

Universidade de Brasília  
Pós-Graduação em Biologia Molecular

Caracterização da variabilidade antigênica do gene  
do envelope (*env*) em amostras de HIV-1 circulantes  
no Distrito Federal.

Verônica Sales Veras

2006

Universidade de Brasília  
Pós-Graduação em Biologia Molecular

Caracterização da variabilidade antigênica do gene  
do envelope (*env*) em amostras de HIV-1 circulantes  
no Distrito Federal.

Dissertação apresentada ao Departamento de  
Biologia Celular do Instituto de Ciências  
Biológicas da Universidade de Brasília, como  
requisito parcial à obtenção do grau de mestre  
em Biologia Molecular.

Verônica Sales Veras

Orientadora  
Cláudia Renata Fernandes Martins

2006

Universidade de Brasília  
Pós-Graduação em Biologia Molecular

Caracterização da variabilidade antigênica do gene  
do envelope (*env*) em amostras de HIV-1 circulantes  
no Distrito Federal.

Banca examinadora

Profa. Dra. Andréa Queiroz Maranhão - Universidade de Brasília

Profa. Dra Cláudia Renata Fernandes Martins - Universidade de Brasília

Profa. Dra. Silviene Fabiana de Oliveira - Universidade de Brasília

Prof. Dr. Marcelo de Macedo Brígido - Universidade de Brasília (suplente)

2006

Aos meus pais, Raimundo e Puri, pelo apoio incondicional em todas as etapas da minha vida e por todo carinho e dedicação que me auxiliaram a enfrentar todos os desafios.

À irmã, Vanessa, com quem compartilho esta conquista, pelo incentivo e apoio neste momento tão especial de minha vida.

Às pessoas que realizaram o exame de carga viral do HIV-1 no serviço de saúde do DF, por, anonimamente, tornarem possível este trabalho.

## Agradecimentos

À **Profª Cláudia Renata Fernandes Martins**, por todos esses anos de amizade, companheirismo e dedicação para a realização deste trabalho.

Aos meus amigos e colegas de laboratório **Eduardo Ramalho, Patrícia Wyant, Daniela Cerqueira, Daniella Moraes, Geni Noceti, Karla Braga, Margareti Medeiros Claudiner Oliveira, Regina Amorim, Nazle Vêras e Tainá Raiol** do grupo de pesquisa em virologia da Universidade de Brasília, pela amizade e colaboração na execução deste trabalho.

À **Profª Mariza Ferreira**, e aos técnicos e funcionários do laboratório de Fitopatologia pela constante colaboração.

Às **Profªs Andréa Queiroz Maranhão e Silviene Fabiana de Oliveira** membros da banca examinadora, pelas suas colaborações para este trabalho.

Ao **Prof. Marcelo Brígido**, por suas valiosas observações durante a realização deste trabalho.

Às plataformas de seqüenciamento automático da EMBRAPA-CENARGEN e do Laboratório de Biologia Molecular da Universidade de Brasília, que possibilitaram a realização deste trabalho.

Às minhas inseparáveis amigas **Mayra Amboni, Mariana Letti, Cristiana Carvalho, Juliana Ribeiro, Gisele Duarte, Valéria Madeira**, por todos esses anos de companheirismo, amizade e incentivo.

Aos meus amigos e colegas de graduação que sempre me apoiaram e me incentivaram nesta conquista.

Aos meus pais, **Raimundo Veras e Maria da Purificação Veras** pela infinita dedicação, carinho e apoio incondicional.

À minha irmã **Vanessa Veras**, pelo amor e exemplo de dedicação.

Ao meu namorado **Zezinho**, por sempre me apoiar com muito amor, respeito e compreensão.

À **Ana**, funcionária da Secretaria do Departamento de Biologia Celular, por me receber sempre com presteza e eficiência.

Ao programa de Pós-Graduação em Biologia Molecular da Universidade de Brasília, pelo incentivo e apoio financeiro.

Ao CNPq, pelo apoio financeiro.

## Sumário

<b>Apresentação.....</b>	<b>8</b>
<b>Resumo.....</b>	<b>9</b>
<b>Abstract.....</b>	<b>10</b>
<b>Lista de Figuras.....</b>	<b>11</b>
<b>Lista de Tabelas.....</b>	<b>13</b>
<b>Lista de siglas e abreviaturas.....</b>	<b>14</b>
<b>1. Revisão Bibliográfica.....</b>	<b>15</b>
1.1. Introdução.....	15
1.2. Histórico.....	16
1.3. Classificação e Morfologia.....	17
1.4. Organização do genoma do HIV-1.....	19
1.5. As proteínas do envelope.....	23
1.5.1. O <i>loop</i> V3.....	26
1.6. Ciclo de replicação do vírus.....	28
1.7. Epidemiologia.....	31
1.8. A variabilidade genética e a distribuição geográfica do HIV-1.....	32
1.9. Resposta imune ao HIV.....	38
1.9.1. Desenvolvimento de vacinas anit-HIV.....	40
1.9.2. Anticorpos neutralizantes.....	42
<b>2. Objetivos.....</b>	<b>45</b>
<b>3. Metodologia.....</b>	<b>46</b>
3.1. População estudada.....	46
3.2. Coleta das Amostras.....	46

3.3. Transcrição Reversa.....	47
3.4. Reação em cadeia da polimerase.....	47
3.5. Purificação e Precipitação dos produtos de PCR	48
3.6. Genotipagem.....	49
3.6.1. Seqüenciamento automático.....	49
3.6.2. Análise da seqüências.....	49
3.6.3. Análise Filogenética.....	50
<b>4. Resultados.....</b>	<b>52</b>
4.1. Amostragem.....	52
4.2. Amplificação da região C2-C3 de <i>env</i> .....	52
4.3. Definição dos subtipos do HIV-1.....	52
4.4. Análise Filogenética.....	55
4.5. Caracterização da variabilidade antigênica da região do <i>loop</i> V3 nos isolados do HIV-1 do Distrito Federal.....	58
4.6. Análise dos resíduos de aminoácidos localizados nas posições 11 e 25 do <i>loop</i> V3 dos isolados do HIV-1.....	61
<b>5. Discussão.....</b>	<b>64</b>
5.1. Definição das formas genéticas do HIV-1.....	65
5.2. Caracterização da diversidade genética dos isolados do HIV-1 no Distrito Federal.....	67
5.3. Caracterização da diversidade antigênica dos isolados do HIV-1 no Distrito Federal.....	70
5.4. Caracterização dos resíduos de aminoácidos nas posições 11 e 25 no <i>loop</i> V3 dos isolados do HIV-1 no Distrito Federal.....	72
<b>6. Conclusões.....</b>	<b>74</b>
<b>7. Referências Bibliográficas.....</b>	<b>76</b>
<b>8. Anexo.....</b>	<b>84</b>

## **Apresentação**

Meu interesse pela ciência despertou-se cedo, durante a infância, na forma de livros, revistas e documentários que me auxiliaram a descobrir e compreender melhor o mundo ao meu redor. Toda a minha curiosidade e nostalgia influenciaram a minha escolha pelo curso de graduação em Ciências Biológicas. Já dentro da Universidade, uma área intrigante e apaixonante me chamou atenção, a Virologia.

Em Agosto de 2000, iniciei um estágio em Virologia, com a professora Cláudia Renata Fernandes Martins, na Universidade de Brasília. A realização de três projetos de iniciação científica pelo programa PIBIC (CNPq), em Virologia Molecular, todos sob orientação da professora Cláudia Renata Fernandes Martins, foi essencial para a minha decisão em me especializar nessa área.

Este trabalho faz parte de uma das linhas de pesquisa em Virologia desenvolvidas pelo nosso grupo na Universidade de Brasília. Nossos trabalhos têm o objetivo de traçar o perfil da variabilidade genética de vírus como HIV, HCV, HPV e Rotavírus no Distrito Federal. Os dados obtidos em nossas pesquisas, apesar de modestos, podem ser bastante relevantes para o controle das infecções causadas por esses vírus.

Os objetivos deste trabalho consistem em caracterizar a variabilidade das formas genéticas do HIV-1 que circulam no Distrito Federal e a variabilidade dos epitopos antigênicos encontrados nessa região. Essas informações podem ser relevantes para o desenvolvimento de vacinas efetivas e para a compreensão da epidemiologia do HIV-1.



## Resumo

O HIV-1 é caracterizado por uma ampla diversidade genética, que possui implicações na epidemiologia da AIDS em uma dada região geográfica. O objetivo deste estudo foi descrever a variabilidade intersubtipo e intra-subtipo dos isolados do HIV-1 circulantes no Distrito Federal, região Central do Brasil. Cinquenta e três amostras de RNA do ano de 2002, cedidas pelo Laboratório Central de Saúde Pública do Distrito Federal, foram transcritas reversamente. Os cDNAs foram amplificados por nested PCR para a obtenção de um fragmento de 564 pb, equivalente a região C2/C3 de *env*. Os produtos de PCR foram submetidos ao seqüenciamento automático e as seqüências obtidas foram analisadas pelo programa REGA. Os subtipos identificados foram o B, presente em 51 (96,2%) das 53 amostras, um isolado F1 (1,9%) e um isolado B/F (1,9%). As análises filogenéticas realizadas pelo método de máxima verossimilhança foram 100% concordantes com os resultados obtidos pelo programa REGA. A caracterização das formas genéticas do HIV-1 foi também realizada a partir das regiões genômicas: RT, PR, *nef* e *gag*, para melhor avaliar a diversidade genética dos isolados do HIV-1. Foi observada a ocorrência de oito (15%) amostras discordantes, todas recombinantes B/F. As seqüências de aminoácidos preditas de 41 isolados foram analisadas para a caracterização da variabilidade genética e antigênica intra-subtipo. Dentre os isolados do subtipo B, foi encontrada uma alta heterogeneidade do tetrapeptídeo localizado no topo do *loop* V3. O mais prevalente foi o motivo GPGR (50%) típico dos isolados norte-americanos/europeus, seguido pelo variante brasileiro B'' GWGR (10%), pelos motivos GFGR (7,5%) e GPGK (7,5%) e GLGR (5%). Outros 8 polimorfismos foram descritos: GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH. Os motivos GPGY e ALGR não estão relatados na literatura. Aminoácidos de carga positiva nas posições 11 e/ou 25, do *loop* V3, também foram avaliados. Resíduos positivos nessas posições foram observados em 22% dos isolados, sugerindo o fenótipo X4 e indutor de sincícios. O *loop* V3 de *env* possui importância na indução de anticorpos neutralizantes e no reconhecimento do co-receptor e a presença de variabilidade neste domínio pode ter implicações no desenvolvimento de vacinas e na resposta imune do HIV-1. Este estudo pretendeu contribuir para a compreensão dos polimorfismos genéticos do HIV-1 no Brasil, e pode ser útil para o planejamento do uso de vacinas anti-HIV-1 no Distrito Federal.

## Abstract

The HIV-1 is characterized by a high genetic diversity. Such diversity has potential implication on the local epidemiology of AIDS. This work aimed to describe the inter-subtype and intra-subtype variability of HIV-1 isolates circulating in the Federal District, Central Brazil. Fifty-three RNA samples, of the year 2002, from a Public Health Laboratory of Federal District were reversely transcribed. Complementary DNA were amplified by nested PCR to obtain a fragment of 564 bp, corresponding to the *env* C2/C3 region. The PCR products were automatically sequenced. The sequences obtained were analyzed by the REGA. Subtypes identified were B 53 (96.2%), F 1 (1.9%) and B/F 1 (1.9%). Phylogenetic analysis based on the C2/C3 *env* DNA sequence from the HIV-1 isolates analyzed was 100% concordant with REGA results. These data were complemented with the subtype characterization based on *gag*, *pol* and *nef*, to evaluate the genetic diversity of HIV-1 isolates in this geographic region. We observed discordance in 8 samples (15%), all presented mosaic genomic B/F forms. The amino acid sequences of 41 HIV-1 isolates were analyzed for the genetic and antigenic intra-subtype diversity. High heterogeneity was observed at the tetrapeptide on tip of the V3 loop in the subtype B variants. The most prevalent was the GPGR (50%) typical in the U.S./European strains, followed by the Brazilian variant B''(GWGR), (10%), the GFGR (7.5%) and GPGK (7.5%) motifs and the GLGR motif (5%). Other V3 variants were found in 8 isolates (20%): GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH. The tetrapeptide sequences GPGY and ALGR were not previously described in the literature. Amino acids with positively charge at the position 11 and/or 25, at the V3 loop, were also analyzed. It was observed that 22% of the HIV-1 isolates presented charged residues at these positions, suggesting a X4 and syncytium-inducing profile. As the V3 loop is considered to be important for neutralizing antibodies and coreceptor recognition, the presence of such variability in this genomic region may have considerable implications on vaccine design and immune response. This study may contribute to the understanding of HIV-1 genetic polymorphism in Brazil, and may prove useful for the development of appropriate anti-HIV vaccines in the Federal District.

**Lista de Figuras**

<b>Figura 1.</b>	Estrutura da partícula do HIV-1.....	19
<b>Figura 2.</b>	Mapa genômico do HIV-1.....	20
<b>Figura 3.</b>	Esquema da estrutura de Env.....	24
<b>Figura 4.</b>	Modelo estrutural do <i>core</i> de gp120, mostrando as posições dos <i>loops</i> variáveis.....	25
<b>Figura 5.</b>	Esquema da seqüência de aminoácidos do <i>loop</i> V3 de isolados de HIV-1 SI (indutor de sincícios) e NSI (não indutor de sincícios).....	28
<b>Figura 6.</b>	Processo de entrada do vírus na célula.....	29
<b>Figura 7.</b>	Distribuição do número de pessoas infectadas pelo HIV no mundo.....	31
<b>Figura 8.</b>	Filogenia dos grupos M, N e O do HIV-1.....	34
<b>Figura 9.</b>	Distribuição geográfica dos subtipos predominantes do grupo M e suas formas recombinantes circulantes (CRFs).....	36
<b>Figura 10.</b>	Distribuição dos subtipos e das formas recombinantes circulantes nas diferentes regiões geográficas do Brasil.....	37
<b>Figura 11.</b>	Seqüências, localização dos <i>primers</i> para a amplificação da região V1-V5 de gp120.....	47
<b>Figura 12.</b>	Análise eletroforética dos produtos de PCR da região de C2/C3 de gp120, em gel de agarose 1% corado com brometo de etídeo.....	52
<b>Figura 13.</b>	Distribuição das formas genéticas do HIV-1 encontradas pela análise da região C2-C3 de gp120.....	53
<b>Figura 14.</b>	Relações filogenéticas pelo método de máxima verossimilhança entre os isolados do HIV-1 circulantes no Distrito Federal (DF 02.N) e seqüências de outros subtipos do HIV-1 do grupo M e de isolados brasileiros do subtipo B.....	56
<b>Figura 15.</b>	Análise por <i>bootscanning</i> da seqüência de nucleotídeos das regiões RT, PR e C2/C3 de gp120 do isolado DF 02.48 B/F, comparando-a aos diferentes subtipos do HIV-1.....	57
<b>Figura 16.</b>	Alinhamento das seqüências de aminoácidos da região do <i>loop</i> V3 obtidas com a seqüência consenso B.....	59
<b>Figura 17.</b>	Distribuição dos variantes do subtipo B do HIV-1 encontrados pela análise das seqüências deduzidas de aminoácidos do <i>loop</i> V3 de gp120.....	60

- Figura 18.** Alinhamento das seqüências de aminoácidos da região do *loop* V3 de gp120 com a seqüência consenso do subtipo B, destacando-se as posições 11 e 25..... 62

## **Lista de Tabelas**

<b>Tabela 1.</b>	Principais funções das proteínas codificadas pelo genoma do HIV-1.....	23
<b>Tabela 2.</b>	Casos de AIDS por região geográfica do Brasil, período de 1980 a 2005.....	32
<b>Tabela 3.</b>	Definição dos subtipos realizada a partir das análises das seguintes regiões genômicas: <i>env gag</i> , RT, PR e <i>nef</i> .....	54

## **Lista de Siglas e Abreviaturas**

AIDS – síndrome da imunodeficiência adquirida	LTRs – long terminal repeat sequences
AZT- zidovidina	NSI- fenótipo não indutor de sincícios
BLAST- basic local alignment search tool	OMS – organização mundial de saúde
cDNA – DNA complementar	Pb- pares de base
CDC- centers for disease control and prevention	PCR – reação em cadeia da polimerase
CRF - forma recombinante circulante	PR- protease
CTL - linfócitos T citotóxicos	RFLP - Restriction Fragment Length Polymorphism
DF- Distrito Federal	R5- vírus que utilizam CCR5 como co-receptor
DNA- ácido desoxirribonucléico	R5/X4- vírus que utilizam CCR5 e CXCR4 como co-receptor
HAART – terapia anti-retroviral altamente ativa	RNA- ácido ribonucléico
HIV- vírus da imunodeficiência humana	RRE- elemento de resposta a Rev
HIV-1 – vírus da imunodeficiência tipo 1	RT- transcriptase reversa
HIV-2 – vírus da imunodeficiência tipo 2	SI- fenótipo indutor de sincícios
HMA- heteroduplex mobility assay	SIV- simian immunodeficiency virus
HTLV - human T-lymphotropic virus	SU- proteína de superfície
IAMI- International AIDS Vaccine Initiative	TM- proteína transmembrana
ICTV – The International Committee on Taxonomy of Viruses	X4- vírus que utilizam CXCR4 como co-receptor
IN- integrase	
kD- kilodalton	
LACEN-DF- Laboratório Central de Saúde Pública do Distrito Federal	

## 1. REVISÃO BIBLIOGRÁFICA

### 1.1. Introdução

O vírus da imunodeficiência humana (HIV) é o agente etiológico da síndrome da imunodeficiência adquirida (*acquired immune deficiency syndrome* – AIDS) em humanos. Essa doença se caracteriza por imunossupressão acompanhada de infecções oportunistas, ou pela ocorrência de um câncer raro, o Sarcoma de Kaposi. A infecção pelo HIV é crônica e a progressão da doença é lenta (Prusiner, 2002).

Os primeiros casos da doença foram relatados em 1981, e logo a epidemia ganhou proporções mundiais. O grau de morbidade e mortalidade causados pelo HIV, o impacto global da infecção sobre a economia e sobre os recursos no atendimento aos portadores do vírus são enormes, e continuam crescendo. Segundo dados do *Joint United Nations Programme on HIV/AIDS* (UNAIDS-WHO) de dezembro de 2005, mais de 25 milhões de pessoas já morreram com AIDS desde seu primeiro relato, no início da década de 80, e, atualmente, cerca de 40 milhões de pessoas vivem infectadas com o vírus em todo mundo.

Tais números apenas enfatizam a importância dos esforços para o controle e combate a AIDS, já que ainda não existe uma imunização profilática ou cura para essa enfermidade. No entanto, avanços importantes já foram conseguidos. Com auxílio da pesquisa básica, foi possível desvendar a causa da AIDS, isolando e identificando seu agente causador, o HIV, em 1983. Atualmente, se conhecem pelo menos dois tipos de HIV: o HIV-1 e o HIV-2. O HIV-1, o tipo mais virulento, é responsável pela pandemia da AIDS. O HIV-2, além de ser menos patogênico, é restrito ao oeste da África (van Regenmortel *et al.*, 2000; Wigg, 2002).

Desde o isolamento do vírus, várias descobertas sobre a biologia do HIV-1 permitiram a elucidação das vias de transmissão, além do desenvolvimento de testes sorológicos e de drogas antivirais, que são cruciais no controle da infecção pelo vírus, até que seja possível a obtenção de uma vacina eficaz (Gallo & Montagnier, 2002).

## 1.2. Histórico

Os primeiros relatos de casos de AIDS surgiram em 1981 e eram relacionados à homens que faziam sexo com outros homens. A imunodeficiência era associada ao aparecimento de um tipo raro de câncer, o Sarcoma de Kaposi, e de uma série de outras infecções oportunistas que afetavam os indivíduos infectados. Logo um grande número de casos da doença foi registrado e a AIDS atingiu proporções epidêmicas. O surgimento de casos em diferentes populações, incluindo homens heterossexuais, usuários de drogas, hemofílicos, crianças e receptores de transfusões de sangue levaram a identificação de um mesmo agente etiológico, que era transferido através de fluídos corporais das pessoas infectadas (Prusiner, 2002).

A identificação e o isolamento do HIV-1 envolveu esforços de dois grandes grupos de pesquisas. O primeiro, liderado por Luc Montagnier e seus colegas, do Instituto Pasteur, em Paris. O segundo, comandado por Robert Gallo, do Centro para Controle e Prevenção de Doenças (*Centres for Disease Control and Prevention - CDC*), dos Estados Unidos. Além disso, várias descobertas da época foram cruciais para o isolamento do HIV, tais como, o isolamento da enzima transcriptase reversa (Baltimore, 1970; Temin & Mizutani, 1970) e a descoberta da interleucina-2 (Morgan *et al.*, 1976), fator de crescimento essencial à cultura de células de linfócitos T necessárias à replicação do HIV *in vitro* (Prusiner, 2002).

Luc Montagnier estudava os retrovírus causadores de câncer em humanos quando do surgimento da AIDS, em 1982, resolveu pesquisar a causa da doença, devido à hipótese levantada por Robert Gallo de que essa enfermidade seria provocada por um retrovírus. A partir de uma amostra da biópsia de um paciente com linfadenopatia, Montagnier e seus colaboradores fizeram um cultivo de linfócitos T. Em janeiro de 1983, foram descobertos no cultivo traços de transcriptase reversa, enzima presente apenas nos retrovírus, o que confirmou a hipótese de Gallo. Esse vírus foi, então, isolado e denominado de vírus associado a linfadenopatia (*lymphadenopathy associated virus - LAV*) (Montagnier, 2002).

Simultaneamente, o grupo de pesquisa de Gallo também estudava os retrovírus e foi o primeiro a isolar o HTLV (*human T-lymphotropic virus - HTLV*), vírus causador de leucemia em humanos. Ele defendia a hipótese de que a AIDS era causada por um vírus da mesma família do HTLV, pois o modo de transmissão era o mesmo e ambos se replicavam em linfócitos T. Sua equipe isolou um novo tipo de vírus linfotrópico de célula T em humanos e o denominou de HTLV do tipo III. No entanto, o agente



etiológico da AIDS possuía várias características distintas dos vírus da família HTLV. Esses não lisam linfócitos e sim provocam a sua imortalização. Assim, era pouco provável que a grande queda no número de células T fosse provocada por um HTLV. Posteriormente, foi comprovado que os dois grupos de pesquisa haviam descrito o mesmo vírus. Em 1984, um outro pesquisador, Jay Levy e seus colaboradores da Universidade da Califórnia, isolaram o vírus a partir de indivíduos assintomáticos, indicando o estado de portador do vírus, e denominando-o de retrovírus associado à AIDS (*AIDS-associated retrovirus* - ARV). Nesse mesmo ano a relação do HIV como agente etiológico da AIDS foi aceita pela comunidade médica e científica (Gallo, 2002; Montagnier, 2002; Prusiner, 2002).

Em 1986, o Comitê Internacional de Taxonomia de Vírus (*International Committee on Taxonomy of Viruses* – ICTV) recomendou a utilização da denominação de vírus da imunodeficiência humana (*Human immunodeficiency virus* – HIV). Ainda em 1986, um outro tipo de HIV foi isolado e denominado vírus da imunodeficiência do tipo 2, *Human immunodeficiency virus type 2* – HIV-2. Esse vírus também causa a AIDS, no entanto, de uma forma mais branda e restrita ao oeste da África. O HIV-1 e o HIV-2 podem ser diferenciados através do genoma e das relações filogenéticas com outros lentivírus de primatas (van Regenmortel *et al.*, 2000).

Apesar da não existência de uma vacina preventiva contra a infecção por HIV, grandes avanços já foram alcançados no controle da AIDS. A partir de 1986, os testes para detecção do HIV se tornaram comercialmente disponíveis, o que reduziu significativamente a transmissão do vírus por transfusão de sangue em países desenvolvidos. Em 1987, foi desenvolvido o AZT, primeira droga anti-HIV. Em 1995, foi disponibilizada a terapia anti-retroviral altamente ativa (HAART) que melhorou substancialmente a vida dos portadores do HIV (Gallo & Montagnier, 2002).

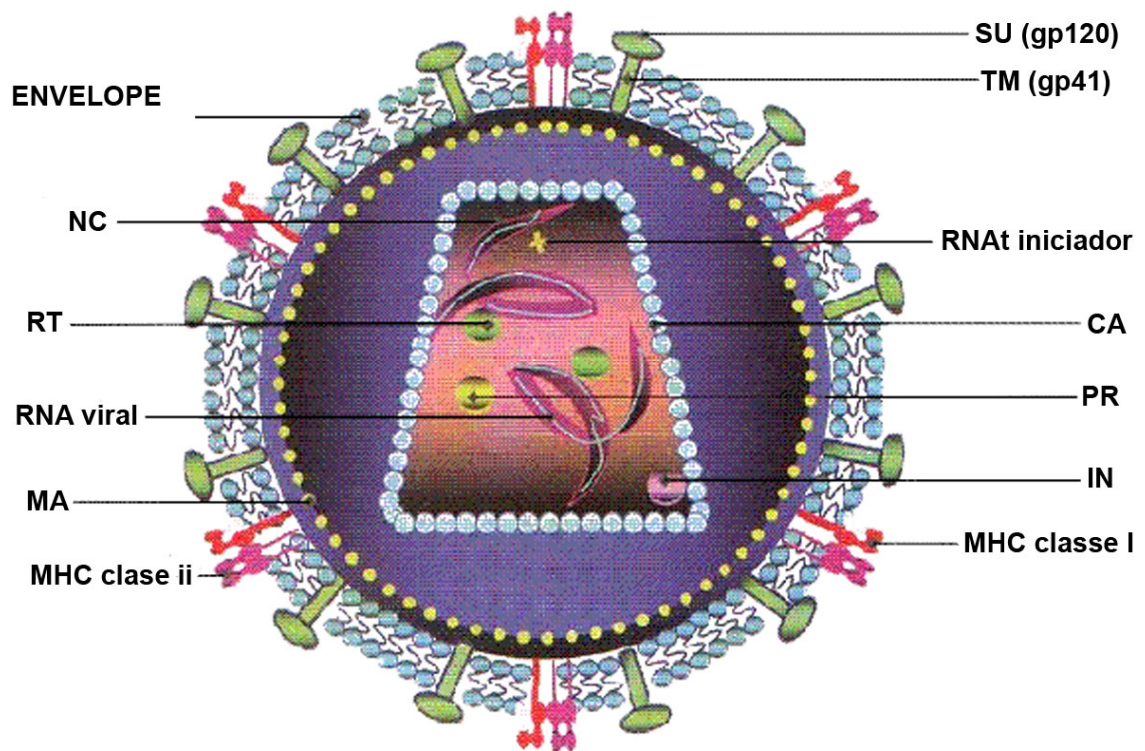
### **1.3. Classificação e Morfologia**

O HIV é um vírus com genoma de RNA de fita simples, pertencente ao gênero *Lentivirus* e à família *Retroviridae*. Os retrovírus utilizam a enzima transcriptase reversa para a transcrição do RNA em DNA. Após a síntese do DNA, o genoma viral se integra ao genoma da célula hospedeira e recebe o nome de provírus. Os retrovírus pertencentes ao gênero *Lentivirus*, do latim *lentus* (lentos), receberam essa denominação porque possuem um longo período de incubação no hospedeiro, podendo persistir e replicar-se por muitos anos antes de causar sinais clínicos de doença. Os lentivírus

produzem efeitos citopáticos, provocam deficiências imunológicas, desordens hepáticas e nervosas, além de doenças autoimunes. Os genomas desses vírus são relativamente grandes, apresentando um número maior de genes do que outros retrovírus. Além dos genes estruturais, *gag*, *pol* e *env*, comuns a todos os retrovírus, os lentivírus codificam proteínas acessórias e regulatórias da replicação viral. A descoberta do HIV levou à caracterização de vários outros lentivírus, como aqueles isolados de primatas não humanos, SIV (*simian immunodeficiency virus*) (Janeway *et al.*, 2002; Flint *et al.*, 2004).

