vírus da Dengue foi isolado em camundongos por Kimura e Hotta, em 1943, e por Sabin e Schelinger, em 1944, sendo neste caso, feito o isolamento das cepas Havaí e de Nova Guiné. Pesquisas realizadas por Sabin, durante a Segunda Guerra Mundial, forneceram a prova da existência de diferentes características antigênicas entre os DENV, sendo a cepa do Havaí caracterizada como sorotipo 1 e a cepa de Nova Guiné como sorotipo 2, as duas consideradas hoje protótipos (Kimura & Hotta, 1944; Sabin, 1952; Sabin & Schlesinger, 1945).

Os sorotipos 3 e 4 foram isolados a partir de pacientes com quadro grave de febre hemorrágica, durante epidemia ocorrida em Manila, nas Filipinas, em 1956, onde se registrou a co-circulação dos quatro sorotipos de Dengue. Foi sugerido que a co-circulação destes sorotipos proporcionou o surgimento de maior número de casos graves da doença (Hammond et al., 1960).

Provavelmente, os primeiros casos de infecção pelo vírus da Dengue surgiram na Ásia, onde os quatro sorotipos de DENV foram demonstrados em ciclos silvestres. Os acontecimentos da Segunda Guerra Mundial, nessa região, criaram condições favoráveis à expansão dos vetores da Dengue, tornando-a epidêmica e proporcionando o surgimento de formas clínicas graves da doença classificadas como: a Febre Hemorrágica por Dengue (FHD) e a Síndrome do Choque por Dengue (SCD) (Rudnick, 1986; Schatzmayr, 2008).

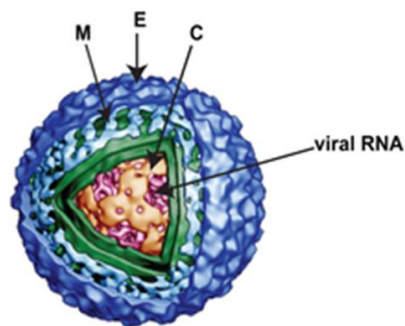
### **1.3 Vírus Dengue**

O vírus Dengue (DENV) pertence à família *Flaviviridae*, gênero *Flavivirus*. Epidemiologicamente são classificados como Arbovírus, pois são transmitidos por artrópodes, sendo neste caso o mosquito vetor do gênero *Aedes* (Gubler, 2002).

### 1.3.1 Estrutura do vírus

Trata-se de um vírus esférico, envelopado, com 40 a 50 nm de diâmetro. O capsídeo de simetria icosaédrica é composto por uma única proteína, a proteína do capsídeo, e é circundado por uma bicamada lipídica a qual se associam as proteínas virais da membrana e do envelope (figura 1.1) (Kuhn et al., 2002).

**Figura 1.1** Estrutura do vírus dengue

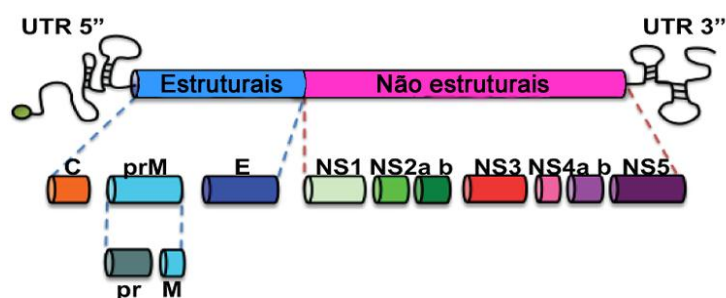


Fonte: imagem adaptada de Angel & Valle, 2013. M: glicoproteína da membrana; E: proteína do envelope; C: capsídeo.

### 1.3.2 Estrutura gênica

O seu genoma é composto por uma fita simples de RNA de polaridade positiva, de aproximadamente 11 kilobases, contendo uma única fase de leitura aberta (ORF, do inglês *open reading frame*), flanqueado por duas regiões não codificantes. A fase de leitura codifica para uma única poliproteína, a qual é clivada por proteases virais e celulares em múltiplos sítios, dando origem às proteínas estruturais (capsídeo, membrana e envelope) e às não estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5) (Figura 1.2) (Chambers et al., 1990).

**Figura 1.2** Estrutura gênica do DENV



Fonte: imagem adaptada de Angel & Valle, 2013. UTR: região não codificante, do inglês “untranslated region”; C: capsídeo; M: glicoproteína de membrana; prM: forma imatura da glicoproteína de membrana; E: envelope; NS: proteína não estrutural.

A posição de cada gene no genoma e das regiões 5' e 3' não codificantes, como também a suas funções podem observar-se a continuação na tabela 1.1:

**Tabela 1.1** Localização e função de cada um dos componentes do genoma viral

Região/Gene	Posição no genoma	Função
<b>5'UTR</b>	1-96	Atua como sítio de ligação e ativação da polimerase viral; contém uma sequência complementar a outra na região 3'UTR; região espaçadora atua promovendo a replicação viral (Lodeiro et al., 2009; Villordo et AL., 2015).
<b>Capsídeo</b>	97-396	Componente estrutural do nucleocapsídeo (Proteína C + RNA genômico) (Murphy, 1980).
<b>preM</b>	397-936	Forma imatura da proteína M (Heinz & Allison, 2001).
<b>Membrana</b>	439-936	Proteína transmembrana que interage com a proteína do envelope.
<b>Envelope</b>	937-2421	Proteína transmembranar. É a principal e maior proteína estrutural. Intervém na montagem da partícula viral; interação com receptores celulares e fusão de membranas; principal alvo para anticorpos neutralizantes; atividade hemaglutinante. (Heinz & Stiasny, 2012).
<b>NS1</b>	2422-3477	Proteína não estrutural envolvida na morfogênese da partícula viral; cofator no processo de replicação viral (Muylaert et al., 1997). Funções de evasão do sistema imune (Avirutnan et al., 2010).
<b>NS2A</b>	3478-4131	Subunidade do complexo de replicação, reconhecendo a região 3'UTR do genoma viral; promove a montagem e a replicação viral (Mackenzie et al.,

		1998).
<b>NS2B</b>	4132-4521	Quando ligada a NS3, exerce uma atividade proteolítica na biossíntese viral (Rice, 1996).
<b>NS3</b>	4522-6375	Atividade proteolítica e helicase. Reconhecimento da porção amino terminal da polimerase viral para a montagem do complexo de replicação (Rice, 1996; Luo et al., 2008).
<b>NS4A</b>	6376-6756	Subunidade do complexo de replicação. Auxilia proteínas não estruturais no reconhecimento do genoma viral (Lindenbach & Rice, 1999; Zou et al., 2015).
<b>Proteína 2K</b>	6757-6825	Porção carboxi-terminal da proteína NS4A. Atua como sinal para a translocação de NS4B no lumem do retículo endoplasmático (Zou et al., 2015).
<b>NS4B</b>	6826-7569	Componente essencial do complexo de replicação associado à membrana do retículo endoplasmático; função pouco conhecida (Zou et al., 2015).
<b>NS5</b>	7570-10269	RNA polimerase RNA dependente. Subunidade do complexo de replicação viral. Metiltransferase (Khromykh et al., 1999).
<b>3'UTR</b>	10270-10723	Contem estruturas essenciais para a replicação viral (Villordo et al., 2015).

### 1.3.3 Replicação dos DENV

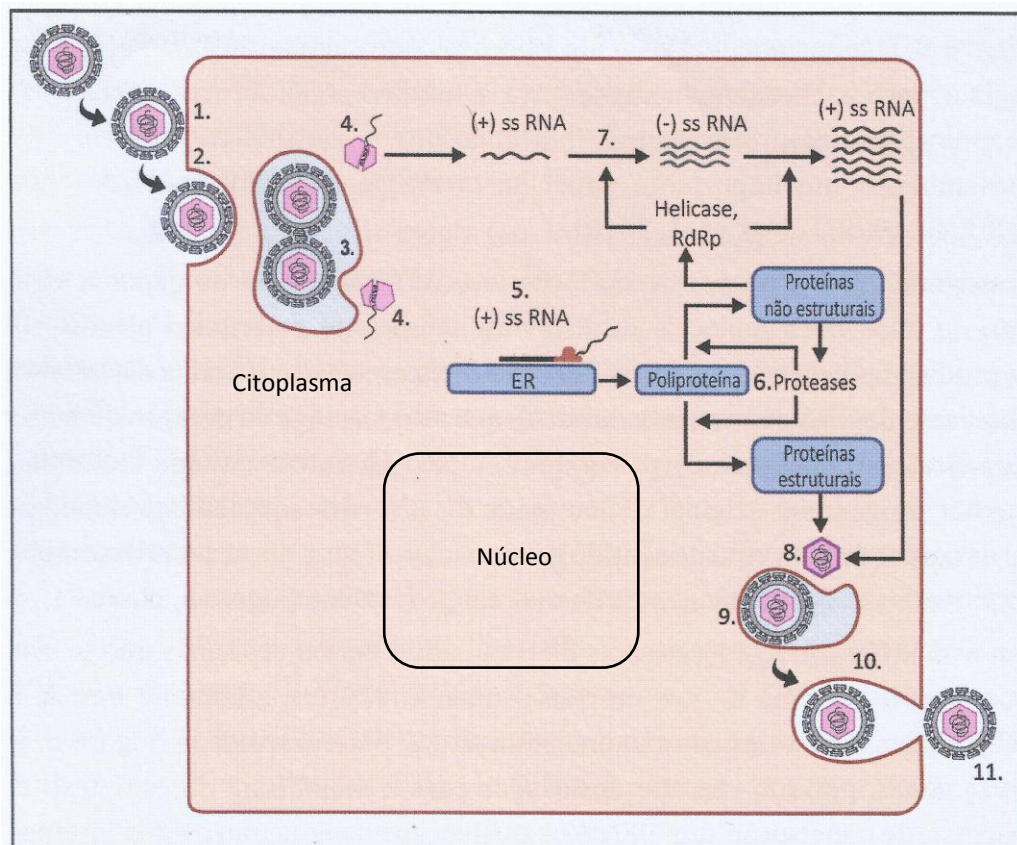
Após a transmissão do vírus que ocorre através do repasto sanguíneo pelo vetor, o vírus infecta inicialmente as células de Langerhans e células dendríticas presentes na pele (Wu et al., 2000). Posteriormente o vírus se replica no citoplasma de células musculares estriadas e/ou lisas, fibroblastos e linfonodos, e após um período de 2 a 7 dias surgem os primeiros sintomas que coincidem com o período de viremia (Kurane & Ennis, 1992).

O processo de interação vírus-célula tem início com a ligação do DENV a receptores presentes na superfície das células do hospedeiro, seguido da endocitose das partículas virais dependente de clatrina, proteínas celulares envolvidas na formação de vesículas membranares nas células eucariontes (Figura 1.3 passos 1 e 2) (Fagnoud et al., 2012). Após a endocitose, uma mudança conformacional dependente de pH permite o escape do RNA viral desde o endossoma ao citoplasma (Figura 1.3 passos 3 e 4). O genoma viral serve como RNA mensageiro, que após ser traduzido pelos ribossomos celulares, da origem a



uma poliproteína que será posteriormente clivada nas diferentes proteínas virais (Figura 1.3 passos 5 e 6). Em seguida, o RNA viral é replicado no retículo endoplasmático perinuclear, intermediado por um RNA de polaridade positiva que serve como molde para a replicação (Figura 1.3 passo 7). O RNA viral associa-se com as proteínas do capsídeo, o qual brota desde a membrana do retículo endoplasmático envolto por uma bicamada lipídica que contem as proteínas virais da membrana (prM/M) e do envelope (E), constituindo assim o envelope viral (Figura 1.3 passos 8 e 9). As partículas virais completas transitam pela via secretora celular e saem finalmente da célula pelo processo de exocitose (Figura 1.3 passos 10 e 11) (Welsch et al., 2009).

**Figura 1.3.3** Ciclo de replicação do DENV



Fonte: imagem disponível em Valle et al., 2015. (+) ssRNA: RNA simples fita polaridade positiva; (-) ssRNA: RNA simples fita polaridade negativa; ER: retículo endoplasmático; RdRp: polimerase viral dependente de RNA. 1: ligação do DENV a receptores presentes na superfície celular; 2: endocitose das partículas virais; 3-4: mudança conformacional dependente de pH e escape do RNA viral desde o endossoma ao citoplasma; 5: tradução do genoma viral; 6: clivagem da poliproteína

nas diferentes proteínas virais; 7: replicação do RNA; 8-9: ensamblagem da partícula viral; 10-11: exocitose das partículas virais maduras.

#### 1.4 Diversidade dos DENV

Com a utilização de análises moleculares e filogenéticas, DENV foi classificado em 4 sorotipos antigênicos e geneticamente diferentes, DENV 1-4 (Holmes & Twiddy, 2003, Vasilakis & Weaver, 2008). Por sua vez, cada um deles foi sub-classificado em diversos genótipos, dada a extensiva variabilidade dentro de cada grupo, tendo sido demonstrado por diversos estudos a existência de cinco genótipos para o DENV-1 (Rico-Hesse, 1990; Weaver & Vasilakis, 2009; Chen & Vasilakis, 2011), seis genótipos para o DENV-2 (Rico-Hesse et al., 1997; Weaver & Vasilakis, 2009; Chen & Vasilakis, 2011), cinco genótipos para DENV-3 (Lanciotti et al., 1994) e quatro genótipos para DENV-4 (Lanciotti et al., 1997; Villabona-Arenas et al., 2011; Chen & Vasilakis, 2011) (Tabela 1.4-1). Os quatro sorotipos compartilham, em termos de sequências de aminoácidos, identidades que variam entre 60-70% para a proteína do envelope, enquanto que dentro de cada sorotipo, as semelhanças atingem 90% (Weaver & Vasilakis, 2009; Pierson & Diamond, 2013). Atualmente, as análises de genotipagem e filogenéticas, baseiam-se no sequenciamento do gene do envelope, o mais amplamente utilizado, uma vez que neste gene estrutural existe uma grande probabilidade de achar alterações de nucleotídeos e de aminoácidos, pois está associado com a imunogenicidade do vírus e a resposta de anticorpos do hospedeiro.

**Tabela 1.4-1** Genótipos disponíveis para cada sorotipo viral

SOROTIPO	GENÓTIPOS	DISTRIBUIÇÃO GEOGRÁFICA
DENV-1	I	Sudeste Asiático, China, Leste da África
	II	Tailândia (1950-1960)
	III	Malásia (cepas selvagens)
	IV	Ilhas do Oeste do Pacífico e Austrália
	V	Américas, Oeste da África Africano, Ásia
DENV-2	Asiático I	Malásia e Tailândia
	Asiático II	Vietnã, da China, Taiwan, Sri Lanka e Filipinas

	Cosmopolita	Austrália, Leste e Oeste Africano, Ilhas dos oceanos Pacífico e Índico Subcontinente Indiano e Oriente Médio
	Americano	América Latina, Caribe (1950-1960)
	Asiático/Americano	Tailândia, Vietnã, Américas (últimos 20 anos) Subcontinente Indiano e Ilhas do Pacífico
	Selvagem	Oeste Africano e Sudeste Asiático (cepas isoladas em humanos, mosquitos silvestres ou macacos sentinelas)
DENV-3	I	Indonésia, Malásia, Filipinas e Sul da Ilhas do Pacífico
	II	Tailândia, Vietnã e Bangladesh
	III	Sri Lanka, Índia, África, Samoa, Tailândia (1962)
	IV	Porto Rico, Américas Latina e Central, Taiti (1965)
	V	Filipinas (1956), Japão (1973), China (1980) América do Sul (2002-2004)
DENV-4	I	Tailândia, Filipinas, Sri Lanka e Japão
	II	Indonésia, Malásia, Taiti, Caribe e América (provenientes do Sudeste Asiático)
	III	Tailândia (cepas recentes)
	Selvagem	Malásia (cepas selvagens)

No Brasil, até 2013, eram detectados somente seis genótipos circulantes: genótipo V de DENV-1 (dos Santos et al., 2011; Drumond et al., 2012), genótipo Asiático/Americano de DENV-2 (Oliveira et al. 2010; Romano et al., 2010; Drumond et al., 2013), genótipos II e III de DENV-3 (Araújo et al., 2009) e os genótipos I e II de DENV-4 (Nunes et al., 2012). Dentro dos diversos genótipos, tem sido descoberta a existência de linhagens, as quais apresentam relações geográficas e temporais diferentes, tanto nas Américas como na Ásia (Myat Thu et al., 2005; Carrillo-Valenzo et al., 2010; Mendez et al., 2010; Duong et al., 2013). Essa diversidade intra-genótipo para DENV é, no entanto, limitada, sendo modulada por dois processos principais, a resposta imune do hospedeiro e o “bottleneck<sup>1</sup>” na transmissão, tanto no hospedeiro vertebrado quanto no invertebrado (Grenfell et al., 2004). Análises de epidemiologia molecular do sorotipo DENV-2, alvo deste estudo, tem demonstrado

---

<sup>1</sup> Bottleneck: eventos estocásticos que limitam a variação genética em uma população (Li & Roossinck, 2004).

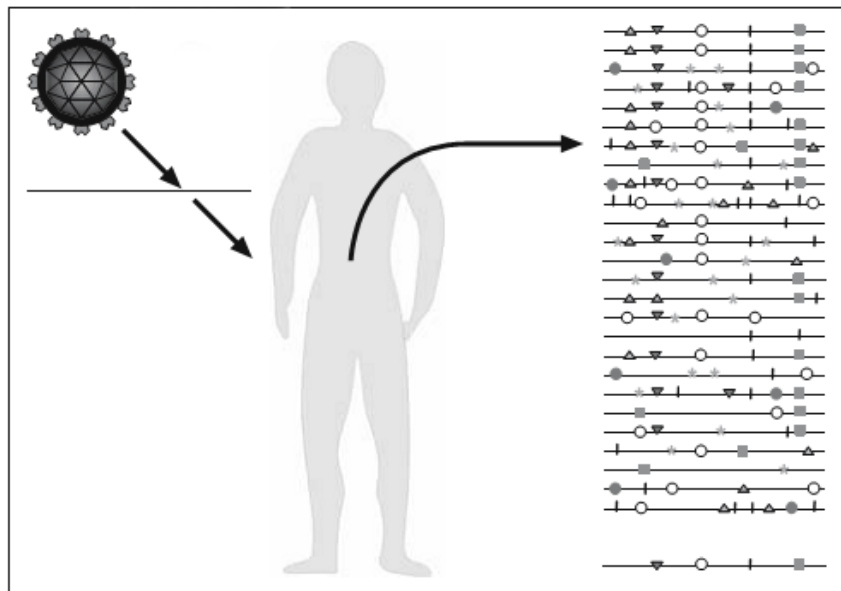
que desde 1990, o genótipo circulante no Brasil é o Asiático/Americano, o qual foi associado com aumentos na gravidade da doença. Na re-emergência deste sorotipo no ano 2007, com a concomitante mudança do padrão clínico-epidemiológico da dengue no Brasil, estava envolvida uma variante viral pertencente ao mesmo genótipo, mas agrupada em um grupo monofilético diferente às variantes até então circulantes. Desta maneira foram identificadas filogenética e epidemiologicamente duas linhagens dentro daquele genótipo: a linhagem I, das variantes virais que circulavam no Brasil desde 1990, e a linhagem II, do período da reemergência (Oliveira et al., 2010; Drumond et al., 2013). No entanto, a análise da região codificante do gene do envelope, das variantes representativas das diferentes manifestações clínicas, não apresentaram mutações no genoma viral que estivessem associadas a severidade da doença, com a exceção da presença de Asparagina (N) na posição E390, o qual representa um marcador genético de virulência (Pryor et al., 2001).

#### **1.4.1 Subpopulações de DENV**

As populações virais existem dentro de cada indivíduo como “pools” de membros relacionados, mas não idênticos, chamados quasiespecies. Por definição, as quasiespecies virais são distribuições dinâmicas de mutantes não idênticas, mas estreitamente relacionadas e genomas virais recombinantes, submetidos a um processo contínuo de variação genética, competição e seleção, porém, atuando como uma unidade de seleção (Domingo et al., 2006) (Figura 1.4.1-1). A geração de quasiespecies que ocorre após a infecção com um vírus RNA, está associada com replicações propensas à introdução de erros no genoma, altas taxas de mutação, eventos de recombinação e tempos de geração curtos na dinâmica evolutiva dos vírus, fatores que permitem o escape à pressão do sistema imune e logo, à persistência delas no hospedeiro. Essa alta variabilidade, geralmente acaba desenvolvendo rápida resistência às vacinas e às drogas antivirais (Lauring & Andino 2010, Poh et al. 2013). Por sua vez, são diversos os estudos que têm descrito que estes espectros de mutantes de genomas estreitamente relacionados, ou quasiespecies, contribui significativamente na patogênese viral, permitindo a modulação da expressão de características fenotípicas distintas por diferentes subpopulações virais. Este espectro pode incluir genomas de memória que reflitam a história evolutiva da linhagem viral, do mesmo jeito em que poderia participar na

extinção viral mediante mutagêneses letais. Além disso, o espectro destes mutantes são o alvo no qual a seleção natural e a deriva genética aleatória agem para moldar a evolução a longo prazo do vírus (Domingo et al., 2006).

**Figura 1.4.1-1** Representação esquemática das quasiespécies virais em um hospedeiro infectado



Fonte: imagem adaptada de Domingo et al., 2006. Os genomas virais são representados nas linhas horizontais, enquanto os símbolos nelas representam as mutações próprias de cada subpopulação.

Particularmente para o DENV, diversos estudos têm demonstrado a presença de distintas subpopulações coexistindo dentro de um mesmo hospedeiro (Lin et al., 2004; Descloux et al., 2009; Puiprom et al., 2011; Takeshi 2011; Chao 2012; Parameswaran et al., 2012; Thai et al., 2012; Takeshi et al., 2013; Romano et al., 2013; Lequime et al., 2015; Rodriguez-Roche et al., 2016). Alguns deles inclusive têm analisado a diversidade dessas quasiespecies em casos correspondentes a infecções primárias e secundárias. Neste contexto, Takeshi e colaboradores acharam uma maior diversidade nos casos das infecções primárias em comparação com os das infecções secundárias, sugerindo que a maior homogeneidade nas infecções secundárias poderia dever-se à presença de anticorpos neutralizantes, o

que provocaria a seleção da variante viral de maior *fitness*<sup>2</sup>, e possível maior virulência. Como consequência, quadros de dengue mais graves poderiam se desenvolver (Takeshi 2011; Takeshi et al., 2013). Rodriguez-Roche e colaboradores pelo contrário, observaram uma diversidade maior nos casos das infecções secundárias (Rodriguez-Roche et al., 2016). Outro estudo em Taiwan, que analisou clusters familiares onde ocorriam subsequentemente casos de dengue e dengue grave (casos na sua maioria primários), propôs que a transmissão mecânica pelo mosquito vetor, de subpopulações menores, porém mais virulentas e com maior *fitness*, possa estar favorecida e causando casos graves entre membros da família, posteriores aos casos de dengue clássico que ocorriam primeiro nessa mesma família (Chao, 2012).

De qualquer maneira, até o momento não foi possível esclarecer a associação entre o nível de diversidade genética do DENV intra-hospedeiro, com a gravidade da doença, refletida nos quadros clínicos.

Continua então como uma incógnita, se a distribuição das quasiespecies de DENV atua como um reservatório de variantes virais, cumprindo um papel essencial na diversificação e seleção, e contribuindo assim na evolução do DENV.

Atualmente, metodologias de sequenciamento de nova geração (NGS) estão sendo aplicadas para identificar e quantificar essas subpopulações virais geradas no hospedeiro, ou seja, no curso de uma única infecção, com elevada precisão e acurácia, sendo uma das ferramentas mais sensíveis no campo da genômica. Inúmeros estudos em diferentes modelos virais têm obtido resultados na análise da biodiversidade de quasiespecies dentro do ecossistema do hospedeiro com a aplicação destas tecnologias de sequenciamento profundo (deep sequencing) (Matranga et al. 2014; Yin et al. 2012; Romano et al., 2013; Grubaugh et al., 2015; Lequime et al., 2015). Essas técnicas de alto rendimento geram um número de sequências muito maior quando comparadas com a técnica tradicional de Sanger, permitindo a detecção de um maior número de alterações com uma cobertura por sua vez também maior e mais precisa (Pho et al. 2013).

---

<sup>2</sup> Fitness: parâmetro importante da genética adaptado por virologistas para quantificar a capacidade de replicação de um vírus e produzir uma progênie infecciosa. Geralmente é determinado em experimentos de competição em cultura de células ou in vivo, junto com um de vírus de referência (Domingo et al., 2012).

## 1.5 Transmissão

Fatores relativos ao ambiente e intrínsecos do mosquito, do DENV e do hospedeiro influenciam a dinâmica da transmissão dessa doença e o seu controle (Valle et al., 2015).

O DENV é mantido na natureza por um ciclo de transmissão que envolve aos hospedeiros vertebrados e aos mosquitos hematófagos do gênero *Aedes* (Gubler, 2002). Na África e na Ásia os DENV se mantêm circulando em ciclos urbanos, silvestres e rurais, enquanto nos outros continentes a circulação é basicamente urbana, envolvendo mosquitos com hábitos domésticos, como é o caso do *Aedes aegypti*. Em algum momento no passado, humanos ou mesmo primatas não humanos adquiriram a infecção por meio da exposição à picada de mosquitos silvestres infectados pelo DENV ao frequentar o ambiente selvagem. Estes hospedeiros, ao circularem virêmicos no ambiente modificado, permitiram a propagação do vírus por mosquitos domésticos (*Ae. aegypti* no caso da África e *Ae. Albopictus* na Ásia), tornando esse ciclo totalmente independente de reservatórios silvestres. A partir do estabelecimento desse ciclo estritamente doméstico e peridomiciliar em portos marítimos, o vírus se espalhou pelo mundo conjuntamente com o seu vetor *Ae. aegypti* (Vasilakis et al., 2011).

A infecção do mosquito vetor (fêmea) pelo DENV se inicia a partir da ingestão de partículas virais infecciosas presentes no sangue de um humano virêmico. Após o repasto sanguíneo, é necessário um período de incubação que varia entre 8 e 12 dias para que o vírus se replique no estômago do mosquito e invada as suas glândulas salivares. Estas partículas virais agora infectantes serão inoculadas em um novo hospedeiro durante o próximo repasto sanguíneo, determinando um novo ciclo (Salazar et al., 2007; Chan & Johansson, 2012).

Existe, ainda, uma transmissão do vírus da fêmea do mosquito infectada para a sua progênie, por via transovariana ou vertical, sem necessidade de outro hospedeiro interveniente (Anderson & Rico-Hesse, 2006).

## 1.6 Manifestações clínicas e patogênese da doença

O homem é o único hospedeiro capaz de desenvolver a doença, cuja apresentação clínica pode variar desde uma infecção assintomática, até a dengue com ou sem sinais de alarme, ou dengue grave, forma que se bem rara, pode atingir

a letalidade (Figura 1.6-1). Aproximadamente 90% dos casos, apresentam-se de forma autolimitada, com uma duração máxima de uma semana (WHO, 2009). A febre, acompanhada pelas dores generalizadas e o possível envolvimento gastrointestinal e aparição do exantema maculopalular, define a dengue clássica. Esta etapa, em termos gerais, costuma ser a única fase da doença, e está associada ao período de viremia. Durante esta etapa, não é possível saber se o paciente vai permanecer com sinais e sintomas da dengue clássica, autolimitando e evoluindo para a cura espontânea, ou se o quadro piorará causando uma dengue grave, com choque e até mesmo hemorragias maciças (manifestações decorrentes da perda de líquidos para o espaço extravascular, pelo aumento da permeabilidade vascular). Essa possível etapa crítica, é marcada geralmente pela queda da febre (Martinez, 2008). No entanto, só uma minoria desenvolve o quadro grave, que pode ser fatal. Existem também formas de apresentação classificadas como atípicas (neurológicas, hepáticas, cardíacas, entre outras), menos frequentes que o dengue clássico ou dengue grave, as quais têm sido relatadas em países do Sudeste Asiático e das Américas, inclusive no Brasil (Gulati & Maheshwari 2007, Valle et al. 2015). Porém, pouco se conhece sobre a incidência destas formas.

**Figura 1.6-1** Classificação clínica da dengue segundo a OMS



Fonte: Imagem adaptada da Organização mundial da saúde (WHO, 2009).

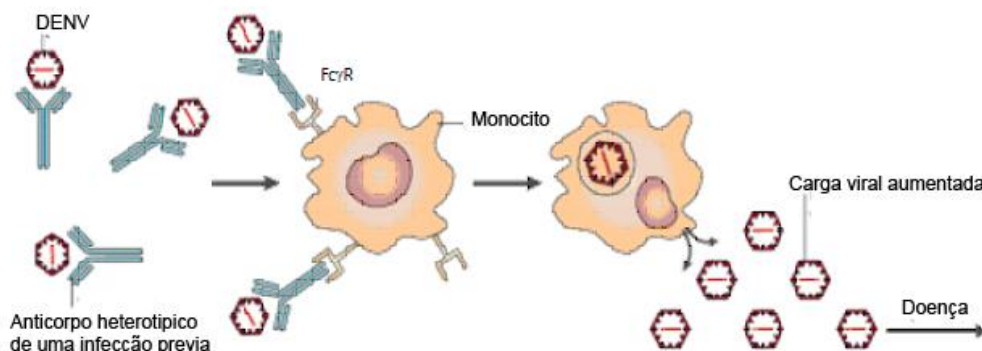


### 1.6.1 Patogênese

Diversas teorias são propostas para explicar o alto grau de variação das manifestações clínicas causadas pelos DENV, no entanto duas são as mais aceitas: a teoria das infecções sequenciais e a da virulência da cepa infectante.

A teoria da infecção sequencial ou da facilitação dependente de anticorpos (do inglês: “ADE - antibody dependent enhancement”), propõe que a resposta imune de um indivíduo sensibilizado é amplificada numa segunda infecção, decorrente da existência prévia de anticorpos heterotípicos e considera que há um aumento da replicação viral em macrófagos via estes anticorpos (acs) heterólogos. Em infecções secundárias com um vírus heterólogo ao da infecção anterior, os anticorpos reagem de forma cruzada com este novo sorotipo, mas não são capazes de neutralizá-los. Esses complexos, ao serem reconhecidos e internalizados por fagócitos mononucleares, resultariam na infecção celular e replicação viral. Essas células infectadas liberam na corrente sanguínea, mediadores vasoativos, aumentando a permeabilidade vascular, ativação do sistema complemento e da tromboplastina tissular, desencadeando os mecanismos responsáveis pelas manifestações clínicas das formas hemorrágicas (Hastead, 1988) (Figura 1.6.1-1).

**Figura 1.6.1-1** Modelo da teoria das infecções sequenciais por DENV



Fonte: imagem adaptada de Whitehead et al. 2007.

A virulência da cepa infectante, sugere que a gravidade da doença se deve às variações genéticas e antigênicas das diferentes cepas de vírus. Uma explicação admitida, é que a evolução genética do vírus dentro de cada sorotipo possa dar origem a cepas epidêmicas ou mais virulentas (Rosen, 1977; Rico-Hesse, 1990).

Apesar do conjunto complexo de fatores que contribuem para a epidemiologia da doença, diversos estudos sugerem que as estruturas virais específicas podem contribuir para o aumento da replicação em células alvo humanas e aumentar a transmissão pelo mosquito vetor (Rico-Hesse, 2003). Esta teoria poderia ser a responsável de explicar a gravidade clínica nas infecções primárias. Sendo assim, nosso estudo tomará como base, a teoria da virulência da cepa viral, considerando que a evolução do DENV resulte da seleção de cepas com maior virulência que terão impacto direto nos seres humanos.

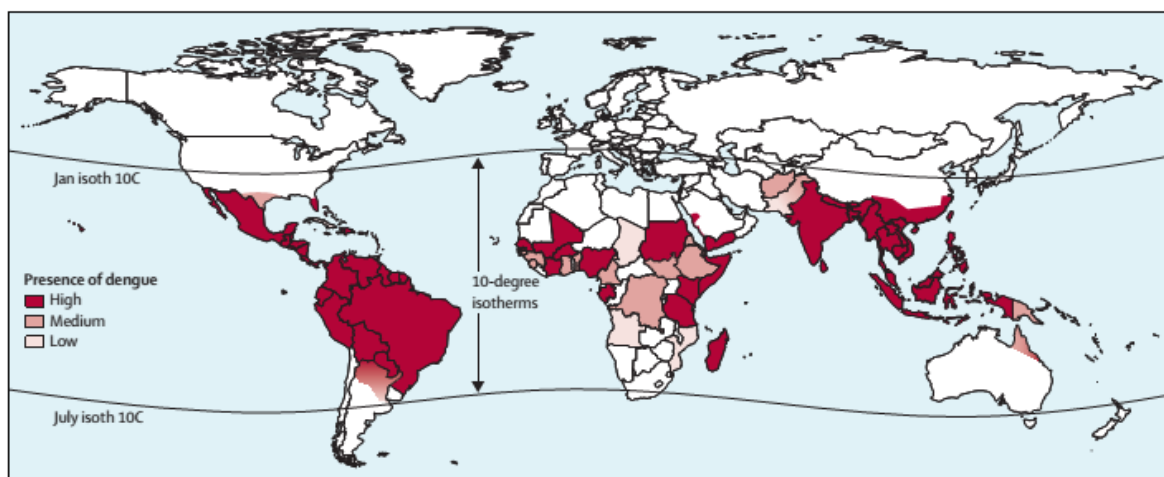
## **1.7 Epidemiologia da doença**

### **1.7.1 No mundo**

Essa arbovirose considerada benigna e sem grandes repercussões, começou a se apresentar já na década de 1950 com epidemias de febre hemorrágica no Sudeste Asiático. Nesse momento, eram só 9 os países envolvidos, e quase 1000 os casos de febre da dengue e febre hemorrágica da dengue reportados à Organização Mundial (OMS) (Gubler, 1997; WHO, 2013). Hoje, já são mais de 70 países que reportam casos, e aproximadamente 100 os que se encontram infestados pelo vetor *Aedes aegypti*, e onde a doença adquire um caráter endêmico (Figura 1.7.1-1). Nos últimos 50 anos, a sua incidência tem aumentado pelo menos 30 vezes com uma crescente expansão geográfica para novos países, e já na década atual, desde áreas urbanas a rurais (WHO, 2012). A OMS estima que sejam cerca de 50-100 milhões de infecções, com 500.000 casos de febre hemorrágica por dengue no mundo a cada ano. Porém, estudos mais recentes, indicam uma notória sub-notificação de casos de dengue com manifestações clínicas, estimando que o número real de casos possa atingir quase os 400 milhões de infecções por ano ocorrendo na América, Ásia, África e Oceania (WHO, 2009; Guzman et al., 2010; Bhatt et al., 2013). Porém, o cenário real da dengue não é tão bem conhecido para

países como Índia, Indonésia, China e da África segundo estudos realizados por Bhatt e colaboradores (Bhatt et al., 2013).

**Figura 1.7.1-1** Distribuição da dengue no mundo



Fonte: Guzman & Harris, 2015.

Nas Américas, o vírus dengue começou a recircular no início dos anos 1960, após um período de silêncio epidemiológico, decorrente possivelmente da eliminação do mosquito vetor em vários países do continente (Pinheiro, 1996). Em 1963, o sorotipo DENV-3 foi isolado na Jamaica de onde começou a se disseminar para as ilhas do Caribe, Venezuela e Colômbia. Nos anos 1968 a 1970, os sorotipos DENV-2 e 3 provocaram epidemias em países da América Central e do Sul, e continuaram circulando pela região durante toda a década de 70, sendo que no final desse período e começo da década de 80, os sorotipos DENV-1 e 4 também foram introduzidos. Desde então, o cenário da dengue foi se agravando, apresentando-se com epidemias consecutivas, atingindo principalmente grandes centros urbanos, e com uma marcada tendência ao aumento da gravidade dos casos, aumentando também a proporção de casos de febre hemorrágica por dengue (San Martin et al., 2010).

### **1.7.2 No Brasil**

Embora os primeiros relatos de dengue no Brasil datem de 1846, com surtos ocorridos simultaneamente nos estados do Rio de Janeiro, Bahia, Pernambuco e em

localidades do norte do país (Mariano, 1917), somente a partir de 1986, a dengue se tornou um problema de Saúde Pública nacional, com a identificação do primeiro caso do DENV-1 no município de Nova Iguaçu, RJ, (Schatzmayr et al., 1986), alcançando um elevado número de notificações em vários estados do Brasil (Miagostovich et al., 1993; Nogueira et al., 1999; Figueiredo, 2000). Em 1990, um novo sorotipo, o DENV-2 foi isolado na cidade de Niterói (Nogueira et al., 1990). A co-circulação de dois sorotipos e a natureza do genótipo de origem asiática de DENV-2 que foi introduzido nas Américas e posteriormente no Brasil, resultou no agravamento do quadro clínico e a notificação dos primeiros casos de dengue hemorrágico (DH), síndrome de choque por dengue (SCD) e óbitos no país. Em dezembro de 2000 mais um sorotipo, o DENV-3, foi detectado no município de Nova Iguaçu (Nogueira et al., 2001), sendo responsável até aquele momento pela maior e mais grave epidemia de dengue já descrita no país e no continente americano, não apenas pelo elevado número de notificações (794.200 casos), assim como pela ocorrência de casos graves e fatais (Araújo et al., 2009). Esse agravamento da expressão clínica da doença, se intensificou ainda mais com o recrudescimento do sorotipo DENV-2 no ano 2007. Foi também característico desse período a mudança da faixa etária acometida, sendo as crianças as mais envolvidas (Cavalcanti et al., 2011). Em 2010, o DENV-4 reemergiu em Roraima e a partir daí, ocorreu a disseminação desse sorotipo para o resto do país (Temporão et al., 2011; Brasil, 2014).

### **1.7.3 No estado do Rio de Janeiro**

Em março de 1986 o sorotipo DENV-1 foi identificado no soro de pacientes residentes da cidade de Nova Iguaçu, Rio de Janeiro. Em maio desse mesmo ano a epidemia se disseminou para outras cidades do estado, incluindo a capital (Schatzmayr et al., 1986). Em abril de 1990 foi isolado o sorotipo DENV-2 em pacientes da cidade de Niterói, novamente na região metropolitana do estado, dando lugar à epidemia dos anos 1990-1991, na qual começaram a observar-se formas mais graves da doença e maior número de hospitalizações (Nogueira et al., 1990). Estes dois sorotipos foram também os responsáveis pela epidemia dos anos 1995-1996, na qual foram notificados 51.456 casos de dengue, e pela epidemia ocorrida em janeiro de 1998 no Vale de Paraíba, a partir da qual o vírus se espalhou

rapidamente para outros municípios. Já no ano 2000, foi isolado o sorotipo DENV-3 de um paciente que apresentou um caso de dengue clássico no município de Nova Iguaçu, assim como também do vetor *Ae. aegypti* coletado do campo. A introdução deste novo sorotipo trouxe associada um aumento no número de casos durante o verão de 2001-2002, provocando a maior epidemia até então ocorrida no estado, com 288.245 casos relatados, dos quais 1.831 foram casos de dengue grave, com quadros hemorrágicos e síndrome de choque, e 91 evoluíram a óbito (Noqueira et al., 2007). Este sorotipo circulou predominantemente no estado até o ano 2007, quando o sorotipo DENV-2 reemergiu ocasionando a grave epidemia do ano de 2008 que acabou afetando todo o país. O estado do Rio de Janeiro, porém, foi o mais acometido, com um total de 198.269 de casos notificados (SVS/MS, 2009).

Em 2011 foi identificado o sorotipo DENV-4 no município de Niterói, o qual se disseminou no ano seguinte para outros municípios do estado, e foi o responsável por grande parte dos casos de dengue do ano 2012 (SVS/MS, 2012). A circulação deste sorotipo começou a diminuir nos anos posteriores, cedendo o predomínio novamente ao sorotipo DENV-1, responsável pelos casos dos últimos dois anos (SVS/MS, 2016).

## **1.8 Diagnóstico laboratorial das infecções por DENV**

Os testes de diagnóstico laboratorial desempenham um papel crucial na assistência ao paciente, confirmando ou descartando o caso por uma infecção suspeita (Peeling et al., 2010).

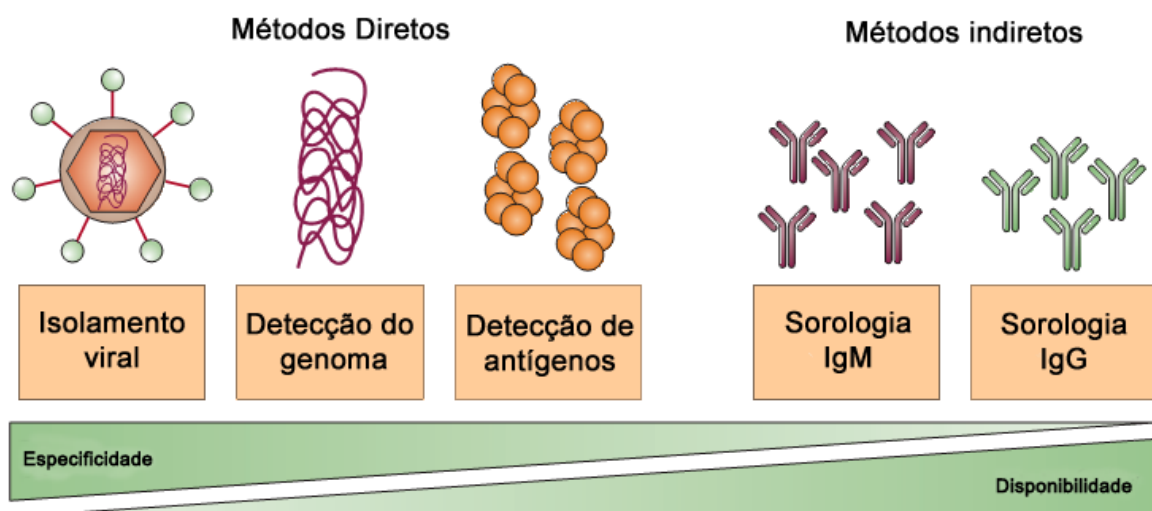
Nas infecções por DENV, o diagnóstico laboratorial de casos suspeitos é importante para a vigilância da doença, o monitoramento dos sorotipos circulantes do vírus e para o diagnóstico diferencial de outras doenças que causam sinais e sintomas clínicos semelhantes à dengue (Nogueira & dos Santos, 2015).

O diagnóstico laboratorial das infecções por DENV pode ser realizado por métodos diretos que visam isolar ou identificar o vírus, ou os seus componentes (antígenos como por exemplo NS1, e o ácido nucleico viral) ou por métodos indiretos, os quais consistem na detecção de anticorpos específicos de tipo IgM e IgG decorrentes da infecção. O período da doença em que o paciente se encontra é fundamental para a apropriada escolha do método diagnóstico a utilizar e para a correta interpretação dos resultados. A viremia é detectável por aproximadamente 4-

5 dias após o início dos sintomas, e correlaciona-se estreitamente com a duração da febre. Nas infecções primárias, os anticorpos de tipo IgG surgem lentamente, com baixos títulos após 8-10 dias desde o começo dos sintomas, enquanto os de tipo IgM são detectados tipicamente após os primeiros 5 dias de iniciados os sintomas e perduram na circulação sanguínea por 2-3 meses. Nas infecções secundárias, no entanto, os anticorpos de tipo IgG surgem rapidamente, e em altos títulos, logo depois do começo da febre, enquanto os de tipo IgM podem em alguns casos ser indetectável (WHO, 2009).

