

Camila Florido

Análise da diversidade genética no gene *nef* do Vírus da
Imunodeficiência Humana, Tipo 1 (HIV-1) em crianças e em
adultos recém infectados.

Tese apresentada à Universidade Federal de
São Paulo – Escola Paulista de Medicina,
para obtenção do Título de Mestre em
Ciências.

São Paulo

2007

Camila Florido

Análise da diversidade genética no gene *nef* do Vírus da
Imunodeficiência Humana, Tipo 1 (HIV-1) em crianças e em
adultos recém infectados.

Tese apresentada à Universidade Federal de
São Paulo – Escola Paulista de Medicina, para
obtenção do Título de Mestre em Ciências.

Orientador: Prof. Dr. Luiz Mario Ramos Janini

Co-orientador: Prof. Dr. Élcio de Souza Leal

São Paulo

2007

Flórido, Camila

Análise da diversidade genética no gene *nef* do Vírus da Imunodeficiência Humana, Tipo 1 (HIV-1) em crianças e em adultos recém infectados. / Camila Flórido -- São Paulo, 2007.

Xii, 110f.

Tese (Mestrado) – Disciplina de Doenças Infecciosas e Parasitárias da Universidade Federal de São Paulo – EPM/UNIFESP.

Título em inglês: Selective pressures on *nef* gene of human immunodeficiency virus type 1 (HIV-1): a comparative analysis between children and adults.

1. HIV-1 2. Crianças 3. Adultos 4. Pressão Seletiva

UNIVERSIDADE FEDERAL DE SÃO PAULO
ESCOLA PAULISTA DE MEDICINA

Departamento de Medicina
Disciplina de Doenças Infecciosas e Parasitárias

Chefe do Departamento:_____.

Coordenador do Curso de Pós-Graduação:_____.

Apoio financeiro da FAPESP

Processo nº 04/12044-3

Camila Florido

Análise da diversidade genética no gene *nef* do Vírus da
Imunodeficiência Humana, Tipo 1 (HIV-1) em crianças e em
adultos recém infectados.

Presidente da banca:

Prof. Dr. _____.

BANCA EXAMINADORA

Prof. Dr.

_____.

Dedico esta tese àquele que deixou saudades

Meu amado pai, Gilmar Rojas Florido, que se tua falta é um silêncio que cruza a minha vida é porque tua existência foi essencial, hoje, mais do que nunca tua presença se faz sentir, pois sou a continuidade do seu brilho, da sua luta.

Dedico àqueles que estão presentes

Minha mãe, Célia de Fátima Mendes Florido, um exemplo de mulher, de simplicidade, humildade, delicadeza, dedicação e perseverança, a luz da minha vida, meu porto seguro.

Meu irmão, Raphael Florido Neto, amor verdadeiro, suporte de afeto, fonte inesgotável de alegria e estímulo constante.

Dedico-lhes o sorriso que trago agora em minha face. Divide comigo os méritos dessa conquista, porque ela nos pertence; ela é tão vossa quanto minha.

Amo vocês!

AGRADECIMENTOS

A Deus por tudo!

Ao Dr. Ricardo Sobhie Diaz por ter concedido a oportunidade para que eu desenvolvesse este trabalho em seu laboratório, por tudo o que fez por mim e por minha família, pelo exemplo de profissional e mestre. Minha total gratidão.

Ao Dr. Luiz Mario Janini pela oportunidade e por estar presente em todos os momentos deste trabalho.

Ao Dr. Élcio Leal por todas as críticas construtivas

A todos do laboratório de retrovirologia, que me ajudaram em momentos críticos, que fizeram parte do meu dia a dia e compartilharam preocupações e alegrias.

Àos meus primos Renata Doracenzi, Vanessa Doracenzi e Thiago Doracenzi Moreira, pelo amor e união.

Àos meus avós Maria de Lourdes Silva Mendes, Lídia Rojas Florido, Antônio Mendes (*in memorian*) e Raphael Florido (*in memorian*) por todo o carinho e apoio, por se fazerem presentes sempre em meu coração.

Aos meus tios Sueli Mendes Doracenzi, Rosiris Rojas Florido, Gilson Rojas Florido e Maximino Tadeu Doracenzi, por todo o amor, dedicação, companheirismo, força.

A toda minha família por estarem presentes em todos os momentos de minha vida

Às amigas Priscilla Rapello, Camilla Estevão França, Iris Campos, pelos conselhos, brincadeiras, por todos os momentos que passamos juntas.

Aos Amigos Elidamar, Rodrigo, Daniel e Michelle, meu braço direito nesta jornada.

A todas as pessoas que fizeram parte da minha vida e torceram pelo meu sucesso.

Dedicatória	vi
Agradecimentos	vii
Listas	viii
Resumo	xii
1. Introdução	1
1.1. Histórico e epidemiologia da infecção pelo HIV/AIDS	2
1.1.2. Histórico	2
1.1.3. Epidemiologia	4
1.2. O HIV	6
1.2.1. Partícula viral	6
1.2.2. Organização genômica do HIV-1	7
1.2.2.1. Genes estruturais	8
1.2.2.2. Genes regulatórios	9
1.2.2.3. Genes acessórios	10
1.2.3. Mecanismo de invasão e ciclo replicativo do HIV-1	13
1.3. Diversidade viral	16
1.3.1. Origem da diversidade	16
1.3.2. Estudo dos subtipos do HIV-1	17
1.3.3. Resposta imune à infecção pelo HIV-1	20
1.3.4. Dinâmica populacional e interação vírus-hospedeiro	24
2. Objetivos	30
2.1. Objetivo geral	31
2.2. Objetivo específico	31
3. Casuística e métodos	32
3.1. Aspectos éticos	33
3.2. Critérios de inclusão no estudo	33
3.3. População estudada	33
3.4. Processamento do material	34
3.4.1. Extração do DNA proviral	34
3.4.2. Reação em cadeia pela polimerase (<i>nested</i> -PCR) para a amplificação do DNA	34
3.4.3. Detecção dos produtos amplificados.	37
3.4.4. Purificação do produto do segundo <i>round</i> da PCR	37
3.4.5. Reação do sequenciamento	38
3.4.6. Precipitação da reação de sequenciamento	40
3.4.7. Sequenciamento automatizado de ácidos nucleicos	40
3.4.8. Análise do sequenciamento	40

3.4.9. Análise molecular dos dados	42
4. Resultados	50
4.1. Extração de DNA proviral	51
4.2. Reação em cadeia pela polimerase (<i>nested-PCR</i>) para amplificação do DNA proviral	52
4.2.2. Gene <i>nef</i>	52
4.2.3. Gene <i>pol</i>	53
4.3. Análise da diversidade genética	53
4.3.1. Identificação do subtipo B do HIV-1	53
4.3.2. Análise da pressão seletiva no gene <i>nef</i>	56
4.3.3. Análise da pressão seletiva no gene <i>pol</i>	68
5. Discussão	71
6. Conclusões	74
7. Referências	76

LISTA DE FIGURAS

- Figura 7.** Comprimentos dos ramos de filogenias do gene *nef* do HIV-1 obtidas de crianças. 57
- Figura 8.** Valores de dN/dS nos códons do gene *nef* do HIV-1 em crianças. 58
- Figura 9.** Valores de dN/dS nos códons do gene *nef* do HIV-1 em crianças e adultos recém infectados. 60
- Figura 10.** Comparação das taxas seletivas (dN/dS) nos códons do gene *nef* do HIV-1 em crianças e adultos recém infectados. 61
- Figura 11.** Códons sob seleção positiva do gene *nef* em epítomos Ab (anticorpos) de crianças e adultos recém infectados pelo subtipo B e C do HIV-1. 66
- Figura 12.** Códons sob seleção positiva do gene *nef* em epitopos CTL (CD8) de crianças e adultos recém infectados pelo subtipo B e C do HIV-1. 67
- Figura 13.** Comparação das taxas seletivas ($dN-dS$) nos códons do gene *pol* do HIV-1 em crianças e adultos. 69

LISTA DE TABELAS

Tabela 1. Descrição das formas recombinantes circulantes (CRFs)	19
Tabela 2. Descrição dos iniciadores externos utilizados na primeira etapa da <i>nested</i> -PCR do gene <i>nef</i> .	35
Tabela 3. Descrição dos iniciadores internos utilizados na segunda etapa da <i>nested</i> -PCR do gene <i>nef</i> .	35
Tabela 4. Descrição dos iniciadores externos utilizados na primeira etapa da <i>nested</i> -PCR do gene <i>pol</i> .	36
Tabela 5. Descrição dos iniciadores internos utilizados na segunda etapa da <i>nested</i> -PCR do gene <i>pol</i> .	36
Tabela 6. Descrição dos iniciadores utilizados na reação de seqüenciamento de DNA viral dos genes <i>nef</i> e <i>pol</i> .	39
Tabela 7. Código IUPAC.	41
Tabela 8. Quantidade de amostras processadas de cada gene estudado, genes <i>nef</i> e <i>pol</i> , dos grupos de crianças, com idades entre 0-3 anos e 4-6 anos.	51
Tabela 9. Análise por Maximum likelihood do gene <i>nef</i> do HIV-1 dos grupos de crianças e adultos recém infectados.	62

LISTA DE ABREVIATURAS E SÍMBOLOS

AIDS	<i>Acquired Immune Deficiency Syndrome</i>
ARV	<i>Associated Retrovírus</i>
AZT	3'-azido-2', 3'-didesoxitimidina: azitromicina, zidovudina ou retrovir
CDC	<i>United States Centers for Disease Control</i>
CRF	Formas Recombinantes Circulantes
CTL	Células T citotóxicas
dN	Taxas de substituições não sinônimas
DNA	Ácido desoxirribonucléico
DNTPs	Desoxirribonucleotídeos.
dS	Taxas de substituições sinônimas
DST	Doença Sexualmente Transmissível
EDTA	Ácido Etilenodiaminotetracético
<i>env</i>	Envelope
FDS	<i>Food and Drug Administration</i>
<i>gag</i>	Grupo Antigênico ou Antígeno Grupo Específico
gp	Glicoproteína
GTR	General Time Reversible
HAART	<i>Highly Active Anti-retroviral Therapy</i>
HBV	Hepatite B
HIV-1	<i>Human Immunodeficiency Vírus Type1</i>
HIV-2	<i>Human Immunodeficiency Vírus Type2</i>
HTLV-III	<i>Human T-Lymphotropic Vírus Type III</i>
ICTV	<i>International Committee of Viral Taxonomy</i>
LAV	<i>Lymphadenopathy Associated Vírus</i>
LCMV	Vírus de Leucemia em Murinos
LTR	<i>Long Terminal Repeats</i>
MHC	<i>Major Histocompatibility Complex</i>
<i>nef</i>	Fator de Regulação Negativa

NNRTIs	Inibidores de Transcriptase Reversa Não Nucleosídicos
NRTIs	Inibidores de Transcriptase Reversa Análogos aos Nucleotídeos
OMS	Organização Mundial de Saúde
PBMC	Células Mononucleares de Sangue Periférico
PCR	Reação em Cadeia pela Polimerase
<i>pol</i>	Polimerase
PR	protease
RNA	Ácido ribonucléico
RT	transcriptase reversa
SIV	Vírus da Imunodeficiência dos Símios
<i>tat</i>	<i>Trans-acting Transcription Transactivation</i>
UNAIDS	Comitê das Nações Unidas para a AIDS
<i>vif</i>	Fator de Infectividade Viral
<i>vpr</i>	Proteína Viral R
<i>vpu</i>	Proteína Viral U

RESUMO

Diversos estudos clínicos têm demonstrado que a dinâmica da AIDS na infecção pediátrica difere da observada em adultos. Em especial, diferenças na carga viral, taxa de depleção de linfócitos e rapidez no aparecimento dos sintomas são mais acentuados em crianças infectadas pelo HIV-1 (Luzuriaga *et al.*, 1999). Contudo, os processos evolutivos que ocorrem nos genes do HIV-1 durante a infecção de crianças ainda são pouco caracterizados. Em razão do sistema imunológico das crianças não estar completamente desenvolvido sugere que a pressão seletiva nos genes do HIV é distinta da observada em adultos. Para explorar a diversificação viral nesse contexto, comparamos a pressão seletiva do gene *nef* do HIV-1 em crianças (transmissão materno-infantil), com idades diferentes (*i.e.* 0-3 e 4-6 anos) e comparamos com a observada em adultos. Para esse propósito comparamos a razão entre o número de substituições não-sinônimas por sítio não-sinônimo (dN) e o número de substituições sinônimas por sítio sinônimo (dS) medidos códon a códon no gene *nef* do HIV-1 entre adultos e crianças. Os resultados mostraram que, independente do tempo de diversificação do HIV-1, não existem diferenças na entendida de seleção e nem na localização dos códons positivamente selecionadas entre crianças e adultos no nível populacional. Em adição, detectamos uma região no gene *nef* que concentra mais os códons sob seleção positiva. Essa região é densa em epitopos descritos para anticorpos e ausente de epitopos para CTL ou CD4-hekper (de acordo com os mapas de epitopos com os mapas de Los Alamos HIV databank). Esses resultados são importantes pois mostram que a pressão mediada por anticorpos é bastante significativa para a diversificação do HIV-1 no nível populacional.



1. INTRODUÇÃO

1.1. HISTÓRICO E EPIDEMIOLOGIA DA INFECÇÃO PELO HIV/AIDS.

1.1.1. HISTÓRICO.

A Síndrome da Imunodeficiência Adquirida (AIDS) foi detectada e definida primeiramente nos Estados Unidos, em 1981, pelo CDC (*United States Centers for Disease Control*), quando pesquisadores de Nova York e da Califórnia observaram entre os pacientes homossexuais jovens do sexo masculino, até então saudáveis, doenças raras como sarcoma de Kaposi, infecções oportunistas como a pneumonia causada por *Pneumocystis carinii* e casos inexplicáveis e persistentes de linfadenopatias. Estes homens tinham em comum um déficit imunológico que resultava numa significativa perda de células T-auxiliares. Estas observações despertaram a preocupação das autoridades de saúde pública nos Estados Unidos (Pinto & Struchiner, 2006; Gottlieb *et al.*, 1981; CDC, 1981; <http://www.aids.gov.br>).

A partir da observação do grupo de doenças que acometiam os pacientes infectados pelo agente causador da AIDS e da ocorrência dessa doença em grupos de risco bastante diversos (usuários de drogas injetáveis, hemofílicos, crianças nascidas de mães portadoras do vírus), em 1982, concluiu-se que a AIDS era causada por um microorganismo infeccioso transmitido de maneira semelhante ao vírus da Hepatite B (HBV): por contato sexual, uso de drogas, exposição a sangue e derivados e transmissão vertical. Neste mesmo ano foi diagnosticado o primeiro caso de AIDS no Brasil, em São Paulo (<http://www.aids.gov.br>; Francis *et al.*, 1983).

Em 1983, o grupo de pesquisa liderado pelo doutor Luc Montagnier (França), isolou um vírus que possuía a enzima transcriptase reversa (enzima que caracteriza os retrovírus), a partir de um paciente com uma doença associada ao novo quadro de comprometimento imunológico, este vírus foi denominado LAV – *Lymphadenopathy Associated Virus* (Barre-Sinoussi *et al.*, 1983) e parecia ser o primeiro indício do agente etiológico da AIDS.

Em 1984 os grupos dos doutores Robert Gallo e Jay Levy (EUA), publicaram também o isolamento de outros dois vírus contendo a enzima transcriptase reversa. Estes foram respectivamente denominados HTLV-III (*Human T-Lymphotropic Virus Type III*), que se tratava de um novo vírus HTLV (*Human T-Lymphotropic Virus*), o qual apresentava proteínas com diferenças significativas quando comparadas com as dos vírus então recentemente descritos

(HTLV-I e HTLV-II), e ARV (*Associated Retrovírus*). Após a identificação da seqüência de nucleotídeos destes vírus, constatou-se que os vírus anteriormente denominados de LAV, HTLV-III e ARV eram variações genéticas do mesmo vírus, estabelecendo-se, portanto, que o agente causador desta nova doença era um retrovírus (Ratner *et al.*, 1985; Marmor *et al.*, 2006; Levy *et al.*, 1984; Blattner, 1991; Gallo *et al.*, 2003, Gallo *et al.*, 1984). Contudo, somente em 1998, foi estabelecido pelo ICTV (*International Committee of Viral Taxonomy*), de acordo com as normas internacionais, baseadas principalmente nas características da estrutura da partícula viral, que o HIV pertencia à família *Retroviridae* e ao gênero *Lentivírus* (Coffin *et al.*, 1986).

A Comissão para Taxonomia de Vírus recomendou, em 1986, a adoção da designação de Vírus da Imunodeficiência Humana – HIV (da sigla em inglês: HIV - *Human Immunodeficiency Vírus*) para o novo vírus e Síndrome da Imunodeficiência Adquirida – SIDA (da sigla em inglês: AIDS - *Acquired Immune Deficiency Syndrome*), para a doença causada pelo mesmo (Coffin, 1996).

Estudos sorológicos do HIV evidenciaram dois tipos antigênicos, HIV-1 e HIV-2, que apesar das suas semelhanças, apresentam diferenças importantes que fornecem dados sobre a evolução, o tropismo e a patogênese viral. A maior diferença entre os tipos do HIV é a patogenicidade: em que o tipo 1 é o mais virulento e o principal responsável pela pandemia da AIDS (Grant, Cock, 2001; Reeves, Doms, 2002), enquanto que o tipo 2, descoberto em 1986 notificado na África central e oeste africano, é aparentemente menos patogênico, apresentando menor carga viral plasmática e/ou melhor resposta imune no controle da replicação viral quando comparada à infecção pelo HIV-1. A diferença da patogenicidade referida pode estar associada ao fato de que o HIV-2 costuma ter deleções no gene *nef*, o que pode reduzir a replicação viral. Por outro lado, no HIV-1, estas deleções não são freqüentes (Reeves &, Doms, 2002; MacNeil *et al.*, 2006). Além disso, estes dois tipos antigênicos do HIV possuem origens zoonóticas distintas, o HIV-1 e o HIV-2 derivam do SIVcpz e do SIVagm (Wang, 2000; Requejo, 2006; Korber *et al.*, 2000). Tanto o HIV-1 quanto o HIV-2, dividem-se em diferentes subtipos que se distinguem por suas organizações genômicas e relações filogenéticas (Hahn *et al.*, 2000; Reeves, Doms, 2002). O HIV-1 possui o gene *vpu* e está dividido em três grupos: grupo M (*main* ou *major*), composto por nove subtipos nomeados A-D, F-H, J e K (as variantes dos subtipos A e F são ainda segregadas como sub-subtipos A1 ou A2 e F1 ou F2, respectivamente); grupo O (*outlier*) e N (*new*). Além destes, 16 formas recombinantes (CRF) circulam na epidemia (Pinto, Struchiner, 2006). O HIV-2 está dividido em sete subtipos, A, B, C, D, E, F e G, e possui tanto o gene *vpu*, também presente no HIV-1, quanto o gene *vpx* (Piroozmand *et al.*, 2006). A variabilidade

genética entre o HIV-1 e o HIV-2 é de aproximadamente 25% de divergência em *gag*, *pol* e *env* (Reeves & Doms, 2002).

Em 1986, a primeira droga “anti-HIV” foi aprovada pelo FDS (*Food and Drug Administration*), órgão americano que regulamenta e autoriza a comercialização de medicamentos nos EUA e posteriormente em todo mundo. A droga foi designada de AZT e também por azitromicina, zidovudina ou retrovir. O AZT, desenvolvido em 1964, era utilizado inicialmente como quimioterápico, mas devido ao alto grau de toxicidade para os pacientes nas doses utilizadas, sua utilização fora abandonada. No início da década de 70, experimentos em laboratório mostraram que o AZT possuía também atividade antiviral e em 1985 foi descrita sua atividade “anti-HIV” (Colson *et al.*, 1991).

Novas drogas foram sendo desenvolvidas e em 1992 é apresentado o primeiro estudo envolvendo terapia combinada a partir da utilização concomitante de duas drogas. Em 1996, é desenvolvida uma nova terapia que combina a utilização concomitante de três ou mais drogas, a denominada Terapia Anti-retroviral Altamente Ativa – HAART (do inglês: *Highly Active Anti-retroviral Therapy*), popularmente chamada de coquetel anti-HIV. Porém, o veloz aparecimento de vírus resistentes a diferentes combinações destes fármacos, impeliu o desenvolvimento de novos fármacos e novas estratégias terapêuticas (De Clerc, 1998).

Atualmente, os fármacos anti-HIV disponíveis tem ação inibindo o sítio de ligação da enzima transcriptase reversa, promovendo a inibição alostérica da enzima transcriptase reversa, promovendo a inibição “competitiva” da protease e ainda inibindo a fusão das proteínas de membrana do HIV-1 com receptores de membrana das células alvo (Tantillo *et al.*, 1994, Wlodawer *et al.*, 1998).

1.1.2. EPIDEMIOLOGIA.

Entre as enfermidades infecciosas emergentes, a AIDS destaca-se pela grande magnitude e extensão dos danos causados às populações, de modo que as características da infecção pelo HIV e suas repercussões têm sido exaustivamente discutidas pela comunidade científica e pela sociedade em geral (Brito *et al.*, 2000; Wang *et al.*, 2000).

Segundo dados divulgados em 2006, pelo panorama geral da epidemia da AIDS da UNAIDS (*Joint United Nations Programme* 2004) divisão da Organização Mundial de Saúde (OMS), até dezembro de 2005, cerca de 38,6 milhões (33,4 – 46,0 milhões) de pessoas encontravam-se infectadas pelo HIV em todo o mundo (figura 1). Estima-se que ocorrem aproximadamente 4,1 milhões (3,4 – 6,2 milhões) de novas infecções pelo HIV e que cerca de 2,8 milhões de pessoas morrem por causa da AIDS.

Os primeiros casos de AIDS no Brasil foram identificados em indivíduos da região Sudeste, em 1982. Ao longo dos anos a AIDS avançou rapidamente em nosso país e, segundo o Boletim Epidemiológico da Coordenação Nacional de DST/AIDS, programa do Ministério da Saúde, o número oficial de casos notificados de AIDS no Brasil até junho de 2006 é de 433.067 pessoas. Contudo, observa-se a redução proporcional dos casos notificados de transmissão vertical. Ainda, segundo o boletim, o número de óbitos em decorrência da AIDS no período de 1980 a 2005 foi de 183.074 mortes. No ano de 2005 foi registrado, no Sistema de Informações sobre Mortalidade, um total de 11.026 óbitos por aids, representando taxa de mortalidade de 6,0/100.000 habitantes, variando de acordo com a região de residência.

A taxa de mortalidade de homens e mulheres aponta para um quadro de estabilidade, nos últimos anos. Esta estabilização nos índices de mortalidade se deve em grande parte ao tratamento anti-retroviral disponibilizado aos portadores do HIV/AIDS em nosso país, porém o índice de novos casos ainda é crescente no Brasil e em todo o mundo.

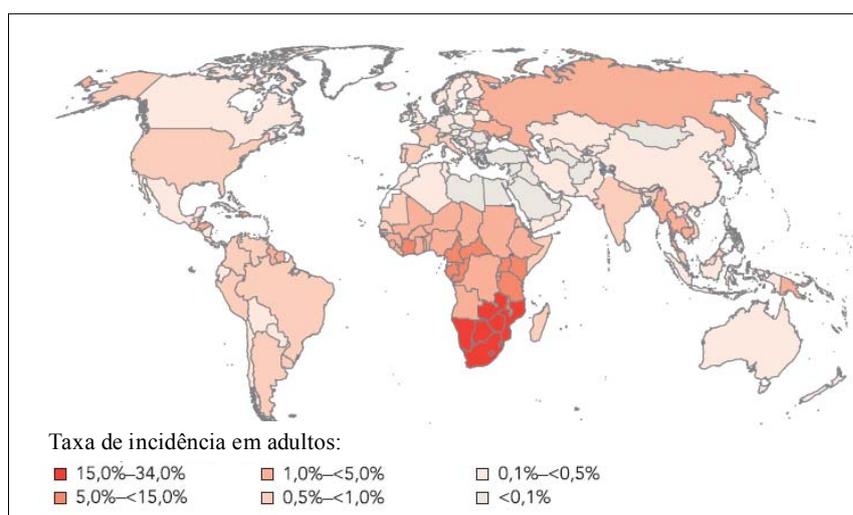


Figura 1. Distribuição global (em porcentagem) de 38,6 milhões de adultos infectados pelo HIV em 2005 (Panorama geral da epidemia de AIDS, UNAIDS, volume 2, página 15, 2006).

1.2. O HIV.

1.2.1. PARTICULA VIRAL.

O HIV-1 é morfologicamente esférico com aproximadamente 100nm de diâmetro e apresenta um envoltório de natureza lipoprotéica denominado envelope que recobre o nucleocapsídeo. O envelope do HIV é composto por uma bicamada lipídica derivada da membrana celular do hospedeiro, na qual se encontram inúmeras proteínas celulares como MHC (*Major Histocompatibility Complex*), actina e ubiquitina, de maneira que possui uma região voltada para o interior do vírus e outra externa, que é responsável pela afinidade com receptores celulares. Nessa região externa, são encontradas 72 projeções contendo trímeros de glicoproteínas. As glicoproteínas, Gp120 e Gp41, derivam de uma molécula precursora, a Gp 160, que é clivada por enzimas do aparelho de Golgi na célula hospedeira. A Gp120 ou proteína externa da superfície prende-se ao vírus através de interações com a Gp41 ou proteína transmembrana. A Gp 120 desempenha um papel fundamental no reconhecimento e ligação do vírus aos receptores celulares (Turner & Summers, 1999; Pinto & Struchiner, 2006).

Internamente, o envelope é sustentado pela p17, uma proteína estrutural da matriz viral, que é importante para a integridade do vírion (partícula viral completa) e para a incorporação das proteínas do envelope nos vírions maduros (Requejo, 2006).

O genoma do HIV é composto por duas fitas simples de RNA, idênticas e de polaridade positiva, estabilizado pela nucleoproteína p7. O genoma é envolto pelo capsídeo viral que possui a forma de tronco cônico, formado por várias unidades da proteína p24, bem como as proteínas p6 e p9, que estão associadas ao material genético. Dentre as proteínas do capsídeo encontram-se a transcriptase reversa (RT, heterodímero de p51 e p66), protease (PR, homodímero de p15), integrase (IN, homotetrâmero de p11), Nef (p27), Rev (p19), Tat (p14), Vif (p23), Vpr (p18) e Vpu (p16) (Pinto & Struchiner, 2006; Arthur *et al.*, 1992; Yu *et al.*, 1992).