Como mencionado anteriormente, estudos genômicos e sorológicos demonstraram a existência de dois tipos distintos de HIV: o HIV-1 e o HIV-2. A caracterização de vários isolados do HIV-1 permitiu a distinção de três grandes grupos: M (*major*), O (*outlier*) e N (non M, non O) (van Regenmortel *et al.*, 2000; Janeway *et al.*, 2002; Flint *et al.*, 2004).

A partícula do HIV-1 possui em seu interior duas cópias de um genoma de RNA de fita simples. O *virion* possui um diâmetro de aproximadamente 100 nm, um genoma de cerca de 9,8 kb e um envelope, oriundo da célula hospedeira. As glicoproteínas virais são derivadas de um precursor gp160, que é clivado por proteases celulares, produzindo as glicoproteínas gp120 de superfície (SU) e gp41, a glicoproteína transmembrana (TM). No interior do envelope encontra-se um segundo envoltório, a matriz, que é constituída pela proteína p17. O capsídeo é formado pela proteína p24, e no interior dele encontram-se as duas cópias do genoma de RNA fita simples, com polaridade positiva, associadas às proteínas p7 e p6, formando o nucleocapsídeo. As enzimas virais transcriptase reversa (RT), integrase (IN) e protease (PR) também são encontradas dentro da partícula viral (Figura 1) (Larder *et al.*, 2001).

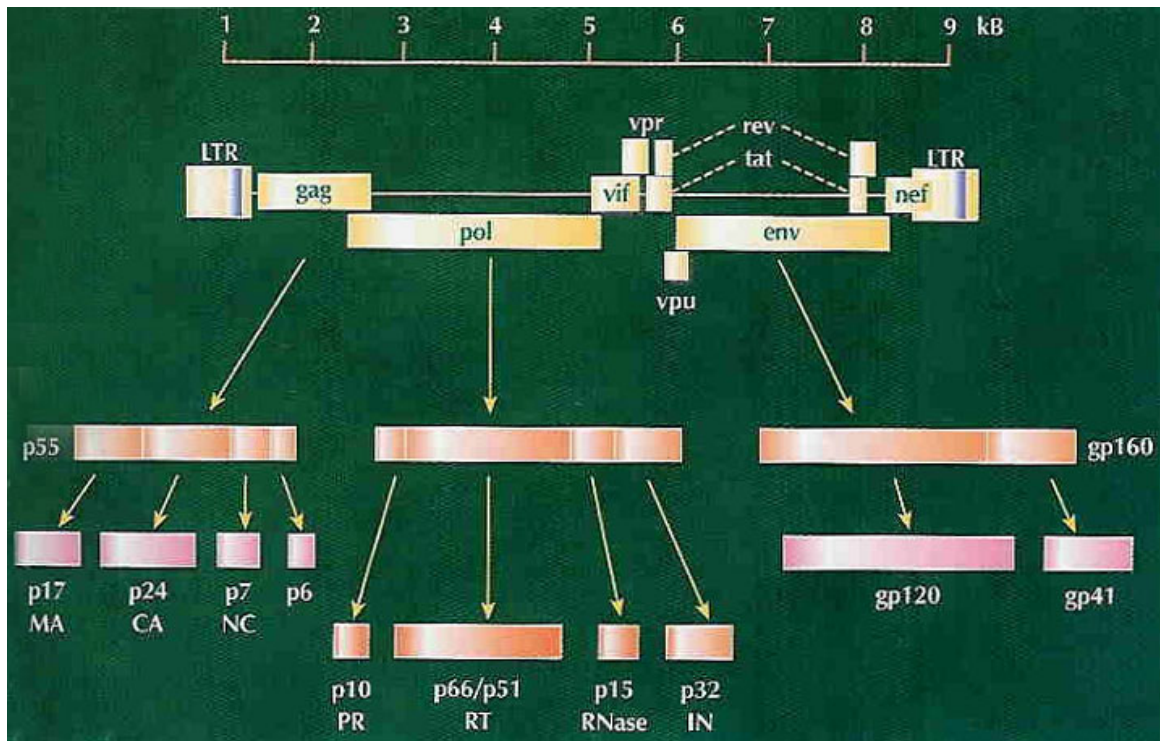


**Figura 1.** Estrutura da partícula do HIV-1 (adaptado de Wigg, 2002).

#### 1.4. Organização do genoma do HIV-1

O genoma do HIV-1 contém pelo menos nove genes diferentes, sendo três deles encontrados em todos os retrovírus, são eles: *gag* (que codifica as proteínas estruturais), *pol* (que codifica as enzimas virais protease [PR], transcriptase reversa [RT] e integrase [IN]) e *env* (que codifica as glicoproteínas do envelope). Esses genes são considerados os genes estruturais. Os genes *gag*, *pol* e *env* são traduzidos em poliproteínas precursoras, que são posteriormente clivadas por proteases celulares e por uma protease viral (Flint *et al.*, 2004).

Há ainda os genes regulatórios, *tat* e *rev*, além de quatro genes acessórios: *nef*, *vif*, *vpr* e *vpu*. A diferença genômica entre o HIV-1 e o HIV-2 está na presença do gene *vpx* no HIV-2. Nas extremidades 3' e 5' do provírus são encontradas longas seqüências repetidas denominadas de LTRs (*long terminal repeats*), que regulam a integração do vírus ao genoma hospedeiro, a expressão gênica e a replicação (Figura 2) (Larder *et al.*, 2001; Wigg, 2002).



**Figura 2.** Mapa genômico do HIV-1 (adaptado de Larder *et al.*, 2001).

O gene *gag* é traduzido em uma poliproteína precursora de 55kD, denominada p55 Gag. Essa é clivada pela protease viral em quatro proteínas menores: p17 que forma a matriz; p24 que constitui o capsídeo; p7 que constitui o nucleocapsídeo; e p6 que medeia a incorporação da proteína Vpr aos *virions*. Essas quatro proteínas são encontradas nas partículas virais maduras do HIV-1 (Göttlinger, 2001; Larder *et al.*, 2001).

O gene *pol* codifica uma poliproteína precursora das enzimas virais PR, RT e IN. A seqüência do gene *pol* não possui um códon de iniciação. Assim, não pode ser traduzida independentemente. A poliproteína Pol é traduzida fusionada a Gag, formando um precursor denominado Gag-Pol. A protease viral cliva o polipeptídeo Pol, separando-o de Gag, e, posteriormente, cliva Pol para gerar PR, RT e a IN. Porém, é somente na meio extracelular, após a liberação do vírus da célula hospedeira, que a protease viral sofre auto-ativação e inicia o processo de clivagem, tornando infecciosa a partícula viral. Todos os produtos do gene *pol* podem ser encontrados dentro das partículas virais do HIV-1 (Larder *et al.*, 2001).

O gene *env* codifica uma poliproteína precursora, que é sintetizada na forma de um trímero, com ligações não covalentes entre as subunidades gp120 e gp41. A função

e a estrutura dessas proteínas serão apresentadas de forma mais detalhada adiante, página 23 (Flint *et al.*, 2004; Moore & Doms, 2003).

O gene *tat* codifica um transativador transcricional que é essencial à replicação do HIV-1. Tat é uma proteína que se liga ao RNA, reconhece uma seqüência em forma de *loop*, denominada elemento de resposta à transativação (*transactivation response element- TAR*), localizada na extremidade 5' de todos os RNAs mensageiros do HIV-1. Como resultado dessa ligação, há o aumento na taxa de produção de transcritos virais primários de pelo menos 1000 vezes (Larder *et al.*, 2001; Flint *et al.*, 2004).

O transcrito completo do HIV possui múltiplos sítios de processamento alternativo. O processamento alternativo dos RNAs mensageiros é necessário para expressão eficiente dos genes virais. As proteínas Tat, Rev e Nef são processadas na fase inicial da infecção e acumulam-se devido à ativação transcricional produzida por Tat. O acúmulo da proteína Rev é responsável pela entrada na fase tardia do ciclo de replicação do HIV-1. Rev é uma proteína que se liga ao RNA e reconhece um elemento estrutural específico localizado em *env*, denominado elemento de resposta a Rev (*Rev Response Element- RRE*). Essa proteína ativa a exportação nuclear de transcritos não processados completos ou parcialmente processados, ou seja, que não sofreram processamento e contém o elemento RRE. Dessa forma, Rev facilita a síntese das proteínas virais estruturais e garante a viabilidade do genoma completo de RNA do HIV-1. Ela também está envolvida no transporte dos transcritos de fase tardia dos genes *vpr*, *vpu* e *vif* para o citoplasma da célula hospedeira (Miller *et al.*, 2000; Larder *et al.*, 2001; Flint *et al.*, 2004).

O gene *nef*, (*negative factor*), codifica um pequeno polipeptídeo, está associado à estimulação da infecção viral, da resistência à ação do sistema imune, e, conseqüentemente, da progressão mais rápida para a AIDS. Diversas funções são atribuídas a *nef*. Algumas delas estão associadas à modulação negativa da expressão de CD4 e de MHC classe I nos linfócitos T, às interações com as proteínas do citoesqueleto, e à inibição dos sinais apoptóticos nas células infectadas pelo HIV (Marsh, 1999; Fackler & Baur, 2002).

A proteína viral U ou Vpu é essencial à maturação e liberação da progênie viral. Nas células infectadas pelo HIV, são formados complexos entre o receptor celular CD4 e a glicoproteína do envelope gp120 no meio intracelular. Tal processo interfere na montagem da partícula viral. Vpu tem a função de liberar a glicoproteína do envelope

gp120 pela clivagem da molécula de CD4 ligada a ela (Larder *et al.*, 2001; Bour & Strebel, 2003; Flint *et al.*, 2004).

A proteína viral R ou Vpr é incorporada à partícula viral pelo peptídeo p6 de Gag. Uma das funções atribuídas a Vpr é facilitar a localização nuclear do complexo de pré-integração, ligando-se a proteínas dos poros nucleares. Vpr também pode bloquear a divisão celular, mantendo o ciclo celular na fase G2, ligar-se a LTR viral e potencializar a transcrição de genes pela interação com fatores transcricionais celulares (Larder *et al.*, 2001; Flint *et al.*, 2004).

A proteína viral Vif, fator de infectividade viral (*viral infective factor*), tem papel importante na infecção e na replicação do vírus, sendo essencial para infecções *in vivo*. Essa proteína também pode bloquear o processamento prematuro da poliproteína Gag pela protease viral ainda no citoplasma (Miller *et al.*, 2000; Flint *et al.*, 2004). As funções desses genes estão resumidas na Tabela 1.

**Tabela 1.** Principais funções das proteínas codificadas pelo genoma do HIV-1 (adaptado de Janeway *et al.*, 2002).

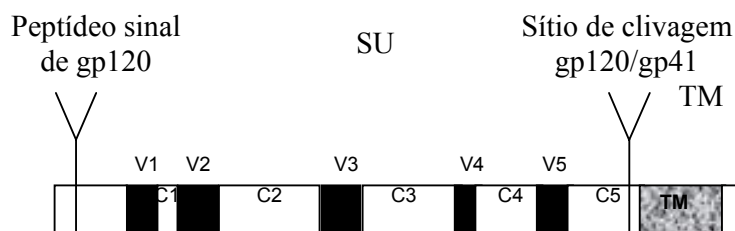
Gene		Função
<i>Gag</i>	Antígeno grupo-específico	Proteínas do <i>core</i> e da matriz.
<i>Pol</i>	Polimerase	Enzimas transcriptase reversa, protease e integrase.
<i>Env</i>	Envelope	Glicoproteínas: gp120 se liga ao CD4 e a receptores de quimiocinas e gp41 intermedeia a fusão do vírus à célula hospedeira.
<i>Tat</i>	Transativador	Regula positivamente a transcrição.
<i>Rev</i>	Regulador da expressão viral	Permite a exportação de transcritos não processados no núcleo da célula hospedeira.
<i>Vif</i>	Fator de infectividade viral	Aumenta a capacidade infectiva das partículas virais.
<i>Vpr</i>	Proteína viral R	Promove a importação nuclear de DNA viral e a parada do ciclo celular da célula hospedeira em G2.
<i>Vpu</i>	Proteína viral U	Promove a degradação intracelular de CD4 e aumenta a liberação de partículas virais na membrana celular da célula hospedeira.
<i>Nef</i>	Fator de regulação negativa	Regula negativamente a expressão de CD4 e de MHC classe I da célula hospedeira e aumenta a infectividade viral.

### 1.5. As proteínas do envelope

As glicoproteínas do envelope possuem papel crucial nos eventos iniciais da infecção pelo HIV, mediando a ligação do vírus à célula hospedeira e a fusão das membranas viral e celular. As proteínas do envelope também são os principais alvos para a resposta imune anti-HIV nos indivíduos infectados, devido a sua localização na superfície da partícula viral. Essas glicoproteínas são formadas por duas subunidades, gp120 e gp41, que são produto da clivagem por proteases celulares da poliproteína precursora gp160. As proteínas gp120 e gp41 formam um complexo trimérico no envelope viral, permanecendo ligadas não covalentemente após a clivagem proteolítica. (Flint *et al.*, 2004; Moore & Doms, 2003).

A glicoproteína gp120 permite o reconhecimento da célula alvo e determina o tropismo celular pela ligação ao receptor celular CD4 e a um dos vários co-receptores membros da família de receptores de quimiocinas, mais frequentemente CCR5 ou CXCR4. A glicoproteína gp41 promove a fusão da membrana viral com a celular, processo que resulta na liberação do nucleocapsídeo viral no interior da célula hospedeira. Durante esse processo, gp41 sofre várias mudanças conformacionais que levam à exposição de um peptídeo de fusão, uma região hidrofóbica rica em resíduos de glicina, essencial ao processo de fusão das membranas viral e celular. A glicoproteína gp41 possui em sua estrutura uma porção citoplasmática, que ancora o complexo das proteínas do envelope na matriz do *virion*, uma região transmembrana hidrofóbica, e um ectodomínio onde se localiza o peptídeo de fusão (Chan *et al.*, 1997; Golding *et al.*, 2002; Kang *et al.*, 2005).

A proteína gp120 é composta por cinco domínios constantes (C1-C5) e por cinco *loops* variáveis (V1-V5) intercalados. Esses domínios V1-V5 possuem grande variabilidade genética. A Figura 3 apresenta um esquema da organização estrutural de gp120 (Wyatt & Sodroski, 1998).

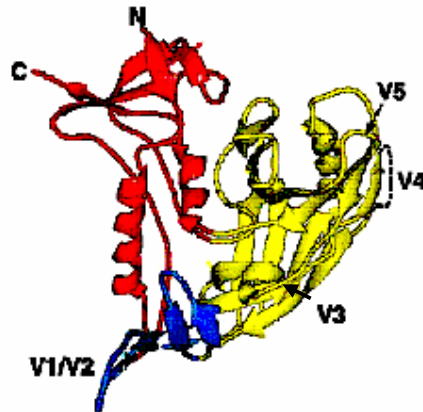


**Figura 3.** Esquema da estrutura de Env. Os *loops* variáveis de gp120 estão em preto, as regiões constantes em branco e em cinza está destacado o domínio transmembrana de gp41 (adaptado de Milich *et al.*, 1993).

Análises estruturais de gp120 revelaram a existência de pontes intramoleculares dissulfeto, que permitem a incorporação das quatro primeiras regiões variáveis em estruturas em forma de *loops* expostos na superfície da glicoproteína. A região mais conservada dessa proteína, o *core*, é formada por dois domínios distintos, um



externo exposto na superfície viral, e um interno que se liga não covalentemente a gp41 (Figura 4). No domínio externo, localizam-se os sítios de ligação à molécula de CD4 e aos co-receptores, principalmente CCR5 e CXCR4 (Wyatt & Sodroski, 1998; Wyatt *et al.*, 1998).



**Figura 4.** Modelo estrutural do *core* de gp120, mostrando as posições dos *loops* variáveis (V1 a V5). O domínio interno está representado em vermelho; o externo em amarelo; a conexão entre eles em azul; C representa a extremidade c-terminal da proteína e N a porção N-terminal. (adaptado de Wyatt *et al.*, 1998).

A função crucial que gp120 desempenha no início da infecção viral e sua constante exposição ao sistema imune, fazem desta proteína um importante alvo para vacinas. No entanto, as primeiras vacinas baseadas em Env, compostas por subunidades monoméricas de gp120 se mostraram ineficientes. Atualmente, a produção de partículas oligoméricas de gp120, que mimetizem de forma mais fiel às proteínas nativas do envelope, é uma das estratégias a serem exploradas para se obter uma imunização mais eficaz. Além disso, a maioria dos anticorpos anti-gp120 não possui capacidade de reação cruzada, sendo específicos apenas para uma linhagem viral, devido a grande variabilidade genética (Burton *et al.*, 2004; Herrera *et al.*, 2005). Entretanto, já foram isolados alguns anticorpos monoclonais de reação cruzada, como, b12, 2G12 e 447-52D, que são capazes de neutralizar diferentes isolados primários de HIV-1. Estes anticorpos serão discutidos mais detalhadamente adiante (página 42).

### 1.5.1. Loop V3

O *loop V3* é uma região hipervariável, localizada na porção externa da glicoproteína do envelope, gp120 (Figura 4). Essa região possui de 35 a 40 aminoácidos e varia muito entre diferentes linhagens virais. Entretanto, o *loop V3* apresenta resíduos de aminoácidos conservados, em seu ápice (GPGR ou GPGQ), que são essenciais ao reconhecimento pelo HIV-1 do co-receptor. Além disso, existe outra região conservada, uma ponte dissulfeto, Cys296-Cys331 na base do *loop V3* (Kwong, 2004; Stanfield *et al.*, 2004).

Vários estudos têm demonstrado a importância biológica da heterogeneidade do *loop V3* em diversas funções virais, tais como, na ligação ao co-receptor tanto CCR5 como CXCR4, e na determinação do tropismo celular. Nessa região também são encontrados muitos epítopos lineares e conformacionais indutores de anticorpos neutralizantes anti-HIV-1 e de resposta de linfócitos T citotóxicos. Devido à capacidade de indução de anticorpos neutralizantes que os peptídeos de V3 possuem, essa região foi conhecida por muito tempo como “determinante principal de neutralização” (*principal neutralizing determinant* - PND) (Casseb *et al.*, 2002; Berzofsky *et al.*, 2004; Stanfield *et al.*, 2004; Zhi-yong *et al.*, 2004).

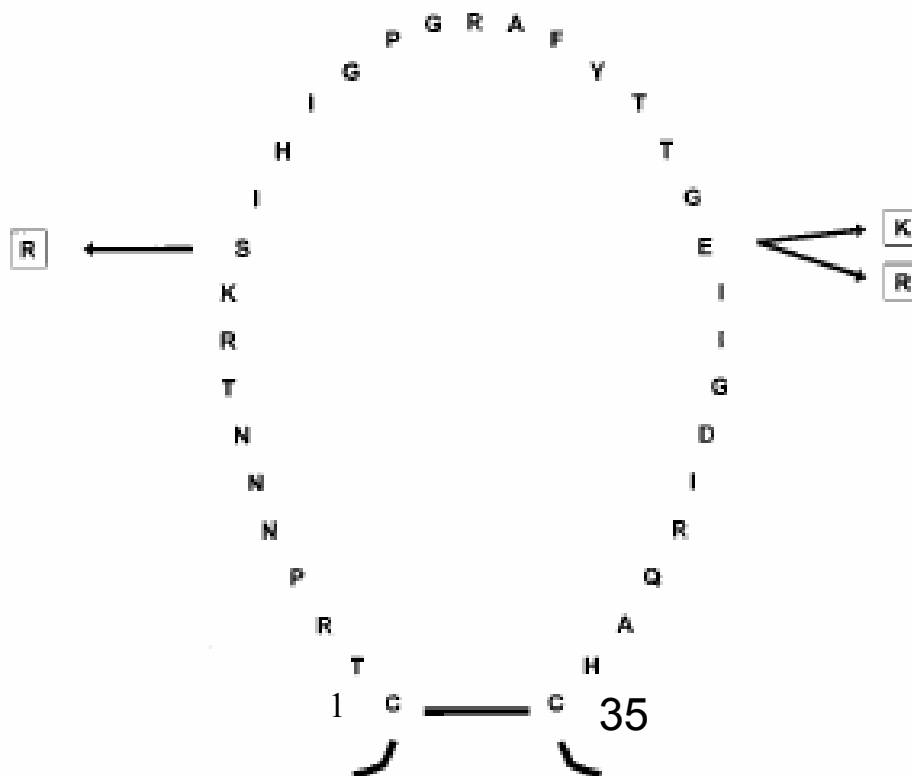
Anticorpos anti-V3 constituem o maior grupo de anticorpos monoclonais humanos existentes. Até 2003, havia 33 anticorpos monoclonais humanos anti-V3 gerados a partir de indivíduos infectados. Tais anticorpos, em sua maioria, são específicos para uma determinada linhagem viral, não sendo capazes de neutralizar diferentes isolados de HIV-1. Recentemente, foram descritos anticorpos anti-V3 com capacidade de reação cruzada e que reconhecem diversas linhagens de HIV-1. Esses anticorpos serão descritos posteriormente (página 42) (Gorny & Zolla-Pazner, 2004; Berzofsky *et al.*, 2004; Burton *et al.*, 2004; Stanfield *et al.*, 2004).

Além do potencial imunogênico, o *loop V3* tem papel importante no fenótipo indutor de sincícios, no tropismo viral e na determinação do tipo de co-receptor a ser utilizado pelo vírus, se CCR5 ou CXCR4. Glicoproteínas gp120 com deleções no *loop V3* não se ligam ao co-receptor, mesmo após a ligação ao CD4. Anticorpos anti-V3 impedem a entrada do vírus na célula porque interferem na ligação gp120-CCR5 (Fouchier, 1995; Trkola *et al.*, 1996; Wyatt & Sodroski, 1998).

Os isolados de HIV-1 podem ser distribuídos em duas categorias, de acordo com os efeitos citopáticos induzidos: indutor de sincícios (SI) e não indutor de sincícios

(NSI). Essa característica está relacionada ao tipo celular infectado e às interações com o co-receptor. Isolados primários de HIV-1, em sua maioria, utilizam o co-receptor CCR5 para a entrada na célula. As linhagens virais que utilizam esse co-receptor são denominadas R5. Para a maioria desses isolados, que são transmitidos e predominam durante a fase inicial da doença, a presença de CCR5 é crucial para a infecção da célula. Há relatos de indivíduos que apresentam resistência à infecção pelo HIV-1 devido às falhas na expressão de CCR5 em suas células. Essas linhagens virais infectam preferencialmente a macrófagos e não induzem a formação de sincícios. Isolados de HIV-1, que surgem na fase mais tardia da infecção utilizam, mais freqüentemente, o co-receptor CXCR4 em lugar de CCR5, ou ambos, e são denominados X4 ou R5/X4, respectivamente. Esses vírus estão associados à rápida progressão da AIDS, infectam os linfócitos T e produzem efeitos citopáticos, formando sincícios (Milich *et al.*, 1993; Fouchier, 1995).

Pouco se sabe sobre o mecanismo de variação desses vírus, como ocorre a transição do uso do co-receptor CCR5 para CXCR4, e as alterações na patogenicidade viral já descritas. No entanto, algumas alterações pontuais no *loop* V3 estão associadas a um determinado fenótipo, ou ao uso de CCR5 ou CXCR4. Aminoácidos nas posições 11, 25 (aminoácidos 306 e 320 na seqüência referência HXB2) carregados positivamente são freqüentemente encontrados nos vírus que utilizam CXCR4. Além disso, o aumento da carga positiva nos aminoácidos do *loop* V3 também está associado aos vírus R4. Essa relação foi encontrada nos *loops* V3 dos subtipos A, B, C, D, CRF01\_AE, F e G. Dessa forma, variações genéticas entre os subtipos podem alterar a patogenicidade viral. A Figura 5 apresenta um esquema da seqüência de aminoácidos do *loop* V3 e as substituições que caracterizam um fenótipo SI (Fouchier, 1995; Koning *et al.*, 2002; Thomson *et al.*, 2002b).

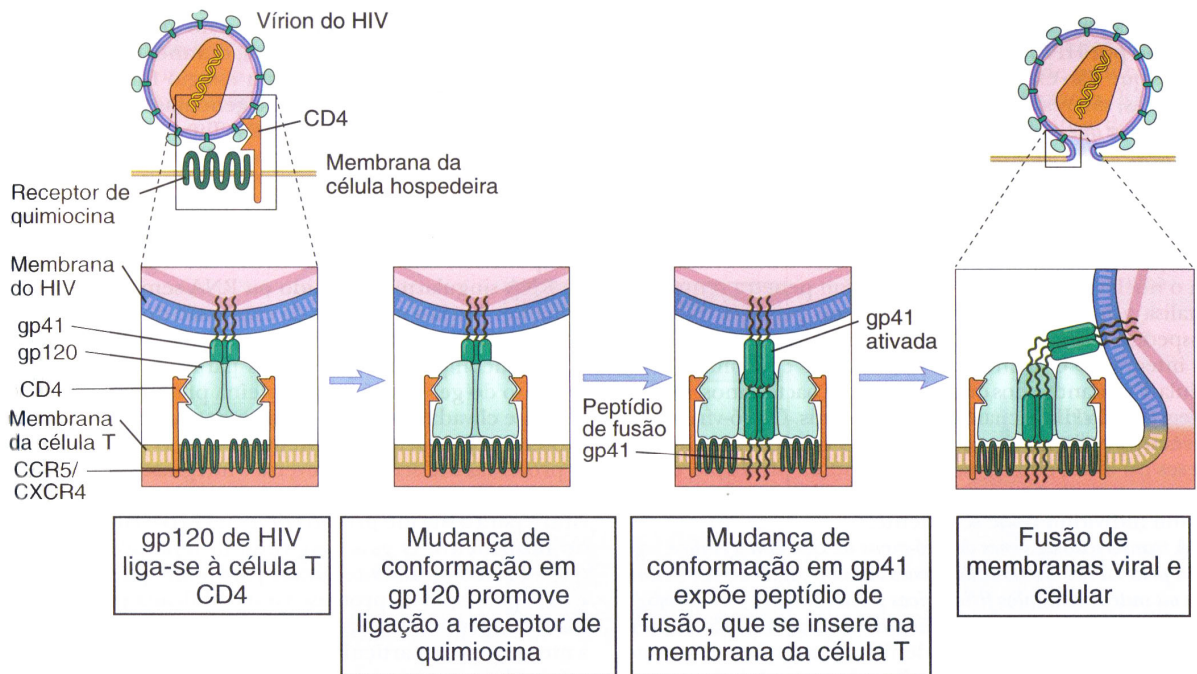


**Figura 5.** Esquema da seqüência de aminoácidos do *loop* V3 de isolados de HIV-1 SI (indutor de sincícios) e NSI (não indutor de sincícios). Isolados com arginina (R) na posição 11 (posição 306 em HXB2) e/ou arginina ou lisina (K) na posição 25 (posição 320 em HXB2) são considerados SI (adaptado de Fouchier, 1995).

## 1.6. Ciclo de replicação do vírus

A infecção das células pelo HIV-1 tem início com a entrada do vírus na célula. Esse processo se dá por interações consecutivas entre as glicoproteínas do envelope, gp120 e gp41, e os receptores de membrana CD4, além de um co-receptor, CCR5 ou CXCR4, que pertencem à família dos receptores de quimiocinas. A expressão conjunta das moléculas de CD4 e de um dos co-receptores de quimiocinas possibilitou a identificação de monócitos, macrófagos, células da glia, células dendríticas, células de Langerhans e células linfóides, como alvos da infecção viral. Na entrada, inicialmente gp120 se liga a molécula de CD4. Tal ligação promove mudanças conformacionais nessa glicoproteína, que expõe novos sítios de ligação e permite a interação de gp120 com o co-receptor (CCR5 ou CXCR4). Essas interações promovem mudanças

conformacionais em gp41 e a exposição de uma região hidrofóbica, chamada de peptídeo de fusão, que se insere na membrana celular e permite a fusão da membrana viral com a da célula-alvo. Esse processo é mediado pela formação de um complexo de seis hélices, três de origem C-terminal e três de origem N-terminal, organizadas em um arranjo antiparalelo. Devido a essa transição, o peptídeo de fusão e o domínio transmembrana de gp41 ficam muito próximos, assim como a membrana celular e a viral. O ganho de energia livre proveniente desse mecanismo permite a fusão das duas membranas. O processo de entrada do vírus na célula está ilustrado na Figura 6 (Moore & Doms, 2003; Abbas & Lichtman, 2005).



