

**FIOCRUZ**

**FUNDAÇÃO OSWALDO CRUZ  
CENTRO DE PESQUISAS GONÇALO MONIZ**

**CURSO DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA EM  
SAÚDE E MEDICINA INVESTIGATIVA**

**TESE DE DOUTORADO**

**DIVERSIDADE MOLECULAR DO VÍRUS DA  
IMUNODEFICIÊNCIA HUMANA TIPO 1 NO ESTADO DA  
BAHIA**

**ADRIANO FERNANDO DA SILVA ARAÚJO**

**Salvador - Bahia - Brasil  
2010**



FIOCRUZ

**FUNDAÇÃO OSWALDO CRUZ  
CENTRO DE PESQUISAS GONÇALO MONIZ**

**CURSO DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA EM  
SAÚDE E MEDICINA INVESTIGATIVA**

**ADRIANO FERNANDO DA SILVA ARAÚJO**

**DIVERSIDADE MOLECULAR DO VÍRUS DA  
IMUNODEFICIENCIA HUMANA TIPO 1 NO ESTADO DA  
BAHIA**

Tese apresentada ao Centro de Pesquisas  
Gonçalo Moniz como parte dos requisitos  
para obtenção do título de Doutor em  
Biotecnologia em Saúde e Medicina  
Investigativa

ORIENTADOR: PROF. DR. LUIZ CARLOS JÚNIOR ALCÂNTARA  
CO-ORIENTADOR: PROF. DR. BERNARDO GALVÃO CASTRO-FILHO

**Salvador - Bahia - Brasil  
2010**

Ficha Catalográfica elaborada pela Biblioteca do  
Centro de Pesquisas Gonçalo Moniz / FIOCRUZ - Salvador - Bahia.

Araújo, Adriano Fernando da Silva.

A663d

**Diversidade molecular do vírus da imunodeficiência humana tipo 1 no estado da Bahia.**

Adriano Fernando da Silva Araújo, Salvador, 2010. 76 f.: il.; 30 cm.

Tese (doutorado) – Fundação Oswaldo Cruz, Centro de Pesquisas Gonçalo Moniz. Programa de Pós-Graduação em Biotecnologia em Saúde e Medicina Investigativa.

Orientador: Prof. Dr. Luiz Carlos Júnior Alcântara.

Laboratório de Avançado de Saúde Pública.

1. HIV-1. 2. Variação genética. 3. Epidemiologia molecular. I. Título.

CDU 578.53

**ADRIANO FERNANDO DA SILVA ARAÚJO**

**DIVERSIDADE MOLECULAR DO VÍRUS DA IMUNODEFICIENCIA  
HUMANA TIPO 1 NO ESTADO DA BAHIA**

**Tese apresentada ao Centro de Pesquisas  
Gonçalo Moniz como parte dos requisitos  
para obtenção do título de Doutor em  
Biotecnologia em Saúde e Medicina  
Investigativa**

**Aprovada pela Banca Examinadora em 05 de abril de 2010.**

**BANCA EXAMINADORA:**

---

**Prof. Dr. Luiz Carlos Junior Alcântara**  
**Orientador**

---

**Prof. Dr. Bernardo Galvão Castro-Filho**  
**Co-orientador**

---

**Prof. Dr. José Carlos Couto Fernandez**

---

**Profa. Dra. Maria Lourdes Farre Vallve**

---

**Profa. Dra. Flora Maria de Campos Fernandes ( Suplente )**

## **AGRADECIMENTOS**

A Deus pelo seu grande amor e graças dispensadas.

Aos meus pais e avós.

As minhas irmãs, cunhados, sobrinhos e amigos.

Ao Dr. Luiz Carlos Junior Alcântara pela amizade e orientação com ensinamentos essenciais, sem os quais não seria possível essa tese.

Ao Dr. Bernardo Galvão Castro-Filho pela amizade, co-orientação e constante disponibilidade em minhas necessidades e dúvidas.

A Dra. Joana Paixão Monteiro pela ajuda nas análises das amostras recombinantes e revisão dos manuscritos.

Ao Dr. Carlos Brites da Faculdade de Medicina da UFBA – LAPIVIR, pelo envio das amostras e dados clínicos dos pacientes de Salvador.

Ao Dr. Edson de Souza Santos do Centro de Referência DST/AIDS da Secretaria Municipal de Saúde de Feira de Santana pelo envio das amostras e dados clínicos.

Aos amigos do laboratório do Centro de HTLV da Escola Bahiana de Medicina Filipe Rego, Luciane Santos, Viviana Olavarria, Leandro Inácio, Gisele de Souza, Diego Frias, Thessika Araújo, Aline Miranda, Marco Mello e Sandra Gadelha Mello pela amizade e ajuda durante os experimentos de Biologia Molecular e Bioinformática.

A todos os amigos do Laboratório Avançado de Saúde Pública do Centro de Pesquisas Gonçalo Moniz.

A todos os amigos do Centro de Pesquisas Gonçalo Moniz.

A Elizabeth Vasconcelos, Eugênia Góes, Cláudio Paulo, Rodrigo Gaspar e Sônia pela ajuda e amizade constantes na administração, secretaria e atendimento.

A Taise Coutinho Caires e Sonia Santos de Oliveira pela prestatividade na secretaria da Coordenação de Ensino de pós-graduação.

Aos pacientes que aceitaram participar destas pesquisas.

***Nossas palavras devem ser ricas em sabedoria, isto é,  
inspiradas pela verdade, bondade e beleza.***

***Bento XVI***

## RESUMO

A análise de sequências genômicas é uma forma poderosa para se estudar a diversidade do HIV-1 e sua expansão nas populações, além de ser crucial para dar suporte à vigilância epidemiológica e ao desenvolvimento de terapias e vacinas eficazes. Neste estudo foram geradas sequências dos genes *gag* e *env* do HIV-1 de 61 pacientes infectados em Salvador e do *pol* de 58 pacientes de Feira de Santana, Bahia, Brasil. A bioinformática foi utilizada para subtipar, genotipar as mutações de resistência, determinar a pressão seletiva e prever o uso de coreceptor. Em Salvador, 92,6% das amostras foram do subtipo B e 7,4% foram recombinantes BF1. Em Feira de Santana, 67,2% foram do subtipo B, 6,9% F1, 1,7% C e 24,1% BF1. Onze (19,0%) destes isolados apresentaram mutações de resistência. Pacientes infectados pelo B tinham em média 0,4 mutações de resistência e nenhuma mutação deste tipo foi observada em BF1. Com base na caracterização da alça V3 do gene *env* em 43 amostras de Salvador, foram encontradas 18,2% de variantes brasileiras (B'-GWGR), 46,5% de GPGR e 34,9% de GXGX. O tempo médio de diagnóstico positivo foi de 13 anos para o subtipo B' e 9 anos para o subtipo B (GPGR + GXGX). Setenta e seis por cento destes vírus fazem uso do coreceptor CCR5, enquanto 24% foram classificados como capazes de usar o CXCR4. Encontramos uma associação entre o subtipo B' e um maior tempo em anos de diagnóstico positivo para HIV-1. As prevalências de subtipo B' e de recombinante B/F1 em Salvador foram menores do que as encontradas em estudos anteriores na Bahia e no Brasil, por outro lado, em Feira de Santana foi encontrada uma alta prevalência de B/F1.

## ABSTRACT

Genomic analysis of sequences is a powerful tool to perform studies in HIV-1 diversity and expansion in the populations and it is crucial to give support to epidemiological surveillance, new therapies and vaccines development. In this study, sequences were generated from *gag* and *env* genes from 61 HIV-1 infected patients sampled in Salvador and from *pol* gene in 58 infected patients sampled in Feira de Santana, Bahia, Brazil. Bioinformatic tools were used to subtype, genotype the resistance mutations, determine the selective pressure and to predict coreceptor use. In Salvador, 92.6% of the samples were subtype B and 7.4% were recombinant BF1. In Feira de Santana, 67.2% of the samples were subtype B, 6.9% were F1, 1.7% was C and 24,1% were recombinant BF1. Eleven (19.0%) of these isolates presented resistance mutations. HIV-1 subtype B infected individuals had, in average, 0.4 resistance mutations by sequence and no mutation was observed in BF1. With regard to V3 loop of gene *env*, it was found 18.2% of Brazilian variants (B'-GWGR), 46.5% of GPGR and 34.9% of GXGX. The *env* sequences showed through to negative pressure. The time, in average, since the diagnosis was 13 years among subtype B' and 9 years among subtype B isolates. Seventy-six percent of these viruses use the CCR5 coreceptor, while 24% were able to use CXCR4. We have found an association between subtype B' and a longer period of time since the HIV positive diagnosis. The prevalence of B' and B/F1 in Salvador were smaller than the found in previous studies in Bahia state and in Brazil, on the other hand, in Feira de Santana it was found a higher prevalence of B/F1.

## LISTA DE FIGURAS E TABELAS

<b>Figura</b>		<b>Página</b>
01	Evolução clínica da infecção pelo HIV-1	9
02	Estrutura esquemática do HIV-1	16
03	Organização genômica do HIV-1	17
04	Estrutura tridimensional da transcriptase reversa	19
05	Ciclo de replicação do HIV-1	22
06	Relações filogenéticas entre HIV e SIV	29
07	Distribuição global dos subtipos, CRFs e URFs do HIV-1.	32
<b>Tabela</b>		
01	Sistema de classificação para adolescentes e adultos infectados pelo HIV	10
02	Principais drogas antiretrovirais aprovadas pelo FDA	39

## LISTA DE ABREVIATURAS

aa	Aminoácido
AIDS	<i>Acquired Immune Deficiency Syndrome</i> (Síndrome da Imunodeficiência Humana Adquirida)
ARV	<i>AIDS-associated retrovirus</i> (Retrovírus Associado a Aids)
ARV	<i>Antiretroviral</i> (Antirretroviral)
AZT	Azidotimidina ou zidovudina
BE	Boletim epidemiológico
CCR5	<i>C-C chemokine receptor type 5</i> (receptor com motivo C-C de quimiocina do tipo 5)
C	Celsius
CXCR4	<i>C-X-C chemokine receptor type 4</i> (receptor com motivo C-X-C de quimiocina do tipo 4)
CD4	<i>Cluster of differentiation 4</i> (agrupamento de diferenciação tipo 4)
CD8	<i>Cluster of differentiation 8</i> (agrupamento de diferenciação tipo 8)
CDC	Centers for Disease Control and Prevention (Centro para o Controle e Prevenção de Doenças)
CN	Coordenação nacional
CRF	<i>Circulation Recombinant Form</i> (forma recombinante circulante)
DNA	<i>Deoxyribonucleic acid</i> (ácido desoxirribonucleico)
dNTP	<i>Deoxy-nucleotide triphosphate</i> (desoxinucleotídeos trifosfatado)
DST	Doenças sexualmente transmissíveis
FDA	<i>Food and Drug Administration</i> (Agência norte-americana de administração de drogas e alimentos)
GTR	<i>General Time Reversible</i> (tempo geral reversível)
HAART	<i>Highly Active AntiRetroviral Treatment</i> (terapia antirretroviral de alta potência)
HIV	<i>Human Immunodeficiency Virus</i> (Vírus da Imunodeficiência Humana)
HTLV	<i>Human T lymphotropic virus</i> (Vírus T-linfotrópico humano)
INNTR	Inibidor não-nucleotídeo da transcriptase reversa
INTR	Inibidor nucleotídeo da transcriptase reversa
IN	Integrase

IP	Inibidor da protease
LAV	<i>Lymphadenopathy Associated Virus</i> (Vírus Associado a Linfadenopatia)
LGP	Linfadenopatia generalizada persistente
LRT	<i>Likelihood Ratio Test</i> (Teste de Razão de Verossimilhança)
LTR	<i>Long terminal repetitive</i> (longos terminais de repetição)
ML	Maximal likelihood (máxima verossimilhança)
MS	Ministério da Saúde
<i>nef</i>	<i>Negative factor</i> (fator negativo)
NJ	<i>Neighbor joining</i> (agrupamento de vizinhos)
NVP	Nevirapina
pb	Pares de base
PCR	<i>Polimerase chain reaction</i> (reação em cadeia da polimerase)
PDB	<i>Protein Data Bank</i> (Banco de dados de proteínas)
PR	Protease
rev	<i>Regulatory factor for expression</i> (Fator regulador da expressão)
RNA	<i>Ribonucleic acid</i> (ácido ribonucleico)
SIV	<i>Simian Immunodeficiency Virus</i> (vírus da imunodeficiência símia)
<i>tat</i>	Transactivator (Transativador)
Th1	T Helper cell type 1 (célula T auxiliar tipo 1)
Th2	T Helper cells type 2 (célula T auxiliar tipo 2)
TR	Transcriptase reversa
UDI	Uso de drogas intravenosas
Unaid	<i>Joint United Nations Program on HIV/AIDS</i> (Programa conjunto das Nações Unidas sobre HIV/AIDS)
URF	<i>Unique Recombinant form</i> (forma única recombinante)
V3	Região variável da gp120 número 3
<i>vif</i>	<i>Viral infectivity factor</i> (fator da infectividade viral)
<i>vpr</i>	<i>Viral protein R</i> (proteína R Viral)
<i>vpu</i>	<i>Viral protein U</i> (proteína U viral)
<i>vpx</i>	<i>Viral protein X</i> (homólogo ao <i>vpr</i> , encontrado no HIV-2)

## SUMÁRIO

	Página
<b>INTRODUÇÃO</b>	12
A Pandemia do HIV-1/AIDS	13
A Epidemiologia do HIV-1/AIDS no Mundo e no Brasil	15
A Epidemiologia do HIV-1/AIDS no Estado da Bahia	17
A Patologia da AIDS	19
A Estrutura Molecular do HIV-1	25
O Ciclo do HIV-1	30
A Geração da Diversidade no HIV-1	33
A Origem e a Classificação do HIV-1	36
A Epidemiologia Molecular do HIV-1	40
<i>A Epidemiologia Molecular do HIV-1 no Mundo</i>	40
<i>A Epidemiologia Molecular do HIV-1 no Brasil e na Bahia</i>	44
As Mutações de Resistência aos Antiretrovirais	46
Análises Moleculares por Ferramentas de Bioinformática	51
<i>Alinhamento e Edição de sequências de DNA e Proteínas</i>	53
<i>Escolha do Melhor Modelo de Substituição Nucleotídica</i>	54
<i>Análises Filogenéticas</i>	56
<i>Teste Estatístico para Confiabilidade das Árvores</i>	59
<i>Cálculo da dN/dS, predição de uso de coreceptor e de mutações de resistência</i>	60
<b>JUSTIFICATIVA</b>	63
<b>OBJETIVOS</b>	67
<b>RESUMO DOS RESULTADOS</b>	69
<b>DISCUSSÃO DOS MANUSCRITOS</b>	76
<b>CONCLUSÕES</b>	85
<b>REFERENCIAS BIBLIOGRÁFICAS</b>	87
<b>MANUSCRITOS</b>	117

## ***INTRODUÇÃO***

## **A Pandemia do HIV-1/AIDS**

A Síndrome da Imunodeficiência Adquirida, AIDS, (sigla do inglês *Acquired Immune Deficiency Syndrome*) manifestada após a infecção do organismo humano pelo Vírus da Imunodeficiência Humana, o HIV (sigla do inglês *Human Immunodeficiency Virus*) é uma grave doença e até hoje sem cura estéril que já chegou a acometer cerca de 1 % da população humana (UNAIDS, 2006). Surgiu em meados de 1981 com elevação de sua incidência nos Estados Unidos da América a partir da identificação de um número elevado de homossexuais, moradores da cidade de São Francisco e de Nova Iorque, que apresentavam sintomas de "sarcoma de Kaposi", pneumonia por *Pneumocystis carinii*, candidíase e outras doenças oportunistas (CDC, 1982). Estudos epidemiológicos posteriores identificaram que a transmissão poderia ocorrer por contato sexual, por administração de drogas intravenosas, por transfusão sanguínea, por transmissão vertical mãe – filho e por doação de órgãos. A evolução para o quadro da AIDS estava associada a uma queda acentuada do número de linfócitos T CD4<sup>+</sup>, o que levava ao comprometimento do sistema imune. Este aspecto, inicialmente, levou o grupo liderado pelo doutor Robert Gallo em meados de 1983 à hipótese de que esta nova doença, ainda não classificada, era causada por um retrovírus mutante muito similar ao HTLV (GALLO *et al.*, 1983), que foi logo nomeado por eles de HTLV-III (POPOVIC *et al.*, 1984). O grupo do doutor Montaignier, também no ano de 1983, isolou um novo vírus de um paciente com linfadenopatia e concluíram que o agente causador da AIDS era um retrovírus, ainda não conhecido,

pertencente à família do HTLV (BARRÉ-SINOUSI *et al.*, 1983) que foi inicialmente nomeado por vírus linfadenopático da AIDS - LAV. De forma independente, o grupo do doutor Levy em 1984 também conseguiu isolar um retrovírus associado a AIDS por eles nomeado de retrovírus associado a AIDS - ARV. Posteriormente, (COFFIN *et al.*, 1986) renomearam o novo retrovírus para a designação atual de HIV.

No Brasil, o primeiro caso foi reconhecido retrospectivamente no Estado de São Paulo, como tendo ocorrido em 1980 (CN-DST/AIDS, 1998). O padrão epidemiológico no Brasil seguiu a transmissão mundial onde estava inicialmente concentrado em homossexuais e numa segunda fase entre os heterossexuais. (BONGERTZ *et al.*, 2000; DOURADO *et al.*, 2007).

Em 28 anos de combate a pandemia da AIDS, apesar de inúmeros avanços, campanhas e recursos aplicados em todo o mundo, ainda não se descobriu uma vacina eficaz, nem uma terapêutica definitiva para sua cura. Apesar da pequena queda em todo o mundo na incidência, o número de pessoas convivendo com o vírus continua crescendo, ainda que em menor avanço que no início da pandemia. Este aumento na prevalência tem-se dado principalmente entre as mulheres, na população de baixa renda e nas pequenas cidades (UNAIDS, 2009; BE - AIDS/DST, 2009).

## **A Epidemiologia do HIV-1/AIDS no Mundo e no Brasil**

Atualmente, estima-se que 33,4 milhões de pessoas convivam com o HIV, uma prevalência de 0,8 (UNAIDS, 2009), sendo 15,7 milhões de mulheres e 2,1 milhões de crianças com menos de 13 anos. Foram notificados 2,7 milhões de novos casos, sendo 16% em crianças. Foram registrados 2,1 milhões de mortes relacionadas ao vírus dos quais 14% em crianças.

No Brasil, de 1980 até junho de 2009, foram registrados 544.846 casos de AIDS, 356.419 (65,4%) do sexo masculino e 188.407 (34,6%) do sexo feminino, além de 217.091 mortes provocadas pela doença. A taxa de transmissão vertical está em torno de 4% após o parto de mulheres infectadas e estima-se que em menores de cinco anos a incidência ocorra numa taxa de 3,9 por 100 mil habitantes. Por ano, são notificados entre 33 mil e 35 mil novos casos. As regiões Norte e Nordeste possuem as taxas de incidência mais baixas do Brasil, no entanto as únicas em crescimento, em torno de 10 e 5 novos casos por ano por 100.000 habitantes respectivamente. A taxa de prevalência da infecção no país é de 0,6%. A região sudeste possui o maior percentual (59,3%), com 323.069 registros da doença. O Sul concentra 19,2% dos casos, com 104.671 notificações; o Nordeste registra 11,9%, com 64.706; o Centro-Oeste 5,7%, com 31.011; e o Norte 3,9%, com 21.389. Em todo o país há uma estimativa de que 630 mil pessoas estejam infectadas pelo HIV-1. A via de transmissão mais importante, e crescente atualmente, é a heterossexual, seguida da homossexual e da ocorrida por uso de drogas

injetáveis, estas duas últimas ocorrendo em declínio gradativo (BE – AIDS/DST, 2009). Notou-se também o crescimento persistente da proporção de óbitos por AIDS nas categorias socialmente menos favorecidas, em ambos os sexos, entre 1998 e 2004 (BE – AIDS/DST, 2007), expondo a dificuldade no acesso aos serviços de saúde para diagnóstico e tratamento precoces e também a hipótese de uma maior susceptibilidade a infecção. Em geral, tem-se observado uma mudança da dinâmica epidemiológica nos últimos anos e estudos mais recentes têm apontado para uma maior interiorização, heterossexualização, femininização, envelhecimento e pauperização da doença (PARKER *et al.*, 1996; SZWARCOWALD *et al.*, 2000; BASTOS *et al.*, 2000; CASTILHO *et al.*, 2000; RIOS 2003; EYER-SILVA *et al.*, 2006; SÁ *et al.*, 2007).

## **A Epidemiologia do HIV-1/AIDS no Estado da Bahia**

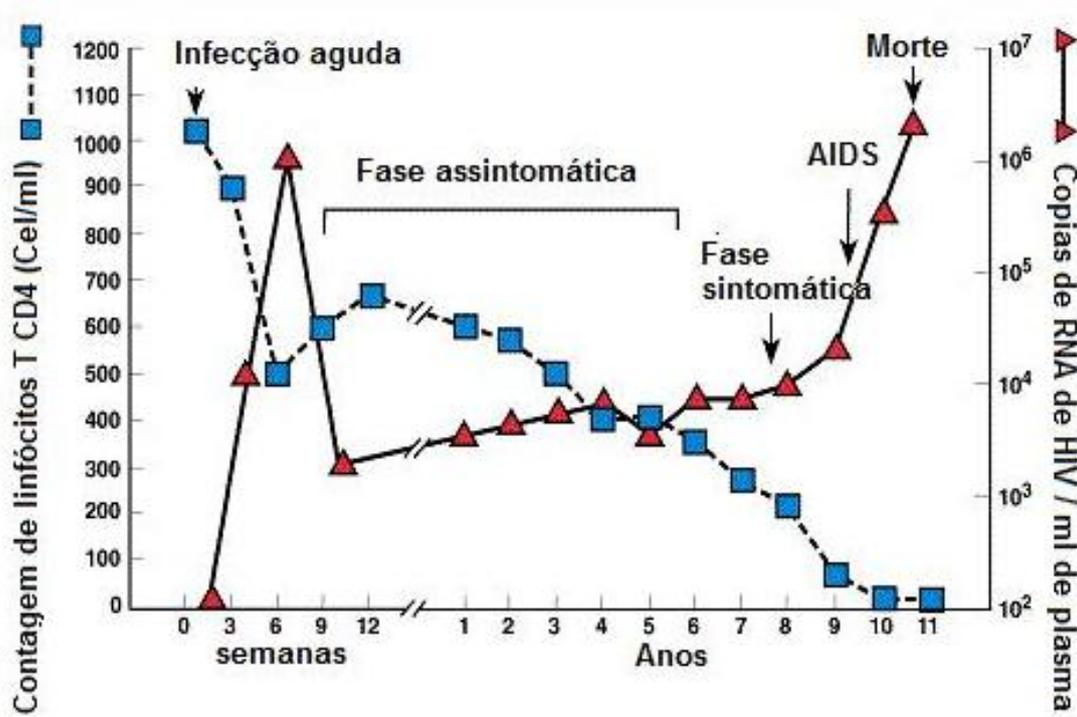
A Bahia é o maior estado da região Nordeste do Brasil com aproximadamente 14 milhões de habitantes, 7,6% da população brasileira. A capital do estado, Salvador, possui 2,9 milhões de habitantes, e Feira de Santana é a segunda maior cidade do estado em número de habitantes, 572 mil (IBGE, 2007). Salvador é o primeiro em produto interno bruto do Nordeste e possui um turismo bem desenvolvido sendo o quinto destino mais visitado pelo turismo de lazer internacional no Brasil (MT, 2007). Enquanto Feira de Santana centraliza uma das maiores redes rodoviária do interior do Nordeste. Desde o registro do primeiro caso de AIDS no Brasil em 1984 até junho de 2009 foram notificados 10.810 casos de AIDS, sendo 3.596 mulheres. Deste total, 1.452 casos ocorreram em adultos e 358 em crianças menores de 13 anos (BE-SUVISA/DIVPEP, 2009). A prevalência de pessoas infectadas pelo HIV-1 na Bahia foi estimada em 0,8% para homens e 0,36% para mulheres, num total de 0,55% (DOURADO *et al.*, 2007). A cidade de Salvador detém 51 % do total de casos do estado (5.341). Há 10 anos o município de Salvador concentrava 63,4 % dos casos notificados de AIDS na Bahia e 209 municípios apresentavam pelo menos um caso de AIDS, representando 50 % dos municípios. Atualmente 338 municípios do Estado, apresentam pelo menos um caso de AIDS, representando 81 % dos municípios do estado, o que denota a capacidade da epidemia de se disseminar para o interior. No município de Feira de Santana, 791 pessoas já foram notificadas no Centro

de Referência em DST/HIV/AIDS deste município, sendo que destas 261 são mulheres e crianças.

## **A Patologia da AIDS**

A AIDS se caracteriza por um grupo de sinais e sintomas de infecções oportunistas, decorrente da incapacidade de proteção do sistema imunológico, adquirida, e não herdada, no paciente pela infecção do HIV. O estágio da doença pelo HIV e os sistemas de classificação são ferramentas críticas para o acompanhamento e monitoramento da epidemia do HIV e para fornecer, aos clínicos e pacientes, informações importantes sobre o estágio da doença pelo HIV e o tratamento clínico adequado. Existem dois grandes sistemas de classificação atualmente em uso: o dos EUA desenvolvido pelo *Centers for Disease Control and Prevention* (CDC) e o sistema de classificação da Organização Mundial da Saúde (OMS) que faz parte do Sistema de Estágio Clínico e Classificação de Doenças.

Em geral a evolução clínica é classificada em quatro estágios ou fases: aguda; assintomática; sintomática inicial; e AIDS (figura 1). A fase aguda ocorre entre 5 e 30 dias após a infecção e caracterizando-se por viremia elevada, aumento de linfócitos T CD8+ e os sintomas de um quadro gripal que podem durar até 14 dias. A fase assintomática dura em mais ou menos 8 anos dependendo da progressão clínica e apresenta um quadro clínico inexistente ou mínimo, que pode ser interrompido por doenças oportunistas ou neoplasias. Na fase sintomática inicial o paciente apresenta os primeiros sinais ou sintomas principalmente sudorese noturna, fadiga e emagrecimento acompanhado de doenças oportunistas tais como: candidíase, gengivite, úlceras aftosas, doença inflamatória pélvica, neuropatia



**Figura 1:** Evolução clínica da infecção pelo HIV-1. Adaptado de Pantaleo *et al.*, 1993.

periférica, diarreia por mais de um mês, sinusites, febres acima de 38,5° C e herpes simples ou Zóster. Na fase de AIDS ocorrem instalações das infecções oportunistas com maior gravidade e agressividade, as principais infecções são: as virais citomegalovirose e herpes simples; as bacterianas tuberculose, pneumonias (recorrente, 2 episódios em um ano) e salmonelose; as fúngicas pneumocistose, candidíase, criptococose, histoplasmose; as parasitárias toxoplasmose, criptosporidiose e isosporíase. Ocorrem ainda carcinoma cervical e neoplasias de sarcoma de Kaposi, linfomas não-Hodgkin, intra-epiteliais anal e cervical. Emagrecimento (mais

que 10% do peso) associado a diarréia crônica e febre (mais de um mês).

**Tabela 01. Sistema de classificação para Adolescentes e Adultos Infectados pelo HIV**

Categorias de células T CD4	Categorias clínicas		
	A	B	C
	Fase assintomática, aguda ou LGP	Sintomáticos ou fase sintomática inicial	Sintomas da fase de AIDS
(1) $\geq 500$ células/ $\mu$ L	A1	B1	B1
(2) 200-499 células/ $\mu$ L	A2	B2	C2
(3) $< 200$ células/ $\mu$ L	A3	B3	C3

LGP = Linfadenopatia generalizada persistente. O destaque em fundo cinza mostra a classificação para a AIDS. Adaptado de <http://aidsetc.org>.

É importante ressaltar que o câncer de

colo do útero compõe o elenco de doenças que pontuam a definição de caso de AIDS na mulher. A sobrevivida, em geral, pode atingir 10 anos sem tratamento com antirretroviral (ARV). A morte por AIDS ocorre devido às infecções oportunistas que normalmente não causariam doença em um paciente saudável (THULER *et al.*, 1998).

O sistema de classificação do CDC baseia-se na gravidade da doença pelo HIV através da contagem de células T CD4 e pela presença de condições relacionadas ao HIV. A definição de AIDS inclui todos os indivíduos infectados com contagens de células T CD4 menor que 200 células / microlitro de sangue (ou porcentagem de T CD4, em linfócitos totais, menor que 14%), bem como certas condições e sintomas relacionadas ao HIV

(tabela 01). Embora os detalhes do sistema de classificação do CDC raramente sejam utilizados na rotina clínica, um conhecimento prático destes critérios (em particular, a definição de AIDS) é útil na assistência ao paciente, além disso, o sistema do CDC é usado em pesquisa clínica e epidemiológica.

Por outro lado, o Sistema de Estágio Clínico e Classificação de Doenças da OMS (HARE, 2006), pode ser facilmente usado em ambientes de recursos limitados, sem acesso às medições de contagem de células T CD4 e outros métodos de diagnóstico e testes de laboratório. O sistema da OMS classifica o infectado pelo HIV com base nas manifestações clínicas que podem ser reconhecidas e tratadas por clínicos em diversos contextos, incluindo com recursos limitados, especialmente aqueles ambientes em que a contagem de linfócitos T CD4 não está disponível, e por clínicos com diferentes níveis de especialização e formação. Esta classificação é utilizada em muitos países para determinar a viabilidade da terapia antirretroviral. Os estágios clínicos são classificados de 1 a 4 de acordo com a progressão da infecção de primária para avançada. Estes estágios são definidos por condições específicas ou sintomas clínicos. Para os propósitos do sistema da OMS, adolescentes e adultos são definidos como indivíduos com idade maior ou igual a 15 anos. A Infecção primária se caracteriza pela fase assintomática e aguda da doença; o estágio clínico 1 se caracteriza por ser assintomático e por LGP; no estágio clínico 2 se encontram alguns sinais da fase sintomática inicial, moderada perda de peso inexplicada (menor que 10% da massa corporal), infecções respiratórias recorrentes (sinusite, amigdalite, otite média e faringite), herpes Zoster, queilite angular, ulceração oral recorrente, erupções papulares pruriginosas, dermatite seborréica e infecções fúngicas

nas unhas; no estágio clínico 3 se encontram sintomas da fase sintomática inicial e alguns sintomas típicos de AIDS, perda de peso inexplicável grave (mais que 10% da massa corporal), diarreia crônica por mais de um mês, febre persistente por mais de um mês (mais que 37,6° C intermitente ou constante), candidíase oral persistente (popularmente conhecido como sapinho), leucoplasia pilosa oral, tuberculose pulmonar, infecções bacterianas graves (pneumonia, empiema, piomiosite, infecção óssea ou articular, meningite bacteriana), estomatite ulcerativa necrosante aguda, gengivite ou periodontite, anemia (hemoglobina menor que 8 g / dL), neutropenia (neutrófilos menor que 500 células / mL), trombocitopenia crônica (contagem de plaquetas menor que 50.000 células / mL); o estágio clínico 4 inclui todas os sintomas graves típicos da AIDS, síndrome de emagrecimento geral tal como definido pelo CDC, pneumonia por *Pneumocystis*, grave pneumonia bacteriana recorrente, infecção por herpes simplex crônica (orolabial, genital, ou anorretal por mais de um mês ou herpes visceral em qualquer lugar), candidíase esofágica (ou candidíase na traquéia, brônquios e pulmões), tuberculose extrapulmonar, sarcoma de *Kaposi*, infecção por citomegalovírus (retinite ou infecção de outros órgãos), toxoplasmose do sistema nervoso central, encefalopatia por HIV, criptococose extrapulmonar (incluindo meningite), infecção disseminada por micobactérias não-tuberculosas, leucoencefalopatia multifocal progressiva, candidíase na traquéia, brônquios e pulmões, criptosporidiose crônica (com diarreia), isosporíase crônica, micose disseminada (histoplasmose, peniciliose), salmonelose recorrente, linfoma (cerebral ou de célula B não-*Hodgkin*), carcinoma invasivo do colo uterino, leishmaniose disseminada atípica, nefropatia sintomática associada ao HIV,

cardiomiopatia sintomática associada ao HIV e reativação de tripanossomíase americana (doença de Chagas, meningoencefalite ou miocardite).

## **A Estrutura Molecular do HIV-1**

O HIV-1 é um Lentivírus (vírus lento) pertencente à família *Retroviridae*, que se caracteriza pela presença de um genoma constituído por duas fitas de ácido ribonucléico (RNA) simples e da enzima transcriptase reversa (TR) que exerce papel chave na transcrição do RNA para uma fita dupla de ácido desoxirribonucléico (DNA) (Figura 2).

