



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO
BIOLOGIA DE AGENTES INFECCIOSOS E PARASITÁRIOS

**CARACTERIZAÇÃO MOLECULAR E RELAÇÃO FILOGENÉTICA DO
GENOMA COMPLETO DOS ARBOVÍRUS BUSSUQUARA, IGUAPE, ILHÉUS
E ROCIO (FAMÍLIA *FLAVIVIRIDAE*, GÊNERO *FLAVIVIRUS*)**

DANIELE BARBOSA DE ALMEIDA MEDEIROS

Belém-Pará

2009

DANIELE BARBOSA DE ALMEIDA MEDEIROS

**CARACTERIZAÇÃO MOLECULAR E RELAÇÃO FILOGENÉTICA DO
GENOMA COMPLETO DOS ARBOVÍRUS BUSSUQUARA, IGUAPE, ILHÉUS
E ROCIO (FAMÍLIA *FLAVIVIRIDAE*, GÊNERO *FLAVIVIRUS*)**

Tese de doutorado apresentada ao Programa de Pós-Graduação Biologia de Agentes Infecciosos e Parasitários, do Instituto de Ciências Biológicas da Universidade Federal do Pará como requisito para a obtenção do grau de Doutor em Biologia de Agentes Infecciosos e Parasitários.

Orientador: Prof. Dr. Pedro Fernando da Costa Vasconcelos

Co-Orientador: Prof. Dr. Márcio Roberto Teixeira Nunes

BELÉM-PA

2009

**CARACTERIZAÇÃO MOLECULAR E RELAÇÃO FILOGENÉTICA DO
GENOMA COMPLETO DOS ARBOVÍRUS BUSSUQUARA, IGUAPE, ILHÉUS
E ROCIO (FAMÍLIA *FLAVIVIRIDAE*, GÊNERO *FLAVIVIRUS*)**

Tese de doutorado apresentada ao Programa de Pós-Graduação Biologia de Agentes Infecciosos e Parasitários, do Instituto de Ciências Biológicas da Universidade Federal do Pará como requisito para a obtenção do grau de Doutor em Biologia de Agentes Infecciosos e Parasitários.

Orientador: Prof. Dr. Pedro Fernando da Costa Vasconcelos
Seção de Arbovirologia e Febres Hemorrágicas, IEC

Co-Orientador: Prof. Dr. Márcio Roberto Teixeira Nunes
Seção de Arbovirologia e Febres Hemorrágicas, IEC

Banca Examinadora: Prof^ª. Dra. Rita Catarina Medeiros Sousa
Núcleo de Medicina Tropical, UFPA

Prof. Dr. Eduardo José de Melo Santos
Instituto de Ciências Biológicas, UFPA

Prof^ª. Dra. Ana Cecília Ribeiro Cruz
Seção de Arbovirologia e Febres Hemorrágicas, IEC

Prof^ª. Dra. Elizabeth Salbé Travassos da Rosa
Seção de Arbovirologia e Febres Hemorrágicas, IEC

Suplente: Prof. Dr. Paolo Marinho de Andrade Zanotto
Instituto de Ciências Biomédicas, USP

Prof. Dr. Ricardo Ishak
Instituto de Ciências Biológicas, UFPA

Belém, 27 de Novembro de 2009

EPIGRAFE

"A ciência humana de maneira nenhuma nega a existência de Deus. Quando considero quantas e quão maravilhosas coisas o homem compreende, pesquisa e consegue realizar, então reconheço claramente que o espírito humano é obra de Deus, e a mais notável."

Galileu Galilei

DEDICATÓRIA

Dedico essa tese de doutorado a minha amada filha, Julia, que veio iluminar à minha vida; aos meus pais, George e Vanda, pelas incansáveis horas de dedicação; á minha irmã Cláudia, pela amizade e carinho; e ao meu companheiro Marcio Levy, por me ensinar o significado da palavra felicidade.

AGRADECIMENTO

Agradeço a Deus pelos momentos de felicidade, que iluminam e me dão força para seguir a minha caminhada, e pelos momentos de dificuldade que me moldam a cada instante para ser um ser humano mais digno a exemplo do Cristo. Agradeço pela graça da minha filha, tão amada, que cresce forte em meu ventre e que me realiza nas menores descobertas.

Agradeço a minha família, o alicerce de minha vida: meus pais, George Medeiros e Vanda Medeiros, pelo eterno cuidado, dedicação e amor; pelo apoio nos momentos difíceis e de inquietantes decisões; por estarem ao meu lado a cada passo, a cada pequena conquista e grandes realizações, pois estes não teriam valor se vocês não estivessem comigo. Agradeço a minha irmã, pela mulher que hoje desabrocha, pelos momentos de orgulho e pelo companheirismo e amizade. Agradeço aos meus tios e primos pelos abraços calorosos de saudade e pelo apoio.

Agradeço ao meu amor, Marcio Levy, pelo companheirismo em todos os momentos, pelos sorrisos, pelo cuidado carinhoso e por mostrar que sonhos podem ser reais. Obrigada pela família maravilhosa que tu me proporcionaste, o nosso amado filho(a) e a família Levy, que me recebeu de braços abertos repletos de carinho. Agradeço em especial aos meus sogros, Salim e Clarice Levy, e a minha cunhada, Claudia Levy, pelo apoio e amizade.

Ao Instituto Evandro Chagas, por me acolher e permitir, desde minha iniciação científica, o desenvolvimento de meu trabalho.

Agradeço ao Dr. Pedro Vasconcelos por ser mais do que meu chefe e orientador, por acreditar na minha capacidade e no meu crescimento profissional e pessoal, pelo apoio em todos os momentos e, principalmente, pela amizade.

Agradeço ao meu Co-orientador Márcio Nunes pelos conhecimentos repassados e por me ensinar a fazer ciência, tendo sempre respeito ao próximo e promovendo a união do grupo.

Agradeço a Dra. Elizabeth Salbé, por ser minha parceira de trabalho, minha amiga, pelo carinho maternal, por ser um grande exemplo de mulher, de mãe e de profissional a ser seguido.

Agradeço a minha terceira família, a Seção de Arbovirologia e Febres Hemorrágicas (IEC), pela luta diária, pelos momentos de descontração e pela parceria a fim de servir à saúde pública e produzir novos conhecimentos. Em especial às equipes: do laboratório de Raiva e Hantavírus - Armando Pereira, Lívia Casseb, Taciana Barbosa, Darlene Simith, Ivy Prazeres, pelo o apoio, compreensão e amizade; do laboratório da biologia molecular - Adriana Carneiro, Ana Cecília, Helena Baldez, Mayra Silva e Samir Casseb pelo auxílio nas etapas de sequenciamento e análise filogenética; do laboratório de cultivo celular, em especial a Eliana Pinto, pelo apoio na infecção dos cultivos celulares utilizados neste trabalho; aos amigos da registro de amostras, Osvaldo Vaz e Jefferson Buna, pelo companheirismo.

Agradeço a Dra Amélia Travassos da Rosa pelo apoio na interpretação dos dados sorológicos.

Agradeço ao Dr. Goro Kuno pela parceria neste trabalho.

Agradeço aos membros do campus Bragança da UFPA, em especial ao Dr. Marcelo Vallinoto, pelos ensinamentos na área de filogenia.

Agradeço aos meus eternos amigos, Aline Lira, Fabíola Brasil, Flávia Biondi, Eduardo Santos, Jannifer Chiang, Leda Guimarães e Perpétua Costa, por estarem sempre ao meu lado, mesmo quando separados pela distância e pela correria da vida diária.

Agradeço aos meus professores da graduação e pós-graduação, a Universidade Federal do Pará, à secretaria da pós-graduação, a todos que contribuíram para o meu crescimento profissional e pessoal, muitíssimo obrigada.

SUMÁRIO

LISTA DE FIGURAS.....	10
LISTAS DE QUADROS E TABELAS.....	14
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.....	16
RESUMO.....	23
ABSTRACT.....	24
1. INTRODUÇÃO.....	25
1.1. CONSIDERAÇÕES GERAIS: FAMÍLIA FLAVIVIRIDAE, GÊNERO FLAVIVIRUS.....	25
1.2. CICLO BIOLÓGICO.....	28
1.3. CLASSIFICAÇÃO DOS FLAVIVÍRUS.....	31
1.4. ESTRUTURA DOS FLAVIVÍRUS.....	35
1.5. REPLICAÇÃO DOS FLAVIVÍRUS.....	37
1.6. GENOMA DOS FLAVIVÍRUS.....	43
1.6.1. Região não-codificante 5'.....	45
1.6.2. Região não-codificante 3'.....	46
1.6.3. Ciclização do RNA viral.....	52
1.7. ESTRUTURAS DAS PROTEÍNAS.....	57
1.7.1. Proteína do Capsídio (C).....	57
1.7.2. Glicoproteína de membrana (M).....	58
1.7.3. Proteína do Envelope (E).....	59
1.7.4. Proteína não-estrutural NS1.....	62
1.7.5. Proteína não-estrutural NS2A.....	63

1.7.6. Proteína não-estrutural NS2B.....	64
1.7.7. Proteína não-estrutural NS3.....	65
1.7.8. Proteína não-estrutural NS4A.....	66
1.7.9. Proteína não-estrutural NS4B.....	66
1.7.10. Proteína não-estrutural NS5.....	67
1.8. FILOGENIA E EVOLUÇÃO DOS FLAVIVÍRUS.....	68
1.8.1. Evolução e Dispersão dos Flavivírus.....	75
1.9. RECOMBINAÇÃO ENTRE FLAVIVIRUS.....	
1.10. FLAVIVÍRUS BRASILEIROS EM ESTUDO.....	79
1.10.1. Vírus Bussuquara.....	79
1.10.2. Vírus Iguape.....	80
1.10.3. Vírus Ilhéus.....	80
1.10.4. Vírus Rocio.....	82
1.11. RELEVÂNCIA.....	84
1.12. OBJETIVOS.....	86
1.12.1. Objetivo Geral.....	86
1.12.2. Objetivos Específicos.....	86
2. MATERIAL E MÉTODOS.....	88
2.1 AMOSTRAS.....	88
2.2. PROPAGAÇÃO VIRAL EM CÉLULAS VERO.....	89
2.3. EXTRAÇÃO DO RNA.....	89
2.4. TRANSCRIÇÃO REVERSA SEGUIDA DA REAÇÃO EM CADEIA MEDIADA PELA POLIMERASE (RT-PCR).....	90
2.5. SEQUENCIAMENTO NUCLEOTÍDICO.....	95

2.6. ALINHAMENTO DAS SEQUÊNCIAS NUCLEOTÍDICAS.....	96
2.7. IDENTIFICAÇÃO DOS MOTIVOS CONSERVADOS, SÍTIOS DE CLIVAGEM, SÍTIOS DE GLICOSILAÇÃO E RESÍDUOS DE CISTEÍNA NA POLIPROTEÍNA.....	98
2.8. ESTRUTURA SECUNDÁRIA DAS 5'RNC/ 3'RNC E CICLIZAÇÃO DO RNA VIRAL.....	99
2.9. DETERMINAÇÃO E ANÁLISE DE SIMILARIDADE (<i>BOOTSCANNING</i>) DA REGIÃO CODIFICANTE.....	100
2.11. ANÁLISE FILOGENÉTICA COMPARATIVA.....	101
3. RESULTADOS	102
3.1. ANÁLISE DO GENOMA VIRAL.....	102
3.2. REGIÕES NÃO CODIFICADORAS.....	104
3.2.1. Sequências Conservadas	104
3.2.2. Estruturas Secundárias das 5' RNC e 3'RNC e Ciclização do RNA viral	107
3.3. ANÁLISE DA POLIPROTEÍNA.....	111
3.3.1. Identificação dos sítios de clivagem, sítios de glicosilação e resíduos de cisteína	111
3.3.2. Identificação de motivos conservados nas proteínas E, NS3 e NS5	114
3.4. ANÁLISE DE SIMILARIDADE.....	117
3.5. FILOGENIA.....	123
3.5.1. Região codificadora	125
3.5.2. Região estrutural	127

3.5.3. Região não estrutural.....	130
3.5.4. Filogenia concatenada E-NS5.....	133
4. DISCUSSÃO.....	135
5. CONCLUSÕES.....	162
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	165
ANEXO I.....	201
ANEXO II.....	202
ANEXO III.....	203
ANEXO IV.....	204
ANEXO V.....	205
ANEXO VI.....	205

LISTA DE FIGURAS

Figura 1 -	(A) Crioeletromicroscopia do VDEN, (B) Mapa eletrondenso dos flavivírus.....	35
Figura 2 -	Formas das partículas virais dos flavivírus (A) partícula viral imatura; (B) partícula viral madura; (C) esquema das formas madura e imatura.....	36
Figura 3 -	(A) Apresentação esquemática do genoma e processamento das poliproteínas dos integrantes do gênero <i>Flavivirus</i> . (B) Topologia e interação dos flavivírus ao nível da membrana do retículo endoplasmático.....	40
Figura 4 -	Esquema da replicação viral dos flavivírus.....	42
Figura 5 -	Estrutura secundária da 5'RNC do VNO.....	46
Figura 6 -	Estrutura secundária preditiva da 3'SL na 3'RNC dos flavivírus VDEN-2 e VNO.....	48
Figura 7 -	Sequências nucleotídicas de alguns motivos conservados na região 3'RNC do genoma de flavivírus transmitidos por mosquitos.....	49
Figura 8 -	Características das 5'- e 3'RNC de flavivírus transmitidos por mosquitos, destacando as sequências envolvidas na ciclização do RNA - CS, CYC, CS1, CS2, RCS2, C3, RCS3 (caixas) – e as repetições em série presentes no genoma do VFA.....	51
Figura 9 -	Sequências nucleotídicas dos terminais 5'- e 3'- de alguns	

	flavivírus transmitido por mosquitos, mostrando alinhamento entre o genoma de RNA sentido positivo e o RNA complementar sentido negativo.....	52
Figura 10 –	Estrutura de ciclização do RNA dos flavivírus.....	54
Figura 11 -	Representação da árvore filogenética dos flavivírus, subdivididas nos quatro principais grupos.....	69
Figura 12 –	Mapa mundial ilustrando a localização aproximada e dispersão dos flavivírus.....	76
Figura 13 -	Esquema da estratégia de amplificação do genoma (RT-PCR) e seqüenciamento nucleotídico.....	90
Figura 14 -	Esquema dos protocolos 5'RACE (A) e 3'RACE.....	93
Figura 15 -	Representação esquemática das 3'RNC dos VBSQ, VIGU, VILH e VROC em comparação a outros flavivírus transmitidos por mosquitos dos grupos: Encefalite japonesa, Kokobera, Kedougou, Spondweni e Dengue.....	105
Figura 16 -	Alinhamento entre as seqüências nucleotídicas conservadas da 3'RNC dos flavivírus transmitidos por mosquitos do gênero <i>Culex</i> , seguindo a disposição 5'→3, comparando com os flavivírus VBSQ, VIGU, VILH e VROC.....	106
Figura 17 -	Estrutura secundária das 5'RNC e 3'RNC dos flavivírus VBSQ, VIGU, VILH e VROC na ciclização do RNA viral....	108
Figura 18 -	Seqüências complementares entre as 5'RNC e 3'RNC (5'CS-3'CS) dos flavivírus VBSQ, VIGU, VILH e VROC.....	109

- Figura 19 -** Alinhamento entre as sequências nucleotídicas correspondente a 5'CYC (5'RNC) dos flavivírus transmitidos por mosquitos do gênero *Culex* comparando com as sequências correspondentes dos VBSQ, VIGU, VILH e VROC..... 110
- Figura 20 -** Análise de similaridade (*Bootscan*) realizada pelo programa SimPlot, utilizando as sequências dos VBSQ (A) e VILH (B) como referência..... 119
- Figura 21 -** Análise de similaridade (*Bootscan*), utilizando o VROC como referência. Notar a maior similaridade com o VBAG e VESL para todo a ORF..... 120
- Figura 22 -** Árvore filogenética pelos métodos MV das sequências da ORF dos flavivírus. Os valores bayesianos foram plotados dentro de parêntese..... 126
- Figura 23 -** Árvore filogenética pelo método MV, com valores bayesianos inseridos em parênteses, das sequências da região estrutural dos 32 flavivírus..... 128
- Figura 24 -** Árvore filogenética gerada pelo método de MV e Bayesiano (valores bayesianos dentro de parênteses) realizada com as sequências nucleotídicas do gene E de 32 flavivírus..... 129
- Figura 25 -** Árvore filogenética pelos métodos MV e Bayesiano (valores inseridos dentro dos parênteses) das sequências da região não estrutural dos 32 flavivírus..... 131

Figura 26 -	Árvore filogenética pelos métodos MV das sequências do gene NS3 (A) e NS5 (B) dos flavivírus.....	132
Figura 27 -	Árvore filogenética pelo Método MV concatenada das sequências parciais dos genes E e NS5 de 38 flavivírus.....	134
Figura 28 -	Número de mutações no peptídeo de fusão dos flavivírus encontrado por Seligman (2008).....	146

LISTAS DE QUADROS E TABELAS

Quadro 1 -	Espécies virais e cepas registradas no CITV.....	33
Quadro 2 -	Comparação dos tamanhos nucleotídicos das 5' e 3'RNCs no genoma de alguns flavivírus.....	44
Quadro 3 -	Localização das posições das sequências de ciclização no genoma de flavivírus.....	56
Quadro 4 -	Amostras virais utilizadas neste estudo.....	88
Quadro 5 -	Iniciadores utilizados para a técnica de RT-PCR e sequenciamento nucleotídico.....	91
Quadro 6 -	Iniciadores específicos para os flavivírus brasileiros utilizados no protocolo 5' RACE e 3'RACE.....	94
Quadro 7 -	Sequências do genoma completo dos 32 flavivírus utilizados para o alinhamento e filogenia neste estudo.....	97
Quadro 8 -	Comparação do genoma dos quatro flavivírus brasileiros em estudo: tamanho do genoma, regiões codificadoras para proteínas estruturais e não-estruturais, RNCs.....	103
Quadro 9 -	Sítios de clivagem dos flavivírus VBSQ, VIGU, VILH e VROC em comparação com membros do complexo da encefalite japonesa e três flavivírus transmitidos por mosquitos procedentes da África	112
Tabela 1 -	Sítios de glicosilação determinado para as proteínas PrM, E e NS1 pelo programa NetNglyc para os flavivírus VBSQ, VIGU, VILH e VROC.....	113

Tabela 2 -	Principais domínios das proteínas E, NS3 e NS5 determinado pelo Pfan para os flavivírus VBSQ, VIGU, VILH e VROC.....	114
Tabela 3 -	Percentual de identidade aminoacídica entre os flavivírus VBSQ, VIGU, VILH e VROC e outros flavivírus transmitidos por <i>Culex</i> ...	122
Tabela 4 -	Modelo de substituição nucleotídica determinado pelo critério de informação de Akaike.....	124
Tabela 5 -	Recíproca das relações antigênicas pelo teste de FC entre os flavivírus do VROC e VILH com os integrantes do grupo Ntaya e encefalite japonesa.....	160

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

%	Porcentagem
α	Alfa
β	Beta
Δ div	Divergência genética
μ L	Microlitro
μ M	Micro molar
3'CYC	Sequência de ciclização conservada na região não-codificadora 3'
3'SL	3' haste-alça (3' <i>stem-loop</i>)
5'CYC	Sequência de ciclização conservada na região não-codificadora 5'
A	Adenina
aa	Aminoácidos
AIC	<i>Akaike Information Criterion</i>
Ala	Alanina
AnchC	Âncora hidrofílica C-terminal (proteína do capsídeo)
Ans	Asparagina
Arg	Arginina
Asp	Ácido aspártico
AV	Agrupamento de vizinhos
C	Citosina; proteína do capsídeo
CDC	Centro para Controle e Prevenção de Doenças (<i>Centers for Disease Control and Prevention</i>)
cDNA	Ácido desoxirribonucléico complementar

Cis	Cisteína
CITV	Comitê Internacional de Taxonomia Viral
CS1	Sequência conservada 1
CS2	Sequência conservada 2
DNA	Ácido desoxirribonucléico (<i>Deoxyribonucleic Acid</i>)
dNTPs	Deoxinucleotídeo trifosfato
DTT	1,4-ditiotreitol
E	Proteína do envelope
ECP	Efeito citopático
EDTA	Etileno-diamino tetra acetato de sódio
FC	Fixação do complemento
Fen	Fenilalanina
FHD	Febre hemorrágica do dengue
G	Guanina
g	Grama
Gli	Glicina
Gln	Glutamina
Glu	Ácido glutâmico
His	Histidina
IEC	Instituto Evandro Chagas
IH	Inibição da hemaglutinação
IL	Interleucina
Ile	Isoleucina
Kcal	Quilocaloria

KDa	Quilodáltons
L	Litros
Leu	Leucina
Lis	Lisina
M	Glicoproteína de Membrana
MEM	Meio Mínimo Essencial (<i>Minimum Essential Medium</i>)
Met	Metionina
MgCl ₂	Cloreto de magnésio
Min	Minutos
mL	Mililitro
mm	Milímetro
mM	Milimolar
MP	Máxima Parcimônia
MTase	Metiltransferase
Nm	Nanômetro
NS1	Proteína Não-estrutural 1
NS2A	Proteína Não-estrutural 2A
NS2B	Proteína Não-estrutural 2B
NS3	Proteína Não-estrutural 3
NS4A	Proteína Não-estrutural 4A
NS4B	Proteína Não-estrutural 4B
NS5	Proteína Não-estrutural 5
nt	Nucleotídeos
NTPase	RNA nucleosídeo trifosfatase

°C	Graus centígrados
OMS	Organização Mundial da Saúde
ORF	Matriz de leitura aberta (<i>Open Reading Frame</i> - ORF)
pb	Pares de bases
PBS	Solução salina tamponada (<i>Phosphate buffer solution</i>)
PCR	Reação em cadeia mediada pela polimerase
pH	Potencial hidrogeniônico
PrM	Precursor da glicoproteína de membrana
Pro	Prolina
RCS2	Repetição da sequência conservada 2
RdRP	RNA polimerase dependente de RNA
RE	Retículo endoplasmático
RNA	Ácido ribonucléico (<i>Ribonucleic Acid</i>)
RNAm	RNA mensageiro
RNC	Região não codificadora
RT	Transcrição reversa
RTPase	RNA trifosfatase
RT-PCR	Transcrição reversa seguida da reação em cadeia mediada pela polimerase
SAM	S-adenosil-metionina
SCD	Síndrome do choque
SCy	Sequência de ciclização
SEARB	Seção de Arbovirologia e Febres Hemorrágicas
seg	Segundos

Ser	Serina
SNC	Sistema nervoso central
SVS	Secretaria de Vigilância em Saúde
T	Timina
Thr	Treonina
Tir	Tirosina
TN	Teste de neutralização
Trp	Triptofano
U	Uracila; Unidade Internacional
UV	Ultravioleta
VAEF	<i>Vírus aedes flavivírus</i>
Val	Valina
VAROA	<i>Vírus Aroa</i>
VBAN	<i>Vírus Banzi</i>
VBDQ	<i>Vírus bussuquara</i>
VBOU	<i>Vírus bouboui</i>
VCFA	<i>Vírus cell fusion agents</i>
VCPC	<i>Vírus cacipacoré</i>
VCXF	<i>Vírus culex flavivírus</i>
VDEN	<i>Vírus dengue</i>
VDEN-1	Vírus dengue sorotipo 1
VDEN-2	Vírus dengue sorotipo 2
VDEN-3	Vírus dengue sorotipo 3
VDEN-4	Vírus dengue sorotipo 4

VEH	<i>Vírus Edge Hill</i>
VEJ	<i>Vírus da encefalite japonesa</i>
VEMV	<i>Vírus encefalite Murray Valley</i>
VENT	<i>Vírus entebbe bat</i>
VERO	Células de rim de macaco verde africano linhagem
VESL	<i>Vírus da encefalite São Luis</i>
VFA	<i>Vírus da febre amarela</i>
VIGU	<i>Vírus Iguape</i>
VILH	<i>Vírus Ilhéus</i>
VJUG	<i>Vírus jugra</i>
VKFD	<i>Vírus Kyasanur Forest disease</i>
VKR	<i>Vírus Kamitii River</i>
VKUN	<i>Vírus kunjin</i>
VLGT	<i>Vírus langat</i>
VLI	<i>Vírus louping ill</i>
VMOD	<i>Vírus Modoc</i>
VNJL	<i>Vírus Naranjal</i>
VNO	<i>Vírus do Nilo ocidental</i>
VNTA	<i>Vírus Ntaya</i>
VOHF	<i>Vírus da febre hemorrágica do Omsk</i>
VPOW	<i>Vírus Powassan</i>
VQB	<i>Vírus quang binh</i>
VRB	<i>Vírus Rio Bravo</i>
VROC	<i>Vírus rocio</i>

VSAB	<i>Virus saboya</i>
VSEP	<i>Virus sepik</i>
VSOK	<i>Virus sokoluk</i>
VTAB	<i>Virus tamana bat</i>
VTBE	<i>Virus tick-borne encephalitis</i>
VTYU	<i>Virus Tyuleniy</i>
VUGS	<i>Virus Uganda S</i>
VYOK	<i>Virus Yokose</i>

RESUMO

Os flavivírus são conhecidos por seu complexo ciclo biológico e importância na saúde pública e na economia mundial. Os aspectos ecológicos e quadros clínicos estão estreitamente relacionados à filogenia e evolução dos flavivírus. Este trabalho objetiva a caracterização molecular dos genomas dos flavivírus Bussuquara (VBSQ), Iguape (VIGU), Ilhéus (VILH) e Rocio (VROC), determinando relações filogenéticas com os demais integrantes do gênero *Flavivirus*. Foi realizado o seqüenciamento completo da região codificadora (ORF) e regiões não codificantes (RNC) 5' e 3'; análise da estrutura secundária do RNA viral e das seqüências conservadas da 3'RNC; determinação dos sítios de clivagem, glicosilação, resíduos Cis e motivos conservados na poliproteína; e as análises de similaridade e filogenética. Os genomas dos VBSQ, VIGU, VILH e VROC apresentaram a mesma organização que os demais flavivírus, medindo 10.815 nt, 10.922 nt, 10.775 nt, 10.794 nt, respectivamente. O padrão das seqüências conservadas da 3'RNC do VBSQ foi RCS2-CS2-CS1, enquanto que para os VIGU, VILH e VROC foram CS3-RCS2-CS2-CS1. As características das estruturas secundárias do RNAs dos flavivírus em estudo foram similares aos demais flavivírus. O número dos sítios de glicosilação das proteínas PrM, E e NS1 foi distinto entre os flavivírus brasileiros, porém o padrão 6,12,12 dos resíduos de Cis e do sítios de clivagem permaneceram conservados. Na proteína E, alterações aminoacídicas pontuais foram observadas no peptídeo de fusão dos VBSQ, VIGU e VROC, e a seqüência do tripeptídeo RGD foi distinta para os quatro vírus em estudo. Os motivos determinantes das atividades de MTase-SAM da NS5, bem como da helicase e protease da NS3, permanecem conservados. Dentre os oito motivos da polimerase viral (NS5), somente os motivos V, VI e VII possuem alguma substituição nucleotídica para o VILH e VROC. As análises de similaridade mostram que VBSQ apresenta maior relação com VIGU enquanto que o VILH e VROC são mais relacionados entre si, porém sendo consideradas espécies virais distintas. Com base nas análises filogenéticas, características moleculares do genoma e biológicas, propõem-se a formação de três grupos genéticos: o grupo Rocio, que agrupa VROC e VILH; o grupo Bussuquara formado pelos VBSQ e *Vírus naranjal* e o grupo Aroa que inclui o *Vírus Aroa* e VIGU.

ABSTRACT

Flaviviruses have been known due their complex biological cycle and their relevance in public health and global economy. The ecologic aspected and clinic pictures are related with phylogeny and evolution of flaviviruses. This work aims the molecular characterization of Bussuquara (BSQV), Iguape (IGUV), Ilheus (VILH) e Rocio (VROC) flaviviruses genomes, determining their phylogenetic relationships with others members of genus *Flavivirus*. The study included: the full-length sequencing of the four Brazilian flaviviruses; analyzes of the predictive secondary structure of RNA and conserved sequences in the 3'NCR; determination of cleavage and glycosilation sites, cysteine residues and conserved motifs in the polyprotein; and similarity and phylogenetic analyzes. The BSQV, IGUV, VILH and VROC genomes present 10815 nt, 10922 nt, 10775 nt, and 10794 nt, respectively. The conserved standard sequences in 3'NCR of BSQV was RCS2-CS2-CS1, while to IGUV, ILHV and ROCV were CS3-RCS2-CS2-CS1. The secondary structure of RNA obtained for the Brazilian flaviviruses were similar to the other flaviviruses. The numbers of the glycosilation sites to PrM, E and NS1 proteins were distinct among the studied Brazilian flaviviruses, therefore the pattern 6,12,12 Cis residues and the cleavage sites were conserved. In the E protein, some singles mutations were observed in fusion peptide of BSQV, IGUV and ROCV, and the RGD motif were distinct for the flaviviruses under study. The motif that determines the MTase-SAM activity in NS5, as well as the helicase and protease activity in NS3 were conserved. Among the eight polimerase motifs in NS5, only the V, VI and VII motifs were observed single mutations in ILHV and ROCV. The similarity analyzes showed that BSQV presents high relationship with VIGU, while ILHV and ROV were more related among themselves, however those viruses were considered distincts species. Based in the phylogenetic analyzes, molecular and biological characteristics, it was proposed the establishment of three distinct genetic groups: the Rocio group, grouping ILHV and ROCV, Bussuquara group formed by BSQV and *Naranjal virus*; and Aroa group, that include Aroa virus and IGUV.

1. INTRODUÇÃO

1.1. CONSIDERAÇÕES GERAIS: FAMÍLIA *FLAVIVIRIDAE*, GÊNERO *FLAVIVIRUS*

A família *Flaviviridae* é composta por vírus cujo genoma é composto de RNA (Ácido Ribonucléico), não segmentado, sentido positivo, cujo protótipo é o *Vírus da febre amarela* (VFA). Essa família compreende três gêneros: *Flavivirus*, *Hepacivirus* e *Pestivirus* (ICTV, 2005; Lindenbach et al., 2007), além de dois grupos de vírus não classificados, GBV-C e GBV-A, que estão aguardando uma classificação formal dentro da família (Lindenbach et al., 2007).

Quanto ao táxon de superfamília, na qual os vírus de RNA são classificados quanto à dependência da RNA polimerase, a família *Flaviviridae* é membro da Superfamília 2 (Lindenbach et al., 2007); em que, os integrantes da família *Flaviviridae* codificam a RNA helicase, que coordena cátions bivalentes dentro do sítio de catálise (Gorbalenya & Koonin, 1993).

O gênero *Flavivirus* apresenta 70 integrantes registrados no Comitê Internacional de Taxonomia Viral (CITV) (ICTV, 2005). A maioria dos flavivírus, cerca de 53 vírus, são considerados Arbovírus (termo derivado da expressão em inglês *Arthropods-borne virus*) por apresentarem um ciclo biológico que envolve como hospedeiros os vertebrados e como vetores/transmissores artrópodes hematófagos. Outros flavivírus são classificados como agentes zoonóticos e são transmitidos entre roedores ou

morcegos, sendo considerados vírus sem vetores ou que o possível vetor artrópode permanece desconhecido (Karabatsos, 1985; Travassos da Rosa et al., 1997; Kuno et al., 1998; ICTV, 2005).

Os flavivírus apresentam uma distribuição cosmopolita. A evolução, o padrão de dispersão e as características epidemiológicas desses vírus são determinados pela combinação de fatores impostos pelo vetor artrópode e também pelo hospedeiro vertebrado. Outros fatores associados à ecologia, quase sempre deletéria, e à atividade humana ao meio ambiente têm, certamente, contribuindo à disseminação e a emergência dos flavivírus (Vasconcelos et al., 1992; Zanotto et al., 1995; Lindenbach et al., 2007; Gubler et al., 2007).

As infecções humanas por flavivírus permitem um espectro clínico variado, envolvendo desde infecções inaparentes, síndromes febris até formas severas potencialmente fatais, caracterizadas por encefalites ou por infecções sistêmicas que culminam no desenvolvimento de síndrome icterohemorrágica (Barrett, 2001).

Os flavivírus transmitidos por mosquitos incluem vários vírus de grande importância epidemiológica por causarem doenças em humanos e/ou animais de interesse econômico e por apresentarem ampla distribuição geográfica. Cerca de 40 flavivírus já foram associados à infecção humana, dentre os quais, além do VFA, se destacam o *Vírus dengue* (VDEN), o *Vírus da encefalite Saint Louis* (VESL), o *Vírus da encefalite japonesa* (VEJ), o *Vírus do Nilo Ocidental* (VNO) e o *Vírus rocio* (VROC), todos capazes de causar epidemias (Travassos da Rosa et al., 1997, Vasconcelos et al., 1998). O VFA é

reconhecido desde o século XVII como responsável por febres hemorrágicas, sendo que, segundo estimativa realizada pela Organização Mundial da Saúde (OMS), o vírus acomete cerca de 200.000 pessoas anualmente em países tropicais e subtropicais da África e América do Sul com a ocorrência de letalidade de 15% (Monath, 2001).

A Dengue é considerada um dos maiores problemas de saúde pública nas regiões tropicais e subtropicais, sendo o arbovírus de maior incidência no mundo. O VDEN, com seus quatro sorotipos reconhecidos - VDEN-1, VDEN-2, VDEN-3 e VDEN-4, é endêmico nas regiões tropicais das Américas, Ásia, África e Oceania e pode causar desde um quadro clínico inaparente, passando por uma síndrome febril - Dengue Clássico - até formas mais severas: a Febre Hemorrágica do Dengue (FHD), Síndrome do Choque do Dengue (SCD) (WHO, 2009). A estimativa da OMS de ocorrência anual de casos de Dengue Clássico é de 50 a 100 milhões de casos registrados em mais de 80 países. Desses, 500.000 casos são de FHD com o registro de 20.000 a 25.000 mortes, a maioria das quais em crianças (Gubler, 2006).

Em 1999, o VNO surgiu como um importante patógeno emergente nas Américas, causando encefalite em equinos e humanos; desde então o vírus vem se tornando endêmico nos EUA e Canadá e, pouco a pouco, vem sendo detectado em países latino-americanos (Lanciotti et al., 1999; Elizondo-Quiroga et al., 2005).

Dentre os flavivírus já identificados, onze vírus já foram isolados no Brasil: *Bussuquara* (BSQV), *Cacipacoré* (CPCV), VDEN (sorotipos 1, 2, 3 e 4), VFA, VROC, *Iguape* (VIGU), *Ilhéus* (VILH) e *Naranjal* (VNJL). Os VCPC, VIGU

e VNJL não foram isolados de seres humanos, nem relacionados a infecções humanas (Travassos da Rosa et al., 1997).

1.2. CICLO BIOLÓGICO

Os flavivírus podem ser transmitidos para humanos e outros vertebrados por meio da picada de artrópodes hematófagos ou entre hospedeiros artrópodes, via transovariana (Karabatsos, 1985). A transmissão para os artrópodes frequentemente necessita que o vírus produza uma elevada viremia no hospedeiro vertebrado e replique nos tecidos dos artrópodes infectados. Dois grupos de artrópodes estão envolvidos na transmissão dos flavivírus: mosquitos e carrapatos (Gubler et al., 2007).

Os artrópodes se tornam infectados durante o repasto sanguíneo. Após a ingestão, os vírus infectam células do intestino médio, estendendo para hemocela e então para os vários tecidos, incluindo a glândula salivar. Uma vez infectada a glândula salivar, o artrópode é capaz de transmitir os vírus através da saliva para os hospedeiros vertebrados, e nesses, após a replicação viral em diversos tipos celulares, atingem a corrente sangüínea (Karabatsos, 1985; Travassos da Rosa et al., 1997; Gubler et al., 2007). O período de tempo entre a infecção do artrópode através do repasto sangüíneo durante a viremia no hospedeiro vertebrado até quando as glândulas salivares do artrópode tornam-se infectadas é chamado de período de incubação extrínseco. Esse período de tempo dura de oito a 14 dias e depende da temperatura ambiente, da cepa viral

e do vetor. O período de incubação extrínseco não se aplica a vírus transmitidos por carrapatos (WHO, 1985; Gubler et al., 2007).

A transmissão vertical ou transovariana é a transferência de vírus entre gerações sem que haja nenhum outro modo de transmissão envolvido. Ambos os flavivírus, transmitidos por mosquitos e carrapatos, apresentam transmissão vertical. A transmissão transovariana ocorre quando os ovários dos artrópodes fêmeas se tornam infectados e o vírus é passado para os ovos. Quando os ovos eclodem, o vírus é mantido nos artrópodes nos diversos estágios de vida até a fase adulta. Na transmissão vertical, os ovos em maturação são infectados quando passam pelo oviduto (Rosen, 1987). Uma grande variedade de flavivírus transmitidos por mosquitos apresentam transmissão vertical em natureza, incluindo VDEN, VFA, VEJ, VNO (Gubler et al., 2007).

A habilidade dos artrópodes de transmitir os flavivírus depende de fatores intrínsecos e extrínsecos (Gubler et al., 2001, Kramer & Ebel, 2003). Um dos principais fatores intrínsecos é a variabilidade genética exibida pelos artrópodes. Dentre os fatores extrínsecos, observa-se que tanto mosquitos quanto carrapatos apresentam seus metabolismo e biologia altamente influenciados pelas condições do ambiente. A temperatura afeta diretamente a taxa de replicação viral e, conseqüentemente, o período de incubação extrínseco. Dentro dos limites de cada espécie, altas temperaturas podem aumentar a transmissão viral ou podem ocasionar um aumento da taxa de mortalidade dos mosquitos. Ademais, umidade, períodos chuvosos e velocidade do vento podem influenciar na taxa de mortalidade, bem como nas

características biológicas e comportamentais dos vetores, tais como extensão de vôo e oviposição (Kramer & Ebel, 2003).

A extensão de hospedeiros vertebrados entre os vetores artrópodes é bastante variável. Carrapatos geralmente são mais específicos, mas infestam diferentes hospedeiros em diferentes estágios do seu ciclo de vida. Sendo assim, carrapatos imaturos podem infectar hospedeiros naturais, tais como os roedores. Nos estágios de ninfa e fase adulta, os carrapatos podem se alimentar de outros hospedeiros, como os humanos, permitindo que vírus zoonóticos de roedores possam infectar o homem, tal como ocorre com o *Vírus da encefalite transmitida por carrapato* (do inglês *Tick borne encephalitis virus* - VTBE) (Gubler et al., 2007).

Quanto aos mosquitos, sua preferência por hospedeiros não é específica. Alguns se alimentam exclusivamente de aves, enquanto outros se alimentam de mamíferos, sendo, portanto, os vetores de maior importância na transmissão e disseminação dos flavivírus. Espécies dos gêneros *Culex* e *Aedes* estão mais relacionadas com a transmissão dos flavivírus, muito embora outros gêneros também possam estar envolvidos na manutenção dos ciclos biológicos desses vírus, tais como *Sabethes* sp. e *Haemagogus* sp. (Travassos da Rosa et al., 1997; Vasconcelos et al., 1998). Há uma estreita relação entre a transmissão de flavivírus que causam encefalites por *Culex* sp., tal como VNO que é mantido em seu ciclo enzoótico por *Cx. pipiens* e *Cx. restuans*, enquanto que os flavivírus que causam febres hemorrágicas são transmitidos por *Aedes* sp., *Aedes* sp. e *Haemagogus* sp., como por exemplo VDEN e VFA, respectivamente (Kuno et al., 1998; Gaunt et al., 2001; Gubler et al., 2007).

1.3. CLASSIFICAÇÃO DOS FLAVIVÍRUS

Dentro do gênero, os flavivírus são categorizados em complexos antigênicos e subcomplexos, baseado em critérios sorológicos clássicos, ou dentro de grupo, clados e espécies de acordo com a filogenia (Calisher & Gould, 2003).

O relacionamento antigênico dos flavivírus é determinado pela reatividade de testes sorológicos de inibição da hemaglutinação (IH), fixação do complemento (FC) e neutralização (TN). Assim, os flavivírus são distribuídos sorologicamente em dois grupos antigênicos: o grupo B e o grupo *Vírus da febre hemorrágica de macaco* (*Simian hemorrhagic fever virus*), existindo ainda flavivírus não sorogrupados (Karabatsos, 1985; Travassos da Rosa et al., 1998; ICTV, 2005). Uma classificação antigênica mais antiga determinou a existência de oito complexos: Os quatro sorotipos do VDEN pertencem ao complexo Dengue; os VEJ, VESL, VNO, *Vírus Kunjin* (VKUN), *Vírus da Encefalite Murray Valley* (VEMV), entre outros, são grupados no sorocomplexo da Encefalite Japonesa; VTBE, *Vírus Louping ill* (VLI), *Vírus langat* (VLGT), *Powassan* (VPOW), *Vírus da Febre hemorrágica do Omsk* (VOHF) e o *Vírus da doença da floresta Kyasanur* (*Kyasanur Forest disease virus* - VKFD) formam o complexo das encefalites transmitidas por carrapato; o VFA é o único membro do complexo da febre amarela; existindo ainda os complexos Tyuleny (*Vírus Tyuleny* - VTYU), Ntaya (*Vírus Ntaya* - VNTA), Uganda S (*Vírus Uganda S* - VUGS) e Rio Bravo (*Vírus Rio Bravo* - VRB) (Calisher et al, 1989; Pugachev et al., 2003).

O gênero *Flavivirus* é composto, atualmente, por 53 espécies de arbovírus, das quais 27 são flavivírus transmitidos por mosquitos, 12 são transmitidos por carrapatos e 14 são agentes zoonóticos com vetor desconhecido (Quadro 1) (ICTV, 2005), além dos vírus que estão em processo de inclusão no gênero: *Vírus Tamana Bat* (VTAB), isolado de morcego (de Lamballerie et al., 2002), e os vírus relacionados como flavivírus de insetos: *Vírus agente fusionador celular* (*Cell fusion agents virus* - VCFA) (Cammisa-Park et al., 1992), *Vírus Rio Kamitii* (*Kamitii River virus* - VKR) (Crabtree et al., 2003), *Vírus Aedes flavivírus* (VAEF) (Hoshino et al., 2009), *Vírus Culex flavivírus* (VCXF) (Hoshino et al., 2007) e *Vírus Quang Binh* (VQB) (Crabtree et al., 2009). Esta classificação é baseada nos conceitos de espécie viral e no conjunto de informações sobre morfologia, organização do genoma, relação das sequências nucleotídicas, associação com vetor e ecologia viral (ICTV, 2005; Gubler et al., 2007).

Considerando a mortalidade e morbidade, 22 cepas de flavivírus transmitidos por mosquitos, 13 dos vírus transmitidos por carrapatos e cinco dos agentes zoonóticos com vetor desconhecido são capazes de causar doença em humanos (Gubler et al., 2007).

Os flavivírus transmitidos por mosquitos têm evoluído em dois grupos maiores distintos pela apresentação clínica em humanos e ecologia. Os flavivírus encefalíticos estão no grupo da encefalite japonesa, incluindo os vírus VEJ, VNO, VEMV, VROC e VESL. Todos são arbovírus zoonóticos que apresentam como hospedeiros vertebrados naturais aves e como vetores mosquitos do gênero *Culex* (Gubler et al., 2007). O outro grupo inclui o VFA e o

VLEN, que são vírus viscerotrópicos e podem causar febres hemorrágicas. Esses vírus apresentam um ciclo silvestre tendo primatas como hospedeiros vertebrados e mosquitos do gênero *Aedes* como principal vetor. Os quatro sorotipos do VLEN são adaptados aos humanos no ambiente urbano e, geralmente, não necessitam de ciclo silvestre para a sua manutenção (Gubler et al., 2007).

Quadro 1 - Espécies virais e cepas registradas no CITV, distribuídos por grupos e sorogrupos (ICTV, 2005)

FLAVIVÍRUS TRANSMITIDOS POR MOSQUITOS				
GRUPO VIRAL	ESPÉCIE VIRAL	CEPAS / SOROTIPOS	SIGLA	GENBANK
Aroa	<i>Vírus Aroa</i>	Vírus Aroa	(VAROA)	AF013362 ^b
		Vírus Bussuquara	(VBSQ)	NC009026*
		Vírus Iguape	(VIGU)	NC009027*
		Vírus Naranjal	(VNJL)	AF013390 ^b
Dengue	<i>Vírus Dengue</i>	Vírus Dengue 1	(VDEN- 1)	NC001477*
		Vírus Dengue 2	(VDEN- 2)	NC001474*
		Vírus Dengue 3	(VDEN- 3)	NC001475*
		Vírus Dengue 4	(VDEN- 4)	NC002640*
Kedougou	<i>Vírus Kedougou</i>		(VKED)	NC012533*
Encefalite Japonesa	<i>Vírus Cacipacore</i>		(VCPC)	AF013367 ^b
	<i>Vírus da Encefalite Japonesa</i>		(VEJ)	NC001437*
	<i>Vírus Koutango</i>		(VKOU)	AF013384 ^b
	<i>Vírus Murray Valley encefalite</i>	Vírus Alfuy	(VALF)	AF013360 ^a
		Vírus Murray Valley encephalitis	(VEMV)	NC000943*
	<i>Vírus St. Louis encefalite</i>		(VESL)	NC007580*
	<i>Vírus Usutu</i>		(VUSU)	NC006551*
	<i>Vírus Nilo do Oeste</i>	Vírus Kunjin	(VKUN)	EU074020 ^a
	Vírus Nilo do Oeste	(VNO)	NC009942*	
Kokobera	<i>Vírus Kokobera</i>	Vírus Kokobera	(VKOK)	NC009029*
		Vírus Stratford	(VSTR)	AF013407 ^b
Ntaya	<i>Vírus Bagaza</i>		(VBAG)	NC012534*
	<i>Vírus Ilhéus</i>	Vírus Ilheus	(VILH)	NC009028*
		Vírus rocio	(VROC)	AF013397*
	<i>Vírus Israel turkey meningoencefalomyelite</i>		(VIT)	AF013377 ^b
	<i>Vírus Ntaya</i>		(VNTA)	AF013392 ^b
Spondweni vírus	<i>Vírus Zika</i>	Vírus Spondweni	(VSPO)	AF013406 ^b
		Vírus Zika	(VZIK)	NC012532*
Febre Amarela	<i>Vírus Banzi</i>	Vírus Banzi	(VBAN)	L40951 ^b
	<i>Vírus Bouboui</i>		(VBOU)	AF013364 ^b
	<i>Vírus Edge Hill</i>		(VEH)	AF013372 ^b
	<i>Vírus Jugra</i>		(VJUG)	AF013378 ^b
	<i>Vírus Saboya</i>	Vírus Potiskum	(VPOT)	AF013395 ^b
		Vírus Saboya	(VSAB)	AF013400 ^b
	<i>Vírus Sepik</i>		(VSEP)	NC008719*
	<i>Vírus Uganda S</i>		(VUGS)	EU074020 ^b
	<i>Vírus Wesselsbron</i>		(VWESS)	NC012735*
<i>Vírus da febre amarela</i>		(VFA)	NC002031*	

Quadro 1 - Espécies virais e cepas registradas no CITV, distribuídos por grupos e sorogrupos (continuação)

FLAVIVÍRUS TRANSMITIDOS POR CARRAPATOS					
GRUPO VIRAL	ESPÉCIE VIRAL	CEPAS / SOROTIPOS	SIGLA	GENBANK	
flavivírus transmitidos por carrapatos com hospedeiros mamíferos	<i>Vírus Gadgets Gully</i>		(VGGY)	AF013374 ^b	
	<i>Vírus Kyasanur Forest disease</i>		(VKFD)	X74111 ^b	
	<i>Vírus Langat</i>		(VLGT)	NC003690*	
	<i>Vírus Louping ill</i>	British subtype		(VLI-Brit)	D12937 ^b
		Irish subtype		(VLI-1r)	X86784 ^b
		Vírus Louping ill		(VLI)	NC001809*
		Spanish subtype		(VLI-Span)	X77470 ^b
	<i>Vírus Omsk hemorrhagic fever</i>	Turkish subtype		(VLI-Turk)	X69125 ^b
	<i>Vírus Powassan</i>			(VOHF)	NC005062*
<i>Vírus Royal Farm</i>	Vírus Karshi		(VKSI)	NC006947*	
	Vírus Royal Farm		(VRF)	AF013398 ^b	
<i>Vírus Tick-borne encefalite</i>	European subtype		(VTBE-Eu)	M27157 ^b	
	European subtype		(VTBE-Eu)	M33668 ^b	
	Far Eastern subtype		(VTBE-FE)	X07755 ^b	
	Siberian subtype		(VTBE-Sib)	L40361 ^b	
	Vírus Tick-borne encefalite		(VTBE)	NC001672*	
flavivírus transmitidos por carrapatos com hospedeiros aves marinhas	<i>Vírus Kadam</i>		(VKAD)	AF013380 ^b	
	<i>Vírus Meaban</i>		(VMEA)	AF013386 ^b	
	<i>Vírus Saumarez Reef</i>		(VSRE)	X80589 ^b	
	<i>Vírus Tyulenyi</i>		(VTYU)	X80588 ^b	
AGENTES ZONÓTICOS COM VETORES DESCONHECIDO					
Grupo do Entebbe bat	<i>Vírus Entebbe bat</i>	Vírus Entebbe bat	(VENT)	NC008718*	
		Vírus Sokoluk	(VSOK)	AF013405 ^b	
	<i>Vírus Yokose</i>		(VYOK)	NC005039*	
	<i>Vírus Apoi</i>		(VAPOI)	NC003676*	
	<i>Vírus Cowbone Ridge</i>		(VCR)	AF013370 ^b	
	<i>Vírus Jutiapa</i>		(VJUT)	AF013379 ^b	
	<i>Vírus Modoc</i>		(VMOD)	NC003635*	
	<i>Vírus Sal Vieja</i>		(VSV)	AF013401 ^b	
	<i>Vírus San Perlita</i>		(VSP)	AF013402 ^b	
	<i>Vírus Bukalasa bat</i>		(VBB)	AF013365 ^b	
	<i>Vírus Carey Island</i>		(VCI)	AF013368 ^b	
	<i>Vírus Dakar bat</i>		(VDB)	AF013371 ^b	
	<i>Vírus Montana miotis leucoencefalite</i>		(VMML)	NC004119*	
	<i>Vírus Phnom Penh bat</i>	Vírus Batu Cave		(VBC)	AF013369 ^b
Vírus Phnom Penh bat			(VPPB)	AF013394 ^b	
<i>Vírus Rio Bravo</i>		(VRB)	NC003675*		
TENTATIVA DE INCLUSÃO NO GÊNERO					
Flavivírus de insetos	<i>Vírus Cell fusing agent</i>		(VCFA)	NC001564*	
	<i>Vírus Kamiti River</i>		(VKR)	NC005064*	
	<i>Vírus Tamana bat</i>		(VTAB)	NC003996*	

* número de acesso ao GenBank referente ao genoma completo; ^a referente à sequência nucleotídica da poliproteína; ^b referente ao gene NS5.

Fonte: ICTV, 2005

1.4. ESTRUTURA DOS FLAVIVÍRUS

Os integrantes do gênero *Flavivirus* possuem partículas virais esféricas, medindo cerca de 50 nm de diâmetro; são vírus com envelope lipoprotéico, que envolve o núcleo protéico de aproximadamente 30 nm: o nucleocapsídeo, que é constituído pela proteína C e protege o genoma viral. A superfície contém duas proteínas virais, as glicoproteínas E (envelope) e PrM/M (membrana) (Figura 1). A glicoproteína E é o principal determinante antigênico da partícula viral, media a ligação e a fusão durante a entrada do vírus na célula. A proteína Pr é um pequeno fragmento proteolítico precursor da proteína M; é produzida durante a maturação da partícula viral, sendo encontrada no meio intracelular (Lindenbach & Rice, 2003; Lindenbach et al., 2007).

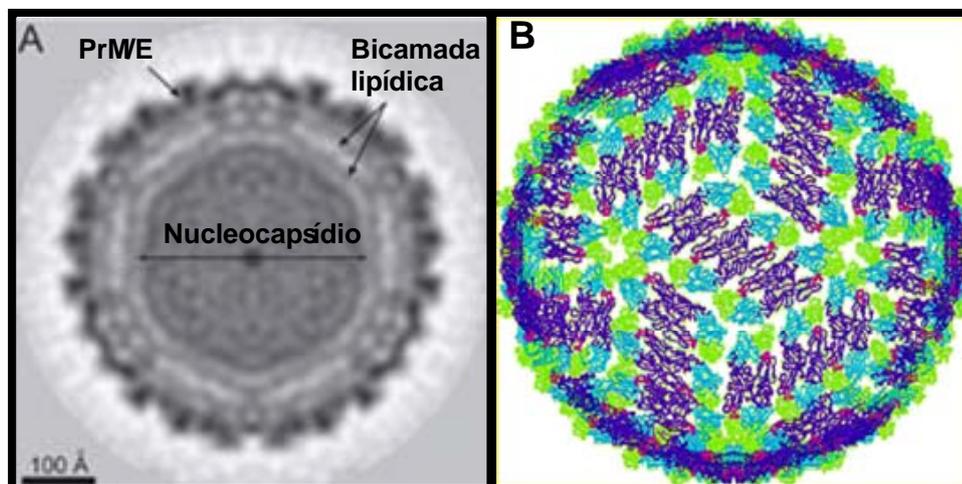


Figura 1 - (A) Crioeletromicroscopia do VDEN, mostrando o envelope viral, no qual estão inseridas as proteínas estruturais PrM/E, e o nucleocapsídeo com aproximadamente 50 nm de diâmetro; (B) Mapa eletron denso dos flavivírus, destacando os domínios I, II e III (verde, roxo e azul, respectivamente) e o peptídeo de fusão (vermelho) da proteína E dos flavivírus. Fonte: (A) <http://www.lightsources.org/cms/?pid=1002693>, (B) Adaptado de Lindenbach & Rice, 2003.

O Genoma dos flavivírus consiste de um RNA de fita simples de polaridade positiva, com média de 11.000 nucleotídeos (nt); possui um *cap* tipo I (m7GppAmpN₂) na extremidade 5', e metilação do resíduo N₂ (*cap* tipo II) e cauda poliadenilada na extremidade 3', que é perdida durante a replicação quando o RNA viral funciona como RNA mensageiro (RNAm) (ICTV, 2005; Lindenbach et al., 2007).

Durante cultivos celulares, duas formas virais podem ser distinguidas: a forma viral madura, encontrada fora das células, possui duas glicoproteínas associadas ao envelope viral, E e M; e a forma imatura, exclusiva do meio intracelular, que apresenta o precursor da proteína M, PrM, o qual é clivado durante o processo de maturação (Figura 2) (Lindenbach et al., 2007).

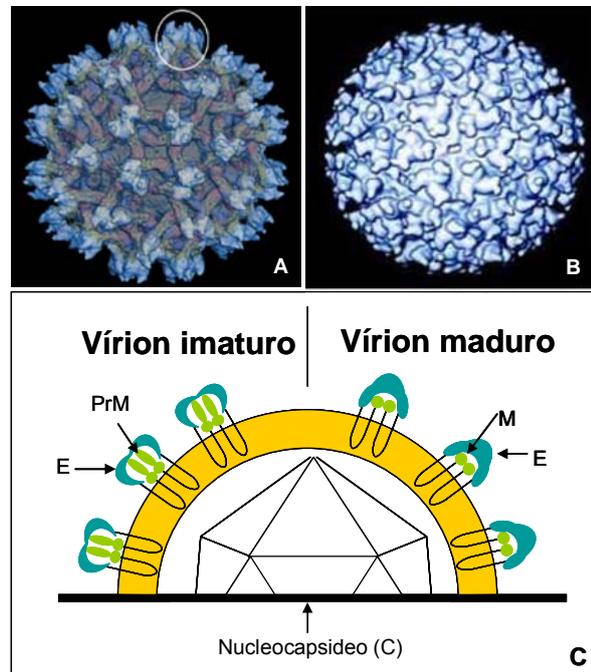


Figura 2 – Formas das partículas virais dos flavivírus (A) partícula viral imatura; (B) partícula viral madura; (C) esquema das formas madura e imatura. Fonte: (A) <http://www.purdue.edu/uns/x/2007b/071019CelHockmeyer.html>; (B) Lindenbach et al., 2007; (C) adaptado de Heinz & Allison, 2003.