Os métodos diretos, que são certamente mais específicos que os indiretos, nem sempre são os mais disponíveis no diagnóstico de rotina das infecções por DENV (Figura 1.8-1).

**Figura 1.8-1** Métodos empregados no diagnóstico das infecções por DENV



Fonte: imagem adaptada de Peeling et al., 2010.

O soro é a amostra por eleição, embora plasma, sangue total e fragmentos de tecidos também sejam úteis. A linhagem celular C6/36 derivada de *Ae. albopictus* é o sistema de isolamento viral de escolha para o diagnóstico, porém, a inoculação intratorácica de mosquitos do gênero *Toxorhynchites* é a ferramenta para isolamento de maior sensibilidade. Ensaio de imunofluorescência com anticorpos monoclonais sorotipo-específicos ou ensaios de transcrição reversa seguida da reação em cadeia da polimerase (RT-PCR) são empregados para a identificação dos sorotipos. Por sua vez, ensaios de RT-PCR ou RT-PCR em tempo real são os métodos de escolha para a detecção do genoma viral. A detecção do antígeno viral NS1, secretado pelas células infectadas, oferece a possibilidade também do diagnóstico inicial, pois pode

ser detectada até 9 dias após o início da febre, e também em amostras de tecidos. Utilizam-se testes rápidos comerciais e ensaios imunoenzimáticos (ELISA). Finalmente, a detecção de anticorpos é o teste mais amplamente utilizado na vigilância laboratorial. Existem diferentes formatos de ELISA que podem ser empregados na detecção dos mesmos (Guzman & Harris, 2015).

## 1.9 Prevenção e controle

O controle da doença tem se tornado cada vez mais problemático. Programas baseados no controle vetorial e programas comunitários que visavam manter o ambiente livre de criadouros trabalharam de maneira constante e ativa, porém, falharam até hoje em erradicar o mosquito vetor. Novas abordagens para o controle vetorial foram desenvolvidas nos últimos 5 anos apresentando progressos interessantes: i) a adaptação endossimbiótica da bactéria *Wolbachia* de *Drosophila* no *Ae. aegypti* mostrou efeitos de encurtamento da vida do mosquito e de bloqueio da transmissão direta do DENV (Hoffmann et al., 2011); ii) avanços com *Ae. aegypti* geneticamente modificados carregando um gene dominante letal e a liberação desses mosquitos machos na natureza também se mostrou como uma medida de controle biológico efetiva na redução de índices entomológicos (Harris et al., 2011); iii) novos óleos essenciais com atividade larvicida e biopesticidas derivados de outros microrganismos estão sendo desenvolvidos (Veerakumar et al., 2013; Dias & Moraes, 2014).

Por outro lado, em consequência das dificuldades encontradas na implementação de programas de controle do mosquito vetor, junto com o crescente número de casos de dengue no mundo, o desenvolvimento de uma vacina tetravalente que estimulasse uma resposta imune balanceada e eficiente para os 4 sorotipos virais tornou-se uma grande necessidade. No entanto, o desenvolvimento de vacinas multivalentes contra a dengue tem sido dificultada pela falta da indução de uma resposta imunitária equilibrada. Vírus atenuados e inativados, proteínas recombinantes e vacinas de DNA encontram-se em desenvolvimento como possíveis candidatos (Tabela 1.9-1) (Thomas & Endy, 2011; Guzman & Harris, 2015).

O primeiro estudo de eficácia de vacina contra a dengue (vacina tetravalente, recombinante, viva e atenuada) mostrou uma eficácia baixa contra o DENV-2,

levantando várias questões importantes. Em primeiro lugar, a discrepância entre os resultados do teste imunológico escolhido como ferramenta e os resultados do estudo sugeriam que este ensaio de neutralização (PRNT) podia não ser suficientemente preditivo do resultado da infecção pelo vírus da dengue. E, alternativamente, o aumento dos títulos dos anticorpos neutralizantes poderia não ter sido suficiente para proteger particularmente contra a cepa epidêmica de DENV-2, ou bem a taxa de ataque clínico da cepa circulante poderia ter sido inusualmente elevada (Sabchareon et al., 2012).

A vacina desenvolvida por Sanofi Pasteur, Dengvaxia®, é a única que até hoje cumpriu estudos de eficácia em fase III (Hadinegoro et al., 2015), foi certificada pela Organização Mundial da Saúde e foi aprovada já em alguns países endêmicos para dengue, como Filipinas, México, Brasil e Paraguai. Esta vacina tetravalente, viva e atenuada, consiste em uma estrutura recombinante do cDNA infeccioso da cepa vacinal do vírus da Febre Amarela (17D) cujos genes preM e envelope foram substituídos pelos correspondentes aos 4 sorotipos do DENV (Guy et al., 2011; Sanofi Pasteur, 2016).

Evidências recentes em populações humanas e modelos de ratos sugerem um papel de proteção exercido pelas células T CD8+, com a maioria dos epítomos localizados nos genes não estruturais (Yauch et al., 2009). Desta forma, a vacina quimérica de Sanofi Pasteur carece assim, de resposta das células T às proteínas não estruturais do vírus dengue. O ensaio em curso de fase 3 a longo prazo será chave para determinar se a vacina confere proteção contra a doença ou não.

**Tabela 1.9-1** Vacinas candidatas contra a dengue

	<b>Descrição</b>	<b>Estado do estudo clínico</b>
<b>Vacina quimérica viva-atenuada</b>		
<b>Vacina quimérica YF17D/dengue</b>	cDNA recombinante infeccioso da cepa vacinal 17D do vírus da febre amarela, com os genes prM e Env substituídos pelos genes prM e Env dos 4 sorotipos do DENV	Fase 3
<b>Vacina viva atenuada</b>		
<b>Vírus atenuado vivo tetravalente</b>	Atenuação por passagem seriada em células PDK	Fase 1/2
<b>Vacina viva-atenuada do clone infeccioso do vírus da dengue</b>		



<b>Vacina atenuada quimérica recombinante</b>	Clone infeccioso atenuado de DENV-2 contendo prM / M e E de DENV-1, DENV-3 e DENV-4	Fase 2
<b>Vacina atenuada com mutante portadora de deleção em 3' UTR</b>	Deleção atenuante de 30 nucleótidos em 3' UTR de DENV-1, DENV-3 e DENV-4; Produção de quimera com prM/M e E de DENV-2 em um esqueleto de DENV 4-atenuado	Fase 1/2
<b>DNA</b>		
<b>D1ME</b>	Genes codificantes das proteínas prM and E	Fase 1
<b>Proteínas</b>		
<b>r80E</b>	Expressão de 80% da proteína do Env (contendo N-terminal) em células de insecto	Fase 1
<b>cEDIII</b>	Domínio III do gene do Env fusionado à proteína p64K de <i>Neisseria meningitides</i> , e expresado <i>Escherichia coli</i>	Pré-clínica
<b>Inativadas</b>		
<b>Vírus dengue purificado e inativado</b>	Vírus total purificado e inativado	Fase 1
<b>Partícula semelhante ao vírus</b>		
<b>EDIII-capsid protein</b>	Proteína quimérica que compreende ao domínio III do ene do Env e a proteína do capsídeo de DENV 1–4	Pré-clínica
<b>Virus vector</b>		
<b>Alphavirus PRV</b>	Partículas de replicação de Alphavirus que expressam duas configurações do antígeno do Env do DENV (Partículas sub-virais [prME] e dímeros solúveis da proteína do Env [E85])	Pré-clínica
<b>Adenovírus</b>	Formulação tetravalente combinando duas construções bivalente de adenovírus	Pré-clínica
<b>Vírus do Sarampo</b>	Expressão do antígeno do vírus da dengue por um vetor derivado do vírus vacinal do sarampo (atenuado/vivo)	Pré-clínica

Fonte: adaptado de Guzman & Harris, 2015. PDK: células de cultura primária de rim de cão.

### 1.10 Justificativa

O cenário hiperendêmico do Brasil, com a co-circulação dos quatro sorotipos de DENV, resultando em padrões epidemiológicos cada vez mais graves, com aumento do número de casos, doença grave em crianças e altas taxas de hospitalização, leva a necessidade de novos estudos que permitam compreender melhor a evolução do DENV e o seu impacto na doença.

Com o objetivo de entender o processo evolutivo do DENV, as variações de linhagem e quanto esses fatores impactam em termos de resposta imune, número

de casos e gravidade da doença, nesse estudo analisamos a relação entre a diversidade genética do DENV-2 inter e intra-hospedeiro com a gravidade do quadro clínico apresentado por pacientes do estado do Rio de Janeiro, durante o período de 1999 a 2011. Para tal, descrevemos as características genéticas virais e sua relação com a gravidade da doença (dengue vs. dengue grave), com o tipo de infecção se primária ou secundária, refletindo uma possível maior virulência e patogenicidade.

Estudos focados no sorotipo DENV-2 revelam a circulação simultânea de diferentes linhagens no país. A existência de uma maior variabilidade genética entre o DENV-2 também tem sido apontada como causa do insucesso em testes com vacinas.

A composição genética das populações virais intra-hospedeiro tem se mostrado essencial para a permanência do poliovírus *in vivo* (Vignuzzi et al., 2006) e por influenciar a patogêneses do vírus da imunodeficiência humano (HIV) (Lee et al., 2008) e do vírus da hepatite C (HCV). Para o caso do HCV, uma diversidade viral elevada foi detectada nos indivíduos que apresentavam hepatite crônica leve, em comparação com aqueles que tinham um dano hepático grave (Sullivan et al., 2007) assim como após o tratamento, a baixa diversidade se associava com uma resposta sustentada à terapia com os antivirais (Moreau et al., 2008). Desta maneira, dada a notória relevância biológica das subpopulações virais é fundamental ressaltar a importância deste estudo.

O atraso observado entre a introdução de um novo sorotipo, genótipo ou linhagem viral em uma área determinada, e a sua detecção, reforça a necessidade de programas de vigilância nessas regiões, a fim de detectar e rastrear a evolução desses vírus o mais rapidamente possível, o qual poderia ajudar a prevenir novos surtos, assim como também reduzir o número de casos de dengue grave e mortes.

Desta forma, a informação gerada neste estudo retrospectivo e analítico, é de suma relevância, porque gera informações importantes que podem auxiliar no controle de uma doença com elevado impacto social e econômico, associado ao conhecimento científico.

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

Analisar a diversidade genética de DENV-2 inter e intra-hospedeiro em amostras de pacientes do estado do Rio de Janeiro, durante o período de 1999-2011, e a sua potencial relação com o quadro clínico.

### **2.2 Objetivos Específicos**

- 1) Determinar o genótipo e linhagem das amostras de DENV-2 selecionadas para o estudo;
- 2) Analisar e comparar a diversidade dos espectros de mutantes virais identificados nas amostras de DENV-2 dos pacientes;
- 3) Relacionar o tipo de resposta imune (primária/secundária), a carga viral, a linhagem envolvida e a diversidade genética nas amostras estudadas com o quadro clínico (dengue com ou sem sinais de alarme e dengue grave).

### 3 MATERIAL E MÉTODOS

#### 3.1 Desenho do estudo

Trata-se de um estudo de tipo retrospectivo, descritivo e observacional.

#### 3.2 Considerações éticas

Foram empregadas para o desenvolvimento deste estudo amostras de soro pertencentes ao banco de amostras do Laboratório de Flavivírus, IOC/FIOCRUZ, Centro de Referência Regional para Dengue e Febre Amarela, após a realização de testes diagnósticos, sem nenhum procedimento adicional para o paciente. As mesmas provêm da demanda espontânea do laboratório, e são mantidas armazenadas a -70°C,

O estudo encontra-se dentro dos objetivos aprovados pelo Comitê de Ética em Pesquisa (CEP 274/05) da Fundação Oswaldo Cruz, Ministério da Saúde.

#### 3.3 Amostragem

Foram selecionadas amostras de soro previamente confirmadas por RT-PCR e/ou isolamento viral como DENV-2. As amostras eram oriundas de pacientes com quadros clínicos classificados como dengue clássico com ou sem sinais de alarme ou dengue grave, de diferentes municípios do estado do Rio de Janeiro e do período de 1999 a 2011.

**\*Critérios de inclusão:** amostras de casos confirmados de dengue (por RT-PCR e/ou isolamento viral) pelo sorotipo viral 2, ocorridos no período compreendido entre os anos 1999 e 2011, de pacientes residentes de diferentes municípios do estado do Rio de Janeiro. Foram selecionadas apenas aquelas amostras que tinham disponível o dado dos dias de doença até a data da coleta e/ou da classificação da resposta imune, com carga viral detectável e cuja ficha epidemiológica informasse os sinais e sintomas descritivos do quadro clínico.

**\*Critérios de exclusão:** amostras de casos suspeitos de dengue não confirmados laboratorialmente para DENV-2, assim como os casos confirmados pelos sorotipos virais 1, 3, ou 4. Foram excluídas aquelas amostras com carga viral não detectável, sem o dado dos dias de doença ou a classificação da resposta imune e sem a ficha epidemiológica ou não preenchida.

### 3.4 Classificação das amostras quanto à gravidade do quadro clínico

A gravidade dos casos foi determinada de acordo com as manifestações clínicas descritas nas fichas epidemiológicas de cada paciente, e seguindo os critérios estabelecidos pela Organização Mundial da Saúde conforme a figura 1.5-1.

### 3.5 Classificação das amostras segundo o tipo de resposta imune (primária ou secundária)

Foi empregada a técnica de ELISA para a determinação do título de anticorpos de tipo IgG específicos contra DENV. O título de anticorpos resultante, em relação ao número de dias de doença após o início dos sintomas determinam o caráter primário ou secundário da resposta imune montada em defesa dessa infecção, segundo proposto por Miagostovich e colaboradores no ano 1999 (Miagostovich et al., 1999), e como pode se observar na seguinte tabela (Tabela 3.5-1):

**Tabela 3.5-1** Critérios para a classificação da resposta imune

	Dias após o início dos sintomas			Resposta imune
	≤ 5 dias	6-9 dias	≥ 10 dias	
Título de IgG	< 1/160	< 1/10.240	< 1/40.960	Primária
	≥ 1/160	≥ 1/10.240	≥ 1/40.960	Secundária

### **3.5.1 Elisa para a determinação do título de anticorpos de tipo IgG anti-DENV.**

Microplacas de poliestireno de 96 poços (Immulon II, Dynatech) foram lavadas uma vez com PBS pH 7,4 e posteriormente sensibilizadas com 100 µl/poço de fluido ascético hiperimune produzido em camundongo (contendo uma mistura de anticorpos anti-DENV-1, DENV-2, DENV-3 e DENV-4) diluído 1/1500 em tampão Carbonato/Bicarbonato 0,1M, pH 9,6. Após 18 horas de incubação à 4°C em câmara úmida, as placas foram lavadas 3 vezes com PBS pH 7,4. Em seguida, os poços foram bloqueados com 200 µl/poço de diluente padrão, composto de PBS pH 7,4, Tween 20 a 0,05% e soro normal de cabra a 3%. As placas foram incubadas 1 hora a 37 °C em câmara úmida, e a seguir o conteúdo dos poços foi trocado por 75 µl/poço de antígeno (mistura equitativa das cepas Hawaii de DENV-1, Nova Guiné de DENV-2 e H87 de DENV-3) diluído previamente 1/15 no diluente padrão. As placas foram incubadas novamente por 1 hora a 37 °C em câmara úmida, e após 3 lavagens com PBS pH 7,4, foram adicionados 100 µl de soro do paciente diluído previamente 1:40 em diluente para soro e conjugado (PBS pH 7,4, Tween 20 a 0,05% e leite em pó desnatado em 3%) na primeira fileira da placa (poço A1 ao A12) enquanto nos poços restantes foram adicionados 75 µl de diluente de soro e conjugado. A continuação, foram transferidos de maneira seriada 25 µl de cada soro (A1-A12 na placa), desde a primeira até a última fileira, resultando cada linha em 8 diluições na razão 4 de cada amostra, começando em 1/40, até 1/655.360. Em cada placa, a primeira e a segunda linha foram reservadas para os controles negativo e positivo respectivamente, os quais consistem basicamente em dois soros humanos de reatividade conhecida. As placas foram incubadas em seguida, por 1 hora a 37 °C em câmara úmida para ser posteriormente submetidas a 3 novas lavagens com PBS pH 7,4 e 1 vez com diluente padrão. Foram adicionados 40 µl de anticorpos anti-IgG humano conjugado com peroxidase (Sigma Chemical Company, St. Luis, EUA), previamente diluído 1/1000 em tampão de diluição para soro e conjugado. As placas foram incubadas durante 1 hora em câmara úmida, a temperatura ambiente, e protegidas da luz. Foram lavadas 6 vezes com PBS pH 7,4 e 100 µl do substrato (resultante da mistura de partes iguais de ABTS e peróxido de hidrogênio; Kirkegaard & Perry Laboratories) foi adicionada em cada poço. As placas foram incubadas por meia hora, novamente em câmara úmida, a temperatura ambiente e protegidas da luz. Finalmente, foi realizada a sua leitura da densidade óptica de

maneira visual e em espectrofotômetro (405nm) (Multiskan MCC, Thermo Electron Corporation). O inverso da última diluição positiva para cada amostra, foi considerada como o título de anticorpos respectivo de cada uma delas.

### **3.6 Extração do RNA viral**

O RNA viral foi extraído a partir de 140 µL de soro empregando o kit QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA-EUA), segundo as instruções do fabricante.

### **3.7 Determinação da carga viral**

Os RNAs extraídos foram submetidos ao RT-PCR em tempo real acompanhados de 5 diluições feitas a partir de um controle positivo para DENV-2 de concentração conhecida, e um controle negativo (água). Para o desenvolvimento desta técnica foi empregado o protocolo proposto por Johnson e colaboradores no ano 2005 (Johnson et al., 2005) e melhorado por Santiago e colaboradores em 2013 (Santiago et al., 2013).

#### **3.7.1 Construção da curva de calibração**

Para se obter a carga viral das amostras expressa em cópias/ml, foi construída uma curva de calibração a partir de diluições seriadas no fator 1:10 do controle positivo de concentração conhecida, em tampão de eluição (EB, Qiagen). A cepa viral empregada como controle positivo para DENV-2 (cepa 40247), foi cultivada em células Vero e titulada por plaqueamento (realizado por Bio-Manguinhos), obtendo-se uma concentração de  $5,01 \times 10^5$  pfu/ml. Foi considerado segundo o modelo descrito por Johnson e colaboradores (Johnson et al. 2005), que 1 RNA infeccioso corresponde a 1 PFU. Foram empregadas as diluições na ordem  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ ,  $10^{-1}$  e  $10^0$  para a construção da curva de calibração.

#### **3.7.2 RT-PCR em tempo real**

A concentração do RNA viral presente nas amostras foi estimada pela quantificação absoluta do DNA complementar (cDNA), realizada pela técnica de RT-PCR em tempo real em uma única etapa, sistema TaqMan, ocorrendo inicialmente a transcrição do RNA viral em cDNA e posteriormente a sua amplificação por PCR. Foi

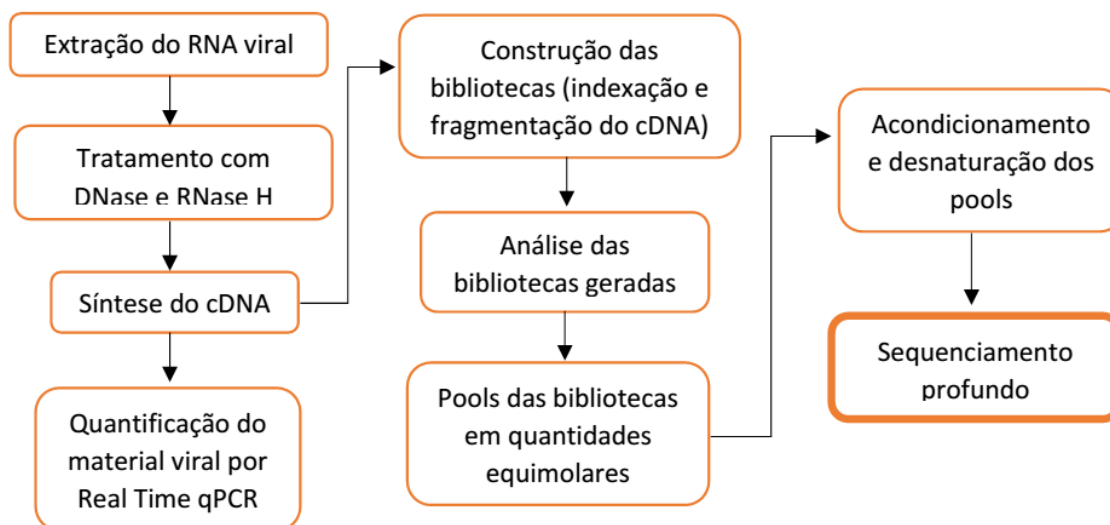
empregado para este fim o kit One-Step RT-PCR da Invitrogen. Os primers e a sonda marcada na posição 5' com o fluoróforo HEX específicos para a região 1426-1504 do gene do envelope do DENV-2, foram previamente descritos por Johnson e colaboradores, 2005.

Os resultados foram analisados com o software do equipamento de RT-PCR em tempo real, LineGene 9660 (Bioer), e também traz a construção manual da curva de calibração, para a qual foram plotadas em um gráfico as concentrações das diluições do controle e os seus valores de Ct obtidos. Os resultados da concentração para as respectivas amostras se extrapolaram desta curva.

### 3.8 Sequenciamento profundo

A pesquisa de subpopulações virais de DENV-2 dentro de uma mesma amostra, foi realizada utilizando a técnica de sequenciamento profundo mediante o sistema NextSeq500 da Illumina. Esta técnica de nova geração tem a capacidade de sequenciar de maneira clonal diferentes fragmentos de DNA de forma massiva e simultânea, assim como também detectar pequenas variações entre os produtos amplificados (Garner & Phil 2011; Illumina 2016a).

Considerando que a sensibilidade deste tipo de ferramenta é consideravelmente maior que a técnica de Sanger, é necessário que as amostras sejam submetidas a um processamento prévio mais rigoroso, conforme detalhado no fluxograma a seguir.





### **3.8.1 Extração do RNA**

O RNA foi purificado a partir das amostras de soro com AVL usando o kit QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA-EUA), seguindo as instruções do fabricante, porém com algumas pequenas modificações introduzidas:

- Foram adicionados 4  $\mu$ l de  $\beta$ -mercaptoetanol a cada amostra com AVL para garantir a desnaturação das ribonucleases que pudessem ser liberadas na lise celular.
- Todas as centrifugações foram realizadas a máxima velocidade para garantir a passagem do volume total das soluções através da coluna.
- A eluição final do RNA da coluna foi feita com uma solução de Acrilamida linear (AL) (Ambion) diluída 1/250 em água (livre de RNase e DNase; Ambion), e previamente aquecida a 70°C para melhorar a recuperação do RNA.

### **3.8.2 Tratamento com DNase**

O produto da extração foi tratado durante 30 minutos a 37°C com DNase Turbo (Ambion) para digerir o DNA contaminante, removendo as moléculas de DNA presentes nas amostras extraídas. Posteriormente se procedeu à lavagem e purificação do RNA utilizando o kit Agencourt RNAClean XP *beads* (1,8X) (Beckman Coulter Genomics), seguindo as instruções do fabricante. As amostras resultantes eluídas em H<sub>2</sub>O-AL foram armazenadas em alíquotas de 5  $\mu$ l cada uma, no freezer a -80°C.

### **3.8.3 Tratamento com RNase H**

Uma alíquota de 5  $\mu$ l de RNA tratado com DNase foi submetida ao tratamento com RNase H (Epicentre) para a remoção do RNA ribossomal e o RNA *carrier*, previamente empregado no processo de extração do RNA. Para isso, sondas que cobrem a região inteira do RNA ribossomal foram inicialmente hibridizadas ao RNA previamente desnaturado (Tabela 3.8.3-1):

**Tabela 3.8.3-1** Condições de desnaturação do RNA e hibridização de sondas específicas para o RNA ribossomal e RNA *carrier*

Volume	Reagente	Condições de reação
2 µl	Tampão de hibridização* 5x	
1,22 µl	DNA Oligo Mix (50 nt cada um, 100 µM)	95°C/2'
1 µl	Oligo (d)T (IDT; 5,6 µg/ul)	95-45°C a 0,1°C/seg
0,78 µl	Água+AL	45°C/5'
5 µl	RNA previamente tratado com DNase	

\*NaCl 1M, Tris-HCl 500 mM pH 7,4.

Em seguida, e o mais rapidamente possível, foi realizada a digestão com RNase H, acrescentando na reação anterior (Tabela 3.8.3-2):

**Tabela 3.8.3-2** Condições de ação para a RNase H

Volume	Reagente	Condições de reação
2 µl	Tampão de reação para a RNase H* 10x	45°C/30'
3 µl	RNase H termoestável (Epicentre; 5U/µl)	seguido de banho de gelo
5 µl	Água+AL	

\*NaCl 1M, Tris-HCl 500mM pH 7,5, MgCl<sub>2</sub> 200mM.

Com o objetivo de remover os oligos DNA empregados anteriormente se prosseguiu imediatamente (para evitar a degradação inespecífica do RNA) com o tratamento das amostras com DNase I (Qiagen), seguindo o seguinte protocolo (Tabela 3.8.3-3):

**Tabela 3.8.3-3** Condições do tratamento com DNase I para remoção das sondas de DNA

Volume	Reagente	Condições de reação
7,5 µl	10x DNase tampão de reação (Qiagen)	37°C/30'
2 µl	RNase free DNase I (Qiagen; 2,72U/µl)	
1 µl	SUPERasa-In RNase inhibitor (20U/µl)	
44,5 µl	H <sub>2</sub> O-AL	
20 µl	RNA tratado	

Foram empregados 5 µl de EDTA 0,5M por amostra para parar a reação.

Finalmente a DNase foi removida e as amostras purificadas empregando o kit Agencourt RNAClean XP *beads* (1,8X) (Beckman Coulter Genomics), seguindo as instruções do fabricante. O RNA foi eluído das *beads* com 11 µl de H<sub>2</sub>O-AL. As amostras foram rotuladas e armazenadas em *strips* no freezer a -20°C.

#### **3.8.4 Síntese do cDNA**

A primeira fita simples de cDNA foi sintetizada empregando iniciadores randômicos hexaméricos (Invitrogen) e a transcriptase reversa de alta-fidelidade Superscript TM III (Invitrogen). Foi adicionado também ao meio de reação o inibidor de RNases SUPERase·In™ (20U/µl; Ambion) para evitar a degradação do RNA (Tabela 3.8.4-1):

**Tabela 3.8.4-1** Condições de síntese da primeira fita simples de cDNA

Volume	Reagente	Condições de reação
1 µl	Primers randômicos hexaméricos	70°C/10' Seguida de Banho de gelo/1 a 5'
6 µl	H <sub>2</sub> O-AL	
4 µl	RNA da amostra	
4 µl	Tampão de reação da 1ª fita 5x	25°C/10'
2 µl	DTT 0,1M	
1 µl	dNTP mix 10mM	55°C/60'
1 µl	SUPERasa-n 20U/µl	Seguida de
1 µl	Superscript™ III RT	Banho de gelo
11 µl	RNA da amostra	

A síntese da segunda fita de cDNA foi realizada empregando as enzimas RNase H, DNA polimerase I e DNA ligase provenientes de *E. coli* (Tabela 3.8.4-2)

**Tabela 3.8.4-2** Condições de síntese da fita complementar do cDNA fita simples

Volume	Reagente	Condições de reação
43 µl	Água livre de RNase	
8 µl	10x tampão de reação da 2ª fita	
3 µl	dNTP mix 10mM	16°C/2hs
1 µl	<i>E.coli</i> DNA Ligase 10U/µl	Seguido de
4 µl	<i>E.coli</i> DNA Polymerase I 10U/µl	Banho de gelo
1 µl	<i>E.coli</i> DNA Rnase H 2U/µl	
20 µl	cDNA fita simples	

Foram empregados 5 µl de EDTA 0,5M por amostra para deter a reação. Finalmente, se realizou a purificação do cDNA empregando novamente o kit Agencourt RNAClean XP *beads* (1,8X) (Beckman Coulter Genomics), seguindo as instruções do fabricante, eluindo o cDNA em 9 µl de tampão de eluição (EB, Qiagen). O cDNA foi desta forma rotulado e armazenado no freezer a -20°C até o seu próximo uso.

### 3.8.5 PCR quantitativa em tempo real (qPCR)

Foi empregado um ensaio de PCR em tempo real “in house” (Tabela 3.8.5-1) para determinar a quantidade de material viral recuperado após o processamento anterior. A marcação dos produtos da PCR gerados fez o uso de Sybr green. Foram utilizados primers específicos para dengue vírus que mapeiam na região 9702-10125 do gene NS5 que codifica para a polimerase viral (Primer forward 5´ GAAAGACCAGAGATCCTGCTGTCT 3´ e primer reverse 5´ ACCATTCCATTTTCTGGCGTT 3´). O equipamento empregado foi o “QuantStudio™ 7 Flex Real-Time PCR System” (Applied Biosystems).

**Tabela 3.8.5-1** Condições do ensaio *in-house* PCR em tempo real

Reagentes (1x)	Condições de reação
10µl de Master mix Sybr green plus (Applied Biosystems)	95°C/10´
1µl de Primer forward (10 pmol/µl)	40 ciclos de: 95°C/15´´
1µl de Primer reverse (10 pmol/µl)	60°C/1´
7µl de agua	
1µl de cDNA	

Para a determinação da concentração de cópias do RNA viral presentes nas amostras, foi empregada uma curva de calibração gerada com sete diluições seriadas (1:10) de um controle positivo interno do laboratório para DENV-2. O mesmo foi feito a partir de um fragmento de cDNA da polimerase viral específico de DENV-2, de 109 bp, inserto em um plasmídeo, com posterior replicação em bactérias (*Escherichia coli*), purificação, digestão e quantificação, resultando numa concentração de 17ng/µl. Esse valor, transformado em cópias/ml foi de  $1,44 \times 10^{11}$  cópias/ml. Para tal conversão foi empregada a seguinte fórmula (CELS, 2004): Número de cópias = (quantidade \*  $6,022 \times 10^{23}$ )/(comprimento \*  $1 \times 10^9$  \* 650), representando  $6,022 \times 10^{23}$  ao número de Avogadro (moléculas/mol),  $1 \times 10^9$  ao fator de conversão de gramas a nanogramas, e 650 ao peso estimado em gramas de um mol de pares de bases.

As diluições do controle positivo foram realizadas com tampão de eluição (EB, Qiagen) resultando em soluções (pontos da curva de calibração) com  $1 \times 10^7$  até 10 cópias/ml.

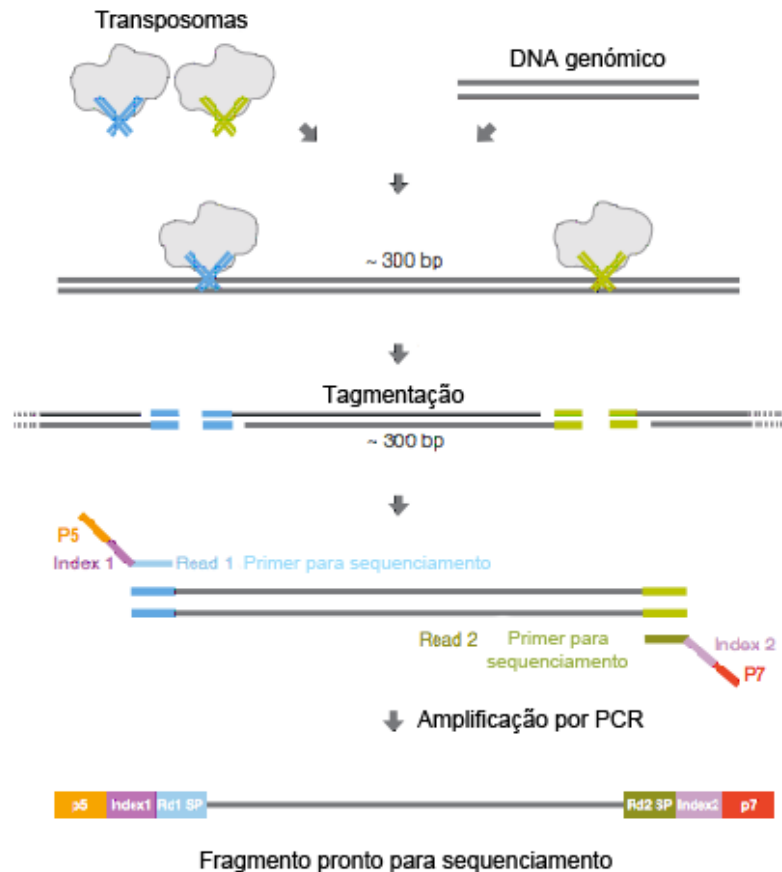
A curva de calibração foi construída com os valores da concentração em cópias/ml de cada diluição do controle, e os seus valores do Ct obtidos da PCR em tempo real. Para cada amostra, o resultado da concentração foi extrapolado dessa curva.

### **3.8.6 Preparo das bibliotecas**

Os cDNAs das amostras foram tratados com o Nextera XT DNA Sample Prep Kit (Illumina), seguindo as instruções dos fabricantes para a tagmentação do cDNA e a posterior adesão, mediante uma amplificação por PCR, de um único par de sequências índices (oligonucleotídeos de 6-8 pares de bases), com o objetivo final da sua posterior identificação. Este procedimento encontra-se esquematizado na figura 3.8.6-1.

Finalmente se procedeu à purificação das bibliotecas empregando novamente o kit Agencourt RNAClean XP *beads* (nesta vez, 0,6X para favorecer a retenção nas *beads* dos fragmentos de tamanhos menores a 1000 pares de bases) (Beckman Coulter Genomics), seguindo as instruções do fabricante. O cDNA agora fragmentado e indexado, ou biblioteca, foi eluído em 10  $\mu$ l de tampão de eluição (EB, Qiagen).

**Figura 3.8.6-1** Processo de fragmentação do cDNA e marcação com sequências índices



Fonte: imagem adaptada de “Nextera XT DNA Library Prep Reference Guide” (Illumina 2016c).

### 3.8.7 Análises das bibliotecas geradas: Bioanalyzer 2100

As bibliotecas foram posteriormente analisadas com o kit Agilent High Sensitivity DNA no Agilent 2100 Bioanalyzer seguindo as instruções dos fabricantes. O fundamento da análise no Bioanalyzer é a separação dos fragmentos de cDNA, neste caso, por eletroforese em gel, sendo a detecção deles efetuada através do sinal de fluorescência que emite o “Dye” ao ser estimulado por um laser, e fica intercalado entre as fitas do DNA. Com a ajuda do “Ladder” que contém componentes de tamanhos conhecidos, é construída uma curva padrão de tempo de migração relacionando com o tamanho dos fragmentos, sendo os valores das

amostras extrapolados dessa curva. Desta maneira obteve-se a concentração média dos fragmentos para cada amostra, assim como também o seu tamanho médio (Agilent, 2016). Desta forma, foi possível estimar quão efetiva foi a construção das bibliotecas.

### **3.8.8 Construção dos pools das amostras selecionadas**

As bibliotecas geradas foram então classificadas segundo a sua concentração, e misturadas em pools de maneira equimolar (valor determinado pela amostra que apresente a biblioteca em menor concentração), segundo à similaridade das suas concentrações. Este critério foi utilizado para evitar gerar um viés no sequenciamento propriamente dito, devido ao fato de que amostras em baixa concentração poderiam ficar sub-representadas se misturadas com outras de elevada concentração.

### **3.8.9 Acondicionamento das bibliotecas**

O *pool* das bibliotecas foi posteriormente tratado com o kit Agencourt RNAClean XP *beads* (0,7X para favorecer novamente a limpeza e retenção nas *beads* dos fragmentos menores de 1000 pares de bases) (Beckman Coulter Genomics), seguindo as instruções do fabricante, eluindo-as em 10 µl de tampão de eluição (EB, Qiagen). Em seguida o *pool* purificado foi novamente analisado no Bioanalyzer empregando o kit Agilent High Sensitivity DNA para conhecer qual a sua concentração e qualidade.

### **3.8.10 Sequenciamento profundo**

Dependendo da concentração obtida para cada *pool* das bibliotecas, o mesmo foi diluído para atingir uma das 4 possíveis concentrações iniciais propostas no “Protocolo A: Standard Normalization Method” presente na guia “NextSeq System Denature and Dilute Libraries Guide” (4nM, 2nM, 1nM ou 0,5nM) (Illumina, 2016b).



Seguindo as instruções do mesmo se diluíram e desnaturaram os pools das bibliotecas. A diluição final, no entanto, diferente do que sugere o protocolo, foi feita de maneira a obter uma concentração final de 1,5 pM. Esta modificação do protocolo teve como objetivo evitar a saturação do processo de sequenciamento, evento que tem sido previamente observado no desenvolvimento de outros projetos, ao trabalhar com a concentração proposta de 1,8 pM. Foi empregado como controle interno para o sequenciamento a biblioteca “PhiX” (fornecida pela Illumina), a qual foi processada da mesma maneira que os pools de interesse. As bibliotecas desnaturadas e diluídas na concentração 1,5 pM junto ao controle PhiX 1% foram finalmente sequenciadas no sistema NextSeq 500® da Illumina, sendo geradas sequências curtas de aproximadamente 75-80 nucleotídeos, com síntese dos fragmentos em ambos sentidos (*pair-ended*).

### 3.9 Análise bioinformática

Os dados gerados no sequenciamento profundo foram analisados mediante o emprego de “Galaxy”, uma plataforma computacional pública disponível online em <https://usegalaxy.org> (Afgan et al., 2016). Diversos softwares foram empregados consecutivamente para analisar as sequências obtidas para cada amostra, e as posições nucleotídicas envolvidas na diversidade (Apêndice I).

### 3.10 Análise da diversidade genética intra- e inter-hospedeiro

A complexidade genética das populações virais foi estimada empregando o cálculo da Entropia de Shannon normalizada para cada posição nucleotídica (Grubaugh et al., 2015; Gregori et al., 2016; Lequime et al., 2015):

$$S_n = \frac{-(p \times \ln(p)) + ((1-p) \times \ln(1-p))}{\ln(4)}$$

O valor de p representa a frequência obtida para cada SNV (do inglês “Single Nucleotide Variation”) respectivamente, enquanto o Ln(4) representa a máxima

complexidade que pode se achar para uma determinada posição nucleotídica (A, T, G e C simultaneamente). Os valores de  $S_n$  para uma posição nucleotídica dada, podem variar de 0 (sem diversidade, ou seja, sem polimorfismo) a 1, o que representa a máxima complexidade possível (por exemplo, números iguais de nucleotídeos alternativos). Para cada amostra, se obteve a média deste índice, considerando a somatória do valor obtido para cada posição (incluindo aquelas conservadas, ou seja, onde a complexidade é zero) dividida pelo comprimento do genoma total obtido para cada amostra, respectivamente.

Por outro lado, foi estimada também a diversidade genética ( $\pi$ ) para cada posição nucleotídica (Gregori et al., 2016; Lequime et al., 2015):

$$\pi = \frac{D}{D - 1} \times (1 - (p^2 + (p - 1)^2))$$

Sendo D o valor que expressa a profundidade do sequenciamento para cada posição, e o p, como já exposto, a frequência da variante. Como acontece com a  $S_n$ , os valores para  $\pi$  variam entre 0 e 1. Novamente foi calculada a média sobre todas as posições nucleotídicas que compõem o genoma viral para cada amostra. Este índice fornece uma medida do valor médio do número de diferenças nucleotídicas que existem entre duas subpopulações virais quaisquer, que coexistem dentro de um indivíduo (Gregori et al., 2016).

Se bem ambos índices representam a diversidade genética, este último é menos sensível às variantes de baixa frequência (Lequime et al., 2016).

Em contrapartida, foi analisado o impacto das variações achadas para cada amostra, com respeito à região do genoma, e a sua natureza sinônima ou não sinônima. Desta forma, as sequências nucleotídicas consenso obtidas das amostras foram empregadas como plataformas onde se montaram todas as variantes achadas para cada uma delas. Posteriormente tanto as sequências consenso como as portadoras das variantes foram traduzidas em aminoácidos e foram assim detectadas quais dessas variantes apresentam mudanças de impacto nas diversas proteínas virais (substituições não sinônimas).

### **3.11 Análise filogenética**

As sequências consenso obtidas do sequenciamento profundo das amostras foram alinhadas com sequências de referência de DENV-2 obtidas do Genbank,

utilizando o programa Bioedit. O modelo de evolução foi obtido com o software JModeltest v2.1.6, e a árvore filogenética de máxima verossimilhança construída com o software RaxML v8.2.8. A árvore foi visualizada com os softwares Figtree v1.4.2 e Mega v7.0. O objetivo desta análise foi determinar qual o genótipo e a linhagem correspondente de cada amostra e a relação filogenética entre as cepas virais analisadas neste estudo.

### **3.12 Análise estatística**

Os pacotes estatísticos EpiInfo<sup>TM</sup>7 v7.1.1.14 e GraphPad Prism v7.01 foram empregados para o processamento e análise de dados. A significância estatística da relação entre variáveis categóricas (como por exemplo o gênero dos pacientes, a classificação clínica ou do status imune do paciente, entre outras) foi determinada empregando os testes Chi quadrado ou exato de Fisher conforme o caso. Variáveis quantitativas agrupadas foram analisadas com os testes t-student para dados não pareados ou Mann-Whitney nos casos onde a variável não apresentava um comportamento de tipo normal ou gaussiano, ou análises de variâncias (ANOVA) ou teste de Kruskal-Wallis, nos casos onde os grupos eram mais de dois. Exemplos destas variáveis quantitativas são a idade dos pacientes, a carga viral presente nas amostras, resultados de cobertura do sequenciamento, valores adotados pelos índices  $S_n$  e  $\pi$ , entre outras. Finalmente, o teste de correlação linear de Pearson foi aplicado na avaliação da relação entre variáveis quantitativas não agrupadas.