Estruturalmente o HIV é uma partícula icosaédrica medindo cerca de 110 nm e possui externamente um envelope oriundo da bicamada lipídica da membrana externa da célula do hospedeiro. Este envelope envolve o nucleocapsídeo, onde estão abrigadas duas fitas simples de RNA. O genoma de RNA possui cerca de 9 kb de comprimento e é ocupado em grande parte pelos três genes que codifica proteínas estruturais e enzimáticas: *gag* que codifica as proteínas da matriz, do capsídeo e do nucleocapsídeo; *pol* que codifica as proteínas enzimáticas; *env* que codifica a glicoproteína precursora do envelope; e seis outros genes regulatórios, sendo cinco: *vif*, *vpr*, *tat*, *rev* e *nef* comuns ao HIV-1 e HIV-2; um *vpu*, presente apenas no HIV-1; e um outro *vpx*, presente apenas no HIV-2. Este genoma é delimitado em cada ponta por duas regiões de repetições terminais longas (LTR), onde estão as principais sequências promotoras para a transcrição dos genes virais (Figura 3).

O HIV-1 produz a maior parte de suas proteínas virais, alvo de terapias ARV e pesquisa em vacinas, como proteínas precursoras de grande porte, incluindo a proteína Gag, as proteínas precursoras GagPol e Env

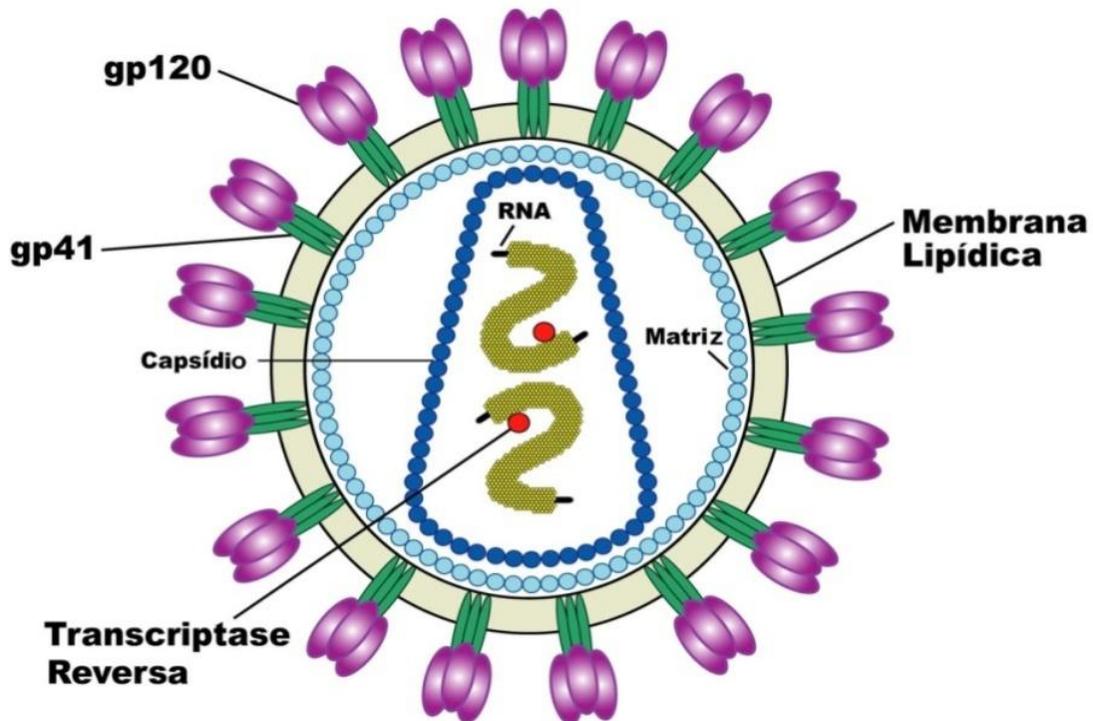


Figura 2: Estrutura esquemática do HIV. Adaptado de [www.cdc.gov](http://www.cdc.gov).

(glicoproteína 160 ou gp160). A precursora Gag após clivagem gera proteínas estruturais do vírus, enquanto GagPol produz três enzimas virais: protease (PR), TR e integrase (IN), e a Env as glicoproteínas 40 (gp40) e 120 (gp120). A PR é responsável pela clivagem das primeiras duas proteínas precursoras em sítios de clivagem diferentes numa ordem definida, e a gp160 é clivada durante a migração e processamento no retículo endoplasmático rugoso e depois pelo complexo de Golgi até a montagem do vírus e brotamento na superfície celular (BIRD *et al.*, 1990).

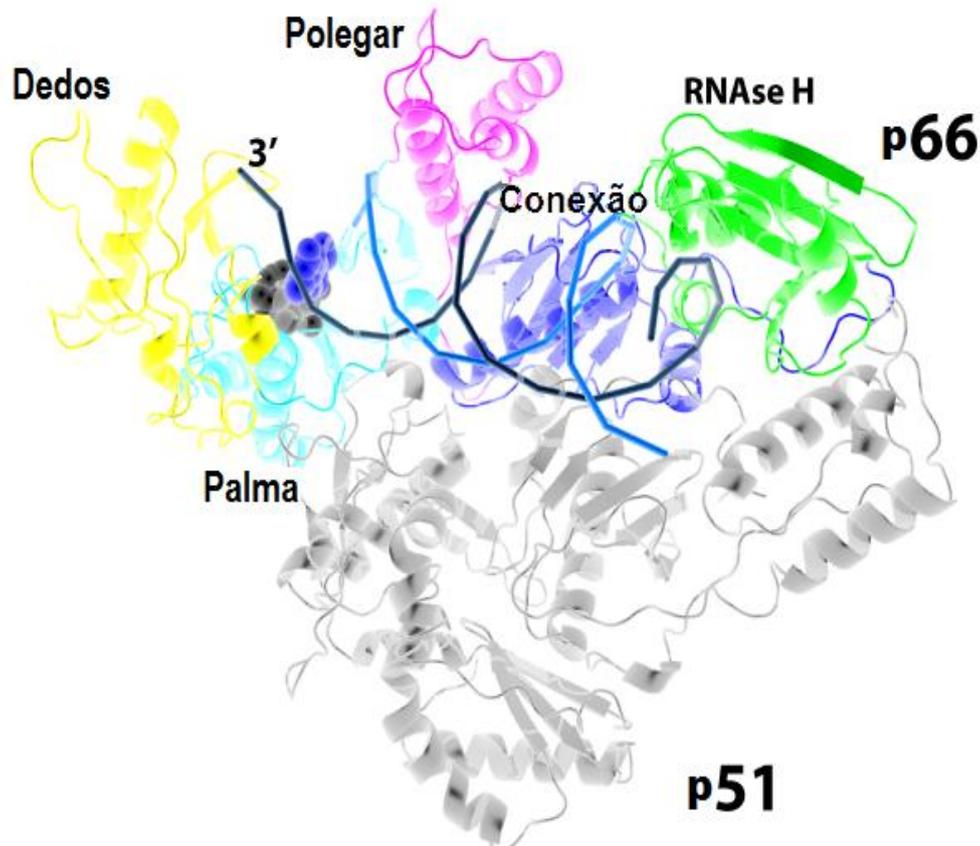
O gene ***gag*** de 1295 pares de bases (pb) codifica uma proteína precursora miristilada de 55 quilo Daltons (kDa) (p55) com funções muito importantes na montagem viral. Esta é clivada pela protease em etapas que



marcadamente diferente. A subunidade p66 contém uma fenda para DNA no sítio ativo enquanto a subunidade p51 não exibe atividade enzimática e funciona apenas como um andaime para a subunidade p66 enzimaticamente ativa (SARAFIANOS *et al.*, 2009). A subunidade p66 tem subdomínios nomeados de dedos, palma e polegar que participam da polimerização e conexão de ácidos nucleicos (Figura 4).

**A gp120** é capaz de iniciar o processo de ligação, fusão e entrada na célula alvo. Esta glicoproteína se apresenta no formato de um trímero que se associa a outro trímero de gp41 com localização transmembranar, formando a coroa ou espícula do envelope viral (LU *et al.*, 1995). Cada gp120 apresenta cinco regiões conservadas (C1-C5) intercaladas por cinco regiões variáveis (V1-V5).

Uma dessas regiões variáveis, designada de alça V3, está envolvida na fusão do vírus à célula. Esta é composta por 35 aminoácidos onde o primeiro e o último resíduo cisteína se unem por pontes dissulfeto formando uma estrutura do tipo grampo beta. Deleções na alça V3 bloqueiam a infecção viral (CHIOU *et al.*, 1992; CAO *et al.*, 1997; YANG *et al.*, 2004) e variações na eficiência da entrada são principalmente devido a grande variabilidade no gene do envelope que, por sua vez, está relacionada com as fortes pressões seletivas do ambiente hospedeiro (RICHMAN *et al.*, 2003). Apesar de muito variável, o número de aminoácidos e as características estruturais são bem conservados (CARDOZO *et al.*, 2007).

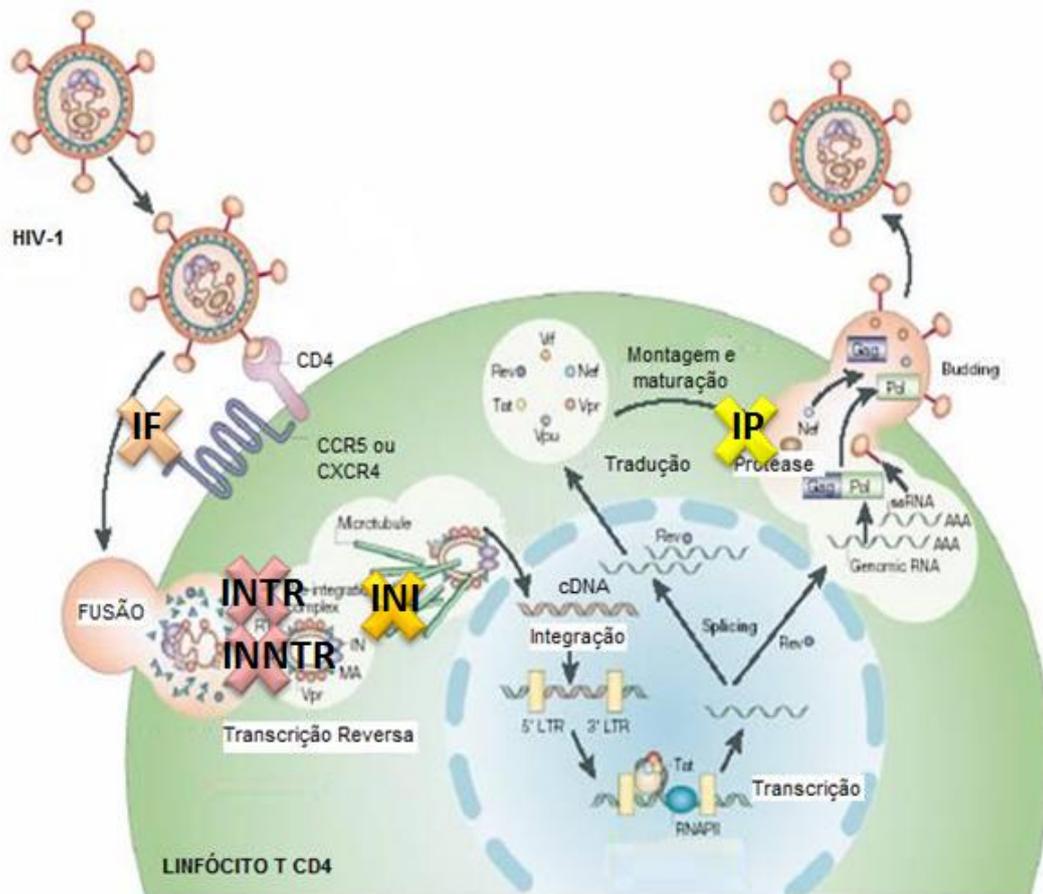


**Figura 4:** Estrutura tridimensional esquemática da transcriptase reversa. Cristalografia por raios-X da p66/p51 mostrando seus principais domínios. O preenchido em preto se refere ao sítio ativo formado por três aspartatos: D110, D185 e D186 e em azul o dNTP de entrada na cadeia de DNA. Em verde a porção catalítica que cliva o RNA molde (linha em azul escuro) ao DNA nascente (linha em azul claro). A conexão se refere a fenda que abriga a ligação RNA-DNA (adaptado de HUANG *et al.*, 1998 e PDB = 1rtd).

## **O Ciclo do HIV-1**

Como característica de todo vírus, o HIV utiliza a maquinaria celular para se replicar (figura 5). O ciclo retroviral se inicia com a ligação da partícula, através da região ligante de CD4 da espícula do envelope (gp120) com o receptor CD4 na superfície das células linfocitárias T, monócitos, macrófagos, células dendríticas, células de Langerhans ou microgliais (KEDZIERSKA *et al.*, 2003), o qual, após as mudanças conformacionais, segue a ligação da alça V3 do envelope a uma das moléculas receptoras de quimiocina, denominadas de coreceptores CCR5 ou CXCR4 (UGOLINI *et al.*, 1999). A formação do complexo ternário resultante (gp120/CD4/coreceptor) leva a uma segunda mudança conformacional das glicoproteínas da espícula e a exposição de sítios hidrofóbicos da gp41 que se ancora na membrana plasmática. Ocorre então a fusão entre o envelope viral e a membrana celular permitindo a entrada do capsídeo na célula. No citoplasma da célula o capsídeo libera além da dupla fita de RNA, as enzimas retrovirais: IN, PR e a TR. A TR possui duas regiões ativas importantes, já descritas anteriormente, que transforma o RNA retroviral em DNA complementar (cDNA): a polimerase e a ribonuclease. O RNA retroviral passa pela região polimerase da TR produzindo uma hélice dupla RNA/DNA, esse complexo passa pela região da ribonuclease que separa essa dupla hélice, em seguida a hélice simples de DNA volta a passar pela região da polimerase que produz uma dupla hélice de DNA. Em seguida este cDNA se liga a IN que o transporta para o núcleo celular e o incorpora ao genoma da célula hospedeira. O provírus pode então

ser replicado e transcrito juntamente com o genoma celular. A ativação da célula produzirá um RNA proviral mensageiro que migra para o citoplasma e é traduzido em proteínas virais. Estas proteínas longas são clivadas pela PR. Por fim, duas fitas simples de RNA juntamente com enzimas provirais são unidas e montadas em um capsídio que em seguida se associa a membrana plasmática com as proteínas do envelope que brotam da célula. Esses novos vírus estão prontos para reinfectar novas células T CD4 com os coreceptores apropriados, reiniciando assim o ciclo (UGOLINI *et al.*, 1999). Este ciclo se repete gerando em média até  $10^{10}$  partículas virais por dia (MARKOWITZ *et al.*, 2003), culminando na destruição de todas os linfócitos T CD4 do paciente, no comprometimento do sistema imune e surgimento de doenças oportunistas.



**Figura 5:** Ciclo de replicação do HIV (adaptado de STEVESON *et al.*, 1990). As marcas coloridas em X representam inibidores do ciclo: IF = inibidor de fusão; INTR = inibidor nucleosídico da transcriptase reversa; INNTR = inibidor não nucleosídico da transcriptase reversa; INI = inibidor de integrase; IP = inibidor da protease.

## **A Geração da Diversidade no HIV-1**

A combinação de alguns fatores pode explicar a variabilidade do vírus no hospedeiro durante a longa evolução da infecção. A capacidade do vírus em escapar da pressão do sistema imunológico e das drogas ARV, aumenta a variabilidade das cepas circulantes, sendo estas espalhadas por varias populações nas epidemias.

A base molecular da extrema variabilidade do HIV-1 está principalmente na falha da TR em não possuir reparo na transcrição inversa do RNA em DNA durante o processo de replicação viral, o que resulta numa alta taxa de erro na incorporação dos diferentes nucleotídeos que pode chegar aproximadamente a 1 em  $10^4$ , taxa muito maior da observada na polimerase celular de 1 em  $10^8$  (OVERBAUGH *et al.*, 2001).

Outro processo reconhecido como recombinação homóloga também contribui para a diversidade do HIV. Este ocorre quando uma pessoa está infectada com dois subtipos diferentes ou mais. Normalmente nesta situação uma célula pode ser infectada por dois ou mais retrovírus de sequência ou subtipos diferentes, ocorrem a transcrição reversa de seus RNA em DNA e estes são incorporados ao genoma nuclear do hospedeiro. Quando a célula transcreve seu DNA genômico junto com os DNA provirais, os RNA mensageiros produzidos dos dois subtipos diferentes passam para o citoplasma onde ocorre a tradução para proteínas retrovirais e montagem das partículas, estas uma vez montadas podem ser homozigotas, se possuírem dois RNA de mesma sequência ou heterozigotas se possuir sequência de

RNA diferentes, quando uma destas partículas heterozigota infecta uma nova célula, o par de RNA é liberado e ocorre a transcrição reversa, ou seja, de RNA para DNA, é nesta fase do ciclo que ocorre a recombinação das duas sequências de RNA diferentes e a geração do DNA recombinante que é integrado ao DNA da célula hospedeira que irá produzir assim, partículas recombinantes. (HU & TEMIN, 1990; PINTO & STRUCHINER, 2006). Obviamente que um mesmo subtipo pode se recombinar, no entanto, como sua sequência é similar ou com pouca variação, fica muito difícil, por exemplo, determinar em que ponto ocorreu a recombinação.

Outros dois fatores de diversidade têm sido identificados mais recentemente. Os linfócitos expressam uma família de genes que codificam enzimas capazes de converter guanossina em adenosina no DNA viral (APOBEC3G), causando hipermutação no código genético viral. No entanto, o vírus possui o gene *vif* cujo produto pode se ligar ao RNA mensageiro da APOBEC3G e neutralizar este efeito, tornando viável o DNA viral, sem evitar mutações completamente (SHEEHY *et al.*, 2002; DONAHUE *et al.*, 2008; MERCENNE *et al.*, 2009).

Um último fator de variabilidade é o estresse oxidativo, que causa mutações no RNA e no DNA proviral antes da integração no DNA do hospedeiro. Em pacientes com AIDS pode-se detectar níveis plasmáticos de hiperoxidação de lipídios e oxidação da guanina. Este efeito é causado pela proteína *Tat* que inibe a expressão em linfócitos do gene *Mn-superóxido dismutase* aumentando a proliferação e a apoptose (SEVE *et al.*, 1999). Esta enzima é essencial para a transformação do ânion superóxido, altamente oxidante, em peróxido de hidrogênio. Moléculas altamente reativas derivadas

de oxigênio podem oxidar as bases, em particular a guanina, modificando assim sua capacidade de codificação ou induzindo uma substituição errada no reparo.

Em geral, esses fatores promovem um acúmulo de mutações que faz com que o vírus possa escapar da pressão do sistema imune e da terapia antirretroviral gerando no indivíduo infectado uma população heterogênea de vírus chamada de 'quasispecies' que pode ter até 15% de diversidade entre eles (STEAIN *et al.*, 2004), criando assim um grande desafio para a geração de uma vacina ou terapia efetiva.

## A Origem e a Classificação do HIV-1

Vários estudos têm demonstrado que a infecção pelo HIV e a consequente AIDS nos seres humanos é proveniente de múltiplos episódios de transmissão zoonótica de lentivírus de macacos africanos com tropismo para células T CD4<sup>+</sup> (GAO *et al.*, 1999; HANH *et al.*, 2000 ), os quais são coletivamente chamados de vírus da imunodeficiência símia (SIV, sigla do inglês *Simian Immunodeficiency Virus*). O HIV-1 é responsável pela grande maioria dos casos de AIDS em humanos, originária de um lentivirus de chimpanzés da espécie *Pan troglodytes* e por isso denominado de SIVcpz (PEETERS *et al.*, 1989; GAO *et al.*, 1999), enquanto o HIV-2, bem menos pandêmico e patogênico, deriva do lentivírus que infecta naturalmente macacos mangabeis da espécie *Cercocebus atys* SIVsmm (GAO *et al.*, 1992; MARX *et al.*, 1991; HIRCSH *et al.*, 1989). O HIV-2 foi utilizado para gerar as linhagens adaptadas SIVmac239 e SIVmac251 ao macaco rhesus (*Macaca mulatta*) com o objetivo de estudos de patologia e vacinas (NAIDU *et al.*, 1988; JONHSON *et al.*, 1992). Todos os lentivírus de primatas identificados até hoje se agrupam numa única linhagem filogenética distinta de lentivírus de outras espécies de mamíferos (TALBOTT *et al.*, 1989; BAILES *et al.*, 2002). A análise filogenética dos lentivírus de primatas mostra que cada espécie em geral é infectada com um grupo específico de SIV (Figura 6), ou seja, cada clado monofilético corresponde a uma espécie hospedeira com exceção do mandril, *Mandrillus sphinx*, que abriga dois tipos divergentes de SIV (SOUQUIERE *et al.*, 2001; GORDON *et al.*, 2009).

Estudos filogenéticos da diversidade genômica do HIV permitiram ir além à classificação deste vírus. O HIV-2 é formado por seis subtipos ordenados de A a F e com distribuição mais restrita ao Oeste da África (ALAEUS *et al.*, 2000). Enquanto o HIV-1 se apresenta dividido em três grupos: M (main), O (outlier) e N (neither). O grupo M é o responsável em grande parte pela pandemia sendo dividido em nove subtipos: A, B, C, D, F, G, H, J e K, além de formas recombinantes circulantes (CRF, do inglês *Circulation Recombinant Forms*) e formas recombinantes únicas (URF, do inglês *Unique Recombinant Forms*) (ROBERTISON *et al.*, 1999).

Em geral, se considerou que os subtipos puros estavam prevalecendo no início dos anos 80 quando começou a pandemia do HIV-1, enquanto as CRF passaram a aumentar a partir dos anos 90. Assim, quando os subtipos foram geneticamente caracterizados no início dos anos 90, os primeiros vírus identificados passaram a representar subtipos puros e os vírus posteriormente achados foram então comparados a estes protótipos. No entanto, estudos mais recentes incluindo amostras de soro coletados em meados da década de 80, em Kinshasa, na República Democrática do Congo (KALISH *et al.*, 2004), mostraram que recombinações intersubtipo já eram significativamente altas em meados de 80, quando o HIV-1 ainda não era conhecido. Desta forma, pelo menos alguns dos recombinantes, principalmente da África central, foram classificados como puros depois de ser exportados da África para as demais regiões do mundo.

Ainda há muitas controvérsias sobre quando começou a infecção pelo HIV-1 e a atual filogenia de genótipos. Estudos usando correlações linear de tempo e distâncias estimam que o HIV-1 primeiro divergiu do HIV-2 ao

redor do ano de 1930 e então divergiu para sua filogenia atual há 40 anos (GOJOBORI *et al.*, 1990). Por outro lado, estudos filogenéticos sobre a origem do HIV e do SIV estimam que o nó de divergência da origem dos HIV-1 e HIV-2 está em aproximadamente 1.800 anos e o nó mais velho ligando todos retrovírus humano e símios data entre 2.400 a 3.000 anos, enquanto o HIV-1 está estimado para aproximadamente 100 anos (MYERS *et al.*, 1994; KORBER *et al.*, 2001; LOUWAGIE *et al.*, 2005).



## **A Epidemiologia Molecular do HIV-1**

### *A Epidemiologia Molecular do HIV-1 no Mundo*

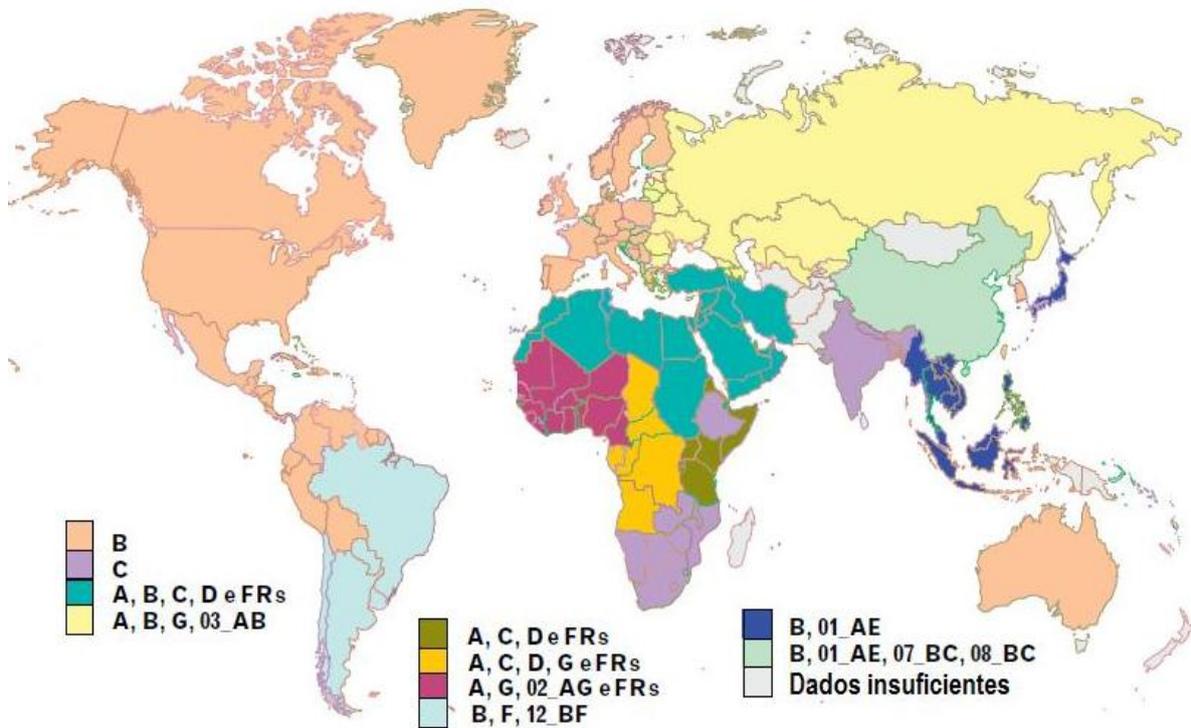
A classificação e subtipagem do HIV têm sido um efetivo marcador molecular e uma ferramenta poderosa para traçar o curso da infecção e monitorar as mudanças geográficas da pandemia. É possível buscar associações entre as linhagens circulantes em expansão e a patogenicidade e transmissibilidade da cepa. A pesquisa do perfil da diversidade molecular de uma determinada região geográfica fornece informações preciosas sobre o conteúdo antigênico que poderá ser considerado numa preparação antigênica de uma vacina específica. Saber se uma vacina ou terapia ARV baseada num determinado subtipo produz uma proteção cruzada contra outro subtipo também tem sido um estudo importante (GAO *et al.*, 2004; MARTINEZ-CAJAS *et al.*, 2009).

Os diferentes subtipos puros e CRFs do grupo M têm uma distribuição geográfica particular (Figura 7). A África possui todos os subtipos, sendo assim consistentes com estudos que mostram este continente como origem de disseminação pandêmica (HAHN *et al.*, 2000). O subtipo A tem sido responsável por 80% das infecções na África Ocidental, e por 30% na África Oriental (HOLGUIN *et al.*, 2002). Na Europa Oriental o subtipo A tem sido disseminado desde 1995 nos países da Antiga União Soviética, principalmente a Rússia e a Ucrânia (NABATOV *et al.*, 2002).

O subtipo B é a forma circulante mais distribuída entre os continentes e por isso, a mais estudada. Configura-se como a forma mais prevalente nas Américas, na Austrália (HERRING *et al.*, 2003), Coreia do Sul, Índia e Singapura (SE-THOE *et al.*, 1998). Na Europa Ocidental é o componente epidêmico principal (60%), junto com subtipo A (11%), C (5%) e outros subtipos (11%). No Japão, onde o subtipo B tem sido predominante (74%), outros subtipos como C (3%), A (2%), F (1%) e CRF01\_AE (20%) também são circulantes (KITSUTANI *et al.*, 1998).

Na América do Sul, o subtipo B tem sido predominante no Brasil (MORGADO *et al.*, 1998; COUTO-FERNANDEZ *et al.*, 1999; RAMOS *et al.*, 1999; BORGERTZ *et al.*, 2000; CARIDE *et al.*, 2000; MORGADO *et al.*, 2002; GADELHA *et al.*, 2003), seguidos pelos subtipos F, e C, com proporção pequena de subtipo D. Na Argentina, Bolívia, Peru, Paraguai, Uruguai e na Venezuela e também nas Ilhas caribenhas estes mesmos subtipos foram encontrados desde os meados dos anos 90 (RIOS *et al.*, 2003; VAUGHAN *et al.*, 2003; CASTRO *et al.*, 2003).

O subtipo C é a forma mais prevalente e em expansão no mundo (ESPARZA *et al.*, 2000; YERLY *et al.*, 2004; MONTEIRO *et al.*, 2009), constituindo cerca de 60 % das infecções globais. É predominante no oriente e sul da África e sul da Ásia (KOCH *et al.*, 2001; WAINBERG, 2004). Foi encontrado na Malásia e no Sudoeste da China, refletindo uma difusão deste vírus pela Índia (CASSOL *et al.*, 1996). Na Tanzânia há 50% de prevalência desta classe seguida por subtipos A e B (Arroyo *et al.*, 2004). Foi verificado epicentro na África Meridional envolvendo a Botsuana, Zimbábue, Malauí, Zâmbia, Namíbia, África do Sul que se espalhou para a Índia, Nepal e China.



**Figura 7.** Distribuição global dos subtipos, das CRFs e das URFs do HIV-1.

FR = formas recombinantes. Adaptado de Woodman & Williamson 2009.

Na Europa também houve um aumento encontrado na Escócia desde 2000, principalmente devido a transmissão de indivíduos com exposição no Reino Unido, países africanos e asiáticos (YERLY *et al.*, 2004).

O subtipo D está prevalente em até 40% em alguns países do leste e do centro da África (PEETERS & SHARP, 2000). O F tem sido encontrado na Romênia e na América do Sul (BELLO *et al.*, 2006; SANABANI *et al.*, 2009). Os subtipos G, H, J e K são encontrados mais frequentemente na região central da África e tem se espalhado pelo sul da Europa e Ásia (REQUEJO, 2006).

Existem atualmente 43 CRFs que se distribuem regionalmente de acordo com a sua expansão ou prevalência dos subtipos e que representam atualmente 18% das infecções globais (HEMELAAR *et al.*, 2006). Dentre estes alguns se destacam por serem encontrados em um maior número de infectados. O CRF01\_AE, antigo subtipo E (GAO *et al.*, 1996), foi primeiro identificado na Tailândia (CARR *et al.*, 1996), onde é predominante com mais de 80% de prevalência (SRISUPHANUNT *et al.*, 2004) e depois tem sido comum encontrá-lo no Vietnã e países vizinhos na maioria associados ao uso de drogas intravenosa (UDI) (CHAKRABARTI *et al.*, 2000). O CRF02\_AG é a CRF mais prevalente no mundo e no continente africano representa até 70% das formas circulantes. Também é encontrada na Europa, principalmente na França, Itália, Bélgica e Reino Unido (MONNO *et al.*, 2005; SNOECK *et al.*, 2004; VACHOT *et al.*, 2004). Na Argentina, Uruguai, e Brasil a CRF12\_BF ficou prevalente principalmente em população heterossexual, e em crianças verticalmente infectadas (CARR *et al.*, 2001).

O grupamento O do HIV-1 corresponde em torno de 10% das infecções mundiais, sendo endêmico no Oeste e na África Central, principalmente na República do Camarões, Gabão, Nigéria, e Guiné Equatorial. Este grupo teve seu início na África e depois se expandiu para a Europa e os Estados Unidos, no entanto, a maior parte das infecções devido ao grupo O está diretamente associadas a pessoas que tiveram algum contato de risco ou habitam a África Central ou Oeste (JONASSEN *et al.*, 1997; YAMAGUCHI *et al.*, 2002; YAMAGUCHI *et al.*, 2004).

*A Epidemiologia Molecular do HIV-1 no Brasil e na Bahia*

No Brasil, país de extensão continental, com o maior número de casos na América Latina e maior diversidade filogenética das Américas (BASTOS *et al.*, 2008), a distribuição de subtipos e formas recombinantes assume padrões diversificados de acordo com as regiões geográficas. Em geral, o subtipo B é a forma circulante predominante, no entanto também é encontrado F, C, e D (KUPEK *et al.*, 1990; TANURI *et al.*, 1999; BONGERTZ *et al.*, 2000; MORGADO *et al.*, 2002; COUTO-FERNANDEZ *et al.*, 2006; BARRETO *et al.*, 2006). Outras formas recombinantes já encontradas no Brasil como B/C, C/B, B/F, F/D, e o triplo B/C/F estão associadas com os quatro subtipos circulantes (RAMOS *et al.*, 1999).

Nos Estados brasileiros da região sul, principalmente Rio Grande do Sul e Santa Catarina, predominam o subtipo C seguido do B e das formas recombinantes BC (BRINDEIRO *et al.*, 2003; SOARES MA *et al.*, 2003; SOARES EA *et al.*, 2003; SALEMI *et al.*, 2005; MONTEIRO *et al.*, 2007; RODRIGUES *et al.*, 2007). No sudeste, assim como no resto do país, o subtipo B predomina seguido do F1, C, BF e BC (SABINO *et al.*, 1996; GONSALEZ *et al.*, 2007). Na região centro-oeste e na região amazônica predominam os subtipos B e F, mas também C, D e o CRF02\_AG foram encontrados (VICENTE *et al.*, 2003; VÉRAS *et al.*, 2007; MACHADO *et al.*, 2009; CARDOSO *et al.*, 2009).