1.5. REPLICAÇÃO DOS FLAVIVÍRUS

Para o processo de replicação dos flavivírus, diferentes tipos celulares e de espécies hospedeiras estão envolvidos, sendo utilizados diversos receptores que proporcionam a adsorção e, em seguida, a entrada e replicação viral. A infecção em células dendríticas é de particular importância, por serem os alvos iniciais no processo de infecção (Lindenbach et al., 2007). Estudos com o VDEN mostraram que a infecção das células dendríticas depende do receptor DC-SIGN (CD209) - um tipo de Lectina tipo C -, que participa somente do processo de adsorção viral (Navarro-Sanchez et al., 2003; Tassaneetrithep et al., 2003). Outras proteínas foram identificadas como receptores para flavivírus, incluindo as $\alpha_v\beta_3$ integrina, GRF78 (BiP), a CD14, glicosaminoglicanas altamente sulfatadas (exemplo: heparan sulfato) (Chen et al., 1997b; Kroschewski et al., 2003). Ademais, a opsonização das partículas virais por imunoglobulinas aumentam a infecção de células que expressam receptores para a porção Fc das imunoglobulinas (Barba-Spaeth et al., 2005).

A internalização das partículas virais é mediada via interação com proteína tipo Clatrina, que são encaminhadas em direção aos compartimentos endocíticos perilisossomal, na qual o pH ácido induz a fusão entre o envelope viral e a membrana da célula hospedeira, liberando o nucleocapsídeo para o citoplasma. Devido à fluidez do capsídeo, o RNA viral é disponibilizado para tradução logo após a fusão entre as membranas. A tradução do RNA viral produz proteínas que permitem a nova síntese do RNA viral e a montagem de novas partículas virais (Lindenbach & Rice, 2003; Lindenbach et al., 2007).

Uma vez no citoplasma, o RNA viral desempenha três papéis importantes para a replicação viral: (i) como RNAm para a tradução de todas as proteínas virais; (ii) molde para a síntese de RNA; e (iii) material genômico empacotado dentro das partículas virais recém sintetizadas (Lindenbach et al., 2007).

A eficiência da tradução pode ser o primeiro determinante da infectividade dos flavivírus. Portanto, esses vírus utilizam diversos mecanismos para o processo de tradução, incluindo estruturas especializadas dentro das regiões não codificadoras (RNC) 3' e 5'. A tradução é dependente da estrutura *cap* tipo I e inicia com o *scanner* ribossomal. O códon de iniciação para muitos flavivírus transmitidos por mosquitos não apresentam o motivo Kozak, onde se observa frequentemente a presença de múltiplas sequências AUG próximas umas das outras (sequências *in-frame*). Para ajudar a seleção do sítio de iniciação, o VDEN aparentemente utiliza uma pequena estrutura em grampo (do inglês *hairpin*) no gene C para induzir a parada do ribossomo no códon de parada AUG (Clyde & Harris, 2006)

Uma única cadeia de leitura codifica uma grande poliproteína de aproximadamente 3.400 aminoácidos (aa), que é clivada co- e pós-traducionalmente por proteínas de origem celular e viral em dez proteínas, sendo três estruturais e sete não-estruturais (ICTV, 2005; Lindenbach et al., 2007). A sequência da tradução da poliproteína dos flavivírus foi confirmada com base no mapeamento da tradução de células infectadas pelo *Vírus kunjin* (VKUN) (Scheader & Westaway, 1988), na qual o processamento completo se

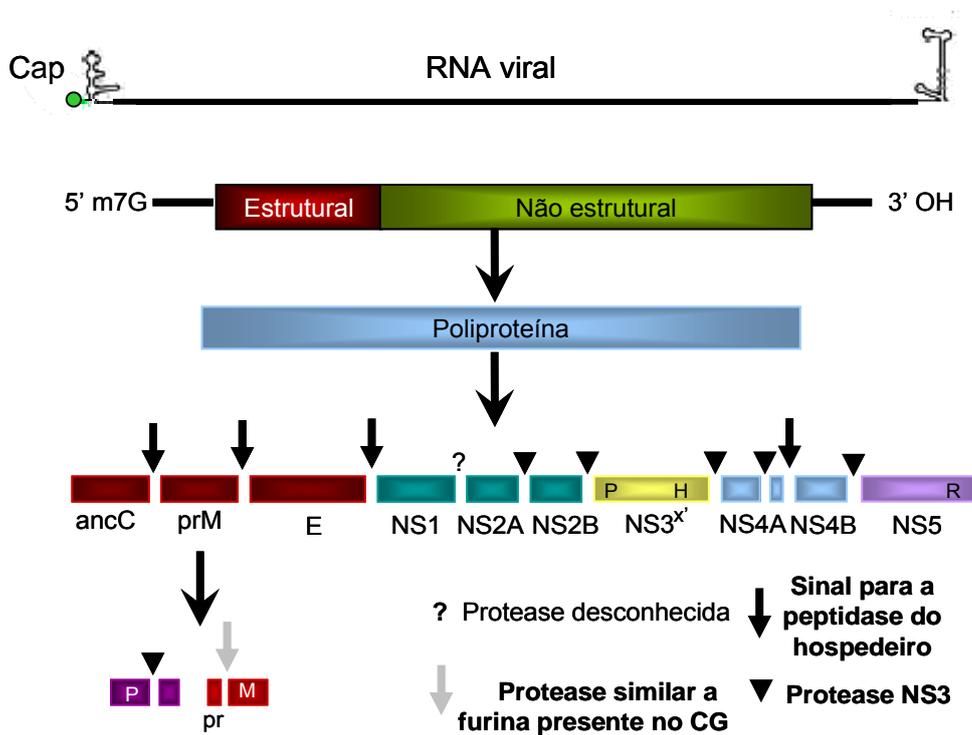
dá em torno de 17 minutos estendido a 30 minutos devido aos processos de clivagem pós-traducional (Westaway et al., 2003).

Na poliproteína, as proteínas estruturais C-prM-E localizam-se na porção N-terminal e são seguidas das proteínas não-estruturais NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 (Chambers et al., 1990; Lindenbach & Rice, 2003). A peptidase do hospedeiro é responsável pela clivagem entre as junções C/PrM, PrM/E, E/NS1 e 2K-NS4B. A protease serina codificada pelo vírus é responsável pela clivagem entre as junções NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/2K e NS4B/NS5. A enzima responsável pela clivagem NS1/NS2A ainda permanece desconhecida. Uma protease similar a furina presente no complexo de Golgi (CG) faz a clivagem da Pr/M durante o processo de maturação da partícula viral (Figura 3) (Lindenbach et al., 2007).

A síntese do RNA viral ocorre na membrana do retículo endoplasmático (RE) perinuclear, e pode ser detectada três horas pós-infecção (Lindenbach & Rice, 1997). Após a tradução do RNA genômico, a síntese do RNA começa com a formação da fita negativa complementar que é usada como molde para a síntese do RNA viral de polaridade positiva. Estas fitas são sintetizadas por um mecanismo semiconservativo envolvendo intermediários replicativos, contendo regiões de fitas duplas, bem como moléculas de fita simples nascente e formas replicativas (moléculas de RNA duplex) (ICTV, 2005). A síntese do RNA viral é assimétrica, com o acúmulo da fita RNA positiva em relação à fita molde (RNA negativa) (Cleaves et al., 1981; Muylaert et al., 1996). Para a maioria dos flavivírus esse RNA viral metabólico tem sido descrito como: RNA genômico; uma forma intermediária constituída por dupla-

fita; e uma população heterogênea da síntese intermediária, que comumente apresenta regiões duplex e o RNA padrão recém-sintetizado. Essas duas últimas formas são reconhecidas como precursores do RNA genômico (Chun & Westaway, 1985).

(A) **Genoma dos *Flavivirus***



Topologia das proteínas virais na membrana do retículo endoplasmático

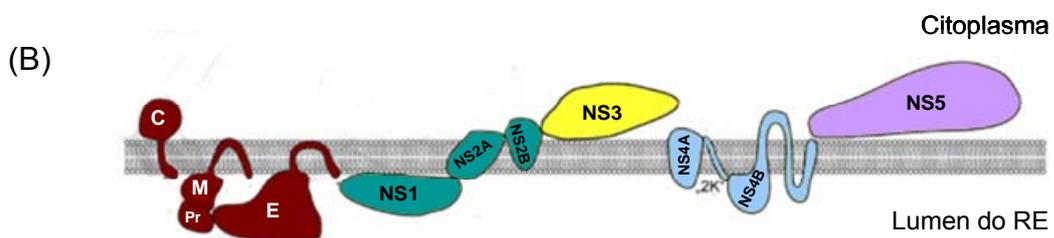


Figura 3 – (A) Apresentação esquemática do genoma e processamento das poliproteínas dos integrantes do gênero *Flavivirus*. (B) Topologia e interação dos flavivírus ao nível da membrana do retículo endoplasmático. Fonte: (A) adaptado de ICTV, 2005; (B) Lindenbach et al., 2007.

Durante a montagem das partículas virais, são observadas ao microscópio eletrônico mudanças nas membranas das células infectadas, principalmente, na região perinuclear. Inicialmente, ocorre a proliferação da membrana do RE, seguido da formação de estruturas de membranas densas, que são grupamentos de vesículas de 70 a 200 nm de diâmetro, contendo material eletrondenso dentro do lúmen do RE rugoso. Essa estrutura continua acumulando ao longo do processo de infecção, quando elas tornam-se adjacentes dando origem a uma nova estrutura: aglomeração de membranas, que parece ser contínua com o RE. O RNA viral acumula em associação às vesículas empacotadas na região perinuclear (Lindenbach et al., 2007).

As proteínas virais NS2B e NS3 localizam-se nas membranas aglomeradas para o processo de reorganização da membrana, muito embora estruturas subcelulares distintas estejam dedicadas ao processamento da poliproteína versus síntese da replicação, resultando no acúmulo de partículas virais dentro dos compartimentos do RE. Os vírions nascentes são, então, transportados, através das vias secretórias, até a superfície celular. O estágio tardio da maturação viral está associado ao processamento da PrM e da proteína C, bem como à glicosilação das proteínas E e PrM num mecanismo similar da glicosilação das proteínas celulares. Finalizando o processo, as partículas virais maduras são liberadas por exocitose (Figura 4) (Lindenbach et al., 2007).

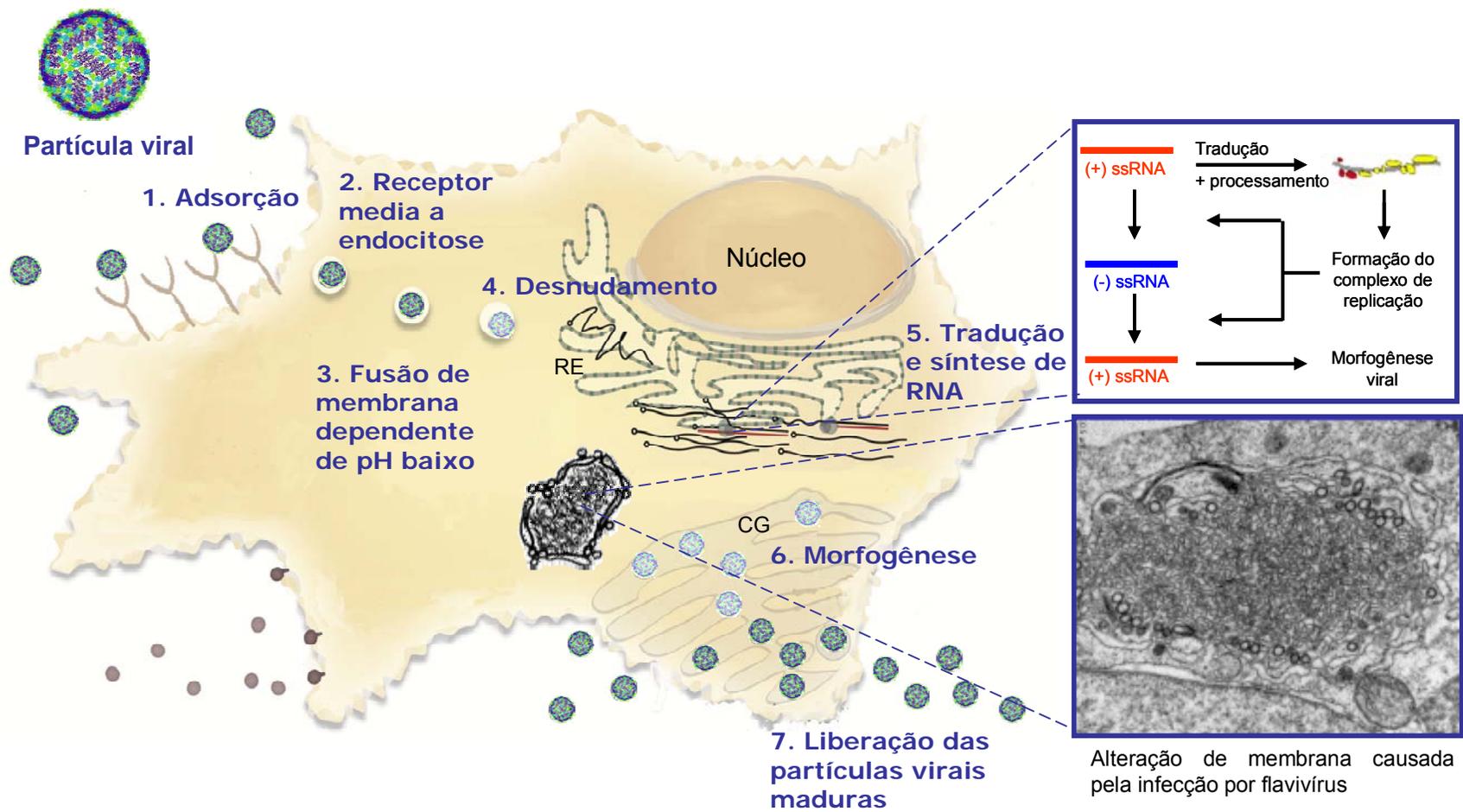


Figura 4 - Esquema da replicação viral dos flavivírus. Fonte: adaptado de Clyde et al., 2006 e www.klinikum.uni-heidelberg.de/2-Replication-Cycle.104948.0.html

1.6. GENOMA DOS FLAVIVÍRUS

1.6. GENOMA DOS FLAVIVÍRUS

O genoma dos flavivírus apresenta uma única matriz de leitura aberta (*Open Reading Frame* - ORF), que é flanqueada por duas RNC, 5'RNC e 3'RNC, com aproximadamente 100 nucleotídeos (nt) e 350 a 700 nt, respectivamente. São reconhecidos dez genes seguidos no sentido 5'-3': C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 (Figura 3) (Rice et al., 1985; Chambers et al., 1990; Lindenbach et al., 2007).

As RNCs no RNA genômico de flavivírus podem estar envolvidas na iniciação da síntese de RNA de sentido negativo e em muitos processos podem estar relacionadas com a mudança na síntese do RNA de sentido negativo para a produção de RNA viral da progênie (sentido positivo). Ademais, as RNCs podem também desempenhar um importante papel no empacotamento dos vírions nascentes e conter sequências nucleotídicas requeridas para iniciar a tradução (Markoff, 2003).

Quando se compara a extensão das RNCs dos flavivírus, nota-se que algumas 5'RNC são extremamente pequenas. De modo geral, a 5'RNC dos flavivírus transmitidos por mosquitos são menores quando comparadas com as 5'RNC de flavivírus transmitidos por carrapatos (Quadro 2) (Markoff, 2003).

A extensão das 3'RNCs dos flavivírus está diretamente ligada à presença ou ausência de uma cauda poli(A). A perda dessa cauda poli(A) em

algumas espécies virais é causada por deleções das sequências que flanqueiam essa estrutura (Markoff, 2003).

Quadro 2 - Comparação dos tamanhos nucleotídicos das 5' e 3'RNCs no genoma de alguns flavivírus (Adaptado de Markoff, 2003).

Vírus	5'RNC (nt)	3'RNC (nt)	Genoma (nt)
VDEN-1 cepa WP74	97	466	10.736
VDEN-2 cepa NGC	89	451	10.723
VDEN-3 cepa H87	93	432	10.696
VDEN-4 cepa 814669	101	384	10.649
VEJ cepa JaOAR5982	95	585	10.976
VNO CEPA NY99	96	631	11.029
VEMV	95	614	11.014
VFA cepa 17D	118	511	10.862
VTBE cepa Sofjin	131	518	10.894
VTBE cepa Neudorfl	132	764	10.469
VPOW	111	480	10.839
VCFA	113	559	10.695

1.6.1. Região não-codificante 5'

A região 5'RNC tem por característica não ser completamente conservada entre os flavivírus, embora possam ser encontradas sequências

nucleotíficas com significativa homologia entre subgrupos distintos, entre vírus de um mesmo grupo sorológico e entre cepas de uma mesma espécie viral (Brinton & Dispoto, 1988), além da formação de uma estrutura secundária com tamanho, forma e estabilidade termodinâmica similar aos integrantes do gênero (Markoff, 2003).

A forma secundária consiste de uma haste com uma pequena alça no topo da estrutura e uma ramificação lateral. Essa estrutura trata-se do pareamento de bases nucleotídicas por pontes de hidrogênio, que é invariavelmente interrompida por alças formadas pela aposição de nucleotídicos que são incapazes de formar tais ligações intermoleculares. A ramificação lateral é formada por sequências 3' terminal da 5'RNC, incluindo o códon de iniciação da ORF e por nucleotídeos da porção inicial do gene C. A estabilidade termodinâmica dessa estrutura fica em torno de -17,3 kcal (com base na sequência do VESL) a -27,8 kcal (para a cepa E101 do VNO). A conservação dessa estrutura entre os flavivírus evidencia a possível participação e relevância da 5'RNC na replicação viral (Markoff, 2003) (Figura 5). Por vez, uma segunda haste-alça pode ser observada (Mandl et al., 1993) .

Dentre os processos funcionais, a 5'RNC pode estar envolvida na tradução do genoma e na complementaridade da cadeia negativa, que serve de sítio de iniciação da cadeia positiva durante a síntese do RNA (Brinton & Dispoto, 1988; Lindenbach & Rice, 2003). Ademais, estudos sugerem que fatores hospedeiro-específicos interagem com a 5'RNC (Li et al., 2002; Yocupicio-Monroy et al., 2003).

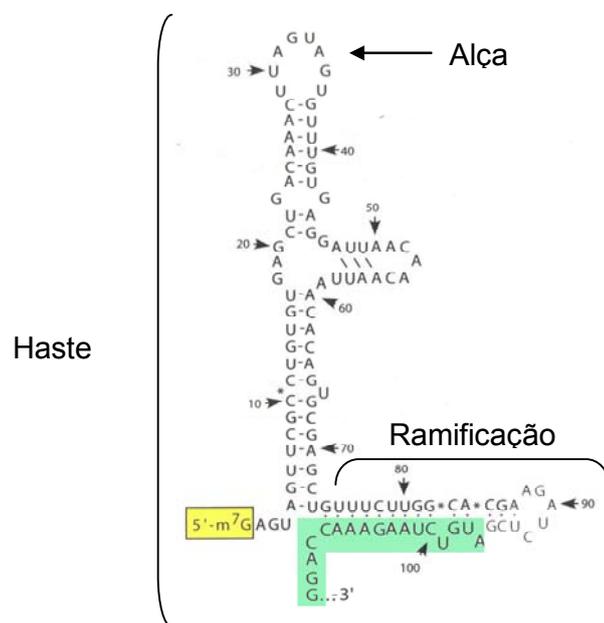


Figura 5 - Estrutura secundária da 5'RNC do VNO. Nucleotídeos são numerados (seta) a partir do terminal 5', excluindo o cap (5'-m⁷G) (amarelo). Nucleotídeos da ORF começam a partir do nucleotídeo número 96 e são mostrados em verde. Fonte: Brinton & Dispoto, 1988.

1.6.2. Região não-codificante 3'

A região 3'RNC exibe uma grande variabilidade no gênero, no entanto, apresenta sequências conservadas entre os flavivírus transmitidos por mosquitos, sendo estas diferenciadas das sequências observadas no genoma dos flavivírus transmitidos por carrapatos (ICTV, 2005). Essa variabilidade é determinada por diferenças no tamanho da região 3'RNC, que é determinada pela inserção ou deleção de sequências repetidas na porção inicial, o que contrasta com a região altamente conservada da porção final, quando se comparam vírus do mesmo sorocomplexo/grupo (Poindinger et al., 1996).

A maior estrutura de similaridade entre os flavivírus é reconhecida como 3' haste-alça (do inglês 3' *stem loop* - 3'SL) - formada por aproximadamente 30 pontes de hidrogênio -, que difere na sequência primária entre os flavivírus transmitidos por mosquitos e os transmitidos por carrapatos (Tilgner et al., 2005; Markoff, 2003; Yu & Markoff, 2005) e apresenta uma estabilidade térmica em torno de 40 a 45,2 kcal (Khromykh et al., 2001b; Markoff, 2003). Estudos com VNO e VDEN-2 revelaram regiões funcionais vírus-específicas e hospedeiro-específicas dentro da 3'SL (Tilgner et al., 2005; Yu & Markoff, 2005). Outrossim, a região 3'SL pode interagir com muitas proteínas de relevância funcional, ativando as polimerases virais NS3 e NS5 (Chen et al., 1997a; Cui et al., 1998). Uma pequena haste-alça adjacente, característica da 3'RNC dos flavivírus, é observada entre os nucleotídeos 90 a 100, sentido 3'→5', a partir do terminal 3' do genoma (Markoff, 2003) (Figura 6).

Na porção mais próxima do terminal 3' localiza-se o pentanucleotídeo 5'-CACAG, que faz parte de uma região não-pareada dentro da estrutura 3'SL (Hahn et al., 1987). A função do pentanucleotídeo no ciclo replicativo dos flavivírus ainda não é conhecida, mas há hipóteses que sugerem sua participação na ligação celular ou de proteínas virais na estrutura 3'SL durante a síntese do RNA (Khromykh et al., 2001a) (Figura 6).

No sentido 5' terminal, uma região de 25 nt, reconhecida como CS1, é bem conservada entre os flavivírus transmitidos por mosquitos e apresenta uma região pareada complementar (5'CYC) no início do gene do capsídeo (gene C). A complementaridade entre essas sequências para a

ciclização é essencial para a replicação dos flavivírus e é requerida para a seleção do molde de RNA para a sua síntese, no entanto a ciclização do genoma aparentemente não está envolvida no processo de tradução (Lindenbach et al., 2007). A CS1 não está presente no genoma dos flavivírus transmitidos por carrapatos e não é totalmente conservada entre os vírus de um mesmo grupo (Markoff, 2003).

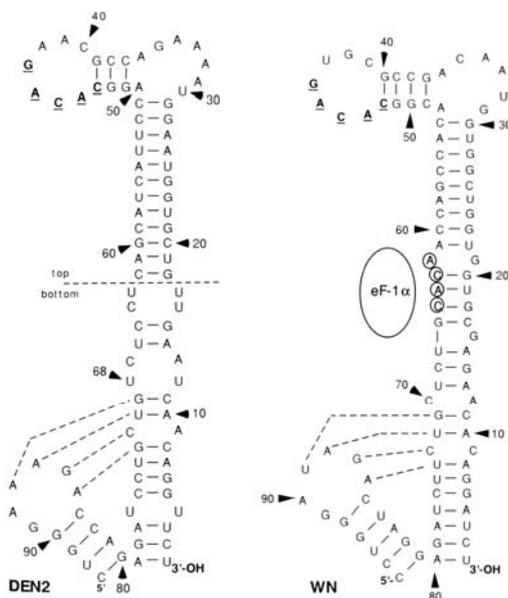


Figura 6 – Estrutura secundária preditiva da 3'SL na 3'RNC dos flavivírus VDEN-2 e VNO. Em negrito, destaca-se o pentanucleotídeo 5'-CACAG-3'. Fonte: Markoff, 2003.

Uma ou duas cópias de uma segunda sequência conservada - CS2 e RCS2 - também podem ser encontradas em alguns flavivírus transmitidos por mosquitos, mas não está presente na 3'RNC de flavivírus transmitidos por carrapatos (Hahn et al., 1987). Deleções na CS2 ou RCS2 tornam os vírus viáveis, mas altamente atenuados (Alvarez et al., 2005a; Lo et al., 2003).

A CS2 localiza-se próximo ao terminal 3', entre os nucleotídeos 12 a 22 a frente da CS1 (sentido acima, 3'→5') referente aos genomas do VDEN, VFA e dos vírus do grupo VEJ. A sequência de 20 nt da CS2 é completamente conservada entre os vírus do grupo da encefalite japonesa (VEMV, VNO e VKUN) e entre os quatro sorotipos do VDEN, exceto por uma substituição simples (G→U) no genoma do VDEN-3 (Figura 7) (Markoff, 2003).

<u>CS1</u>		<u>CS2</u>	
Consenso	AGC AUAUUGAC ACCUGGGAA-AGAC	Consenso	GGACUAGAGGUAGAGGAGACCC
VEMV A	VEMV
VNO U -.....	VNO
VEJ U	VEJ
VDEN-1 G -..... G -.....	VDEN-1
VDEN-2 G -..... G -.....	VDEN-3 U
VFA	C G . A-	VFA	U U

<u>CS3 e RCS3 NA 3'RNS DO VEMV</u>	
CS3	CCCCAGGAGGACUGGGUUACCAAAGCUG
RCS3 A . A C .

<u>REPETIÇÕES EM SÉRIE NA 3'RNC DO VFA</u>	
1	* UAACCGGGAUACAAACCACGGGUGGAGAACCGGACUCCCCACA
2	GA..... U C G G U
3	GA..... A U A A

Figura 7 – Sequências nucleotídicas de alguns motivos conservados na região 3'RNC do genoma de flavivírus transmitidos por mosquitos. Deleções na CS1 são representadas por ponto (.). Substituições nucleotídicas encontram-se destacadas em azul e sublinhadas. Em vermelho, destaca-se a 3'CYC na CS1. Legenda: VDEN-1 = *Vírus do dengue sorotipo 1*; VDEN-2 = *Vírus do dengue sorotipo 2*; VEMV = *Vírus da encefalite Murray Valley*, VEJ = *Vírus da encefalite japonesa*; VFA = *Vírus da febre amarela*; VNO = *Vírus do Nilo Ocidental*. Fonte: Adaptado de Markoff, 2003.

Não há repetições da CS2 (RCS2) no genoma do VFA. Ao invés, existem três repetições em série próximas da extremidade 5' da 3'RNC. A primeira dessas repetições apresenta 40 nt no genoma do VFA cepa 17D e inicia no sentido 5'→3' do terceiro códon de parada encontrado no terminal 3' de sua ORF (Figura 7). Muito embora a segunda repetição seja quatro nucleotídeos maior que a primeira, apenas cinco nucleotídeos diferem entre os dois segmentos. Similarmente, o terceiro segmento difere do primeiro em quatro nucleotídeos (Markoff, 2003).

Nos vírus do sorogrupo da encefalite japonesa existe um par adicional de repetições no sentido 3'→5' da CS2 e RCS2 na 3'RNC, designadas de CS3 e RCS3. Esses segmentos apresentam 28 nt de comprimento. O segmento RCS3 situa-se a partir do nucleotídeo 153 (referente ao terminal 5' da 3'RNC), após o códon de parada da ORF do VEMV. São observados apenas dois nucleotídeos diferentes entre as CS3 e RCS3 do VEMV (Markoff, 2003) (Figura 8). Segundo Proutski e colaboradores (1997), as 3'RNCs dos flavivírus podem formar uma estrutura secundária, descrevendo três regiões dentro dessa estrutura caracterizada como elementos centrais: (i) A região I consiste de sequência nucleotídica, que apresenta grande variabilidade nucleotídica, localizada no sentido 3'→5' da CS1, CS2 e RCS2 no genoma dos flavivírus transmitidos por mosquitos; corresponde, em todos os flavivírus, a um grande grampo com uma grande ramificação lateral haste-alça, ou, em alguns casos, uma alça no lado 5' do grampo principal; (ii) A região II inclui a CS2 e RCS2 e apresenta uma estrutura secundária distinta entre

grupos de flavivírus; (iii) A região III inclui a CS1 e a 3'SL (Proutski et al., 1997; Markoff, 2003).

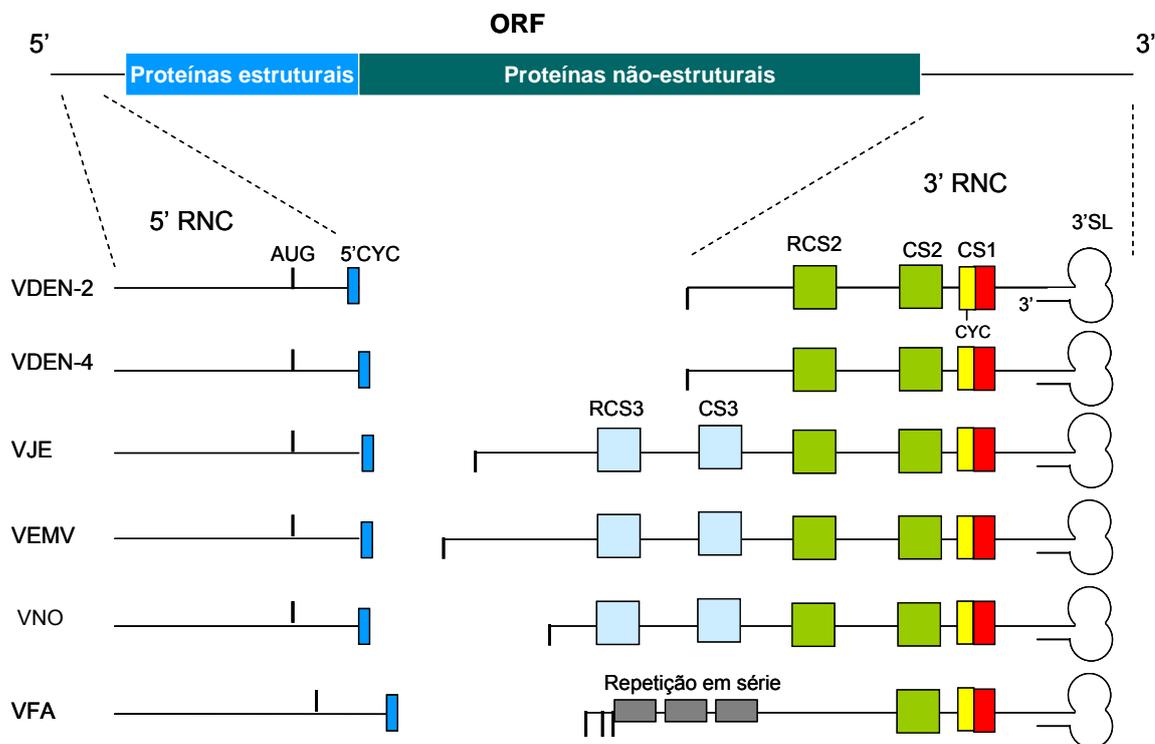


Figura 8 – Características das regiões não codificantes (RNCs) dos flavivírus transmitidos por mosquitos, destacando as seqüências conservadas na 3'RNC (CS1, CS2, RCS2, C3, RCS3), as seqüências envolvidas na cicização do RNA – 5'CYC, 3'CYC – e as repetições em série presentes no genoma do VFA. Linhas horizontais representam as seqüências nucleotídicas alinhadas das 5'RNC (lado esquerdo) e 3'RNC (lado direito) do vírus VDEN-2, VDEN-4, VEJ, VEMV, VNO e VFA. Traços verticais voltados para cima representam o códon de iniciação, AUG, para a tradução da ORF, enquanto que os traços verticais voltados para baixo representam os códons de parada. A estrutura 3' haste-alça (3'SL) está representada no terminal 3'. Legenda: VDEN-2 = *Vírus do dengue sorotipo 2*; VDEN-4 = *Vírus do dengue sorotipo 4*; VEMV = *Vírus da encefalite Murray Valley*; VEJ = *Vírus da encefalite japonesa*; VFA = *Vírus da febre amarela*; VNO = *Vírus do Nilo Ocidental*. Fonte: Adaptado de Markoff, 2003.

1.6.3. Ciclização do RNA viral

A primeira sequência genômica completa de um flavivírus foi publicada em 1985, por Rice e colaboradores. Esses autores notaram que as extremidades 5'- e 3'-terminal do genoma do VFA cepa 17D eram parcialmente homólogas a uma região complementar no VNO. Em ambos os genomas foi observado que; (i) dois nucleotídeos 5'-terminal (5'-AG) eram complementares a dois nucleotídeos 3'-terminal (3'UC); e (ii) a presença de uma pequena região de cinco nucleotídeos idêntica entre os segmentos proximais do 3'-terminal do RNA sentido positivo e o 3'-terminal RNA sentido negativo de cada um dos vírus: as sequências 3' UGUGU na 3'RNC e a sequência 5'-ACACA na 5'RNC (Figura 9).

VFA	(+) 5' AGJAAAUCCUGUGGC . . . A ACACAAAACCA CU 3'
	(-) 3' UCAUUUAGG ACACAC G . . . UUGUGUUUUGGUGA 5'
VWN	(+) 5' AGJAGUUCGCCUGUGUG . . . A ACACAGGAUCU 3'
	(-) 3' UCAUCAAGCGG ACACAC . . . UUGUGUCCUAGA 5'
VEMV	(+) 5' AGACGUUCAUCUGCGUGAGC . . . GAGAAGACCACAGGU CU 3'
	(-) 3' UCUGCAAGUAGACGCACUCG . . . CUCUUCUGGUGUCCUGA 5'
VDEN-3	(+) 5' AGJUGUUAGUCUACGUG . . . AUCAACAGGU UCU 3'
	(-) 3' UCAACAAUCAGAUGCAC . . . UAGUUGUCCA AGA 5'
VKUN	(+) 5' AGJAGUUUAUCUGUGUGAAC . . . GGUUCGAGA ACACAGGAUCU 3'
	(-) 3' UCAUCAAAUAG ACACA CUUG . . . CCAAGCUCUUGUGUCCUAGA 5'
VJE	(+) 5' AGAAGUUUAUGUGUGUGAAC . . . GAGGAAGA ACACAGGAUCU 3'
	(-) 3' UCUUCAAAUAG ACACA CUUG . . . CUCCUUCUUGUGUCCUAGA 5'

Figura 9 – Sequências nucleotídicas dos terminais 5'- e 3'- de alguns flavivírus transmitidos por mosquitos, mostrando alinhamento entre o genoma de RNA sentido positivo e o RNA complementar sentido negativo. Dinucleotídeos conservados nas extremidades 5' e 3' são mostrados nas caixas em vermelho. A sequência semiconservada pentanucleotídica, ACACA, é mostrada em destaque (negrito) nas extremidades 3'-terminal dos RNA sentido positivo e negativo. Fonte: Markoff, 2003.

A análise do genoma de flavivírus mostrou que os pares dinucleotídicos complementares das extremidades do RNA viral (5'-AG/UC-3') são altamente conservados entre os flavivírus transmitidos por mosquitos e por carrapatos, mas não são conservados para o flavivírus de insetos, tal como o VCFA (5'-GU e GC-3'). Por outro lado, a sequência pentanucleotídica 5'-ACACA-3' e sua sequência complementar não é conservada em posições análogas nas 3'RNC e 5'RNC nos genomas de diversos flavivírus (Markoff, 2003). Além de participar da formação da ciclização do RNA viral, a sequência 5'-ACACA também pode atuar como sítio de reconhecimento para a polimerase viral que pode ser usada durante a síntese do RNA sentido positivo ou negativo (Rice et al., 1985).

A estrutura secundária do RNA dos flavivírus foi inicialmente descrita por Hahn e colaboradores (1987), observando a complementaridade, entre sequências conservadas da 5'RNC e a região conservada que precede o 3'SL, designadas de sequência de ciclização (CYC). Com isso, foi sugerido que o pareamento entre as 5'CYC e 3'CYC pode, em teoria, permitir a formação da ciclização do RNA, que ocorre provavelmente durante as fases de sínteses do RNA viral (Figura 10). Essa conformação supostamente permitiria a polimerase viral se ligar em ambas as extremidades do RNA molde simultaneamente, gerando cópias do genoma viral (Hahn et al., 1987). No entanto, a função dessa estrutura secundária dos RNA genômicos dos flavivírus durante a replicação viral só foi possível ser comprovada com o estudo de Khromykh & Westaway (1997).

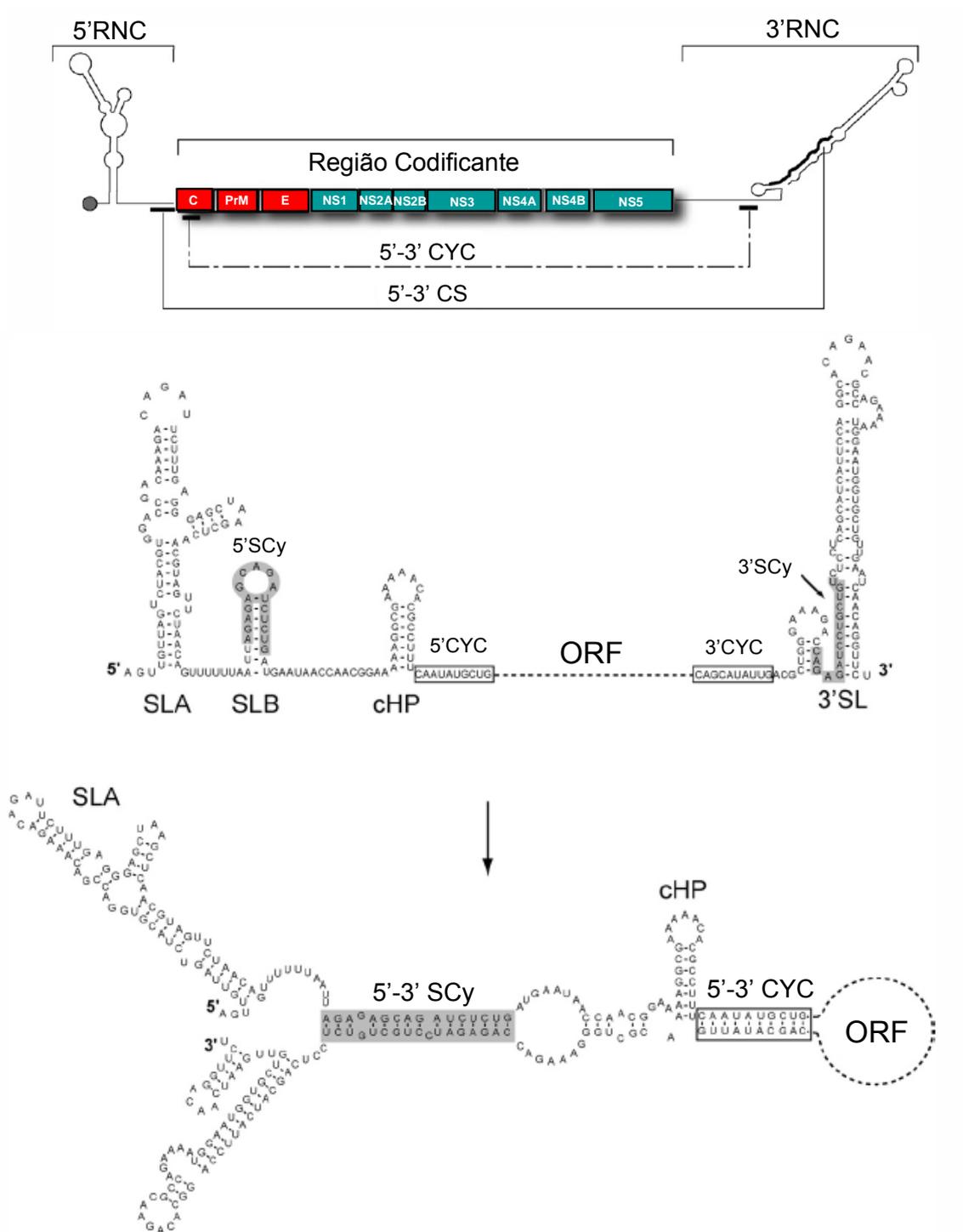


Figura 10 – Estrutura de ciclização do RNA dos flavivírus, mostrando as sequências conservadas de ciclização (5'CYC, 5'SCy, 3'CYC e 3'SCy) e estrutura secundária das 5'RNC (SLA, SLB e cHP) e 3'RNC (3' SL). Legenda: SLA=estrutura principal da 5'RNC, SLB=segunda haste-alça da 5'RNC; cHP=*Harpain* da região do capsídeo. Fonte: Adaptado de Villordo & Gamarnik, 2009.

A 3'CYC é comum a todos os flavivírus e apresenta localização variável entre as posições 99 a 112 em relação ao terminal 3' do genoma (Quadro 3) e está incluída dentro da CS1 (Figura 8) que também é semiconservada entre os genomas dos flavivírus transmitidos por mosquitos (Mandl et al., 1993). Estudos utilizando programas computacionais de análise de genoma demonstraram que o pareamento de bases ocorre entre a 5'CYC (5'-UCAAUAUG-3') e a 3'CYC é através de ligações covalentes e esse pareamento é essencial para a replicação viral (Westaway et al., 2003). Para o genoma dos flavivírus transmitidos por carrapatos, duas 5'CYC em potencial foram identificadas, sendo designadas C1 (5'-GGAGAACAAGA-3') e C2 (5'-GGGGCGGUCCC-3'), mas essas são distintas das 5'CS dos flavivírus transmitidos por mosquitos (Quadro 3) (Khromykh et al., 2001b; Mandl et al., 1993).

Mathews e colaboradores (1999) analisando a sequência dos VKUN e outros flavivírus transmitidos por mosquitos, além dos VTBE e VCFA, pelo programa MFOLD, revelou que o genoma de todos os flavivírus contém complementaridade, com polaridade *trans*, entre as 5'CYC e 3'CYC. Essa estrutura pode ser ainda mais estabilizada pela presença de nucleotídeos adicionais não-conservados de um segmento rico em pirimidinas, localizado no sentido 3'→5' da 5'CYC, que são complementares a uma região rica em purinas, situadas próximas a 3'CYC, totalizando 11 ou 12 pareamento de bases que possibilita a ciclização do RNA dos VFA, VEMV, VNO e VDEN-2, formada por uma estabilidade termodinâmica de -9,1 kcal (para o VEMV) a -2,3 kcal (para o VFA). Essa energia livre é suficiente para a formação da estrutura

secundária do RNA viral, tendo por base a ciclização dos flavivírus que requer condições fisiológicas mais baixas (Hsu et al., 1973).

Quadro 3 - Localização das posições das sequências de ciclização (CS) no genoma de flavivírus (Markoff, 2003).

VÍRUS	CÓDON DE INICIAÇÃO*	LOCALIZAÇÃO DA CS**	GENBANK	
Flavivírus Transmitido por mosquitos	VKUN	97	5' CYC (137) UCAAUAUG 3' CYC (-104) AGUUAUAC	D00246, L24511
	VEJ	96	5' CYC (136) UCAAUAUG 3' CYC (-111) AGUUAUAC	M10370
	VEMV	96	5' CYC (136) UCAAUAUG 3' CYC (-111) AGUUAUAC	NC000943
	VNO	97	5' CYC (137) UCAAUAUG 3' CYC (-105) AGUUAUAC	M12294
	VDEN-1	81	5' CYC (118) UCAAUAUG 3' CYC (-103) AGUUAUAC	M87512
	VDEN-2	97	5' CYC (134) UCAAUAUG 3' CYC (-103) AGUUAUAC	M20558
	VDEN-3	94	5' CYC (132) UCAAUAUG 3' CYC (-103) AGUUAUAC	M93130
	VDEN-4	102	5' CYC (136) UCAAUAUG 3' CYC (-99) AGUUAUAC	M14931
	VFA	119	5' CYC (156) UCAAUAUG 3' CYC (-112) AGUUAUAC	NC002031
	VCFA	114	5' (169) CCCCGUUCUGG 3'(-134) GGGGCA-GGUC	M91671
114		5' CS "D" (149) GCCAGGG 3' CS "C" (-471) CGGUCCC	M91671	
Flavivírus transmitidos por carrapatos	VTBE	133	5' C2 (164) GGGGC^{GG}UCCC 3' C2' (-193) CCCC _G AGGGG	U27495
	VPOW	112	5' C2 (136) GGGGGC^{GG}UCC 3' C2' (-189) CCCCCG-AGG	L06436
	VLI	130	5' C2 (160) GGGGGC^{GG}UCC 3' C2' (-193) CCCCC _G -AGG	Y07863
	VTBE	133	5' C1 (115) GGAGAACAAGA 3' C1' (-81) CCUCUUGUUCU	U27495
	VPOW	112	5' C1 (88) GGAGAACAAGA 3' C1' (-81) CCUCUUGUUCU	L06436
	VLI	130	5' C1 (112) GGAGAACAAGA 3' C1' (-81) CCUCUUGUUCU	Y07863

* – Número da posição do primeiro nucleotídeo do códon de iniciação AUG;

** – 5' e 3' sequências de ciclização (CYC). Em negrito, o primeiro nucleotídeo de cada CYC. Dentro dos parênteses está indicada a posição do primeiro nucleotídeo de cada CYC em relação aos terminais 5' e 3' (sinal negativo) de cada genoma.

Alvarez e colaboradores (2005b) mostraram uma segunda sequência de ciclização na 5'RNC (5'SCy) que provavelmente participa da estrutura de secundária do RNA viral, que está localizada no sentido 3'→5' do códon de iniciação AUG, na segunda haste-alça. Essa 5' SCy é complementar a uma região presente dentro da 3'SL, a 3'SCy (Figura 10).

Baseado na localização conservada das sequências de ciclização, foi proposto que a estrutura secundária do RNA viral apresenta duas funções importantes no mecanismo de síntese do RNA: (i) promotor do complexo da polimerase próximo do sítio iniciação 3'; (ii) e promover a conformação correta do 3'RNC durante o processo de iniciação (Villordo & Gamarnik, 2009)

1.7. ESTRUTURAS DAS PROTEÍNAS

1.7.1. Proteína do Capsídio (C)

A proteína do capsídio é altamente básica, com massa molecular de aproximadamente 11 KDa, sendo responsável pela formação do nucleocapsídio junto com o RNA viral. Resíduos carregados positivamente estão inseridos nas terminações -N e -C, separados por uma região hidrofóbica que media a associação da membrana (Lo et al., 2003). A proteína C recém sintetizada apresenta uma âncora hidrofílica C-terminal (AnchC) que serve como peptídeo-sinal para a translocação da PrM. Esse domínio hidrofílico é

clivado para a maturação da proteína C pela protease serina viral (Lobigs, 1993). A proteína C dobra-se dentro de um dímero compacto com cada monômero contendo quatro α -hélices. No entanto, não está claro como o dímero da proteína C é organizado dentro do nucleocapsídeo, mas a interação com o RNA pode induzir dímeros de proteínas C isolados à montagem junto a partículas ainda não formadas (Kiermayr et al., 2004).

1.7.2. Glicoproteína de membrana (M)

A proteína precursora da glicoproteína M (PrM) com 26 KDa é translocada para dentro do RE pelo domínio AnchC. A clivagem desse peptídeo sinal ocorre tardiamente até que a proteína serina viral seja clivada. Este processo de coordenação da clivagem da AnchC/Pr serve para atrasar o processamento das proteínas estruturais e a produção viral até que os níveis da protease serina viral estejam suficientemente elevados (Chambers *et al.*, 1990; Nowak & Wengler, 1987). A região N-terminal da PrM contém um dos três sítios de glicosilação e seis resíduos de cisteína conservados (Chiu et al., 2005), nos quais há a ligação dissulfeto (Nowak et al., 1989).

A principal função da PrM é prevenir que a proteína E sofra rearranjo por catálise ácida durante o trânsito através da via secretória. A conversão da partícula viral imatura para o vírus maduro ocorre na via secretória e coincide com a clivagem da PrM nos fragmentos Pr e M pela protease furina encontrada no CG ou por enzimas relacionadas (Stadler et al., 1997; Elshuber et al., 2003).

A resposta imune humoral contra a PrM/M inclui anticorpos neutralizantes e conferem proteção em camundongos. Este fato sugere que a clivagem entre Pr/M pode ser incompleta em uma subpopulação de vírus maduros e/ou que epítomos protetores podem estar localizados no ectodomínio da proteína M nas partículas virais maduras (Bray & Lai, 1991)

1.7.3. Proteína do Envelope (E)

A proteína E (53 KDa) é a maior proteína de superfície dos flavivírus e media a ligação com o receptor celular e a fusão intermembrana. É sintetizada como uma proteína de membrana tipo II, contendo 12 resíduos conservados de *Cis*, que formam as ligações dissulfetos. Para alguns flavivírus, a proteína E é N-glicosilada (Heinz & Allison, 2003; Lindenbach et al., 2007).

Vem sido proposto que a proteína E dos flavivírus está envolvida em inúmeras atividades biológicas, incluindo adsorção do vírus ao receptor celular, fusão de membrana (Heinz & Allison, 2003; Heinz et al., 2004) e montagem da partícula viral (Li et al., 2006; Limjindaporn et al., 2009).

A proteína E é composta de duas subunidades, cada uma constituída por três domínios: domínios I, II e III. O contato entre os monômeros que formam o dímero não é contínuo ao longo de todo o comprimento da molécula, resultando em um sulco ao longo do eixo do dímero. O contato no centro envolve sítios hidrofóbicos pertencentes apenas ao domínio II, enquanto que o contato distal é monopolar e composto por resíduos dos domínios I e III

de uma subunidade e a ponta do domínio II da outra subunidade (Heinz & Allison, 2003)

Com exceção de duas pequenas cadeias α -hélices no domínio II, as estruturas β -pregueadas constituem a estrutura secundária predominante em toda a molécula. O domínio I é formado por uma estrutura β -pregueada com o seu eixo orientado paralelamente à membrana viral. Esse domínio N-terminal contém duas ligações dissulfetos e exibe na superfície uma única cadeia de carboidrato (Heinz & Allison, 2003).

O domínio II é projetado ao longo da superfície viral por entre as regiões transmembranas da subunidade homodímero (Lindenbach et al., 2007). O domínio II é constituído por duas alças descontínuas estendidas a partir do domínio I. Uma dessas alças é estabilizada por três ligações dissulfetos e constitui uma estrutura que interage com o sulco hidrofóbico constituído por resíduos adjacentes da junção dos domínios I e III da segunda subunidade. Essa alça funciona como peptídeo de fusão (Heinz & Allison, 2003).

O peptídeo de fusão, que media a inserção dentro da membrana da célula-alvo é localizado na ponta distal do domínio II com relação à região transmembrana (Lindenbach et al., 2007), correspondendo aos aminoácidos 98 a 110 referente à proteína E do VTBE. Essa sequência é altamente conservada entre os flavivírus com exceção da posição 104, que é uma Histidina (His) nos flavivírus transmitidos por carrapatos e uma Glicina (Gli) para os vírus transmitidos por mosquitos. Ademais, outras variações afetam o aminoácido 107, que é uma Leucina (Leu) para a maioria dos flavivírus, mas que pode ser

a Fenilalanina (Fen) no VPOW (Kuno et al., 2001; Ebel et al., 2001), ou uma *Cis* em cepas do VEJ (Nitayaphan et al., 1990) e VDEN-2 (Blok et al., 1989).

O domínio III contém a porção C-terminal da proteína e uma estrutura em prega similar à observada nas imunoglobulinas. Uma única ligação dissulfeto a conecta ao domínio I através de 15 aa (Heinz & Allison, 2003). Entre os ectodomínios da proteína E e a membrana existe uma pequena região composta por duas α -hélices que atravessa a membrana e que atua como âncora para a síntese da proteína E (Lindenbach et al., 2007). No domínio III da proteína E do VEMV (Hurrelbrink & McMinn, 2001) e VEJ (Lee & Lobigs, 2000), bem como de outros integrantes do complexo da Encefalite Japonesa (Gubler et al., 2007), é reconhecido à sequência tripeptídica RGD (*Arg-Gly-Asp*) remanescente de um motivo integrina-ligante. Esse motivo parece estar envolvido na ligação ao receptor celular e é o maior alvo de anticorpos neutralizantes (Heinz & Allison, 2003; Crill & Roehrig, 2001). Sequências similares ao motivo RGD também ocorrem no genoma dos vírus VFA (Van der Most et al., 1999), VDEN (Chen et al., 1997b) e VTBE (Bhardwaj et al., 2001).

A proteína E contém antígenos conservados grupos-específicos, que podem ser identificados pelos testes de IH e TN (Vasconcelos, 2003), determinada, principalmente pela propriedade de hemaglutinina (Casals & Brown, 1954). O teste de neutralização de redução em placa (PRNT), bem como o TN é mais específico e utilizado para identificar espécies virais (Calisher et al., 1989). Quanto à indução da resposta imune, 19 epítomos já foram caracterizados, 16 dos quais estão contidos no domínio II. A exposição

da maioria dos epítomos é dependente da conformação da proteína (Roehrig, 2003). Seis regiões de neutralização foram mapeadas, utilizando anticorpos monoclonais, em todos os três domínios da proteína E do VTBE, muito embora as regiões presentes no domínio III sejam fortemente bloqueadores da adsorção celular (Crill & Roehrig, 2001).

1.7.4. Proteína não-estrutural NS1

A glicoproteína NS1 (46 KDa) é translocada para dentro do RE durante a síntese e clivagem da proteína E por uma peptidase do hospedeiro (Falgout et al., 1989; Falgout & Markoff, 1995). A NS1 é formada por homodímeros e possui um ou dois sítios de glicosilação e 12 resíduos de Cis conservados que formam as pontes dissulfetos. Não se sabe ao certo como a NS1 desempenha o seu papel na síntese do RNA, entretanto, é nesta proteína que estão localizados os sítios da síntese do mesmo (Mackenzie et al., 1996). Ademais, mutações nos sítios N-ligantes da glicosilação ocasionam efeitos drásticos na síntese do RNA (Muylaert et al., 1996).

A NS1 pode ser facilmente encontrada dentro das células infectadas, bem como na superfície celular e é lentamente secretada pelas células de mamíferos. Uma vez que a NS1 é altamente hidrofílica e perde o domínio transmembrana, a natureza de sua associação com a membrana ainda não é compreendida. A função da forma extracelular ainda não foi esclarecida, no entanto uma forte resposta humoral é ativada contra essa proteína e os anticorpos contra a forma da superfície celular pode induzir a lise

mediada por complemento, sendo por isso designada como antígeno fixador do complemento solúvel (Lindenbach & Rice, 2003), sendo detectada pelo teste de FC (Vasconcelos, 2003). Ademais, a forma extracelular foi caracterizada como o antígeno fixador de complemento presente no soro e tecidos de animais infectados, tendo o ápice de produção durante a fase aguda da infecção (Alcon et al., 2002). Anticorpos anti-NS1 específicos podem produzir a fosforilação da tirosina em proteína de células infectadas pelo VDEN, o que sugere que a NS1 pode mimetizar importantes moléculas humanas (Chang et al., 2002).