## 4 RESULTADOS

### 4.1 Amostragem

Desde a introdução do DENV-2 no estado do Rio de Janeiro em 1990 até o ano de 2011 o Laboratório de Flavivírus confirmou DENV-2 por RT-PCR e/ou isolamento viral em 913 amostras de casos suspeitos de dengue do estado do Rio de Janeiro. Informações de sinais e sintomas que permitissem classificar os casos de dengue com ou sem sinais de alarme e dengue grave, estavam disponíveis nas fichas epidemiológicas de 58% dos pacientes (533/913). Foram pré-selecionadas 52 amostras que reunissem os critérios desejados, agrupadas nos seguintes intervalos temporais: 1999-2001, 2007-2009, 2010 e 2011. Teve-se em conta também a sua origem geográfica, se nesse ano em particular houve um surto, ou tratava-se de casos isolados e qual a sua classificação clínica, deixando como parâmetros aleatórios a idade e sexo dos pacientes. Procedeu-se em seguida à extração do RNA segundo a técnica já descrita, cujo produto foi submetido ao RT-PCR em tempo real, segundo o protocolo descrito por Johnson e colaboradores. Conforme foi descrito por Johnson e Santiago (Johnson et al., 2005; Santiago et al., 2013), amostras que apresentaram valores de Ct menores que 37 são consideradas positivas. A nossa seleção final só incluiu aquelas cujo valor de Ct não superasse de 34. É recomendado para estudos de sequenciamento profundo, amostras com cargas virais apresentando Cts menores que 30, pois o resultado final é diretamente afetado por este parâmetro. A concentração das amostras foi calculada a partir da curva de calibração construída com as diluições seriadas do controle para DENV-2. Desta maneira, considerando os resultados obtidos no RT-PCR em tempo real, 29 amostras foram selecionadas para serem incluídas no estudo de sequenciamento profundo (Tabela 4.1-1). Das 23 amostras não incluídas, 5 apresentaram Ct maior que 35 e as 18 restantes foram negativas para o PCR em tempo real. As amostras selecionadas foram então aliqüotadas, acondicionadas segundo descrito previamente e enviadas em gelo seco ao Instituto Pasteur, França, onde foram armazenadas a -80°C até o seu posterior processamento. As amostras com Cts mais altos (maior ou igual a 26) e volume suficiente disponível, foram aliqüotadas e enviadas com o dobro do volume com objetivo de tentar, após a extração do RNA, compensar parcialmente a menor carga viral presente nelas. Por outro lado, foi

determinado o título de IgG anti-DENV nas amostras que não tivessem sido já classificadas (23/29), de forma que se conseguiu posteriormente conhecer o tipo de resposta imune de cada paciente (primária ou secundária).

O painel de amostras selecionadas, as suas características e o seu resultado do RT-PCR em tempo real podem se observar na tabela 4.1-1.

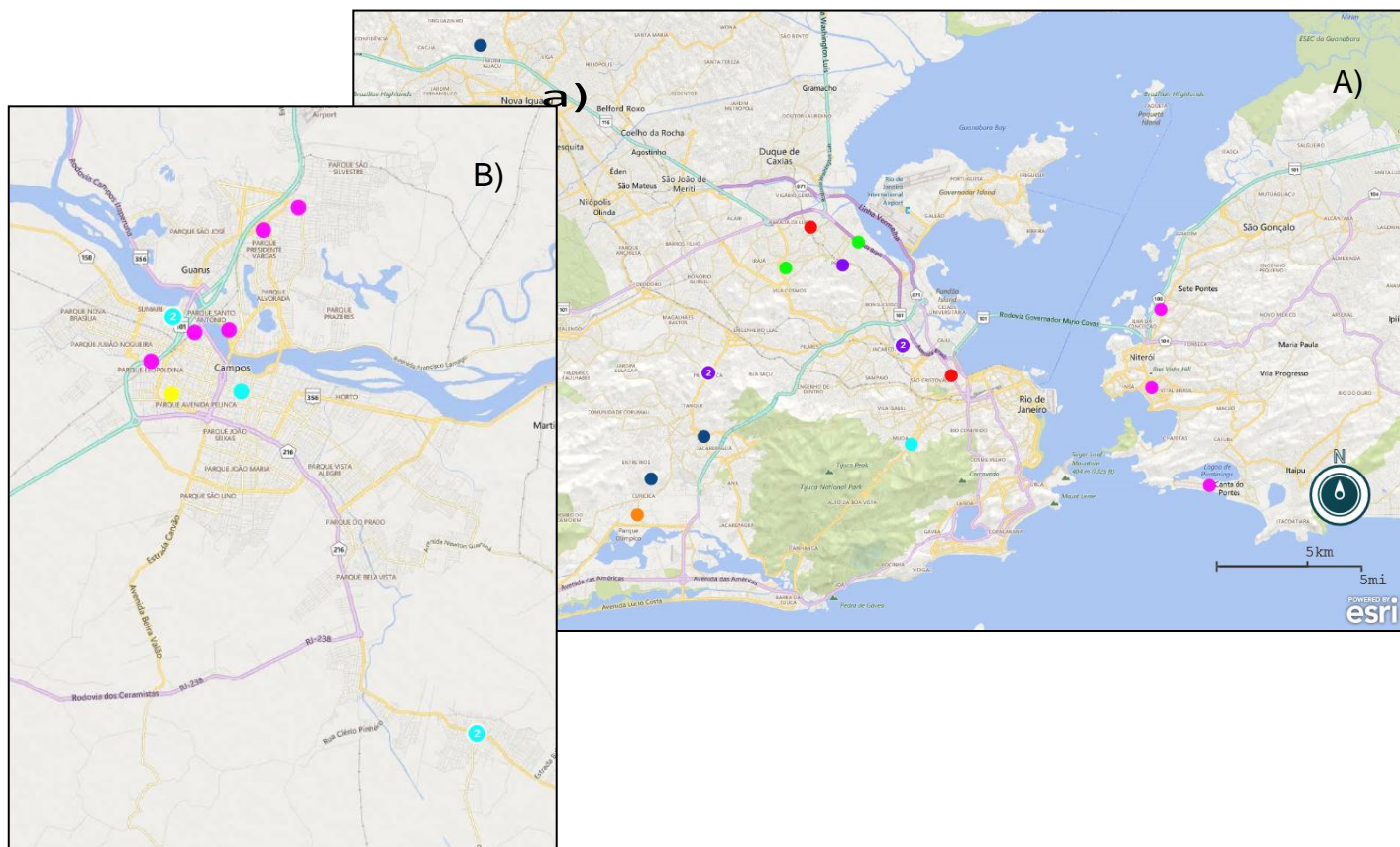
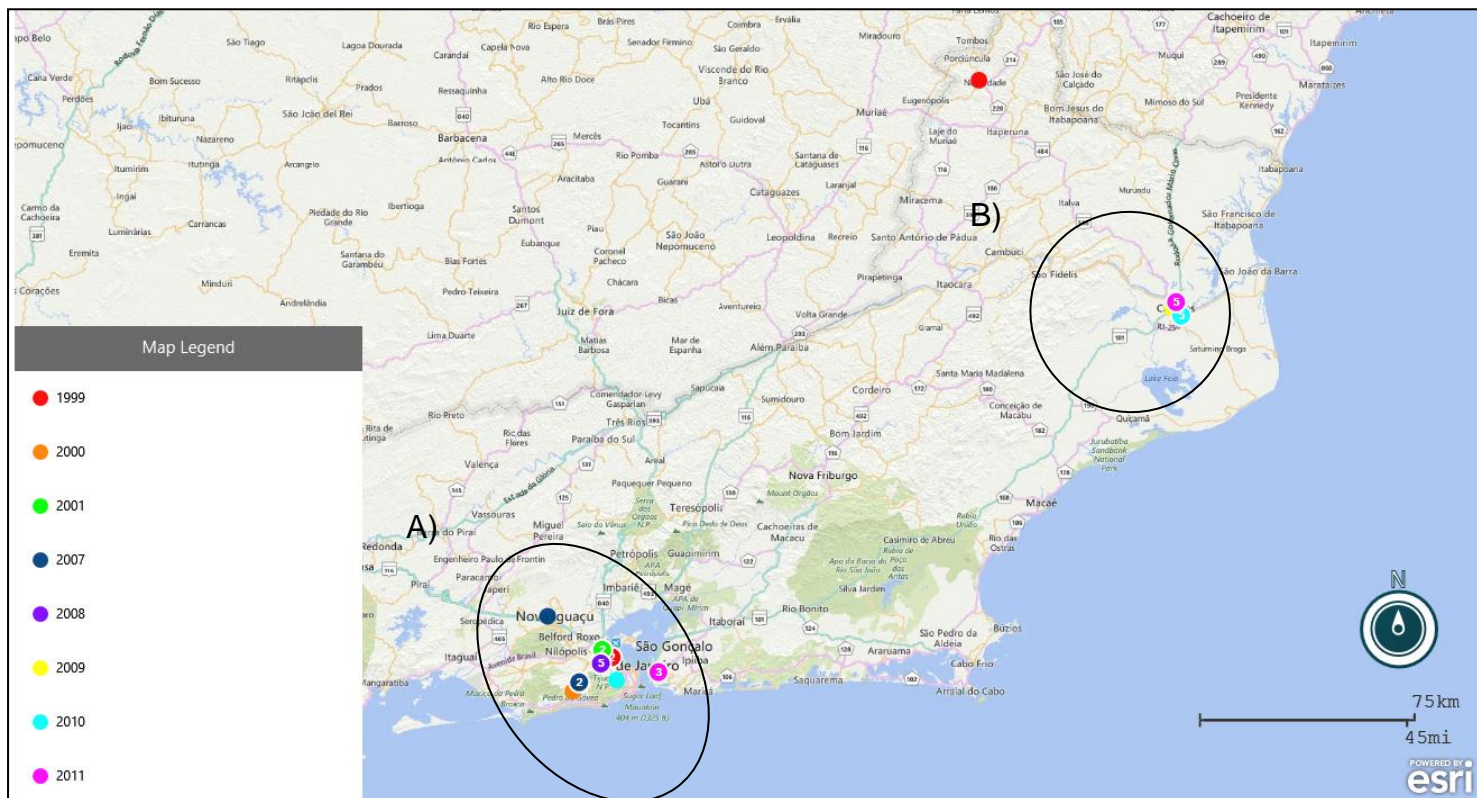
**Tabela 4.1-1** Características das amostras selecionadas

Amostra	Ano	Mês	Dias de sint.	Infecção	Classificação	Gênero	Idade	Bairro/Munic.	CT	CC (cópias/ml)	Volume amostra
131	1999	4	2	PRI	clássico	M	-	São Cristovão/RJ	23	8,69E+02	140ul
132	1999	4	2	PRI	clássico	M	-	Cordovil/RJ	32	2,87E+00	280ul
133	1999	1	2	PRI	sinais alarme	M	22	Santo Antonio de Padua	30	2,33E+02	280ul
134	2000	5	2	PRI	clássico	F	39	Jacarepaguá/RJ	30	1,42E+02	280ul
135	2001	4	1	PRI	clássico	M	27	Penha/RJ	31	1,23E+02	280ul
136	2001	5	1	PRI	sinais alarme	F	-	Vila da Penha/RJ	24	3,53E+03	140ul
137	2007	8	-	PRI	grave (óbito)	-	-	Pechincha/RJ	30	8,81E+00	280ul
138	2007	11	2	PRI	sinais alarme	M	8	Cerâmica/Nova Ig	19	1,01E+04	140ul
139	2007	12	5	PRI	grave	F	8	Curicica/RJ	33	3,42E+01	280ul
140	2008	3	10	PRI	grave (óbito)	F	75	Engenho da penha/RJ	29	3,61E+02	280ul
141	2008	3	5	SEC	clássico	F	10	Manguinhos/RJ	19	1,16E+04	140ul
142	2008	4	5	PRI	clássico	M	10	Manguinhos/RJ	22	1,61E+03	140ul
143	2008	3	-	SEC	grave (óbito)	F	14	Praça SECa/RJ	29	2,81E+02	280ul
144	2008	4	-	PRI	sinais alarme (óbito)	M	85	Praça SECa/RJ	27	7,60E+02	280ul
145	2009	12	3	SEC	clássico	F	29	Pq. Dom Bosco	29	2,47E+02	280ul
146	2010	4	2	SEC	clássico	F	50	Tijuca/RJ	17	2,29E+05	140ul
147	2010	3	2	SEC	sinais alarme	F	62	Caju/Campos	26	1,73E+03	280ul
148	2010	3	1	SEC	sinais alarme	F	17	Caju/Campos	26	1,06E+02	140ul
149	2010	6	5	PRI	grave	F	5	Centro/Campos	15	4,94E+05	140ul
150	2010	3	2	SEC	sinais alarme	F	31	Goitacazes/Campos	22	1,66E+04	140ul
151	2010	3	2	SEC	clássico	M	41	Goitacazes/Campos	24	3,82E+03	140ul
152	2011	2	2	SEC	sinais alarme	M	36	Bandeirantes/Campos	28	4,22E+02	280ul
153	2011	2	12	PRI	sinais alarme	F	45	Novo campos/Campos	31	1,06E+02	280ul
154	2011	5	2	SEC	clássico	M	21	Pq.Leopoldina/Campos	14	1,56E+05	140ul
155	2011	4	2	SEC	clássico	M	54	Barreto/Niteroi	14	1,66E+05	140ul
156	2011	4	3	PRI	sinais alarme	F	8	Icaraí/Niteroi	31	4,68E+00	280ul
157	2011	4	3	SEC	sinais alarme	M	28	Piratininga/Niteroi	30	1,59E+02	280ul
158	2011	5	1	SEC	clássico	M	63	Jardim carioca/Campos	20	3,38E+04	140ul
159	2011	5	3	SEC	clássico	M	36	Pte.Vargas/Campos	19	7,04E+03	140ul

Nota: Dias de sint: dias de doença após o início dos sintomas; PRI: primaria; SEC: secundaria; sinais alarme: dengue clássico com sinais de alarme; M: masculino; F: feminino; CC: concentração.

Em termos gerais, o grupo de amostras selecionadas se compunha por homens e mulheres em igual proporção (14 de cada um; 1 caso não determinado), com uma idade média de 33 anos, variando de 5 até 85 anos. Na figura 4.1-1 pode observar-se a localização geográfica dos casos e o ano de ocorrência.

**Figura 4.1-1** Localização geográfica e temporal dos casos incluídos neste estudo.



Nota: cada ponto colorido representa pelo menos um caso ocorrido nessa localização geográfica. Se mostram ampliados os focos de casos da região metropolitana (A) e do município de Campos dos Goytacazes (B).

Do total das amostras, 52% (n=15) correspondem a casos com respostas imunes primárias, enquanto 48% (n=14) secundárias. No que diz respeito à classificação clínica destes casos, 45% (n=13) foram casos de dengue clássico, 38% dengue com sinais de alarme (n=11) e 17% (n=5) casos de dengue grave.

Das amostras selecionadas, 3 apresentaram o número de dias de doença (desde o início dos sintomas até a coleta da amostra) desconhecido. Porém, a classificação da resposta imune para essas três amostras já constava na ficha epidemiológica. Das 26 amostras restantes, 92% (n=24) pertencem a casos com até 5 dias de doença (23% com 1 dia, 39% com 2 dias, 15% com 3 dias e 15% com 5 dias), enquanto o 8% restante (amostras 140 e 153) apresentou 10 e 12 dias de doença.

O estudo da relação destas variáveis entre si, e com a carga viral obtida para cada amostra, não encontrou nenhuma tendência ou diferença significativa quando analisada estatisticamente ( $p\text{-value}>0,05$ ) com a exceção da carga viral e o tipo de resposta imune, observando-se maiores cargas virais nas infecções de tipo secundário (Mediana= $5,43 \times 10^3$ ) quando comparadas com as primárias (Mediana= $2,33 \times 10^2$ ) ( $p\text{-value}$ : 0,02; teste T de student para dados não pareados). Curiosamente, nesta casuística não parece haver uma diferença significativa entre a carga viral atingida e os dias de doença ( $p\text{-value}>0,05$ ; análise da variância de um fator). As amostras com 2 e 5 dias de doença apresentam as maiores cargas virais (com médias na ordem de  $10^3$ ), seguidas das amostras com 1, 3, 10 e 12 dias, com valores médios na ordem de  $10^2$ .

Outra observação estatisticamente significativa foi que conforme aumenta a gravidade do caso clínico (desde dengue clássico até grave) aumenta a proporção de pacientes mulheres (desde 31% nos casos clássicos, passando por 54,5% nos casos com sinais de alarme até 100% nos casos graves) e diminui a de homens (69% nos casos clássicos, 45,5% nos casos com sinais de alarme e 0% nos casos graves) ( $p\text{-value}$ :0,05 para o teste do Chi quadrado aplicado sobre a frequência absoluta de homens e mulheres segundo a classificação clínica).

Quando analisada a idade dos pacientes com respeito ao quadro clínico, não existe uma diferença significativa na média da idade entre os três grupos, porém, os quatro casos graves com o dado da idade disponível, pertencem a 3 casos de crianças (de 5, 8 e 14 anos de idade) e um caso a um paciente idoso de 75 anos. Estes quatro casos graves aconteceram entre os anos 2007 e 2010.

## 4.2 Aspectos técnicos

### 4.2.1 Tratamento com DNase e RNase H, e posterior síntese do cDNA

Para avaliar o efeito do processamento das amostras nesta primeira fase, foi determinado o valor da carga viral mediante o emprego do qPCR em tempo real descrita anteriormente no item 3.8.5. Os resultados podem-se observar na tabela 4.2.1-1, enquanto o efeito deste tratamento sobre a carga viral pode se observar com clareza na figura 4.2.1-1.

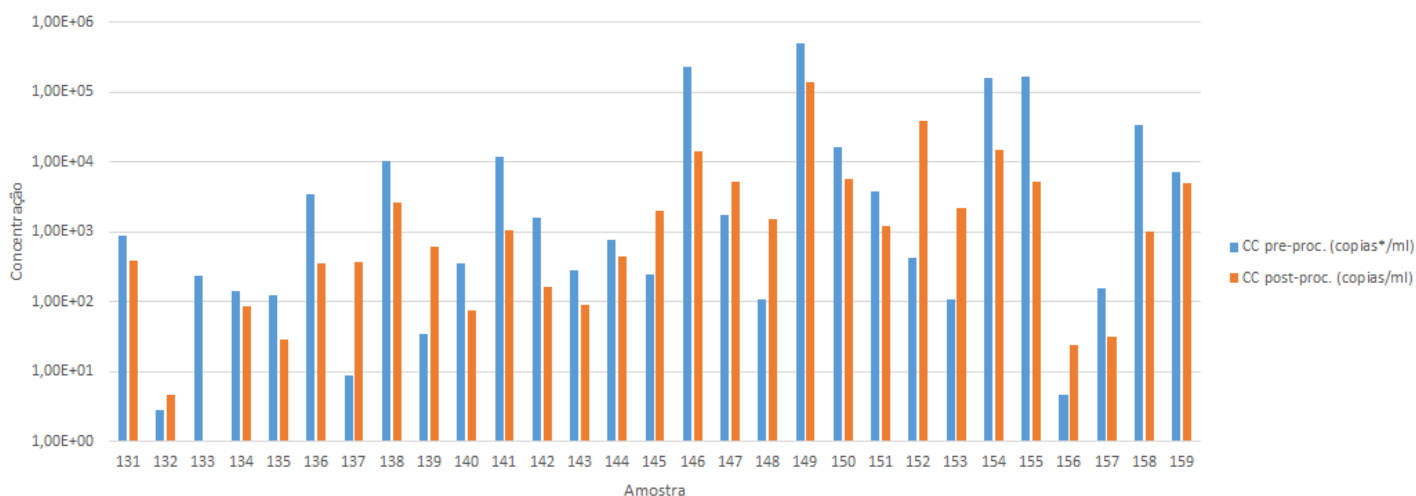
**Tabela 4.2.1-1** Resultado do PCR em tempo real das amostras processadas, pré e pós-tratamento com DNase, RNase H e síntese do cDNA.

Amostra	Vol. Amostra	Ct pré-proc.	Ct pós-proc.	CC pré-proc. (cópias*/ml)	CC pós-proc. (cópias/ml)
131	140ul	22,78	32,84	8,69E+02	3,97E+02
132	280ul	32,04	39,44	2,87E+00	4,71E+00
133	280ul	29,55	Undet.	2,33E+02	0,00E+00
134	280ul	30,48	34,83	1,42E+02	8,69E+01
135	280ul	30,74	36,56	1,23E+02	2,91E+01
136	140ul	24,46	32,63	3,53E+03	3,49E+02
137	280ul	30,23	32,51	8,81E+00	3,77E+02
138	140ul	18,79	26,116	1,01E+04	2,6E+03
139	280ul	33,14	31,75	3,42E+01	6,10E+02
140	280ul	28,73	35,03	3,61E+02	7,67E+01
141	140ul	18,57	30,89	1,16E+04	1,05E+03
142	140ul	21,78	33,84	1,61E+03	1,63E+02
143	280ul	29,2	34,79	2,81E+02	8,91E+01
144	280ul	27,33	32,65	7,60E+02	4,49E+02
145	280ul	29,44	30,28	2,47E+02	2,01E+03
146	140ul	16,65	23,588	2,29E+05	1,4E+04
147	280ul	25,79	25,054	1,73E+03	5,3E+03
148	140ul	26,19	30,75	1,06E+02	1,49E+03
149	140ul	15,21	20,24	4,94E+05	1,4E+05
150	140ul	21,56	24,958	1,66E+04	5,7E+03
151	140ul	24,31	31,10	3,82E+03	1,19E+03
152	280ul	28,44	25,57	4,22E+02	3,95E+04
153	280ul	31,02	30,12	1,06E+02	2,21E+03
154	140ul	14,36	23,512	1,56E+05	1,5E+04
155	140ul	14,25	25,105	1,66E+05	5,1E+03
156	280ul	31,25	37,23	4,68E+00	2,46E+01
157	280ul	30,27	36,83	1,59E+02	3,18E+01
158	140ul	20,23	27,489	3,38E+04	1,0E+03
159	140ul	19,38	28,84	7,04E+03	4,99E+03

Nota: vol: volume; pre-proc: pré-processamento; pós-proc: pós-processamento; CC: concentração; cópias\*/ml: cópias infectivas/ml.



**Figura 4.2.1-1** Variação da carga viral (cópias/ml) com o acondicionamento das amostras para o sequenciamento profundo.



Nota: cc: concentração.

A carga viral após o tratamento com DNase, RNase H e a síntese do cDNA, foi realizada sob condições diferentes ao RT-PCR em tempo real inicial, mostrou um aumento em 9 das 29 amostras (132, 137, 139, 145, 147, 148, 152, 153 e 156). Casualmente, 8 delas foram extraídas a partir do dobro do volume (280 µl). No entanto, as 7 amostras restantes que também se processaram a partir de 280 µl (133, 134, 135, 140, 143, 144 e 157), apresentaram uma diminuição da carga viral, como as processadas a partir do volume inicial de 140 µl. O teste de Wilcoxon para dados pareados mostrou uma significância estatística na comparação dos dois grupos (p-value 0,026), implicando que o processamento efetuado sobre as amostras desde a sua extração, impactou na sua carga viral.

#### **4.2.2 Construção dos pools das bibliotecas de sequenciamento profundo**

Após a construção e purificação das bibliotecas das amostras, dois *pools* de quantidades equimolares de cada biblioteca foram feitos. O primeiro (#1), incluiu as bibliotecas das amostras 146, 147, 149, 154, 155, 158 e outras não pertencentes a este projeto, enquanto o segundo *pool* (#2) incluiu todas as mostras restantes do nosso conjunto, com a exceção das amostras 150 e 152. Seguiu-se o critério de excluir tais amostras pelo fato de delas apresentarem bibliotecas em concentrações muito elevadas (superando àquelas de menor concentração em até 1000 vezes), e como já exposto, isso poderia gerar um viés no sequenciamento final. As mesmas serão analisadas numa corrida separada no futuro. Os pools foram posteriormente processados como sugerido na guia “NextSeq System Denature and Dilute Libraries Guide” (Illumina 2016b), e sequenciados por separado, sendo geradas duas corridas de sequenciamento profundo.

#### **4.2.3 Sequenciamento profundo**

Obtivemos duas corridas com bons parâmetros; a primeira com 200 clusters/mm<sup>2</sup>, 90,2% dos clusters passaram pelo filtro, enquanto a segunda teve 253 clusters/mm<sup>2</sup> e 72,5% dos clusters passaram pelo filtro. Em ambos casos, mais do 90% dos *reads*<sup>3</sup> apresentou um “score” de qualidade<sup>4</sup> (Q) igual ou superior a 30, ou seja, com uma probabilidade de 1 em 1000 de que a base assignada durante o sequenciamento esteja errada.

No que diz respeito aos resultados individuais para cada amostra, pode se observar na tabela 4.2.3-1 a quantidade de *reads* obtidos para cada amostra, quantos deles mapearam efetivamente à sequência de referência de DENV-2 (Genbank, número de acesso GU131879) e qual a porcentagem do genoma viral que conseguiu-se reconstruir.

---

<sup>3</sup> Entende-se por *read* a cada sequência curta de 35-38 pares de bases sintetizada no processo de sequenciamento profundo, a partir de cada fragmento da biblioteca.

<sup>4</sup> Um score de qualidade (Q-score) é uma previsão da probabilidade de um erro na assignação de uma base no processo de sequenciamento por síntese. Ele serve para informar probabilidades de erro muito pequenas.

**Tabela 4.2.3-1** Número de *reads* e cobertura do genoma das amostras sequenciadas no sistema NextSeq 500.

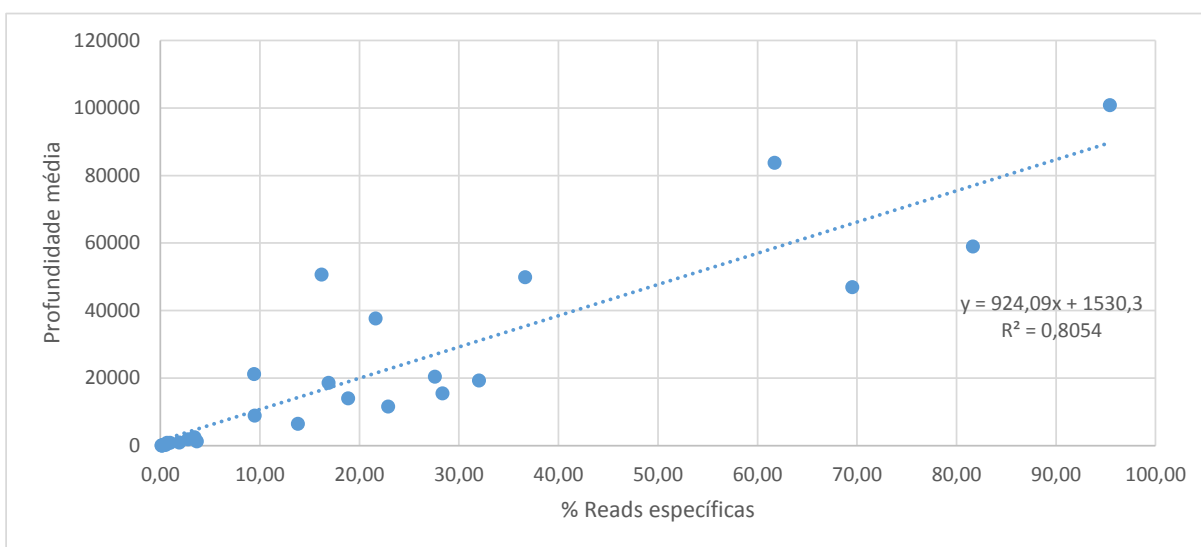
Amostra	Reads	Reads mapeadas (%)	Cobertura (%)	Profundidade média
131	25227692	244119 (1,0%)	99,91	841,33
132	28919580	30876 (0,1%)	99,31	100,07
133	37445334	246927 (0,7%)	99,41	848,08
134	10487688	382234 (3,6%)	99,7	1323,8
135	14170470	269255 (1,9%)	99,57	925,62
136	21400050	4038197 (18,9%)	99,87	14005,5
137	65074044	6126413 (9,4%)	99,89	21208,99
138	15772886	4468485 (28,3%)	99,87	15486,82
139	27131548	2567233 (9,5%)	99,76	8889,18
140	16123670	82735 (0,5%)	99,74	280,97
141	90147034	14585511 (16,2%)	99,97	50682,8
142	19833998	553352 (2,8%)	99,65	1920,03
143	14534780	26364 (0,2%)	99,68	84,54
144	20879494	713934 (3,4%)	99,78	2468,12
145	21399866	5899985 (27,6%)	99,98	20461,56
146	20844826	17017638 (81,6%)	100,00	59036,14
147	39112700	24142377 (61,7%)	99,95	83755,68
148	50487294	10917824 (21,6%)	99,97	37718,33
149	30491666	29091993 (95,4%)	100	100818,72
151	17411510	5574129 (32,0%)	99,82	19340,09
153	13627456	1883074 (13,8%)	99,98	6534,61
154	39170300	14355388 (36,6%)	99,95	49874,57
155	19430178	13506042 (69,5%)	99,89	46950,29
156	31177520	102702 (0,3%)	99,65	298,05
157	23385758	34415 (0,1%)	99,22	114,49
158	31661364	5352185 (16,9%)	99,98	18585,83
159	14666016	3355195 (22,9%)	99,87	11646,87

Nota: Entenda-se por “Cobertura” a percentagem do genoma que conseguiu ser mapeado, enquanto por “Profundidade média” o número médio de bases que mapeiam uma única posição nucleotídica determinada do genoma.

A percentagem de *reads* específicos apresentou valores diferentes entre as amostras, distribuídos em uma escala de valores que vai desde 0,11 até 95,41% ( $\bar{x}$  = 21,36 e SD=26,55), sendo que somente 4 amostras apresentaram mais de 40% de *reads* específicos (146, 147, 149 e 155). Existe uma clara correlação (p-value

<0,0001) entre o número de *reads* específicos e a profundidade média obtida para cada amostra, como pode se observar na figura 4.2.3-1. Por sua vez, existe uma correlação muito forte entre a carga viral inicial das amostras com a profundidade média atingida no sequenciamento (p-value 0,0009), relação que não tem suporte estatístico quando considerada a carga viral post-processamento (p-value>0,05).

**Figura 4.2.3-1** Relação entre a porcentagem dos *reads* específicos para DENV-2 e a profundidade média atingida no sequenciamento profundo.



De qualquer maneira, pode observar-se que para todos os casos se atingiu uma cobertura do genoma superior ao 99% ( $\bar{x}= 99,79$  e  $SD=0,21$ ), o que significa que o genoma viral foi sequenciado praticamente na sua totalidade.

### 4.3 Análise filogenética

Como parte do fluxo da análise bioinformática, foram obtidas as sequências consenso representativas do genoma completo das amostras virais em questão. A partir delas, foi construída uma árvore filogenética, empregando a sequência montada do genoma completo e as disponíveis no Genbank para DENV-2. Foi empregado o método de máxima verossimilhança (maximum likelihood; ML),

utilizando o modelo de substituição nucleotídica GTR+I+ $\Gamma$  (General Time Reversible com distribuição de substituições gamma e sítios invariantes) e 1000 réplicas como medida de suporte (bootstrap) (Figura 4.3-1a-d).

Todas as sequências obtidas para o vírus DENV-2 deste conjunto agrupam dentro do genótipo asiático/americano, circulante no Brasil desde a sua introdução nos anos 1988-1989 (Drumond et al., 2013). Porém, como se mostra na figura 4.3-1a existem dentro deste genótipo 3 linhagens, agrupados em dois clusters diferentes, que coincidem com os já descritos com Drumond e colaboradores no ano 2013. A linhagem I, no clado superior, pode ser observada detalhadamente no painel b) da figura 4.3-1, enquanto as linhagens II e III integradas no clado inferior podem se visualizar nos painéis c) e d) respectivamente desta figura.

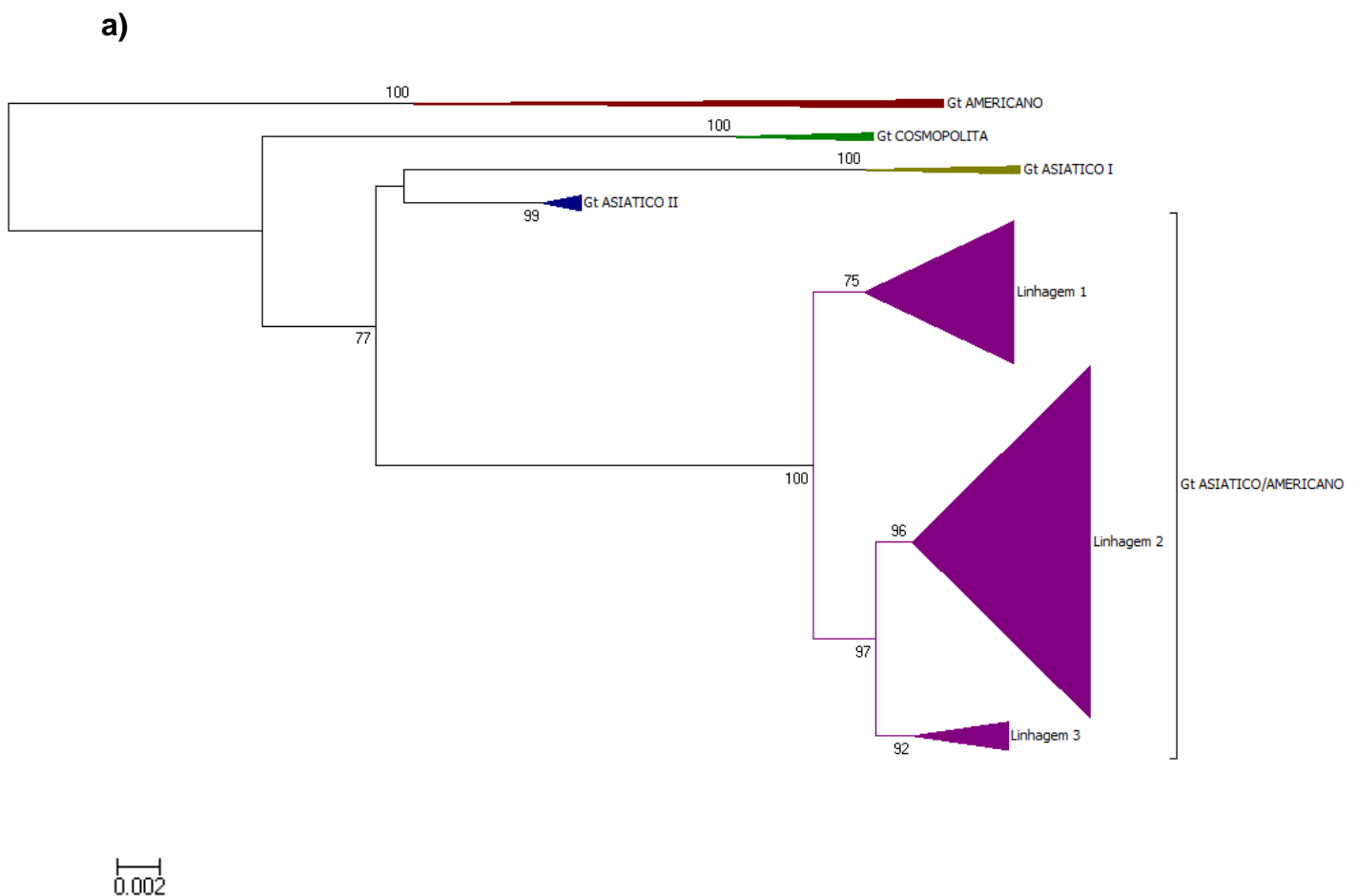
As amostras 131 a 136, correspondentes aos casos dos anos 1999 até 2001 se agruparam em diferentes clusters dentro da linhagem I, junto com outras sequências do Brasil que datam dos anos 1990 a 2003 e da Venezuela, Colômbia e Porto Rico. As amostras 132, 135 e 136 se localizam na árvore em um cluster separado das amostras 131, 133 e 134, porém, sem formar um grupo monofilético. A amostra 136, se mostrou mais próxima a sequências originárias do estado de Espírito Santo que às do Rio de Janeiro.

As amostras restantes (137-159) agruparam todas dentro da linhagem II, junto com outras sequências brasileiras dos anos 2007 a 2011 e sequências de diferentes países da latino-americano, inclusive Jamaica, de onde se acredita que essa linhagem pode ter tido origem (Drumond et al., 2013). Todas as amostras deste estudo parecem agrupar juntas dentro de um único cluster com outras sequências do Brasil e do Perú, existindo vários subgrupos dentro do mesmo. Chamou a atenção os seguintes eventos: entre as amostras do ano 2008, todas pertencentes a casos do município de Rio de Janeiro, 4 delas agruparam juntas (140, 141, 142 e 144) enquanto a amostra 143 do mesmo ano e originária do bairro de Praça Seca como a 144, agrupou separada, num subgrupo diferente. Por sua vez, as amostras originárias do município de Campos dos Goytacazes, pertencentes aos casos dos anos 2010 e 2011, agruparam juntas (147, 148, 149, 153, 154 e 159), com a exceção da amostra 158, que se localizou mais próxima de outras sequências brasileiras de anos anteriores. E finalmente, de maneira similar, as 3 amostras provenientes do surto ocorrido no ano 2011 no município de Niterói, agruparam todas separadas; a 155 junto com a 146 do bairro da Tijuca (município de Rio de Janeiro) do ano 2010, a 156 junto com outras amostras do município de Rio de

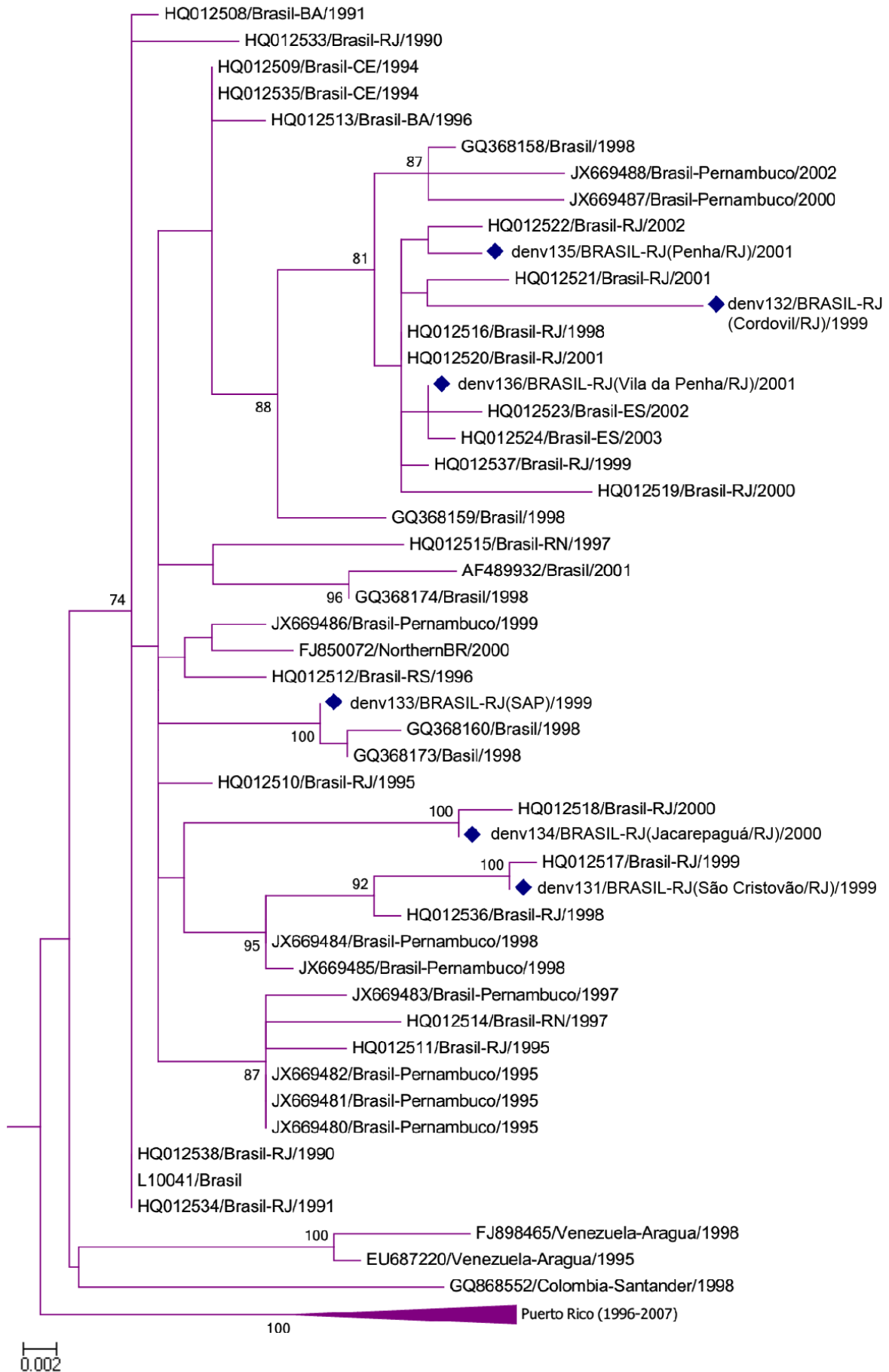
Janeiro também do ano 2010 e a 157 que agrupou junto com sequências de um surto de São Paulo no ano 2010.

Na linhagem III, dentro da qual não se achou nenhuma das amostras deste conjunto, agruparam isolados de casos acontecidos no nordeste do Brasil, durante os anos 2001 a 2006 e um grupo monofilético de Porto Rico. Acredita-se que esta linhagem tenha sido de circulação espaço-temporal limitada (Drumond et al., 2013).