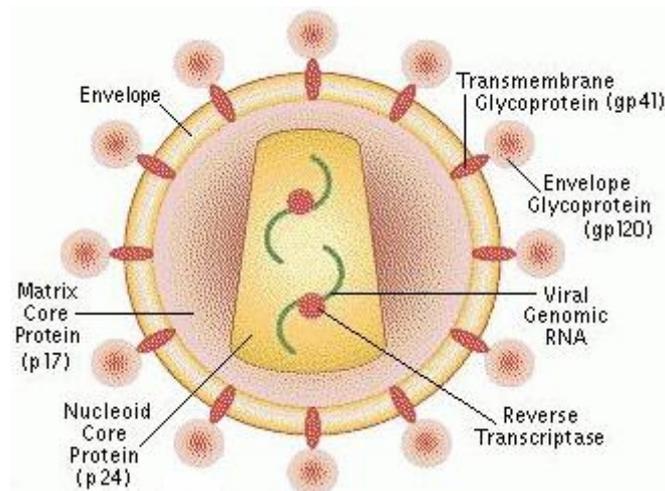


Figura 2. Partícula viral do HIV-1. As setas indicam as estruturas da partícula viral do HIV-1.

1.2.2. ORGANIZAÇÃO GENÔMICA DO HIV-1.

O genoma do HIV consiste em duas fitas simples de RNA com aproximadamente 10 Kb, o qual contém genes estruturais comuns a todos os retrovírus, como *gag*, *pol* e *env*, que codificam enzimas e proteínas participantes diretas da estrutura do vírus ou de seu ciclo viral. Além desses genes, o vírus possui genes regulatórios (*tat* e *rev*) e acessórios (*nef*, *vif*, *vpr* e *vpu*). Nas extremidades 5' e 3' encontram-se longas seqüências terminais repetidas - LTRs (do inglês: *Long Terminal Repeats*) (Arora *et al.*, 2002).

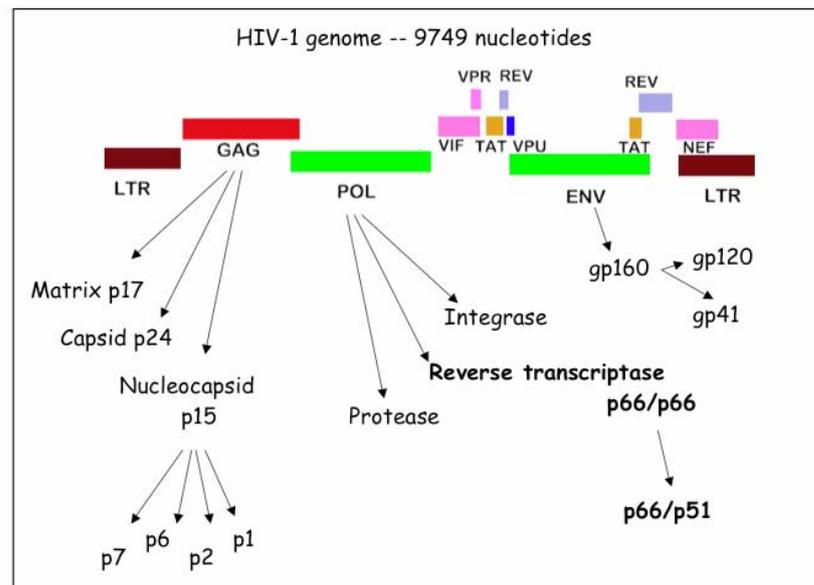


Figura 3. Organização genômica do HIV-1. As setas indicam, de maneira esquemática, as proteínas codificadas pelos genes do HIV-1.

1.2.2.1. GENES ESTRUTURAIIS:

- **Gene *gag*:** É um componente da partícula viral que constitui o capsídeo. O gene codifica uma proteína precursora de 9.2 Kb, a p55, que sofre *splicing*, a partir da ORF (*Open Reading Frame*) localizada na porção inicial do gene *pol*, e é clivada pela protease viral dando origem a quatro proteínas que irão constituir o cerne do HIV: p17 (proteína da matriz), p24 (principal proteína da capsídeo viral), p6 (proteína rica em prolina, envolvida na maturação da partícula viral e auxilia na inclusão da proteína acessória Vpr nos vírions) e p7 (proteína de ligação com o ácido nucléico). As proteínas do gene *gag* executam diversas funções principais durante a montagem viral incluindo: 1. formação do esqueleto estrutural do virion; 2. encapsidação do genoma viral; 3. taxiando a partícula nascente para exportação da célula e 4. aquisição da bicamada lipídica e glicoprotéica associadas ao envelope. Esses processos requerem que as proteínas interajam em processos proteína-proteína, proteína-RNA e proteína-membrana (Wellensiek *et al.*, 2006).

- **Gene *pol*:** Região gênica cuja sobreposição se dá no início do gene *pol* e término do gene *gag*. Codifica três enzimas funcionais necessárias para a replicação viral: 1) protease: responsável pela formação do gene *gag* que ocorre pelo processo de autocatálise do *gag/pol*; 2) transcriptase reversa: possui duas atividades catalíticas que incluem a atividade de DNA polimerase responsável pela polimerização do DNA a partir do RNA e a atividade da ribonuclease H (Rnase H), que degrada o componente RNA das moléculas híbridas RNA-DNA que foram formadas; 3) integrase: responsável pela integração do genoma viral ao genoma da célula hospedeira (Turner & Summers, 1999; Gottlinger, 2001; Hoffman *et al.*, 2002; Hill *et al.*, 2005).
- **Gene *env*:** Codifica uma proteína precursora, a Gp160, que ao ser clivada origina duas glicoproteínas, Gp120 e Gp41. Estudos das seqüências de nucleotídeos da região que codifica a Gp120, mostram que esta região é composta de cinco regiões ou domínios variáveis intercalados por regiões mais conservadas (C1-V1-V2 - C2 - V3 - C3 - V4 - C4 - V5). A Gp120 é a proteína de superfície do HIV, responsável pela interação do vírus com o receptor CD4 e os co-receptores CCR5 e/ou CXCR4, presentes na membrana das células alvo (Deng *et al.*, 1996; Roux, Taylor, 2007; Rits-Volloch *et al.*, 2006). Ao ocorrer a ligação entre a Gp120 e o receptor da célula alvo, as glicoproteínas virais sofrem alterações estruturais expondo a alça V3 da Gp120 e o domínio de fusão da Gp41 (Moore *et al.*, 1991). Esta região é respectivamente alvo de anticorpos neutralizantes e de fármacos anti-retrovirais inibidores de fusão (Nakamura *et al.*, 1993; Pinter *et al.*, 1998; Vancott *et al.*, 1995; Stiegler *et al.*, 2001). Além disso, foi demonstrado que o tropismo viral por co-receptores celulares é influenciado pelas características físico-químicas de aminoácidos na posição 11 e 25 do loop da região V3 do envelope viral (De Jong *et al.*, 1992; Hoffman *et al.*, 2002).

1.2.2.2. GENES REGULATÓRIOS:

- **Gene *tat*:** Codifica a proteína de regulação Tat (do inglês: *trans-acting transcription transactivator*) que é a principal reguladora transcricional do LTR e é essencial para a replicação viral *in vitro* e *in vivo*. A Tat é uma proteína nuclear do HIV-1, que contém 101 resíduos de aminoácidos, codificada por dois exons. O segundo exon extremamente conservado entre os isolados de HIV-1 e o primeiro imbuído das funções de ativação
-

transcricional. por meio da região transativadora responsiva (TAR), localizada a montante do sítio de iniciação de transcrição do RNA, a Tat aumenta a processividade da RNA polimerase II. (Sheridan *et al.*, 1991; Brady, Kashanchi, 2005).

- **Gene *rev***: Codifica a proteína de regulação com 116 resíduos aminoácidos denominada Rev, necessária para a replicação viral (Invernizzi *et al.*, 2006). Essa proteína atua após a transcrição e promove o transporte das moléculas de RNAm das proteínas estruturais e acessórias do núcleo celular para o citoplasma para que ocorra a tradução, controla a proporção entre RNAm processado e não processado através da ligação nos elementos de resposta a Rev (RRE) presente em todos os transcritos virais que ainda não sofreram *splicing* e é responsável pela regulação da síntese de proteínas virais, isto é, regula a proporção de produção de proteínas reguladoras em relação às estruturais (Melo *et al.*, 2006; Goff, 2003; Petelin & Trono, 2003).

1.2.2.3. GENES ACESSÓRIOS:

Os genes acessórios produzem proteínas que podem alterar drasticamente o curso e a severidade da infecção viral, replicação e progressão da doença.

- **Gene *nef***: Codifica a proteína Nef, que contém aproximadamente 29 kDa, expressa durante todo o ciclo de vida do HIV na célula infectada, concomitantemente a Tat e Rev. Ademais, outra característica importante, é a produção de anticorpos e CTL contra este gene, pois os níveis de ambos também são altos durante toda a infecção no paciente.

A proteína Nef está envolvida em diversos mecanismos utilizados na manipulação da célula infectada para auxiliar na replicação viral, no escape das pressões imunológicas, na patogênese e expansão viral. Além disso, Nef, interage com diversos fatores celulares (Das, Jameel, 2005; Arora *et al.*, 2002; Roeth & Collins, 2006), tais como:

1. A redução na expressão de proteínas de superfície celular que fazem parte da resposta imune do hospedeiro (Argañaraz *et al.*, 2003; Das, Jameel, 2005). Nef se liga às moléculas de CD4, CD28, MHC-I e MHC-II promovendo a internalização e transporte destas moléculas para os lisossomos levando a degradação destas, alterando assim a apresentação dos peptídeos virais e promovendo um escape do reconhecimento de células
-

T citotóxicas (Lindwasser *et al.*, 2007; Roeth & Collins, 2006). Estudos demonstraram que Nef utiliza diferentes maneiras de interagir com estas moléculas. Em resumo, Nef promove a formação de complexos de multiproteínas (multiprotein complexes), recrutando proteínas adaptadoras do hospedeiro para promover o transporte intracelular vesicular das moléculas. Alteração do transporte intracelular de outras moléculas tem sido relatada, e sugere-se que Nef gera efeitos pleiotrópicos nos caminhos exocitóticos e endocitóticos que alteram o transporte intracelular de proteínas e pode auxiliando no ciclo de vida do HIV-1 ligação com a superfície interna da membrana plasmática (Roeth & Collins, 2006).

2. Interage com sinalizações celulares através da agregação e incorporação de *lipid rafts* da membrana celular. Nesta interação de Nef com a membrana celular a presença do domínio N-terminal (aa1-61) é necessária. Este domínio possui resíduos altamente conservados de lisina e arginina que contribuem na localização e associação de Nef com a membrana celular. As lisinas localizadas na posição 4 e 7 não são essenciais na agregação de Nef com a membrana, no entanto, contribuem na incorporação de *lipid rafts*, um mecanismo regulatório que determina a atividade biológica das subpopulações distintas de Nef nas células infectadas. A associação de Nef com a membrana é controlada por um padrão de assinaturas (signature motifs) que contribuem nas atividades individuais de Nef. Em resumo, Nef recruta proteínas de sinalização celular para os *lipid rafts* e age como um regulador interferindo na transcrição de diversos genes celulares (Balog & Minarovits, 2006; Giese *et al.* 2006).
 3. Possui atividade catalítica. Nef aumenta a transcrição de genes que codificam enzimas e fatores que catalizam a biossíntese de glicolípídeos e fosfolípídeos de membrana (Ndolo *et al.*, 2006).
 4. Modula a PAK 2 (p21-ativated Kinase 2) na mediação de sinais celulares induzindo a reorganização do citoesqueleto, para facilitar a transcrição reversa ou empacotamento viral (O'Neil *et al.*, 2006);
 5. Liga-se à PI3K, uma enzima regulatória e com atividades catalíticas, que auxilia no processo de redução da expressão de moléculas MHC-I. Nef interage com uma subunidade da PI3K (phosphatidylinositol-3-kinase), p85 (Blagoveshchenskaya *et al.*, 2002; Linnemann *et al.*, 2002). Além de auxiliar no processo de diminuição de expressão de MHC, a PI3K é necessária para ativar a interação de Nef ao PAK (p21-activated
-

kinase), o que sugere um importante papel na ativação de células T e progressão da doença durante a infecção (Agopian *et al.*, 2006).

6. Interage com a p53, inibindo a apoptose (Das & Jameel, 2005).
7. Prejudica o transporte do colesterol celular. Diversas etapas da replicação do HIV-1 dependem de colesterol. A infecção por HIV está associada a alterações no metabolismo lipídico e lipoprotéico. O HIV-1 prejudica ABCA1 (ATP-binding cassette transporter A1) dependente de colesterol, resultando no acúmulo de lipídios nas células infectadas. Esta ação é mediada pela proteína Nef, através da sua interação e transporte do colesterol para o sítio de montagem do virion na membrana plasmática. Sugere-se que este mecanismo assegura o acesso do virions nascentes ao colesterol intracelular, o qual é importante para a infectividade do HIV (Mujawar *et al.*, 2006).

Nef interage com diversas proteínas celulares, que em parte se relacionam com sinalização intracelular. Estudos estruturais desta proteína identificaram alguns domínios envolvidos na interação proteína-proteína e favorece um modelo onde Nef age como um adaptador para acumular complexos de sinalização na célula infectada. Além disso, vírus que apresentam deleção no gene *nef* são menos virulentos e estão associados com a progressão lenta para AID. Estas características promovem a associação de Nef à patogenicidade do HIV (Das & Jameel, 2005).

- **Gene *vif*:** O gene *vif* (Fator de infectividade viral) codifica uma proteína de 23 KDa (Vif), que é produzida tardiamente durante a replicação viral. Além disso, esse gene interage com a proteína celular APOBEC-3G, impedindo que a mesma seja encapsulada junto com as novas partículas virais e ocasione destruição de parte do genoma viral via hipermutação (Goff, 2003).
 - **Gene *vpu*:** Codifica uma proteína de 16KDa (Vpu). Esta proteína aumenta a taxa de liberação do vírus na célula infectada, facilitando a maturação e montagem de novos vírus e auxilia na degradação das moléculas de CD4 recém-sintetizadas, antes que estas sejam transportadas para a membrana citoplasmática (Bour & Strebel, 2003).
 - **Gene *vpr*:** Codifica a proteína Vpr, de 15KDa, que é traduzida a partir de um RNAm que sofreu um único evento de *splicing*. É encapsulada com o vírion, participando no transporte do complexo de pré-integração do genoma viral para o núcleo da célula infectada. Participa no processo de parada do ciclo celular na fase G2 do ciclo, beneficiando o vírus por garantir
-

maior taxa de expressão protéica. Está envolvida nos processos de apoptose, localização nuclear do complexo de pré-integração, ativação do LTR e de outros promotores heterólogos (Tungaturthi *et al.*, 2003).

1.2.3. MECANISMO DE INVASÃO E CICLO REPLICATIVO DO HIV-1.

A infecção pelo HIV-1 se inicia quando ocorre a fusão do envelope viral com a membrana celular e liberação de um complexo contendo proteínas do capsídeo viral, RNA e outras proteínas contidas nesse complexo. No processo de fusão ocorre a interação entre o complexo glicoprotéico do envelope viral (Gp120 e Gp41) com a molécula do antígeno de diferenciação celular (CD4) da membrana celular. O CD4 é uma proteína da superfamília das imunoglobinas expressa na superfície de linfócitos T (preferencialmente), macrófagos de sangue periférico e tecidos, células da microglia (Sistema Nervoso Central), células dentríticas nas mucosas (figura 4).

A interação entre as moléculas Gp120 e CD4 promove mudanças estruturais onde ocorre a protrusão do loop V3, este, por sua vez, vai interagir diretamente com os coreceptores (CCR5, CCR2 e CXCR4) para internalização (Huang *et al.*, 2005). O tropismo celular é definido pelas modificações de carga no V3, somente nos vírus subtipo B. Ademais, o tropismo é definido pelo vírus, que vai reconhecer CXCR4 ou CCR5 (Markosyan *et al.*, 2003). Uma vez concluída a ligação Gp120/CD4/co-receptor, a proteína Gp41 dissocia-se da Gp120 (um trímero de um dímero), que sofre modificações na conformação resultando na sua extensão e exposição de seu resíduo peptídico N-terminal hidrofóbico de fusão que é então inserido na membrana celular, concluindo a fusão vírus/célula, permitindo assim, que o capsídeo viral seja liberada no interior do citoplasma celular (Melikyan *et al.*, 2000; Gómez & Hope, 2005; Henderson & Hope, 2006).

Após a fusão ocorre a liberação da capsídeo viral para o citoplasma celular e inicia-se a transcrição do RNA em DNA complementar (cDNA) sob dependência da enzima transcriptase reversa. A transcriptase reversa promove a síntese de uma cópia de DNA de fita dupla e catalisa as reações de polimerização de DNA dependente de RNA e de DNA, além de degradar a porção de RNA do híbrido RNA-DNA formado durante o processo (atividade da ribonuclease H). Esta etapa da replicação é considerada um momento crítico para a diversidade genética do HIV-1. Devido ao alto potencial mutagênico da transcrição, uma vez acontecendo o erro da enzima

transcriptase reversa, este é incorporado ao genoma viral. Desta forma, um DNA diferente do molde original poderá ser formado (Frankel *et al.*, 1998; Peçanha *et al.*, 2002).

Este DNA complementar é levado na forma de um complexo protéico até o núcleo celular em um processo mediado pela proteína Vpr, esse transporte ativo ocorre mesmo quando a carioteca está intacta. Uma vez no núcleo o DNA viral pode permanecer em forma circular, isoladamente ou pode haver integração no genoma celular (provírus), dependente da enzima integrase, que integra de forma estável o DNA do genoma viral no DNA cromossômico através de várias etapas como endonuclease, transferência de fita, desintegração e ligação, estabelecendo um provírus e completando desta forma a fase pré-integrativa (Cohen, 1996; Peçanha *et al.*, 2002).

Após a integração, a replicação do provírus é feita por enzimas celulares, contudo, a expressão dos genes provirais é dependente do estado de ativação celular. Nas infecções humanas por retrovírus, como no caso do HIV-1, a expressão dos provírus integrados é medida por proteínas virais e celulares. Os RNAs virais transcritos são expressos a partir de um promotor localizado na região 5' do LTR e a proteína acessória Tat aumenta consideravelmente a velocidade de transcrição. O conjunto de RNAs transcritos, de tamanho genômico ou divididos, é transportado para o citoplasma, onde ocorrerá a tradução (Hope, 1997; Frankel *et al.*, 1998; Peçanha *et al.*, 2002).

O processo de tradução e formação de novas partículas virais é regulado pela enzima Rev. Este processo produz as proteínas Vif, Vpr, Nef, além das poliproteínas Gag e Gag-Pol que serão processadas posteriormente em matriz, capsídeo, nucleocapsídeo, p6, protease, integrase e transcriptase reversa. As proteínas Gp41 e Gp120, no entanto, são formadas a partir da poliproteína Gp160, que é co-expressa com os receptores CD4 no retículo endoplasmático. Nesse momento, a proteína Vpu degrada os receptores CD4 recém sintetizados assim que as proteínas do Env sejam transportadas até a membrana celular. A porção p17 da Gag promove a agregação da Gag e da Gag-Pol no citoplasma e a p7 (nucleocapsídeo) liga-se e carrega o RNA, iniciando desta forma a montagem do vírus (Gonda *et al.*, 1985; Veronese *et al.*, 1988; Peçanha *et al.*, 2002).

O processamento, ou montagem dos vírions ocorre inicialmente próximo à membrana celular. As partículas nascentes (imaturas) são compostas de envelope glicoprotéico, RNA genômico e poliproteínas virais que passam por uma modificação morfológica (maturação) durante ou após o “brotamento”. O evento da maturação ocorre com a clivagem das poliproteínas

Gag e Gag-Pol pela protease viral produzindo enzimas e proteínas estruturais do capsídeo. A maturação da partícula viral ocorre no estágio avançado do ciclo replicativo e requer altas quantidades de substrato, provavelmente devido à baixa atividade enzimática da protease não processada; este estado é alcançado somente durante o “brotamento”. Após uma hidrólise inicial, a protease liberada pode então exercer sua atividade máxima, de acordo com os processos supracitados, processando de forma rápida as poliproteínas, completando assim o ciclo replicativo do HIV-1. Uma vez maduros e livres os novos vírions são potencialmente capazes de infectar novos linfócitos adjacentes (Vaishnan *et al.*, 1991; Meek, 1992; Peçanha *et al.*, 2002).

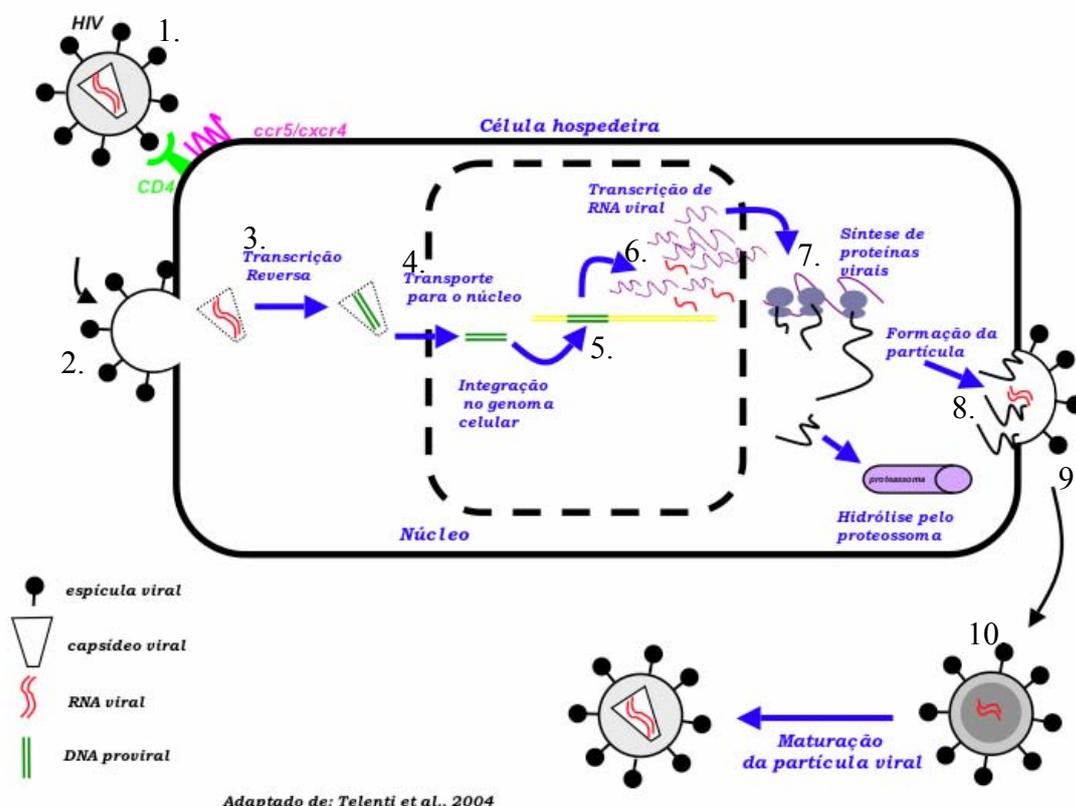


Figura 4. Ciclo replicativo do HIV-1. 1.) Adsorção da partícula viral por meio de glicoproteínas de superfície do envelope viral (espícula viral, representado por uma estrutura preta); 2.) fusão e liberação do capsídeo no citoplasma celular; 3.) transcrição reversa; 4. formação do complexo DNA viral/integrase e transporte do DNA viral para o núcleo; 5.) integração do DNA viral no genoma celular; 6.) transcrição de RNA viral; 7.) síntese de proteínas virais; 8.) formação da partícula; 9.) montagem e liberação da partícula viral; 10. maturação da partícula viral.

1.3. DIVERSIDADE VIRAL.

1.3.1. ORIGEM DA DIVERSIDADE.

O HIV possui elevada diversidade genética resultado de sua alta taxa de mutação e adaptação ao hospedeiro (Holmes *et al.*, 1992; Leigh Brown, 1997). Esta diversidade ocorre devido a uma combinação de fatores (tais como altas taxas de recombinação, elevado número de partículas virais produzidos durante o ciclo replicativo) entre eles a peculiaridade da enzima transcriptase reversa não possuir a propriedade de correção durante o processo de replicação viral. A ausência de correção resulta em uma taxa de erro na incorporação de nucleotídeos de aproximadamente 10^{-3} a 10^{-5} por sítio, por ciclo de replicação. Considerando-se que, a cada dia, são produzidas em média 10^{10} novas partículas virais, é possível estimar a diversidade potencialmente gerada durante o longo curso da infecção. Esta característica de formação de formas virais variantes (formação de quase-espécies) pode aumentar o potencial para o sucesso da adaptação do hospedeiro do HIV-1 (Lemey *et al.*, 2006; Mullins & Jensen, 2006; Fares & Travers, 2006; Pinto & Struchiner, 2006; Frost *et al.*, 2005). Além disso, a geração da diversidade também é impulsionada pelas pressões seletivas exercidas pelo sistema imunológico, ou seja, anticorpos e células citotóxicas (Jung *et al.*, 2002; Bazykin *et al.*, 2006). A combinação destes fatores contribui para que em cada 25 gerações surja uma substituição no envelope (*env*) do HIV-1, que confere vantagem adaptativa, ou seja, a cada 2,5 meses uma nova mutação é fixada por seleção (Willianson, 2003). As implicações dessa diversificação constituem um dos maiores obstáculos para a implementação de estratégias vacinais (Gaschen *et al.*, 2002; Rambaut *et al.*, 2004), além de constituírem um fator muito importante na geração de variantes de resistência aos anti-retrovirais (Ribeiro & Bonhoeffer, 2000; Charpentier *et al.*, 2006).

1.3.2 ESTUDO DOS SUBTIPOS DO HIV-1.

A partir da tentativa de organizar a diversidade observada entre as cepas de HIV-1 que circulam na epidemia, até 1992, as variantes do vírus foram identificadas como "americanas" ou "africanas", com base em seu local de origem. Com a descoberta de novos espécimes, esta separação tornou-se inadequada e foi proposta uma classificação em subtipos com base na análise dos genes *env* ou *gag* (Robertson *et al.*, 2000).

A análise filogenética de fragmentos do genoma do vírus possibilitou a subdivisão do HIV-1 em três grupos: M (*main* ou *major*), O (*outlier*) e N (*new*), que podem ter origem zoonótica em eventos independentes (Hahn *et al.*, 1984; Srinivasan *et al.*, 1987; Myers, 1992). O grupo M é predominante na epidemia global de AIDS e engloba a maioria dos HIV-1. Este grupo é representado por linhagens geneticamente distintas (subtipos), mas que são agrupadas filogeneticamente. Isso permite que o grupo M possa ser classificado em diversos subtipos (A1, A2, B, C, D, F1, F2, G, H, J e K) distantes entre si entre 10 e 30%, dependendo da região gênica analisada (Myers *et al.*, 1997).