A NS1 vem sendo alvo de ensaios sorológicos (ELISA) desenvolvidos para sorotipagem a nível de espécie específica, subcomplexo específica, bem como complexo-específico (Hall et al., 1995; Shu et al., 2004).

1.7.5. Proteína não-estrutural NS2A

A NS2A é uma proteína pequena hidrofóbica com aproximadamente 22 KDa. A extremidade N-terminal é gerada pela clivagem NS1/NS2A por uma enzima do RE não identificada (Falgout & Markoff, 1995). A clivagem citosólica da junção NS2A/NS2B pela protease serina viral, indica uma topologia transmembrana dessa proteína. Estudos sugerem que a NS2A esteja envolvida no processo de montagem da partícula viral (Liu et al., 2003). Sítios subcelulares da NS2A interagem com a polimerase das proteínas NS3 e NS5, bem como com a 3'RNC do genoma (Mackenzie et al., 1998). Portanto, a NS2A está possivelmente associada com a coordenação da mudança do RNA

empacotado para o RNA a ser replicado (Khromykh et al., 2001a). Ademais, a NS2A pode agir como antagonista de interferon pela inibição de sua sinalização (Muñoz-Jordán et al., 2003).

1.7.6. Proteína não-estrutural NS2B

A NS2B é uma proteína associada à membrana (12 KDa), que forma um complexo estável com a NS3, agindo como um cofator NS2B-NS3 protease serina (Falgout et al., 1991). A atividade de cofator encontra-se numa região hidrofílica central conservada com 40 resíduos (Chambers *et al.*, 1993; Falgout et al., 1993; Leung et al., 2001), flanqueada por regiões hidrofóbicas que mediam a associação com membranas (Clum et al., 1997).

Em estudos com o VKUN, um sítio de clivagem inesperado foi descoberto. Um peptídeo sinal precede o sítio de clivagem da NS2B, mas não da NS2A, entretanto ambos são clivados depois da sequência conservada Val-X-Ala, onde X pode ser qualquer aminoácido, e que revela uma sequência octapeptídica relativamente conservada. Essa clivagem não-usual pode ser mediada por uma peptidase celular com uma especificidade bem definida (Speight et al., 1988). Ademais, durante a expressão da proteína do vírus amarelo 17D (VFA vacinal), a clivagem do peptídeo sinal precedente de NS4B é dependente da NS2B-NS3 protease no sítio NS2A/2K (Lin et al., 1993).

1.7.7. Proteína não-estrutural NS3

A NS3 é uma proteína multifuncional (70 KDa), exercendo várias atividades durante o processamento da poliproteína e da síntese do RNA. O terço N-terminal dessa proteína é o domínio catalítico do complexo NS2A-NS3 protease serina, que está localizada dentro dos primeiros 180 aa, relacionado à tríade *His-Asp-Ser* nas posições aminoacídicas 51, 75 e 137, respectivamente (com base na sequência da NS3 do VKUN), como sítios catalíticos (Li et al., 1999). As clivagens das junções NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4B/NS5, C/AnchC e NS4A/2k ocorre num sítio interno do complexo NS2B-NS3, no qual a protease serina cliva preferencialmente após resíduos básicos adjacentes (Lindenbach et al., 2007).

A região C-terminal da NS3 apresenta significativa homologia com a RNA helicase do supergrupo 2. Apesar da atividade helicase ter sido demonstrada ou sugerida para muitos vírus de RNA sentido positivo, seu papel preciso na síntese do RNA ainda permanece desconhecido. Possíveis funções incluem ligações com regiões da estrutura secundária do RNA envolvidas no modelo de reconhecimento; aumento da atividade da polimerase pela eliminação da estrutura secundária, resultando em formas duplex durante o processo de síntese; e ativação com translocação para a remoção ou mudanças de ligações de proteínas para o RNA viral. Produtos clivados contidos nessa região acreditam-se formar componentes enzimáticos da RNA polimerase (Lindenbach et al., 2007).

A estrutura da RNA helicase apresenta três subdomínios, dois dos quais são conservados em todos os membros da família helicase e está envolvido na hidrólise de nucleosídeos trifosfato e um subdomínio C-terminal, que pode estar envolvido no reconhecimento de proteínas e RNA vírus específicos (Lindenbach et al., 2007).

A atividade RNA nucleosídeo trifosfatase (NTPase) e a atividade de extensão do RNA têm sido demonstradas para muitos flavivírus (Warrener et al., 1993), do mesmo modo que a atividade RNA trifosfatase (RTPase), envolvida na desfosforilação da terminação 5' do RNA genômico antes da adição do *cap* tipo II. Ademais, a NS3 se liga a 3'SL em associação a NS5, tendo sua atividade RTPase aumentada na presença da NS5 (Chen et al., 1997a; Cui et al., 1998).

1.7.8. Proteína não-estrutural NS4A

A NS4A é uma proteína hidrofóbica (16 KDa). Seu papel na síntese do RNA é sustentado pela co-localização dessa proteína com complexos de síntese. Ademais, esta proteína está envolvida no rearranjo da membrana (Roosendaal et al., 2006).

1.7.9. Proteína não-estrutural NS4B

A NS4B (27 KDa) é uma proteína hidrofóbica, resultante da clivagem mediada pela NS2A-NS3 protease serina entre as junções 2K/NS4B e

NS4B/NS5. A associação da membrana com a NS4B pode ser independente do peptídeo sinal 2K, indicando que esta proteína seja uma proteína de membrana politópica (Miller et al., 2006). A NS4B co-localiza-se juntamente com a NS3 e com a fita dupla de RNA na estrutura de membrana derivada do RE, que é presumidamente o sítio de síntese do RNA. De modo similar, a NS4A, a NS4B pode bloquear a sinalização do tipo I de interferon (Muñoz-Jordán et al., 2005).

1.7.10. Proteína não-estrutural NS5

A NS5 é uma proteína multifuncional (103 KDa) altamente conservada tanto em tamanho como na sequência nucleotídica entre os flavivírus; apresenta atividades metiltransferase (MTase) e de RNA polimerase dependente de RNA (RdRP). A região C-terminal da NS5 exibe homologia com a MTase dependente de S-adenosil-metionina (SAM), sugerindo que essa proteína esteja envolvida na modificação do *cap* 5' (Koonin, 1993). A atividade de polimerase da NS5 tem sido confirmada com NS5 recombinante (Ackermann & Padmanabhan, 2001; Guyatt et al., 2001). O complexo NS3-NS5 pode estimular as atividades NTPase e RTPase da NS3 (Cui et al., 1998) e se liga ao 3'SL do RNA genômico (Chen et al., 1997a). A NS5, especialmente a forma fosforilada, é encontrada no núcleo das células infectadas, podendo estar envolvida na ativação e secreção da interleucina-8 (IL-8), citocina envolvida com o recrutamento de células inflamatórias para o sítio de infecção, o que agrava a infecção viral (Lindenbach et al., 2005).

1.8. FILOGENIA E EVOLUÇÃO DOS FLAVIVÍRUS

Diversos estudos filogenéticos têm possibilitado uma melhor compreensão sobre a natureza evolutiva dos flavivírus, bem como colaborado para o entendimento da epidemiologia molecular e dispersão global desses vírus (Kuno et al., 1998; Gubler et al., 2007).

A comparação de sequências do gene NS5 de 71 flavivírus permitiu avaliar a relação filogenética entre esses vírus com base na divergência genética (Δ div). Um grupo é designado com base num valor de *bootstrap* acima de 95% e associação comum entre hospedeiro vertebrado e vetor. Um clado é definido como um grupo de flavivírus que apresenta um valor de *bootstrap* a partir de 69% ou uma elevada identidade nucleotídica (Calisher et al., 1989; Zanotto et al., 1996). A espécie viral é definida como uma classe de vírus com identidade nucleotídica maior que 84%, baseada nas diferenças entre duas cepas de uma mesma espécie viral (Kuno et al., 1998; Gaunt et al., 2001). Árvores filogenéticas baseadas na sequência aminoacídica da proteína NS5 apresentaram resultados similares da análise nucleotídica com relação a subdivisões em grupos, clados e espécies virais (Kuno et al., 1998; Gubler et al., 2007).

As análises filogenéticas do gene NS5 e ORF demonstraram a associação de quatro grandes grupos de associação entre os flavivírus: os flavivírus transmitidos por mosquito, por carrapatos, os agentes zoonóticos com vetores ainda desconhecidos e o flavivírus de insetos (Figura 11) (Kuno et al., 2009).

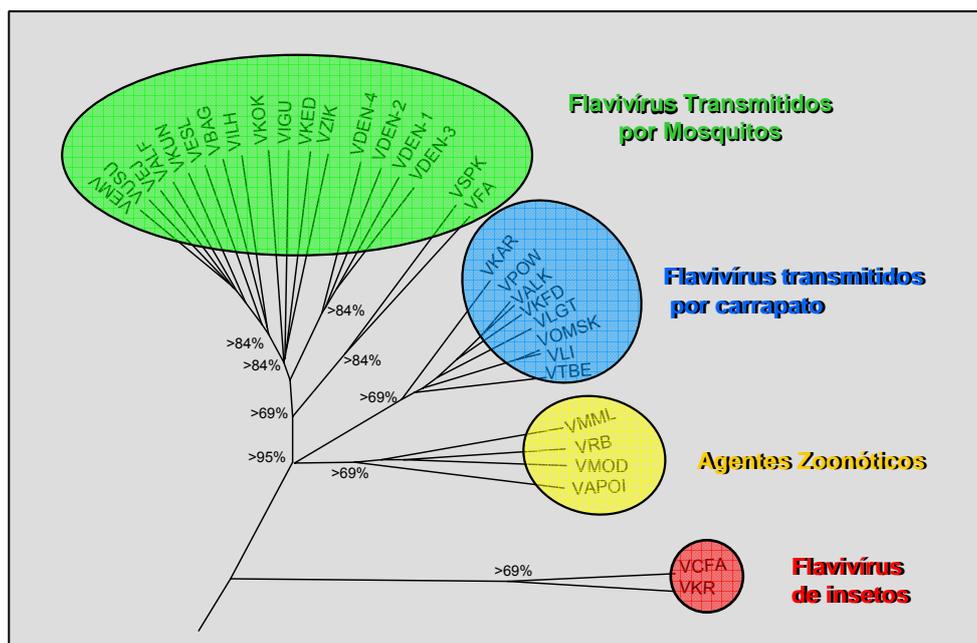


Figura 11 - Representação da árvore filogenética dos flavivírus, subdivididas nos quatro principais grupos: flavivírus transmitidos por mosquitos (verde), flavivírus transmitidos por carrapato (azul), agentes zoonóticos com vetor desconhecido (amarelo) – e os flavivírus de insetos (vermelho) Fonte: Adaptado de Kuno et al., 1998; Kuno et al., 2009.

O grupo dos flavivírus transmitidos por mosquitos inclui nove clados: (i) grupo da encefalite japonesa: VKUN, VNO, VEJ, VEMV, Cacipacoré - VCPC; (ii) o grupo representado pelos vírus Kokobera (VKOK) e Strateor; (iii) o grupo constituído pelos flavivírus VBSQ, VIGU, VNJL e Aroa (VAROA); (iv) grupo Ntaya, no qual estão incluídos os flavivírus brasileiros VROC e VILH; (v) grupo do vírus Zika (VZIK); (vi) grupo do VDEN; (vii) o grupo do VFA representada pelas cepas do VFA e o vírus Sepik (VSEP); (viii) o clado que inclui três vírus de vetor desconhecido, que perderam a capacidade de serem transmitidos por mosquitos ou que não foram completamente estudados: Entebbe bat (VENT), Sokoluk (VSOK) e Yokose (VYOK); (ix) e o clado que contém os vírus Banzi (VBAN), Bouboui (VBOU), Edge Hill (VEH), Jugra

(VJUG), Saboya (VSAB), Wesselsbron (VWESS) e Uganda S (VUGS) (Gubler et al., 2007).

Gaunt e colaboradores (2001) realizaram uma análise conjunta da sequência parcial do gene NS5 e do gene E, sugerindo a subdivisão dos flavivírus transmitidos por mosquitos, correlacionando os gêneros dos vetores *Culex* sp. e *Aedes* sp., os principais hospedeiros vertebrados (aves, roedores e primatas) e o tropismo viral nas infecções humanas (neurotrópicos e viscerotrópicos). Os flavivírus pertencentes ao grupo *Aedes* estão associados a doenças hemorrágicas e a hospedeiros mamíferos, enquanto que a maioria dos integrantes do grupo *Culex* (20 dentre os 21 vírus) é responsável por encefalites e estão associados às aves, como principal hospedeiro, e equinos.

Os flavivírus transmitidos por carrapatos encontram-se divididos em dois clados (Kuno et al., 1998; Gubler et al., 2007). Estudos mostram associação dos clados dos flavivírus transmitidos por carrapatos com possíveis hospedeiros vertebrados: roedores e aves marinhas, respectivamente, no qual o VTBE encontra-se no clado associado aos roedores (Gaunt et al., 2001).

Os flavivírus com vetor artrópode desconhecido apresentam relação filogenética estreita com seus hospedeiros. Portanto, o grupo dos flavivírus de vetor desconhecido foi subdividido em três clados, nos quais os dois primeiros agrupam os vírus associados a roedores e o terceiro clado agrupa vírus que estão associados a morcegos (Kuno et al., 1998; Kuno & Chang, 2005; Gaunt et al., 2001).

A característica antigênica também pôde ser correlacionada à filogenia baseada nas sequências dos genes E e NS5, sendo demonstrado o

mesmo tipo de agrupamento observado anteriormente para a classificação dos sorocomplexos (Kuno et al., 1998; Billoir et al., 2000; Gaunt et al., 2001; Gould et al., 2001). Esses estudos enfatizam a ligação existente entre métodos filogenéticos e antigênicos (Calisher & Gould, 2003).

Resultados similares quanto à subdivisão filogenética dos flavivírus baseado na análise da NS5 (Kuno et al., 1998; ICTV, 2005) foram obtidos através da análise das sequências do gene NS3, proteína E e genoma completo viral, demonstrando uma considerável conservação genética entre os flavivírus (Billoir et al., 2000; Charrel et al., 2001; Gaunt et al., 2001; Marin et al., 1995; Zanotto et al., 1995). Com relação a 3'RNC, foi demonstrada a mesma disposição organizacional para o complexo da febre amarela e o complexo dos flavivírus transmitidos por carrapatos. Entretanto, foi evidenciada uma maior proximidade genética dos diferentes genótipos do VFA com o vírus *Sepik* (VSEP), sugerindo uma provável origem ancestral comum desses com o grupo do VUGS (Mutebi et al., 2003).

Ao se sugerir teorias evolutivas dos flavivírus devem ser levadas em consideração as relações genéticas observáveis, a ecologia individual das espécies virais em natureza e a distribuição geográfica (Gubler et al., 2007). As análises da base da relação genética entre os flavivírus sugerem que o processo evolutivo, no que tange a segregação de grupos virais, está provavelmente associada ao acúmulo de mutações ao longo do tempo, relacionadas à adaptação viral a um dado hospedeiro, culminando a uma divisão evolutiva primária (Blok et al., 1992; Monath et al., 1996; Zanotto et al., 1996; Kuno et al., 1998; Billoir et al., 2000; Cook & Holmes, 2006). O principal

fator que age sobre o direcionamento na evolução dos flavivírus é, provavelmente, o mecanismo de escape da resposta imune. Entretanto, deve ser ressaltado que muitos flavivírus são constantemente transferidos entre vertebrados e artrópodes e podem se readaptar em cada passagem durante as alternâncias de hospedeiros (Lin et al., 2004).

Estudos mostram que a taxa de mutação ao nível de aminoácido dos vírus transmitidos por mosquitos e carrapatos é baseada no conceito de uma taxa evolutiva lenta nos vírus transmitidos por carrapatos. De fato, a estimativa da taxa de divergência aminoacídica entre os flavivírus transmitidos por mosquitos é estimado ser de 2 a 2,5 vezes mais rápida do que no grupo dos vírus transmitidos por carrapatos (Zanotto et al., 1995; Torny et al., 2005).

Os primeiros trabalhos sobre filogenia e evolução dos flavivírus basearam-se na sequências dos genes E, NS3 e NS5, e evidenciaram a subdivisão desses agentes virais em dois grandes grupos: os flavivírus transmitidos por mosquitos e os transmitidos por carrapatos (Blok & Gibbs, 1995; Lewis et al., 1993; Gibbs & Blok, 1989). Posteriormente, as análises filogenéticas incluíram sequências nucleotídicas de diferentes flavivírus isolados de vertebrados com vetores artrópodes desconhecidos, e revelam a presença de um terceiro grupo filogenético: o grupo dos flavivírus zoonóticos com vetores desconhecidos (Kuno et al., 1998)

De acordo com a análise do gene NS5, Kuno e colaboradores (1998) demonstraram que, primeiramente, um ancestral comum do gênero *Flavivirus*, não identificado, deu origem a dois grupos de vírus: os transmitidos por vetores e os flavivírus de vetor desconhecido. Posteriormente, os flavivírus

transmitidos por vetores foram subdivididos em flavivírus transmitidos por mosquitos e flavivírus transmitidos por carrapatos. Gould e colaboradores (2003a) complementaram essa teoria, sugerindo que os flavivírus transmitidos por vetores originaram-se a partir dos agentes zoonóticos, que se adaptaram aos artrópodes.

Alguns dados suportam a hipótese de que os vírus transmitidos por carrapatos são mais primitivos do que os flavivírus transmitidos por mosquitos (Kuno & Chang, 2005), tais como os repetidos isolamentos dos flavivírus transmitidos por mosquitos, os vírus Koutango, Saboya, VNO e VFA, em carrapatos (Monath et al., 1996; Billoir et al., 2000), além do fato observado de que alguns vírus transmitidos por carrapatos podem se replicar em mosquitos (Lawrie et al., 2004a; 2004b).

Zanotto e colaboradores (1996), analisando o gene E, revelaram importantes diferenças no modo de evolução dos dois grupos de flavivírus transmitidos por vetores. Árvores filogenéticas que comparam flavivírus transmitidos por mosquitos e por carrapatos mostram uma topologia assimétrica entre os dois grupos. Os flavivírus transmitidos por carrapatos segregam-se continuamente através do tempo, contrastando com a estrutura de ramificação mais balanceada dos flavivírus transmitidos por mosquitos, em que, num período relativamente longo de tempo, observou-se a intensa ramificação, seguido pela extinção de linhagens e conseqüente formação de subclados que passaram a evoluir separadamente, suportando o modelo evolutivo de "*Boom and Bust*".

Por outro lado, o estudo evolutivo realizado por Cook & Holmes (2006), utilizando diversos métodos de alinhamento de sequências aminoacídicas (Clustal W, T-Coffee e Muscle) e análise filogenética pelo método de máxima verossimilhança (MV), demonstraram que as sequências do gene da proteína NS3 e da ORF completas indicam uma segregação precoce dos flavivírus transmitidos por mosquitos, e em separado dos vírus transmitidos por carrapatos e de vetores desconhecidos, que, por sua vez, evoluíram em associação. Anteriormente, mas seguindo a mesma linha de raciocínio, Billor e colaboradores (2000), com base nas análises filogenéticas da NS3, NS5 e ORF completas, propuseram que a posição dos clados dos agentes zoonóticos com vetores desconhecidos, flavivírus transmitidos por mosquitos e por carrapatos pode diferir de acordo com o gene que é usado na filogenia.

Numa análise evolutiva envolvendo os três gêneros da família *Flaviviridae*, observou-se que apenas o gênero *Flavivirus* contém arbovírus e que alguns dos vírus mais divergentes desse gênero não são transmitidos por vetores. Este fato sugere que a condição ancestral dentro da família *Flaviviridae* é de que a transmissão independe de vetor e que a transmissão por vetores é uma recente adaptação evolutiva, o que suporta a teoria de que os agentes zoonótico com vetores desconhecidos foram os primeiros a divergirem dentro do gênero *Flavivirus* (Gould et al., 2003b). Tem sido também sugerido que esse processo evolutivo tenha ocorrido no início da Era do Gelo e que vem acompanhando o desenvolvimento da civilização humana há aproximadamente 10.000 anos (Zanotto et al., 1996; Gould et al., 2003b)

Quanto à evolução dentro do clado dos flavivírus transmitidos por mosquitos, análises filogenéticas mostraram que a segregação dos flavivírus associados a mosquitos do gênero *Aedes* é parafilética enquanto que dos flavivírus associados a mosquitos do gênero *Culex* é monofilética (Billor et al., 2000; Jenkins et al., 2001). Finalmente, Gaunt e colaboradores (2001) propuseram que os vírus transmitidos por *Culex* surgiram no Velho Mundo a partir dos flavivírus transmitidos por *Aedes*.

1.8.1. Filogenia, Evolução e Dispersão dos Flavivírus

É sabido que os flavivírus ocorrem em todos os continentes, com exceção da Antártida. Nos continentes afetados, eles podem ter sido introduzidos muitas vezes e permanecido circulantes e estáveis em condições ambientais adversas. Entretanto, as espécies virais apresentam uma distribuição geográfica distinta (Figura 12) e sua dispersão está estritamente relacionada a: (i) infecção de diferentes hospedeiros vertebrados e/ou invertebrados; (ii) ecologia do habitat natural em que cada vírus é encontrado; (iii) condições climáticas; e (iv) impacto da urbanização e do sistema de transporte. Os dados filogenéticos ilustram claramente como a adaptação de cada vírus em hospedeiros vertebrados e invertebrados específicos influenciam na evolução viral, dispersão, epidemiologia e, possivelmente, na patogênese dos flavivírus (Gould et al., 2003b).

mais espécies recentes no noroeste da Eurásia (Gubler et al., 2007). Tem sido sugerido que esse complexo do VTBE tenha dado origem aos flavivírus transmitidos por carrapatos que causam encefalomielite em ovelhas e cabras e que são propagados, principalmente, via transporte de animais domésticos (Charrel et al., 2001). A ausência de vírus transmitidos por carrapatos no Novo Mundo, com uma única exceção (VPOW), se acredita ser devido à recente introdução dos flavivírus nas Américas (Gould et al., 2003b).

Apesar dos flavivírus transmitidos por carrapatos apresentarem o mesmo ancestral, observa-se que esses vírus recentemente segregaram dando origem a dois clados estreitamente associadas à manutenção desses vírus entre carrapatos e roedores, bem como entre carrapatos e aves marinhas. Muitos desses flavivírus são transmitidos por diversas e diferentes espécies de carrapatos, muito embora eles possam estar mais comumente associados a uma determinada espécie. Isso explica, pelo menos em parte, como esses vírus são capazes de apresentar, simultaneamente, ampla distribuição geográfica no hemisfério norte (Gould et al., 2003b).

Em contrapartida, a filogenia dos flavivírus transmitidos por mosquitos mostra uma descontinuidade no padrão evolutivo, com muito menos confinamento geográfico e freqüente extinção de linhagens (Cook & Holmes, 2006; Gould et al., 2003b; Zanotto et al., 1996). Como consequência, os humanos foram expostos a uma crescente diversidade de cepas de flavivírus transmitidos por mosquitos. Esse padrão da dispersão viral pode ser derivado de um período de vida relativamente curto dos mosquitos (mensurado em dias ou semanas), bem como de outros fatores que diferenciam os mosquitos dos

carrapatos, tais como: (i) o número de repasto sanguíneo durante o período de vida e o volume de sangue ingerido; (ii) o número de hospedeiros vertebrados envolvidos; (iii) a mobilidade desses vetores, que é maior para os mosquitos em relação aos carrapatos; e (iv) a frequência da transmissão vertical (Zanotto et al., 1996; Gould et al., 2003b; Lin et al., 2004). Ademais, as atividades humanas têm demonstrado um importante papel para a dispersão dos flavivírus transmitidos por mosquitos, tal como foi visto para o vírus amarelo e a sua dispersão do Velho Mundo para o Novo Mundo, através do tráfico de escravos (Bryant et al., 2007), além da disseminação global do VDEN (Twiddy et al., 2003).

No grupo dos agentes zoonóticos com vetores artrópodes desconhecidos, tanto os flavivírus associados a roedores quanto os associados a morcegos são equivalentemente encontrados no Velho e no Novo Mundo, (Karabatsos, 1985). A presença desses flavivírus tanto no Novo Mundo quanto no Velho Mundo implica que esses vírus foram amplamente dispersos logo após a emergência dos mesmos (Gubler et al., 2007).

A emergência e a dispersão dos agentes zoonóticos têm sido claramente ditadas pelos roedores e morcegos, que estão associados a esses vírus e o sucesso de sua permanência em natureza, vem do fato que tanto os roedores quanto os morcegos se agrupam em grande número de indivíduos, promovendo excelentes condições para a transmissão entre os hospedeiros. Apesar do número limitado de isolamentos desses vírus, evidências sugerem que muitos desses vírus ocupam áreas geográficas distintas, provavelmente, devido à estreita dependência das espécies de roedores ou de morcegos.

Ademais, roedores e morcegos infectados pelos agentes zoonóticos são capturados aparentemente saudáveis, sugerindo que esses vírus infectam o hospedeiro persistentemente, com baixos níveis de replicação viral e, conseqüentemente, que o movimento global dos vírus resulta ser relativamente baixo (Gould et al., 2003b).

1.9. RECOMBINAÇÃO ENTRE FLAVIVIRUS

A suspeita da ocorrência de recombinação entre os flavivírus surgiu a partir de análises individuais de diferentes genes, onde se observa frequentes mudanças na topologia da árvore filogenética para alguns vírus (Billor et al., 2000; Gaunt et al., 2001; Gould et al., 2003b). Uma segunda evidência da recombinação é a cerca da possibilidade de que fenótipos virais podem ser alterados como resultado de recombinações, o que pode ter permitido o aumento da virulência para hospedeiros mamíferos ou mesmo resultar na alteração da especificidade do vetor artrópode (Gould et al., 2003b).

Ao se avaliar a possibilidade da ocorrência de recombinação, os flavivírus exibem alguns fatores que propiciam esse processo: (i) a dispersão global dos flavivírus e a sobreposição geográfica e ecológica de diversos desses vírus em diferentes regiões do mundo; (ii) carrapatos e mosquitos permanecem infectados por flavivírus durante todo o ciclo de vida, realizando o repasto sanguíneo, e isso possibilita a exposição a um segundo vírus; e (iii) ocorrência de epidemias e infecções crônicas dos hospedeiros vertebrados -

por exemplo, infecção crônica de humanos pelo VTBE (Gritsun et al., 2003) e infecção de aves pelo VNO ou VESL (Gubler et al., 2007).

Alguns estudos de filogenia em que genótipos são identificados e às vezes correlacionados com doenças severas sugerem que a evolução dos flavivírus seja clonal e que a divergência é gerada pelo acúmulo de mutações (Chen et al., 1990; Leitmeyer et al., 1999; Rico-Hesse, 1990). Por outro lado, as análises das árvores filogenéticas individuais utilizando genes de qualquer parte do genoma ou mesmo o genoma completo são relativamente congruentes, o que implica em sugerir que a recombinação entre os flavivírus ocorre entre vírus relativamente próximos (Gould et al., 2003b). De fato, recombinações homólogas foram demonstradas para os quatro sorotipos do VDEN (Holmes et al., 1999; Worobey et al., 1999; Tolou et al., 2001; Uzcategui et al., 2001), bem como para cepas de VNO, VFA, VEJ, VESL e complexo VTBE (Twiddy & Holmes, 2003), mostrando o importante papel da recombinação no processo evolutivo dos flavivírus.

1.10. FLAVIVÍRUS BRASILEIROS EM ESTUDO

1.10.1. Vírus Bussuquara

O VBSQ foi isolado em 1956, na floresta do antigo Instituto Agrônomo do Norte (IAN), hoje EMBRAPA, município de Belém, Estado do Pará, a partir do sangue de um macaco sentinela da espécie *Alouatta belzebul*,

tendo sido caracterizado antigenicamente como integrante do Grupo B dos flavivírus, com base nos resultados obtidos por testes sorológicos de FC e IH (Gomes & Causey, 1959; Karabatsos, 1985).

O VBSQ tem sido freqüentemente isolado na Amazônia brasileira em roedor silvestre *Proechimys guyannensis*, bem como em mosquitos culicídeos no Panamá e na Colômbia (Galindo et al., 1966; Vasconcelos et al., 1998). Apenas um relato de infecção humana foi descrita por Srihongse & Johnson (1971), no qual o paciente apresentou um quadro febril agudo, acompanhado por anorexia, cefaléia e artralgia. Até o presente não houve registro de surtos (Figueiredo, 2000).

1.10.2. Vírus Iguape

O Vírus Iguape (VIGU) foi isolado a partir de camundongo sentinela exposto à picada de artrópodes na mata Atlântica do município de Iguape, Estado de São Paulo, sudeste brasileiro, em 1979 (Coimbra et al., 1999). Antigenicamente, foi classificado como integrante do Grupo B, sendo portanto um flavivírus. Nenhum caso de infecção humana foi associado ao vírus até o momento (Figueiredo, 2000).

1.10.3. Vírus Ilhéus

O VILH foi originalmente isolado a partir do lote misto de *Aedes* sp. e *Psorophora* sp. capturados no município de Ilhéus, estado da Bahia, em

1944 (Laemmert & Hughes, 1947; Karabatsos, 1985). Em 1957, o primeiro isolamento do VILH em humanos foi obtido a partir de uma paciente que apresentava quadro febril (Causey et al., 1961). Subsequentemente, esse vírus foi isolado de humanos e de vertebrados silvestres (morcegos e aves) em diferentes países da América do Sul e Central, no entanto nenhum surto atribuído ao VILH foi registrado (Vasconcelos et al., 1998), embora haja uma considerável soroprevalência para o VILH em muitos países tropicais, o que sugere que a maioria das infecções pelo VILH é inaparente ou subclínicas (Pinheiro et al., 1986; Pinheiro, 1994; Figueiredo, 2000).

Nos ciclos biológicos de manutenção onde atuam diversas espécies de artrópodes, têm papel como vetor a espécie *Psorophora ferox*, identificada como vetor primário na Amazônia brasileira, Colômbia, Panamá e Trinidad. Outras espécies do gênero *Psorophora* e o *Aedes serratus* parecem atuar como vetores secundários. Vários grupos de animais, tais como morcegos, aves silvestres, macacos e eqüinos parecem atuar como hospedeiros vertebrados. A infecção de humanos parece ser fortuita e decorrente da exposição em áreas silvestres (Travassos da Rosa et al., 1997; Vasconcelos et al., 1998).

VILH e VROC são antigenicamente relacionados, observando-se frequentemente reação sorológica cruzada entre estes vírus nos testes de IH. A infecção por um deles só pode ser determinada com precisão utilizando o TN (Vasconcelos et al., 1998). Sendo relacionado ao VROC, o VILH apresenta potencial encefalitogênico. O espectro clínico varia desde infecções assintomáticas até quadro de encefalite. No entanto, a maioria absoluta dos

casos de VILH registrados caracterizou-se por quadros febris, com evolução média de três dias (Vasconcelos et al., 1992; 1998). A doença apresenta início súbito, com febre moderada ou elevada, por vezes acompanhada de cefaléia, calafrios, fotofobias, artralgia, mialgias e astenia, sendo a recuperação completa sem sequelas (Vasconcelos et al., 1992; Travassos da Rosa et al., 1997; Figueiredo, 2000).

Depois dos vírus VDEN e VFA, o VILH é o flavivírus com maior soroprevalência pelo teste de IH na Amazônia brasileira. Ressalta-se, no entanto, que o isolamento viral em amostras humanas é muito difícil em virtude da maioria dos casos ocorrerem como infecção assintomática ou oligoassintomática e apresentar curto período de viremia (Pinheiro et al., 1986; Vasconcelos *et al.*, 1998). A distribuição geográfica do vírus é ampla no Brasil e anticorpos específicos para o VILH têm sido detectados na população do Vale do Ribeira, estado de São Paulo (Iversson et al., 1982)

1.10.4. Vírus Rocio

O VROC foi isolado pela primeira vez em 1975 no litoral sul do estado de São Paulo, na região do Vale do Ribeira, a partir de tecidos de cerebelo e medula espinhal de caso fatal de encefalite (Lopes et al., 1978a; 1978b; Karabatsos, 1985). Sorologicamente, o VROC está relacionado aos vírus VILH, VESL, VEJ, e VEMV (Lopes et al., 1978a; 1978b; Iversson et al., 1988; Vasconcelos et al., 1998).

O VROC foi amplamente distribuído no Vale do Ribeira, tendo sido endêmico na região no período de 1973 a 1980. Desde a epidemia ocorrida entre 1975-1976 (Tiriba et al., 1976; Lopes et al., 1978a; 1978b), vários casos fatais adicionais foram registrados em 1985, caracterizados por encefalite severa e associados à baixa resposta imune (Iversson, 1988; Iversson et al., 1989; Figueiredo, 2000). Fora da área epidêmica, cinco casos de infecção pelo VROC foram diagnosticados no estado do Paraná, numa região próxima ao Vale do Ribeira (Iversson, 1988) e oito casos com sorologia positiva para o VROC foram diagnosticados em quatro municípios do estado da Bahia (Straatmann et al., 1997).

Os principais sintomas gerais associados à infecção pelo VROC são febre elevada de início súbito, cefaléia, anorexia, náuseas, vômito e mialgia. O quadro encefalítico aparece tardiamente, acompanhado por convulsões, distúrbios dos reflexos superficiais, flacidez motora, irritação meníngea e síndrome cerebelar. Alguns pacientes apresentam distensão abdominal e retenção urinária. A infecção pode causar sérias seqüelas neurológicas tais como distúrbios visuais, olfatórios e auditivos, falta de coordenação motora, distúrbio do equilíbrio e da marcha, inconsciência e graves apresentaram seqüelas severas (Figueiredo, 2000).

O ciclo de manutenção viral com base em raros isolamentos parece incluir aves silvestres (*Zonothrichia capensis*), roedores e o mosquito *Psorophora ferox*, no entanto, devido à falta de consistência de isolamentos e dados sorológicos, necessita ser melhor esclarecido (Lopes et al., 1978a; 1978b; Vasconcelos et al., 1998). Por outro lado, a susceptibilidade do

mosquito *Aedes scapularis* à infecção oral com o VROC, sua habilidade de transmitir o vírus experimentalmente, e a sua abundância e estrita associação com o homem na zona epidêmica, torna essa espécie suspeita de ter também atuado como vetor durante as epidemias ocorridas entre 1975-1976, embora nenhum isolamento do VROC tenha sido obtido em natureza dessa espécie de mosquito (Forattini et al., 1985).

1.11. RELEVÂNCIA

Muitos flavivírus são agentes virais, que em um contexto mundial, são considerados importantes patógenos humanos e com grande impacto em termos de saúde pública. De fato, vários flavivírus são responsáveis por considerável morbidade e mortalidade em humanos, bem como uma diversidade clínica que compreende desde quadros febris agudos, encefalites e síndromes hemorrágicas (Vasconcelos et al, 1998; Barrett, 2001).

A grande variedade de vetores hematófagos e hospedeiros vertebrados silvestres naturais, aliado às intensas atividades antrópicas com alterações importantes no ambiente natural, tem proporcionado o aparecimento de casos isolados, surtos, epidemias, bem como a emergência e reemergência de flavivírus patogênicos para seres humanos (Srihongse & Johnson, 1971; Travassos da Rosa et al, 1997; Vasconcelos et al, 1998; Lanciotti et al., 1999; Monath, 2001; Gubler, 1997).

A complexidade das interações dos flavivírus com os diferentes ecossistemas existentes, bem como o amplo espectro clínico impulsionaram vários estudos voltados para o entendimento da epidemiologia molecular desses agentes virais. Os resultados até então obtidos vêm permitindo determinar a origem evolutiva, a correlação entre os dados epidemiológicos e sorológicos com as características genéticas (Kuno et al., 1998; Gaunt et al., 2001), bem como diferenciar genotipicamente as espécies virais e, também, associar as divergências genéticas à gravidade das doenças (Gubler, 1997). Entretanto, esses estudos têm sido realizados pontualmente e com poucos flavivírus.

De fato, a identificação de pelo menos sete genótipos do vírus amarelo circulantes na África (cinco) e nas Américas (dois) e as diferenças genéticas observadas entre esses genótipos têm gerado informações, que podem explicar, por exemplo, a maior ou menor patogenicidade para o homem (Mutebi et al., 2002; Vasconcelos et al., 2004). A existência de diferentes populações genotípicas do VDEN, exibindo diferenças filogenéticas, também já foi associada à maior ou menor gravidade das manifestações clínicas durante epidemias de FHD (Gubler, 1997; King et al., 2008).

Portanto, a exemplo dos intensos estudos genéticos sobre os VDEN, VFA e VNO, desenvolver um estudo molecular amplo com os flavivírus brasileiros VBSQ, VIGU, VILH e VROC, envolvendo o sequenciamento completo do genoma, a análise das sequências nucleotídicas e aminoacídicas, bem como análise evolutiva, provavelmente, proporcionará um melhor entendimento da estrutura genômica desses vírus e, ao mesmo tempo, gerar

informações sobre possíveis sítios antigênicos presentes, patogênese e aspectos evolutivos. Finalmente, os dados obtidos poderão ser utilizados em estudos sistemáticos, bem como para o desenho de iniciadores específicos que poderão ser empregados no desenvolvimento de técnicas moleculares para o diagnóstico rápido dos mesmos, contribuindo para futuros programas de monitoramento desses flavivírus.

1.12. OBJETIVOS

1.12.1. Objetivo Geral

Realizar a caracterização molecular do genoma completo dos flavivírus brasileiros Bussuquara, Iguape, Ilhéus e Rocio.

1.12.2. Objetivos Específicos

- Sequenciar a região codificadora, bem como as regiões não codificantes 5' e 3' do genoma dos flavivírus Bussuquara, Iguape, Ilhéus e Rocio, identificando as semelhanças e diferenças genéticas entre os vírus em estudo;
- Demonstrar a estrutura secundária do RNA dos vírus Bussuquara, Iguape, Ilhéus e Rocio e comparar suas estruturas e sequências conservadas com as de outros flavivírus;

- Identificar os sítios de clivagem, resíduos de cisteína e sítio de glicosilação nas sequências de aminoácidos das poliproteínas dos flavivírus Bussuquara, Iguape, Ilhéus e Rocio;
- Identificar os domínios e os motivos conservados das proteínas E, NS3 e NS5 dos flavivírus Bussuquara, Iguape, Ilhéus e Rocio
- Realizar a análise filogenética com base nos genes estruturais e não-estruturais, promovendo a comparação da região codificante, bem como gene por gene desses vírus com outros flavivírus;

2. MATERIAL E MÉTODOS

2.1 AMOSTRAS

Para este estudo foram utilizadas as amostras dos VBSQ, VIGU, VILH e VROC (Quadro 4) provenientes do acervo da Seção de Arbovirologia e Febres Hemorrágicas (SEARB) do Instituto Evandro Chagas (IEC), conforme o termo de consentimento em anexo (ANEXO I), em parceria com a equipe do Dr. Goro Kuno, da Divisão de Doenças Infecciosas Transmitidas por Vetores do Centro de Controle e Prevenção de Doenças (CDC), Fort Collins, Colorado, EUA.

Quadro 4 – Amostras virais utilizadas neste estudo, mostrando o número de registro da cepa, amostra biológica de origem e número de passagens.

Vírus	Cepa	Amostra de origem	Ano de isolamento	Estoque inicial*	Passagem*
VBSQ	BE AN 4073	Sangue de <i>Alouatta belzebul</i>	1956	Liofilizado de cérebro de camundongo	9 ^a
VIGU	SP AN 71686	Vísceras de camundongo sentinela	1979	Cérebro de camundongo	5 ^a
VILH	BE H 7445	Sangue de caso humano febril**	1957	Liofilizado de cérebro de camundongo	4 ^a
VROC	SP H 34675	Encéfalo de caso de encefalite humana	1975	Cérebro de camundongo	5 ^a

* amostras utilizadas no estudo;

** paciente procedente de uma colônia japonesa do bairro do Guamá em Belém, Pará.

2.2. PROPAGAÇÃO VIRAL EM CÉLULAS VERO

As amostras virais liofilizadas ou de suspensão de cérebro de camundongos infectados foram propagadas em cultivos celulares mediante a inoculação das suspensões virais em garrafas de 175 cm² contendo monocamadas confluentes de células VERO (células de rim de macaco verde africano) mantidas em meio de manutenção *Minimum Essential Medium* (MEM, Sigma) enriquecido com soro bovino fetal a 2%. As culturas infectadas foram examinadas diariamente até a observação de efeito citopático (ECP) em 70 a 75% da área da monocamada celular, quando então o sobrenadante das culturas infectadas foram colhidos e usados para a obtenção do RNA viral.

2.3. EXTRAÇÃO DO RNA VIRAL

A obtenção do RNA viral a partir das culturas infectadas foi feita mediante a aplicação do método do reagente TRIZOL LS (Invitrogen) e, alternativamente, empregando o kit comercial Qlamp Viral RNA Mini kit (Qiagen), seguindo estritamente as instruções dos fabricantes. O RNA obtido foi então usado imediatamente para a obtenção do cDNA ou mantido em estoque à -70°C até o uso.

2.4. TRANSCRIÇÃO REVERSA SEGUIDA DA REAÇÃO EM CADEIA MEDIADA PELA POLIMERASE (RT-PCR)

A técnica de RT-PCR foi feita em duas etapas, primeiro a Transcrição reversa (RT) seguida da reação em cadeia mediada pela polimerase (PCR), com base nos protocolos descritos por Chang (1997) e Kuno & Chang (2005), utilizando os iniciadores descritos no Quadro 5. O cDNA referente à região entre o término 5' do genoma e a região conservada CS2 da 3'RNC dos flavivírus foi obtido pela técnica de RT, utilizando o iniciador VD8 (Pierre et al. , 1994) (Figura 13). Iniciadores adicionais específicos para os quatro vírus em estudo foram desenhados, com base nas sequências geradas pelo protocolo de Chang (1997) e Kuno & Chang (2005), para complementar as sequências genômicos dos flavivírus, bem como para confirmar as sequências obtidas.

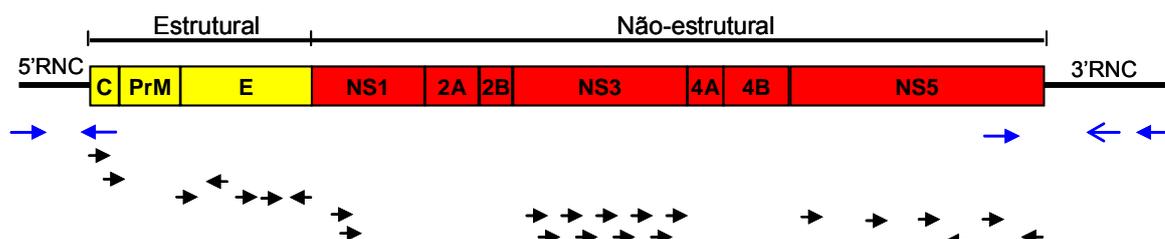


Figura 13 – Esquema da estratégia de amplificação do genoma (RT-PCR) e seqüenciamento nucleotídico estabelecida por Chang (1997) e Kuno & Chang (2005). Em azul, destacam-se os iniciadores utilizado pelo protocolo 5'RACE e 3'RACE, a seta “←” indica o iniciador VD8 descrito por Pierre e colaboradores (1994).

Quadro 5 - Iniciadores utilizados para a técnica de RT-PCR e sequenciamento nucleotídico (Chang, 1997; Kuno & Chang, 2005).

Localização genômica	Motivos (sequência de aminoácido)	Direção do Iniciador	Sequência dos Iniciadores (5' → 3')
Capsídio		Senso Senso	ATGWCTAARAAACCAGGA TCAATATGCTRAAACGCGG
Envelope	DRGWGNGC GLFGKGS GHLKCRV PFGDSYIV DTAWDFGS	Senso Reverso Senso Senso Reverso Senso	GAYMGWGGVTGGGGHAAAYGGVTG GGMYTKTTYGGDAARGGRAGC GGMCAYSTBWMYTGMTGVSTG CCSTTYGGWGAYTCNTACATHGT TCCYCTKCCHACYACDATGTA GAYACDGCBTGGGAYTTTGGMTC
NS1	GCWYGMEI YGMEIRP	Senso Senso	GGHTGTTGGTAYGGMATGGA TAYGGMATGGARAYHMGRCC
NS3	GTSGSPI GLYGNG LAPTRVV DVMCHATF MDEAHF SIAARGY MTATPPG ISEMGAN SAAQRRGR	Senso Senso Senso Senso Senso Senso Senso Senso Senso	GGMACDTCNGGHTCNCCHAT GGNCTNTAYGGNAAYGG YTRGCDCCNACNMGRGTNGT GAYGTSATGTGYCAYGCHAC ATGGAYGARGCHCAYTT AGYATMGCNGCHMGAGG ATGACNGCVACNCCNCCNGG ATHTCNGARATGGGDGCVAA WSYGCWGCBCARAGRMGRGGVMG
NS5	DLGCGRG SRNSTHEMY NMMGKREKK ADDTAGWDT WMTTEDML	Senso Senso Senso Senso Reverso Reverso	GAYCTNNGNTGYGGNMGNNGG TCAAGGAACTCCACACATGAGATGTACT TACAACATGATGGGAAAGAGAGAGAA GCTGATGACACCGCCGGCTGGGACAC GTGTCCCAGCCGGCGGTGTCATCAGC AGCATGTCTTCCGTGGTCATCCA
3RNC/ CS2*		Reverso	GGGTCTCCTCTAACCTCTAG

Legenda : * iniciador VD8 descrito por Pierre et al., 1994 ; bases degeneradas D,B, H, K, M,N, W, Y,R,T ver Anexo III

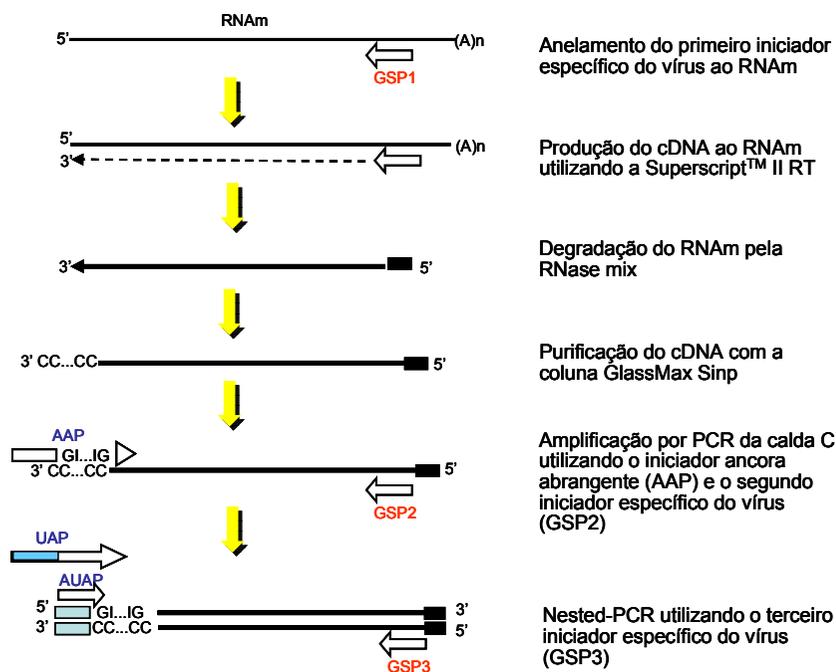
A desnaturação do RNA viral a 94°C por 2 minutos, foi seguida da realização da etapa de RT, cuja mix conteve 5 µL do tampão de RT (5X), 1 µL de mix de deoxinucleotídeo (dNTPs - 10mM), 1 µL (6 U) da enzima transcriptase reversa (Superscript II, Invitrogen), 1 µL do iniciadores reversos (100 µM), 5 µL do RNA viral desnaturado, completando a reação para 20 µL com água deionizada livre de RNase e DNase (Invitrogen). A RT ocorreu em um único ciclo a 42°C por 1 hora.

A PCR foi realizada empregando o kit *Expand Long Template PCR System* (Roche), no qual uma alíquota de 10 µL do cDNA obtido da etapa de RT foi processada em uma reação contendo: 10 µL do Tampão de PCR (10X), 2 µL de dNTP mix, 1 µL de cada iniciador, senso e reverso (50 µM), 2 µL da DNA polimerase (*Platinum Taq polimerase*, Invitrogen), para uma reação final para 100 µL. O termociclador foi programado para um ciclo de 94°C-1 min/50°C-1 min/68°C-5 min; três ciclos de 94°C-20 sec/50°C-1 min/68°C-4 min; dez ciclos de 94°C-20 sec/50°C-30 sec/68°C-4 min e um ciclo de extensão final a 68°C por 7 min.

As 5'RNC e 3'RNC foram amplificadas com auxílios do kits 5'RACE e 3'RACE (Invitrogen, Life Technologies), respectivamente, segundo o protocolo do fabricante (Figura 14), para os quais foram desenhados iniciadores específicos para os flavivírus em estudo (Quadro 6).

Os amplicons obtidos foram observados em gel de agarose 1,5% corado com brometo de etídio (10 mg/mL), utilizando transiluminador com luz UV (Vilber Lourmat).

(A) 5' RACE



(B) 3'RACE

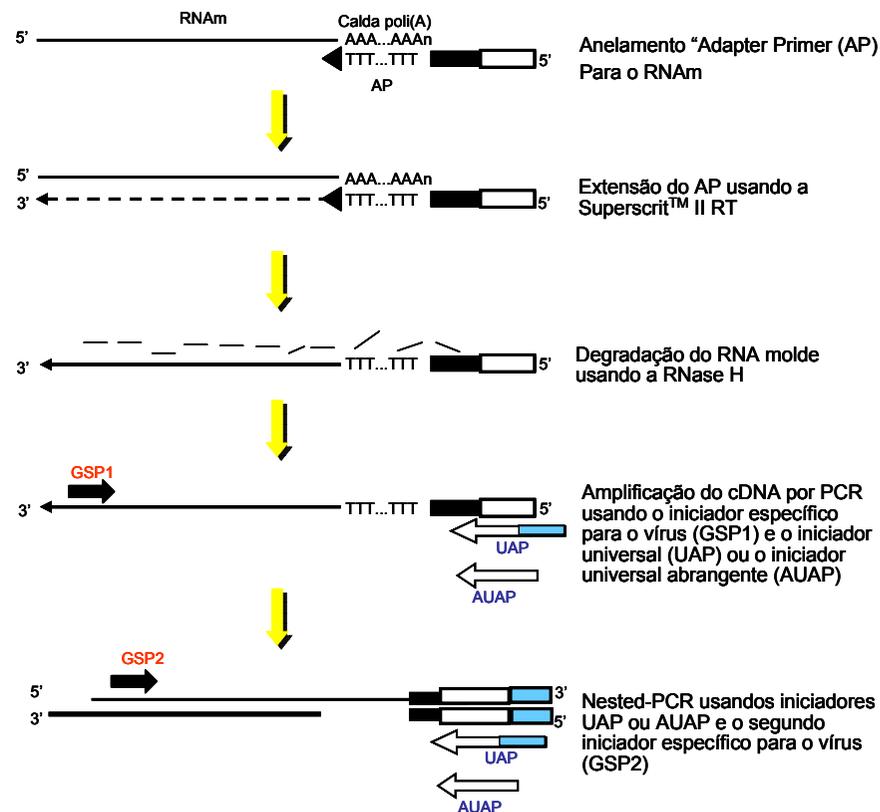


Figura 14 – Esquema do protocolos 5'RACE (A) e 3'RACE. Fonte: adaptado do protocolo do fabricante (Invitrogen). Os iniciadores em azul são iniciadores disponíveis nos kits, enquanto os iniciadores em azul estão demonstrado no Quadro 6.

Quadro 6 – Iniciadores específicos para os flavivírus brasileiros utilizados no protocolo 5' RACE e 3'RACE

5'RACE			
Vírus	iniciador	Sequência	Tm (°C)
VBSQ	GP1	CAGACAAGCGGTCAACAGCA	63,59
	GP2	TGCTCTCAACGGAGCGTTTT	63,29
	GP3	CCAGCATGCTTCCAATCTCC	63,01
VIGU	GP1	TTCCGGTTTGGAAACACAATG	62,9
	GP2	AGTCCGTGCAGTCCACGTTT	63,5
	GP3	TGGATCATCGAGCCAAGTTGT	62,91
VILH	GP1	CGACCGCTGTGATCATGGTA	63,06
	GP2	TCCTTGCGCTTGTTGACTGA	63,07
	GP3	TTGTTGACAGAACGCCATCG	63,19
VROC	GP1	GATTGACATCAAACTTTGCCTTG	62,01
	GP2	AAGCGCGATCCAAAGTAGCA	63,18
	GP3	TGGCCTTCCGTCTATTGATGA	62,78
3'RACE			
Vírus	iniciador	Sequência	Tm(°C)
VBSQ	GP1	GTCTGGAACAGAGTGTGGAT	54,89
	GP2	ATATTTGGGAAAAAGGGAAG	54,64
VIGU	GP1	TGAGAGACTGGAAGGAAGTG	55,42
	GP2	GAAGCCTTATTGGTTACCG	54,92
VILH	GP1	GTCTGGATTGAGGACAATGA	55,92
	GP2	GGACTGACATCCCCTACAT	54,57
VROC	GP1	GTGGAACAGAGTGTGGATCT	54,89
	GP2	GACTGACGTCCCCTACATT	54,85

Tm = temperatura de melt

2.5. SEQUENCIAMENTO NUCLEOTÍDICO

Os cDNAs obtidos foram purificados usando o kit comercial *GFX™ PCR DNA and Gel Band Purification kit* (GE healthcare). As sequências dos flavivírus estudados foram obtidas em sequenciador automático modelo ABI 377 (Applied Biosystems) empregando o método de terminação de cadeia descrito por Sanger e colaboradores (1977). A estratégia de sequenciamento foi a mesma descrita por Kuno & Chang (2005).

Na reação de sequenciamento foram utilizados 4 μL de *Terminator ready reaction mix*, 2 μL (80 ng) do produto de PCR purificado e 3,2 pmoles de oligonucleotídeos, ajustados com água livre de RNase e DNase para um volume final de 10 μL . A amplificação foi realizada em termociclador (GeneAmp PCR Systems 9600, Applied Biosystems), usando um programa específico que consiste de 25 ciclos, cada um composto por etapas de desnaturação a 96°C por 10 segundos, anelamento dos iniciadores a 50 °C por 5 segundos e extensão a 60 °C por 4 minutos.

O produto final foi precipitado com a adição de 80 μL de isopropanol 75% (álcool isopropílico, Merck). A mistura foi agitada, incubada por 20 minutos a temperatura ambiente e centrifugada a 14.000 rpm por 25 minutos. Ao final, o sobrenadante foi retirado e desprezado, e em seguida o precipitado foi lavado com 250 μL de etanol a 70% e centrifugado a 12.000 rpm por 5 minutos. O sobrenadante foi descartado e o precipitado foi incubado a 90°C por 2 minutos para a completa evaporação do etanol. O precipitado foi então reconstituído em 5 μL de solução de formamida (Amresco) deionizada,

na proporção de 5:1 com azul dextran/EDTA (Applied Biosystem), aquecido a 95 °C por 3 minutos e submetido à eletroforese no sequenciador automático.

2.6. ALINHAMENTO DAS SEQUÊNCIAS NUCLEOTÍDICAS

As sequências nucleotídicas foram montadas pelo programa SeqMan™ II, versão 5.03 (DNASTAR ® Lasergene, EUA). Este software também foi utilizado para indicar a ORF relacionada à poliproteína dos vírus em estudo, bem como determinar as regiões 5'RNC e 3'RNC. O alinhamento das sequências obtidas juntamente com sequências de outros flavivírus (Quadro 7) foi feito pelo programa Clustal W (Thompson et al., 1997), implementado pelo programa BioEdit versão 5.0 (Hall, 2001), e ajustadas manualmente.