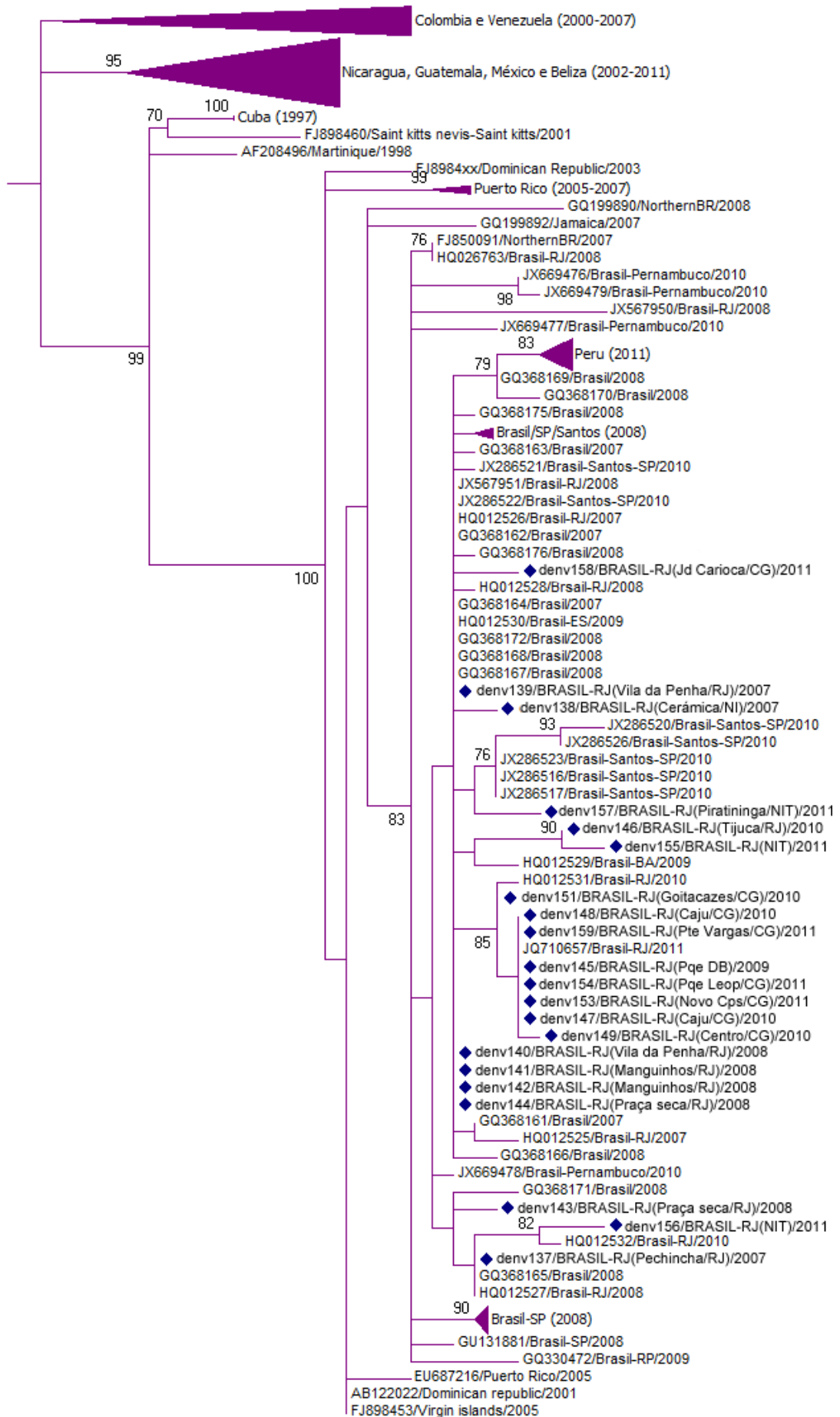
**Figura 4.3-1(a)** Árvore filogenética de máxima verossimilhança obtida para o genoma completo de DENV-2. **(b-d)** Ampliação das sub-árvores correspondentes às linhagens I (b), II (c) e III (d) do genótipo asiático/mericano. Identificadas com um losango azul encontram-se as sequências analisadas neste estudo. Os números nos nodos representam a medida de suporte (bootstrap) para cada um deles. Por consenso se mostram apenas os maiores a 70%. BR: Brasil; RJ: Rio de Janeiro; SP: São Paulo; ES: Espírito Santo.



b) Sub-árvore da linhagem I

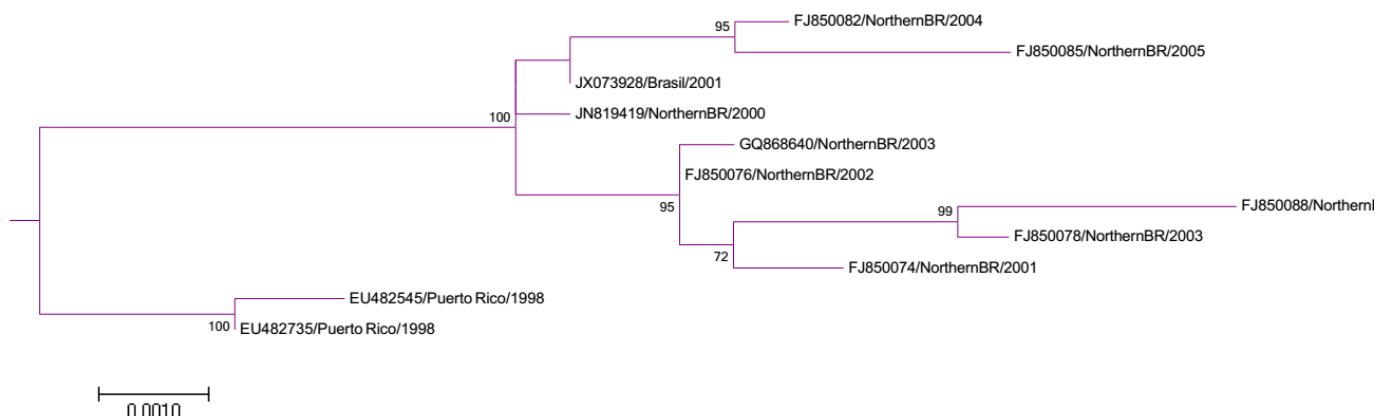


c) Sub-árvore da linhagem II





### c) Sub-árvore da linhagem III



## 4.4 Análise da diversidade viral

A procura de variantes (do inglês *variant-calling*) realizada com o software Lofreq v2.1.2 permitiu determinar o espectro de mutantes presente em cada uma das amostras, detectando-se qual a substituição nucleotídica, a posição do genoma envolvida, e a sua frequência (Apêndice II).

### 4.4.1 Índices de complexidade da população viral e da diversidade nucleotídica

A partir desses dados, foram calculados os índices da entropia de Shannon ( $S_n$ ) e índice  $P_i$  ( $\pi$ ) para cada posição nucleotídica, e posteriormente calculadas as suas médias respectivas para o genoma inteiro reconstruído de cada amostra (Tabela 4.4.1-1). As amostras 132 e 133 não apresentaram variação nenhuma, fato que indicaria que não existe diversidade viral nessas amostras clínicas. A média para o índice da entropia de Shannon resultou  $8,90 \times 10^{-4}$  ( $SD = 9,73 \times 10^{-4}$ ), enquanto para o índice  $P_i$   $6,48 \times 10^{-4}$  ( $SD = 6,10 \times 10^{-4}$ ).

**Tabela 4.4.1-1** Número de variantes e índices de complexidade e diversidade nucleotídica calculados para cada amostra

Amostra	Nº SNVs	Sn	$\pi$
131	18	6,11E-04	5,84E-04
132	0	-	-
133	0	-	-
134	3	1,09E-04	1,09E-04
135	21	6,04E-04	5,57E-04
136	150	1,94E-03	1,32E-03
137	22	7,36E-04	7,01E-04
138	96	1,24E-03	8,47E-04
139	6	1,92E-04	1,83E-04
140	3	1,28E-04	1,32E-04
141	160	1,84E-03	1,20E-03
142	18	5,21E-04	4,64E-04
143	22	8,78E-04	8,97E-04
144	33	8,22E-04	6,80E-04
145	221	2,46E-03	1,61E-03
146	99	6,46E-04	4,34E-04
147	119	8,10E-04	5,60E-04
148	488	4,62E-03	2,89E-03
149	100	6,70E-04	4,95E-04
151	168	1,90E-03	1,25E-03
153	18	3,56E-04	2,79E-04
154	76	5,15E-04	3,46E-04
155	80	4,17E-04	2,53E-04
156	17	6,95E-04	7,11E-04
157	4	1,71E-04	1,79E-04
158	68	5,12E-04	3,27E-04
159	43	6,48E-04	4,70E-04

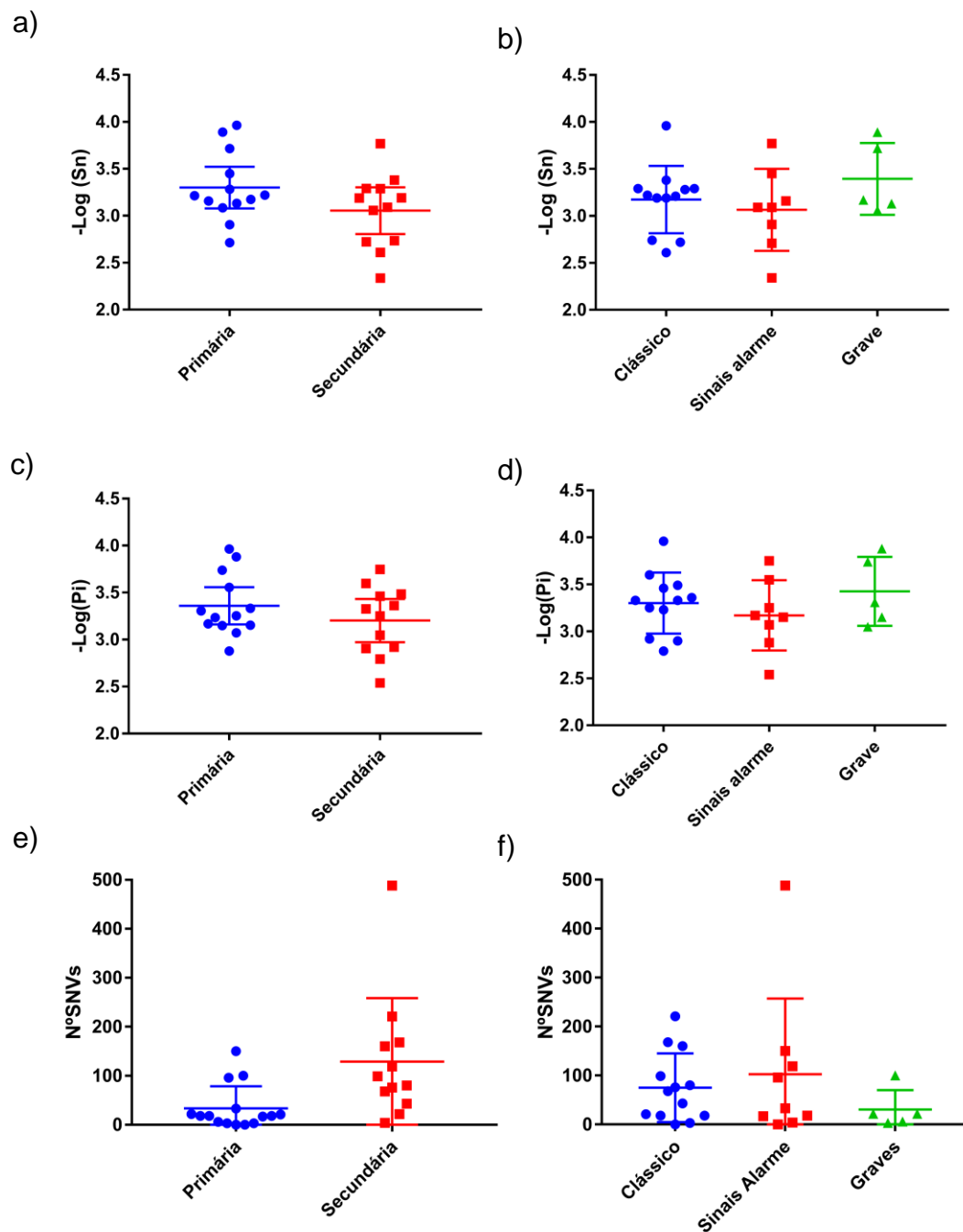
Nota: Nº SNVs: número de substituições de um único nucleotídeo; do inglês “single nucleotide variation”; Sn: índice da entropia de Shannon;  $\pi$ : índice Pi de diversidade nucleotídica.

Na figura 4.4.1-1(a-f) se observa a variação destes índices (representados como  $-\log$  do índice) quanto ao tipo de resposta imune, primária ou secundária (figura 4.4.1-1a e c), e segundo a classificação do quadro clínico (figura 4.4.1-1b e d). O comportamento dos índices em questão parece ser o mesmo, fato que se

distingue também ao observar os valores que eles adotam para cada amostra na tabela 4.4.1-1. Os resultados obtidos para os casos correspondentes às infecções primárias apresentam uma média maior que a dos secundários (figura 4.4.1-1a e c), o que implicaria menores índices de complexidade e diversidade nucleotídica. Com respeito às diferenças na gravidade do quadro clínico (figura 4.4.1-1b e d) os casos graves apresentam índices menores de diversidade. No entanto, nenhuma destas diferenças é estatisticamente significativa ( $p\text{-value}>0,05$ ). No que tange o número de SNV achadas para cada amostra se detectou uma diversidade estatisticamente menor para os casos de infecções primárias quando comparados com as secundárias (teste de Mann-Whitney;  $p\text{-value}=0,0027$ ) (figura 4.4.1-1e). No entanto, com relação à variação deste parâmetro nos grupos com diferentes apresentações clínicas, não houve uma diferença significativa entre as suas médias, utilizando o teste de Kruskal-Wallis ( $p\text{-value}=0,64$ ) (figura 4.4.1-1f).

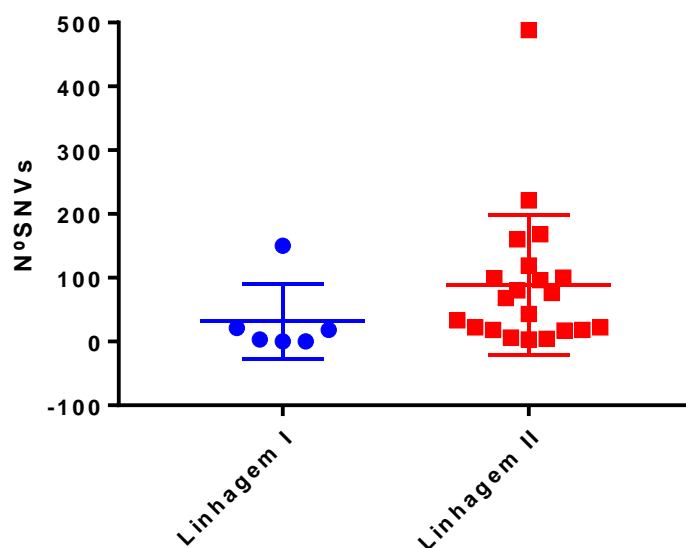
O número médio de SNV obtido para cada amostra foi de 82 ( $\bar{x}= 82,12$ ;  $SD=104,16$ ). Uma observação interessante é a comparação entre o número de SNV que apresentam as amostras de cada linhagem. Existe uma diferença significativa ( $p\text{-value}=0,04$ ) entre a média do número de SNVs da linhagem I ( $\bar{x}=32$ ) quando comparada com a média para a linhagem II no teste de Mann-Whitney, resultando maior nesta última ( $\bar{x}=88,6$ ) (figura 4.4.1-2). São somente 6 as amostras que pertencem à linhagem I, e duas delas não apresentaram diversidade nenhuma (132 e 133). No entanto, a profundidade média só foi baixa (100x) apenas para a amostra 133, enquanto para as restantes (131, 132, 134, 135 e 136) os valores foram superiores às 800 vezes a média em que cada posição nucleotídica foi sequenciada. É importante identificar estes fatores, para não acabar rejeitando uma observação que pode indicar uma informação interessante, por uma escassa cobertura na amostragem.

**Figura 4.4.1-1(a-f)** Número de variações nucleotídicas e medidas de diversidade viral segundo o tipo de resposta imune e a gravidade do quadro clínico.



Nota: Se representam os valores adoptados por cada amostra, classificadas em grupos segundo a gravidade do quadro clínico, as médias para cada grupo e os seus respectivos desvios padrões.

**Figura 4.4.1-2** Frequência das SNV achadas para os casos pertencentes à Linhagem I ou II



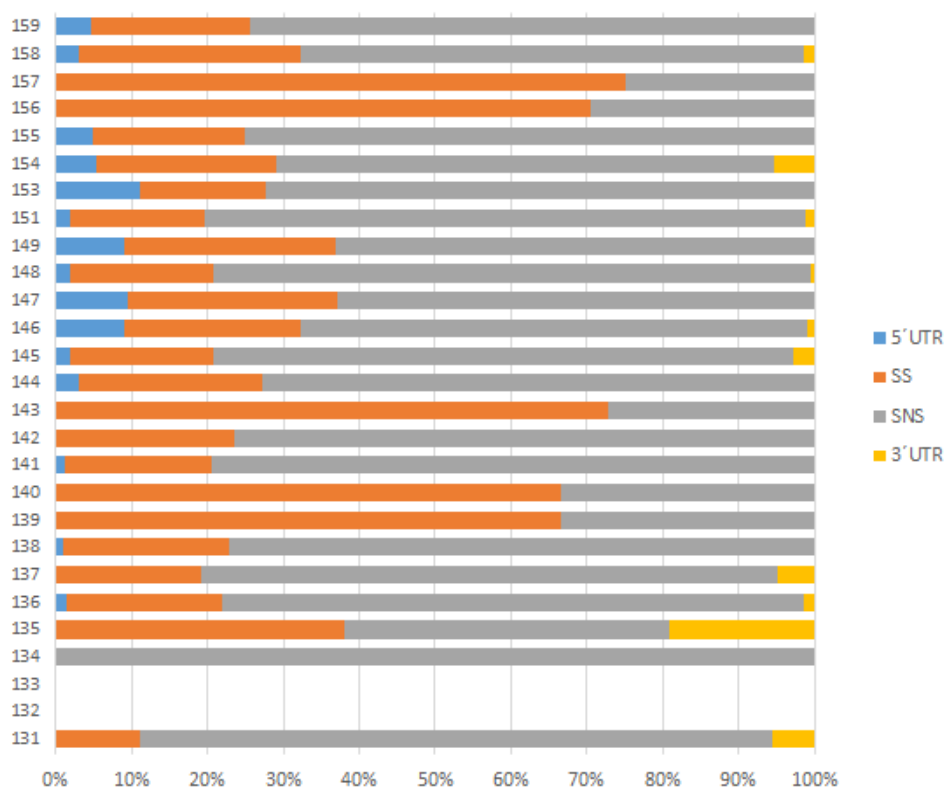
#### 4.4.2 Espectro de mutantes

Foram analisadas todas as variantes que as amostras apresentaram, e separadas segundo a sua região no genoma viral. Por sua vez, todas as que se encontraram dentro da fase aberta de leitura da poliproteína, foram traduzidas e determinado o seu impacto nas proteínas, ou seja, se tratavam-se de substituições sinónimas (SS) ou não sinónimas (SNS).

Das 27 amostras sequenciadas, 25 apresentaram SNV como visualizado na tabela 4.4.1-1, das quais 15 (60%) apresentaram substituições na região 5' não codificante enquanto apenas 10 (40%) na região 3' não codificante. Das mutações achadas na fase aberta de leitura da poliproteína, todas as amostras apresentaram entre 25 e 100% de substituições não sinónimas ( $\bar{x}= 67,8\%$ ;  $SD=21$ ).

Na figura 4.4.2-1 encontra-se representada para cada amostra, a frequência relativa das SS e SNS assim como também das SNV que ficaram fora do marco de leitura, ou seja, nas regiões 5' e 3' não codificantes. As percentagens foram calculadas com respeito ao número total de SNV para cada amostra.

**Figura 4.4.2-1** Frequência do tipo de SNV com respeito à região do genoma viral



Nota: UTR: região não codificante, do inglês “untranslated region”; SS: substituições sinónimas; SNS: substituições não sinónimas.

Na figura 4.4.2-2 está representada a frequência intra-hospedeiro de todas as SNV achadas na fase aberta de leitura da poliproteína para cada amostra. A linha contínua por cima da linha de pontos, mostra o nível de frequência de 1%.

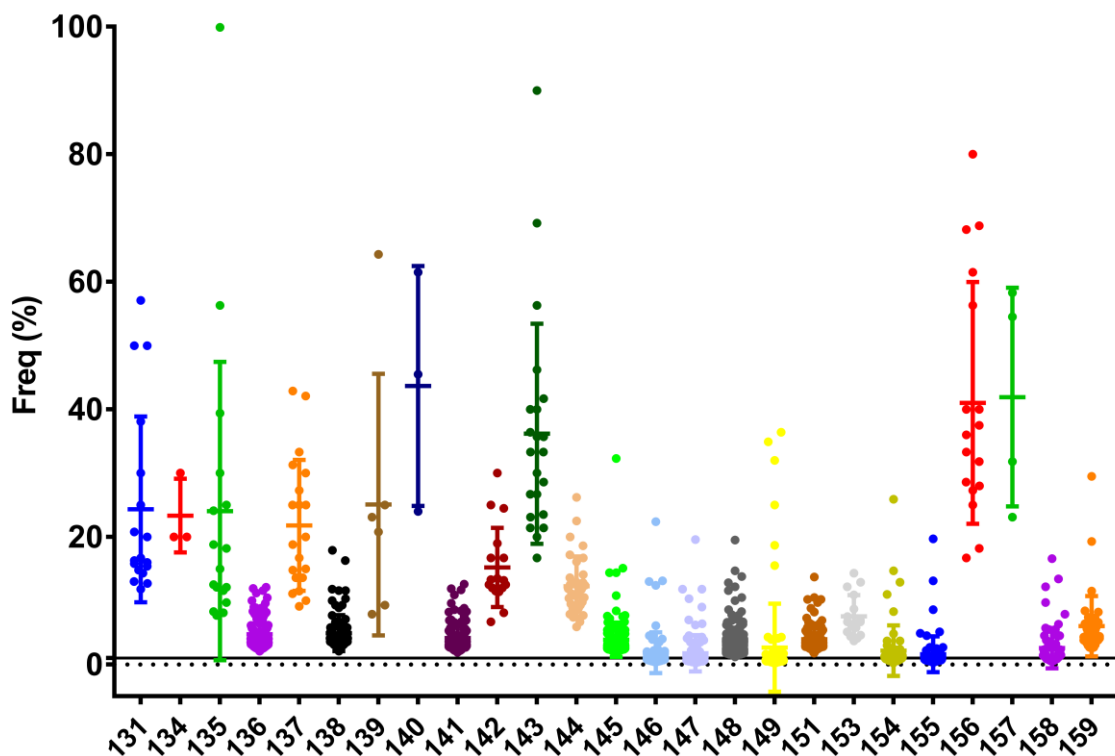
Pode observar-se que as amostras 131, 134, 135, 137, 139, 140, 143, 156 e 157 apresentaram uma frequência média superior a 20%, com valores mínimos superiores a 8%, sendo 137, 139, 140 e 143 pertencentes a casos graves (representando 80% do total dos casos graves desta amostragem), 156 e 157 a casos com sinais de alarme e 131, 134 e 135 a casos de dengue clássico. Já para as amostras 142 (dengue clássico), 144 e 153 (dengue com sinais de alarme) a frequência média resultou maior que 7% com valores mínimos superiores a 3,5%.

Estas 12 amostras apresentam as menores profundidades médias de todo o conjunto, porém, com a exceção da 143, todas elas superam a média de 100, logo, frequências de até 1% poderiam ter sido achadas.

A média da frequência intra-hospedeiro das SNV para as amostras restantes (136, 138, 141, 145, 146, 147, 148, 149, 151, 154, 155, 158 e 159) foi menor que

5%, e em termos gerais, as SNV achadas nelas estão presentes em baixas proporções, podendo atingir até 0,2%. Oito destas últimas amostras pertencem a casos de dengue clássico (141, 145, 146, 151, 154, 155, 158 e 159), 4 a casos com sinais de alarme (136, 138, 147 e 148) e 1 a um caso grave (149). As diferenças observadas não foram estatisticamente significativas quando analisadas com o teste da variância de um fator; porém, quando comparada à média da frequência intra-hospedeiro destas SNV para cada amostra com o quadro clínico, o teste “T de student” apresentou uma diferença significativa entre as médias dos casos de dengue clássico e os de dengue grave (*p-value*: 0,016).

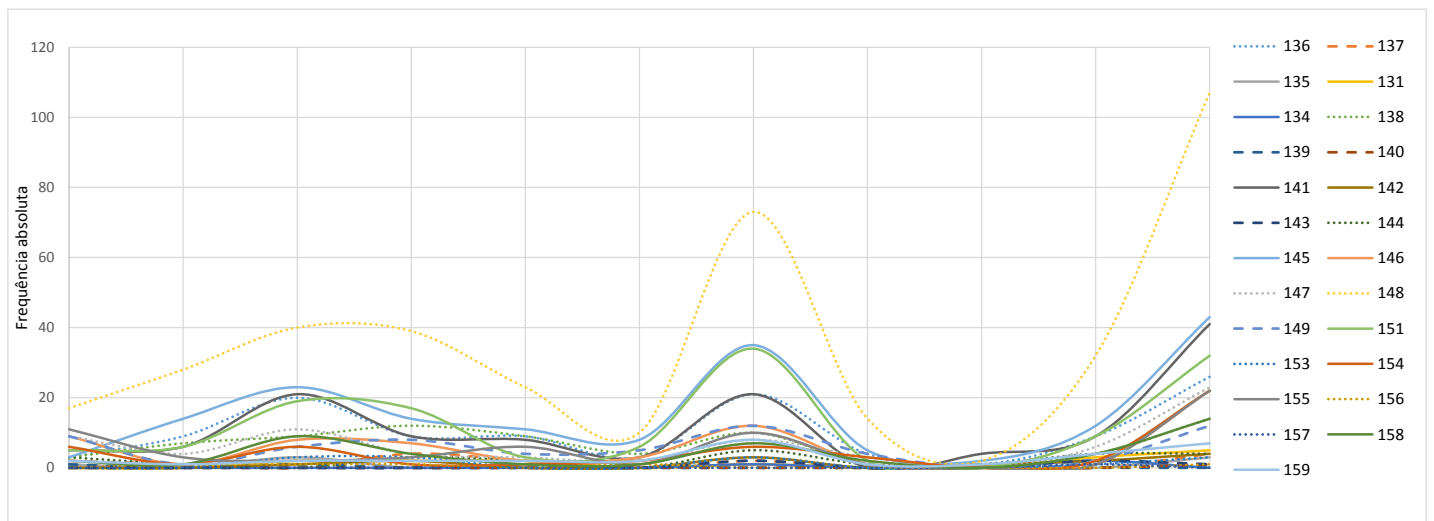
**Figura 4.4.2-2** Frequência percentual das SNV localizadas na fase aberta de leitura da poliproteína para cada amostra



As mutações não sinônimas que foram encontradas dentro da fase aberta de leitura da poliproteína, foram classificadas de acordo com o gene envolvido (Figura 4.4.2-3). Em termos gerais, o maior número de mutações foi encontrado no gene NS5, seguido do NS3 e do envelope, enquanto nos genes NS4A, NS2B e do que

codifica para as glicoproteínas da membrana, o menor. Porém, quando observada a frequência relativa destas mutações levando em conta o cumprimento de cada um destes genes, esse padrão de variabilidade muda, indicando que os genes que codificam para as proteínas do capsídeo, da membrana e da proteína 2k, apresentam uma variabilidade maior (Figura 4.4.2-4). Ao analisar esses 3 genes em particular, se destaca que para os dois primeiros, a elevada frequência ocorre principalmente entre os casos das infecções secundárias, enquanto para a proteína 2k não existe uma diferença entre as infecções primárias ou secundárias.

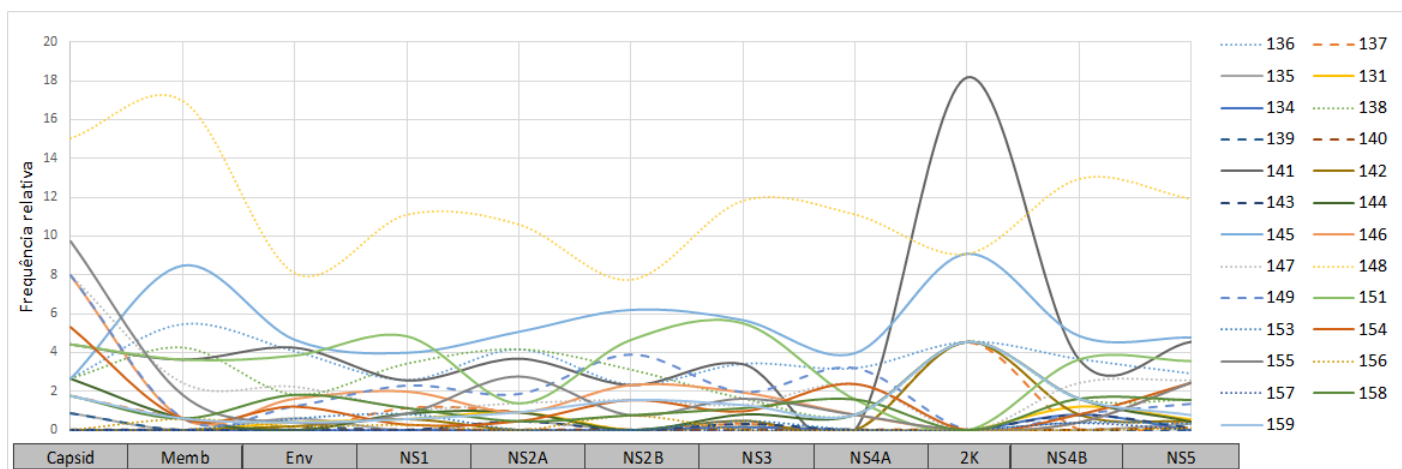
**Figura 4.4.2-3** Frequência absoluta das mutações não sinônimas segundo a sua posição no genoma



Nota: Os casos de dengue clássico estão representados nos gráficos com uma linha contínua (—), os casos de dengue com sinais de alarme com uma linha pontilhada (...) e os casos de dengue grave uma linha interrompida (- - -).



**Figura 4.4.2-4** Frequência das mutações não sinônimas relativa ao comprimento dos genes

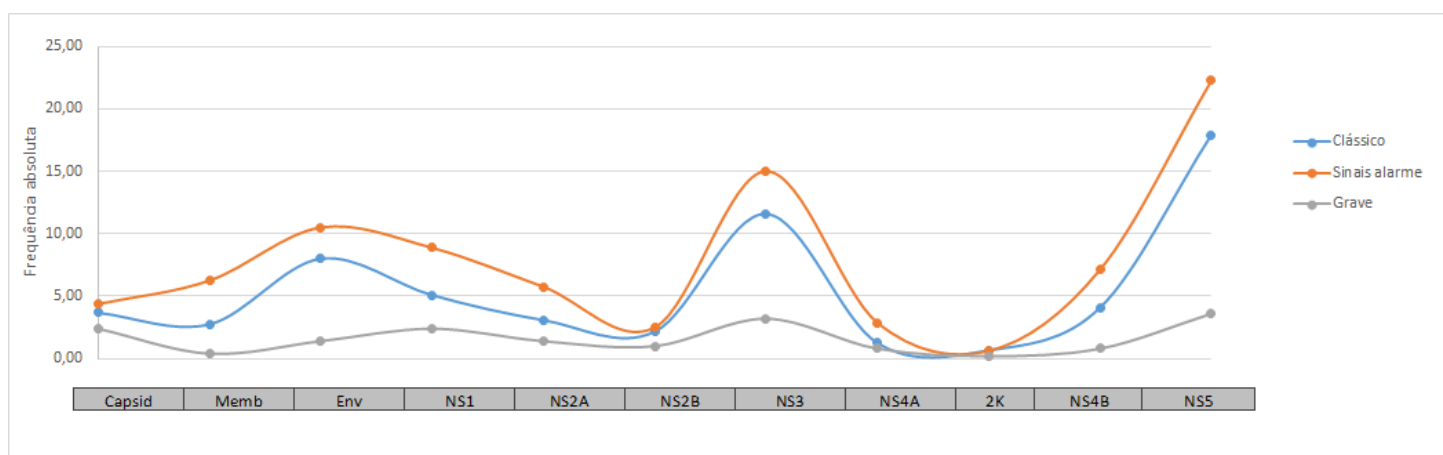


Nota: Os casos de dengue clássico estão representados nos gráficos com uma linha contínua (—), os casos de dengue com sinais de alarme com uma linha pontuada (...) e os casos de dengue grave uma linha interrompida (- - -).

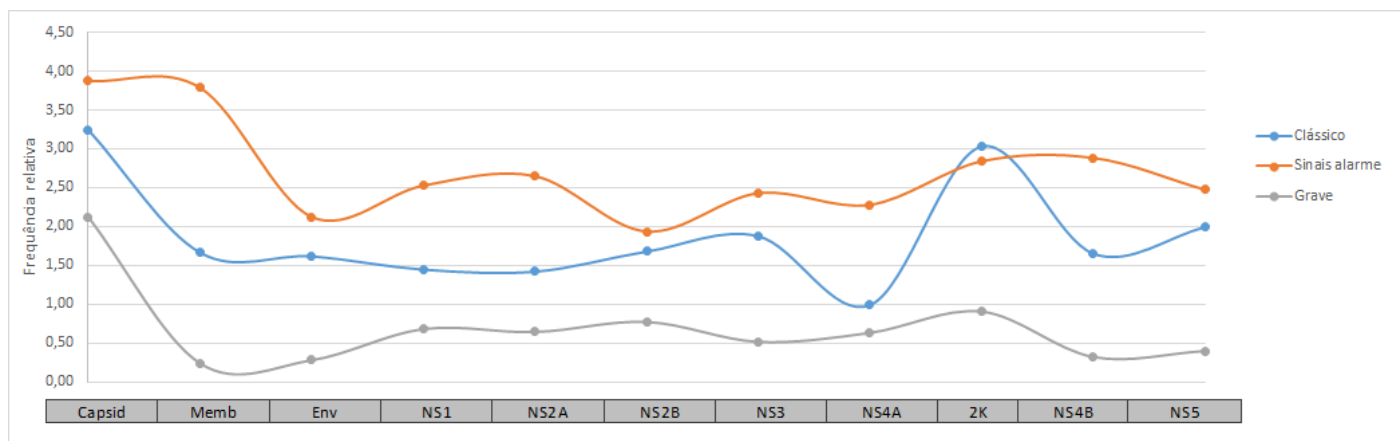
Tendo em vista a classificação clínica dos casos, foram analisadas as variantes de maneira relativa a cada grupo (Figura 4.4.2-5a), e ao tamanho dos genes (Figura 4.4.2-5b).

**Figura 4.4.2-5(a-b)** Distribuição das substituições não sinônimas ao longo do genoma segundo a classificação do quadro clínico

a)



b)



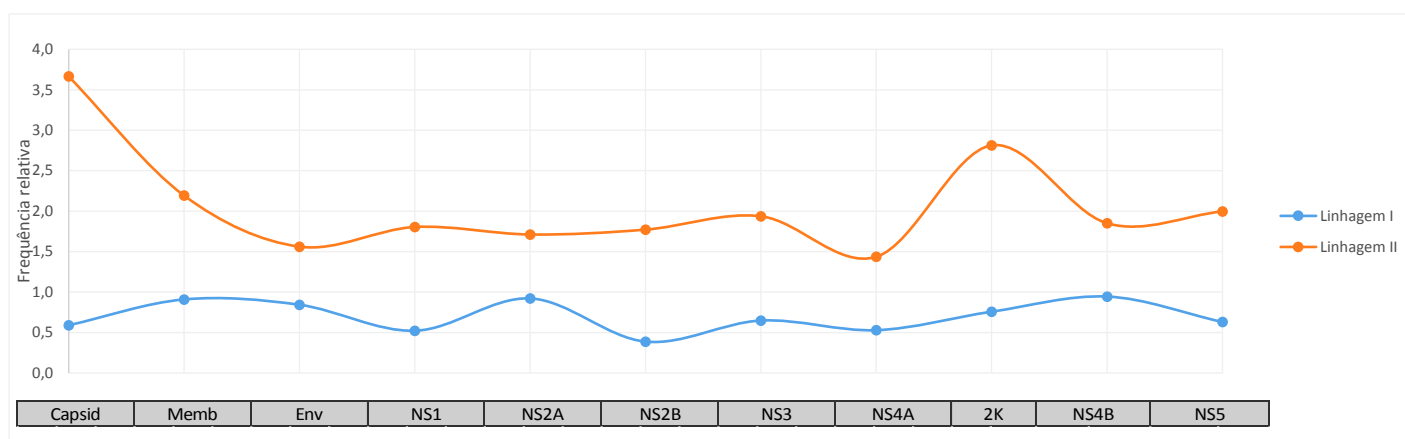
Segundo se distingue na figura 4.4.2-5(a e b), foi o grupo formado pelos casos de dengue clássico com sinais de alarme o qual apresentou a maior frequência absoluta de SNV não sinónimas, enquanto os casos de dengue grave a menor. No entanto, não há um suporte estatístico que confirme a diferença ( $p\text{-value} > 0,05$ ). Quando analisada a frequência relativa destas variantes com respeito ao comprimento dos genes, essa diferença foi mantida, porém, os padrões de variabilidade não foram exatamente iguais: os genes com maior variabilidade dentre os casos de dengue clássico foram o gene do capsídeo, e os não estruturais 2k e NS5; para os casos de dengue com sinais de alarme os genes do capsídeo, o codificante para as glicoproteínas da membrana, 2k e NS4B, e finalmente para os casos de dengue grave, somente o gene do capsídeo apresentou uma variabilidade comparativamente maior ao resto dos genes.

Da mesma forma, não foi achada uma diferença estatisticamente significativa quando se analisou a presença/ausência de substituições nas regiões 5' e 3' não codificantes para os três grupos em questão.

Por outro lado, foi estudada como se distribuíam estas SNV não sinónimas nos genes (frequência relativa ao comprimento dos genes) conforme se tratasse de uma infecção primária ou secundária. Como foi mencionado anteriormente, os casos primários apresentaram uma diversidade menor em termos de número total de SNV, porém, não foi encontrada uma diminuição paralela em nenhum dos genes em particular, sendo a variabilidade dos genes estruturais comparável com a dos genes não estruturais.

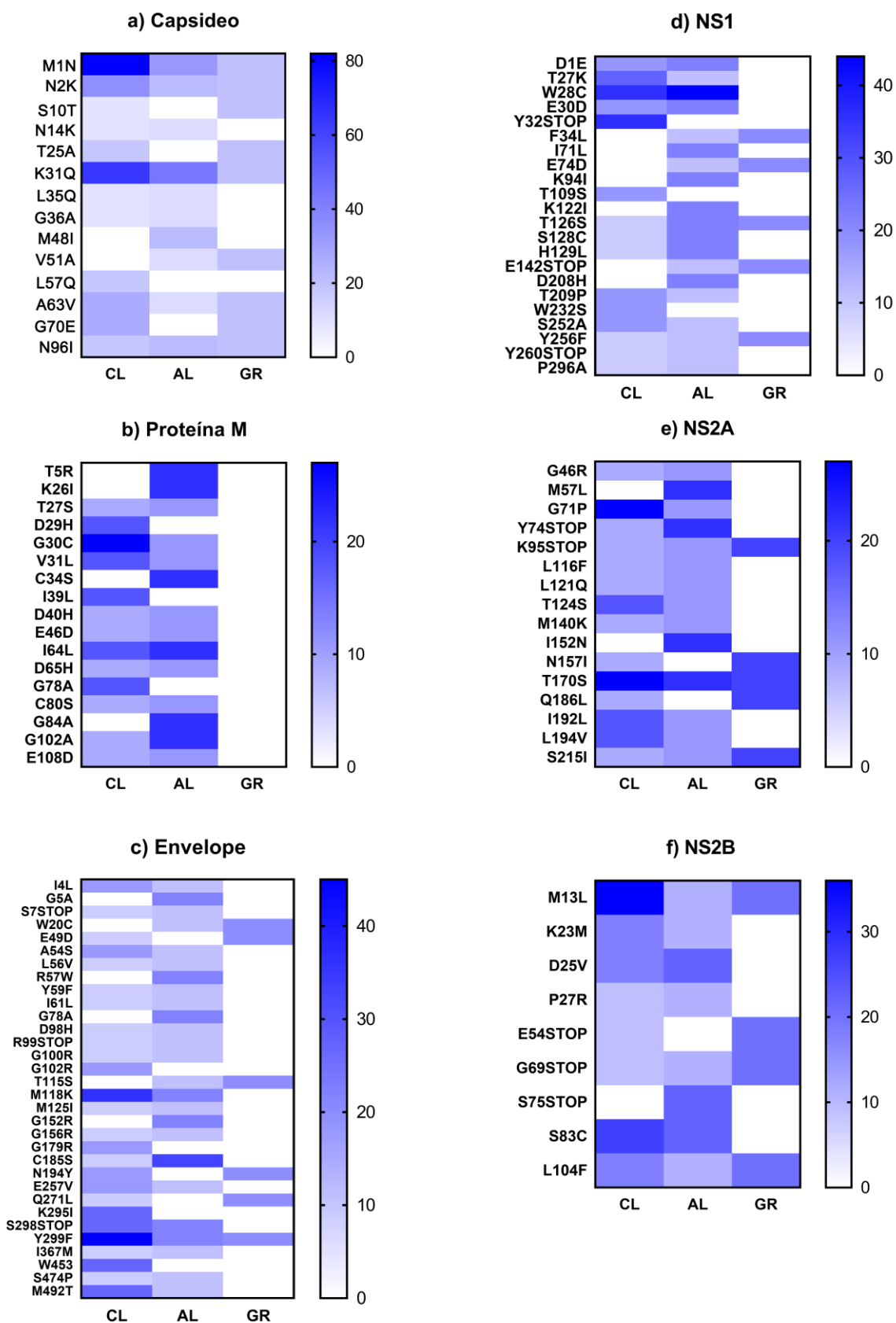
Da mesma maneira, essa abordagem foi feita com respeito às duas linhagens do genótipo asiático/americano circulantes no Brasil, como pode ser observado na figura 4.4.2-6. Existe uma notória diferença entre as variabilidades de ambas, sendo que para a linhagem I a frequência relativa das SNV não sinônimas adotou valores mais homogêneos ao longo do genoma viral, com uma pequena diferença em aumento para o gene que codifica para as glicoproteínas da membrana viral, NS2A e NS4B, enquanto para a linhagem II detectou-se uma variabilidade maior nos genes codificantes do capsídeo e das proteínas da membrana assim como também no gene codificante para a proteína 2k.