**Figura 6.** Processo de entrada do vírus na célula. Nesse modelo, mudanças conformacionais sequenciais em gp120 e gp41 ocorrem devido à ligação ao receptor CD4. Essas mudanças promovem a ligação do vírus ao co-receptor (um receptor de quimiocina, CCR5/CXCR4) e a fusão das membranas do HIV-1 e da célula hospedeira. O peptídeo de fusão de gp41 ativado contém resíduos hidrofóbicos que medeiam a inserção na membrana plasmática da célula hospedeira (adaptado de Abbas & Lichtman, 2005).

Após a entrada do *vírion* na célula, as enzimas presentes no nucleocapsídeo se tornam ativas e iniciam o ciclo de reprodução viral. O genoma do vírus é transcrito pela enzima viral transcriptase reversa. O produto desta enzima é um DNA linear de dupla fita que penetra no núcleo. A integrase viral também entra no núcleo e catalisa a inserção do DNA viral no genoma da célula hospedeira. (Flint *et al.*, 2004; Abbas & Lichtman, 2005).

O início da transcrição dos genes do HIV-1 nas células T está associado à ativação fisiológica das células TCD4<sup>+</sup> por antígenos e citocinas. A ativação dessas células induz a produção de diversos fatores de transcrição, que se ligarão não somente aos promotores no DNA celular, mas também, a sítios localizados na LTR do provírus, iniciando a transcrição pela RNA polimerase celular. O produto da transcrição é processado de várias maneiras para produzir mRNAs que codificam todas as proteínas virais (Janeway *et al.*, 2002; Abbas & Lichtman, 2005).

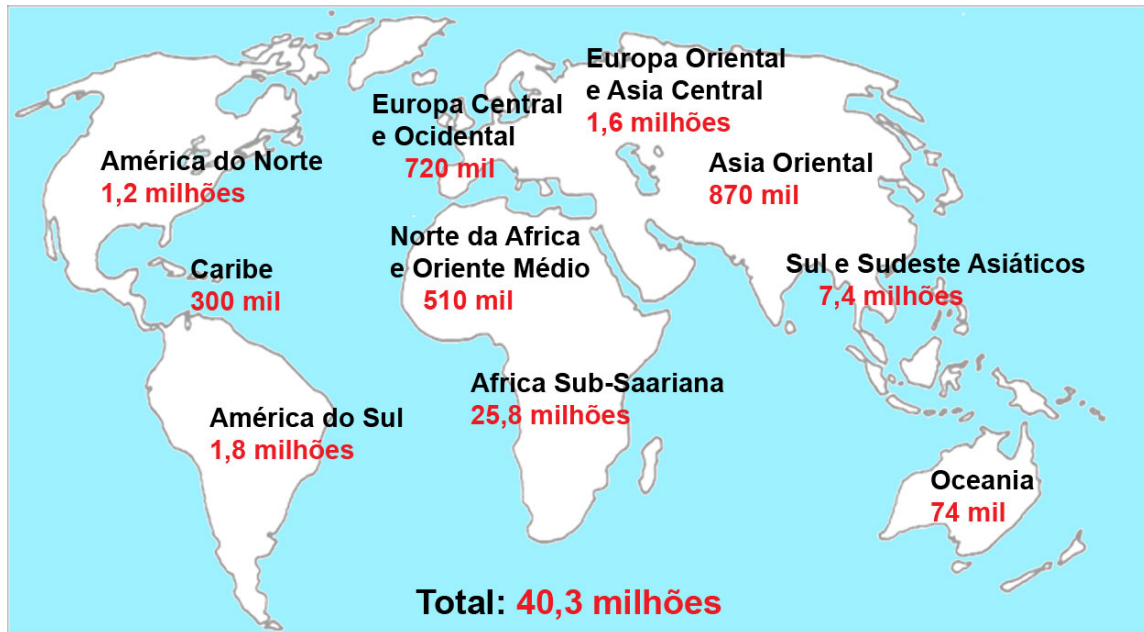
As proteínas virais Tat e Rev são essenciais à replicação viral nas células T ativadas. Tat é um potente regulador transcricional, que funciona como um fator de alongação, ligando-se ao mRNA incipiente e melhorando a transcrição deste pela RNA polimerase. A proteína Rev permite a exportação para fora do núcleo de mRNAs não processados para a tradução de diversas proteínas virais (Flint *et al.*, 2004).

A expressão gênica do HIV pode ser dividida em fase precoce em que os genes regulatórios *rev*, *tat* e *nef* são expressos, e em uma fase tardia, em que os genes estruturais são expressos e genomas completos são produzidos. A mudança de fase de expressão gênica é iniciada pela produção e acúmulo da proteína Rev, que promove a exportação dos mRNAs tardios para fora do núcleo (Flint *et al.*, 2004).

A montagem dos *vírions* começa pelo acúmulo de transcritos completos do genoma de RNA do provírus dentro de um complexo de nucleoproteínas, formado pelas proteínas básicas de *gag* e pelas enzimas codificadas por *pol*, necessárias ao ciclo seguinte de integração. Esse nucleocapsídeo é envolvido por um envelope, oriundo da membrana plasmática da célula no qual estão presentes tanto proteínas de membrana celulares como glicoproteínas gp120 e gp41, e liberado por brotamento da célula hospedeira (Abbas & Lichtman, 2005; Janeway *et al.*, 2002).

## 1.7. Epidemiologia

De acordo com *Joint United Nations Programme on HIV/AIDS* (UNAIDS/WHO), cerca de 40,3 milhões de pessoas vivem infectadas pelo HIV em todo mundo e mais de 5 milhões foram infectadas somente no ano de 2005. Nesse mesmo ano, mais de 3,1 milhões de óbitos foram provocados por essa doença. A Figura 7 ilustra a distribuição de casos de HIV em todo mundo, segundo relatório da UNAIDS/WHO de 2005.



**Figura 7.** Distribuição do número de pessoas infectadas pelo HIV no mundo (adaptado de UNAIDS/WHO, 2005).

O número de pessoas infectadas pelo HIV tem crescido em todo mundo. No entanto, a área geográfica mais afetada é a África Sub-saariana, com cerca de 25,8 milhões de pessoas infectadas pelo vírus, e 2,4 milhões de mortes provocadas pela epidemia no ano de 2005 (UNAIDS/WHO, 2005).

No Brasil, desde a identificação dos primeiros casos de AIDS, de 1980 até 2005, já foram notificados cerca de 372 mil casos da doença. Apesar da redução na mortalidade devido à adoção da política de acesso universal ao tratamento anti-retroviral, o impacto da AIDS ainda é grande. Até o ano de 2004, cerca de 172 mil óbitos foram registrados como sendo causados pela AIDS. A Tabela 2 apresenta a

distribuição de casos de AIDS nas regiões do Brasil, no período entre 1980 e 2005 (Brasil, 2005).

**Tabela 2.** Casos de AIDS por região geográfica do Brasil, período de 1980 a 2005 (Fonte: Boletim Epidemiológico AIDS –DST 2005).

<i>Região</i>	<i>Número</i>
Norte	10.149
Nordeste	38.837
Centro-Oeste	21.099
Sudeste	234.736
Sul	67.006
<b>TOTAL</b>	<b>371.827</b>

A epidemia tem demonstrado uma tendência à estabilização em algumas regiões do país, como é o caso da região Sudeste. No entanto, o número de casos tem aumentado nas regiões Norte e Centro-Oeste. Segundo dados do Programa Nacional de DST e AIDS do Ministério da Saúde, houve um aumento crescente nas taxas de incidência da doença até meados dos anos 90, chegando a alcançar, em 1998, índices de 18 casos por 100 mil habitantes. Atualmente, a taxa de incidência é de 17,2 casos por 100 mil habitantes.

No Distrito Federal foram registrados cerca de 5.000 casos de AIDS, e aproximadamente 2.000 óbitos desde 1985. No entanto, a partir do ano 2000, se observa uma estabilização na taxa de mortalidade nessa região geográfica (Brasil, 2005). Considerando-se que a epidemia continua a crescer no Brasil e no Distrito Federal é essencial o monitoramento da infecção pelo HIV-1 com relação à distribuição geográfica, subpopulações em risco e prevalência dos subtipos virais (Sabino *et al.*, 1996).

### **1.8. A variabilidade genética e a distribuição geográfica do HIV-1**

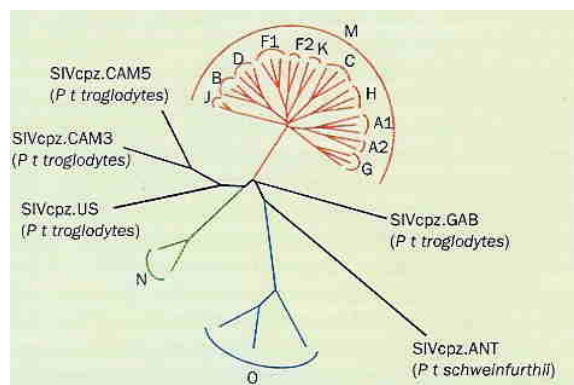
O HIV-1 apresenta uma alta variabilidade genética, resultante da taxa elevada de geração de novos *virions* (10 bilhões de partículas virais por dia), de substituições nucleosídicas incorretas durante a transcrição reversa (aproximadamente  $10^{-4}$  por



nucleotídeo) e de mutações ( $3 \times 10^{-5}$  mutações por nucleotídeo por dia). Além disso, a transcriptase reversa (RT) possibilita a recombinações dos genomas virais em indivíduos infectados por subtipos virais distintos e é incapaz de reparar as seqüências nucleotídicas durante o processo de replicação. Tais fatores geram uma grande diversidade de formas genéticas distintas, porém relacionadas entre si, denominadas quasiespécies. A existência de quasiespécies, que representam praticamente todos os mutantes possíveis, gera uma resistência natural à ação dos medicamentos e do sistema imune (Hirsch *et al.*, 1998; Larder *et al.*, 2001; Thomson *et al.*, 2002b).

Os primeiros estudos para a classificação do HIV-1 eram baseados em seqüências parciais dos genes *gag* e *env*. No entanto, a identificação de novas formas genéticas que foi possível com a caracterização de todo genoma viral, mostrou que essa metodologia para a classificação deveria ser revisada. Atualmente, análises filogenéticas de linhagens virais de diferentes regiões geográficas permitiram a classificação do HIV-1 em três grandes grupos: M (*major*), O (*outlier*) e N (non M, non O). O grupo M é o mais disseminado pelo mundialmente e o responsável pela pandemia da AIDS. É nesse grupo que se encontra a maioria dos HIVs do tipo 1. O grupo O é endêmico da República de Camarões e de outros países vizinhos da África Central. Porém, mesmo nessas regiões geográficas, esses vírus representam a minoria das linhagens do HIV-1 existentes, apresentando uma prevalência de 2 a 5% entre as linhagens do HIV-1 circulantes. O grupo N foi identificado mais recentemente e está restrito a um pequeno número de casos registrados na República de Camarões (Peeters, 2000; Gaschen *et al.*, 2002; <http://hiv-web.lanl.gov>).

Em relação ao grupo M, existem informações suficientes para a classificação das linhagens virais pertencentes a esse grupo em subtipos, subsubtipos e CRFs. São identificados nove subtipos distintos, denominados por meio das letras A-D, F-H-J e K. Esses subtipos formam grupos equidistantes entre si nas análises filogenéticas. Alguns subtipos podem ainda ser separados em subsubtipos distintos. O subtipo A pode ser dividido nos subsubtipos A1 e A2 e o subtipo F em F1 e F2. Os subtipos B e D são mais próximos entre si em relação aos outros subtipos. Mas devido a razões históricas, a separação entre os subtipos B e D permanece sendo usada (Figura 8) (Peeters, 2000; Thomson *et al.*, 2002b).



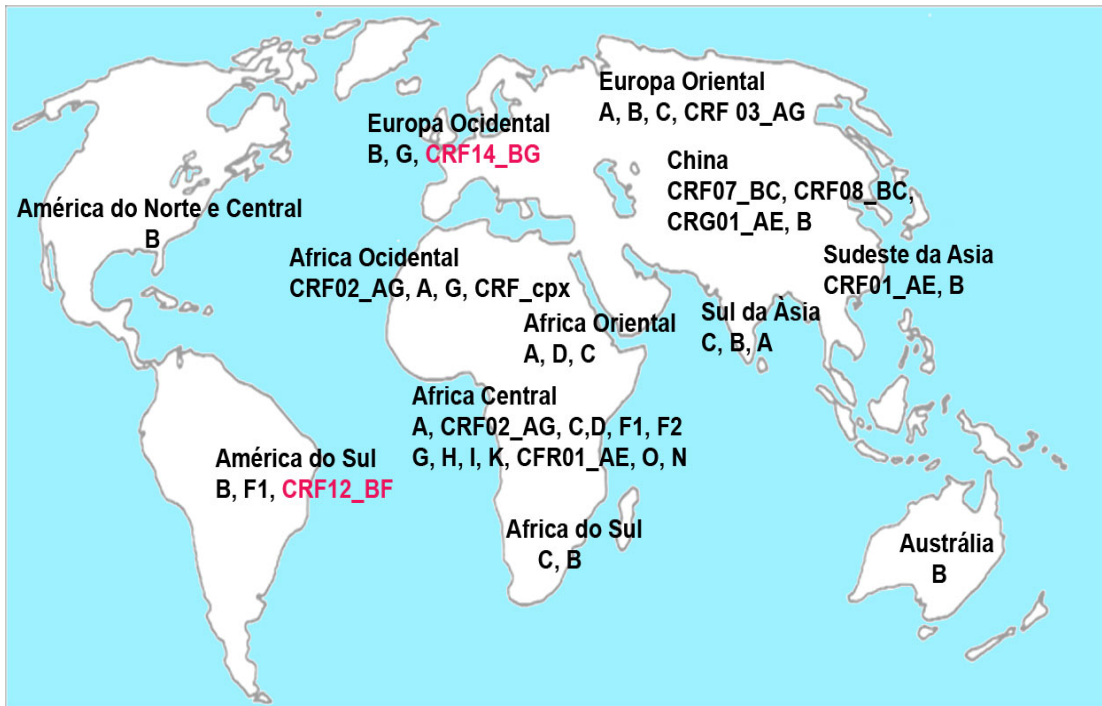
**Figura 8.** Filogenia dos grupos M, N e O do HIV-1. O grupo M é formado por grupos equidistantes entre si, o que possibilitou sua classificação em subtipos e subsubtipos (adaptado de Thomson, 2002b).

Após a classificação do grupo M em subtipos, foi observado que alguns isolados, quando analisados filogeneticamente, eram classificados em diferentes grupos se seqüências distintas fossem analisadas. Esses genomas foram denominados formas recombinantes circulantes (CRFs) ou mosaicos. A formação das CRFs tem origem nos eventos de recombinação ocorridos entre fragmentos genômicos de duas linhagens virais de subtipos diferentes que infectam, simultaneamente, a mesma célula. Atualmente, 16 formas recombinantes circulantes já foram descritas. As CRFs já representam as linhagens de HIV-1 predominantes em diversas regiões do mundo. Estima-se que de 10 a 20% das novas linhagens virais caracterizadas sejam CRFs (Peeters, 2000; Thomson *et al.*, 2002b; Perrin *et al.*, 2003; <http://hiv-web.lanl.gov>).

A variabilidade nucleotídica entre os diferentes subtipos do HIV-1 é de até 25% na região do envelope, de 16% para *gag* e de até 13% para a região da polimerase. A divergência genética entre os subsubtipos é de 11-16% para o gene do envelope e de 7-12% para o gene *gag*. O efeito de tal variabilidade genética na transmissão e patogênese das linhagens virais ainda não é bem definido. Vários estudos epidemiológicos sugerem que os subtipos do grupo M possuam propriedades diferentes, podendo ter influência direta na constituição antigênica dos vírus. Conseqüentemente, é importante determinar se as linhagens do HIV-1 que circulam nos países onde serão realizados testes com vacinas estão genética e antigenicamente relacionadas ao protótipo da vacina em teste. Para se obter esse tipo de informação, é necessário manter programas de monitoramento

da variabilidade genética do HIV que avaliem, sistematicamente, a distribuição dos subtipos do HIV-1 e a introdução de novas variantes (Peeters, 2000; Arens, 2001; Gaschen *et al.*, 2002; Morgado *et al.*, 2002; Centlivre *et al.*, 2005).

Dados epidemiológicos e filogenéticos mostraram que as linhagens do HIV-1 do grupo M não estão distribuídas de maneira uniforme pelo mundo. As formas genéticas mais prevalentes em todo mundo são C, A (A1), B e CRF02\_AG. A maior diversidade de formas genéticas do HIV-1 está na África Central, com prevalência mais elevada dos subtipos A e C e da CRF02\_AG. Nessa região geográfica encontram-se todos os subtipos do grupo M, o que é consistente com o fato desse continente ser possivelmente o local de origem da epidemia da AIDS. O subtipo C possui, atualmente, uma prevalência de 50% em todo mundo, circulando na região sudeste da África, Índia, China e Etiópia. Recentemente, o subtipo C tem se disseminado para diversas regiões no mundo, ocorrendo, de forma minoritária, também no Brasil e na Rússia. O subtipo A é predominante na porção central e leste da África, enquanto o subtipo B é predominante em parte da Ásia, Caribe, América do Norte, Europa, América do Sul e Oceania. O subtipo D é encontrado, principalmente no leste da África. O subtipo F (F1) predomina na Romênia e também possui uma pequena incidência no Brasil. O subtipo G circula na região ocidental e central da África e em Portugal (Peeters, 2000; Thomson *et al.*, 2002b; Centlivre *et al.*, 2005). A figura 9 mostra a distribuição geográfica mundial dos subtipos e das formas recombinantes circulantes do HIV-1, modificada a partir de Thomson *et al.*, 2002b.



**Figura 9.** Distribuição geográfica dos subtipos predominantes do grupo M e suas formas recombinantes circulantes (CRFs) (adaptado de Thomson *et al.*, 2002b).

Os países da América do Sul apresentam prevalência mais elevada do subtipo B, seguido pela presença dos subtipos F e C. Há ainda uma pequena prevalência do subtipo A no Chile, Peru e Guiana Francesa e da CRF\_01AE no Uruguai. Existem também relatos da ocorrência de formas recombinantes B/F e de CRF\_012 na Argentina, Uruguai e Venezuela (Morgado *et al.*, 2002; Peeters, 2000).

O Brasil, por ser um país de grandes dimensões, apresenta um padrão de distribuição de subtipos divergente entre as várias regiões geográficas. Estudos sobre a epidemiologia molecular do HIV-1 no Brasil mostraram a presença de múltiplos subtipos do grupo M, com predominância do subtipo B de cerca de 90% e uma prevalência menor dos subtipos F1, C, D e A. Esses dados refletem especialmente a distribuição de subtipos nas regiões Sul e Sudeste do Brasil, com algumas informações preliminares sobre as regiões Norte, Nordeste e Centro-Oeste. O subtipo F1 é o variante não-B mais prevalente e representa aproximadamente 18% dos isolados do HIV-1 no Brasil (Morgado *et al.*, 2002). Na região Norte, foi descrita a prevalência de 38,7% do subtipo B nas amostras analisadas, seguida por uma alta porcentagem de recombinantes B/F (35,5%) e pelo subtipo F (16%) (Vicente *et al.*, 2000). No Centro-Oeste, o subtipo

B (89,5%) é o mais prevalente, seguido por uma pequena porcentagem de recombinantes B/F (10,5%) (Cerqueira *et al.*, 2004a). Na região Nordeste, o subtipo B também é o mais prevalente seguido pelo subtipo F, e pela forma recombinante B/F (Couto-Fernandez *et al.*, 1999; Morgado *et al.*, 2002; Gadelha *et al.*, 2003). Além disso, uma prevalência crescente de infecções pelo subtipo C tem sido observada na maioria dos estados do Sul e do Sudeste do Brasil. Formas recombinantes circulantes B/F, B/D e B/C também já foram descritas, além de infecções com dois subtipos. A Figura 10 apresenta dados de 2002 para a distribuição dos subtipos e das formas recombinantes circulantes do grupo M do HIV-1 nas diferentes regiões do Brasil (Morgado *et al.*, 2002; Tanuri *et al.*, 2002; Castro *et al.*, 2003; Soares *et al.*, 2003).



**Figura 10.** Distribuição dos subtipos e das formas recombinantes circulantes nas diferentes regiões geográficas do Brasil (adaptado de Morgado *et al.*, 2002).

Além dos subtipos já descritos foi identificado ainda um variante do subtipo B, denominado B'', que difere do subtipo B tradicional pela presença do motivo GWGR no topo do *loop* V3 da proteína do envelope gp120, em substituição à assinatura GPGR, presente nos isolados do subtipo B americanos e europeus. Assim, os isolados do subtipo B no Brasil podem ser divididos em dois grupos distintos, um americano - europeu (GPGR), e o outro típico do Brasil, que possui o motivo GWGR, B''. Apesar da alta prevalência de B'' no Brasil, o variante do HIV-1 contendo o motivo GWGR no topo do *loop* V3 foi descrito pela primeira vez no Japão, e pode ser encontrado em alguns países da América Latina e na França. Esse variante do subtipo B apresenta propriedades genéticas e antigênicas distintas. A presença de isolados B'' deve ser considerada, já que sua prevalência é alta em algumas áreas do Brasil, tendo importância epidemiológica e relevância nos estudos de testes de vacinas nesse país (Covas *et al.*, 1998; Morgado *et al.*, 1998; Bongertz *et al.*, 2000; Vicente *et al.*, 2000; Morgado *et al.*, 2002).

### 1.9. Resposta imune ao HIV

A infecção pelo HIV induz uma resposta imunológica tanto humoral como celular, mas sem completa eliminação do vírus. A caracterização da resposta imune ao HIV é de grande importância considerando-se a sua relevância em estudos de triagem, já que anticorpos anti-HIV são marcadores diagnósticos da infecção. A produção de anticorpos anti-HIV pode ser detectada, em alguns casos, até poucos dias após a infecção, mas na maioria dos casos ocorre em um período que varia de 1 a 3 meses depois da exposição ao vírus. Os anticorpos são secretados no sangue e nas mucosas e são detectados em testes de diagnóstico sorológico. O desenvolvimento de vacinas contra o HIV também exige conhecimentos sobre a biologia do vírus e da resposta imune de seu hospedeiro (Flint *et al.*, 2004; Abbas & Lichtman, 2005).

A resposta imune celular inclui a ativação de linfócitos T citotóxicos (CTLs) e dos linfócitos T auxiliares. A grande maioria dos CTLs é de células TCD8<sup>+</sup>, que possuem papel importante na diminuição da replicação viral e no controle da viremia. Os CTLs são capazes de reconhecer todas as proteínas virais produzidas na célula

infectada pelo vírus. Eles respondem aos fragmentos de peptídeos que são expostos pelas moléculas de MHC classe I na superfície das células infectadas. No entanto, a resposta das células TCD8<sup>+</sup> tem mostrado ser dependente das células TCD4<sup>+</sup>, que são extremamente afetadas durante a infecção pelo HIV, tornando a resposta imune celular menos eficiente (Berzofsky *et al.*, 2004; Flint *et al.*, 2004).

Anticorpos neutralizantes possuem papel importante no controle da replicação viral durante a fase assintomática e inicial da infecção. No entanto, geralmente, tais anticorpos encontram-se em títulos baixos, o que favorece a seleção de linhagens virais resistentes. Anticorpos neutralizantes se ligam a sítios específicos das proteínas do envelope gp120 e gp41, podendo impedir a entrada do vírus na célula hospedeira. Entretanto, o HIV apresenta mecanismos adaptativos que dificultam a atividade do sistema imune, como a presença de muitos sítios de glicosilação em Env, dificultando a exposição de motivos antigênicos e a grande variabilidade dos *loops* de gp120, que compromete a ação de anticorpos de reação cruzada (Burton *et al.*, 2004).

No curso normal de uma infecção, o indivíduo apresenta ativação da resposta imune inata, uma primeira barreira ao estabelecimento da infecção, seguida por uma resposta imune adaptativa, que elimina a infecção e estabelece uma imunidade protetora. No entanto, nem sempre isso acontece. Os vertebrados desenvolveram muitas defesas estratégicas contra patógenos. Esses também desenvolveram uma série de mecanismos para escapar dessas defesas. O HIV possui várias estratégias para escapar da resposta imunológica, impedindo não só o reconhecimento das células infectadas pelos linfócitos T citotóxicos, mas também, a produção de anticorpos neutralizantes (Janeway *et al.*, 2002).

A diminuição da resposta imune celular é em grande parte provocada pela perda dos linfócitos T CD4<sup>+</sup>. A infecção pelo HIV leva a destruição das células T CD4<sup>+</sup> de diversas formas. Uma delas é provocada pela própria infecção viral, que resulta na lise dos linfócitos infectados, além de aumentar a suscetibilidade destes à indução de apoptose. Além disso, há ainda a morte dessas células que são reconhecidas pelos linfócitos T CD8<sup>+</sup> (Janeway *et al.*, 2002).

A alta taxa de mutação do genoma do HIV que codifica epitopos capazes de induzir uma resposta imune, contribui para a progressão da infecção e também para a ineficácia das estratégias de imunização anti-HIV (Gzyl *et al.*, 2004).

Outros mecanismos envolvem várias características peculiares das proteínas do envelope, que são um dos principais alvos da ação do sistema imune. Essas proteínas, além de apresentarem alta variabilidade, são muito glicosiladas. Os açúcares diminuem consideravelmente o potencial imunogênico dos epitopos expostos. Outro fator a ser considerado é a organização conformacional de gp120 e gp41 na forma de trímeros que pode tornar vários epitopos, antes expostos nas formas monoméricas dessas proteínas, inacessíveis à ação do sistema imune. Além disso, mudanças conformacionais ocorridas nas duas proteínas, durante o processo de entrada do vírus na célula provocam a exposição transitória de sítios vulneráveis à ação de anticorpos, impondo, assim, barreiras cinéticas e espaciais a ação dos mesmos (Berzofsky, *et al.*, 2004; Burton *et al.*, 2004; Gzyl *et al.*, 2004).

### **1.9.1 Desenvolvimento de vacinas anti-HIV**

As vacinas são uma eficiente estratégia no controle de doenças infecciosas. Durante o século 20, várias vacinas contra enfermidades provocadas por diferentes vírus e bactérias foram desenvolvidas com sucesso e permitiram a erradicação de doenças extremamente contagiosas como a varíola e a poliomielite em alguns países. Apesar do HIV-1 ter sido identificado como o agente etiológico da AIDS há mais de duas décadas, ainda não foi possível a produção de uma vacina contra esse vírus. O desenvolvimento de uma vacina contra o HIV tem enfrentado muitas dificuldades e parece exigir o entendimento de mecanismos complexos da replicação viral e da resposta imune do hospedeiro.

A maioria das vacinas bem sucedidas foi desenvolvida contra infecções naturais agudas e que produzem uma imunização duradoura se o paciente sobreviver à infecção inicial. A infecção pelo HIV é crônica e, na maioria dos casos, a resposta imune natural do indivíduo não é suficiente para conter a infecção. Assim, dentro desse contexto, conhecimentos fundamentais em biologia molecular, virologia e imunobiologia serão essenciais no desenvolvimento de vacinas contra o HIV, que gerem uma resposta imune mais eficiente do que a natural e que protejam contra as infecções crônicas (Berzofsky *et al.*, 2004).

Considerando o alto custo da terapia antiviral e o contínuo aumento de pessoas infectadas pelo vírus, principalmente em países em desenvolvimento, há grande interesse no desenvolvimento tanto de vacinas preventivas, que impeçam a instalação da



infecção do vírus nas células hospedeiras, como de vacinas terapêuticas, que auxiliem no combate à infecção de indivíduos já infectados pelo HIV.