Estudos realizados na região nordeste, especificamente nos estados do Ceará, Piauí, Rio Grande do Norte, Paraíba, Pernambuco e

Alagoas, demonstraram algumas variações na distribuição da prevalência de subtipos. O subtipo B de 80 a 82%, BF de 2,7 a 4,6% e F1 de 2,7 a 12% (GADELHA *et al.*, 2003; CAVALCANTI *et al.*, 2007). Em Salvador foi demonstrado um perfil de 84% para o subtipo B, 13,2% para BF, 2,3% para F e 0,6% para FD (MONTEIRO *et al.*, 2009). Outros estudos mostram que a região nordeste possui uma prevalência alta de B (>80%) seguida de uma baixa frequência (<3%) de subtipo F e BC recombinante (BRINDEIRO *et al.*, 2003; BRENNAN *et al.*, 2007). Em geral observa-se sempre o predomínio do subtipo B, das formas recombinantes BF e mais raramente de outras formas.

Aproximadamente metade das amostras de subtipo B brasileiras apresentam na alça V3 o tetrapeptídeo GWGR, designado assim como subtipo brasileiro B'. Esta assinatura molecular é rara no subtipo B de outras partes do mundo. A forma comumente encontrada é a GPGR, mais prevalente na Europa e EUA e as outras assinaturas GXGX (DA COSTA *et al.*, 1995; MORGADO *et al.*, 1998; BONGERTZ *et al.*, 2000; DIAZ 2009).

## **As Mutações de Resistência aos Antiretrovirais**

A partir do surgimento da epidemia de AIDS, os esforços foram direcionados para o desenvolvimento não só de vacinas, mas também de substâncias capazes de combater esta doença. Em 1986 surgiu a primeira droga terapêutica contra o HIV, conhecida como zidovudina (AZT ou azidotimidina), um análogo nucleosídico da timidina que age inibindo preferencialmente a atividade da enzima transcriptase reversa e, portanto, impedindo a replicação do vírus (YARCHOAN *et al.*, 1986). Apesar de essa droga ter se mostrado eficaz inicialmente, a mesma não alterou significativamente o tempo de sobrevivência dos pacientes. Anos depois, surgiram novas substâncias que impediam a atividade de outras enzimas, como os inibidores da protease que impedem a maturação de novas partículas virais. Estas substâncias, associadas ao AZT (terapia dupla), aumentaram discretamente a sobrevida das pessoas afetadas. Em 1996 surgiu uma nova proposta terapêutica conhecida como coquetel anti-HIV ou terapia antirretroviral de alta potência (HAART, do inglês *Highly Active AntiRetroviral Treatment*), baseada na combinação de três classes de agentes inibidores: um inibidor nucleosídico da transcriptase reversa (INTR) e um inibidor de protease (IP) ou um inibidor não nucleosídico da transcriptase reversa (INNTR). Esta terapia tem proporcionado uma drástica redução na carga viral plasmática de pacientes infectados pelo HIV-1 com consequente aumento do número de células T CD4, não aparecimento de doenças oportunistas, neoplasias, diminuição da taxa de mortalidade, morbidade e

hospitalizações (PERELSON *et al.*, 1996; SEPKOWITZ, 2001; PALELLA *et al.*, 1998). Porém, trouxe também alguns problemas como: nova resistência viral, toxicidade, complexidade de esquemas terapêuticos e alto custo, o que dificulta sua distribuição universal. Além disso, está longe de ser perfeita, por isso as pesquisas continuam buscando novos alvos, objetivando torná-la mundialmente efetiva e acessível (POMERANTZ & TRONO, 1995).

Outra classe de ARV, que utiliza inibidores de entrada, tem melhorado a qualidade de vida dos pacientes com falha terapêutica. Dentre as duas drogas em teste desta categoria se destaca o Maraviroc, aprovado em 2007 para comercialização, que inibe a entrada do HIV-1 na célula ao se ligar no coreceptor CCR5. Em média, 80% dos pacientes se beneficiaram deste ARV que demonstrou ser eficiente em diminuir a carga viral numa média de 2 logs e aumentar em média o número de células T CD4+ por mL de sangue em 100 após 24 meses de tratamento (RAY, 2009). Atualmente, há cerca de 25 drogas que compõem o arsenal contra o HIV (Tabela 02).

O início do tratamento ARV em indivíduos soropositivos é realizado considerando a contagem de células T CD4+, a carga viral e os aspectos clínicos, que vão determinar o grau de imunodeficiência já existente e o risco para a progressão da AIDS (MS, 2008). De acordo com as recomendações do CDC, todos aqueles que apresentem os sintomas da infecção devem iniciar o uso da HAART. Para os assintomáticos, devem ser analisados os riscos e benefícios do tratamento. Em geral, se recomenda que estes pacientes preencham os seguintes critérios: número de linfócitos T CD4 < 350 e carga viral > 100.000 cópias por ml de sangue periférico. O resultado da terapia deve ser avaliado pela carga viral do paciente. Espera-se que ela

diminua em 1,0 log após oito semanas de tratamento e se torne indetectável (< 50 cópias) em quatro a seis meses.

No Brasil, de modo inusitado, em comparação a outros programas de controle mundiais, tem-se enfrentado a epidemia com campanhas de distribuição de preservativos e de medicamentos ARV sem custo adicional para os pacientes. Além disso, houve implantação de laboratórios para diagnóstico na rede pública, acompanhamento dos pacientes e suporte para pesquisas. Essas medidas, se não são suficientes para interromper a disseminação da epidemia, foram capazes de aumentar a sobrevivência e melhorar a qualidade de vida de pessoas com AIDS, além de possibilitar a melhoria das condições técnico-científicas dos profissionais de saúde. (TEIXEIRA *et al.*, 2004; NUNN *et al.*, 2007; DO LAGO *et al.*, 2009; CAMPOS *et al.*, 2009).

Estima-se que 10 a 20% dos pacientes que iniciam o tratamento não conseguem suprimir a viremia de forma satisfatória, após alguns meses de terapia (falha terapêutica primária), e cerca de 20 a 50% dos que apresentam boa resposta inicial, apresentarão falha terapêutica após um ano de tratamento (falha terapêutica secundária), sendo a taxa de resposta virológica, a tratamentos subsequentes, progressivamente menores (BE, 2007).

**Tabela 02** - Principais drogas antiretrovirais aprovadas pelo *Food and drug administration* (FDA) nos EUA.

<b>Categoria</b>	<b>Nome Genérico</b>	<b>Nome Comercial</b>
<b>INTR</b>	ZIDOVUDINA (AZT, ZDV)	Retrovir
	DIDANOSINA (DDI)	Videx
	ESTAVUDINA (D4T)	Zeritavir
	LAMIVUDINA (3TC)	Ziagenavir
	ABACAVIR (ABC)	Trivir
	ZIDOVUDINA/LAMIVUDINA	Biovir
	EMTRICITABINE (FTC)	Emtriva
	TENOFOVIR DF (TDF)	Viread
<b>INNTR</b>	DELAVIRDINA (DLV)	Rescriptor
	EFAVIRENZ (EFZ)	Sustiva
	NEVIRAPINA (NVP)	Viramune
	ETRAVIRINE (ETR)	Tibotec
<b>IP</b>	AMPRENAVIR (FPV)	Agenerase
	RITONAVIR (RTV)	Norvir
	RITONAVIR/LOPINAVIR (LPV)	Kaletra
	SAQUINAVIR (SQV)	Invirase
	NELFINAVIR (NFV)	Viracept
	ATAZANAVIR (ATV)	Reyataz
	DARUNAVIR (DRV)	Prezista
	INDINAVIR (IDV)	Crixivan
<b>INI</b>	TIPRANAVIR (TPV)	Aptivus
	RALTEGRAVIR (RGV)	Isentress
<b>IF / INE</b>	ELVITEGRAVIR (EGV)	Truvada
	ENFUVIRTIDE (T-20)	Fuzeon
	MARAVIROC (MAR)	Celsentri

Abreviações: INTR = inibidor nucleosídico da transcriptase reversa; INNTR = inibidor não nucleosídico da transcriptase reversa; IP = inibidor da protease; INI = inibidor da integrase; IF = inibidor de fusão; INE = inibidor de entrada. Adaptado de hivdb.stanford.edu.

Está resistência aos ARV pode ocorrer naturalmente devido às mutações e polimorfismos dos subtipos do HIV-1, ou devido aos fatores ambientais do hospedeiro, que exercem uma pressão seletiva sobre o vírus como resposta adaptativa (COFFIN, 1995). No entanto há controvérsias se os subtipos do grupo M têm diferentes sensibilidades aos medicamentos ARV atualmente utilizados. Essa diferença já foi observada tanto *in vitro* quanto em genotipagem (PALMER *et al.*, 1998; MARTINEZ-CAJAS *et al.*, 2009). Alguns estudos têm demonstrado diferenças significativas entre sequências de RT e PR de subtipo B e não-B na geração de mutações de resistência à droga ARV após o início do tratamento (PIENIAZEK *et al.*, 2000; CARIDE *et al.*, 2000), e também no que se refere a diferença de aquisição de mutações de resistência entre B e C (SOARES EA *et al.*, 2007) e entre B e BF (CAROBENE *et al.*, 2004). No entanto, outros estudos mostraram nenhum efeito entre subtipos ou forma recombinantes durante o tratamento ARV (FRATER *et al.*, 2002; DE SÁ-FILHO *et al.* 2008)

A identificação de resistência aos medicamentos ARV tornou-se uma questão importante no manejo clínico das doenças relacionadas ao HIV. Desta forma, a vigilância epidemiológica utilizando subtipagem e genotipagem das cepas circulantes é imprescindível para se determinar se os subtipos do HIV diferem em sua susceptibilidade primária às drogas ARV ou qual o grau de sua capacidade em desenvolver resistência ARV.

## **Análises Moleculares por Ferramentas de Bioinformática**

A bioinformática surge principalmente com a necessidade de se manter banco de dados para armazenamento e manutenção das informações biológicas contidas nas sequências de nucleotídeos e aminoácidos. Com o objetivo de se estudar como as atividades normais de um organismo são alteradas em diferentes estados da doença, os dados biológicos são associados para formar um quadro completo dessas atividades. Desta forma, a bioinformática se desenvolveu nesta combinação, análise e interpretação de vários tipos de dados, incluindo sequências de nucleotídeos e aminoácidos, os domínios de proteínas e suas estruturas. Estudar as mutações nucleotídicas de grandes e várias sequências, traduzir uma sequência de nucleotídeos *in silico* para uma sequência de aminoácidos, prever a estrutura da proteína e/ou função, já fazem parte da rotina e pesquisa de milhares de laboratórios no mundo.

O genoma do HIV-1 contém uma valiosa fonte de informações que estão registradas através do acúmulo de alterações hereditárias no DNA das sequências que podem ser comparadas para mostrar diferenças entre as funções da proteína ou usadas em análises filogenéticas para explicar o aparecimento de subtipos e formas recombinantes. Também é útil para predição de epítomos em regiões conservadas potencialmente eficazes no desenvolvimento de vacinas, para predição do uso de coreceptores cujo conhecimento é importante para medidas de terapêutica e estudo do desenvolvimento da doença, para a genotipagem de mutações de resistência

aos antiretrovirais e para saber em qual tipo de pressão seletiva estão as amostras.

*Alinhamento e edição de sequências de DNA e Proteínas*

Com o intuito de se estudar grande número de sequências genômicas geradas é preciso relacioná-las em forma de alinhamento múltiplo. Este é uma hipótese de homologia posicional entre bases nucleotídicas ou aminoácidos em sequência de duas ou mais espécies. O alinhamento pode ser feito igualando uma sequência sobre outra e utilizando espaços ou traços para representar as deleções ou inserções de modo a minimiza a diferença entre elas. Alguns programas de informática como, por exemplo, Clustal (THOMPSON *et al.*, 1994; THOMPSON *et al.*, 1997) e Muscle (EDGAR, 2004), podem carregar sequências em diversos formatos que possuem sinais gráficos ou pequenas escrituras informativas para serem lidos, o formato fasta é o mais comum destes. Após o programa executar o alinhamento, o mesmo pode ser editado em programas como o GENEDOC (NICHOLAS, 1997) ou BIOEDIT (HALL, 1999) e exportado em outros formatos para posteriores análises moleculares.

*Escolha do Melhor Modelo de Substituição Nucleotídica*

Antes de iniciar as análises filogenéticas dos dados é preciso determinar o melhor modelo de substituição nucleotídica que descreva o processo a partir do qual uma sequência de caracteres se transforma em outro conjunto e ser capaz de incorporar heterogeneidade, pois nem todos os sítios modificam na mesma taxa. A distribuição mais usada é a gama e o parâmetro alfa. Valores baixos de alfa indicam que a maioria das sequências tem sítios que mudam menos (SCHNEIDER, 2003).

O primeiro modelo foi proposto em 1969 por Jukes e Cantor, assumindo que as taxas de transição (mudanças no sítio entre duas bases púricas ou duas pirimídicas) e de transversão (mudanças entre uma púrica e outra pirimídica e vice-versa) são iguais, bem como as frequências de equilíbrio de todas as bases, ou seja, mudanças nucleotídicas ocorrem na mesma taxa em qualquer sítio. Em 1980, Kimura introduziu um modelo com dois parâmetros: um para a taxa de transição e outro para a transversão e em 1981, Felsenstein fez um modelo em que a taxa de substituição corresponde à frequência de equilíbrio do nucleotídeo alvo. Hasegawa, Kishino e Yano em 1985 unificaram os dois últimos modelos para um modelo com seis parâmetros, o modelo HKY. Rodriguez em 1990 desenvolveu um modelo com seis parâmetros independentes, para calcular a taxa de substituição nucleotídica, denominado *General Time Reversible* (GTR). Tamura e Nei em 1993 desenvolveram um modelo para uma solução do cálculo de distâncias no modelo de HKY.

O programa Modeltest (POSADA & CRANDAL 1998), implementado no programa PAUP (SWOFFORD 1999), seleciona o melhor modelo de substituição nucleotídica, dentre 56, que se ajuste melhor aos dados através de testes de probabilidade hierárquicas.

Este programa usa o teste da razão de verossimilhança (LTR, do inglês *likelihood ratio test*) para escolha entre dois modelos. Um modelo relativamente mais complexo é comparado com um modelo mais simples para ver se ele se ajusta a um conjunto particular significativamente melhor. Desta forma, especialmente no estudo de alinhamentos de sequências de HIV, os parâmetros do modelo mais complexo são mais utilizados nas análises. Assim, seguindo uma hierarquia de comparações de modelos, a adição de parâmetros sempre resultará em uma pontuação maior de probabilidade. No entanto, chega um momento em que a adição de parâmetros adicionais não se justifica em termos de melhora significativa no ajuste de um modelo para um conjunto de dados particular. O LRT fornece um critério objetivo para a seleção entre modelos possíveis em que começa com a comparação dos escores de probabilidade de dois modelos, seguindo uma distribuição aproximada ao qui-quadrado (FELSENSTEIN 1981; SWOFFORD *et al.*, 1996. HUELSENBECK JP & RANNALA B, 1997)

### *Análises filogenéticas*

O acúmulo e fixação de mutações, ou seja, a taxa de substituições nucleotídicas, é considerada como um processo de ramificação, onde as populações são alteradas ao longo do tempo e podem se agrupar em ramos distintos. Isto pode ser visualizado em uma árvore filogenética que é baseada em uma hipótese a ser testada para reconstruir a topologia mais próxima da realidade. Estas árvores são representações constituídas de nós e ramos na qual um nó conecta dois ramos. A extensão de cada ramo representa o número de substituições por sítio ou uma estimativa de tempo de divergência de um táxon e cada nó representa um táxon hipotético de um ancestral comum. Os taxa agrupados por semelhança podem formar grupos monofiléticos (SCHNEIDER, 2003). Desta forma, podemos usar árvores para visualizar a topologia e comparar a diversidade de um grupo monofilético com outro a partir da sequência nucleotídica de um gene dentro de uma espécie. Os métodos mais comumente usados para inferir filogenias incluem aqueles baseados em distância, principalmente: agrupamento de pares não ponderados baseados na média aritmética (UPGMA, do inglês: *Unweighted Pair Group Method with Arithmetic Mean*) e agrupamento de vizinhos (NJ, tradução de *neighbor-joining*). Ou em análises de estado de caráter: máxima parcimônia, máxima verossimilhança (ML, do inglês: *Maximal likelihood*) e a inferência Bayesiana com análises de cadeias markovianas de Monte Carlo (MCMC do inglês, Markov chain Monte Carlo). Todos estes métodos dependem de um modelo matemático que descreve a mudança dos

caracteres observados a partir de alinhamentos de sequências de nucleotídeos ou de aminoácidos.

**NJ** é um método cujo algoritmo precisa da distância entre o par de cada táxons. O algoritmo NJ trabalha da seguinte forma: a) com base na matriz de distância ele calcula uma matriz Q para a junção dos nós; b) em seguida encontra o par de táxons em Q com valores mais baixos; c) criar um nó na árvore que une estes dois taxa, ou seja, juntar os vizinhos mais próximos; d) calcula a distância de cada um dos táxons no par de cada novo nó; e) calcula a distância de todos os táxons fora deste par para este novo nó; f) começa o algoritmo novamente, considerando o par de vizinhos como um táxon e usando as distâncias calculadas na etapa anterior (SAITOU, 1987). NJ baseia-se no critério de modificações mínimas, ou seja, a topologia que dá o menor comprimento de ramos é escolhida em cada etapa do algoritmo. No entanto, NJ pode não encontrar a topologia da árvore verdadeira com o comprimento de ramos esperados porque a reconstrução ocorre de maneira gradual. Mesmo assim, este método tem sido extensivamente testado e geralmente é encontrada a árvore que é bastante próxima à ideal quando não se opta por métodos que não dependem de medidas de distância e oferece precisão superior na maioria das condições. A principal virtude do NJ em relação a outros métodos é a sua eficiência computacional. É possível usar um conjunto de dados grandes onde outros métodos de análise filogenética não suportam. NJ não assume que todas as linhagens se modificam na mesma taxa, como na hipótese do relógio molecular que produz uma árvore sem raízes. Além disso, NJ é estatisticamente consistente com muitos modelos de substituição nucleotídica (GASCUEL, 2006). Assim, tendo em

conta os dados de comprimento, NJ irá reconstruir a verdadeira árvore com alta probabilidade. Dois programas podem ser utilizados com bastante eficiência para análises de NJ: PAUP\* (SWOFFORD, 1999) e SEAVIEW 4 (GOUY *et al.*, 2010).

*Teste estatístico para confiabilidade das árvores*

Nenhum dos métodos descritos acima garante a detecção de uma árvore ótima. Para estimar o quanto uma árvore resultante do algoritmo é confiável, os métodos de reconstrução buscam a melhor árvore e atribuem valores estatísticos de  $p$ , obtido pelo método de ML e o de *bootstrap*. O valor  $p$  é adquirido a partir da probabilidade de se obter uma taxa tão grande ou maior do que a razão observada sob a hipótese nula de que um ramo tem comprimento zero (FELSENSTEIN, 1981). Já o método de *bootstrap* é um processo de cálculos de uma nova árvore que é repetido várias vezes (FELSENSTEIN, 2004). Os valores de *bootstrap* são então atribuídos em porcentagens nos grupos que aparecem na topologia da árvore resultante. Valores de *bootstrap* indicam que uma filogenia tem suporte quando um ramo é ótimo em 90% das subamostras e aqueles com valor abaixo de 70% são vistos como suspeitos. Isso mostra que algumas filogenias são menos confiáveis porque os dados são igualmente consistentes com diversos padrões de ramificação. Só as que apresentam forte suporte representam uma descrição realista e rigorosa da descendência comum e de classificação.

*Cálculo da  $dN/dS$ , predição de uso de coreceptor e de mutações de resistência*

As sequências do HIV mostram em geral uma alta diversidade nucleotídica e divergência de formas. A comparação entre sequências de DNA são úteis para determinar se genes específicos estão em mudança ou são mais conservados. Um parâmetro que pode ser estimado a partir das sequências de DNA é a relação entre a taxa substituição não-sinônimas ( $d_N$ ) e a taxa de substituição sinônima ( $d_S$ ). A relação  $d_N/d_S$  ou  $\omega$  pode refletir a restrição seletiva em um gene, especialmente aquelas que agem para remover substituições de aminoácidos. Desta forma podemos comparar conjuntos de sequências com alguma qualidade particular ou em associação com algum fator de seleção em pacientes tais como, uso de terapia, resposta imune e outros fatores do hospedeiro. Foi demonstrado, utilizando o programa SNAP (KORBER, 2000), num estudo de vacina de fase III envolvendo pacientes imunizados com rgp120, que o gene gp120 do HIV-1 de pacientes afrodescendentes apresentavam nível de pressão seletiva maior (positiva) de 1,24 em relação às outras etnias que variaram de 0,66 a 0,81 ( $p < 0,001$ ), independentemente de terem sido imunizados com a vacina em teste ou placebo (PÉREZ-LOZADA *et al.*, 2009). Revelando assim a necessidade de se considerar, não só genes mais conservados, mas populações diferentes no desenvolvimento de vacinas.

Os correceptores do HIV mais importantes são os receptores de quimiocinas CCR5 e CXCR4. As partículas do HIV-1 se encontram em três

categorias de acordo com qual deles se ligam para se inserir numa célula: alguns só podem usar CCR5 (vírus R5), enquanto outros só podem utilizar o CXCR4 (vírus X4), e alguns podem usar qualquer um deles (vírus R5X4). Antes e durante o tratamento medicamentoso com um antagonista do coreceptor, é importante saber sobre o uso de coreceptor da população de vírus no paciente. A aprovação de uma nova droga antirretroviral inibidora de entrada, antagonista do coreceptor CCR5, para aqueles pacientes infectados com a variante R5, aliado com a falha terapêutica do uso convencional de HAART traz a necessidade de predição do uso de coreceptores por estes vírus. Um dos melhores programas de predição de coreceptor é o Geno2Pheno (SING *et al.*, 2007) disponível de graça e *online*. Ele ainda tem a vantagem de dispensar a utilização de ensaios experimentais caros e demorados ao utilizar sequências nucleotídicas do gene *env* com a região V3. O programa é baseado num método de aprendizagem estatística em duas classes, no qual os vírus X4 e R5X4 são considerados como positivos e R5, negativos. Assim, após o programa considerar as cargas das posições 11 e 25 da alça V3 envolvidas no tropismo do vírus, ou seja na sua ligação aos dois diferentes coreceptores, o programa prediz se o vírus é positivo ou capaz de utilizar CXCR4 como coreceptor ou negativo como capaz de utilizar o CCR5 e sugerir ao paciente o medicamento antagonista de CCR5, o Maraviroc. Além da sequência é possível incluir no programa os dados clínicos dos pacientes no sentido de aumentar a acurácia no diagnóstico de pacientes virgens de tratamento antirretroviral.

A predição de mutações de resistência a retrovirais pode ser realizada em ferramentas de bioinformática disponíveis na internet. O mais

amplamente utilizado é o programa *online* HIVdb da Universidade de Stanford. Este programa realiza subtipagem e prediz todas as mutações nos genes da TR e PR em relação a sequência referência HXB2. O programa prediz, destaca e fornece as referências bibliográficas principalmente das mutações de resistência de vigilância epidemiológica, ou seja, aquelas que causam alta resistência às drogas antiretrovirais. O HIVdb utiliza um algoritmo inteligente que permite após a inclusão das sequências pol do HIV-1 a saída dos níveis de resistência a 20 medicamentos antirretrovirais aprovados pela FDA, incluindo 8 IP, 7 INRT, 4 INNRT, e um INI. No programa cada mutação de resistência é atribuída uma pontuação para a droga e um comentário com citações atualizadas, a pontuação da droga é derivada da soma da pontuação de cada mutação associada à resistência a essa droga. Usando a pontuação total da droga, o programa informa um dos seguintes níveis de resistência: suscetíveis, resistência potencial de baixo nível, resistência de baixo nível, resistência intermediária, e resistência de alto nível. Em geral, os trabalhos utilizam as mutações de resistência de alto nível que causam forte resistência às drogas ARV e por isso mais importantes na vigilância epidemiológica.

É importante ressaltar que em um paciente infectado pelo HIV-1 existe uma população viral com uma taxa de diversidade que pode chegar a 15% (STEAIN *et al.*, 2004). Desta forma, para efeito de estudo em ferramentas de predição, se utiliza apenas um clone por paciente correspondendo assim à amplificação por PCR do espécime pró-viral mais prevalente no doente.

***JUSTIFICATIVA***

A transmissão e a infecção pelo HIV-1 ocorrem a mais de 25 anos no mundo desde que nos Estados Unidos identificaram os primeiros casos da síndrome que veio a ser conhecida como AIDS. A pandemia encontra-se, no momento, estabilizada, porém em níveis inaceitáveis. Atualmente, cerca de 33,4 milhões de pessoas vivem com o HIV, uma prevalência de 0,8 por cento, 2,7 milhões de novos casos e 2,1 milhões de mortes relacionadas a esses vírus no ano passado foram registrados (UNAIDS, 2009). Enquanto a incidência diminui nos grandes centros e entre as pessoas de melhor nível cultural ela aumenta nas pequenas cidades do interior e entre as pessoas de baixa escolaridade (BE, 2009). A maioria das infecções por HIV atualmente são devido à transmissão heterossexual, enquanto a transmissão vertical representa a principal forma de infecção entre as crianças. A transmissão vertical pode ocorrer durante a gravidez, parto ou amamentação e está relacionada a vários fatores que inclui diversidade viral, fatores maternos comportamentais e obstétricos (MARTINEZ, 2006). No Brasil há mais de 506 mil infecções por HIV e foram documentados cerca de 13.000 casos de transmissão vertical (BE, 2008). A heterossexualização da infecção coloca em alerta o risco pela transmissão vertical.

A capacidade deste retrovírus em causar essa doença incurável está associada ao fato do seu genoma não possuir sistema de reparo durante a replicação e da recombinação homóloga entre os diferentes subtipos. Estes aspectos levam a uma grande diversificação genética. Assim, ao mesmo tempo em que este vírus replica e escapa da pressão do sistema imune, ele depleta as células T CD4 (PERELSON *et al.*, 1996; MELLORS *et al.*, 1996; GALVANI, 2005). Esta diversidade genética se distribui diferencialmente em

todo o mundo. Especificamente no Brasil, país de extensão continental, que não só possui o maior número de infectados na América Latina como um elevado nível de diversidade retroviral (BASTOS *et al.*, 2008), o subtipo B é o genótipo predominante seguido por recombinantes B/F1, F1 e C (MORGADO *et al.*, 2002).

Desde 1993, verifica-se a predominância de cepas brasileiras do subtipo B (B') que contém um motivo GWGR na ponta da alça V3 em cerca de metade das amostras (POTTS *et al.*, 1993; MORGADO *et al.*, 1994). Os estudos do subtipo B' têm sido importantes uma vez que há uma relação entre sua presença e a lentidão na progressão para a AIDS (DE BRITO *et al.*, 2006). A determinação do tropismo do HIV-1 também tem sido crucial para as opções na escolha de medicamentos antirretrovirais alternativos e estudo da patogênese. A variante do HIV-1 R5, com tropismo para o correceptor CCR5, é encontrado principalmente nos pacientes na fase aguda ou assintomática e pode ser uma alternativa às drogas antagonistas do CCR5 em pacientes com falha terapêutica, enquanto as variantes R5X4 e X4, com tropismo para CXCR4 estão associadas com uma maior carga viral, contagem baixa de células T CD4 e aumento da progressão da doença e da resistência aos inibidores de CCR5 (CAVARELLI & SCARLATTI 2009).

No Brasil, os pacientes infectados pelo HIV possuem acesso gratuito para todos os ARV desde 1992, antes do advento da HAART em 1996. Embora tenha contribuído para um aumento na sobrevida e qualidade de vida, o uso sequencial do ARV também tem contribuído para o aparecimento de resistência viral, que tem como consequências: falha virológica; aumento da transmissão; aumento de resistência primária,

desenvolvimento de resistência secundária; e aumento da transmissão vertical. (HACKER *et al.*, 2007).

A vigilância epidemiológica molecular, a fim de monitorar a transmissão e expansão dos subtipos e as mutações de resistência em regiões definidas, utilizando biologia molecular e ferramentas de bioinformática, tem sido fundamental para dar suporte a políticas de saúde pública, para um melhor entendimento dos fatores da transmissão e para o desenvolvimento de novas terapias e vacinas (DE QUEIROZ *et al.* 2007).

Desta forma, o estudo em Feira de Santana tem a grande vantagem de dar acesso a prevalência de subtipos, formas circulantes e de mutações de resistência numa região do interior do Brasil com livre acesso ao HAART, alto trânsito rodoviário, baixo nível sócio econômico e de educação. E em Salvador, a vantagem de uma atualização da diversidade molecular do HIV-1 circulante nesta grande cidade brasileira, capital do Estado da Bahia, Nordeste do Brasil.

## **OBJETIVOS**

## **Gerais**

Caracterizar a diversidade molecular do HIV-1 nas cidades de Salvador e Feira de Santana associando com dados clínicos e epidemiológicos.

## *Específicos*

1. Determinar a prevalência de subtipos e formas recombinantes do HIV-1;
2. Caracterizar as formas recombinantes identificadas no estudo;
3. Caracterizar a variante brasileira do subtipo B, calcular sua pressão seletiva, comparar com o subtipo B os dados clínicos;
4. Determinar a prevalência do uso do coreceptor por predição em ferramenta de bioinformática;
5. Determinar a prevalência de mutações de resistência aos antiretrovirais;
6. Comparar a diversidade genética de Salvador e Feira de Santana.

## ***RESUMO DOS RESULTADOS***

Com o objetivo de avaliar e caracterizar a diversidade genética do HIV-1 em Salvador e Feira de Santana, realizamos um estudo de corte transversal em cada cidade.

Em Salvador foram analisadas 61 amostras (estudo aprovado pelo CEP/CPqGM 84/2006) coletadas de pacientes em acompanhamento no Hospital Universitário Professor Edgar Santos. Estes pacientes eram encaminhados a partir de vários ambulatórios e hospitais de Salvador após a detecção de anticorpos para HIV-1 ou HIV-2 e confirmação com um teste de Western-Blot e/ou após os primeiros sinais de doenças oportunistas para a unidade de especialização em retrovírus. Quarenta e um destes pacientes eram do sexo masculino e 20 do sexo feminino (proporção de 2:1). A média de idade foi de 43,3 anos para os homens e 39,1 anos para as mulheres. Não houve diferença estatística significativa quando comparamos entre homens e mulheres as médias de idade, do tempo de infecção, da carga viral e da contagem de linfócitos T CD4 e T CD8 no programa PASW 18. O percentual de grupos de acordo com a etnia foi de 23%, 47,5% e 29,5% entre Européia/Latina, mestiços (Européia/Latina e afrodescendentes) e afrodescendentes respectivamente. Da mesma forma, não houve diferenças significativas quando comparamos a média dos dados laboratoriais destes grupos entre si.

As análises filogenéticas de 42 isolados no gene *gag* e 46 no gene *env* revelaram uma alta concentração do subtipo B e poucos BF1. Nas 27 sequências que foram amplificadas em ambos os genes, 25 (92,6 %) agruparam com as referências do subtipo B nas duas árvores, duas amostras

(7,4%) não agruparam com nenhum subtipo de referência puro na árvore *gag* mas agruparam com o subtipo B na árvore *env*.

As duas sequências recombinantes em *gag* apresentaram no *bootscanning* padrão de recombinação entre os subtipos B e F1. Ambas as sequências são URF, pois os pontos de recombinação não foram encontrados nas CRF conhecidas.

Com base na caracterização da alça V3, encontramos poucas variantes brasileiras. A distância genética média, nos grupos GWGR, GPGR e GXGX, ficou bem próxima: 0,11, 0,13 e 0,12 respectivamente. A comparação da pressão seletiva a partir da razão  $d_N/d_S$ , realizadas no programa SNAP (<http://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>) e avaliadas estatisticamente no PASW 18, de cada grupo mostrou não haver diferença significativa entre GWGR e os outros dois grupos: GPGR e GXGX ( $p > 0,05$ ). Apenas encontramos diferença significativa entre o grupo GPGR e o GXGX ( $p = 0,018$ ) revelando um valor de pressão negativa maior no primeiro grupo. Foram comparadas as características clínicas dos indivíduos infectados pelo subtipo B' (GWGR) e subtipo B (GPGR + GXGX). O tempo desde o diagnóstico inicial foi maior no grupo de B' (12,71 anos) do que no grupo B (8,83 anos), a média de idade também foi maior entre o subtipo B' (48,9 anos) em comparação B (39,1 anos). Quanto ao uso do coreceptor, utilizando o programa de predição Geno2Pheno (<http://coreceptor.bioinf.mpi-inf.mpg.de>), 35 (76%) tinham preferência pelo coreceptor CCR5 (vírus R5) e onze (24%) pelo CXCR4 (vírus R5x4 / X4), sendo estas, desta forma, resistentes ao uso de inibidores de fusão. A média de idade entre os indivíduos com o vírus X4 (32,4

anos) foi menor que a com vírus R5 (43,3 anos) ( $p < 0,05$ ). Detalhes deste estudo podem ser encontrados no **manuscrito 01**.