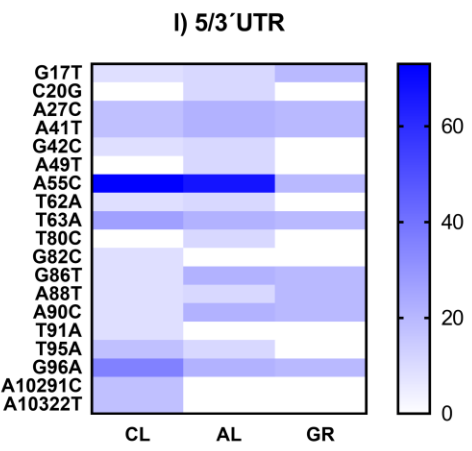
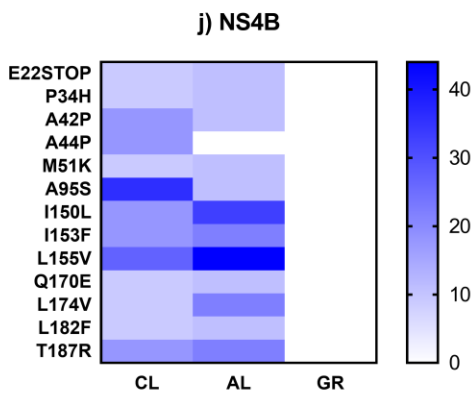
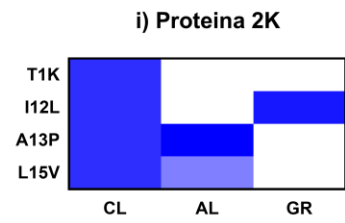
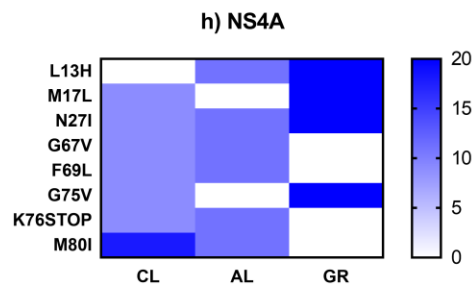
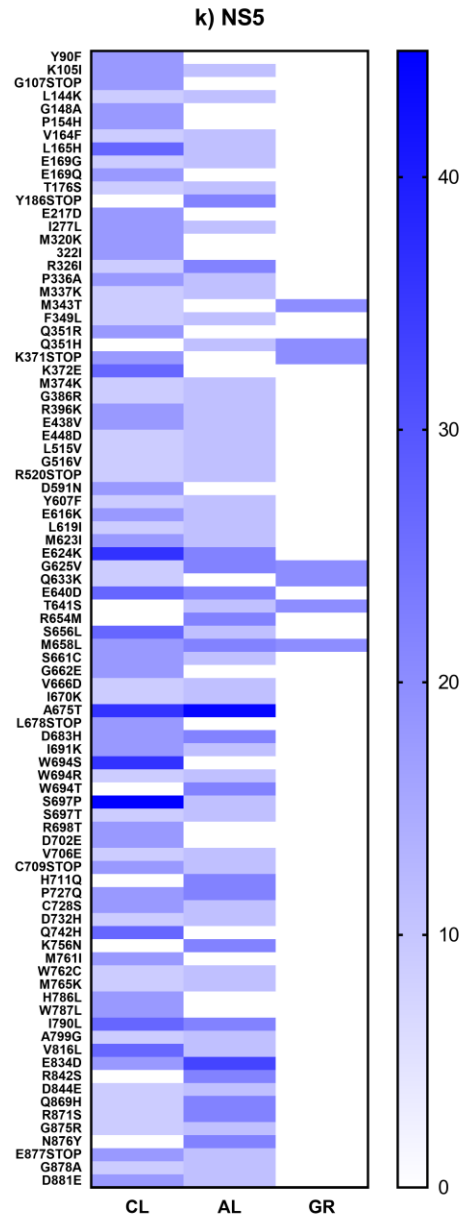
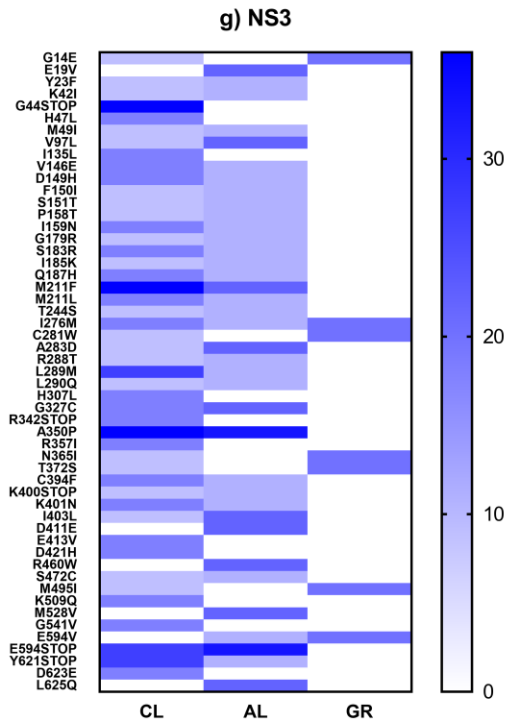
**Figura 4.4.2-6** Frequência relativa das SNV não sinônimas para cada gene com respeito à linhagem do genótipo asiático/americano



Em seguida, foram comparadas as SNV presentes em cada amostra. Curiosamente, do total das substituições diferentes achadas entre as 27 amostras (n=950), 29% (n=250) se repetiram em pelo menos duas amostras; essa frequência inter-hospedeiro varia em um intervalo de 8 a 60%. Sendo assim, tentando visualizar se existia algum padrão nessas repetições, as variantes foram classificadas por localização genômica e foi determinada a sua frequência relativa ao grupo de gravidade clínica ao qual cada amostra pertence (Figura 4.4.2-7).

**Figura 4.4.2-7** Frequência relativa das variantes compartilhadas segundo a sua localização no genoma e à classificação clínica dos casos



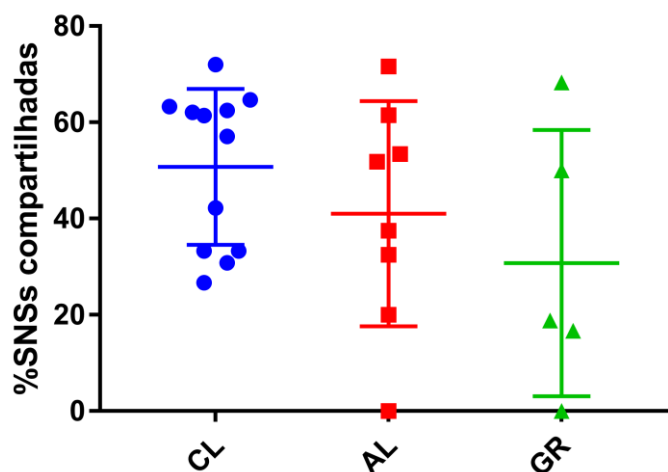


Nota: CL: dengue clássico; AL: dengue com sinais de alarme; GR: dengue grave.

De maneira geral, não foi detectado nenhum padrão específico entre a presença destas substituições e o quadro clínico apresentado pelos pacientes, porém, no grupo dos casos de dengue grave, foram observadas menores percentagens de variantes compartilhadas (figura 4.4.2-8), o que se traduziria de maneira implícita, em maiores percentagens de SNV únicas dentre essas amostras. No entanto, os valores de percentagem das variantes compartilhadas dentro deste grupo se apresentam em um intervalo amplo (0-68%). Também neste caso não existe uma medida estatística que suporte a diferença observada ( $p\text{-value}>0,05$ ).

Dentro do grupo das SNV não sinónimas, se observou que 25 delas geram na sua tradução um códon de parada (stop) prematuro na poliproteína, o que interrompe a tradução da mesma. Estes códons foram achados no gene do Envelope, e em todos os genes não estruturais. Por sua vez, foram encontrados em 70% dos casos de dengue clássico (9), 67% dos casos de dengue clássico com sinais de alarme (6) e 20% nos casos de dengue grave, percentagem que representa somente à amostra 149.

**Figura 4.4.2-8** Percentagens de variantes compartilhadas nas amostras segundo a sua classificação clínica



Nota: CL: casos de dengue clássico; AL: casos de dengue com sinais de alarme; GR: casos de dengue grave. São representados os valores individuais para cada amostra, e a média de cada grupo com o desvio padrão.

É importante notar, que a maioria das SNS compartilhadas para o grupo dos casos graves foram detectadas na amostra 149. Das quatro restantes, duas (140 e

143) atingiram profundidades médias baixas no sequenciamento profundo (tabela 4.2.4-1), o qual como se observou, tem uma associação direta com número de SNV.

Notavelmente, algumas SNV se apresentaram numa frequência entre as amostras superior ao 20%, ou seja, em pelo menos 6 das 25 que exibiram um espectro de mutações (Tabela 4.4.2-1). Foi analisada a frequência intra-hospedeiro (em cada amostra individual) em que se achou cada uma destas mutações, e observou-se que em alguns casos atinge aproximadamente 50%, como por exemplo a substituição nucleotídica A55C (região 5' não codificante) nas amostras 153, 154 e 159, ou G99T (capsídeo) na 135. Isso poderia implicar que a subpopulação portadora dessa mutação estaria presente e competindo em igual número com a subpopulação consenso majoritária. De qualquer maneira, não se conhece ainda qual é o impacto de todas estas mutações no genoma viral, na sua interação com o hospedeiro, no seu ciclo de replicação e na sua dinâmica populacional.

Por outro lado, o fato de algumas das SNV se apresentarem numa baixa frequência (<1%), reforça a observação feita inicialmente, da relação entre a profundidade média obtida no sequenciamento e o número de SNV reportados. Desta forma, não é ilógico que aquelas amostras com profundidade média baixa apresentem menos mutações.

**Tabela 4.4.2-1** Mutações de elevada frequência inter-hospedeiro

Região no genoma	Posição nucleotídica	Mudança no aa	Frequência inter-hospedeiro (%)	Frequência intra-hospedeiro (%)		
				$\bar{x}$	Mín.	Máx.
5'UTR	A41T		24,0	10,9	1,3	33,0
	A55C		60,0	31,8	13,0	51,0
	T63A		24,0	11,7	4,0	15,0
	G96A		28,0	5,4	2,0	17,0
Capsideo	G99T	M1I	52,0	18,7	7,0	56,0
	T102A	N2K	28,0	11,7	5,1	15,5
	A187C	K31Q	48,0	11,4	5,0	20,0
Envelope	T1289A	M118K	32,0	1,9	0,4	4,3
	A1832T	Y299F	24,0	5,4	3,4	9,0
NS1	G2505C	W28C	32,0	5,5	3,3	9,0
NS2A	A3985T	T170S	24,0	4,7	2,2	14,0
NS2b	A4168T	M13L	24,0	1,0	0,4	3,0
NS3	A5092T	M191L	24,0	16,8	11,0	30,0
	G5531T	R337I	28,0	4,4	2,4	7,0
	T6324A	Y601STOP	24,0	6,3	4,2	8,0
NS4B	C7288G	L155V	28,0	5,8	3,0	12,0
NS5	G9592A	A675T	32,0	10,4	6,6	15,0
	G9439A	E624K	24,0	5,6	4,0	8,0
	T9658C	S697P	24,0	1,1	0,5	3,0

Nota: aa: aminoácido;  $\bar{x}$ : média da frequência intra-hospedeiro; Mín: frequência mínima achada entre as amostras; Máx: frequência máxima achada entre as amostras; 5'UTR: região 5' não codificante.

#### 4.4.3 Mutações fixas no genoma viral

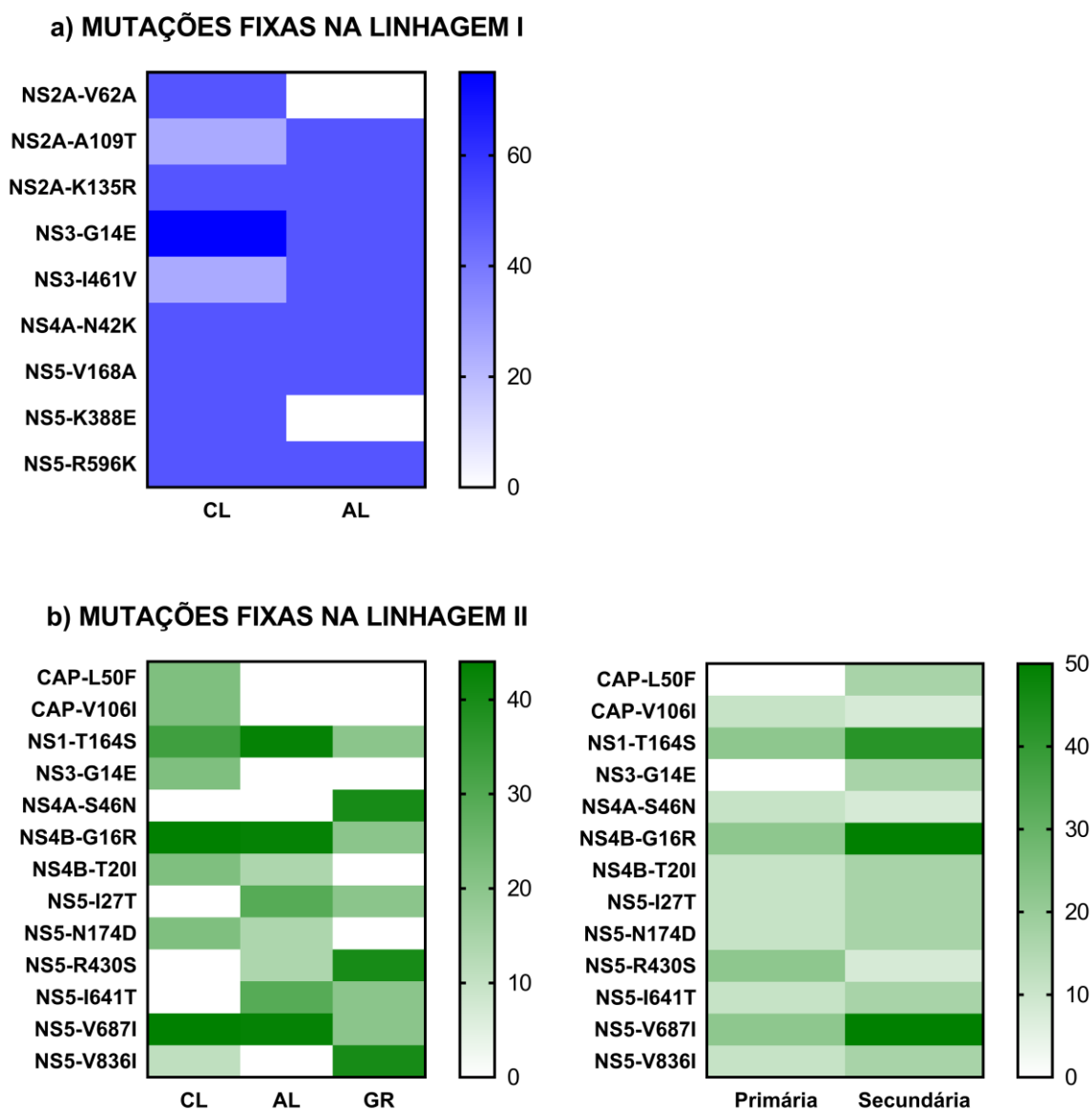
Ao alinhar as sequências consenso montadas para cada amostra com as sequências de DENV-2 obtidas do Genbank, foram detectadas mutações fixas nos genomas em questão, as quais foram confirmadas nos alinhamentos dos *reads* à sequência de referência inicialmente usada para o mapeamento dos mesmos. No apêndice III pode se achar a listagem de todas elas, qual a região do genoma viral onde ocorreram, e qual a mudança de aminoácido que geraram na proteína correspondente.

Novamente, se detectou que muitas delas se repetiam dentre as amostras, e particularmente, o conjunto de mutações das amostras correspondentes à linhagem I era consideravelmente diferente do conjunto correspondente às amostras da linhagem II. Foi analisada a relação entre a presença das mesmas e a classificação do quadro clínico, assim como também da resposta imune. Esta última, só no caso



da linhagem II, já que as amostras que pertencem a linhagem I correspondem todas a casos de infecções primárias (Figura 4.4.3-1). Para a linhagem I, é importante ressaltar que se trata somente de um grupo formado por seis casos, motivo pelo qual deve se tomar muito cuidado com as observações que podem ser feitas. Nos casos clínicos de dengue com sinais de alarme, as mutações NS2A-V62A e NS5-K388E estão ausentes, e NS3-G14E foi encontrada somente numa das amostras. As diferenças observadas não foram estatisticamente significativas quando analisadas com o teste do Chi quadrado. No caso da linhagem II, grupo composto por 21 amostras, se detectam tendências mais marcadas. Na sua relação com a classificação clínica dos casos, as mutações no capsídeo foram detectadas unicamente nos casos de dengue clássico, assim como também NS3-G14E. Por outro lado, NS4A-S46N, NS5-R430S e NS5-V836I apresentaram uma frequência notavelmente maior nos casos com dengue grave. As mutações NS5-I27T, NS5-I641T e NS5-R430S foram encontradas só nos casos de dengue com sinais de alarme ou dengue grave. Com respeito à classificação da resposta imune, todas estas mutações se encontraram em maior frequência nas infecções secundárias, com a exceção de 3: CAP-V106I, NS4A-S46N e NS5-R430S, apresentando-se estas últimas duas casualmente em casos de dengue com sinais de alarme ou grave. Porém, como já mencionado anteriormente, não se conhece ainda o impacto real da presença destas mutações na clínica da infecção. As observações efetuadas com respeito à linhagem II também não contam com suporte estatístico quando analisadas com o teste do Chi quadrado ( $p\text{-value}>0,05$ ).

**Figura 4.4.3-1** Frequência relativa das mutações fixas achadas nos genomas virais com respeito à classificação clínica e à resposta imune



Nota: CL: dengue clássico; AL: dengue com sinais de alarme; GR: dengue grave.

Analisou-se também, se alguma das mutações fixas que foram encontradas nos genomas virais, estava presente em alguma amostra diferente e anterior temporalmente, porém, como SNV. Neste caso, poderia suspeitar-se de uma transmissão efetiva da subpopulação portadora dessa mutação e a sua posterior fixação no genoma viral. Contudo, não foi identificada amostra nenhuma que

cumprisse esses requisitos, mas sim foram identificadas amostras que apresentavam alguma dessas mutações fixas como SNV (Tabela 4.4.3-1).

**Tabela 4.4.3-1** Mutações fixas nos genomas virais achadas como SNV

	Mutação					
	Fixa			SNV		
	Nº amostra	Ano	Local	Nº amostra	Ano	Local
NS2A-M57L	132	1999	Cordovil/RJ	136	2001	Vila da Penha/RJ
				144	2008	Praça seca/RJ
NS3-G14E	132	1999	Cordovil/RJ	143	2008	Praça seca/RJ
	134	2000	Jacarepaguá/RJ	154	2011	Pq.Leopoldina/Campos
	135	2001	Penha/RJ			
	136	2001	Vila da Penha/RJ			
	158	2011	Jardim carioca/Campos			
	159	2011	Pte.Vargas/Campos			
NS4B-G16R	145	2009	Pq. Dom Bosco	157	2011	Piratininga/Niteroi
	147	2010	Caju/Campos			
	148	2010	Caju/Campos			
	149	2010	Centro/Campos			
	151	2010	Goitacazes/Campos			
	153	2011	Novo campos/Campos			
	154	2011	Pq.Leopoldina/Campos			
	159	2011	Pte.Vargas/Campos			
NS5-G523S	157	2011	Piratininga/Niteroi	140	2008	Engenho da penha/RJ

## 5 DISCUSSÃO

Atualmente é indiscutível que a presença de diferentes cepas virais na população tem um impacto importante na resposta imune natural dos hospedeiros, assim como também sobre as terapias com antivirais ou vacinas. Da mesma maneira, o espectro de quasiespecies ou subpopulações geradas intra-hospedeiro impacta na resposta imune do mesmo, e na resposta que ele possa ter ante a uma estratégia terapêutica ou profiláctica. Portanto, qualquer informação sobre esses espectros de mutantes ou quasiespecies, apresenta um aspecto crítico a ser considerado no desenho de drogas e vacinas antivirais, aportando também benefícios significativos na pesquisa e conhecimento da evolução viral, a progressão da doença, a biologia e a patogênese viral (Jayasundara et al., 2015).

### 5.1 Amostragem

Por se tratar de uma temática em ativa pesquisa, com um número pequeno de estudos que a abordam (os quais por sua vez, empregam diversas metodologias e reportam medidas de diversidade intra-hospedeiro diferentes), não existe uma regra que defina qual o tamanho amostral mínimo para desenvolver este tipo de estudo, sendo que também, por cada amostra a analisar, são na verdade múltiplas as subpopulações virais envolvidas, as quais de fato são o alvo deste tipo de pesquisa. Desta forma, se considerou uma amostra de conveniência de  $n=29$ . Um fator limitante na seleção das amostras, foi o fato de dados clínicos e/ou epidemiológicos de várias amostras mais antigas não estarem disponíveis. Tratando-se de amostras que chegam ao laboratório por demanda espontânea para diagnóstico, muitas vezes as fichas são incompletas e é difícil retornar na fonte para coletar dados complementares dos casos mais antigos.

Nas infecções por dengue vírus, estima-se que o paciente permaneça virêmico por 5 a 8 dias após o início dos sintomas atingindo uma carga viral de  $10^3$  cópias/ml nas infecções primárias, enquanto pode se elevar até  $10^6$  cópias/ml nas secundárias (Johnson et al. 2005). Em nossa amostragem, observamos essa relação somente nas amostras da linhagem I, sendo todos casos primários que possuem cargas virais que atingem no máximo a ordem de  $10^3$  cópias/ml. Porém, já nos casos da linhagem II, existem casos primários que atingem valores nas ordens de  $10^4$ - $10^5$ , como por

exemplo as amostras 138 e 149. Por sua vez, dois dos casos selecionados nesta amostragem apresentaram 10 e 12 dias de sintomas, com cargas virais detectáveis e na ordem de  $10^2$  (140 e 153). Estas observações demonstram que os limites acima citados são arbitrários, e devem ser considerados cuidadosamente.

É interessante notar, que as infecções primárias com cargas virais elevadas, pertenciam a casos de dengue com sinais de alarme ou de dengue grave, o que poderia levar à suspeita da presença e seleção de subpopulações mais virulentas e patogênicas de DENV-2.

A carga viral, como mencionada anteriormente é um fator importante na seleção das amostras que serão incluídas na análise de sequenciamento profundo. No entanto, em estudos retrospectivos e com um intervalo de tempo de mais de 10 anos, conforme o proposto, é bastante difícil reunir um conjunto de amostras que preencham critérios únicos e ideais. Sendo assim, foram selecionadas também amostras que ultrapassavam 8 dias após o surgimento dos primeiros sintomas, porém, com carga viral detectável e por tratar-se de casos de interesse (casos de óbito).

Na tentativa de apresentar o maior número de fatores envolvidos nestes casos de infecção por DENV-2, foram utilizados também como critérios de seleção o bairro onde o paciente morava e o mês da infecção. Levando em consideração que a infecção por este arbovírus é multifatorial, e que não depende apenas de fatores do vírus e hospedeiro, mas também de fatores do vetor e fatores ecológicos, procuramos obter uma amostragem a mais representativa e criteriosa possível. Durante o período de tempo sob estudo, não houve registro de uma quantidade constante de casos anuais de dengue, sendo que existiram anos onde ocorreram importantes surtos localizados em diferentes pontos geográficos. Por sua vez, as populações de mosquitos provavelmente não são constantes nem homogêneas ao longo do tempo, nem a sazonalidade da doença foi estritamente associada ao verão, pois houve anos com clima mais seco que outros, contribuindo assim a uma potencial diminuição das populações do mosquito vetor (Valle et al., 2015). O mosquito vetor atinge uma vida média de 30 dias e a distância de deslocamento de vôo dos mosquitos pode atingir um máximo de 800 metros aproximadamente (Nelson, 1986). Desta maneira, foram selecionadas amostras pertencentes a casos de infecções por DENV-2 que tivessem acontecido no mesmo mês ou em meses sucessivos, e em bairros próximos entre si, para tentar evidenciar a transmissão de subpopulações virais entre os hospedeiros.

De acordo ao observado na correlação entre o sexo e a idade dos pacientes desta coorte com a classificação clínica dos casos, aumenta a proporção de mulheres conforme a gravidade da clínica da infecção. Nos estudos que descrevem fatores envolvidos na gravidade da doença por DENV-2, o gênero dos pacientes não foi um fator associado, ou de risco para o desfecho clínico (Parameswaran et al., 2012; Romano et al., 2013; Hapuarachchi et al., 2015; Chee Fu Yung et al., 2016). No entanto, no que diz respeito à idade dos pacientes, os resultados não coincidem para todos os estudos. Enquanto Hapuarachchi e colaboradores observaram maior predisposição para a doença entre os adultos jovens, Romano e colaboradores não acharam influencia nenhuma da idade sobre a gravidade da doença. Neste estudo, a média da idade dos pacientes também não diferiu entre os casos agrupados com diferentes apresentações clínicas. Porém, o fato de termos achado só crianças e um paciente idoso entre os casos graves deste estudo que ocorreram casualmente entre os anos 2007 e 2011, poderia estar vinculado com a mudança do padrão clinico-epidemiológico ocorrida no Rio de Janeiro. Como já foi descrito, a mesma trouxe associada um aumento na gravidade dos casos assim como também uma mudança da faixa etária acometida, incluindo-se desde esse então, às crianças e aos idosos (Cavalcanti et al., 2011).

## **5.2 Aspectos técnicos**

No que diz respeito à variação de carga viral observada nas amostras após o pré-processamento necessário para o sequenciamento, existem várias questões que precisam ser consideradas: I) a natureza dos diferentes controles positivos para DENV-2 empregados em ambos ensaios de RT-PCR em tempo real, sendo a concentração do controle inicial expressa em pfu/ml (cópias infectivas/ml) e a do controle empregado no ensaio “in house” expressa em cópias/ml. II) a diferença nas sensibilidades dos modelos de RT-PCR em tempo real empregados, utilizando o primeiro deles a marcação por sonda de tipo Taqman e o segundo a marcação por Sybr Green, sendo esta de conhecida menor sensibilidade. Por este motivo, não descartamos aquelas amostras que apresentaram cargas virais menores após o seu processamento. III) O tratamento com RNase H, que visa eliminar partículas de RNA ribossômico (RNAr) ou RNA carrier que possam atuar como interferentes no sequenciamento final, porém, poderia estar se eliminando também RNA viral de interesse. IV) A síntese do cDNA foi realizada empregando primers randômicos.

Conseqüentemente, não só o RNA viral de interesse foi transformado em cópias de DNA, mas também qualquer outra molécula de RNA presente na amostra, podendo levar a uma diminuição na eficiência da transcrição reversa do genoma viral.

É com efeito evidente que o pré-processamento das amostras conduz a uma diminuição da carga viral, porém, mesmo partindo de amostras com carga viral baixa, o genoma do vírus de interesse pôde ser reconstruído praticamente na sua totalidade. No entanto, o número de reads total e a profundidade média si foram fatores influenciado pela baixa carga viral, impactando de maneira indireta na detecção de variantes.

### **5.3 Análise filogenética**

A análise filogenética demonstrou que as amostras pertencentes aos casos ocorridos nos anos 1999 a 2001 agruparam dentro do cluster da linhagem I enquanto as amostras dos anos posteriores a 2007 pertenciam ao cluster da linhagem II reforçando a teoria da circulação de diferentes linhagens do genótipo asiático/americano no território brasileiro, conforme descrito anteriormente por Puiprom et al., 2011, Drumond et al., 2013, Romano et al., 2013 e Barcelos Figueiredo et al., 2014. Não obstante, pelo fato de termos encontrado amostras provenientes do mesmo período e região geográfica agrupando separadamente entre si, como foi o caso das sequências pertencentes aos surtos dos anos 2010 e 2011 no município de Campos dos Goytacazes e das sequências do município de Niterói pertencentes aos casos de 2011, entre outras, suspeitamos que tenham ocorrido diversas introduções de cepas provenientes de diferentes locais, dado o caráter hiperendêmico do dengue no Brasil, da presença do mosquito em todas as áreas e as facilidades de locomoção que existem hoje em dia para o homem. Esta hipótese precisará ser confirmada em um estudo futuro utilizando-se uma análise de coalescência.

### **5.4 Análise da diversidade viral**

O emprego de índices de diversidade é uma boa abordagem para quantificar e comparar a diversidade microbiana entre as amostras (Bonilla-Rosso et al., 2012), sendo neste caso a diversidade microbiana representada pela diversidade das

subpopulações virais ou espectro de mutantes presentes nas amostras. Existem diversos índices que medem a diversidade genética e são ferramentas fundamentais para definir as quasiespecies virais em nível molecular. Os comumente usados para comparar a diversidade das quasiespecies virais são a frequência de mutação mínima ( $M_f \text{ min}$ ), a frequência de mutação máxima ( $M_f \text{ max}$ ) a entropia de Shannon normalizada ( $S_n$ ) e a diversidade nucleotídica ou índice  $P_i$  ( $\pi$ ). Seria interessante poder contar com a informação de vários índices de diversidade, mas a estratégia de sequenciamento empregada, com a síntese de *reads* no sequenciamento de 35 pares de bases, limita a reconstrução dos haplotipos<sup>5</sup> que representariam cada subpopulação presente nas amostras. Do ponto de vista bioinformático, existem ferramentas que seriam úteis para essa finalidade, porém, o erro que acompanharia tal cálculo seria o suficientemente grande para perder a certeza sobre o haplotipo montado. Desta forma, escolhemos o índice da entropia de Shannon como medida da complexidade genética das amostras e o índice  $P_i$  da diversidade nucleotídica, como parâmetros para a caracterização do espectro de mutantes. O cálculo dos mesmos pode ser adaptado para cada posição nucleotídica e posteriormente, mediante o cálculo da sua média, à população viral presente em cada amostra. A avaliação da complexidade da população viral em relação ao impacto biológico, deve-se considerar cuidadosamente, pois o perfil da diversidade das quasiespecies virais não é um parâmetro constante (Gregori et al., 2016). Desta maneira, a análise aqui realizada da diversidade viral só estaria representando um momento definido e pontual na história da infecção por DENV-2 destes pacientes.

Poucos são os estudos que, até então, focaram em analisar a diversidade intra-hospedeiro de DENV, e ainda com consideráveis limitações: ou foram poucas as sequências analisadas, ou apenas determinados genes os estudados, ou limitadas as metodologias empregadas (Lin et al., 2004; Descloux et al., 2009; Romano et al., 2013; Drummond et al., 2013; Takeshi et al., 2014). Nestes estudos, uma ampla variedade de resultados foi obtida para a diversidade intra-hospedeiro, tanto em humanos, como em mosquitos (Lin et al., 2004; Lequime et al., 2015; Lambrechts & Lequime, 2016).

Os primeiros estudos de diversidade intra-hospedeiro de DENV reportavam valores de diversidade nucleotídica na ordem de  $10^{-3}$ , e se baseavam em técnicas de amplificação por PCR e clonagem seguido de sequenciamento pela técnica de

---

<sup>5</sup> Haplotipos: sequências genômicas com idêntico padrão de mutações (Gregori et al., 2016).



Sanger. Thai e colaboradores, analisando somente o gene do envelope, e da mesma maneira Parameswaran e colaboradores estudando o genoma completo do vírus, incorporaram nas suas análises algoritmos que minimizavam a contribuição dos erros introduzidos pelo próprio processamento, e acharam uma diversidade nucleotídica na ordem  $10^{-5}$ . Estes autores sugerem que estudos prévios poderiam estar representando medidas de diversidade nucleotídica sobrestimadas. Além disso, a diversidade neste tipo de estudos é calculada de maneira simples como o número de substituições/comprimento do fragmento x 100, não levando em consideração o número de clones analisados ou mesmo dizer a profundidade atingida. Já com os aportes do sequenciamento de nova geração, mediante os quais se obtêm coberturas amplamente maiores para cada posição nucleotídica, em conjunto com algoritmos bioinformáticos que dão lugar à correção de erros na detecção de variantes, Romano e colaboradores reportaram uma diversidade média intra-hospedeiro de  $4,82 \times 10^{-5}$  (considerando somente SNV de frequência >1%). Comparativamente, neste estudo obteve-se uma diversidade nucleotídica média representada pelo índice  $\pi$ , de  $6,48 \times 10^{-4}$ . Essa diferença de uma ordem poderia ser explicada pelas metodologias empregadas nestes estudos visto que a cobertura para cada posição nucleotídica foi menor que a obtida no nosso estudo. Seja pela limitação metodológica ou pelo valor de corte escolhido, estes autores só reportaram SNV com frequência > 1%. Acreditamos que o valor da diversidade reportado por nosso estudo represente com maior fidelidade à diversidade das populações virais dentro de um hospedeiro.

Descloux e colaboradores foram pioneiros em correlacionar a diversidade viral intra-hospedeiro de DENV-1 com a evolução clínica da doença, encontrando uma menor diversidade intra-hospedeiro em pacientes com dengue grave, quando comparados com quadros clínicos clássicos (Descloux et al., 2009). Estes resultados coincidem com a tendência observada neste estudo para DENV-2, no qual os casos de dengue grave apresentaram menor diversidade nucleotídica, complexidade da população viral e número total de SNV comparados com os casos de dengue clássico com ou sem sinais de alarme. Não obstante, estas diferenças não contam com suporte estatístico. Estudos mais recentes, no entanto, não conseguiram demonstrar nenhuma associação entre as medidas de diversidade intra-hospedeiro e a gravidade da doença (Thai et al., 2012; Parameswaran et al., 2012).

Rodriguez-Roche e colaboradores estudaram a diversidade genética do DENV-3 em pacientes, durante uma epidemia ocorrida entre os anos 2001 e 2002

em Cuba. Embora se trate de um sorotipo diferente, e de um cenário epidemiológico também diferente ao nosso, eles detectaram uma variabilidade maior nos pacientes que apresentavam infecções secundárias. Neles, era maior o número de SNV-SNS nos genes estruturais (PrM e ENV), enquanto aqueles com resposta primária apresentavam a maioria das SNV-SNS nos genes não estruturais (Rodriguez-Rocha et al., 2016). Takeshi e colaboradores, pelo contrário, observaram uma diversidade média maior para os pacientes com infecções primárias. No entanto, este último resultado deve se considerar de maneira cuidadosa, pois eles analisaram apenas o gene do envelope (Takeshi et al., 2013). Os resultados obtidos neste estudo se assemelham mais com os achados de Rodriguez-Rocha e colaboradores, observando-se um número de SNV significativamente menor nos casos de infecções primárias quando comparadas com as secundárias. Assim mesmo, porém, sem suporte estatístico, se observou maior variabilidade nos genes estruturais para as infecções secundárias, enquanto que as primárias não mostraram uma tendência clara na diversidade seletiva de genes estruturais ou não estruturais. Segundo Domingo e colaboradores, a vantagem de um amplo espectro de mutantes reside na maior capacidade de que a população viral encontre no espaço de todas as sequências, regiões do genoma nas quais poder aumentar o *fitness* (Domingo et al., 2012). Deve se considerar a possibilidade de que a pressão imunológica exercida pelos anticorpos heterólogos nas infecções secundárias atue como um fator indutor do aumento da diversidade intra-hospedeiro, como foi observado nos resultados deste estudo.

No seu estudo da diversidade intra-hospedeiro de DENV-2 em casos de infecções secundárias de pacientes na Nicaragua, ocorridos durante o período 1998-2012, Parameswaran e colaboradores reportaram um agrupamento filogenético das suas amostras em dois clados bem diferenciados, os quais particularmente apresentavam uma diversidade diferente nos genes codificantes da proteína do Envelope e das proteínas não estruturais NS4A e NS4B (Parameswaran et al., 2012). Não obstante, apesar de também serem analisadas no nosso estudo casos de infecções primárias, o fato destes autores terem encontrado uma diversidade intra-hospedeiro diferencial entre cepas virais de DENV-2 filogeneticamente distantes, é comparável com os achados neste estudo para as linhagens I e II. Neste estudo encontrou-se uma diferença entre os genes apresentando maior variabilidade na linhagem II com respeito à linhagem I. Tanto no nosso caso como no estudo de Parameswaran e colaboradores, estes resultados devem ser revistos

cautelosamente, pois a capacidade par fazer tais associações fica limitado devido ao tamanho amostral para cada cluster ou linhagem correspondente.

Com respeito às SNV-SNS achadas consistentemente entre várias das amostras analisadas, podem sugerir-se diferentes explicações: I) sítios variáveis comuns entre as populações podem ser consequência de mutações “de novo” em sítios de menor restrição, ou também chamados “hotspots”. II) possivelmente, as subpopulações menores não sejam eliminadas de maneira imediata pelo efeito da deriva gênica e possam ser mantidas durante a transmissão a um novo hospedeiro, independentemente do seu *fitness* e até por longos períodos de tempo (Romano et al., 2031). O fato dessas SNV serem compartilhadas principalmente pelas amostras dos casos de dengue clássico com ou sem sinais de alarme, sugeriria que as mesmas não estariam envolvidas nos processos da patogenia grave da infecção por DENV. Os mecanismos que levam aos diferentes padrões de diversidade genética intra-hospedeiro de DENV-2 e a sua relação com a gravidade clínica permanecem ainda indeterminados. Da mesma maneira que no estudo de Descloux e colaboradores, não sabemos se as diferenças observadas são a causa ou a consequência da gravidade da doença. É preciso desenvolver uma análise mais aprofundada de variação viral e do impacto dela dentro do hospedeiro.

Um dos parâmetros críticos das quasiespecies virais é o número de mutações num vírus RNA que se necessitam para se obter uma mudança fenotípica no mesmo. Evidências experimentais sugerem que somente uma ou poucas substituições de aminoácidos são suficientes para modificar as características biológicas de um vírus. Logo, mutações fenotípicas relevantes podem ser frequentemente representadas no espectro de mutantes, e adquirir o potencial de converter-se em dominantes em resposta às demandas do meio-ambiente. Paradoxalmente, a presença de um mutante com uma característica fenotípica específica pode não ser suficiente para garantir a sua posição dominante, mesmo quando uma restrição seletiva favoreça essa característica fenotípica. Isso resulta em uma consequência inevitável da presença do espectro de mutantes em torno à mutante relevante. De maneira geral, o comportamento de uma subpopulação menor acaba influenciado pelas outras subpopulações ao redor dela (Domingo et al., 2006). Em função destas considerações, e do amplo espectro de frequências nas quais se encontraram todas as SNV, deverá analisar-se, junto com o impacto de cada SNV nas proteínas virais, a frequência com que cada uma delas se apresenta, para determinar finalmente se existem SNV que possam atuar como fatores de

virulência, ou permitirem mudanças adaptativas e vantajosas, porém estejam camuflados entre o espectro de todas as variantes.

Vírus defeituosos, infecciosos ou não, podem interferir com a evolução viral, porém, a transmissão a longo prazo de cepas com códons “stop” já tem sido descrita dentro dos seres humanos e mosquitos infectados com DENV-1. A transmissão a longo prazo destes vírus defeituosos na natureza foi conseguida através de complementação por co-infecção de células hospedeiras com vírus funcionais (Myat Thu et al., 2005). O genoma defeituoso geralmente apresenta uma afinidade maior que o não defeituoso por uma proteína viral ou do hospedeiro, ou por outros componentes do hospedeiro (Domingo et al., 2006; Li et al., 2011). Neste estudo a presença de mutações que dariam lugar a uma poliproteína truncada, e em decorrência a genomas defectivos, foi frequentemente detectada entre as amostras. No entanto, a variante selvagem para cada uma dessas mutações também estava presente. Fica como uma incógnita se esses genomas defectivos foram produto de mutações “de novo” ou foram efetivamente transmitidos. Novos estudos serão necessários para elucidar qual a vantagem evolutiva que esta transmissão pode conceder aos DENV.

As quasiespecies virais possuem no seu espectro de variantes, genomas em proporções minoritárias que refletem aqueles que foram dominantes em fases anteriores da evolução destas quasiespecies. Esta memória nas quasiespecies virais atua proporcionando um reservatório molecular capaz de reagir rapidamente ante uma restrição seletiva que foi previamente experimentada pela mesma população (Domingo et al., 2006). Portanto, quando analisadas as mutações fixas nas sequências consenso representativas das amostras deste estudo, a presença das mesmas em amostras de anos posteriores e em qualidade de SNV poderia estar representando subpopulações de memória, que evidentemente foram dominantes nesses anos anteriores, e perderam posteriormente o seu domínio ante algum evento de tipo *bottleneck*. Curiosamente, quando as mutações fixas foram detectadas, esperava-se achar que as mesmas fossem consequência de antigas SNV que tivessem sido selecionadas e fixadas, e não no sentido inverso.

Tendo em vista que a produção de vacinas assim como de drogas antivirais é baseada no conhecimento prévio das diferentes etapas no processo de interação vírus-hospedeiro, e no processo de replicação viral, a presença de diferentes cepas virais na população vai alterar a resposta imune natural do hospedeiro, como também à terapia com antivirais e se coloca como desafio para as vacinas em

estudo. Portanto, qualquer informação sobre essas cepas, ou espectro de quasiespecies, é de crítica importância no desenho de drogas e vacinas antivirais, gerando também benefícios significativos na pesquisa e conhecimento da evolução viral, a progressão da doença, a biologia e na patogênese viral (Jayasundara et al., 2015). Estudos de diversidade genética viral deveriam ser incluídos nos programas de vigilância da dengue, uma vez que a avaliação da eficácia da vacina deve considerar diferentes situações epidemiológicas conjuntamente com a ampla variabilidade do DENV, que pode se apresentar com variantes de distinta virulência e *fitness* viral segundo o cenário, tanto durante epidemias como no contexto endêmico.

O desenho do nosso estudo contou com a limitação de não incluir experimentos que demonstrem efetivamente o efeito de subpopulações particulares na clínica e patogenia da infecção por DENV-2, porém, acredita-se que as evidências obtidas neste estudo e o anteriormente descrito por outros autores reforcem o efeito da diversidade do DENV-2 intra-hospedeiro sobre a clínica da doença.

## 6 CONCLUSÕES

- A estratégia experimental utilizada neste estudo, quando aplicada a amostras de baixa carga viral, influencia negativamente na análise da diversidade.
- Na amostragem estudada, as infecções secundárias apresentaram uma carga viral mais elevada e maior número de variantes (SNV) em comparação com as infecções primárias.
- Casos de dengue com clínicas mais graves apresentaram menor diversidade viral. Assim, subpopulações virais mais virulentas poderiam ter menos vantagens adaptativas.
- SNV-SNS encontradas nos casos de dengue “clássico” também foram encontradas nos casos de dengue com sinais de alarme, mas não nos casos de dengue grave. A ausência dessas nos casos graves poderia ser um marcador indicativo de progressão à gravidade.
- A frequência média das SNV intra-hospedeiro, foi significativamente maior nos casos de dengue grave do que nos casos de dengue clássico, o que reafirma a maior homogeneidade observada nas subpopulações virais dos casos de clínica mais grave.
- Neste estudo, todas as amostras analisadas corresponderam ao genótipo asiático/americano de DENV-2,
- Confirmamos a circulação de duas linhagens no estado do Rio de Janeiro. A linhagem I, nas amostras correspondente ao período 1999 a 2001 e a linhagem II, nas amostras posteriores ao ano 2007.
- Neste estudo, a linhagem I apresentou uma frequência intra-hospedeiro menor de SNV-SNS, em comparação com a linhagem II. Os genes

submetidos à maior variabilidade para a linhagem I incluíam genes estruturais e não estruturais, enquanto para a linhagem II foram apenas os estruturais.

- Diferentes padrões de mutações fixas nos genomas virais foram identificados nas duas linhagens. Se desconhece ainda se essas mutações possam ter tido algum efeito no curso da infecção. Algumas dessas mutações, foram encontradas em outras amostras posteriores na escala temporal, apresentando-se como SNV.

## 7 PERSPECTIVAS

Até hoje, apenas poucos estudos pesquisaram a dinâmica da diversidade genética intra-hospedeiro do DENV. A maioria deles assim como o nosso, utilizou uma abordagem descritiva para correlacionar a extensão da diversidade genética viral dentro do hospedeiro com características epidemiologicamente relevantes. No entanto, novas perguntas surgem com a análise dos resultados obtidos neste estudo.

Foram observadas diversas tendências entre os comportamentos das variáveis sob estudo, porém que necessitam de suporte estatístico. Surge como consequência a incógnita de se essas tendências representam realmente a realidade ou são meramente um efeito do acaso. Desta forma, seria interessante poder ampliar a nossa amostragem. Casualmente, o mesmo fenômeno foi observado em outros estudos, dado que as amostragens deste tipo de estudo costumam ser baixas, como mencionado anteriormente.

As análises da evolução viral intra-hospedeiro precisam contar com os padrões observados de diversidade genética para inferir os processos evolutivos subjacentes. As forças evolutivas que agem dentro das populações virais intra-hospedeiro não são fundamentalmente diferentes daquelas agindo em qualquer população (Pybus & Rambaut, 2009). Elas incluem mecanismos genéticos (ou seja, mutação e recombinação), processos demográficos (isto é, a deriva genética) e seleção natural. Logo, é importantíssimo que os resultados apresentados neste trabalho sejam complementados por estudos que avaliem as forças evolutivas que estão agindo sobre estas subpopulações virais. Desta maneira, análises da evolução do DENV dentro do hospedeiro terão o potencial para facilitar o entendimento dos processos que acontecem dentro dos hospedeiros, assim como também fornecer importantes conhecimentos sobre a dinâmica epidemiológica desta virose.