A distribuição geográfica e o padrão de disseminação dos diferentes subtipos do HIV-1 são bastante heterogêneos. O subtipo B é o principal responsável pela disseminação da AIDS fora da África (<http://www.unaids.org>). Entretanto, o subtipo C é o mais freqüente sendo o principal responsável causador da epidemia na Índia e na África (Novitsky *et al.*, 1999; Shankarappa *et al.*, 2001). Estima-se que aproximadamente 39.5 milhões de pessoas estão infectadas pelo HIV no mundo e que a maioria dos casos (aproximadamente 24.7 milhões de pessoas) concentram-se nos países da região subsaariana. (<http://www.unaids.org>). Notavelmente, na África ocorrem todos subtipos já descritos além de inúmeras formas recombinantes. Além disso, a diversidade genética do subtipo C no Congo é tão grande quanto à diversidade global estimada para o subtipo B (Rambaut *et al.*, 2001). No Brasil, estima-se que aproximadamente 600.000 pessoas estejam infectadas pelo HIV-1, sendo que existem aproximadamente 240.000 casos registrados e que só no estado de São Paulo estão localizados 108.000 dos casos (<http://www.aids.gov.br>). Os subtipos comumente encontrados no Brasil são B, C e F1 sendo que o subtipo B é o de maior prevalência, cerca de 75% das infecções, o subtipo F (aproximadamente 10%) e o subtipo C com uma prevalência maior nos estados do sul e sudeste, predominantemente Rio Grande do Sul, com cerca de 30% das infecções. Além disso,

são encontradas formas recombinantes entre os subtipos B e F (Brindeiro *et al.*, 2003; Louwagie *et al.*, 1994). Na Argentina, a distribuição de subtipos é bastante peculiar pelo fato que entre 40 a 65% dos indivíduos infectados deve-se à ocorrência de infecções por formas recombinantes BF (CRF12-B/F) e pela quase ausência de infecção pelo subtipo F (Thomson *et al.*, 2000). Proporções tão altas de infecção por CRFs também são observadas na África (<http://www.bioafrica.net>) e no sul da Ásia (McCutchan *et al.*, 2000). No Brasil, a CRF_BC representa 11% dos vírus circulantes na região sul do país (Brindeiro *et al.*, 2003; Soares *et al.*, 2003; Santos *et al.*, 2006). Entretanto, outras formas recombinantes circulantes têm sido descritas no Brasil, como, CRF28_BF e CRF29_BF, que são formadas exclusivamente pelos subtipos B e F e podem desempenhar um importante papel na epidemia brasileira e da América do Sul (Sanabani *et al.*, 2006; De Sa Filho *et al.*, 2006). Ademais, foi relatada a primeira ocorrência de um recombinante africano, CRF_02/AG, no Rio de Janeiro (Couto-Fernandez *et al.*, 2005). Diversas formas recombinantes foram relatadas no mundo e estão descritas na tabela que segue (Tabela 1).

Tabela 1. Descrição das formas recombinantes circulantes (CRFs) (Leitner *et al.*, 2005).

CRF	Mix	Seqüência	Acc . N°	Ano	Origem da Amostra	Referência
01	AE	CM240	U54771	1990	Thailand	Carr, J.K. <i>J Virol</i> 70 :5935–43 (1996)
02	AG	IBNG	L39106	NA	Nigeria	Howard, T.M. <i>ARHR</i> 10 :1755–57 (1994)
03	AB	KAL153	AF193276	NA	Russia	Salminen, M.O., <i>Unpublished</i> (1998)
04	AGHKU	CY032.3	AF049337	1994	Cyprus (Greece)	Gao, F. <i>J Virol</i> 72 :10234–41 (1998)
05	DF	VI310	AF193253	NA	Belgium (?DRC)	Laukkanen, T. <i>Virology</i> 269 :95-104 (2000)
06	AGJK	BFP90	AF064699	1996	Australia (Burkina Faso)	Oelrichs, R.B. <i>ARHR</i> 14 :1495–500 (1998)
07	BC	CN54	AX149771	1997	China	Guan, Q. <i>J. Clin. Microbiol.</i> 42 :4261–7 (2004)
08	BC	GX-6F	AY008715	1997	China	Piyasinilp, S. <i>J Virol</i> 74 :11286–95 (2000)
09		96GH2911	AY093605	1996	Ghana	McCutchan, F.E. <i>ARHR</i> 20 :819–26 (2004)
10	CD	TZBF061	AF289548	1996	Tanzania	Kouliniska, I.N. <i>ARHR</i> 17 :423–31 (2001)
11	A01GJ	GR17	AF179368	NA	Greece (?DRC)	Paraskevis, D. <i>ARHR</i> 16 :845–55 (2000)
12	BF	ARMA159	AF385936	1991	Argentina	Carr, J.K. <i>AIDS</i> 15 :F41–7 (2001)
13	A01GJU	96CM-1849	AF460972	1996	Cameroon	Wilbe, K. <i>ARHR</i> 18 :849–56 (2002)
14	BG	X397	AF423756	1999	Spain	Delgado, E. <i>JAIDS</i> 29 :536–43 (2002)
15	01B	99TH.MU2079	AF516184	1999	Thailand	Viputtigul, K. <i>ARHR</i> 18 :1235–7 (2002)
16	A2D	KISII5009	AF457060	NA	Kenya	Dowling, W.E. <i>AIDS</i> 16 :1809–20 (2002)
17	BF1	BF	PSP0096			Jean Carr unpublished
18		CU14	AY586540	NA	Cuba	Thomson, M. <i>AIDS</i> 19 :1155–63 (2005)
19		CU38	AY588970	1999	Cuba	Cuevas, M.T. <i>AIDS</i> 16 :1643–53 (2005)
20	BG					Michael Thomson unpublished
21	A2D					Francine McCutchan unpublished
22	01A1	3097MN				Jean Carr unpublished
23	01A1					Jean Carr unpublished
24	BG					Michael Thomson unpublished
25	BG					Michael Thomson unpublished
26	AG					Jean Carr unpublished
27						Martine Peeters unpublished
28						Martine Peeters unpublished
29	BF	UFRJ1				Janini unpublished
30	BF	BREMP12313				Janini unpublished
31	0206					Martine Peeters unpublished
32	02A1					Jean Carr unpublished

1.3.3. RESPOSTA IMUNE À INFECÇÃO PELO HIV-1.

A infecção por HIV é caracterizada por uma viremia crônica que pode resultar na progressiva deterioração da função do sistema imune, progredir para a AIDS e, posteriormente, para o óbito. Na maioria dos adultos e em alguns recém-nascidos evidencia uma resposta imunológica humoral e celular intensa, relevante para o controle da infecção pelo HIV (Geffin *et al.*, 2003). Contudo, a dinâmica viral na infecção por HIV-1 em crianças é distinta quando comparada a dos adultos. Em contraste com a resposta imunológica dos adultos, nas crianças, a resposta mediada por células T CD8 é menor e rara antes dos 6 meses de idade (Chandwani *et al.*, 2004). No entanto, há controvérsias, pois existe relato de que a resposta imunológica mediada por CTL, respondem tão intensamente quanto adultos (Feeney *et al.*, 2003). Ademais, estudos demonstraram que crianças com idade inferior a 4 anos possuem uma resposta imunológica menor quando comparada às crianças mais velhas, portanto, a resposta imune contra o HIV-1 pode ser influenciada pela idade (Sandberg *et al.*, 2003).

A infecção aguda por HIV-1 em adultos é caracterizada por um alto pico de carga viral que declina rapidamente durante aproximadamente 3 a 4 meses após a infecção até alcançar a estabilidade (Sandberg *et al.*, 2003). Este declínio na viremia coincide com a emergência da resposta por meio de células T citotóxicas (Mann, 1999; Musey *et al.*, 1997; Phillips, 1999). Em contraste, a carga viral das crianças é muito alta e o declínio da viremia é lento. Esta elevada carga viral encontrada em crianças recém-nascidas pode estar relacionada com a alta produção do timo, uma vez que crianças possuem um elevado número de células T CD4 circulantes e porque a resposta de CTL ocorre somente após 6 meses de idade. Além disso, estudos demonstraram que a sobrevivência das crianças infectadas é menor do que a dos adultos infectados, o que pode ser explicado por alguns fatores incluindo: imaturidade do sistema imune, falha na aquisição da imunidade materna, momento da transmissão ou carga viral materna, taxa de replicação viral, indução de sincício, diminuição de células T CD4 e a infecção no timo (Zang *et al.*, 2006).

Na infecção primária demonstrou-se a presença de mutações nos epítopos de célula T citotóxica (CTL) no HIV-1, que podem ser maior e/ou mais comuns em determinados genes como *gag*, *nef* e *tat* (Frost *et al.*, 2005). Estas mutações ocorrem devido a uma intensa pressão seletiva na população viral por meio de CTLs, o que promove uma rápida emergência de

variantes que escapam do reconhecimento destas células, acarretando numa viremia persistente pós soroconversão, além de importante papel na patogênese viral. Em contraste, a pressão seletiva exercida pela resposta dos anticorpos neutralizantes pode ser branda, podendo afetar o *fitness* da população viral e contribuir para a rápida evolução dos vírus, entretanto, altos níveis de anticorpos estão correlacionados com progressão lenta da doença. Os anticorpos neutralizantes podem exercer uma pressão seletiva e contribuir na origem da diversidade viral, esta diversidade pode ser observada na infecção recente (Frost *et al.*, 2005; Phillay & Phillips, 2005; O'Connor *et al.*, 2006). Desta forma, a resposta imune do hospedeiro influencia a seleção de formas virais capazes de escapar do sistema imune, promovendo assim, a evolução viral. (Price *et al.*, 1997; Phillay & Phillips, 2005).

Uma potente resposta celular (células T citotóxicas circulantes) e humoral foram demonstradas na infecção primária, esta resposta está associada à manutenção da carga viral em níveis baixos durante a infecção aguda e crônica e aos indivíduos que progridem lentamente para a doença (Geffin *et al.*, 2003). Ademais, alguns estudos sugerem que a progressão lenta para AIDS também esteja associada à presença de resposta citotóxica direcionada aos diferentes epitopos do gene *gag* (Klein *et al.*, 1995; Riviere *et al.*, 1995; Betts *et al.*, 1999). Por outro lado, estudos realizados *in vitro*, evidenciaram que a resposta de CTL está direcionada principalmente contra poucos epitopos dos genes *env* e *pol*, sugerindo que a progressão tende a ser rápida (Goulder *et al.*, 1997; Hay *et al.*, 1999; Zhang *et al.*, 1994). Alguns códonos do gene *nef* também são considerados importantes alvos da seleção positiva, além disso, estudos demonstraram que indivíduos que progridem lentamente para a AIDS, estão associados com deleções no gene *nef*, dando suporte ao papel deste gene com a patogênese viral e com a progressão da doença (Zanotto *et al.*, 1999; O'Neill *et al.*, 2006; Yang *et al.*, 2006). Dessa forma, não somente a presença de atividade CTL, bem como a especificidade, duração e abrangência da resposta contra os diferentes tipos de epitopos, são de fundamental importância na manutenção da replicação do HIV em níveis baixos.

Entretanto, apesar de o sistema imunológico ser capaz de reconhecer várias proteínas do HIV e apresentar uma resposta celular e humoral vigorosa, principalmente durante a fase assintomática da infecção, o vírus não é completamente controlado e a progressão da doença é inevitável (Brander *et al.*, 1999). A habilidade de o vírus persistir, mesmo com a presença da resposta imune celular e humoral, é atribuída às constantes mutações e a geração de novas variantes virais que são resistentes à resposta imune do hospedeiro (Geffin *et al.*, 2003). Estudos

demonstraram que a transmissão materna de variantes virais pode limitar a resposta aos epítopos alvo, quando o padrão haplotípico é idêntico na criança (Feeny *et al.*, 2003).

Diversos fatores têm sido descritos que ajudam a explicar os mecanismos de resistência ao desenvolvimento e progressão para a AIDS. Em alguns casos, foi observado que a resistência pode ser resultante da infecção por vírus HIV-1 com mutações em um ou mais genes (Deacon *et al.*, 1995). Em outros casos, existe a associação entre mutações em genes de receptores celulares de quimiocinas e a resistência à infecção pelo HIV-1 (Smith *et al.*, 1997). Também foi observada uma correlação entre a heterogeneidade do complexo de histocompatibilidade principal (MHC) e a progressão para AIDS em indivíduos infectados pelo HIV-1 (Carrington *et al.*, 1999). Contudo, algumas estratégias utilizadas pelo HIV tornam a resposta imune pouco eficaz no controle da replicação viral e da infecção, assim como:

- Interferência do vírus na apresentação e reconhecimento das proteínas virais. Sob esse aspecto, observou-se que uma das atividades da proteína Nef é induzir a redução específica da expressão de algumas classes de MHC na superfície de células infectadas pelo HIV, interferindo na apresentação normal de peptídeos virais (Collins *et al.*, 1998). Ainda nesse sentido, observou-se a ocorrência de variantes do HIV (mutantes) capazes de escapar do reconhecimento feito pelo sistema imunológico (Goulder *et al.*, 1997).
 - Colonização rápida de diferentes tipos de tecidos. Durante a infecção pela mucosa, o HIV usa as células apresentadoras de antígenos (macrófagos e células de Langerhans). Estas células infectadas migram para os tecidos linfóides para processamento e apresentação de antígenos, onde o HIV passa a replicar e a colonizar rapidamente outros tecidos (Kacani *et al.*, 1998). A colonização rápida de diferentes tipos de tecidos pode resultar no surgimento de variantes virais do HIV, que influenciam na progressão para AIDS (Shankarappa *et al.*, 1999).
 - Infecção de células latentes. Concomitante à colonização dos tecidos linfóides, o HIV pode infectar linfócitos T CD4⁺ inativos, permanecendo nestas células na forma de provírus latente. Os linfócitos com HIV latente constituem um reservatório viral, visto que não ocorre a replicação do provírus enquanto as células permanecem inativas e conseqüentemente não ocorre apresentação de antígenos e indução de resposta imune (Finzi *et al.*, 1999).
 - Depleção de alguns tipos específicos de células T. O desenvolvimento de uma resposta imune intensa depende da integridade dos diferentes tipos de células envolvidas na indução e
-

manutenção da resposta imune. Na infecção pelo HIV, ocorre uma depleção dos linfócitos T auxiliares (Th), o que causa uma diminuição acentuada na manutenção da resposta imune (Brander *et al.*, 1999; Kalams *et al.*, 1999; Rosenberg *et al.*, 1999). Além do comprometimento da manutenção da resposta imune, durante a infecção pode ocorrer eliminação de determinados clones de células CTL específicos para alguns variantes do HIV (McKinney *et al.*, 1999).

Todas essas estratégias combinadas possibilitam que o HIV cause um comprometimento gradual e aparentemente irreversível do sistema imunológico (Chun *et al.*, 1999).

Em resposta à pressão exercida pelo sistema imunológico ou pelos diferentes regimes anti-retrovirais, ocorre seleção de mutações que proporcionam algum tipo de vantagem adaptativa ao vírus (Mammano *et al.*, 1998; Kosalaraksa *et al.*, 1999). Esse fenômeno foi extensivamente demonstrado em pacientes infectados pelo HIV, submetidos a tratamento anti-retroviral (Molla *et al.*, 1996). O vírus responde à pressão exercida pelas drogas fixando mutações que conferem resistência aos compostos utilizados (Sheldon *et al.*, 1999).

Por outro lado, os efeitos resultantes da pressão exercida pelo sistema imunológico no genoma do HIV são menos óbvios. Observou-se que a presença de mutações em algumas regiões dos genes do HIV, as quais correspondem a epitopos reconhecidos pelas células T citotóxicas, possibilitam a evasão desses mutantes à vigilância feita pelo sistema imunológico. A presença de mutantes do HIV não reconhecidos pelo sistema imunológico foi observada nas diferentes fases da infecção e está associada com a rapidez na progressão para a AIDS (Phillips, *et al.*, 1991; Borrow *et al.*, 1997; Goulder *et al.*, 1997; Price *et al.*, 1997; Phillay & Phillips, 2005). No entanto, alguns trabalhos sugerem que a presença de mutantes aptos a evadir a vigilância feita pelo sistema imunológico seja um fenômeno importante e comum durante todas as fases da infecção pelo HIV (Brander *et al.*, 1999; Hay *et al.*, 1999).

Possivelmente, o padrão de evolução e a diversidade que são observados no genoma do HIV possam ser consequência da combinação de diferentes fatores como: seleção natural, mutação, deriva gênica e colonização de diferentes tecidos (Rambaut *et al.*, 2004). Além disso, estudos filogenéticos sugerem que estas variantes virais também possam ser transmitidas, entretanto, não existem dados conclusivos de que estas variantes, com adaptações às respostas imunes celulares, sejam preferencialmente transmitidas (Pillay & Phillips, 2005). Em adição, assinaturas genômicas decorrentes da pressão seletiva do sistema imunológico têm sido descritas

em diversos vírus, no *Epstein-Barr* (EBV) (Campos-Lima *et al.*, 1993), no vírus da *Influenza A* (Lindsay *et al.*, 2001) e mesmo no HIV-1 (Kelleher *et al.*, 2002; Moore *et al.*, 2002). Além disso, padrões distintos de seleção natural foram estabelecidos em seqüências de DNA viral e correlacionados com a pressão atuante do sistema imunológico dos hospedeiros em diferentes genes virais (Leal *et al.*, 2004; Ross & Rodrigo, 2002; Sheridan *et al.*, 2004). Os mecanismos evolutivos que atuam nos genes do HIV-1, decorrentes das interações com o sistema imunológico, têm sido estudados extensivamente (Kelleher *et al.*, 2002; Leal *et al.*, 2004; Moore *et al.*, 2002; Phillips *et al.*, 1991; Price *et al.*, 1997; Ross & Rodrigo, 2002; Sheridan *et al.*, 2004).

Entretanto, em pacientes com o sistema imunológico ainda em desenvolvimento, o perfil evolutivo do HIV-1 e a extensão da resposta precisam ser mais explorados. Neste trabalho propomos a exploração de alguns aspectos da genética evolutiva do HIV, através de uma abordagem comparativa entre pacientes com o sistema imunológico imaturo (crianças) e pacientes com sistema imunológico hígido (adultos recém infectados). Especificamente, analisamos o impacto causado pela infecção do HIV na habilidade do organismo responder aos epitopos virais. Isso foi feito de forma comparativa avaliando a diversidade genética e a extensão da pressão seletiva no gene *nef* em crianças e adultos recém infectados.

1.3.4. DINÂMICA POPULACIONAL E INTERAÇÃO VÍRUS-HOSPEDEIRO.

A dinâmica da população inferida por filogenia do HIV-1 no paciente é muito diferente daquela ao nível populacional (Rambaut *et al.*, 2004; Grenfell *et al.*, 2004). Isso é bastante evidente em filogenias inferidas de genes altamente imunogênicos, como por exemplo, o gene *env* do HIV (Grenfell *et al.*, 2004; Yusim *et al.*, 2002). Em contrapartida, o gene *pol* é menos variável com uma taxa de substituição menor que a do gene *env* (Palmer *et al.*, 2002). Portanto, as mutações de resistência surgidas em um hospedeiro provavelmente são menos transmitidas (Leigh Brown, 1999; Ribeiro & Bonhoeffer, 2002; Yusim *et al.*, 2002). Além disso, o gene *pol* é pouco imunogênico, de forma que a pressão seletiva do sistema imunológico tende a ser menos intensa (Phillips, 1999). Talvez por isso, a distância genética do gene *pol*, estimada dentro do indivíduo é similar à estimada na população (entre indivíduos não correlacionados), com isso, a estrutura filogenética no paciente é similar à observada na população (Palmer *et al.*, 2002;

Grenfell *et al.*, 2004). Em razão destes fenômenos e do predomínio de seleção purificadora no gene *pol*, o surgimento e extinção de linhagens são menos frequentes (Ribeiro & Bonhoeffer, 2002; Rouzine & Coffin 1999). No entanto, as altas taxas de recombinação observadas no HIV podem favorecer o surgimento de populações virais com perfil de mutações de resistência complexo (Morris *et al.*, 1999; Rambaut *et al.*, 2004).

O processo de substituição alélica (fixação de mutações) é dinâmico e sofre a influência de mecanismos determinísticos (não influenciados por efeitos do acaso) e estocásticos (influenciados por efeitos do acaso) (Rouzine & Coffin, 1999). Particularmente, nos eventos de transmissão do vírus entre indivíduos, ocorre uma redução bastante acentuada na população viral que é transmitida para o novo indivíduo infectado. Com isso, os eventos de substituição são altamente afetados por mecanismos influenciados por efeitos do acaso (Haase *et al.*, 1996; Leigh Brown *et al.*, 1997; Perelson *et al.*, 1996; Rouzine & Coffin, 1999). Por outro lado, mesmo nestas circunstâncias existe a pressão de seleção natural, principalmente no envelope viral (em especial, nas glicoproteínas de superfície (Gp120)), resultante do processo de adaptação ao ambiente celular do novo hospedeiro (Shankarappa *et al.*, 1999).

O processo de co-evolução decorrente das interações entre vírus e hospedeiro pode ser evidenciado em diversos sistemas de infecção natural, no entanto, a interpretação desses achados envolve considerações acerca dos diversos fatores associados a este mecanismo. Este processo resulta de uma combinação bastante complexa de fatores inerentes ao vírus e à população infectada, que pode ser analisado através das variações genéticas adaptativas surgidas no HIV-1 decorrente de fatores seletivos (pressão do sistema imunológico ou terapia anti-retroviral) durante a interação do parasita com seu hospedeiro e a da manutenção nas populações. Além disso, a virulência do HIV pode ser influenciada por esta adaptação molecular e contínua seleção ao longo do tempo (Moore *et al.*, 2002; Yusim *et al.*, 2002; Anderson & May, 1979; Woolhouse *et al.*, 2002; Arien *et al.*, 2005).

A biologia evolutiva tem abordado os mecanismos genéticos envolvidos na interação entre parasita e hospedeiro sob diferentes aspectos. Estes mecanismos de co-evolução têm sido estudados destacando o padrão de mutações em genes virais e do hospedeiro que influenciam na susceptibilidade à infecção pelo HIV. Nesse contexto, foi observado que existe uma relação entre a baixa susceptibilidade à infecção causada pelo HIV e uma deleção no receptor de quimiocinas CCR5 (Dean *et al.*, 2002; Schlieckelman *et al.*, 2001). Além disso, alguns autores sugeriram que existe uma relação entre a presença de determinados haplótipos no complexo de

histocompatibilidade principal (MHC) com a gravidade da infecção (Gao *et al.*, 2001; Carrington *et al.*, 1999). Outro aspecto da co-evolução é a possibilidade de que determinados padrões polimórficos do vírus favoreçam a sua disseminação numa determinada população. Evidências desse tipo de escape têm sido descritas na população viral existente no paciente (Kelleher *et al.*, 2001) e mesmo em populações de hospedeiros que compartilham um mesmo padrão haplotípico (Campos-Lima *et al.*, 1993).

A pressão que o sistema imunológico exerce no HIV deve-se principalmente ao reconhecimento de epítomos virais, os quais são apresentados na superfície de células infectadas. Estas células apresentam os peptídeos resultantes da hidrólise de proteínas virais feita pelo complexo proteassoma no citoplasma celular (Janeway *et al.*, 1997). Os epítomos apresentados na superfície das células infectadas ligam-se ao MHC e este complexo, MHC-peptídeo, é então reconhecido pelos receptores de linfócitos T CD8 (TCR), iniciando a resposta imune humoral e celular. O controle da infecção pelo HIV é feito principalmente por linfócitos T CD8 que destroem as células infectadas por meio da atividade citotóxica (CTL) (Janeway *et al.*, 1997). Esta atividade CTL específica é mediada pelo TCR das células CD8 e pelo complexo MHC-peptídeo da superfície das células infectadas (Janeway *et al.*, 1997; Litman *et al.*, 1999) e é extremamente importante no controle da replicação viral durante toda a fase de infecção pelo HIV-1 (Rinaldo *et al.*, 1995; Phillips, 1999; Pantaleo *et al.*, 1997).

Os efeitos da interação do HIV com o sistema imunológico do paciente também podem ser observados a nível populacional. A frequência de determinados haplótipos (MHC) em algumas populações pode induzir um determinado padrão de mutações no HIV-1 (Moore *et al.*, 2002; Yusim *et al.*, 2002). Nos casos em que a população possui uma composição de MHC pouco diversificada, pode ocorrer que o padrão de mutações surgido no parasita possibilite a sua evasão do sistema imunológico (Campos-Lima *et al.*, 1993; Hill, 1998). O surgimento de mutações que favorecem a evasão já foi observado ao longo da infecção no paciente (Kelleher *et al.*, 2002; Price *et al.*, 1997). Por exemplo, determinados padrões de mutações em alguns sítios da RT estão diretamente associados com a presença de certos haplótipos na população infectada (Moore *et al.*, 2002; Yusim *et al.*, 2002). Isso talvez explique o escape viral que ocorre em populações que têm um padrão haplotípico pouco diversificado (Campos-Lima *et al.*, 1993; Carrington *et al.*, 1999).

A dinâmica populacional e a filogenética do HIV são fortemente influenciadas pelas interações do vírus com o hospedeiro (Crandall *et al.*, 1999; Frost *et al.*, 2001; Ross & Rodrigo,

2002). Os eventos de infecção e a malha de transmissão explorada também afetam a dinâmica do HIV (Grenfell *et al.*, 2004; Pybus *et al.*, 2000; Rambaut *et al.*, 2004). Sendo assim, a evolução do HIV resulta de uma combinação de fatores relativos à biologia do vírus (taxa de substituição, recombinação) e fatores relacionados com o hospedeiro (estrutura genética dos indivíduos, malha de transmissão, pressão do sistema imunológico). A pressão evolutiva nos genes do HIV-1 é distinta, de modo que alguns genes são mais imunogênicos e outros menos. Com isso, o acúmulo de diversidade pode variar bastante nos genes do HIV-1 (Nájera *et al.*, 1995; Shankarrapa *et al.*, 1999; Yoshimura *et al.*, 1996). Em consequência disso, o padrão filogenético observado nos diferentes genes presentes no genoma viral é distinto (Grenfell *et al.*, 2004; Rambaut *et al.*, 2004; Shankarrapa *et al.*, 1999). Por exemplo, o gene *env* é altamente imunogênico, de forma que a pressão exercida pelo sistema imunológico ocasiona o surgimento e a extinção de várias linhagens virais durante o período de infecção (Shankarrapa *et al.*, 1999). Portanto, a dinâmica evolutiva do gene *env* do HIV-1 no paciente é muito diferente daquela observada a nível populacional. A estrutura das filogenias feitas com seqüências do gene *env* obtidas num mesmo paciente caracteriza-se pela presença de ramos longos e pouca estrutura de agrupamentos. Contudo, este mesmo padrão não é observado quando a filogenia é construída com seqüências de indivíduos não relacionados, onde existe bastante resolução no padrão de agrupamentos em razão do isolamento e evolução de linhagens independentes (Rambaut *et al.*, 2004). Com isso, as inferências feitas de análises de seqüências obtidas de um paciente não refletem o processo evolutivo da população.