Em Feira de Santana foi realizado um estudo de corte transversal em 96 mulheres acompanhadas no Centro de Referência em DST/AIDS desta cidade. A idade das pacientes variou de 14 a 49 anos; (média de 33,7 e desvio padrão de  $\pm 15$ ). A maioria das mulheres da amostra era solteira (51%), de classe baixa (97%) e com escolaridade primária (79,2%). A via de transmissão foi 96,2 % heterossexual, 2,6 % pelo uso de drogas injetáveis e 1,3 %. Doze delas tinham filhos HIV-1 positivos (SANTOS, 2010).

Quanto aos dados clínicos e laboratoriais, a média da carga viral foi de  $\log_{10}$  2,3 cópias/mL, a contagem dos linfócitos T CD4 variou de 50 a 1.784 cél/ $\mu$ L, com média de 494 cél/ $\mu$ L, o tempo médio de doença foi de 3,7 anos. A média do tempo de terapia ARV foi 2,5 anos (SANTOS, 2010).

Sobre a evolução clínica das mulheres estudadas, 30 estavam no estágio A (assintomáticas) do CDC; outras 47 no estágio B (sintomas leves), enquanto 19 no estágio C (AIDS). As coinfeções mais encontradas durante o acompanhamento foram as infecções respiratórias, seguida das intestinais, da tuberculose, da hepatites e do HTLV (4 casos). As patologias genitais mais encontradas foram as vulvo-vaginítes, seguida da sífilis, da herpes e do HPV, e de neoplasias intraepiteliais. Interessante observar que houve associação entre o tempo de doença, de tratamento e o estágio C, e ainda a prevalência de doenças genitais e o grupo de mulheres com subtipo B (SANTOS, 2010). Detalhes deste estudo podem ser lidos no **manuscrito 02**.

Um outro estudo em Feira de Santana de epidemiologia molecular do HIV-1 foi realizado em 8 crianças e 50 mulheres provenientes das amostras

do estudo anterior. A idade entre as mulheres variou de 20 a 73 anos (média = 37,9 anos) e entre as crianças de 2 a 11 (média = 7 anos). Cinco (62,5%) crianças eram do sexo masculino e 3 (37,5%) do feminino. Quanto ao estado civil das mulheres, 28,0% eram solteiras, 56% eram casadas ou tinham uma relação estável e 16,0% viúvas. A maioria tinha renda baixa (83,0%) e tinham menos de 4 anos de escolaridade (65,0%). A maioria das infecções foi relacionada à transmissão sexual (96,6%), enquanto uma relatou uso de droga intravenosa (UDI) (1,7%) e outra transfusão de sangue (1,7%). Todas as crianças foram infectadas por transmissão vertical durante a gravidez, parto ou amamentação. Quatro pacientes foram pares mãe-filho (BR88FSM131-BR131FSC88; BR91FSM83-BR83FSC91; BR100FSM145-BR145FSC100; BR110FSM111-BR111FSC110). De acordo com o CDC, 13 (22,4%) pacientes foram incluídos no estágio A, 39 (67,2%) no estágio B e 6 (10,3 %) no estágio C. Entre as mulheres, 10 (20%) foram incluídos no estágio A, 34 (68%) no estágio B e 6 (12%) no estágio C. Três crianças (37,5) estavam no estágio A e 5 (62,5%) no estágio B. Quarenta e oito (82,8%) pacientes estavam sob terapia antirretroviral, enquanto 10 (17,2%) eram virgens de tratamento.

A análise filogenética das sequências *pol* revelou que 39 (67,2%) eram isolados do subtipo B, 4 (6,9%) do F1 e 1 (1,7%) do C (Fig. 1). Quatorze (24,1%) amostras não ficaram dentro de nenhum cluster de subtipo puro. Doze (20,7%) dessas sequências formaram um grupo monofilético entre os subtipos B e F (Bootstrap = 70). As 14 amostras foram submetidos à análise de recombinação. As 12 amostras do cluster monofilético apresentaram o mesmo padrão de recombinação enquanto as outras duas amostras (BR35FS e BR50FS) apresentaram cada um, um único padrão de recombinação. Os

pontos de recombinação entre B e F1 foram investigados e a análise revelou que 10 das 12 amostras com o padrão BF1 partilhavam o mesmo ponto de recombinação entre as posições 2553 e 2584 (em relação ao genoma HXB2), enquanto as outras duas amostras (BR69FS e BR70FS) apresentaram interrupção entre as posições 2542-2572. As análises de NJ e ML dos fragmentos divididos em alinhamentos B e F1 confirmaram os resultados de recombinação pelo *Bootscreening*.

A fim de melhor caracterizar o vírus recombinante circulante na população, nós comparamos o seu padrão de recombinação genética com as sequências previamente caracterizadas no Brasil, bem como com as CRFs, e sequências puras B e F. A árvore NJ mostrou que as 10 amostras com o mesmo padrão de recombinação BF formaram um grupo monofilético separado dos puros subtipos B e F e dos clusters previamente caracterizados de CRFs.

Quanto à prevalência das mutações de resistências aos ARV dos isolados do HIV-1 de Feira de Santana, onze amostras (19,0%) apresentavam mutações associadas a alto nível de resistência dos quais: 10 (17,2%) aos INTR, 4 (6,9%) aos INNTR e 2 (3,4%) aos IP. As mutações relacionadas à resistência aos inibidores da TR foram: M184V (6), M41L (6), T69N (1), G190A (2), G190S (1) K103N (2), K101E (1) e Y181C (1); estas mutações conferem resistência à maioria dos ARV. Mutações na região da PR foram detectados nas sequências de 2 pacientes (BR100FSM145 e BR51FS): o N88D (1) e I50L (1) conferem resistência ao Nelfinavir (NFV) e Atazanavir (ATV), respectivamente. Entre os pacientes que tinham mutações de resistência, 10 (90,9%) pacientes apresentaram ao menos uma dos ARV no seu esquema, 2 (18,2%) tinham mutações para dois ARV e 2 (18,2%) tinham para três. O

número médio de mutações associadas com alto nível de resistência aos ARV por amostra foi de 0,4 entre as mulheres e 0,5 entre as crianças. Nós comparamos as características clínicas, epidemiológicas e genéticas entre o subtipo B e os recombinantes BF1. Pacientes infectados pelo subtipo B tinham, em média, 0,4 mutações de alto nível por sequência, enquanto recombinantes BF nenhuma mutação foi observada ( $p = 0,005$ ). Detalhes deste estudo podem ser encontrados no **manuscrito 03**.

## ***DISCUSSÃO DOS MANUSCRITOS***

No Brasil, seguindo uma tendência mundial, o início da pandemia foi restrito aos grandes centros urbanos, afetando grupos de risco específicos como: homossexuais, usuários de drogas injetáveis e hemofílicos, depois, a pandemia atingiu novas tendências. O HIV-1 difundiu-se para as cidades de médio a pequeno porte do Brasil, fenômeno este descrito como interiorização (MARTINEZ *et al.*, 2006). Neste contexto ainda, todas as categorias e faixas populacionais foram atingidas, ocorrendo uma feminilização da epidemia através da transmissão sexual (EYER-SILVA *et al.*, 2006), uma juvenilização pela transmissão vertical de mãe para filho e uma pauperização, tomando um contorno da epidemia verificada na África (RIOS 2003).

Estudos moleculares do HIV-1 são fundamentais para compreender melhor os mecanismos envolvidos na transmissão, diversidade e patogenia da AIDS e, assim, apoiar o desenvolvimento de vacinas e tratamentos mais eficazes. Porém, o número de sequências de HIV na região nordeste ainda é pequeno, bem como informações precisas associando a diversidade do HIV-1 com os dados clínicos.

Neste sentido, considerando a expansão do HIV-1 no Brasil, estudamos a diversidade molecular de sequências coletadas de pacientes infectados e randomizados em Salvador e Feira de Santana buscando associações com os dados clínicos e laboratoriais.

A cidade de Salvador apresenta um turismo empresarial e de lazer em desenvolvimento sendo a terceira maior cidade do Brasil. A cidade de Feira de Santana está no eixo de uma das malhas rodoviárias mais movimentadas do nordeste do Brasil, ligando todas as regiões do interior do país. Desta forma este estudo representa duas dinâmicas: a relacionada ao

turismo, especialmente ao de lazer, como o que ocorre nas altas temporadas e durante o carnaval, que ainda fornece um alto risco de transmissão; e o relacionado a interiorização, onde o empobrecimento, a desestruturação das famílias e a prostituição nas estradas do interior do Brasil esta sempre presente.

No Brasil, especialmente as regiões Sudeste, Centro-Oeste, Norte e Nordeste, o subtipo B continua a ser predominante, no entanto, na última década, diversos estudos têm relatado um aumento da prevalência de outros genótipos, especialmente recombinantes BF (BRENNAN *et al.*, 2007; MONTEIRO *et al.*, 2009). Em Salvador, encontramos entre as 61 amostras estudadas uma menor prevalência de BF1 (3,3%) em comparação com estudos recentes que têm demonstrado 10% a 13% de prevalência BF1 no Nordeste do Brasil (BRENNAN *et al.*, 2007; MONTEIRO *et al.*, 2009). Isto pode ser explicado pelo fato de que entre os dois fragmentos sequenciados, *gag* e *env*, poderíamos ter perdido outras regiões do genoma onde os eventos de recombinação poderiam ter ocorrido. Isto pode ser verificado no estudo de Feira de Santana onde as formas recombinantes BF1 identificadas no gene *pol* tiveram uma prevalência de 24%, a mais alta prevalência de BF1 encontrada até hoje no Brasil. Cabe ressaltar que apesar da presença de apenas uma sequência de subtipo C em Feira de Santana, esta pode ter vindo de outras regiões, especificamente a Sul, onde a prevalência deste subtipo é bem alta. No entanto, para se confirmar esta hipótese, é preciso incluir esta sequência num estudo de subtipo C circulante brasileiro.

Dentre as 14 formas recombinantes BF1 do estudo de Feira de Santana, 10 agruparam no mesmo cluster tanto na reconstrução filogenética

utilizando alinhamento com sequências de referências de subtipos puros quanto na reconstrução filogenética realizada com as formas recombinantes circulantes já catalogadas no banco de dados de Los Alamos. Estudos que envolvam o sequenciamento completo do genoma destes isolados de formas recombinantes BF1 que apresentam expansão nesta região geográfica podem contribuir para um melhor conhecimento da diversidade molecular circulante do HIV-1 neste local. Este estudo poderá responder se estas formas recombinantes encontradas em Feira de Santana fazem parte de um novo CRF em expansão no interior do Brasil e se esta região genômica é um *hotspot*, de ocorrência de recombinações como já relatado em outros estudos (NDEMBI *et al.*, 2004; MININ *et al.*, 2007).

No estudo de Salvador, a prevalência do subtipo B' foi inferior aos 38% encontrados na região amazônica (MACHADO *et al.*, 2009) e 50% na região Sudeste (DIAZ *et al.*, 2008), mas ainda superior ao encontrado na região Centro-Oeste de 6,7% (VÉRAS *et al.*, 2007). O estado da Bahia e a região Centro-Oeste aparecem assim como tendo uma provável introdução ou efeito fundador diferente das outras regiões Norte e Sudeste. Além disso, estratificamos a amostra em subtipo B e B', comparamos os dados clínicos e demonstramos que a média de idade de 49 anos e o tempo de diagnóstico de 13 anos dos pacientes infectados pelo subtipo B' foi superior à média de idade de 39 anos e tempo de diagnóstico de 9 anos das pessoas infectadas com o subtipo B, podendo estar relacionado a uma maior taxa de replicação e uma progressão mais rápida da doença de pacientes infectados pelo HIV-1 com motivo GPGR (SANTORO-LOPES *et al.*, 2000; CASSEB *et al.*, 2002; DE BRITO *et al.*, 2006). A lenta progressão para AIDS em si significaria maior

aptidão viral (*fitness*), devido a um longo período sem sintomas para a variante GWGR. Neste sentido, períodos mais longos, de fase assintomática, significariam maiores chances para os vírus infectar novos hospedeiros e, portanto, sobreviver e se reproduzir. No entanto observamos que a prevalência desta variante não vem crescendo. Um estudo de base populacional desta prevalência cruzando com dados clínicos ao longo do tempo pode responder melhor esta aparente discordância. Observamos também que não há distinção filogenética na topologia das árvores *gag* e *env* entre o subtipo B' e B, que é consistente com relato anterior (DIAZ *et al.*, 2008). Os valores da razão  $d_N/d_S$  no gene *env* do HIV-1 estavam de acordo com os encontrados em estudos anteriores (DIAZ *et al.*, 2008; PÉREZ-LOZADA *et al.*, 2009). As sequências do grupo GPGR apresentaram também neste estudo uma maior razão  $d_N/d_S$ , revelando assim uma pressão seletiva negativa menor se comparada aos outros grupos GWGR e GXGX avaliados.

Além da diferença de prevalência de subtipos observada entre as cidades estudadas. Essas análises mostraram um padrão diferenciado ao encontrado em estudos anteriores. Um estudo realizado na região nordeste do país (BRINDEIRO *et al.*, 2003) mostrou uma distribuição do subtipo B de 80%, BF de 10%, F1 de 6% e C de 4%. Outro estudo realizado em Salvador (BRENNAN *et al.*, 2007) mostrou um perfil mais próximo de 82,3% de B, 13% de BF, 3,5% de F e 1,1% de C.

Realizamos a prevalência de mutações de resistência aos antiretrovirais controladores da infecção pelo HIV-1 no gene *pol* do mesmo vírus caracterizada em Feira de Santana, além de demonstrarmos a mais alta prevalência de formas recombinantes BF1 e a mais baixa de subtipos B

encontrada no Brasil. Quarenta e oito dos 58 pacientes estavam em uso de terapia antirretroviral. A análise do padrão de resistência deste estudo encontrou diversas mutações nos *genes* da TR e PR em 95% dos isolados, sendo que a maioria delas não altera a sensibilidade dos vírus às drogas.

Um isolado de um paciente virgem de terapia (1/10, 10%) apresentou uma mutação primária de alta resistência aos INTR. Embora consistente com estudos anteriores para mutações de resistência na TR, não foi encontrado mutações de resistência aos INNTR e IP, no entanto este resultado pode representar um limite deste estudo, visto que o número de amostras virgens de tratamento foi pequeno, sendo assim, um maior número de pacientes recém infectados poderia ajudar a elucidar ou confirmar a prevalência de transmissão de mutações primárias de alta resistência. Estudos de prevalência de mutações primárias, ou seja, aquelas provenientes de vírus de pacientes sem terapia antirretroviral, mostraram variar em torno de 14 a 18% (PEDROSO *et al.*, 2007; EYER-SILVA *et al.*, 2008; DE SÁ FILHO *et al.*, 2009)

Dez isolados (21%) apresentaram mutações secundárias de alta resistência aos inibidores da TR e dentre estes 2 (4%) apresentaram mutações secundárias de alta resistência também aos inibidores da PR. A prevalência de resistência aos ARV encontrada neste estudo é considerada a esperada se comparada com a média nacional e mundial de estudos de amostras similares, que é de aproximadamente 14% para os ITR e 1,5% para os IP (COUTO-FERNANDEZ *et al.*, 2005; GIFFORD *et al.*, 2007; EYER-SILVA *et al.*, 2008).

Dentre os pacientes que apresentaram mutações de alta resistência, dois (49FS e 145FS) estavam em uso do mesmo esquema

terapêutico aos quais tinham resistência. No entanto, a carga viral e a contagem de linfócitos T CD4 permaneceram controladas. Isto pode ser explicado pelo fato de que o *fitness* viral cai após a aquisição de mutações de resistência, fazendo com que, as cepas se tornarem menos eficientes (NIJHUIS *et al.*, 2001). No entanto, para se confirmar esta hipótese com mais segurança no presente estudo, seria preciso um número amostral maior, sendo assim, nosso trabalho apresenta um fator limitante quanto ao número de sequências de pacientes sem tratamento.

Nós comparamos os indivíduos com subtipos B e formas recombinantes BF1 para encontrar diferenças entre a frequência média de idade, carga viral, da quantidade de linfócitos T CD4 e T CD8 e tempo de diagnóstico. Interessante constatar que a única diferença significativa foi na frequência média de mutações por sequência que foi de 0,41% entre o subtipo B e 0% entre as formas recombinantes BF1, ou seja, das 14 amostras recombinantes BF nenhuma apresentou mutações de alta resistência aos antiretrovirais. É importante ressaltar que não houve diferença entre a frequência média de idade, tempo de terapia ou de diagnóstico. Desta forma as diferenças estatísticas da frequência média de mutações de resistência não estão associadas a viés amostral, uma vez que uma frequência média de idade menor no grupo de pacientes infectados pela forma recombinante BF1 em comparação ao grupo subtipo B poderia explicar a falta de mutações de resistência em função do tempo, ou seja, quanto menor for a idade do grupo, menor a probabilidade de se encontrar o evento de resistência. Da mesma forma, a frequência média de tempo de tratamento e de idade. A associação entre os subtipos e surgimento de mutações de resistência é bem controversa.

Um estudo avaliando a emergência de mutações de resistência na cidade de Santos, sudeste do país (DE SÁ-FILHO *et al.*, 2008) demonstrou a ocorrência de prevalências igualmente semelhantes nos subtipos B, F e formas recombinantes BF, devido as pressões seletivas dos antiretrovirais. Por outro lado, num estudo de corte transversal com um número grande de pacientes com falha terapêutica na cidade de São Paulo, encontrou variação de mutações de resistência de acordo com o subtipo em todas as três principais classes de antirretrovirais, INTR, INNTR e IP (MUNERATO *et al.*, 2010). A expansão de formas recombinantes BF1 associada a não emergência de mutações de resistência nestas formas pode sugerir uma maior aptidão destas formas. Estudos de pressão seletiva *in sílica*, de aptidão viral (*fitness*) *in vitro* e de outras mutações de aminoácidos associados à resistência aos antirretrovirais ainda desconhecidas na literatura poderão elucidar melhor as características moleculares desses recombinantes. Nosso grupo já tem demonstrado inclusive uma expansão de possíveis novas formas recombinantes BF1 na cidade de Salvador, Bahia (MONTEIRO *et al.*, 2009). Além disso não se pode descartar a possibilidade destas formas recombinantes terem um risco maior de transmissão heterossexual que o subtipo B, mais prevalente no início da pandemia que tinha transmissão marcadamente homossexual.

Os resultados dos dois estudos aqui apresentados mostram uma menor diversidade genética na cidade de Salvador quando comparado a cidade de Feira de Santana, sobretudo considerando as formas recombinantes BF1. Um melhor esclarecimento se esta discrepância está relacionada à presença de pontos de recombinação em regiões genômicas diferentes das

analisadas, dependerá sequenciamento da mesma região *pol* nas amostras de Salvador e de *gag* e *env* nas de Feira de Santana. De fato, um estudo já em preparação pelo nosso grupo (SANTOS *et al.*, 2010) em 56 sequências do gene *pol* de amostras de pacientes atendidos em Salvador demonstrou a prevalência de subtipo B de 80,4%, subtipo F1 de 5,4% e de BF1 de 14,3%.

A diversidade nestes estudos demonstram no seu conjunto à necessidade de uma contínua vigilância epidemiológica através principalmente da subtipagem e genotipagem. Este monitoramento molecular de subtipos, das mutações de resistência, da expansão da infecção pelo HIV-1 e suas formas de transmissão é crucial para dar suporte a políticas de saúde pública e desenvolvimento de novas terapias e vacinas.

## **CONCLUSÕES**

No estudo em Salvador, a prevalência de subtipos mostrou menor diversidade do que estudos anteriores na mesma cidade.

A prevalência da variante brasileira do subtipo B (B') em Salvador ficou abaixo da média encontrada no Sudeste.

Os pacientes com a variante brasileira B' apresentaram idade e tempo de diagnóstico médio maiores que o subtipo B.

Os pacientes com a variante R5X4 ou X4 tiveram idades menores que aqueles com R5.

As sequências do motivo GWGR apresentaram pouca diversidade em relação as sequências de outros motivos da alça V3 e pressão seletiva negativa.

No estudo em Feira de Santana, a prevalência de formas recombinantes BF1 do HIV-1 revelou ser a mais alta encontrada no Brasil.

A prevalência de mutações de resistência aos antirretrovirais encontradas em Feira de Santana estão de acordo com estudos anteriores realizados na Bahia e no Brasil.

Pacientes com recombinantes BF1 do HIV-1 não apresentaram mutações de resistência.

Em geral, a diversidade molecular de Salvador é menor do que a de Feira de Santana: a prevalência de subtipo B do HIV-1 foi mais alta e a de formas recombinantes BF1 mais baixa em Salvador quando comparada a cidade de Feira de Santana.

## ***REFERÊNCIAS BIBLIOGRÁFICAS***

- Alaeus A. Significance of HIV-1 genetic subtypes. *Scand J Infect Dis.* 2000.32(5): 455-63.
- Arroyo MA, Hoelscher M, Sanders-Buell ES, Herbinger KH, Samki E, Maboko L, Hoffmann O, Robb MR, Birx DL, McCutchan FE.. HIV type 1 subtypes among blood donors in the Mbeya region of southwest Tanzania. *AIDS Res Hum Retroviruses.* 2004;20(8):895-901.
- Bailes E. The evolution of primate lentiviruses and origins of AIDS, in *The molecular epidemiology of human viruses*, T. A. Leitner, Editor. 2002, Kluwer Academic Publishers: Boston, Mass. p. 65-96.
- Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. Isolation of a T-Lymphotropic Retrovirus from a patient at risk for Acquired Immune Deficiency Syndrome (AIDS). *Science.* 1983.Apr; 220: 868-70.
- Barreto CC, Nishyia A, Araujo LV, Ferreira JE, Busch MP, Sabino EC. Trends in antiretroviral drug resistance and clade distributions among HIV-1–infected blood donors in Sao Paulo, Brazil. *J Acquir Immune Defic Syndr* 2006, 41:338-341.
- Bastos FI & Szwarcwald CL. AIDS and pauperization: principal concepts and empirical evidence *Cad. Saúde Pública*, 2000. Rio de Janeiro. 16(Sup. 1): 65-76.

- Bastos FI, Cáceres C, Galvão J, Veras MA, Castilho EA. AIDS in Latin America: assessing the current status of the epidemic and the ongoing response. *Int J Epidemiol* 2008; 37(4): 729-737.
- BE - Boletim Epidemiológico - Aids e DST. Ano V - nº 1 - 01<sup>a</sup> - 26<sup>a</sup> - semanas epidemiológicas - janeiro a junho de 2008.
- BE - Boletim Epidemiológico AIDS/DST - Ano IV - nº 1 - 01<sup>a</sup> - 26<sup>a</sup> - semanas epidemiológicas - janeiro a junho de 2007.
- BE - Boletim Epidemiológico AIDS/DST - Ano VI - nº 1 - 01<sup>a</sup> - 29<sup>a</sup> - semanas epidemiológicas - janeiro a junho de 2009.
- BE - Boletim Epidemiológico. Situação Epidemiológica da AIDS Estado da Bahia. *Suvisa/Divep*. Ano 1, Nº2, 1º de Agosto de 2009.
- Bello G, Eyer-Silva WA, Couto-Fernandez JC, Guimarães ML, Chequer-Fernandez SL, Teixeira SL, Morgado MG. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect Genet Evol*. 2007. 7(2):263-70.
- Bird C, Burke J, Gleeson PA, McCluskey J. Expression of human immunodeficiency virus 1 (HIV-1) envelope gene products transcribed from a heterologous promoter. Kinetics of HIV-1 envelope processing in transfected cells. *J Biol Chem*. 1990. 5;265(31):19151-7.
- Bongertz V, Bou-Habib DC, Brigido LF, Caseiro M, Chequer PJN, Couto-Fernandez JC, Ferreira PC, Galvão-Castro B, Greco D, Guimarães ML, Linhares de Carvalho MI, Morgado MG, Oliveira CA, Osmanov S, Ramos CA, Rossini M, Sabino E, Tanuri A, Ueda M. HIV-1 diversity in Brazil: genetic, biologic, and immunologic characterization of HIV-1 strains in

three potential HIV vaccine evaluation sites. *J Acquir Immune Defic Syndr*. 2000. 23(2):184-93.

Brennan CA, Brites C, Bodelle P, Golden A, Hackett J Jr, Holzmayer V, Swanson P, Vallari A, Yamaguchi J, Devare S, Pedroso C, Ramos A, Badaro R.. HIV-1 strains identified in Brazilian blood donors: significant prevalence of B/F1 recombinants. *AIDS Res Hum Retroviruses*. 2007. 23(11):1434-41.

Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, Dantas MC, Barreira D, Teixeira PR, Tanuri A, Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet). A survey of chronically infected individuals. *AIDS* 2003.17: 1063–1069.

Campos LN, César CC, Guimarães MD. Quality of life among HIV-infected patients in Brazil after initiation of treatment. *Clinics (Sao Paulo)*. 2009. 64(9):867-75.

Cao J, Sullivan N, Desjardin E, Parolin C, Robinson J, Wyatt R, Sodroski J. Replication and neutralization of human immunodeficiency virus type 1 lacking the V1 and V2 variable loops of the gp120 envelope glycoprotein. *J Virol* 1997. 71(12): 9808-12.

Cardoso LP, Queiroz BB, Stefani MM. HIV-1 pol phylogenetic diversity and antiretroviral resistance mutations in treatment naïve patients from Central West Brazil. *J Clin Virol*. 2009 Oct;46(2):134-9.

- Cardozo T, Kimura T, Philpott S, Weiser B, Burger H, Zolla-Pazner S. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS Res Hum Retroviruses* 2007. 23(3): 415-26.
- Caride E, Brindeiro R, Hertogs K, Larder B, Dehertogh P, Machado E, de Sá CA, Eyer-Silva WA, Sion FS, Passioni LF, Menezes JA, Calazans AR, Tanuri A. Drug-resistance reverse transcriptase genotyping and phenotyping of B and non-B subtypes (F and A) of human immunodeficiency virus type 1 found in brazilian patients failing HAART. *Virology*. 2000;275(1)107-15.
- Carr JK, Avila M, Gomez-Carrillo M, Salomon H, Hierholzer J, Pando M, Pando MA, Negrete M, Russell KL, Sanchez J, Birx DL, Andrade R, Vinales J, McCutchan FE. Diverse BF recombinants have spread widely since introduction of HIV-1 into South America. *AIDS*. 2001;15(15):F41-7.
- Carr JK, Salminen MO, Koch C, Gotte D, Artenstein AW, Hegerich PA, St Louis D, Burke DS, McCutchan FE.. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. *J Virol* 1996; 70:5935–43.
- Casseb J, Komninakis S, Abdalla L, Brigido LF, Rodrigues R, Araújo F, Veiga AP, de Almeida A, Flannery B, Hendry RM, Duarte AJ. HIV disease progression: is the Brazilian variant subtype B' (GWGR motif) less pathogenic than US/European subtype B (GPGR)? *Int J Infect Dis* 2002. 6: 164-169.
- Cassol S, Weniger BG, Babu PG, Salminen MO, Zheng K, Htoon MT, Delaney A, O'Shaughnessy M, Ou CY.. Detection of HIV type 1 env subtypes A,

B, C, and E in Asia using dried blood spots: a new surveillance tool for molecular epidemiology. *AIDS Res Hum Retroviruses*. 1996;12(15):1435-41.

Castilho EA, Bastos FI, Scwarcwald CL, Fonseca MGM. AIDS in Brazil: a changing epidemic. *Cad. Saúde Pública*. Rio de Janeiro. 16(Sup. 1): 4-5. 2000.

Castro E, Echeverría G, Deibis L, González de Salmen B, Dos Santos Moreira A, Guimarães ML, Bastos FI, Morgado MG.. Molecular epidemiology of HIV-1 in Venezuela: high prevalence of HIV-1 subtype B and identification of a B/F recombinant infection. *J Acquir Immune Defic Syndr* 2003. 32(3):338-44.

Cavalcanti AM, Lacerda HR, Brito AM, Pereira S, Medeiros D, Oliveira S. Antiretroviral resistance in individuals presenting therapeutic failure and subtypes of the human immunodeficiency virus type 1 in the Northeast Region of Brazil. *Mem Inst Oswaldo Cruz*. 2007 102(7):785-92.

Cavarelli M & Scarlatti G. Phenotype variation in human immunodeficiency virus type 1 transmission and disease progression. *Dis Markers*. 2009; 27(3):121-36.

*Centers for Disease Control and Prevention - CDC*. Task Force on Kaposi's Sarcoma and Opportunistic Infections *Engl J Med* 1982. 306: 248.

Chakrabarti S, Panda S, Chatterjee A, Sarkar S, Manna B, Singh NB, Naik TN, Detels R, Bhattacharya SK.. HIV-1 subtypes in injecting drug users &

their non-injecting wives in Manipur, India. *Indian J Med Res* 2000; 111:189-94.

Chiou SH, Freed EO, Panganiban AT, Kenealy WR. Studies on the role of the V3 loop in human immunodeficiency virus type 1 envelope glycoprotein function. *AIDS Res Hum Retroviruses* 1992. 8(9): 1611-8.

CN-DST/AIDS. Coordenação Nacional de DST e AIDS. Aconselhamento em DST, HIV e AIDS: diretrizes e procedimentos básicos. / 2ª ed. Brasília: Ministério da Saúde, 1998. 25 p.

Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, Temin H, Toyoshima K, Varmus H, Vogt P, Weiss R. Human immunodeficiency viruses. *Science* 1986. 9; 232(4751): 697.

Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 1995. 483-9.

Couto-Fernandez JC, Eyer-Silva WA, Guimarães ML, Chequer-Fernandez SL, Grinsztejn B, Delaporte E, Peeters M, Morgado MG. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: tracing the origin of this subtype in Brazil. *AIDS Res Hum Retroviruses* 2006. 22(2):207-11.

Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, Oliveira SM, Arakaki-Sanchez D, Chequer PJ, Morgado MG. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio de Janeiro, Brazil: assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active

antiretroviral therapy. Mem Inst Oswaldo Cruz. 2005 Feb;100(1):73-8.  
Epub 2005 Apr 12.

Couto-Fernandez JC, Morgado MG, Bongertz V, Tanuri A, Andrade T, Brites C, Galvão-Castro B. HIV-1 subtyping in Salvador, Bahia, Brazil: a city with african sociodemographic characteristics. J Acquir Immune Defic Syndr 1999. 22(3):288-93.

Da Costa SM, Schechter M, Shindo N, Vicente AC, Oliveira EF, Pinto ME, Tanuri A. Sequence and phylogenetic analysis of glycoprotein 120 of an HIV type 1 variant (GWGR) prevalent in Brazil. AIDS Res Hum Retroviruses 1995. 11(9):1143-5.

De Brito A, Komninakis SC, Novoa P, de Oliveira RM, Fonseca LA, Duarte AJ, Casseb J. Women infected with HIV type 1 Brazilian variant, subtype B (B'-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). Clin Infect Dis 2006; 43(11): 1476-1481.

De Queiroz AT, Mota-Miranda AC, de Oliveira T, Moreau DR, Urpia Cde C, Carvalho CM, Galvão-Castro B, Alcantara LC. Re-mapping the molecular features of the human immunodeficiency virus type 1 and human T-cell lymphotropic virus type 1 Brazilian sequences using a bioinformatics unit established in Salvador, Bahia, Brazil, to give support to the viral epidemiology studies. Mem Inst Oswaldo Cruz 2007 May; 102(2):133-9.

De Sa-Filho DJ, Ambar RF, Duarte NB, Matias RB, Candido V, Gagliani LH, Caseiro MM. HIV type 1 diversity from newly diagnosed patients in

Santos metropolitan area/Brazil. *AIDS Res Hum Retroviruses*. 2009 Sep;25(9):925-9.

Diaz RS, Leal E, Sanabani S, Sucupira MC, Tanuri A, Sabino EC, Janini LM. Selective regimes and evolutionary rates of HIV-1 subtype B V3 variants in the Brazilian epidemic. *Virology* 2008. 381(2):184-93.

Do Lago RF, Costa Ndo R. Antiretroviral manufacturers and the challenge of universal access to drugs through the Brazilian National STD/AIDS Program. *Cad Saude Publica* 2009. 25(10):2273-84.

Donahue JP, Vetter ML, Mukhtar NA, D'Aquila RT. The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology* 2008. 377 (1): 49–53.

Dourado I, Milroy CA, Mello MA, Ferraro GA, Castro-Lima-Filho H, Guimarães ML, Morgado MG, Teixeira MG, Barreto ML, Galvão-Castro B. HIV-1 seroprevalence in the general population of Salvador, Bahia State, Northeast Brazil. *Cad de Saúde Pública* 2007. 23(1).

Edgar, RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research* 2004. 32(5), 1792-97.

Esparza J & Bhamarapavati N. Accelerating the development and future availability of HIV-1 vaccines: Why, when, where, and how? *Lancet* 2000. 355:2061–2066.