Quadro 7- Sequências do genoma completo dos 32 flavivírus utilizados para o alinhamento e filogenia neste estudo.

VÍRUS	SIGLA	ISOLAMENTO			GENBANK	
		ANO	HOSPEDEIRO	LOCAL	CEPA	NUMERO DE ACESSO
<i>Vírus Alfuy</i>	VALFUY	1966	<i>Centropus plasinius</i>	Austrália	MRM-3929	AF013360
<i>Vírus Apoi</i>	VAPOI	1954	<i>Apodimus</i> sp <i>Clethrionomys</i> sp	Japão	ApMAR	NC003676
<i>Vírus Bagaza</i>	VBAG	1966	<i>Mixed cullex</i>	Bagaza República Central	DakAr B209	NC012534
<i>Vírus Bussuquara</i>	VBSQ	1956	<i>Alouatta belzebul</i>	Brasil	BeAn 4073	NC009026
<i>Vírus cell fusion agent</i>	VCFA	1999	mosquitos	Kenya	?	NC001564
<i>Vírus da encefalite japonesa</i>	VEJ	1935	Humano	Tokyo, Japão	?	NC001437
<i>Vírus encefalite Murray Valley</i>	VEMV	1951	Humano	Victoria/AS	?	NC000943
<i>Vírus da encefalite Saint Louis</i>	VESL	1933	Humano	Missouri/EUA	Kern217	NC007580
<i>Vírus da febre amarela</i>	VFA	1927	Humano	África	?	NC002031
<i>Vírus dengue sorotipo 1</i>	VDEN1	1944	Humano	Hawai	?	NC001477
<i>Vírus dengue sorotipo 2</i>	VDEN2	1944	Humano	Nova guiné	16681	NC001474
<i>Vírus dengue sorotipo 3</i>	VDEN3	1956	Humano	Tanda Filipinas	?	NC001475
<i>Vírus dengue sorotipo 4</i>	VDEN4	1956	Humano	Quezon Filipinas	?	NC002640
<i>Vírus Iguape</i>	VIGU	1979	Camundongo sentinela	São Paulo	SP AN 71686	NC009027
<i>Vírus Ilhéus</i>	VILH	1944	<i>Ae. psorophora</i>	Ilhéus/BA	BE H 7445	NC009028
<i>Vírus Kamiti river</i>	VKR	?	<i>Ae. macintoshi</i>	Kenya	SR-82	NC005064
<i>Vírus kedougou</i>	VKED	?	?	Senegal	DakAar D1470	NC012533
<i>Vírus kokobera</i>	VKOK	1960	<i>Culex annulirostris</i>	Norte Queensland/AS	AusMRM 32	NC009029
<i>Vírus Langat</i>	VLGT	1956	<i>Ixodes granulatus</i>	Malásia	TP21	NC003690
<i>Vírus Louping ill</i>	VLI	1929	<i>Ovis ovis</i> (ovelha)	Escócia	369/T2	NC001809
<i>Vírus Modoc</i>	VMOD	1958	<i>Peromyscus mampulatus</i>	California	M544	NC003635
<i>Vírus Montana miotis leucoencefalite</i>	VMML	1958	<i>Myotis lucifugus</i>	Montana	?	NC004119
<i>Vírus do Nilo Ocidental</i>	VNO	1937	Humano	Uganda	Original	EU074020
<i>Vírus omsk hemorrhagic fever</i>	VOHF	1947	Humano	OMSK oblast Rússia	Bogoluvovska	NC005062
<i>Vírus powassan</i>	VPOW	1958	Humano	Ontário Canadá	LB	NC003687
<i>Vírus Rio Bravo</i>	VRB	1954	<i>Tadarida brasiliensis mexicana</i>	Califórnia	RiMAR	NC003675
<i>Vírus rocío</i>	VROC	1975	Humano	São Paulo	SP H 34675	AF013397
<i>Vírus sepike</i>	VSEP	?	mosquitos	Papua Nova Guiné	MK7148	NC008719
<i>Vírus tick-borne encefalite</i>	VTBE	1975	?	?	?	NC001672
<i>Vírus Usutu</i>	VUSU	2001	Pássaro negro	Áustria	Vienna 2001	NC006551
<i>Vírus Wesselsbron</i>	VWES	?	?	África do Sul	SAH177	NC012735
<i>Vírus zika</i>	VZIK	?	Macaco sentinela	Uganda	MR 766	NC012532

2.7. IDENTIFICAÇÃO DOS MOTIVOS CONSERVADOS, SÍTIOS DE CLIVAGEM, SÍTIOS DE GLICOSILAÇÃO E RESÍDUOS DE CISTEÍNA NA POLIPROTEÍNA,

Os sítios de clivagem da poliproteína dos flavivírus em estudo foram identificados seguindo o esquema de processamento proteolítico em cascata para ORF dos flavivírus previamente descrito por Rice e Strauss (1990). As junções do capsídeo intracelular e pre-membrana (Ci/prM), prM e envelope (prM/E), proteína E e a proteína não estrutural NS1 (E/NS1) foram determinadas com base no escore de maior potencial de clivagem para sinalase da célula hospedeira, utilizando o programa Signal PNN (<http://www.cbs.dtu.dk/services/>) (Chang et al., 2000). Os sítios de glicosilação e resíduos de Cisteína foram determinados pelos programas NetNglyc (versão 1.0) (<http://www.cbs.dtu.dk/services/>) e Protean (versão 5.03) (Lasergene, DNASTAR®), respectivamente.

Para a identificação dos principais domínios das proteínas E, NS3 e NS5 foram realizadas análises de similaridade em bancos de dados biológicos secundários (também conhecidos como bancos de dados de padrões ou de assinaturas) que determinam as funções específicas de proteínas. Tais análises de domínios protéicos foram realizadas no banco de dados INTERPRO através da ferramenta Interproscan (<http://www.ebi.ac.uk/Tools/InterProscan>). O INTERPRO alberga vários bancos de dados de domínios, motivos e famílias protéicas, tais como: ProDom, SMART, TIGRFAMS, Pfam, SUPERFAMILY, PANTHER e SignalPHMM, diminuindo assim a redundância de dados e aumentando a acurácia da predição. Os

resultados dessas análises corroboram a identificação de proteínas não-redundantes como também direcionam na inferência de função de proteínas espécie-específicas.

Os motivos conservados das proteínas E, NS3 e NS5 foram verificados no alinhamento as sequências aminoacídicas entre os vírus em estudo e os demais flavivírus, com base nos achados da literatura (Kamer & Argos, 1984; Koonin, 1991; Chang, 1997; Gorbalenya et al., 1989; Matusan et al., 2001; Zhou et al., 2007; Seligman, 2008).

2.8. ESTRUTURA SECUNDÁRIA DAS 5'RNC/ 3'RNC E CICLIZAÇÃO DO RNA VIRAL

A estrutura secundária ao nível da 3'RNC e formação de alças entre os terminos 3'RNC e 5'RNC foram determinadas pelo programa MFOLD 3.0 (Zuker et al., 1999) (<http://mfold.burnet.edu.au/>). Os primeiros 200 nt no terminal 5' e toda a sequência nucleotídica do terminal 3'RNC após o códon de parada do gene NS5 foram utilizados na análise, bem como os dois terminais 3' RNC e 5' RNC para a construção da estrutura secundária de ciclização do RNA viral, similar ao estudo proposto por Khromykh e colaboradores (2001b).

2.9. DETERMINAÇÃO E ANÁLISE DE SIMILARIDADE (*BOOTSCANNING*) DA REGIÃO CODIFICANTE

Inicialmente, as regiões codificantes dos flavivírus VBSQ, VIGU, VILH e VROC foram identificadas utilizando o programa *ORF finder* disponível no portal NCBI (National Center for Biological Information; <http://www.ncbi.nlm.nih.gov/projects/gorf/>).

Posteriormente, o programa Bootscan (SimPlot versão 3.5.1. - <http://sray.med.som.jhmi.edu/SCRoftware/simplot/>) foi empregado para examinar as sequências nucleotídicas dos flavivírus em estudo ao longo de toda a região codificante (Felsenstein, 2009; Ray, 1999) e comparadas com sequências de outros flavivírus previamente selecionados. Árvores filogenéticas pelo método de Agrupamento de Vizinhos (AV) foram obtidas para valores de *bootstrap* fixados em 1.000 réplicas, usando os programas SEQBOOT e DNAPARS (PHYLIP software versão 3.69) (Felsenstein, 2009), utilizando o modelo de substituição nucleotídica* Kimura 2 parâmetros e alternativamente F84. A análise foi repetida de 20 em 20 nt ao longo de todo o genoma. Os valores obtidos para *bootstrap* indicaram o grau de relação filogenética entre os vírus em estudo e outros flavivírus, sendo tabulados e plotados a cada passo.

* O modelo de substituição nucleotídica escolhido para a análise de bootscan foi determinado pela disponibilidade do programa.

2.10. ANÁLISE FILOGENÉTICA COMPARATIVA

Para a análise filogenética foram adotados os métodos de MV utilizando o programa PHYML (Guindon & Gascuel, 2003) e Bayesiano, implementado pelo programa MrBayes versão 3 (Ronquist & Huelsenbeck, 2003). O programa Moldtest versão 3.7 (Posada & Crandall, 1998) foi utilizado para determinar o melhor modelo de substituição nucleotídica a ser usado baseado no critério de informação Akaike (*Akaike Information Criterion*, AIC), com o auxílio do programa PAUP, versão 4 (Swofford, 2002).

Para o método MV, a análise de *bootstrap* foi realizada usando 1.000 pseudoréplicas (valor de confiabilidade de 95%) obtidas a partir do alinhamento original entre as sequências para a obtenção de resultados mais confiáveis (Felsenstein *et al.*, 1985).

A análise Bayesiana foi realizada empregando o modelo de Markov Monte Carlo (*Markov Chain Monte Carlo*, MCMC). A análise foi feita para dois milhões de réplicas, sendo a amostragem fixada a cada 1.000 árvores geradas. Os valores Bayesianos foram estimados em 50% das árvores consenso geradas (Huelsenbeck & Ronquist, 2003). O programa TRACER (www.evolve.zoo.ox.ac.uk) foi utilizado para verificar se a análise realizada pelo programa MrBayes alcançou a convergência apropriada.

3. RESULTADOS

3.1. ANÁLISE DO GENOMA VIRAL

O tamanho dos genomas completos obtidos para os flavivírus brasileiros em estudo ficaram em torno de 10.800 nt, sendo determinado 10.815 nt para o VBSQ, 10.922 nt para o VIGU, 10.775 nt para o VILH e 10.794 nt para o VROC. Para cada vírus foi reconhecida uma única ORF, flanqueada pelas 5' RNC e 3'RNC. Em cada ORF foram reconhecidos os três genes que codificam as proteínas estruturais, C-prM-E, e os setes genes das proteínas não-estruturais, NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5. A 5'RNC dos vírus VBSQ e VIGU apresentaram 92 nt, enquanto que para VILH e VROC foi de 104 nt. Maior variação em tamanho foi observada ao nível da 3'RNC, sendo que para o VILH determinou-se a de menor tamanho com 383 nt, seguido do VBSQ (418 nt), VROC (424 nt) e VIGU (564 nt). As características dos genomas dos quatro flavivírus em estudo encontram-se sumarizadas no Quadro 8.

Os tamanhos dos genes Pr/M, NS2A, NS2B e NS4A/2K foram similares para os quatro vírus: 92/75 aa, 227 aa, 131 aa e 126/23 aa, respectivamente. Para a proteína C, o VBSQ apresentou o maior comprimento, 125 aa, seguido do VIGU (119 aa) e dos VILH e VROC (118 aa); a posição de clivagem da AnchC variou em um anti-códon para os VIGU e VROC, determinando um tamanho de 17 aa. Quanto a NS3, com exceção do VILH (618 aa), os demais vírus apresentaram 619 aa. A variabilidade no tamanho do genoma também foi observado quanto a NS4B (variação de 253 a 256 aa) e NS5 (904 aa para VIGU e VROC; 905 aa para VBSQ e VILH) (Quadro 8).

Quadro 8 – Comparação do genoma dos quatro flavivírus brasileiros em estudo: tamanho do genoma, regiões codificadoras para proteínas estruturais e não-estruturais, RNCs.

Componentes do genoma	VBSQ			VIGU			VILH			VROC			
	nt	posição	aa	NT	posição	aa	nt	posição	aa	nt	posição	aa	
5'RNC	104	1 a 104	-	104	1 a 104	-	92	1 a 92	-	92	1 a 92	-	
Proteínas estruturais	VirC	321	105 a 425	107	306	105 a 410	102	300	93 a 392	100	303	93 a 395	101
	AnchC	54	426 a 479	18	51	411 a 461	17	54	393 a 446	18	51	396 a 446	17
	Pr		480 a 755	92	276	462 a 737	92	276	447 a 722	92	276	723 a 947	92
	M	225	756 a 980	75	225	738 a 962	75	225	723 a 947	75	225	948 a 2.450	75
	E	1.503	981 a 2.483	501	1.482	963 a 2.444	494	1.503	948 a 2.450	501	1.503	948 a 2.450	501
Proteínas não-estruturais	NS1	1.059	2.484 a 3.542	352	1.059	2.445 a 3.503	353	1.059	2.451 a 3.509	353	1.059	2.451 a 3.509	353
	NS2A	681	3.543 a 4.223	227	681	3.504 a 4.184	227	681	3.509 a 4.190	227	681	3.510 a 4.190	227
	NS2B	393	4.224 a 4.616	131	393	4.185 a 4.577	131	396	4.191 a 4.586	131	393	4.191 a 4.583	131
	NS3	1.857	4.617 a 6.473	619	1.857	4.578 a 6.434	619	1.854	4.587 a 6.440	618	1.857	4.584 a 6.440	619
	NS4A	378	6.474 a 6.851	126	378	6.435 a 6.812	126	378	6.441 a 6.818	126	378	6.441 a 6.818	126
	2K	69	6.852 a 6.920	23	69	6.813 a 6.881	23	69	6.819 a 6.887	23	69	6.819 a 6.887	23
	NS4B	759	6.921 a 7.679	253	762	6.882 a 7.643	254	765	6.887 a 7.652	255	768	6.888 a 7.655	256
	NS5	2.718	7.680 a 10.397	905	2.715	7.644 a 10.358	904	2.718	7.653 a 10.370	905	2.715	7.656 a 10.370	904
3'NCR	418	10.398 a 10.815	-	564	10.359 a 10.922	-	385	10.371 a 10.755	-	424	10.371 a 10.794	-	
ORF	10.290	-	3.429	10251	-	3416	10.272	-	3424	10.275	-	3.425	
Genoma Total	10.815	-	-	10922	-	-	10.775	-	-	10.794	-	-	

Legenda: VBSQ: vírus Bussuquara; VIGU: vírus Iguape; VILH: vírus Ilhéus; nt: nucleotídeos; aa: aminoácido; NCR: Região não-codificadora; ORF: *Open*

Reading Frames

3.2. REGIÕES NÃO CODIFICADORAS

3.2.1. Sequências Conservadas

Dentre as sequências conservadas da 3'RNC dos flavivírus transmitidos por mosquitos (CS1[3'CYC], CS2, RCS2, CS3, RCS3), as sequências localizadas dentro do 3'SL e o pentanucleotídeo 5'-CACAG-3', mostraram a presença das CS1[3'CYC], CS2, RCS2 e 5'-CACAG-3' nos genomas dos quatro vírus em estudo (Figura 15), muito embora a CS1 do VIGU seja considerada imperfeita por apresentar três alterações, sendo uma delas uma inserção de "A" (Figura 16). Quanto às sequências CS3 e RCS3, comumente observadas nos representantes do grupo da encefalite japonesa, as mesmas encontram-se ausentes no genoma do VBSQ, enquanto que a CS3 foi encontrada na 3'RNC dos VIGU, VILH e VROC. Comparando-se com a RNC dos demais flavivírus, o VBSQ apresentou disposição das sequências conservadas similar aos vírus causadores de doença icterohemorrágica do grupo dengue, enquanto que os VIGU, VILH e VROC apresentaram padrão similar aos vírus encefalitogênicos do grupo da encefalite japonesa (VEJ, VEMV, VNO), excetuando a RCS3 (Figura 15)

Em relação à sequência consenso gerada pelo alinhamento dos quatro flavivírus em estudo com os flavivírus transmitidos por *Culex*, na região análoga da CS3 do VBSQ foram observadas oito substituições nucleotídicas, que descaracterizaram a CS3. Algumas alterações pontuais foram observadas: CS3 do VIGU, posição 10.614 (A→T); CS2 do VIGU, posição 10.771(A→T); CS2 do VILH, posição 10.616 (A→T); CS1 do VIGU, substituição nucleotídica

nas posições 10.817(G→A) e 10.823(T→C) – esta dentro da 3'CYC e uma inserção de “A” na posição 10.818; CS1 do VILH, posição 10.660 (A→G); pentanucleotídeo da 3'SL do VIGU, posição 10.880(A→C) (Figura 16).

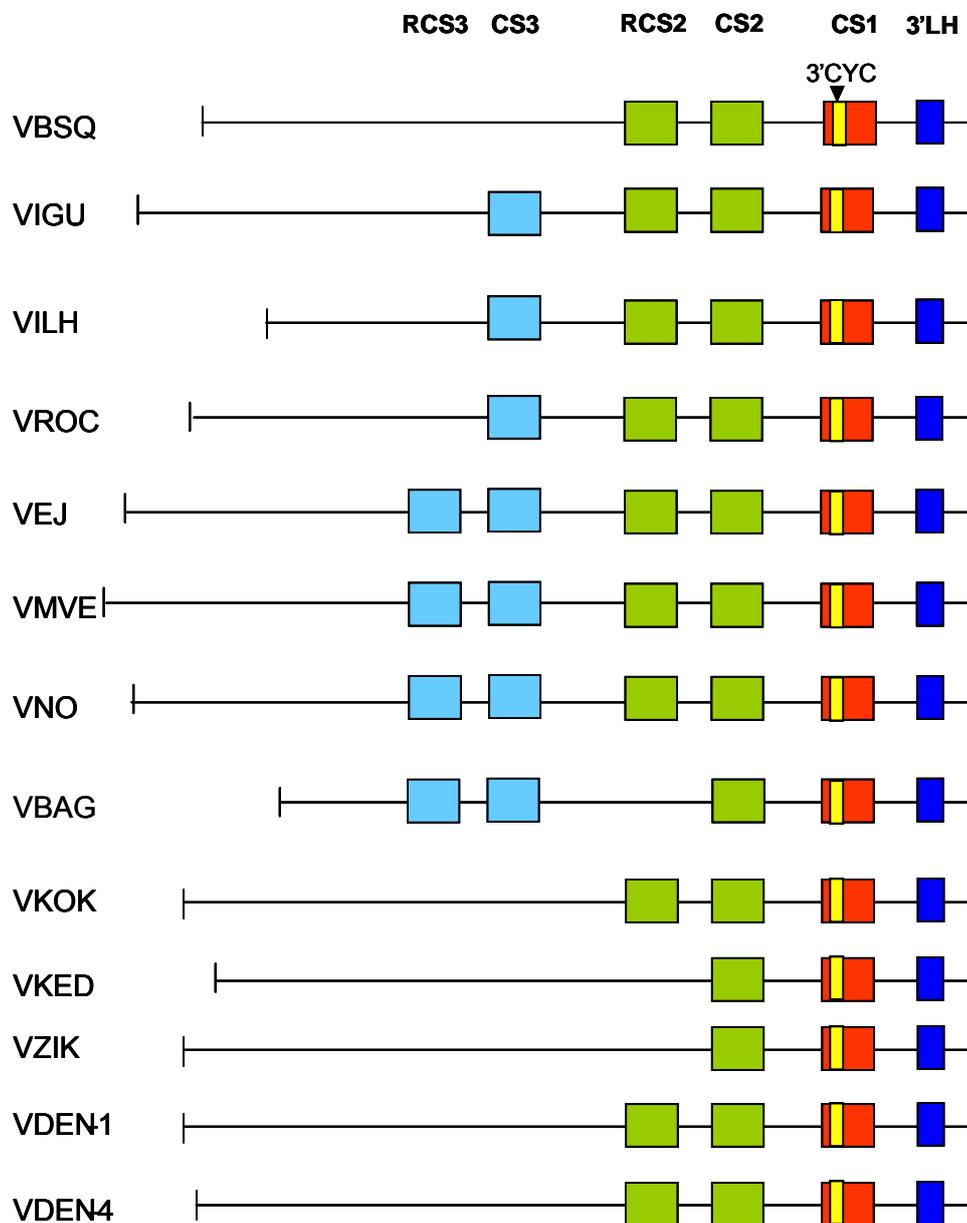


Figura 15 – Representação esquemática das 3'RNC dos *Vírus bussuquara* (VBSQ), *Vírus Iguape* (VIGU), *Vírus Ilhéus* (VILH) e *Vírus rocio* (VROC) em comparação a outros flavivírus transmitidos por mosquitos dos grupos: Encefalite japonesa [*Vírus da encefalite japonesa* (VEJ)], *Vírus da encefalite Murray Valley* (VEMV) e *Vírus Nilo Ocidental* (VNO)], Ntaya [*Vírus Bagaza* (VBGA)], Kokobera [*Vírus kokobera* (VKOK)], Kedougou [*Vírus kedougou* (VKED)], Spondweni [*Vírus zika* (VZIK)] e Dengue [*Vírus dengue sorotipo 1* (VDEN1) e *Vírus dengue sorotipo 4* (VDEN4)]. Observar que a CS3 e RCS3 está ausente no VBSQ, bem como a RCS3 está ausente no VIGU, VILH e VROC.

CS3			RCS2		
Consenso	GCCCCAGGTGGACTGG	posição	GACTAGAGGTTAGAGGAGACCC	posição	
VBSQ	T.T..G.AA...A.AA	10477-10492	10579-10600	
VIGUT.....	10606-10621	10681-10702	
VILH	10451-10466	10529-10550	
VROC	10289-10304	10364-10385	
VALFUYT.....	10644-10659	10724-10745	
VEJ	10667-10682	10748-10769	
VUSU	10750-10765	10832-10853	
VEMV	10703-10718	10783-10804	
VNO	10718-10733	10801-10822	
VKUN	10711-10726	10794-10815	
VESL	10640-10655	10716-10737	
VBAGT.....	10710-10725	-----		
VKOK	CT.....	10567-10582	10646-10667	
VKED	C...TG...CT.T...	10457-10471	GG..G.T.CG..A.AT-...A.	10508-10528	
VZIK	CG..AT....G...G	10481-10496	10564-10585	

CS2			CS1			3'SL		
Consenso	GACTAG-AGGTTAGAGGAGACCC	posição	CAG-CATATTGACACCTGGGA	posição	CACAG	posição		
VBSQ	10656-10677	10709-10728	CACAG	10770-10774		
VIGUT.....	10760-10781	..AA...C.....	10815-10835	CACCG	10877-10881		
VILHT.....	10603-10624G.....	10648-10667	CACAG	10709-10713		
VROC	10441-10462	10488-10507	CACAG	10549-10553		
VALFUY	10801-10822C.....	10850-10869	CACCG	10914-10918		
VEJ	10826-10847	10863-10882	CACAG	10928-10931		
VUSU	10908-10929	10952-10971	CACAG	11018-11022		
VEMV	10857-10878	10901-10920	CACCG	10966-10970		
VNO	10875-10896	10922-10941	CACAG	10983-10987		
VKUN	10794-10815	10915-10934	CACAG	10976-10980		
VESL	10791-10812	10833-10852	CACAG	10893-10897		
VBAG	10787-10808	10834-10835	CACAG	10896-10900		
VKOK	10723-10744	10771-10790	CACAG	10831-10835		
VKEDG.....	10567-10589	10614-10633	CACAG	10677-10681		
VZIK	10646-10668	...-.....--.....	10685-10702	CACAG	10746-10750		

Figura 16 - Alinhamento entre as sequências nucleotídicas conservadas da 3'RNC dos flavivírus transmitidos por mosquitos do gênero *Culex*, seguindo a disposição 5'→3', comparando com os flavivírus VBSQ, VIGU, VILH e VROC. A sequência de ciclização, 3'CYC está representada em negrito dentro da CS1. Em amarelo, destacam-se as sequências que apresentam mais de cinco alterações, o que descaracteriza as sequências conservadas. **Legenda:** VALFUY = *Vírus alfuy*; VBAG = *Vírus Bagaza*; VBSQ = *Vírus bussuquara*; VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*; VESL = *Vírus encefalite Saint Louis*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VKUN = *Vírus kunji*; VNO = *Vírus Nilo ocidental*; VROC = *Vírus rocio*; VUSU = *Vírus usutu*; VZIK = *vírus zika*.

3.2.2. Estruturas Secundárias das 5' RNC e 3'RNC e Ciclização do RNA viral

As estruturas secundárias preditivas para a ciclização entre as regiões 5' e 3'RNC dos flavivírus brasileiros, gerada pelo programa MFOLD, apresentaram similar organização para a 5'RNC e diferentes padrões para a 3'RNC. A Figura 17 representa os modelos da estrutura secundária de ciclização para os respectivos flavivírus brasileiros, que apresentaram baixos níveis de energia e distintos entre os flavivírus em estudo (VBSQ: -106,4 Kcal; VIGU: -261,7 Kcal; VILH: -101,1 Kcal; VROC: -97,0 Kcal).

Nas estruturas de ciclização, os nucleotídeos terminais das 5' RNC e 3'RNC estão localizados próximos, no entanto não estão ligados. Foram encontradas duas sequências complementares entre 5' RNC e 3'RNC: (i) as sequências complementares, 5'SCy e 3'SCy; e (ii) as sequências conservadas entre os flavivírus transmitidos por mosquito 5'CYC e 3'CYC (Figura 17).

As estruturas secundárias peculiares das regiões não codificantes também puderam ser identificadas na estrutura de ciclização dos flavivírus brasileiros, tais como: na 5'RNC a estrutura secundária da 5' (5'ES), próximo ao terminal 5', e o grampo na região do gene C (5'CapHS), localizada entre o códon de iniciação e a 5'CYC; e na 3'RNC, o 3'SL, observando na alça distal a presença do pentanucleotídeo 5'-CACAG-3'. Ademais, as sequências conservadas da 3'RNC – CS2, RCS2 e CS3 – foram representadas nas estruturas secundárias, observando a formação de grampos, com exceção da RCS2 do VIGU (Figura 17).

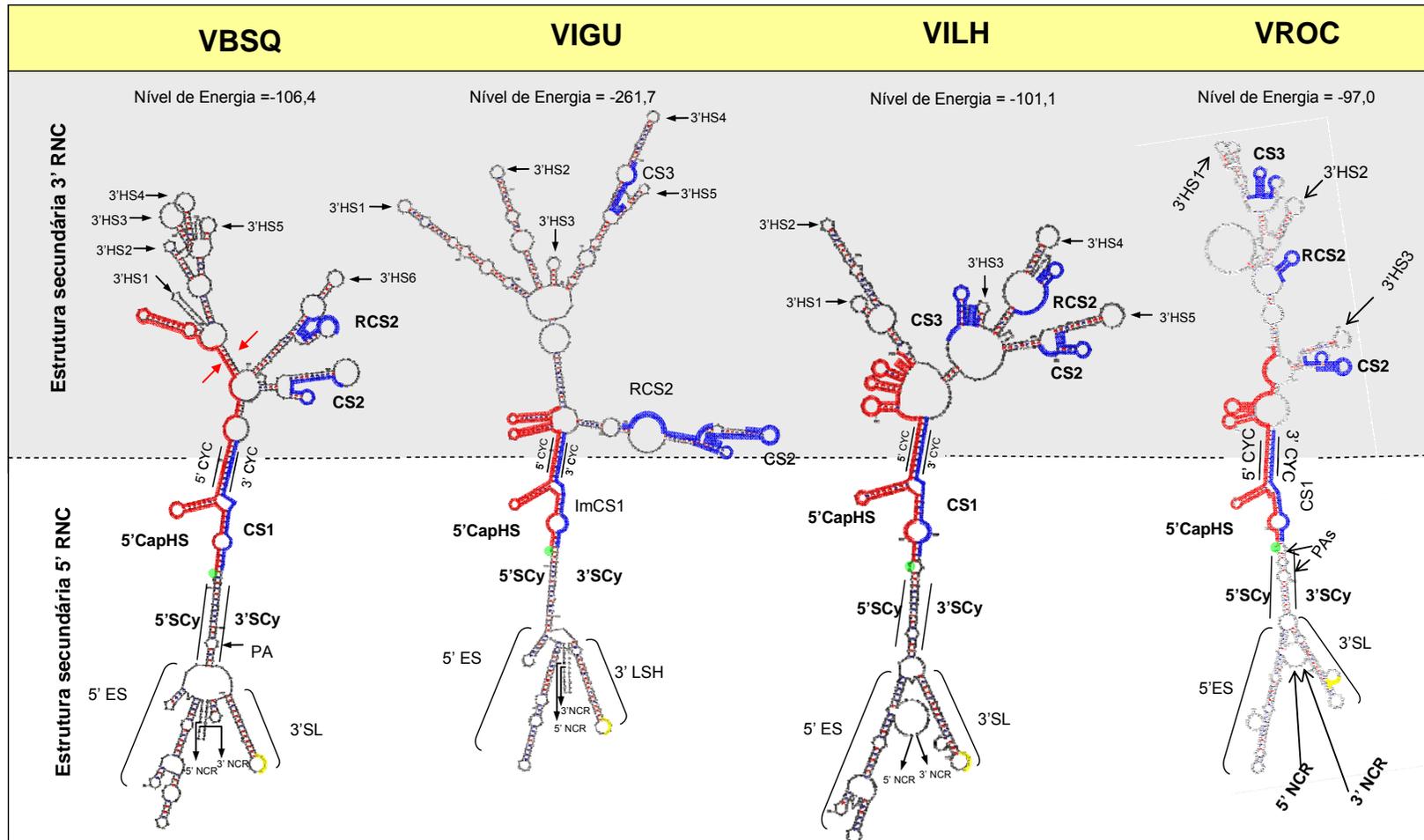


Figura 17 – Estrutura secundária das 5'RNC e 3'RNC dos flavivírus VBSQ, VIGU, VILH e VROC na ciclagem do RNA viral. Em destaque região do gene C (vermelho), as sequências conservadas da 3'RNC (azul), o códon de iniciação (verde) as 5'CYC e 3'CYC pareadas, e o penta nucleotídeo CACAG (amarelo)

Na região que antecede o códon de iniciação da ORF dos flavivírus, encontram-se, em posições análogas do genoma dos flavivírus VBSQ, VIGU, VILH e VROC, as sequências complementares, 5'CS e 3'CS, com tamanhos variando entre 15 a 17 nt, sendo, porém não conservadas. A complementaridade entre as 5'SCy-3'SCy nos VBSQ e VROC não são completas, observando a formação de pequenas alças (PA). No VBSQ, esse fato é determinado pela inserção de duas bases, C e G, na 3'CS que é complementar à adenina na posição 91 da 5'CS; enquanto que no VROC observa-se a não complementaridade entre duas bases consecutivas, nas posições 86 e 87 (referente à região 5'CS) (Figura 18).

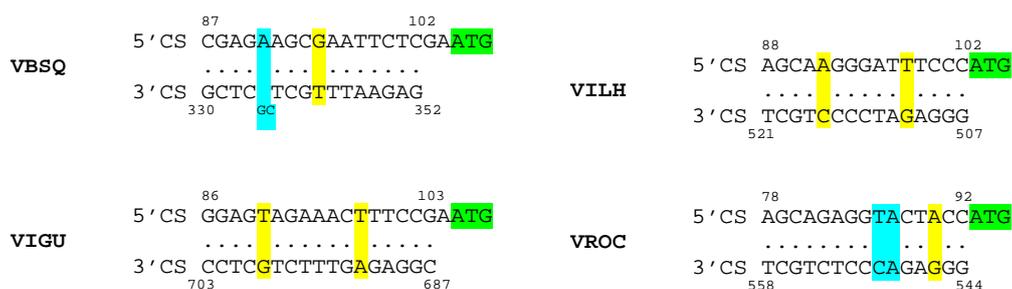


Figura 18 – Sequências complementares entre as 5'RNC e 3'RNC (5'CS-3'CS) dos flavivírus Bussuquara (VBSQ), Iguape (VIGU), Ilhéus (VILH) e Rocio (VROC). **Legenda:** verde = códon de iniciação; amarelo = bases não pareadas; azul = bases que formam pequenas alças.

Na porção 5', a sequência conservada 5'CYC, localizada dentro do gene C e complementar à região 3'CYC (5'-CATATTGA-3'), apresenta-se conservada nos VBSQ e VILH, porém no VIGU observa-se uma substituição no nucleotídeo 148(A→G), cujo nucleotídeo correspondente na 3'CYC (5'-CATACTGA-3' Figura 16) também foi alterado de T→C, permanecendo a complementaridade completa entre essas sequências. No VROC, percebe-se duas substituições em duas posições: uma na posição análoga ao nucleotídeo

148 do VIGU, 134(A→C), e outra no nucleotídeo 137(T→C) (Figura 29), muito embora a 3'CYC tenha se mantido conservada (Figura 16).

Aparentemente, uma terceira sequência de complementaridade entre a região do capsídio e a 3'RNC foi observada na estrutura de ciclização do VBSQ (seta vermelha, Figura 17), sendo visualizada em todos os modelos gerados pelo programa MFOLD. Essa região corresponde a uma sequência de sete nucleotídeos: 5'-GGCGGCC-3' (gene C, posição 167 a 170) complementar a 5'-GGCCGCC-3' (3'RNC, posição 10.536 a 10.542).

Consenso	5'CYC TCAATATG	Posição
VBSQ	145-152
VIGU	...G....	145-152
VILH	133-140
VROC	...C..C.	131-138
VALFUY	...G....	136-143
VEJ	136-143
VUSU	137-144
VEMV	136-143
VNO	137-144
VKUN	137-144
VESL	139-146
VZIK	147-154
VBGA	138-145
VKOK	124-131
VKED	141-148

Figura 19 - Alinhamento entre as sequências nucleotídicas correspondente a 5'CYC (5'RNC) dos flavivírus transmitidos por mosquitos do gênero *Culex* comparando com as sequências correspondentes dos VBSQ, VIGU, VILH e VROC. Em destaque as posições de substituição nucleotídica observadas nos vírus VIGU e VROC. **Legenda:** VALFUY = *Vírus alfuy*; VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*; VESL = *Vírus encefalite Saint Louis*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VKUN = *Vírus kunjí*; VNO = *Vírus Nilo ocidental*; VROC = *Vírus Rocio*; VUSU = *Vírus usutu*; VZIK = *vírus zika*.

3.3.. ANÁLISE DA POLIPROTEÍNA

3.3.1. Identificação dos sítios de clivagem, sítios de glicosilação e resíduos de cisteína

As sequências dos cinco aminoácidos correspondentes aos sítios de clivagens dos flavivírus brasileiros foram comparadas aos respectivos sítios dos flavivírus VESL e VNO, do complexo do da Encefalite Japonesa e três flavivírus transmitidos por mosquitos procedentes da África, os vírus Bagaza (VBAG), Kedougou (VKED) e Zika (VZIK) (Quadro 9). Uma ou duas alterações de aminoácidos foram observadas, em comparação à sequência consenso, sendo mais evidente entre as junções VirC/AnchC, enquanto que os sítios de clivagem entre M/E, NS2A/NS2B, NS2B/NS3 e NS4A/2K foram mais conservados.

O número de sítios em potencial para glicosilação N-ligantes (N-LGlyS) ao nível das proteínas prM, E e NS1 foram diferentes para os quatros flavivírus brasileiros, correspondendo, respectivamente: 2,0,2 para VBSQ; 0,1,3 para VIGU; 1,1,2 para VILH e 1,1,2 para VROC. Todos os sítios foram encontrados nos domínios hidrofóbicos das referidas proteínas (Tabela 1).

Seis resíduos de Cys foram encontrados no domínio N-terminal do precursor (Pr) da proteína de membrana e 12 no ectodomínio das proteínas E e NS1.

Quadro 9 – Sítios de clivagem dos flavivírus VBSQ, VIGU, VILH e VROC em comparação com membros do complexo da encefalite japonesa (VSLE e VNO) e três flavivírus transmitidos por mosquitos procedentes da África (VBAG, VKED e VZIK).

Vírus	Sítios de clivagem da poliproteína					
	VirC/Anch C	Anch C/PR	Pr/M	M/E	E/NS1	NS1/NS2A
VBSQ	NK G KR / SV E SS	CLAL G / FK V V T	K R SRR / NAVFP	APAY N / MRC I G	INV S A / DTG C V	SKV S A / GD G NE
VIGU	NR K KK / RGL H G	PIT L C / FQ T GN	RR S KR / E T HLP	APAY S / LR C IG	LG V GA / D S G C G	SKV S A / GK G SP
VILH	R K E K K / K S F S T	TAV A G / L K ISS	RR G RR / A I NIP	APAY S / LN C L G	V N V H A / DT G CA	SKV S A / G N G Q T
VROC	R K AKR / G N GS V	T G S M A / L R L G T	RR G RR / SV N IP	APAY S / IN C L G	M N V H A / DT G CA	SKV T A / G T G N D
VNO	K Q K K R / G G K T G	AS V GA / VT L SN	RR S RR / SL T V Q	APAY S / FN C L G	V N V H A / DT G CA	S Q V N A / Y N AD M
VESL	P S K K R / G G TR S	GL A SS / L Q L S T	RR S RR / S I S V Q	APAY S / FN C L G	TS V Q A / D S G C A	SR V T A / G V AG G
VBAG	G K K K R / G G TT V	G V A Q A / I K I G S	RR S RR / S I T V H	APAY S / FN C L G	T N V H A / DT G CA	SR V T A / Y D G A G
VKED	N K R K R / S P V N W	G V L T A / V K I G D	RR S RR / S V S L P	APAY S / I R C I G	TS V N G / D Q G C A	AR V S A / G S G S G
VZIK	ER K RR / G A D T S	TT A M A / A E I T R	RR S RR / A V T L P	APAY S / I R C I G	T A V S A / D V G C S	SM V T A / G S T D H
Consenso	NKKKK / GGXTS	XXXAXA / XXXIXS	RRSRR / SSXXXP	APAYS / XNCLG	TXVHA / DTGCA	SKVXA / GXGXG

Vírus	NS2A/NS2B	NS2B/ NS3	NS3/NS4A	NS4A/2K	2K/NS4B	NS4B/NS5
VBSQ	K H G K R / S W P P S	K S G R R / A G AL W	G G R R S / V T T G L	PE K Q R / S I Q D N	G G I A A / N E M G M	N T G K R / G G A D G
VIGU	N L K K R / S W P A S	K S G R R / A G AL W	S G R R S / G V P S I	PE K Q R / T I Q D N	L A V A A / N E M G Y	P N K R G / G G A D G
VILH	S L G K R / G W P A S	K I H K R / G G V M W	A A G K R / S A G V S	PE K Q R / S Q T D N	G A V S A / N E M G W	P K L K R / G G G S A
VROC	C A T K R / G W P A S	K I H K R / G G V L W	A A G K R / S A G S M	PE K Q R / S Q T D N	S A V S A / N E M G W	P K V K R / G G I A A
VNO	P N R K R / G W P A T	Q Y T K R / G G V L W	A S G K R / S Q I G L	PE K Q R / S Q T D N	S A V A A / N E M G W	P G L K R / G G A K G
VESL	P N G K R / S W P A S	K H S K R / G G AL W	A A G K R / S A L G M	PE K Q R / S Q T D N	G V V A A / N E M G L	P K G K R / G G G K G
VBAG	P S N R R / G W P V S	H S P K R / S G A I W	A C G K R / S A I G V	PER Q R / S Q T D S	G T V A S / N E M G W	G S M R R / G G G K G
VKED	G E K R R / S W P P T	K K A K R / S G AL W	A A G R R / G A T A G	PER Q R / S V Q D N	G L V A A / N E A G L	S S T K R / S N R G L
VZIK	R S G K R / S W P P S	K T G K R / S G AL W	A A G K R / G A A L G	PE K Q R / S P Q D N	G L I T A / N E L G W	L V K R R / G G G T G
Consenso	PXKKR / SWPAS	KSGKR / GGALW	AAGKR / SAXGX	PEKQR / SQTDN	GAVAA / NELGW	PXKKR / GGGKG

Legenda: Em destaque estão marcados os aminoácidos que diferem da sequência consenso; VirC =Capsídeo; AnchC = peptídeo-sinal no domínio hidrofóbico terminal; prM=precursores da Proteína M; E = proteína de envelope; NS=proteínas não-estrutural; x=qualquer aminoácido. VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VESL = *Vírus encefalite Saint Louis*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VNO = *Vírus Nilo ocidental*; VROC = *Vírus rocio*; VZIK = *vírus zika*.

Tabela 1 – Sítios de glicosilação determinado para as proteínas PrM, E e NS1 pelo programa NetNglyc para os flavivírus VBSQ, VIGU, VILH e VROC O valor limítrofe considerado é de 0,5.

Vírus	Proteína	Posição*	e-valor	
VBSQ	PrM	31	NNTC	0,6650
		148	NRSQ	0,5066
	E	-	-	-
	NS1	130	NATF	0,5148
		207	NLTW	0,6152
VIGU	PrM	-	-	-
	E	154	NDTS	0,7330
	NS1	85	NLSV	0,5899
		130	NNTF	0,5187
		147	NRTW	0,6178
VILH	PrM	15	NKTD	0,8014
	E	154	NYTA	0,7407
	NS1	130	NSTF	0,6133
		316	NCTL	0,5131
VROC	PrM	15	NKTD	0,7059
	E	154	NYTT	0,7133
	NS1	106	NTSD	0,6348
		130	NSTF	0,6186

Legenda: * em relação ao primeiro aminoácido da proteína; em negrito o aminoácido Asparagina (N) que é glicosilado.

3.3.1. Identificação de motivos conservados nas proteínas E, NS3 e NS5

Dentre do banco de dados do Interproscan, foi escolhido os parâmetros Pfam para determinação dos domínios nas sequências das proteínas E, NS3 e NS5, muito embora os demais bancos tenham apresentado resultados similares. Nas referidas sequências aminoacídicas dos flavivírus VBSQ, VIGU, VILH e VROC foram encontrados: na proteína E os domínios central e dimerização (ectodomínio), que representa o domínio II; e o domínio III, que apresenta estrutura similar à imunoglobulina; na proteína NS3, os domínios da protease serina e helicase (DEAD); e na proteína NS5, os domínios MTase e RdRP. As localizações de cada domínio estão descritos na Tabela 2.

Tabela 2 - Principais domínios das proteínas E, NS3 e NS5 determinado pelo Pfam para os flavivírus VBSQ, VIGU, VILH e VROC.

Vírus	Proteína	Domínio	Posição*	Tamanho	Sítios ativos Resíduo/posição	e-value
VBSQ	E	Central e dimerização	1-301	300 aa	-	6,1e-197
		Domínio III	303-399	96 aa	-	6,5e-54
	NS3	Protease Serina	7-177	170 aa	H51, D75, S135	1,5e-104
		DEAD (Helicase)	184-331	147 aa	-	4,5e-102
	NS5	MTase	54-277	223 aa	-	2,7e-05
		RdRP	251-900	649 aa	-	0
VIGU	E	Central e dimerização	1-294	293 aa	-	4,3e-202
		Domínio III	296-392	96 aa	-	1e-46
	NS3	Protease Serina	6-173	167 aa	H50, D74, S134	1,7e-90
		DEAD (Helicase)	183-330	147 aa	-	4,6e-101
	NS5	MTase	54-228	174 aa	-	8,6e-05
		RdRP	252-899	647 aa	-	0
VILH	E	Central e dimerização	1-299	298 aa	-	7,4e-216
		Domínio III	301-399	98 aa	-	2,1e-53
	NS3	Protease Serina	7-179	172 aa	H51, D75, S135	9,4e-101
		DEAD (Helicase)	184-331	147 aa	-	1,1e-103
	NS5	MTase	54-227	173 aa	-	1,9e-05
		RdRP	252-900	648 aa	-	0
VROC	E	Central e dimerização	1-299	298 aa	-	4,1e-196
		Domínio III	301-399	98 aa	-	9,7e-53
	NS3	Protease Serina	7-178	171 aa	H51, D75, S135	2,4e-99
		DEAD (Helicase)	184-331	147 aa	-	1,4e-106
	NS5	MTase	54-227	172 aa	-	1,1e-05
		RdRP	252-900	648 aa	-	0

* Em relação ao primeiro códon da proteína NS5.

A região do ectodomínio da proteína E (domínio II), constituída pelos domínios central e de dimerização, apresentaram tamanhos de 300 aa, 293 aa, 298 aa e 298 aa para os flavivírus VBSQ, VIGU, VILH e VROC, incluindo a sequência conservada relativa ao peptídeo de fusão, DRGWGNGCGLFGKGS (posições 99 a 114), para os flavivírus transmitidos por *Culex* sp. (Seligman, 2008). Nos VBSQ e VIGU foram observadas três e duas alterações, respectivamente, em relação à sequência consenso, como segue: VBSQ: **NRGW**NNGCGLFGK**GDI** (D99N; C103N; S113D); VIGU: DRGW**N**NNGCGLFGK**GSL** (C103N; I114L). Dentre os seis resíduos de Gli existentes, os VBSQ e VIGU apresentaram somente cinco em virtude da mutação na posição 103, reconhecida por C103N. O VILH apresentou a sequência do peptídeo de fusão similar a sequência consenso, enquanto que para o VROC foi observada uma única substituição nucleotídica referente a I114L.

O domínio III apresenta 96 resíduos para VBSQ e VIGU e 98 para VILH e VROC. O motivo RGD reconhecido na proteína E dos VALFUY (387-389), VEJ (387-389), VEMV (388-390) e VUSU (387-389), correspondem aos VBSQ, VIGU, VILH e VROC, ao tripeptídeo VGD (388-390), SYD (381-383), QEN (388-390) e TGP (388-390), respectivamente.

Na proteína NS3, os sítios catalíticos da tríade H-D-S (*His-Asp-Ser*) do domínio catalítico da protease serina estão preservados nos quatro flavivírus em estudo, relacionados às posições 51, 75 e 135, respectivamente, para VBSQ, VILH e VROC e posições 50, 74, 134 para o VIGU (Tabela 2). Ainda no domínio da protease serina, o motivo conservado XXGLYGNG que está relacionado ao sulco de ligação do substrato (Chang, 1997), é

representado pela sequência VIGLYGNG para o VROC (posições 145 a 152) e VBSQ (posições 145 a 152) bem como para os VKUN, VZIK, VUSU e VNO, mas apresentam duas substituições nucleotídicas nas duas primeiras posições do motivo para o VIGU (**T**IGLYGNG) e VILH (**I**TGLYGNG).

Os motivos conservados entre os flavivírus associados à família DEAD da RNA helicase [GAGKTRK (motivo I), MDEAH (motivo II) e SAAQRRGR(V/I)GR (motivo VI) (Gorbalenya et al., 1989; Matusan et al., 2001)], se mantiveram conservados nos quatro flavivírus em estudo.

Na proteína NS5, dentre os oito motivos determinantes da atividade RdPd no genoma dos flavivírus: motivo I (MGKREKCLG), motivo II (KAKGSRAIWLGARFLEF), motivo III* [(G/A/S)GVEGXGI(Q/H)**X**LG Y (V/I)LX(E/D)**X**], motivo IV (**X**GG**X**IYADD), motivo V* [**X**DQRGSGQV(V/G)TY (A/G)LNTFTN**X**(V/A)QL(V/I)**RX**(M/A)], motivo VI [R(M/L)(A/L)ISGDDCVV], motivo VII (VPFCS(H/N)HFX), motivo VIII (RX(L/I)VVPCRQDEL) (Koonin, 1991), somente os motivos I, II, III, IV e VIII apresentam-se conservados para os quatro flavivírus em estudo, bem como os sítios da MTase: K₆₁, D₁₄₆, K₁₈₂ e E₂₁₈ (posições referentes ao VNO) (Zhou et al., 2007).

No motivo V, o VROC apresenta uma mudança de aminoácido na posição 614, Y→F (NDQRGSGQVVTYALNTY⁶¹⁴TNLAVQLIRCM). Ainda no motivo V, o VILH apresenta uma mudança de aminoácido na posição 619, V→A, (EDQRGSGQVVTYALNTFTNLAA⁶¹⁹QLVRCM). No motivo VI uma única mudança aminoacídica R→P, na posição 661 (**P**⁶⁶¹MAVSGDDCVV). No motivo VII, o VILH apresenta uma mudança de C→G na posição 713 (VPFG⁷¹³SHHFN). Para os demais flavivírus em estudo os motivos V, VI e VII apresentam-se conservados.

* Sequência consenso baseada apenas nos flavivírus transmitidos por mosquitos; **X** = locus com variação de aminoácido.

O motivo V(V/I)DLGCGRGGW associado à atividade MTase-SAM (Koonin, 1993), permanece conservado.

3.4. ANÁLISE DE SIMILARIDADE

O grau de relação filogenética das sequências nucleotídicas entre os quatro flavivírus brasileiros foi demonstrado pelo método de plotagem de similaridade utilizando o valor de permutação de 50% como valor limiar com relação ao eixo Y, que corresponde ao percentual de permutação filogenética (ppt) baseado nos valores de *Bootstrap*, e o eixo X representa a posição no genoma dos flavivírus. As análises tiveram VBSQ e VILH como parâmetro de referência em relação VBAG, VIGU, VESL, VKED, VKOK, VROC VNO e VZIK os quais mostram maior similaridade filogenética entre si (Kuno et al., 1998; Gubler et al., 2007; Kuno et al., 2009).

Para a ORF do VBSQ observou-se uma maior similaridade com o VIGU, mostrando um forte sinal para os genes C, PrM, porção inicial e N terminal da proteína E, NS4A e NSAB; e de moderada a forte para as regiões MTase e RdRp no gene NS5. Próximo a região C-terminal da proteína E, há uma maior similaridade com o VKOK (ppt>85%). Para o gene NS1, a porção central é similar com o VKOK (intensidade moderada; ppt=80%) e a porção N-terminal com o VKED (intensidade forte; ppt>85%). A região da protease no gene NS3 apresenta forte similaridade com o VZIK (ppt>85%) e a região N-terminal com o VKOK (Figura 20 A).

Para o VILH, de modo geral, um forte sinal de similaridade com o VROC foi observado para os genes C-PrM-E, NS2A-2B, NS3 e NS4, muito

embora as regiões C-terminal da NS1 e N-terminal da PrM apresentem similaridade forte (ppt>90%) e moderada (ppt>70%) com os VNO e VESL, respectivamente (Figura 20 B).

Para verificar a relação do VROC com os demais flavivírus, excetuando o VILH, foi realizada a análise de similaridade genética pelo método de *Bootscan*, tendo o VROC como referência e em comparação com os vírus do grupo da encefalite japonesa e com o representante do grupo Ntaya, o VBAG. No decorrer da ORF, o VROC obteve maior similaridade com os VBAG e VESL. Para os genes C e NS2A, o VROC apresenta similaridade moderada a intensa com o VESL, enquanto que para os genes NS2B, NS4B (moderada a intensa) e PrM (baixa) apresentam maior similaridade com o VBAG. Os genes E, NS1, NS3, NS4A e NS5, a similaridade do VROC em relação ao VESL e VBAG foi mais balanceada, mas tendendo a maior similaridade, quanto à intensidade, com o VBAG (Figura 21)

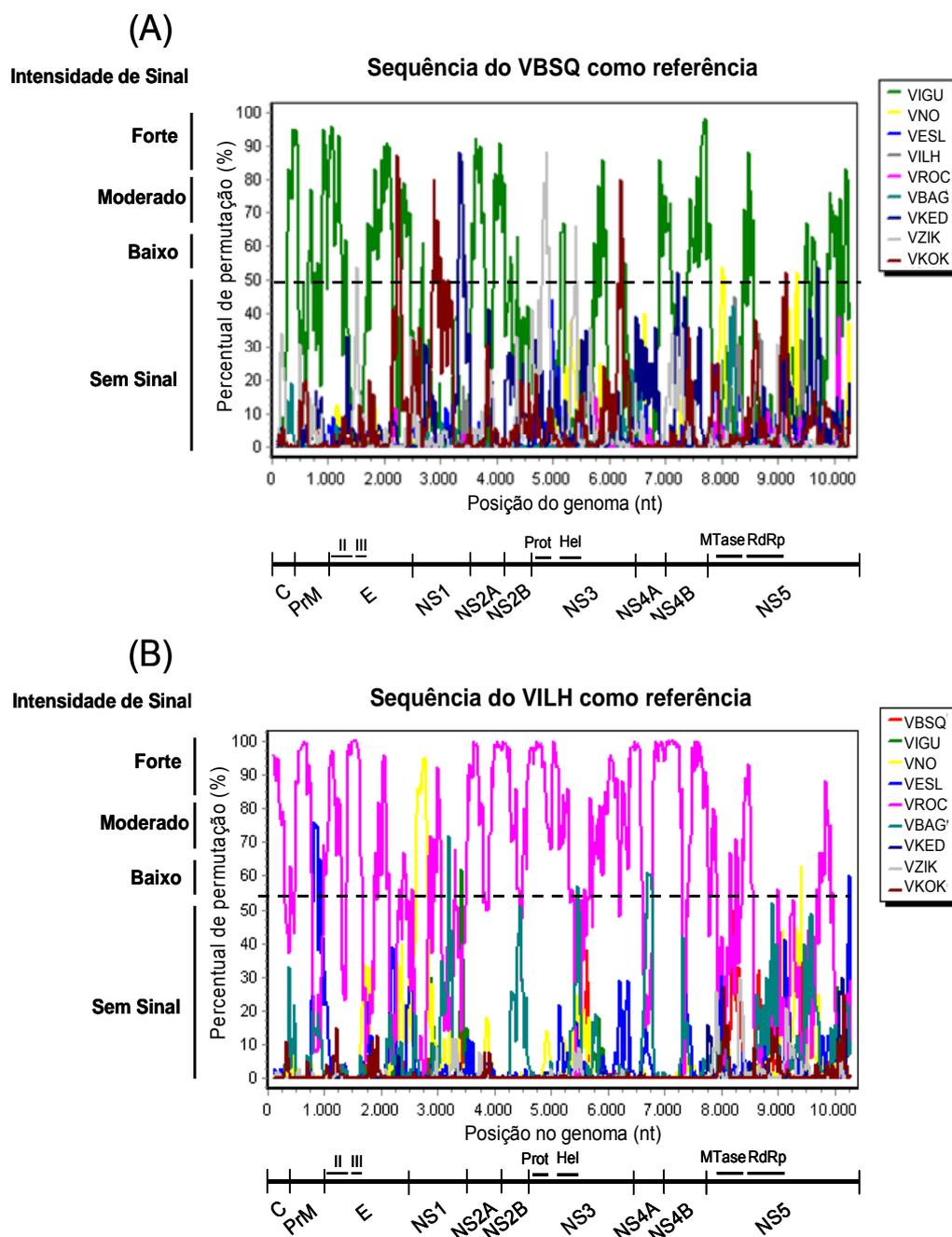


Figura 20 – Análise de similaridade (*Bootscan*) realizada pelo programa SimPlot, utilizando as sequências dos VBSQ (A) e VILH (B) como referência, O eixo Y representa percentual de permutação filogenética (ppt); o eixo X corresponde às diferentes regiões da ORF, O parâmetro vertical expressa às intensidades da relação filogenética (Forte, Moderado, Baixo e sem sinal de relação filogenética), tendo o valor de *bootstrap* igual a 50% corresponde ao valor limiar, **Legenda:** Prot= região da protease; Hel=região da Helicase; MTase=região da metiltransferase; RdRp = região da Polimerase. VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VESL = *Vírus encefalite Saint Louis*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VNO = vírus do Nilo Ocidental; VROC = *Vírus rocio*; VZIK = *vírus zika*

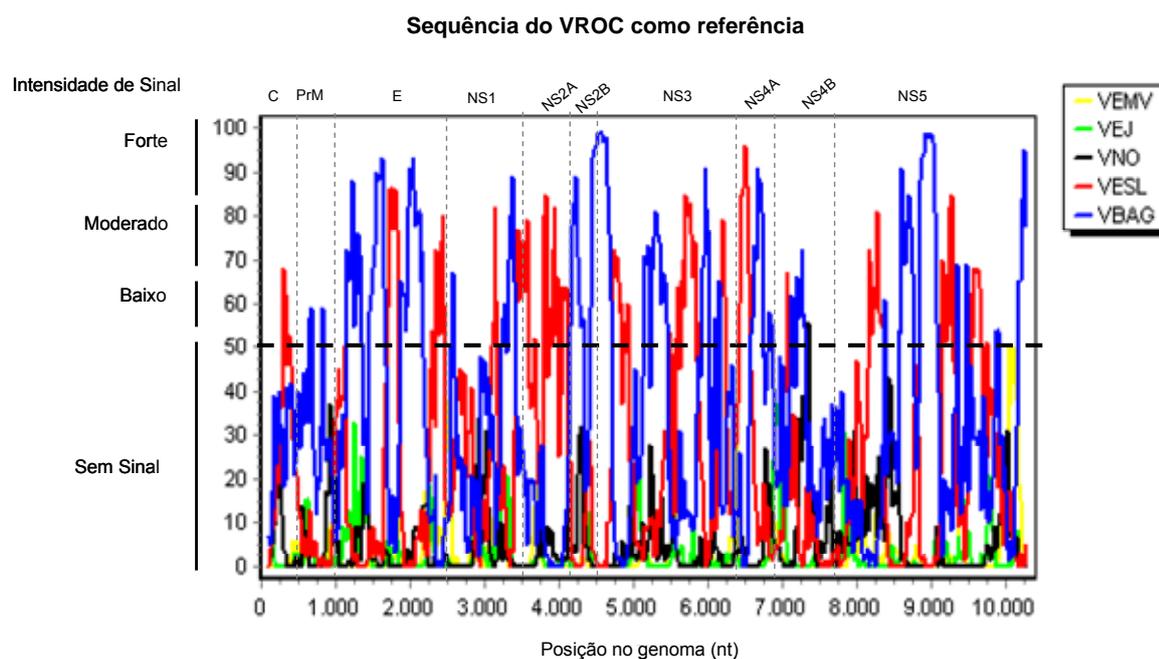


Figura 21 – Análise de similaridade (*Bootscan*), utilizando o VROC como referência. Notar a maior similaridade com o VBAG e VESL para toda a ORF. **Legenda:** VBAG = *Vírus Bagaza*; VEJ = *Vírus da encefalite japonesa*; VESL = *Vírus encefalite Saint Louis*; VNO = *Vírus Nilo Ocidental*; VROC = *Vírus rocio*.