A dinâmica evolutiva do gene *nef* depende da combinação de diversos fatores como: seleção natural, efeitos estocásticos associados à redução da população durante os eventos de transmissão e a malha de transmissão explorada pelo vírus. A complexa interação entre o vírus e o sistema imunológico e a capacidade deste sistema exercer pressão seletiva por meio de CTL e anticorpos neutralizantes contra o vírus pode gerar padrões específicos de evolução. Esta pressão seletiva decorrente da resposta imune do hospedeiro face às infecções ocorre tanto nas crianças quanto nos adultos (Pillay & Phillips, 2005). No entanto alguns estudos mostraram que em crianças com menos de 4 anos a resposta é menos intensa que nos adultos (Pillay & Phillips, 2005; Sandberg *et al.*, 2003). Visto que a pressão seletiva difere de acordo com a idade, neste estudo abordaremos, através da utilização de dados moleculares e de ferramentas de análise filogenética e, também, de genética de populações, os processos evolutivos do HIV-1 em populações de crianças e adultos recém infectados.

Evidências filogenéticas indicam, que grande parte das mutações que surgem nas mães são transmitidas às crianças (transmissão vertical) e que somente com o transcorrer do tempo de infecção é que essas mutações podem desaparecer, deste modo, a transmissão vertical de variantes capazes de evadir da resposta de células T CD8 específica podem comprometer a montagem de uma resposta imunológica, quando crianças compartilharem o mesmo perfil de haplótipo das mães, resultando na redução do controle da replicação viral e rápida progressão para AIDS (Pillay & Phillips, 2005; Sanchez-Merino *et al.*, 2005). Além disso, estudos sugerem que a resposta de CTL nas crianças somente se inicie por volta dos 6 meses de idade (Sandberg *et al.*, 2003), contudo nos adultos esta resposta surge logo no início da infecção. Por outro lado, a resposta por anticorpos neutralizantes pode ser gerada rapidamente e contribuir para a diversidade genética do vírus nos primeiros meses da infecção, tanto nos adultos quanto nas crianças (McAdam, Gotch, 1999; Moore *et al.*, 2002; Nájera *et al.*, 1995; Rouzine & Coffin, 1999a; Yusim *et al.*, 2002; Pillay & Phillips *et al.*, 2005; Lemey *et al.*, 2007).

A resposta imunológica por meio de CTL desempenha um importante papel no controle da replicação viral. Esta resposta pode induzir mutações que permitem a evasão do HIV-1 às pressões seletivas exercidas pela resposta imunológica. Além disso, estas mutações de evasão, descritas tanto na fase aguda quanto na crônica, foram evidenciadas nos epítopos dos genes *gag*, *nef* e *pol* do HIV-1, no entanto, foi demonstrado que o gene *nef* é um importante alvo da resposta imunológica na infecção primária (Fenney *et al.*, 2005; Karlsson *et al.*, 2007).

Na fase inicial da infecção, observou-se que a resposta imunológica mediada por células T citotóxicas (CTL) contra os peptídeos do gene *nef* é igualmente intensa (Novitsky *et al.*, 2001; Yang *et al.*, 2006). Dessa maneira, a intensa pressão seletiva no gene *nef*, pode estar correlacionada com a extensão e duração da infecção no hospedeiro (Feeney *et al.*, 2005; Noviello *et al.*, 2007).

Conseqüentemente, a análise comparativa entre crianças e adultos recém infectados é relevante sob o aspecto que, contrário aos adultos, crianças apresentam um sistema imunológico imaturo. Neste cenário o HIV encontra pouca resistência imposta pelos mecanismos celulares de defesa, com isso o estabelecimento da infecção pode ter uma dinâmica distinta. Essas diferenças têm sido caracterizadas em crianças através de estudos clínicos (níveis de linfócitos e carga viral) e imunológicos (resposta CTL a determinados epítopos virais) e, além disso, possuem importantes implicações no controle da viremia do HIV-1, visto que, isolados virais obtidos nos estágios iniciais da infecção têm bastante influência na dinâmica da doença (Perelson *et al.*,

1996; Noviello *et al.*, 2007; Luzuriaga *et al.*, 1999; Wu *et al.*, 2004, Feeney, 2004). Portanto, devido à necessidade de explorar em mais detalhes os mecanismos evolutivos concorrentes entre adultos e crianças, abordaremos a genética evolutiva do processo de interação entre vírus-hospedeiro durante a infecção recente pelo HIV-1.



2. OBJETIVOS

2.1. OBJETIVO GERAL:

- Estudar comparativamente os mecanismos evolutivos atuantes especificamente no gene *nef* do HIV-1 entre populações de indivíduos com perfil imunológico distinto (*i.e.* crianças e adultos).

2.2. OBJETIVOS ESPECÍFICOS:

- Avaliar a diversidade genética, através da estimativa da razão entre o número de mutações não-sinônimas e mutações sinônimas no genes *nef* do HIV-1 em crianças e comparar com a diversidade genética do HIV em adultos recém infectados.
 - Comparar os resultados obtidos no gene *nef*, com resultados parciais da pressão seletiva no gene *pol*.
-



3. CASUÍSTICAS E MÉTODOS

3.1. ASPÉCTOS ÉTICOS.

O presente estudo foi submetido à apresentação do Comitê de Ética em Pesquisa do Hospital São Paulo/UNIFESP sob o nº 04/12044-3, diretamente ligado ao projeto também submetido e aprovado pelo mesmo Comitê sob o nº 0288/05.

3.2. CRITÉRIOS DE INCLUSÃO NO ESTUDO.

Foram considerados os seguintes critérios para a inclusão dos pacientes nos diferentes grupos:

Grupo 1.

1. Ser infectada pelo subtipo B do HIV-1,
2. Possuir 0 a 3 anos de idade.

Grupo 2.

1. Ser infectada pelo subtipo B do HIV-1,
2. Possuir 4 a 6 anos de idade.

3.3. POPULAÇÃO ESTUDADA.

Foram utilizadas as amostras de creme leucocitário (*buffy coat*) que estavam disponíveis no Laboratório de Retrovirologia da Escola Paulista de Medicina – UNIFESP, previamente identificadas. Estas amostras foram classificadas em dois grupos, um com amostras de crianças com idade entre 0-3 anos e outro com amostras de crianças com idade entre 4-6 anos.

Grupos	Extração de DNA	PCR	Sequenciamento
0-3	65	65	54
4-6	65	65	55
Total	130	130	109

3.4. PROCESSAMENTO DO MATERIAL.

3.4.1. EXTRAÇÃO DE DNA PROVIRAL.

A extração do DNA proviral das alíquotas do creme leucocitário (*buffy coat*) congeladas foi realizada utilizando-se o QIAamp DNA Blood Mini Kit (Quiagen Inc. Santa Clarita, CA, EUA). Segundo recomendações do fabricante

3.4.2. REAÇÃO EM CADEIA PELA POLIMERASE (PCR) EM DUAS ETAPAS (NESTED) PARA A AMPLIFICAÇÃO DO DNA PROVIRAL.

Após a extração de DNA, foi utilizada a técnica de *nested*-PCR para amplificar os genes *nef* e *pol* do HIV-1. Esta técnica é uma modificação da reação em cadeia pela polimerase (PCR) que consiste em duas etapas da PCR, envolvendo dois pares de iniciadores (*primers*). Na primeira etapa da reação um segmento de DNA é amplificado com os iniciadores (um par de *primers*). Uma alíquota do produto amplificado é então submetida à nova amplificação usando-se novo par de *primers*, localizados internamente em relação à posição do par de *primers* inicialmente utilizados. A segunda amplificação com *primers* diferentes, garante a especificidade da reação. Essa técnica apresenta sensibilidade superior à do PCR.

Gene nef:

Após a extração de DNA foi realizada a *nested*-PCR das amostras pertencentes ao grupo de crianças entre 0-3 anos e ao de crianças entre 4-6 anos de idade, de acordo com POP (procedimento operacional padrão) do Laboratório de Retrovirologia da UNIFESP. Esta reação ocorre em duas etapas:

1. A primeira etapa da *nested*-PCR foi realizada por meio da utilização de uma solução *master mix* que contém: Tampão 1X, MgCl₂ 2.5mM, dNTP 0.2mM, Taq DNA polimerase 0.25 U e 0.2 pmol de cada um dos iniciadores (*primers*) externos descritos por Shugars *et al.*, 1993 (Nef-1 e Nef-2) descritos na tabela (2):

Tabela 2. Descrição dos iniciadores externos utilizados na primeira etapa da *nested*-PCR do gene *nef*.

<i>Primers</i>	Posição (HXB2)	Sequência	Tamanho do Fragmento
Nef-1	8528 - 8547	5' ccg ccg ctt gag aga ctt ac 3'	1105pb
Nef-2	1314 – 1331	5' atc tga ggg tta gcc act cc 3'	

2. Para a segunda etapa foi utilizada uma solução de *master mix* nas mesmas condições da primeira etapa da *nested*-PCR, entretanto, para a segunda etapa, foi adicionado 0.2 pmol de cada um dos iniciadores (*primers*) internos descritos por Shugars *et al.*, 1993 (Nef-3 e Nef-4 – tabela 3) e foi transferido 5 µL do produto amplificado na primeira etapa para a segunda etapa da *nested*-PCR:

Tabela 3. Descrição dos iniciadores internos utilizados na segunda etapa da *nested*-PCR do gene *nef*.

<i>Primers</i>	Posição (HXB2)	Sequência	Tamanho do Fragmento
Nef-3	8688 – 8708	5' gat gct gag ggg aca gat agg 3'	872pb
Nef-4	2108 – 2127	5' cct ccc cgg aaa gtc ccc ag 3'	

Como controle negativo da reação de PCR foi utilizada água milli-Q e como controle positivo foi usado o genoma do HXB2 clonado em plasmídio. As condições físicas para esta

reação foram as mesmas descritas na primeira reação, com exceção da temperatura de anelamento dos *primers* que foi de 58°C.

Gene *pol*:

Após a extração de DNA foi realizada a *nested*-PCR das amostras pertencentes ao grupo de crianças entre 0-3 anos e ao de crianças entre 4-6 anos de idade, utilizando o protocolo do POP do Laboratório de Retrovirologia da UNIFESP, para amplificar a região da protease e da transcriptase reversa do gene *pol*. Esta reação ocorre em duas etapas:

1. A primeira etapa da *nested*-PCR foi realizada com a utilização de uma solução *master mix*: Tampão 1X, MgCl₂ 2.5mM, dNTP 0.4mM, Taq DNA polimerase 0.25 U e 0.2 pmol de cada um dos iniciadores (*primers*) externos (Pol-1 e Pol-2 - tabela 4):

Tabela 4. Descrição dos iniciadores externos utilizados na primeira etapa da *nested*-PCR do gene *pol*.

<i>Primers</i>	Posição (HXB2)	Seqüência	Tamanho do Fragmento
Pol-1	1842-1865	5' ggg agt ggg ggg acc cgg cca taa 3'	1709pb
Pol-2	3534-3558	5' cat tgg cct tgc ccc tgc ttc tgt 3'	

2. Para a segunda etapa foi utilizada uma solução de *máster mix* nas mesmas condições da primeira etapa da *nested*-PCR, entretanto, para a segunda etapa, foi adicionado 0.2 pmol de cada um dos iniciadores (*primers*) internos (Kozal-1 e Kozal-2 – tabela 5) (Kozal *et al.*, 1996) e foi transferido 5 µL do produto amplificado na primeira etapa para a segunda etapa da *nested*-PCR:

Tabela 5. Descrição dos iniciadores internos utilizados na segunda etapa da *nested*-PCR do gene *pol*.

<i>Primers</i>	Posição (HXB2)	Seqüência	Tamanho do Fragmento
Kozal-1	2147-2166	5' cag agc caa cag ccc cac ca 3'	1200pb
Kozal-2	5745-5776	5' ttt ccc cac taa ctt ctg tat gtc att gac a 3'	

Como controle negativo da reação de PCR foi utilizada água milli-Q e como controle positivo foi usado o genoma do HXB2 clonado em plasmídeo. As condições físicas para esta reação serão as mesmas descritas na primeira reação.

3.4.3. DETECÇÃO DOS PRODUTOS AMPLIFICADOS.

Para verificar a presença do produto amplificado pela reação de *nested*-PCR dos genes *nef* e *pol* com o tamanho esperado, de aproximadamente 872 e 1200 pb respectivamente, foi realizado o fracionamento eletroforético utilizando o produto de *nested*-PCR com tampão de carregamento (40% de sacarose; 0,25% de azul de bromofenol) que constitui na separação do produto amplificado em gel de agarose 1,5% em tri-borato/ EDTA (TBE) 1X (0,89 M Tris; 0,89 ácido Bórico e 0,02 M EDTA pH 8,0), contendo brometo de etídeo (0,5 µg/mL). A separação dos fragmentos foi efetuada a 100 Volts (V) e 55 mili-amperes (mA) por 45 minutos em tampão TBE 1X. Para a comparação do tamanho molecular em pares de base dos fragmentos amplificados foi utilizado um marcador de tamanho molecular padrão de 100 pares de base (bp) (GIBCO/BRL, Gaithersburg, MD, EUA). O gel foi visualizado sob iluminação ultravioleta (U.V.) (320 nm) e foi fotografado utilizando câmara Polaroid (Pittsburgh).

3.4.4. PURIFICAÇÃO DO PRODUTO DO SEGUNDO ROUND DA PCR.

Os produtos da *nested*-PCR foram purificados utilizando-se o ChargeSwitch® PCR Clean-Up Kit (Invitrogen Corporation, 1600 Faraday Ave. Carlsbad, CA 92008) que retira todos os reagentes utilizados na reação de PCR, deixando somente o produto amplificado. A tecnologia do ChargeSwitch® utiliza grânulos magnéticos (*beads* magnéticos) que é carga dependente de pH. Em baixas condições de pH os *beads* magnéticos possuem carga positiva e liga-se ao ácido nucléico que possui carga negativa. Após esta ligação o tubo contendo o produto da PCR, o tampão e os *beads* magnéticos são colocados numa rack magnética, onde os tubos ficarão posicionados na frente de um ímã e os *beads* ligados ao ácido nucléico irão migrar para a parede do tubo. O sobrenadante da reação é retirado e descartado e os ácidos nucléicos serão lavados por meio de tampões de lavagem fornecidos pelo Kit. Para eluir os ácidos nucléicos a carga

magnética é neutralizada pela utilização de um tampão com pH de 8,5, com isso os *beads* magnéticos vão se desligar dos ácidos nucleicos e tubo contendo os ácidos nucleicos purificados, o tampão e os *beads* magnéticos são colocados numa rack magnética, onde os tubos ficarão posicionados novamente na frente de um imã, com isso os *beads* irão migrar para a parede do tubo e o sobrenadante será o produto de PCR purificado.

Após a obtenção do produto de *nested*-PCR purificado, foi corrido um gel de agarose contendo tampão de carregamento (40 % de sacarose; 0,25 % de azul de bromofenol) e 5 µL do produto de purificação. Como padrão foi aplicado 5 µL do marcador de massa molecular para comparar a quantidade de DNA da amostra (produto de amplificação) com a intensidade de luminescência do marcador de massa molecular, esta luminescência é visualizada com a utilização de uma lâmpada ultravioleta (quanto maior a intensidade da luminescência, maior a quantidade de DNA presente naquela amostra).

3.4.5. REAÇÃO DE SEQUENCIAMENTO.

Os produtos purificados da reação de *nested*-PCR dos genes *nef* e *pol* foram submetidos à reação de sequenciamento. Esta reação é similar a uma reação de PCR, onde uma amostra de DNA é amplificada *in vitro*, porém, de forma diferente da reação de PCR, apenas um primer é utilizado para síntese das fitas, no sequenciamento de DNA.

A quantificação do produto da PCR purificado em nanogramas (ng) por meio da sua comparação a um padrão de peso molecular “Low DNA Mass Ladder” (200mL, 4mL/appl - Invitrogen, Califórnia, USA), observado em gel de agarose, foi previamente realizada, para determinar a quantidade de amostra utilizada na reação de sequenciamento. A solução para cada amostra, teve volume final de 6µL com água deionizada q.s.p, e o volume de cada amostra para esta solução, foi definido por meio da quantificação realizada, como descrito acima, onde cada amostra deve conter aproximadamente 40ng.

Após a quantificação e diluição das amostras para 40ng, foi realizada a reação utilizando-se o Kit “Ampli-Taq - Big-Dye Terminator Cycle Sequencing” (Applied Biosystems, CA, USA), onde uma solução é preparada, segundo as especificações do fabricante. Para a solução de reagentes foram utilizados tampão 5X, Big-Dye Terminator versão 3.0 (que contém

nucleotídeos livres marcados com fluorescência (ddntp's), nucleotídeos não marcados e enzima polimerase) e iniciadores na concentração de 2 pM/ μ L, para um volume final de 9 μ L, a este volume foi adicionado 6 μ L de amostras na concentração de aproximadamente 40ng. O sequenciamento é determinado bi-direcionalmente, usando-se um conjunto de *primers*, sendo um senso e outro anti-senso. Os iniciadores utilizados para o sequenciamento dos genes *nef* e *pol* foram (tabela 6):

Tabela 6. Descrição dos iniciadores utilizados na reação de seqüenciamento de DNA viral dos genes *nef* e *pol*.

Genes		
<i>Primers</i>	Posição (HXB2)	Seqüência
<i>Gene nef</i>		
Nef-3	8688 – 8708	5' gat gct gag ggg aca gat agg 3'
Nef-4	2108 – 2127	5' cct ccc cgg aaa gtc ccc ag 3'
<i>Gene pol</i>		
Kozal-1	2147-2166	5' cag agc caa cag ccc cac ca 3'
DP-11	1585-1611	5' cca ttc ctg gct tta att tta ctg gta 3'
F-1	2519-2539	5' gtt gac tca gat tgg ttg cac 3'
F-2	3305-3324	5' gta tgt cat tga cag tcc agc 3'

A reação de sequenciamento foi incubada em um termociclador (ABI 9700 - Applied Biosystems, CA, EUA) e aquecida por 5 minutos a 94°C. Após o aquecimento seguirão 25 ciclos consistindo de: desnaturação do DNA a 95°C por 30 segundos, anelamento dos *primers* a 55°C por 30 segundos e extensão a 60°C por 4 minutos. Durante esta reação o primer se liga à seqüência complementar e os nucleotídeos são incorporados de acordo com a fita molde. Os produtos gerados nesta reação foram precipitados para eliminar os resíduos e posteriormente analisados por um seqüenciador automático modelo ABI 3100 (ABI-Perkin-Elmer, EUA).

3.4.6. PRECIPITAÇÃO DA REAÇÃO DE SEQÜENCIAMENTO.

Após o termino da reação de sequenciamento são gerados fragmentos de tamanhos diferentes e também há sobra de *primers*, dDNTPs e alguns sais. A precipitação da reação de sequenciamento visa eliminar tais resíduos aumentando a qualidade das seqüências produzidas.

3.4.7. SEQUENCIAMENTO AUTOMATIZADO DE ÁCIDOS NUCLÉICOS.

Após a reação de sequenciamento, os fragmentos de DNA foram submetidos à separação eletroforética. A leitura destes fragmentos foi realizada por meio de um feixe de laser do seqüenciador automático de DNA modelo ABI PRISM 3100 (Applied Biosystems, CA, EUA).

3.4.8. ANÁLISE DO SEQUENCIAMENTO.

As edições dos fragmentos seqüenciados foram feitas utilizando-se o programa Sequencer 4.2 (Gene Code Corporation, EUA) por meio da análise do eletroferograma gerado a partir da leitura dos fragmentos de DNA no sequenciamento automático.

Na análise do gene *nef* foi realizado o sequenciamento do mesmo trecho duas vezes, com a utilização de *primers* diferentes para cada reação, estes *primers* foram os mesmos utilizados na segunda etapa reação de PCR. Para a análise do gene *pol*, cada região do gene (protease e transcriptase reversa) foi seqüenciada duas vezes para cada amostra analisada com a utilização dos primers Kozal 1, DP11, F1 e F2, já citados anteriormente na reação de *nested-PCR*.

Os sinais fluorescentes gerados no sequenciamento automático são transferidos para um computador que analisa a posição e a força do sinal, produzindo um eletroferograma. Por meio da análise do eletroferograma gerado a partir da leitura dos fragmentos de DNA no

sequenciamento automático as edições dos fragmentos seqüenciados foram realizadas utilizando-se o programa Sequencer 4.2 (Gene Code Corporation, EUA).

A área sob o pico do eletroferograma representa a intensidade do sinal e a cor do pico é especificada de acordo com a base nucleotídica em questão. Sendo assim, o programa nomeia a base A, C, T ou G para cada posição, ou então N. Nos casos, onde a determinação das bases não é clara (ex. múltiplos picos sobrepostos), a nomeação das bases foi realizada de forma arbitrária onde os picos com maior intensidade são considerados. Assim usam-se os códigos da IUPAC (*International Union of Pure and Applied Chemistry* - União Internacional de Química Pura e Aplicada) citados na tabela abaixo (Tabela 7). Se numa mesma posição do eletroferograma, por exemplo, existem dois picos com cores designando as bases A e G, então, usa-se a letra R na seqüência de DNA para designar a ambigüidade A/G.

Tabela 7. Código IUPAC.

Código	Tradução
A	Adenina
C	Citosina
G	Guanina
T	Timina
B	C, G ou T
D	A, G ou T
H	A, C ou T
R	A ou G
Y	C ou T
K	G ou T
M	A ou C
S	G ou C
W	A ou T
N	Qualquer base
V	A, C ou G

3.4.9. ANÁLISE MOLECULAR DOS DADOS.

Compilação e alinhamento das seqüências.

As seqüências que contém inserções e deleções (*indels*), códons indeterminados, e códons de terminação foram editadas usando o programa Se-Al (Rambaut *et al.*, 2000). As edições foram realizadas a fim de preservar a fase de leitura da seqüência codificantes. Nos casos onde as seqüências continham grande número de indels e codons indeterminados, estas foram excluídas das análises. Essa forma de edição é necessária para que o computo dos valores de likelihood e para a determinação das taxas das taxas evolutivas. A seguir, as seqüências foram alinhadas com referências obtidas de bancos de dados, utilizando o programa Clustal W (Thompson *et al.*, 1994).

Análise dos subtipos.

Após o alinhamento, as seqüências selecionadas foram subtipadas por meio de duas metodologias para que nenhuma amostra pertencente a outros subtipos não B fosse analisada, pois esta poderia superestimar os cálculos de pressão seletiva do gene. Inicialmente foi realizada uma triagem inicial utilizando o programa “*HIV blast search*” (http://HIV-web.lanl.gov/content/HIV-db/BASIC_BLAST/basic_blast.html), devido a rapidez da obtenção de uma identificação de subtipos. No entanto, houve a necessidade de uma análise mais detalhada, portanto, as seqüências, dos genes *nef* e *pol* do HIV-1, pertencentes ao subtipo B, segundo o programa Blast, foram submetidas a uma análise filogenética usando um método de verossimilhança, que utiliza matrizes de probabilidade de mudança entre nucleotídeos que explicitam modelos evolutivos para a inferência dos estados ancestrais dos caracteres a partir do estados observados nos táxons apicais.

Inferência filogenética.

O método de inferência filogenética foi feito por máxima verossimilhança que utiliza matrizes de probabilidade de mudança entre nucleotídeos que explicitam modelos evolutivos para a inferência dos estados ancestrais dos caracteres a partir do estados observados nos táxons apicais. As reconstruções ótimas maximizam a verossimilhança para todos os caracteres ao longo de uma topologia. No entanto, o número de substituições observadas em um alinhamento tende a subestimar o número real de eventos evolutivos para determinado conjunto de seqüências (Yang, 1994a; Li, 1997). Por isso, é necessário o uso de modelos evolutivos de substituição de nucleotídeos (ou aminoácidos) que permitem a correção do número de eventos evolutivos (substituições) num determinado conjunto de dados (seqüências). Os modelos utilizados para descrever o processo de evolução das substituições de nucleotídeos, ou transição de estados (*ex.* A→G) em seqüências de DNA (RNA) seguem as seguintes premissas: 1. as taxas de mutações para cada nucleotídeo não se alteram com o tempo, 2. disponibilidade celular de nucleotídeos é constante e 3. o estado atual de um nucleotídeo só depende do seu estado anterior, e não de estados intermediários. As probabilidades de transição (P) de nucleotídeos (T, C, A, G) de um estado (i) para outro estado (j) em um intervalo de tempo (dt), dadas às premissas anteriores, podem ser expressas de forma matricial (4x4).

$$P_t = \begin{pmatrix} p_{TT} & p_{CT} & p_{AT} & p_{GT} \\ p_{TC} & p_{CC} & p_{AC} & p_{GC} \\ p_{TA} & p_{CA} & p_{AA} & p_{GA} \\ p_{TG} & p_{CG} & p_{AG} & p_{GG} \end{pmatrix} \quad (1)$$

Este modelo de representação matricial das probabilidades, expresso em termos de estados descontínuos, é conhecido como cadeia de Markov. Cada elemento p_{ij} da matriz corresponde a uma probabilidade de mudança de estado i pelo j (*ex.* A→C), num intervalo de tempo. Vários modelos evolutivos propostos diferem na complexidade dos elementos da matriz, indo de um modelo onde todos os elementos têm a mesma probabilidade (25%) de transição (modelo Jukes e Cantor) até uma matriz mais geral, onde cada elemento tem uma probabilidade distinta.

Neste trabalho, a matriz de probabilidade de substituição foi escolhida de acordo com o modelo evolutivo indicado pelo teste de razão de verossimilhança LRT (“Likelihood Ratio

Test”). O teste LRT é usado em inferências estatísticas e deve ser aplicado somente quando os modelos são aninhados. Dois modelos são aninhados quando o modelo mais complexo (aquele com um número maior de parâmetros livres) derivou de um modelo mais simples (com menos parâmetros livres) e no mais complexo existe a adição de pelo menos um parâmetro. Assim, após a estimativa dos valores de verossimilhança do modelo mais simples (L_0) e do mais complexo (L_1) eles são tratados no teste LRT como hipótese nula (H_0) e hipótese alternativa (H_1), respectivamente. Isso é feito pela diferença de valores do logaritmo da verossimilhança entre dois modelos aninhados na seguinte equação:

$$LRT=(2[\ln L_1-\ln L_0]) \quad (2)$$

Os valores de LRT seguem uma distribuição qui-quadrado assintótica (χ^2), onde os graus de liberdade são dados pela diferença de parâmetros livres entre os modelos. A hipótese nula é descartada quando o valor do teste for menor que o nível de significância escolhido (valor de alfa). As estimativas dos parâmetros e de verossimilhança para 56 modelos aninhados foram feitas com o programa PAUP. O teste LRT foi executado pelo programa MODELTEST 3.06b (Huelsenbeck, Crandall, 1997; Posada, Crandall, 1998). A heterogeneidade das taxas de substituição entre os sítios foi incorporada através da distribuição gama (Γ), onde o parâmetro alfa (α) e a proporção de sítios que não variam foram estimados a partir dos dados. Da mesma forma, os valores das matrizes de transição e a frequência de bases também foram estimados de forma iterativa durante o processo de reconstrução. A reconstrução filogenética foi feita em etapas:

1. Iniciou-se com a escolha do modelo de evolução de nucleotídeos e obtenção de uma árvore inicial pelo método de distâncias de agrupamento de vizinhos (“neighbor joining”).
2. Por fim, foram feitos rearranjos (perturbações) na árvore inferida pelo método dos quartetos, através do método de busca heurística com o algoritmo “*stepwise addition*” e o método “*nearest neighbor interchange (NNI)*” (algumas vezes foi usado o método “*tree bisection and reconnection*” (TBR)) de rearranjo das árvores. As árvores com maior verossimilhança que a inicial foram escolhidas.