Foram achadas neste estudo um total de 950 SNV não sinônimas distintas que mapeiam na fase aberta de leitura da poliproteína, e 39 SNV diferentes nas regiões 5 e 3' não codificantes. É fundamental realizar uma análise mais abrangente do impacto dessas variações menores aqui apresentadas. Será que alteram a estrutura e em decorrência a funcionalidade da proteína viral em questão? Aumentarão ou diminuirão a afinidade por fatores virais ou do hospedeiro, facilitando a replicação do vírus nas células? Terão um impacto na resposta imune? Qual o



impacto das mutações nas regiões não codificantes? Para responder a estas questões é indispensável identificar detalhadamente se estas variações observadas dentro do espectro de mutantes criam alterações nas estruturas secundárias e terciárias das proteínas, para inferir posteriormente mudanças de função. Da mesma maneira, precisa ser estudado minuciosamente o efeito que as mutantes achadas poderiam exercer sobre as estruturas das regiões não codificantes do genoma viral que intervêm na replicação e na modulação da tradução dos genomas virais. Um aspecto crítico para a pesquisa no futuro será dar o passo a partir dos padrões de diversidade genética do DENV à análise das características fenotípicas subjacentes ao *fitness* dessas populações previamente detectadas, avaliando por exemplo a resistência aos medicamentos, a taxa de replicação em cultura, entre outras. Isso irá requerer da associação de análises robustas da evolução viral intra-hospedeiro com experimentos cuidadosamente desenhados de genética reversa viral.

Apenas conhecendo e compreendendo verdadeiramente o comportamento do DENV, seremos capazes de diminuir o impacto dessa infecção.

## 8 REFERÊNCIAS BIBLIOGRÁFICAS

Afgan E, Baker D, Van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. ***Nucleic Acids Research***.

Agilent. 2016. Agilent High Sensitivity DNA Kit Guide, presente em: [http://www.agilent.com/cs/library/usermanuals/Public/G2938-90321\\_SensitivityDNA\\_KG\\_EN.pdf](http://www.agilent.com/cs/library/usermanuals/Public/G2938-90321_SensitivityDNA_KG_EN.pdf)

Anderson CR & Rico-Hesse R. 2006. Aedes aegypti vectorial capacity is determined by the infecting genotype of dengue virus. ***Am J Trop Med Hyg***. 75(5):886-92.

Angel RMD, Valle JRd. 2013. Dengue Vaccines: Strongly Sought but Not a Reality Just Yet. ***PLoS Pathog***. 9(10): e1003551.

Araújo JM, Nogueira RM, Schatzmayr HG, Zanotto PM, Bello G. 2009. Phylogeography and evolutionary history of dengue virus type 3. ***Infect Genet Evol***, 9(4), 716-25.

Ashburn PM & Craig CF. 1907. Experimental Investigations Regarding the Etiology of Dengue. ***J Infect Dis***. 4:440–75.

Avirutnan P, Fuchs A, Hauhart RE, et al. Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. 2010. ***J Exp Med***. 207(4):793-806.

Bancroft, TL. 1906. On the etiology of dengue fever. ***Aust. med. Gaz***. 25:17-18.

Barcelos Figueiredo L, Sakamoto T, Leomil Coelho LF, de Oliveira Rocha ES, Gomes Cota MM, et al. 2014. Dengue Virus 2 American-Asian Genotype Identified during the 2006/2007 Outbreak in Piauí, Brazil Reveals a Caribbean Route of Introduction and Dissemination of Dengue Virus in Brazil. ***PLoS ONE***. 9(8): e104516.

Bonilla-Rosso G, Eguiarte LE, Romero D, Travisano M, Souza V. 2012. Understanding microbial community diversity metrics derived from metagenomes: performance evaluation using simulated data sets. ***FEMS Microbiol Ecol***. 82, 37–49.

BRASIL, 2014. Ministério da Saúde. Incidência de Dengue. Brasil, Grandes Regiões e Unidades Federadas, 1990 a 2014. Disponível em:

<http://portalsaude.saude.gov.br/images/pdf/2015/julho/29/Incidencia-dengue-at-2014.pdf>

Carrillo-Valenzo E, Danis-Lozano R, Velasco-Hernández JX, Sánchez-Burgos G, Alpuche C, López I, et al. 2010. Evolution of dengue virus in Mexico is characterized by frequent lineage replacement. *Arch Virol*, 155(9), 1401-12.

Cavalcanti LP, et al. 2011. Change in age pattern of persons with dengue, Northeastern Brazil. *Emerging Infectious Diseases*, 17: 132-134.

CELS. 2004. College of the environment and life sciences. Center for Biotechnology and Life Sciences, Kingston. Presente em: <http://cels.uri.edu/gsc/cndna.html>

Chambers TJ, Hahn CS, Galler R, Rice CM et al. 1990. Flavivirus genome organization, expression and replication. *Annu Rev Microbiol*. 44:649-688.

Chan M & Johansson MA. 2012. The incubation periods of dengue viruses. *Plos ONE*. 7(11): e50972.

Chao D. 2012. Inter- and Intra-Host Evolution of Dengue Viruses and the Inference to the Pathogenesis. Em InTech, *Viral Genomes - Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus Interactions*, p. 195-218.

Chee Fu Yung, Siew Pang Chan, Tun Linn Thein, Siaw Ching Chai, Yee Sin Leo. 2016. Epidemiological risk factors for adult dengue in Singapore: an 8-year nested test negative case control study. *BMC Infectious Diseases*. 16: 323.

Chen R & Vasilakis N. 2011. Dengue—Quo tu et quo vadis?. *Viruses*. 3(9), 1562-1608

Cleland JB, Bradley B, McDonald W. 1919. Further experiments in the etiology of dengue fever. *J Hyg*. 18:217.

Cornman RS, Boncristiani H, Dainat B, Chen Y, Van Engelsdorp D, Weaver D, et al. 2013. Population-genomic variation within RNA viruses of the Western honey bee, *Apis mellifera*, inferred from deep sequencing. *BMC Genomics*. 14: 154.

Cruz ACR, et al. 2010. Epidemiologia molecular dos sorotipos 2 e 3 do vírus dengue isolados no Brasil de 1991 a 2008. *Rev Pan-Amaz Saude*, 1(3):25-34.

Descloux E, Cao-Lormeau VM, Roche C, De LX. 2009. Dengue 1 diversity and microevolution, French Polynesia 2001-2006: connection with epidemiology and clinics. *PLoS Negl. Trop. Dis*. 3:e493.

Dias CN & Moraes DF. 2014. Essential oils and their compounds as *Aedes aegypti* L. (Diptera: Culicidae) larvicides: review. *Parasitol Res*. 113: 565–92.

Domingo E & Schuster P. 2016. Quasispecies: From Theory to Experimental Systems. ***Curr Top in Microbiol Immunol***. DOI 10.1007/978-3-319-23898-2.

Domingo E, et al. 2012. Viral quasispecies evolution. ***Microbiol Mol Biol Rev***. 76, 159–216

Domingo E, Martín V, Perales C, Grande-Pérez A, García-Arriaza J, Arias A.2006. Viruses as Quasispecies: Biological Implications. ***Curr Top Microbiol Immunol***. 299:51–82.

dos Santos, FB; Nogueira, FB; Castro, MG; Nunes, PC; de, Filippis AM; Faria, NR; et al. First report of multiple lineages of dengue viruses type 1 in Rio de Janeiro, Brazil. ***Virol J.***, 2011, 8, 387.

Drumond BP, Mondini A, Schmidt DJ, Bronzoni RVdM, Bosch I, et al. 2013. Circulation of Different Lineages of Dengue Virus 2, Genotype American/Asian in Brazil: Dynamics and Molecular and Phylogenetic Characterization. ***PLoS ONE***. 8(3): e59422.

Drumond, BP; Mondini, A; Schmidt, DJ; Bosch, I; Nogueira, ML. Population dynamics of DENV-1 genotype V in Brazil is characterized by cocirculation and strain/lineage replacement. ***Arch Virol.***, 2012, 157(11), 2061-73.

Duong V, Simmons C, Gavotte L, Viari A, Ong S, Chantha N, et al. 2013. Genetic diversity and lineage dynamic of dengue virus serotype 1 (DENV-1) in Cambodia. ***Infect Genet Evol.*** 15: 59-68.

Faria NR, Nogueira RM, de Filippis AM, Simões JB, Nogueira FeB, da Rocha Queiroz Lima M, et al. 2013. Twenty years of DENV-2 activity in Brazil: molecular characterization and phylogeny of strains isolated from 1990 to 2010. ***Plos Neglected Tropical Diseases***. 7(3), e2095.

Fragnaud R, Yugueros-Marcos J, Pachot A, Bedin F. 2012. Isotope Coded Protein Labeling analysis of plasma specimens from acute severe dengue fever patients. ***Proteome Sci.*** 10: 60.

Gardella-Garcia CE, Perez-Ramirez G, Navarrete-Espinosa J, Cisneros A, Jimenez-Rojas F, Ramírez-Palacios LR, Rosado-Leon R, Camacho-Nuez M, Munoz MdeL. 2008. Specific genetic markers for detecting subtypes of dengue virus serotype-2 in isolates from the states of Oaxaca and Veracruz, Mexico. ***Biomed Central Microbiology***. 8:117.

Garner C, Phil D 2011. Confounded by Sequencing Depth in Association studies of Rare Alleles. ***Genet Epidemiol.*** 35(4): 261–268.

Gregori J, Perales C, Rodriguez-Frias F, Esteban JI, Quer J, Domingo E. 2016. Viral quasispecies complexity measures. ***Virology***. 493:227-37.

Gregori J, Salicru M, Domingo E, Sanchez A, Esteban JI, Rodriguez-Frias F, et al. 2014. Inference with viral quasispecies diversity indices: clonal and NGS approaches. ***Bioinformatics***. 30: 1104–1111.

Grenfell B, Pybus O, Gog J, Wood J, Daly J, et al. 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. ***Science*** 303: 327–332.

Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, et al. 2015. Experimental Evolution of an RNA Virus in Wild Birds: Evidence for Host-Dependent Impacts on Population Structure and Competitive Fitness. ***PLoS Pathog***. 11(5): e1004874.

Gubler DJ. 1998. Dengue and Dengue Hemorrhagic Fever. ***Clinical Microbiology Reviews***. 11(3):480-496.

Gubler DJ. 2002. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21<sup>st</sup> century. ***TRENDS in Microbiology***. 10:100-103.

Guy B, Barrere B, Malinowski C, Saville M, Teyssou R, Lang J. 2011. From research to phase III: preclinical, industrial and clinical development of the Sanofi Pasteur tetravalent dengue vaccine. ***Vaccine***. 29(42):7229-41.

Guzman MG & Harris E. 2015. Dengue. ***Lancet***. 385(9966):453-65.

Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, Hunsperger E, Kroeger A, Margolis HS, Martínez E, Nathan MB, Pelegrino JL, Simmons C, Yoksan S, Peeling RW. 2010. Dengue: a continuing global threat. ***Nat Rev Microbiol***. 8(120): S7–16. doi:10.1038/nrmicro2460.

Hadinegoro SR, Arredondo-García JL, Capeding MR, Deseda C, Chotpitayasunondh T, Dietze R, Muhammad Ismail HI, Reynales H, Limkittikul K, Rivera-Medina DM, Tran HN, Bouckennooghe A, Chansinghakul D, Cortés M, Fanouillere K, Forrat R, Frago C, Gailhardou S, Jackson N, Noriega F, Plennevaux E, Wartel TA, Zambrano B, Saville M; CYD-TDV Dengue Vaccine Working Group. 2015. Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease. ***N Engl J Med***. 373(13):1195-206.

Halstead SB. 1980. Dengue haemorrhagic fever- a public health problem and a field for research. Bulletin of the World Health Organization. Vol: 58 (1) 1 -21.

Halstead SB. 1988. Pathogenesis of dengue: challenges to molecular biology. ***Science***, 239: 476-481.

Hammond WM, Rudnick A, Sather GE. 1960. Viruses associated with epidemic hemorrhagic fevers on the Philippines and Thailand. **Science**. 131:1102-3.

Hapuarachchi HC, Chua RCR, Shi Y, Thein TL, Lee LK, Lee KS, et al. 2015. Clinical Outcome and Genetic Differences within a Monophyletic Dengue Virus Type 2 Population. **PLoS ONE**. 10(3): e0121696. doi:10.1371/journal.pone.0121696.

Harris AF, Nimmo D, McKemey AR, et al. 2011. Field performance of engineered male mosquitoes. **Nat Biotechnol**. 29: 1034–37.

Heinz FX & Allison SL. 2001. The machinery for flavivirus fusion with host cell membranes. **Curr Op in Microbiol**. 4:450-455.

Heinz FX & Stiasny K. 2012. Flaviviruses and their antigenic structure. **J Clin Virol**. 55(4):289-95. doi: 10.1016/j.jcv.2012.08.024.

Hoffmann AA, Montgomery BL, Popovici J, et al. 2011. Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. **Nature**. 476: 454–57.

Holmes EC & Twiddy SS. 2003. The origin, emergence and evolutionary genetics of dengue virus. **Infect Genet Evol**. 3(1): 19-28.

Illumina 2016a. “Two-channel SBS sequencing technology”, presente em: [http://www.illumina.com/content/dam/illumina-marketing/documents/products/techspotlights/techspotlight\\_two-channel\\_sbs.pdf](http://www.illumina.com/content/dam/illumina-marketing/documents/products/techspotlights/techspotlight_two-channel_sbs.pdf)

Illumina 2016b. “NextSeq System Denature and Dilute Libraries Guide”, presente em: [http://support.illumina.com/content/dam/illumina/support/documents/documentation/system\\_documentation/nextseq/nextseq-denature-dilute-libraries-guide-15048776-02.pdf](http://support.illumina.com/content/dam/illumina/support/documents/documentation/system_documentation/nextseq/nextseq-denature-dilute-libraries-guide-15048776-02.pdf)

Illumina 2016c. “Nextera XT DNA Library Prep Reference Guide” presente em: [http://support.illumina.com/downloads/nextera\\_xt\\_sample\\_preparation\\_guide\\_15031942.html](http://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html)

Jayasundara D, Saeed I, Maheswararajah S, Chang BC, Tang SL, Halgamuge SK. 2015. ViQuaS: an improved reconstruction pipeline for viral quasispecies spectra generated by next-generation sequencing. **Bioinformatics**. 31(6): 886–896.

Johnson BW, Russell BJ and Lanciotti RS. 2005. Serotype-Specific Detection of Dengue Viruses in a Fourplex Real-Time Reverse Transcriptase PCR Assay. **Journal of clinical microbiology**, Vol. 43: 4977–4983.

Khromykh AA, Sedlak PL, Westaway EG. 1999. Trans-Complementation Analysis of the Flavivirus Kunjin ns5 Gene Reveals an Essential Role for Translation of Its N-Terminal Half in RNA Replication. *J Virol.* 73(11):9247-9255.

Kimura R & Hotta S. 1944. Studies on dengue fever (IV) on inoculation of dengue virus into mice. *Nippon Igaku.* 3379:629-33.

Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lenches E, Jones CT, Mukhopadhyay S, Chipman PR, Strauss EG, Baker TS, Strauss JH. 2002. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell.* 108(5):717-25.

Kurane I & Ennis FE. 1992. Immunity and immunopathology in dengue virus infections. *Semin Immunol.* (2):121-7.

Lambrechts L & Lequime S. 2016. Evolutionary dynamics of dengue virus populations within the mosquito vector. *Curr Opin Virology.* 21: 47-53.

Lanciotti RS, Gubler DJ, Trent DW. 1997. Molecular evolution and phylogeny of dengue-4 viruses. *J Gen Virol.* (9):2279-84.

Lanciotti, R. et al. 1994. Molecular evolution and epidemiology of dengue-3 viruses. *J Gen Virol.* (1): 65-75.

Lauring AS & Andino R. 2010. Quasispecies theory and the behaviour of RNA viruses. *PLoS Pathog,* 6: e1001005.

Lee HY, Perelson AS, Park SC, Leitner T. 2008. Dynamic correlation between intrahost HIV-1 quasispecies evolution and disease progression. *PLoS Comput. Biol.* 4:e1000240.

Lequime S, Fontaine A, Ar Gouilh M, Moltini-Conclois I, Lambrechts L. 2015. Genetic Drift, Purifying Selection and Vector Genotype Shape Dengue Virus Intra-host Genetic Diversity in Mosquitoes. *PLoS Genet.* 12(6): e1006111.

Li D, Lott WB, Lowry K, Jones A, Thu HM, et al. 2011. Defective Interfering Viral Particles in Acute Dengue Infections. *PLoS ONE* 6(4): e19447.

Li H & Roossinck MJ. 2004. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. *Journal of virology.* Vol. 78, No. 19: 10582–10587.

Lin SR, Hsieh SC, Yueh YY, Lin TH, Chao DY, Chen WJ, King CC, Wanget WK. 2004. Study of sequence variation of dengue type 3 virus in naturally infected mosquitoes and human hosts: implications for transmission and evolution. *J. Virol.* 78:12717–12721.

Lindenbach BD & Rice CM. 1999. Genetic Interaction of Flavivirus Nonstructural Proteins NS1 and NS4A as a Determinant of Replicase Function. **J Virol.** 73(6):4611-4621.

Lodeiro MF, Filomatori CV, Gamarnik AV. 2009. Structural and functional studies of the promoter element for dengue virus RNA replication. **J Virol** 83:993–1008.

Luo D, Xu T, Hunke C, Gruber G, Vasudevan SG, Lescar J. 2008. Crystal structure of the NS3 protease-helicase from dengue virus. **J. Virol.** 82:173-183.

Mackenzie JM, Khromykh AA, Jones MK, Westaway EG. 1998. Subcellular Localization and Some Biochemical Properties of the Flavivirus Kunjin Nonstructural Proteins NS2A and NS4A. **Virology.** 245: 203–215.

Mariano F 1917. Dengue: Consideração relacionados a sua presença no Rio Grande do Sul, em 1916. **Arch Bras Med,** 7: 272 – 7.

Matranga CB, Andersen KG, Winnicki S, Busby M, Gladden AD, Tewhey R, Stremlau M, Berlin A, Gire SK, England E, Moses LM, Mikkelsen TS, Oda I, Ehiane PE, Folarin O, Goba A, Kahn SH, Grant DS, Honko A, Hensley L, Happi C, Garry RF, Malboeuf CM, Birren BW, Gnirke A, Levin JZ, Sabeti PC 2014. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. **Genome Biology.** 15:519.

Mendez, JA; Usme-Ciro, JA; Domingo, C; Rey, GJ; Sanchez, JA; Tenorio, A; et al. Phylogenetic history demonstrates two different lineages of dengue type 1 virus in Colombia. **Virol J.**, 2010, 7, 226.

Miagostovich MP, Nogueira RM, Cavalcanti SM, Marzochi KB, Schatzmayr HG. 1993. Dengue epidemic in the state of Rio de Janeiro, Brazil: virological and epidemiological aspects. **Rev Inst Med Trop Sao Paulo,** 35(2):149-54.

Miagostovich MP, Nogueira RMR, Araujo ESM, Santos BF, Schatzmayr V. 1999. Evaluation of an IgG enzyme linked immunosorbent assay for dengue diagnosis. **Journal Clin Virol** 14:183-189.

Moreau I, Levis J, Crosbie O, Kenny-Walsh E, Fanning LJ. 2008. Correlation between pre-treatment quasispecies complexity and treatment outcome in chronic HCV genotype 3a. **J Virol.** 5:78.

Murphy FA. 1980. Togavirus: morphology and morphogenesis. In: Schlesinger RW (Ed). *The Togaviruses*. New York: Academic Press.



Muyllaert IR, Galler R, Rice CM. 1997. Genetic analysis of yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. *J. Virol.* 71:291-298.

Myat Thu H, Lowry K, Jiang L, Hlaing T, Holmes EC, Aaskov J. 2005. Lineage extinction and replacement in dengue type 1 virus populations are due to stochastic events rather than to natural selection. *Virology.* 336(2): 163-72.

Nelson MJ. 1986. *Aedes aegypti*: Biología y Ecología. Organización Panamericana de la Salud. Washington, DC. 50 pp.

Nogueira RM & dos Santos F. 2015. Cap 9: Diagnóstico Laboratorial da dengue, em *Dengue: Teorias e práticas*. Valle D, Pimenta DN, Cunha RV. Editora Fiocruz, Brasil.

Nogueira RMR et al. 1990. Isolation of dengue virus type 2 in rio de janeiro. *Memórias do Instituto Oswaldo Cruz*, 85: 253.

Nogueira RMR, Araújo JMG, Schatzmayr HG. 2007. Dengue viroses in Brazil, 1986-2006. *Rev Panam Salud Publica.* 22(5): 358-63.

Nogueira RMR, Miagostovich MP, de Filippis AM, Pereira MA, Schatzmayr HG. 2001. Dengue virus type 3 in Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz* 96: 925–926.

Nogueira RMR, Miagostovich MP, Schatzmayr HG, Araújo ESM, Santos FB, Filippis AMB, Souza RW, Zagne SMO, Nicolai C, Baran M, Teixeira Filho G. 1999. Dengue in the State of Rio de Janeiro, Brazil, 1986-1998. *Mem. Inst. Oswaldo Cruz* 94(3): 297-304.

Nunes MR, Faria NR, Vasconcelos HB, Medeiros DB, Silva de Lima CP, Carvalho VL, et al. 2012. Phylogeography of dengue virus serotype 4, Brazil, 2010-2011. *Emerg Infect Dis.* 18(11), 1858-64.

Oliveira MF, Galvao Araujo JM, Ferreira OC, Ferreira DF, Lima DB, Santos FB, et al. 2010. Two lineages of dengue virus type 2, Brazil. *Emerg Infect Dis.*, 16(3), 576-8.

Parameswaran P, Charlebois P, Tellez Y, Nunez A, Ryan EM, Malboeuf CM, Levin JZ, Lennon NJ, Balmaseda A, Harris E et al. 2012. Genome-wide patterns of intrahuman dengue virus diversity within-host dengue virus evolution reveal associations with viral phylogenetic clade and interhost diversity. *J Virol.* 86:8546-8558.

Peeling RW, Artsob H, Pelegrino JL, Buchy P, Cardoso MJ, Devi S, Enria DA, Farrar J, Gubler DJ, Guzman MG, Halstead SB, Hunsperger E, Kliks S, Margolis HS,

Nathanson CM, Nguyen VC, Rizzo N, Vázquez S, Yoksan S. 2010. Evaluation of diagnostic tests: dengue. **Nat Rev Microbiol.** 8(12): S30-8.

Perez-Ramirez et al. 2009. Multiple recombinants in two dengue virus, serotype-2 isolates from patients from Oaxaca, Mexico. **Biomed central Microbiology**, 9:260.

Pierson T & Diamond M. 2013. Flaviviruses. In: Knipe D, Howley P, editors. Fields Virology. 1. 6<sup>a</sup> ed. Philadelphia: Lippincott Williams and Wilkins, p. 747-94.

Poersch CO et al. 2005. Dengue virus infections: comparison of methods for diagnosing the acute disease. **J Clin Virol.** 32(4):272-7.

Poh WT et al. 2013. Viral quasispecies inference from 454 pyrosequencing. **Biomed central Bioinformatics.** 14:355.

Pryor MJ, Carr JM, Hocking H, Davidson AD, Li P, et al. 2001. Replication of dengue virus type 2 in human monocyte-derived macrophages: comparisons of isolates and recombinant viruses with substitutions at amino acid 390 in the envelope glycoprotein. **Am J Trop Med Hyg.** 65:427–434.

Puiprom O et al. 2011. Co-existence of major and minor viral populations from two different origins in patients secondarily infected with dengue virus serotype 2 in Bangkok. **Biochemical and Biophysical Research Communications**, 413: 136-142.

Pybus OG, Rambaut A. 2009. Evolutionary analysis of the dynamics of viral infectious disease. **Nat Rev Genet.** 10:540-550.

Rice CM. 1996. Flaviviridae: the viruses and their replication, p.931–959. In B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields virology, 3rd ed. Lippincott-Raven, Philadelphia, Pa.

Rico-Hesse R. 1990. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. **Virology**, 174: 479-493.

Rico-Hesse R. 1997. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. **Virology.** 2:244-51

Rico-Hesse R. 2003. Microevolution and virulence of dengue viruses. **Adv Virus Res**, 59: 315-41.

Rodriguez-Roche R, Blanc H, Bordería AV, Díaz G, Henningsson R, Gonzalez D, Santana E, Alvarez M, Castro O, Fontes M et al. 2016. Increasing clinical severity during a dengue virus type 3 Cuban epidemic: deep sequencing of evolving viral populations. **J Virol**, 90:4320-4333.

Romano CM, Lauck M, Salvador FS, Lima CR, Villas-Boas LS, et al. 2013. Inter- and Intra-Host Viral Diversity in a Large Seasonal DENV2 Outbreak. **PLoS ONE**. 8(8): e70318.

Rosen L. 1977. The emperor's new clothes revisited, or reflections on the pathogenesis of dengue haemorrhagic fever. **American Journal of Tropical Medicine and Hygiene**. 26:337-343.

Rudnick A. 1986. Dengue fever epidemiology in Malaysia 1901-1980. *In: Dengue Fever Studies in Malaysia*. 23:9-38.

Sabchareon A, Wallace D, Sirivichayakul C, et al. 2012. Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. **Lancet**. 380: 1559–67.

Sabin AB & Schlesinger RW. 1945. Production of immunity to Dengue with virus modified by propagation in mice. **Science**. (2634):640-2.

Sabin AB. 1952. Research on dengue during World War II. **Am J Trop Med Hyg**. 1:30-50.

Salazar MI, Richardson JH, Sánchez-Vargas I, Olson KE, Beaty BJ. 2007. Dengue virus type 2: replication and tropisms in orally infected *Aedes aegypti* mosquitoes. **BMC Microbiology**. 7:9.

Salvador FS, Amorim JH, Alves RPS, Pereira SA, Ferreira LCS, Romano CM. 2015. Complete genome sequence of an atypical dengue virus serotype 2 lineage isolated in Brazil. **Genome Announc**. 3(4):e00779-15.

Sanofi Pasteur. 2016. Sanofi Pasteur dengue vaccine: frequently asked questions. Disponível em: <http://www.dengue.info/sites/default/files/media-faqs-dengueinfo-july2016.pdf>.

Santiago GA, Vergne E, Quiles Y, Cosme J, Vazquez J, Medina JF, Medina F, Colon C, Margolis H, Muñoz-Jordan JL. 2013. Analytical and Clinical Performance of the CDC Real Time RT-PCR Assay for Detection and Typing of Dengue Virus. **PLoS Negl Trop Dis**. 7(7): e2311.

Schatzmayr HG, 2008. Aspectos Históricos da Dengue e seus Vetores. Dengue Diagnóstico, Tratamento e Prevenção - 2ª Edição, Editora Rubio Ltda p. 1-10.

Schatzmayr HG, Nogueira RMR, Travassos da Rosa APA. 1986. An outbreak of dengue virus at Rio de Janeiro - 1986. **Mem. Inst. Oswaldo Cruz**, 81 (2): 245-6.

Secretaria de vigilância em saúde, Ministério de Saúde. 2009. Informe epidemiológico da Dengue: semanas de 1 a 52 de 2009. Disponível em:

[http://portal.saude.gov.br/portal/arquivos/pdf/informe\\_epidemiologico\\_semana\\_1a52\\_09\\_revisado.pdf](http://portal.saude.gov.br/portal/arquivos/pdf/informe_epidemiologico_semana_1a52_09_revisado.pdf)

Secretaria de vigilância em saúde, Ministério de Saúde. 2012. Casos de dengue. Brasil, grandes regiões e Unidades Federadas. 1997 a 2011. Disponível em:

[http://portal.saude.gov.br/portal/arquivos/pdf/dengue\\_classica\\_90\\_11\\_10\\_02\\_12.pdf](http://portal.saude.gov.br/portal/arquivos/pdf/dengue_classica_90_11_10_02_12.pdf)

Secretaria de vigilância em saúde, Ministério de Saúde. 2016. Dengue: semana epidemiológica 32. Disponível em:

<http://portalsaude.saude.gov.br/images/pdf/2016/setembro/16/2016-028---Dengue-SE32.pdf>

Sullivan DG, Bruden D, Deubner H, McArdle S, Chung M, Christensen C, Hennessy T, Homan C, Williams J, McMahan BJ, Gretch DR. 2007. Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. **J. Infect. Dis.** 196:239–248.

Takeshi K et al. 2014. Sequence variation on Dengue Type 2 virus isolated from Clinical Cases in Thailand. **Japan Journal of Infectious Diseases**, 67: 132-134.

Takeshi K. 2011. Quasispecies of dengue virus. **Tropical Medicine and Health**, 39: 29-36.

Temporão JG, Penna GO, Carmo EH, Coelho GE, do Socorro Silva Azevedo R, Nunes MR, et al. 2011. Dengue virus serotype 4, Roraima State, Brazil. **Emerg Infect Dis.** 17:938–40.

Thai KT, Henn MR, Zody MC, Tricou V, Nguyet NM, Charlebois P, Lennon NJ, Green L, de Vries PJ, Hien TT et al. 2012. High-resolution analysis of intrahost genetic diversity in dengue virus serotype 1 infection identifies mixed infections. **J Virol.** 86:835-843.

Thomas SJ & Endy TP. 2011. Critical issues in dengue vaccine development. **Curr Opin Infect Dis.** 24: 442–50.

Valle D, Pimenta DN, Cunha RV 2015. Dengue: teorias e práticas. Editora Fiocruz, Brasil.

Vasilakis N & Weaver SC. 2008. The history and evolution of human dengue emergence. **Adv Virus Res** 72: 1-76.

Vasilakis N, Cardosa J, Hanley KA, Holmes EC, Weaver SC. 2011. Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. **Nature Reviews Microbiology.** 9:532-541.

Veerakumar K, Govindarajan M, Rajeswary M. 2013. Green synthesis of silver nanoparticles using *Sida acuta* (Malvaceae) leaf extract against *Culex quinquefasciatus*, *Anopheles tephensi*, and *Aedes aegypti* (Diptera: Culicidae). ***Parasitol Res.*** 112: 4073–85.

Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. ***Nature.*** 439:344–348.

Villabona-Arenas CJ & de Andrade Zanotto PM. 2011. Evolutionary history of Dengue virus type 4: Insights into genotype phylodynamics. ***Infection, Genetics and Evolution,*** 11(5), 878-885.

Villordo SM, Filomatori CV, SánchezVargas I, Blair CD, Gamarnik AV. 2015. Dengue Virus RNA Structure Specialization Facilitates Host Adaptation. ***PLoS Pathog.*** 11(1): e1004604.

Weaver SC & Vasilakis N. 2009. Molecular evolution of dengue viruses: Contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. ***Inf Gen and Evol.*** 4:523-540.

Welsch S, Miller S, Romero-Brey I, et al. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. ***Cell Host Microbe.*** 5: 365–75.

Whitehead SS, Blaney JE, Durbin AP, Murphy BR. 2007. Prospects for a dengue virus vaccine. ***Nature Reviews Microbiology.*** 5: 518–528.

World Health Organization (WHO) 2009. Dengue: Guías para el diagnóstico, tratamiento, prevención y control. Bolivia. 170 pp. Disponible em: <http://www.who.int/denguecontrol/en/>.

World Health Organization (WHO). 2012. Global strategy for dengue prevention and control. Geneva: World Health Organization.

Wu SJL, Grouard-Vogel G, Sun W, Mascola JR, Brachtel E, Putvatana R, Blauvelt A. 2000. Human skin Langerhans cells are targets of dengue virus infection. ***Nature medicine.*** 6(7), 816-820.

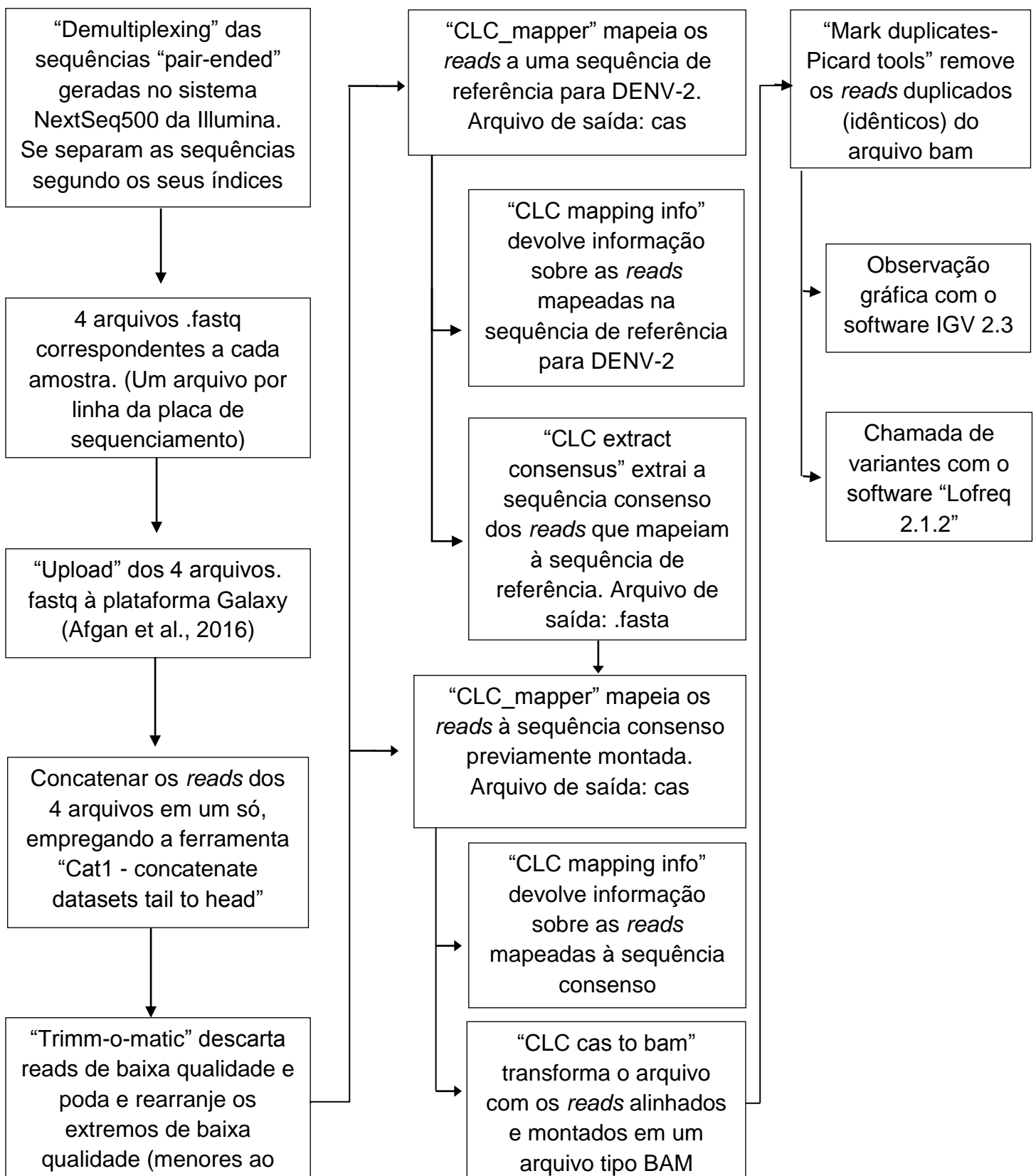
Yauch LE, Zellweger RM, Kotturi MF, et al. 2009. A protective role for dengue virus-specific CD8+ T cells. ***J Immunol.*** 182: 4865–73

Yin L, Liu L, Sun Y, Hou W, Lowe AC, Gardner BP, Salemi M, Williams WB, Farmerie WG, Sleasman JW, Goodenow MM. 2012. High-resolution deep sequencing reveals biodiversity, population structure, and persistence of HIV-1 quasispecies within host ecosystems. ***Retrovirology.*** 9:108.

Zou J, Xie X, Wang Q-Y, Dong H, Lee MY, Kang C, Yuan Z, Shi P-Y. 2015. Characterization of dengue virus NS4A and NS4B protein interaction. ***J Virol.*** 89:3455–3470.

## 9 APÊNDICES E/OU ANEXOS

**Apêndice I** – Fluxograma das análises bioinformáticas das sequências obtidas por sequenciamento profundo no sistema NextSeq 500 da Illumina (se coloca como exemplo para uma única amostra)



**Apêndice II – SNV reportadas para cada amostra.** POS: posição; REF: referência; ALT: alterada, Freq %: frequência porcentual; aa: aminoácido.

131						134					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
1489	T	A	12,7	ENV	C185S	5092	A	T	30,0	NS3	M211L
2408	C	G	16,3	ENV	A491G	6895	G	C	20,0	NS4B	E24Q
2481	C	T	57,1	NS1	-	7057	C	G	20,0	NS4B	L78V
2517	T	A	15,8	NS1	Y32STOP						
3022	T	C	25,0	NS1	W201R						
3518	G	T	13,0	NS2A	G14V						
3870	G	C	11,8	NS2A	M131I						
5092	A	T	50,0	NS3	M211F						
5094	G	C	50,0	NS3	-						
6866	G	C	30,0	NS4B	G14A						
7004	C	G	16,0	NS4B	S60STOP						
7010	A	T	14,8	NS4B	N62I						
8491	A	C	20,8	NS5	S308R						
8602	G	C	16,7	NS5	D345H						
8808	A	T	20,0	NS5	T414S						
9795	G	T	38,1	NS5	Q742H						
9853	T	A	14,3	NS5	W762R						
10369	T	A	15,4	3'UTR							

135						136					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
99	G	T	56,3	CAP	MI1	27	A	C	72,7	5'UTR	
405	C	A	11,6	preMEM	-	55	A	C	22,5	5'UTR	
952	A	C	7,7	ENV	I6L	187	A	C	11,9	CAP	K31Q
1277	T	A	15,0	ENV	V114E	246	T	C	11,2	CAP	-
1281	C	A	12,5	ENV	-	337	T	A	5,1	CAP	L81M
1289	T	A	8,1	ENV	M118K	341	G	T	7,1	CAP	R82I
1578	G	A	18,2	ENV	-	514	A	T	2,6	MEM	K26STOP
3609	T	C	39,4	ORF	-	539	G	C	3,5	MEM	C34S
3801	C	A	30,0	NS2A	-	556	G	C	2,1	MEM	D40H
3818	G	C	25,0	NS2A	S114T	562	G	C	2,8	MEM	G42R
5460	T	C	99,9	NS3	-	628	A	T	5,6	MEM	I64L
5568	A	T	12,1	NS3	E369D	677	G	C	3,4	MEM	C80S
5574	C	G	24,1	NS3	-	689	G	C	3,4	MEM	G84A
6606	T	A	12,1	NS4A	-	743	G	C	3,2	MEM	G102A
8767	A	T	18,8	NS5	T400S	929	C	G	2,5	MEM	S164STOP
9650	G	C	8,3	NS5	W694S	959	A	T	4,8	ENV	N8I
10209	C	A	9,7	NS5	T880K	983	C	T	2,4	ENV	S16L
10332	G	A	63,6	3'UTR		1002	C	G	2,5	ENV	D22E
10467	C	A	21,1	3'UTR		1003	A	T	3,0	ENV	I23L
10488	G	C	95,2	3'UTR		1008	C	G	2,8	ENV	-
10638	G	A	81,8	3'UTR		1169	G	C	5,4	ENV	G78A



POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
202	G	C	9,1	CAP	-	3314	T	A	5,6	NS1	L298STOP
234	T	A	11,5	CAP	-	3485	G	C	12,1	NS2A	G3A
498	A	T	10,0	MEM	K21STOP	3532	C	G	6,0	NS2A	L19V
1544	T	C	42,9	ENV	-	3646	A	C	6,9	NS2A	M57L
1598	G	A	42,1	ENV	-	3896	T	A	3,0	NS2A	M140K
2796	A	T	11,1	NS1	T126S	3932	T	A	3,3	NS2A	I152N
3015	G	A	25,0	NS1	G199S	3938	G	C	3,1	NS2A	-
3459	A	G	31,3	NS1	N347D	3939	T	A	3,0	NS2A	C154S
3471	A	G	27,3	NS1	T351A	3958	C	A	3,1	NS2A	L161I
3782	G	A	25,0	NS2A	M102I	4038	G	T	3,9	NS2A	Q187H
3984	A	T	13,6	NS2A	T170S	4109	T	A	5,1	NS2B	L211H
5203	C	G	33,3	NS3	A228G	4215	G	C	2,6	NS2B	M28I
5424	A	G	25,0	NS3	T302A	4355	C	G	4,3	NS2B	S75STOP
6789	A	T	14,8	P2K	I12L	4378	A	T	9,6	NS2B	S83C
7790	A	G	13,6	NS5	-	4491	A	T	3,6	NS2B	-
8578	T	A	16,7	NS5	M337K	4527	A	T	3,8	NS3	-
8847	G	C	30,0	NS5	E427Q	4528	G	T	3,5	NS3	V3F
9795	G	A	18,8	NS5	G743R	4530	A	T	3,5	NS3	-
9887	G	A	15,0	NS5	-	4646	A	T	4,0	NS3	K42I
10014	G	T	20,0	NS5	V816L	4810	G	C	5,5	NS3	V97L
10303	G	A	25,0	3'UTR	-	4862	G	C	3,4	NS3	G114A
1228	G	C	2,5	ENV	D98H	4865	T	A	2,6	NS3	L115H
1234	G	C	4,2	ENV	G100R	4989	A	T	2,4	NS3	-
1237	T	A	2,9	ENV	W101R	4993	A	T	2,9	NS3	S158C
1276	G	C	6,2	ENV	V114L	5015	T	A	3,8	NS3	I165K
1289	T	A	3,4	ENV	M118K	5022	G	T	3,1	NS3	Q167H
1291	T	A	4,4	ENV	F119I	5092	A	T	7,7	NS3	M187F
1311	G	A	9,0	ENV	M125I	5094	G	C	6,1	NS3	-
1347	T	A	3,6	ENV	Y137STOP	5235	T	A	3,2	NS3	-
1363	C	A	2,5	ENV	P143T	5324	G	C	3,3	NS3	R268T
1595	G	T	4,1	ENV	K220L	5381	A	T	2,4	NS3	H287L
1677	A	T	3,8	ENV	K247N	5405	C	A	2,3	NS3	A295E
1719	C	G	3,3	ENV	H261Q	5531	G	T	5,5	NS3	R337I
1725	A	T	4,8	ENV	-	5619	A	T	4,5	NS3	K366N
1779	A	T	4,0	ENV	-	5625	A	T	4,5	NS3	-
1804	G	T	2,2	ENV	D290Y	5699	A	T	2,8	NS3	E393V
1829	C	G	6,9	ENV	S298STOP	5718	C	G	8,1	NS3	-
1832	A	T	4,3	ENV	Y299F	5873	A	G	8,7	NS3	H451R
2424	T	A	5,1	NS1	D1E	5955	T	A	4,1	NS3	-
2430	T	A	4,2	NS1	-	6083	G	T	10,0	NS3	G521V
2484	T	A	4,5	NS1	-	6179	A	T	7,9	NS3	E553V
2505	G	C	3,3	NS1	W28C	6188	A	T	7,9	NS3	N556I
2632	A	T	4,0	NS1	I71L	6324	T	A	6,4	NS3	Y601STOP
2695	G	T	5,9	NS1	D92Y	6367	G	C	5,7	NS3	G616R
2702	A	T	5,3	NS1	K94I	6478	G	A	8,6	NS4A	E35X
2804	C	G	3,9	NS1	S128C	6478	G	T	7,0	NS4A	E35X
2986	A	T	8,9	NS1	K189STOP	6575	G	T	4,1	NS4A	G67V
3043	G	C	4,5	NS1	D208H	6618	G	T	3,5	NS4A	M80I