O percentual de identidade aminoacídica entre os flavivírus brasileiros VBSQ, VIGU, VILH e VROC, e os flavivírus encefalitogênicos dos grupos da Encefalite Japonesa (VJE, VWN e VESL), Ntaya (VBAG), Kedougou (VKED), Kokobera (VKOK) e Spondweni (VZIK) foi realizado pelo programa MegAlign (DNASTar) comparando-se tanto a ORF quanto a identidade de cada gene. Nota-se que os maiores percentuais de identidade entre os vírus transmitidos por *Culex* foram para as proteínas NS5 e NS3, enquanto que os menores percentuais foi com relação à proteína C.

Entre VBSQ e VIGU a maior identidade aminoacídica foi observada para os genes NS2A, NS2B, NS4B e NS5, e para a ORF. Quanto aos genes C, PrM, NS1, NS3 e NS4A, o VBSQ apresentou maior similaridade com VKOK e VKED, e em maior percentual com o VZIK. O VIGU, para os

genes C, PrM, E, NS1 e NS4A, apresenta maior similaridade com VKOK, VESL, VKOK, e VROC, respectivamente (Tabela 3).

Observou-se que entre VILH e VROC tanto a ORF quanto os genes estruturais (C-PrM-E) e os não estruturais (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) apresentaram maior percentual de identidade entre si do que com qualquer outro flavivírus. Para a ORF, a identidade aminoacídica entre VILH e VROC foi de 77,5% (71% para nucleotídeo). Dentre as proteínas, os maiores percentuais de identidade entre os VILH e VROC foram para as proteínas E (78,6%), NS2B (85,6%), NS3 (85,4%) e NS5 (78,5%). A proteína C foi a que apresentou o menor percentual de identidade (55,1%), seguido da PrM (67,1%) e NS2A (70%).

Quanto aos demais grupos, o VILH mostrou maior similaridade com o grupo da encefalite japonesa para os genes C, E, NS1, NS2A e NS5, e para a ORF, e grupo Ntaya para os genes PrM, NS2B, NS3, NS4A e NS4B. Para o VROC maior similaridade deu-se para o grupo da encefalite japonesa quanto aos genes C, PrM, NS2A, NS3, NS4A, NS4B e NS5, e para a ORF, e quanto aos genes E, NS1 e NS2B para o grupo Ntaya (Tabela 3).

Tabela 3 – Percentual de identidade aminoacídica entre os flavivírus VBSQ, VIGU, VILH e VROC e outros flavivírus transmitidos por *Culex*.

VÍRUS BUSSUQUARA											
	C	PrM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	ORF
VIGU	35,2	48,5	57,5	56,8	46,7	62,6	65,4	40,7	56,5	72,7	60,7
VILH	38,4	41,3	52,9	56,5	34,4	39,7	64,9	37,3	43,9	70,9	56,2
VROC	46,4	47,9	52,1	56,8	37,4	38,2	64,7	42,7	44,7	69,0	56,8
VEJ	38,4	47,3	54,9	54,5	38,3	38,2	63,9	46,0	45,5	68,1	56,5
VNO	44,0	48,5	53,9	53,7	38,3	37,4	65,5	45,3	44,3	69,6	57,1
VESL	36,8	50,3	54,9	53,4	36,6	36,6	63,3	45,3	50,2	70,3	56,9
VBAG	40,0	46,1	50,7	54,0	33,9	36,6	64,1	42,7	45,5	67,6	55,6
VKED	33,6	50,3	53,9	56,5	32,2	46,6	62,8	36,7	53,0	64,8	55,5
VKOK	46,4	46,7	55,1	56,8	38,3	45,0	61,7	37,3	52,2	68,0	57,3
VZIK	39,2	45,5	53,5	57,7	39,2	55,0	67,6	44,0	56,5	68,0	58,7
VÍRUS IGUAPE											
	C	PrM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	ORF
VBSQ	35,2	48,5	57,5	56,8	46,7	62,6	65,4	37,3	56,5	72,7	60,7
VILH	42,9	40,1	54,7	56,9	35,2	42,4	61,8	41,6	44,5	66,0	55,3
VROC	41,2	47,3	51,2	56,9	37,9	41,7	63,2	52,4	42,5	65,8	55,8
VEJ	42,9	42,5	55,9	56,4	41,9	37,9	62,6	47,6	45,3	68,6	56,9
VNO	47,9	50,9	56,5	57,2	35,2	37,1	62,9	47,6	49,6	68,5	57,2
VESL	42,0	45,5	55,3	58,4	34,4	37,9	62,9	50,8	51,2	68,3	57,4
VBAG	46,2	47,9	54,3	53,8	33,5	33,3	61,8	50,8	43,3	66,0	55,5
VKED	35,3	45,5	51,6	56,7	31,7	41,7	64,0	51,6	51,2	64,9	55,0
VKOK	52,1	41,9	60,1	57,8	35,7	47,7	60,8	45,2	50,8	67,0	57,4
VZIK	37,8	41,3	51,2	56,4	38,3	54,5	65,3	42,1	52,0	67,0	56,9
VÍRUS ILHÉUS											
	C	PrM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	ORF
VBSQ	38,4	41,3	57,5	56,5	34,4	39,7	64,9	37,3	43,9	70,9	56,2
VIGU	42,9	40,1	54,7	56,9	35,2	42,4	61,8	41,6	44,5	66,0	55,3
VROC	55,1	67,1	78,6	73,4	70,0	85,6	85,4	77,9	72,9	78,5	77,5
VEJ	48,3	50,9	64,5	62,0	45,4	51,5	71,4	57,7	51,4	72,8	63,8
VNO	49,2	53,3	64,3	64,9	49,3	50,0	70,9	55,7	51,4	74,0	65,7
VESL	45,8	41,9	66,1	62,0	47,6	44,7	71,4	59,1	56,1	76,2	64,5
VBAG	39,8	54,5	64,7	53,3	42,7	53,0	72,0	60,4	56,5	73,6	64,4
VKED	39,8	41,9	52,1	61,2	27,8	35,6	63,8	34,2	43,5	65,7	53,5
VKOK	44,1	44,3	60,1	55,5	30,0	34,1	60,5	30,9	45,5	67,2	55,2
VZIK	41,5	40,7	54,3	55,8	34,4	46,2	68,6	40,9	45,1	69,7	57,3
VÍRUS ROCIO											
	C	PrM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	ORF
VBSQ	46,4	47,9	52,1	56,8	37,4	38,2	64,7	42,7	44,7	69,0	56,8
VIGU	41,2	47,3	51,2	56,9	37,9	41,7	63,2	52,4	42,5	65,8	55,8
VILH	55,1	67,1	78,6	73,4	70,0	85,6	85,4	77,9	72,9	78,5	77,5
VEJ	48,3	59,3	64,3	62,9	48,5	49,6	71,4	59,1	51,2	71,2	64,6
VNO	52,5	58,7	62,5	63,7	47,6	49,6	72,1	62,4	54,7	71,8	65,0
VESL	50,0	47,9	65,3	62,9	50,7	46,6	70,0	61,1	58,2	75,0	66,1
VBAG	46,6	57,5	64,7	64,9	48,9	50,4	70,6	59,7	53,5	73,9	65,4
VKED	39,8	47,9	50,5	54,7	30,0	31,3	65,8	38,3	44,1	64,2	54,1
VKOK	44,1	51,5	59,1	57,2	31,3	33,6	60,4	34,2	46,9	65,4	55,8
VZIK	42,4	48,5	54,3	56,1	33,9	46,6	68,8	38,9	44,9	67,8	57,4

Legenda: cinza escuro: maior percentual de identidade; Dentro de parêntese são mostrados as identidades nucleotídicas de cada gene e da ORF entre VILH e VROC. VBSQ = *Vírus Bussuquara*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VROC = *Vírus Rocio*; VEJ = *Vírus da encefalite japonesa*; VNO = *Vírus Nilo ocidental*; VESL = *Vírus encefalite Saint Louis*; VBAG = *Vírus Bagaza*; VKOK = *Vírus kokobera*; VKED = *Vírus kedougou*; VZIK = *vírus zika*

3.5. FILOGENIA

O alinhamento entre as sequências nucleotídicas dos 32 flavivírus, incluindo os quatro flavivírus brasileiros em estudo, foi realizado de modo a coincidir os sítios de clivagem de cada gene, o que possibilitou a mesma organização ao se fazer a análise filogenética da ORF, bem como dos genes individuais.

As análises filogenéticas pelos métodos MV e Bayesiano foram realizadas utilizando os quatro flavivírus em estudo (VBSQ, VIGU, VILH e VROC) juntamente com representantes dos quatro grupos principais de flavivírus: flavivírus transmitidos por mosquitos, flavivírus transmitidos por carrapatos, agentes zoonóticos com vetores desconhecidos e flavivírus de insetos, tendo o grupo dos flavivírus de insetos (VCFA e VKR) como grupo externo (Quadro 7, página 98). Primeiramente, foi analisada a ORF e depois foram avaliadas as regiões estrutural e não estrutural, separadamente, e os genes individuais.

A topologia das árvores geradas pelos métodos de MV e Bayesiano foi similar, sendo mostradas as árvores da MV e plotados os valores bayesianos dentro de parênteses.

Os modelos de substituição nucleotídica determinados pelo AIC para cada região do genoma analisada estão mostrados na Tabela 4.

Tabela 4 – Modelo de substituição nucleotídica determinado pelo critério de informação de Akaike (AIC), mostrando os dados para frequência das bases, taxa de matriz, parâmetros de distribuição gamma e proporção de sítios invariáveis.

REGIÃO		MODELO DE SUBSTITUIÇÃO NUCLEOTÍDICA (AIC)												
DO GENOMA	Modelo	AIC	Frequências de base				Taxa de Matriz						Gamma	pinvar
			A	C	G	T	A-C	A-G	A-T	C-G	C-T	G-T		
ORF	GTR+I+G	567.200,06	0,2890	0,2310	0,2710	0,2090	2,1834	2,9118	1,3606	1,3010	5,0231	1	0,8556	0,0889
Estrutural	GTR+I+G	396.648,84	0,2982	0,2265	0,2742	0,2011	2,1450	2,8328	1,4038	1,3553	5,2646	1	1,0139	0,1472
C	GTR+I+G	22.575,54	0,2865	0,2268	0,2873	0,1994	2,0575	2,4500	0,7382	1,0046	3,8151	1	1,2036	0,0174
PrM	GTR+I+G	29.593,50	0,2752	0,2298	0,2766	0,2184	2,6715	2,8679	1,2088	1,0215	4,7975	1	0,9292	0,0405
E	GTR+I+G	82.980,60	0,2868	0,2338	0,2584	0,2209	2,1395	3,2755	1,5763	1,3814	5,6451	1	0,7038	0,0375
Não Estrutural	GTR+I+G	430.515,03	0,2904	0,2306	0,2724	0,2066	2,1940	2,9326	1,3912	1,3233	5,1501	1	0,8574	0,1047
NS1	GTR+I+G	59.114,26	0,3119	0,2251	0,2601	0,2029	0,3119	3,9679	2,0758	1,8510	7,2012	1	0,7801	0,0826
NS2A	GTR+I+G	45.840,84	0,2446	0,2254	0,2923	0,2376	1,9575	2,3360	0,8375	1,1144	3,3479	1	1,2679	0,0079
NS2B	GTR+I+G	24.615,41	0,2527	0,2192	0,2974	0,2307	2,7358	2,3802	1,6002	1,3769	4,1828	1	1,1961	0,0203
NS3	GTR+I+G	94.532,51	0,3058	0,2292	0,2598	0,2052	2,9864	4,3029	1,7208	1,5777	7,2007	1	0,7251	0,1104
NS4A	GTR+G	22.832,62	0,2474	0,2256	0,3010	0,2260	2,1262	2,4339	1,0148	1,2554	4,4887	1	0,7176	0
NS4B	GTR+I+G	49.464,71	0,2600	0,2411	0,2729	0,2260	1,6425	2,5814	1,1524	1,1426	4,0935	1	0,7537	0,0207
NS5	GTR+I+G	132.819,01	0,3164	0,2354	0,2692	0,1790	2,1258	2,9775	1,6953	1,3421	6,2232	1	0,8760	0,1812
3'RNC	GTR+G	189.46,02	0,2577	0,2756	0,2707	0,1959	0,9646	2,6194	2,2307	0,5773	3,2512	1	1,8053	0

3.5.1. Região Codificante

A análise filogenética da região codificante (ORF) de 32 flavivírus, mostrou a formação de quatro grupos, os quais apresentam relação eco-epidemiológica: flavivírus transmitidos por mosquitos, suportado por um valor de *bootstrap* de 100% (VEJ, VUSU, VALFUY, VEMV, VNO, VESL, VROC, VILH, VBAG, VIGU, VBSQ, VKOK, VZIK, VKED, VDEN1, VDEN2, VDEN3, VDEN4, VSEP, VWES, VFA), flavivírus transmitidos por carrapatos, com *bootstrap* de 100% (VLI, VTBE, VOHF, VLGT e VPOW), os agentes zoonóticos com vetores desconhecidos, suportado pelo *bootstrap* de 100% (VMML, VVRB, VMOD e VAPOI) e os flavivírus de insetos, representados pelos VCFA e VKR (*bootstrap* = 100%).

O grupo dos flavivírus transmitidos por mosquitos foi subdividido em nove clados: o clado I dos flavivírus do grupo da encefalite japonesa (VEJ, VUSU, VALFUY, VEMV, VNO, VESL); o clado II que agrupa VILH e VROC; o clado III representado pelo VBAG; clado IV que inclui os VBSQ e VIGU; o clado V, VI e VII representados pelos VKOK, VZIK e VKED, respectivamente; o clado VIII que inclui o grupo do dengue (VDEN1, VDEN2, VDEN3 e VDEN4); o clado IX que inclui os VFA, VSEP e VWES (Figura 22).

Analisando o tropismo viral e vetores, associam-se os clados I, II, III, IV e V, como vírus neurotrópicos transmitidos por *Culex* sp. e os clados VI, VII, VIII e IX com vírus viscerotrópicos transmitidos por *Aedes* sp.

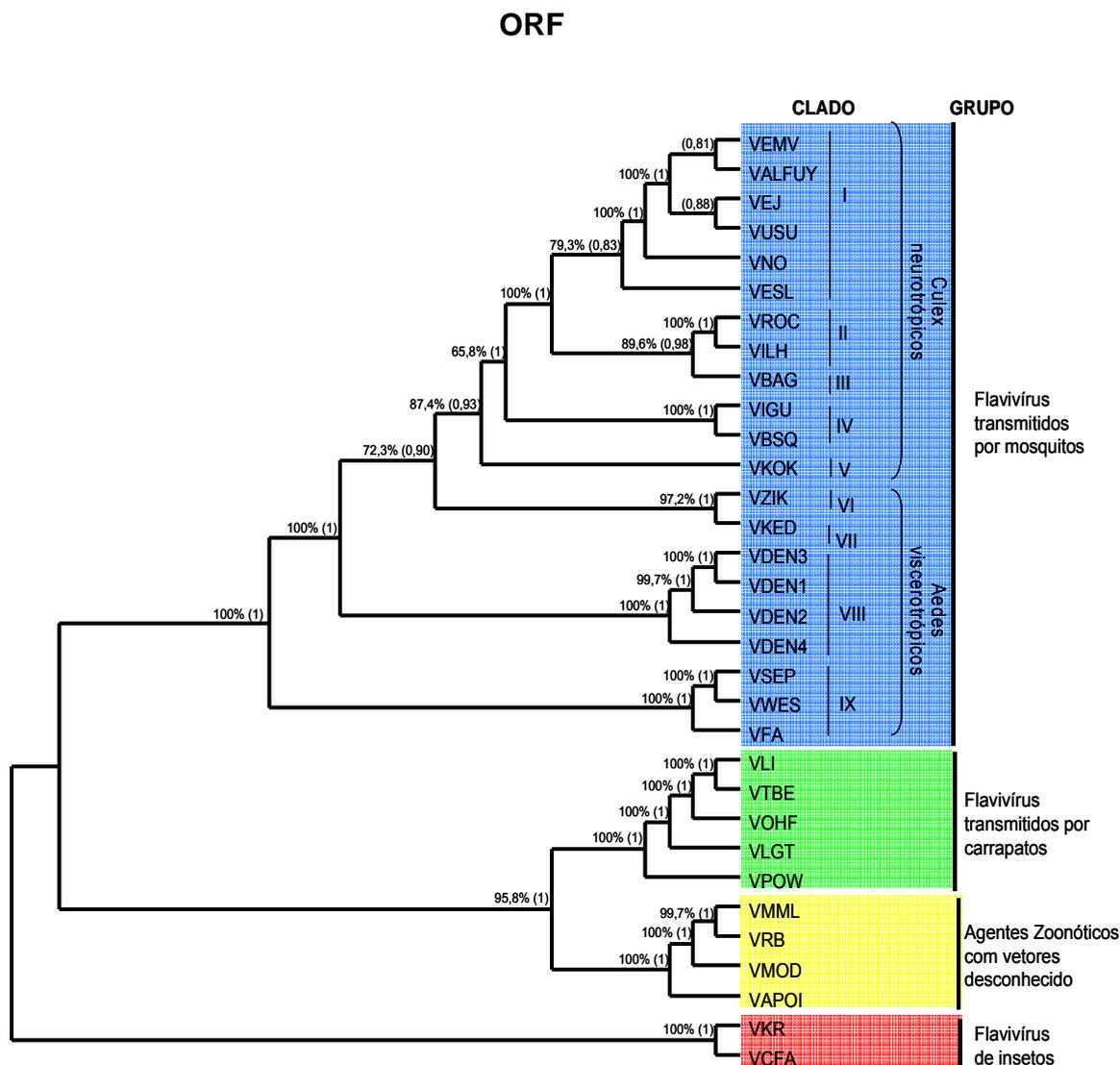


Figura 22 - Árvore filogenética pelos métodos MV das sequências nucleotídica da ORF dos flavivírus. Os valores bayesianos foram plotados dentro de parêntese. Os grupos dos flavivírus transmitidos por mosquitos, flavivírus transmitidos por carrapatos, agentes zoonóticos com vetores desconhecidos e flavivírus de insetos, estão representados pela cores azul, verde, amarelo e vermelho, respectivamente. O grupo dos flavivírus transmitidos por mosquitos foi subdividido em nove clados. **Legenda:** VALFUY = *Vírus Alfuy*, VAPOI = *Vírus apoi*, VBAG = *Vírus Bagaza*, VBSQ = *Vírus Bussuquara*, VCFA = *Vírus Cell fusion agent*, VEJ = *Vírus da encefalite japonesa*, VEMV = *Vírus encefalite Murray Valley*, VESL = *Vírus encefalite Saint Louis*, VFA = *Vírus da febre amarela*, VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*, VLI = *Vírus louping ill*, VMOD = *Vírus modoc*; VMML = *Vírus Montana miotis leukoencephalitis*; VNO = *Vírus do Nilo ocidental*; VOHF = *Vírus Omsk hemorrhagic fever*, VPOW = *Vírus powassan*; VRB = *Vírus Rio Bravo*; VROC = *Vírus Rocio*; VSEP = *Vírus Sepike*; VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*, VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

3.5.2. Região estrutural

A análise da região estrutural, compreendendo os genes C-PrM-E, mostrou topologia similar para a formação dos quatro grupos principais dos flavivírus, assim como observado na análise da ORF. No grupo dos flavivírus transmitidos por mosquitos, nove clados foram reconhecidos: o clado I, o qual inclui os vírus da encefalite japonesa (*bootstrap* de 76,6%); o clado II formado pelos VILH e VROC (*bootstrap* de 100%), que aparenta ser mais relacionado ao grupo da encefalite japonesa (*bootstrap* de 91,3%); o clado III representado pelo VBAG, do grupo Ntaya (*bootstrap* de 100%); o clado V formado pelos flavivírus brasileiros VBSQ e VIGU (*bootstrap* de 91,3%); o clado VI formado pelos quatro sorotipos do VDEN (*bootstrap* de 100%); e associados a este os clados VII e VIII, representado pelo VZIK e VKED, respectivamente; e o clado IX formado pelo grupo da febre amarela (*bootstrap* de 100%) (Figura 23).

A análise individual dos genes estruturais mostrou a conservação da topologia dos quatro principais grupos dos flavivírus, no entanto a organização dos clados do grupo dos flavivírus transmitidos por mosquitos para os genes C e PrM apresentou topologia distinta da apresentada para a região estrutural e por vezes apresentando valores de *bootstrap* baixos para suportar a inclusão dos vírus em uma clado, porém com topologia similar ao observado na análise da região estrutural (Figura 23). Por sua vez, o gene E, apresentou valores de *bootstraps* mais elevados para a formação dos grupos e clados, porém discordando da análise da região estrutural apenas com relação aos VBAG, VILH e VROC, que neste caso, foram agrupados no mesmo clado, porém suportados por valores de *bootstrap* inferiores a 70% (Figura 24).

Região Estrutural

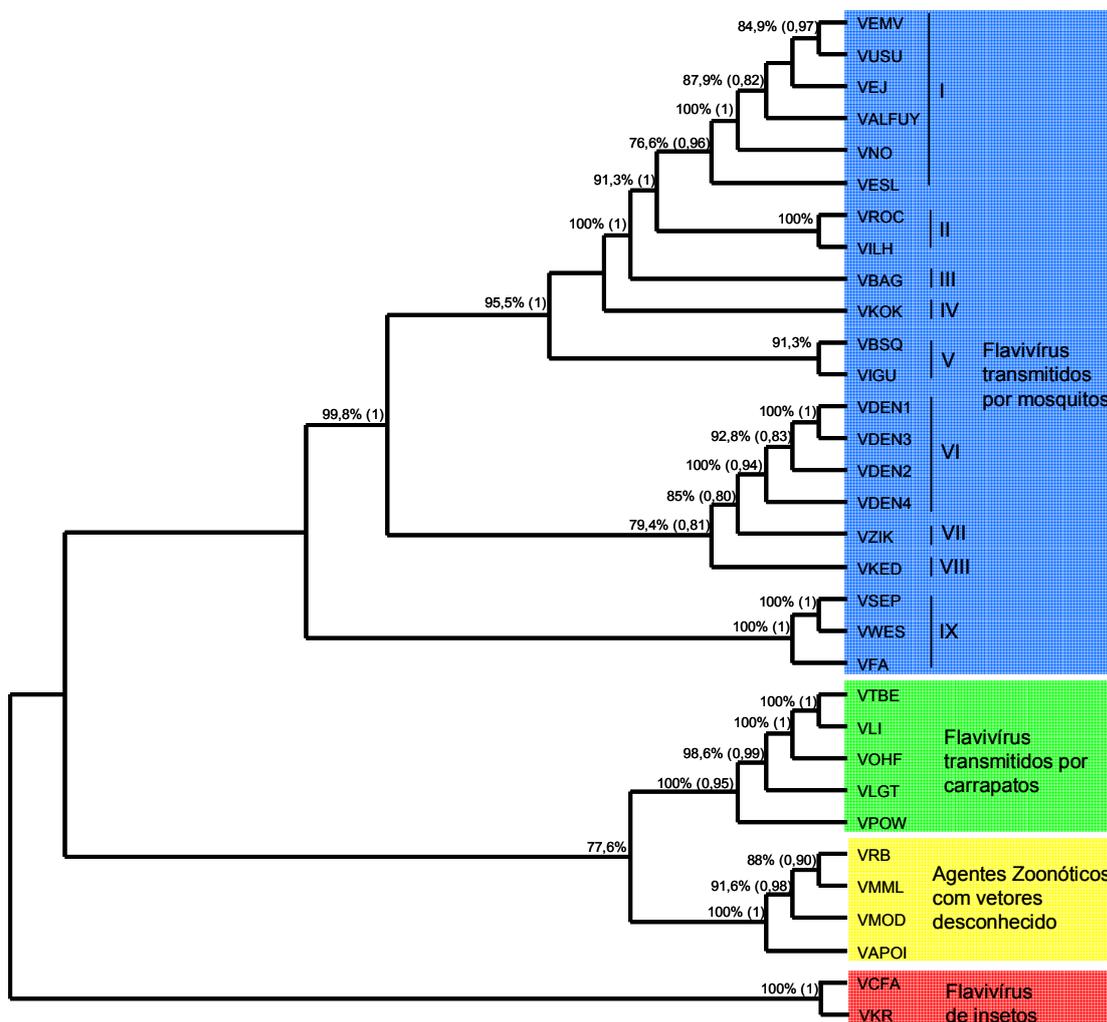


Figura 23 - Árvore filogenética pelo método MV, com valores bayesianos inseridos em parênteses, das sequências nucleotídicas da região estrutural dos 32 flavivírus. Observar que VILH e VROC encontram-se associados ao grupo da encefalite japonesa. **Legenda:** VALFUY = *Vírus Alfuy*; VAPOI = *Vírus apoi*; VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VCFA = *Vírus Cell fusion agent*; VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*; VESL = *Vírus encefalite Saint Louis*; VFA = *Vírus da febre amarela*; VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*; VLI = *Vírus louping ill*; VMOD = *Vírus modoc*; VMML = *Vírus Montana miotis leucoencephalitis*; VNO = *Vírus do Nilo ocidental*; VOHF = *Vírus Omsk hemorrhagic fever*; VPOW = *Vírus powassan*; VRB = *Vírus Rio Bravo*; VROC = *Vírus Rocio*; VSEP = *Vírus sepike*; VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*; VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

Gene E

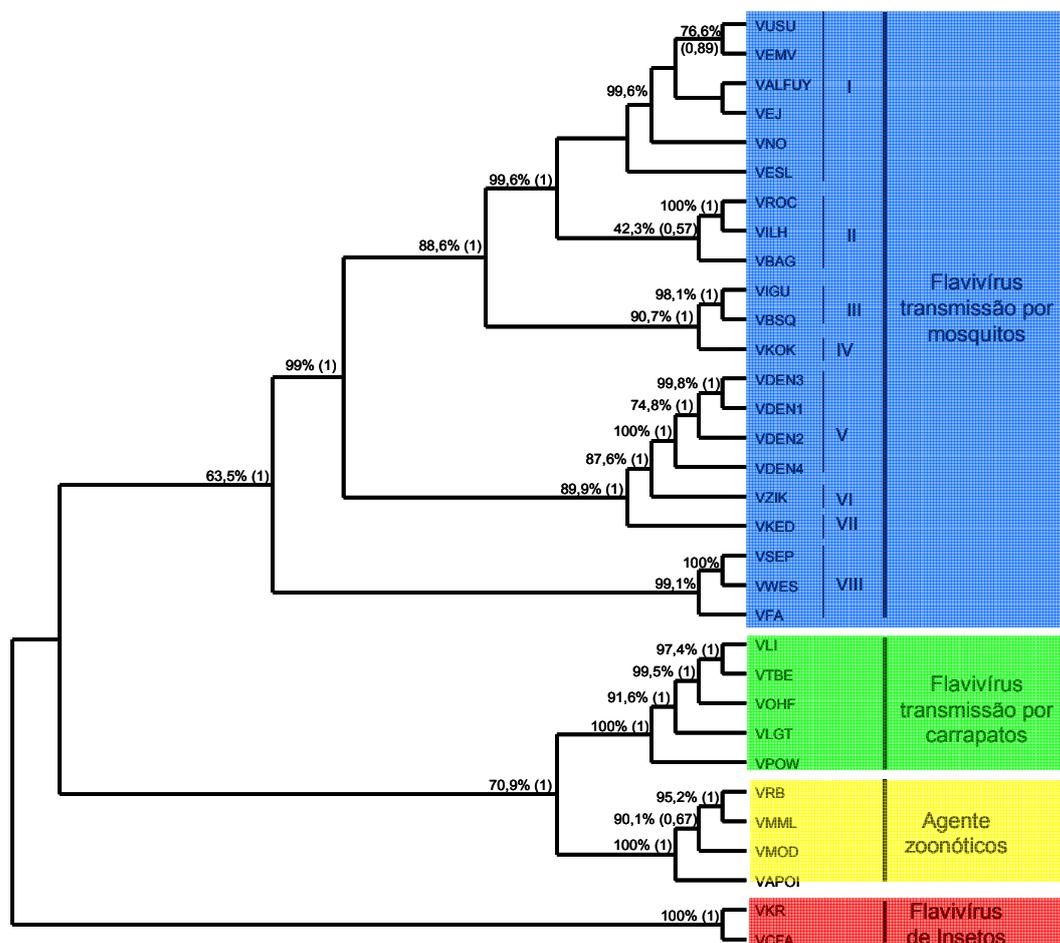


Figura 24 – Árvore filogenética gerada pelo método de MV e Bayesiano (valores bayesianos dentro de parênteses) realizada com as sequências nucleotídicas do gene E de 32 flavivírus. Notar a organização com a formação da topologia conservada para os grupos e a inclusão dos VBAG, VILH e VROC no mesmo clado, porém com baixo suporte de *bootstrap*. **Legenda:** VALFUY = *Vírus Alfuy*, VAPOI = *Vírus apoi*, VBAG = *Vírus Bagaza*, VBSQ = *Vírus Bussuquara*, VCFA = *Vírus Cell fusion agent*, VEJ = *Vírus da encefalite japonesa*, VEMV = *Vírus encefalite Murray Valley*, VESL = *Vírus encefalite Saint Louis*, VFA = *Vírus da febre amarela*, VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*, VKED = *Vírus kedougou*, VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*, VLI = *Vírus louping ill*, VMOD = *Vírus modoc*, VMML = *Vírus Montana miotis leukoencephalitis*, VNO = *Vírus do Nilo ocidental*, VOHF = *Vírus Omsk hemorrhagic fever*, VPOW = *Vírus powassan*, VRB = *Vírus Rio Bravo*, VROC = *Vírus Rocio*, VSEP = *Vírus Sepike*, VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*, VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

3.5.3. Região não estrutural

A diferença na topologia dos flavivírus entre as regiões estrutural e não estrutural (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) também ficou a critério da formação dos clados do grupo dos flavivírus transmitidos por mosquitos. Neste caso, para a região não estrutural, oito clados foram identificados: o clado I representado pelo grupo da encefalite japonesa (*bootstrap* de 100%); considerando que os flavivírus VILH e VROC como espécies virais distintas, estes formam um clado (clado II, *bootstrap* de 100%), que está associado ao VBAG (clado III; *bootstrap* de 96,5%); clado IV que agrupa os flavivírus VBSQ e VIGU; os clados interrelacionados V, VI e VII, representados pelos VZIK, VKED e VKOK, respectivamente; o clado VIII representado pelo grupo Dengue (*bootstrap* de 100%); e o clado IX representado pelo vírus do grupo da febre amarela (*bootstrap* de 100%) (Figura 25).

A exemplo do que foi observado anteriormente, para a região não estrutural a subdivisão em grupos foi mantida para todos os genes, porém a subdivisão em clados dos flavivírus transmitidos por mosquito apresentou algumas diferenças. As árvores geradas para os genes NS3 e NS5 foram as que apresentaram melhor topologia, similar à árvore gerada para o conjunto da região não-estrutural, com os valores de *bootstraps* mais elevados (Figura 26). Os demais genes apresentaram topologia similar a árvore filogenética para toda a região não estrutural, com exceção da NS1, no qual o VILH e VROC estão associados com o grupo da encefalite japonesa, tal como observado para a região estrutural.

Para o gene NS3, os VBAG, VILH e VROC agrupam-se no clado II. Os VIGU e VBSQ estão mais relacionados aos VZIK e VKED (Figura 26 A). Quanto ao gene NS5, a topologia foi similar a observado na análise da ORF dos flavivírus (Figura 26 B).

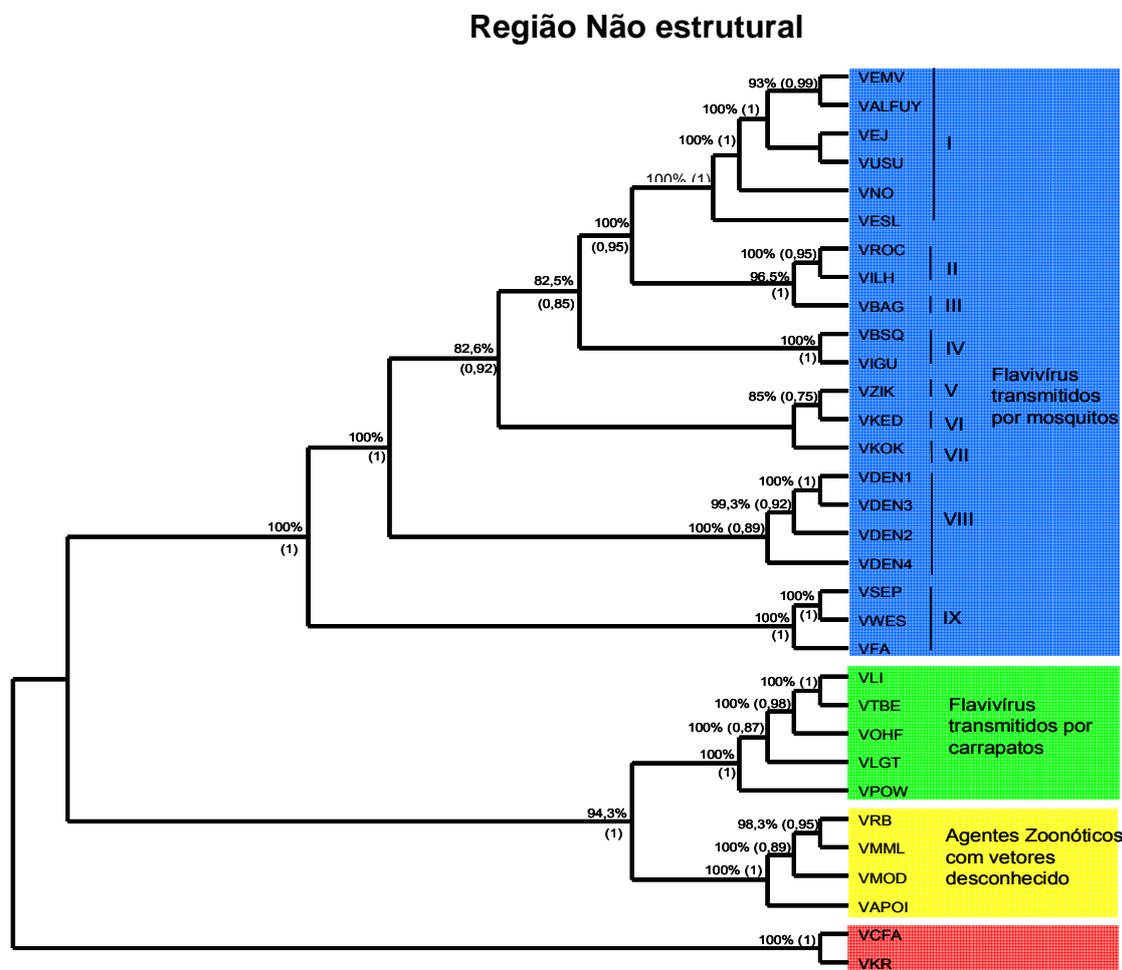


Figura 25 - Árvore filogenética pelos métodos MV e Bayesiano (valores inseridos dentro dos parênteses) das sequências nucleotídicas da região não estrutural dos 32 flavivírus. **Legenda:** VALFUY = *Vírus Alfuy*; VAPOI = *Vírus apoi*; VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VCFE = *Vírus Cell fusion agent*; VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*; VESL = *Vírus encefalite Saint Louis*; VFA = *Vírus da febre amarela*; VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*; VLI = *Vírus louping ill*; VMOD = *Vírus modoc*; VMML = *Vírus Montana miotis leukoencephalitis*; VNO = *Vírus do Nilo ocidental*; VOHF = *Vírus Omsk hemorrhagic fever*; VPOW = *Vírus Powassan*; VRB = *Vírus Rio Bravo*; VROC = *Vírus Rocio*; VSEP = *Vírus Sepike*; VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*; VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

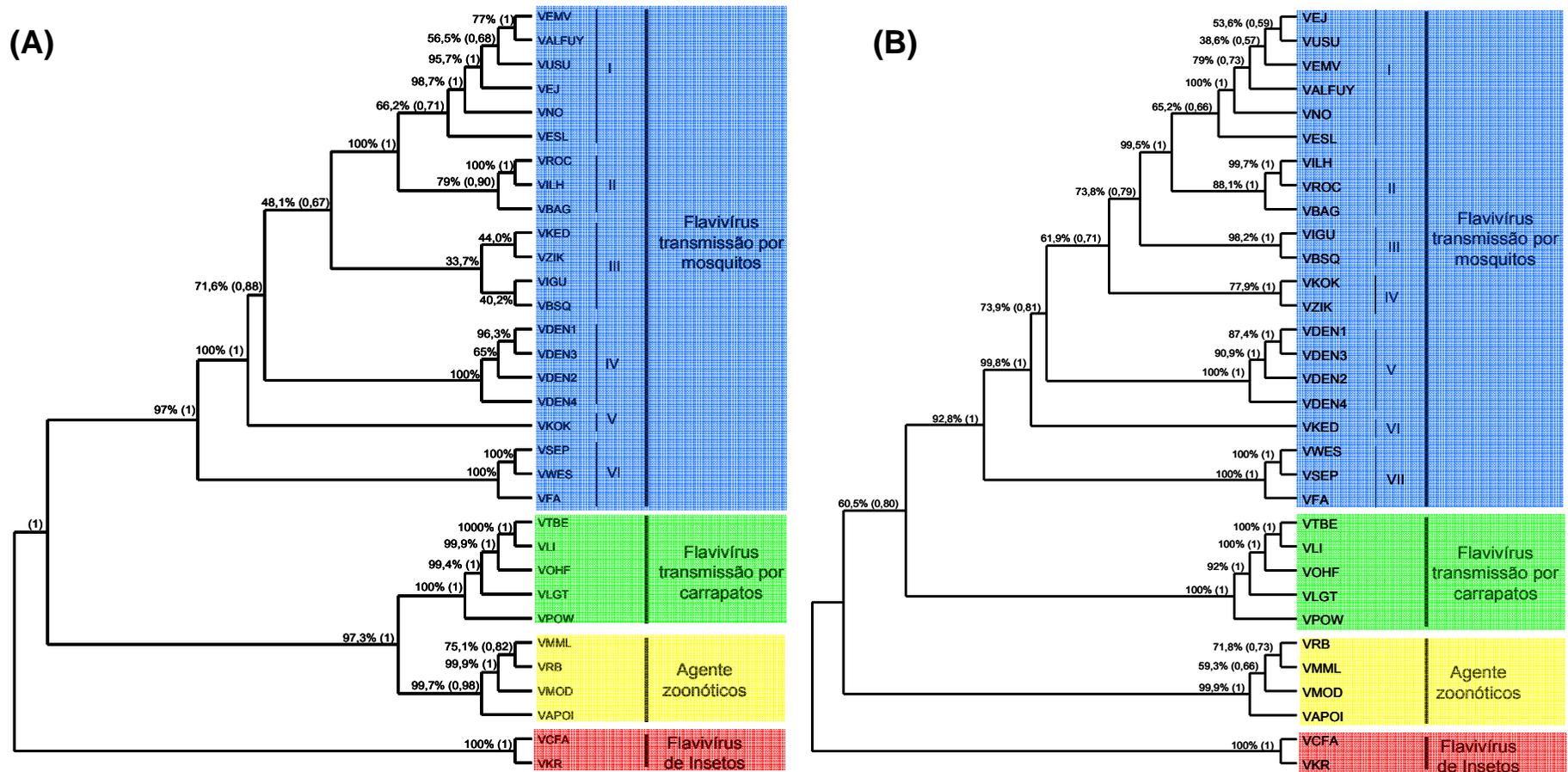


Figura 26 – Árvore filogenética pelos métodos MV das sequências nucleotídicas do gene NS3 (A) e NS5 (B) dos flavivírus. Os valores bayesianos foram plotados dentro de parêntese. Perceber as mudanças de topológica dos flavivírus transmitidos por mosquitos. **Legenda:** VALFUY = *Vírus Alfuy*; VAPOI = *Vírus apoi*; VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VCFA = *Vírus Cell fusion agent*; VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*; VESL = *Vírus encefalite Saint Louis*; VFA = *Vírus da febre amarela*; VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*; VLI = *Vírus louping ill*; VMOD = *Vírus modoc*; VMML = *Vírus Montana miotis leukoencephalitis*; VNO = *Vírus do Nilo ocidental*; VOHF = *Vírus Omsk hemorrhagic fever*; VPOW = *Vírus powassan*; VRB = *Vírus Rio Bravo*; VROC = *Vírus Rocio*; VSEP = *Vírus Sepike*; VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*; VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

3.5.5. Filogenia concatenada E-NS5

Com intuito de observar a topologia dos flavivírus, utilizando sítios informativos, simultaneamente das regiões estrutural e não estrutural, foi realizada a análise filogenética concatenada das sequências parciais dos genes E e NS5 para que se pudessem acrescentar outros flavivírus do grupo Ntaya: o *Vírus Israel turkey meningoencefalomyelite* (VIT; E = AF372415; NS5 = AF013377), *Vírus Ntaya* (VNTA; E = AF372416; NS5 = AF013392) e *Vírus Tembusu* (VTMU; E = AB110494; NS5 = AB110489), bem com o VAROA (E = AF372413; NS5 = AF013362) e VNJL (E = AF372411; NS5 = AF013390). Também foram acrescentados nesta análise outros flavivírus: VAEF (AB488408), VCXF (EU879060) e VQB (NC012671).

O modelo de substituição nucleotídica determinado pelo AIC foi o GTR+I+G [Base=(A=0,2965; C=0,2462; G=0,2533) Nst=6 Rmat=(A-C=1,9077; A-G=3,0328; A-T=1,5385; C-G=1,2199; C-T=5,5299) gamma =0,8723 Pinvar=0,1350].

A subdivisão nos quatro grupos principais dos flavivírus se manteve, observando a inserção dos VAEF, VCXF e VQB juntamente com o VCFA e o VKR, no grupo dos flavivírus de insetos, tendo sido este grupo utilizado com grupo externo para o enraizamento da árvore. O grupo dos flavivírus transmitidos por mosquitos foi subdividido em dez clados: clado I, que inclui os vírus do grupo da encefalite japonesa (*bootstrap* = 100%); o clado II (*bootstrap* = 100%) representados pelos integrantes do grupo Ntaya (VBAG, VNTA, VIT e VTMU); associado ao clado II, porém distinto deste (suportado por valor de *bootstrap* = 70%), o clado III, que agrupa os flavivírus VILH e VROC (*bootstrap* = 99,9%); o clado IV representado pelos VBSQ e VNJL; o clado V que inclui o VAROA e o VIGU; o clado VI representado pelo VKOK; o clado VII que agrupa

os quatro sorotipos do VDEN; e associado ao VII os clados VIII e IX, representado pelos VZIK e VKED, respectivamente; e o clado X representado pelo vírus associados ao VFA (Figura 27).

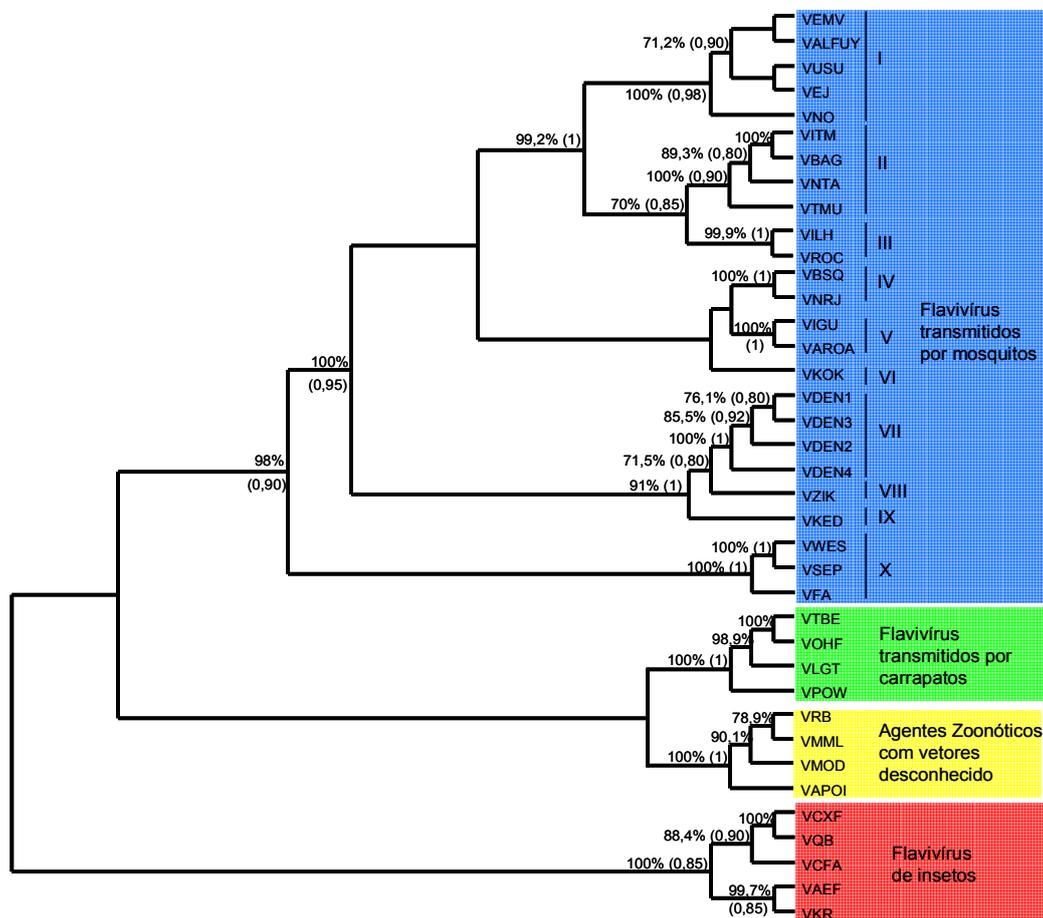


Figura 27 – Árvore filogenética pelo Método MV concatenada das sequências nucleotídicas parciais dos genes E e NS5 de 38 flavivírus. Os valores Bayesianos estão representados dentro de parênteses. Os grupos dos flavivírus transmitidos por mosquitos, flavivírus transmitidos por carrapatos, agentes zoonóticos com vetores desconhecidos e flavivírus de insetos, estão representados pela cores azul, verde, amarelo e vermelho, respectivamente. No grupo dos flavivírus transmitidos por mosquitos, notar a distinção entre os clados II (grupo Ntaya) e clado III (VILH e VROC), bem com a associação entre os VBSQ e VNLJ distinta dos VAROA e VIGU. **Legenda:** VALFUY = *Vírus Alfuy*, VAPOI = *Vírus apoi*, VBAG = *Vírus Bagaza*; VBSQ = *Vírus Bussuquara*; VCFA = *Vírus Cell fusion agent*, VEJ = *Vírus da encefalite japonesa*; VEMV = *Vírus encefalite Murray Valley*, VESL = *Vírus encefalite Saint Louis*; VFA = *Vírus da febre amarela*; VDEN1 = *Vírus dengue sorotipo 1*; VDEN2 = *Vírus dengue sorotipo 2*; VDEN3 = *Vírus dengue sorotipo 3*; VDEN4 = *Vírus dengue sorotipo 4*; VIGU = *Vírus Iguape*; VILH = *Vírus Ilhéus*; VKR = *Vírus Kamitti river*; VKED = *Vírus kedougou*; VKOK = *Vírus kokobera*; VLGT = *Vírus Langat*, VLI = *Vírus louping ill*; VMOD = *Vírus modoc*; VMML = *Vírus Montana miotis leukoencephalitis*; VNO = *Vírus do Nilo ocidental*; VOHF = *Vírus Omsk hemorrhagic fever*, VPOW = *Vírus Powassan*; VRB = *Vírus Rio Bravo*; VROC = *Vírus Rocio*; VSEP = *Vírus Sepike*; VTBE = *Vírus Tick-borne encephalitis*; VUSU = *Vírus Usutu*, VWES = *Vírus Wesselsbron*; VZIK = *vírus Zika*.

4. DISCUSSÃO

Os membros do gênero *Flavivirus* apresentam incontestável importância em saúde pública, bem como para a economia, isto devido às inúmeras epidemias ocorridas em comunidades humanas, com expressiva morbidade e mortalidade, bem como o comprometimento da produtividade de rebanhos, que são fundamentais para manutenção da economia e subsistência global (Gritsun et al., 2003; Halstead & Jacobson, 2003; Staples & Monath, 2008; Ulloa et al., 2009). Tendo em vista essa importância, valiosos esforços da comunidade científica brasileira vêm gerando importantes informações para o melhor entendimento dos aspectos clínicos, ecológicos e de distribuição geográfica dos flavivírus no país (Travassos da Rosa et al., 1997; Vasconcelos et al., 1998; Figueiredo, 2000).

Os estudos de caracterização genética vieram complementar as lacunas do conhecimento deixadas pelos estudos sorológicos, infecções experimentais e epidemiologia dos flavivírus, contribuindo com informações no âmbito molecular, o que estreitou as relações entre as características biológicas e genéticas, além de contribuir com a taxonomia viral e constituir a base para o desenvolvimento de ferramentas diagnósticas mais precisas e rápidas (Kuno et al., 1998; Scaramozzino et al., 2001), bem como para a produção de vacinas (Schlesinger et al., 1992; Chang et al., 2004) e drogas antivirais (Ray & Shi, 2006; Sampath & Padmanabhan, 2009).

Um dos maiores obstáculos para o estudo dos flavivírus é o número limitado de sequências completas do genoma, uma vez que as características únicas para os subgrupos é frequentemente revelada apenas

quando se analisa vários membros de um dado grupo. Tal dificuldade pode ser vista nas análises envolvendo os grupos dos flavivírus transmitidos por mosquitos: Ntaya (VBAG - Kuno & Chang, 2007) – grupo Kokobera (VKOK) e Spondweni (VZIK - Kuno et al., 2005), onde apenas um membro foi completamente seqüenciado ou do grupo Aroa, cujas primeiras sequências completas são apresentadas neste trabalho.

Por outro lado, nos últimos anos com o avanço dos métodos de sequenciamento e de filogenia, o número de sequências completas dos flavivírus vem aumentando consideravelmente, bem como, novos vírus têm sido geneticamente caracterizados (Crabtree et al., 2003; Kuno et al., 2005; Kuno & Chang, 2007; Hoshino et al., 2007; Crabtree et al., 2009; Hoshino et al., 2009), proporcionando: (i) um melhor entendimento das relações filogenéticas e evolutivas entre os integrantes do gênero *Flavivirus* (Zanotto et al., 1996; Kuno et al., 1998; Billoir et al., 2000; Gould et al., 2003b; Cook & Holmes, 2006); (ii) permitir a análise da organização genômica (Kuno et al. 2009); (iii) identificar motivos conservados nas sequências aminoacídica associados com a função da proteína (Gorbalenya et al., 1989; Chang, 1997; Koonin, 1991; 1993; Matusan et al., 2001; Stiasny et al., 2006; Seligman, 2008); (iv) descrever a estrutura secundária do RNA e sua relação com a síntese do RNA (Alvarez et al., 2005a; 2005b; Villordo & Gamarnik, 2009). Assim, este trabalho vem contribuir com a descrição e análises das sequências completas de quatro flavivírus brasileiros classificados segundo o ICTV como segue: os VBSQ e VIGU ditos como pertencentes ao grupo Aroa e os VILH e VROC, ditos como integrantes do grupo Ntaya (ICTV, 2005).

A caracterização genética dos genomas dos flavivírus VBSQ, VIGU, VILH e VROC mostraram uma organização genômica similar aos demais integrantes do gênero: uma única região codificante (ORF) flanqueada por duas RNC (5'RNC e 3'RNC). Na ORF, as dez proteínas virais, C-PrM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5, também foram observadas, como anteriormente descritas (Rice et al., 1985; Chambers et al., 1990).

As diferenças no tamanho do genoma, e conseqüentemente, no tamanho dos genes, observados entre os flavivírus brasileiros é também similar ao padrão apresentado ao se comparar aos demais flavivírus (ANEXO II). Os tamanhos dos genes Pr/M, NS2A, NS2B, NS4A/2K dos flavivírus brasileiros em estudo são iguais aos de outros flavivírus transmitidos por mosquitos do gênero *Culex* e associados à encefalite, observando-se pequenas variações quanto aos tamanhos dos genes NS1, NS3, NS4 e NS5. Em relação ao gene E, no entanto, o VIGU, com 1.482 nt/ 494aa, apresenta tamanho compatível com os vírus do grupo dengue e febre amarela (493 a 495 aa) que são associados com febres hemorrágicas e transmissão por *Aedes* e distinto dos flavivírus encefalitogênicos (500 a 501 aa), incluindo os VBSQ, VILH e VROC (Quadro 8). Nota-se que o tamanho da poliproteína dos VILH e VROC difere em apenas um aminoácido, que provavelmente é determinado por uma deleção na NS3 do VILH (Quadro 8). Entre os flavivírus, a maior variação em tamanho (ANEXO II), bem como em variabilidade genética (Tabela 3) é observada para o gene C.