Esse método de reconstrução filogenética é bastante apropriado para ajuste de comprimentos de ramos, visto que em cada uma das etapas é feito um refinamento dos parâmetros da matriz de transição (modelo escolhido pelo MODELTEST).

Neste trabalho, foram utilizadas matrizes de probabilidade de mudança entre nucleotídeos que explicitam modelos evolutivos para a inferência dos estados ancestrais dos caracteres a partir dos estados observados nos táxons apicais. As reconstruções ótimas maximizam a verossimilhança (*maximum likelihood*) para todos os caracteres ao longo de uma topologia. As árvores de verossimilhança foram feitas com o modelo GTR (*general time reversible*), utilizando o software PAUP* (Swofford *et al.*, 2001).

Métodos de detecção de seleção.

Os métodos mais sensíveis de detecção de seleção natural utilizam a razão entre o número de substituições não-sinônimas por sítio não-sinônimo (dN) e o número de substituições sinônimas por sítio sinônimo (dS) ($\omega = dN/dS$). Nos casos onde as substituições não-sinônimas enfrentam restrições funcional-estruturais espera-se $dN < dS$ (seleção negativa ou purificadora). Por outro lado, na presença de seleção positiva (ou diversificadora) espera-se obter $dN > dS$. O regime de seleção é determinado pela razão entre dN/dS (ω). Ou seja:

$\omega > 1$ para seleção positiva.

$\omega < 1$ para seleção negativa ou purificadora.

$\omega = 1$ para um regime neutro de substituições.

Determinação de dN e dS pelo método de distancias par-a-par

Nessa abordagem, são calculadas as distâncias genéticas normalizadas pelo número de sítios ao longo das seqüências. As distâncias são computadas par-a-par para um alinhamento composto por n seqüências de DNA de mesmo comprimento (método de Nei-Gojobori) (Nei & Kumar, 2000).

Especificamente, no método de Nei-Gojobori, dN e dS são estimadas como a razão do número total de substituições sinônimas (S) e não-sinônimas (N), divididos respectivamente pelo número de sítios potencialmente sinônimos (s) e potencialmente não-sinônimos (n). Estes

valores são computados para cada códon sob a premissa de que as mudanças de nucleotídeos têm a mesma probabilidade (modelo de Jukes-Cantor). Assumindo que f_i designa a proporção de mudanças sinônimas possíveis na posição i -ésima de um determinado códon ($i=1, 2, 3$), o número de sítios potencialmente sinônimos (s) pode ser obtido por:

$$s = \sum_{i=1}^3 f_i \quad (3)$$

O número de sítios potencialmente não-sinônimos (n) é dado por $n=3-s$ (Nei & Kumar, 2000). Para exemplificar, temos o caso da trinca da fenilalanina (TTT), onde as substituições na primeira e na segunda posição são não-sinônimas e somente 1 em 3 substituições são sinônimas (TTC). Assim, temos que o número de sítios potencialmente sinônimos (s) é igual a: $0+0+1/3$. Por outro lado, o número de sítios potencialmente não-sinônimos (n) é igual a: $3-1/3=8/3$. Quando uma mudança resulta num códon de terminação este não é computado. Para obter os valores do número total de substituições sinônimas (S) na seqüência inteira usa-se a expressão:

$$S = \sum_{j=1}^C s_j \quad (4)$$

Para os valores do número de substituições não-sinônimas (N) usa-se a expressão $N=3C-S$, onde C é o número total de códons e s_j é o valor do número de sítios potencialmente sinônimos para o j -ésimo códon (Nei & Kumar, 2000). Nota-se que $S+N=3C$ é igual ao número total de nucleotídeos comparados. O cálculo das diferenças de nucleotídeos sinônimos e não-sinônimos é feito par-a-par, códon por códon. Por exemplo, na comparação dos pares de códons TTT (F) e GTA (V) existem duas possibilidades neste caminho:

$$(1) \text{ TTT (F)} \leftrightarrow \text{GTT (V)} \leftrightarrow \text{GTA (V)}$$

$$(2) \text{ TTT (F)} \leftrightarrow \text{TTA (L)} \leftrightarrow \text{GTA (V)}$$

Se designarmos os símbolos s_d e n_d para as diferenças sinônimas e não sinônimas por códon, respectivamente, e assumindo que ambos os caminhos (1) e (2) têm as mesmas probabilidades, temos que $s_d = 1/2$ e que $n_d = 3/2$. Nos casos onde as substituições resultarem em códons de terminação, os passos não são computados. Com base nestas premissas, é possível estimar a proporção das diferenças sinônimas (p_s) e não-sinônimas (p_n) por: $p_s = s_d / S$ e $p_n = n_d / N$, respectivamente, onde S e N são os números médios de sítios sinônimos e não-sinônimos para duas seqüências. Para estimar o número de substituições sinônimas (dS) e o número de não sinônimas (dN) por sítio usa-se a expressão:

$$dX = -(3/4) \ln[1 - (4/3)p_x] \quad (5)$$

do modelo de Jukes-Cantor, substituindo-se p por p_s ou p_n . Assim, o número de substituições sinônimas é calculado por: $dS = -(3/4) \ln[1 - (4/3)p_s]$, e o número de substituições não sinônimas é calculado por $dN = -(3/4) \ln[1 - (4/3)p_n]$. Com isso, o valor de ω é calculado a partir das médias dos valores de dN e dS . Este método fornece uma estimativa aproximada de ω visto que, o processo de substituição nos sítios sinônimos e não sinônimos, não segue o modelo de Jukes-Cantor (Nei & Kumar, 2000). Além de ser sensível ao comprimento e ao número de seqüências analisadas, outros fatores podem influenciar nos resultados. O excesso de polimorfismos e as taxas de heterogeneidade entre os sítios também podem influenciar nas estimativas.

Vários programas usam o método de Nei-Gojobori (1986). Em nossas análises usamos o programa Mega2 que calcula a média geral entre os valores de dN e dS e seus erros associados para todos os pares de seqüências (Nei & Kumar, 2000).

Determinação de dN e dS pelo método de verossimilhança máxima

Goldman e Yang (1994) e MUSE e Gaut (1994) propuseram uma estimação da relação entre as taxas de substituições não-sinônimas e sinônimas, ω , dentro de um contexto da verossimilhança máxima (*maximum likelihood*), utilizando um modelo de substituição de códon. Nielsen e Yang (1998) propuseram um método onde os sítios (códon) de uma proteína podem assumir valores de dN e dS distintos para cada sitio. Com isso, a detecção de seleção positiva pode ser feita individualmente para cada um dos sítios (códon) para uma determinada proteína. Este método está implementado no programa CODEML, do pacote PAML v.3.154 (Yang, 1997). O CODEML utiliza vários modelos evolutivos que diferem entre si na distribuição de dN/dS (ω) entre os códon.

Dentre os vários modelos de substituição de códon implementados no programa CODEML alguns deles são mais ajustáveis às seqüências de DNA de vírus e serão discutidos abaixo (Nielsen & Yang, 1998; Yang *et al.*, 2000):

1. M0 é o mais simples, pois assume um único valor de dN/dS para todos os sítios (códon).
 2. M1 (neutro) assume duas categorias de códon: os conservados ($dN/dS = 0$) e os neutros ($dN/dS = 1$), distribuídos em diferentes proporções (p_0 e p_1).
-

3. M2 (seleção) comparado ao modelo anterior, adiciona uma nova categoria de códons (p_2) a qual pode assumir valores maiores que 1 para dN/dS (ω_2).
4. M3 (discreto) permite n categorias de códons para um determinado número de classes (K). Os valores de dN/dS (ω) por categoria são estimados a partir dos dados e seguem uma distribuição gama, subdividida em K categorias. Em nossos estudos utilizamos $K=8$, de acordo com Yang *et al.* (2000) sendo que obtivemos valores de ω_0 até ω_7 e suas respectivas proporções (p_0 até p_7) de códons pertencentes às diferentes classes. A proporção p_8 de sítios com valores de dN/dS mais altos avalia os códons com seleção positiva mais intensa.

A abordagem de detecção de seleção positiva é inicialmente feita pela estimativa de parâmetros para todos os modelos acima citados; a seguir, o teste LRT é usado para verificar qual o modelo que melhor se ajusta aos dados. O teste é feito pela diferença dos valores de verossimilhança entre os modelos aninhados. Visto que essa diferença segue uma distribuição qui-quadrado assintótica isso permite determinar a significância estatística entre os modelos, assumindo que os graus de liberdade são dados pela diferença entre os parâmetros livres de cada modelo.

A detecção de seleção positiva ocorre se o modelo escolhido assumir valores de dN/dS maiores que um ($\omega > 1$) e se os valores estimados forem significativos. Assim, se o teste LRT indicar que o modelo M2 é mais adequado que os modelos M0 e M1, então a hipótese de seleção positiva é aceita. Se acaso o teste indicar que o modelo M3 é melhor, quando comparados os modelos M1 *vs.* M3 e M2 *vs.* M3, isso permite a identificação dos códons sob seleção positiva. A melhor comparação entre: M0 *vs.* M2 e M1 *vs.* M2, é mais conservadora na estimativa para detectar seleção positiva (Yang, 2000).

Em adição utilizamos um método de verossimilhança que avalia se o número de substituições não-sinônimas por sítio não-sinônimo (dN) e o número de substituições sinônimas por sítio sinônimo (dS) diferem entre duas populações (Pond *et al.*, 2006). A significância das diferenças é avaliada por LRT assumindo um grau de liberdade.

A taxa de substituição não-sinônima/sinônima ($\omega=dN/dS$) fornece uma medida sensível da pressão seletiva pois possibilita a adequação de modelos de substituição de códons (uso de diferentes matrizes de transição de aminoácidos) e o ajuste das taxas de heterogeneidade nas estimativas de ω para cada códon, também possibilita o uso da informação genealógica dos

dados para descrever o processo evolutivo de substituição de códons (Felsenstein, 1973; Yang, 1994).



4. RESULTADOS

4.1. EXTRAÇÃO DE DNA PROVIRAL.

O número de amostras proposto para o estudo foi de 60 amostras, sendo 30 de crianças de 0-3 anos de idade e 30 de crianças de 4-6 anos de idade.

As extrações de DNA proviral do *buffy coat* de sangue total foram feitas em 130 amostras, sendo que 65 amostras fazem parte do grupo de crianças de 0-3 anos de idade e as outras 65 amostras fazem parte do grupo de crianças de 4-6 anos de idade. Esta extração de DNA foi realizada utilizando-se o Kit QIAamp DNA Blood Mini Kit (Quiagen Inc. Santa Clarita, CA, EUA).

Tabela 8. Quantidade de amostras processadas de cada gene estudado, genes *nef* e *pol*, dos grupos de crianças, com idades entre 0-3 anos e 4-6 anos.

Grupos	Genes	Extração de DNA	PCR	Sequenciamento*
0 - 3	<i>Nef</i>	65	65	54
	<i>Pol</i>	30	30	30
4 - 6	<i>Nef</i>	65	65	55
	<i>Pol</i>	30	30	30
Total	<i>Nef</i>	130	130	109
	<i>Pol</i>	60	60	60

*Correspondentes somente às amostras com amplificação positiva

4.2. REAÇÃO EM CADEIA PELA POLIMERASE (NESTED-PCR) PARA A AMPLIFICAÇÃO DO DNA PROVIRAL.

4.2.2. GENE *NEF*.

O DNA proviral de 109 amostras foi amplificado, sendo que 54 pertenciam ao grupo de crianças entre 0 - 3 anos e 55 ao grupo de crianças entre 4 - 6 anos de idade.

O produto amplificado na segunda etapa da PCR foi visualizado pela separação eletroforética em gel de agarose e as amostras que apresentaram produto amplificado foram purificadas. Após a purificação do produto da *nested*-PCR foi realizado o sequenciamento de DNA das amostras dos grupos de crianças com idade entre 0-3 e 4-6 anos e posteriormente estas foram analisadas.

As seqüências selecionadas foram subtipadas por meio de duas metodologias para que nenhuma amostra pertencente a outros subtipos não B fosse analisada, pois esta poderia superestimar os cálculos de pressão seletiva do gene. Inicialmente foi realizada uma triagem inicial utilizando o programa “*HIV blast search*” (http://HIV-web.lanl.gov/content/HIV-db/BASIC_BLAST/basic_blast.html), pois este fornece rapidamente a análise dos subtipos. No entanto, para uma análise mais detalhada dos subtipos estas amostras foram submetidas a uma metodologia mais sensível por meio da construção de uma árvore filogenética de verossimilhança feita com o modelo GTR (*general time reversible*). Desta forma, somente as amostras pertencentes ao subtipo B foram selecionadas para análise.

As amostras pertencentes ao subtipo B foram analisadas para verificar a presença de pressão seletiva no gene *nef*, visto que a presença de pressão seletiva foi detectada na maioria dos genes do HIV-1 e sugere-se que a intensidade da seleção está relacionada com o resultado da infecção (Yang *et al.*, 2003; Ross & Rodrigo, 2002), portanto a ação da pressão seletiva no gene *nef* do HIV-1 foi analisada através da comparação de indivíduos com perfil imunológico distinto (crianças e adultos).

4.2.3.GENE *POL*.

As amostras foram analisadas a partir do gene *nef* e as que pertenciam ao subtipo B do HIV-1 foram selecionadas para uma futura análise do gene *pol*.

O produto amplificado na segunda etapa da *nested*-PCR do gene *pol* das amostras selecionadas, foi visualizado pela separação por eletroforese em gel de agarose, onde as bandas apresentavam o tamanho esperado, o produto amplificado foi purificado e seqüenciado.

As seqüências selecionadas foram subtipadas por meio do programa “*HIV blast search*” (http://HIV-web.lanl.gov/content/HIV-db/BASIC_BLAST/basic_blast.html) e assim foi possível

selecionar as amostras do subtipo B. Contudo, com a finalidade de uma análise mais detalhada dos subtipos, as amostras pertencentes ao subtipo B foram novamente analisadas por uma metodologia mais sensível por meio da construção de uma árvore filogenética de verossimilhança feita com o modelo GTR (*general time reversible*).

4.3. ANÁLISE DA DIVERSIDADE GENÉTICA.

4.3.1. IDENTIFICAÇÃO DO SUBTIPO B DO HIV-1.

Após o alinhamento das seqüências foi realizada a subtipagem das amostras por meio da análise filogenética usando um método de verossimilhança para identificar os subtipos do HIV-1 dos genes *nef* e *pol* nas amostras dos dois grupos de crianças (Figura 5 e 6). Essa análise mostrou que entre as 60 seqüências do gene *nef*, maioria das seqüências do gene *nef* pertencem ao subtipo B (azul na figura 5). Três seqüências agruparam com referências F, as quais podem ser F ou recombinantes BF (verde na figura 5). Outra seqüência foi classificada pertencente ao subtipo C (amarelo na figura 5). Todas as seqüências não identificadas como subtipo B foram excluídas das análises seguintes.

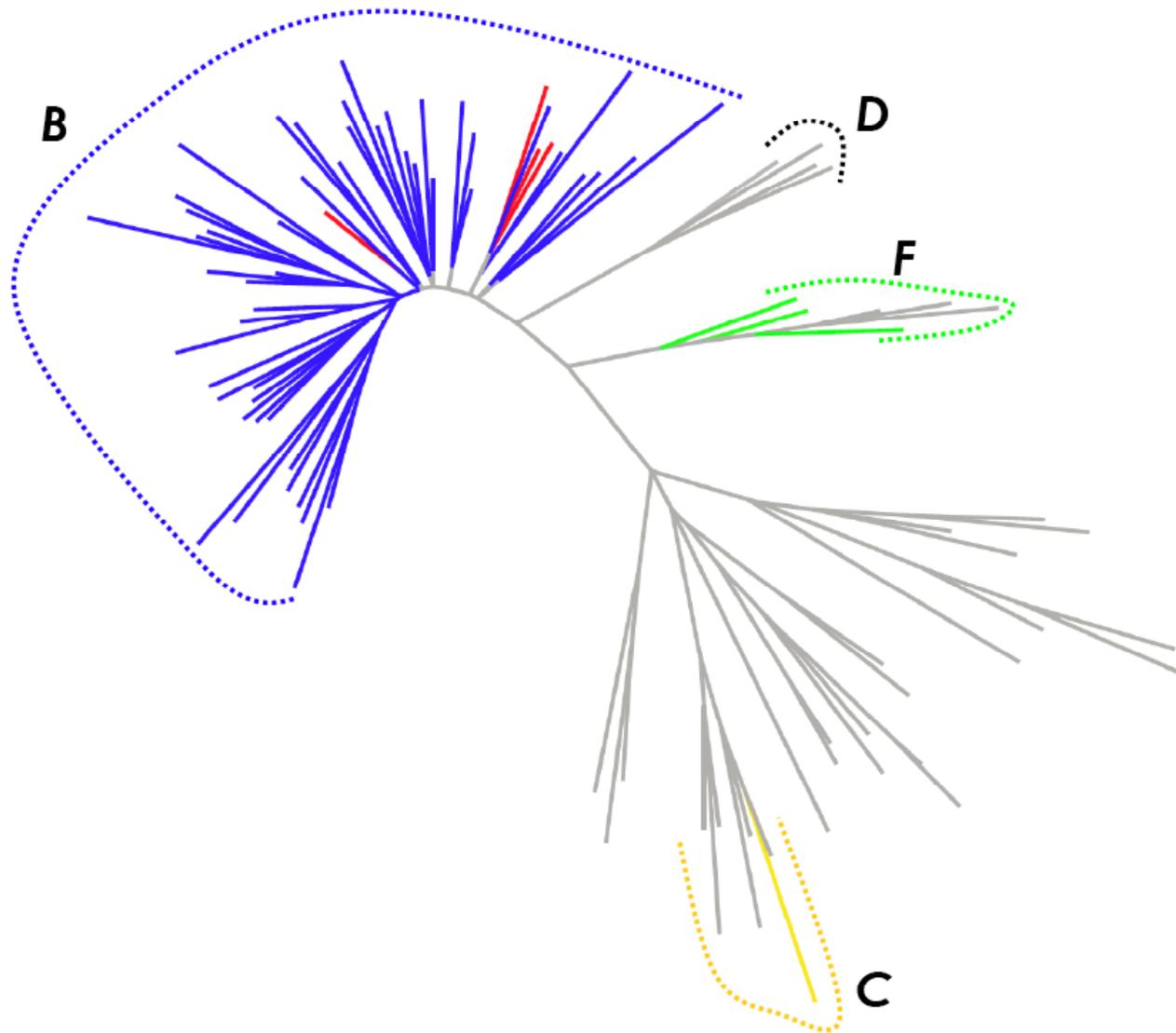


Figura 5. Genotipagem das seqüências do gene *nef* do HIV-1 obtidas das crianças. Os agrupamentos representativos dos principais subtipos do HIV-1 estão identificados na figura (*i.e.* B, C, D e F). A árvore mostra que maior parte das amostras das crianças (azuis) agrupa com as amostras referencias do subtipo B (vermelho) do HIV-1. Três amostras de crianças agrupam com as referencias do subtipo F (verde) e uma amostra (amarelo) agrupa com o subtipo C.

O mesmo procedimento de identificação de subtipos por análise filogenética foi feito com 18 seqüências do gene *pol*. Essa análise revelou que a maioria das seqüências do gene *pol* pertenciam ao subtipo B (azul na figura 6). Quatro amostras agruparam com as referencias do subtipo F, sendo prováveis recombinantes BF. Seis amostras apresentaram padrão indeterminado de agrupamento, no entanto é provável que essas amostras sejam formas recombinantes.

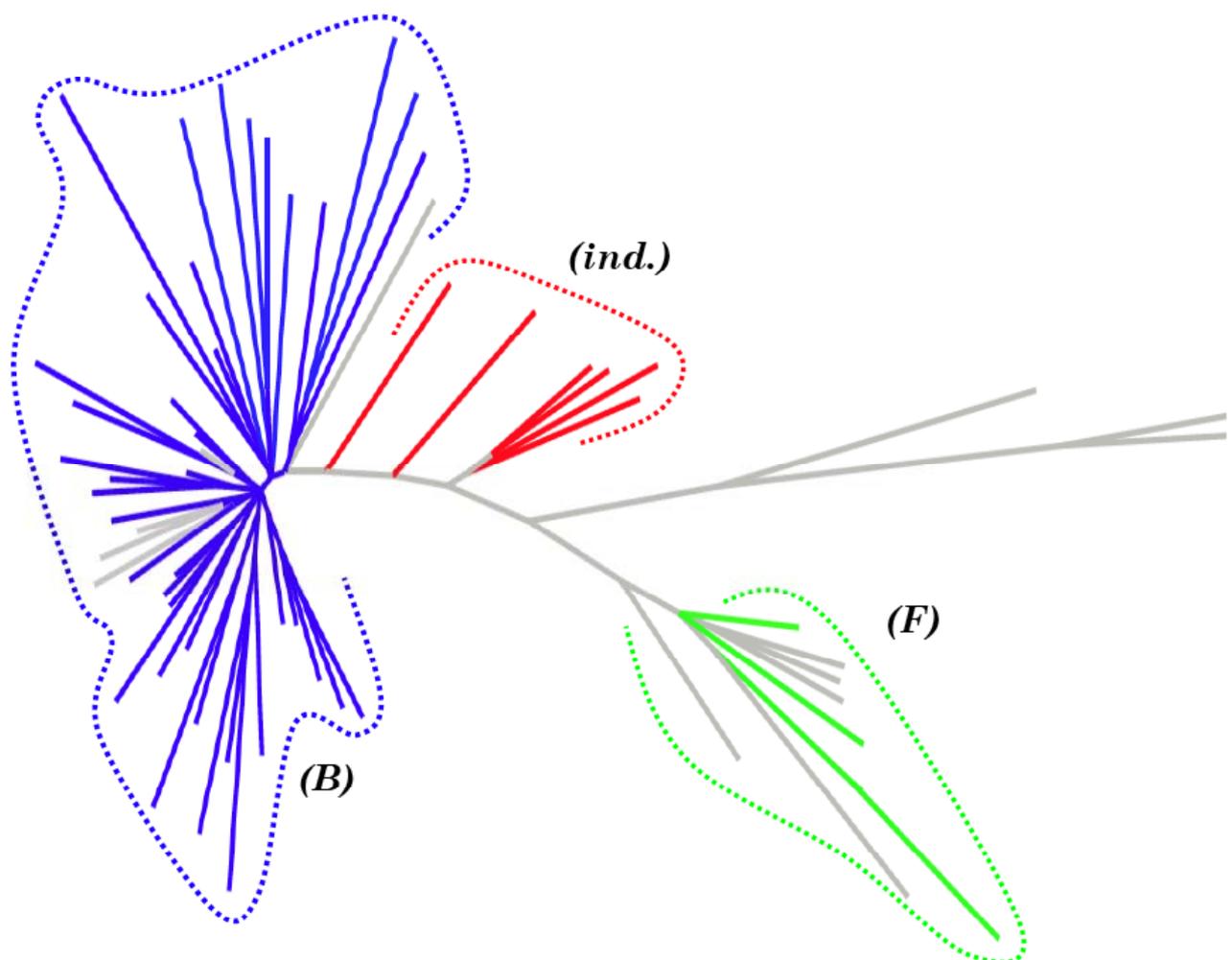


Figura 6. Genotipagem das seqüências do gene *pol* do HIV-1 obtidas das crianças. Os agrupamentos representativos dos principais subtipos do HIV-1 estão identificados na figura (*i.e.* B, F). A árvore mostra que, maior parte das amostras das crianças agrupam com as amostras referencias do subtipo B (azuis na árvore) do HIV-1. Quatro amostras de crianças agrupam com as referencias do subtipo F (verde) e seis amostra com padrão de agrupamento indefinido (provavelmente recombinantes BF).

A análise filogenética permitiu a classificação das seqüências dos genes *pol* e *nef*, de acordo com os subtipos referencias do HIV-1. Dessa forma temos uma idéia da proporção de subtipos encontrados nas seqüências crianças. Os resultados mostraram que existe uma prevalência do subtipo B nas seqüências analisadas. Apesar dessa amostra ter sido escolhida ao acaso a prevalência do subtipo B reflete a tendência da epidemia Brasileira observada caracterizada em adultos.

4.3.2. ANÁLISE DA PRESSÃO SELETIVA NO GENE *NEF*.

Após a determinação do genótipo das amostras passamos a análise da pressão seletiva nas seqüências de crianças do HIV-1 subtipo B. Dessa forma, foram excluídas seqüência com indels e/ou ambigüidades, as remanescentes foram separadas em dois grupos de acordo com a idade: 1. crianças entre 0 e 3 anos (n=14); 2. crianças entre 4 a 6 anos (n=39). Inicialmente inferimos filogenias para cada grupo. Essas filogenias devem conter uma estimativa razoável de comprimentos de ramos. Portanto, usamos o método de verossimilhança e assumimos o modelo GTR para estimar os comprimentos dos ramos das filogenias. Visto que os dois grupos de crianças apresentavam idades distintas e que essa idade corresponde ao tempo de diversificação do HIV-1 (infecção materno-infantil), avaliamos o efeito desse tempo nas taxa nas topologias. Para isso estimamos o número de substituição por sítio. Essa análise mostrou que o maior tempo de diversificação do HIV-1 no grupo 2 (crianças de 4 a 6 anos) implica na inflação das taxas evolutivas nos ramos externos da filogenia, comparado a taxas observadas em crianças de 0 a 3 anos (figura 7), no entanto, esta diferença no comprimento dos ramos não interferem no valor de ω (omega).