## 137 continuação

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
6667	G	C	2,8	NS4A	A97P	9156	A	T	8,3	NS5	-
6775	T	A	4,4	P2K	-	9325	A	C	8,0	NS5	-
6777	G	C	4,0	P2K	L7I	9438	G	C	8,9	NS5	M623I
6878	T	A	4,3	NS4B	I18N	9541	A	T	8,9	NS5	M658L
6894	T	C	4,0	NS4B	-	9592	G	A	11,6	NS5	A675T
6909	G	T	4,1	NS4B	-	9648	A	C	4,1	NS5	Q693H
6926	C	A	3,6	NS4B	P34H	9649	T	A	4,1	NS5	W694R
6971	C	A	2,7	NS4B	T49K	9658	T	A	2,6	NS5	S697T
7057	C	A	2,9	NS4B	L78I	9695	G	C	5,3	NS5	C709STOP
7116	A	C	3,1	NS4B	-	9697	T	C	3,9	NS5	S710P
7212	A	T	4,8	NS4B	-	9702	C	A	4,0	NS5	H711Q
7273	A	T	5,2	NS4B	I150L	9752	G	C	3,9	NS5	C728S
7282	A	T	3,3	NS4B	I153F	9837	G	T	2,4	NS5	K756N
7288	C	G	3,7	NS4B	L155V	9950	C	G	4,0	NS5	T794S
7296	A	T	4,0	NS4B	-	10071	A	T	3,8	NS5	E834D
7338	A	T	3,8	NS4B	-	10077	C	G	4,8	NS5	-
7345	C	G	3,0	NS4B	L174V	10083	C	G	4,0	NS5	Y838STOP
7385	C	G	3,9	NS4B	T187R	10095	A	T	3,2	NS5	R842S
7890	A	T	10,4	NS5	-	10126	C	G	2,8	NS5	L853V
7920	A	C	8,2	NS5	-	10161	A	T	2,5	NS5	Q864H
7999	T	A	5,8	NS5	L144K	10167	A	T	2,7	NS5	-
8127	T	A	5,6	NS5	Y186STOP	10182	A	T	2,5	NS5	R871S
8546	G	T	4,0	NS5	R326I	10195	A	T	2,4	NS5	-
8548	C	G	7,5	NS5	L327V	10196	A	T	4,1	NS5	N876Y
8878	A	T	6,6	NS5	R437W	10202	A	T	4,9	NS5	E878V
8913	A	T	5,1	NS5	E448D	10430	A	T	7,6	3'UTR	
8916	A	T	6,6	NS5	-	10660	C	G	9,0	3'UTR	

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	2,1	5'UTR		3935	C	G	4,1	NS2A	S153W
181	C	T	2,2	CAP	-	3969	A	T	4,1	NS2A	-
187	A	C	2,3	CAP	K31Q	3985	A	T	4,2	NS2A	T170S
200	T	A	2,4	CAP	L35H	4051	A	T	4,2	NS2A	I192L
240	G	C	2,5	CAP	M48I	4199	A	T	4,2	NS2B	K23M
452	C	G	2,6	MEM	T5R	4205	A	T	4,2	NS2B	D25V
458	A	T	2,6	MEM	N7I	4211	C	G	4,3	NS2B	P27R
515	A	T	2,7	MEM	-	4378	A	T	4,3	NS2B	S83C
517	A	T	2,7	MEM	K26I	4577	A	T	4,3	NS3	E19V
519	A	T	2,7	MEM	T27S	4810	G	C	4,4	NS3	V97L
529	G	C	2,8	MEM	V31L	4942	G	C	4,5	NS3	D141H
628	A	T	2,9	MEM	I64L	5022	G	C	4,5	NS3	Q167H
677	G	T	2,9	MEM	C80F	5092	A	T	4,5	NS3	M191L
735	G	T	2,9	MEM	-	5100	C	A	4,5	NS3	-
950	G	C	3,1	ENV	G5A	5371	G	C	4,5	NS3	D284H
956	C	G	3,1	ENV	S7STOP	5409	T	C	4,6	NS3	-
1105	A	T	3,1	ENV	R57W	5531	G	T	4,7	NS3	R337I
1109	A	T	3,1	ENV	K58M	5694	C	G	4,9	NS3	D391E
1112	A	T	3,1	ENV	Y59F	5876	C	G	4,9	NS3	S452C
1221	C	G	3,1	ENV	-	6083	G	T	5,0	NS3	G521V
1390	G	C	3,1	ENV	G152R	6700	A	T	5,2	NS4A	I108L
1489	T	A	3,2	ENV	C185S	6793	G	C	5,5	P2K	A13P
2037	A	G	3,2	ENV	I367M	7016	C	A	5,6	NS4B	S64Y
2411	T	C	3,3	ENV	M492T	7242	A	T	5,7	NS4B	-
2424	T	A	3,4	NS1	D1E	7273	A	T	5,8	NS4B	I150L
2438	T	A	3,4	NS1	V6E	7288	C	G	5,8	NS4B	L155V
2505	G	C	3,4	NS1	-	7385	C	G	5,8	NS4B	T187R
2511	G	C	3,4	NS1	W28C	8409	A	T	5,9	NS5	R280S
2515	T	A	3,4	NS1	E30D	8575	C	G	6,0	NS5	P336A
2517	T	A	3,5	NS1	Y32K	9437	T	A	6,1	NS5	M623K
2598	A	T	3,5	NS1	-	9489	A	T	6,2	NS5	E640D
2630	A	T	3,5	NS1	Q70L	9592	G	A	6,5	NS5	A675T
2686	A	T	3,5	NS1	M89L	9622	G	C	6,6	NS5	G685R
2786	A	T	3,6	NS1	K122I	9649	T	A	6,7	NS5	W694T
2797	A	T	3,6	NS1	T126S	9650	G	C	6,7	NS5	-
2807	A	T	3,6	NS1	H129L	9690	T	A	7,1	NS5	-
2856	C	G	3,6	NS1	N145K	9696	T	A	7,7	NS5	C709STOP
2962	C	T	3,7	NS1	S181P	9736	G	C	8,9	NS5	V723L
3579	G	C	3,7	NS2A	-	9813	G	C	9,3	NS5	L748F
3600	G	T	3,7	NS2A	-	9939	A	T	9,3	NS5	-
3613	G	C	3,8	NS2A	G46R	9945	A	T	10,0	NS5	-
3696	T	A	3,8	NS2A	-	9958	A	T	10,3	NS5	I797L
3699	T	A	3,8	NS2A	Y74STOP	10071	A	T	11,6	NS5	E834D
3797	G	C	3,9	NS2A	G107A	10077	C	G	11,6	NS5	-
3839	T	A	3,9	NS2A	L121Q	10101	C	G	11,8	NS5	D844E
3845	T	A	4,0	NS2A	L123Q	10176	A	T	16,3	NS5	Q869H
3855	G	T	4,0	NS2A	-	10179	C	G	17,8	NS5	-
3932	T	A	4,1	NS2A	I152N	10212	C	G	17,9	NS5	D881E

139						140					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
298	G	C	25,0	CAP	R68T	2298	T	A	45,5	ENV	-
449	C	T	7,9	MEM	-	8790	A	T	24,0	NS5	-
1278	A	T	20,8	ENV	T115S	9136	A	G	61,5	NS5	G523S
2655	C	T	23,1	NS1	-						
5363	T	C	64,3	NS3	-						
8531	G	A	9,3	NS5	-						

141											
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	1,9	5'UTR		1829	C	G	2,9	ENV	S298STOP
62	T	A	2,0	5'UTR		1832	A	T	2,9	ENV	Y299F
99	G	T	2,1	CAP	M1I	1843	A	T	2,9	ENV	T303S
140	T	A	2,1	CAP	-	1885	C	G	2,9	ENV	H317D
141	G	C	2,1	CAP	M15N	2037	A	G	2,9	ENV	I367M
181	C	T	2,1	CAP	-	2295	G	C	2,9	ENV	W453
187	A	C	2,2	CAP	K31Q	2411	T	C	3,0	ENV	M492T
198	A	T	2,2	CAP	-	2501	C	A	3,0	NS1	T27K
248	T	A	2,2	CAP	V51E	2505	G	C	3,0	NS1	W28C
351	G	C	2,2	CAP	R85S	2511	G	C	3,1	NS1	E30D
519	A	T	2,2	MEM	-	2517	T	A	3,1	NS1	Y32STOP
546	C	G	2,2	MEM	-	2864	G	T	3,1	NS1	R148I
553	A	T	2,2	MEM	I39L	2911	A	T	3,1	NS1	-
628	A	T	2,3	MEM	I64L	2920	A	T	3,1	NS1	I167L
655	A	T	2,3	MEM	T73S	3046	A	C	3,1	NS1	T209P
671	G	C	2,3	MEM	G78A	3116	G	C	3,1	NS1	W232S
735	G	T	2,3	MEM	-	3201	C	A	3,1	NS1	Y260STOP
872	G	C	2,4	MEM	G145A	3309	C	A	3,2	NS1	-
933	G	C	2,4	MEM	M165I	3498	C	G	3,2	NS2A	N7K
940	C	G	2,5	ENV	R2G	3579	G	C	3,2	NS2A	-
942	C	A	2,5	ENV	-	3688	G	C	3,2	NS2A	G71P
1063	T	A	2,6	ENV	F43I	3689	G	C	3,3	NS2A	-
1112	A	T	2,6	ENV	Y59F	3838	C	A	3,3	NS2A	L121I
1117	A	T	2,7	ENV	I61L	3857	T	A	3,3	NS2A	L127STOP
1127	A	T	2,7	ENV	K64M	3985	A	T	3,3	NS2A	T170S
1289	T	A	2,7	ENV	M118K	3995	C	A	3,3	NS2A	A173E
1403	G	C	2,7	ENV	G156A	4051	A	T	3,3	NS2A	I192L
1471	G	C	2,7	ENV	G179R	4057	C	G	3,4	NS2A	L194V
1475	C	A	2,7	ENV	T180N	4205	A	T	3,4	NS2B	D25V
1516	A	T	2,7	ENV	N194Y	4339	A	T	3,4	NS2B	S70C
1533	G	T	2,7	ENV	-	4378	A	T	3,4	NS2B	S83C
1641	G	C	2,7	ENV	E235D	4437	A	T	3,4	NS2B	-
1686	C	G	2,7	ENV	-	4585	G	T	3,4	NS3	A22S
1692	C	G	2,7	ENV	-	4758	T	A	3,4	NS3	Y79STOP
1695	A	T	2,8	ENV	L253F	4889	T	A	3,5	NS3	I123K
1706	A	T	2,8	ENV	E257V	4985	T	A	3,6	NS3	V155D
1779	A	T	2,8	ENV	-	5002	T	A	3,6	NS3	Y161N
1820	A	T	2,9	ENV	-	5192	C	G	3,6	NS3	T224S
1821	A	T	2,9	ENV	K295I	5234	T	A	3,6	NS3	L238H

## 141 continuação

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
5330	T	A	3,6	NS3	L270Q	8398	A	T	4,8	NS5	I277L
5387	C	G	3,7	NS3	T289R	8403	A	T	4,8	NS5	-
5485	A	T	3,7	NS3	R322STOP	8528	T	A	4,9	NS5	M320K
5501	A	G	3,7	NS3	Q327R	8534	A	T	4,9	NS5	N322I
5509	G	C	3,8	NS3	A330P	8690	T	A	4,9	NS5	M374K
5531	G	T	3,8	NS3	R337I	8756	G	A	5,0	NS5	R396K
5567	A	T	3,8	NS3	E349V	8807	T	A	5,1	NS5	F413Y
5632	A	T	3,9	NS3	I371L	8916	A	T	5,1	NS5	-
5668	A	T	3,9	NS3	I383L	8947	A	T	5,2	NS5	K460STOP
5743	A	T	3,9	NS3	T408S	8966	G	T	5,2	NS5	G466V
5946	G	C	3,9	NS3	M475I	8991	A	T	5,2	NS5	-
5971	G	C	3,9	NS3	D484H	9107	A	T	5,3	NS5	H513L
6324	T	A	3,9	NS3	Y601STOP	9112	C	G	5,3	NS5	L515V
6335	T	A	3,9	NS3	L605Q	9116	G	T	5,5	NS5	G516V
6657	C	G	3,9	NS4A	-	9127	A	T	5,6	NS5	R520STOP
6726	T	C	3,9	NS4A	-	9566	T	A	5,6	NS5	V666D
6758	C	A	4,0	P2K	T1K	9578	T	A	5,6	NS5	I670K
6790	A	T	4,0	P2K	I12L	9592	G	A	5,6	NS5	A675T
6793	G	C	4,1	P2K	A13P	9616	G	C	5,7	NS5	D683H
6799	C	G	4,1	P2K	L15V	9641	T	A	6,3	NS5	I691K
6928	G	C	4,1	NS4B	A35P	9650	G	C	6,3	NS5	W694S
7052	T	A	4,1	NS4B	M76K	9658	T	A	6,6	NS5	S697T
7170	A	T	4,1	NS4B	-	9662	G	C	6,7	NS5	R698T
7273	A	T	4,1	NS4B	I150L	9675	T	A	6,7	NS5	D702E
7282	A	T	4,2	NS4B	I153F	9686	T	A	6,7	NS5	V706E
7288	C	G	4,2	NS4B	-	9732	T	A	6,8	NS5	-
7333	C	G	4,2	NS4B	Q170E	9741	C	A	6,9	NS5	-
7368	A	T	4,3	NS4B	-	9751	T	A	6,9	NS5	C728S
7371	A	T	4,3	NS4B	L182F	9795	G	T	7,0	NS5	Q742H
7385	C	G	4,3	NS4B	T187R	9814	A	C	7,6	NS5	K749Q
7545	G	C	4,3	NS4B	M239I	9852	G	C	8,2	NS5	M761I
7548	G	C	4,4	NS4B	K240N	9855	G	C	8,3	NS5	W762C
8009	T	A	4,4	NS5	I147K	9861	G	C	8,4	NS5	-
8012	G	C	4,5	NS5	G148A	9863	T	A	8,6	NS5	M765K
8016	G	T	4,5	NS5	E149D	9879	T	A	8,8	NS5	-
8074	G	C	4,6	NS5	E169Q	9965	C	G	9,6	NS5	A799G
8096	C	G	4,6	NS5	T176S	10009	A	T	10,9	NS5	N814Y
8105	G	C	4,7	NS5	C179S	10015	G	T	11,7	NS5	V816L
8160	A	T	4,7	NS5	-	10101	C	G	11,9	NS5	D844E
8220	A	T	4,7	NS5	E217D	10249	G	C	12,6	NS5	E894L
8312	A	T	4,8	NS5	K248M	10250	A	T	14,3	NS5	-

142						143					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	46,2	5'UTR		1599	G	A	33,3	ENV	-
1269	A	T	13,5	ENV	-	1842	T	C	16,7	ENV	-
1432	C	G	13,3	ENV	P166A	2010	G	A	40,0	ENV	-
2627	A	T	12,5	NS1	K69I	2064	C	T	23,1	ENV	-
3175	T	G	25,0	NS1	S252A	2556	A	T	21,4	NS1	-
5092	A	T	16,7	NS3	-	3114	C	A	20,0	NS1	-
5094	G	C	12,5	NS3	M191F	3129	A	G	36,4	NS1	-
5097	C	T	24,5	NS3	-	3584	T	A	35,7	NS2A	V36D
5140	A	T	12,0	NS3	I207L	3876	C	T	40,0	NS2A	-
5341	A	G	6,7	NS3	R274G	4562	G	A	35,7	NS3	G14E
6758	C	A	16,7	P2K	T1K	4854	G	A	26,7	NS3	-
7087	G	T	8,1	NS4B	D88Y	4974	C	T	28,6	NS3	-
7134	C	T	13,2	NS4B	-	5142	A	C	21,4	NS3	-
7279	G	T	11,4	NS4B	V152L	5757	A	C	30,0	NS3	E412D
7808	T	C	12,0	NS5	L80P	7232	C	A	33,3	NS4B	R136S
8625	G	T	19,0	NS5	Q352H	7248	G	A	46,2	NS4B	M142V
8756	G	A	30,0	NS5	R396K	7707	A	T	23,5	NS5	R47STOP
8899	G	C	11,5	NS5	E444Q	9033	C	T	26,7	NS5	-
						9149	G	A	90,0	NS5	-
						9839	T	C	56,3	NS5	-
						9848	A	G	69,2	NS5	-

144					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	5,9	5'UTR	
99	G	T	6,7	CAP	M1I
187	A	C	7,4	CAP	K31Q
305	G	C	7,7	CAP	G70A
309	G	T	7,9	CAP	-
539	G	C	8,5	MEM	C34S
1785	C	G	8,8	ENV	-
2637	A	T	8,8	NS1	-
3079	G	T	8,8	NS1	V220F
3262	G	C	9,3	NS1	E281Q
3448	G	C	9,5	NS1	E343Q
3501	T	C	9,8	NS2A	-
3641	G	C	10,4	NS2A	R55T
3646	A	C	10,5	NS2A	M57L
5092	A	T	10,8	NS3	-
5094	G	C	10,9	NS3	M191F
5097	C	T	11,1	NS3	-
5101	C	T	11,8	NS3	H194Y
5278	G	C	11,9	NS3	G253R
5876	C	G	12,4	NS3	S452C
6324	T	A	12,5	NS3	Y601STOP
9908	T	C	41,7	NS5	-

## 145

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
41	A	T	18,9	5'UTR		556	G	C	2,3	MEM	D40H
42	G	C	19,6	5'UTR		582	A	T	2,9	MEM	-
55	A	C	27,3	5'UTR		631	G	C	3,8	MEM	D65H
91	T	A	2,6	5'UTR		671	G	C	4,2	MEM	G78A
99	G	T	7,5	CAP	M1I	6481	G	C	12,8	NS4A	V36L
102	T	A	5,1	CAP	N2K	6823	G	A	12,9	P2K	A23T
165	G	T	3,2	CAP	-	7147	A	T	13,6	NS4B	T108S
187	A	C	8,4	CAP	K31Q	7150	G	C	14,1	NS4B	A109P
363	A	T	3,0	CAP	-	7335	A	T	16,1	NS4B	Q170H
418	T	C	5,1	preMEM	-	7408	A	G	16,7	NS4B	T194A
517	A	T	3,5	MEM	T27STOP	8494	T	C	17,2	NS5	Y309H
518	C	G	2,0	MEM	-	9006	T	C	18,6	NS5	-
523	G	C	1,9	MEM	D29H	9616	G	C	20,0	NS5	D683H
529	G	C	3,1	MEM	V31L	9834	G	A	22,5	NS5	-
541	A	T	3,4	MEM	T35S	9863	T	A	26,2	NS5	M765K
547	A	T	2,3	MEM	M37L	9872	A	C	27,1	NS5	H768P
553	A	T	3,0	MEM	I39L						

## 146

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
17	G	T	36,8	5'UTR		1164	A	T	3,5	ENV	-
27	A	C	11,9	5'UTR		1228	G	C	3,9	ENV	D98H
41	A	T	3,6	5'UTR		1231	A	T	2,6	ENV	R99STOP
55	A	C	38,6	5'UTR		1234	G	C	3,2	ENV	G100R
63	T	A	13,0	5'UTR		1240	G	C	4,0	ENV	G102R
86	G	T	0,5	5'UTR		1243	A	T	3,2	ENV	N103Y
88	A	T	0,5	5'UTR		1269	A	T	2,8	ENV	-
90	A	C	0,5	5'UTR		1273	A	T	2,5	ENV	I113F
96	G	A	3,9	5'UTR		1289	T	A	4,1	ENV	M118K
99	G	T	22,4	CAP	M1I	1396	G	C	2,4	ENV	D154H
102	T	A	12,4	CAP	N2K	1468	T	A	3,1	ENV	Y178N
125	G	C	0,5	CAP	S10T	1471	G	C	3,6	ENV	G179R
135	C	T	2,1	CAP	T25A	1477	G	C	2,9	ENV	V181L
165	G	T	1,9	CAP	-	1535	A	T	4,1	ENV	Q200L
169	A	G	1,0	CAP	K31Q	1563	C	G	3,1	ENV	H209Q
181	C	T	3,9	CAP	-	1692	C	G	3,1	ENV	-
187	A	C	13,0	CAP	-	1760	G	T	2,9	ENV	G275V
198	A	T	2,2	CAP	-	1779	A	T	3,1	ENV	-
266	T	A	0,5	CAP	L57Q	1785	C	G	3,6	ENV	-
284	C	T	13,1	CAP	A63V	1798	A	T	2,9	ENV	R288STOP
305	G	A	3,5	CAP	G70E	1805	A	T	3,1	ENV	D290V
677	G	C	3,3	MEM	C80S	1820	A	T	3,5	ENV	K295I
735	G	C	3,3	MEM	-	1829	C	G	4,1	ENV	S298STOP
737	G	C	3,1	MEM	G100A	1904	G	T	2,7	ENV	R323I
743	G	C	4,3	MEM	G102A	1956	A	C	5,4	ENV	-
769	A	T	3,2	MEM	M111L	2295	G	C	3,3	ENV	W453C
869	T	A	2,4	MEM	I144K	2311	A	C	5,3	ENV	-
946	A	T	2,2	ENV	I4L	2312	T	C	3,3	ENV	I459P

146 continuação												
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa	
2406	A	T	4,7	ENV	-	4986	C	A	1,8	NS3	-	
2411	T	C	4,2	ENV	M492T	4996	G	C	2,4	NS3	G159R	
2424	T	A	2,9	NS1	D1E	5010	T	A	2,0	NS3	S163R	
2451	C	T	32,3	NS1	-	5015	T	A	3,0	NS3	I165K	
2501	C	A	2,7	NS1	T27K	5092	A	T	14,4	NS3	M191L	
2505	G	C	4,8	NS1	W28C	5152	G	T	3,1	NS3	A211S	
2511	G	C	2,5	NS1	E30D	5192	C	G	3,3	NS3	T224S	
2517	T	A	2,3	NS1	Y32STOP	5224	C	G	2,7	NS3	L235V	
2520	G	C	1,7	NS1	K33N	5289	C	G	1,9	NS3	I256M	
2607	A	T	4,5	NS1	R62S	5330	T	A	2,4	NS3	L270Q	
2687	T	A	3,1	NS1	M89K	5381	A	T	2,5	NS3	H287L	
2738	C	G	3,5	NS1	P106R	5424	A	T	3,4	NS3	-	
2745	C	T	4,9	NS1	-	5430	A	T	3,5	NS3	-	
2747	C	G	5,4	NS1	T109S	5435	A	T	3,3	NS3	E305V	
2797	A	T	5,2	NS1	T126S	5440	G	T	3,6	NS3	G307C	
2804	C	G	3,2	NS1	S128C	5486	G	C	2,8	NS3	R322T	
3049	T	A	3,1	NS1	W210R	5531	G	T	4,4	NS3	R337I	
3321	G	C	4,3	NS1	-	5534	A	T	2,1	NS3	E338V	
3365	C	G	2,7	NS1	S315C	5555	A	T	4,0	NS3	N345I	
3375	A	T	3,3	NS1	-	5562	A	T	3,2	NS3	-	
3689	G	C	2,7	NS2A	G71A	5617	A	T	3,2	NS3	K366STOP	
3699	T	A	2,3	NS2A	Y74STOP	5659	A	T	4,4	NS3	K380STOP	
3825	G	C	4,0	NS2A	L116F	5664	A	T	3,8	NS3	K381N	
3839	T	A	2,5	NS2A	L121Q	5668	A	T	4,0	NS3	I383L	
3846	G	T	2,2	NS2A	-	5718	C	G	4,8	NS3	-	
3848	C	G	2,4	NS2A	T124S	5783	G	T	2,9	NS3	R421M	
3896	T	A	2,2	NS2A	M140K	5788	A	T	3,5	NS3	I423L	
3911	T	A	2,2	NS2A	L145STOP	5821	C	G	3,0	NS3	L434V	
3947	A	T	2,3	NS2A	N157I	5880	T	A	4,5	NS3	S453R	
3985	A	T	4,3	NS2A	T170S	5986	A	T	5,3	NS3	K489STOP	
4051	A	T	3,4	NS2A	I192L	6242	A	T	3,7	NS3	E574V	
4057	C	G	2,8	NS2A	L194V	6324	T	A	7,7	NS3	Y601STOP	
4168	A	T	2,7	NS2B	M13L	6330	T	A	4,8	NS3	D603E	
4199	A	T	2,3	NS2B	K23M	6341	T	A	3,0	NS3	L607Q	
4205	A	T	2,9	NS2B	D25V	6395	T	A	3,0	NS4A	I7N	
4211	C	G	3,0	NS2B	P27R	6516	A	G	6,6	NS4A	-	
4345	C	G	3,5	NS2B	P72A	6600	A	T	2,6	NS4A	-	
4378	A	T	7,6	NS2B	S83C	6601	A	T	4,3	NS4A	K76STOP	
4407	A	T	4,2	NS2B	E92D	6607	A	T	6,3	NS4A	-	
4516	C	G	3,5	NS2B	Q129E	6614	A	T	6,0	NS4A	K79M	
4527	A	T	2,1	NS3	-	6652	A	T	2,1	NS4A	I92F	
4557	A	T	4,7	NS3	-	6657	C	G	3,5	NS4A	-	
4589	A	T	5,9	NS3	Y23F	6793	G	C	2,7	P2K	A13P	
4651	G	T	4,1	NS3	G44STOP	6799	C	G	3,3	P2K	L15V	
4661	A	T	2,9	NS3	H47L	6894	T	A	3,3	NS4B	-	
4891	G	C	3,2	NS3	G124R	6923	G	C	3,7	NS4B	R33P	
4906	G	C	3,0	NS3	D129H	6926	C	A	3,5	NS4B	P34H	
4937	T	A	3,1	NS3	I139N	6946	T	A	3,0	NS4B	Y41N	



## 146 continuação ii

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
6949	G	C	2,8	NS4B	A42P	9650	G	C	5,4	NS5	W694S
6955	G	C	3,4	NS4B	A44P	9662	G	C	2,1	NS5	R698T
6977	T	A	3,7	NS4B	M51K	9675	T	A	1,8	NS5	D702E
7030	G	C	3,9	NS4B	A69P	9691	T	A	2,0	NS5	F708I
7053	G	C	2,8	NS4B	M76I	9696	T	A	3,8	NS5	C709STOP
7229	A	G	4,7	NS4B	K135R	9749	C	A	2,7	NS5	P727Q
7288	C	G	5,9	NS4B	L155V	9751	T	A	2,7	NS5	C728S
7385	C	G	3,4	NS4B	T187R	9774	T	A	2,4	NS5	-
7536	T	A	5,8	NS4B	F236L	9795	G	T	2,7	NS5	Q742H
7999	T	A	5,4	NS5	-	9812	T	A	2,6	NS5	L748STOP
8000	T	A	3,9	NS5	L144K	9852	G	C	3,0	NS5	M761I
8012	G	C	5,3	NS5	G148A	9958	A	T	3,2	NS5	I797L
8063	T	A	3,4	NS5	L165H	9974	A	T	2,4	NS5	E802V
8220	A	T	3,7	NS5	E217D	9999	G	C	2,3	NS5	-
8398	A	T	4,7	NS5	I277L	10005	C	G	4,5	NS5	-
8528	T	A	3,3	NS5	M320K	10015	G	T	3,9	NS5	V816L
8534	A	T	3,1	NS5	N322I	10065	A	T	2,0	NS5	-
8575	C	G	6,0	NS5	P336A	10071	A	T	3,1	NS5	E834D
8610	T	A	4,8	NS5	-	10077	C	G	3,9	NS5	-
8725	G	C	4,4	NS5	G386R	10080	A	T	3,4	NS5	-
8882	A	T	5,4	NS5	E438V	10167	A	T	3,2	NS5	-
8913	A	T	5,3	NS5	E448D	10168	A	T	2,5	NS5	I867L
9082	A	T	3,2	NS5	S505C	10176	A	T	2,8	NS5	Q869H
9172	A	T	7,2	NS5	T535S	10182	A	T	2,3	NS5	R871S
9369	T	C	14,4	NS5	-	10191	A	T	2,3	NS5	G875R
9424	T	A	6,0	NS5	L619I	10196	A	T	3,3	NS5	N876I
9438	G	C	5,4	NS5	M623I	10198	G	T	2,5	NS5	E877STOP
9489	A	T	10,8	NS5	E640D	10202	G	C	2,8	NS5	G878A
9541	A	T	3,2	NS5	M658L	10212	C	G	3,0	NS5	D881E
9556	G	C	3,0	NS5	D663H	10291	A	C	6,7	3'UTR	
9569	T	A	2,4	NS5	V667E	10311	C	T	7,4	3'UTR	
9592	G	A	15,1	NS5	A675T	10322	A	T	5,6	3'UTR	
9602	T	A	3,3	NS5	L678STOP	10363	G	C	8,0	3'UTR	
9641	T	A	3,4	NS5	I691K	10578	A	T	6,7	3'UTR	
						10586	G	T	7,9	3'UTR	

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
17	G	T	51,9	5'UTR		5304	T	G	0,5	NS3	C261W
20	C	G	10,2	5'UTR		5424	A	T	0,8	NS3	-
27	A	C	20,3	5'UTR		5555	A	T	1,3	NS3	N345I
41	A	T	6,1	5'UTR		5575	A	T	0,7	NS3	T352S
49	A	T	0,6	5'UTR		5649	G	A	1,4	NS3	-
55	A	C	40,8	5'UTR		5887	C	T	1,9	NS3	Q456L
63	T	A	14,7	5'UTR		5888	A	T	1,1	NS3	-
80	T	C	1,0	5'UTR		5915	C	A	1,4	NS3	P465Q
86	G	T	1,7	5'UTR		5986	A	C	1,2	NS3	K489Q
88	A	T	2,0	5'UTR		5994	T	C	1,1	NS3	-
90	A	C	1,9	5'UTR		6043	A	G	1,8	NS3	M508V
96	G	A	4,1	5'UTR		6102	A	T	1,2	NS3	-
99	G	T	19,6	CAP	M1I	6236	T	G	0,6	NS3	I572R
102	T	A	9,0	CAP	N2K	6258	G	T	1,0	NS3	-
135	C	T	1,7	CAP	-	6305	G	T	1,2	NS3	W595L
165	G	T	1,0	CAP	-	6424	A	T	1,9	NS4A	M17L
181	C	T	3,5	CAP	-	6732	C	T	6,1	NS4A	-
187	A	C	11,8	CAP	K31Q	6801	C	T	1,3	NS4B	-
195	C	A	0,4	CAP	F33L	7108	G	T	0,9	NS4B	A95S
198	A	T	1,7	CAP	-	7883	A	T	2,4	NS5	K105I
248	T	C	2,2	CAP	V51A	7888	G	T	1,2	NS5	G107STOP
383	A	T	0,8	CAP	N96I	8010	A	G	0,6	NS5	I147M
526	G	T	0,6	MEM	G30C	8032	A	C	0,8	NS5	T155P
1096	G	T	1,2	ENV	A54S	8059	G	T	1,1	NS5	V164F
1491	C	T	0,4	ENV	-	8063	T	A	0,9	NS5	L165H
1516	A	T	0,5	ENV	N194Y	8074	G	C	0,9	NS5	-
1748	A	T	0,5	ENV	Q271L	8075	A	G	0,7	NS5	E169R
1832	A	T	0,7	ENV	Y299F	8613	A	G	1,7	NS5	-
2041	G	T	0,7	ENV	A369S	8683	A	G	1,4	NS5	K372E
2098	A	T	1,2	ENV	K388STOP	9304	C	T	2,1	NS5	R579C
2312	T	C	0,6	ENV	I459T	9340	G	A	1,3	NS5	D591N
2315	G	T	0,7	ENV	G460V	9371	G	C	1,2	NS5	S601T
2530	G	C	0,4	NS1	E37Q	9430	A	T	1,4	NS5	R621STOP
2534	C	G	0,4	NS1	S38C	9439	G	A	4,5	NS5	E624K
2610	G	T	0,9	NS1	L63F	9444	A	C	1,6	NS5	-
2747	C	G	0,7	NS1	T109S	9500	A	C	0,6	NS5	Q644Q
3188	A	T	1,4	NS1	Y256F	9536	C	T	4,0	NS5	S656L
3212	C	T	0,5	NS1	T264I	9540	A	G	2,2	NS5	-
3221	C	T	1,1	NS1	P267L	9554	G	A	0,7	NS5	G662E
3645	G	T	0,5	NS2A	-	9559	G	T	0,9	NS5	-
3760	A	T	0,5	NS2A	K95STOP	9560	A	C	0,9	NS5	D664S
3768	C	T	0,8	NS2A	-	9589	T	C	2,0	NS5	F674L
4034	A	T	0,4	NS2A	Q186L	9658	T	C	0,7	NS5	S697P
4168	A	T	1,0	NS2B	M13L	9699	A	C	0,8	NS5	-
4291	G	T	0,7	NS2B	E54STOP	9929	G	T	0,6	NS5	W787L
4336	G	T	0,6	NS2B	G69STOP	9999	G	T	0,5	NS5	-
4651	G	T	1,4	NS3	G44STOP	10019	G	T	0,6	NS5	W817L
4864	A	C	0,7	NS3	I115L	10241	A	G	1,0	NS5	R891K
4981	G	T	0,4	NS3	V154F	10395	C	T	1,4	3'UTR	
5034	C	T	0,5	NS3	-						

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
41	A	T	32,5	5'UTR		4168	A	T	0,4	NS2B	E49V
42	G	C	6,4	5'UTR		4577	A	T	0,9	NS3	-
55	A	C	41,6	5'UTR		4803	A	G	1,0	NS3	-
62	T	A	2,4	5'UTR		5263	A	T	0,5	NS3	R248STOP
86	G	T	8,8	5'UTR		5304	T	G	0,6	NS3	C261W
88	A	T	8,1	5'UTR		5424	A	T	0,8	NS3	-
90	A	C	8,8	5'UTR		5470	A	T	0,5	NS3	T317S
95	T	A	2,3	5'UTR		5547	T	A	1,3	NS3	-
96	G	A	17,3	5'UTR		5577	G	T	1,0	NS3	-
99	G	T	10,0	CAP	M1I	5785	G	T	0,7	NS3	V422F
102	T	A	11,6	CAP	N2K	5899	A	T	1,9	NS3	R460STOP
119	C	A	1,8	CAP	A8E	5946	G	T	1,7	NS3	M475I
138	T	A	2,7	CAP	N14K	6104	A	C	1,2	NS3	E528A
165	G	C	2,6	CAP	-	6258	G	T	0,7	NS3	-
171	T	A	3,2	CAP	-	6260	A	G	0,5	NS3	E580G
181	C	T	2,0	CAP	-	6285	G	A	0,7	NS3	-
187	A	C	12,2	CAP	K31H	6413	T	A	1,7	NS4A	L13H
189	G	C	2,2	CAP	-	6418	A	G	1,2	NS4A	T15A
200	T	A	2,0	CAP	L35Q	6455	A	T	0,7	NS4A	N27I
201	T	A	2,2	CAP	-	6888	G	T	0,6	NS4B	P21H
203	G	C	3,1	CAP	G36A	6889	G	T	1,0	NS4B	E22STOP
249	G	A	1,9	CAP	-	6913	A	G	0,5	NS4B	I30V
252	G	A	4,5	CAP	-	7064	A	G	0,9	NS4B	K80R
284	C	T	4,3	CAP	A63V	7108	G	T	0,7	NS4B	A95S
285	A	T	0,5	CAP	-	7239	A	T	0,7	NS4B	-
308	C	T	0,8	CAP	T71M	7386	A	T	1,1	NS4B	-
383	A	T	0,6	CAP	N96I	7453	G	T	0,7	NS4B	G209C
403	G	T	0,4	preMEM	C3G	7455	G	T	0,6	NS4B	-
526	G	T	0,6	MEM	G30C	7855	A	T	0,8	NS5	N96Y
701	G	C	0,6	MEM	R88T	7865	A	C	0,8	NS5	E99A
702	A	T	0,5	MEM	-	7883	A	T	3,2	NS5	K105I
761	A	T	0,3	MEM	E108V	7909	A	T	0,6	NS5	I114F
803	G	T	0,4	MEM	R122I	7952	G	C	0,8	NS5	S128T
996	G	T	0,6	ENV	W20C	8096	C	G	0,7	NS5	T176S
1039	G	T	0,6	ENV	A35S	8247	G	T	0,9	NS5	-
1096	G	T	0,9	ENV	A54S	8358	C	T	0,6	NS5	-
1117	A	T	0,4	ENV	I61L	8616	T	A	1,1	NS5	F349L
1280	C	G	0,5	ENV	T115S	8622	A	C	1,0	NS5	Q351H
1495	C	T	0,3	ENV	P187S	8700	G	A	1,5	NS5	-
1771	T	C	0,4	ENV	F279L	9133	G	A	0,7	NS5	V522M
1775	C	T	0,5	ENV	T280M	9389	A	T	3,8	NS5	Y607F
1832	A	T	0,7	ENV	Y299F	9393	C	T	3,5	NS5	-
2286	G	T	0,9	ENV	-	9415	G	A	7,0	NS5	E616K
2349	T	C	11,8	ENV	-	9439	G	A	6,2	NS5	E624K
2365	G	A	0,5	ENV	V477M	9443	G	T	1,5	NS5	G625V
2386	G	C	0,5	ENV	V484L	9481	A	C	1,4	NS5	T638P
2523	C	A	0,4	NS1	F34L	9489	A	T	6,4	NS5	E640D
2590	A	T	0,6	NS1	I57F	9504	C	T	1,6	NS5	-
2805	T	A	0,3	NS1	-	9530	G	T	0,7	NS5	R654M
2845	G	T	0,4	NS1	E142STOP	9541	A	T	4,1	NS5	M658L
3188	A	T	0,7	NS1	Y256F	9550	A	T	2,5	NS5	S661C
3513	C	T	2,0	NS2A	-	9642	A	C	0,5	NS5	-
3624	T	C	10,3	NS2A	-	9656	C	T	0,4	NS5	P696L
3645	G	T	0,5	NS2A	-	9658	T	C	0,8	NS5	S697P
3760	A	T	0,5	NS2A	K95STOP	9749	C	A	0,8	NS5	P727Q
3855	G	T	0,5	NS2A	-	9837	G	T	0,5	NS5	K756N
3974	A	T	0,4	NS2A	K166M	9999	G	T	0,4	NS5	-
4121	G	T	1,0	NS2A	S215I	10057	G	T	0,8	NS5	V830L
4161	A	T	0,3	NS2B	M13L	10357	G	T	1,5	3'UTR	
						10711	A	T	34,3	3'UTR	

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
17	G	T	58,537	5'UTR		689	G	C	2,8	MEM	G84A
27	A	C	10,723	5'UTR		698	G	C	2,0	MEM	R87T
41	A	T	2,2592	5'UTR		732	T	A	2,4	MEM	H98Q
55	A	C	41,298	5'UTR		740	T	A	2,9	MEM	M101K
63	T	A	9,1716	5'UTR		743	G	C	2,7	MEM	G102A
86	G	T	0,4531	5'UTR		762	A	T	2,3	MEM	E108D
88	A	T	0,5111	5'UTR		764	C	G	2,4	MEM	T109R
90	A	C	0,5493	5'UTR		821	T	A	1,8	MEM	L128STOP
96	G	A	3,2461	5'UTR		845	T	A	1,4	MEM	M136K
99	G	T	24,968	CAP	M1I	883	T	A	1,9	MEM	F149N
102	T	A	15,542	CAP	N2K	884	T	A	1,9	MEM	-
125	G	C	0,4562	CAP	S10T	892	G	T	2,4	MEM	A152S
135	C	T	1,561	CAP	-	914	C	G	2,4	MEM	T159R
165	G	T	1,8516	CAP	-	917	C	G	2,8	MEM	A160G
169	A	G	0,9145	CAP	T25A	923	C	G	2,7	MEM	A162G
181	C	T	4,3215	CAP	-	950	G	C	2,5	ENV	G5A
187	A	C	18,72	CAP	K31Q	962	G	C	2,0	ENV	R9T
198	A	T	1,8851	CAP	-	986	G	T	2,5	ENV	G17V
248	T	C	1,6375	CAP	V51A	1021	A	T	3,1	ENV	S29C
249	G	A	1,4318	CAP	-	1102	C	G	1,8	ENV	L56V
252	G	A	3,8513	CAP	-	1105	A	T	2,3	ENV	R57W
209	T	A	3,0	CAP	-	1108	A	T	1,6	ENV	K58L
214	G	C	2,7	CAP	G40R	1109	A	T	1,9	ENV	K58L
235	T	A	2,1	CAP	F47I	1111	T	A	2,1	ENV	Y59I
240	G	C	2,2	CAP	M48I	1112	A	T	2,8	ENV	-
246	T	A	2,2	CAP	-	1123	G	C	1,8	ENV	A63P
252	G	A	5,0	CAP	-	1164	A	T	2,9	ENV	-
263	T	A	2,1	CAP	F56Y	1169	G	C	2,7	ENV	G78A
343	G	T	2,5	CAP	G83W	1174	C	G	2,2	ENV	P80A
378	C	G	2,5	CAP	I94M	1180	C	G	2,8	ENV	L82V
383	A	T	2,3	CAP	N96I	1197	C	G	2,6	ENV	D87E
396	G	T	1,9	CAP	R100S	1200	A	T	3,0	ENV	K88N
409	A	T	2,6	preMEM	F5I	1231	A	T	4,3	ENV	R99STOP
452	C	G	2,8	MEM	T5R	1290	G	C	2,8	ENV	M118I
488	A	T	2,4	MEM	Q17L	1295	C	G	2,1	ENV	T120R
495	A	T	2,0	MEM	K19N	1354	G	C	3,8	ENV	V140L
502	A	T	1,6	MEM	S22C	1367	A	T	2,5	ENV	H144L
515	A	T	1,4	MEM	K26I	1391	G	C	2,2	ENV	G152R
517	A	T	2,3	MEM	T27C	1402	G	C	2,0	ENV	G156R
518	C	G	1,6	MEM	-	1442	C	A	2,5	ENV	S169Y
519	A	T	1,6	MEM	-	1486	G	C	2,6	ENV	E184L
524	A	T	1,4	MEM	D29V	1487	A	T	3,9	ENV	-
538	T	A	2,1	MEM	-	1489	T	A	2,5	ENV	C185S
539	G	C	2,3	MEM	C34T	1519	G	C	2,7	ENV	E195Q
549	G	C	2,2	MEM	M37I	1522	A	G	2,6	ENV	M196E
576	A	T	3,0	MEM	E46D	1523	T	A	2,1	ENV	-
580	A	T	3,0	MEM	T48S	1571	G	T	2,6	ENV	W212L
582	A	T	4,1	MEM	-	1613	C	G	3,2	ENV	T226R
590	A	T	3,0	MEM	Y51F	1619	G	C	2,7	ENV	G228A
631	G	C	2,7	MEM	D65H	1625	A	T	4,5	ENV	N230I
676	T	C	3,3	MEM	C80R	1696	G	C	1,6	ENV	G254R
678	T	A	3,8	MEM	-	1698	G	T	1,7	ENV	-