Essa variação no tamanho da ORF e de genes individuais entre os grupos dos flavivírus transmitidos por mosquito foi observada,

estatisticamente pelos testes não paramétricos por Kuno e colaboradores (2009). Nessa análise, os autores comprovaram uma variação significativa crescente entre os grupos: flavivírus de insetos, agentes zoonóticos com vetores desconhecidos e flavivírus transmitidos por vetores (mosquitos e carrapatos), bem como que entre os flavivírus transmitidos por mosquitos, os vírus relacionados com encefalites apresentam a ORF maior; os VKOK, VIGU, VKED e VZIK apresentam tamanho intermediário; e grupo do VDEN apresenta ORF com menor tamanho.

Quanto à variação no tamanho das RNCs, as 5'RNCs dos flavivírus brasileiros em estudo apresentaram tamanho compatível com a média dos flavivírus transmitidos por mosquitos (Markoff, 2003), notando-se que os VBSQ e VIGU apresentaram tamanhos equivalentes, o que também é observado entre os VILH e VROC (Quadro 8). A maior variação no comprimento foi observada para a 3'RNC, fato este também descrito para outros flavivírus (Hahn et al., 1987; Markoff, 2003).

As sequências conservadas da 3'RNC dos flavivírus transmitidos por mosquitos RCS3-CS3-RCS2-CS2-CS1[3'CYC] (sentido genômico 5'→3'), também foram analisadas nos genomas dos flavivírus brasileiros, que apresentaram tamanho compatível com o descrito na literatura (Hahn et al., 1987; Markoff, 2003). O padrão das sequências conservadas da 3'RNC do VBSQ foi similar ao apresentado pelos sorotipos do VDEN e pelo VKOK, ao passo que para os VIGU, VILH e VROC o padrão foi mais relacionado ao grupo da encefalite japonesa, observando-se, no entanto, a ausência da RCS3

(Figura 15), o que reforça a similaridade desses vírus com o grupo da encefalite japonesa, a exemplo do que ocorre na filogenia (Kuno et al., 1998).

Sabe-se que o padrão das sequências conservadas da 3'RNC está estritamente relacionado à subdivisão dos flavivírus em grupos/sorocomplexos: todos os integrantes do grupo da encefalite japonesa apresentam o padrão RCS3-CS3-RCS2-CS2-CS1, enquanto que os membros do complexo dengue possuem um padrão RCS2-CS2-CS1 (Hahn et al., 1987; Poindinger et al., 1996; Markoff, 2003). A presença de diferentes padrões entre os flavivírus transmitidos por mosquitos e flavivírus transmitidos por carrapatos (Markoff, 2003), sugere a importância dessas regiões para a classificação taxonômica desses agentes virais (Hahn et al., 1987; Charlier et al., 2002; Leyssen et al., 2002; Markoff, 2003; Kuno & Chang, 2006; Kuno & Chang, 2007; Hoshino et al., 2009).

Apesar dos VILH e VROC serem incluídos no grupo Ntaya (ICTV, 2005), a análise mostrou que estes vírus apresentam um padrão CS3-RCS2-CS2-CS1 (Figuras 15 e 16) e que é distinto do VBAG (RCS3-CS3-CS2-CS1) (Kuno & Chang, 2007). Da mesma forma, O VBSQ não apresenta a CS3, que por sua vez está presente no genoma do VIGU e ambos os vírus são ditos com membros do grupo Aroa, de acordo com o CITV (ICTV, 2005).

É interessante notar que ao se comparar o padrão das sequências conservadas da 3'RNC com os diferentes quadro clínicos associados aos flavivírus, observa-se a presença das CS3 e/ou RCS3 no genoma dos vírus encefalitogênicos, tal como aqueles pertencentes ao grupo da encefalite japonesa (Gubler et al., 2006), como por exemplo o VBAG

(Bondre et al., 2009) e o VROC (Lopes et al., 1978a; 1978b). Por outro lado, o padrão RCS2-CS2-CS1 está associado aos vírus que causam manifestações exantemáticas e/ou icterohemorrágicas aqui representados pelos sorotipos do VDEN (Travassos da Rosa et al., 1997), VKOK (Nisbet et al., 2005) e VZIK (Hayes, 2009). Ressalta-se que o VFA apresenta um padrão das sequências conservadas que é distinto dos demais flavivírus transmitidos por mosquitos, possuindo três repetições em série de uma sequência conservada, ao invés da RCS2-CS2 (Markoff, 2003; Bryant et al., 2005). No entanto, para verificar essa hipótese seria necessário realizar estudos infecção experimental com cepas mutantes (sem as CS3 e/ou RCS3) para avaliar o papel dessa região em relação ao tropismo viral e ao tipo de alterações patológicas.

Quanto à conservação das CS1[3'CYC], CS2, RCS2, CS3, RCS3, autores como Kuno e Chang (2007) consideram como sequências conservadas imperfeitas (ImCS) aquelas que apresentam entre duas a cinco alterações nucleotídicas, quando comparadas com a sequência consenso. Sendo assim, a análise do alinhamento nucleotídico das 3'RNC dos flavivírus transmitidos por mosquitos mostrou que a CS1 do VIGU é imperfeita (ImCS1), por apresentar três alterações nucleotídicas (duas substituições do tipo transições e uma inserção). Quanto a CS2 ser completamente conservada entre os integrantes dos grupos da encefalite japonesa e do dengue (exceto o VDEN3), VBAG, VKOK e VZIK (Markoff, 2003; Figura 16), para VIGU e VILH foi observado uma substituição em posições distintas. Ademais, na CS3 do VIGU observou-se uma substituição do tipo transversão (A→T) na posição 10.485, que também é observada na CS3 dos VALFUY e VBAG (Figura 16).

Estudos anteriores mostraram que mudanças pontuais nas sequências conservadas da 3'RNC estão relacionadas a importantes funções biológicas dos flavivírus, tais como sobrevivência e/ou patogenicidade (Bryant *et al.*, 2005; Khromykh *et al.*, 2001) e que deleções das CS2, RCS2 e CS1 estão associadas à redução considerável na competência da replicação viral (Men *et al.*, 1996). Entretanto, estudos adicionais são necessários para correlacionar a não ocorrência de infecção humana pelos VIGU e VBSQ, bem como as infecções subclínicas pelo VILH, com as alterações observadas nas sequências conservadas da 3'RNC.

O papel da estrutura secundária de ciclização do RNA dos flavivírus tem sido associada com os processos de síntese do RNA e tradução das proteínas virais, evidenciando a importância da conservação das sequências de ciclização 5'CYC e 3'CYC (Proutski *et al.*, 1997; Khromykh & Westaway, 1997; Villordo & Gamarnik, 2009). Apesar da CS1 do VIGU ter sido considerada imperfeita (ImCS1) onde uma única alteração foi observada na região 3'CYC, a estrutura secundária do RNA foi conservada observando-se o pareamento com a região 5'CYC.

As estruturas secundárias preditivas para flavivírus brasileiros foram similares às observadas nos estudos realizados por Proutski e colaboradores (1997) e Kuno & Chang (2006; 2007). As estruturas secundárias peculiares da 5'RNC - 5'ES e grampo do capsídeo (5'CapHS) –, bem como da 3'RNC (3'SL) foram mantidas na estrutura de ciclização dos flavivírus brasileiros (Figura 18), corroborando a análise de Villordo & Gamarnik (2009). Como esperado, o pentanucleotídeo 5'-CACAG-3', que é extremamente

importante para a replicação dos flavivírus (Hahn et al., 1987), foi encontrado ao nível da 3'SL dos quatro flavivírus em estudo (Figura 17).

Comparando a energia livre necessária para manter a ciclização do RNA viral dos VBSQ, VIGU, VILH e VROC, observa-se que a energia necessária para a formação da estrutura secundária do VIGU foi menor quando comparada com os demais flavivírus em estudo. Isso possivelmente se deve ao reduzido número de grampos observados na conformação secundária do RNA do VIGU (Figura 17).

As sequências responsáveis pela ciclização do RNA, também estão presentes na estrutura preditiva dos flavivírus brasileiros: pareamento entre as sequências conservadas 5'CYC e 3'CYC (Hahn et al., 1987) e das sequências de ciclização, 5'SCy e 3'SCy (Alvarez et al., 2005b). Apesar do pareamento 5'SCy-3'SCy dos VBSQ e VROC não serem perfeitas (Figuras 18 e 19), observa-se a presença de pequenas alças (PA). Da mesma forma, a 5'CyC do VROC não é completamente pareada em virtude da presença de substituições nucleotídicas (Figura 19). Estes fatos parecem não interferir na formação da estrutura de ciclização desses vírus, tão pouco na infectividade viral, uma vez que o VROC foi capaz de causar lesões intensas no SNC de hamsters dourados (Henriques, 2009). Uma terceira sequência complementar entre a 5'RNC e a 3'RNC (com sete nucleotídeos) foi observada no VBSQ (Figura 17), que até o presente momento ainda não foi descrita para outros flavivírus.

Muito embora as mutações não-silenciosas estejam ligadas diretamente à mudança conformacional das proteínas, o que pode resultar em

alterações bioquímicas e biológicas das proteínas e, conseqüentemente, podendo modificar a virulência, bem como a patogênese viral (atenuar um vírus ou mesmo modificar o seu tropismo tecidual/celular, etc.), as mutações silenciosas podem alterar a conformação da estrutura secundária dos vírus e com isso, interferir em suas características biológicas (Markoff, 2003; Villordo & Gamarnik, 2009). Sendo assim, as alterações genéticas (sinônimas e/ou não sinônimas) que distanciam o VILH do VROC, podem estar relacionadas com a distinta estrutura secundária do RNA do VILH de modo que tal vírus tornou-se menos virulento que o VROC, tal como mostrou o estudo de Henriques (2009).

Com relação à poliproteína, os flavivírus brasileiros basicamente seguiram o mesmo padrão de clivagem de proteína estabelecido para os demais flavivírus isolados de mosquitos (Rice & Strauss, 1990; Billoir et al., 2000; Kuno & Chang, 2007). Como esperado, o padrão da distribuição dos resíduos de Cis nas proteínas PrM, E e NS1 (6,12,12) foi o mesmo para os demais flavivírus (Mackenzie et al., 1996; Heinz & Allison, 2003; Chiu et al., 2005), provavelmente devido à grande importância desses resíduos na formação das ligações dissulfeto responsáveis pela estabilização da estrutura secundária das proteínas (Nowak et al., 1989).

O papel dos carboidratos não é completamente compreendido em relação à função biológica, imunogenicidade e estabilidade das glicoproteínas virais. Anticorpos monoclonais contra a proteína E de diferentes cepas do VESL mostraram que a glicosilação não é essencial para conformação dos epítomos ou atividade biológica, uma vez que cepas não glicosiladas mantêm a mesma antigenicidade (Vorndam et al., 1993). Ademais, a completa

deglicosilação da proteína E do VTBE, pela endoglicosidase F, não impede a infectividade viral nem a atividade hemaglutinante da proteína E, mostrando que a glicosilação não exerce um papel fundamental na estrutura antigênica do VTBE desde que a reatividade das partículas deglicosiladas é idêntica quando analisada por anticorpos policlonais e monoclonais (Winkler et al., 1987). Por outro lado, a importância dos sítios de glicosilação da proteína E nos processos de replicação e maturação viral (Goto et al., 2005; Li et al., 2006), bem como no aumento da eficiência replicativa é evidente (Scherret et al., 2001; Shirato et al., 2004). Ademais, os sítios de glicosilação também estão associados à neurovirulência (Halevy et al., 1994; Chambers et al., 1998; Shirato et al., 2004; Beasley et al., 2005).

A presença de potenciais sítios de glicosilação ao longo das proteínas PrM, E e NS1 dos flavivírus brasileiros, objeto deste trabalho, mostrou que o número de tais sítios difere entre esses vírus, assim como para os demais flavivírus transmitidos por mosquitos (Kuno & Chang, 2006). É interessante notar que os sítios de glicosilação descritos nas posições 154 da proteína E (exceto VBSQ) e 130 da NS1, permanecem conservados entre os quatro flavivírus. A glicosilação na posição 154 da proteína E também foi reconhecida no VKUN (Adams et al., 1995), VDEN (Johnson et al., 1994) e VTBE (Heinz & Allison, 2003); enquanto que a glicosilação na posição 130 da proteína NS1, bem como na posição 207 (vista na NS1 do VBSQ), foi detectada no VDEN-2, mostrando a participação desses sítios na estabilidade da partícula viral e no processamento/transporte da NS1 (Crabtree et al., 2005). Ademais, a deglicosilação de um ou ambos os sítios de glicosilação da NS1

reduz a replicação do VDEN-2 em células de mosquitos clone C6/36, bem como a neurovirulência em camundongos para o VDEN-2 (Crabtree et al., 2005) e VFA (Muylaert et al., 1996). Entretanto, estudos adicionais de infecção experimental fazem-se necessários para tentar correlacionar a função dos sítios de glicosilação das proteínas virais com as características biológicas e antigênicas dos flavivírus.

As análises dos domínios e dos motivos conservados foram realizadas para as proteínas E, NS3 e NS5 devido à relevância dessas proteínas no processo replicativo e na patogenicidade dos flavivírus (Gubler et al., 2007; Lindenbach et al., 2007), além de serem constantemente alvos de análises filogenéticas (Kuno et al., 1998; Gaunt et al., 2001; Kuno et al., 2009).

O tamanho da proteína E varia entre 493 a 501 aa entre os flavivírus transmitidos por vetores (mosquitos e carrapatos) e os agentes zoonóticos. Dois, dos três domínios da proteína E foram identificados pelos bancos de dados do Interproscan: o domínio II (ectodomínio) e III (similar à imunoglobulina). As posições e o tamanho desses domínios dentre as proteínas E dos flavivírus em estudo foram mais ou menos equivalentes (Tabela 1). No domínio II, foi verificada a conservação do peptídeo de fusão e no domínio III, a do tripeptídeo RGD.

Sabe-se que apenas 18% da proteína E é conservada entre os flavivírus patogênicos (Seligman & Bucher, 2003). Apesar do peptídeo de fusão corresponder a 3,2% dos aminoácidos conservados, esse motivo contém cerca de 13% dos aminoácidos conservados da proteína E. A necessidade requerida para as mudanças conformacionais durante o processo da fusão da proteína E

com a membrana celular faz com que o peptídeo de fusão apresente alto grau de conservação entre os flavivírus, sugerindo que a ocorrência de mutações sejam raras, muito embora cinco sítios (N104, L108, F109, S113 e I114) sejam considerados mais variáveis quando se compara diferentes integrantes do gênero *Flavivirus* (Seligman, 2008) (Figura 28, Anexo V).

Os peptídeos de fusão dos VBSQ, VIGU, VILH e VROC estão em posição análoga aos peptídeos de fusão dos demais flavivírus. Ademais, a sequência aminoacídica para os flavivírus transmitidos por *Culex* (DRGWGNGCGLFGKGS) apresenta-se completamente conservada para o VILH (Seligman, 2008). Notou-se que os seis resíduos de Gli necessários para que ocorra a rotação em torno das pontes C-C e C-N, facilitando as mudanças conformacionais da proteína E necessárias para a fusão com a membrana do hospedeiro (Seligman, 2008), permanecem conservada para VILH e VROC, havendo uma substituição de G103N nos VBSQ e VIGU. Dentre os resíduos de Gli, destaca-se o resíduo G105, que influencia diretamente a fusão de membrana e a replicação viral e é altamente conservado entre os flavivírus transmitidos por mosquitos (Seligman, 2008), assim como foi observada para os flavivírus brasileiros.

Quanto aos VBSQ, VIGU e VROC, três, duas e uma alterações aminoacídicas foram detectadas ao nível dos peptídeos de fusão, respectivamente. Comparando-se o número de mutações por sítio (Figura 28), observa-se que mutações no sítio G103, não são muito frequentes entre os flavivírus, muito embora tenham sido observadas tanto no VBSQ quanto no VIGU. O mesmo foi observado para a posição D99, que no VBSQ é N99. Por

outro lado, as alterações nas posições S113 e I114, observada no VBSQ, VIGU e VROC, apresentam as maiores divergências entre os flavivírus (Seligman, 2008).

A perda de uma das seis Gli do peptídeo de fusão do VBSQ e VIGU pela substituição de G→N, precisa ser melhor estudada. Nessa mudança, apesar da polaridade do aminoácido ter se mantido (G e N são ambos polares – ANEXO III), não se sabe o quanto tal mudança pode influenciar a infectividade dos VBSQ e VILH, sendo necessário realizar estudos de proteômica e infecção experimental com cepas geneticamente alteradas, objetivando elucidar essa questão.

O tripeptídeo RGD (resíduos 388 a 390) encontrado no domínio III da proteína E dos VEJ (Lee & Lobigs, 2000), VALFUY (May et al., 2006), VUSU (Bakonyi et al., 2004) e VEMV (Hurrelbrink & McMinn, 2001) aparentemente envolvido no processo de adsorção viral à célula hospedeira, corresponde à RGE no VNO (Castle et al., 1985) e RPG no VESL (Ciota *et al.*, 2007). O tripeptídeo TPG, visto na proteína E do VROC, é também encontrado no VKED, e difere no VBAG no último aminoácido (TGE), enquanto que o tripeptídeo VGD do VBSQ é encontrado no VZIK (Kuno & Chang, 2007). Em nenhum outro flavivírus foi encontrado a sequência QEN, observada no VILH. A função na interação vírus-célula hospedeira em relação aos diferentes tripeptídeos dos flavivírus brasileiros análogos RGD do grupo da encefalite japonesa, precisa ser melhor estudada.

A atividade de protease da NS3 está localizada dentro dos primeiros 180 aa (Li et al., 1999) tal como foi encontrada para os VBSQ, VIGU,

VILH e VROC (Tabela 1). Os sítios catalíticos *His-Asp-Ser* nas posições aminoacídicas 51, 75 e 137, respectivamente (com base na sequência da NS3 do VKUN), também permanecem conservados entre os vírus em estudo, muito embora a localização tenha situado um aminoácido antes no VIGU (Tabela 1). O motivo **VIGLYGNG**, entre as posições 145 a 152 da NS3 do VKUN corresponde ao sulco de ligação do substrato (Chang, 1997) e está conservado para VROC e VBSQ. Ampliando a análise para os demais flavivírus, observa-se a variação nos dois primeiros aminoácidos, o que é compartilhado para o VIGU e VILH. Como esperado, os motivos são relacionados à atividade de protease serina da NS3 (Gorbalenya et al., 1989; Matusan et al., 2001).

Os domínios MTase-SAM e RdRP da NS5 dos flavivírus (Koonin, 1993), apresentam posições análogas no genoma dos flavivírus brasileiros. Dentre os oito motivos determinantes para a atividade de polimerase (RdRP) da NS5 dos flavivírus, observam-se mutações nos motivos V, VI e VII dos flavivírus em estudo. A mesma alteração observada no motivo V do VROC (Y614F) foi também observada para os integrantes do grupo da febre amarela, enquanto que a alteração para o VILH (V619A) foi descrita para os sorotipos do VDEN. A única alteração do motivo VI foi observada exclusivamente para o VILH, assim como a mutação C713G no motivo VII. Ademais, o sítio DDC (também denominado de motivo B), que faz parte do sítio NTP-ligante para inúmeras enzimas ATP- e GTP-ligantes (Kramer & Argos, 1984; Koonin, 1993), também permanece conservado para os quatro flavivírus em estudo.

Sabe-se que alterações nas proteínas que fazem parte do complexo de síntese do RNA viral (NS1-NS2A-NS3-NS4B-NS5) podem

interferir a infectividade viral e promover a uma melhor eficiência da replicação em células do SNC (Gubler et al., 2007), o que chama a atenção para as alterações nos motivos VI e VII do VILH, sugerindo a necessidade de realização de estudos adicionais para estudar essas características desses vírus.

A análise gene por gene utilizando o *Bootscan*, proporcionou informações valiosas para a melhor compreensão das relações filogenéticas entre os quatro flavivírus em estudo com outros flavivírus transmitidos por *Culex*. Essa análise corroborou os achados da literatura (Kuno et al, 1998; Kuno & Chang, 2005; Kuno et al., 2009) que mostra a estreita relação genética entre VBSQ e VIGU (Figura 20 A) e uma proximidade desses vírus com os VKOK (E, NS1 e NS3), VZIK (NS3), VKED (NS1). O VILH e VROC são estreitamente relacionados entre si e próximos dos VNO (NS1) e VBAG (PrM) (Figura 20 B), corroborando as análises filogenéticas com base em estudos prévios com NS5 (Kuno et al., 1998; Gaunt et al., 2001), E (Gaunt et al., 2001) e ORF (Kuno et al., 2009), que mostram que os flavivírus brasileiros apresentam topologia posicionada entre os membros do grupo da encefalite japonesa e o VBAG do grupo Ntaya. Essa relação fica melhor evidenciada quando o VILH foi retirado da análise do *Bootscan* (Figura 21), em que se observa que ao longo da ORF, o VROC apresenta maior similaridade com os VBAG e VESL, corroborando com a filogenia de diferentes genes.

Ademais, o programa *Bootscan* também já servir como ferramenta para avaliar recombinação entre cepas virais, tal como observado para *Vírus da imunodeficiência humana* (Lole et al., 1999). Apesar dos

flavivírus exibirem alguns fatores ecológicos bem como a ocorrência de epidemias e de infecções crônicas dos hospedeiros que propiciam a recombinação (Gritsun et al., 2003; Gubler et al., 2007), as análises das árvores filogenéticas individuais utilizando genes de qualquer parte do genoma ou mesmo o genoma inteiro sugerem recombinação entre os flavivírus de uma mesma espécie (Gould et al., 2003b). Portanto, a análise do *Bootscan* utilizando espécies distintas apenas mostra que um dado flavivírus pode apresentar maior similaridade com diferentes vírus, dependendo da região da ORF que está sendo analisada, afastando a possibilidade de recombinação entre esses agentes virais.

A exemplo da similaridade nucleotídica entre os flavivírus vista na análise do *Bootscan*, o cálculo de identidade aminoacídica reflete a mesma proximidade genética: VBSQ com VIGU e VILH com VROC. É válido ressaltar que as identidades aminoacídica entre VILH e VROC (77,5%) e nucleotídica (72,5%) para a ORF, assim como para o gene NS5 (78,5% aa / 70,3% nt) demonstradas neste estudo são inferiores ao valor limítrofe de 84% para identidade nucleotídica utilizado como critério de inclusão de espécie viral dentro do gênero *Flavivirus*, como previamente descrito (Kuno et al., 1998; Gaunt et al., 2001). Portanto, esse dado suporta a hipótese de que os VILH e VROC correspondem a espécies distintas. Este raciocínio também se aplica para a situação dos VBSQ e VIGU (Tabela 2).

Com base no critério de inclusão de grupo, que suporta valores \geq 71% para a identidade aminoacídica (Kuno et al., 1998; Gaunt et al., 2001), os VILH e VROC não apresentaram valores suficientes para a inclusão no grupo

Ntaya (média da identidade aminoacídica da ORF do VILH e VROC em relação ao VBAG: 64,9%) nem com o grupo da encefalite japonesa (média de 64,9%).

Ademais, o gene C apresenta os menores valores de identidade aminoacídica entre os flavivírus (Tabela 2), o que explica a dificuldade de estabelecer relações filogenéticas utilizando tal gene. Ao se analisar o alinhamento das sequências nucleotídicas e aminoacídicas do gene C, percebe-se a grande variabilidade genética existente, com um número de *gaps* representativos capazes de gerar valores de *bootstraps* de baixa confiabilidade para dar suporte à relação filogenética. Por outro lado, a análise de 62 cepas virais do VDEN1, mostrou que o gene C é extremamente conservado dentro de uma dada espécie (Carneiro, 2009). Sendo assim, sugere-se que gene C seja melhor analisado e que este possa ser um bom alvo para diferenciar espécies virais do gênero *Flavivirus* para fins de diagnóstico.

A avaliação de sequências genômicas completas e parciais para um crescente número de flavivírus tem se mostrado importante para estimar a relação filogenética e a base para hipóteses evolutivas. Análises filogenéticas tem sido feitas com base nas seqüências genômicas completas, sequências nucleotídicas parciais ou em sequências aminoacídicas das proteínas virais, principalmente dos genes E e/ou NS5 (Zanotto et al., 1996; Billoir et al., 2000; Gaunt et al., 2001; Gould et al., 2003a; Kuno & Chang, 2005; Cook & Holmes, 2006; Kuno et al., 2009).

As subdivisões nos quatro grupos relacionados às características ecológicas/filogenéticas: flavivírus transmitidos por mosquitos, flavivírus transmitidos por carrapatos, agentes zoonóticos com vetor desconhecido e

flavivírus de insetos (Kuno et al., 1998; Billoir et al., 2000; Gaunt et al., 2001; Gould et al., 2003a; Kuno et al., 2009), foram mantidas em todas as análises filogenéticas apresentadas neste trabalho. Essa subdivisão reflete largamente a seleção adaptativa natural imposta ao vírus pelos hospedeiros vertebrados, vetores invertebrados e pelas associações ecológicas (Gould et al., 2003a).

No que se refere à cladogênese dos flavivírus transmitidos por artrópodes, nota-se que a organização genômica do grupo dos flavivírus transmitidos por carrapatos é monofilética, ao passo que para flavivírus transmitidos por mosquitos é parafilética. De fato, essa peculiaridade também foi observada por Zanotto e colaboradores (1996), que relata intensa ramificação, seguido por extinção de linhagens e conseqüente formação de subclados para o grupo dos flavivírus transmitidos por mosquitos que se encontra associado com a teoria de “*Boom and Bust*”, enquanto que os flavivírus transmitidos por carrapatos segregam-se continuamente através do tempo.

Percebe-se que a análise individual de cada gene, bem como das regiões estrutural e não estrutural, o subgrupo dos flavivírus transmitidos por *Culex* apresenta uma maior variação da topologia, contrastando com o padrão estável dos grupos do VDEN e VFA (flavivírus transmitidos por *Aedes*). Esse aspecto dá subsídios para a análise evolutiva de Gaunt e colaboradores (2001) que sugere que os eventos evolutivos dos flavivírus transmitidos por *Culex* sejam mais recentes, e por isso apresentam relações filogenéticas mais instáveis quando comparadas aos flavivírus transmitidos por *Aedes*. Foi observado que o grupo da febre amarela constitui uma linhagem distinta dentre

os flavivírus transmitidos por mosquitos, que, provavelmente está associado às próprias características genéticas desses grupo que o separa dos sorotipos do VDEN.

Dependendo da região gênica a ser analisada, o grupo constituído pelos VILH e VROC pode estar mais associado com o grupo da encefalite japonesa do que com o grupo Ntaya, que é o caso da análise da região estrutural, que, por sua vez, corrobora os dados sorológicos. Por outro lado, os VILH e VROC podem estar mais associados ao grupo Ntaya, como é observado na árvore construída pela região não estrutural. Sabendo que cerca de 80% da ORF é composta da região não estrutural, é evidente que a árvore filogenética da ORF seja similar a da região não estrutural.

A associação filogenética entre VILH e VROC com o grupo da encefalite japonesa, e distinto do grupo Ntaya (tal como mostra a topologia da Figura 24) também foi observada na análise filogenética realizada por Bronde e colaboradores (2009), utilizando o método de AV. Sendo assim, variando-se o método filogenético, a associação entre VILH e VROC com os referidos grupos também pode ser observada.

A relação peculiar entre os flavivírus transmitidos por *Aedes* (VDEN e VFA) e o viscerotropismo, bem como os flavivírus transmitidos por *Culex* com o neurotropismo (grupo da encefalite japonesa) foi primeiramente associado por Sabin (1959). Posteriormente, Gaunt e colaboradores (2001) confirmaram essa relação com base nas análises filogenéticas do gene NS5, muito embora algumas exceções a essa teoria tenha sido mostrada para alguns flavivírus transmitidos por mosquitos (Mutebi et al., 2004).

Com tudo, as análises filogenéticas da ORF, região estrutural e região não estrutural mostraram uma subdivisão com relação ao tropismo viral um pouco distinta da apresentada por Gaunt e colaboradores (2001). Atualmente, já se sabe que o VZIK é transmitido por *Aedes* sp. e está associado com síndromes exantemáticas (Kuno e Chang, 2005), tal como os VDEN e o VFA (Travassos da Rosa et al., 1997). O VKED está associado filogeneticamente ao VZIK e também é transmitido por *Aedes* sp., muito embora as manifestações clínicas do VKED sejam desconhecidas (Kuno e Chang, 2005). Na análise da região estrutural, que está diretamente associada às características biológicas dos flavivírus, observa-se claramente a associação do VZIK, VKED, com o grupo do VDEN (Figura 24).

O grupo dos flavivírus transmitidos por *Culex* é formado pelos membros do grupo da encefalite japonesa, os VILH, VROC, VBAG, VBSQ e VKOK. Apesar do VILH e VROC terem sido isolados de *Psorophora* sp. (Vasconcelos et al., 1992; 1998), o VROC pode se replicar em mosquitos do gênero *Culex* sp., pelo menos em condições laboratoriais (Mitchell et al., 1981). A maioria desses flavivírus está associada à encefalite, com exceção dos vírus: VBSQ, cuja clínica foi associada há apenas um caso febril (Srihongse & Johnson, 1971); VIGU, para o qual não há registro de casos humanos (Figueiredo, 2000); e o VKOK, que causa doença febril exantemática (Nisbet et al., 2005).

Essas lacunas da associação entre as características ecológica e tropismo viral com a filogenia vista no grupo dos flavivírus transmitidos por mosquitos é reflexo da própria organização parafilética desse grupo.

Provavelmente, essa questão só poderá ser elucidada com a descoberta de novos flavivírus. Sendo assim, quanto mais conhecimento se produz a respeito da filogenia dos flavivírus, mais sequências estarão disponíveis e, conseqüentemente, se novos isolados são identificados, eles podem ser inseridos em um grupo já existente ou mudar a posição de alguns vírus dentro da classificação atual determinada pelo CITV (ICTV, 2005).

Esta plasticidade da organização filogenética dos flavivírus pode ser observada na análise do vírus Kadam (KADV) por Gaunt e colaboradores (2001), no qual este flavivírus transmitido por carrapatos era classificado, primeiramente, no grupo dos hospedeiros mamíferos, sendo sugerido a sua maior aproximação genética com o grupo dos hospedeiros de aves marinhas, formando um forte grupo filogenético juntamente com os vírus Tyuleniy (TYUV), Saumarez Reef (SREV) e Meanban (MEAV). A posição filogenética dos vírus VESL, VZIK, VEJ, VKED, Spondweni (VSPON) são constantemente contestadas e devem ser revistas (Gould et al., 2003b), a exemplo desta revisão quanto à classificação e posições filogenéticas dos VBSQ, VIGU, VILH e VROC.

A análise filogenética concatenada das sequências parciais dos genes E-NS5 agrupou informações contidas tanto das regiões estrutural quanto não estrutural. Nessa análise ficou evidente a distinção entre VBSQ e VIGU, e que estes vírus pertencem a grupos genéticos distintos: sendo o VBSQ associado ao VNJL e o VIGU associado ao VAROA. Ademias, os VILH e VROC estão separados do grupo Ntaya, constituído pelos VBAG, VNTA, VITM e VTMU, mostrando serem dois grupos distintos, porém ambos associados ao

grupo da encefalite japonesa. Esses dados corroboram as análises filogenéticas das sequências parciais do gene NS5 realizada por Gould e colaboradores (2003a; 2003b) e Cook & Holmes (2006).

Comprando todas as análises realizadas neste trabalho, acredita-se que a filogenia da ORF ainda apresente dados mais fidedignos com relação à verdadeira associação entre os flavivírus do que utilizar análises de genes individuais, muito embora as análises dos genes E, NS3 e, principalmente NS5, possam resolver a topologia de alguns flavivírus.

O VROC é classificado como subtipo do VILH e ambos incluídos no grupo grupo Ntaya. De maneira similar, o VIGU e VBSQ são classificados como subtipos VAROA. Com base nos próprios critérios do ICTV para a classificação de espécie viral, analisaram-se as diferenças desses vírus quanto a: organização do genoma, identificação das proteínas, propriedades antigênicas, relações filogenéticas e características biológicas (ICTV, 2005).

Os critérios de morfologia e propriedades físico-químicas que são frequentemente utilizados para a inclusão dentro da família viral, que no caso dos vírus em estudo não há dúvidas que pertençam à família *Flaviviridae*. Os critérios de organização do genoma, bem como, características das proteínas (número, tamanho e atividade funcional de proteínas estruturais; número, tamanho e atividade funcional de proteínas não-estruturais; sequências de aminoácidos; glicosilação), que relacionam uma dada espécie viral a um gênero, já foi discutida anteriormente, confirmando que os VBSQ, VIGU, VILH e VROC pertencem ao gênero *Flavivirus*.

As relações filogenéticas, como discutidas anteriormente, mostram que VILH e VROC são vírus distintos, relacionados entre si e formando num grupo filogenético a parte do grupo Ntaya: o grupo Rocio. Já o VBSQ e o VIGU não são subtipos do VAROA, tratando-se de espécies viaris distintas, sendo então, proposta a formação de dois grupos genéticos: o grupo Bussuquara, que agrupa VBSQ e VNJL e o grupo Aroa, constituído por VAROA e VIGU.

Quanto às propriedades antigênicas os quatro flavivírus em estudo foram relacionados ao grupo B segundo a classificação sorológica de Casals (Karabatsos, 1985). Entretanto, algumas peculiaridades sobre a relação antigênica desses vírus devem ser relevadas.

Como relatado anteriormente, VILH e VROC são antigenicamente relacionados entre si com os membros do grupo da encefalite japonesa (Lopes et al., 1978a; 1978b; Iversson et al., 1988; Vasconcelos et al., 1998), sendo diferenciados apenas pelo TN (Vasconcelos et al., 1998). De fato, as relações dos títulos dos anticorpos homólogos versus heterólogos obtidos quando o VROC é utilizado com antígeno ou com seu anticorpo específico é de $>640/<10$ e $1280/<10$, respectivamente (Karabatsos, 1985).

A reação cruzada existente entre os flavivírus, sobretudo observada por IH, já foi relacionada à existência de diversos epítomos da proteína E que são compartilhados por diferentes flavivírus (Gould et al., 1985; Heinz, 1986), e este fato deve estar diretamente relacionado à emergência recente dos flavivírus transmitidos por *Culex* (Gould et al., 2003b). Ademais, tendo em vista que a proteína E é de suma importância para a adsorção viral à

célula hospedeira, tropismo viral e patogênese, uma vez que mutações nessa proteína possam resultar em neurotropismo de vírus viscerotrópicos (Gubler et al., 2007), a existência de epítomos compartilhados entre os flavivírus deve ser reflexo de regiões conservadas importantes para a funcionalidade da proteína E.

Quanto à relação antigênica não existem dados na literatura que comprovem a relação desses vírus com o grupo Ntaya. Baseado no teste de FC realizado na *University of Texas Medical Branch (UTMB)* (Travassos da Rosa, A.P.A., dados não publicados), não há reatividade entre o VROC com o VNTA enquanto que o VILH apresenta baixa reatividade (Tabela 5, Anexo V). O VNTA está mais relacionado antigenicamente com o grupo da encefalite japonesa (Tabela 5, Anexo V), assim como descrito para o VBAG (Kuno & Chang, 2007; Bondre et al., 2009). O que se deduz que a inclusão de VILH e VROC deve ter sido feita com base das análises filogenéticas do que NS5 e da ORF.

As relações antigênicas por teste de FC e IH entre VAROA e VIGU são mais evidentes entre si do que com os VBSQ e VNJL, que por sua vez são mais correlacionados (Travassos da Rosa, A.P.A., comunicação pessoal; Karabatsos, 1985).

Do ponto de vista ecoepidemiológico, VILH e VROC apresentam possivelmente o mesmo vetor (*Psorophora* sp.); a distribuição geográfica dos desses flavivírus é distinta, sendo o VILH detectado em regiões diferentes do Brasil e o Rocio mais restrito ao sudeste do país.

Com relação às características biológicas, muito embora o VILH e VROC apresentem potencial encefalitogênico, o VILH está mais associado a casos febris esporádicos com menor comprometimento neurológico, ao passo que para o VROC foram registrados casos de encefalites graves e por vezes fatais, o que claramente indica uma diferença na neurovirulência e no modo de transmissão desses dois vírus. De fato essas diferenças de patogenicidade foram observadas por Henriques (2009) ao analisar o padrão de resposta imune e características das lesões teciduais e persistência da infecção experimental em hamsters dourados inoculados via intraperitoneal. No referido estudo, VILH apresentou-se mais imunogênico ao passo que o VROC apresentou maior capacidade de neurovirulência; e o tempo de persistência da infecção para o VROC (três meses) foi maior do que para o VILH (30 dias). Estudos de infecção experimental com os VAROA, VBSQ, VIGU e VNJL são necessários e possivelmente elucidativas para a compreensão das diferenças e semelhanças das características biológicas desses vírus.

Portanto, baseados nas características genéticas apresentadas neste trabalho, aliados às propriedades antigênicas e biológicas dos VILH e VROC, sugere-se que esses vírus sejam considerados espécies virais distintas e pertencentes a um grupo genético a parte do grupo Ntaya, designado de grupo Rocio, devido ao fato desse vírus ter maior importância médica que o VILH.

Por outro lado, saber se os VBSQ e VNRJ, bem como VAROA e VIGU tratam-se de um mesmo vírus ou espécies diferentes, é uma questão que só poderá ser completamente respondida pelo sequenciamento completo do

genoma do VAROA e VNJL e consequente análise genética, aliada aos estudos de infecção experimental, a exemplo dos estudos conduzidos para VILH e VROC.

Para finalizar, este trabalho estimula a comunidade científica a desenvolver estudos de caracterização genética e análise da patogenicidade dos VAROA, VNJL, além do VCPC e outros membros do grupo Ntaya, e ressalta a importância de concatenar informações genéticas, relações antigênicas, características biológicas e epidemiológicas para elucidar questões controversas.

5. CONCLUSÕES

- Os Flavivírus VBSQ, VIGU, VILH e VROC apresentaram organização gênômica e tamanho da ORF, 5'RNC e 3'RNC similar aos demais integrantes do gênero *Flavivirus*, sendo mais estritamente relacionados ao grupo dos flavivírus transmitidos por mosquitos.
- O padrão das sequências conservadas da 3'RNC do VBSQ assemelhou-se com o padrão VDEN e VKOK, ao passo que os VIGU, VILH e VROC apresentaram um padrão similar aos vírus do grupo da encefalite japonesa, com exceção da ausência da RCS3;
- As estruturas secundárias dos flavivírus brasileiros VBSQ, VIGU, VILH e VROC são similares à estrutura dos demais flavivírus, apresentando a conservação das estruturas secundárias da 5'RNC e 3'RNC e a presença das sequências de ciclização 5'SCy-3'SCy e 5'CYC-3'CYC;
- Foi observado para o VBSQ um terceiro par de sequências pareadas entre a porção 5' e a 3'RNC inexistente para os demais flavivírus;
- Os flavivírus brasileiros basicamente seguiram o mesmo padrão de clivagem de proteína estabelecido para os demais flavivírus transmitidos por mosquitos;

- O padrão da distribuição dos resíduos de Cis nas proteínas PrM, E e NS1 para os flavivírus em estudo é similar aos dos demais membros do gênero *Flavivirus*;
- O número de sítios de glicosilação para as proteínas PrM, E e NS1 é distinto entre os flavivírus brasileiros em relação aos demais flavivírus, muito embora tenha sido observada a conservação nos sítios de glicosilação na posição 154 da proteína E e 130 na proteína NS1;
- A sequência do peptídeo de fusão (proteína E) para o VILH é similar a sequência consenso para os flavivírus transmitidos por mosquitos ao passo que para os VBSQ, VIGU e VROC, observam-se pequenas alterações aminoácídicas pontuais.
- O tripeptídeo RGD, visto entre os membros do grupo da encefalite japonesa, apresenta diferentes sequências para os VBSQ, VIGU, VILH e VROC;
- Os motivos conservados que determinam a atividade de helicase da NS3 apresentam-se conservados para os flavivírus transmitidos por mosquitos, enquanto que para os motivos relacionados à atividade RdPd da NS5, o VROC apresentou mutações pontuais no motivo V, e o VILH nos motivos V, VI e VII;

- As análises filogenéticas, associadas ao padrão das sequências conservadas da 3'RNC, a divergência aminoacídica e as características sorológicas e biológicas sugerem que o VILH e VROC são espécies virais distintas e que formam um grupo relacionado mas separado dos grupos Ntaya e da encefalite japonesa;
- É proposta a formação de três grupos adicionais aos existentes para o gênero *Flavivirus*: o grupo Rocio, que agrupa o VILH e VROC; o grupo Bussuquara, composto por VBSQ e VNRJ; e o grupo Aroa, que inclui os VAROA e VIGU.

6. REFERENCIAS BIBLIOGRÁFICAS

- ACKERMANN, M.; PADMANABHAN, R. De novo synthesis of the RNA by Dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. **The Journal of Biological Chemistry**, **276**: 39926-39937, 2001.
- ADAMS, S.C.; BROOM, A.K.; SAMMELS, L.M.; HARTNETT, A.C.; HOWARD, M.J.; COELEN, R.J.; MACKENZIE, J.S.; HALL, R.A. Glycosylation and antigenic variation among Kunjin virus isolates. **Virology**, **206**(1):49-56, 1995.
- ALCON, S.; TALARMIN, A.; DEBRUYNE, M.; FALCONAR, A.; DEUBEL, V.; FLAMAND, M. Enzyme-linked immunosorbent assay specific to Dengue virus type 1 nonstructural protein NS1 reveals circulation of the antigen in the blood during the acute phase of disease in patients experiencing primary or secondary infections. **Journal of Clinical Microbiology**, **40**(2):376-81, 2002.
- ALVAREZ, D.E.; LELLA EZCURRA, A.L.; FUCITO, S., GAMARNIK, A.V. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. **Virology**, **339**(2):200-12, 2005a.
- ALVAREZ, D.E. LODEIRO, M.F., LUDUENA, S.J.; PIETRASANTA, L.I.; GAMARNIK, A.V. Longe-range RNA-RNA interaction circularize dengue virus genome. **Journal of Virology**, **79**(11): 6631-6643, 2005b.
- BARBA-SPAETH, G; LONGMAN, R.S.; ALBERT, M.L.; RICE, C.M. Live attenuated yellow fever 17D infects human DCs and allows for presentation

- of endogenous and recombinant T cell epitopes. **The Journal of Experimental Medicine**, **202**(9):1179-84, 2005.
- BARRETT, A.D. Current status of flaviviruses vaccines. **Annals of the New York Academy of Sciences**, **951**: 262-271, 2001.
- BAKONYI, T.; GOUL, D. E. A.; KOLODZIEJEK, J.; WEISSENBOCK, H.; NOWOTNY, N. Complete genome analysis and molecular characterization of Usutu virus that emerged in Austria in 2001: comparison with the South African strain SAAR-1776 and other flaviviruses. **Journal of Virology**, **328**: 301-310, 2004.
- BEASLEY, D.W.; WHITEMAN, M.C.; ZHANG, S.; HUANG, C.Y.-H.; SCHNEIDER, B.S.; SMITH, D.R.; GROMOWSKI, G.D.; HIGGS, S.; KINNEY, R.M.; BARRETT, A.D.T. Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. **Journal of Virology**, **79**: 8339-8347, 2005.
- BHARDWAJ, S.; HOLBROOK, M.; SHOPE, R. E.; BARRETT, A. D. T.; WATOWICH, S. J. Biophysical characterization and vector-specific antagonist activity of domain III of the tick-borne flavivirus envelope protein. **Journal of Virology**, **75**: 4002–4007, 2001.
- BILLOIR, F.; CHESSE, R.; TOLOU, H.; MICCO, P.; GOULD, E.A.; LAMBALLERIE. Phylogeny of the genus *Flavivirus* using complete coding sequences of arthropod-borne viruses and viruses with no known vector. **Journal of General Virology**, **81**: 781-790, 2000.
- BLOK, J.; GIBBS, A.J. Molecular systematic of the flaviviruses and their relatives. In: **Molecular basis of virus evolution**. Gibbs, A.; Calisher, C.H.;

- Garcia-Arenal, F. (eds). Cambridge University Press, Cambridge, England, 1995. p. 270-289.
- BLOK, J.; MCWILLIAM, S.M.; BUTLER, H.C.; GIBBS, A.J.; WEILLER, G.; HERRING, B.L.; HEMSLEY, A.C.; AASKOV, J.G.; YOKSAN, S.; BHAMARAPRAVATI, N. Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence relationships with the flaviviruses and other viruses. **Virology**, **187**(2): 573-90, 1992.
- BLOK, J.; SAMUEL, S.; GIBBS, A.J.; VITARANA, U.T. Variation of the nucleotide and encoded amino acid sequence of the envelope gene from eight dengue-2 viruses. **Archive of Virology**, **105**: 39-53, 1989.
- BRAY, M.; LAI, C.J. Dengue virus premembrane and membrane proteins elicit a protective immune response. **Virology**, **185**: 505–508, 1991.
- BRINTON, M.A.; DISPOTO, J.H. Sequence and secondary structure analysis of the 5'-terminal region of flavivirus genome RNA. **Virology**, **162**(2): 290-9 1988.
- BONDRE, V. P.; SAPKAL, G. N.; YERGOLKAR, P. N.; FULMALI, P. V.; SANKARARAMAN, V.; AYACHIT, V. M.; MISHRA, A. C., GORE, M. M. Genetic characterization of Bagaza virus isolated in India and evidence of anti-Bagaza virus antibodies in sera collected from encephalitis patients. **Journal General Virology**, publicado em Julho, 2009 (in Press).
- BRYANT, J.E.; VASCONCELOS P.F.C.; RIJNBRAN, D.R.C.A.; MUTEBI, J.P.; HIGGS, S.; BARRETT, A.D.T. Size heterogeneity in the 3' non-coding region of South American isolates of yellow fever virus. **Journal of Virology**, **79**: 3807-3821, 2005.

- CALISHER, C.H.; GOULD, E.A. Taxonomy of the virus family *Flaviviridae*. **Advances in Virus Research**, **59**:1-19, 2003.
- CALISHER, C.H.; KARABATSOS, N.; DALRYMPLE, J.M.; SHOPE, R.E.; PORTERFIELD, J.S.; WESTAWAY, E.G.; BRANT, W.E. Antigenic relationship between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. **Journal of General Virology**, **70**:37-43, 1989.
- CAMMISA-PARK, H.; CISAR, L.A.; KANE, A.; STOLLAR, V.; The complete nucleotide sequence of the cell fusion agent (CFA): Homology between the nonstructural proteins encoded by CFA and the nonstructural proteins encoded by arthropod-borne flaviviruses. **Virology**, **189**: 511-524, 1992.
- CARNEIRO, A. R. **Caracterização molecular de cepas do vírus dengue 1 isoladas no Brasil entre os anos de 1994 a 2008**. Dissertação (Mestrado em Biologia de Agentes Infecciosos e Parasitários) – Belém, Universidade Federal do Pará, 2009. 103p.
- CASALS, J.; BROWN, L.V. Hemagglutination with arthropod-borne viruses. The **Journal of Experimental Medicine**, **99**:429–449, 1954.
- CASTLE, E.; NOWAK, T.; LEIDNER, U.; WENGLER, G.; WENGLER, G. Sequence analysis of the viral core protein and the membrane-associated proteins V1 and NV2 of the flavivirus West Nile virus and of the genome sequence for these proteins. **Virology**, **145**: 227-236, 1985.
- CAUSEY, O.R.; CAUSEY C.C.; MAROJA, O.M.; MACEDO, D.G. The isolation of arthropod-borne viruses, including members of two hitherto undescribed serological groups, in the Amazon region of Brazil. **American Journal of Tropical Medicine and Hygiene**, **10**: 227-249, 1961.

- CHAMBERS, T. J.; HALEVY, M.; NESTOROWICZ, A.; RICE, C. M.; LUSTIG, S. West Nile virus envelope proteins: nucleotide sequence analysis of strains differing in mouse neuroinvasiveness. **Journal of General Virology**, **79**, 2375–2380, 1998.
- CHAMBERS, T.J.; NESTOROWICZ, A.; AMBERG, S.G.; RICE, C.M. Mutagenesis of the yellow fever virus NS2B protein: Effects on proteolytic processing, NS2B-NS3 complex formation, and viral replication. **Journal of Virology**, **67**: 6797-6807, 1993.
- CHAMBERS, T.J.; HAHN, C.S.; GALLER, R.; RICE, C.M. Flavivirus genome organization, expression, and replication. **Annual Review of Microbiology**, **44**: 649-688, 1990.
- CHANG, G-J. J. Molecular biology of dengue viruses. In: **Dengue and dengue hemorrhagic fever**. Gubler, D.J.; Kuno, G. (eds). CAB International, Wallingford, UK, 1997. p. 175-198.
- CHANG, G.-J.; KUNO, G.; PURDY, D. E.; DAVIS, B. S. Recent advancement in flavivirus vaccine development. **Expert Review of Vaccines**, **3**(2):199-220, 2004.
- CHANG, H.H.; SHYU, H.F.; WANG, Y.M.; SUN, D.S.; SHYU, R.H.; TANG, S.S.; HUANG, Y.S. Facilitation of cell adhesion by immobilized dengue viral nonstructural protein 1 (NS1): arginine-glycine-aspartic acid structural mimicry within the dengue viral NS1 antigen. **Journal of Infectious Disease**, **186**(6): 743-51, 2002.

- CHANG, G-J. J.; HUNT, A.R.; DAVIS, B. A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice. **Journal of Virology**, **74**: 4244-4252, 2000.
- CHARLIER, N.; LEYSSEN, P.; PLEIJ, C.W.; LEMEY, P.; BILLOIR, F.; VAN LAETHEM, K.; VANDAMME, A.M.; DE CLERCQ, E.; DE LAMBALLERIE, X.; NEYTS, J. Complete genome sequence of Montana Myotis leukoencephalitis virus, phylogenetic analysis and comparative study of the 3' untranslated region of flaviviruses with no known vector. **Journal of General Virology**, **83**(8):1875-85, 2002.
- CHARREL, R.N.; ZAKI, A.M.; ATTOUI, H.; FAKEEH, M.; BILLOIR, F.; YOUSEF, A.I.; DE CHESSE, R.; DE MICCO, P.; GOULD, E.A.; DE LAMBALLERIE, X. Complete coding sequence of the Alkhurma virus, a tick-borne flavivirus causing severe hemorrhagic fever in humans in Saudi Arabia. **Biochemical and Biophysical Research Communications**, **287**(2):455-61, 2001.
- CHEN, C.J.; KUO, M.D.; CHIEN, L.J.; HSU, S.L.; WANG, Y.M.; LIN, J.H. RNA-protein interaction: involvement of NS3, NS5, and 3' noncoding regions of Japanese encephalitis virus genomic RNA. **Journal of Virology**, **71**(5): 3466-3473, 1997a.
- CHEN, Y.; MAGUIRE, T.; HILEMAN, R.E.; FROMM, J.R.; ESKO, J.D.; LINHARDT, R.J.; MARKS, R.M. Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. **Nature Medicine**, **3**(8):866-71, 1997b.

- CHEN, W.R.; TESH, R.B.; RICO-HESSE, R. Genetic variation of Japanese encephalitis virus in nature. **Journal of General Virology**, **71**(12):2915–2922, 1990.
- CHIU, W.W.; KINNEY, R.M.; DREHER, T.W. Control of translation by the 5'- and 3'-terminal regions of the dengue virus genome. **Journal of Virology**, **79**(13): 8303-15, 2005.
- CHUN, P.W.G.; WESTAWAY, E.G. Replication strategy of Kunjin virus: evidence for recycling role of replication form RNA as template in semiconservative and asymmetric replication. **Virology**, **140**: 68-79; 1985.
- CIOTA, A.T.; LOVELACE, A.O.; NGO, K.A.; LE, A.N.; MAFFEI, J.G.; FRANKE, M.A.; PAYNE, A.F.; JONES, S.A.; KAUFFMAN, E.B.; KRAMER, L.D. Cell-specific adaptation of two flaviviruses following serial passage in mosquito cell culture. **Virology**, **357**: 165-174, 2007.
- CLEAVES, G.R.; RYAN, T.E.; SCHLESINGER, R.W. Identification and characterization of type 2 dengue virus replicative intermediate and replicative form RNAs. **Virology**, **111**(1):73-83, 1981.
- CLUM, S.; EDNER, K.E.; PADMANABHAN, R. Cotranslational membrane insertion of the serine proteinase precursor NS2B-NS3 (Pro) of the dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. **The Journal of Biological Chemistry**, **272**(49): 30715-30723, 1997.
- CLYDE, K.; HARRIS, E. RNA secondary structure in the coding region of dengue virus type 2 directs translation start codon selection and is required for viral replication. **Journal of Virology**, **80**(5):2170-82, 2006.

- CLYDE, K. KYLE, J.L.; HARRIS, E. Recent advances in deciphering viral and host determinants of Dengue virus replication and pathogenesis. *Journal of Virology*, **80**(23): 11418-11431, 2006.
- CRABTREE, M. B.; NGA, P. T.; MILLER, B. R. Isolation and characterization of a new mosquito flavivirus, Quang Binh virus, from Vietnam. *Archive of Virology*, **154**: 857–860, 2009.
- CRABTREE, M.B., KINNEY, R.M, MILLER, B.R. Deglycosylation of the NS1 protein of dengue 2 virus, strain 16681: construction and characterization of mutant viruses. *Archive of Virology*, **150**(4): 771-86, 2005.
- CRABTREE, M.B.; SANG, R.C.; STOLLAR, V.; DUNSTER, L.M.; MILLER, B.R. Genetic and phenotypic characterization of the newly described insect flavivirus, Kamiti River virus. *Archive of Virology*, **148**(6):1095-118, 2003.
- CRILL, W.D.; ROEHRIG, J.T. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *Journal of Virology*, **75**(16): 7769-73, 2001.
- COIMBRA, T.L., NASSAR, E.S., NAGAMORI, A.H., FERREIRA, I.B., PEREIRA, L.E., ROCCO, I.M., UEDA-ITO, M., ROMANO, N.S. Iguape: a newly recognized flavivirus from São Paulo State, Brazil. *Intervirology*, **36**(3):144-52, 1999.
- COOK, S.; HOLMES, E.C. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. *Archive of Virology*, **151**(2):309-25, 2006.
- CUI, T. SUGRUE, R.J.; XU, Q.; LEE, A.K.; CHAN, Y.C.; FU, J. Recombinant dengue virus type 1 NS3 protein exhibits specific viral RNA binding and

- NTPase activity regulated by the NS5 protein. **Virology**, **246**(2):409-17, 1998.
- DE LAMBALLERIE, X.; CROCHU, S.; BILLOIR, F.; NEYTS, J.; DE MICCO, P.; HOLMES, E.C.; GOULD, E.A. Genome sequence analysis of Tamana bat virus and its relationship with the genus Flavivirus. **Journal of General Virology**, **83**(10): 2443-54, 2002.
- EBEL, G.D.; SPIELMAN, A.; TELFORD, S.R. Phylogeny of North American Powassan virus. **Journal of General Virology**, **82**:1657-1665, 2001.
- ELIZONDO-QUIROGA, D.; DAVIS, C.T.; FERNANDEZ-SALAS, I.; ESCOBAR-LOPEZ, R.; VELASCO OLMOS, D.; SOTO GASTALUM, L.C.; AVILES ACOSTA, M.; ELIZONDO-QUIROGA, A.; GONZALEZ-ROJAS, J.I.; CONTRERAS CORDERO, J.F.; GUZMAN, H.; TRAVASSOS DA ROSA, A.; BLITVICH, B.J.; BARRETT, A.D.; BEATY, B.J.; TESH, R.B. West Nile Virus isolation in human and mosquitoes, Mexico. **Emerging Infectious Diseases**, **11**(9): 1449-52, 2005.
- ELSHUBER, S.; ALLISON, S.L.; HEINZ, F.X.; MANDL, C.W. Cleavage of protein prM is necessary for infection of BHK-21 cells by tick-borne encephalitis virus. **Journal of General Virology**, **84**(1):183-91, 2003.
- FALGOUT, B.; MARKOFF, L. Evidence that flavivirus NS1-NS2A cleavage is mediated by membrane-bound host protease in endoplasmic reticulum. **Journal of Virology**, **69**(11):7232-7243, 1995.
- FALGOUT, B.; MILLER, R.H.; LAI, C.J. Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B-NS3 protease activity. **Journal of Virology**, **67**(4):2034-42, 1993.