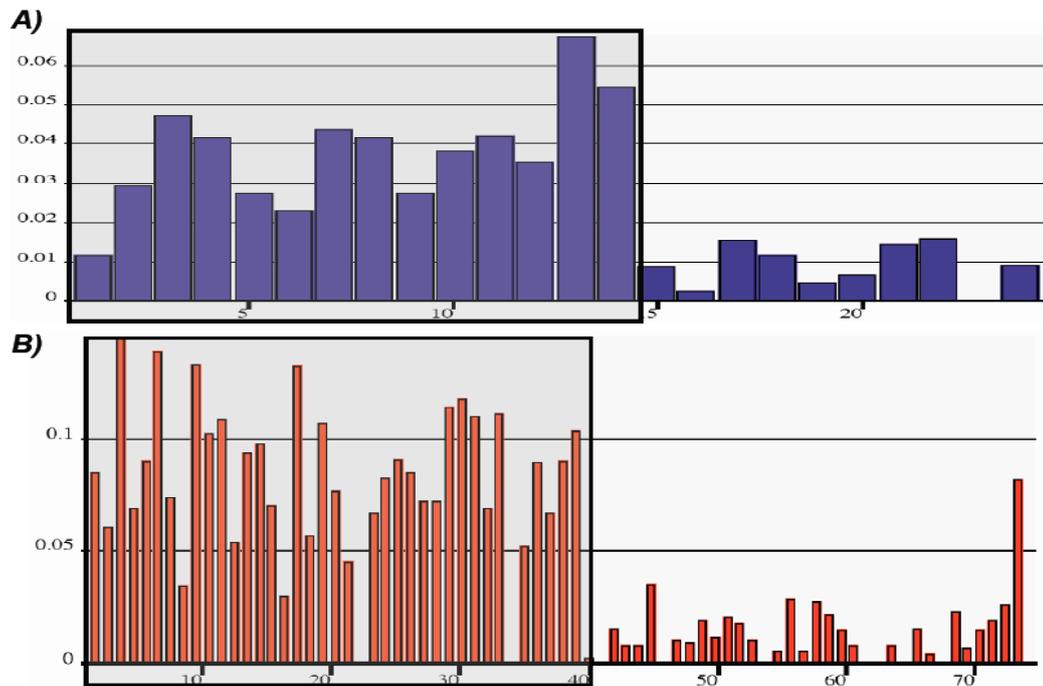


Figura 7. Comprimentos dos ramos de filogenias do gene *nef* do HIV-1 obtidas de crianças. A) Comprimentos dos ramos da filogenia de crianças entre 0 a 3 anos. B) Comprimento dos ramos da filogenia de crianças entre 4 a 6 anos. Os quadros cinza mostram os valores das taxas evolutivas dos ramos externos das filogenias.

Esses resultados mostram que existe uma influência no tempo de diversificação do HIV-1 que é refletido diretamente no comprimento dos ramos das filogenias obtidas por verossimilhança. As taxas evolutivas são mais altas quanto maior o tempo de diversificação do vírus na população. Em nossos dados observamos quanto maior for a idade das crianças maior é o número de substituição por sítio do vírus (taxa evolutiva) (figura 7).

Visto que os dois grupos de crianças apresentaram distinções quanto às taxas evolutivas (numero de substituição de nucleotídeos por sítio entre as seqüências), passamos a analisar as taxas de pressão seletiva entre as crianças. As filogenias e os alinhamentos editados, ou seja, excluídos os códons de terminação, inserções e deleções, foram analisados para determinar a pressão seletiva. Inicialmente a abordagem foi feita através dos métodos de Nielsen & Yang (NY) (1998) que utiliza maximum likelihood para estimar o número de substituições não sinônimas (dN) e sinônimas (dS) códon a códon. Este método está implementado no programa PAML v. 3.14. Os resultados mostraram um predomínio de seleção purificadora ($dN/dS < 1$) no gene *nef*. Assim, 31% e 35% dos códons no gene *nef*, respectivamente em crianças entre 0 a 3 e

4 a 6 anos, estão sob pressão purificadora. Seleção positiva ($dN/dS > 1$) foi detectada em 8% a 9% dos códons, respectivamente em crianças de 0 a 3 e 4 a 6 anos. A distribuição dos códons sob seleção positiva em ambos os grupos de crianças é idêntica (figura 8). Estes códons estão presentes ao longo das seqüências, no entanto, existe uma concentração de códons sob seleção positiva na região inicial do gene *nef* (área cinza na figura 8) em ambos os grupos de crianças.

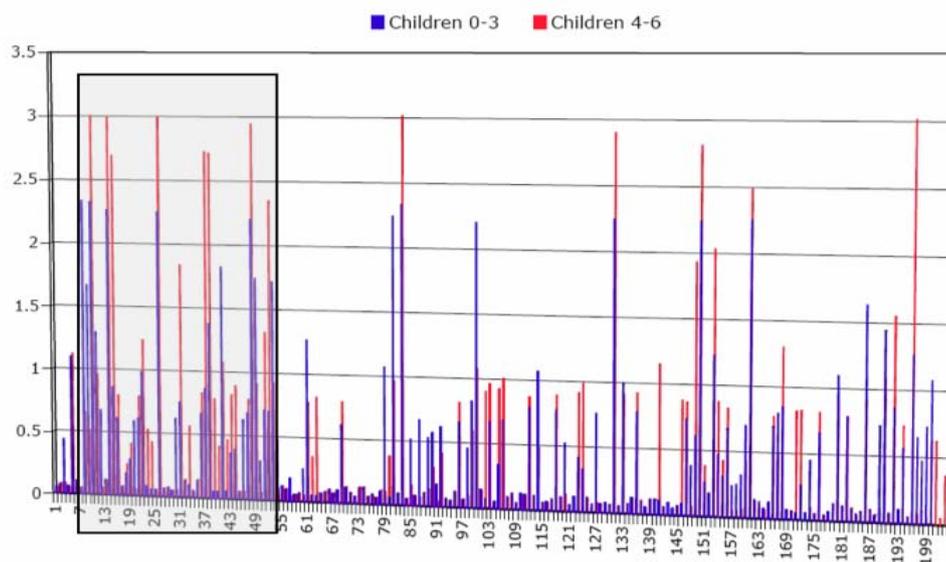


Figura 8. Valores de dN/dS nos códons do gene *nef* do HIV-1 em crianças. A área cinza marca a região no gene onde existe uma concentração de códons sob seleção positiva ($dN/dS > 1$).

A comparação entre a pressão seletiva de crianças com diferentes idades mostrou que a tanto a intensidade (valores de dN/dS) quanto a localização dos códons sob seleção positiva são coincidentes entre os dois grupos. Isso sugere que não existem diferenças significativas nas pressões seletiva em crianças de acordo com o tempo de diversificação do HIV-1. Além disso, as diferenças nas taxas evolutivas observadas principalmente nos ramos externos das filogenias não afetam as estimativas das taxas seletivas.

Uma vez constatadas similaridades na pressão seletiva em crianças de acordo com as idades, passamos as análises comparativas entre crianças e adultos. Para isso agrupamos as seqüências de crianças em um único grupo ($n=53$) e comparamos este

grupo com seqüências obtidas de adultos (n=78) durante a infecção primária pelo HIV-1. Igualmente, essa análise feita pelo método NY e mostrou que a maioria dos sítios no gene *nef* estão sujeitos a uma seleção purificadora ($dN/dS < 1$), respectivamente 30% nas crianças e 31% nos adultos. A proporção de sítios sob pressão seletiva positiva ($dN/dS > 1$) foram detectados 9% em crianças e 6% em adultos. Essas proporções não diferem e mostram que o gene *nef* está igualmente sob seleção em adultos e em crianças. Essa análise também indicou que a distribuição das pressões seletivas no gene *nef* é similar entre adultos e crianças. Ademais, uma análise complementar e comparativa entre os subtipos C e B, das seqüências de adultos recém infectados, mostrou relevante diferença entre os códons sob pressão seletiva no gene *nef* (figura 9). Na figura 9 podemos observar que os valores de dN/dS ao longo do gene *nef* é similar entre os dois grupos. Em adição, os valores individuais de dN/dS não diferem estatisticamente entre crianças e adultos (t - $test=1.02$, p - $value=0.435$). Da mesma forma que foi observado anteriormente com a comparação entre os grupos de crianças, a análise feita comparando adultos e crianças mostraram uma concentração de códons sob seleção positiva na região inicial do gene *nef* (figura 9).

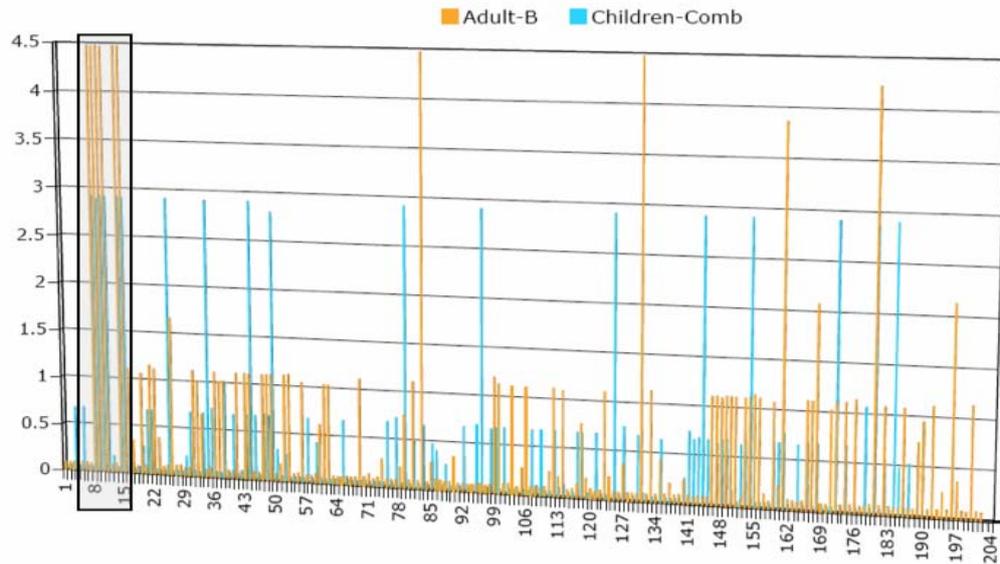


Figura 9. Valores de dN/dS nos códons do gene *nef* do HIV-1 em crianças e adultos recém infectados. A área cinza marca a região onde existe uma concentração de códons sob seleção positiva ($dN/dS > 1$) nos dois grupos.

Apesar das análises feitas individualmente com as seqüências de adultos e crianças indicarem similaridades na posição dos códons sob seleção positiva e no montante global da razão dN/dS . Essas análises não exploram formalmente se existem diferenças códon-por-códon entre duas populações distintas. Em razão disso, fizemos um teste para determinar se existem codons no gene *nef* sob pressões distintas de seleção entre adultos e crianças. Esse teste usa filogenias e indica (baseado no LRT e no p -value) quais codons estão diferentemente selecionados entre os grupos. O resultado dessa análise mostrou que, grande parte dos códons do gene *nef*, estão igualmente sob as mesmas pressões seletivas em adultos e em crianças (figura 10 e Tabela 9). No entanto, seis códons estão diferentemente selecionados entre os grupos analisados (setas na figura 10).

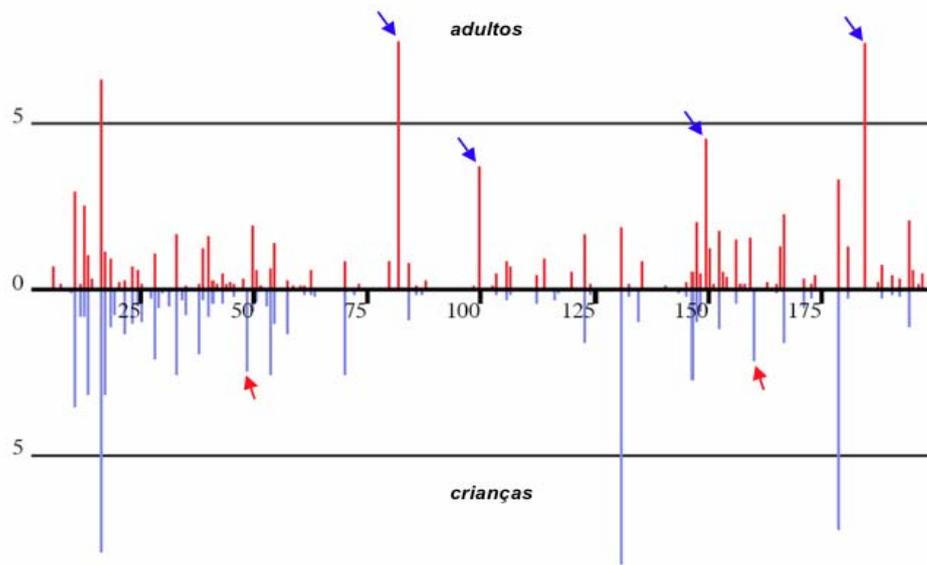


Figura 10. Comparação das taxas seletivas (dN/dS) nos códons do gene *nef* do HIV-1 em crianças e adultos. As barras indicam os valores de dN/dS nos códons do alinhamento de adultos (vermelho) e no alinhamento das crianças (azul). As setas indicam as posições onde o teste LRT (p -value) identificou códons com valores de dN/dS distintos entre as populações (adultos e crianças).

Tabela 9. Análise por Maximum likelihood do gene *nef* do HIV-1 dos grupos de crianças e adultos recém infectados.

Modelo	$-n\ln L$	Parâmetros Estimados	Seleção Positiva códons com $p > 0.999$
<i>Grupos de Crianças A (0 – 3 anos de idade) (n=14, l=612)</i>			
M0 (<i>one ratio</i>)	3127.32011	$\omega_0 = 0.4327$ κ (ts/tv)= 2.11437	<i>None</i>
M1 (<i>nearlyneutral</i>)	3018.35208	$p_0 = 0.65225$ $p_I = 0.34775$ $\omega_0 = 0.00000$ $\omega_1 = 1.00000$ κ (ts/tv)= 2.01824	<i>Not allowed</i>
M2 (<i>selection</i>) *	3004.46263 ¹	$p_0 = 0.63301$ $p_I = 0.30331$ $p_2 = 0.0637$ $\omega_0 = 0.04273$ $\omega_1 = 1.000$ $\omega_2 = 3.43841$ κ (ts/tv)= 2.59750	10, 14, 26, 83, 196
M3 (<i>discrete</i>)	3003.98168	$p_0 = 0.60233$ $p_I = 0.3184$ $p_2 = 0.07928$ $\omega_0 = 0.03146$ $\omega_1 = 0.8169$ $\omega_2 = 3.02336$ κ (ts/tv)= 2.25216	10, 14, 26, 48, 83, 196
<i>Grupos de Crianças B (4 – 6 anos de idade) (n=39, l=612)</i>			
M0 (<i>one ratio</i>)	3351.23052	$\omega_0 = 0.4067$ κ (ts/tv)= 2.92624	<i>None</i>
M1 (<i>nearlyneutral</i>)	3253.6001	$p_0 = 0.64067$ $p_I = 0.35933$ $\omega_0 = 0.00000$ $\omega_1 = 1.00000$ κ (ts/tv)= 2.48571	<i>Not allowed</i>
M2 (<i>selection</i>) *	3244.95198 ²	$p_0 = 0.6270$ $p_I = 0.31451$ $p_2 = 0.0584$ $\omega_0 = 0.0688$ $\omega_1 = 1.000$ $\omega_2 = 2.96394$ κ (ts/tv)= 2.67378	8, 83

M3 (<i>discrete</i>)	3242.81107	$p_0= 0.5582$ $p_1= 0.3452$ $p_2= 0.0965$	8, 10, 14, 83
		$\omega_0=0.0420$ $\omega_1=0.6729$ $\omega_2= 2.3454$	
		κ (ts/tv)= 2.61661	

Grupos de Crianças Combinados (A eB) (n=53, l=612)

M0 (<i>one ratio</i>)	8580.184763	$\omega_0=0.4202$ κ (ts/tv)= 2.20087	None
M1	8073.326331	$p_0= 0.69820$ $p_1= 0.30180$	Not allowed
(<i>nearlyneutral</i>)		$\omega_0=0.00000$ $\omega_1=1.00000$	
		κ (ts/tv)= 2.11126	
M2	7963.27659 ²	$p_0= 0.64267$ $p_1= 0.26784$ $p_2= 0.0895$	8, 10, 11, 14, 26,
(<i>selection</i>)		$\omega_0=0.07113$ $\omega_1=1.000$ $\omega_2= 3.43496$	48, 83, 100, 131,
		κ (ts/tv)= 2.39112	180
M3	7951.84511 ³	$p_0= 0.59467$ $p_1= 0.30683$ $p_2= 0.0985$	8, 10, 11, 14,15,
(<i>discrete</i>) *		$\omega_0=0.04524$ $\omega_1=0.67408$ $\omega_2= 2.9115$	26,37, 48, 83,
		κ (ts/tv)= 2.27345	100, 131, 150,
			161, 180, 196

Grupos de Crianças e Adultos do Subtipo C (n=78, l=612)

M0 (<i>one ratio</i>)	10353.50682	$\omega_0=0.4401$ κ (ts/tv)= 2.31434	None
M1	9726.004577	$p_0= 0.67323$ $p_1= 0.32677$	Not allowed
(<i>nearlyneutral</i>)		$\omega_0=0.06092$ $\omega_1=1.00000$	
		κ (ts/tv)= 2.19269	
M2	9644.09090 ⁴	$p_0= 0.66561$ $p_1= 0.25875$ $p_2= 0.0756$	15, 47, 118, 149,
(<i>selection</i>) *		$\omega_0=0.06974$ $\omega_1=1.000$ $\omega_2= 2.76348$	186, 190, 192
		κ (ts/tv)= 2.41124	
M3	9643.55413	$p_0= 0.66264$ $p_1= 0.25354$ $p_2= 0.0838$	15, 38, 47, 118,
			149, 166, 186,

<i>(discrete)</i>		$\omega_0=0.06727$ $\omega_1=0.92411$ $\omega_2= 2.5922$	190, 192
		κ (ts/tv)= 2.39518	
<i>Grupos de Crianças e Adultos do Subtipo B (n=30, l=612)</i>			
M0 (<i>one ratio</i>)	5513.034008	$\omega_0=0.5139$ κ (ts/tv)= 2.34002	<i>None</i>
M1	5258.007758	$p_0= 0.65566$ $p_I= 0.34434$	<i>Not allowed</i>
<i>(nearlyneutral)</i>		$\omega_0=0.07071$ $\omega_1=1.00000$	
		κ (ts/tv)= 2.14265	
M2	5200.17740 ⁵	$p_0= 0.63352$ $p_I= 0.31150$ $p_2= 0.0549$	9, 10, 11, 12, 14,
<i>(selection) *</i>		$\omega_0=0.07635$ $\omega_1=1.000$ $\omega_2= 4.17993$	15, 83, 131
		κ (ts/tv)= 2.34630	
M3	5199.472421	$p_0= 0.64368$ $p_I= 0.30386$ $p_2= 0.0524$	9, 10, 11, 12, 14,
<i>(discrete)</i>		$\omega_0=0.08384$ $\omega_1=1.11840$ $\omega_2= 4.4797$	15, (...), 192
		κ (ts/tv)= 2.37786	

*) Best-fit model according to the LRT. p =proportion of codons under a defined category of ω (d_N/d_S) value κ (ts/tv)=transitions/transversion ratio. 1) LRT M2vsM3 ratio=0.961914, $p=0.618191$. 2) LRT M2vsM3 ratio=4.281738 $p=0.117553$. 3) LRT M2vsM3 ratio=22.862305 $p=0.000011$, 4) LRT M2vsM3 ratio=1.074219 $p=0.584435$. 5) LRT M2vsM3 ratio=1.409180 $p=0.494311$

A análise comparativa da pressão seletiva positiva exercida entre crianças e adultos mostrou características interessantes sobre a pressão adaptativa no gene *nef* durante a infecção pediátrica. Notavelmente, a intensidade de seleção, medida por dN/dS , foi equivalente nas crianças e nos adultos. A localização dos sítios sobre pressão seletiva positiva foi mapeada por meio da utilização de uma seqüência consenso do gene *nef* e estes foram principalmente localizados em regiões com alta densidade (high density) anticorpos e epítomos CTL (Figura 11 e 12). Estes dados demonstram que o gene *nef* na infecção pediátrica é um importante alvo do sistema imune imaturo. Além disso, mostram que os sítios sob seleção estão dispostos ao longo da seqüência, no entanto, encontram-se concentrados na região inicial do gene *nef* para seqüências do subtipo B, exatamente como foi visualizado pelas metodologias descritas anteriormente. Na análise dos sítios sob seleção do subtipo C verificou-se uma concentração maior na porção final do gene *nef*. A utilização de seqüências do subtipo C foi necessária para excluir a hipótese de que os sítios sob pressão seletiva positiva estariam relacionados o padrão geográfico dos subtipos, com isso, foi realizado o mapeamento de seqüências consenso de crianças e adultos subtipo B em comparação com amostras de adultos com subtipo C do HIV-1. Nesta análise verificou-se que os epítomos sob seleção estão relacionados com as diferenças entre os subtipos e não com as regiões geográficas. Ademais, verificou-se uma diferença relevante nos sítios sob pressão seletiva positiva entre os subtipos B e C no gene *nef*, o que coincide com dados relatados por Pond e colaboradores em outros genes do HIV-1 (figura 11 e 12) (Pond, *et al.* 2006).

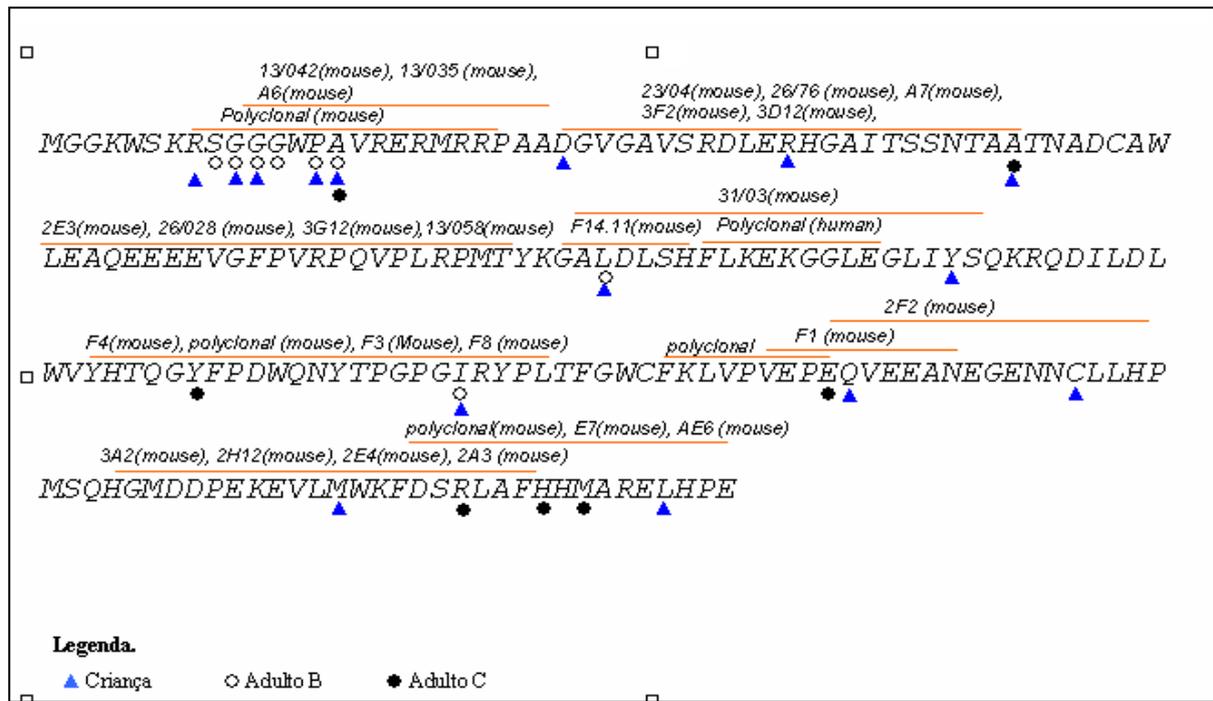


Figura 11. Códons sob seleção positiva do gene *nef* em epítomos Ab (anticorpos) de crianças e adultos recém infectados pelo subtipo B e C do HIV-1.

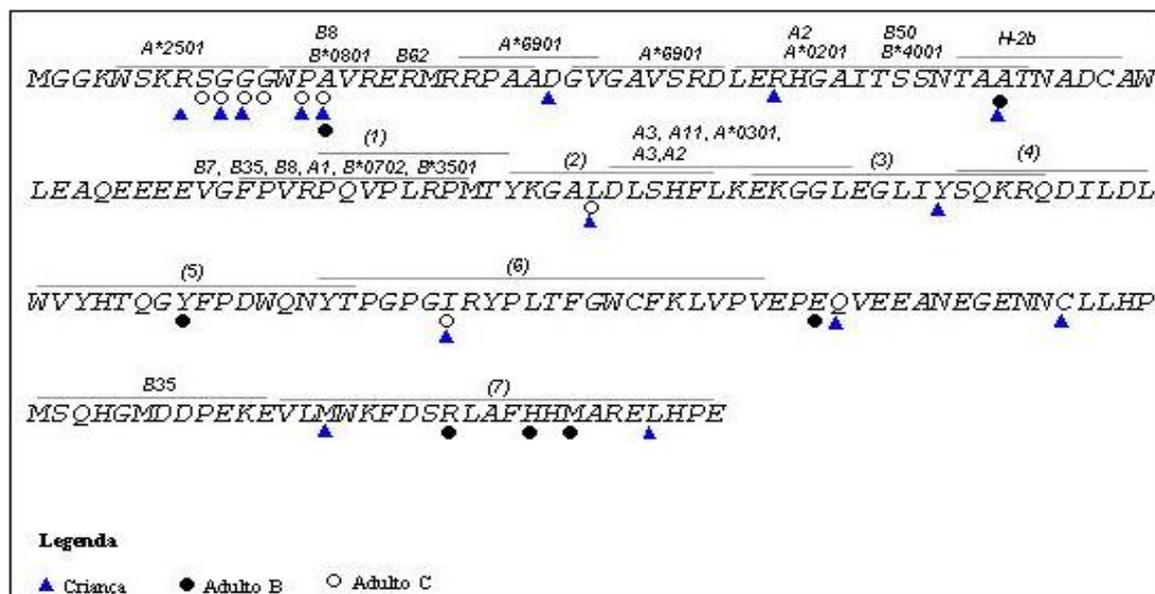


Figura 12. Códons sob seleção positiva do gene *nef* em epitopos CTL (CD8) de crianças e adultos recém infectados pelo subtipo B e C do HIV-1. O números indicam os epitopos CTL (CD8): 1.) B*4201, B*4202, B*8101, B35, B7, B81, B*3501, B51, A*0301, A*1101, A2, A2, B27, Cw4, B3; 2.) B*5801, B57, A2, Cw*0802, Cw8, A*0201, A*0205, A2, B60, B62, Cw*0802, Cw3, Cw8; 3.) A2, B8, A1, B61, B60, B*4002, B84001; 4.) CW*0602, CW*0401, B37, A24, B7, B37, CW*0602, B*2705, B27, B18, CW*0701, CW*0702, CW7, B13, B37, CW*0602, B18, CW*0701, CW7, A3; 5.) CW*0602, CW*0401, B35, A1, B17, B37, B62, B57, A1, A29, B*3701, B*5701, B15, B37, B51, B57, B63, Cw6, A1, B*5801, A29, A*3002, A24, B*5801; 6.) B7, B57, B58, B63, B*0702, B7, B*0702, B*3501, B*4201, B88101, B35, B7, B35, B57, A*2402, A24, A33, B27, A*2402, A24, B35, B18, A2, B*0702, B*1801, B*5301, B*1517, B57, B63, A*6901, A2; 7.) A*0201, A2, A2, B8, B*1503, A1, A1, A24, B51, B52, B35, Cw4, A*6901, A*0201, A*0202, A2, A24, A3, B51, B52.