## 149 continuação

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
1700	C	G	2,3	ENV	S255C	3529	T	A	3,2	NS2A	F18I
1702	C	G	2,2	ENV	Q256E	3586	G	C	1,7	NS2A	A37P
1706	A	T	2,1	ENV	E257V	3600	G	T	2,2	NS2A	-
1752	G	C	1,8	ENV	M272I	3679	A	T	1,9	NS2A	I68L
1758	A	T	2,9	ENV	-	3688	G	C	1,8	NS2A	-
1829	C	G	3,3	ENV	S298STOP	3689	G	C	2,0	NS2A	G71P
1863	G	C	2,2	ENV	-	3695	C	G	2,0	NS2A	T73S
1916	A	T	2,5	ENV	E327V	3699	T	A	2,0	NS2A	Y74STOP
2020	G	C	6,7	ENV	D362H	3740	C	A	2,7	NS2A	A88E
2035	A	T	3,3	ENV	I367L	3743	C	A	2,4	NS2A	A89D
2356	T	C	1,9	ENV	S474P	3756	G	C	4,6	NS2A	L93F
2464	T	A	2,2	NS1	C15S	3758	G	C	2,0	NS2A	R94T
2492	A	T	3,8	NS1	N24I	3782	T	A	2,1	NS2A	M102K
2495	T	A	2,6	NS1	V25E	3825	G	C	2,4	NS2A	L116F
2501	C	A	2,1	NS1	T27K	3838	C	A	2,3	NS2A	L121K
2505	G	C	3,7	NS1	W28C	3839	T	A	1,8	NS2A	L121K
2511	G	C	2,6	NS1	E30D	3842	A	T	2,4	NS2A	E122V
2551	T	C	3,2	NS1	S44P	3848	C	G	2,2	NS2A	T124S
2557	A	G	3,0	NS1	I46V	3928	G	C	2,3	NS2A	A151P
2598	A	T	2,6	NS1	-	3973	A	T	1,4	NS2A	K166L
2605	A	T	3,1	NS1	R62C	3974	A	T	1,7	NS2A	K166L
2607	A	T	4,0	NS1	-	3979	A	T	1,4	NS2A	S168C
2612	A	T	3,1	NS1	E64V	3984	C	G	1,6	NS2A	C169W
2615	A	T	3,8	NS1	N65I	3985	A	T	2,5	NS2A	T170S
2632	A	T	3,1	NS1	I71L	4057	C	G	1,9	NS2A	L194V
2637	A	T	4,0	NS1	-	4107	T	A	2,6	NS2A	-
2643	A	T	4,3	NS1	E74D	4112	C	G	3,0	NS2A	S212C
2665	A	T	4,0	NS1	N82Y	4113	G	C	3,6	NS2A	-
2702	A	T	2,1	NS1	K94I	4205	A	T	3,4	NS2B	D25V
2722	A	T	2,6	NS1	K101STOP	4219	G	T	2,1	NS2B	G30C
2786	A	T	1,7	NS1	K122I	4336	G	T	2,6	NS2B	G69STOP
2788	A	T	2,4	NS1	M123L	4355	C	G	2,1	NS2B	S75STOP
2797	A	T	2,9	NS1	T126S	4381	A	T	2,6	NS2B	M84L
2804	C	G	3,6	NS1	S128C	4384	T	A	1,9	NS2B	S85T
2807	A	T	2,6	NS1	H129L	4435	G	T	3,0	NS2B	G102STOP
2834	C	G	3,4	NS1	P138R	4443	G	T	2,3	NS2B	L104F
2840	C	G	2,4	NS1	T140R	4462	C	G	3,8	NS2B	P111G
2987	A	T	2,4	NS1	K189I	4463	C	G	3,8	NS2B	P111G
3043	G	C	2,5	NS1	D208H	4469	C	G	2,6	NS2B	S113W
3046	A	C	3,2	NS1	T209P	4538	A	T	2,4	NS3	D6V
3054	G	C	2,3	NS1	K211N	4544	C	G	3,3	NS3	P8R
3057	G	C	2,9	NS1	M212I	4577	A	T	3,9	NS3	E19V
3124	G	C	4,0	NS1	G235R	4589	A	T	2,8	NS3	Y23F
3201	C	A	3,6	NS1	Y260STOP	4668	G	C	2,5	NS3	M49I
3215	C	A	3,7	NS1	A265E	4728	G	C	2,5	NS3	W69C
3272	C	G	3,1	NS1	T284R	4786	T	A	3,2	NS3	W89R
3296	G	C	2,4	NS1	G292A	4795	G	C	3,3	NS3	G92R
3299	A	T	3,4	NS1	N293I	4798	G	C	3,2	NS3	-
3307	C	G	2,9	NS1	P296A	4860	T	A	6,6	NS3	-
3338	A	T	2,5	NS1	K306M	4896	T	A	1,9	NS3	-
3388	T	A	1,7	NS1	Y323N	4898	T	A	2,3	NS3	V126E
3442	A	T	2,3	NS1	K341STOP	4904	T	A	1,2	NS3	L128Q
3447	A	T	3,7	NS1	E342D	4906	G	C	2,1	NS3	D129H
3459	C	G	4,1	NS1	-	4909	T	A	1,6	NS3	F130I
3524	C	G	2,9	NS2A	A16G	4912	T	A	1,5	NS3	S131T

## 149 continuação ii

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
4931	C	A	2,4	NS3	S137Y	5826	A	T	4,4	NS3	-
4933	C	A	3,7	NS3	P138T	5832	C	G	3,4	NS3	-
4937	T	A	3,9	NS3	I139N	5835	A	T	3,4	NS3	E438D
4941	C	A	3,1	NS3	-	5852	C	G	3,8	NS3	A444G
4976	A	T	1,8	NS3	N152I	5870	C	G	3,3	NS3	T450S
4991	G	C	1,6	NS3	R157T	5931	C	G	6,9	NS3	D470E
4994	G	T	1,4	NS3	S158I	5940	A	T	5,5	NS3	-
5010	T	A	1,6	NS3	S163R	6132	G	C	3,8	NS3	M537I
5038	G	C	1,8	NS3	E173Q	6139	G	T	5,3	NS3	G540C
5041	G	C	1,9	NS3	D174H	6141	G	C	4,8	NS3	-
5047	C	G	2,3	NS3	P176A	6147	A	T	3,7	NS3	L542F
5086	A	T	3,4	NS3	T189S	6171	G	C	3,5	NS3	-
5088	C	A	2,2	NS3	-	6194	C	T	3,3	NS3	A558V
5092	A	T	14,7	NS3	M191L	6241	G	T	2,5	NS3	E574STOP
5103	T	A	2,0	NS3	H194Q	6324	T	A	6,1	NS3	Y601STOP
5136	A	T	5,4	NS3	-	6325	T	A	2,7	NS3	S602T
5183	T	A	1,8	NS3	L221Q	6327	T	A	3,0	NS3	-
5186	C	G	2,9	NS3	A222G	6330	T	A	3,8	NS3	D603E
5192	C	G	3,5	NS3	T224S	6334	C	T	2,3	NS3	-
5228	G	T	1,5	NS3	R236I	6351	C	A	3,0	NS3	F610L
5233	C	A	1,8	NS3	L238I	6381	G	C	5,1	NS4A	L2F
5283	A	T	2,2	NS3	-	6385	C	A	6,2	NS4A	L4M
5289	C	G	2,2	NS3	I256M	6400	G	C	3,6	NS4A	E9Q
5306	A	T	1,8	NS3	H262L	6434	A	T	6,6	NS4A	K20M
5309	C	A	3,3	NS3	A263E	6457	T	A	3,3	NS4A	L28M
5310	C	G	2,3	NS3	-	6579	C	G	3,1	NS4A	I68M
5312	C	G	1,9	NS3	T264R	6582	C	G	4,2	NS4A	F69L
5321	T	A	2,0	NS3	M267K	6595	A	T	2,9	NS4A	S74C
5324	G	C	2,0	NS3	R268T	6601	A	T	3,6	NS4A	K76STOP
5326	C	A	2,1	NS3	L269M	6607	A	T	5,0	NS4A	I78L
5330	T	A	2,1	NS3	L270Q	6637	A	T	1,9	NS4A	I87L
5380	C	G	1,5	NS3	H287D	6645	A	T	2,2	NS4A	S91R
5401	A	T	1,5	NS3	I294L	6651	C	G	3,0	NS4A	-
5439	G	C	3,4	NS3	M306I	6657	C	G	3,7	NS4A	-
5484	C	G	2,1	NS3	S321R	6660	A	T	1,9	NS4A	-
5485	A	T	3,2	NS3	R322STOP	6682	C	A	3,8	NS4A	H102N
5516	T	A	2,6	NS3	I332N	6736	C	G	4,7	NS4A	P120A
5531	G	T	2,4	NS3	R337I	6793	G	C	2,9	P2K	A13P
5566	G	T	2,5	NS3	E349STOP	6799	C	G	4,9	P2K	L15V
5576	C	G	3,5	NS3	T352R	6831	G	C	3,0	NS4B	E2D
5614	A	T	3,1	NS3	I365L	6839	T	A	3,4	NS4B	F5Y
5625	A	T	6,2	NS3	-	6867	A	T	19,5	NS4B	-
5630	A	T	2,6	NS3	D370V	6879	T	A	3,0	NS4B	-
5654	A	T	2,9	NS3	N378I	6894	T	A	1,9	NS4B	-
5659	A	T	2,3	NS3	K380STOP	6922	C	G	3,2	NS4B	R33G
5664	A	T	2,9	NS3	K381N	6926	C	A	2,8	NS4B	P34Q
5668	A	T	3,6	NS3	I383L	6927	T	A	4,5	NS4B	A35G
5673	A	T	2,8	NS3	Q384H	6929	C	G	3,8	NS4B	-
5684	A	T	3,1	NS3	K388M	6931	T	A	2,0	NS4B	S36T
5694	C	G	3,7	NS3	D391E	6935	C	T	2,7	NS4B	A37V
5699	A	T	2,4	NS3	E393V	6948	T	A	2,6	NS4B	Y41STOP
5700	G	C	3,0	NS3	-	6949	G	C	2,3	NS4B	A42P
5794	C	G	3,8	NS3	P425A	6965	T	A	2,5	NS4B	F47Y
5820	A	T	3,0	NS3	-	6969	C	G	4,8	NS4B	-
5824	A	T	3,1	NS3	T435S	6977	T	A	2,4	NS4B	M51K

149 continuação iii											
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
6982	C	T	4,8	NS4B	R53STOP	8713	T	A	3,7	NS5	W382R
6992	T	A	2,8	NS4B	I56N	8725	G	C	10,5	NS5	G386R
7012	G	C	3,3	NS4B	V63L	8731	G	C	4,0	NS5	E388Q
7041	T	A	2,1	NS4B	-	8750	G	T	3,4	NS5	C394F
7186	A	T	4,6	NS4B	I121L	8754	C	T	3,0	NS5	-
7196	G	T	3,0	NS4B	G124V	8806	T	A	3,4	NS5	F413I
7218	A	T	4,4	NS4B	R131S	8838	T	A	4,0	NS5	-
7251	G	T	3,2	NS4B	M142I	8882	A	T	4,5	NS5	E438V
7273	A	T	3,1	NS4B	I150L	8916	A	T	4,1	NS5	-
7282	A	T	3,2	NS4B	I153F	8985	A	T	3,8	NS5	R472S
7288	C	G	3,0	NS4B	L155V	8988	C	G	3,7	NS5	-
7290	A	T	3,2	NS4B	-	9008	G	C	5,3	NS5	G480A
7333	C	G	2,2	NS4B	Q170E	9106	C	G	1,9	NS5	H513D
7339	A	T	3,2	NS4B	M172L	9112	C	G	3,3	NS5	L515V
7345	C	G	3,4	NS4B	L174V	9116	G	T	2,6	NS5	G516V
7371	A	T	3,3	NS4B	L182F	9127	A	T	4,6	NS5	R520STOP
7372	A	T	3,7	NS4B	M183L	9143	A	T	6,6	NS5	K525M
7378	A	T	2,9	NS4B	R185R	9180	A	T	5,8	NS5	-
7380	G	T	2,4	NS4B	-	9192	A	T	12,8	NS5	R541S
7382	C	G	2,4	NS4B	T186S	9195	C	G	13,8	NS5	I542M
7384	A	T	4,3	NS4B	T187STOP	9349	T	A	3,6	NS5	S594T
7385	C	G	4,6	NS4B	-	9424	T	A	2,7	NS5	L619I
7391	C	G	3,2	NS4B	A189G	9428	T	A	3,4	NS5	I620N
7439	G	C	4,5	NS4B	W204S	9431	G	C	6,3	NS5	R621T
7472	C	G	3,4	NS4B	T215S	9435	G	T	3,2	NS5	Q622H
7791	A	T	5,3	NS5	E74D	9439	G	A	7,8	NS5	E624K
7807	C	A	4,9	NS5	L80I	9443	G	T	3,9	NS5	G625V
7841	G	C	5,4	NS5	C91S	9445	G	T	3,5	NS5	E626Y
7844	G	C	4,9	NS5	G92A	9447	A	T	2,6	NS5	-
7847	G	C	5,2	NS5	G93A	9448	G	C	2,1	NS5	G627R
7850	T	A	3,9	NS5	L94Q	9490	A	T	4,0	NS5	T641S
7885	G	C	4,7	NS5	G106R	9530	G	T	4,2	NS5	R654M
7982	A	T	4,5	NS5	E138V	9536	C	T	8,3	NS5	S656L
7983	A	T	5,4	NS5	-	9563	G	C	2,8	NS5	C665S
7992	T	A	4,3	NS5	D141E	9564	T	A	3,8	NS5	-
7995	A	T	4,0	NS5	-	9566	T	A	3,0	NS5	V666D
8019	G	C	8,6	NS5	-	9568	G	C	2,2	NS5	V667L
8053	C	G	3,5	NS5	L162V	9571	A	T	2,3	NS5	K668STOP
8057	G	C	3,4	NS5	R163T	9575	C	A	2,1	NS5	P669Q
8063	T	A	3,8	NS5	L165H	9576	T	A	3,1	NS5	-
8075	A	G	5,7	NS5	E169G	9578	T	A	1,9	NS5	I670K
8112	G	C	2,6	NS5	K181N	9580	G	C	3,1	NS5	D671H
8115	C	A	3,2	NS5	-	9585	C	G	2,1	NS5	D672E
8127	T	A	2,5	NS5	Y186STOP	9592	G	A	7,7	NS5	A675T
8175	A	T	3,4	NS5	-	9597	T	A	1,6	NS5	S676R
8248	A	T	4,1	NS5	N227Y	9616	G	C	4,9	NS5	D683H
8251	A	T	3,0	NS5	I228L	9641	T	A	3,2	NS5	I691K
8398	A	T	5,5	NS5	I277L	9649	T	A	2,5	NS5	W694T
8403	A	T	2,8	NS5	-	9650	G	C	3,8	NS5	-
8546	G	T	2,9	NS5	R326I	9651	G	C	2,1	NS5	-
8562	T	A	2,9	NS5	-	9659	C	G	1,6	NS5	S697STOP
8598	G	C	4,3	NS5	M343I	9686	T	A	2,4	NS5	V706E
8659	C	A	4,0	NS5	Q364K	9692	T	A	1,9	NS5	F708Y
8690	T	A	5,7	NS5	M374K	9702	C	A	2,9	NS5	H711Q
8708	G	C	7,0	NS5	W380S	9706	T	A	3,1	NS5	F713I

## 149 continuação iv

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
9712	G	C	2,5	NS5	E715Q	10092	A	T	2,4	NS5	-
9745	G	C	2,4	NS5	V726L	10095	G	T	2,6	NS5	R842S
9749	C	A	2,1	NS5	P727Q	10097	A	T	1,9	NS5	E843V
9751	T	A	2,3	NS5	C728R	10100	A	T	2,6	NS5	D844V
9753	C	G	1,8	NS5	R729T	10103	A	T	3,5	NS5	Q845L
9755	G	C	1,7	NS5	-	10104	A	T	3,6	NS5	-
9763	G	C	2,8	NS5	D732H	10116	A	T	1,9	NS5	-
9839	C	A	1,8	NS5	S757Y	10120	A	T	2,0	NS5	I851F
9843	C	G	2,0	NS5	Y758STOP	10130	C	G	2,4	NS5	T854R
9847	C	A	2,2	NS5	Q760N	10132	A	T	2,6	NS5	S855C
9849	G	C	1,9	NS5	-	10138	G	C	2,1	NS5	A857P
9855	G	C	2,6	NS5	W762C	10172	A	T	2,5	NS5	N868I
9880	G	T	1,4	NS5	D771STOP	10176	A	T	1,7	NS5	Q869H
9882	C	G	2,0	NS5	-	10182	A	T	2,0	NS5	R871S
9884	T	A	1,3	NS5	L772H	10192	G	C	2,2	NS5	G875R
9901	G	C	2,4	NS5	A778P	10195	A	T	2,2	NS5	N876Y
9919	C	G	3,1	NS5	P784A	10198	G	T	2,0	NS5	E877STOP
9936	A	T	1,7	NS5	-	10202	G	C	1,9	NS5	G878A
9939	A	T	1,6	NS5	-	10205	A	T	1,8	NS5	Y879F
9945	A	T	2,1	NS5	-	10214	A	T	3,2	NS5	Y882L
9958	A	T	2,3	NS5	I797L	10215	C	G	2,9	NS5	-
9965	C	G	2,9	NS5	A799G	10222	T	A	2,6	NS5	S885T
9979	A	T	2,4	NS5	M804L	10239	A	T	2,2	NS5	R890S
10005	C	G	1,8	NS5	-	10254	G	C	5,0	NS5	E895D
10071	A	T	3,1	NS5	E834D	10257	A	T	3,8	NS5	-
10077	C	G	4,6	NS5	-	10496	C	G	5,9	3'UTR	
10091	A	T	3,0	NS5	K841I	10641	C	G	3,6	3'UTR	



POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	27,2	5'UTR		3043	G	A	10,2	NS1	D208H
82	G	C	3,4	5'UTR		3046	A	C	8,2	NS1	T209P
95	T	A	2,7	5'UTR		3307	C	G	3,6	NS1	P296A
99	G	T	6,9	CAP	M1I	3317	G	C	3,8	NS1	R299T
138	T	A	2,7	CAP	N14K	3400	G	C	3,0	NS1	D327H
187	A	C	8,8	CAP	K31Q	3432	T	A	3,9	NS1	-
200	T	A	2,8	CAP	L35Q	3688	G	C	1,9	NS2A	G71P
201	T	A	2,6	CAP	-	3689	G	C	2,4	NS2A	-
203	G	C	3,9	CAP	G36A	3985	A	T	2,9	NS2A	T170S
246	T	A	2,9	CAP	-	4121	G	T	6,0	NS2A	S215L
497	G	T	3,7	MEM	G20V	4199	A	T	2,6	NS2B	K23M
517	A	T	3,5	MEM	T27S	4202	A	T	2,8	NS2B	N24I
519	A	T	3,6	MEM	-	4216	A	T	3,0	NS2B	T29S
523	G	C	1,9	MEM	D29H	4264	A	T	3,4	NS2B	T45S
529	G	C	2,8	MEM	V31L	4270	C	G	2,4	NS2B	R47G
532	A	T	3,5	MEM	N32Y	4378	A	T	5,7	NS2B	S83C
628	A	T	10,2	MEM	I64L	4413	A	T	3,6	NS2B	-
946	A	T	3,2	ENV	I4L	4646	A	T	3,4	NS3	K42I
960	T	A	2,4	ENV	N8K	4661	A	T	2,5	NS3	H47L
1067	A	T	4,4	ENV	E44V	4668	G	C	2,6	NS3	M49I
1102	C	G	3,8	ENV	L56V	4743	G	A	9,7	NS3	-
1164	A	T	5,3	ENV	-	4810	G	C	3,2	NS3	V97L
1168	G	C	3,5	ENV	G78P	4816	G	C	3,1	NS3	A99P
1169	G	C	3,3	ENV	-	4842	A	T	4,8	NS3	R107S
1235	G	C	2,9	ENV	G100A	4848	C	G	3,6	NS3	-
1240	G	C	3,9	ENV	G102R	4898	T	A	3,5	NS3	V126E
1271	G	T	2,8	ENV	G112V	4906	G	C	3,4	NS3	D129H
1289	T	A	3,8	ENV	M118K	4909	T	A	2,5	NS3	F130I
1323	C	T	3,8	ENV	-	4912	T	A	2,3	NS3	S131T
1355	T	A	2,7	ENV	V140E	4933	C	A	3,8	NS3	P138T
1377	A	T	2,4	ENV	E147D	4937	T	A	4,5	NS3	I139N
1478	T	A	2,8	ENV	V181D	4943	A	T	3,6	NS3	D141V
1706	A	T	2,3	ENV	E257V	5010	T	A	2,4	NS3	S163R
1820	A	T	4,3	ENV	K295I	5043	C	A	2,7	NS3	D174E
1829	C	G	4,9	ENV	S298STOP	5065	A	T	2,8	NS3	I182F
1832	A	T	4,1	ENV	Y299F	5092	A	T	10,5	NS3	M191L
2295	G	T	2,9	ENV	W453C	5191	A	T	3,5	NS3	T224C
2301	G	C	3,7	ENV	M455I	5192	C	G	4,3	NS3	-
2411	T	A	2,8	ENV	M492K	5324	G	C	2,7	NS3	R268T
2424	T	A	2,8	NS1	D1E	5326	C	A	2,1	NS3	L269M
2484	T	A	5,0	NS1	-	5330	T	A	3,4	NS3	L270Q
2501	C	A	3,4	NS1	T27K	5485	A	T	3,2	NS3	R322STOP
2505	G	C	6,4	NS1	W28C	5504	G	C	4,1	NS3	S328T
2509	G	T	2,3	NS1	E30Y	5509	G	C	2,8	NS3	A330P
2511	G	C	3,8	NS1	-	5531	G	T	3,1	NS3	R337I
2517	T	A	3,0	NS1	Y32STOP	5562	A	T	4,1	NS3	-
2563	A	T	4,5	NS1	K48STOP	5659	A	T	3,0	NS3	K380STOP
2795	C	G	1,8	NS1	S125C	5694	C	G	3,3	NS3	D391E
2807	A	T	3,2	NS1	H129L	5717	C	G	4,4	NS3	A399G
2812	C	G	2,5	NS1	Q131E	5718	C	G	5,3	NS3	-
2815	A	T	4,1	NS1	T132S	5781	G	T	3,1	NS3	E420D
2824	A	T	2,8	NS1	I135F	5809	A	T	4,0	NS3	K430STOP
2966	A	T	7,3	NS1	K182I	5839	C	T	5,0	NS3	R440W

151 continuação											
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
5886	G	C	6,2	NS3	-	8991	A	T	6,2	NS5	-
6324	T	A	4,8	NS3	Y601STOP	9069	A	T	3,1	NS5	R500S
6330	T	A	3,4	NS3	D603E	9078	C	G	3,6	NS5	-
6335	T	A	3,4	NS3	L605Q	9340	G	C	6,5	NS5	D591H
6404	T	A	5,1	NS4A	M10K	9405	C	G	4,9	NS5	F612L
6600	A	T	3,6	NS4A	-	9489	A	T	4,6	NS5	E640D
6651	C	G	3,4	NS4A	S91R	9497	T	A	4,1	NS5	V643E
6657	C	G	5,2	NS4A	-	9564	T	A	3,2	NS5	C665STOP
6789	C	G	3,4	P2K	-	9583	G	C	2,2	NS5	D672H
6922	C	T	4,2	NS4B	R33C	9592	G	A	6,6	NS5	A675T
6946	T	A	3,6	NS4B	Y41K	9602	T	A	5,5	NS5	L678STOP
6948	T	A	3,1	NS4B	-	9616	G	C	2,7	NS5	D683H
6949	G	C	3,5	NS4B	A42P	9649	T	A	2,6	NS5	W694R
6955	G	C	4,0	NS4B	A44P	9658	T	C	2,9	NS5	S697P
6969	C	G	3,7	NS4B	-	9684	G	T	2,5	NS5	Q705H
7028	T	A	3,5	NS4B	I68K	9777	C	G	2,9	NS5	Q742H
7046	T	A	5,1	NS4B	V74E	9795	G	T	4,0	NS5	-
7273	A	T	3,8	NS4B	I150L	9875	G	T	2,5	NS5	R769I
7282	A	T	3,0	NS4B	I153F	9926	A	T	3,1	NS5	H786L
7288	C	G	5,4	NS4B	L155V	9946	A	T	3,4	NS5	T793S
7368	A	T	3,0	NS4B	-	9958	A	T	2,3	NS5	I797L
7835	A	T	5,9	NS5	Y89F	9993	C	G	3,8	NS5	D808E
7995	A	T	6,1	NS5	-	10005	C	G	5,8	NS5	-
8030	C	A	4,3	NS5	P154H	10015	G	T	3,5	NS5	V816L
8103	C	A	4,9	NS5	F178L	10071	A	T	2,9	NS5	E834D
8570	C	T	13,7	NS5	V334A	10198	G	T	2,8	NS5	E877STOP
8575	C	G	3,1	NS5	P336A	10212	C	G	4,5	NS5	D881E
8819	A	T	4,0	NS5	N417I	10217	T	A	3,1	NS5	M883K
8882	A	T	3,4	NS5	E438V	10224	C	G	5,2	NS5	-
8912	A	T	3,6	NS5	E448V	10339	A	T	7,4	3'UTR	-
8913	A	T	4,5	NS5	-	10656	C	G	24,0	3'UTR	-

153											
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	49,0	5'UTR	-	4996	G	C	4,4	NS3	G159R
63	T	A	15,4	5'UTR	-	5447	C	T	12,9	NS3	A309V
946	A	T	5,8	ENV	I4L	5485	A	T	6,0	NS3	R322STOP
1289	T	A	8,7	ENV	M118K	6645	A	T	4,6	NS4A	-
1489	T	A	3,7	ENV	C185S	6657	C	G	6,7	NS4A	-
2505	G	C	5,1	NS1	W28C	7288	C	G	12,2	NS4B	L155V
2650	C	G	7,5	NS1	H77D	8059	G	T	10,9	NS5	V164F
3083	A	T	4,1	NS1	K221I	8756	G	A	6,9	NS5	R396K
3501	T	C	6,7	NS2A	-	9592	G	A	14,3	NS5	A675T

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
27	A	C	8,9	5'UTR		6198	C	A	0,7	NS3	D559E
55	A	C	46,3	5'UTR		6242	A	T	0,6	NS3	E574V
63	T	A	13,8	5'UTR		6258	G	T	1,1	NS3	-
96	G	A	2,4	5'UTR		6336	G	A	12,9	NS3	-
99	G	T	25,9	CAP	M1I	6439	A	T	0,6	NS4A	R22STOP
102	T	A	14,7	CAP	N2K	6455	A	T	0,9	NS4A	N27I
135	C	T	0,8	CAP	-	6618	G	T	1,0	NS4A	M80I
165	G	T	1,2	CAP	-	6889	G	T	0,5	NS4B	E22STOP
181	C	T	3,0	CAP	-	7108	G	T	0,4	NS4B	A95S
187	A	C	8,3	CAP	K31Q	7239	A	T	0,6	NS4B	-
198	A	T	0,7	CAP	-	7386	A	T	0,6	NS4B	-
248	T	C	0,6	CAP	-	7750	A	T	0,8	NS5	K61STOP
249	G	A	0,6	CAP	-	7838	A	T	1,0	NS5	Y90F
252	G	A	1,9	CAP	V52A	7883	A	T	1,2	NS5	K105I
284	C	T	3,7	CAP	A63V	7886	G	T	1,4	NS5	G106V
285	A	T	0,7	CAP	-	7888	G	T	1,4	NS5	G107STOP
305	G	A	0,9	CAP	G70E	8075	A	G	1,8	NS5	E169G
336	T	C	2,3	CAP	-	8137	G	A	1,3	NS5	V190I
526	G	T	0,6	MEM	G30C	8546	G	T	0,8	NS5	R326I
1096	G	T	0,5	ENV	A54S	8621	A	G	0,9	NS5	Q351R
1323	C	A	0,6	ENV	-	8680	A	T	1,2	NS5	K371STOP
1385	C	T	2,4	ENV	A150V	8683	A	G	1,2	NS5	K372E
1679	A	T	0,5	ENV	Q248L	8990	T	C	0,9	NS5	I474T
1832	A	T	0,6	ENV	Y299F	9000	G	A	0,8	NS5	M477I
1936	A	T	0,5	ENV	K334STOP	9389	A	T	2,1	NS5	Y607F
2356	T	C	1,1	ENV	S474P	9393	C	T	1,7	NS5	-
2967	A	G	1,3	NS1	-	9415	G	A	2,3	NS5	E616K
3504	A	T	0,7	NS2A	-	9439	G	A	3,7	NS5	E624K
3654	G	A	11,0	NS2A	M59I	9443	G	T	0,9	NS5	G625V
3858	A	G	0,2	NS2A	-	9489	A	T	4,8	NS5	E640D
4121	G	T	1,3	NS2A	S215I	9536	C	T	1,1	NS5	S656L
4168	A	T	0,6	NS2B	M13L	9658	T	C	0,7	NS5	S697P
4443	G	T	1,3	NS2B	L104F	9749	C	A	0,5	NS5	P727Q
4562	G	A	2,1	NS3	G14E	10214	A	T	0,5	NS5	Y882F
5440	G	T	1,0	NS3	G307C	10322	A	T	1,0	3'UTR	
5642	G	T	0,6	NS3	C374F	10326	T	A	1,2	3'UTR	
5649	G	A	1,1	NS3	-	10541	G	T	1,3	3'UTR	
5839	C	T	1,1	NS3	R440W	10577	G	T	1,3	3'UTR	

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
41	A	T	1,3	5'UTR		5309	C	A	0,7	NS3	A263D
55	A	C	18,1	5'UTR		5342	G	T	0,5	NS3	R274I
63	T	A	3,8	5'UTR		5589	G	C	0,5	NS3	-
96	G	A	1,9	5'UTR		5722	G	C	0,6	NS3	D401H
99	G	T	19,7	CAP	M1I	5833	G	T	0,8	NS3	E438STOP
102	T	A	13,1	CAP	N2K	6026	A	T	1,1	NS3	E502V
135	C	T	0,7	CAP	-	6043	A	G	2,0	NS3	M508V
165	G	T	1,5	CAP	-	6374	A	T	0,8	NS3	K618I
169	A	G	0,6	CAP	T25A	6575	G	T	0,8	NS4A	G75V
181	C	T	2,5	CAP	-	7062	G	T	0,6	NS4B	-
187	A	C	8,6	CAP	K31Q	7108	G	T	0,4	NS4B	A95S
198	A	T	2,7	CAP	-	7386	A	T	0,5	NS4B	-
206	T	C	0,4	CAP	M37T	7631	G	T	2,7	NS5	G21V
213	A	C	0,3	CAP	Q39H	7838	A	T	0,8	NS5	Y90F
266	T	A	0,6	CAP	L57Q	8030	C	A	1,5	NS5	P154H
284	C	T	5,1	CAP	A63V	8034	G	A	0,8	NS5	-
293	T	A	0,5	CAP	L66Q	8063	T	A	1,3	NS5	L165H
305	G	A	1,2	CAP	G70E	8074	G	C	1,2	NS5	E169Q
383	A	T	0,6	CAP	N96I	8597	T	C	1,2	NS5	M343T
526	G	T	0,5	MEM	G30C	8613	A	G	1,4	NS5	-
762	A	T	0,5	MEM	E108D	8616	T	A	2,2	NS5	F349L
832	G	T	0,8	MEM	G132C	8621	A	G	0,7	NS5	Q351R
1083	A	T	0,7	ENV	E49D	8680	A	T	1,4	NS5	K371STOP
1992	T	C	0,7	ENV	-	8683	A	G	1,5	NS5	K372E
2145	G	A	0,8	ENV	-	8817	G	A	0,5	NS5	-
2216	G	T	0,6	ENV	G427V	9074	A	G	1,1	NS5	N502S
3053	A	C	1,3	NS1	K211T	9075	C	T	1,1	NS5	-
3159	T	A	1,0	NS1	N246K	9091	G	T	0,7	NS5	E508STOP
3475	G	T	0,7	NS1	A352S	9326	G	T	1,0	NS5	R586I
3517	G	T	0,6	NS2A	G14STOP	9340	G	A	1,0	NS5	D591N
3560	G	T	0,5	NS2A	G28V	9415	G	A	1,7	NS5	E616K
3601	A	C	0,4	NS2A	T42P	9439	G	A	4,9	NS5	E624K
3613	G	C	0,6	NS2A	G46R	9444	A	C	1,8	NS5	-
3743	C	T	0,7	NS2A	A89V	9466	C	A	0,6	NS5	Q633K
3890	G	T	0,4	NS2A	R138I	9540	A	G	1,1	NS5	-
4434	G	A	4,5	NS2B	-	9541	A	T	0,9	NS5	M658L
4443	G	T	1,3	NS2B	L104F	9550	A	T	0,9	NS5	S661C
4651	G	T	0,9	NS3	G44STOP	9554	G	A	0,5	NS5	G662E
5214	G	T	1,1	NS3	M231I	9658	T	C	0,5	NS5	S697P
5304	T	G	0,6	NS3	C261W	9954	G	T	0,4	NS5	W795C

156						157					
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
431	T	C	18,2	ORF (CAP)	V112A	6828	T	C	31,8	NS4B	-
1544	T	C	16,7	ENV	-	6863	C	T	58,3	NS4B	-
1586	G	A	61,5	ENV	-	6870	G	A	54,5	NS4B	G16R
1598	G	A	56,3	ENV	-	9910	A	G	23,1	NS5	-
1625	C	T	68,8	ENV	-						
1641	A	T	33,3	ENV	T236S						
2375	A	G	31,8	ENV	-						
2579	C	T	27,3	NS1	-						
2676	T	C	28,6	NS1	-						
2678	A	G	37,5	NS1	-						
3174	T	G	40,0	NS1	S252A						
3578	G	A	40,0	NS2A	-						
4248	G	A	28,0	NS2B	V40I						
4970	C	T	68,2	NS3	-						
6957	C	T	25,0	NS4B	-						
7455	G	A	36,0	NS4B	-						
9080	A	T	80,0	NS5	L504Q						

158											
POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	12,8	5'UTR		4797	A	G	0,6	NS3	-
96	G	A	5,9	5'UTR		4799	A	C	1,3	NS3	E93A
99	G	T	16,6	CAP	I1M	4864	A	C	2,0	NS3	I115L
135	C	T	5,5	CAP	-	4897	G	T	1,4	NS3	V126L
181	C	T	4,5	CAP	-	5681	G	T	1,9	NS3	R387M
187	A	C	13,4	CAP	Q31K	5722	G	C	1,4	NS3	D401H
237	C	T	2,3	CAP	-	5886	A	G	2,3	NS3	-
252	G	A	4,5	CAP	-	6599	G	T	1,4	NS4A	G75V
601	C	A	0,9	MEM	L55I	6618	G	T	0,7	NS4A	M80I
735	G	T	0,9	MEM	-	6885	C	G	0,8	NS4B	-
1090	C	A	0,7	ENV	Q52K	6890	A	C	2,9	NS4B	E22A
1301	A	C	0,9	ENV	K122T	7108	G	T	0,5	NS4B	A95S
1304	A	T	0,9	ENV	K123M	7257	C	A	0,7	NS4B	N144K
1311	G	T	1,0	ENV	M125I	7401	G	T	0,9	NS4B	E191D
1468	T	C	0,6	ENV	Y178H	7425	C	G	0,9	NS4B	-
1754	C	T	1,1	ENV	S273L	8040	A	C	1,0	NS5	E157D
1832	A	T	0,6	ENV	Y299F	8589	G	T	2,8	NS5	Q340H
2000	A	T	1,7	ENV	N355I	8615	T	C	2,6	NS5	F349S
2411	T	C	1,4	ENV	M492T	9133	G	A	9,7	NS5	V522K
2435	T	G	1,4	NS1	I5S	9134	T	A	12,2	NS5	-
2526	A	G	0,5	NS1	-	9397	A	G	3,2	NS5	N610D
2598	A	T	0,7	NS1	-	9438	G	T	7,9	NS5	M623I
2709	C	A	1,1	NS1	-	9439	G	A	5,7	NS5	E624K
2805	T	C	0,6	NS1	-	9520	G	T	1,3	NS5	G651W
2985	T	C	0,9	NS1	-	9536	C	T	2,0	NS5	S656L
3116	G	C	2,7	NS1	W232S	9540	A	G	2,1	NS5	-
3175	T	G	4,4	NS1	S252A	9550	A	T	3,8	NS5	S661C
3242	A	T	6,3	NS1	E274V	9577	A	T	1,2	NS5	I670L
3748	C	T	0,8	NS2A	-	9624	A	G	1,5	NS5	-
3855	G	T	0,9	NS2A	-	9647	A	C	1,7	NS5	Q693P
3881	A	T	0,5	NS2A	K135I	9648	A	C	2,0	NS5	-
4168	A	T	0,6	NS2B	M13L	9658	T	C	0,7	NS5	S697P
4651	G	T	1,2	NS3	G44STOP	9929	G	T	1,5	NS5	W787L
4757	A	C	1,2	NS3	Y79S	10291	A	C	5,5	3'UTR	

POS	REF	ALT	Freq %	Região no genoma	Mudança aa	POS	REF	ALT	Freq %	Região no genoma	Mudança aa
55	A	C	51,0	5'UTR		5092	A	T	19,3	NS3	M191L
95	T	A	3,2	5'UTR		5298	A	T	2,6	NS3	-
99	G	T	8,2	CAP	M1I	5307	T	A	2,6	NS3	H262Q
172	G	T	5,2	CAP	V26L	5531	G	T	7,1	NS3	R337I
576	A	T	4,8	MEM	E46D	6330	T	A	4,2	NS3	D603E
956	C	G	3,8	ENV	S7STOP	6576	A	T	5,5	NS4A	-
1402	G	C	3,8	ENV	G156R	6582	C	G	5,2	NS4A	F69L
2505	G	C	4,6	NS1	W28C	6790	A	T	6,4	P2K	I12L
3239	T	C	11,5	NS1	L273P	7175	A	T	6,3	NS4B	H117L
3579	G	C	8,4	NS2A	-	7276	A	T	4,7	NS4B	T151S
3688	G	C	3,4	NS2A	G71P	7288	C	G	5,0	NS4B	L155V
3689	G	C	3,7	NS2A	-	7345	C	G	6,6	NS4B	L174V
3848	C	G	4,1	NS2A	T124S	8579	T	A	5,6	NS5	M337K
3991	T	C	29,5	NS2A	-	9067	A	T	6,1	NS5	R500STOP
4148	A	T	3,1	NS2B	E6V	9592	G	A	7,4	NS5	A675T
4175	G	T	3,8	NS2B	S15I	9650	G	C	4,4	NS5	W694S
4437	A	T	5,0	NS2B	-	9696	T	A	4,7	NS5	C709STOP
4662	T	A	4,3	NS3	H47Q	9926	A	T	2,7	NS5	H786L
4702	A	G	7,3	NS3	R61G	9958	A	T	5,9	NS5	I797L
4898	T	A	4,7	NS3	V126E	10179	C	G	4,1	NS5	-
5009	G	C	2,7	NS3	S163T	10188	T	A	4,5	NS5	-
5010	T	A	3,0	NS3	-						

**Apêndice III – Mutações fixas achadas no genoma consenso representante da população viral de cada amostra. Em vermelho, encontram-se destacadas aquelas que foram achadas em pelo menos duas amostras.**

<b>Amostra</b>	<b>Mutações fixas no genoma viral</b>								
131	CAP-R97K	CAP-A112V	NS1-D208G	NS2A-R51K	NS2A-A109T	NS3-I461V	NS5-K388E		
132	CAP-I94T	ENV-V129I	ENV-L131Q	ENV-M201I	ENV-M340T	ENV-G395V	NS1-T164S	NS2A-M57L	NS2A-V62A
	NS2A-K135R	NS2B-V8I	NS3-G14E	NS3-T583I	NS4A-N42K	NS5-V354L	NS5-K388E	NS5-I412V	NS5-R596K
	NS5-H786Y	NS5-W795G							
133	ENV-F43L	ENV-I432T	NS1-I19M	NS2A-A109T	NS2A-A164G	NS3-I461V	NS4A-A36T	NS4B-T73A	NS5-D521E
	NS5-V637A	NS5-T880I	ENV-K160E						
134	MEM-R122K	ENV-E202G	NS2A-V62A	NS3-G14E	NS3-S171N	NS4B-L15F	NS5-S23N	NS5-V168A	
135	MEM-A112V	NS2A-K135R	NS3-G14E	NS3-I567V	NS4A-N42K	NS4B-A197V	NS5-V168A	NS5-R596K	
136	MEM-R16K	NS2A-L60V	NS2A-P63T	NS2A-K135R	NS3-G14E	NS4A-N42K	NS5-V168A	NS5-R596K	
137	NS4A-S46N	NS5-R430S	NS5-V836I						
138	ENV-K122R	ENV-K291R	NS3-V/A395I						
139	-								
140	NS1-T126I								
141	CAP-V106I								
142	CAP-V106I								
143	ENV-E383G	NS1-T29I	NS3-S452F	NS3-A454S	NS4A-S46N	NS4B-E206D	NS4B-G207V	NS5-R430S	NS5-V836I
144	-								
145	MEM-V161I	NS1-T164S	NS4B-G16R	NS5-V687I					
146	CAP-L50F								
147	NS1-T164S	NS4B-G16R	NS5-I27T	NS5-I641T	NS5-V687I				
148	NS1-T164S	NS4B-G16R	NS5-L20M	NS5-I27T	NS5-I641T	NS5-V687I			
149	NS1-T164S	NS4B-G16R	NS5-I27T	NS5-I641T	NS5-V687I				
151	NS4B-G16R	NS5-V687I							
153	NS1-T164S	NS4B-G16R	NS4B-T20I	NS5-N174D	NS5-V687I				
154	NS1-T164S	NS4B-G16R	NS4B-T20I	NS5-N174D	NS5-V687I				
155	CAP-L50F								
156	MEM-D29N	ENV-I46M	ENV-M183I	ENV-H346Y	NS1-D190N	NS5-V5I	NS5-R430S	NS5-G511W	NS5-E653D
157	ENV-T180I	NS3-T407R	NS5-G501E	NS5-E521D	NS5-G523S				
158	CAP-R85K	NS1-V177A	NS2A-A173V	NS3-G14E	NS4B-N245S	NS5-K29R	NS5-P136Q	NS5-V836I	
159	NS1-T164S	NS3-G14E	NS4B-G16R	NS4B-T20I	NS5-N174D	NS5-V687I			