Eyer-Silva WA & Morgado MG. Molecular epidemiology of HIV-1 infection in a small Brazilian county: usefulness of envelope and polymerase

sequences to epidemiologic studies. *J Acquir Immune Defic Syndr* 2006. 41:664-70.

Eyer-Silva WA, Couto-Fernandez JC, Silva-de-Jesus C, Morgado MG. Prevalence of HIV type 1 drug resistance mutations in treatment-naïve and experienced patients from resource-limited settings with universal access to antiretroviral therapy: a survey in two small Brazilian cities. *Mem Inst Oswaldo Cruz*. 2008 Mar;103(2):143-9.

Felsenstein J, *Inferring phylogenies*. Sinauer Associates, Massachusetts, 2004. P. 580.

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985. 39:783-791.

Felsenstein, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 1981. 17:368-376.

Frater AJ, Dunn DT, Beardall AJ, Ariyoshi K, Clarke JR, McClure MO, Weber JN. Comparative response of African HIV-1-infected individuals to highly active antiretroviral therapy. *AIDS* 2002; 16:1139–1146.

Gadelha SR, Shindo N, Monte Cruz JN, Morgado MG, Galvão-Castro B. Molecular epidemiology of human immunodeficiency virus-1 in the state of Ceará, northeast, Brazil. *Mem Inst Oswaldo Cruz* 2003. 98(4):461-3.

Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalianaraman VS, Mann D, Sidhu GD, Stahl RE, Zolla-Pazner S, Leibowith J, Popovic M. Isolation of Human T-Cell Leukemia Virus in

- Acquired Immune Deficiency Syndrome (AIDS). Science 1983. 220: 865-7.
- Galvani AP. The role of mutation accumulation in HIV progression. Proc Biol Sci 2005; 272(1574): 1851-8.
- Gao F, Korber BT, Weaver E, Liao HX, Hahn BH, Haynes BF. Centralized immunogens as a vaccine strategy to overcome HIV-1 diversity. Expert Rev Vaccines 2004. 3(4 Suppl):S161-8.
- Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, Fultz PN, Girard M, Shaw GM, Hahn BH, Sharp PM.. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J Virol 1996;70:7013–29.
- Gao F., Yue L, White AT, Pappas PG, Barchue J, Hanson AP, Greene BM, Sharp PM, Shaw GM, Hahn BH. Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. Nature 1992. 358(6386): p. 495-9.
- Gao F., et al., Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. Nature 1999. 397(6718): p. 436-41.
- Gascuel O, Steel M. "Neighbor-joining revealed". Mol Biol Evol 2006. 23(11): 1997–2000.
- Gifford RJ, de Oliveira T, Rambaut A, Pybus OG, Dunn D, Vandamme AM, Kellam P, Pillay D; UK Collaborative Group on HIV Drug Resistance. Phylogenetic surveillance of viral genetic diversity and the evolving

molecular epidemiology of human immunodeficiency virus type 1. *J Virol* 2007. 81(23):13050-6.

Gonzalez CR, Alcalde R, Nishiya A, Barreto CC, Silva FE, de Almeida A, Mendonça M, Ferreira F, Fernandes SS, Casseb J, Duarte AJ. Drug resistance among chronic HIV-1-infected patients naïve for use of anti-retroviral therapy in Sao Paulo city. *Virus Res.* 2007. 129(2):87-90.

Gordon S, Pandrea I, Dunham R, Apetrei C, Silvestri G. The Call of the Wild: What Can Be Learned from Studies of SIV Infection of Natural Hosts? pp. 2-29 in *HIV Sequence Compendium 2005*. Edited by: Thomas Leitner T, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. LA-UR 06-0680.

Göttlinger HG. 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.

Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 2010. 27(2):221-4.

Hacker MA, Kaida A, Hogg RS, Bastos FI. The first ten years: achievements and challenges of the Brazilian program of universal access to HIV/AIDS comprehensive management and care, 1996-2006. *Cad Saude Publica.* 2007; 23 Suppl 3:S345-59.

- Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. *Science* 2000. 287(5453): p. 607-14.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser* 1999. 41:95-98.
- Hare CB. WHO Staging System for HIV Infection and Disease in Adolescents and Adults, Table 3; Clinical Overview of HIV Disease. In: Peiperl L, Coffey S, Volberding PA, eds. *HIV InSite Knowledge Base*[textbook online]. San Francisco: UCSF Center for HIV Information; 2006. Accessed March 23, 2006. [www.hivinsite.ucsf.edu](http://www.hivinsite.ucsf.edu).
- Hasegawa, M., Kishino, H and Yano, T. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 1985. 22:160-174.
- Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006. 24;20(16):W13-23.
- Herring BL, Ge YC, Wang B, Ratnamohan M, Zheng F, Cunningham F, Saksena NK, Dwyer DE.. Segregation of human immunodeficiency virus type 1 subtypes by risk factor in Australia. *J Clin Microbiol* 2003;41(10):4600-4.
- Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* 1989. 339(6223): p. 389-92.

Holguin A, Álvarez A, Soriano V. HIV-1 subtype J recombinant viruses in Spain.

AIDS Res Hum Retroviruses 2002;18(7):523-9.

Hu WS, Temin HM. Genetic consequences of packaging two RNA genomes in

one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc Natl Acad Sci U S A 1990. 87(4):1556-60.

Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped

catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 1998. 82(5394):1669-75.

Huelsenbeck JP & Rannala B. Phylogenetic methods come of age: testing

hypotheses in an evolutionary context. Science. 1997 Apr 11;276(5310):227-32.

IBGE, 2007, em parceria com os Órgãos Estaduais de Estatística, Secretarias

Estaduais de Governo e Superintendência da Zona Franca de Manaus - SUFRAMA.

Johnson PR & Hirsch VM. SIV infection of macaques as a model for AIDS

pathogenesis. Int Rev Immunol 1992. 8(1): p. 55-63.

Jonassen TO, Stene-Johansen K, Berg ES, Hungnes O, Lindboe CF, Frøland

SS, Grinde B. Sequence analysis of HIV-1 group O from norwegian patients infected in the 1960s. Virology 1997;231(1):43-7.

Jukes TH & Cantor CR. Evolution of protein molecules. Pp. 21-123 in H. N.

Munro, ed. Mammalian protein metabolism 1969. Academic Press, New York.

- Kalish ML, Robbins KE, Pieniazek D, Schaefer A, Nzilambi N, Quinn TC, St Louis ME, Youngpairoj AS, Phillips J, Jaffe HW, Folks TM. Recombinant viruses and early global HIV-1 epidemic. *Emerg Infect Dis.* 2004 Jul;10(7):1227-34.
- Kedzierska K, Crowe SM, Turville S, Cunningham AL. The influence of cytokines, chemokines and their receptors on HIV-1 replication in monocytes and macrophages. *Rev Med Virol* 2003.13(1):39-56.
- Kimura M. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* 1980. 16:111-120.
- Kitsutani PT, Naganawa S, Shiino T, Matsuda M, Honda M, Yamada K, Taki M, Sugiura W.. HIV type 1 subtypes of nonhemophiliac patients in Japan. *AIDS Res Hum Retroviruses* 1998;14(12):1099-103.
- Koch N, Ndiokubwayo JB, Yahi N, Tourres C, Fantini J, Tamelet C. Genetic analysis of HIV type 1 strains in Bujumbura (Burundi): predominance of subtype C variant. *AIDS Res Hum Retroviruses* 2001;17(3):269-73.
- Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. *Brit. Med. Bull* 2001. 58:19-42.
- Korber B. HIV Signature and Sequence Variation Analysis. *Computational Analysis of HIV Molecular Sequences*, Chapter 4, pages 55-72. Allen G. Rodrigo and Gerald H. Learn, eds. Dordrecht, Netherlands: Kluwer Academic Publishers 2000.

- Kupek EM. HIV seroprevalence among blood donors in southern Brazil in the decade of 1990. *Braz J Infect Dis* 2000. 4:217-25.
- Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 2007. 31;8:424.
- Levy JA, Hoffman AD, Kramer SM, Landis JA, Shimabukuro JM, Oshiro LS. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 1984. 225(4664): 840-2.
- Louwagie J, Janssens W, Mascola J, Heyndrickx L, Hegerich P, McCutchan FE, Burke DS. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. *J Virol* 1995. 69(1):263-71.
- Lu M, Blackow S, Kim P. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature Struct Biol.* 1995. 2: 1075–82.
- Machado LF, Ishak MO, Vallinoto AC, Lemos JA, Azevedo VN, Moreira MR, Souza MI, Fernandes LM, Souza LL, Ishak R. Molecular epidemiology of HIV type 1 in northern Brazil: identification of subtypes C and D and the introduction of CRF02\_AG in the Amazon region of Brazil. *AIDS Res Hum Retroviruses* 2009. 25(10):961-6.
- Markowitz M, Louie M, Hurley A, Sun E, Di Mascio M, Perelson AS, Ho DD. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *J Virol* 2003. 77:5037-8.

- Martinez AMB, Hora VP, Santos AL, Mendonza-Sasi R, Groll AZ, Soarez EAJM, D'Ávila N, Silveira J, Leal RG, Tanuri A, Soares MA. Determinants of HIV-1 Mother-to-Child transmission in Southern Brazil. *Annals of the Brazilian Academy of Sciences* 2006. 78(1): 113-121.
- Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008). *J Int AIDS Soc* 2009.12(1):11.
- Marx PA, Li Y, Lerche NW, Sutjipto S, Gettie A, Yee JA, Brotman BH, Prince AM, Hanson A, Webster RG, Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a west African pet sooty mangabey. *J. Virol* 1991. 65(8): p. 4480-5.
- Mellors JW, Muñoz A, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, Kingsley LA, Todd JA, Saah AJ, Detels R, Phair JP, Rinaldo CR Jr. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Annals of Internal Medicine* 1997. 946-54.
- Mercenne G, Bernacchi S, Richer D, Bec G, Henriët S, Paillart JC, Marquet R. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res* 2009.
- Minin VN, Dorman KS, Fang F, Suchard MA. Phylogenetic mapping of recombination hotspots in human immunodeficiency virus via spatially smoothed change-point processes. *Genetics*. 2007 Apr;175(4):1773-85. Epub 2006 Dec 28.

- Monno L, Brindicci G, Caputo SL, Punzi G, Scarabaggio T, Riva C, Di Bari C, Pierotti P, Saracino A, Lagioia A, Mazzotta F, Balotta C, Angarano G. HIV-1 subtypes and circulating recombinant forms (CRFs) from HIV infected patients residing in two regions of central and southern Italy. *J Med Virol.* 2005;75(4):483-90.
- Monteiro JP, Ferraro GA, Oliveira T, Goldani LZ, Kashima S, Alcantara LC, Morgado MG, Bou-Habib DC, Galvão-Castro B. Genetic and biologic characterization of HIV type 1 subtype C isolates from south Brazil. *AIDS Res Hum Retroviruses.* 2007. 23(1):135-43.
- Morgado MG, Guimarães ML, Galvão-Castro B. HIV-1 polymorphism: a challenge for vaccine development - a review. *Mem Inst Oswaldo Cruz* 2002. 97(2): 143-50.
- Morgado MG, Guimarães ML, Gripp CB, Costa CI, Neves I Jr, Veloso VG, Linhares-Carvalho MI, Castello-Branco LR, Bastos FI, Kuiken C, Castilho EA, Galvão-Castro B, Bongertz V.. Molecular epidemiology of HIV-1 in Brazil: high prevalence of HIV-1 subtype B and identification of an HIV-1 subtype D infection in the city of Rio de Janeiro, Brazil. *J Acquir Immune Defic Syndr Hum Retrovirol.* 1998.18(5):488-94.
- MT – Ministério do Turismo. Anuário Estatístico de Turismo 2009. Vol. 36, ano base 2009.
- Munerato P, Sucupira MC, Oliveros MP, Janini LM, de Souza DF, Pereira AA, Inocencio LA, Diaz RS. HIV type 1 antiretroviral resistance mutations in subtypes B, C, and F in the City of São Paulo, Brazil. *AIDS Res Hum Retroviruses.* 2010 Mar;26(3):265-73.

Myers G. HIV: between past and future. *AIDS Res Hum Retroviruses* 1994.10(11):1317-24.

Nabatov AA, Kravchenko ON, Lyulchuk MG, Scherbinskaya AM, Lukashov V. Simultaneous introduction of HIV type 1 subtype A and B viruses into injecting drug users in southern Ukraine at the beginning of the epidemic in the former Soviet Union. *AIDS Res Hum Retroviruses* 2002. 18(12):891-5.

Naidu YM, Kestler HW 3rd, Li Y, Butler CV, Silva DP, Schmidt DK, Troup CD, Sehgal PK, Sonigo P, Daniel MD. Characterization of infectious molecular clones of simian immunodeficiency virus (SIVmac) and human immunodeficiency virus type 2: persistent infection of rhesus monkeys with molecularly cloned SIVmac. *J Virol* 1988. 62(12):4691-6.

Ndembi N, Takehisa J, Zekeng L, Kobayashi E, Ngansop C, Songok EM, Kageyama S, Takemura T, Ido E, Hayami M, Kaptue L, Ichimura H. Genetic diversity of HIV type 1 in rural eastern Cameroon. *J Acquir Immune Defic Syndr*. 2004 Dec 15;37(5):1641-50

Nicholas KB, Nicholas Jr HB, Deerfield II DW. GeneDoc: Analysis and Visualization of Genetic Variation *EMBNET news* 1997. 4:1-4.

Nijhuis M, Deeks S, Boucher C. Implications of antiretroviral resistance on viral fitness. *Current Opinion in Infectious Disease* 2001. 23 – 28.

Nunn AS, Fonseca EM, Bastos FI, Gruskin S, Salomon JA. Evolution of antiretroviral drug costs in Brazil in the context of free and universal access to AIDS treatment. *PLoS Med* 2007. 13; 4(11): 305.

- Overbaugh J, Bangham CRM. Selection forces and constraints on retroviral sequence variation. *Science* 2001. 292:1106-9.
- Palmer S, Alaeus A, Albert J, Cox S. Drug susceptibility of subtypes A,B,C,D, and E human immunodeficiency virus type 1 primary isolates. *AIDS Res Human Retroviruses* 1998. 14:157–162.
- Pantaleo G, Graziosi C, Fauci AS. The immunopathogenesis of human immunodeficiency virus infection. *N Engl J Med* 1993a;328(5):327-35
- Parker R, Galvão J. Quebrando o silêncio: mulheres e Aids no Brasil. Rio de Janeiro: Relume-Dumará: ABIA:IMS/UERJ 1996. 1(2): 33.
- Pedroso C, Queiroz AT, Alcântara LC, Drexler JF, Diaz RS, Weyll N, Brites C.. High Prevalence of Primary Antiretroviral Resistance Among HIV-1–Infected Adults and Children in Bahia, a Northeast State of Brazil. *J Acquir Immune Defic Syndr* 2007. 45:251-3.
- Peeters M, Honoré C, Huet T, Bedjabaga L, Ossari S, Bussi P, Cooper RW, Delaporte E. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *AIDS* 1989. 3(10): p. 625-30.
- Peeters M & Sharp PM. Genetic diversity of HIV-1: the moving target. *AIDS* 2000.14 Suppl 3:S129-240.
- Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 1996. 1582-6.

- Pieniazek D, Rayfield M, Hu DJ, Nkengasong J, Wiktor SZ, Downing R. Protease sequences from HIV-1 group M subtypes A–H reveal distinct amino acid mutation patterns associated with protease resistance in protease inhibitor-naive individuals worldwide. *AIDS* 2000. 14:1489–1495.
- Pinto ME & Struchiner CJ. A diversidade do HIV-1: uma ferramenta para o estudo da pandemia. *Caderno de Saúde Pública* 2006. 473-84.
- Pomerantz RJ, Trono D. Genetic therapies for HIV infections: promise for the future. *AIDS* 1995. 985-93.
- Popovic M, Sarngadharan MG, Read E, Gallo RC. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 1984. 224(4648): 497-500.
- Posada D & Krandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998. 14: 817-818.
- Ramos A, Tanuri A, Schechter M, Rayfield MA, Hu DJ, Cabral MC, Bandea CI, Baggs J, Pieniazek D.. Dual and recombinant infections: an integral part of the HIV-1 epidemic in Brazil. *Emerg Infect Dis* 1999. 5(1):65-71.
- Requejo HIZ. Worldwide molecular epidemiology of HIV. *Rev Saúde Pública* 2006. 40(2):331-45.
- Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* 2003. 100(7):4144-9.

- Rios LF Juvenilização do HIV/AIDS por via sexual. *Boletim ABIA*, Rio de Janeiro, v. 49, 2003.
- Rios M, Villanueva C, Ramirez E. Identification of B and F human immunodeficiency virus subtypes in Chilean patients. *Rev Med Chil* 2003.131(7):711-8.
- Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Kalish ML, Salminen M, Sharp PM, Wolinsky S and Korber B. HIV-1 Nomenclature Proposal. pp. 492-505 in *Human Retroviruses and AIDS 1999*. Edited by: Kuiken CL, Foley B, Hahn B, Korber B, McCutchan F, Marx PA, Mellors JW, Mullins JI, Sodroski J, and Wolinsky S. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
- Rodrigues R, Scherer LC, Oliveira CM, Franco HM, Sperhacke RD, Ferreira JL, Castro SM, Stella IM, Brigido LF Low prevalence of primary antiretroviral resistance mutations and predominance of HIV-1 clade C at polymerase gene in newly diagnosed individuals from south Brazil. *Virus Res* 2006. 116:201–207.
- Rodríguez F, Oliver JL, Marín A, Medina JR. The general stochastic model of nucleotide substitution. *J Theor Biol* 1990.142(4):485-501.
- Sá MS, Sampaio J, Haguihara T, Ventin FO, Brites C. Clinical and laboratory profile of HIV-positive patients at the moment of diagnosis in Bahia, Brazil. *Braz J Infect Dis* 2007. 11(4): 395-8. 2007.

- Sabino EC, Diaz RS, Brigido LF, Learn GH, Mullins JI, Reingold AL. Distribution of HIV-1 subtypes seen in an AIDS clinic in São Paulo city, Brazil. *AIDS* 1996.10(13):1579-84.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987. 4(4):406-25.
- Salemi M, de Oliveira T, Soares MA, Pybus O, Dumans AT, Vandamme AM, Tanuri A, Cassol S, Fitch WM. Different epidemic potentials of the HIV-1B and C subtypes. *J Mol Evol* 2005. 60:598–605.
- Sanabani SS, Pastena ERS, Neto WK, Barreto CC, Ferraria KT, Kalmar EMN, Ferreira S, Sabino EC. Near full-length genome analysis of low prevalent human immunodeficiency virus type 1 subclade F1 in São Paulo, Brazil. *Virol J* 2009. 6:78.
- Santoro-Lopes G, Harrison LH, Tavares MD, Xexeo A, Dos Santos AC, Schechter M.. HIV disease progression and V3 serotypes in Brazil: is B different from B-Br? *AIDS Res Hum Retroviruses* 2000. 16: 953-958.
- Santos AF, Sousa TM, Soares EA, Sanabani S, Martinez AM, Sprinz E, Silveira J, Sabino EC, Tanuri A, Soares MA. Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS* 2006. 20:2011–2019.
- Santos ES, Araújo AF, Galvão-Castro B, Alcantara LC Jr. Genetic diversity of human immunodeficiency virus type-1 (HIV-1) in infected women from a northeast city of Brazil. *Rev Bras Ginecol Obstet* 2009. 31(12):609-14.

- Santos ES. Variabilidade Genética dos Isolados do HIV-1 em Mulheres e Crianças de Feira de Santana. Salvador 2010. Dissertação apresentada à Fundação Bahiana para Desenvolvimento da Ciência – Escola Bahiana de Medicina e Saúde Pública.
- Santos LA, Araujo AF, Brites C, Galvão-Castro B, Alcantara LC. Diversidade e mutações de resistência de sequências do HIV-1 de pacientes atendidos num centro de referência em retrovírus na cidade de Salvador, Brasil. 2010. (em preparação).
- Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol.* 2009 Jan 23;385(3):693-713. Epub 2008 Nov 3. Review.
- Schneider H. Métodos de Análise Filogenética: Um Guia Prático. Segunda Edição. Ribeirão Preto: 2003. P. 114.
- Sepkowitz KA. AIDS: the first 20 years. *N Engl J Med* 2001 Jun 7;344(23):1764-72..
- Se-Thoe SY; Foley BT; Chan SY; Lin RV; Oh HM; Ling AE; Chew SK; Snodgrass I; Sng JE. Analysis of sequence diversity in the C2V3 regions of the external glycoproteins of HIV type 1 in Singapore. *AIDS Res Hum Retroviruses* 1998.14(17):1601-4.
- Seve M, Favier A, Osman M, Hernandez D, Vaitaitis G, Flores NC, McCord JM, Flores SC. The human immunodeficiency virus-1 Tat protein increases cell proliferation, alters sensitivity to zinc chelator-induced apoptosis, and

changes Sp1 DNA binding in HeLa cells. Arch Biochem Biophys 1999. 361(2):165-72.

Sharp PM, Bailes E, Chaudhuri RR, Rodenburg CM, Santiago MO, Hahn BH. Evolutionary origin of human and simian immune deficiency viruses. Proc Natl Acad Sci USA. 1990. 87(11):4108-11.

Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 2002. 418 (6898): 646–50.

Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues F, Büch J, Däumer M, Kaiser R, Lengauer T, Harrigan PR. Predicting HIV co-receptor usage based on genetic and clinical covariates. Antiviral Therapy, 2007.

Snoeck J, Van Laethem K, Hermans P, Van Wijngaerden E, Derdelinckx I, Schrooten Y, van de Vijver DA, De Wit S, Clumeck N, Vandamme AM. Rising prevalence of HIV-1 non-B subtypes in Belgium: 1983-2001. J Acquir Immune Defic Syndr. 2004 Mar 1;35(3):279-8

Soares EA, Martínez AM, Souza TM, Santos AF, Da Hora V, Silveira J, Bastos FI, Tanuri A, Soares MA. HIV-1 subtype C dissemination in southern Brazil. AIDS 2005. 19(Suppl. 4): S81–86.

Soares EA, Santos AF, Sousa TM, Sprinz E, Martinez AM, Silveira J, Tanuri A, Soares MA. Differential drug resistance acquisition in HIV-1 of subtypes B and C. PLoS One 2007. 15;2(1):e730.

- Soares EA, Santos RP, Pellegrini JA, Sprinz E, Tanuri A, Soares MA: Epidemiologic and molecular characterization of human immunodeficiency virus type 1 in southern Brazil. *J Acquir Immune Defic Syndr* 2003. 34:520–526.
- Soares MA, De Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brígido L, Pires IL, Morgado MG, Dantas MC, Barreira D, Teixeira PR, Cassol S, Tanuri A. A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. *AIDS* 2003. 17:11–21.
- Souquière S, Bibollet-Ruche F, Robertson DL, Makuwa M, Apetrei C, Onanga R, Kornfeld C, Plantier JC, Gao F, Abernethy K, White LJ, Karesh W, Telfer P, Wickings EJ, Maucière P, Marx PA, Barré-Sinoussi F, Hahn BH, Müller-Trutwin MC, Simon F. Wild *Mandrillus sphinx* are carriers of two types of lentivirus. *J Virol* 2001. 75(15): p. 7086-96.
- Srisuphanunt M, Sukeepaisarnchareon W, Kucherer C, Pauli G. The epidemiology of HIV-1 subtypes in infected patient's from northeastern Thailand. *Southeast Asian J Trop Med Pub Health* 2004. 35(3):641-8.
- Steain MC, Wang B, Dwyer DE, Saksena NK. HIV-1 co-infection, superinfection and recombination. *Sexual Health*. 2004. 1: 239-250.
- Stevenson M, Stanwick TL, Dempsey MP, Lamonica CA. HIV-1 replication is controlled at the level of T cell activation and proviral integration. *EMBO J* 1990. 9(5):1551-60.

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM. Phylogenetic Inference em Molecular Systematics, Second Edition Edited by Hillis DM, Moritz C and Mable BK. 1996, 655 páginas.

Swofford DL. PAUP 4.0. Phylogenetic analysis using parsimony (and other methods), 8 4.0b2a. Sunderland, MA, USA: Sinauer Associates, Inc. 1999.

Szwarcwald CL, Bastos FI, Esteves MA, de Andrade CL. The spread of the AIDS epidemic in Brazil from 1987 to 1996: a spatial analysis. Cad Saude Publica 2000. 16: 7-19.

Talbott RL, Sparger EE, Lovelace KM, Fitch WM, Pedersen NC, Luciw PA, Elder JH. Nucleotide sequence and genomic organization of feline immunodeficiency virus. Proc Natl Acad Sci U S A 1989. 86(15):5743-7.

Tamura K & Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 1993. 10:512-526.

Tanuri A, Swanson P, Devare S, Berro OJ, Savedra A, Costa LJ, Telles JG, Brindeiro R, Schable C, Pieniazek D, Rayfield M. HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil. J Acquir Immune Defic Syndr Hum Retrovirol. 1999 1;20(1):60-6.

Teixeira PR, Vitória MA, Barcarolo J. Antiretroviral treatment in resource poor settings: the Brazilian Experience. AIDS 2004. 18 Suppl 3:S5-7.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL X windows interface: flexible strategies for multiple sequence

alignment aided by quality analysis tools. *Nucleic Acids Res* 1997. 15;25(24): 4876-82.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994. 11;22(22):4673-80.

Thuler LC, Hatherly AL, Góes PN, Almeida e Silva JR. Mortality descriptors in HIV in patients. *Rev Saude Publica* 1998. 32(6):572-8.

Ugolini S, Mondor I, Sattentau QJ. HIV-1 attachment: another look. *Trends in microbiology* 1999. 7(4):144-9.

Un aids, *AIDS epidemic update*: Dezembro de 2006.

Un aids. *AIDS epidemic update*: Novembro de 2009.

Vachot L, Ataman-Onal Y, Terrat C, Durand PY, Ponceau B, Biron F, Verrier B. Short communication: retrospective study to time the introduction of HIV type 1 non-B subtypes in Lyon, France, using env genes obtained from primary infection samples. *AIDS Res Hum Retroviruses* 2004. 20(7):687-91.

Vaughan HE, Cane P, Tedder RS. Characterization of HIV-1 clades in the Caribbean using *pol* gene sequences. *AIDS Res Hum Retroviruses* 2003. 19(10):929-32.

- Véras NM, Véras VS, Ramalho ED, Kyaw C, Silva RR, Brígido MM, Martins CR. HIV type 1 genetic variability in central Brazil AIDS Res Hum Retroviruses 2007. 23(12):1481-90.
- Vicente AC, Otsuki K, Silva NB, Castilho MC, Barros FS, Pieniazek D, Hu D, Rayfield MA, Bretas G, Tanuri A. The HIV epidemic in the Amazon Basin is driven by prototypic and recombinant HIV-1 subtypes B and F. J Acquir Immune Defic Syndr 2000. 23(4):327-31.
- Wainberg MA. HIV-1 subtype distribution and the problem of drug resistance. AIDS 2004. 18 Suppl 3:S63-8.
- Woodman Z & Williamson C. HIV molecular epidemiology: transmission and adaptation to human populations. Curr Opin HIV AIDS 2009. 4(4):247-52. Review.
- Yamaguchi J, Bodelle P, Vallari AS, Coffey R, McArthur CP, Schochetman G, Devare SG, Brennan CA. HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and O infections. AIDS Res Hum Retroviruses 2004. 20(9):944-57.
- Yamaguchi J, Vallari AS, Swanson P, Bodelle P, Kaptué L, Ngansop C, Zekeng L, Gürtler LG, Devare SG, Brennan CA. Evaluation of HIV type 1 group O isolates: identification of five phylogenetic clusters. AIDS Res Hum Retroviruses 2002.18(4):269-82.
- Yang ZY, Chakrabarti BK, Xu L, Welcher B, Kong WP, Leung K, Panet A, Mascola JR, Nabel GJ. Selective modification of variable loops alters

tropism and enhances immunogenicity of human immunodeficiency virus type 1 envelope. *J Virol* 2004 78(8): 4029-36.

Yarchoan R, Klecker RW, Weinhold KJ, Markham PD, Lyerly HK, Durack DT, Gelmann E, Lehrman SN, Blum RM, Barry DW, et al. Administration of 3'-azido-3'-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. *Lancet* 1986. 1(8481): 575-80.

Yerly S, Jost S, Monnat M, Telenti A, Cavassini M, Chave JP, Kaiser L, Burgisser P, Perrin L; Swiss HIV Cohort Study. HIV-1 co/super-infection in intravenous drug users. *AIDS* 2004. 18(10):1413-21.

Yu X, Yuan X, Matsuda Z, Lee TH, Essex M. The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J Virol* 1992. 66(8):4966-71.

***MANUSCRITOS***

**Manuscrito nº 1:**

**Araujo A F**, Brites C, Monteiro JP, Santos LA, Galvão-Castro B, Alcantara LC.

LOWER PREVALENCE OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1  
BRAZILIAN SUBTYPE B FOUND IN NORTHEASTERN BRAZIL WITH  
SLOWER PROGRESSION TO AIDS.

Submetido ao *AIDS Research and Human Retroviruses*.

**Manuscrito nº 2:**

Santos E de S, **Araujo AF**, Galvão-Castro B, Alcantara LC.

GENETIC DIVERSITY OF HUMAN IMMUNODEFICIENCY VIRUS TYPE-1  
(HIV-1) IN INFECTED WOMEN FROM A NORTHEAST CITY OF BRAZIL.

*Rev Bras Ginecol Obstet*. 2009 Dec;31(12):609-14.

**Manuscrito nº 3:**

Monteiro JP \*, **Araújo AF** \*, Santos E de S, Galvão-Castro B, Alcantara LC.

LACK OF HIGH-LEVEL RESISTANCE MUTATIONS IN HIV-1 BF  
RECOMBINANT STRAINS CIRCULATING IN NORTHEAST BRAZIL.

Submetido ao *AIDS Research and Human Retroviruses*.

**Manuscrito 01**

LOWER PREVALENCE OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1  
BRAZILIAN SUBTYPE B FOUND IN NORTHEASTERN BRAZIL WITH  
SLOWER PROGRESSION TO AIDS

RUNNING TITLE: **HIV-1 DIVERSITY IN BAHIA, BRAZIL**

AUTHORS: ADRIANO F. ARAUJO<sup>1\*</sup>, CARLOS BRITES<sup>2\*</sup>, JOANA P.  
MONTEIRO<sup>1</sup>, LUCIANE A. SANTOS<sup>1</sup>, BERNARDO GALVÃO-CASTRO<sup>1,3</sup>, LUIZ  
CARLOS JUNIOR ALCANTARA<sup>1,3,4</sup>.

AUTHORS' AFFILIATION:

*<sup>1</sup>Advanced Public Health Laboratory, Gonçalo Moniz Research Center,  
Oswaldo Cruz Foundation, Salvador, BA, Brazil; <sup>2</sup>Federal University of Bahia;  
Salvador, BA, Brazil; <sup>3</sup>Bahia School of Medicine and Public Health/Foundation  
for Development of Science, Salvador, BA, Brazil; <sup>4</sup>Animal Models and  
Retroviral Vaccines Section, NCI/NIH, Bethesda, MD, USA \* These authors  
contributed equally to this work*

CORRESPONDING AUTHOR: LUIZ CJ ALCANTARA

Address: National Cancer Institute/NIH, Building 41, Room C-303, Bethesda,  
Maryland, USA 20892-5065. Tel. +1(301)4026158 (USA) / Fax +1(301)402-  
0055. Email: [alcantaralc@mail.nih.gov](mailto:alcantaralc@mail.nih.gov)

**Abstract**

Besides being extremely useful in measuring the level of HIV-1 diversity and prevalence in populations, the molecular analysis of genomic sequences provides crucial surveillance support and aids in the development of new therapies and effective vaccines. The present study focused on *gag* and *env* DNA and amino acid sequences that were generated from samples taken from 61 infected patients in the City of Salvador, Bahia, located in northeastern Brazil. In order to determine selective pressure and predict coreceptor usage, Bioinformatics tools were employed in phylogeny reconstruction. Fifty-six (91.8%) viruses were classified as belonging to subtype B, three (4.9%) from F1 and two (3.3%) from BF1 recombinants. Based on the characterization of the V3 region, the subtype B strains were represented by eight (18.2%) Brazilian variants (B'-GWGR), 20 (46.5%) European/EUA B variants (GPGR) and 15 (34.9%) GXGX variants. The mean time elapsed since diagnosis was 13 years among subtype B' and 9 years in subtype B. The mean  $d_N/d_S$  ratios from the GWGR, GPGR and GXGX groups, when compared to an HXB2 reference, were 0.72, 0.77 and 0.67 respectively. Seventy-six percent of the viruses studied were predicted to use the CCR5 coreceptor for cell entry (R5 viruses) while 24% were predicted to use the CXCR4 or were classified as dual tropic viruses. The prevalence of subtypes B' and recombinant B/F1 was shown to be lower than findings from previous studies performed both in Brazil (B') and in Bahia (B/F1). The association between subtype B' and a lengthy period of time since diagnosis can be correlated with a slower disease progression in infected patients, when compared with those infected with subtype B.