Embora grandes esforços estejam sendo realizados no desenvolvimento de novas estratégias de vacinas contra o HIV, vários obstáculos vêm sendo enfrentados para a produção de uma vacina eficiente e segura contra esse vírus. Um dos problemas está no tipo de vacina que deve ser utilizada contra o HIV. A utilização de vírus vivos atenuados não é viável para uma vacina anti-HIV. Nesses modelos de imunização, se mimetiza a infecção viral. Entretanto, como a resposta imune natural contra o HIV não é capaz de erradicar a doença, essa estratégia além de não ser eficiente pode ser perigosa, pois tais vírus poderiam, eventualmente, provocar uma infecção e a AIDS. Vacinas que utilizam vírus mortos não têm apresentado bons resultados em modelos animais. Conseqüentemente, as estratégias mais utilizadas, atualmente, são baseadas em proteínas e peptídeos do HIV, ou fragmentos de seu DNA que é levado às células através de vetores (Berzofsky *et al.*, 2004; Gallo, 2005).

Um segundo obstáculo é a não existência de modelos de animais de pequeno porte para estudos da infecção do HIV. Embora não exista um modelo animal perfeito, os experimentos para o desenvolvimento de uma vacina contra o HIV são realizados com o vírus da imunodeficiência de símios (SIV), que é relacionado ao HIV e infecta macacos. O SIV causa uma doença similar a AIDS nesses animais. Apesar de ser um modelo viável, ele envolve a infecção de primatas não humanos, o que demanda grandes investimentos e está disponível para poucos pesquisadores (Janeway *et al.*, 2002; Gallo, 2005).

Outro problema no desenvolvimento de vacinas contra o HIV é a variabilidade que esse vírus apresenta e a sua elevada taxa de mutação em um mesmo indivíduo, além da diversidade dos subtipos e da freqüente recombinação entre eles, gerando as formas recombinantes (CRFs). Essa variabilidade facilita o escape do vírus à ação do sistema imune. Alguns dos epitopos de maior potencial neutralizante do HIV são muito variáveis, dificultando ainda mais o desenvolvimento de anticorpos de reação cruzada. No entanto, alguns anticorpos monoclonais, que são capazes de neutralizar várias linhagens virais primárias, já foram descritos (Berzofsky *et al.*, 2004).

Além de todos os obstáculos acima relatados, há outro problema importante a ser considerado. Sendo o HIV um retrovírus, ele é capaz de integrar o seu genoma no DNA da célula hospedeira, podendo persistir de forma latente, uma vez estabelecida a

infecção. Nesse contexto, a ação da resposta imune celular é de extrema importância para a remoção das células infectadas (Janeway *et al.*, 2002; Gallo, 2005).

Segundo dados do *International AIDS Vaccine Initiative* (IAVI), atualmente, cerca de 30 vacinas anti-HIV estão em fase de testes em 19 países, nos seis continentes, incluindo no Brasil. No entanto, somente um estudo em larga escala, fase III, foi concluído. Nesse estudo, completado em 2004 na Tailândia e nos Estados Unidos, foram testadas a capacidade de produção de anticorpos neutralizantes e de uma resposta T citotóxica eficiente a partir de subunidades de gp120 dos subtipos B e E. No entanto, ele não obteve sucesso em demonstrar a eficácia dessa vacina (Berzofsky *et al.*, 2004; <http://www.iavi.org/>).

A maioria das vacinas em desenvolvimento estimula uma resposta imune celular. Porém, atualmente, há consenso na comunidade científica que a resposta imune induzida por uma vacina anti-HIV deve estimular a produção de anticorpos neutralizantes de ação cruzada e, ao mesmo tempo, induzir uma resposta celular eficiente e duradoura (Berzofsky *et al.*, 2004; Bongertz *et al.*, 2005; <http://www.iavi.org/>).

### **1.9.2 Anticorpos neutralizantes**

Considerando que este trabalho tem como objetivo o estudo de epitopos do *loop* V3 de gp120, abordaremos os principais anticorpos neutralizantes relacionados a esta proteína. A recente caracterização de anticorpos monoclonais de reação cruzada, b12, 2G12 e 447-52D, que são capazes de neutralizar diferentes isolados primários de HIV-1, trouxe esperança para a produção de vacinas anti-HIV-1. A existência desses anticorpos evidencia que há regiões em gp120 suscetíveis ao desenvolvimento de uma imunidade humoral. As interações entre os anticorpos e as proteínas do envelope devem ser bem descritas, para que se possa avançar no desenho de novas vacinas (Burton *et al.*, 2004).

Esses anticorpos monoclonais ligam-se a regiões distintas de gp120: o domínio de ligação ao receptor CD4, ao sítio de ligação ao receptor de quimiocinas (o co-receptor), a regiões do *loop* V3, que apesar de toda a variabilidade genética possui estruturas conformacionais conservadas e muito imunogênicas (Berzofsky *et al.*, 2004; Burton *et al.*, 2004).

## ➤ b12

O primeiro desses anticorpos a ser identificado foi o b12. Esse anticorpo reconhece epitopos sobrepostos ao sítio de ligação ao CD4, impossibilitando a interação gp120-CD4. A interação do vírus com essa molécula tem como sítio ativo o aminoácido Phe-43 em CD4, que é reconhecido por b12. A obstrução desse sítio torna a interação da proteína viral com o receptor muito fraca, impedindo o processo de fusão do vírus com a célula. Muitos outros anticorpos com a mesma especificidade de interação ao CD4 já foram identificados. No entanto, ao contrário de b12, eles se ligam a subunidades monoméricas de gp120, mas não possuem a habilidade de neutralizar isolados de HIV-1. Ainda é preciso se descobrir como se produzir outros anticorpos com a mesma especificidade de b12 (Wyatt & Sodroski, 1998; Saphire *et al.*, 2002; Burton *et al.*, 2004).

## ➤ 2G12

2G12 reconhece um conjunto de resíduos de oligomanose ao redor das regiões V3/C3/V4 de gp120. Uma vez ligado à partícula viral ele atua impedindo, de maneira não específica, a entrada do vírus na célula. Os resíduos reconhecidos por esse anticorpo fazem parte do revestimento de açúcar presente em gp120, que dificulta a ação do sistema imune. A relativa conservação desses açúcares permite que esse anticorpo reconheça várias linhagens de HIV-1 (Burton *et al.*, 2004; Kang *et al.*, 2005).

## ➤ 447-52D

O anticorpo 447-52D foi produzido a partir de heterohibridomas derivados de células periféricas mononucleadas do sangue de um indivíduo infectado com linhagens do subtipo B, reconhece um epitopo no topo do *loop* V3. É o anticorpo anti-V3 que possui maior capacidade de reação cruzada. Ele reconhece resíduos que estão sobrepostos a região central do topo do *loop* V3, os resíduos GPxR (resíduos 312-315, em HXB2), que são altamente conservados entre as linhagens do subtipo B. Ao contrário da maioria dos anticorpos anti-V3, 447-52D é capaz de neutralizar isolados virais que utilizam como co-receptor tanto CXCR4 como CCR5, sendo capaz de neutralizar linhagens de V3 com grande variabilidade em suas seqüências. Estudos demonstraram que esse anticorpo foi capaz de neutralizar praticamente todas as linhagens que possuíam o epitopo GPGR, e cerca de 50% dos isolados do subtipo

B analisados. O anticorpo parece não reconhecer o motivo GPGQ no topo do *loop* V3, presente na maioria dos isolados não pertencentes ao subtipo B (Stanfield *et al.*, 2004; Zolla-Pazner *et al.*, 2004).

## 2. OBJETIVOS

Considerando a pequena quantidade de informações referentes aos isolados do HIV-1 que circulam no Distrito Federal, neste trabalho pretendemos:

- ✓ Realizar a caracterização dos subtipos das amostras analisadas com base na região C2-C3 de gp120;

- ✓ Verificar a concordância do subtipo caracterizado com base em C2-C3 de gp120 com a definição de subtipo feita a partir de outras regiões do genoma, a saber: *gag*, *nef*, PR e RT;

- ✓ Caracterizar, nas amostras do Distrito Federal, os domínios antigênicos já descritos na literatura para a região do *loop* V3 de gp120 capazes de induzir anticorpos neutralizantes;

- ✓ Analisar a distribuição de aminoácidos carregados positivamente nas posições 11 e 25 do *loop* V3 de gp120, que possam sugerir o tipo fenótipo produzido pelos vírus nas amostras do Distrito Federal;

- ✓ Realizar a análise filogenética das seqüências obtidas.

### **3. METODOLOGIA**

#### **3.1. População Estudada**

A população alvo foi composta de indivíduos HIV-1 positivos, maiores de 18 anos, de ambos os sexos e que apresentavam carga viral para o RNA do HIV-1 superior a 50.000 cópias/ml. As amostras analisadas são provenientes do ano de 2002 e foram obtidas a partir do material biológico (RNA) estocado no Laboratório Central de Saúde Pública do Distrito Federal (LACEN-DF). Este estudo é anônimo e tem a aprovação do comitê de ética. Todas as amostras foram previamente codificadas, não havendo qualquer identificação dos pacientes por nome, sexo, idade ou qualquer outro dado pessoal.

A representatividade da população foi calculada para um nível de confiança de 99%, utilizando-se o programa Estatcalc Epi-6 (CDC), retirado do endereço eletrônico: <http://www.cdc.gov/epiinfo/>. Para tanto, foi considerada a população de indivíduos infectados pelo HIV no DF (n = 4024).

#### **3.2. Coleta das Amostras**

Para a realização do exame de carga viral do HIV-1 no serviço público de assistência à saúde no Distrito Federal é coletada uma amostra de sangue do indivíduo que será submetido ao teste. Essa amostra é processada para a obtenção do plasma no LACEN-DF. O RNA viral é então extraído, de acordo com as instruções do kit para carga viral, e mantido em um volume final de 50 µl. Deste volume de RNA obtido, 5 µl são usados na reação de carga viral e os 45 µl restantes são armazenados em freezer a -80° C.

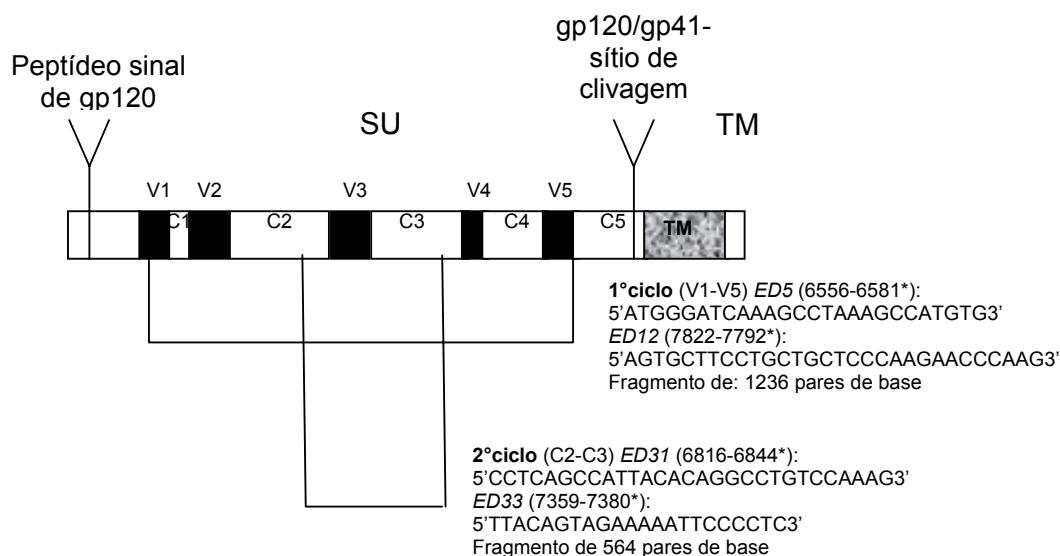
Nos experimentos realizados foram coletadas 60 alíquotas de 45 µl de RNA que foram encaminhadas ao nosso laboratório e armazenadas a -80°C para a realização dos procedimentos descritos a seguir. Os isolados analisados neste estudo foram denominados por DF 02.n, de acordo com o ano de coleta da amostra e o número dado ao isolado.

### 3.3. Transcrição Reversa

A reação de transcrição reversa foi realizada em um volume final de 20 $\mu$ L, a partir das amostras de RNA anteriormente mencionadas. Inicialmente, uma solução contendo 5 $\mu$ L de RNA viral, 0,1 $\mu$ g de iniciadores reandômicos e 6,2 $\mu$ L de água Milli Q foi submetida a 70°C por 5 minutos. Adicionou-se, em seguida, 2,5x de tampão *First Strand* (*Invitrogen*), 0,5mM de dNTP, 0,02M de DTT (*Invitrogen*), 40 unidades (U) de inibidor de RNase (*Invitrogen*) e 200 unidades (U) da enzima transcriptase reversa Super Script III (*Invitrogen*). Esta reação foi incubada a 25°C por 5 minutos e, em seguida, a 50°C por 60 minutos, sendo inativada por aquecimento a 70°C por 15 minutos. O cDNA sintetizado foi armazenado a -20°C até o momento do uso.

### 3.4. Reação em Cadeia da Polimerase

O cDNA do HIV-1 foi amplificado por meio de *nested PCR*, utilizando-se um conjunto de iniciadores específico para região V1-V5 de gp120. No primeiro ciclo foram usados os iniciadores externos ED5 (senso)/ED12 (reverso) (*Integrated DNA Technologies*) e no segundo ciclo os iniciadores internos ED31 (senso)/ED33 (reverso) (*Integrated DNA Technologies*) (Rodrigo & Mullins, 2001) (Figura 11).



**Figura 11.** Sequências, localização dos iniciadores para a amplificação da região V1-V5 de gp120.\* As indicações das posições dos iniciadores de gp120 estão baseadas no genoma do HIV-1 HXB2.

A PCR foi realizada num volume final de 50µL, contendo 5µL de cDNA, 0,25mM de dNTPs (*Invitrogen*), 0,8µM de *primers* ED5/ED12 (*Integrated DNA Technologies*) (Rodrigo & Mullins, 2001) 2U da enzima *Taq* polimerase (*Invitrogen*), 10x de tampão da enzima (*Invitrogen*) e 2,5mM de cloreto de magnésio (*Invitrogen*). As amostras foram amplificadas em um termociclador (PTC-100 Research INC) usando-se os seguintes ciclos térmicos: 94°C 1min; 58°C 1min; 72°C 1min; 2X 94°C 1min, 58°C 1min, 72°C 1 min; 31x 94C 15seg, 49°C 45seg, 72°C 5 min. Três microlitros do produto de PCR gerado no 1º ciclo foram submetidos a uma segunda amplificação com os *primers* ED31/ED33 (*Integrated DNA Technologies*) (Rodrigo & Mullins, 2001), nas mesmas condições e usando-se os mesmos ciclos de amplificação (Figura 11).

As amostras amplificadas foram submetidas à eletroforese em gel de agarose a 1%, posteriormente corado com brometo de etídeo. Os fragmentos de DNA esperados deveriam apresentar um tamanho em torno de 564 pares de bases (pb), correspondente à região C2-C3 de gp120. Controles negativos foram incluídos durante os procedimentos de extração e amplificação, para monitorar possíveis contaminações.

Considerando-se a importância da caracterização genética dos subtipos do HIV-1 circulantes no Distrito Federal, dados baseados em outras regiões do genoma do vírus foram gerados pelo nosso grupo de pesquisa (dados gerados por Vêras, Nazle Colaço e Ramalho, Eduardo). Dessa maneira, no presente trabalho, apresentaremos também, para fins de comparação, os resultados dos subtipos que foram caracterizados com base nas regiões: *gag*, *nef*, PR e RT, dados obtidos por nosso grupo de pesquisa.

### 3.5. Purificação e Precipitação dos produtos de PCR

Os produtos de PCR que apresentaram bandas inespecíficas quando visualizados em gel de agarose foram purificados utilizando-se o *Kit GFX™ PCR DNA and gel purification (Amersham)*, de acordo com o protocolo sugerido pelo fabricante, e ressuspendidos em 30µL de água Mili Q.

Os demais produtos foram precipitados pelo tratamento com acetato de amônio e etanol 100%. Para isso, foram adicionados para cada volume de PCR meio volume de acetato de amônio 7,5 M e 2,5 volumes de etanol 100%. A reação foi mantida a -80°C por 30 minutos e submetida à centrifugação a 13.000 rpm por 30 minutos, a 4°C. Após o descarte do sobrenadante, o precipitado foi lavado com 100µL de etanol 70% e centrifugado a 5.000 rpm por cinco minutos, à temperatura ambiente. O sobrenadante



foi, mais uma vez, descartado e o precipitado foi mantido a 37°C por uma hora e, em seguida, ressuspendido em 15µL de água Mili Q.

### 3.6. Genotipagem

#### 3.6.1. Seqüenciamento automático

Os produtos de PCR foram seqüenciados automaticamente, em ambas as direções, utilizando-se os iniciadores internos ED31 (senso)/ED33 (reverso) (Rodrigo & Mullins, 2001), para o seqüenciamento do segmento C2-C3 de gp120. As reações de seqüenciamento nucleotídico foram realizadas no laboratório de Biologia Molecular da Universidade de Brasília, utilizando-se o kit “*Taq Dye-terminator*”, no Sistema *Megabace Molecular Dynamics (Amersham, Pharmacia)* e na plataforma de seqüenciamento da EMBRAPA (Cenargen) utilizando-se o kit *ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit* versão 3.1 (*Applied Biosystems*).

#### 3.6.2. Análise das seqüências

A qualidade das seqüências obtidas nas reações de seqüenciamento automático foi verificada pelo programa PHRED (Ewung *et al.*, 1998). As homologias das seqüências nucleotídicas foram determinadas pelo programa BLAST (*Basic Local Alignment Search Tool*) (Altschul *et al.*, 1990), em que as mesmas são alinhadas com seqüências do HIV-1 disponíveis no banco genômico dos Los Alamos (<http://hiv-web.lanl.gov/>). O alinhamento das seqüências em ambas as direções foi realizado pelo programa CLUSTAL W (Thompson *et al.*, 1994) e otimizado por inspeção visual. Após a correção das divergências, prosseguiu-se com a tradução e alinhamento das seqüências obtidas com aquelas seqüências de referência, também retiradas do banco de genomas do *Los Alamos* (<http://hiv-web.lanl.gov/>).

A determinação dos subtipos foi feita pela análise de homologia entre as seqüências dos isolados avaliados neste estudo e as seqüências consenso do grupo M disponíveis no banco de genomas do *Los Alamos* (<http://hiv-web.lanl.gov/>). O programa *REGA HIV Subtyping Tool* (de Oliveira *et al.*, 2005) foi utilizado para confirmar a caracterização dos subtipos. Esse programa utiliza a análise filogenética associada ao método de *bootscanning* para a determinação de subtipos e de formas genéticas recombinantes de seqüências genômicas completas e parciais do HIV-1.

As seqüências que poderiam apresentar pontos de recombinação foram analisadas pelo método de *bootscanning*, utilizando-se o programa *SimPlot* (Lole *et al.*, 1999). Esse programa utiliza o alinhamento da seqüência analisada, que provavelmente possui pontos de recombinação, com seqüências referências dos subtipos do grupo M do HIV-1. Os resultados gerados refletem a similaridade da seqüência analisada com uma das seqüências referências utilizadas. As análises são realizadas a partir de pequenos intervalos (janelas) das seqüências e os resultados são plotados em um gráfico, possibilitando o reconhecimento de pontos de recombinação. Para a determinação de um subtipo foi utilizado um valor de *bootstrap* de 70%, segundo Hillis & Bull, 1993. Para a realização do teste de *bootscanning* foram utilizados os seguintes parâmetros: janela = 300 pb, passo = 20pb, GapStrip = on, número de repetições = 1000, Ts/Tv = 2, modelo de distância dois parâmetros de Kimura e modelo de árvore *neighbor joining* do programa *SimPlot*, versão 2,5 (Ray, 1999).

Para analisar a variabilidade genética da região do *loop* V3 dos isolados do HIV-1 avaliados neste estudo, foi realizada a dedução de suas seqüências de aminoácidos a partir do alinhamento das suas seqüências de nucleotídeos obtidas com os iniciadores senso (ED31) e reverso (ED33). As seqüências deduzidas dos isolados do subtipo B foram alinhadas com a seqüência referência do subtipo B, HXB2, e aquelas pertencentes aos isolados do subtipo F foram alinhadas com a seqüência referência F1 BE93, para tais alinhamentos utilizou-se o programa CLUSTAL W (Thompson *et al.*, 1994).

### 3.6.3 Análise Filogenética

A análise filogenética foi realizada utilizando-se o conjunto de programas PHYLIP (Felsenstein, 1993). As seqüências de nucleotídeos de cada isolado analisado neste estudo (designadas por DF 02.n, de acordo com o ano de coleta e o número do isolado), foram alinhadas com seqüências referência de vários subtipos do grupo M do HIV-1, retiradas do sitio *Los Alamos* (<http://hiv-web.lanl.gov>), utilizando o programa CLUSTAL W (Thompson *et al.*, 1994). Foi utilizado um fragmento de 330 pares de base que era comum a todas as seqüências analisadas. Seqüências de isolados brasileiros da região C2/C3 de gp120 também foram adicionadas ao conjunto de dados analisado (Bongertz *et al.*, 2000).

A distância evolutiva foi calculada utilizando-se o programa *DNAdist*, pelo método de correção de dois parâmetros de Kimura (Kimura, 1980). As relações filogenéticas foram determinadas pelo método de *Maximum likelihood* (máxima verossimilhança). A reprodutibilidade de cada ramo foi estimada por um valor de *bootstrap* de 1000 réplicas. O programa *Consensus Tree* foi utilizado para gerar uma árvore consenso. A árvore foi desenhada pelo programa *Tree-View*. A seqüência consenso do grupo O foi utilizada como grupo externo. Um valor de *bootstrap* de 70% foi considerado confiável para determinação de um determinado subtipo (Hillis & Bull, 1993).

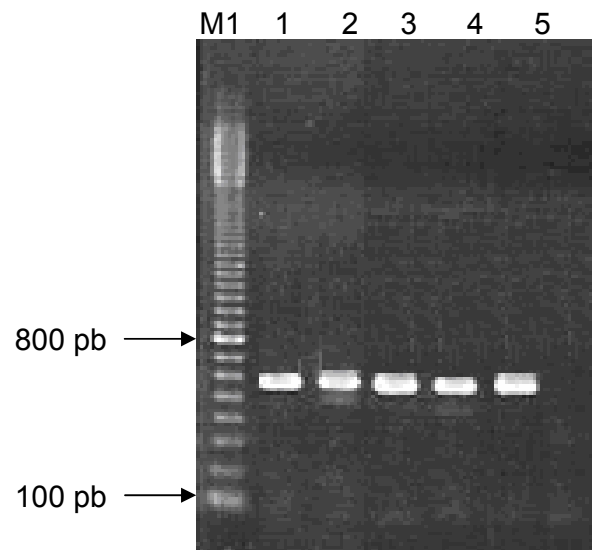
## 4. RESULTADOS

### 4.1. Amostragem

Neste estudo foram analisadas 60 amostras de RNA. O cálculo da representatividade, realizado pelo programa Epi-6 (CDC), mostrou que essa amostragem é representativa da população de indivíduos infectados pelo HIV no Distrito Federal, para um nível de confiança de 99%.

### 4.2. Amplificação da região C2-C3 de *env*

Todas as 60 amostras coletadas foram transcritas reversamente e foram testadas para amplificação da região C2-C3 de gp120, conforme descrito na metodologia. Todas as amostras foram amplificadas para a região C2-C3 de gp120, obtendo-se um fragmento de DNA de cerca de 560 pb (Figura 12).



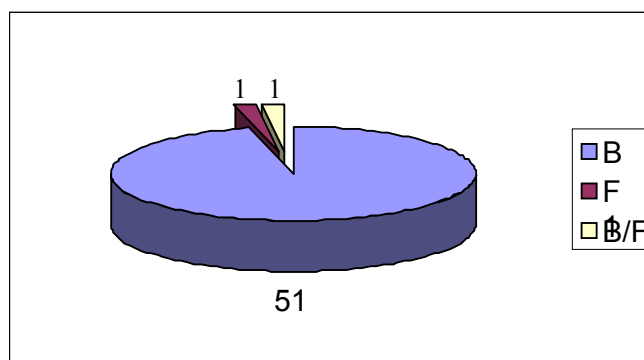
**Figura 12.** Análise eletroforética dos produtos de PCR da região de C2/C3 de gp120, em gel de agarose 1% corado com brometo de etídeo. M1 indica o marcador de peso molecular de 100pb (Promega); os números 1 a 5 indicam amostras de isolados do HIV-1; CN é o controle negativo.

### 4.3. Definição dos subtipos do HIV-1.

A caracterização das formas genômicas do HIV-1 foi feita através da análise das seqüências geradas para os produtos de PCR. As amostras foram seqüenciadas em

ambos os sentidos para a obtenção da região equivalente a C2-C3 de gp120 (*primers* internos ED31/ED33). As seqüências obtidas foram submetidas a análises por homologia pelo BLAST (*Basic Local Alignment Search Tool*) e pelo *REGA HIV Subtyping Tool* para a determinação de seus subtipos (Altschul *et al.*, 1990; de Oliveira *et al.*, 2005).

Das 60 amostras amplificadas, 53 (83%) apresentaram seqüências de tamanho suficiente para serem caracterizadas com base na região C2-C3 de gp120. As sete amostras restantes não apresentaram seqüências de qualidade e tamanho suficientes para a realização da caracterização, e foram excluídas deste estudo. As seqüências de nucleotídeos das amostras caracterizadas estão dispostas no Anexo, página 84. A determinação dos subtipos revelou a prevalência do subtipo B, presente em 51 (96,2%) das 53 amostras, seguida pela presença de um isolado F1 (1,9%) e um isolado B/F (1,9%). A Figura 13 representa a distribuição das formas genéticas encontradas pela análise das seqüências da região do envelope.



**Figura 13.** Distribuição das formas genéticas do HIV-1 encontradas pela análise da região C2-C3 de gp120. A cor azul representa as amostras do subtipo B, a vermelha a do subtipo F1 e a amarela a forma genética B/F. A determinação dos subtipos de cada amostra foi baseada nas análises de homologia, realizadas utilizando-se o programa REGA (de Oliveira *et al.*, 2005), e de filogenia. Os eventos de recombinação foram confirmados por *SimPlot* (Ray, 1999).

Os procedimentos mais comuns para a determinação dos subtipos do HIV-1 são baseados na análise de seqüências parciais de *env* e no ensaio de HMA (*Heteroduplex Mobility Analysis*), em que também se analisam fragmentos do gene do envelope

(Rodrigo & Mullins, 2001; Gadelha *et al.*, 2003). No entanto, a descrição de novas formas genéticas e da existência de formas recombinantes circulantes revelaram que tal metodologia não era tão eficiente, surgindo a necessidade da caracterização de outras regiões genômicas (Thomson *et al.*, 2002b).

Assim, os resultados obtidos foram complementados com os encontrados pela caracterização de subtipos baseada em outras regiões genômicas: *gag*, PR, RT e *nef*. A Tabela 3 mostra os subtipos encontrados pela análise da região C2-C3 de gp120 e aqueles obtidos a partir da caracterização de outras regiões do genoma do HIV-1, dados que foram gerados pelo nosso grupo de pesquisa.

**Tabela 3.** Definição dos subtipos realizada a partir das análises das seguintes regiões genômicas: *env gag*, RT, PR e *nef*. A determinação do subtipo foi baseada em análises de homologia e filogenia. As formas genéticas discordantes aparecem destacadas em vermelho e verde.