- FALGOUT, B.; PETHEL, M.; ZHANG, Y.M.; LAI, C.J. Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. **Journal of Virology**, **65**(5):2467-75, 1991.
- FALGOUT, B.; CHANOCK, R.; LAI, C-J. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2A. **Journal of Virology**, **63**(5):1852-1860, 1989.
- FELSENSTEIN, J. **PHYLIP**, Versão 3.57c Departamento de Genética, Universidade de Washington, Seattle, W.A. 2009.
- FELSENSTEIN, J. Confidence limits on phylogenies: an approach using the bootstrap. **Evolution**, **39**: 783-791, 1985.
- FIGUEIREDO, L.T.M. The Brazilian flaviviruses. **Microbes and Infection**, **2**:1643-1649, 2000.
- FORATTINI, O.P.; KAKITANI, I.; MASSAD, E.; MARUCCI, D. Studies on mosquitoes (Diptera: *Culicidae*) and anthropic environment. Synanthropy and epidemiological vector role of *Aedes scapularis* in south-eastern Brazil, **Revista da Saúde Pública**, **29**: 199-207, 1985.
- GALINDO P.; SRIHONGE, S.; DE RODANSHINCHE, E.; GRAYSON, M.A.; An ecological survey for arboviruses in Almirante, Panamá, 1959-1962. **American Journal of Tropical Medicine and Hygiene**, **15**(3): 385-400, 1966.
- GAUNT, M.W.; SALL, A.A.; LAMBALLERIE, X; FALCONAR, A.K.I.; DZHIVANIAN T.I.; GOULD, E.A. Phylogenetic relationships of flaviviruses

correlate with their epidemiology, disease association and biogeography.

Journal of General Virology, **82**:1867-76, 2001.

GIBBS, A.J.; BLOK, J.S.M. Molecular Systematic flaviviruses. In: **Symposium of Arbovirus Research in Australia**. Uren, M.F; Mernderson, L.H. (eds). Brisbane: CSIRO and Queensland Inst Med Res; 1989. p.49-51.

GRITSUN, T.S.; FROLOVA, T.V.; ZHANKOV, A.I.; ARMESTO, M.; TURNER, S.L.; FROLOVA, M.P.; POGODINA, V.V.; LASHKEVICH, V.A.; GOULD, E.A. Characterization of a Siberian virus isolated from a patient with progressive chronic tick-borne encephalitis. **Journal of Virology**, **77**:25–36, 2003.

GOMES, G; CAUSEY, O.R. Bussuquara, a new arthropod-borne virus. **Revista do Serviço Especial de Saúde Pública**, p. 103-109, 1959.

GORBALENYA, A.E.; KOONIN, E.V.; Helicases: Amino acid sequence comparisons and the structure-function relationship. **Current Opinion in Structural Biology**, **3**(3): 419-429, 1993.

GORBALENYA, A.E.; KOONIN, E.V.; DONCHENKO, A.P.; BLINOV, V.M. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RBA genomes. **Nucleic Acids Research**, **17**: 4713-4730, 1989.

GOULD, E.A.; LAMBALLERIE, X.; ZANOTTO, P.M.A.; HOLMES, C. Evolution, epidemiology and dispersal of flaviviruses revealed by molecular phylogenies. **Advances Virus Research**, **57**: 71-103, 2003a.

- GOULD, E.A.; DE LAMBALLERIE, X.; ZANOTTO, P.M.; HOLMES, E.C.
Origins, evolution, and vector/host co adaptations within the genus *Flavivirus*. **Advances in Virus Research**, **59**: 277-314, 2003b.
- GOULD, E.A.; DE LAMBALLERIE, X.; ZANOTTO, P.M.; HOLMES, E.C.
Evolution, epidemiology, and dispersal of flaviviruses revealed by molecular phylogenies. **Advances in Virus Research**, **57**: 71-103, 2001.
- GOULD, E.A.; BUCKLEY, A.; CAMMACK, N.; BARRETT, A.D.; CLEGG, J.C.;
ISHAK, R.; VARMA, M.G. Examination of the immunological relationships between flaviviruses using yellow fever virus monoclonal antibodies. **Journal of General Virology**, **66**(7):1369-82, 1985.
- GOTO, A.; YOSHII, K.; OBARA, M.; UEKI, T.; MIZUTANI, T.; KARIWA, H.;
TAKASHIMA, I. Role of the N-linked glycans of the prM and E envelope proteins in tick-borne encephalitis virus particle secretion. **Vaccine**, **23**: 3043-3052, 2005.
- GRITSUN, T.S.; FROLOVA, T.V.; ZHANKOV, A.I.; TURNER, S.L.; FROLOVA,
M.P. POGODINA, V.V.; LASHKEVICH, V.A.; GOULD, E.A. Characterization of a Siberian virus isolated from a patient with progressive chronic tick-borne encephalitis. **Journal of Virology**, **77**: 25-36, 2003
- GUBLER, D.J. Dengue/dengue haemorrhagic fever: history and current status. **Novartis Foundation Symposium**, **277**: 3-16, 2006.
- GUBLER, D.J. Dengue and Dengue hemorrhagic fever: its history and resurgence as a global public health problem. In: **Dengue and Dengue hemorrhagic fever**. Gubler, D.J., Kuno, G. (eds). New York, University Press, Cambridge, 1997. p.1-22.

- GUBLER, D.J.; KUNO, G.; MARKOFF, L. Flavivirus In: **Fields Virology**. Knipe, D.M. & Howley, P.M. (eds), Philadelphia, USA, Wolters Kluwer Health, 2007, p. 1153- 1252.
- GUBLER, D.J.; REITER, P.; EBI, K.L.; YAP, W.; NASCI, R.; PATZ, J.A. Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases. **Environmental Health Perspective**, **109**(2): 223-233, 2001.
- GUINDON, S.; GASCUEL, O. A Simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. **Systematic Biology**, **52**(5): 696-704, 2003.
- GUYATT, K.J.; WESTWAY, E.G.; KHROMYKH, A.A. Expression and purification of enzymatically active recombinant RNA-dependent RNA polymerase (NS5) of the flavivirus Kunjin. **Journal of Virology Methods**, **92**(1): 37-44, 2001.
- HAHN, C.S.; HAHN, Y.S.; RICE, C.M.; LEE, E. DALGARNO, L.; STRAUSS, E.G.; STRAUSS, J.H. Conserved elements in the 3' untranslated of flavivirus RNAs and potential cyclization sequences. **Journal of Molecular Biology**, **198**: 33-41, 1987.
- HALEVY, M.; AKOV, Y.; BEN-NATHAN, D.; KOBILER, D.; LACHMI, B.; LUSTIG, S. Loss of active neuroinvasiveness in attenuated strains of West Nile virus: pathogenicity in immunocompetent and SCID mice. **Archive of Virology**, **137**: 355–370, 1994.
- HALL, R.A.; BROOM, A.K.; HARTNETT, A.C.; HOWARD E, M.J.; MACKENZIE, J. S. Immunodominant epitopes on the NS1 protein of MVE and KUN

- viruses serve as targets for a blocking ELISA to detect virus-specific antibodies in sentinel animal serum. **Journal of Virological Methods**, **51**: 201-210, 1995.
- HALL, T.A. **BioEdit**. Departamento de Microbiologia, Estado da Carolina do Norte, Universidade, Raleigh, NC. 2001.
- HALSTEAD, S.B., JACOBSON, J. Japanese encephalitis. **Advances in Virus Research**, **61**:103-38, 2003.
- HAYES, E.B. Zika virus outside Africa. **Emerging Infectious Disease**, **15**(9): 1347-50, 2009.
- HEINZ, F.X. Epitope mapping of flavivirus glycoproteins. **Advances in virus Research**, **31**: 103-163, 1986
- HEINZ, F.X.; ALLISON, S.L. Flavivirus structure and membrane fusion. **Advances Virus Research**, **59**:63-67, 2003.
- HEINZ FX, STIASNY K, ALLISON SL. The entry machinery of flaviviruses. **Archives of Virology**, **18**:133-137. 2004
- HENRIQUES, D.F. **Infecção persistente pelos flavivírus Ilhéus e Rocio em hamsters dourados jovens (*Mesocricetus auratus*)**. Dissertação (Mestrado em Doenças Tropicais) – Belém, Universidade Federal do Pará, 2009. 99p.
- HOLMES, E.C.; WOROBEY, M.; RAMBAUT, A. Phylogenetic evidence for recombination in dengue virus. **Molecular Biology and Evolution**, **16**:405–409, 1999.
- HOSHINO, K.; HARUHIKO, I.; TSUDA, Y.; SAWABE, K.; KOBAYASHI, M.. Isolation and characterization of a new insect flavivirus from *Aedes*

- albopictus and Aedes flavopictus mosquitoes in Japan. **Virology**, **391**: 119–129, 2009.
- HOSHINO, K.; ISAWA, H.; TSUDA, Y.; YANO, K.; SASAKI, T.; YUDA, M.; TAKASAKI, T.; KOBAYASHI, M.; SAWABE, K. Genetic characterization of a new insect flavivirus isolated from Culex pipiens mosquito in Japan. **Virology**, **359**: 405–414, 2007.
- HSU, M.T.; KUNG, H.J.; DAVIDSON, N. Secondary structure of tRNA. **Cold Spring Harbor Symposia on Quantitative Biology**, **38**:943-950, 1973.
- HUELSENBECK, J. P.; RONQUIST, F. MrBAYES: Bayesian inference of phylogeny and its impact on evolutionary biology. **Science**, **14**(294): 2310-2314, 2001.
- HURRELBRINK, R. J.; MCMINN, P. C. Attenuation of Murray Valley encephalitis virus by site-directed mutagenesis of the hinge and putative receptor-binding regions of the envelope protein. **Journal of Virology**, **75**: 7692–7702, 2001.
- ICTV. **Virus Taxonomy: Classification and Nomenclature of Virus. Eighth Report of the International Committee on the Taxonomy of Virus** In: FAUQUET, C.M.; MAYO. M.A., MANILOFF, J.; DESSEBERGER, U.; BALL, L.A. (eds), San Diego: Elsevier Academic press, 2005, p. 981-998.
- IVERSSON L.B. Rocio Encephalitis, In: **The arboviruses: Epidemiology and ecology**. Monath TP, (ed), CRC Press, Boca Raton, USA, 1988. p-14.
- IVERSSON L.B.; TRAVASSOS DA ROSA, A.P.A.; ROSA M.D.B. Ocorrência recente de infecção humana por vírus Rocio na região do Vale do Ribeira. **Revista Instituto Medicina Tropical de São Paulo**, **21**: 28-31, 1989.

- IVERSSON, L.B.; TRTAVASSOS DA ROSA, A.P.A.; TRAVASSOS DA ROSA, J.F.S.; ELEUTÉRIO, G.C.; PRADO J.A. Estudos sorológicos para pesquisa de anticorpos de arbovírus na população humana d região do Vale do Ribeira. I. Seguimento sorológico de grupo populacional residente em ambiente silvestre. In: Simpósio Internacional sobre Arbovírus dos Trópicos e Febres Hemorrágicas, Belém-PA, 14-18 de abril de 1980. Academia Brasileira de Ciências, Rio de Janeiro. 1982. p.229-243.
- JENKINS, G.M.; PAGEL, M.; GOULD, E.A.; DE, A.; ZANOTTO, P.M.; HOLMES, E.C. Evolution of base composition and codon usage bias in the genus Flavivirus. **Journal of Molecular Evolution**, **52**(4):383-90, 2001.
- JOHNSON, A. J.; GUIRAKHOO, F.; ROEHRIG, J. T. The envelope glycoproteins of dengue 1 and dengue 2 viruses grown in mosquito cells differ in their utilization of potential glycosylation sites. **Virology**, **203**:241–249, 1994.
- KAMER, G.; ARGOS, P. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. **Nucleic Acids Research**, **12**: 7269-7282, 1984.
- KARABATSOS, N. **International Catalogue of Arboviruses, including certain other viruses of vertebrates**, 3rd. ed. Publ. The American Society for Tropical Medicine and Hygiene, San Antonio, Texas, p. 1041, 1985.
- KHROMYKH, A.A.; WESTAWAY, E.G. Subgenomic replication of the flavivirus Kunjin: Construction and amplification. **Journal of Virology**, **71**: 1497-1505, 1997.

- KHROMYKH, A.A.; VARNAVSKI, A.N.; SEDLAK, P.L.; WESTAWAY, E.G.
Coupling between replication and packaging of flavivirus RNA: evidence derived from the use of DNA-based full-length cDNA clones of Kunjin virus. **Journal of Virology**, **75**(10): 4633-4640, 2001a.
- KHROMYKH, A.A.; MEKA, H.; GUYATT, K.J.; WESTAWAY, E.G. Essential role of cyclization sequences in flavivirus RNA replication. **Journal of Virology**, **75**: 6719-6728, 2001b.
- KIERMAYR, S.; KOFLER, R.M.; MANDL, C.W.; MESSNER, P.; HEINZ, F.X.
Isolation of capsid protein dimers from the tick-borne encephalitis flavivirus and in vitro assembly of capsid-like particles. **Journal of Virology**, **78**(15): 8078-84, 2004.
- KING, C-C.; CHAO, D-Y.; CHIEN, L-J. , CHANG, G-J. J.; LIN, T-H.; WU, Y-C.
HUANG, J-H. Comparative analysis of full genomic sequences among different genotypes of dengue virus type 3. **Virology Journal**, **5**(63): 1-13, 2008.
- KOONIN, E.V. Computer assisted identification of a putative methyltransferase domain in NS5 of flaviviruses and lambda of reovirus. **Journal of General Virology**, **74**: 733-740, 1993.
- KOONIN, E.V. The phylogeny of RNA-dependent RNA polymerase of positive strand RNA virus. **Journal of General Virology**, **72**: 2197-2206, 1991.
- KRAMER, L.D.; EBEL, G.D. Dynamics of flavivirus infection in mosquitoes. **Advances in Virus Research**, **60**:187-232, 2003.

- KROSCHEWSKI, H.; ALLISON, S.L.; HEINZ, F.X.; MANDL, C.W. Role of heparan sulfate for attachment and entry of tick-borne encephalitis virus. **Virology**, **308**(1): 92-100, 2003.
- KUNO, G; CHANG, GJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. **Archive of Virology**, **152**: 687-696, 2007.
- KUNO, G.; CHANG, G.J. Characterization of Sepik and Entebbe bat viruses closely related to yellow fever virus. **American Journal of Tropical Medicine and Hygiene**, **75**(6): 1165-1170, 2006.
- KUNO, G.; CHANG, G.J. Biological Transmission of Arboviruses reexamination of the new insights into Components, Mechanisms, and unique traits as well as their evolutionary trends. **Clinical Microbiology Reviews**, **18**: 608-637, 2005.
- KUNO, G.; CHANG, G-J. J.; CHIEN, L-J. Correlation of phylogenetic relation with host range, length of ORF or genes, organization of conserved sequences in the 3' Noncoding region, and viral classification among the members of the genus flavivirus In: **Viral Genomes: Diversity, Proprieties and Parameters**. Feng, Z. Long, M. (eds). Nova Science Publeshers, Inc (in press), 2009.
- KUNO,G., CHANG,G.-J.J. AND TSUCHIYA,K.R. Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. **Clinical Microbiology Review**, **18** (4), 608-637, 2005.

- KUNO, G.; ARTSOB, H.; KARABATSOS, N.; TSUCHIYA, K.R.; CHANG, G.J.
Genomic sequencing of deer tick virus and pathogeny of powassan-related virus of north America. **American Journal of Tropical Medicine and Hygiene**, **65**: 671-676, 2001;
- KUNO, G., CHANG, G.-J. J., TSUCHIYA, K. R., KARABATSOS, N. & CROPP, S. B. Phylogeny of genus *Flavivirus*. **Journal of Virology**, **72**:73-83, 1998.
- LAEMMERT, H.W.; HUGHES, T.P. The virus of Ilheus encephalitis. Isolation, serological specificity, and transmission. **Journal of Immunology**, **55**: 61–67, 1947.
- LANCIOTTI, R.S.; ROEHRIG, J.T.; DEUBEL, V.; SMITH, J.; PARKER, M.; STEELE, K.; CRISE, B.; VOLPE, K.E.; CRABTREE, M.B.; SHERRET, J.H.; HALL, R.A.; MACKENZIE, J.S.; CROPP, C.B.; PANIGRAHY, B.; OSTLUND, E.; SCHIMITT, B.; MARKINSON, M.; BANET, C.; WEISSMAN, J.; KOMAR, N.; SAVAGE, H.M.; STORE, W.; MCNAMARA, T.; GLUBBER, D.J.
Origin of the West Nile virus responsible for an outbreak of encephalitis in the Northeastern United State. **Science**, **286**: 2333-2337, 1999.
- LAWRIE, C.H.; UZCÁTEGUI, N.Y.; ARMESTO, M.; BELL-SAKYI, L.; GOULD E.A. Susceptibility of mosquito and tick cell lines to infection with various flaviviruses. **Medical and Veterinary Entomology**, **18**(3): 268-74, 2004a.
- LAWRIE, C.H.; UZCÁTEGUI, N.Y.; GOULD, E.A.; NUTTALL, P.A. Ixodid and argasid tick species and West Nile virus. **Emerging Infectious Disease**, **10**(4): 653-657, 2004b.

- LEE, E.; LOBIGS, M. Substitutions at the putative receptor-binding site of an encephalitic flavivirus alter virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. **Journal of Virology**, **74**:8867–8875, 2000.
- LEITMEYER, K.C.; VAUGHN, D.W.; WATTS, D.M.; SALAS, R.; VILLALOBOS, I.; DE CHACON; RAMOS, C.; RICO-HESSE, R. Dengue virus structural differences that correlate with pathogenesis. **Journal of Virology**, **73**:4738–4747, 1999.
- LEUNG, D.; SCHRODER, K.; WHITE, H.; FANG, N.X.; STOERMER, M.J.; ABBENANTE, G.; MARTIN, J.L.; YOUNG, P.R.; FAIRLIE, D.P. Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. **The Journal of Biological Chemistry**, **276**(49): 45762-45771, 2001.
- LEYSSSEN, P.; CHARLIER, N.; LEMEY ,P.; BILLOIR, F.; VANDAMME, A.M.; DE CLERCQ, E.; DE LAMBALLERIE, X.; NEYTS, J. Complete genome sequence, taxonomic assignment, and comparative analysis of the untranslated regions of the Modoc virus, a flavivirus with no known vector. **Virology**, **293**(1):125-40, 2002.
- LEWIS, J.G.; CHANG, G.J.; LANCIOTTI, R.S.; KINNEY, R.M.; MAYER, L.W.; TREND, D.W. Phylogenetic relationship of Dengue-2 virus. **Virology**, **197**: 216-224, 1993.
- LI, J.; BHUVANAKANTHAM, R.; HOWE, J.; NG, M-L. The glycosylation site in the envelope protein of West Nile virus (Sarafend) plays an important role in replication and maturation processes. **Journal of General Virology**, **87**: 613-622, 2006.

- LI, W.; LI, Y.; KEDERSHA N.; ANDERSON, P.; EMARA, M.; SWIDEREK, K.M.; MORENO, G.T.; BRINTON, M.A. Cell Protein TIA-1 and TIAR interact with the 3' stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. **Journal of Virology**, **76**(23):11989-2000, 2002.
- LI, H.; CLUM, S. YOU, S.; EBNER, K.E.; PADMANABHAN, R. The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of the dengue virus type 2 NS3 converge within a region of 20 amino acids. **Journal of Virology**, **73**:3108-3116, 1999.
- LIMJINDAPORN T, WONGWIWAT W, NOISAKRAN S, SRISAWAT C, NETSAWANG J, PUTTIKHUNT C, KASINRERK W, AVIRUTNAN P, THIEMMECA S, SRIBURI R, SITTISOMBUT N, MALASIT P, YENCHITSOMANUS PT. Interaction of dengue virus envelope protein with endoplasmic reticulum-resident chaperones facilitates dengue virus production. **Biochemical and Biophysical Research Communications**. **379**(2): 196-200, 2009.
- LIN, S.R.; HSIEH, S.C.; YUEH, Y.Y.; LIN, T.H.; CHAO, D.Y.; CHEN, W.J.; KING, C.C.; WANG, W.K. Study of sequence variation of dengue type 3 virus in naturally infected mosquitoes and human hosts: implications for transmission and evolution. **Journal of Virology**, **78**(22):12717-21, 2004.
- LIN, C. AMBERG, S.M.; CHAMBERS, T.J.; RICE, C.M. Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A/4B signalase site. **Journal of Virology**, **67**:2327-2335, 1993.

- LINDENBACH B.D.; RICE, C.M. Molecular biology of flaviviruses. **Advance of Virus Research**, **59**: 23-61, 2003.
- LINDENBACH, B.D.; RICE, C.M. Trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. **Journal of Virology**, **71**(12): 9608-9617, 1997.
- LINDENBACH, D.B.; THIEL; H-J.; RICE, C.M. Flaviviridae: The viruses and their Replication In: **Fields Virology**. Knipe, D.M. & Howley, P.M. (eds), Philadelphia, USA, Wolters Kluwer Health. 2007. p.1101-1152.
- LINDENBACH, B.D.; EVANS, M.J.; SYDER, A.J.; WÖLK, B.; TELLINGHUISEN, T.L.; LIU, C.C.; MARUYAMA, T.; HYNES, R.O.; BURTON, D.R.; MCKEATING, J.A.; RICE, C.M. Complete replication of hepatitis C virus in cell culture. **Science**, **309**(5734):623-6, 2005.
- LIU, W.J.; CHEN. H.B.; KHROMYKH, A.A. Molecular and functional analyses of Kunjin virus infectious cDNA clones demonstrate the essential role of NS2A in virus assembly and for a nonconservative residue for NS3 in RNA replication. **Journal of Virology**, **77**(14): 7804-7813, 2003.
- LO, M.K.; TILGNER, M.; BERNARD, K.A.; SHI, P.Y. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3' untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. **Journal of Virology**, **77**(18):10004-10014, 2003.
- LOBIGS, M. Flavivirus premembrane protein cleavage and spike heterodimer secretion requires the function of the viral proteinase NS3. **Proceedings of**

the National Academy of Sciences of the United States of America, 90:
6218-6222, 1993.

LOPES O.S.; COIMBRA, T.L.M.; SACCHETTA L.A. CALISHER C.H.
Emergence of a new arbovirus disease in Brazil: I. Isolation and
characterization of the etiologic agent, Rocio virus, **America Journal
Epidemiology, 108:** 444-449, 1978a.

LOPES O.S.; SACCHETTA L.A.; COIMBRA, T.L.M. PINTO, G.H. GLASSNER
C.M. Emergence of a new arbovirus disease in Brazil: II. Epidemiologic
studies on 1975 epidemic. **America Journal Epidemiology, 108:** 394-401,
1978b.

LOLE, K.S.; BOLLINGER, R.C.; PARANJAPE, R.S.; GADKARI, D.; KULKARNI,
S.S.; NOVAK, N.G.; INGERSOLL, R.; SHEPPARD, H.W.; RAY, S.C. Full-
Length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-
Infected Seroconverters in India, with Evidence of Intersubtype
Recombination. **Journal of Virology, 73(1):** 152-160, 1999.

MACKENZIE, J.M.; KHROMYKH, A.A.; JONES, M.K.; WESTAWAY, E.G.
Subcellular localization and some biochemical properties of the flavivirus
Kunjin nonstructural proteins NS2A and NS4A. **Virology, 245(2):**203-15,
1998.

MACKENZIE J.M.; JONES, M.K.; YOUNG, P.R. Immunolocalization of the
dengue virus nonstructutal glycoprotein NS1 suggest a role in RNA
replication. **Virology, 220:** 232-240, 1996.

- MANDL, C.W.; HOLZMANN, H.; KUNZ, C.; HEINZ, F. Complete genomic sequence of Powassan virus: Evaluation of genetic elements in tick-borne versus mosquito-borne flavivirus. **Virology**, **194**:173-184, 1993.
- MARIN, M.S.; ZANOTTO, P.M.; GRITSUN, T.S.; GOULD, E.A. Phylogeny of TYU, SRE, and CFA virus: different evolutionary rates in the genus Flavivirus. **Virology**, **206**(2):1133-9, 1995.
- MARKOFF, L. 5'- and 3'-Noncoding Regions in Flavivirus RNA. **Advances Virus Research**, **59**:177-228, 2003.
- MATHEWS, D.H.; SARBINA, J.; ZUKER, M.; TURNER, D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. **Journal of Molecular Biology**, **288**:911-940, 1999.
- MATUSAN, A.E.; PRYOR, M.L.; DAVIDSON, A.D.; WRIGHT, P.J. Mutagenesis of the Dengue virus type 2 NS3 protein within and outside helicase motifs: Effects on enzyme activity and virus replication. **Journal of Virology**, **75**:9633-9643, 2001.
- MAY, F.J.; LOBIGS, M.; LEE, E.; GENDLE, D.J.; MACKENZIE, J.S.; BROOM, A.K.; CONLAN, J.V.; HALL, R.A. Biological, antigenic and phylogenetic characterization of the flavivirus Alfuy. **Journal of General Virology**, **87**: 329-337, 2006.
- MEN, R.; BRAY, M.; CLARK, D.; CHANOCK, R.; LAI, C. Dengue type 4 virus mutants containing deletions in the 30 noncoding region of the RNA genome: Analysis of growth restriction in cell culture and altered viremia

- pattern and immunogenicity in rhesus monkeys. **Journal of Virology**, **70**: 3930–3937, 1996.
- MILLER, S.; SPARACIO, S.; BARTENSCHLAGER, R. Subcellular localization and membrane topology of the dengue virus type 2 nonstructural protein 4B. **The Journal of Biological Chemistry**, **281**(13): 8854-8863, 2006.
- MITCHELL, C.; MONATH, T. P.; CROPP, C. B. Experimental transmission of Rocio virus by mosquitoes. **American Journal of Tropical Medicine and Hygiene**, **30**: 465-472, 1981.
- MONATH, T.P. Yellow Fever: an update (Reviews). **Lancet Infectious Diseases**, **1**:11-20, 2001.
- MONATH, T.P.; HEINZ, F.X. The Flaviviruses In: **Fields Virology**, Knipe D.M.; Howlay, P.M. (eds), 3rd ed. Vol 1. Philadelphia, Pa: Lippincott-Raven; 1996. p. 961-1034.
- MUÑOZ-JORDÁN, J.L.; LAURENT-ROLLE, M.; ASHOUR, J.; MARTÍNEZ-SOBRIDO, L.; ASHOK, M.; LIPKIN, W.I.; GARCÍA-SASTRE, A. Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. **Journal of Virology**, **79**(13):8004-13, 2005.
- MUÑOZ-JORDÁN, J.L.; SÁNCHEZ-BURGOS, G.G.; LAURENT-ROLLE, M.; GARCÍA-SASTRE, A. Inhibition of interferon signaling by dengue virus. **Proceedings of the National Academy of Sciences of the United States of America**, **100**(24):14333-14338, 2003.
- MUTEBI, J.P.; BARRETT, A.D. The epidemiology of yellow fever in Africa. **Microbes and Infection**, **4**:1459–1468, 2002.

- MUTEBI, J. P.; RIJNBRAND, R. C. A.; WANG, H.; RYMAN, K. D.; WANG, E.; FULOP, L. D.; TITBALL, R.; BARRETT, A. D. T. Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the *Flavivirus* genus based on the 3' noncoding region. **Journal of Virology**, **78**: 9652-9665, 2004.
- MUTEBI, J.P.; RIJNBRAND, R.C.A.; WANG, H.; RYMAN, K.D.; WANG, E.; FULOP, L.D.; TITBALL, R.; BARRET, A.D.T. Genetic relationships and evolution virus of genotypes of Yellow fever virus and others members of the Yellow fever group within *Flavivirus* genus based on the 3' Noncoding region. **Journal of Virology**, **78(18)**: 9652-9665, 2003.
- MUYLAERT, I.R.; CHAMBERS, T.J.; GALLER, R.; RICE, C.M. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. **Virology**, **222(1)**:159-68, 1996.
- NAVARRO-SANCHEZ, E.; ALTMAYER, R.; AMARA, A.; SCHWARTZ, O.; FIESCHI, F.; VIRELIZIER, J.L.; ARENZANA-SEISDEDOS, F.; DESPRÈS, P. Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. **EMBO Reports**, **4(7)**: 723-728, 2003.
- NISBET, D.J.; LEE, K.J.; VAN DEN HURK, A.F.; JOHANSEN, C.A.; KUNO, G.; CHANG, G-J.; MACKENZIE, J.S.; RITCHIE, S.A.; HALL, R.A. Identification of new flaviviruses in the Kokobera virus complex. **Journal of General Virology**, **86**: 121–124, 2005.

- NITAYAPHAN, S.; GRANT, J.A.; CHANG, G.J.; TRENT, D.W. Nucleotide sequence of the virulent AS-14-14-2. **Virology**, **177**: 541-552, 1990.
- NOWAK, T.; FÄRBER, P.M.; WENGLER, G.; WENGLER, G. Analyses of the terminal sequences of West Nile virus structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. **Virology**, **169**(2): 365-76, 1989.
- NOWAK, T.; WENGLER, G. Analysis of disulfides present in the membrane proteins of the West Nile flavivirus. **Virology**, **156**(1): 127-37, 1987.
- PINHEIRO, F.P. **Ilhéus fever** In: Beran, G.W. (ed). Handbook of Zoonosis, Doca Raton, CRC Press, 1994.
- PINHEIRO, F.P.; TRAVASSOS DA ROSA, A.P.A; FREITAS, R.B.; TRAVASSOS DA ROSA, J.F.S.; VASCONCELOS, P.F.C. Aspectos clínico-epidemiológicos. **50 anos do Instituto Evandro Chagas**, Belém, Fundação Serviço de Saúde Pública, **1**:375-407, 1986.
- PIERRE, V.; DROUET, M-T.; DEUBEL, V. Identification of mosquito-borne flavivirus sequences using universal primers and reverse transcription/polymerase chain reaction. **Research of Virology**, **145**: 179-188, 1994.
- POINDINGER, M.; HALL, R.A.; MACKENZIE, J.S. Molecular characterization of the Japanese encephalitis serocomplex of the *Flavivirus* genus. **Virology**, **218**: 417-421, 1996.
- POSADA, D.; CRANDALL, K .A. Modeltest: testing the model of DNA substitution. **Bioinformatics**, **14**(9): 817-818, 1998.

- PROUTSKI, V.; GOULD, E.A.; HOLMES, E.C. Secondary structure of the 3' untranslated region of the flaviviruses: similarities and differences. **Nucleic Acids Research**, **25**: 194-1202, 1997.
- PUGACHEV, K.V.; GUIRAKHOO, F. TRENT, D.W.; MONATH, T.P. Traditional and novel approaches to flavivirus vaccine. **International Journal of Parasitology**, **33**: 567-582, 2003.
- RAY, S.C. **Simplot**. School of Medicine, Johns Hopkins University, 1999.
- RAY, D.; SHI, P.Y. Recent advances in flavivirus antiviral drug discovery and vaccine development. **Recent Patents on Anti-Infective Drug Discovery**, **1**(1):45-55, 2006.
- RICE, C.M.; STRAUSS, J.H. Production of flavivirus polypeptides by proteolytic processing. **Seminars in Virology**, **1**, 357-367, 1990.
- RICE, C.; LENCHER, E.; EDDY, S.; SHIN, S.; SHEETS, R.; STRAUSS, J. Nucleotides sequence of the Yellow fever virus. Implications for flavivirus gene expression and evolution. **Science**, **229**: 726-733, 1985.
- RICO-HESSE, R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. **Virology**, **174**:479-493, 1990.
- ROEHRIG, J.T. Antigenic structure of flavivirus proteins. **Advances in Virus Research**, **59**:141-175, 2003.
- RONQUIST, F.; HUELSENBECK, J.P.; MrBayes: Bayesian Inference of Phylogeny, version 3. **Bioinformatics Application Note**, **17**(8): 754-755, 2003.
- ROOSENDAAL, J.; WESTAWAY, E.G.; KHROMYKH, A.; MACKENZIE, J.M. Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a

- major role in rearranging cytoplasmic membranes and Golgi trafficking of the NS4A protein. **Journal of Virology**, **80**(9): 4623-4632, 2006.
- ROSEN, L. Sur le mechanism de la transmission verticale du virus de la dengue chez les moustiques. **Compters Rendus de l'Academie des Sciences**, **304**: 357-350, 1987
- SABIN, A.B. Survey of knowledge and problems in field of arthropod-borne virus infections. **Archiv für Gesant Virusforschung**, **9**: 1-10, 1959.
- SANGER, F., NICKLEN, S., COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, **74**: 5463-5467, 1977.
- SAMPATH, A., PADMANABHAN, R. Molecular targets for flavivirus drug Discovery. **Antiviral Research**, **81**: 6–15, 2009.
- SCARAMOZZINO, N.; CRANCE, J-M.; JOUAN, A., DEBRIEL, D. A.; STOLL, F.O., GARIN, D. Comparison of Flavivirus Universal Primer Pairs and Development of a Rapid, Highly Sensitive Heminested Reverse Transcription-PCR Assay for Detection of Flaviviruses Targeted to a Conserved Region of the NS5 Gene Sequences. **Journal of Clinical Microbiology**, **39**(5): 1922–1927, 2001.
- SELIGMAN, S.J. Constancy and diversity in the flavivirus fusion peptide. **Virology Journal**, **5**(27) 1-10, 2008.
- SELIGMAN, S.J.; BUCHER, D.J. The importance of being outer: consequences of the distinction between the outer and inner surfaces of flavivirus glycoprotein E. **Trends Microbiology**, **11**: 108-110, 2003.

- SCHEIDER, A.P.; WESTAWAY, E.G. Translation mapping with the flavivirus Kunjin: Gene order and anomalies in translation of NS5. **Virus & Research**, **9**: 323-333, 1988.
- SCHERRET, J. H.; MACKENZIE, J. S.; KHROMYKH, A. A. HALL, R. A. Biological significance of glycosylation of the envelope protein of Kunjin virus. **New York Academy of Sciences**, **951**, 361–363. 2001.
- SCHLESINGER, J.J.; PUTNAK, J.R.; ECKELS, K.H. New approaches to flavivirus vaccine development. **Biotechnology**, **20**:289-307, 1992.
- SHIRATO, K.; MIYOSHI, H.; GOTO, A.; AKO, Y.; UEKI, T.; KARIWA, H.; TAKASHIMA, I. Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus. **Journal of General Virology**, **85**: 3637–364, 2004.
- SHU, P-Y., CHEN, L-K.; CHANG, S-F.; SU, C-L.; CHIEN, L-J.; CHIN, C., LIN, T-H.; HUANG, J-H. Dengue Virus Serotyping Based on Envelope and Membrane and Nonstructural Protein NS1 Serotype-Specific Capture Immunoglobulin M Enzyme-Linked Immunosorbent Assays. **Journal of Clinical Microbiology**, **42**(6): 2489-2494, 2004.
- SPEIGHT, G.; COIA, G.; PARKER, M.D.; WESTAWAY, E.G. Gene mapping and positive identification of the non-structural proteins NS2A, NS2B, NS3, NS4B and NS5 of the flavivirus Kunjin and their cleavage sites. **Journal of General Virology**, **69**:23-34, 1988.
- SRIHONGSE, S.; JOHNSON, C.M. The first isolation of Bussuquara virus from man. **Transactions of the Royal Society of Tropical Medical Hygiene**, **65**: 541-542, 1971.

- STADLER, K.; ALLISON, S.L.; SCHALICH, J.; HEINZ, F.X. Proteolytic activation of tick-borne encephalitis virus by furin. **Journal of Virology**, **71**(11): 8475-81, 1997.
- STAPLES JE, MONATH TP. Yellow fever: 100 years of discovery. **JAMA: The Journal of American Medical Association**, **300**(8): 960-962, 2008.
- STIASNY, K.; KIERMAYR, S.; HOLZMANN, H.; HEINZ, F.X. Cryptic Properties of a Cluster of Dominant Flavivirus Cross-Reactive Antigenic Sites. **Journal of Virology**, **80**(19): 9557–9568, 2006.
- STRAATMANN, A.; SANTOS-TORRES, S.; VASCONCELOS, P.F.C.; TRAVASSOS DA ROSA, A.P.A.; RODRIGUES, S.G.; TAVARES-NETO, J. Evidências sorológicas da circulação do arbovírus Rocio (Flaviviridae) na Bahia. **Revista da Sociedade Brasileira de Medicina Tropical**, **30**(6): 511-515, 1997.
- SWOFFORD, D. L. **PAUP: Phylogenetic Analysis Using Parsimony and Other Methods**, Version 4. Sunderland, MA: Sinauer Associates, 2002.
- TASSANEETRITHEP, B.; BURGESS, T.H.; GRANELLI-PIPERNO, A.; TRUMPFHELLER, C.; FINKE, J.; SUN, W.; ELLER, M.A.; PATTANAPANYASAT, K.; SARASOMBATH, S.; BIRX, D.L.; STEINMAN, R.M.; SCHLESINGER, S.; MAROVICH, M.A. DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. **The Journal of Experimental Medicine**, **197**(7): 823-829, 2003.
- THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.J. The Clustal X window interface: flexible strategies for

- multiple sequence alignment aided by quality analysis tools. **Nucleic Acidic Research**, **25**:4876-4882, 1997.
- TILGNER, M.; DEAS, T.S.; SHI, P.Y.; The flavivirus conserved pentanucleotide in the 3' stem-loop of the West Nile virus genome requires a specific sequence and structure for synthesis, but not viral translocation. **Virology**, **331**(2): 375-386, 2005.
- TIRIBA, A.C.; MIZIARA, A.M.; LOURENÇO, R.; COSTA, C.R.B.; COSTA, C.S.; PINTO, G.H. Encefalite humana primária epidêmica por arbovírus observada no litoral sul de São Paulo. **Revista da Associação Médica Brasileira**, **22**:415, 1976.
- TONRY, J.H.; BROWN, C.B.; CROPP, C.B.; CO, J.K.; BENNETT, S.N.; NERURKAR, V.R.; KUBERSKI, T.; GUBLER, D.J. West Nile virus detection in urine. **Emerging Infectious Disease**, **11**(8):1294-6, 2005.
- TOLOU, H.J.; COUISSINIER-PARIS, P.; DURAND, J.P., MERCIER, V.; DE PINA, J.J.; DE MICCO, P.; BILLOIR, F.; CHARREL, R.N.; DE LAMBALLERIE, X. Evidence for recombination in natural populations of dengue virus type 1 based on the analysis of complete genome sequences. **Journal of General Virology**, **82**:1283–1290, 2001.
- TRAVASSOS DA ROSA, J.F.S., TRAVASSOS DA ROSA, A.P.A., VASCONCELOS, P.F.C., PINHERO, F.P., RODRIGUES, S.G., TRAVASSOS DA ROSA, E.S., DIAS, L.B., CRUZ, A.C.R. Arboviruses isolated in Evandro Chagas Institute, including some described from the first time in Brazilian Amazon Region: their known hosts, and their pathology for man. In: **An overview of Arbovirology in Brazil and neighboring**

- countries.** Travassos da Rosa, A.P.A., Vasconcelos, P.F.C. Travassos da Rosa, J.F.S. (ed.), 1998, p.19-31.
- TRAVASSOS DA ROSA, A.P.A., TRAVASSOS DA ROSA, J.F.S., VASCONCELOS, P.F.C., PINHERO, F.P. Arboviroses. *In: Doenças Infecciosas e Parasitárias.* LEÃO, R.N.Q. (ed.), Belém: Cejup: UEPA: Instituto Evandro Chagas, 1997, p. 207-226.
- TWIDDY, S.S.; HOLMES, E.C. The extent of homologous recombination in members of the genus *Flavivirus*. **Journal of General Virology**, **84**: 429–440, 2003.
- TWIDDY, S.S.; HOLMES, E.C., RAMBAUT, A. Inferring the rate and time scale of dengue virus evolution. **Molecular Biology and Evolution**, **20**: 122–129, 2003.
- ULLOA, A., FERGUSON, H.H., MÉNDEZ-SÁNCHEZ, J.D., DANIS-LOZANO, R., CASAS-MARTÍNEZ, M., BOND, J.G., GARCÍA-ZEBADÚA, J.C., OROZCO-BONILLA, A., JUÁREZ-ORDAZ, J.A., FARFAN-ALE, J.A., GARCÍA-REJÓN, J.E., ROSADO-PAREDES, E.P., EDWARDS, E., KOMAR, N., HASSAN, H.K., UNNASCH, T.R., RODRÍGUEZ-PÉREZ, M.A. West Nile Virus Activity in Mosquitoes and Domestic Animals in Chiapas, México. **Vector Borne Zoonotic Disease**, **9**(5): 555-560, 2009.
- UZCATEGUI, N.Y.; CAMACHO, D.; COMACH, G.; CUELLO DE UZCATEGUI, R.; HOLMES, E.C.; GOULD, E.A. Molecular epidemiology of dengue type 2 virus in Venezuela: evidence for in situ virus evolution and recombination. **Journal of General Virology**, **82**: 2945–2953, 2001.

- VAN DER MOST, R. G.; CORVER, J.; STRAUSS, J. H. Mutagenesis of the RGD motif in the yellow fever virus 17D envelope protein. **Virology**, **265**: 83–95, 1999.
- VASCONCELOS, P.F.C Febre Amarela. **Revista da Sociedade Brasileira de Medicina Tropical**, **36**: 275-293, 2003.
- VASCONCELOS, P.F.C.; BRYANT, J.E.; TRAVASSOS DA ROSA, A.P.A.; TESH, R.B.; RODRIGUES, S.G.; BARRETT, A.D.T. Genetic divergence and dispersal of yellow fever virus, Brazil. Atlanta, EUA. **Emerging Infectious Diseases**, **10**(9): 1578-1584, 2004.
- VASCONCELOS, P.F.C.; TRAVASSOS DA ROSA, A.P.A.; PINHEIRO, F.P.; SHOPE, R.E.; TRAVASSOS DA ROSA, J.F.S.; DEGALLIER, N.; TRAVASSOS DA ROSA, E.S. Arboviruses pathogenic for man in Brazil In: **An Overview of Arbovirology in Brazil and Neighboring Countries**. 1998. p. 72-96.
- VASCONCELOS, P.F.C., TRAVASSOS DA ROSA, A.P.A., DÉGALLIER, N., TRAVASSOS DA ROSA, J.F.S., PINHEIRO, F.P. Clinical and ecoepidemiological situation of human arboviruses in Brazilian Amazonia. **Ciência e Cultura** (Journal of the Brazilian Association for the Advancement of Science), **44**: 117-124, 1992.
- VORNDAM, V.; MATHEWS, J.H.; BARRETT, A.D.T.; ROEHRIG, J.T.; TRENT D.W. Molecular and biological characterization of a non-glycosylated isolate of St. Louis encephalitis virus. **Journal of General Virology**, **74**: 2653-2660, 1993.

- VILLORDO, S.M.; GAMARNIK, A.V. Genome cyclization as strategy for flavivirus RNA replication. **Virus Research**, **139**: 230-239, 2009.
- WARRENER, P.; TAMURA, J.K.; COLLETT, M.S. An RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. **Journal of Virology**, **67**: 989-996, 1993.
- WESTAWAY ET AL, E.G.; MACKENZIE, J.M.; KHROMYKH, A.A. Kunjin RNA replication and applications of the KUJIN replicons. **Advances Virus Research**, **59**: 99-140, 2003.
- WHO. Dengue guidelines for diagnosis, treatment, prevention and control. **Technical Report Series**, World Health Organization, Geneva, 2009, p. 157
- WHO Scientific Group. Arthropod-borne and rodent-borne viral disease. **Technical Report Series**, N° 719. World Health Organization, Geneva, 1985.
- WINKLER, G.; HEINZ, F.X.; KUNZ, C. Studies on the glycosylation of flavivirus E proteins and the role of carbohydrates in antigenic structure. **Virology**, **159**: 237-243, 1987.
- WOROBAY, M.; RAMBAUT, A.; HOLMES, E.C. Widespread intraserotype recombination in natural populations of dengue virus. **Proceedings of the National Academy of Sciences of the United States of America**, **96**:7352–7357, 1999.
- YOCUPICIO-MONROY, R.M.; MEDINA, F.; REYES-DEL VALLE, J.; DEL ANGEL, R.M. Cellular proteins from human monocytes bind to dengue 4 virus minus-strand 3' untranslated region RNA. **Journal of Virology**, **77**(5): 3067-76, 2003.

- YU, L.; MARKOFF, L. The topology of bulges in the long stem of the flavivirus 3' stem-loop is a major determinant of RNA replication competence. **Journal of Virology**, **79**(4): 2309-24, 2005.
- ZANOTTO, P.M.A., GOULD, E.A., GAO, G.F., HARVEY, P.H.; HOLMES, E.C. Population dynamics of flaviviruses revealed by molecular phylogenies. **Proceedings of the National Academy of Sciences of the United States of America**, **93**: 548-553, 1996.
- ZANOTTO, P.M.A.; GAO, G.F.; VENUGOPAL, K.; GRITSUN, T.S.; MARIN, S.; JIANG, W.R.; REID, H.W.; GOULD, E.A. An RNA virus cline across the Northern Hemisphere. **Virology**, **210**: 152-159, 1995.
- ZHOU, Y.; RAY, D.; ZHAO, Y.; DONG, H.; REN, S. LI, Z.; GUO, Y.; BERNARD, K. A.; SHI, P-Y. ; LI, H. Structure and Function of Flavivirus NS5 Methyltransferase. **Journal of Virology**, **81**(8):3891–3903, 2007.
- ZUKER, M.; MATHEWS D.H.; TUNER, D.H. Algorithms and therdynamics for RNA secondary structure prediction: a practical guide. In: **RNA biochemistry and biotechnology**. Barciszewski, J. Clark, B.F.C. (eds). Kluwer Academic Publishers, Amsterdam, 1999. p.11-43.

ANEXOS

ANEXO I

Documento de Autorização para a utilização das amostras dos vírus Bussuquara, Iguape, Ilhéus e Rocio fornecida pela direção do Instituto Evandro Chagas.

AUTORIZAÇÃO ¹

Autorizamos a utilização das amostras dos vírus Bussuquara (BE AN 4073), Iguape (SP AN 71686), Ilhéus (BE H 7445) e Rocio (SP H 34675) pertencentes à Seção de Arbovirologia e Febres Hemorrágicas a fim de desenvolver a tese de doutorado intitulada "**CARACTERIZAÇÃO MOLECULAR DO GENOMA COMPLETO DOS ARBOVÍRUS BUSSUQUARA, IGUAPE, ILHÉUS E ROCIO (FAMÍLIA FLAVIVIRIDAE, GÊNERO FLAVIVIRUS)**" pela aula de pós-graduação **Daniele Barbosa de Almeida Medeiros**, sob a orientação do pesquisador Prof. Dr. Pedro Fernando da Costa Vasconcelos e co-orientação do Prof. Dr. Márcio Roberto Teixeira Nunes.

Ananindeua, Pará, 20 de Agosto de 2006


ELISABETH C. DE OLIVEIRA SANTOS
Diretora da Instituto Evandro Chagas
Secretaria de Vigilância em Saúde / MS

ANEXO II

Distribuição dos tamanhos da sequências aminoacídicas entre os genes individuais e ORF dos quatro principais grupos dos flavivírus.

Grupos	Vírus		genes e tamanho das proteínas									ORF		
			Estrutural			Não-estruturais								
			C	PrM	E	NS1	NS2A	NS2B	NS3	NS4A/2K	NS4B		NS5	
Flavivírus Transmitidos por mosquitos	BSQV		125	167	501	352	227	131	619	149	253	905	3.429	
	IGUV		119	167	494	353	227	131	619	149	254	904	3.417	
	ILHV		118	167	501	353	227	131	618	149	255	905	3.424	
	ROCV		118	167	501	353	227	131	619	149	256	904	3.425	
	JEV*	SLEV		121	167	501	352	227	131	618	149	258	906	3.430
		WNV		123	167	501	352	231	131	619	149	255	905	3.433
		JEV		127	167	500	352	227	131	619	149	255	905	3.432
	DENV*	DENV-1		113	166	495	352	218	130	619	150	249	899	3.391
		DENV-2		114	166	495	352	218	130	615	150	248	900	3.388
		DENV-3		114	166	493	352	218	130	619	150	248	901	3.391
		DENV-4		113	166	495	352	218	130	618	150	245	900	3.387
	YFV		121	164	493	352	224	130	623	149	250	905	3.411	
	Flavivírus Transmitidos por Carrapatos	TBEV		112	168	496	352	230	131	621	149	252	903	3.414
POWV		110	168	497	353	230	131	622	149	252	904	3.416		
Agentes Zoonóticos	MODV		110	162	482	353	221	132	618	144	254	898	3.374	
	RBV		102	165	484	353	229	130	617	144	258	897	3.379	
Flavivírus de insetos	KRV		143	143	432	390	232	124	577	168	261	887	3.357	

* subgrupos; em cinza os vírus em estudo.

ANEXO III

Código Genético

		2ª BASE													
		U		C		A		G							
1ª BASE	U	UUU	Fen	UCU	Ser	UAU	Tir	UGU	Cis	3ª BASE	U				
		UUC		UCC				UAC				UGC		C	
		UUA	Leu	UCA				UAA			Códon de Finalização	UGA	Códon de Finalização	A	
		UUG		UCG				UAG				UGG	Trp	G	
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg		U				
		CUC				CCC					CAC		CGC		C
		CUA				CCA					CAA	Gln	CGA		A
		CUG				CCG					CAG		CGG		G
	A	AUU	Ile	ACU	Tre	AAU	Asn	AGU	Ser		U				
		AUC		ACC				AAC				AGC		C	
		AUA	Met	ACA				AAA			Lis	AGA	Arg	A	
		AUG		ACG				AAG				AGG		G	
	G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gli		U				
		GUC				GCC					GAC		GGC		C
		GUA				GCA					GAA	Glu	GGA		A
		GUG				GCG					GAG		GGG		G

Legenda: Vermelho – códon de iniciação; azul – códon de finalização

Simbologia para substituição nucleotídica		
B	C ou G ou T	
D	A ou G ou T	
H	A ou C ou T	
V	A ou C ou G	
R	A ou G	(puRine)
Y	C ou T	(pYrimidine)
K	G ou T	Keto
M	A ou C	aMino
S	G ou C	Strong – 3 pontes de H
W	A ou T	Weak – 2 pontes de H
N	qualquer base	

ANEXO IV

Nomes e Abreviaturas dos Aminoácidos comuns

Aminoácido	Abreviatura	Carga	Hidropacidade	Estrutura	
Alanina	Ala	A	Neutro	Apolar	Acíclica
Arginina	Arg	R	Básico	Polar	Acíclica
Asparagina	Ans	N	Neutro	Polar	Acíclica
Ácido aspártico	Asp	D	Ácida	Polar	Acíclica
Cisteína	Cis / Cys	C	Neutro	Polar	Acíclica
Fenilalanina	Fen /Phe	F	Neutro	Apolar	Cíclica
Glicina	Gli / Gly	E	Neutro	Polar	Acíclica
Glutamina	Gln	G	Neutro	Polar	Acíclica
Ácido glutâmico	Glu	Q	Ácida	Polar	Acíclica
Histidina	His	H	Básico	Polar	Cíclica
Isoleucina	Ile	I	Neutro	Apolar	Acíclica
Leucina	Leu	L	Neutro	Apolar	Acíclica
Lisina	Lis / Lys	K	Básico	Polar	Acíclica
Metionina	Met	M	Neutro	Apolar	Acíclica
Prolina	Pro	P	Neutro	Apolar	Cíclica
Serina	Ser	S	Neutro	Polar	Acíclica
Treonina	Ter / Thr	T	Neutro	Polar	Acíclica
Tirosina	Tir / Tyr	Y	Neutro	Polar	Cíclica
Tritofano	Trp	W	Neutro	Apolar	Cíclica
Valina	Val	V	Neutro	Apolar	Acíclica

* No caso de aminoácidos indeterminados ou não padrão usa-se a letra "X".

ANEXO V

Mutações no peptídeo de fusão

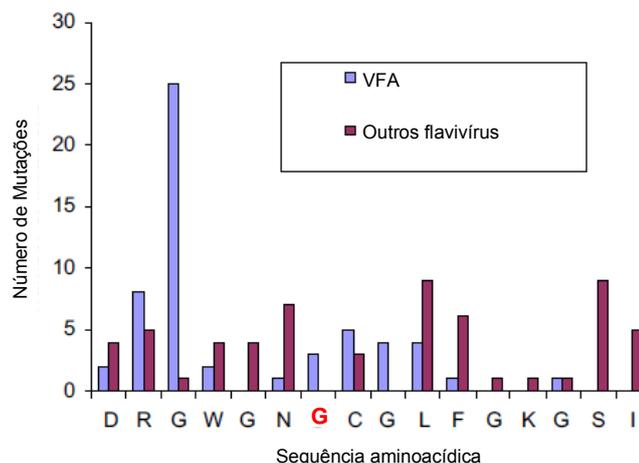


Figura 28 – Número de mutações no peptídeo de fusão dos flavivírus segundo Seligman (2008). Nota-se que as posições N104, L108, F109, S113 e I114 apresentam maior variação entre os flavivírus. Em vermelho, destaca-se o sétimo aminoácido, representado por G105 nos flavivírus transmitidos por mosquitos e que corresponde a H105 nos flavivírus transmitidos por carrapatos, e que são altamente conservados entre esses dois grupos.

Tabela 5 – Recíproca das relações antigênicas pelo teste de FC entre os flavivírus do VROC e VILH com os integrantes do grupo Ntaya (VNTA) e encefalite japonesa (VEJ, VNO e VESL).

Antígeno	Teste de fixação do complemento						Controle (veronal)
	Anticorpo						
	VROC	VILH	VNTA	VEJ	VESL	VNO	
VROC	256+ ≥512	64 ≥512	0	16 64	128 ≥512	512 ≥512	0
VILH	64 128	256 ≥512	8 32	32 128	128 128	512 ≥512	0
VNTA	32 8	32 32	512 128	16 8	256 128	512 128	0
VEJ	32 8	32 32	0	512 128	256 128	>1024 128	0
VESL	128 32	128 32	8 8	128 128	>1024 ≥512	>1024 ≥512	0
VNO	0	0	0	8 8	32 32	>1024 128	0

Legenda: Em vermelho, destacam-se os títulos recíprocos entre antígenos e anticorpos homólogos. Os quadrantes marcados mostram: a relação antigênica do VILH e VROC entre si (verde); destes com o grupo encefalite japonesa (em amarelos); e destes com o VNTA (cinza).