## **Introduction**

The remarkable genomic variability of HIV is one of the main obstacles preventing immunologic infection control and impeding antiretroviral therapy efficacy as well as vaccine development. This great variation has enabled virus classification and the study of viral genetic prevalence through phylogenetic analysis. HIV-1 group M variants, mainly responsible for the current HIV/AIDS pandemic, are divided into nine subtypes, circulating and unique recombinant forms (CRF and URF). These strains present distinctive geographic distribution worldwide, as well as in Brazil, with the highest number of HIV infections and greatest level of viral diversity found in Latin America<sup>1</sup>. The subtype B is the predominant genotype found in Brazil, followed by F1, C and recombinants B/F1<sup>2</sup>. Since 1993, a viral strain containing a GWGR signature motif at the tip of the V3 loop, named Brazilian subtype B (B'), has been detected in about half the samples analyzed<sup>3,4</sup> and a slower AIDS progression has been associated with subtype B' when compared to subtype B<sup>5</sup>. To gain entry into human cells, an HIV-1 particle interacts with a CD4 cellular receptor and a coreceptor, most commonly CCR5 or CXCR4. CCR5 usage (R5 virus) is associated with acute or asymptomatic patients, while CXCR4 (X4 virus) is associated with an elevated viral load, low CD4 T-cell counts, an increased disease progression rate and a pronounced resistance to CCR5 inhibitors. HIV-1 tropism is frequently determined during clinical practice and antiretroviral therapy management. The aim of this study was to characterize the molecular diversity of HIV-1 in Salvador, the state capital of Bahia which is located in the northeastern region of Brazil, and to compare genomic variability with the epidemiological and clinical characteristics of HIV-1-infected patients.

## Materials and Methods

Sixty-one HIV-infected individuals, patients from the Professor Edgar Santos University Hospital (HUPES), located in the City of Salvador, Brazil, were recruited in 2006. All patients signed a letter of informed consent. Two to five ml of whole blood samples were collected and transported to the Advanced Laboratory of Public Health (LASP) at CPqGM/FIOCRUZ where they were stored at -20°C until use. Clinical data were also obtained from patient medical records. This study was approved by the CPqGM/FIOCRUZ Ethics Committee. DNA samples were extracted from 200µl of blood using a QIAamp DNA kit (Qiagen Inc., Valencia, California, USA) in accordance with manufacturer's directions. Fragments of *gag* (positions 898 to 1968) and *env* (positions 6945 to 8183) genes, both relative to the HXB2 reference sequence, were PCR-amplified using nested primers. The amplified DNA was purified using a purification kit (Qiagen Inc.) and sequenced using a BigDye Terminator v.3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, US) and an automated ABI 3100 Genetic Analyzer (Applied Biosystems, US). Primers GAG2, G17, H1G777 and MZ14 were used for *gag* sequencing reactions, while primers ED31, MM4, ES7 and ED14 were used for the *env* region. Electropherograms were obtained, analyzed and exported in FASTA format using SeqScape v2.1.1 software (Applied Biosystems, US). Sequences were subjected to the BLAST search algorithm (<http://blast.ncbi.nlm>) to detect any evidence of contamination. Using a subtype reference set from the Los Alamos database (<http://hiv-web.lanl.gov>), an alignment was created using MUSCLE software<sup>6</sup> and manually edited using GENEDOC software<sup>7</sup>.

Phylogenetic analyses were performed using PAUP\* 4.0b10 software<sup>8</sup> to generate neighbor-joining (NJ) and maximum likelihood (ML) trees using the GTR model of nucleotide substitution<sup>9</sup>. Node reliability was assessed using bootstrap analysis (1,000 replicates) and a likelihood-ratio test was used to calculate statistical support for tree branches. The trees were drawn using Figtree software (<http://tree.bio.ed.ac.uk>). Sequences that did not cluster within any pure subtype groups were submitted to recombination analysis using the Simplot software bootscanning tool<sup>10</sup> and Genedoc software to determine the specific crossover point using visual inspection.

Coreceptor usage was predicted *in silico* based on V3-sequences and clinical data using the geno2pheno [coreceptor] online tool (<http://coreceptor.bioinf.mpi-inf.mpg.de>) which is based on a statistical learning method that predicts CXCR4 usage by HIV-1<sup>11</sup>.

Selective pressure was assessed in the 1155-bp envelope fragment using the SNAP<sup>12</sup> web-based tool (<http://www.hiv.lanl.gov>) which calculates nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitution rates. Using these rates, the  $d_N/d_S$  ratio was calculated. Considering that a high  $d_N/d_S$  ratio suggests a trend towards positive selection<sup>13</sup>, this ratio indicates the level of selective pressure through which the retrovirus isolates pass. Clinical, epidemiological and molecular data were analyzed using PASW software, version 18.0 for Windows (SPSS Inc, US). A statistically significant difference was assumed when  $p < 0.05$  and 95% CI did not cross zero.

## Results

Out of 61 HIV-1 positive samples analyzed, 41 were obtained from males and 20 were from females (2:1 ratio). Among males, a mean age of 43.3 years was observed versus 39.1 years in females. Mean infection time was 9.94 years, mean viral load was 75,587.87 copies/mL and mean TCD4 cell counts were 421.61 cells/mL while TCD8 cell counts were 1,262.24 cells/mL. Regarding ethnicity, 23% of the patients reported being of European or Latin descent while 47.5% indicated Mixed-race and 29.5% African descent. No statistically significant differences regarding time of infection, viral load or TCD4 and TCD8 cell counts were observed between ethnic groups. Out of 61 samples, 42 *gag* and 46 *env* gene sequences were amplified with twenty-seven sequences amplified in both genes. Phylogenetic analysis of the *gag* gene showed 40 (95.2%) sequences clustered within the subtype B reference group, while two (4.8%) did not cluster within any pure subtypes. Recombination analysis of these two sequences revealed recombination events between subtypes B and F1. The BAS026 sequence showed two breakpoints (B/F1/B): one in positions 1357-1398 and another at positions 1482-1615 (relative to HXB2). The BAS096 sequence presents a breakpoint in positions 1187-1228 (relative to HXB2). In the *env* sequences, 43 (93.5%) were shown to cluster inside subtype B and 3 (6.5%) inside F1 (figure 1). The two *gag* recombinants were subtyped as B in the *env* tree. With respect to the 27 samples that were characterized in both genes, 25 (92.6 %) belonged to subtype B while only two (7.4 %) were found to be BF1/B. Considering all 61 samples, 56 (91.8%) were shown to be subtype B, three (4.9%) were F1 and two (3.3%) were BF1 recombinants. The 43 subtype B *env* sequences were translated into amino acid sequences. Based on V3

characterization, eight (18.2%) Brazilian (B'-GWGR), 20 (46.5%) European/EUA B (GPGR) and 15 (34.9%) GXGX variants were found. The mean pairwise distance was 0.11, 0.13 and 0.12 within the GWGR, GPGR and GXGX groups, respectively. To test whether selective pressure was altered due to V3 loop substitution, the  $d_N/d_S$  ratio of each sequence was determined under comparison with an HXB2 reference sequence and the sequences were subsequently grouped in categories: GWGR, GPGR and GXGX. The mean  $d_N/d_S$  ratio of all sequences was 0.725 when compared to HXB2. The mean GWGR, GPGR and GXGX  $d_N/d_S$  group ratios were 0.72, 0.77 and 0.67, respectively, when compared to HXB2. Statistical testing revealed no statistically significant differences between the GWGR and GPGR groups ( $p>0.05$ ). A comparison between GPGR and GXGX group sequences showed a significant difference ( $p=0.018$ ), revealing a higher rate of positive selection in the GPGR group. We compared the clinical characteristics of individuals harboring B' (GWGR) and B (GPGR + GXGX) viruses. Table 1 shows that the mean time period since diagnosis was higher in the subtype B' group than in the subtype B group and that the mean age was also higher in the former group. With respect to coreceptor usage, eleven (24%) out of the 46 V3 sequences were predicted to use the CXCR4 coreceptor (X4 virus). The mean age among subjects infected with X4 virus (32.4 years) was lower than the mean age among subjects infected with R5 virus (43.3 years) ( $p<0.05$ ).

## **Discussion**

Molecular studies of HIV-1 are critical to furthering the understanding of mechanisms involved in AIDS pathogeny and to support the development of vaccines and efficacious therapies. The number of HIV sequences from Northeastern Brazil remains scarce, as well as precise information associating HIV-1 diversity with clinical data. To this end, the authors studied samples and laboratory information from randomized patients in Salvador, the capital of the Northeastern Brazilian state of Bahia and the third most populous city in the country.

In the northeastern region of Brazil, as well as the country as a whole, HIV-1 subtype B remains predominant. However, in the past decade, several studies have reported on the increasing prevalence of other genotypes, notably BF recombinants<sup>14</sup>. In the City of Salvador, a lower prevalence of B/F1 recombinant forms (3.3%) was observed in our samples when compared with recent studies showing a 10% and 13% prevalence of BF forms in the Brazilian Northeast<sup>14,15</sup>. This may be explained by the fact that our sequencing involved two fragments within the *gag* and *env* genes, implying that other genomic regions where recombination could have occurred may have been overlooked.

This study found the prevalence of subtype B' to be much lower than 50%, which was observed in the country's southeastern region<sup>16</sup>. This leads us to suggest that Salvador has experienced different introduction(s) and founder effect(s) from the Brazilian Southeast. In addition, no phylogenetic distinctions were observed between subtypes B' and B, findings consistent with previous reports<sup>16</sup>. However, this study demonstrated that mean age (49 years) and time

period elapsed since diagnosis (13 years) in subtype B'-infected patients was higher than mean age (39 years) and time period elapsed since diagnosis (9 years) for subtype B-infected individuals. This could be related to an increased replication rate and accelerated disease progression characteristic of subtype B sequences<sup>5</sup>. In fact, the GPGR sequences exhibited a slightly higher mean pairwise distance which may explain the clinical difference between subtype B' and subtype B.

The results presented in this study demonstrate a lower level of genetic diversity and, specifically, a lower prevalence of BF recombinant forms than those previously found in Bahia and in the Brazilian Northeast region<sup>14,15</sup>. This may be related to the presence of recombination points in genomic regions other than those analyzed by this study, and, therefore, further studies involving the full genome sequencing of HIV isolates from this geographic region could contribute to a better understanding of this region's HIV epidemiology.

### **Acknowledgments**

The authors are grateful to the individuals who donated blood for the purposes of this study, to the FIOCRUZ-PDTIS sequence platforms, Mr. Augusto Santana and Mrs. Maurina Alcantara from HUPES for providing access to patient biomedical records, to Mrs. Elisabeth Deliege Vasconcelos for editing and revising this manuscript.

### **Sequence Data**

The new sequences in this study were reported to GenBank under the accession numbers GU595197-281 and GU722093-95.

## References

01. Bastos FI, Cáceres C, Galvão J, Veras MA, Castilho EA. AIDS in Latin America: assessing the current status of the epidemic and the ongoing response. *Int J Epidemiol* 2008; 37(4): 729-37.
02. Morgado MG, Guimarães ML, Galvão-Castro B. HIV-1 polymorphism: a challenge for vaccine development - a review. *Mem Inst Oswaldo Cruz* 2002; 97(2): 143-50.
03. Potts KE, Kalish ML, Lott T, Orloff G, Luo CC, Bernard MA, Alves CB, Badaro R, Suleiman J, Ferreira O, et al. Genetic heterogeneity of the V3 region of the HIV-1 envelope glycoprotein in Brazil. Brazilian Collaborative AIDS Research Group. *AIDS* 1993; 7(9): 1191-7.
04. Morgado MG, Sabino EC, Sphaer EG, et al. V3 region polymorphisms in HIV-1 from Brazil: prevalence of subtype B strains divergent from North American/European prototypes and detection of subtype F. *AIDS Res Hum Retroviruses* 1994; 10: 569-76.
05. De Brito A, Komninakis SC, Novoa P, de Oliveira RM, Fonseca LA, Duarte AJ, Casseb J. Women infected with HIV type 1 Brazilian variant, subtype B (B'-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). *Clin Infect Dis* 2006; 43(11): 1476-1481.
06. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004; 5(1): 113-32.

07. Nicholas KB, Nicholas Jr HB, Deerfield II DW. GeneDoc: Analysis and Visualization of Genetic Variation. *Embnet News* 1997; 4:1-4.
08. Swofford D. PAUP 4.0: Phylogenetic analysis using parsimony (and other methods), 8 4.0b2a. Sunderland, MA, USA: Sinauer Associates, Inc. 1999.
09. Posada D & Krandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998; 14: 817-8.
10. Salminen MO, Carr JK, Burke DS, McCutchan FE. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retrovir* 1995; 11: 1423-1425.
11. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues F, Büch J, Däumer M, Kaiser R, Lengauer T, Harrigan PR: Predicting HIV co-receptor usage based on genetic and clinical covariates. *Antiviral Therapy* 2007; 12(7): 1097-106.
12. Korber B. (2000). HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences, Chapter 4, pages 55-72. Allen G. Rodrigo and Gerald H. Learn, eds. Dordrecht, Netherlands: Kluwer Academic Publishers.
13. Mindell DP. Positive selection and rates of evolution in immunodeficiency viruses from humans and chimpanzees. *Proc Natl Acad Sci U S A*. 1996 Apr 16; 93(8):3284-8.
14. Monteiro JP, Alcantara LC, de Oliveira T, Oliveira AM, Melo MA, Brites C, Galvão-Castro B. Genetic variability of human immunodeficiency virus-1 in

Bahia state, Northeast, Brazil: high diversity of HIV genotypes. *J Med Virol* 2009; 81(3): 391-9.

15. Brennan CA, Brites C, Bodelle P, Golden A, Hackett J Jr, et al. HIV-1 strains identified in Brazilian blood donors: significant prevalence of B/F1 recombinants. *AIDS Res Hum Retroviruses*. 2007 Nov;23(11): 1434-41.

16. Diaz RS, Leal E, Sanabani S, Sucupira MC, Tanuri A, Sabino EC, Janini LM. Selective regimes and evolutionary rates of HIV-1 subtype B V3 variants in the Brazilian epidemic. *Virology* 2008 Nov 25; 381(2): 184-93.

**Figure Legend**

**Fig. 1.** NJ tree based on *gag* (A) and *env* (B) sequences showing phylogenetic relationships between HIV-1 samples and group M reference sequences. A sequence from group O was used as an outgroup sequence. The GTR+I+G nucleotide substitution model was used in both trees. Branches supported by the ML method are indicated by an asterisk (\*) when significant ( $p < 0.05$ ), and marked with a double asterisk (\*\*) when highly significant ( $p < 0.001$ ). Bootstrap values for 1,000 replicates are shown in percentages. The HIV-1 sequences which were generated in the present study are indicated in **bold**. Subtype B' sequences are marked by a plus sign (+). Two sequences remained outside the subtype B and subtype D (indicated by a bullet point (●)).

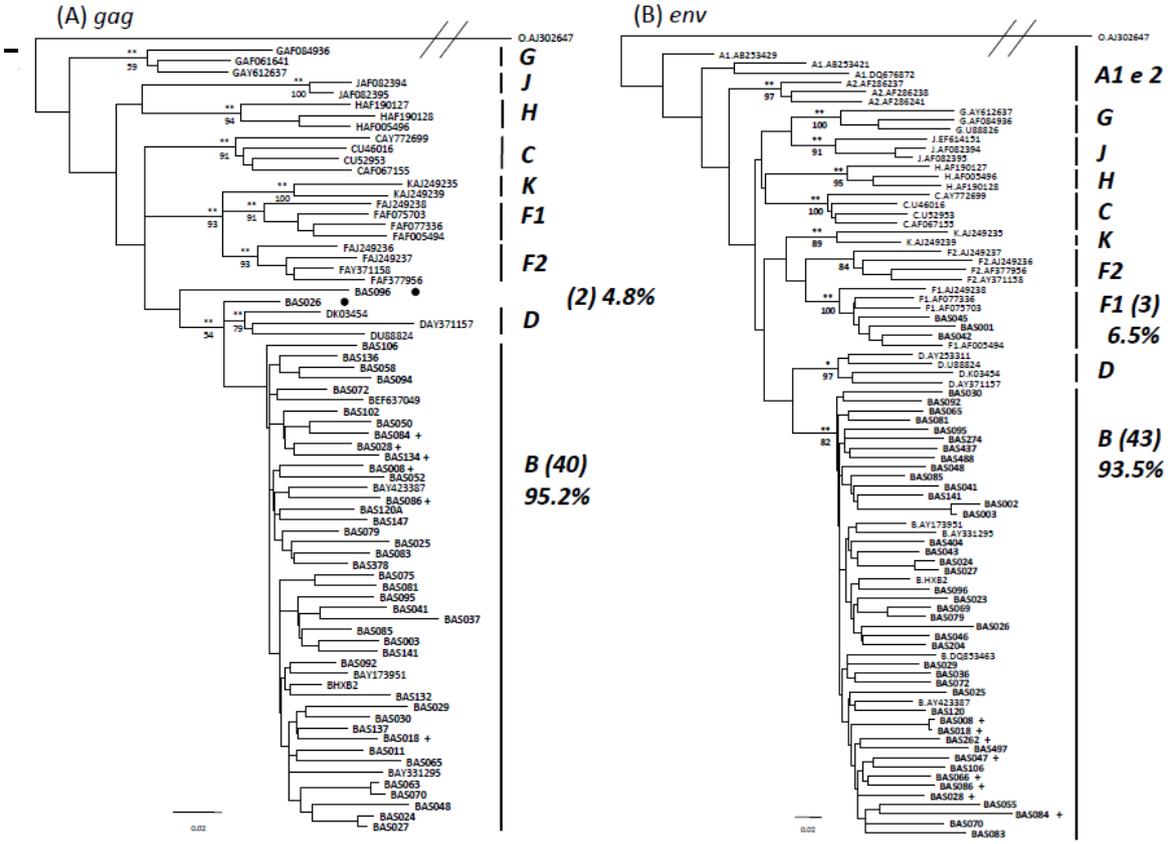


Table 1. Comparison between subtype B and B' and clinical characteristics of HIV-1-positive patients from Salvador. Brazil.

	B x B'	N	Mean	Std. Deviation	p*	95% (C.I.)	
						Lower	Upper
<b>Age</b>	B	34	39.09	12.983	0.057	-19.876	0.303
	B'	8	48.88	11.294			
<b>Viral load</b>	B	34	32202.35	81,611.235	0.857	-55584.779	66556.985
	B'	8	26716.25	48,895.262			
<b>CD4</b>	B	34	420.97	248.905	0.690	-232.868	155.559
	B'	8	459.63	222.836			
<b>CD8</b>	B	28	1439.32	1,391.531	0.538	-759.012	1428.798
	B'	7	1104.43	436.400			
<b>Time since diagnose</b>	B	29	8.83	3.536	<b>0.019</b>	<b>-7.087</b>	<b>-0.686</b>
	B'	7	12.71	4.572			

\*, Test t for equality of means.

**Manuscrito 02**

Edson de Souza Santos<sup>1</sup>Adriano Fernando Araújo<sup>2</sup>Bernardo Galvão-Castro<sup>3</sup>Luiz Carlos Junior Alcantara<sup>4</sup>

# Diversidade genética do vírus da imunodeficiência humana tipo 1 (HIV-1) em mulheres infectadas de uma cidade do nordeste do Brasil

*Genetic diversity of human immunodeficiency virus type-1 (HIV-1) in infected women from a northeast city of Brazil*

Artigo original

## Palavras-chave

HIV-1  
Variação genética  
Genótipo  
Resistência a medicamentos  
Evolução clínica

## Keywords

HIV-1  
Genetic variation  
Genotype  
Drug resistance  
Clinical evolution

## Resumo

**OBJETIVO:** descrever a diversidade genética dos isolados de HIV-1 de mulheres soropositivas acompanhadas em um centro de referência. **MÉTODOS:** estudo transversal, no qual foram incluídas 96 mulheres com dois testes sorológicos ELISA e um teste confirmatório Western Blot. Das amostras de sangue periférico, foram determinadas a carga viral pelo kit b-DNA e a contagem de linfócitos T CD4 e T CD8 pela citometria de fluxo excalibur. A extração e purificação do DNA pró-viral foi realizada pela reação em cadeia da polimerase (PCR), utilizando o kit QIAamp Blood (Qiagen Inc., Chatsworth, CA, USA). O sequenciamento da região pol foi realizado em 52 isolados com o (3100 Genetic Analyzer, Applied Biosystems Inc., Foster City, CA) e a genotipagem foi investigada pela ferramenta Rega (Rega Subtyping Tool). O padrão de resistência aos antirretrovirais (ARV) foi inferido pelo algoritmo do banco de dados Stanford HIV Resistance. Os estágios clínicos das participantes foram definidos como A, B ou C segundo os critérios do Center for Diseases Control (CDC). Para a análise estatística dos dados, foram utilizados os testes do  $\chi^2$  para as variáveis categóricas e o teste *t* de Student para as variáveis numéricas. **RESULTADOS:** a média de idade da amostra, o tempo médio de doença e de tratamento foram: 33,7; 3,8 e 2,5 anos, respectivamente. A média da carga viral foi  $\log_{10}$  2,3 cópias/mL; a dos linfócitos T CD4 e T CD8 foi 494,9 células/ $\mu$ L e 1126,4 células/ $\mu$ L. Sobre o estágio clínico, 30 mulheres estavam no estágio A, 47 no B e 19 no C. O sequenciamento dos 52 isolados encontrou 33 do subtipo B, quatro do F, um do C e 14 do recombinante BF. A análise da resistência aos ARV mostrou 39 (75,0%) isolados susceptíveis, 13 (25,0%) resistentes aos inibidores da transcriptase reversa (INTR) e três (5,7%) aos inibidores da protease (IP). **CONCLUSÕES:** Houve grande diversidade do HIV-1 e elevado percentual de isolados resistentes aos ARV na amostra estudada.

## Abstract

**PURPOSE:** to describe the genetic diversity of HIV-1 isolates from serum positive women followed up at a reference center. **METHODS:** transversal study, including 96 women with two ELISA serological tests and a Western Blot confirmatory test. The viral charge was determined by the b-DNA kit, and the counting of T CD4 and T CD8 lymphocytes, by the Excalibur flow cytometry, from the samples of peripheral blood. The extraction and purification of pro-viral DNA was performed by the polymerase (PCR) chain reaction, using the QIAamp Blood kit (Qiagen Inc., Chatsworth, CA, U.S.A.). Sequencing of the pol region was done in 52 isolates with the 3100 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA), and the genotyping was assessed by the Rega Subtyping Tool. The resistance pattern to anti-retrovirals (ARV) was inferred by the algorithm from the Stanford HIV Resistance data bank. Participants' clinical stages were defined as A, B or C, according to the criteria established by the Center for Diseases Control (CDC). For statistical analysis, the  $\chi^2$  test was used for the categorical variables and the Student's *t* test, for the numerical variables. **RESULTS:** The average age of the sample, the disease and treatment average duration were respectively: 33.7 years old, 3.8 and 2.5 years. The viral charge average was  $\log_{10}$  2.3 copies/mL; the T CD4 e T CD8 lymphocytes, 494.9 cells/ $\mu$ L and 1126.4 cells/ $\mu$ L. Concerning the clinical stage, 30 women were in stage A, 47 in B and 19 in C. Sequencing from the 52 isolates found 33 of B subtype, 4 of F, 1 of C and 14 of BF recombinant. The analysis of resistance to ARV has shown 39 (75.0%) susceptible isolates, 13 (25.0%) resistant to reversal transcriptase inhibitors (RTIN), and 3 (5.7%) resistant to protease inhibitor (PI). **CONCLUSIONS:** There has been a large variety of HIV-1 and a high percentage of isolates resistant to ARV in the studied sample.

## Correspondência:

Edson de Souza  
Santos Centro de Referência  
DST/HIV/Aids Rua Prof.  
Germínio Costa, s/n  
CEP 44025-070 – Feira de Santana (BA), Brasil  
Tel./Fax: (75) 3221-2088/3625-2125  
E-mail: edsongin@hotmail.com

## Recebido

29/6/09

## Aceito com modificações

19/11/09

Centro de Referência em DST/HIV/Aids da Secretaria de Saúde de Feira de Santana (BA), rasil.

<sup>1</sup> Médico do Centro de Referência em DST/HIV/Aids da Secretaria de Saúde de Feira de Santana; Mestrando da Escola Bahiana de Medicina e Saúde Pública – EBMSP – da Fundação Bahiana para o Desenvolvimento das Ciências – FBDC – Feira de Santana (BA), Brasil.

<sup>2</sup> Pós-graduando (Doutorado) em Biotecnologia do Laboratório Avançado de Saúde Pública do Centro de Pesquisas Gonçalo Moniz da Fundação Oswaldo Cruz – FIOCRUZ – Salvador (BA), Brasil.

<sup>3</sup> Coordenador do Curso de Pós-Graduação em Medicina e Saúde Humana da Escola Bahiana de Medicina e Saúde Pública – EBMSP – Feira de Santana (BA), Brasil; Coordenador de Ensino e Supervisor do Laboratório Avançado de Saúde Pública do Centro de Pesquisas Gonçalo Moniz da Fundação Oswaldo Cruz – FIOCRUZ – Salvador (BA), Brasil.

<sup>4</sup> Professor do Curso de Pós-Graduação em Medicina e Saúde Humana da Escola Bahiana de Medicina e Saúde Pública – EBMSP – Feira de Santana (BA), Brasil; Chefe do Laboratório Avançado de Saúde Pública do Centro de Pesquisas Gonçalo Moniz da Fundação Oswaldo Cruz – FIOCRUZ – Salvador (BA), Brasil. Não houve conflitos de interesse na realização deste trabalho.

## Introdução

A Síndrome da Imunodeficiência Adquirida (Aids) causada pela infecção pelo vírus da imunodeficiência humana tipo 1 (HIV-1) constitui um sério problema de Saúde Pública em todo o mundo. Convivem com o HIV-1 33,2 milhões de pessoas, das quais 15,4 milhões de mulheres e 2,5 milhões de crianças com menos de 13 anos. Entre as mulheres, observa-se um aumento de casos em todas as regiões do globo<sup>1</sup>.

No Brasil, existiam, em 2007, 159.793 mulheres infectadas, o que representa uma relação de 1,5 homens/mulher. Na faixa dos 13 aos 19 anos de idade, esta relação foi invertida desde 1998, porém a mais atingida foi de 25 a 39 anos. A via de transmissão viral mais importante foi a heterossexual e foi observada uma redução gradativa da transmissão por uso de drogas injetáveis<sup>2</sup>.

Há evidências de que as secreções vaginais de mulheres portadoras têm elevadas concentrações do vírus, o que favorece a contaminação durante o intercurso sexual e no momento do parto<sup>3</sup>. Outros fatores relacionados à mãe, como a carga viral plasmática elevada, a diversidade do vírus circulante e a resistência aos antirretrovirais (ARV) representam riscos adicionais responsáveis por 30 a 35% dos casos de transmissão vertical<sup>4</sup>. No entanto, há relatos com resultados controversos sobre o papel dos subtipos e recombinantes nesta forma de transmissão viral<sup>5-7</sup>.

O HIV-1 tem um genoma relativamente pequeno, que mede aproximadamente 9 kilobytes (Kb) de comprimento ocupados em grande parte pelos três genes estruturais: gag (grupo antígeno), pol (polimerase) e env (envelope), que são delimitados por duas regiões contendo sequências repetitivas, denominadas LTR (longas regiões terminais repetidas)<sup>8</sup>. Sua principal característica é apresentar um alto grau de diversidade de sequências genéticas em decorrência de vários fatores biológicos, como as altas taxas de erros de transcrição da transcriptase reversa e de replicação viral que, atingindo uma frequência de mais de 300 ciclos por ano, dão origem às inúmeras mutações<sup>9</sup>.

Esta intensa mutabilidade do vírus afeta não só a interação com a célula hospedeira, mas também os genes que codificam as enzimas: transcriptase reversa, protease e integrase, que são os sítios de atuação dos ARV. A pressão seletiva exercida sobre o vírus induz o aparecimento de resistência, principalmente aos inibidores nucleosídicos da transcriptase reversa (INTR) e aos inibidores da protease (IP)<sup>10</sup>.

A análise de sequências conhecidas do HIV-1 permitiu classificá-lo em três grupos distintos denominados M, O

e N. O grupo M (major) envolve a maioria das infecções em todo o mundo e é composto por nove subtipos filogeneticamente distintos, denominados A, B, C, D, F, G, H, J e K, cujas sequências de genes diferenciam entre si em cerca de 20%<sup>11</sup>. O subtipo C é o mais prevalente, sobretudo na Índia, enquanto o subtipo B é o mais difundido em todos os continentes. Foram descritas 43 formas recombinantes circulantes (CRF) que são responsáveis por 18% do total das infecções. O CRF BF surgiu na América do Sul, tendo como origem o subtipo F da África. No Brasil, predominam os Subtipos B, C, D e F e os CRF: BF, BD e BC<sup>12</sup>.

Na Bahia, estudos recentes revelam uma profunda modificação do perfil epidemiológico da difusão do HIV-1 no Estado devido à circulação de cepas virais recombinantes com alto potencial para desenvolverem resistência e progressão da doença<sup>13</sup>. Feira de Santana é a segunda cidade do estado, centro de grande fluxo migratório e entroncamento rodoviário o que pode favorecer o processo de interiorização da epidemia da Aids. Neste município, já haviam sido registradas, no Centro de Referência em DST/HIV/Aids, 616 pessoas soropositivas até o ano de 2007, das quais 180 eram mulheres e 32 eram crianças.

O objetivo deste estudo foi descrever a diversidade genética e a resistência aos ARV de isolados do HIV-1 e correlacionar os resultados com os dados epidemiológicos, estágios clínicos da doença e a transmissão vertical dos sujeitos envolvidos.

## Métodos

Foi realizado um estudo de corte transversal, de caráter observacional, descritivo e analítico em uma amostra de conveniência de 96 mulheres adultas na faixa reprodutiva, acompanhadas no Centro de Referência em DST/HIV/Aids, da Secretaria Municipal de Saúde de Feira de Santana (BA), Brasil, no período de março de 2008 a junho de 2009.

Todas pacientes assinaram o termo de consentimento livre e esclarecido e o projeto foi aprovado pela comissão de ética do Centro de Referência em DST/HIV/Aids e pelo Comitê de Ética em Pesquisa (CEP) da Escola Bahiana de Medicina e Saúde Pública (EBMSP).

Os critérios para inclusão no estudo foram: mulheres que aceitaram participar, com idades entre 14 e 49 anos, notificadas como portadoras do HIV-1 há pelo menos um ano. As pacientes foram submetidas a dois testes sorológicos de triagem pelo método ELISA (HIV-1 rp21, enhanced, EIA, Cambridge Biotech Corporation) e a um

teste confirmatório Western Blot (HIV Blot 2.4, Genelabs Diagnostics, Singapore).

As informações sobre estágios clínicos, evolução da doença, tratamento com ARV e comorbidades foram obtidas dos registros de prontuários, objetivando-se correlacioná-las aos subtipos virais e ao padrão de resistência dos isolados. A classificação dos estágios clínicos obedeceu aos critérios do CDC e do Ministério da Saúde do Brasil. Foram enquadradas no estágio A, as mulheres com dois testes de triagem e um teste confirmatório positivos, porém assintomáticas. No estágio B, foram alocadas aquelas que apresentavam uma doença indicativa de Aids ou imunodeficiência moderada, enquanto aquelas com linfócitos T CD4 inferiores a 350 células/ $\mu$ L foram classificadas como estágio C.

Foram coletados 15 mL de sangue periférico de cada paciente por punção venosa a vácuo com seringa e anti-coagulante EDTA, para determinação da carga viral pelo método b-DNA e contagem dos linfócitos T CD4 e CD8 pelo método de citometria de fluxo escalibur. Os exames de biologia molecular foram realizados no Laboratório Avançado de Saúde Pública (LASP) da FIOCRUZ (BA), obedecendo às seguintes etapas: reação em cadeia da polimerase (PCR) para extração, amplificação e purificação do DNA de células mononucleares utilizando 200  $\mu$ L de concentrado provenientes de 2 mL de sangue periférico utilizando o kit QIAamp Blood (Qiagen Inc., Chatsworth, CA, USA). Em seguida, os produtos amplificados da região pol parcial, correspondentes a 817 nucleotídeos entre as posições 2.343 e 3.160 na sequência referência HXB2 de 52 novos isolados do HIV-1 em estudo, foram sequenciados automaticamente pelo (3100 Genetic Analyzer, Applied Biosystems Inc., Foster City, CA). O restante das coletas apresentava pouco material do DNA devido à baixa carga viral, e por isso não foi possível realizar o sequenciamento dessas amostras. Os subtipos do HIV-1 foram investigados através da ferramenta de subtipagem Rega (Rega Subtyping Tool); c) o nível de suscetibilidade das sequências aos antirretrovirais foi inferido através do algoritmo do banco de dados Stanford HIV Resistance. O arquivo contendo as sequências em formato fasta foi submetido ao sítio eletrônico <http://hivdb.stanford.edu/>.