4.3.3. ANÁLISE DA PRESSÃO SELETIVA NO GENE *POL*

Com o intuito de complementar os dados obtidos da análise do gene *nef*, foi realizada uma análise da pressão seletiva no gene *pol*. No entanto, esta análise foi pouco refinada quando comparada com as análises realizadas no gene *nef*, já que o gene *pol*, é mais conservado. Deste modo, as amostras pertencentes ao subtipo B do HIV-1 foram analisadas para a determinação do padrão de pressão seletiva. Para tanto estimamos os valores de dN e dS sítio-a-sítio usando o método de Muse & Gaut. Foram analisadas 40 sequencias obtidas das crianças dos dois grupos concomitantemente, pois como foi demonstrado por metodologias anteriores do gene *nef*, não houve diferença no perfil de seleção entre estes grupos, portanto, sugere-se que o gene *pol* apresente um perfil semelhante ao que foi demonstrado no gene *nef*. Esta análise foi comparada com as estimativas feitas em 80 sequencias provenientes de adultos recém infectados (obtidas no banco de dados do HIV DataBase-Los Alamos, disponível em: (<http://HIV-web.lanl.gov/content/HIV-db>)). Os resultados desta análise mostram que, os sítios sob pressão seletiva positiva ($dN > dS$) estão presentes ao longo da seqüência e que este padrão de seleção é similar entre crianças e adultos. Além disso, semelhante aos resultados obtidos por meio da análise no gene *nef*, o padrão de pressão seletiva no gene *pol* em adultos e em crianças é predominante impulsionado por seleção purificadora ($dN/dS < 1$). Isso coincide com as características desse gene codificadores de enzimas virais (RT e PR) e grandes partes das mutações observadas são deletérias para as características funcionais e manutenção da estrutura funcional da enzima (figura 13).

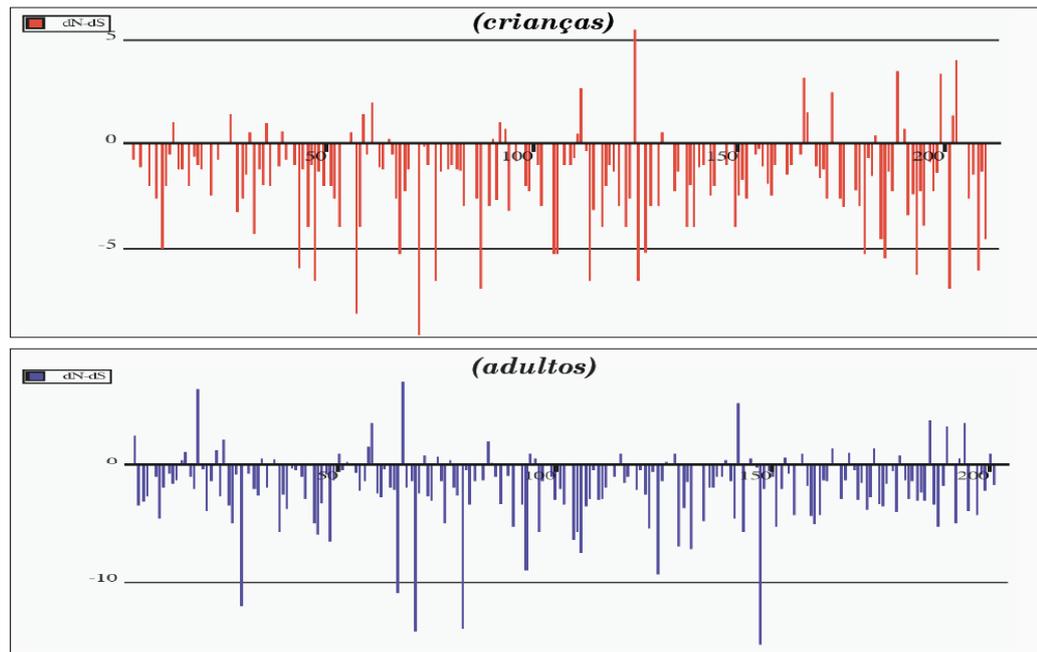


Figura 13. Comparação das taxas seletivas ($dN-dS$) nos códons do gene *pol* do HIV-1 em crianças e adultos. As barras indicam os valores de $dN-dS$ nos códons do alinhamento de crianças (vermelho) e no alinhamento das adultos (azul).



5. DISCUSSÃO

A dinâmica evolutiva do HIV é fortemente influenciada pelas interações do vírus com o hospedeiro (Crandall *et al.*, 1999; Frost *et al.*, 2001; Ross & Rodrigo, 2002; Van Opijnen *et al.*, 2007). Essa dinâmica resulta de uma combinação bastante complexa de fatores inerentes ao vírus e à população infectada, que pode ser analisado através das variações genéticas adaptativas surgidas no HIV-1 decorrente de fatores seletivos (pressão do sistema imunológico ou terapia anti-retroviral) durante a interação do vírus com o hospedeiro e a da manutenção nas populações. Além disso, a virulência do HIV pode ser influenciada por esta adaptação molecular e contínua seleção ao longo do tempo (Moore *et al.*, 2002; Yusim *et al.*, 2002; Anderson & May, 1979; Woolhouse *et al.*, 2002; Arien *et al.*, 2005). A complexa interação entre o vírus e o sistema imunológico e a capacidade deste sistema exercer pressão seletiva, tanto nas crianças quanto nos adultos, por meio de CTL e anticorpos neutralizantes contra o vírus pode gerar padrões específicos de evolução. Em razão do sistema imunológico das crianças não estar completamente desenvolvido, sugere que a pressão seletiva nos genes do HIV é distinta da observada em adultos, o que implica numa força de pressão seletiva diferente entre estes grupos (Buseyne *et al.*, 2006).

Trabalhos anteriores demonstram que, crianças com idade entre 0 e 3 anos, possuem um sistema imunológico pouco desenvolvido, contudo crianças com idade superior a 4 anos são capazes de montar uma resposta eficaz contra o HIV (Sandberg *et al.*, 2003; Pillay & Phillips, 2005). Nesse sentido, estes dados incitaram um estudo comparativo cuja intenção foi analisar a diversidade genética do gene *nef*, em indivíduos com perfil imunológico distintos. Desta forma comparamos a razão entre o número de substituições não-sinônimas por sítio não-sinônimo (dN) e o número de substituições sinônimas por sítio sinônimo (dS) medidos códon a códon no gene *nef* do HIV-1 entre adultos e crianças.

Neste trabalho, fizemos uma abordagem centralizada na observação do regime seletivo em crianças e comparadas com adultos recém infectados, assumindo que entre eles existe uma diferença na maturidade do sistema imunológico (Buseyne *et al.*, 2006). Com isso, tínhamos uma hipótese (nula) de que, as respostas imunológicas, de crianças e adultos, provocariam condições de evolução para o vírus no processo de pressão seletiva gerada por meio da interação entre vírus-hospedeiro durante a infecção recente pelo HIV-1 subtipo B. Deste modo, foi analisado o regime seletivo no gene *nef* em adultos recém infectados e crianças (transmissão materno-infantil) que foram divididas em dois grupos de acordo com a idade (Grupo 1= 0-3 e Grupo 2 = 4-6 anos de idade), visto que, estudos demonstraram que, em contraste com o grupo de crianças

com idades entre 0-3 anos, crianças acima de 4 anos de idade possuem um sistema imunológico capaz de montar uma resposta imunológica eficaz contra o vírus (Sandberg *et al.*, 2003). Com o objetivo de complementação, avaliamos o regime seletivo no gene *pol* entre crianças e adultos. Portanto, para abordar os processos evolutivos do HIV-1 em crianças e adultos recém infectados, foram utilizados dados moleculares (seqüências de DNA viral) e ferramentas de análise filogenética. Isso foi feito através da análise comparativa da intensidade de seleção (medida pelas taxas de substituição sinônimas e não-sinônimas) no gene *nef* e *pol* do HIV por meio da utilização do método de inferência filogenética foi feito por máxima verossimilhança com a utilização de modelos evolutivos de substituição de nucleotídeos (ou aminoácidos) que permitem a correção do número de eventos evolutivos (substituições) num determinado conjunto de dados (seqüências) Além disso, para estimar o número de substituições não sinônimas (dN) e sinônimas (dS) códon a códon foi realizada a análise da pressão seletiva através dos métodos de Nielsen & Yang (NY) (1998) que utiliza maximum likelihood. No entanto, essas análises não exploram formalmente se existem diferenças códon-por-códon entre duas populações distintas. Em razão disso, fizemos um teste para determinar se existem codons no gene *nef* sob pressões distintas de seleção entre adultos e crianças. Esse teste usa filogenias e indica (baseado no LRT e no p -value) quais codons estão diferentemente selecionados entre os grupos. O resultado dessa análise mostrou que, grande parte dos códon do gene *nef* estão igualmente sob as mesmas pressões seletivas em adultos e em crianças (figura 11 e Tabela 9). No entanto, seis códon estão diferentemente selecionados entre os grupos analisados (setas na figura 11).

Os resultados obtidos mostram seleção positiva nos genes *nef* e *pol* tanto em adultos recém infectados, quanto em crianças. A análise comparativa entre crianças (Grupo 1 e Grupo 2) mostrou que, com o aumento do tempo de infecção (crianças com idade mais avançada, Grupo 2), ocorre também um aumento na substituição de nucleotídeos por sítio (taxa evolutiva), observado nos ramos externos das filogenias (figura 7) (Noviello *et al.*, 2007). Esse aumento decorre do maior tempo (divergência) de infecção nos indivíduos e pode ser observado nas filogenias pelo acúmulo de mutações nos ramos externos (apomorfismos). Essas mutações não serão necessariamente transmitidas na população, deste modo, para minimizar o efeito das mutações acumuladas intrapaciente, utilizados seqüências de pacientes recém infectados.

As mutações deletérias normalmente são purgadas da população, de forma que mutações apicais (recentes) tendem a ser levemente deletéria. Em contrapartida, mutações em nós internos (antigas) são altamente adaptativas (Nielsen & Weinreich, 1999). Somente as mutações que favorecem o escape imune é que apresentarão vantagem seletiva na população,

comparada as mutações surgidas nos indivíduos (apomorfismos). Nosso achados, corroboram isso visto que o aumento observado nas taxas evolutivas dos ramos externos (crianças entre 4 e 6 anos), não afetou o regime seletivo entre crianças e adultos, visto que a intensidade de seleção em ambos os grupos foi similar. Em detalhes, ao compararmos os grupos de crianças (Grupo 1 e 2), observamos que, onde há uma distinção entre o tempo de infecção (divergência viral) a razão omega (dN/dS), não foi afetada. Deste modo, a divergência das seqüências decorrentes do acúmulo de mudança de nucleotídeos não implica em seleção por adaptação, portanto, as diferenças nos comprimentos dos ramos observadas neste estudo, não afetaram as estimativas de dN/dS , visto que ambos grupos de crianças tiveram valores parecidos. Estudos mostram que, a variação de dN/dS (que é o que foi estimado neste estudo) não oscila com a divergência (tempo de infecção). No entanto, existe um acúmulo de divergência ao longo do tempo. Isso mostra que, se tomarmos as medidas de dN/dS ao longo do tempo em um paciente, não vão ocorrer variações significativas, exceto no final da infecção onde a variabilidade (formas distintas) do vírus reduz. A comparação da pressão seletiva entre crianças mostrou ainda que, códons sob seleção positiva, foram detectados em toda a extensão dos genes *nef* e *pol*. Após as análises é possível concluir que, a pressão seletiva é ampla em ambos os grupos de crianças (grupo 1 e 2). A identificação de uma área densa em códons positivamente selecionados indica que, existe uma região preferencial do gene *nef* onde a pressão imune é mais intensa, no entanto, no gene *pol* os códons sob pressão seletiva positiva aparecem distribuídos ao longo da seqüência (figura 10 e 13).

Igualmente as comparações feitas, com os genes *nef* e *pol*, entre crianças e adultos mostraram características importantes no processo de seleção do HIV-1. A similaridade na intensidade de seleção e na localização dos códons sob seleção positiva mostra que, não existe diferença (estatisticamente significativa) na diversificação do HIV no nível populacional entre crianças e adultos. Esse achado é muito importante, pois contrasta com os resultados obtidos em ensaios imunológicos que mostram uma redução de intensidade de amplitude no reconhecimento de epitopos virais em crianças. Vale a pena salientar que o gene *nef* é bastante imunogênico de forma que a pressão do sistema imunológico será intensa em todas as fases da infecção, especialmente durante a infecção recente, ao contrario de outros genes (*gag* e *pol*) que apresentam oscilações na resposta ao longo da infecção (principalmente mediada por CTL). Talvez essa peculiaridade do *nef* faz com que ocorra resposta imune intensa durante todas as fases da infecção pelo HIV-1. Possivelmente essa pressão é suficiente para diversificar o gene *nef* rapidamente nos primeiros estágios da infecção e dado a alta imunogenicidade isso ocorre mesmo em crianças onde o sistema imunológico ainda não esta completamente desenvolvido.

Uma observação interessante foi a identificação de uma área (codons 10 a 15), que apresenta uma concentração de codons sob pressão seletiva positiva intensa no gene *nef*. Essa região é densa em epítomos que são reconhecidos grandemente por anticorpos (resposta Humoral) e pouco densa em epítomos reconhecidos por CTL ou CD4-helper, o que contrapõe os achados de Feeney *et al.*, 2005. Portanto, os epítomos expressos pela célula durante a infecção, desencadeiam uma resposta imunológica que promove citotoxicidade celular gerando desta forma, uma alteração na seqüência do HIV. Assim é tentador sugerir que a diversificação do *nef* decorre grandemente da geração de anticorpos para essa região. A diversificação do HIV-1 observada nesse grupo reflete, preferencialmente, as mutações selecionadas antes da transmissão, ao invés daquelas surgidas nas crianças. Nesse contexto, a concentração de códons na região do *nef* rica em epítomos para anticorpos provavelmente reproduz essa característica do grupo de adultos. Em razão da coincidência na localização dos códons entre crianças e adultos, a maior intensidade de seleção positiva na região do *nef* de crianças talvez decorra das pressões seletivas mediadas pelos anticorpos maternos.

Contudo, no gene *pol*, o padrão de pressão seletiva em adultos e em crianças é predominante impulsionado por seleção purificadora. Isso coincide com as características funcionais desse gene (codificador das enzimas RT e PR) e, grandes partes das mutações observadas, são prejudiciais para as características funcionais e manutenção da estrutura da enzima. Além disso, os resultados desta análise mostram que, os sítios sob pressão seletiva positiva estão presentes ao longo da seqüência e que este padrão de seleção é similar entre crianças e adultos (figura 13).

Estes resultados mostraram que, apesar dos valores médios de *dN/dS* não se diferenciarem para um mesmo gene entre crianças e adultos no nível populacional, sistematicamente os valores de alguns códons sob seleção positiva em crianças foram mais elevados. Ademais, na maior parte dos genes a proporção de códons sob seleção positiva em crianças é menor. Isso sugere que o sistema imunológico em crianças responde de forma menos ampla e mais intensa aos epítomos virais do HIV-1.



6. CONCLUSÃO

A análise entre crianças de idades distintas mostrou que com o aumento do tempo de infecção (crianças com idade mais avançada), ocorre também um aumento na substituição de nucleotídeos por sítio (taxa evolutiva), decorrente do aumento de mutações nos ramos externos (apomorfismos). A análise entre os grupos de crianças mostrou que, o acúmulo de apomorfismos não influi no padrão de seleção do HIV-1 na população.

A comparação da pressão seletiva entre crianças mostrou que códons sob seleção positiva foram detectados ao longo dos genes *nef* e *pol*. Entretanto a identificação de uma área densa em códons positivamente selecionados indica que, existe uma região preferencial do *nef* onde a pressão imune é mais intensa, no entanto, no gene *pol* os códons sob pressão seletiva positiva aparecem distribuídos ao longo da seqüência.

A similaridade na intensidade de seleção e na localização dos códons sob seleção positiva, entre crianças e adultos, mostra que não existe diferença, estatisticamente significativa, na diversificação do HIV em crianças e adultos, tanto no gene *nef* quanto no gene *pol*.

Foi identificado, no gene *nef*, uma área (codons 10 a 15), que apresenta uma concentração de códons sob pressão seletiva positiva intensa. Essa região é densa em epitopos que são reconhecidos grandemente por anticorpos (resposta Humoral) e em contrapartida, pouco eficiente em suscitar uma resposta mediada por CTL ou CD4-helper.

No gene *pol*, o padrão de pressão seletiva em adultos e em crianças é predominante impulsionado por seleção purificadora. Além disso, os resultados desta análise mostram que, os sítios sob pressão seletiva positiva ($dN > dS$) estão presentes ao longo da seqüência e que este padrão de seleção é similar entre crianças e adultos



7. REFERÊNCIAS BIBLIOGRÁFICAS

AGOPIAN K, WEI BL, GARCIA JV, GABUZDA D. CD4 and MHC-I downregulation are conserved in primary HIV-1 Nef alleles from brain and lymphoid tissues, but Pak2 activation is highly variable. *Virology*. 358 (1): 119-35, 2006.

ANDERSON, R. M.; MAY, R. M. Population biology of infectious diseases. *Nature*, v. 280, p. 361-455, 1979.

ARGANARAZ ER, SCHINDLER M, KIRCHHOFF F, CORTES MJ, LAMA J. Enhanced CD4 down-modulation by late stage HIV-1 nef alleles is associated with increased Env incorporation and viral replication. *J Biol Chem*. 278 (36): 33912-9, 2003.

ARIEN KK, TROYER RM, GALI Y, COLEBUNDERS RL, ARTS EJ, VANHAM G. Replicative fitness of historical and recent HIV-1 isolates suggests HIV-1 attenuation over time. *AIDS*. 19 (15): 1555-64, 2005.

ARORA VK, FREDERICKSEN BL, GARCIA JV. Nef: agent of cell subversion. *Microbes Infect*. 4(2): 189-99, 2002.

ARTHUR LO, BESS JW JR, SOWDER RC 2ND, BENVENISTE RE, MANN DL, CHERMANN JC, HENDERSON LE. Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. *Science*. 258 (5090):1935-8, 1992.

BALOG K, MINAROVITS J. Nef: a pleiotropic modulator of primate lentivirus infectivity and pathogenesis. *Acta Microbiol Immunol Hung*. 53 (1): 51-75, 2006.

BARRE-SINOUSSE F, CHERMANN JC, REY F, NUGEYRE MT, CHAMARET S, GRUEST J, DAUGUET C, AXLER-BLIN C, VEZINET-BRUN F, ROUZIQUX C, ROZENBAUM W, MONTAGNIER L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 220(4599): 868-71, 1983.

BAZYKIN GA, DUSHOFF J, LEVIN SA, KONDRASHOV AS. Bursts of nonsynonymous substitutions in HIV-1 evolution reveal instances of positive selection at conservative protein sites. *Proc Natl Acad Sci U S A*. 103 (51): 19396-401, 2006.

BETTS, M. R. et al. Human immunodeficiency virus type 1-specific cytotoxic T lymphocyte activity is inversely correlated with HIV type 1 viral load in HIV type 1-infected long-term survivors. *AIDS Res Hum Retrov*, v. 15, p. 1219-1228, 1999.

BLAGOVESHCHENSKAYA AD, THOMAS L, FELICIANGELI SF, HUNG CH, THOMAS G. HIV-1 Nef downregulates MHC-I by a PACS-1- and PI3K-regulated ARF6 endocytic pathway. *Cell*. 111 (6): 853-66, 2002.

BORROW, P. et al. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus, *Nat Med*, v. 3, p. 205-121, 1997.

BOUR S, STREBEL K. The HIV-1 Vpu protein: a multifunctional enhancer of viral particle release. *Microbes Infect*. 5 (11): 1029-39, 2003.

BRANDER C, WALKER BD. T lymphocyte responses in HIV-1 infection: implications for vaccine development. *Curr Opin Immunol*. 11 (4): 451-9, 1999.

BRINDEIRO, R. et al. Sequence diversity of the reverse transcriptase of human immunodeficiency virus type 1 from untreated brazilian individuals. *Antimicrob Agents Chemother*, v. 43, p. 1674-1680, 1999.

BRINDEIRO RM, DIAZ RS, SABINO EC, MORGADO MG, PIRES IL, BRIGIDO L, DANTAS MC, BARREIRA D, TEIXEIRA PR, TANURI A; BRAZILIAN NETWORK FOR DRUG RESISTANCE SURVEILLANCE. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *AIDS*. 17 (7): 1063-9, 2003.

BRITO AM, CASTILHO EA, SZWARCOWALD CL. AIDS and HIV infection in Brazil: a multifaceted epidemic. *Revista da Sociedade Brasileira de Medicina Tropical*. 34 (2): 207-217, 2000.

BUSEYNE F, SCOTT-ALGARA D, CORRE B, PORROT F, MONCHATRE E, BELLAL N, BURGARD M, ROUZIUX C, BLANCHE S, RIVIERE Y. Poor recognition of HIV-1 Nef protein by CD8 T cells from HIV-1-infected children: impact of age. *Virology*. 354 (2): 271-9, 2006.

CAMPOS-LIMA, P. O. et al. HLA-A11 epitope loss isolates of Epstein-Barr virus from a highly A11+ population. *Science*, v. 260, p. 98-100, 1993.

CARRINGTON M., et al. HLA and HIV-1: Heterozygote advantage and B35-Cw4 disadvantage. *Science*, v. 283, p. 1748-1752, 1999.

CHANDWANI R, JORDAN KA, SHACKLETT BL, PAPASAVVAS E, MONTANER LJ, ROSENBERG MG, NIXON DF, SANDBERG JK. Limited magnitude and breadth in the HLA-A2-restricted CD8 T-Cell response to Nef in children with vertically acquired HIV-1 infection. *Scand J Immunol*. 59 (1): 109-14, 2004.

CHARPENTIER C, NORA T, TENAILLON O, CLAVEL F, HANCE AJ. Extensive recombination among human immunodeficiency virus type 1 quasispecies makes an important contribution to viral diversity in individual patients. *J Virol*. 80 (5): 2472-82, 2006.

CHUN TW, STUYVER L, MIZELL SB, EHLER LA, MICAN JA, BASELER M, LLOYD AL, NOWAK MA, FAUCI AS. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A*. 94 (24): 13193-7, 1997.

CHUN, T. W. et al. Re-emergence of HIV after stopping therapy. **Nature**, v. 401, p. 874-875, 1999.

COFFIN J, HAASE A, LEVY JA, MONTAGNIER L, OROSZLAN S, TEICH N, TEMIN H, TOYOSHIMA K, VARMUS H, VOGT P, ET AL. Human immunodeficiency viruses. **Science**. 232 (4751): 697, 1986.

COFFIN JM. HIV viral dynamics. **AIDS**. 3: S75-84, 1996.

COFFIN, J. M. HIV population dynamics in vivo: implications for genetic variation with pathogenesis and therapy. **Science**, v. 267, p. 483-489, 1995.

COHEN EA, SUBBRAMANIAN RA, GOTTLINGER HG. Role of auxiliary proteins in retroviral morphogenesis. **Curr Top Microbiol Immunol**. 214: 219-35, 1996.

COLLINS, K.L. et al. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. **Nature**, v. 391, p. 397-401, 1998.

COLSON ER, HORWITZ RI, BIA FJ, VISCOLI CM. Zidovudine (AZT) for treatment of patients infected with human immunodeficiency virus type 1. An evaluation of effectiveness in clinical practice. **Arch Intern Med**. 151(4):709-13, 1991.

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.

DAS SR, JAMEEL S. Biology of the HIV Nef protein. **Indian J Med Res**.121 (4): 315-32, 2005.

DE CLERCQ E. New perspectives for the treatment of HIV infections. *Verh K Acad Geneesk Belg.* 60 (1): 13-45, 1998.

DE JONG JJ, DE RONDE A, KEULEN W, TERSMETTE M, GOUDSMIT J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. *J Virol.* 66 (11): 6777-80, 1992.

DE SA FILHO DJ, SUCUPIRA MC, CASEIRO MM, SABINO EC, DIAZ RS, JANINI LM. Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS. Res Hum Retroviruses.* 22 (1): 1-13, 2006.

DE S LEAL E, HOLMES EC, ZANOTTO PM. Distinct patterns of natural selection in the reverse transcriptase gene of HIV-1 in the presence and absence of antiretroviral therapy. *Virology.* 325 (2): 181-91, 2004.

DEACON NJ, TSYKIN A, SOLOMON A, SMITH K, LUDFORD-MENTING M, HOOKER DJ, MCPHEE DA, GREENWAY AL, ELLETT A, CHATFIELD C, LAWSON VA, CROWE S, MAERZ A, SONZA S, LEARMONT J, SULLIVAN JS, CUNNINGHAM A, DWYER D, DOWTON D, MILLS J. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science*, v. 10, p. 988-991, 1995.

DEAN M, CARRINGTON M, O'BRIEN SJ. Balanced polymorphism selected by genetic versus infectious human disease. *Annu Rev Genomics Hum Genet.* 3: 263-92, 2002.

DENG H, LIU R, ELLMEIER W, CHOE S, UNUTMAZ D, BURKHART M, DI MARZIO P, MARMON S, SUTTON RE, HILL CM, DAVIS CB, PEIPER SC, SCHALL TJ, LITTMAN DR, LANDAU NR. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 381 (6584): 661-6, 1996.

FARES MA, TRAVERS SA. A novel method for detecting intramolecular coevolution: adding a further dimension to selective constraints analyses. *Genetics*. 173 (1): 9-23, 2006.

FEENEY ME. HIV and children: the developing immune system fights back. *West Indian Med J*. 53 (5): 359-62, 2004.