Fonte: Travassos da Rosa, A.P.A., UTMB.

ANEXO VI

MEDEIROS, D.B.; NUNES, M.R.; VASCONCELOS, P.F.; CHANG, G-J.; KUNO, G. Complete genome characterization of Rocio virus (Flavivirus: Flaviviridae), a Brazilian flavivirus isolated from a fatal case of encephalitis during an epidemic in Sao Paulo state. **Journal of General Virology**, **88**(8):2237-46, 2007.

Complete genome characterization of Rocio virus (*Flavivirus: Flaviviridae*), a Brazilian flavivirus isolated from a fatal case of encephalitis during an epidemic in São Paulo state

Daniele B. A. Medeiros,¹ Márcio R. T. Nunes,¹ Pedro F. C. Vasconcelos,¹ Gwong-Jen J. Chang² and Goro Kuno²

Correspondence

Pedro F. C. Vasconcelos
pedrovasconcelos@iec.pa.gov.br

¹Seção de Arbovirologia e Febres Hemorrágicas, Instituto Evandro Chagas, Secretaria de Vigilância em Saúde, Ministério da Saúde, Belém, Pará, Brasil

²Arboviral Diseases Branch, Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention (CDC), Fort Collins, CO, USA

The flaviviruses of major medical importance in South American countries are yellow fever, dengue, Saint Louis encephalitis, West Nile and Rocio viruses. Rocio virus (ROCV) has been responsible for epidemics of severe encephalitis in Brazil with a case-fatality rate of 10% and development of sequelae in 20% of the survivors. We have sequenced and characterized the entire genome of ROCV for the first time, by determining the general traits of the open reading frame and the characteristics of viral genes including the potential cleavage sites, conserved or unique motifs, cysteine residues and potential glycosylation sites. The conserved sequences in the 3'-non-coding region were identified, and the predicted secondary structures during cyclization between 5'- and 3'-non-coding regions were studied. Multiple protein and phylogenetic analyses based on antigenically important and phylogenetically informative genes confirmed a close relationship between ROCV and Ilheus virus (ILHV), together constituting a unique and distinct phylogenetic subgroup as well as the genetic relationship of ROCV with several members of the Japanese encephalitis group. Although ROCV is phylogenetically related to ILHV, our study shows that it is still a virus distinct from the latter virus. This is the first flavivirus uniquely indigenous to Brazil that has been sequenced completely and the genome characterized. The data should be useful for further studies at the molecular level, including construction of infectious clone, identification of gene function, improved disease surveillance based on molecular diagnostic tools and vaccine development.

Received 25 January 2007

Accepted 7 April 2007

INTRODUCTION

In South America, until the mid-1970s, the two neurotropic flaviviruses had been Saint Louis encephalitis virus (SLEV) and Ilheus virus (ILHV). However, in contrast to North America, epidemic of encephalitis in humans by SLEV in South America was not recorded for many years until the first outbreak in Argentina in 2005 (Diaz *et al.*, 2006), despite accumulated records of virus isolation from mosquitoes and vertebrate hosts throughout the continent (Spence, 1980). As for the other virus, ILHV, human infection was sporadic and development of neurotropic manifestations has been even more infrequent. Thus, the

sudden emergence of an outbreak of encephalitis by a flavivirus in 1974–75 in Brazil was unexpected.

Rocio virus (ROCV) belongs to the genus *Flavivirus* in the family *Flaviviridae*. Its prototype (strain SPH 34675) was isolated in Ribeira Valley in the south-eastern São Paulo state of Brazil in 1975 from the cerebellum of a fatal human case during an epidemic of encephalitis that could not be readily diagnosed with the reagents then available for known encephalitic viruses. The epidemic, which began in 1974, spread to more than 20 municipalities over the following 2 years causing approximately 1000 diagnosed cases.

Members of the genus *Flavivirus* (hereafter called flaviviruses) are positive-sense, single-stranded RNA viruses that infect humans and other vertebrates. They are mainly transmitted by bites of haematophagous arthropods (either mosquitoes or ticks), except for a group of vertebrate

The GenBank/EMBL/DBJ accession number for the sequence obtained in this study is AY632542-R1.

Supplementary tables are available with the online version of this paper.

D. B. A. Medeiros and others

viruses that do not need arthropod-borne mode of transmission for survival (hereafter called no-vector viruses) and an increasing number of arthropod flaviviruses that replicate only in arthropods (mosquitoes). Currently, in this genus, more than 60 distinct viruses have been recognized worldwide with several others being classified as either tentative species or genotypes of the other viruses (Gould *et al.*, 2003; ICTV, 2005). The genomic RNA consists of a unique open reading frame (ORF) that is flanked by a type 1 capped 5'-terminal non-coding region (NCR) and a 3'-terminal NCR that, together, form specific secondary stem-loop structures required for RNA translation, replication and/or expression of biological traits most likely including pathogenetic determinants (Lindenbach & Rice, 2003; Rice *et al.*, 1985). The flavivirus ORF encodes 10 proteins including three structural [capsid (C), pre-membrane (prM) and envelope (E)] and seven non-structural (NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5) proteins (Chambers *et al.*, 1990).

Although ROCV was found to be antigenically related to ILHV and grouped in the Japanese encephalitis virus (JEV) antigenic complex (Chamberlain, 1980), a phylogenetic study based on partial NS5 gene revealed a close but distinct relation of ROCV and ILHV from the JEV complex viruses (Kuno *et al.*, 1998). However, fundamental studies based on the full-genome sequence are desirable for identifying in more depth the phylogenetic relationship, genetic traits and other molecular traits of these neurotropic viruses. In this report, we selected ROCV as the first virus for such investigations. The data obtained from such studies should, in turn, facilitate vaccine preparation and designing specific primers and probes for a better diagnosis and surveillance of Rocio encephalitis. Recently, evidence of human infection by ROCV was again reported (Straatmann *et al.*, 1997); and another outbreak in the future is highly possible. Thus, the study undertaken was timely because the diagnosis of flaviviral infections in South America has become more complicated due to recent introduction of the West Nile virus (WNV) and occurrence of unusual haemorrhagic manifestations in laboratory-confirmed SLE cases that resulted in confusion with dengue in Brazil (Mondini *et al.*, 2007).

In order to meet such an epidemiological threat effectively, we have determined and characterized the full-length genome sequence of ROCV (strain SPH 34675) that was obtained from the cerebellum of a fatal encephalitis case in 1975 (Iversson, 1988; de Sousa Lopes *et al.*, 1978, 1981).

METHODS

Virus. Sequencing of ROCV was conducted in two laboratories [CDC and Instituto Evandro Chagas (IEC)] as a collaborative research study, using the identical strain (SPH 34675) at similar mouse brain passage levels (level 4 in CDC and level 5 in IEC).

RNA extraction, reverse-transcription, PCR and nucleotide sequencing. Viral RNA was extracted directly from a suspension

of infected mouse brain using the QIAamp Viral RNA Mini kit (Qiagen) according to the manufacturer's instruction. A cDNA spanning from the 5' terminus to the conserved sequence (CS2) in the 3'-NCR was prepared by reverse transcribing viral RNA using reverse transcriptase and a reverse primer VD8 (5'-GGGTCTCTCTAAC-CTCTAG-3') previously designed from the CS2 sequence (Pierre *et al.*, 1994). PCR was performed using the Expand Long Template PCR System kit (Roche Applied Science). The thermocycling programme set up in the Gene Amp PCR System 9600 thermocycler (Perkin-Elmer) was 1 cycle of 94 °C for 1 min, 50 °C for 5 min; 3 cycles of 94 °C for 20 s, 50 °C for 1 min, 68 °C for 4 min; 10 cycles of 94 °C for 20 s, 50 °C for 30 s, 68 °C for 4 min with an increment of 20 s per cycle; and a final cycle of extension at 68 °C for 7 min. PCR products (amplicons) were purified with a Centriprep column (Princeton Separations), and aliquots of approximately 60–160 ng of the purified DNA templates were used for direct cycle sequencing by using Prism DNA sequencing kit (Big Dye) for dye terminator cycle sequencing with Ampli-Taq FS enzyme (ABI) and CEQ8000 Genetic Analysis System (Beckman Coulter).

Because the full account of the sequencing strategy based on primer walking as well as a list of degenerate primers used were described earlier (Kuno & Chang, 2007) and to avoid redundancy, only the essential features of the technique are briefly presented here. For sequencing with this cDNA, a number of degenerate primers were designed either in one or in both directions, based on the conserved amino acid sequences in the non-structural genes (NS), previously identified in a multiple sequence alignment of flaviviral sequences available in GenBank. The gene and amino acid sequences selected for such degenerate primers are as follows: NS3 (GTSGSPI, GLYGNG, LAPTRVV, DVMCHATF, MDEAHF, SIAARGY, MTATPPG, ISEMGAN and SAAQRRGR) and NS5 (DLGCGRG, SRNSTHEMY, NMMGKREKK, ADDTAGWDT and WMTTEDML). Degenerate primers were also designed from less conserved genes: envelope (E) (DRGWGNGC, GLFGKGS, GHLKCRV, PFGDSYIV and DTAWD-FGS), NS1 (GCWYGMEI and YGMEIRP). The three forward primers in the 5'-NCR that were empirically determined are 5'-ATG(A/T)CTAA(A/G)AAACCAGGA-3', 5'-AAACCAGGA(G/A)(G/A)(G/A)CCCGG-3' and 5'-TCAATATGCT(A/G)AAACGCGG-3'.

In the first stage of sequencing, using pairs of forward and reverse degenerate primers, multiple amplicons spanning any regions within the cDNA were generated. Primers were selected so that the lengths of the amplicons were preferably less than 3.5 kb. The generated amplicons were sequenced first using the same amplification primers. Specific internal primers in both directions were designed from the sequences obtained. Those internal primers were used for further sequencing of the amplicons. In the second stage, non-overlapping gaps between sequenced regions or between amplified genomic regions were amplified and sequenced. The 5'- and 3'-ends of the genome were first poly-C-tailed or poly-A-tailed, respectively, and then amplified using 5'- or 3'-RACE kits, according to the manufacturer's instructions (Invitrogen).

Assembly of overlapping sequences for full-length genome and multiple sequence alignment. Partial sequences of ROCV were first inspected for quality and then assembled for generating a contiguous full-genome sequence, by using the SeqMan II program (v. 5.03) in LaserGene program (DNA Star). This software was also used to predict the entire ORF for the polyprotein of ROCV as well as to determine the 5'- and 3'-NCRs. The full-length genome sequence of ROCV was deposited in GenBank as revised version 1 (AY632542-R1) of the ORF sequence previously deposited (ROCV: AY632542), by including the 5'- and 3'-NCR sequences. Multiple sequence alignment was obtained first with the use of CLUSTAL W (Thompson *et al.*, 1994), followed by manual adjustment with BioEdit (version 5.0.0) (Hall, 2001), and then realigned by applying the 'removing gap

function' of MEGA align software (v. 5.03) in the LaserGene program (DNA Star) to generate correctly aligned multiple sequence files (for nucleotide and deduced amino acid sequences).

Determination of cleavage sites, glycosylation sites, cysteine residues and conserved motifs. Potential cleavage sites for ROCV polyprotein were primarily determined, according to the proteolytic processing cascade pattern for the flavivirus ORFs developed by Rice & Strauss (1990). The highest cleavage potential scores obtained by SignalP-NN computer program (<http://www.cbs.dtu.dk/services/>) (Chang *et al.*, 2000) were used for determining the sites cleaved by the host cell-encoded signalase. Predicted glycosylation and cysteine residue sites were determined using the NetNGlyc (v.1.0) (<http://www.cbs.dtu.dk/services/>) and Protean (v. 5.03) of the LaserGene program (DNA Star), respectively. The MEGA align software was used to assess the presence of flaviviruses highly conserved motifs in the ROCV genome.

Scanning of ROCV ORF and gene-by-gene analyses. For the identification of related viral proteins, the amino acid sequence of ROCV ORF was scanned against other flavivirus sequences available in GenBank using the protein-protein BLAST program (<http://www.ncbi.nlm.nih.gov/>). Parameters were set as follows: database-filter, only virus sequences; matrix, BLOSUM62; gap cost, existence 11 and extension 1; expected [e-value threshold=0.005]. Scanning was performed on multiple amino acid sequence alignment prepared by the method above. Bootstrap values were obtained by generating 100 replicas, and the sequences most closely related to ROCV (query sequence) were displayed based on the lowest e-value (<0.005) and the highest p-score (p-score value that corresponds to a determined percentage identity). Sequences represented by the lowest e-value and the highest genome p-score were used for multiple sequencing alignments (gene-by-gene) using the CLUSTAL W software (Thompson *et al.*, 1994), and percentage of similarities for nucleotide and deduced amino acid sequences were calculated with the MEGA align software (LaserGene program; DNA Star).

Phylogeny. For phylogenetic analyses, the sequences of the genes possessing antigenically important traits (E and NS1 genes) and conserved genes (NS3 and NS5) were concatenated in that order for ROCV, and other selected flaviviruses representing mosquito-borne, tick-borne, no-vector and arthropod (mosquito) virus groups. Neighbour-joining (NJ) (Saitou & Nei, 1987) and maximum-parsimony (MP) (Swofford, 1999) methods were performed. For NJ, a distance matrix calculated from the aligned sequences by Kimura two-parameter formula (Kimura, 1980) was used, and a weight of four for transitions versus one for transversion was selected. In MP, in order to obtain the most parsimonious tree, the heuristic algorithm method was performed; and for determining the reliability of tree topology bootstrap analysis (Felsenstein, 1995) was carried out on 1000 replicas. Bootstrap resampling technique (Efron, 1982) was then used to further evaluate the reliability of the bootstrap analysis with a confidence value of 0.95 (95%). Two distantly related arthropod flaviviruses, cell fusing agent virus (CFAV) and Kamiti River virus (KRV), were used as outgroups. For further analysis of the relationship between ROCV and Ntaya virus (NTAV), the E-NS5 concatenated amino acid sequences of Bagaza virus (BAGV), NTAV and six members of the JEV group [Murray Valley encephalitis virus (MVEV), Usutu virus (USUV), JEV, WNV, Kunjin virus (KUNV) and SLEV] were used. Confidence values used as criteria for group inclusion or exclusion were calculated based on the mean of the concatenated E-NS1-NS3-NS5 [71% ± 2 standard deviations (SD)] and E-NS5 (71.5% ± 2 SD) amino acid sequence identities within and among selected members of the JEV, dengue virus (DENV), tick-borne and no-vector flavivirus groups.

Computer-generated secondary structures in the 3'-NCR and genome cyclization between the 5'- and 3'-NCRs. The secondary

structures in the 3'-NCR and cyclization between the 3'- and 5'-terminal regions were analysed with the mfold program (Zuker *et al.*, 1999). The first 200 nt of the genome representing the entire 5'-NCR followed by a part of the C gene and the entire 3'-NCR were used for the analysis, as these two terminal regions of flaviviruses contain the domains for genome cyclization between them as well as highly conserved sequences in the 3'-NCR (Khromykh *et al.*, 2001).

RESULTS

Genome analyses

The complete genome of ROCV was determined to be 10 794 nt in length with an ORF of 10 275 nt that is flanked by 5'- and 3'-NCRs of 92 and 427 nt, respectively. The complete ORF was predicted to encode a large polyprotein of 3425 aa, that, after co- and post-translational cleavage processes, gives rise to 10 viral proteins, typical of flaviviruses. The lengths and positions of the genes or genomic regions are shown in Table 1. To further examine the lengths of ROCV genes, we compared their nucleotide and amino acid sequences with selected mosquito-borne [ILHV, SLEV, WNV, JEV, DENV-1-4 and yellow fever virus (YFV)] and tick-borne (Langat virus, Powassan virus and Tick-borne encephalitis virus) flaviviruses, as well as with representative members of no-vector flaviviruses (Modoc virus and Rio Bravo virus). Although differences in genome length observed among flaviviruses is mainly due to size variability in the 3'-NCRs (Bryant *et al.*, 2005; Wallner *et al.*, 1995), size heterogeneity was also observed along the coding regions. Generally, ROCV genes are more similar in length with those of ILHV, SLEV, WNV and JEV than with DENV serotypes 1-4 and tick-borne flaviviruses; but the length differences with no-vector viruses were pronounced (Supplementary Table S1 available in JGV Online).

Cleavage sites, glycosylation sites and cysteine residues

The predicted pentapeptide cleavage sites of ROCV were compared with the corresponding sites of representative members of the JEV (WNV, SLEV, JEV and MVEV) and Ntaya (BAGV) groups as shown in Supplementary Table S2 (available in JGV Online). The comparison revealed more differences (two or more amino acids in either peptide) at VirC/AnchC, AnchC/pr, pr/M, NS1/NS2A, NS2A/NS2B and NS4B/NS5 potential cleavage sites where amino acid substitutions were found in the ROCV genome. Single residue substitutions were also observed in the M/E, E/NS1 and NS2B/NS3.

The numbers of N-linked potential glycosylation sites (N-LGlyS) in the prM, E and NS1 genes for ROCV were two, one and four, respectively. No N-LGlyS was found in the hydrophobic domain. This pattern of three site numbers is compared with two, one and three for Argentine 66 strain of SLEV; one, zero and two for JEV and one, one and three for WNV. However, it should be noted that strain variation in the number of sites has been reported. Thus, in the E

Table 1. Full-length genome organization of ROCV

Proteins	Genes	Size (nt)	Genome position	Protein size (aa)	
Structural proteins	5'-NCR	92	1-92	-	
	C	VirC	303	93-395	101
		AnchC	51	396-446	17
	prM	pr	276	447-722	92
		M	225	723-947	75
	E		1 503	948-2450	501
Non-structural proteins	NS1	1 059	2 451-3 509	353	
	NS2A	681	3 510-4 190	227	
	NS2B	393	4 191-4 583	131	
	NS3	1 857	4 584-6 440	619	
	NS4A	378	6 441-6 818	126	
	2K	69	6 819-6 887	23	
	NS4B	768	6 888-7 655	256	
	NS5	2 715	7 656-10 370	904	
	3'-NCR	424	10 371-10 794	-	

gene of SLEV, while the Argentine 66 has one, strain MSI-7 has two, but strain Tr 9464 has no glycosylation sites (Kuno & Chang, 2005; Vorndam *et al.*, 1993). Regarding cysteine residues, as strictly required for stabilizing disulfide bridges in all mosquito-borne viruses, 12 residues were found in the ectodomain of E and in NS1 genes of ROCV. As previously observed among other flaviviruses, the six cysteine residues in prM are clustered in the N-terminal hydrophobic pr domain.

Motifs

In prM, a few peptides are conserved among flaviviruses. Although, the amino acid W (at residue 68) of ROCV is shared by all mosquito-borne viruses, it is replaced with F in all tick-borne viruses. Many amino acid motifs in the E, NS3 and NS5 proteins are shared between ROCV and other mosquito-borne, tick-borne and no-vector flaviviruses. In the E protein, the tripeptide in the domain III region (amino acid residues 388-390), which corresponds to the RGD motif in JEV, Alfuy virus (ALFV), USUV and MVEV (Bakonyi *et al.*, 2004; Hurrelbrink *et al.*, 1999; May *et al.*, 2006; Sumiyoshi *et al.*, 1987), RGE in WNV and KUNV (Castle *et al.*, 1985; Coia *et al.*, 1988) and RGP in SLEV (Ciota *et al.*, 2007), is TGP in ROCV. The TGP peptide is also found in Kedougou virus (KEDV) and is very similar to those of Kokobera virus (KOKV) and BAGV (TGE) (Kuno & Chang, 2007). Regarding the conserved motif corresponding to the fusion peptide region (residues 98-110), ROCV also shares with other mosquito-borne flaviviruses the same amino acid sequence, differing from tick-borne flavivirus members at the amino acid residue 104 (G versus H). In the NS3 gene, the catalytic triad (H-D-S) of protease is well preserved in ROCV. However, a notable difference from the mosquito-borne viruses is found in the hexapeptide motif MDEAHF (residues 284-289), where tick-borne and no-vector groups replace F by W. Among

most mosquito-borne viruses including ROCV, the heptapeptide at residues 294-300 is SIAARGY, but Y is replaced by W among yellow fever group viruses (YFV, Sepik virus and Entebbe bat virus) and by H among tick-borne viruses. In addition, at residue 225 in the conserved motif LAPTRVV (residues 222-228), a single amino acid change (T→S) was observed in ROCV. In the NS5 gene, like in NS3 gene, many conserved motifs are shared between ROCV and other flaviviruses. However, at the highly conserved peptide AMTDTTPFGQQRVFKEKVDV (residues 343-362) shared by all mosquito-borne viruses, including ROCV, P is replaced by A in both tick-borne and no-vector groups, while it is replaced by T for two insect flaviviruses (CFAV and KRV). At residues 550-551, the dipeptide amongst most mosquito-borne viruses, including ROCV, is NE; however, it is DE among KEDV, YFV, tick-borne, no-vector (except TBV) and arthropod flaviviruses.

Genome scanning

Protein-protein BLAST analysis using the entire ORF of ROCV as a query sequence revealed protein identity with several flaviviruses (*e*-values=0.0; *p*-scores ranging from 1083 to 5691). The highest similarity was observed with ILHV (*p*-score: 5681, 88%) and members of the more broadly defined JEV group: SLEV (*p*-score: 4910; 81%), WNV (*p*-score: 4799, 80%), MVEV (*p*-score: 4784, 80%), and JEV (*p*-score: 4784, 80%). As shown in Table 2, the close genetic relationship between ROCV and ILHV determined by CLUSTAL W software (77.5% for amino acid sequences and 71% for nucleotides in ORF) is reflected in higher amino acid identities for NS2B (85.6%), NS3 (85.4%) and NS5 (79.9%). The identities in the E, NS4A and NS4B proteins are slightly lower but still close to the level of NS5 protein. The reduced identity (55.9%) found only with capsid protein of ILHV is peculiar because low level identities below 55% is found in multiple proteins

Table 2. Amino acid identities of ROCV with ILHV and JEV complex viruses (nucleotide identities with ILHV in parentheses)

GenBank accession nos: ILHV (AY632539), JEV (M18370), WNV (M12294), MVEV (AF161266) and SLEV (DQ525916).

Virus	Amino acid (nucleotide) identity (%) of ROCV										Entire ORF (%)
	C	pr/M	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	
ILHV	55.9 (61.4)	67.1 (65.7)	78.6 (70.0)	73.4 (68.6)	70.0 (66.5)	85.6 (74.0)	85.4 (75.0)	77.9 (68.2)	72.7 (70.3)	79.9 (72.5)	77.5 (71.0)
JEV	45.2	61.3	64.4	63.5	51.5	51.9	71.4	61.7	53.1	71.0	64.5
WNV	50.0	56.0	62.5	64.0	48.5	51.9	72.2	56.4	55.8	71.4	65.0
MVEV	54.2	58.7	62.3	64.5	53.3	55.0	71.2	61.1	51.2	70.9	64.7
SLEV	53.6	69.3	65.4	63.2	52.8	48.1	70.0	61.1	57.0	74.8	66.1

besides capsid when four JEV complex viruses are compared (Table 2). The identity (67.1 %) in prM of ILHV is also low but higher than that of the capsid protein. Another peculiarity of ROCV is that the identity in NS5 protein of ILHV is lower than those of NS2B and NS3 proteins, because generally NS5 gene is more conserved than the other genes among flaviviruses. Other than with ILHV, however, identities of ROCV NS3 and NS5 proteins with other viruses compared are generally similar.

Phylogeny

The phylogenetic analysis based on the concatenated 2.3 kb E-NS1-NS3-NS5 amino acid sequences of ROCV and other flaviviruses using MP and NJ methods resulted in two trees with a similar topology, although the bootstrap values in the MP consensus tree were slightly lower than those in the NJ tree (data not shown). Accordingly, in this report, the NJ tree was selected to represent the phylogeny of the selected viruses. Based on reliable bootstrap values (>75 %) and confidence values (≥ 71 and $71.5\% \pm 2$ SD for E-NS1-NS3-NS5 and E-NS5, respectively) as the criteria for group inclusion, the phylogram clearly demonstrates major cluster segregation with 100 % bootstrap values among the mosquito-borne, tick-borne and no-vector flaviviruses, and subclassification of the mosquito-borne cluster to nine subclades (Fig. 1a). In the mosquito-borne cluster, divergence of the two major clades (predominantly *Culex*-borne viruses and *Aedes*-borne viruses) is evident. As shown in Fig. 1(a), the encephalitogenic ROCV is found in the predominantly *Culex*-borne virus clade and is most closely affiliated with ILHV. When the concatenated sequence is limited to E and NS5, as in Fig. 1(b), segregation of the lineage consisting of ROCV and ILHV from JEV complex viruses and also from Ntaya group (BAGV and NTAV) is more clearly demonstrated.

5'- and 3'-NCRs

The analysis of the computer-generated secondary structures of ROCV revealed a similar pattern found in other mosquito-borne flaviviruses. The conserved sequences (CS3, RCS2, CS2 and CS1) and associated secondary

structures (Chambers *et al.*, 1990; Hahn *et al.*, 1987; Proutski *et al.*, 1997) are evident in Fig. 2. The predicted cyclization between the cyclization domain (5'-CYC) of the 5'-NCR and the complementary sequence (3'-CYC) of the 3'-NCR is also illustrated in Fig. 2. As expected, the pentanucleotide (CACAG) that is critically important for the replication of all flaviviruses is found in the loop of the long, stable hairpin structure (3'-LSH) near the 3'-terminal sequence (Fig. 2).

The CS1 (GCATATTGACACCTGGGAAAGAC) and CS2 (GGACTAGAGGTTAGAGGAGACCC) of ROCV are nearly identical to the corresponding conserved sequences among YFV, DENV and JEV (Chambers *et al.*, 1990; Hahn *et al.*, 1987; Khromykh *et al.*, 2001). On the other hand, the 3'-terminal sequence of the CS3 of ROCV (CCCCAGGAGGACTGGGTTTCAAAGC) is slightly degenerate or different. The pattern of CS organization (in 5'→3' direction) in the 3'-NCR of ROCV is CS3-CS2-RCS2-CS1. The structure illustrated in Fig. 2 suggests that CS2, RCS2 and CS3, as well as 3'-LSH are involved in the secondary structure. Generally, a high level of stability is observed along the predicted secondary structures in the lower left region below the dashed line in Fig. 2, where minor variation in the number of bulges and/or small loops (SmLs) is noted at different holding energy levels. By contrast, the region in the upper right of Fig. 2 above the dashed line including the CS2, RCS2 and CS3, as well as the hairpin sequences (HS) (5'-CapHS, 3'-HS1-3) revealed a greater variation in possible folding patterns generated at different energy levels.

DISCUSSION

During the encephalitis outbreak in humans in the São Paulo state located in the south-eastern region of Brazil in the 1970s, an enormous public health impact of ROCV was felt, because the outbreak affected at least 23 municipalities in the Ribeira Valley area of the São Paulo state. The encephalitic disease was characterized by an acute onset of fever, headache, and several central nervous system (CNS) signs and symptoms, resulting in a case-fatality rate of over 10 %, and, importantly, development of CNS sequelae in

<<paper no. vir82883 charlesworth ref: vir82883>>

D. B. A. Medeiros and others

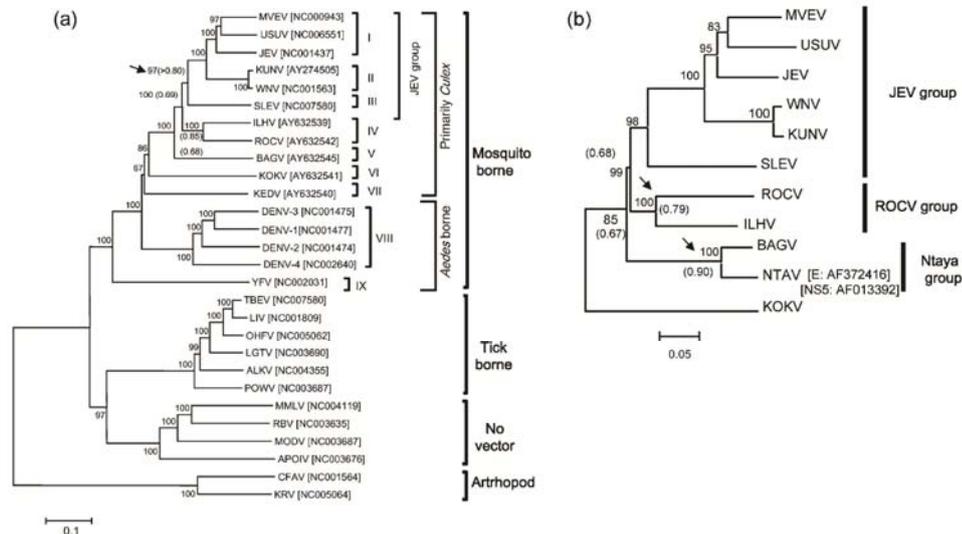


Fig. 1. NJ trees of ROCV and selected flaviviruses based on amino acid sequences of (a) concatenated E-NS1-NS3-NS5 or (b) E-NS5. Two distant arthropod flaviviruses (KRV and CFAV) and KOKV were used as outgroups in (a) and (b), respectively. The roman numerals refer to subclades described in the Discussion. The numbers at nodes indicate bootstrap support values and arrows indicate high supporting bootstrap value for JEV (97%), ROCV (100%) and Ntaya (100%) groups. Confidence values used as criteria for group inclusion or exclusion are shown in parentheses. The horizontal scale bar represents 10 and 5% nucleotide sequence divergences for (a) and (b), respectively. GenBank accession nos are shown in square brackets. MVEV (Murrey valley encephalitis virus), USUV (Usutu virus), JEV (Japanese encephalitis virus), KUNV (Kunjin virus), WNV (West Nile virus), SLEV (Saint Louis encephalitis virus), ILHV (Ilheus virus), ROCV (Rocio virus), BAGV (Bagaza virus), NTAV (Ntaya virus), KOKV (Kokobera virus), KEDV (Kedougou virus), DENV-1 (dengue virus type 1), DENV-2 (dengue virus type 2), DENV-3 (dengue virus type 3), DENV-4 (dengue virus type 4), YFV (yellow fever virus), TBEV (tick-borne encephalitis virus), LIV (louping ill virus), OHFV (Omsk hemorrhagic fever virus), LGTV (Langat virus), ALKV (Alkhurma virus), POWV (Powassan virus), MMLV (Montana myotis leukoencephalitis virus), RBV (Rio Bravo virus), MODV (Modoc virus), APOIV (Apoi virus), KRV (Kamiti River virus) and CFAV (cell fusing agent virus).

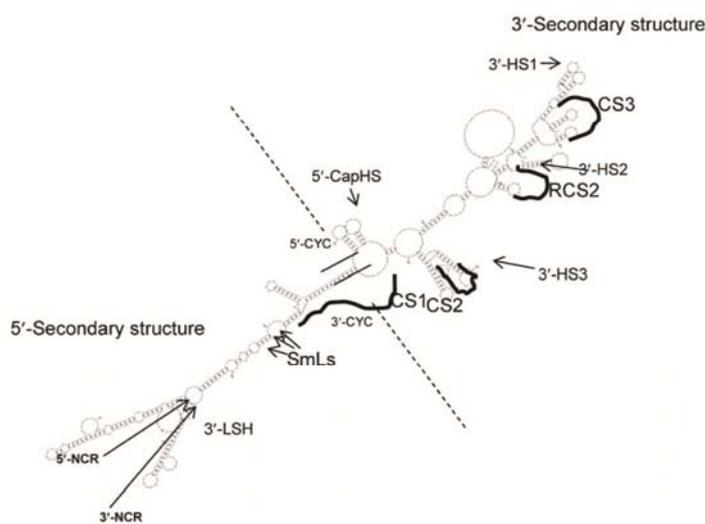


Fig. 2. Predicted secondary structures of the 5'- and 3'-NCRs of ROCV in cyclization. The folding pattern was generated by the mfold program (energy level = -97.0). Abbreviations: CS, conserved sequences; RCS, repeated conserved sequence; SmLs, small loop; LSH, long stable hairpin; 3'-CYC, cycling sequence within CS1 of the 3'-NCR; 5'-CYC, cycling sequence within capsid (Cap) gene; HS, hairpin sequence and NCR, non-coding region.

about 20 % of the survivors (de Sousa Lopes *et al.*, 1978, 1981; Iversson, 1988). The most frequent CNS symptoms were meningeal irritation (57.3 %), alteration of consciousness (51 %) and motor abnormalities (49.65 %) especially gait and impaired equilibrium, but blindness and deafness were also reported (Iversson, 1988). Accordingly, full genome characterization and determination of the exact phylogenetic and taxonomic relationships of ROCV with other related flaviviruses were undertaken, as such data will be essential for achieving the ultimate goals of designing better molecular probes and primers for improved surveillance and diagnosis, determination of the neurovirulence markers at a molecular level, and development of attenuated vaccine and antiviral drugs.

As a result of accelerated sequencing studies, more than 40 flaviviruses have been fully sequenced thus far (Kuno *et al.*, 1998; Billoir *et al.*, 2000; Gaunt *et al.*, 2001; Grard *et al.*, 2007). Unfortunately, with an exception of SLEV, the paucity of such full-length genomic sequences for the viruses indigenous to South America has hampered the understanding of the evolution of ROCV and its genetic characteristics. Thus, our gene-by-gene genome scanning and phylogenetic analysis provide information for a better understanding of the genetic relationship of ROCV among related flaviviruses.

The genomic organization of ROCV shares the same pattern found in all other flaviviruses (Rice *et al.*, 1985; Chambers *et al.*, 1990). This study also reveals that ROCV shares much of the genetic traits with other mosquito-borne flaviviruses. Thus, as expected, its cysteine residue distribution pattern in prM, E and NS1 genes (6, 12 and 12, respectively) is identical to the pattern observed in other flaviviruses, because of the critical importance in proper conformation afforded by disulfide bridges.

Regarding the presence of potential N-LGlyS along the prM-E-NS1 genes in ROCV, differences in the number of sites was observed in the E and NS1 genes of the mosquito-borne, tick-borne and no-vector flaviviruses. Furthermore, as mentioned earlier, strain variation exists at least for SLEV (Spence, 1980; Vorndam *et al.*, 1993; Diaz *et al.*, 2006). Thus, basically the function of the potential sites in ROCV remains unknown in the absence of experimental data; and, clearly, further experimental studies are necessary to clarify the functional significance. Here, it is important to mention contradictory findings in the past. In one study, the correlation between the presence of such sites and antigenic properties of flavivirus was not established because deglycosylated viruses could maintain the same antigenicity (Winkler *et al.*, 1987). On the other hand, in other more recent studies, glycosylation was found to be very important for viral replication, virulence, maturation or release of viral-like particles (Beasley *et al.*, 2005; Crabtree *et al.*, 2005; Goto *et al.*, 2005; Li *et al.*, 2006).

Our analysis of the conserved motifs showed that ROCV shares numerous conserved amino acid and/or peptides with several flaviviruses across different groups each with a

varied host range. Interestingly, in the E protein, the tripeptide (TGP) of ROCV within domain III was found to be different from those observed in all other JEV members (RGX, where X are D, E or T). As the RGD motif is speculated to be involved in the virus adsorption to host cells for JEV and MVEV, two important JEV group flaviviruses (Bhardwaj *et al.*, 2001; Lee & Lobigs, 2000) and as it was also found in ALFV and USUV, further experimental studies for ROCV are needed to investigate the role of the TGP tripeptide in the mechanisms that involve the virus-cell interaction.

The functional roles of the 3'-NCR of flaviviruses have been recently studied and the roles in the genomic cyclization between the 5'- and 3'-NCRs, viral replication and translation were elucidated. Moreover, the putative changes in the highly conserved regions related to important biological functions of the flaviviruses (survival and/or pathogenicity) were recognized (Bryant *et al.*, 2005; Khromykh *et al.*, 2001). Khromykh *et al.* (2001) demonstrated that at least one base-pairing between the 5'-3' termini, rather than the nucleotide sequences per se, is essential for the replication of viral RNA of the mosquito-borne Kunjin virus, a member of the JEV group, and that more than one pair of cyclization sequences might be involved in the replication of the tick-borne flaviviruses.

In terms of CS organization in the 3'-NCR, our results revealed a CS3-CS2-RCS2-CS1 pattern for ROCV. This CS organization for ROCV is very similar to but distinct from that of the JEV group viruses (RCS3-CS3-RCS2-CS2-CS1). Thus, regarding the CS organization in the 3'-NCR, the relationship of ROCV to JEV group viruses is similar to the previously established phylogenetic relationship in which ROCV is closer to the JEV complex viruses but is still distinct (Kuno *et al.*, 1998). Although the CS pattern in the 3'-NCR has not been used for subgrouping within the mosquito-borne flaviviruses, as shown in this report, it deserves further studies to determine its utility as a useful supplementary taxonomic marker at viral population level rather than at individual strain level.

According to the recent taxonomic classification, ROCV is classified as a genotype of ILHV in the Ntaya virus group (ICTV, 2005). Numerous phylogenetic studies (Billoir *et al.*, 2000; Cook & Holmes, 2005; Gould *et al.*, 2003; Kuno *et al.*, 1998, 2005) have been conducted using sequences of individual genes and/or ORF to investigate the flavivirus genetic relationship. These studies generated basically two contrasting phylogenies, NS5 gene tree and NS3/ORF tree. In this study, we tried to determine how the application of concatenated sequences of selected genes would affect phylogeny, by focusing only on antigenically important or phylogenetically more informative genes but by dropping highly variable genes, such as NS2 and NS4. The phylogeny of flaviviruses obtained using the concatenated sequences (E-NS1-NS3-NS5) is similar to the topology previously obtained with NS3 gene or ORF (Billoir *et al.*, 2000; Cook & Holmes, 2005; Kuno & Chang, 2006). Our phylogenetic

results confirmed a close genetic relationship between ROCV and the encephalitogenic JEV group viruses. Using the criteria for group inclusion (amino acid identity $\geq 71\%$), ROCV did not show enough identity to be included either in the JEV or in the Ntaya group, which is represented in our study by NTAV and BAGV because of insufficient (69%) identity. Furthermore, the phylogeny obtained in this study suggests a strong possibility that the two Brazilian flaviviruses (ROCV and ILHV) segregated earlier independently, suggesting that they represent a lineage distinct from the lineage leading to JEV and Ntaya groups.

Currently, ROCV is classified as a subtype of ILHV (ICTV, 2005). However, traditionally, ROCV has been readily distinguished from ILHV serologically, as the ratios of homologous versus heterologous titres obtained when ROCV was used as antigen or antibody ($>640/<10$ and $1280/<10$, respectively) were far higher and significant than the minimal threshold ($>fourfold$) used to distinguish flaviviral species (Karabatsos, 1985). In terms of phylogenetic classification, between the two phylogenetically related viruses, ROCV and ILHV, the nucleotide sequence identity (72.5%) of the full-length NS5 gene nucleotide sequence (Table 2) demonstrated in this study strongly supports recognition of those viruses as two distinct species, if the species criterion (species defined by $>84\%$ nucleotide sequence identity based on the 3'-terminal 1 kb sequence of NS5) previously established for all flavivirus species (Kuno *et al.*, 1998) could be applied to complete NS5 gene sequences as well. Actually, quantitative criteria based on evolutionary distance of a particular gene, such as above, have been used to improve classification of tick-borne flaviviruses (Charrel *et al.*, 2001) and for demarcation of strains and species (Van Regenmortel *et al.*, 1997). Also, biologically, although both ROCV and ILHV are neurotropic, the former has been often associated with epidemic of encephalitis, while the latter has been associated only with sporadic febrile illness with minor neurological abnormalities, clearly reflecting significant differences in neurovirulence and transmissibility (Vasconcelos *et al.*, 1992, 1998). Accordingly, our conclusion differs from the assignment of ROCV as subtype (or genotype) of ILHV by ICTV (2005) and adds to the list of recent studies in which taxonomic classification of a particular flavivirus was found different from that by ICTV (Grard *et al.*, 2007; Kuno & Chang, 2007; May *et al.*, 2006).

The peculiar vector-virus relationships between *Aedes* vectors and the viscerotropic flaviviruses (such as DENV and YFV) and between *Culex* vectors and predominantly neurotropic viruses (such as JEV group viruses) were recognized early by Sabin (1959). In the more recent phylogenetic study, the relationship was basically confirmed phylogenetically (Gaunt *et al.*, 2001). However, exceptions to this generalization have been reported for several mosquito-borne flaviviruses (Mutebi *et al.*, 2004). Although both ROCV and ILHV are neurotropic and together share a subclade in the phylogenetic tree, unlike

the members of the neurotropic JEV group viruses transmitted by *Culex* mosquitoes, both viruses have been most often isolated from *Psorophora* mosquitoes (de Sousa Lopes *et al.*, 1981; Vasconcelos *et al.*, 1992, 1998). Several other members of the *Culex*-borne virus clade that do not meet this specific vector-disease syndrome relationship (Gaunt *et al.*, 2001) include Koutango and Saboya viruses that have been isolated more often from ticks and sandflies, respectively, and which are not known to be neurotropic. Nonetheless, the above generalization (Gaunt *et al.*, 2001) can still be applied to ROCV, because it is a neurotropic virus that can replicate in *Culex* mosquitoes at least under laboratory conditions (Mitchell *et al.*, 1981).

Complete genome sequencing is a pre-requisite to genomic characterization because it provides the critically needed information for an accurate viral classification, for a better choice of informative genes to conduct phylogenetic analysis, and for designing specific and/or broadly reactive diagnostic primers and/or probes in order to improve disease surveillance and diagnosis. Furthermore, the knowledge of the genomic sequence is indispensable for preparing infectious clones for further applications to such investigations as the identification of gene function and of virulence marker and development of antiviral drugs, as well as vaccine development. Our data on ROCV represent the first complete genomic characterization of an indigenous Brazilian flavivirus of public health importance and should be useful for elucidating the genetic relationship with other South American flaviviruses transmitted by mosquitoes, such as Cacipacore virus, ILHV, Bussuquara virus and Iguape virus.

ACKNOWLEDGEMENTS

This work is dedicated in memory of Dr Oscar de Souza Lopes who isolated ROCV and contributed greatly to our understanding of the medical importance of the virus and natural transmission cycle. This study was supported by the IEC/SVS and the CNPq grants 300460/2005-8 and 402859/2005-7.

REFERENCES

- Bakonyi, T., Gould, E. A., Kolodziejek, J., Weissenböck, H. & Nowotny, N. (2004). Complete genome analysis and molecular characterization of Usutu virus that emerged in Austria in 2001: comparison with the South African strain SAAR-1776 and other flaviviruses. *Virology* 328, 301–310.
- Beasley, D. W., Whiteman, M. C., Zhang, S., Huang, C. Y.-H., Schneider, B. S., Smith, D. R., Gromowski, G. D., Higgs, S., Kinney, R. M. & Barrett, A. D. T. (2005). Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. *J Virol* 79, 8339–8347.
- Bhardwaj, S., Holbrook, M., Shope, R. E., Barrett, A. D. T. & Watowich, S. J. (2001). Biophysical characterization and the vector-specific antagonist activity of the domain III of the tick-borne flavivirus envelope protein. *J Virol* 75, 4002–4007.
- Billoir, F., de Chesse, R., Tolou, H., de Micco, P., Gould, E. A. & de Lamballerie, X. (2000). Phylogeny of the genus *Flavivirus* using

- complete coding sequences of arthropod-borne viruses and viruses with no known vector. *J Gen Virol* **81**, 781–790.
- Bryant, J. E., Vasconcelos, P. F. C., Rijnbran, D. R. C. A., Mutebi, J. P., Higgs, S. & Barrett, A. D. T. (2005).** Size heterogeneity in the 3' non-coding region of South American isolates of yellow fever virus. *J Virol* **79**, 3807–3821.
- Castle, E., Nowak, T., Leidner, U., Wengler, G. & Wengler, G. (1985).** Sequence analysis of the viral core protein and the membrane-associated proteins V1 and NV2 of the flavivirus West Nile virus and of the genome sequence for these proteins. *Virology* **145**, 227–236.
- Chamberlain, R. W. (1980).** Epidemiology of arthropod-borne togaviruses: the role of arthropods as hosts and vectors and of vertebrate hosts in natural transmission cycles. In *The Togaviruses: Biology, Structure, Replication*, pp. 175–239. Edited by R.W. Schlesinger. New York, London, Toronto, Sydney, San Francisco: Academic Press.
- Chambers, T. J., Hahn, C. S., Galler, R. & Rice, C. M. (1990).** Flavivirus genome organization, expression, and replication. *Annu Rev Microbiol* **44**, 649–688.
- Chang, G.-J. J., Hunt, A. R. & Davis, B. (2000).** A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice. *J Virol* **74**, 4244–4252.
- Charrel, R. N., Zaki, A. M., Attoui, H., Fakeeh, M., Billoir, F., Yousef, A. I., de Chesse, R., de Micco, P., Gould, E. A. & de Lamballerie, X. (2001).** Complete coding sequence of the Alkhurma virus, a tick-borne flavivirus causing severe hemorrhagic fever in humans in Saudi Arabia. *Biochem Biophys Res Commun* **287**, 455–461.
- Ciota, A. T., Lovelace, A. O., Ngo, K. A., Le, A. N., Maffei, J. G., Franke, M. A., Payne, A. F., Jones, S. A., Kauffman, E. B. & Kramer, L. D. (2007).** Cell-specific adaptation of two flaviviruses following serial passage in mosquito cell culture. *Virology* **357**, 165–174.
- Coia, G., Parker, M. D., Speight, G., Byrne, M. E. & Westaway, E. G. (1988).** Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene order and characteristics of the virus-specified proteins. *J Gen Virol* **69**, 1–21.
- Cook, S. & Holmes, E. C. (2005).** A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: *Flaviviridae*) and the evolution of vector transmission. *Arch Virol* **151**, 309–325.
- Crabtree, M. B., Kinney, R. M. & Miller, B. R. (2005).** Deglycosylation of the NS1 protein of dengue 2 virus, strain 16681: construction and characterization of mutant viruses. *Arch Virol* **150**, 771–786.
- de Sousa Lopes, O., Coimbra, T. L. M., Sacchetta, L. A. & Calisher, C. H. (1978).** Emergence of a new arbovirus in Brazil. I. Isolation and characterization of the etiologic agent. *Am J Epidemiol* **107**, 444–449.
- de Sousa Lopes, O., Sacchetta, L. A., Francy, D. B., Jakob, W. & Calisher, C. H. (1981).** Emergence of a new arbovirus disease in Brazil. III. Isolation of Rocio virus from *Psorophora ferox* (Humboldt, 1819). *Am J Epidemiol* **113**, 122–125.
- Diaz, L. A., Re, V., Almiron, W. R., Farias, A., Vazquez, A., Sanchez-Seco, M. P., Aguilar, J., Spinsanti, L., Konigheim, B. & other authors (2006).** Genotype III Saint Louis encephalitis virus outbreak, Argentina, 2005. *Emerg Infect Dis* **12**, 1752–1754.
- Efron, B. (1982).** The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia.
- Felsenstein, J. (1995).** PHYLIP, version 3.57c. Department of Genetics, University of Washington, Seattle, WA.
- Gaunt, M. W., Sall, A. A., de Lamballerie, X., Falconar, A. K. I., Dzhanian, T. I. & Gould, E. A. (2001).** Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J Gen Virol* **82**, 1867–1876.
- Goto, A., Yoshii, K., Obara, M., Ueki, T., Mizutani, T., Kariwa, H. & Takashima, I. (2005).** Role of the N-linked glycans of the prM and E envelope proteins in tick-borne encephalitis virus particle secretion. *Vaccine* **23**, 3043–3052.
- Gould, E. A., de Lamballerie, X., Zanotto, P. M. & Holmes, E. C. (2003).** Origins, evolution, and vector/host coadaptations within the genus *Flavivirus*. *Adv Virus Res* **59**, 277–314.
- Grard, G., Moureau, G., Charrel, R. N., Lemasson, J.-J., Gonzalez, J.-P., Gallian, P., Gritsun, T., Holmes, E. C., Gould, E. A. & de Lamballerie, X. (2007).** Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. *Virology* **361**, 80–92.
- Hahn, C. S., Hahn, Y. S., Rice, C. M., Lee, E., Dalgarno, L., Strauss, E. G. & Strauss, J. H. (1987).** Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. *J Mol Biol* **198**, 33–41.
- Hall, T. A. (2001).** BioEdit. Department of Microbiology, North Carolina State University, Raleigh, NC.
- Hurrelbrink, R. J., Nestorowicz, A. & McMinn, P. C. (1999).** Characterization of infectious Murray Valley encephalitis virus derived from a stably cloned genome-length cDNA. *J Gen Virol* **80**, 3115–3125.
- ICTV (2005).** Virus Taxonomy: Eighth Report of the International Committee on the Taxonomy of Viruses. p. 1259. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. London: Elsevier.
- Iversson, L. B. (1988).** Rocio encephalitis. In *The Arboviruses – Epidemiology and Ecology*, vol. 4, pp. 77–92. Edited by T. P. Monath. Boca Raton, FL: CRC Press.
- Karabatsos, N. (1985).** International catalogue of arboviruses, 3rd edn. San Antonio, TX: American Society of Tropical Medicine and Hygiene.
- Khromykh, A. A., Meka, H., Guyatt, K. J. & Westaway, E. G. (2001).** Essential role of cyclization sequences in flavivirus RNA replication. *J Virol* **75**, 6719–6728.
- Kimura, M. (1980).** A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Kuno, G. & Chang, G.-J. J. (2005).** Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. *Clin Microbiol Rev* **18**, 608–637.
- Kuno, G. & Chang, G.-J. J. (2006).** Characterization of Sepik and Entebbe bat viruses closely related to yellow fever virus. *Am J Trop Med Hyg* **75**, 1165–1170.
- Kuno, G. & Chang, G.-J. J. (2007).** Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. *Arch Virol* **152**, 687–696.
- Kuno, G., Chang, G.-J. J., Tsuchiya, K. R., Karabatsos, N. & Cropp, C. B. (1998).** Phylogeny of the genus *Flavivirus*. *J Virol* **72**, 73–83.
- Lee, E. & Lobigs, M. (2000).** Substitutions at the putative receptor-binding site of an encephalitic flavivirus alter virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. *J Virol* **74**, 8867–8875.
- Li, J., Bhuvanankantham, R., Howe, J. & Ng, M.-L. (2006).** The glycosylation site in the envelope protein of West Nile virus (Sarafenid) plays an important role in replication and maturation processes. *J Gen Virol* **87**, 613–622.
- Lindenbach, B. D. & Rice, C. M. (2003).** Molecular biology of flaviviruses. *Adv Virus Res* **59**, 23–61.
- May, F. J., Lobigs, M., Lee, E., Gendle, D. J., Mackenzie, J. S., Broom, A. K., Conlan, J. V. & Hall, R. A. (2006).** Biological, antigenic and

<<paper no. vir82883 charlesworth ref: vir82883>>

D. B. A. Medeiros and others

- phylogenetic characterization of the flavivirus Alfuy. *J Gen Virol* 87, 329–337.
- Mitchell, C. J., Monath, T. P. & Cropp, C. B. (1981).** Experimental transmission of Rocio virus by mosquitoes. *Am J Trop Med Hyg* 30, 465–472.
- Mondini, A., Cardeal, I. L. S., Lazaro, E., Nunes, S. H., Moreira, C. B., Rahal, P., Maia, I. L., Franco, C., Gongora, D. V. N. & other authors (2007).** Saint Louis encephalitis virus, Brazil. *Emerg Infect Dis* 13, 176–178.
- Mutebi, J. P., Rijnbrand, R. C. A., Wang, H., Ryman, K. D., Wang, E., Fulop, L. D., Titball, R. & Barrett, A. D. T. (2004).** Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the *Flavivirus* genus based on the 3' noncoding region. *J Virol* 78, 9652–9665.
- Pierre, V., Drouet, M.-T. & Deubel, V. (1994).** Identification of mosquito-borne flavivirus sequences using universal primers and reverse transcription/polymerase chain reaction. *Res Virol* 145, 93–104.
- Prutski, V., Gould, E. A. & Holmes, E. C. (1997).** Secondary structure of the 3' untranslated region of flaviviruses: similarities and differences. *Nucleic Acids Res* 25, 1194–1202.
- Rice, C. M. & Strauss, J. H. (1990).** Production of flavivirus polypeptides by proteolytic processing. *Semin Virol* 1, 357–367.
- Rice, C. M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R. L. & Strauss, J. H. (1985).** Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science* 229, 726–733.
- Sabin, A. B. (1959).** Survey of knowledge and problems in field of arthropod-borne virus infections. *Arch Gesamte Virusforsch* 9, 1–10.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstruction phylogenetic trees. *Mol Biol Evol* 4, 406–425.
- Spence, L. P. (1980).** St. Louis encephalitis in tropical America. In *St. Louis Encephalitis*, pp. 451–471. Edited by T. P. Monath. Washington, DC: American Public Health Association.
- Straatmann, A., Santos-Torres, S., Vasconcelos, P. F., Travassos da Rosa, A. P. A., Rodrigues, S. G. & Tavares-Neto, J. (1997).** Serological evidence of the circulation of the Rocio arbovirus (*Flaviviridae*) in Bahia. *Rev Soc Bras Med Trop* 30, 511–515.
- Sumiyoshi, H., Mori, C., Fuke, I., Morita, K., Kuhara, S., Kondou, J., Kikuchi, Y., Nagamatu, H. & Igarashi, A. (1987).** Complete nucleotide sequence of the Japanese encephalitis virus genome RNA. *Virology* 161, 497–510.
- Swofford, D. L. (1999).** PAUP. Phylogenetic analysis using parsimony (and other methods), version 4. Sunderland, MA: Sinauer Associates.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994).** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673–4680.
- Van Regenmortel, M. H. V., Bishop, D. H. L., Fauquet, C. M., Mayo, M. A., Maniloff, J. & Calisher, C. H. (1997).** Guidelines to the demarcation of virus species. *Arch Virol* 142, 1505–1518.
- Vasconcelos, P. F. C., Travassos da Rosa, A. P. A., Dégallier, N., Travassos da Rosa, J. F. S. & Pinheiro, F. P. (1992).** Clinical and ecoepidemiological situation of human arboviruses in Brazilian Amazonia. *Cienc Cult* 44, 117–124.
- Vasconcelos, P. F. C., Travassos da Rosa, A. P. A., Pinheiro, F. P., Shope, R. E., Travassos da Rosa, J. F. S., Rodrigues, S. G., Dégallier, N. & Travassos da Rosa, J. F. S. (1998).** Arboviruses pathogenic for man in Brazil. In *An Overview of Arbovirology in Brazil and Neighbouring Countries*, pp. 72–99. Edited by A. P. A. Travassos da Rosa, P. F. C. Vasconcelos & J. F. S. Travassos da Rosa. Belem, Brazil: Instituto Evandro Chagas.
- Vorndam, V., Mathews, J. H., Barrett, A. D., Roehrig, J. T. & Trent, D. W. (1993).** Molecular and biological characterization of a non-glycosylated isolate of St Louis encephalitis virus. *J Gen Virol* 74, 2653–2660.
- Wallner, G., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995).** The flavivirus 3'-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. *Virology* 213, 169–178.
- Winkler, G., Heinz, F. X. & Kunz, C. (1987).** Studies on the glycosylation of flavivirus E proteins and the role of carbohydrates in antigenic structure. *Virology* 159, 237–243.
- Zuker, M., Mathews, D. H. & Turner, D. H. (1999).** Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In *RNA Biochemistry and Biotechnology*, pp. 11–43. Edited by J. Barciszewski & B. F. C. Clark. Amsterdam: Kluwer Academic Publishers.