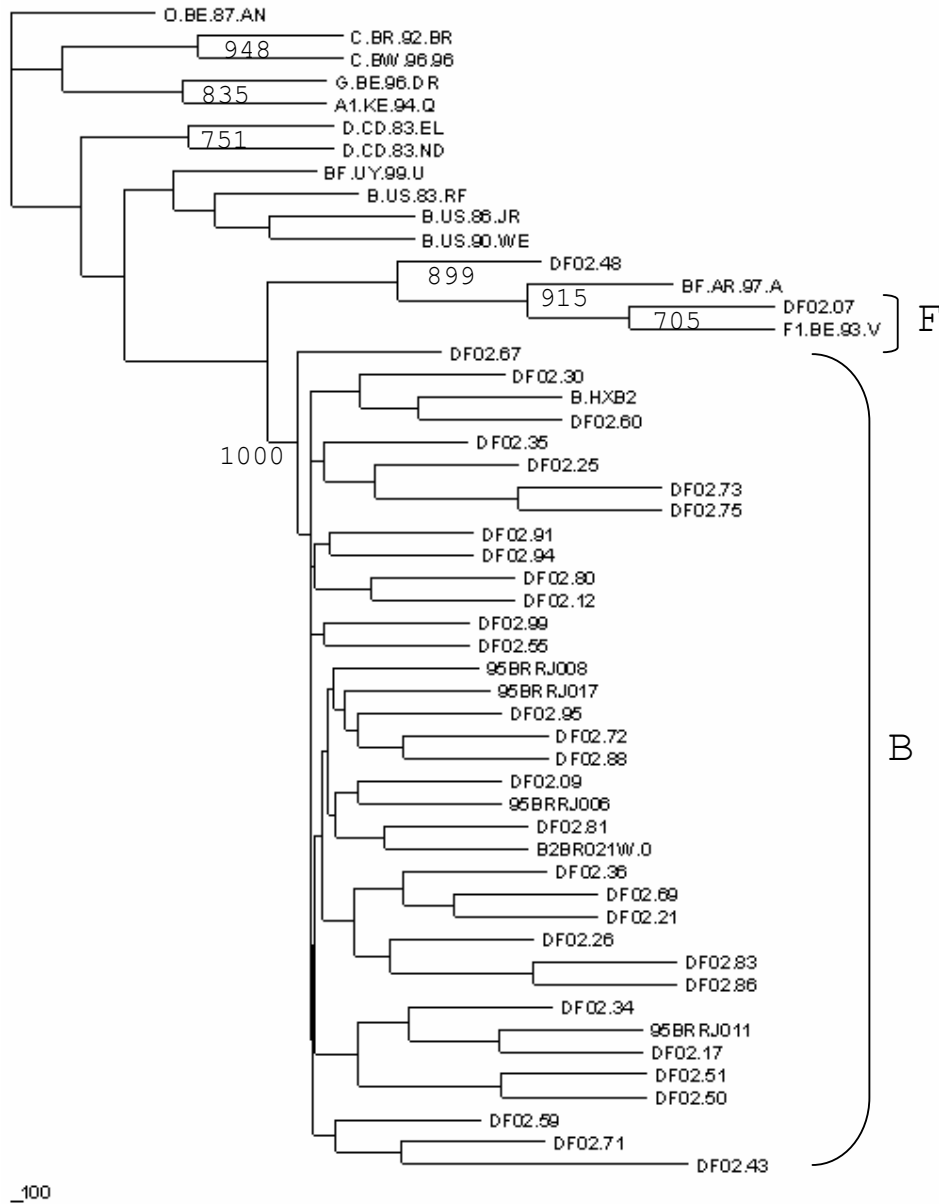
Amostras	Subtipo <i>env</i> (6816-7380)	Subtipo <i>gag</i> (837-1690)	Subtipo PR (2253-2549)	Subtipo RT (2604-3250)	Subtipo <i>Nef</i> (8274-9025)
DF02.02	B	B	B	B	-
DF02.03	B	B	B	B	B
DF02.05	B	B	B	B	-
DF02.06	B	B	B	B	-
DF02.07	<b>F1</b>	-	B	B	-
DF02.09	B	B	B	B	B
DF02.12	B	B	B	B	-
DF02.14	B	B	B	B	-
DF02.15	B	-	B	B	-
DF02.17	B	<b>B/F</b>	B	<b>F1</b>	B
DF02.20	B	-	B	B	B
DF02.21	B	B	B	B	B
DF02.23	B	B	B	B	B
DF02.25	B	<b>B/F</b>	B	B	-
DF02.26	B	B	B	B	-
DF02.30	B	B	B	B	B
DF02.32	B	B	<b>B/F</b>	B	B
DF02.33	B	B	B	B	B
DF02.34	B	B	B	B	B
DF02.35	B	B	B	B	B
DF02.36	B	B	B	B	-
DF02.37	B	B	B	B	-
DF02.38	B	<b>B/F</b>	<b>B/F</b>	B	-
DF02.43	B	B	B	B	-
DF02.48	<b>B/F</b>	B	B	B	-
DF02.50	B	B	B	B	B
DF02.51	B	-	B	B	B

DF02.53	B	-	-	-	-
DF02.55	B	B	B	B	-
DF02.58	B	B	B	B	-
DF02.59	B	B	B	B	-
DF02.60	B	-	B	B	-
DF02.65	B	B	B	-	-
DF02.67	B	-	B	B	B
DF02.69	B	B	B	B/F	-
DF02.71	B	-	B	B	-
DF02.72	B	-	B	B	B
DF02.73	B	-	B	B	B
DF02.75	B	-	B	B	-
DF02.76	B	B	B	B	B
DF02.80	B	-	B	B	B
DF02.81	B	B	B	B	B
DF02.83	B	B/F	B/F	B	B
DF02.86	B	-	B	B	-
DF02.88	B	B	B	B	B
DF02.89	B	-	B	B	B
DF02.91	B	B	B	B	-
DF02.94	B	-	B	B	-
DF02.95	B	B	B	B	B
DF02.97	B		B	B	
DF02.98	B		B	B	
DF02.99	B	-	B	-	B
DF02.102	B	-	B	B	B

A comparação desses dados mostrou que a ocorrência de oito (15%) amostras discordantes, todas recombinantes B/F, com os seguintes perfis genéticos:  $env^{F1}/gag^B/PR^B/RT^B$ ,  $env^B/gag^{B/F}/PR^B/RT^{F1}/nef^B$ ,  $env^B/gag^{B/F}/PR^B/RT^B$ ,  $env^B/PR^{CRF12-BF}/RT^B/nef^B$ ,  $env^B/gag^{B/F}/PR^{B/F}/RT^B$ ,  $env^{B/F}/gag^B/PR^B/RT^B$ ,  $env^B/gag^B/PR^B/RT^{B/F}$ .

#### 4.4. Análise Filogenética

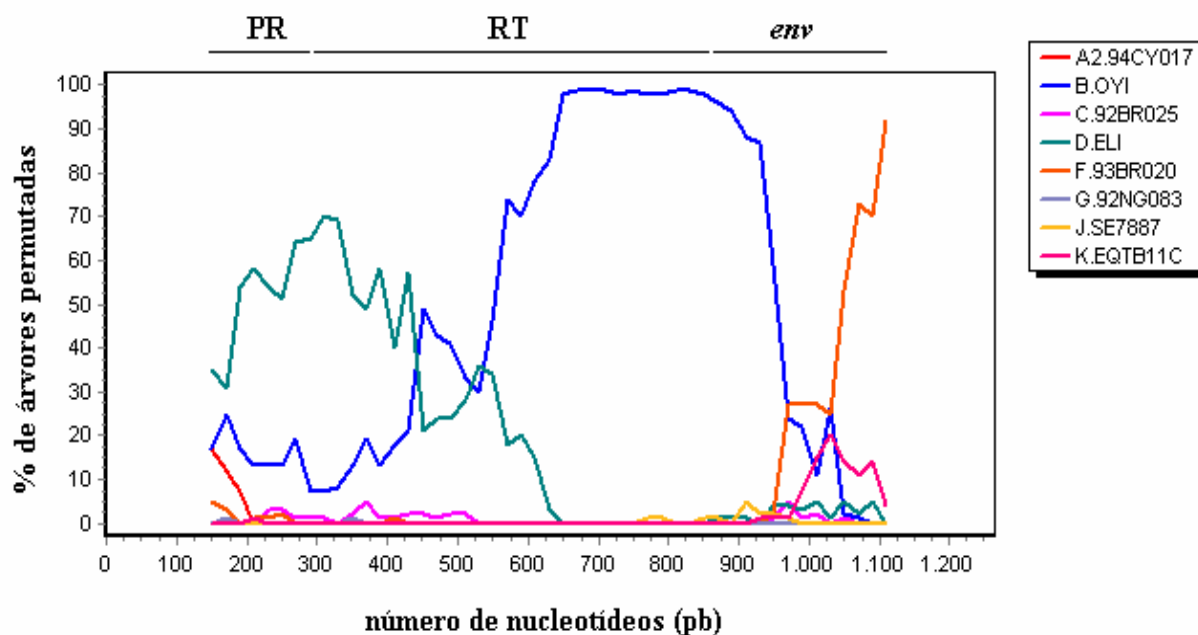
Para definir as relações filogenéticas entre os isolados analisados, foi utilizado o método de máxima verossimilhança através do conjunto de programas PHILIP (Felsenstein, 1993). A reprodutibilidade de cada ramo foi calculada por um valor de *bootstrap* de 1000 replicações. Um valor de *bootstrap* confiável foi estabelecido acima de 70% (Hillis & Bull, 1993). Desse modo, 50 das 53 amostras que tiveram suas formas genéticas caracterizadas com base na região C2/C3 de gp120 apresentaram seqüências de tamanho suficiente para serem submetidas à análise filogenética, 330 pb (Figura 14).



**Figura 14.** Relações filogenéticas pelo método de máxima verossimilhança entre os isolados do HIV-1 circulantes no Distrito Federal (DF 02.N) e seqüências de outros subtipos do HIV-1 do grupo M e de isolados brasileiros do subtipo B (B2BR021W, 95BRRJ006, 95BRRJ008, 95BRRJ009, 95BRRJ011). O grupo O foi utilizado como grupo externo. O método de dois parâmetros de Kimura foi utilizado para estimar a distância genética entre as seqüências. Valores de *bootstrap* significativos estão indicados ao lado de cada ramo (1000 replicações).



A análise filogenética confirmou os resultados encontrados a partir da caracterização das formas genéticas por homologia. Os isolados do subtipo B caracterizados neste estudo foram agrupados com outros isolados brasileiros do subtipo B e com a seqüência referência HXB2 com um valor de *bootstrap* de 1000 vezes. O isolado DF 02.07 está agrupado com o subtipo F1, com um valor de *bootstrap* de 705, confirmando tratar-se de uma amostra do subtipo F1. O isolado DF 02.48 forma um ramo a parte, e está agrupado com as seqüências referência B/F e F1, com valor de *bootstrap* de 899. Para confirmar a presença de um ponto de recombinação neste isolado, a seqüência foi submetida ao teste de *bootscanning*. Para a realização desse teste, as seqüências de outras regiões genômicas, RT e PR, foram adicionadas com a finalidade de aumentar a precisão da determinação dos pontos de recombinação (Figura 15).



**Figura 15.** Análise por *bootscanning* da seqüência de nucleotídeos das regiões RT, PR e C2/C3 de gp120 do isolado DF 02.48 B/F, comparando-a aos diferentes subtipos do HIV-1. O eixo horizontal representa o número de nucleotídeos analisados, o eixo vertical representa a porcentagem de árvores permutadas que suportam um determinado subtipo. Os seguintes parâmetros: janela = 300 pb, passo = 20pb, GapStrip = on, número de repetições = 1000, Ts/Tv = 2, modelo de distância dois parâmetros de Kimura e modelo de árvore *neighbor joining*.

A análise por *bootscanning* confirmou, por um valor de *bootstrap* acima de 70%, a presença de um ponto de recombinação B/F na região C2/C3 de gp120 do isolado DF 02.48.

#### **4.5. Caracterização da variabilidade antigênica da região do *loop* V3 nos isolados do HIV-1 do Distrito Federal**

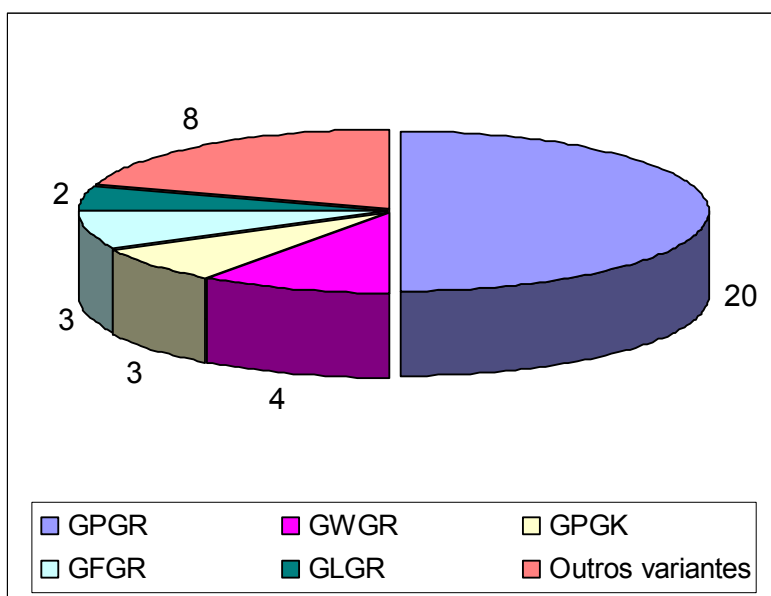
Para analisar a variabilidade genética da região do *loop* V3 dos isolados do HIV-1, foi realizada a dedução de suas seqüências de aminoácidos a partir do alinhamento das suas seqüências de nucleotídeos obtidas com os iniciadores senso (ED31) e reverso (ED33). Devido a problemas na qualidade das seqüências, foi possível a dedução de seqüências de aminoácidos de 41 amostras das 60 analisadas. As 19 amostras restantes não puderam ser caracterizadas.

Dentre os isolados que tiveram suas seqüências de aminoácidos caracterizadas, 40 eram do subtipo B e uma do subtipo F1, o isolado DF 02.07. O tretrapeptídeo GPGR localizado no topo do *loop* V3 foi avaliado para a caracterização da variabilidade intra-subtipo e antigênica desse motivo. A análise da região do *loop* V3, realizada a partir das seqüências de aminoácidos deduzidas, mostrou a grande variabilidade dessa região (Figura 16).

CONS-B	1	CTRPNNNTRKSIHI	<b>GPGR</b>	AFYTTGEIIGDIRQAHC 35
DF02.03		.....M	GPGR	T..A.-DE.....K.Y.
DF02.06		.....M	<b>GWGR</b>	.I.MP.G.....
DF02.07		.....P.	<b>GPGR</b>	...A..D.....
DF02.09		.....G..L	GFGR	TL.A..R.....
DF02.12		....K...R.PM	GLGR	.....S.....KK..
DF02.15		.....T...R.QR	GPGN	...V..G.T.-----
DF02.17		.....S.	<b>GPGR</b>	...A..S...N.....
DF02.21		.I.....G...	GPGA	S..A.....
DF02.25		.....L	GLGR	...A.....S.....
DF02.26		.I..G.T...R.RG	GPGY	...V...T...K...
DF02.30		.....T.	<b>GPGR</b>	...A.....
DF02.32		.....N.	<b>GPGR</b>	...A...M.....
DF02.33		.....N.	<b>GPGR</b>	...A..D.....
DF02.34		...G.....R...	<b>GPGR</b>	...ASEG.V.....
DF02.35		.....S.	<b>GPGR</b>	..FA..D.....
DF02.36		.I..S...R...L	GQGR	.W.....K.Y.
DF02.37		.....YV	<b>GPGR</b>	T.....-K..N.....
DF02.43		.....	GPGR	...A.-D.....
DF02.48		...G.....G...	<b>GPGR</b>	...A.-KT...K...
DF02.50		...S.....G...	<b>GPGR</b>	...A.EA.T.....
DF02.51		.....	<b>GPGR</b>	.....A.....
DF02.55		.I.....LP.	<b>GPGR</b>	.LFA..A.....
DF02.58		.I.....G...	<b>GPGR</b>	...A..D...N.....
DF02.59		.....I.G...	GPGS	...AR.Q.....R.Y.
DF02.60		.....Q.	<b>GPGR</b>	...AI.K.....
DF02.67		...G.....	<b>GPGR</b>	...A..D.....
DF02.69		.I..G...G.R.	<b>GPGR</b>	TI.A..R.....
DF02.71		.....R.G...	<b>GPGR</b>	.....T.....E..
DF02.72		.I.....R...M	<b>GFGR</b>	.L..N...V.....
DF02.73		.....G...	<b>GPGR</b>	T..A..D...N.....
DF02.75		.I.....R...M	<b>GPGR</b>	T..AE.D.....
DF02.80		.....M	GPGR	...A..D...N.....
DF02.81		.....R...M	<b>GWGR</b>	...A...V.....
DF02.83		...G.....PL	<b>GWGR</b>	VL.A..D.....
DF02.86		...G...I...M	ALGR	.L.A.ER.V.N..N..
DF02.88		.....	GFGR	.L.A.A.Q.....
DF02.91		.....N.	GPGR	...A..D.....
DF02.94		.....	APGG	...A.....
DF02.95		.....I.R..RM	<b>GWGR</b>	T..A..Q.K.S.....
DF02.99		.....PM	GPGH	..HA.D.....
DF02.102		.....R.N.	<b>GPGR</b>	..LA..A.....

**Figura 16.** Alinhamento das seqüências de aminoácidos da região do *loop* V3 obtidas com a seqüência consenso B. Em destaque a região do topo do *loop* V3, em laranja está representado o motivo GPGR, em verde está representado o motivo GWGR (B''). Aminoácidos idênticos são indicados por pontos, deleções são indicadas por traços. Em destaque o motivo GPGR.

Ao analisarmos as seqüências obtidas em relação à apresentação de motivos antigênicos e a sua variabilidade, notamos que o principal motivo de V3 (GPGR nas posições 312-315 de gp160 de HXB2) em 20 das 40, (50%), das amostras. O epitopo GWGR característico do subtipo B'' do Brasil foi encontrado em 4 das 40 amostras, (10%). Os motivos GFGR e GPGK foram encontrados em 3 amostras, (7,5%), cada um, e o motivo GLGR estava presente em 2 das 40 amostras (5%). Outras variações foram descritas em apenas uma amostra: GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH, apresentando uma freqüência de 20%. Os motivos GPGY (DF 02.26) e ALGR (DF 02.86), ainda não descritos na literatura, foram descritos neste estudo. O isolado DF 02.07, do subtipo F, apresentou o motivo GPGR, característico de outros isolados desse subtipo já descritos no Brasil (Banda *et al.*, 1995). A Figura 17 mostra a distribuição dos motivos encontrados no topo do *loop* V3 das amostras do subtipo B analisadas.



**Figura 17.** Distribuição dos variantes do subtipo B do HIV-1 encontrados pela análise das seqüências deduzidas de aminoácidos do *loop* V3 de gp120. A cor azul escuro representa os isolados contendo o motivo GPGR; a roxa os que apresentam o motivo GWGR; a amarela os que apresentam o motivo GPGK; a azul os que possuem o motivo GFGR; a verde os que apresentam GLGR; a laranja representa as outras variações desse motivo observadas (GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG e GPGH).

#### **4.6. Análise dos resíduos de aminoácidos localizados nas posições 11 e 25 do *loop* V3 dos isolados do HIV-1.**

A presença de aminoácidos carregados positivamente nas posições 11 e 25 do *loop* V3 de gp120 possui relevância para a determinação do fenótipo SI. Aminoácidos de carga positiva, arginina (R) ou lisina (K), em uma das posições 11 e/ ou 25 sugere a utilização de CXCR4 como co-receptor e a indução de um fenótipo SI. A presença de aminoácidos não básicos nessas posições está relacionada ao fenótipo NSI e a utilização de CCR5 como co-receptor nas células infectadas (Koning *et al.*, 2002). Diante desse contexto, analisamos a presença de resíduos de R e K nas 41 seqüências de aminoácidos deduzidas para as amostras analisadas neste estudo. A Figura 18 mostra o alinhamento das seqüências dos isolados do Distrito Federal com o consenso do subtipo B, e mostra em destaque as posições 11 e 25 do *loop* V3.

		11		25	35	
CONS-B	1	CTRPNNNTRK	S	IHIGPGRAFYYTTG	E	IIGDIRQAHC
3		.....	.	..M...KT..A.-	D	E.....K.Y.
6		.....	.	..M.W...I.MP.	G	.....
7		.....	.	.P.....A..	D	.....
9		.....	G	..L.F..TL.A..	<b>R</b>	.....
12		....K....	<b>R</b>	.PM.L.....	S	....KK...
15		....T....	<b>R</b>	.QR...N...V..	G	.T.-----
17		.....	.	.S.....A..	S	..N.....
21		.I.....	G	.....AS..A..	.	.....
25		.....	.	..L.L.....A..	.	...S.....
26		.I..G.T...	<b>R</b>	.RG...Y...V..	.	.T...K...
30		.....	.	.T.....A..	.	.....
32		.....	.	.N.....A..	M	.....
33		.....	.	.N.....A..	D	.....
34		....G....	<b>R</b>	.....ASE	G	.V.....
35		.....	.	.S.....FA..	D	.....
36		.I..S...R	.	..L.Q...W....	.	.....K.Y.
37		.....	.	.YV...T.....	-	K..N.....
43		.....	.	.....K...A.-	D	.....
48		....G....	G	.....A..	-	KT...K...
50		....S....	G	.....A.E	A	.T.....
51		.....	.	.....	A	.....
55		.I.....	.	LP.....LFA..	A	.....
58		.I.....	G	.....A..	D	..N.....
59		.....I.	G	.....S...AR.	Q	.....R.Y.
60		.....	.	.Q.....AI.	<b>K</b>	.....
67		....G....	.	.....A..	D	.....
69		.I..G....	G	.R....TI.A..	<b>R</b>	.....
71		.....R.	G	.....	.	.T...E..
72		.I.....R	.	..M.F...L..N.	.	.V.....
73		.....	G	.....T..A..	D	..N.....
75		.I.....R	.	..M...T..AE.	D	.....
80		.....	.	..M...K...A..	D	..N.....
81		.....R	.	..M.W...A..	.	.V.....
83		....G....	.	.PL.W..VL.A..	D	.....
86		....G...I.	.	..MAL...L.A.E	<b>R</b>	.V.N..N..
88		.....	.	...F...L.A..	A	.Q.....
91		.....	.	.N.....A..	D	.....
94		.....	.	..AP.G...A..	.	.....
95		.....I.R	.	.RM.W..T..A..	Q	.K.S.....
99		.....	.	.PM...H...HA.D	.	.....
102		.....	<b>R</b>	.N.....LA..	A.	.....

**Figura 18.** Alinhamento das seqüências de aminoácidos da região do *loop* V3 de gp120 com a seqüência consenso do subtipo B, destacando-se as posições 11 e 25. Em vermelho estão estacados os resíduos de R e K encontrados nas posições 11 e 25 do *loop* V3 de gp120. Aminoácidos idênticos são indicados por pontos, deleções são indicadas por traços.

Ao analisarmos a disposição de aminoácidos nas posições 11 e 25 do *loop* V3 de gp120, encontramos que 9 das 41, (22%) apresentaram resíduos carregados positivamente em pelo menos uma das posições avaliadas. Nenhum dos isolados apresentou aminoácidos de carga positiva nas duas posições analisadas. As deleções encontradas nessas posições estão descritas na literatura e não estão associadas a um fenótipo SI (Milich *et al.*, 1993; Fouchier, 1995). Este resultado sugere que 22% dos isolados analisados possuem fenótipo indutor de sincício e utilizam como co-receptor CXCR4 e que o restante das amostras analisadas, provavelmente, utilizam CCR5 como co-receptor e produzem um fenótipo NSI.

## 5. DISCUSSÃO

A genotipagem do HIV-1 é de grande importância para a caracterização dos subtipos e seus polimorfismos existentes em uma dada região geográfica. Tais aspectos possuem relevância para compreensão das características epidemiológicas da infecção pelo HIV-1, assim como as formas de transmissão e patogênese do vírus, além de possuir implicações na eficácia das vacinas anti-HIV-1 que venham a ser adotadas (Morgado *et al.*, 2002; Thomson *et al.*, 2002b).

A diversidade genética de V3 tem implicações na resistência aos anti-retrovirais, na resposta imune, na patogenicidade das linhagens virais e no tropismo celular (Koning *et al.*, 2002; Brumme *et al.*, 2004). A emergência do fenótipo indutor de sincícios (SI), na fase tardia da doença, é associada com um rápido decréscimo dos linfócitos T CD4<sup>+</sup> e rápida progressão da doença. Além disso, o fenótipo SI também é associado a uma resposta mais fraca às drogas anti-retrovirais (Brumme *et al.*, 2004). Esses aspectos ressaltam a importância de se conhecer e monitorar a variabilidade genética de *env*.

O Brasil é um país de proporções continentais, conseqüentemente pode-se observar diferentes padrões de distribuição dos subtipos nas diferentes regiões geográficas. O subtipo predominante no Brasil é o subtipo B, seguido pelo subtipo F1, mas observa-se também a presença dos subtipos C, D e A. Há ainda relatos de formas recombinantes circulantes (CRFs), sendo as mais prevalentes aquelas entre os subtipos B e F. O subtipo C tem apresentado uma prevalência crescente nos estados do Sul e do Sudeste do Brasil. O subtipo C é, atualmente, o de maior prevalência em todo mundo. Vários estudos têm demonstrado que a prevalência do subtipo C não apenas tem aumentado nas regiões Sul e Sudeste, mas também está se expandindo para as regiões mais ao norte do país (Morgado *et al.*, 2002; Soares *et al.*, 2003).

É importante destacar que a distribuição dos subtipos do HIV-1 em uma localidade tem determinantes próprios e não pode ser generalizada para todo país, ou até mesmo para uma região geográfica (Rossini *et al.*, 2001). O conhecimento da diversidade do HIV-1 em uma região tem implicações importantes sobre a epidemiologia do HIV-1, tanto para o diagnóstico, como na eficácia do tratamento, na avaliação da resistência às drogas anti-retrovirais, e no desenvolvimento de vacinas anti-HIV. Essa ferramenta também pode prover informações relevantes sobre a origem e a propagação dos variantes do HIV-1 em distintas regiões geográficas, as formas de



transmissão e a patogenicidade dessas linhagens virais (Thomson *et al.*, 2002b; Perrin *et al.*, 2003).

Os primeiros estudos de caracterização e epidemiologia molecular do HIV-1 eram realizados com base na região do envelope, devido a interesses no desenvolvimento futuro de vacinas contra HIV-1. Mais recentemente, análises baseadas no gene *gag* também foram utilizadas para determinar as formas genéticas do HIV-1 (Morgado *et al.*, 2002; Thomson *et al.*, 2002b). No entanto, a existência de recombinação entre diferentes subtipos torna necessária a análise de diferentes regiões do genoma para que se possa caracterizar as formas genéticas recombinantes, as infecções mistas causadas por diferentes subtipos e para a melhor compreensão das formas genéticas circulantes em uma região (Castro *et al.*, 2003; Thomson *et al.*, 2002a).

Para melhor conhecer as características epidemiológicas do HIV-1 no Brasil vários métodos são utilizados na definição das formas genéticas circulantes desse vírus. Dentre as técnicas utilizadas está o sequenciamento de isolados do HIV-1, que é sensível e reflete melhor os polimorfismos das linhagens virais. No entanto, há outras técnicas, como RFLP (*Restriction Fragment Length Polymorphism*) e o EIA (*Enzyme Immunoassay*), um método sorológico e o HMA (*heteroduplex mobility assay*), um dos métodos mais utilizados para monitorar a distribuição dos subtipos no Brasil e na América do Sul (Casseb *et al.*, 2002; Castro *et al.*, 2003; Gadelha *et al.*, 2003).

### **5.1. Definição das formas genéticas do HIV-1**

Neste estudo, sobre a caracterização antigênica do HIV-1 no Distrito Federal, foram testadas 60 amostras de indivíduos HIV-1 positivos com carga viral acima de 50.000 cópias de RNA/mL de soro. Dessas 60 amostras, 53 (83%) foram amplificadas com os conjuntos de iniciadores utilizados ED5/ED12 –ED31/ED33. Essas amostras tiveram suas formas genéticas definidas. Os subtipos foram determinados por análises de homologia, filogenia e em caso de formas genéticas recombinantes, também foi utilizada a análise de *bootscanning*, pelo programa *SimPlot*. Todas as amostras apresentaram resultados concordantes entre a análise de homologia e a filogenia. A maioria das seqüências analisadas, 51 das 53 (96,2%), foi classificada como do subtipo B, que é o mais prevalente em todo Brasil (Vicente *et al.*, 2000; Morgado *et al.*, 2002; Gadelha *et al.*, 2003; Soares *et al.*, 2003; Cerqueira *et al.*, 2004b). As formas genéticas

não B apresentaram uma prevalência de 3,8% (2 de 53), sendo encontrado um subtipo F1 e uma forma recombinante B/F. O isolado DF 02.48, que apresentou como genótipo a forma B/F, teve o resultado encontrado na homologia e na filogenia confirmado por meio de *bootscanning*, que detectou o ponto de recombinação na região C2/C3 de gp120. A ocorrência do subtipo F1 e da forma recombinante B/F foi observado em um caso cada um (1,8%).