A abordagem estatística do estudo contemplou a descrição das variáveis categóricas como: estado civil (solteira e estável, incluindo as casadas); escolaridade (fundamental médio e superior); classe social inferida pela profissão e renda salarial referida (baixa, média-baixa e alta); etnia definida pela cor da pele autorreferida (branca e não-branca); vias de transmissão (sexual e -sexual). As comorbidades, os subtipos virais e as mutações encontradas também foram descritas através de proporções e frequências e analisadas pelo teste do  $\chi^2$ . Foi considerado significativo  $p < 0,05$ .

As variáveis numéricas, como idade, carga viral, razão T CD4/T CD8, tempo de doença e tempo de tratamento, foram expressas em médias, desvio padrão e variação.

A comparação das médias entre os grupos de mulheres portadoras do subtipo (MSB) e as portadoras de outros subtipos ou recombinantes (MSNB) foi feita pelo Teste *t* de Student para as variáveis de distribuição normal ou pelo teste não-paramétrico de Mann-Whitney para as variáveis de distribuição anormal. Foi considerado significativo  $p < 0,05$ . Os dados foram alocados em uma planilha do programa Microsoft Office Excel 2007 (Windows XP) e analisados por intermédio do programa Statistical Package for the Social Sciences (SPSS) versão 12.0 para Windows.

## Resultados

A idade da população em estudo variou de 14 a 49 anos; (média de 33,7 e desvio padrão de  $\pm 15$ ). As características gerais da população estão descritas na Tabela 1. A maioria das mulheres da amostra era solteira (51%) de classe baixa (97%) e com escolaridade primária (79,2%). Quanto à etnia, a maioria (54%) era não-branca. Sobre a via de transmissão, 100% das mulheres declararam ser heterossexuais, porém houve relatos de uso de drogas injetáveis por duas delas e transfusão sanguínea por uma. Entre as mulheres estudadas, 12 delas tinham filhos soropositivos, nascidos de partos naturais e que haviam sido amamentados. Nenhuma destas fez uso do Protocolo do Aids Clinical Trial Group (PACTG 076), que previne a transmissão vertical do HIV-1.

**Tabela 1** - Características gerais da amostra associadas aos subtipos de 52 isolados do HIV-1 de mulheres soropositivas de Feira de Santana (BA)

Variáveis	MSB	MSNB	Valor de p*
<b>Idade/anos (média<math>\pm</math>DP)</b>	33,5 $\pm$ 11,4	33,6 $\pm$ 12,8	
<b>Estado civil (%)</b>			
Solteira	18 (34,7)	11 (21,2)	0,17
Estável	15 (28,8)	8 (15,3)	0,15
<b>Etnia (%)</b>			
Não-branca	20 (38,5)	13 (25,0)	0,16
Branca	13 (25,0)	6 (11,5)	0,09
<b>Classe de renda (%)</b>			
Baixa	30 (57,6)	17 (32,7)	0,08
Média-baixa	3 (5,8)	2 (3,9)	0,24
<b>Escolaridade</b>			
Primária	29 (55,7)	12 (23,3)	0,07
Secundária	4 (7,6)	7 (13,4)	0,17
<b>Via de transmissão (%)</b>			
Sexual	32 (61,6)	18 (34,6)	0,47
Não-sexual	1 (1,9)	1 (1,9)	0,67

MSB: mulheres com subtipo B; MSNB: mulheres com subtipos não-B (C, Fe BF)

\*valor para o teste  $\chi^2$  de Pearson; DP: desvio padrão.

Os resultados relativos aos aspectos clínicos e laboratoriais da amostra em estudo são descritos na Tabela 2. A média da carga viral foi  $\log_{10}$  2,3 cópias/mL, desvio padrão  $\pm 1,9$  e variação de 0,1 a 5,6. A contagem dos linfócitos T CD4 variou de 50 a 1.784  $\text{cél}/\mu\text{L}$ , com média de 494  $\text{cél}/\mu\text{L}$ , desvio padrão de  $\pm 184,3$ . A média da proporção entre o TCD4/TCD8 foi 0,64. O tempo médio presumido de doença foi 3,7 anos, desvio padrão  $\pm 2,5$  e variação de 1 a 9. A média do tempo de terapia com antirretroviral (ARV) da amostra foi 2,5 anos, desvio padrão  $\pm 2,0$  e variação de 0 a 9 anos.

Sobre a evolução clínica das mulheres estudadas, 30 (31,3%) foram enquadradas como estágio A do CDC; outras 47 (48,9%) como estágio B, enquanto 19 (19,8%) foram consideradas estágio C. A gravidade do quadro esteve relacionado com maior tempo de doença, maior carga

viral e menor proporção TCD4/TCD8. As coinfeções mais encontradas na história progressa foram: infecções respiratórias (62 ocorrências), intestinais (47), tuberculose (9), hepatites (8) e HTLV (4 casos). As patologias genitais mais encontradas foram: vulvo-vaginites (25 casos), sífilis (14 casos), herpes (12 casos) e HPV (12 casos), sendo que quatro casos estiveram associados com neoplasia intraepitelial do colo uterino (NIC) (Tabela 2).

Nos estudos de variabilidade genética dos 52 isolados, foram encontrados 19 (36,5%) subtipos não-B em circulação. As frequências das espécies virais encontradas foram: 63,5% do subtipo B (33 sequências), 7,6% do subtipo F (4 sequências), 1,9% do subtipo C (1 sequência) e 26,9% da forma recombinante BF (14 sequências).

Comparando o grupo de mulheres portadoras do subtipo B (MSB) ao grupo de mulheres portadoras de subtipos não-B (MSNB), não foram observadas diferenças significativas concernentes à idade, etnia, escolaridade, classe de renda, estado civil ou vias de transmissão viral (Tabela 1). Não houve diferença entre os grupos na determinação da carga viral e na contagem dos linfócitos T CD4 e T CD8. No entanto, houve diferença significativa, com  $p < 0,05$  quando foram comparados o tempo de doença, o tempo de tratamento, a evolução para Aids (estágio C) e a prevalência de doenças genitais no grupo MSB (Tabela 2).

A análise de resistência destes 52 isolados mostrou que 39 (75,0%) deles portavam diversas mutações nas cadeias do gen pol que, no entanto, não alteravam a suscetibilidade aos ARV. Dos 13 isolados (25,0%) que apresentaram resistência aos ARV, três amostras foram exclusivamente ao lamivudina (3TC), duas à zidovudina (AZT), duas à nevirapina (NVP), uma à didanosina (ddI) e uma ao efavirenz (EFV). Três isolados tinham resistência múltipla a mais de um ARV em três isolados. Um isolado mostrou resistência primária à didanosina (ddI). Dos três isolados (5,7%) resistentes aos inibidores da protease (IP), dois foram ao nelfinavir (NFV) e um ao fosamprenavir (FPV). Todos eles também tinham resistência aos INTR.

A análise estatística que comparou as mulheres com filhos soropositivos àquelas sem filhos soropositivos não constatou diferenças significativas nas variáveis estudadas, com exceção da resistência aos ARV, o que é limitado pelo número examinado (Tabela 3).

**Tabela 2** - Comparação dos dados imunológicos, estágios clínicos e comorbidades entre os subgrupos de mulheres com o subtipo B e aquelas com subtipos não-B em Feira de Santana

Variáveis (médias)	MSB	MSNB	Valor de p	IC95%
Idade (anos $\pm$ DP)	33,5 $\pm$ 11,4	33,6 $\pm$ 12,8	0,96 <sup>†</sup>	-7,0 a 6,7
Carga viral ( $\log_{10}$ )	2,9 $\pm$ 1,8	2,1 $\pm$ 1,9	0,17 <sup>*</sup>	-0,35 a 1,8
T CD4 células/ $\mu\text{L}$	431,1 $\pm$ 194,3	484,2 $\pm$ 168,8	0,12 <sup>*</sup>	-32 a 162
Relação T CD4/T CD8	0,4 $\pm$ 0,3	0,5 $\pm$ 0,2	0,39 <sup>*</sup>	-0,26 a 1,0
Tempo de doença (anos)	4,8 $\pm$ 2,6	3,2 $\pm$ 2,1	0,02 <sup>*</sup>	0,21 a 2,9
Tempo de TARV	3,2 $\pm$ 2,7	2,1 $\pm$ 1,4	0,05 <sup>*</sup>	-0,04 a 2,2
Estágio clínico (%)				
A	6 (11,5)	5 (9,6)	0,47 <sup>**</sup>	
B	22 (42,3)	13 (25,0)	0,46 <sup>**</sup>	
C	5 (9,6)	1 (1,9)	0,01 <sup>**</sup>	
Comorbidades (%)				
Sífilis	5 (9,6)	3 (5,7)	0,53 <sup>**</sup>	
Herpes	4 (7,6)	2 (3,8)	0,09 <sup>**</sup>	
HPV/NIC	4 (7,6)	2 (3,8)	0,09 <sup>**</sup>	
HTLV	3 (5,7)	1 (1,9)	0,04 <sup>**</sup>	
Vaginites	8 (15,2)	5 (9,6)	0,02 <sup>**</sup>	

MSB: mulheres com o Subtipo B; MSNB: Mulheres com Subtipos Não-B; DP = desvio padrão; IC: intervalo de confiança; TARV: terapia antirretroviral; HPV: papiloma vírus humano; NIC: neoplasia intraepitelial; HTLV: vírus linfotrófico humano.

<sup>†</sup>Teste de Mann-Whitney; <sup>\*</sup>valor do teste t de Student; <sup>\*\*</sup>teste do  $\chi^2$ .

**Tabela 3** - Resultados da análise genotípica e das mutações de resistência aos ARV dos 52 isolados do HIV-1, comparando mães de filhos soropositivos e mães sem filhos soropositivos de Feira de Santana (BA)

Subtipos e mutações	MCFH		MSFH		Valor de p
	n	%	n	%	
Subtipo B	6	11,5	27	51,9	0,61 <sup>*</sup>
Subtipos NB	2	3,8	17	32,6	0,47 <sup>*</sup>
Sem mutações	6	11,5	33	63,4	0,07 <sup>*</sup>
Mutações ao INRT	2	3,8	11	21,1	0,08 <sup>*</sup>
Mutações ao INTR e IP	2	3,8	1	1,9	0,02 <sup>*</sup>

MCFH: mães com filhos soropositivos; MSFH: mães sem filhos soropositivos;

<sup>\*</sup>Teste  $\chi^2$  de Pearson; <sup>\*\*</sup>Teste Exato de Fisher; subtipos NB: (Subtipos C, F e recombinante BF); INTR: inibidores nucleosídeos da transcriptase reversa; IP: inibidores da protease.

## Discussão

Os resultados obtidos mostram uma tendência atual da dispersão do HIV-1, às cidades do interior, atingindo preferencialmente as mulheres e as camadas sociais mais pobres, como foi verificado no Estado do Rio de Janeiro, onde foi encontrada uma proporção elevada de mulheres infectadas (1,1 homem/ mulher)<sup>14</sup>. Em nosso Centro,

ainda encontramos uma proporção de 2,3 homens/ mulher infectada.

A pauperização e a baixa escolaridade predominante nessa amostra foram associadas à transmissão heterossexual e vertical do HIV-1 devido à falta de seguimento efetivo no serviço de pré-natal e acesso ao protocolo do Aids Control Trial Group (PACTG) adotado pelo Ministério da Saúde. Esta medida, em nosso país, vem reduzindo a transmissão vertical a taxas inferiores a 4%, de acordo com relatos recentes<sup>15</sup>.

A prevalência das doenças do trato genital inferior associada ao HIV-1 (24%) foi menor em nosso estudo em relação a outro realizado em Salvador<sup>16</sup>. Isto é decorrente da sistemática diagnóstica e do nível de assistência adotado neste município. A prevalência do vírus linfotrópico humano (HTLV) foi muito elevada (4/96=4,1%) provavelmente devido à sua difusão a partir de Salvador, a capital de maior prevalência (1,8 %) deste vírus no país<sup>17</sup>. Em nosso estudo, ele esteve associado às portadoras do Subtipo B, o que pode ser explicado pelas condições socioeconômicas, baixa escolaridade, evolução mais lenta da doença e baixa imunidade encontrada nestas pacientes, portadoras da infecção dual HIV-1/HTLV.

O perfil da diversidade genética e do padrão de resistência aos ARV encontrado no presente estudo difere dos primeiros estudos de sequenciamento do HIV-1 na Bahia que mostraram uma proporção maior do subtipo B e bem inferior do recombinante BF, não tendo sido identificados os subtipos C e F<sup>18</sup>. Estudo recente realizado em Salvador mostrou uma maior diversidade e elevadas taxas de resistência primária aos ARV<sup>19</sup>.

O elevado percentual de 25,0% de isolados com resistência secundária encontrado em nosso estudo é muito superior aos relatos nacionais de 7,0% para os inibidores

da transcriptase reversa (INTR) e 1,5% para os inibidores da protease (IP). A taxa de resistência primária de 2% porém foi inferior aos mesmos relatos<sup>20</sup>.

Este padrão de resistência teve uma associação maior ao subtipo B do que ao recombinante BF e aos subtipos C e F. Isto provavelmente se deve ao tempo de inserção desta variedade na população estudada, sua evolução lenta para Aids e sua exposição a diversos esquemas terapêuticos que induzem o surgimento da resistência.

A baixa escolaridade, a questão do preconceito e as condições socioeconômicas que levam à não-adesão e interrupção do tratamento fecham o ciclo vicioso que causa a falência terapêutica e o desenvolvimento de novas mutações adaptativas do vírus. A dispersão de cepas virais mutantes pelo fluxo migratório e intercâmbio comercial torna cada vez mais difícil o controle da epidemia e erradicação do HIV-1<sup>21</sup>.

O tempo presumido de doença, o maior tempo de tratamento, o estágio clínico mais grave e a maior prevalência de doenças ginecológicas também foram associados às portadoras do subtipo B, provavelmente pelos mesmos motivos anteriormente referidos.

Por se tratar de uma população ainda sexualmente ativa, esta proporção elevada de cepas recombinantes e subtipos oriundos de outras regiões pode causar no futuro um profundo impacto na evolução da doença e no prognóstico reprodutivo destas mulheres.

## Agradecimentos

Ao Doutor Ênio Oliveira, do Centro de Referência DST/HIV/Aids, ao Laboratório Central do Estado (LACEN), ao Laboratório Avançado em Saúde Pública (LASP) e à FIOCRUZ (BA), que colaboraram com a realização deste estudo.

## Referências

1. Joint United Nations, Programme on HIV/AIDS (UNAIDS). World Health Organization (WHO). AIDS epidemic update: December 2007 [Internet]. Geneva: UNAIDS/WHO; 2007 [citado 2008 Jun 12]. Available from: [http://data.unaids.org/pub/EPISlides/2007/2007\\_epiupdate\\_en.pdf](http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf)
2. Brasil. Ministério da Saúde. Boletim Epidemiológico AIDS e DST [Internet]. 2006 jul-dez/2007 jan-jun;4(1) [citado 2008 Jun 12]; Disponível em: [http://www.aids.gov.br/data/documents/storedDocuments/%7BB8EF5DAF-23AE-4891-AD36-1903553A3174%7D/%7B721527B6-FE7A-40DF-91C4-098BE8C704E0%7D/Boletim2007\\_internet090108.pdf](http://www.aids.gov.br/data/documents/storedDocuments/%7BB8EF5DAF-23AE-4891-AD36-1903553A3174%7D/%7B721527B6-FE7A-40DF-91C4-098BE8C704E0%7D/Boletim2007_internet090108.pdf)
3. Hudson CN. Eletive caesarean section for the prevention of vertical HIV-1 infection. *Lancet*. 1999;353(9158):1030-1.
4. Duarte G, Quintana SM, ElBeitune P. Fatores que influenciam a transmissão vertical do vírus da imunodeficiência humana. *Rev Bras Ginecol Obstet*. 2005;27(11):698-705.
5. Renjifo G, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, et al. Preferencial in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS*. 2004;18(12):1629-36.
6. Tápia N, Franco S, Puig-Bassagoiti F, Menéndez C, Alonso PL, Mshinda H, et al. Influence of human immunodeficiency virus type 1 subtype on mother-to-child transmission. *J Gen Virol*. 2003;84(Pt 3):607-13.
7. Martinez AM, Hora VP, Santos AL, Mendoza-Sassi R, Von Groll A, Soares EA, et al. Determinants of HIV-1 mother-to-child transmission in Southern Brazil. *An Acad Bras Cienc*. 2006;78(1):113-21.

Santos ES, Araújo AF, Galvão-Castro B, Carlos Junior Alcântara L.

8. Barré-Sinoussi F. HIV as the cause of AIDS. *Lancet*. 1996;348(9019):31-5.
9. Robertson DL, Sharp PM, McCutchan FE, Hahn BH. Recombination in HIV-1. *Nature*. 1995;374(6518):124-6.
10. Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol*. 2007;81(19):10209-19.
11. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med*. 2008;358(15):1590-602.
12. Morgado MG, Guimarães ML, Neves Júnior I, dos Santos VG, Linhares-de-Carvalho MI, Castello-Branco LR, et al. Molecular epidemiology of HIV in Brazil: polymorphism of the antigenically distinct HIV-1 B subtype strains. The Hospital Evandro Chagas AIDS Clinical Research Group. *Mem Inst Oswaldo Cruz*. 1998;93(3):383-6.
13. Dourado I, Milroy CA, Mello MAG, Ferraro GA, Castro-Lima Filho H, Guimarães ML, et al. HIV-1 seroprevalence in the general population of Salvador, Bahia State, Northeast Brazil. *Cad Saúde Pública*. 2007;23(1):25-32.
14. Eyer-Silva WA, Basílio-de-Oliveira CA, Morgado MG. HIV infection and AIDS in a small municipality in Southeast Brazil. *Rev Saúde Pública*. 2005;39(6):950-5.
15. Kakehasi FM, Pinto JA, Romanelli RMC, Carneiro M, Cardoso CS, Tavares MCT, et al. Determinants and trends in perinatal human immunodeficiency virus type 1 (HIV-1) transmission in the metropolitan area of Belo Horizonte, Brazil: 1998-2005. *Mem Inst Oswaldo Cruz*. 2008;103(4):351-7.
16. Oliveira PM, Mascarenhas RE, Ferrer SR, Oliveira RPC, Travessa IEM, Gomes MVC, et al. Vulvovaginites em mulheres infectadas pelo vírus da imunodeficiência humana. *Rev Bras Ginecol Obstet*. 2008;30(3):121-6.
17. Bittencourt AL, Dourado I, Bastos Filho P, Santos M, Valadão E, Alcântara LC, et al. Human T-cell lymphotropic virus type 1 infection among pregnant women in northeastern Brazil. *J Acquir Immune Defic Syndr*. 2001;26(5):490-4.
18. Couto-Fernandez JC, Morgado MG, Bongertz V, Tanuri A, Andrade T, Brites C, et al. HIV-1 subtyping in Salvador, Bahia, Brazil: a city with African sociodemographic characteristics. *J Acquir Immune Defic Syndr*. 1999;22(3):288-93.
19. Pedroso C, Queiroz AT, Alcântara LC, Drexler JF, Diaz RS, Brites C. High prevalence of primary antiretroviral resistance among HIV-1-infected adults and children in Bahia, a northeast state of Brazil. *J Acquir Immune Defic Syndr*. 2007;45(2):251-3.
20. Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, et al. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio de Janeiro, Brazil: assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active antiretroviral therapy. *Mem Inst Oswaldo Cruz*. 2005;100(1):73-8.
21. Rozenbaum W, Katlama C, Massip P, Bentata M, Zucman D, Delfraissy JF, et al. Treatment intensification with abacavir in HIV-infected patients with at least 12 weeks previous lamivudine/zidovudine treatment. *Antivir Ther*. 2001;6(2):135-42.

**Manuscrito 03**

**LACK OF HIGH-LEVEL RESISTANCE MUTATIONS IN HIV-1 BF  
RECOMBINANT STRAINS CIRCULATING IN NORTHEAST BRAZIL**

Running title: **HIV-1 diversity in Northeast Brazil**

Authors: Joana P Monteiro<sup>1\*</sup>, Adriano F S Araújo<sup>1\*</sup>, Edson S Santos<sup>2</sup>, Bernardo Galvão-Castro<sup>1,2</sup>, Luiz C J Alcântara<sup>1,2</sup>.

Author's affiliation:

<sup>1</sup>Laboratório Avançado de Saúde Pública (LASP), Centro de Pesquisa Gonçalo Moniz (CPqGM), Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Bahia, Brasil;

<sup>2</sup> Fundação Bahiana para o Desenvolvimento das Ciências (FBDC), Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador, Bahia, Brasil.

\* These authors contributed equally to this work

Corresponding author:

Luiz C J Alcantara

LASP/CPqGM/FIOCRUZ. Rua Waldemar Falcão 121, Candeal, Salvador, Bahia, Brazil 40296-710. Telephone # 55 71 3176 2255. Fax # 55 71 3176 2300. E-mail: lalcan@bahia.fiocruz.br

## **ABSTRACT**

The high genetic diversity of HIV-1 influences virus transmission, resistance to antiretroviral drugs and disease progression. This work aimed to study the genetic variability and the prevalence of drug resistance-associated mutations (DRAM) of HIV-1 isolates from 50 women and 8 children from Feira de Santana, Bahia, Brazil. DNA samples were obtained and *pol* sequences were generated by PCR and direct sequencing. Phylogenetic analysis showed that 39 (67.2%) samples were subtype B, 4 (6.9%) F, 1 (1.7%) C and 14 (24.1%) BF recombinants. Four different BF recombination patterns were detected. Ten (17.2%) samples shared the same breakpoint within the reverse transcriptase (RT) sequence. Fifty-five (94.8%) isolates showed several resistance-associated mutations in the RT and the protease (PR) genes. Eleven (19.0%) isolates presented mutations associated with high-level of resistance: 10 (17.2%) to nucleoside RT inhibitors (NRTI), 4 (6.9%) to non-nucleoside RT inhibitors (NNRTI) and 2 (3.4%) to PR inhibitors (PIs). One therapy-naïve patient (1.7%) harbored primary resistance mutations to five NRTIs. Subtype B infected patients had, on average, 0.4 high-level DRAM per sequence while no mutations was observed in BF recombinants ( $p=0.005$ ), although the two groups were under ARV for a similar period of time. Our data show an increased prevalence of BF recombinants in this region and the predominance of a well-disseminated recombinant strain in Bahia, which could be related to adaptive advantages of these variants over the predominant subtype B.

Keywords: HIV-1, BF recombinant, drug resistance mutation

## **INTRODUCTION**

It is estimated that 33.2 million people are living with the human immunodeficiency virus-1 (HIV-1) worldwide. Out of these, 15.4 million are women and 2.5 million are children under 13 years. Every year, approximately 2.5 million people are contaminated, 18% of whom are children.<sup>1</sup> The majority of HIV infections result from heterosexual transmission, while mother-to-child transmission (MTCT) represents the main form of HIV infection among children. MTCT can occur during pregnancy, labor, delivery or breastfeeding and is related to multiple factors. This includes viral, maternal, behavioral, obstetric and newborn factors.<sup>2</sup> In Brazil, more than 506 thousands of HIV infections and roughly 13,000 cases of MTCT have been documented.<sup>3</sup>

The HIV is highly diverse and has been classified into two main types: HIV-1 and HIV-2. The HIV-1 is largely disseminated throughout the globe and is divided into three groups: M, O and N. The group M is further divided into 9 subtypes (A, B, C, D, F, G, H, J, and K), 43 Circulating Recombinant Forms (CRF) and several Unique Recombinant Forms (URF).<sup>4-7</sup> These various genotypes present differential geographic distribution worldwide. The subtype B is more disperse while the subtype C is more prevalent.<sup>8</sup> In Brazil, subtype B is the predominant genotype followed by B/F recombinants and subtypes F and C.<sup>7,9-12</sup> However, due to the vast territory and to an intense social and cultural heterogeneity, the Brazilian geographic regions have been characterized by striking differences concerning subtype distribution.

The extreme variability of HIV-1 is related to the high error rate of the reverse transcriptase enzyme and to the occurrence of homologous recombination of the viral RNA during replication.<sup>13</sup> The accumulation of mutations favors the

emergence of virus mutants that escape the immune response and ARV therapy. In Brazil, HIV-infected patients are provided with full and free access to all U.S. Food and Drug Administration-approved ARV since 1996. Although it has contributed to an increase in survival and quality of life, the sequential use of ARV has also contributed to the emergence of viral resistance, which has consequences at the individual level, such as virological failure and at the community level, such as primary/secondary resistance and increased transmission, including MTCT.

We conducted this study in order to characterize the prevalence of DRAM and the molecular diversity of HIV isolates in a population composed by women and children from Feira de Santana, an inner city of Bahia State, in the Northeast of Brazil. Our results indicate a need for expanding genotyping tests.

## **METHODS**

### **Population**

During 2007, blood samples were collected from 50 women and 8 children randomly selected at the Reference Center for Sexual Transmitted Disease (STD)/AIDS of Feira de Santana. Samples were processed at the Advanced Laboratory of Public Health (LASP)/CPqGM/FIOCRUZ in Bahia, Brazil. All participating patients gave written informed consent. Children (2-11 years), infected through MTCT, were included with the consent of parents or guardians. Clinical and epidemiological data were obtained from medical records.

### **DNA extraction and PCR amplification**

Genomic DNA was extracted using the Qiagen extraction kit (QIAGEN, Valencia, CA). The amplification of the *pol* fragment (positions 2344-3160 relative to HXB2 genome) was obtained through nested polymerase chain reaction (PCR) in a Perkin Elmer 9600 Thermal Cycler (Waltham, MA) using 500ng of DNA. DP10 and LR54 and DP16 and RT12 were used as outer and inner primers respectively. The PCR cycling conditions were as follow: three cycles of denaturing at 95°C for 3 min, annealing at 55°C for 1 min and primer extension at 72°C for 1 min; another 35 cycles of denaturing at 95°C for 1 min, annealing at 55°C for 45 sec and primer extension at 72°C for 1 min and a final extension at 72°C for 10 min.

### **Sequencing and Phylogenetic Analysis**

PCR products were purified with Qiagen columns (QIAGEN, Valencia, CA) and sequenced in an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) using a Big Dye Terminator kit (Applied Biosystems, Foster City, CA), the

same PCR inner primers and additional primers (LR49 and LR51).<sup>14</sup> Sequences were assembled using SeqScape v2.1.1 software (Applied Biosystems, Foster City, CA) and checked for contamination by BLAST search against the HIV-1 sequences database [<http://blast.ncbi.nlm.nih.gov>].

Sequences were aligned with a subtype reference set from Los Alamos database [<http://hiv-4web.lanl.gov>] and with other sequences previously characterized from Brazil and from other regions using the MUSCLE software<sup>15</sup> and edited manually in GENEDOC.<sup>16</sup> Phylogenetic analyses were performed using the PAUP\* 4.0b10 software<sup>17</sup> to generate neighbor-joining (NJ) and maximum likelihood (ML) trees with the General Time Reversible (GTR) model of nucleotide substitution.<sup>18</sup> The reliability of the nodes was assessed by bootstrap analysis (1000 replicates). Furthermore, the likelihood ratio test method was used to calculate statistical support for the branches. The trees were drawn with the TreeView program.<sup>19</sup> Recombination was identified by the Bootscanning method implemented in the SIMPLOT software<sup>20</sup> and the GENEDOC software was used to determine the specific crossover point by visual inspection of the alignment. Each part of the mosaic structure was confirmed by the reconstruction of bootstrapped NJ trees as described above. The BF recombinant forms determined by Simplot were aligned with reference sequences of CRFs from the Los Alamos database and analyzed by phylogenetic reconstruction as described above.

### **Drug Resistance analysis**

The level of susceptibility to antiretroviral drugs was inferred by the Stanford University HIV Drug Resistance Database HIVdb program, version 6.0.5 [<http://hivdb.stanford.edu>]. The Stanford algorithm assigns to each mutation a

score of resistance. We considered only the mutations associated with high levels of drug resistance to calculate the prevalence. This data was correlated with the treatment that each patient was undergoing.

### **Statistical analysis**

Epidemiological, clinical and laboratorial (TCD4 cells and viral load) data were analyzed in the SPSS software, version 14.0 for Windows [Chicago: SPSS Inc].

## RESULTS

### General characteristics of the study population

The studied sample consisted of 50 women and 8 children. The age ranged from 20 to 73 years among women (mean = 37.9 years) and from 2 to 11 among children (mean age = 7 years). Out of the children, 5 (62.5%) were male and 3 (37.5%) were female. Regarding the marital status, the women were either single (28.0%), married and/or with a stable relationship (56%) or widow (16.0%). Most of them was within the low-income class (83.0%) and had less than 4 years of education (65.0%). Also, the majority of the infections were related to sexual transmission (96.6%), while one patient was intravenous drug user (IDU) (1.7%) and one reported blood transfusion (1.7%). All the children were infected through vertical transmission during pregnancy or childbirth. Four patients were mother-child pairs (BR88FSM131-BR131FSC88; BR91FSM83-BR83FSC91; BR100FSM145-BR145FSC100; BR110FSM111-BR111FSC110). Clinical data such as viral load, CD4 T cell counting, time since the diagnosis and time of ARV treatment are summarized in Table 1. According to the CDC Classification System (CDC, 1993), 13 (22.4%) patients were in category A, 39 (67.2%) were in category B and 6 (10.3%) were in category C. Among women, 10 (20.0%) were in category A, 34 (68.0%) were in category B and 6 (12.0%) were in category C. Three (37.5%) children were in category A and 5 (62.5%) were in category B. Forty-eight (82.8%) patients were under ARV therapy, while 10 (17.2%) were therapy-naïve.

The phylogenetic analysis of the *pol* sequences revealed that 39 (67.2%) isolates were subtype B, 4 (6.9%) were subtype F1 and 1 (1.7%) was subtype C (Fig. 1). Fourteen (24.1%) samples did not cluster inside of any of the pure-

subtype clusters. Conversely, twelve (20.7%) of these sequences formed a monophyletic cluster in the three located between subtypes B and F (Bootstrap = 70). Therefore, these 14 samples were further submitted to recombination analysis through the Bootscanning tool, using Simplot software.<sup>20</sup> The twelve sequences inside the monophyletic cluster presented the same recombination pattern between subtypes F and B, while the other two samples (BR35FS and BR50FS) presented each one, a unique recombination pattern (Fig. 2a). The crossover points of the B/F recombinant sequences were investigated by visual inspection of the dataset containing reference sequences of both subtypes. This analysis showed that 10 out of the 12 samples with the same BF pattern in the Bootscanning plot also shared breakpoint at the same location (between positions 2553 and 2584 relative to the HXB2 genome), while the other two samples (BR69FS and BR70FS) presented breakpoints between positions 2542-2572. The sample BR50FS, showed a F/B/F profile within the *pol* fragment, with two different breakpoints: the first located between positions 2518-2536, which was also present in the sample BR35FS and the second located between positions 2645-2661. NJ and ML analyses of the individual fragments using the GTR model of nucleotide substitution confirmed the Bootscanning recombination results (Fig. 2b).

In order to further characterize the recombinant viruses circulating in this population, we compared their genetic pattern with recombinant sequences previously characterized in Brazil as well as with the established CRFs and pure B and F sequences. The NJ (Fig. 3) analysis shows that 10 samples with the same BF recombination pattern also formed a monophyletic cluster separated from the pure subtypes B and F and the previously characterized CRFs

clusters. The sample BR50FS clustered as an out-group of the CRF40 cluster<sup>21</sup> and samples BR69FS and BR70FS clustered inside the CRF28/CRF29 group. The sample BR35FS did not cluster with any characterized CRF or pure subtype sequence.

The prevalence of DRAM was investigated in the *pol* sequences of the HIV isolates circulating in Feira de Santana. We found several resistance-associated mutations in the reverse transcriptase (RT) and the protease (PR) regions in 55 (94.8%) isolates. Eleven samples (19.0%) presented mutations associated with high-level of resistance to HIV drugs: 10 (17.2%) to NRTI, 4 (6.9%) to NNRTI and 2 (3.4%) to PIs (Table 2). The mutations related with resistance to RT inhibitors were: M184V (6), M41L (6), T69N (1), G190A (2), G190S (1), K103N (2), K101E (1) and Y181C (1); these mutations confer resistance to most ARV. Mutations within the PR region were detected in the sequences of 2 patients (BR100FSM145 and BR51FS): the N88D (1) and I50L (1) substitutions, which confer resistance to Nelfinavir (NFV) and Atazanavir (ATV) PIs respectively. Among the patients who had high resistance mutations, 10 were using the treatment regimen that they were resistant to. Ten (90.9%) patients presented DRAM to at least one of the drug of their ARV regimen, 2 (18.2%) had DRAM to two of the drug of their ARV regimen and 2 (18.2%) had DRAM to three of the drug of their ARV regimen (Table 2). The mean number of mutations associated with high-level of resistance per virus was 0.4 among women and 0.5 among children. We have compared the clinical, epidemiological and genetic features between the subtype B and the BF recombinants groups (Table 3). There were no significant difference between the groups regarding age, viral load, TCD4 cell count, time under ARV therapy

and time of the first HIV positive diagnostic. Among subtype B-infected individuals, 8 (20.5%) were in clinical category A of the CDC classification, 26 (66.7%) were in category B and 5 (12.8%) were in category C. Among subtype F-infected patients, 4 (28.6%) were in category A, 9 (64.3%) were in category B and 1 (7.1%) was in category C. Subtype B infected patients had, on average, 0.4 high-level DRAM per sequence while no mutations was observed in BF recombinants ( $p = 0.005$ ) (Table 3).