No nosso estudo, a prevalência de formas genéticas não B e de recombinantes foi menor que as relatadas por Gadelha *et al.* (2003) para o Ceará, que encontraram a prevalência de 2,7% para os recombinantes B/F e para o subtipo F. Cerqueira *et al.* (2004a) relataram a prevalência de 10,5% de formas recombinantes no Distrito Federal, com base na análise dos genes da protease e da transcriptase reversa. Na região Norte Vicente *et al.* (1999), por meio da análise da região C2/V3 de gp120 e do gene da protease, relataram a prevalência das formas genéticas recombinantes B/F de 35,5% de e do subtipo F de 16%. Outros estudos conduzidos, principalmente, nas regiões Sudeste e Sul do Brasil apresentaram um maior percentual de subtipos não B, o subtipo F possui uma prevalência entre 9 e 23%. O subtipo C está se expandindo no sentido Sul-Norte do Brasil, e já apresenta uma prevalência de mais de 40% no Rio Grande do Sul, e as formas recombinantes também têm aumentado o seu percentual de contribuição para epidemiologia da AIDS em todo país (Morgado *et al.*, 1998; Bongertz *et al.*, 2000; Martinez *et al.*, 2002; Brindeiro *et al.*, 2003; Soares *et al.*, 2003).

A recombinação B/F na região C2/C3 de gp120 relatada neste estudo, ainda não foi descrita no Brasil. Os estudos realizados para a caracterização das formas genéticas do HIV-1 no Brasil, em sua maioria, utilizam o método de HMA, baseado em C2/C3 de gp120. A análise das seqüências dessa região é realizada para uma pequena parcela das amostras (Gadelha *et al.*, 2003; Bongertz *et al.*, 2000). Talvez por isso, a variabilidade genética desse segmento genômico pode estar sendo subestimada. Além disso, é importante ressaltar que a análise somente da região C2/C3 de gp120 pode não caracterizar precisamente a diversidade genética de uma região. Já que a detecção de formas genéticas recombinantes não é muito eficiente quando se analisa somente uma região genômica. Portanto, para melhor caracterizar a diversidade genética do HIV-1, outras regiões do genoma devem ser analisadas (Bongertz *et al.*, 2000; Thomson *et al.*, 2002a; Gadelha *et al.*, 2003).

Para avaliar a diversidade genética do HIV-1 no Distrito Federal, os resultados encontrados pela análise da região C2/C3 de GP120 foram complementados com os obtidos pela análise das regiões genômicas de RT, PR, *gag* e *nef*, dados que foram obtidos pelo nosso grupo de pesquisa. A comparação desses dados mostrou uma prevalência do subtipo B em 45 amostras (85%) e a ocorrência de oito (15%) amostras discordantes, todas recombinantes B/F. A prevalência de 15% de formas genéticas discordantes entre os genes *env*, *gag*, *nef*, da RT e da PR é semelhante à encontrada no Brasil por Brindeiro *et al.* (2003) que analisaram os genes da RT e PR de diversas regiões do Brasil (Rio Grande do Sul, Paraná, São Paulo, Rio de Janeiro, Mato Grosso do Sul, Pará, Bahia e Ceará) e encontraram uma prevalência de formas genômicas divergentes de 14,48%. Estudos realizados, anteriormente no Distrito Federal, por Cerqueira *et al.* (2004a, 2004b) relataram a prevalência de formas genéticas divergentes entre os genes da PR e da RT de 10,5%.

Thomson *et al.* (2002a) descreveram o aumento da prevalência das formas genéticas recombinantes B/F na Argentina quando se avalia o genoma completo de isolados do HIV-1. Tais resultados evidenciam a importância da análise de outras regiões genômicas para melhor caracterizar as formas genéticas recombinantes. A análise de múltiplas regiões do genoma pode evidenciar pontos de recombinação não acessíveis quando se utiliza apenas uma porção genômica para a determinação das formas genéticas circulantes do HIV-1 em uma dada região geográfica (Bongertz *et al.*, 2000; Thomson *et al.*, 2002a).

## **5.2. Caracterização da diversidade genética dos isolados do HIV-1 no Distrito Federal**

O subtipo B do HIV-1 é predominante em todo Brasil. Dentre os isolados brasileiros do subtipo B pode-se descrever uma série de variações intra-subtipo genéticas e antigênicas, que podem ter várias implicações na patogenicidade desses isolados nos testes de vacinas para essa região. Muitos variantes do subtipo B apresentam no topo do *loop* V3 de gp120 o tetrapeptídeo GWGR, no lugar da assinatura GPGR, muito conservada nos isolados norte-americanos/europeus (Morgado *et al.*, 1998; Morgado *et al.*, 2002).

Com o objetivo de melhor caracterizar a variabilidade dos isolados do HIV-1 circulantes no Distrito Federal foi realizada a predição das seqüências de aminoácidos

da região do *loop* V3 de gp120 em 41 das 53 amostras analisadas neste trabalho. Foram avaliadas 40 amostras do subtipo B, incluindo a amostra BSB02.48 que é um variante B/F, e uma amostra do subtipo F, a BSB02.07. Dentre as 40 amostras do subtipo B, foi encontrada a prevalência da assinatura GPGR, comum nas linhagens norte americanas e européias, em 20 das 40 amostras analisadas, (50%). Em seguida, o motivo GWGR, variante B'', foi detectado em 4 amostras; os motivos GFGR e GPGK foram encontrados em 3 amostras, cada um; o motivo GLGR em 2. Outras variações desse motivo foram encontradas em 8 amostras, (20%), são elas: GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH.

A prevalência do variante B'' encontrada nesse estudo difere das relatadas em outras regiões do Brasil, que variam de 30 a 57% (Covas *et al.*, 1998; Bongertz *et al.*, 2000; Gadelha *et al.*, 2003). A prevalência de 10% do variante B'' encontrada nos isolados do subtipo B do Distrito Federal é a menor já relatada no Brasil. Por outro lado, foram encontrados outros 11 variantes nas seqüências do topo do *loop* V3 analisadas. Os mais prevalentes foram os motivos GFGR e GPGK que foram encontrados em 3 amostras cada um e o motivo GLGR, que teve uma prevalência de 5% (2 das 40). Bongertz *et al.* (2000), em um estudo realizado nos estados da região Sudeste (Rio de Janeiro, São Paulo e Minas Gerais), relataram a prevalência de 10,5% para cada um dos motivos GFGR e GPGS. Covas *et al.* (1998) também relataram a ocorrência dos motivos GFGR, GLGR e GPGK em um estudo realizado em São Paulo.

Foi observada uma heterogeneidade nas seqüências analisadas. Além dos variantes descritos acima, outros 8 motivos, 20% das amostras, foram encontrados em menor prevalência (apenas em um isolado). Os motivos GPGA, GQGR, GPGS, GPGN, APGG e GPGH já haviam sido relatados na literatura. Os tetrapeptídeos GPGY e ALGR relatados neste estudo, não estão descritos anteriormente na literatura (<http://hiv-web.lanl.gov>).

Novamente, é importante ressaltar que a maioria dos estudos sobre a variabilidade intra-subtipo B no Brasil é realizada a partir do método de RFLP (*Restriction Fragment Length Polymorphism*), baseado na digestão do fragmento de DNA amplificado pelos *primers* ED31/ED33, para a região C2/C3 de gp120, pela enzima restrição *FOK I*, que reconhece a seqüência de DNA (GGA TGG GGA) correspondente ao tripeptídeo GWG (Morgado *et al.*, 1998). Essa técnica é capaz de

distinguir com bastante eficiência o variante brasileiro B'' do subtipo B norte-americano/europeu (Morgado *et al.*, 1998; Covas *et al.*, 1998). No entanto, o uso desse método pode subestimar a heterogeneidade dessa região, já que as amostras negativas para *FOK I*, em sua maioria, não são seqüenciadas e podem apresentar seqüências distintas da assinatura GPGR, encontrada nos isolados norte-americanos/europeus (Covas *et al.*, 1998).

No isolado DF 02.07, do subtipo F, descrito nesse trabalho, foi encontrado o motivo GPGR ao invés da assinatura GPGQ característica dos isolados pertencentes aos subtipos não B. No entanto, este resultado está de acordo com outros já relatados no Brasil e em outros países da América do Sul, sugerindo que a presença do motivo GPGR em isolados não B é típica da América do Sul (Bandeia *et al.*, 1995; Apetrei *et al.*, 1998; Bongertz *et al.*, 2000; Castro *et al.*, 2003).

O papel dessa heterogeneidade encontrada no *loop* V3 de gp120 na epidemia da AIDS no Brasil, ainda não está bem definido. O variante B'', que possui prevalência elevada em algumas regiões do país, pode ser diferenciado genética e antigenicamente do variante do subtipo B predominante no Brasil (Casseb *et al.*, 1998; Morgado *et al.*, 2002). Há estudos que demonstram a existência reação cruzada entre ambos variantes do subtipo B. Bongertz *et al.* (2005), em um estudo para avaliar a capacidade de neutralização cruzada entre os subtipos do HIV-1 predominantes na região Sudeste, relataram a capacidade de reação cruzada dos anticorpos de indivíduos infectados com o isolado B norte-americano/europeu (GPGR), o isolado B'' (GWGR) e isolados do subtipo F. Mas relataram uma menor freqüência de reação cruzada entre os isolados do subtipo F pelos anticorpos anti-GWGR, sugerindo que esse anticorpo seja mais específico. Casseb *et al.* (1998) estudaram alguns aspectos epidemiológicos do variante B'' em São Paulo, e não encontraram distinções entre as formas de transmissão e a distribuição desse variante na população estudada. Tais fatos mostram a importância que o isolado B'' possui para a epidemiologia do HIV-1 no Brasil, principalmente em relação ao desenvolvimento e aos testes de vacinas neste país (Morgado *et al.*, 2002).

Em relação aos outros variantes encontrados, não há estudos de suas implicações sobre a epidemia do HIV/AIDS no Brasil. Além, disso, existem relatos sobre a neutralização cruzada entre variantes do mesmo subtipo e entre subtipos diferentes no

Brasil. Portanto, mais estudos devem ser realizados para a melhor caracterização da diversidade do *loop* V3 de gp120 no país.

### **5.3. Caracterização da diversidade antigênica dos isolados do HIV-1 no Distrito Federal**

O *loop* V3 é uma região que apresenta alta variabilidade. No entanto, algumas características estruturais e conformacionais conservadas. É uma região de alto potencial imunogênico e alvo de interesse de muitos grupos de pesquisa em todo mundo (Gorny & Zolla-Pazner, 2004; Binley *et al.*, 2004). O *loop* V3 de gp120 possui várias regiões antigênicas e anticorpos anti-V3 podem bloquear a entrada do vírus na célula, além de influenciarem na infectividade e no tropismo celular da partícula viral (Covas *et al.*, 1998; Burton *et al.*, 2004).

Entre os isolados do subtipo B há a predominância do variante norte-americano/europeu, e, apesar da variabilidade genética há a presença do motivo GPGR, no topo do *loop* V3, relativamente conservado nesses isolados. No entanto esse motivo é pouco conservado entre as linhagens virais dos outros subtipos. Essa região é relacionada a ligação ao co-receptor, função essencial às linhagens virais. Conseqüentemente, apesar da maioria dos anticorpos anti-V3 ser específica a uma linhagem viral, existem alguns anticorpos que apresentam grande capacidade de reação cruzada (Kwong, 2004).

O anticorpo monoclonal 447-52D é o anticorpo anti-V3 de maior reação cruzada. Reconhece o motivo GPGR, é capaz de neutralizar cerca de 50% dos isolados do subtipo B, além de não fazer distinção entre os isolados R5 e X4. A capacidade de neutralização desse anticorpo está sendo analisada por diferentes grupos de pesquisa, para melhor compreender os aspectos envolvidos na interação desse anticorpo com o *loop* V3 (Kwong, 2004). Esse anticorpo interage com 14 resíduos no topo do *loop* V3, mas vários estudos demonstram que três resíduos são fundamentais para a interação entre o anticorpo 447-52D e o *loop* V3, os resíduos de glicina (G), prolina (P) e arginina (R). Assim, o epitopo do anticorpo 447-52D é mais bem representado pela seqüência GPxR ( Gorny & Zolla-Pazner, 2004; Zolla-Pazner *et al.*, 2004).

Zolla-Pazner *et al.* (2004) analisaram a capacidade de reação cruzada ente o anticorpo 447-52D e 57 isolados primários pertencentes aos subtipos A, B, C, D, E, F,

G e H do HIV-1 contendo os motivos GPGR, GPGQ (típico dos isolados do subtipo C), e outras variantes (GPGA, GPGH, APGR, ARGR, AVPG, GGGR), no topo do *loop* V3. Nesse estudo, foi relatada a neutralização de 92% dos isolados do HIV-1 que apresentavam o motivo GPGR. Esses isolados pertenciam a diferentes subtipos: B, F, A e H. Outros isolados primários, que não possuíam motivo GPGR no topo do *loop* V3, não foram neutralizados por 447-52D. Binley *et al.* (2004) também relataram resultado semelhante, a neutralização de todas as seqüências que continham o motivo GPGR e de nenhum dos isolados que apresentaram variações nessa região.

A presença do resíduo de R no motivo GPGR é essencial à ação neutralizante do anticorpo 447-52D. Estudos estruturais da ligação deste anticorpo ao *loop* V3 mostraram uma série de interações eletrostáticas entre o resíduo de R do motivo GPGR e a cadeia pesada de 447-52D, essas interações contribuem para a estabilização do complexo anticorpo/vírus, e permitem a neutralização viral. Quando o resíduo de R é substituído por outro, essas interações se tornam menos favoráveis e dificultam a neutralização do vírus (Gorny & Zolla-Pazner, 2004; Stanfield *et al.*, 2004; Zolla-Pazner *et al.*, 2004).

Neste estudo foi encontrada uma grande variabilidade na região do *loop* V3 de gp120 dos isolados do HIV-1 analisados. Uma alta prevalência do variante norte-americano/europeu, referente a 50% das amostras do subtipo B, apresentando o epítipo GPGR foi verificada, além da presença desta assinatura, GPGR, no isolado DF 02.07 do subtipo F1. Possivelmente, tais isolados podem ser neutralizados por 447-52D. Outros variantes também foram encontrados, em menor proporção, GWGR, GFGR, GPGK, GLGR, GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH, representando 48,78% dos isolados analisados neste trabalho. Tal variabilidade tem conseqüências importantes para testes de vacinas na região Central do Brasil.

O epítipo GPGR é utilizado em estudos clínicos para testes de vacinas eficazes contra o HIV-1 (Pialoux *et al.*, 2001). Assim, a caracterização e a descrição desses domínios antigênicos nas amostras de HIV-1 circulantes é de grande importância para futuros testes de vacinas anti-HIV na região do Centro-Oeste, onde poucos dados sobre a variabilidade antigênica do HIV-1 estão disponíveis. Entretanto, é relevante ressaltar que uma vacina anti-HIV deve ser composta por uma série de epítopos que induzam a ação de distintos anticorpos neutralizantes. Por isso, a caracterização antigênica de

outras regiões genômicas do HIV-1 é importante para se conhecer a distribuição dos principais epítopos neutralizantes circulantes em uma região geográfica. Outros estudos sobre a caracterização antigênica do HIV-1 no Distrito Federal estão sendo realizados por nosso grupo de pesquisa em Virologia, com o objetivo de traçar um perfil pré-vacinas dessa região.

#### **5.4. Caracterização dos resíduos de aminoácidos 11 e 25 no *loop* V3 dos isolados do HIV-1 circulantes no Distrito Federal**

A diversidade genética do *loop* V3 possui papel importante na capacidade de indução de sincícios, patogenicidade, na resistência aos anti-retrovirais, no tropismo celular e no uso do co-receptor. As linhagens virais indutoras de sincícios, SI, utilizam como co-receptor CXCR4 ou CCR5 e CXCR4. Estas linhagens estão presentes na fase tardia da infecção pelo HIV-1, são associadas à progressão mais rápida da AIDS e possuem maior virulência. Já as linhagens não indutoras de sincícios, NSI, utilizam predominantemente CCR5 como co-receptor, são prevalentes na fase inicial da infecção do HIV-1 e estão associadas à progressão mais lenta da AIDS (Thomson *et al.*, 2002b).

Na interação do *loop* V3 com os co-receptores CCR5 e CXCR4 o efeito de mutações em apenas um aminoácido, apesar de diferir entre os isolados virais, pode influenciar diretamente na interação dessa região com o co-receptor, e alterar a preferência e a afinidade do uso de um dos co-receptores. Em relação a isso, mutações nesse segmento genômico podem levar a uma forma viral mais patogênica. Esse processo de transição ainda não é bem conhecido, mas alterações na seqüência do *loop* V3 estão associadas com o fenótipo SI (Fouchier, 1995; Koning *et al.*, 2002; Dragic, 2001).

A presença de aminoácidos carregados positivamente em uma ou em ambas as posições 11 e 25 do *loop* V3 possui relevância para a determinação do fenótipo SI. Já que a presença de um ou ambos resíduos positivamente carregados nessas posições é encontrada em mais de 90% dos isolados X4. De uma forma geral, os variantes SI possuem alta carga positiva em suas seqüências, e emergem ao longo do curso da doença (Koning *et al.*, 2002). Esta associação é encontrada em diferentes subtipos do HIV-1: A, B, C, D, E (CRF01\_AE), F e G (Thomson *et al.*, 2002b). Não é claro que a simples presença de aminoácidos de carga positiva é suficiente, ou até mesmo



necessária para o surgimento de vírus X4, nem mesmo como um acúmulo gradual de mutações pode levar a tal fenótipo (Jensen *et al.*, 2003).

A determinação do fenótipo dos isolados do HIV-1 pelos métodos convencionais, por meio da infecção de células MT2 e observação de fenótipo induzido, não é ideal para ser realizada como teste de clínico de rotina, devido à alta demanda de custo e de tempo. Assim, a genotipagem pode facilitar o monitoramento das linhagens virais durante a infecção do HIV-1. A caracterização dos aminoácidos localizados nas posições 11 e 25 é um método simples e viável para a determinação do fenótipo e do co-receptor utilizado pelos isolados do HIV-1. Além da determinação do fenótipo 11/25, outros métodos de bioinformática estão sendo desenvolvidos para realizar a predição, de maneira mais precisa, do uso de co-receptor a partir das seqüências virais (Jensen *et al.*, 2003; Brumme *et al.*, 2004).

Neste estudo, verificou-se a presença de resíduos de aminoácidos carregados positivamente (K, R) em uma ou em ambas as posições 11 e 25. A presença desses resíduos pode sugerir um fenótipo formador de sincícios, de maior patogenicidade e que pode levar a uma progressão mais rápida da doença. Resíduos carregados positivamente nessas posições foram encontrados em 9 das 41, (22%), amostras analisadas, sugerindo um fenótipo não indutor de sincícios na maior parte das amostras deste estudo (78%), que não apresentaram aminoácidos de carga positiva em nenhuma das posições avaliadas. Não foram encontrados resíduos carregados positivamente em ambas as posições, 11 e 25, em nenhum dos isolados do HIV-1 avaliados.

Os resultados encontrados neste trabalho contribuem para a melhor compreensão da epidemiologia do HIV-1 e também para o futuro planejamento de vacinas para o Distrito Federal. A heterogeneidade descrita para a região do *loop* V3 entre os isolados do HIV-1 circulantes nessa região tem implicações genéticas e imunogênicas importantes, sendo relevante o seu monitoramento. Assim, este estudo, que foi realizado com amostras coletadas no ano de 2002, constitui parte de uma série histórica sobre a caracterização dos subtipos do HIV-1 no Distrito Federal, que pretende analisar amostras coletadas no período de 2002 a 2007. Através desse estudo, iniciado a partir deste trabalho, será possível conhecer a história evolutiva do HIV-1 na região do Distrito Federal.

## 6. CONCLUSÕES

✓ O subtipo B do HIV-1 foi o mais prevalente, estando presente em 96,2% das amostras, caracterizadas a partir da região C2/C3 de *env*, seguido pelo subtipo F1 e pela forma genética recombinante B/F que apresentaram prevalência de 1,8% cada um.

✓ Nossos resultados estavam de acordo com a caracterização do HIV-1 em outras regiões do Brasil, em que o subtipo B é a forma genética mais prevalente, seguido do subtipo F1.

✓ Em nosso estudo pudemos descrever uma forma genética recombinante B/F caracterizadas a partir da região C2/C3 de *env*.

✓ A comparação da caracterização realizada a partir da região C2/C3 de *env* com a definição dos subtipos realizada a partir das regiões: *gag*, PR, RT e *nef* do genoma do HIV-1 apresentou resultados discordantes em 15% das amostras, todas recombinantes B/F.

✓ A prevalência de 15% de formas genéticas recombinantes encontrada em nosso estudo, é superior à anteriormente relatada para o Distrito Federal de 10,5% (Cerqueira *et al.*, 2004a). Esse aumento se deve ao maior número de regiões genômicas analisadas neste estudo, o que permitiu a melhor caracterização das formas genéticas do HIV-1 circulantes em uma região.

✓ A análise da variabilidade intra-subtipo revelou a prevalência de 50% do variante norte-americano/europeu nos isolados do subtipo B analisados.

✓ O variante típico brasileiro B'' apresentou uma prevalência de 10%. Prevalência bem menor que as relatadas em outras regiões do Brasil.

✓ Outros variantes do motivo GPGR, característico das amostras norte-americanas/europeias, foram relatados neste estudo. São eles: GFGR, GPGK, GLGR, GPGN, GPGA, GPGY, GQGR, GPGS, ALGR, APGG, GPGH.

✓ Os motivos GPGY e ALGR ainda não tinham sido relatados na literatura.

✓ O isolado DF 02.07, subtipo F, apresentou o motivo GPGR, na região do topo do *loop* V3, resultado semelhante a outros relatados no Brasil e em outros países da América do Sul.

✓ A variabilidade genética e antigênica encontrada nos isolados do HIV-1 analisados neste estudo não é relatada em outras regiões do país. Isso se deve ao fato da

caracterização da variabilidade intra-subtipo ser, em sua maioria, realizada por RFLP, o que pode estar subestimando a diversidade genética dessa região.

✓ A presença do epítipo GPGR essencial à ação neutralizante do anticorpo anti-V3 447-52D foi relatada em 51,2% das amostras analisadas.

✓ A presença de aminoácidos carregados positivamente nas posições 11 e 25, sugerindo um fenótipo indutor de sincício e de maior virulência dos isolados do HIV-1, foi encontrada em 22% das amostras analisadas.

✓ A caracterização da variabilidade genética intersubtipo e intra-subtipo têm implicações para o desenvolvimento de vacinas anti-HIV-1, que serão testadas em uma região geográfica, além de contribuir para a melhor compreensão da epidemiologia do HIV-1 no Distrito Federal.

✓ Este estudo constitui parte de uma série histórica sobre a caracterização dos subtipos do HIV-1 na região Central do Brasil, que pretende analisar amostras coletadas no período de 2002 a 2007 para poder conhecer a história evolutiva do HIV-1 nessa região.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- Abbas Ak, Lichtman AH. 2005. Vírus da Imunodeficiência Humana e Síndrome da Imunodeficiência Adquirida. *In: Imunologia Celular e Molecular*. 5° ed. Elsevier, p 478-489.
- Altschul SF *et al.* 1990. Basic Local alignment search tool. *J Mol Biol*, v.215, n.3, p.403-410.
- Apetrei C, Necula A, Holm –Hansen C, *et al.* 1998. HIV Diversity in Romania. *AIDS* 12, 1079-1085.
- Arens M. 2001. Clinically relevant sequence-based genotyping of HBV, HCV, CMV, and HIV. *Journal of Clinical Virology* 22:11-29.
- Bandea CL, Ramos A, Pieniazek D. *et al.* 1995. Epidemiological and evolutionary relationship between Romanian and Brazilian HIV-1 subtypes F strains. *Emerg Infect Dis* 1, 91-93.
- Baltimore D. 1970. RNA- dependent DNA polymerase in virions of RNA tumor viruses. *Nature* 226, 1209-1211.
- Berzofsky J. *et al.* 2004. Progress on new vaccine strategies against chronic viral infections. *Journal of clinical investigation* 114 (4), 450-461.
- Binley JM, Wrin T, Korber B, Zwick MB, *et al.* 2004. Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78(23), 13232 –13252.
- Bongertz V, Bou-Habib CD, Brígido MFL, *et al.* 2000. HIV-1 Diversity in Brazil: Genetic, Biologic, and Immunologic Characterization of HIV-1 Strains in Three Pontencial HIV Vaccine Evaluation Sites. *JAIDS* 23: 184-193.
- Bongertz V, Teixeira SLM, Grinztejn B, Pilotto JH, Veloso VG, *et al.* 2005. Human immunodeficiency virus type 1 neutralization by plasma from B and F genotype infected individuals. *Mem Int Oswaldo Cruz* 100 (1), 85-89.
- Bour S, Strebel K. 2003. The HIV-1 Vpu protein: a multifunctional enhancer of viral particle release. *Microbes Infect* 5, 1029-1039.

- Brasil, Ministério da Saúde, Coordenação Nacional de DST e AIDS. Dados e Pesquisas em DST e AIDS. Website, <http://www.aids.gov.br/final/dados/documentos.htm>. Acessado em 20 de dezembro de 2005.
- Brasil, Ministério da Saúde, Coordenação Nacional de DST e AIDS. 2005. Boletim Epidemiológico AIDS DST 2005 – ano II nº 01.
- Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, *et al.* 2003. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *AIDS* 17, 1063-1069.
- Brumme ZL, Dong WYW, Yip B, *et al.* 2004. Clinical and Immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy. *AIDS* 18 (4), 1-9.
- Burton, DR, Desrosiers RC, Doms RW, *et al.* 2004. HIV vaccine design and the neutralizing antibody problem. *Nature Immunology* 5 (3), 233-236.
- Casseb J, Hong MA, Gonzalez C, *et al.* 1998. Two variants of HIV-1 B serotype are transmitted heterosexually in São Paulo, Brazil. *Braz J Med Biol Res* 31 1243-1246.
- Casseb J, Katzenstein D, Winters M, Brígido LFM, *et al.* 2002. Serotyping HIV-1 with V3 peptides: detection of high avidity antibodies presenting clade-specific reactivity. *Braz J Med Biol Res* 35(3), 369-375.
- Castro E, Echeverría G, Deibis L, de Salmen BG, *et al.* 2003. Molecular Epidemiology of HIV-1 in Venezuela: High Prevalence of Subtype B and Identification of B/F Recombinant Infection. *JAIDS* 32, 338-344.
- Centlivre M, Sommer P, Michel M, *et al.* 2005. HIV-1 clade promoters strongly influencespatial and temporal dynamics of viral replication in vivo. *Journal Clinical Investigation* 115 (2), 348-358.
- Cerqueira DM, Amorim MSR, Silva PR, Camara NLG, Brígido MM, Martins CRF. 2004a. Antiretroviral Resistance and Genetic Diversity of Human Immunodeficiency Virus Type 1 Isolates from the Federal District, Central Brazil. *Mem Inst Oswaldo Cruz*, 99(8), 877-882.
- Cerqueira DM, Ramalho ED, Oliveira CP, Silva PR, Franchini M, Felipe MSS, Martins CRF. 2004b. HIV-1 subtypes and mutations associated to antiretroviral drug resistance in human isolates form Central Brazil. *Braz J Microbiol* 35, 187-192.

- Chan DC, Fass D, Berger JM, Kim PS 1997. Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* 89, 263–273.
- Couto-Fernandez JC, Morgado MG, Bongertz V, *et al.* 1999. HIV-1 Subtyping in Salvador, Bahia, Brazil: A City With African Sociodemographic Characteristics. *JAIDS* 22 (3), 288-293.
- Covas DT, Biscaro TA, Kashima S, Duarte G, Machado AA. 1998. High Frequency of the GWG (Pro Trp) Envelope Variant of HIV-1 in Southeast Brazil. *J. Acquired Immune Defic. Syndr. Hum. Retrovirol.* 19(1), 74-79.
- de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, Snoeck J, van Rensburg EJ, Wensing AMJ, van de Vijver DA, Boucher CA, Camacho R, Vandamme AM. 2005. An Automated Genotyping System for Analysis of HIV-1 and other Microbial Sequences. *Bioinformatics* 2005 (*in press*).
- Dragic T. 2001. An overview of the determinants of CCR5 e CXCR4 co-receptor function. *Journal of General Virology*, vol 82, 1807-1814.
- Ewung B. *et al.* 1998. Base-calling of automated sequencer traces using PHRED. I. Accuracy assessment, *Genome Res*, 8, 175-185.
- Fackler OT, Baur AS. 2002. Live and Let Die: *Nef* Functions beyond HIV Replication. *Immunity* 16, 493-497.
- Felsenstein, 1993. PHYLIP (Phylogenetic Interference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Flint, SJ, Enquist, LW, Krug RM, Racaniello VR, Skalka AM. 2004. Principles of Virology: Molecular Biology, Pathogenesis, and Control. Ed. AMS Press, p-627.
- Fouchier, R.A.M. 1995. Simple determination of Human Immunodeficiency Virus type 1 syncytium-inducing V3 genotype by PCR. *Journal of Microbiology*, vol 44, nº3, p-906-911.
- Gadelha SR, Shindo N, Cruz JNM, *et al.* 2003. Molecular Epidemiology of Human Immunodeficiency Virus-1 in the State of Ceará, Northeast, Brazil. *Mem Inst Oswaldo Cruz* 98(4), 461-464.
- Gallo RC. 2002. The early years of HIV/AIDS. *Science*, 298, 1728-1730.
- Gallo RC. 2005. The end or the beginning of the drive to an HIV-preventive vaccine: a view from over 20 years. *Lancet*, 366, 1894-1898.

- Gallo RC & Montagnier L. 2002. Prospects for the Future. *Science*, 298, 1730-1731.
- Gaschen B *et al.* 2002. Diversity Considerations in HIV-1 Vaccine selection. *Science* 296, 2354-2360.
- Golding H, Zaitseva M, de Rosny E, *et al.*, 2002. Dissection of Human Immunodeficiency Virus Type 1 Entry with Neutralizing Antibodies to gp41 Fusion Intermediates *J. Virol.* 76, 6780-6790.
- Gorny MK, Zolla-Pazner S. 2004. Human Monoclonal Antibodies that Neutralize HIV-1. pp 37-51 in *HIV Immunology and HIV/SIV Vaccine Databases 2003*, Edited by: Korber BT, Brander C, Haynes BF, Koup R, Moore JP, Walker BD, and Watkins DI. Published by: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR 04-8162.
- Göttlinger HG. 2001. HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. pp. 2-28 in *HIV Sequence Compendium 2001*. Edited by: Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877.
- Gzyl J, Bolesta E, Wierzbicki A, *et al.* 2004. Effect of partial and complete variable loop deletions of the human immunodeficiency virus type 1 envelope glycoprotein on the breadth of gp160-specific immune responses. *Virology* 318, 493-506.
- Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 42, 182-192.
- Herrera C, Klasse PJ, Michael E, *et al.* 2005. The impact of envelope glycoprotein cleavage on the antigenicity, infectivity, and neutralization sensitivity of Env-pseudotyped human immunodeficiency virus type 1 particles. *Virology*. 338(1), 154-172.
- Hirsch MS, Conway B, D'Aquila RT, Johnson VA, Brun-Vézinet F, Clotet B, Demeter LM, Hammer SM, Jacobsen DM, Kuritzkes DR, Loveday C, Mellors JW, Vella S & Richman DD. 1998. Anti-retroviral Drug Resistance Testing in Adults with HIV Infection. *JAMA* 279 (24), 1984-1991.
- Janeway CA, Travers P, Walport M, Shlomchik M. 2002. *Imunobiologia – O sistema imune na saúde e na doença*. 5° ed., Artmed, São Paulo.
- Jensen MA, Li F, van't Wout AB, Nickle DC, *et al.* 2003. Improved Coreceptor Usage Prediction and Genotypic Monitoring of R5-to-X4 Transition by Motif Analysis of

- Human Immunodeficiency Virus Type 1 *env* V3 Loop Sequences. *J Virol* 77 (24), 13376-13388.
- Kang S., Quan F.S., Huang C. *et al.* 2005. Modified HIV proteins with enhanced binding to neutralizing monoclonal antibodies. *Virology* 331, 20-32.
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequence. *J Mol Evol* 16, 111-120.
- Koning F, van Rij R, Schuitemaker H. 2002. Biological and Molecular Aspects of HIV-1 Coreceptor Usage. pp. 24-42 in HIV Sequence Compendium 2002. Edited by: Kuiken C, Foley B, Freed E, Hahn B, Korber B, Marx PA, McCutchan F, Mellors, JW, and Wolinsky S. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. LA-UR 03-3564
- Kwong PD. 2004. The 447-52D Antibody: Hitting HIV-1 Where Its Armor Is Thickest. *Structure* 12, 173-174.
- Larder B, Richman D; Vella S. 2001. *HIV Resistance and Implications for Therapy*, 2nd ed., MediCom Inc., Atlanta.
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol.*73(1), 152-60.
- Marsh, JW. 1999. The Numerous Effector Functions of Nef. *Archives of Biochemistry and Biophysics* 365 (2),192-198.
- Martinez AMB, Barbosa ED, Ferreira PCP, *et al.* 2002. Molecular epidemiology of HIV-1 in Rio Grande, RS, Brasil. *Revista da Sociedade Brasileira de Medicina Tropical* 35(5), 471-476.
- Milich L, Margolin B, Swanstrom R.1993. V3 Loop of the Human Immunodeficiency Virus Type 1 Env Protein: Interpreting Sequence Variability. *J Virol* 67(9): 5623-5634.
- Miller RJ, Cairns JS, Bridges S, Sarver N. 2000. Human Immunodeficiency Virus and AIDS: Insights from Animals Lentiviruses. *J Virol* 74(16), 7187-7195.
- Montagnier L. 2002. A history of HIV discovery. *Science*, 298, 1727-1728.



- Moore JP, Doms RW. 2003. The entry of entry inhibitors: a fusion of science and medicine. *PNAS* 100(19), 10508-10602.
- Morgado MG, Guimarães ML, Neves Jr, et al. 1998. Molecular Epidemiology of HIV-1 in Brazil: Polymorphism of the Antigenically Distinct HIV-1 B Subtype Stains. *Mem Inst Oswaldo Cruz* 93 (3), 383-386.
- Morgado MG, Guimarães ML, Galvão-Castro B. 2002. HIV-1 Polymorphism: a Challenge for Vaccine Development - A Review. *Mem Inst Oswaldo Cruz* 97 (2): 143-150.
- Morgan DA, Ruscetti FW, Gallo RC. 1976. Selective in vitro growth of T lymphocytes from normal human bone marrows. *Science* 193, 1007-1008.
- Peeters M. 2000. Recombinant HIV Sequences: Their Role in The Global Epidemic. In C Kuiken, B Foley, B Hahn et al. *HIV sequence Compendium 2000*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 01-3860.
- Perrin L, Kaiser L, Yerly S. 2003. Travel and the spread of HIV-1 genetic variants. *Lancet Infect Dis* 3, 22-27.
- Pialoux G, et al. 2001. Lipopeptides induce cell-mediated anti-HIV immune responses in seronegative volunteers. *AIDS* 15 (10): 1239-1249.
- Prusiner SR. 2002. Discovering the Cause of AIDS. *Science* 298, 1726-1727.
- Ray SC. 1999. SimPlot for Windows (version 2.5). Baltimore, MD. Distributed by the author. Available at <http://sray.med.som.jhmi.edu/RaySoft/SimPlot/>.
- Rodrigo GA & Mullins JI. 2001. Heteroduplex Mobility Analysis. Protocol Version 3.1.
- Rossini MAA, Diaz RS, Caseiro M, Turcato C, et al. 2001. HIV-1 subtypes among intravenous drug users from two neighboring cities in São Paulo State, Brazil. *Braz J Med Biol Res* 34, 45-47.
- Sabino EC, Diaz RS, Brígido LF, Learn GH, Mullins JI, Reingold AL, Duarte AJS, Mayer A, Busch MP. 1996. Distribution of HIV-1 subtypes seen in an AIDS clinic in São Paulo City, Brazil. *AIDS* 10: 1579-1584.
- Saphire EO, Parren PW, Pantophlet R, et al. 2002. Crystal structure of a neutralizing human ICC against HIV-1: a template for vaccine design. *Science*, 293, 1155-1159.

- Soares EAJM, Santos RP, Pellegrini JA, Sprinz E, *et al.* 2003. Epidemiologic and Molecular Characterization of Human Immunodeficiency Virus Type 1 in Southern Brazil. *JAIDS* 34(5): 520-526.
- Stanfield RL, Gorny MK, Williams C, Zolla-Pazner S, Wilson IA. 2004. Structural Rationale for the Broad Neutralization of HIV-1 by Human Monoclonal Antibody 447-52D. *Structure* 12: 193-204.
- Tanuri A, Caridea E, Dantas MC, Morgado MG, Mello DLC, Borges S, Tavares M, Ferreira SB, Santoro-Lopes G, Martins CRF, Esteves ALC, Diaz RS, Andreo SMS, Ferreira LAP, Rodrigues R, Reuter T, Cavalcanti AMS, Oliveira SM, Barbosa HB, Teixeira PR, Chequer PN. 2002. Prevalence of mutations related to HIV-1 anti-retroviral resistance in Brazilian patients failing HAART. *J Clin Virol* 25: 39-46.
- Temin HM, Mizutani S. 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211-1213.
- Thompson J, Higgins D & Gibson T. 1994. CUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res.* 22: 4673-4680.
- Thomson MM, Delgado E, Herrero I, *et al.* 2002a. Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. *J Gen Virol* 83:107-119.
- Thomson MM, Pérez-Álvarez L, Nájera R. 2002b. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect Dis* 2: 461-471.
- Trkola A, Dragic T, Arthos J, *et al.* 1996. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature* 384, 184 – 187.
- UNAIDS/WHO. AIDS epidemic update: December 2005. 2005. Geneva, Switzerland: UNAIDS/WHO.
- van Regenmortel MHV, *et al.* 2000. Virus Taxonomy. Seventh Report of International Committee on Taxonomy of Viruses. San Diego; Academic Press, p-1162.

- Vicente ACP, Otsuki K, Silva NB, Castilho MC, *et al.* 2000. The HIV epidemic in the Amazon basin is driven by prototypic and recombinant HIV-1 subtypes B and F. *JAIDS* 23, 327-331.
- Zhi-yong Y, Chakrabarti BK, Xu L, *et al.* 2004. Selective Modification of Variable Loops Alters Tropism and Enhances Immunogenicity of Human Immunodeficiency Virus Type 1 Envelope. *Journal of Virology* 78 (8), 4029-4036.
- Zolla-Pazner S, Zhong P, Revesz K, *et al.* 2004. The Cross-Clade Neutralizing Activity of a Human Monoclonal Antibody is Determined by the GPGR V3 Motif of HIV Type 1. *AIDS Research and Human Retroviruses* 20 (11), 1254-1258.
- Wigg MD. 2002. Vírus da imunodeficiência humana. In: Santos NSO, Romanos MTV, Wigg MD. *Introdução à Virologia Humana*. Guanabara Koogan, Rio de Janeiro, p.183-197.
- Wyatt R, Kwong PD, Hendrickson WA, Sodroski JG. 1998. Structure of the core of the HIV-1 gp120 Exterior Envelope Glycoprotein. pp. III-3-9 in *Human Retroviruses and AIDS 1998*. Edited by: Korber B, Kuiken CL, Foley B, Hahn B, McCutchan F, Mellors JW, and Sodroski J. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
- Wyatt R, Sodroski J. 1998. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. *Science* 280, 1884-1888.

## ANEXO

Seqüências da região C2/C3 de *env* dos isolados do HIV-1 do Distrito Federal, coordenadas baseadas em HXB2.

- **DF02.03 (nucleotídeos 6821 a 7217 ): 391 pb seqüenciados**

AGCCATTACACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATACAT  
TATTGTGCCCCGGCTGGTTTTGCGATTCTAAAGTGTAATGATAAGAAGTTCA  
ATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAA  
TTAGGCCAGTAGTGTCAACTCAACTGCTGTTAAATGGCAGTTTATCAGAAGG  
AGAGGTAATAATTAGATCTGAAAATATCACAAACAATGTTAAAACCATAAT  
AGTACAGCTGAATGAACTGTAGAAATTAAGTACCAGACCCAATAACAA  
TACAAGAAAAAGTATACATATGGGACCAGGGAAAAACATTTTATGCAACAGA  
CGAAATAGGAGATATAAGAAAAGCATATTGT

- **DF02.06 (nucleotídeos 6875 a 7270 ): 393 pb seqüenciados**

TTGTGCCCCGGCTGGTTTTGCGATTTTAAAGTGTAATAATAAAACATTCAAT  
GGAAAAGGACCATGTAAAAATGTCAGCACAGTACAATGTACACATGGAAAT  
AGGCCAGTAGTGTCAACTCAACTACTGTTAAATGGCAGTCTAGCAGAAGAA  
GAGATAGTAATTAGATCTGAAAACCTCACAGACAATGCTAAAACCATAATA  
GTACAGCTGAACAGCTCTGTAGAAATTAAGTACAAGACCCAACAACAAT  
ACAAGAAAAAGCATAACACATGGGATGGGGAAGAGCAATTTATATGCCAGG  
AGGCATAATAGGAGATATAAGACAAGCACATTGTAACATTAGTAGAGCAAA  
ATGGAATAACACTTTAAAACAGATAGCTAGCAAATT

- **DF02.07 (nucleotídeos 6887 a 7409): 508 pb seqüenciados**

TGGGTATGCGATTCTAATGTGCAATGATGAAAAGTTCAATGGGACGGGGCC  
ATGCACGAATGTCAGCACAGTACAATGTACACATGGCATTAAACCAATGGT  
ATCAACTCAATTGTTGTTAAATGGCAGCCTAGTTATAATAATCAGATCTCAA  
AATATCTCAGATAATGCAAAAACCATAATAGTACACCTTAATGAATCTGTG  
CAGATTAATTGTACAAGACCTAACAATAATACAAGGAAAAGTATACCAATA  
GGACCAGGACGAGCATTATGCAACAGGAGATATAATAGGAGACATCAGA  
CAGGCACATTGTAACATTAGTGGAAAACAATGGAATAAAACATTAGATAGG  
GTAAAGGCAAAGTTAAGCCTTTATTTCCCTAATGCAACAATAAGATTTAACT  
CATCCTCAGGAGGGGACCTAGAAGTCACAATGCATAGTTTTACTTGTAGAG  
GAGAATTTTCTACTGCAATACAACAATACTGTTTAATGGCACAAAA

- **DF02.09 (nucleotídeos 6831 a 7380): 541 pb seqüenciados**

CAGGCCTGTCCAAAGGTAACCTTTGAACCAATTCCCATACATTATTGTGCCC  
CAGCTGGTTTTGCGATTCTCAAGTGTAACGATAAGAAATTCAATGGAACAG  
GACCATGTACAAATGTCAGCACAAATACAATGTACACATGGAATTAGACCAG  
TAGTGTCAACTCAACTGCTGTTAAATGGCAGTTTATAGCAGAAGAAGATATAA

TAATTAGATCTGAAAATCTCACAAACAATGTTAAAACCATAATAGTACAGC  
 TGAATGAACTGTACAAATTAATTGTACAAGACCCAACAACAATACAAGAA  
 AAGGTATACATCTAGGATTTGGGAGAACATTGTATGCAACAGGAAGAATAA  
 TAGGCGATATAAGACAAGCACACTGTAACATAACTAGAGCAAAATGGAAC  
 AACACTTTAAAACGGATAGGTATAAAATTAAGAGAACAATTTAATAAAAACA  
 ATAATATTTGATCACTCCTCAGGAGGGGATCCAGAAATTATAATGCACACCT  
 TTAATTGTGGAGGGGAATTTTTCTACTGTA

- **DF02.12 (nucleotídeos 6825 a 7390): 560 pb seqüenciados**

ATTACACAGGCCTGTCCAAAGGTAACCTTTGAGCCAATTTCCCATACATTATT  
 GTGCCCGGCTGGTTTTGCGATTCTAAAAGTGAATAATAAAAAGTTTGATGG  
 AAAAGGACCATGTACAAACGTCAGCACAGTACAATGTACACATGGAATTAA  
 ACCAATAGTGTCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGA  
 AATGGTAATTAGATCTGACAATTTACGGATAATACTAAAACCATAATAGT  
 ACAGCTAAAGGAACCTGTAGAAATTAATTGTACAAGACCCAATAAAAATAC  
 AAGAAAAAGGATACCCATGGGACTAGGGAGAGCATTATATACAACAGGGA  
 GTATAATAGGAGATATAAAAAAAGCACACTGTACCATTAATACAACAAAAT  
 GGAATGAACTCTAAGGCAGATAGCTAAAAAATTAAGAAAAGTTTGAG  
 AACAGAACAAATAGCCTTTAATCGATCCTCAGGAGGAGATCCAGAAATTGTA  
 ATGCACAGTTTTAATTGTAGAGGGGAATTTTTCCACTGTAATACAACAAA

- **DF02.15 (nucleotídeos 6864 a 7390): 494 pb seqüenciados**

CCCATACATTATTGTGCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATA  
 AGCAGTTCAATGGATCAGGACCATGCACAAATGTCAGCACAGTACAATGTA  
 CACATGGAATCAGGCCAGCAGTGTCTACTCAATTACTGTTAAATGGCAGTCT  
 AGCAGAGGAAGAAATAAAAATAGATCTGAAAATTTACAAACAATGCTAA  
 AATCATAATAGTACAGCTGAATGAACTGTAACAATTAATTGTACAAGACC  
 CAACAACACTACAAAGAAAAGGATACAGAGAGGACCGGGAATGCATTTTA  
 TGTAACAGGAGGAATAACAGGAGATATAAGAAAATCGGAATAAACTTTA  
 CAACAGATAGCTAGGAAATTACAAGAACAATTTGGTAATAAAACAATAGTC  
 TTTAATCAATCATCAGGAGGGGACCCAGAGATTGTGATGCACAGTTTTAATT  
 GTGGAGGGGAATTTTTCTACTGTAATTCGTCAA

- **DF02.17 (nucleotídeos 6820 a 7373): 548 pb seqüenciados**

CAGCCATTACACAGGCCTGTCCAAAGGTATCCTTTACTCCAATTTCCATACA  
 TTATTGTGCCCGGCTGGTTTTGCGATTCTAAAATGTAGAGACAAAAGTTC  
 AATGGAACAGGTCCATGTACAAATGTCAGCACAGTACAATGTACACATGGA  
 ATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAA  
 GAAGAGATAGTAATTAGATCCGCCAATTTACGAATAATGCTAAAATCATA  
 ATAGTACAACCTGAATGAATCTGTAGAAATTAATTGTACAAGACCCAACAAC  
 AATACAAGAAAAAGTATATCTATAGGACCAGGCAGAGCATTATGCAACA  
 GGAAGCATAATAGGAAACATAAGACAGGCACATTGTAACATTAGTAAAC  
 AAAATGGCATAACACTCTAAAACAATAAGTAGAAAATTAAGAAATCAATT  
 TGGGAACAAGACAATAGTCTTTAATCACTCCTCAGGAGGAGACCCAGAGAT  
 AGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTTC

- **DF02.21 (nucleotídeos 6875 a 7390): 510 pb seqüenciados**

TTGTACCCCGGCTGGTTTTGCGATTCTAAAATGTAATGATAAGAAGTTCAAT  
 GGAACAGGACCATGTACAAATGTCAGCACAAATACAATGTACACATGGGATT  
 AGGCCAGTGGTGTCAACTCAATTACTGTTAAATGGCAGTTTGGCAGAGGAA  
 GAGGTAGTAATTAGATCTGAAAACTTTACAAACAATTTTAAAAATCATAATA  
 GTACAGCTGAATGAATCTGTACAAATTAATTGTATAAGGCCCAACAACAAT  
 ACAAGAAAAGGTATACATATAGGACCGGGGCATCATTTTATGCAACAGGA  
 GAAATAATAGGAGATATAAGACAGGCACATTGTAACATTAGTACAGGAGA  
 ATGGGAAAAAACTTTAAAACAGATAGCTATGAAATTAATAGAACAATTTAA  
 AAATAAAACAATATCCTTTAATCAATCCTCAGGAGGGGATCCAGAAGTTGT  
 AATGCACACTTTTAATTGTGGAGGGGAATTTTCTACTGTAATTCAACACA

- **DF02.25 (nucleotídeos 6838 a 7390): 547 pb seqüenciados**

GTCCAAAGGTATCCTTTGAACCAATCCCATAACATTATTGTGCTCCAGCTGG  
 TTTTGCCTTCTAAAGTGTAATGATAAGACATTCAATGGATCAGGACTGTGT  
 ACAAATGTTAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTGTCA  
 ACCCAGCTACTGTTAAATGGCAGCCTAGCAGAAGAAGAGATAGTAATTAGA  
 TCTGAAAATTTACGGACAATGCTAAAATCATAATAGTACAGCTGAAAGAA  
 GCTGTAGGAATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATA  
 CATCTAGGATTAGGAAGAGCATTTTATGCAACAGGAGAAATAATAGGAAGT  
 ATAAGACAAGCACATTGTAACCTTAGTAGAGCACATTGGAATAACACTTTA  
 CAGCAGATAGTTACAAAATTAAGAGACCAATTTAATAATAGAACAATAGTC  
 TTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATT  
 GTGGAGGGGAATTTTCTACTGTAATTCAACACA

- **DF02.26 (nucleotídeos 6876 a 7392): 511 pb seqüenciados**

TGTGCCCCCGGCTGGTTTTGCGATTCTAAAGTGTAATAATAAGCAGTTCAATG  
 GATCAGGACCATGCACAAATGTCAGCACAGTACAATGTACACATGGAATCA  
 GGCCAGCAGTGTCTACTCAATTACTGTTAAATGGCAGTCTAGCAGAAGGAG  
 ATGTGATAATTAGATCTGAAAATTTACAAACAATGCTAAAACCATAATAG  
 TACAGCTGAATGAAACTGTAACAATTAATTGTATAAGACCCGGCAACACTA  
 CAAAGAAAAGGATACGTGGAGGACCGGGGTATGCATTTTATGTAACAGGAG  
 GAATAACAGGAGATATAAGAAAAGCACATTGTAACATTAGTAAAGAAAAA  
 TGGGAAAAAACTTTACAGCAGGTAGCTAAAAAATTACAAGAACAATTTGGT  
 AATAAAACAATAGTCTTTAATCAATCATCAGGAGGGGACCCAGAGATTGTG  
 ATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAATTCGTCAAAGC

- **DF02.30 (nucleotídeos 6833 a 7390): 552 pb seqüenciados**

GGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATAACATTATTGTGCCCCG  
 GCTGGTTTTGCGATTCTAAAGTGTAATGATAATAAGTTCAATGGAACAGGA  
 CCATGTAAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTG

GTGTCAACTCAACTGCTGTTAAATGGCAGCTTAGCAGAGGAAGAGGTAGTA  
 ATTAGATCCGAAAATTTACAAACAATGCTAAAACCATAATAGTACAGCTG  
 AACAACTCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAA  
 AGTATAACTATTGGACCAGGGAGAGCATTCTATGCAACAGGAGAAATAATA  
 GGAGACATAAGACAAGCACATTGTAACATTAGTAGGGCAAAGTGGAAAGGA  
 CACCTTAAATAAGATAGTTAGCAAATTAAGAGAACAATTTGGGAATAAAAC  
 AATAACCTTTAACCAGTCCTCAGGAGGGGACCCAGAAATTATAATGCACAC  
 TTTTAATTGTGGAGGGGAATTTTTCTACTGTAATACAACAAA

- **DF02.32 (nucleotídeos 6827 a 7372): 545 pb seqüenciados**

TACACAGGCCTGTCCAAAGACAACCTTTGAACCTATTCCCATACATTATCGT  
 ACCCCGGCTGGGTTTGGCATTCTAAAGTGTAAGGATAAAGATTTCAATGGA  
 ACAGGAGCATGTCAAATGTCAGCACAGTACAATGTACACATGGGATTAGG  
 CCAGTAGTGTCAACTCAACTGTTGTTAAATGGCAGTCTATCAGAAGGGGAG  
 ATAGTAATTAGATCTGAAAATCTCACGAACAATGCTAAAACATAATAGTA  
 CAACTGAATGAACTGTAGAAATCAATTGTACAAGACCCAATAACAACACA  
 AGAAAAAGTATAAATATAGGACCAGGGAGAGCATTCTATGCAACAGGAGA  
 GATGATAGGAGATATAAGACAAGCACATTGCAACATTAGTACAGAAAAATG  
 GAATAAACTTTAGGACAGATAACTATAAAATTAAGGAACAATTTGGGAA  
 TAAAACAATAGTCTTTAATCGATCCTCAGGAGGGGACCCAGAAATTGAAAG  
 GCACAGTTTTAATTGTGGAGGGGAATTTTTCTACT

- **DF02.33 (nucleotídeos 6875 a 7385): 510 pb seqüenciados**

TTGTACCCCGGCTGGTTTTGCGATTCTAAAGTGTAAGGATAAAGATTTCAAT  
 GGAACAGGAGCATGTCAAATGTCAGCACAGTACAATGTACACATGGGATT  
 AGGCCAGTAGTGTCAACTCAACTGTTGTTAAATGGCAGTCTATCAGAAGGG  
 GAGATAGTAATTAGATCTGAAAATCTCACGAACAATGCTAAAACATAATA  
 GTACAACTGAATGAACTGTAGAAATCAATTGTACAAGACCCAATAACAAC  
 ACAAGAAAAAGTATAAATATAGGACCAGGGAGAGCATTCTATGCAACAGG  
 AGATATAATAGGAGATATAAGACAAGCACATTGCAACATTAGTACAGAAAA  
 ATGGAATAAACTTTAGGACAGATAGCTATAAAATTAAGGAACAATTTGG  
 GAATAAAACAATAGTCTTTAATCGATCCTCAGGAGGGGACCCAGAAATTGA  
 AAGGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACA

- **DF02.34 (nucleotídeos 6876 a 7377): 495 pb seqüenciados**

TTGTGCCCGGCTGGTTTTGCAATTCTGCAATGTAAAGATAAGAAGTTCAAT  
 GGAACAGGAATATGTCACAATGTCAGCACAGTACAATGTACACATGGAATT  
 AGACCAGTGGTATCAACTCAACTGCTGTTAAATGGCAGTTTAGCAGAAGAA  
 GAAATAGTAGTTAGATCTGCCAATTTACAGACAATACTAAAACCATAATA  
 GTACAGCTGAAGGAAGCTGTAGAAATTAATTGTACAAGACCCGGCAACAAT  
 ACAAGAAAAAGGATACATATAGGACCAGGGAGAGCATTTTATGCATCAGA  
 AGGTATAGTAGGAGACATAAGACAAGCACATTGTAACATTAGTAGAGCAAA  
 ATGGGATAACACTTTAAAAGGGGTAGTTAAAAAATTAAGAAAACAATTTAT

AAATAAAACAATAGCCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGT  
GATGCACAATTCTAATTGTGGAGGGGAATTTTGCTCT

- **DF02.35 (nucleotídeos 6875 a 7390): 513 pb seqüenciados**

TTGTGCCCCAGCTGGTTTTGCGATTCTAAAGTGTAGAGATCATAATTTCAAT  
GGAACAGGACTATGTAAAAATGTCAGCACAGTACAATGTACACATGGAATC  
AGGCCAGTAGTGTCAACTCAACTACTGTAAATGGCAGCCTAGCAGAAGGA  
GGACAGGTAGTAATTAGGTCTGAAAATTTACAAACAATGCTAAAACCATA  
ATAGTACAGCTGAGTGAAGCTGTAGTAATTAATTGTACAAGACCCAACAAC  
AACACAAGAAAAAGTATAAGTATAGGACCAGGGAGAGCATTTTTTTGCAACA  
GGAGACATAATAGGAGACATAAGACAAGCACATTGTAACATTAGTGGAGC  
AAAATGGAATAACACTTTACAACAGATAGTTGCAAATTAAGAGGACAATT  
TGGGAATAAAACAATAGTCTTTAAGAACTCCTCAGGAGGGGATCCAGAAAT  
TGTAATGCACAGTTTTAATTGTGGAGGAGAATTTTTCTACTGTAATACAACA  
GA

- **DF02.36 (nucleotídeos 6875 a 7367): 487 pb seqüenciados**

TTGTGCCCCGGCTGGTTTTGCGATTCTAAAGTGTAAATGATAAGAAGTTCAAT  
GGATCAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATT  
AGGCCAGTAGTGTCAACTCAACTACTGTAAATGGCAGTCTAGCAGAAGGA  
GAGGTAATGATTAGATCTGAAAACTTTACGGAGAATAAGTACAACATAATA  
GTACAACTGAACAAATCTGTAAAAATTAATGTATAAGGCCAGCAACAAT  
ACAAGAAGAAGTATACATTTAGGACAGGGGAGAGCATGGTATACAACAGG  
AGAAATAATAGGAGATATAAGAAAGGCATATTGTAACATTAGTAAATCAGA  
ATGGGAAGAAGCTTTAAACCAGACAGTTATAAAATTAAGAGAACAATTTGG  
GAATAAAACAATAGCCTTTTATCCATCCTCAGGAGGGGATCCAGAAATTGT  
AACGCACAGTTTTAATTGTGGAGGGGAA

- **DF02.37 (nucleotídeos 6968 a 7284): 308 pb seqüenciados**

AAATGTCAGCGcAGTAgtATcaACTCAACTGCTGTAAATGGCAGTCTAGCAG  
AAGAAGAGGTAGTAATTAGGTCTAAGAATTTACAAACAATGTTAAAACATA  
TAATAGTACAGCTGAATGAATCTGTAATAATTAATTGTACAAGACCCAATA  
ATAATACAAGAAAAAGTATATATGTAGGACCAGGGAGAACATTTTATACAA  
CAGGAAAAATAGGAAATATAAGACAAGCACATTGTAACATTTGGAGGACA  
AATTGGAATAAAACTTTACATCAGATAGTTAAACAATTAAGGGAAGAATTT  
G

- **DF02.43 (nucleotídeos 6824 a 7367): 535 pb seqüenciados**

CATTACACAGGCCTGTCCAAAGATATCCTTTGAGCCAaTTCCCATACATTTCT  
GCGCCCCAGCTGGTTTTGCGATTCTAAAGTGCAACAATAAGACATTCAATG  
GAACAGGACCATGTACAAATGTCAGTACAGTACAATGTACACATGGAATTA  
AGCCAGTGGTGTCAACTCAACTGCTATTAATGGCAGTCTAGCAGAAGAGG  
ATGTAATGATTAGATCAAAAAATTTCTCGGACAATGTTAAAAGCATAATAG



TACAGCTGAAGGAACCAGTAGAAATTAATTGTACAAGACCCAACAACAATA  
 CAAGAAAAAGTATACATATAGGACCAGGGAAAGCATTATGCAACAGACA  
 TAATAGGAGATATAAGACAAGCACATTGCAACCTTAGTGAGACAAAATGGA  
 ATAACACTTTAAAGCAAATAGTTCAAAAATTAAGAGAACAGTTTGGGAACA  
 AAACAATAGTCTTTAATCGATCCTCAGGAGGGGACCCAGAACTTGTAATGC  
 ACATTTTAACTGTGGAGGGGAA

- **DF02.48 (nucleotídeos 6815 a 7264): 447 pb seqüenciados**

CACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCC  
 ATACATTATTGTACCCCGGCTGGTTTTGCGCTTCTAAAGTGTAATGATAAGA  
 AATTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACAC  
 ATGGAATTAGGCCAGTAGTATCAACCCAACTGCTGTTAAATGGCAGTCTAG  
 CAGAAGAAGAGATAGGAATTAGGGCTGAAAATATCTCAGACAATGCAAAA  
 ACCATAATAGTACACTTTAATGAATCTATACAGATTAATTGTACAAGACCCG  
 GCAACAATACAAGAAAAGGTATACATATAGGACCTGGGAGAGCATTATG  
 CAACAGGAAAAACAGGGGACATCAGAAAGGCACATTGTAACGTTAGTAGA  
 CCACAATGGAATAAAACATTAGAACAGGTAAGGGCAAG

- **DF02.50 (nucleotídeos 6824 a 7347): 521 pb seqüenciados**

CATTACAACAGGCCTGTCCAAAGATATCCTTTCAGCCAATTCCCATACATTA  
 TTGTGCCCGGCTGGTTTTGCGATTCTACAATGTAAAGATAAAGGGTTCAA  
 GGAAAAGGACAATGTACAAATGTAAGCACAGTACAATCTACACATGGAATT  
 AGGCCAGTAGTATCAACTCAATTGCTGTTAAATGGCAGTCTAGCAGAAGAA  
 GAGGTAGTAATTAGATCTGACAATTTACGAACAATGCTAAAATATACTAG  
 TACAGCTGAAGGAATCTGTAGAAATTAATTGCACAAGACCCAGCAACAATA  
 CAAGAAAAGGTATACATATAGGACCAGGTAGAGCATTATGCAACAGAAG  
 CCATAACAGGAGATATAAGGCAAGCACATTGTAACATTAGTAGAGCAAAAT  
 GGAATAACACTTTAAAACAGATAGTTATAAAATTAAGAACAATTTGGGT  
 ATAATAAAACAATAGCCTTTAATCAATCCTGAGGAGGGGACCCAGAAATTG  
 TAATGCACA

- **DF02.51 (nucleotídeos 6851 a 7283): 430 pb seqüenciados**

CTTTGAGCCAATTCCCATACATTATTGTGCCCGGCTGGTTTTGCGATTCTAA  
 AGTGTAATGATAAGAAGTTCAATGGAACAGGACCATGTACAAATGTCAGCA  
 CAGTACAATGTACACATGGAATTAGGCCAGTAGTGTCAACCCAACTGCTGT  
 TAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCA  
 CGGACAATGCTAAAACCATAATAGTACAGCTGAACGAATCTGTAGAAATTA  
 ATTGTAACAAGACCCAACAACAATACAAGAAAAAGTATACATATAGGACCAG  
 GGAGAGCATTATACAACAGGCGCAATAATAGGAGATATAAGACAAGCAC  
 ATTGTAACATTAGTAGAGAAAAATGGAATAACACTTTAAAACAGATAGCTA  
 AAAAATTAAGAGAACAATTT

- **DF02.55 (nucleotídeos 6820 a 7286): 464 pb seqüenciados**

CAGCCATTACACAGGCCTGTCCAAAGACATCCTTTGAACCAATTCCCATACA  
 TTATTGTACTCCGGCTGGTTATGCGATTCTAAAGTGTAATAATAAGGAGTTC  
 AATGGAACAGGACCATGTAAAAATGTCAGTACAGTACAATGTACTCATGGA  
 ATTAGGCCAGTAGTGTCAACTCAACTGCTGTAAATGGCAGTCTAGCAGAA  
 GGAGAGATAATACTTAGGTCTGAAAATTTACAAACAATGCTAAAACCATA  
 ATAGTACATTTGAATGAATCTATAGAAATTAATTGTATAAGACCCAACAAC  
 AATACAAGAAAAAGTCTACCTATAGGACCGGGGAGAGCATTATTTGCAACA  
 GGAGCCATAATAGGAGATATAAGACAAGCACATTGTAACCTTAATAGTGCA  
 AAATGGGAAAACACTTTAAAACGGGTAGTAGGAAAATTAAGAGAACAATTT  
 GAA

- **DF02.58 (nucleotídeos 6894 a 7277): 381 pb seqüenciados**

GCGATTCTAATATGTAATGATAAAAAGTTCAATGGAAAAGGACCATGTACA  
 AATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTGTCAACT  
 CAATTACTGTAAATGGTAGTCTAGCAGAAGAAGAGGTAGAAATTAGATCT  
 GAAAACCTTTACAAACAATGCTAAAACCATAATAGTACAGCTGCATCAATCT  
 GTAGGAATTGATTGTATAAGGCCCAATAACAATACAAGAAAAGGTATACAT  
 ATAGGACCGGGGAGAGCATTTTATGCAACAGGAGACATAATAGGAAATATA  
 AGACAAGCACATTGTAACCTTAGCTGAGTAGGGTGAAACAAAACCTTTAAGT  
 CAGATAGTCACAAAATTAAGAGAC

- **DF02.59 (nucleotídeos 6885 a 7369): 479 pb seqüenciados**

GCTGGTTTTGCGATTCTAAAATGTAATGATAAGAAGTTCAATGGAACAGGA  
 CCATGTAAAATGTCAGCACAGTACAGTGTACACATGGGATTAGGCCAGTA  
 GTGTCAACCCAACTACTGTAAATGGCAGCCTAGCAGAAGAAGATGTAGTA  
 ATTAGATCTGAAAATTTCTCGAACAATGCTAAAACCATAATAGTACAGCTG  
 AATAAACCTGTAGTAATTAATTGTACAAGACCCAACAACAATACAATAAAA  
 GGTATACATATAGGGCCGGGAGTGCCTTTTATGCAAGAGGACAAATCATA  
 GGAGACATAAGAAGAGCATATTGTGTCCTTAATGGAATAAAAATGGAATAAC  
 ACTTTACAACAAATAGTTAAAAAATTAAGAGAACAATTTAGGAATAAAAACA  
 ATAGTCTTTAATCACTCCTCAGGAGGGGACCCAGAAATTATAACGCACAGT  
 TTTAATTGTGGAGGGGAATT

- **DF02.60 (nucleotídeos 6828 a 7309): 476 pb seqüenciados**

ACACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGTG  
 CCCC GGCTGGTTTTGCGATTCTAAAATGTAATGATAAGAAGTTCAATGGAAC  
 AGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCC  
 AGTAGTGTCAACTCAACTACTGTAAATGGCAGTCTAGCAGAAGAAGAGGT  
 AGTAATTAGATCTGAAAATTTACGGACAATGCTAAAACCATAATAGTACA  
 GCTGAATAAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAG  
 AAAAAGTATACAGATAGGACCAGGGAGAGCATTTTTATGCAATAGGAAAAA

TAATAGGAGATATAAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGG  
ATAACACTTTAAAACAGATAGTTAGAAAATTAAGAGAACAATTTGGGAATA  
AAACAATAGTCTTTAA

- **DF02.67 (nucleotídeos 6875 a 7366): 486 pb seqüenciados**

TTGTGCCCCGGCTGGTTTTGCGATTCTAAGGTGTAACAATAAGACGTTCAAT  
GGAAAAGGAGAATGTAACGATGTCAGCACAGTACAATGTACACATGGAATT  
AGGCCAGTAGTGTCAACTCAACTGCTGTAAATGGCAGCCTAGCAGAAGAA  
GATGTAGTAATTAGATCTGAAAATTTCTCAAACAATGCTAGAACCATAATA  
GTACAGCTGAAAGAAGCTGTACAAATTAATTGTACAAGACCAGGCAACAAT  
ACAAGAAAAAGTATACATATAGGACCAGGGAGAGCATTTTATGCAACAGG  
AGACATAATAGGAGACATAAGACAAGCACATTGCAACATTAGTGGAACAC  
AATGGAATAACACTTTAAGACAGATAGTTACAAAATTAAGACAACAATTTG  
GAAATAAAACAATAGTTTTTAGCAACTCCTCAGGGGGGGACCCAGAAATTG  
TAACGCACAGTTTTAATTGTGGAGGGGA

- **DF02.69 (nucleotídeos 6830 a 7390): 555 pb seqüenciados**

ACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATATATTATTGTGCC  
CCGGCTGGTTTTGCGATTCTAAAATGTAATGATAAGAAGTTCAATGGAACA  
GGACCATGTAAAAATGTCAGCACAGTACAATGCACACATGGAATTAACCA  
GTAGTATCAACTCAACTACTGTAAATGGCAGTCTAGCAGAAGGAGAGGTA  
GTAATTAGATCTGAAAACTTTACAAACAATGTAAAACCATAATAGTACAG  
CTGAATGAAACTGTAAAAATTAATTGTATAAGGCCCGGCAACAATACAAGA  
AAAGGTATACGTATTGGACCGGGGAGAACAATTTATGCAACAGGAAGAATA  
ATAGGAGATATAAGACAAGCACATTGTAACATTAGTAAAGCAGACTGGGAA  
AGAACTTTAAAAGGGTAGCTATAAAATTAAGAGAACAATTTGTGAATAAA  
ACAATAATCTTTAATCAATCCTCAGGAAGGGATCCAGAAATTGTAAAGCAC  
ACTTTTAATTGTGGAGGGGAATTTTTTTACTGTAATACAACACA

- **DF02.71 (nucleotídeos 6850 a 7390): 534 pb seqüenciados**

CCTTTGAGCCAATTCCTATACATTATTGTGCCCCAGCTGGTTTTGCGCTTCTC  
AAGTGTAAGATAAGAAGTTCAATGGAACAGGACCATGCACAAATGTCAGC  
ACAGTACAATGTACACATGGAATTAAGCCAGTAGTGTCAACCCAAGTCTG  
TTAAATGGCAGTCTAGCAGAAGAAGATGTAGTACTTAGATCTGCCAATTTCT  
CGGACAATACTAAAAGCATAATAGTACAGCTGAAGGAACCTGTAGTAATTA  
ATTGTACAAGACCCAACAACAATACAAGAAGAGGTATACATATAGGACCAG  
GGAGGCATTTTATACAACAGGCGAAATAACAGGAGATATAAGACAAGAAC  
ATTGTAACCTTAGTAGAAAAGATTGGAATAACACTTTAAAGGAAATAGTTA  
AAAAATTAGGAGAACAGTTTAAGAACAAAACAATAGTCTTTAACCAATCCT  
CAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAACTGTGGAGGGGAAT  
TTTTCTACTGTAGTACAACACA

- **DF02.72 (nucleotídeos 6831 a 7390): 554 pb seqüenciados**

CAGGCCTGTCCAAAGATATCCTTTGAACCAATTCCCATACATTATTGTGCCC  
 CGGCTGGTTTTGCGATTCTAAAGTGTAGAGATGAGAAATTCAATGGAACAG  
 GACCATGTACCAATGTCAGCACAGTACAATGCACACATGGAATTAGACCAG  
 TAGTGTCAACTCAACTGCTGTTAAATGGCAGTCTAGCAAAAGAAGAGATAG  
 TAATTAGATCTGAAAATCTCACGAACAATGCTAAAACCATAATAGTACAAC  
 TGAACAAATCTGTAGAAATTAATTGTATAAGACCCAACAACAATACAAGAA  
 GAAGCATACATATGGGATTTGGGAGAGCACTTTATACAAATGGAGAAATAG  
 TAGGAGACATCAGACAAGCACATTGTAACCTTAGTAGAACAGATTGGGGAG  
 ACATGTTAACACAGGTAGCTATAAAATTAAGAGAAAAATTTGAGAATAAAA  
 CAATAGTCTTTAATCAAACCTGAGGAGGGGACCCAGAAATTGTAATGCACA  
 CTTTTAATTGTGGAGGGGAATTTTTCTACTGTAAGACAACACA

- **DF02.73 (nucleotídeos 6831 a 7390): 554 pb seqüenciados**

CAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGTGCCC  
 CGGCTGGTTTTGCGCTTCTAAAATGTAAGGATAACAAATTCAATGGAACAG  
 GACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAG  
 TAGTATCAACTCAACTACTGTTAAATGGCAGTCTAGCAGAAGAAGAGATAG  
 TAGTTAGATCTGAAAATTTACAGACAATGCTAAAATCATAATAGTACAGC  
 TGAATAAACTGTAGAAATTAATTGTACAAGACCTAACACAATACAAGAA  
 AAGGTATACATATAGGACCAGGGAGAACATTTTATGCAACAGGAGACATAA  
 TAGGAAATATAAGACAAGCACATTGTAACCTTAGTAGAGCAGAATGGAATG  
 ACACTTTAAAAAAGGTAGCTGGAGAATTAAGAAAACAATTTAAGAATAAAA  
 CAATAATCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACA  
 GTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACA

- **DF02.75 (nucleotídeos 6834 a 7390): 551 pb seqüenciados**

GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGG  
 CTGGTTTTGCGATTCTAAAGTGTAAAGGATAAGACATTCAATGGAACAGGAC  
 CATGTCCAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAG  
 TGTC AACCCA ACTGCTGTTGAATGGCAGTTTAGCAGAAGGGGAGATAATAA  
 TTAGATCTGAAAATCTCACAAACAATGCCAAAACCATAATAGTGCATCTGA  
 ACGAGACTGTAAAATTAAT TGTATCAGACCCAACAACAATACAAGAAGA  
 AGTATACATATGGGACCGGGGAGAACATTTTATGCAGAAGGAGACATAATA  
 GGAGATATAAGACAAGCACATTGTAACCTTAATGCAACAGAATGGAACAAC  
 ACTTTGACACGGGTAGCAATAAAATTTAAAGAACTATTTAATAATAAAA  
 CAATAATCTTTAACCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCAC  
 AGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAA CACA

- **DF02.80 (nucleotídeos 6831 a 7362): 529 pb seqüenciados**

CAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCC  
 CGGCTGGTTTTGCGATTCTAAAGTGTAAATGATAAAAAGTTTAATGGAAGCG

GATCATGTACAAACGTCAGCACAGTACAATGTACACATGGAATTAGACCAG  
 TAGTGTCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAAGTAG  
 TAATTAGATCTGAAAATTTACGAATAATGCTAAAACCATAATAGTACAGC  
 TGAAGGAATCTGTAAAAATTAATTGTACAAGACCCAATAACAATACAAGAA  
 AAAGTATACATATGGGACCAGGGAAAGCATTTTATGCAACAGGAGATATAA  
 TAGGAAATATAAGACAAGCACATTGTAACATTAGTAGAAGAGACTGGAATA  
 ACACTCTAAGCCAGATAGCTGTAAAATTAAGGGAAAAGTTTGGGAACAACA  
 AAACAATAGTCTTTAATCAATCCTCAGGAGGGGATCCAGAAATTGTAATGC  
 ACAGTTTTAATTGTGGAG

- **DF02.81 (nucleotídeos 6832 a 7381): 541 pb seqüenciados**

AGGCCTGTCCAAAGGTATCTTTTGAACCAATTCCCATCCATTATTGTGCCCC  
 TGCTGGTTTTGCAATTCTAATATGTAATGATAAGAAATTCAATGGAACAGGA  
 CCATGTAAAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCGGTA  
 GTGTCAACTCAACTACTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTTAGTA  
 ATTAGATCTGAAAATCTCACGAACAATGCCAAAACCATAATAGTGCAGTTG  
 AATGAAACTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAGA  
 AGTATACACATGGGCTGGGGGAGAGCATTTTATGCAACAGGAGAAATAGTA  
 GGAGATATAAGACAAGCACATTGTAACCTTAGTGTAAACAAAATGGAACAAC  
 ACGTTAAAACACATAGCTAAAAAATTAAGAACAATTTAATAAAACAATA  
 ATCTTTAATAAATCCTCAGGAGGGGACCCAGAAGTTGTAATGCACACTTTTA  
 ATTGTGGAGGGGAATTTTTCTACTGTAA

- **DF02.83 (nucleotídeos 6821 a 7377 ): 548 pb seqüenciados**

AGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATACAT  
 TATTGTGCCCCGGCTGGATTTGCGCTTCTAAAATGTAAGGATAACAAATTCA  
 ATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAA  
 TTAGGCCAGTAGTATCAACTCAACTACTGTTAAATGGCAGCCTAGCAGAAG  
 AAGAGATAGTAATTAGATCTGAAAATTTACGAACAATGCTAAAACCATAA  
 TAGTACAGCTGAAGGAACCTGTAGCAATTAATTGCACAAGACCCGGCAACA  
 ATACAAGAAAAAGTATACCTCTAGGATGGGGGAGAGTGCTTTATGCAACAG  
 GAGATATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTAGAACAG  
 AATGGAATAACACTTTAGAACAGGTAGCTGGAAAATTAAGAACAATTTA  
 ATAAAACAATAACCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAA  
 TGCACACTTTTAATTGTGGAGGGGAATTTTTCTACT

- **DF02.86 (nucleotídeos 6832 a 7367): 533 pb seqüenciados**

AGGCCTGTCCAAAGATATCCTTTGAACCAATTCCCATACATTTTTGTGCCCC  
 GGCTGGTTTTGCGATTCTAAGGTGCAATGATAAAAAGTTTAATGGATCAGG  
 ACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAAACCAGT  
 AGTGTCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAGGAAGAGGTAAT  
 AATTAGATCTGAAAACATCTCAAACAATGCTAAAACCATAATAGTACAGTT  
 GAACGAAACTGTACAAATTAATTGTACAAGACCCGGCAACAACAATAAA  
 GAGTATACATATGGCATTAGGGAGAGCATTATATGCAACAGAAAGAATAGT  
 AGGAAATATAAGAAATGCACATTGTAACATTAGTGAGGGAGAATGGGAAA

AAACTTTACACAAGGTAGTTGTA AAAATTAAGAGAACAATTCGAGGAAAACA  
GCACAATAGTCTTTACACAACCCTCAGGAGGGGACCCAGAAATTATGTCAC  
ACAGTTTTAATTGTGGAGGGGAA

- **DF02.88 (nucleotídeos 6829 a 7366): 535 pb seqüenciados**

CACAGGCCTGTCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGTGC  
CCCGGCTGGTTTTGCGATTCTAAAGTGTAACAATAAGACATTCAATGGAAC  
AGGACCATGTAACAACGTCAGCACAGTACACTGTACACATGGAATTAAGCC  
AGTAGTGTCAACTCAACTGTTGCTAAATGGCAGTTTAGCAGAAGAAGAGAT  
AATAATTAGATCTGAAAATTTACAGACAATGCTAAAACCATAATAGTACA  
GCTGAACAGA ACTGTAACAATTAATTGTACAAGACCCAACAACAATACAAG  
AAAAAGTATACATATAGGATTTGGGAGAGCACTTTATGCAACAGGAGCAAT  
ACAAGGAGATATAAGACAAGCACATTGTAATCTTAGTGAATCAGAGTGGGG  
AGAACTTTAAAACAGATAGCTATAAAAATTAAGAGAAAAATTTGACAATAA  
TGCATCAATAGTCTTTAAGCAATTCTCAGGAGGGGACCCAGAAATTGTAAC  
GCACAGTTTTAATTGTGGAGGGGA

- **DF02.91 (nucleotídeos 6841 a 7293): 448 pb seqüenciados**

CAAAGATATCCTTTGAGCCAATTTCCATACATTATTGTGCCCTAGCTGGTTTT  
GCGATTCTAAAGTGTAAGATACAAAGTTCAACGGAACCGGGACCTGTAAC  
AATGTCAGCACAGTACAATGTACACATGGCATCAAGCCAGTAGTGTCAACT  
CAACTGCTGTAAATGGCAGTCTAGCAGAAGAAGTAGTAATTAGATCTGAA  
AATTCACAAACAATGCTAAAACATAATAGTACAGCTGAAGGACCTGTGA  
GAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATAAATATA  
GGGCCAGGGAGAGCATT TTTATGCAACAGGAGATATAATAGGAGATATAAG  
ACAGGCACATTGTAACATTAGTACAGCAAAATGGAACAACACTTTAAAGCA  
AATAGTTGGGCAAATTAAGAAGACAGTTTGGGAACAATA

- **DF02.94 (nucleotídeos 6830 a 7361): 520 pb seqüenciados**

CTGTCCAAAGGTATCCTTTGAGCCAATTACCATACATTATTGTGCCCCAGCT  
GGTTTTGCGATTCTAAAGTGTAATGATAAAAATTTCAATGGAACAGGACCA  
TGTCACAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTG  
TCAACTCAACTGCTGTAAACGGCAGTCTAGCAGAAGAAGAGGTAGTAATT  
AAATCTGAAAATCTCACAACAATGCAAAAAGCATAATAGTACAGCTGAAC  
CAAGCTGTAAAATTA ACTGTACAAGACCCAACAACAATACAAGGAAAAGT  
ATACATATAGCACCAGGGGGCGCATTTTATGCAACAGGAGAAATAATAGGA  
GATATAAGACAAGCACATTGTAACATTAGTAGAACAGACTGGAATAACT  
TTAAAGAATATAAGTGAAAAGTTAAGAGAAAAGTTTGGGAACAAAACAAT  
AATCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTT  
AATTGTGGA

- **DF02.95 (nucleotídeos 6875 a 7342): 465 pb seqüenciados**

TTGTGCCCCAGCTGGTTTTGCGATTCTAAAGTGTAATGATAAAAAGTTCAAT  
GGAACAGGACCATGTAGAAATGTCAGCACAGTACAATGTACACATGGGATT  
AGGCCAGTAGTGTCAACTCAACTGCTGTAAATGGCAGTCTAGCAGAAGAA  
GAGATAGTAATTAGATCTGAAAATTTTCAGTAACAATGCTAAAACCATAATA  
GTACAGCTGAACGAATCTGTAGCAATTAATTGTACAAGACCCAACAACACT  
ATAAGACGAAGTATCCGTATGGGATGGGGGAGAACATTTTATGCAACAGGA  
CAGATAAAAGGAAGTATAAGACAAGCACATTGTAACCTTAGTAGAGCAGA  
ATGGGAAAAAACCTTAAATCGGATAAGTATTAATTAAGAACAATTTAA  
GAACACTACAAAATTATCTTTAATCAATCCTCAGGAGGGGACCCAGAAAT  
TGTAAC

- **DF02.99 (nucleotídeos 6875 a 7361): 481 pb seqüenciados**

TTGCGCCCCAGCTGGTTTTGCGATTCTAAAATGTAACAATGAGACATTTAAA  
GGAAAAGGACCATGTAAAAATGTCAGCGTAGTACAATGTACACATGGAATT  
AGGCCAGTAGTGTCAACTCAACTGCTGTAAATGGCAGCCTAGCAGAAAAG  
GAGATAGCAATTAGGTCTGAAAATTTTCACGAATAATGCTAAAAACATAATA  
GTACAGCTGAAAGAACCTGTAAAAATTAATTGTACAAGACCCAACAACAAT  
ACAAGGAAAAGTATACCTATGGGACCAGGGCATGCATTTTCATGCAACAGGG  
GATATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTGCAGCACAA  
TGGAATAACACTTTAAAGCAGATAGTTGGAAAATTAAGAGAACAGTTTAAA  
AACAAAACAATAGTCTTTAAGCAATCCTCAGGAGGGGACTCAGAAATTGTA  
ATGCACAGTTTTAATTGTGGA

- **DF02.102 (nucleotídeos 6994 a 7375): 377 pb seqüenciados**

CTCAATTACGGTTAAATGGCAGTCTAGCAGAGGAAGAGGTAATAATTAGAT  
CTGAAAATTTACAAACAATGCTAAAACCATAATAGTACAGCTGAATGAAA  
CTGTACAAATTAATTGTACCAGACCCAACAACAATACAAGAAAAAGAATAA  
ATATAGGACCAGGGAGAGCATTTTTAGCAACAGGAGCAATAATAGGAGATA  
TAAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAACAACACTTTAA  
AACAGATAGTTATGAAATTAAGAGAACAATTTAGTAATTAACAATAATC  
TTTAAGGAATCCTCAGGAGGGGACCCAGAAATTGAAACGCACAGTTTTAAT  
TGTGGAGGGGAATTTTTCTC