## **DISCUSSION**

In Brazil, following a worldwide trend, the beginning of the epidemic was restricted to large urban centers, affecting specific risk groups such as homosexuals, IDUs and hemophiliacs. However, within the last two decades, the epidemic has reached new trends, spreading to cities of medium and small sizes, a phenomenon described as internalization.<sup>2</sup> In this context, this work represent an effort to characterize the molecular epidemiology and the drug resistance profile of the HIV-1 circulating among women and children in Feira de Santana, an inner city in the Northeast of Brazil, the region with the highest level of poverty in the country.

Feira de Santana is located 107 km from Salvador, the capital of Bahia State and is the second most populous city in the state. It is also situated at one of the major road junctions of the Northeast of the country, connecting it with the Midwest, the South and the rest of the Northeast regions. Until March 2007, the Reference Center for STD/AIDS in Feira de Santana has reported 781 cases of HIV infections: 569 (72.9%) men, 180 (23.0%) women (male/female ratio = 3.2:1) and 32 (4.1%) children infected through vertical transmission.

We have studied the clinical, epidemiological and virological characteristics of 58 HIV-infected individuals, including 50 women and 8 children. Similar to what has been observed in other parts of the country, the majority was in the low-income class (83.0%) and reported low level of education (65.0%). Sexual transmission was by far the predominant route of contamination. Similar data were previously observed in Salvador.<sup>22,23</sup> We compared the viral load, T CD4 cell count, time since diagnosis (in years) and time of ARV treatment (in years) between women and children (Table 1). Although the children means were

higher for all categories, there were no statistical difference between the two groups and we believe that these higher numbers are related, in part, to the small number of the group. None of the children were in the CDC clinical category C, which is characterized by AIDS-indicative conditions. A higher proportion of children (37.5%) were in the category A, the asymptomatic stage, compared with women (20.0%). On the other side, a higher proportion of women were in the clinical categories B (68.0%) and C (12.0%) compared with children (62.5% and 0% respectively).

The phylogenetic analyses of the *pol* sequences showed the presence of 3 distinct subtypes (B, C and F) and 4 different patterns of BF recombinant forms in this population. The subtype B was the most prevalent (67.2%) followed by BF recombinants, which accounted for 24.1% of the viruses, F1 (6.9%) and C (1.7%). This data show an increased prevalence of BF recombinants if compared with previous studies in Bahia and Brazil. In 1999, Couto-Fernandez and cols., reported 3.4% of BF recombinants among a group of individuals infected through sexual contact in Salvador.<sup>24</sup> In 2003, Brindeiro and cols studied samples from Northeast Brazil and showed 80% B, 10% BF, 6% F1 and 4% C.<sup>25</sup> A recent report from our group has shown 84.0% B, 13.1% BF, 2.3% F1 and 0.6% F/D<sup>11</sup> and another study conducted in Salvador showed, in 2007, 82.3% of B, 13% of BF, 3.5% of F and 1.1% of C.<sup>26</sup> Therefore, the present work confirms the expansion of BF recombinant forms in despite of the pure subtype F<sup>11,27,28</sup> and a decrease of the prevalence of subtype B in this region. The factors involved in the observed changes on subtype prevalence are still not understood but might involve the acquisition of genetic features that confer an increased viral fitness. In fact, the overtaking of a HIV epidemic by one genetic

strain over another has been described in other regions. For example, in South Brazil and in Thailand, the subtype C<sup>29</sup> and the CRF01\_AE,<sup>30,31</sup> respectively, overtook the subtype B and became the predominant genotypes. These observations reflect the complexity of the dynamics and evolution of HIV. The high prevalence of BF recombinants (24.1%) among our population mostly composed by women that reported sexual transmission (96.6%), also suggests an association between these variants and heterosexual transmission.

Among our samples, eight comprised 4 mother-child pairs (BR88FSM131-BR131FSC88, BR91FSM83-BR83FSC91, BR100FSM145-BR145FSC100, BR110FSM111-BR111FSC110). As expected, the sequences of each pair grouped together in both NJ and ML trees (Fig. 1) with high bootstrap levels (varying from 83 to 100).

Four different recombination patterns between subtypes B and F were observed among our samples (Fig. 2). The first pattern was represented by 10 samples that formed a monophyletic cluster in the subtyping tree (Fig. 1), as well as, in the trees based on the fragment of each subtype separately (Fig. 2b). These ten samples sharing the same breakpoint also clustered separated from the cluster formed by CRF28 and CRF29 sequences, which have both a very similar pattern of BF recombination with breakpoints in a very close genomic position. In this regard, we have previously demonstrated,<sup>11</sup> based on *gag* sequences, that a group of 11 BF recombinants from Salvador also clustered as a monophyletic group and shared a common breakpoint not found in other recombinant strains. Taken together, our results strongly indicate the presence of a new, well-disseminated HIV recombinant form in Bahia state. The second pattern of recombination that we found was represented by two samples

[BR69FS and BR70FS] that clustered inside the CRF28/CRF29 group. In fact, the breakpoints of these variants were at the same position as the CRFs described in the southeast of the country.<sup>7</sup> The third BF pattern was represented by sample BR35FS, characterized by an initial, small F fragment (192 bp) followed by a longer B fragment (624 bp). However, in its B fragment, this sample presented a region (~100 bp) that could not be assigned to any subtype. Finally, sample BR50FS represented the fourth BF recombination pattern and clustered close to the CRF40<sup>21</sup> sequences. This could be explained by the fact that this sample showed breakpoints close to the ones found in CRF40 (Fig. 2a). It is worth noting that all four recombinant patterns presented breakpoint in the beginning of the RT or at the end of the PR. Likewise, various CRFs also have crossovers at this genomic region [[www.hiv.lanl.gov](http://www.hiv.lanl.gov)], confirming the existence of recombination hotspots on the HIV genome<sup>32</sup>.

Forty-eight (82.8%) patients of this cohort were taking ARV. Although 55 (94.8%) individuals presented mutations in the TR and PR mutations associated with drug resistance, most of these mutations do not alter the sensitivity of the viruses to the drugs. However, 11 (19.0%) sequences showed mutations associated with high-level of resistance to RT inhibitors and 2 (4%) to PR inhibitors (Table 2).

From the individuals with high-level resistance mutations, 90.9% were taking at least one drug they were resistant to, 18.2% were taking 2 drugs they were resistant to and 18.2% were taking 3 drugs they were resistant to. However, their viral load and CD4 T remained controlled. This is consistent with the fact that the viral fitness drops after the acquisition of resistance mutations, making the strains become less efficient<sup>33</sup>. On the contrary, 9.1% of the individuals who

had mutations associated with high-level of resistance were not taking the drugs associated with the mutations in their sequences, and therefore, might have been infected by resistant viral strains. Indeed, one patient was therapy-naïve (BR22FS – Table 2) and presented the T69N mutation, which is associated with high resistance to 5 different drugs. Another study involving a larger number of naïve patients from Bahia showed high rates of primary resistance to NRTIs (9.8%), NNRTIs (11.4%) and PIs (5.0%).<sup>34</sup> Similarly, De Sa-Filho and cols demonstrated that 18.2% of samples from newly diagnosed patients in Southeast Brazil had primary antiretroviral resistance mutations.<sup>35</sup> These numbers reflect an ongoing transmission of resistant strains, an important obstacle to control the HIV/AIDS epidemic worldwide and point to the need of expanding genotyping tests for clinical management of the best therapy combination. This could be particularly important in Brazil, where full and free access to all ARV make the issue of antiretroviral resistance especially critical. An efficient transmission of a resistant strain is possibly the case of sample BR145FSC100, a 7-year old child with the DRAM M184V (high resistance to 3TC and FTC) that is also present in the sequence of his mother (BR100FSM145 – Table 2). However, BR145FSC100 did not present the RT mutations D67N and V118I or the PR mutations E35D, Q58E and N88D, all present in the mother. In this case, either a less prevalent strain without the mutations was transmitted and gained fitness in the new host environment, or the mutations were reverted after the transmission. On the contrary, BR145FSC100, but not his mother, had the M41L mutation, indicating the acquisition after transmission. In fact, M41L confer high resistance to D4T and AZT, drugs that were part of the child regimen.

From all 6 individuals in CDC category C (AIDS symptoms), only 1 presented high-level DRAMs. The other 5 patients at AIDS stage were under ARV therapy, however, did not present high-level DRAMs. This is the contrary of what we would expect, since the development of AIDS symptoms is usually related to therapy failure and with the emergence of DRAMs. However, the absence of detectable viral resistance after treatment failure may result from factors such as the presence of drug-resistant minority viral populations, nonadherence to medications, lack of current knowledge of the association of certain mutations with drug resistance, and others.<sup>36</sup> The majority (9) of the patients with high-level DRAMs (Table 2) were in the B category (81.8%) and only 1 (9.1%) was in category A.

The presence of high-level DRAM was not identified among BF sequences, while among the subtype B group, the mean number of DRAM per patient was 0.41, in despite of the fact that these two groups were under the selective pressure of ARV therapy for a very similar period of time (Table 3). This difference was supported by  $p = 0.005$ . This could be related to an attenuation of the BF strains. On the contrary, De Sá Filho and cols showed high-level of resistance to reverse transcriptase inhibitors in both subtypes B and F and in recombinant forms B/F in samples from therapy-naïve<sup>35</sup> and ARV-treated patients.<sup>37</sup> We believe that further studies involving a higher number of samples might contribute to better understand the evolution of these recombinant strains in the local HIV epidemic.

This study demonstrates that the prevalence of BF recombinants is growing in Brazil, which could be related to adaptive advantages of these strains over the predominant subtype B and point to the importance of continuous surveillance

of the HIV epidemic through molecular genotyping. The monitoring of molecular subtypes, resistance mutations, expansion of HIV-1 and its transmission is crucial to support public health policies and the development of new therapies and vaccines.

## **ACKNOWLEDGMENTS**

This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), the Brazilian Ministry of Health STD/AIDS Program, CNPq, FBDC and FIOCRUZ.

## **SEQUENCE DATA**

The new sequences were reported to GenBank under the accession numbers GU121166 to GU12122.

## REFERENCES

1. UNAIDS. *Report on the global AIDS epidemic*. Geneva, UNAIDS. 2009.
2. Martinez AMB, Hora VP, Santos AL, Mendonza-Sasi R, Groll AZ, Soares EAJM, *et al.* Determinants of HIV-1 Mother-to-Child transmission in Southern Brazil. *Annals of the Brazilian Academy of Sciences* 2006;78:113-121.
3. AIDS – Epidemiological Bulletin Year V - # 1 - Epidemiological Week - 01 to 26 - January / June, 2008
4. McCutchan FE, Carr JK, Albert J, Essex M, Goudsmit J, Asjö B, *et al.* Simplified strategy for detection of recombinant HIV-1 group M isolates by *gag/env* heteroduplex mobility assay. *J Virol* 2000;74:363-370.
5. Casado G, Thomson MM, Sierra M, Najera R. Identification of a novel HIV-1 circulating ADG intersubtype recombinant form (CRF19\_cpx) in Cuba. *J Acquir Immune Defic Syndr* 2005;40:532-7.
6. Thomson MM, Casado G, Posada D, Sierra M, Najera R. Identification of a novel HIV-1 complex circulating recombinant form (CRF18\_cpx) of Central African origin in Cuba. *AIDS* 2005;19:1155-63.
7. De Sa Filho DJ, Sucupira MC, Casiero MM, Sabino EC, Diaz RS, Janini LM. Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS Res Hum Retroviruses* 2006;22:1-13.
8. Coetzer M, Cilliers T, Papathanasopoulos M, Ramjee G, Karim SA, Williamson C, *et al.* Longitudinal analysis of HIV type 1 subtype C

- envelope sequences from South Africa. *AIDS Res Hum Retroviruses* 2007;23:316-21.
9. Salemi M, de Oliveira T, Soares MA, Pybus O, Dumans AT, Vandamme AM, *et al.* Different epidemic potentials of the HIV-1B and C subtypes. *J Mol Evol* 2005;60:598-605.
  10. Monteiro JP, Ferraro GA, Oliveira T, Goldani LZ, Kashima S, Alcantara LC, *et al.* Genetic and biologic characterization of HIV type 1 subtype C isolates from south Brazil. *AIDS Res Hum Retroviruses* 2007;23:135-43.
  11. Monteiro JP, Alcantara LC, de Oliveira T, Oliveira AM, Melo MA, Brites C, *et al.* Genetic variability of human immunodeficiency virus-1 in Bahia state, Northeast, Brazil: high diversity of HIV genotypes. *J Med Virol* 2009;81:391-9.
  12. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med* 2008;358:1590-602.
  13. Pinto ME, Struchiner CJ. HIV-1 diversity: a tool for studying the pandemic. *Cad Saude Publica* 2006;22:473-84.
  14. Maia-Teixeira SL, Bastos FI, Hacker MA, Guimarães ML, Morgado MG. Trends in drug resistance mutations in antiretroviral-naïve intravenous drug users of Rio de Janeiro. *J Med Virol* 2006;78:764-9.
  15. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.

16. Nicholas KB, Nicholas Jr HB, Deerfield II DW. GeneDoc: Analysis and Visualization of Genetic Variation EMBNET News 1997;4:1-4.
17. Swofford D: PAUP 4.0: Phylogenetic analysis using parsimony (and other methods), 4.0b2a. Sinauer Associates, Inc., Sunderland, MA, 1999.
18. Posada D and Krandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics 1998;14:817-8.
19. Page RD. TreeView: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357-8.
20. Salminen MO, Carr JK, Burke DS, McCutchan FE. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS Res Hum Retrovir 1995;11:1423-1425.
21. Guimarães ML, Eyer-Silva WA, Couto-Fernandez JC, Morgado MG. Identification of two new CRF\_BF in Rio de Janeiro State, Brazil. AIDS 2008;22:433–438.
22. Galvão-Castro B, Couto-Fernandez JC, Mello MA, Linhares-de-Carvalho MI, Castello-Branco LR, Bongertz V, *et al.* A nationwide effort to systematically monitor HIV-1 diversity in Brazil: preliminary results. Brazilian Network for the HIV-1 Isolation and Characterization. Mem Inst Oswaldo Cruz 1996;9:335-8.
23. Dourado I, Milroy CA, Mello MA, Ferraro GA, Castro-Lima Filho H, Guimarães ML, *et al.* HIV-1 seroprevalence in the general population of Salvador, Bahia State, Northeast Brazil. Cad Saude Publica 2007;23:25-32.

24. Couto-Fernandez JC, Morgado MG, Bongertz V, Tanuri A, Andrade T, Brites C, *et al.* HIV-1 subtyping in Salvador, Bahia, Brazil: a city with African sociodemographic characteristics. *J Acquir Immune Defic Syndr* 1999;22:288-93.
25. Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, *et al.* Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *AIDS* 2003;17:1063-9.
26. Brennan CA, Brites C, Bodelle P, Golden A, Hackett J Jr, Holzmayer V, *et al.* HIV-1 strains identified in Brazilian blood donors: significant prevalence of B/F1 recombinants. *AIDS Res Hum Retroviruses* 2007;23:1434-41.
27. Quarleri JF, Rubio A, Carobene M, Turk G, Vignoles M, Harrigan RP, *et al.* HIV type 1 BF recombinant strains exhibit different pol gene mosaic patterns: descriptive analysis from 284 patients under treatment failure. *AIDS Res Hum Retroviruses* 2004;20:1100-7.
28. Barreto CC, Nishyia A, Araujo LV, Ferreira JE, Busch MP, Sabino EC. Trends in antiretroviral drug resistance and clade distributions among HIV-1--infected blood donors in Sao Paulo, Brazil. *J Acquir Immune Defic Syndr* 2006;41:338-41.
29. Soares EA, Martínez AM, Souza TM, Santos AF, Da Hora V, Silveira J, *et al.* HIV-1 subtype C dissemination in southern Brazil. *AIDS* 2005;19:S81-6.

30. Wasi C, Herring B, Raktham S, Vanichseni S, Mastro TD, Young NL, *et al.* Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide-binding enzyme immunoassay and heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. *AIDS* 1995;9:843-9.
31. Arroyo MA, Phanuphak N, Krasaesub S, Sirivichayakul S, Assawadarachai V, Poltavee K, *et al.* HIV Type 1 Molecular Epidemiology among High-Risk Clients Attending the Thai Red Cross Anonymous Clinic in Bangkok, Thailand. *AIDS Res Hum Retroviruses* 2010; 26:5-12.
32. Minin VN, Dorman KS, Fang F, Suchard MA. Phylogenetic mapping of recombination hotspots in human immunodeficiency virus via spatially smoothed change-point processes. *Genetics* 2007;175:1773-85.
33. Nijhuis M, Deeks S, Boucher C. Implications of antiretroviral resistance on viral fitness. *Curr Opin Infect Dis* 2001;14:23-8.
34. Pedroso C, Queiroz ATL, Alcantara, LCJ, Drexler JF, Diaz RS, Weyll N, *et al.* High Prevalence of Primary Antiretroviral Resistance Among HIV-1–Infected Adults and Children in Bahia, a Northeast State of Brazil. *J Acquir Immune Defic Syndr* 2007;45:251-3.
35. De Sa-Filho DJ, Ambar RF, Duarte NB, Matias RB, Candido V, Gagliani LH, *et al.* HIV type 1 diversity from newly diagnosed patients in Santos metropolitan area/Brazil. *AIDS Res Hum Retroviruses* 2009;25:925-9.

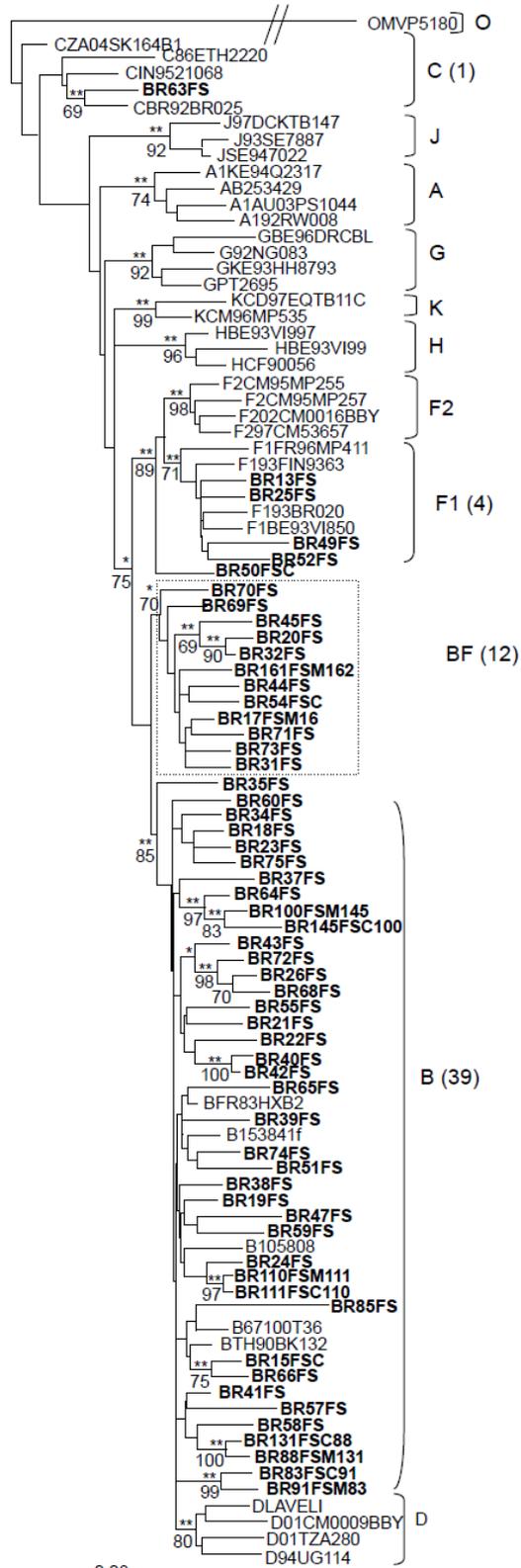
36. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, *et al.* Update of the drug resistance mutations in HIV-1: December 2009. *Top HIV Med* 2009;17:138-45.
  
37. De Sa-Filho DJ, Soares Mda S, Candido V, Gagliani LH, Cavaliere E, Diaz RS, *et al.* HIV type 1 pol gene diversity and antiretroviral drug resistance mutations in Santos, Brazil. *AIDS Res Hum Retroviruses* 2008;24:347-53.

## Figure Legends

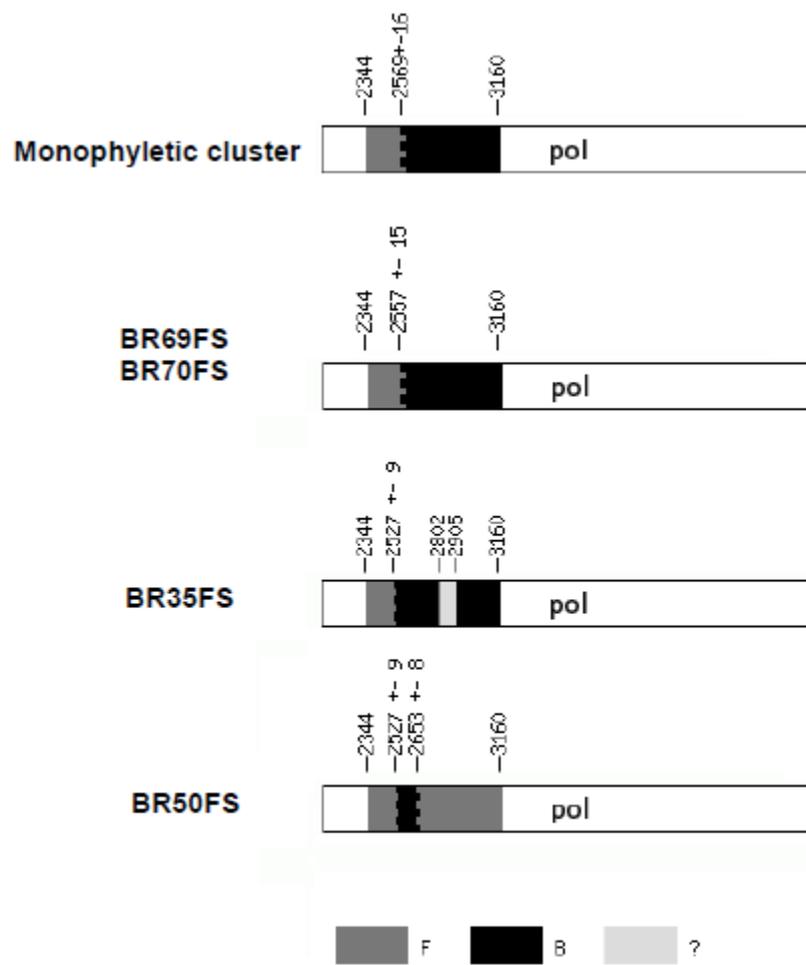
**Fig. 1.** NJ tree based on *pol* sequences showing phylogenetic relations between HIV-1 samples from Feira de Santana and the group M reference sequences from Los Alamos database. A group O sequence was used as outgroup. GTR+I+G was the nucleotide substitution model used. Branches supported by the ML method are indicated as \* ( $p < 0.05$ ) and \*\* ( $p < 0.001$ ). Bootstrap values for 1,000 replicates are indicated in percent. HIV-1 sequences generated in the present study are indicated in bold. 12 sequences formed a monophyletic group outside the subtype F1 cluster (shown within the box).

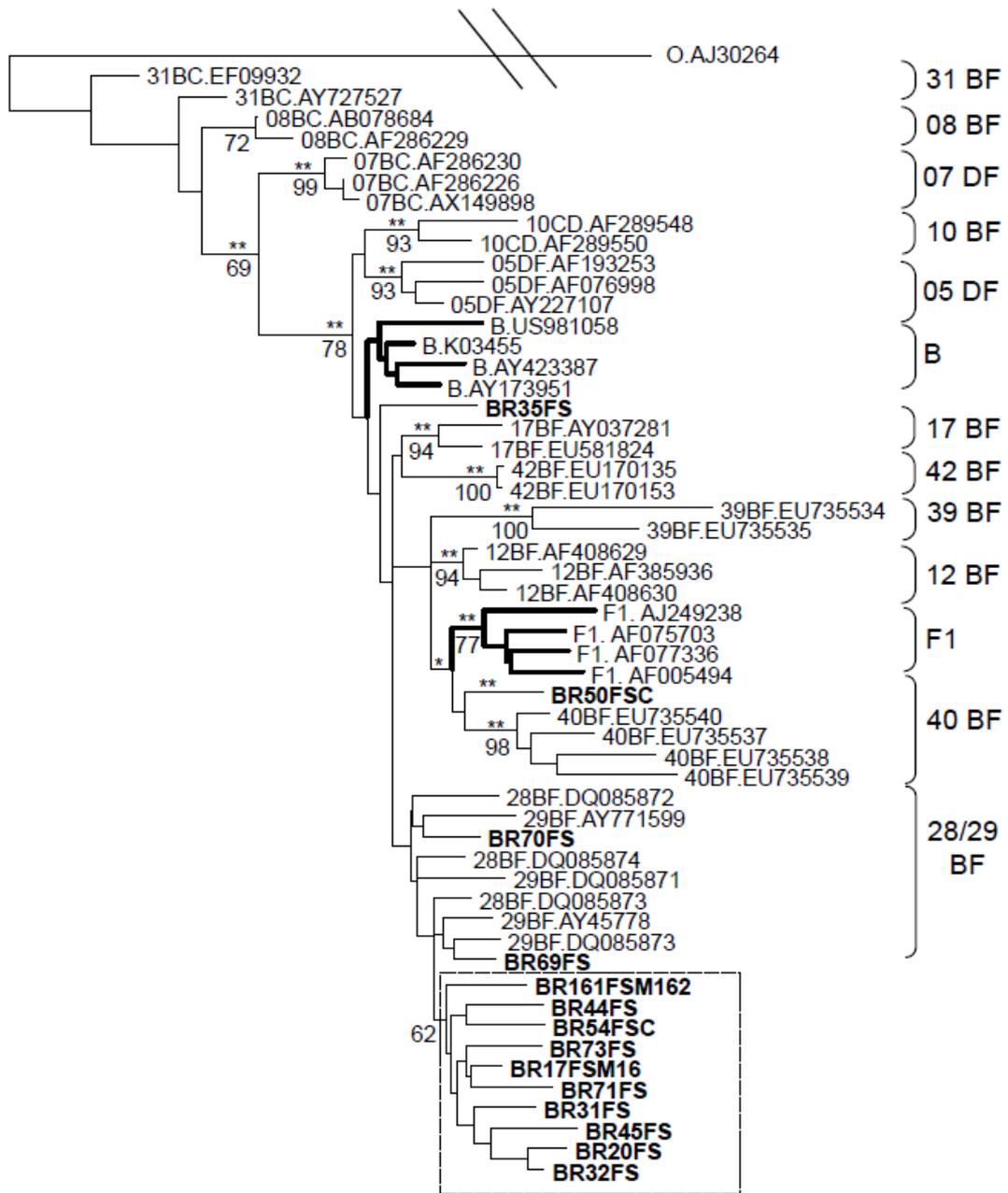
**Fig. 2.** (a) Patterns of recombination between subtypes B and F inside the *pol* gene found among HIV samples from Northeast Brazil. (b) Phylogenetic trees confirming the BF recombination in 10 samples with the same breakpoint: sequences were fragmented at the breakpoint and the segments were submitted to NJ analysis with Los Alamos reference set using the GTR+ G and the GTR+I+G substitution models respectively. Bootstrap values for 1000 replicates are indicated. Branches highly supported by the ML method are indicated as \*\* ( $p < 0.001$ ). HIV-1 sequences generated in the present study are indicated in bold. The same method was used to confirm recombination in the other 3 patterns.

**Fig. 3.** *pol* phylogenetic relationships among HIV BF recombinants from Feira de Santana, compared to other subtype B, F and CRFs sequences. **A** group O sequence was used as outgroup. Trees were constructed based on the NJ method using the GTR+I+G substitution model. Branches supported by the ML method are indicated as \* ( $p < 0.05$ ) and highly supported by \*\* ( $p < 0.001$ ). Bootstrap values for 1,000 replicates are indicated in percent. Recombinant sequences are indicated in bold. Ten sequences formed a unique monophyletic group shown within the box.



(a)





0.09

Table 1. Clinical and Laboratorial data of individuals followed at the Reference Center for Sexual Transmitted Disease (STD)/AIDS of Feira de Santana, Brazil.

	Group	N	Mean	Std. Deviation	Std. Error Mean	p <sup>‡</sup>	95% Confidence Interval of the Difference	
							Upper	Lower
<b>Viral load (copies/ml)</b>	Children	8	113,251.2	171,321.4	60,571.2	<b>0.022</b>	<b>12080.5</b>	<b>152645.2</b>
	Women	50	30,888.4	74,221.0	10,496.4			
<b>CD4 (cells/ml)<sup>*</sup></b>	Children	5	922.4	733.2	327.9	<b>0.012</b>	<b>110.8</b>	<b>831.8</b>
	Women	43	451.1	325.8	49.7			
<b>Time since Diagnosis (years)</b>	Children	8	5.4	3.5	1.2	0.171	-0.6	3.4
	Women	50	4.1	2.4	0.3			
<b>Duration of ARV treatment (years)<sup>§</sup></b>	Children	5	5.2	3.3	1.2	0.608	-1.4	2.4
	Women	43	3,3	2.3	0.3			

\* Only patients who had available data were considered on the mean calculation; § therapy-naïve patients were not included in the analysis; p, p value; ‡, t-test for Equality of Means.

Table 2. Clinical, Epidemiological and Genotypic characteristics of 11 individuals showing mutations associated with high-level of drug resistance.

Id	Age Subtype CDC		Mutation in TR	Mutation in PR	Predicted drug resistance			Drug Regimen	
					NRTI	NNRTI	PI		
BR145FSC100	9	B	B	M41L, M184V, G190S	-	3TC, FTC, D4T, AZT	NVP, EFV	-	AZT, 3TC, EFV
BR75FS	11	B	B	M41L	-	D4T, AZT	-	-	AZT, 3TC, EFV
BR60FS	26	B	B	M184V, K103N	-	3TC, FTC	DLV, EFV, NVP	-	AZT, 3TC, EFV
BR57FS	29	B	B	M41L	-	D4T, AZT	-	-	AZT, 3TC, ATV, ddi
BR51FS	31	B	B	M41L	I50L	D4T, AZT	-	ATV	AZT, 3TC, EFV, RTV
BR49FS	35	F	B	M41L, M184V, K101E, G190A	-	D4T, AZT, 3TC, FTC	NVP, DLV, EFV	-	AZT, 3TC, NVP
BR22FS	40	B	A	T69N	-	ABC, DDI, TDF, D4T, AZT	-	-	None
BR91FSM83	40	B	B	M184V	-	3TC, FTC	-	-	AZT, 3TC, EFV, ATV
BR100FSM145	41	B	C	M184V	N88D	3TC, FTC	-	NFV	ddi, 3TC, AZT, IDV
BR47FS	44	B	B	K103N, Y181C, G190A	-	-	NVP, DLV, EFV, ETR	-	AZT, ddi, EFV
BR52FS	57	F	B	M41L, M184V	-	D4T, AZT, 3TC, FTC	-	-	AZT, 3TC, NFV

Id = patient identification; TR = transcriptase reverse; PR = protease; NRTI = nucleotide reverse transcriptase inhibitors; NNRTI = non-nucleoside reverse transcriptase inhibitor; PI = protease inhibitor.

Table 3. Comparison of clinical, epidemiological and genetic features between the subtype B and the BF recombinants groups.

	Group	N	Mean	Std. Deviation	Std. Error Mean	p <sup>‡</sup>	95% Confidence Interval of the Difference	
							Lower	Upper
Age	B	39	33.28	14.598	2.338	0.987	-9.197	9.047
	BF	14	33.36	14.542	3.887			
Viral load (copies/ml)	B	39	39853.33	83430.148	13359.515	0.481	-84621.793	40399.889
	BF	14	61964.29	137249.756	36681.540			
CD4 (cells/ml) <sup>*</sup>	B	32	454.56	305.872	54.071	0.508	-334.942	168.400
	BF	12	537.83	504.548	145.650			
Duration of ARV <sup>§</sup> treatment (years)	B	32	3.69	2.520	0.445	0.479	-1.089	2.282
	BF	11	3.09	1.921	0.579			
Time since Diagnosis (years)	B	39	4.54	2.703	0.433	0.164	-.496	2.859
	BF	14	3.36	2.620	0.700			
DRAM per sample	B	39	0.41	0.850	0.136	<b>0.005</b>	<b>0.135</b>	<b>0.686</b>
	BF	14	0.00	0.000	0.000			

\* Only patients who had available data were considered on the mean calculation; § therapy-naïve patients were not included in the analysis; p, p value; ‡, t-test for Equality of Means.