FEENEY ME, ROOSEVELT KA, TANG Y, PFAFFEROTT KJ, MCINTOSH K, BURCHETT SK, MAO C, WALKER BD, GOULDER PJ. Comprehensive screening reveals strong and broadly directed human immunodeficiency virus type 1-specific CD8 responses in perinatally infected children. *J Virol*. 77 (13): 7492-501, 2003.

FELSENSTEIN J. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet*. 25 (5): 471-92, 1973.

FINZI D, BLANKSON J, SILICIANO JD, MARGOLICK JB, CHADWICK K, PIERSON T, SMITH K, LISZIEWICZ J, LORI F, FLEXNER C, QUINN TC, CHAISSON RE, ROSENBERG E, WALKER B, GANGE S, GALLANT J, SILICIANO RF. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med*, v. 5, p. 512-517, 1999.

FRANCIS DP, CURRAN JW, ESSEX M. Epidemic acquired immune deficiency syndrome: epidemiologic evidence for a transmissible agent. *J Natl Cancer Inst*. 71(1): 1-4, 1983.

FRANKEL AD, YOUNG JA. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem*. 67: 1-25, 1998.

FROST SD, WRIN T, SMITH DM, KOSAKOVSKY POND SL, LIU Y, PAXINOS E, CHAPPEY C, GALOVICH J, BEAUCHAINE J, PETROPOULOS CJ, LITTLE SJ,

RICHMAN DD. Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection. *Proc Natl Acad Sci U S A*. 102 (51): 18514-9, 2005.

FROST SD, DUMAURIER MJ, WAIN-HOBSON S, BROWN AJ. Genetic drift and within-host metapopulation dynamics of HIV-1 infection. *Proc Natl Acad Sci U S A*. 98 (12): 6975-80, 2001.

GALLO RC, SALAHUDDIN SZ, POPOVIC M, SHEARER GM, KAPLAN M, HAYNES BF, PALKER TJ, REDFIELD R, OLESKE J, SAFAI B, ET AL. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*. 224 (4648): 500-3, 1984.

GALLO RC MONTAGNIER L. The Discovery of HIV as the Cause of AIDS. *N Engl J Med*. 349 (24): 2283-5, 2003.

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*. 70 (10): 7013-29, 1996.

GAO F, BAILES E, ROBERTSON DL, CHEN Y, RODENBURG CM, MICHAEL SF, CUMMINS LB, ARTHUR LO, PEETERS M, SHAW GM, SHARP PM, HAHN BH. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*. 397 (6718): 436-41, 1999.

GAO X, NELSON GW, KARACKI P, MARTIN MP, PHAIR J, KASLOW R, GOEDERT JJ, BUCHBINDER S, HOOTS K, VLAHOV D, O'BRIEN SJ, CARRINGTON M. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N Engl J Med*. 344 (22): 1668-75, 2001.

GASCHEN, B. et al. Diversity considerations in HIV-1 vaccine selection. **Science**, v. 296, p. 2354-2360, 2002.

GEFFIN R, HUTTO C, ANDREW C, SCOTT GB. A longitudinal assessment of autologous neutralizing antibodies in children perinatally infected with human immunodeficiency virus type 1. **Virology**. 310 (2): 207-15, 2003.

CRANDALL KA, VASCO DA, POSADA D, IMAMICHI H. Advances in understanding the evolution of HIV. **AIDS**. 13A: S39-47, 1999.

CHAN DC, KIM PS. HIV entry and its inhibition. **Cell**. 93 (5): 681-4, 1998.

CATANZARO D, PESENTI R, MILINKOVITCH MC. A non-linear optimization procedure to estimate distances and instantaneous substitution rate matrices under the GTR model. **Bioinformatics**. 22 (6): 708 – 15, 2006.

CHENG-MAYER C., SETO D., TATENO M., LEVY J.A., Biologic Features of HIV-1 that correlated with the virulence in the host. **Science**. 240: 80-82, 1998.

CHOE H. Chemokine receptors in HIV-1 and SIV infection. **Arch Pharm Res**. 21 (6): 634-9, 1998.

CLERICI M, et al. Cell-mediated immune response to human immunodeficiency virus (HIV) type 1 in seronegative homosexual men with recent sexual exposure to HIV-1. **J Infect Dis**, v. 65, p. 1012-9, 1992.

DEEKS SG, BARDITCH-CROVO P, COLLIER A, SMITH A, MILLER M, MCGOWAN I, COAKLEY DF. Hydroxyurea does not enhance the anti-HIV activity of low-dose tenofovir disoproxil fumarate. **J Acquir Immune Defic Syndr**. 28 (4): 336-9, 2001.

DOMINGO P, MATIAS-GUIU X, PUJOL RM, FRANCA E, LAGARDA E, SAMBEAT MA, VAZQUEZ G. Subcutaneous adipocyte apoptosis in HIV-1 protease inhibitor-associated lipodystrophy. **AIDS**. 13 (16): 2261-7, 1999.

FLEURY S, RIZZARDI GP, CHAPUIS A, TAMBUSI G, KNABENHANS C, SIMEONI E, MEUWLY JY, CORPATAUX JM, LAZZARIN A, MIEDEMA F, PANTALEO G. Long-term kinetics of T cell production in HIV-infected subjects treated with highly active antiretroviral therapy. **Proc Natl Acad Sci U S A**. 97 (10): 5393-8, 2000.

FU, Y.-X. Estimating mutation rate and generation time from longitudinal samples of DNA sequences. **Mol Biol Evol**, v.18, p. 620-626, 2001.

FURCI, L et al. Human immunodeficiency virus type 1 glycoprotein 120-specific T lymphocytes provide intermolecular help for anti-CD4 autoantibody production in exposed uninfected subjects. **AIDS Res Hum Retrovir**, v. 20, p. 1461-1469, 1997.

GIESE SI, WOERZ I, HOMANN S, TIBRONI N, GEYER M, FACKLER OT. Specific and distinct determinants mediate membrane binding and lipid raft incorporation of HIV-1 (SF2) Nef. **Virology**. 355 (2): 175-91, 2006.

GOFF SP. Death by deamination: a novel host restriction system for HIV-1. **Cell**. 114 (3): 281-3, 2003.

GOLDMAN N, YANG Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. **Mol Biol Evol**. 11 (5): 725-36, 1994.

GOMEZ-CANO M, RUBIO A, PUIG T, PEREZ-OLMEDA M, RUIZ L, SORIANO V, PINEDA JA, ZAMORA L, XAUS N, CLOTET B, LEAL M. Prevalence of genotypic resistance to nucleoside analogues in antiretroviral-naive and antiretroviral-experienced HIV-infected patients in Spain. **AIDS**. 12 (9): 1015-20, 1998.

GONDA MA, WONG-STAAAL F, GALLO RC, CLEMENTS JE, NARAYAN O, GILDEN RV. Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus. **Science**. 227 (4683): 173-7, 1985.

GOULDER, P. et al. Co-evolution of human immunodeficiency virus and cytotoxic T-lymphocyte responses. **Immunol Rev**, v. 159, p. 17-29, 1997.

GOTTLIEB MS, SCHROFF R, SCHANKER HM, WEISMAN JD, FAN PT, WOLF RA, SAXON A. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. **N Engl J Med**. 305: 1425-31, 1981.

GOTTLINGER HG. The HIV-1 assembly machine. *AIDS*. 5: S13-20, 2001.

GOULDER P, PRICE D, NOWAK M, ROWLAND-JONES S, PHILLIPS R, MCMICHAEL A. Co-evolution of human immunodeficiency virus and cytotoxic T-lymphocyte responses. *Immunol Rev*. 159: 17-29, 1997.

GRANT A.D., De COCK K. M. ABC of AIDS. HIV infection and AIDS in the developing world. *BMJ*. 322 (7300):1475-8, 2001.

GRENFELL BT, PYBUS OG, GOG JR, WOOD JL, DALY JM, MUMFORD JA, HOLMES EC. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*. 303 (5656): 327-32, 2004.

GREENOUGH, T. C. et al. Human immunodeficiency virus type 1-specific cytotoxic T lymphocytes (CTL), virus load, and CD4 T cell loss: evidence supporting a protective role for CTL in vivo. *J Infect Dis*, v. 76, p. 118-25, 1997.

HAASE AT, HENRY K, ZUPANCIC M, SEDGEWICK G, FAUST RA, MELROE H, CAVERT W, GEBHARD K, STASKUS K, ZHANG ZHI-QIANG, DAILEY PJ, JR BALFOUR HH, ERICE A, PERELSON AS. Quantitative Image Analysis of HIV-1 Infection in Lymphoid Tissue. *Science*. 274 (5289) 985 – 989, 1996.

HAHN BH, SHAW GM, ARYA SK, POPOVIC M, GALLO RC, WONG-STAAAL F. Molecular cloning and characterization of the HTLV-III virus associated with AIDS. *Nature*. 312 (5990): 166-9, 1984.

HAHN BH, SHAW GM, DE COCK KM, SHARP PM. AIDS as a zoonosis: scientific and public health implications. *Science*. 287 (5453): 607-14, 2000.

HAY, C. M. et al. Lack of viral escape and defective in vivo activation of human immunodeficiency virus type 1-specific cytotoxic T lymphocytes in rapidly progressive infection. **J Virol**, v. 73, p. 5509-5519, 1999.

HENDERSON HI, HOPE TJ. The temperature arrested intermediate of virus-cell fusion is a functional step in HIV infection. **Virol J.** 3: 36, 2006.

HILL, A. V. et al. Genetic analysis of host-parasite co evolution in human malaria. **Philos Trans R Soc Lond**, v. 352, p. 1317-1325, 1997.

HILL AV. The immunogenetics of human infectious diseases. **Annu Rev Immunol.** 16: 593-617, 1998.

HILL M, TACHEDJIAN G, MAK J. The packaging and maturation of the HIV-1 Pol proteins. **Curr HIV Res.** 3 (1): 73-85, 2005.

HOFFMAN NG, SEILLIER-MOISEWITSCH F, AHN J, WALKER JM, SWANSTROM R. Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. **J Virol.** 76 (8): 3852-64, 2002.

HOLMES, E. C. et al. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. **Proc Natl Acad Sci USA**, v. 89, p. 4835-4839, 1992.

HOPE T.J. Viral RNA export. **Chem Biol.** 4 (5): 335-44, 1997.

HUANG CC, TANG M, ZHANG MY, MAJEED S, MONTABANA E, STANFIELD RL, DIMITROV DS, KORBER B, SODROSKI J, WILSON IA, WYATT R, KWONG PD. Structure of a V3-containing HIV-1 gp120 core. **Science.** 310 (5750): 1025-8, 2005.

INVERNIZZI CF, XIE B, RICHARD S, WAINBERG MA. PRMT6 diminishes HIV-1 Rev binding to and export of viral RNA. *Retrovirology*. 18; 3: 93, 2006.

JANEWAY CA, CHERVONSKY AV, SANT'ANGELO D. T-cell receptors: is the repertoire inherently MHC-specific? *Curr Biol*. 7 (5): R299-300, 1997.

JUNG A, MAIER R, VARTANIAN JP, BOCHAROV G, JUNG V, FISCHER U, MEESE E, WAIN-HOBSON S, MEYERHANS A. Multiply infected spleen cells in HIV patients. *Nature*. 418 (6894): 144, 2002.

KACANI, L. et al. Dendritic cells transmit human immunodeficiency virus type 1 to monocytes and monocyte-derived macrophages. *J Virol*, v. 72, p. 6671-6677, 1998.

KALAMS, S. A. et al. Longitudinal analysis of T cell receptor (TCR) gene usage by human immunodeficiency virus 1 envelope-specific cytotoxic T lymphocyte clones reveals a limited TCR repertoire. *J Exp Med*, v. 179, p. 1261-1271, 1994.

KALAMS SA, BUCHBINDER SP, ROSENBERG ES, BILLINGSLEY JM, COLBERT DS, JONES NG, SHEA AK, TROCHA AK, WALKER BD. Association between virus-specific cytotoxic T-lymphocyte and helper responses in human immunodeficiency virus type 1 infection. *J Virol*. 73 (8): 6715-20, 1999.

KARLSSON AC, IVERSEN AK, CHAPMAN JM, DE OLIVIERA T, SPOTTS G, MCMICHAEL AJ, DAVENPORT MP, HECHT FM, NIXON DF. Sequential Broadening of CTL Responses in Early HIV-1 Infection Is Associated with Viral Escape. *PLoS ONE*. 21; 2: e 225, 2007.

KELLEHER, A. D. et al. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27 restricted CTL responses. *J Exp Med*, v. 193, p. 375-385, 2001.

KLEIN, M. R.; MIEDEMA, F. Long-term survivors of HIV-1 infection. **Trends Microbiol**, v. 10, p. 386-391, 1995.

KOSALARAKSA, P. et al. Comparative fitness of multi-dideoxynucleoside-resistant human immunodeficiency virus type 1 (HIV-1) in an In vitro competitive HIV-1 replication assay. **J Virol**, v. 73, p. 5356-5363, 1999.

KOZAL MJ, SHAH N, SHEN N, YANG R, FUCINI R, MERIGAN TC, RICHMAN DD, MORRIS D, HUBBELL E, CHEE M, GINGERAS TR. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. **Nat Med**. 2 (7): 753-9, 1996.

LEAL, E. S.; ZANOTTO, P. M. A. Viral diseases and human evolution. **Mem Inst Oswaldo Cruz**, v. 95, Suppl 1, p. 193-200, 2000.

LEAL, E. S.; HOLMES. E. C.; ZANOTTO, P. M. A. Distinct patterns of natural selection in the reverse transcriptase gene of HIV-1 in the presence and absence of antiretroviral therapy. **Virology**. , v. 325, p. 181-191, 2004.

LEIGH-BROWN, A. J.; RICHMAN, D. D. HIV-1: Gambling on the evolution of drug resistance? **Nat Med**, v. 3, p. 268-271, 1997.

LEIGH BROWN, A. J. et al. Analysis of HIV-1 *env* gene sequences reveals evidence for a low effective number in the viral population. **Proc Natl Acad Sci USA**, v. 94, p. 1862-1865, 1997.

KOUP, R. A. et al. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. **J Virol**, v. 68, p. 4650-4655, 1994.

LAL RB, CHAKRABARTI S, YANG C. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy & vaccine development. *Indian J Med Res.* 121 (4): 287-314, 2005.

LARDER BA, KEMP SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science.* 246 (4934): 1155-8, 1989.

LEARMONT J, COOK L, DUNCKLEY H, SULLIVAN JS. Update on long-term symptomless HIV type 1 infection in recipients of blood products from a single donor. *AIDS Res Hum Retroviruses.* 11 (1): 1, 1995.

LEMEY P, RAMBAUT A, PYBUS OG. HIV evolutionary dynamics within and among hosts. *AIDS Rev.* 8 (3): 125-40, 2006.

LEMEY P, KOSAKOVSKY POND SL, DRUMMOND AJ, PYBUS OG, SHAPIRO B, BARROSO H, TAVEIRA N, RAMBAUT A. Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics. *PLoS Comput Biol.* 3 (2): e29, 2007.

LEVY JA, HOFFMAN AD, KRAMER SM, LANDIS JA, SHIMABUKURO JM, OSHIRO LS. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science.* 225 (4664): 840-2, 1984.

LI YP. Protein B23 is an important human factor for the nucleolar localization of the human immunodeficiency virus protein Tat. *J Virol.* 71 (5): 4098-102, 1997.

LIDSAY, J. et al. The immune response to viruses. In: KNIPE, D. M.; HOWLEY, P. R. (Ed.) *Fields Virology.* 4. ed. Philadelphia: Lippincot Willians & Wilkins. p.285-320, 2001.

LINDWASSER OW, CHAUDHURI R, BONIFACINO JS. Mechanisms of CD4 downregulation by the Nef and Vpu proteins of primate immunodeficiency viruses. *Curr Mol Med.* 7 (2): 171-84, 2007.

LINNEMANN T, ZHENG YH, MANDIC R, PETERLIN BM. Interaction between Nef and phosphatidylinositol-3-kinase leads to activation of p21-activated kinase and increased production of HIV. *Virology.* 294 (2): 246-55, 2002.

LITMAN GW, ANDERSON MK, RAST JP. Evolution of antigen binding receptors. *Annu Rev Immunol.* 17: 109-47, 1999.

LOUWAGIE J, DELWART EL, MULLINS JI, MCCUTCHAN FE, EDDY G, BURKE DS. Genetic analysis of HIV-1 isolates from Brazil reveals presence of two distinct genetic subtypes. *AIDS Res Hum Retroviruses.* 10 (5): 561-7, 1994.

LUZURIAGA, K., WU, H., AND MCMANUS, M. et al. (1999), Dynamics of HIV-1 Replication in Vertically-Infected Infants, *Journal of Virology*, 73, 362-367.

MACNEIL A, SANKALE JL, MELONI ST, SARR AD, MBOUP S, KANKI P. GENOMIC SITES OF human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. *J Virol.* 80 (15): 7316-21, 2006.

MANNANO, F.; PETIT, C.; CLAVEL, F. Resistance-associated loss of viral fitness in human analysis of protease and gag coevolution in protease inhibitor-treated patients. *J Virol*, v. 72, p. 7632-7637, 1998.

MANN J. The future of the global AIDS movement. *Harv AIDS Rev.* 18-21, 1999.

MANSKY LM, TEMIN HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol.* 69 (8): 5087-94, 1995.

MARKOSYAN RM, COHEN FS, MELIKYAN GB. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol Biol Cell.* 14 (3): 926-38, 2003.

MARMOR M., HERTZMARK K., THOMAS S. M., HALKITIS P. N., VOGLER M. Resistance to HIV Infection. *J Urban Health.* 83 (1): 5-17, 2006.

MARTINEZ-PICADO J, SAVARA AV, SHI L, SUTTON L, D'AQUILA RT. Fitness of human immunodeficiency virus type 1 protease inhibitor-selected single mutants. *Virology.* 275 (2): 318-22, 2000.

MAZZOLI, S. et al. HIV-specific mucosal and cellular immunity in HIV-seronegative partners of HIV-seropositive individuals. *Nat Med*, v. 3, p. 1250-1257, 1997.

MCCUTCHAN FE, VIPUTTIGUL K, DE SOUZA MS, CARR JK, MARKOWITZ LE, BUAPUNTH P, MCNEIL JG, ROBB ML, NITAYAPHAN S, BIRX DL, BROWN AE. Diversity of envelope glycoprotein from human immunodeficiency virus type 1 of recent seroconverters in Thailand. *AIDS Res Hum Retroviruses.* 16 (8): 801-5, 2000.

McKINNEY, D. M. et al. The antiviral activity of HIV-specific CD8+ CTL clones is limited by elimination due to encounter with HIV-infected targets. *J Immunol*, v. 163, p. 861-867, 1999.

MEEK T.D. Inhibitors of HIV-1 protease. *J Enzyme Inhib.* 6 (1): 65-98, 1992.

MELLORS JW, RINALDO CR JR, GUPTA P, WHITE RM, TODD JA, KINGSLEY LA. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science*. 272 (5265): 1167-70, 1996.

MELIKYAN GB, MARKOSYAN RM, HEMMATI H, DELMEDICO MK, LAMBERT M, COHEN FS. Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J Cell Biol*. 151 (2): 413-23, 2000.

MOLLA, A. et al. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat Med*, v. 2, p. 760-766, 1996.

MONTAGNIER L, CHERMANN JC, BARRE-SINOUSSE F, KLATZMANN D, WAIN-HOBSON S, ALIZON M, CLAVEL F, BRUN-VEZINET F, VILMER E, ROUZIOUX C, ET AL. Lymphadenopathy associated virus and its etiological role in AIDS. *Princess Takamatsu Symp*. 15: 319-31, 1984.

MOORE JP, MCKEATING JA, NORTON WA, SATTENTAU QJ. Direct measurement of soluble CD4 binding to human immunodeficiency virus type 1 virions: gp120 dissociation and its implications for virus-cell binding and fusion reactions and their neutralization by soluble CD4. *J Virol*. 65 (3): 1133-40, 1991.

MOORE, C. B. et al. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science*, v. 296, p. 1439-1442, 2002.

MORRIS A, MARSDEN M, HALCROW K, HUGHES ES, BRETTE RP, BELL JE, SIMMONDS P. Mosaic Structure of the Human Immunodeficiency Virus Type 1 Genome Infecting Lymphoid Cells and the Brain: Evidence for Frequent In Vivo Recombination Events in the Evolution of Regional Populations. *Journal of Virology*., 73 (10) 8720-8731, 1999.

MUJAWAR Z, ROSE H, MORROW MP, PUSHKARSKY T, DUBROVSKY L, MUKHAMEDOVA N, FU Y, DART A, ORENSTEIN JM, BOBRY SHEV YV, BUKRINSKY M, SVIRIDOV D. Human immunodeficiency virus impairs reverse cholesterol transport from macrophages. *PLoS Biol.* 4 (11): e365, 2006.

MULLINS JI, JENSEN MA. Evolutionary dynamics of HIV-1 and the control of AIDS. *Curr Top Microbiol Immunol.* 299: 171-92, 2006.

MUSE SV, GAUT BS. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Molecular Biology and Evolution.* 11: 715-724, 1994.

MUSEY L, HUGHES J, SCHACKER T, SHEA T, COREY L, MCEL RATH MJ. Cytotoxic-T-cell responses, viral load, and disease progression in early human immunodeficiency virus type 1 infection. *N Engl J Med.* 337 (18): 1267-74, 1997.

MYERS G. AIDS origin. *Science.* 256 (5063): 1502, 1992.

MYERS HF, SATZ P, MILLER BE, BING EG, EVANS G, RICHARDSON MA, FORNEY D, MORGENSTERN H, SAXTON E, D'ELIA L, LONGSHORE D, MENA I. The African-American Health Project (AAHP): study overview and select findings on high risk behaviors and psychiatric disorders in African American men. *Ethn Health.* 2 (3): 183-96, 1997.

NÁJERA, I. et al. Natural occurrence of drug resistance mutations in the reverse transcriptase of human immunodeficiency virus type 1 isolate. *Aids Res Hum Retrovir.* v. 10, p. 1479-1488, 1994.

NAJERA I, HOLGUIN A, QUINONES-MATEU ME, MUNOZ-FERNANDEZ MA, NAJERA R, LOPEZ-GALINDEZ C, DOMINGO E. Pol gene quasispecies of human

immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. *J Virol.* 69 (1): 23-31, 1995.

NAKAMURA GR, BYRN R, WILKES DM, FOX JA, HOBBS MR, HASTINGS R, WESSLING HC, NORCROSS MA, FENDLY BM, BERMAN PW. Strain specificity and binding affinity requirements of neutralizing monoclonal antibodies to the C4 domain of gp120 from human immunodeficiency virus type 1. *J Virol.* 67 (10): 6179-91, 1993.

NEI M., GOJOBORI T. Simple Methods for Estimating the Numbers of Synonymous and Nonsynonymous Nucleotide Substitutions. *Mol. Biol. Evol.* 3 (5) 418-426, 1986.

NIELSEN R, YANG Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics.* 148 (3): 929-36, 1998.

NOVIELLO CM, POND SL, LEWIS MJ, RICHMAN DD, PILLAI SK, YANG OO, LITTLE SJ, SMITH DM, GUATELLI JC. Maintenance of Nef-mediated modulation of major histocompatibility complex class I and CD4 after sexual transmission of human immunodeficiency virus type 1. *J Virol.* 81 (9): 4776-86, 2007.

NOVITSKY, V. A. et al. Molecular cloning and phylogenetic analysis of human immunodeficiency virus type 1 subtype C: a set of 23 full-length clones from Botswana. *J Virol*, v. 73, p. 4427-4432, 1999.

O'NEILL E, BAUGH LL, NOVITSKY VA, ESSEX ME, GARCIA JV. Intra- and intersubtype alternative Pak2-activating structural motifs of human immunodeficiency virus type 1 Nef. *J Virol.* 80 (17): 8824-9, 2006.

PALMER S, VUITTON D, GONZALES MJ, BASSIGNOT A, SHAFER RW. Reverse transcriptase and protease sequence evolution in two HIV-1-infected couples. *J Acquir Immune Defic Syndr.* 31 (3): 285-90, 2002.

PANTALEO G, GRAZIOSI C, FAUCI AS. Virologic and immunologic events in primary HIV infection. *Springer Semin Immunopathol.* 18 (3): 257-66, 1997.

PEÇANHA E.P., ANTUNES O.A.C., TANURI A. Pharmacological strategies for anti-HIV therapy. *Quim Nova.* 25 (6B): 1108-1116, 2002.

PERELSON, A. S. et al. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, v. 271, p. 1582-1586, 1996.

PETERLIN BM, TRONO D. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat Rev Immunol.* 3 (2): 97-107, 2003.

PILLAY T, PHILLIPS RE. Adaptive evolution in perinatal HIV-1. *Best Pract Res Clin Obstet Gynaecol.* 19 (2): 211-29, 2005.

PHILLIPS RE, ROWLAND-JONES S, NIXON DF, GOTCH FM, EDWARDS JP, OGUNLESI AO, ELVIN JG, ROTHBARD JA, BANGHAM CR, RIZZA CR, ET AL. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature.* 354 (6353): 453-9, 1991.

PHILLIPS, A. J. HIV dynamics: lessons from the use of antiretrovirals. In: DALGLEISH, A.; WEIS, R. (Ed.). *HIV and the new viruses.* 50-74, 1999.

PINTO M. E., STRUCHINER C. J. HIV-1 diversity: a tool for studying the pandemic. *Cad Saude Publica.* 22(3): 473-84, 2006.

PIROOZMAND A, KHAMRSRI B, FUJITA M, ADACHI A, UCHIYAMA T. Morphological study on biologically distinct vpx/vpr mutants of HIV-2. *J Med Invest.* 53 (3-4): 271-6, 2006.

POND S. L. K., FROST S. D. W. A Genetic Algorithm Approach to Detecting Lineage-Specific Variation in Selection Pressure. *Mol. Biol. Evol.* 22(3):478–485. 2005

PRICE, D. A. et al. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc Natl Acad Sci USA*, v. 94, p. 1890-1895, 1997.

PYBUS, O. G.; RAMBAUT, A.; HARVEY, P. H. An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, v. 155, p. 1429-1437, 2000.

PYBUS, O.G.; RAMBAUT, A. GENIE. Estimating demographic history from molecular phylogenies. *Bioinformatics*, v. 18, p. 1404-1405, 2002.

RAMBAUT A, ROBERTSON DL, PYBUS OG, PEETERS M, HOLMES EC. Human immunodeficiency virus. Phylogeny and the origin of HIV-1. *Nature*. 410 (6832): 1047-8, 2001.

RAMBAUT, A. et al. The causes and consequences of HIV evolution. *Nat Genet*, v. 5, p. 52-61, 2004.

RATNER L, GALLO RC, WONG-STAAAL F. HTLV-III, LAV, ARV are variants of same AIDS virus. *Nature*. 313 (6004): 636-7, 1985.

REEVES JD, DOMS RW. Human immunodeficiency virus type 2. *J Gen Virol*. 83 (6): 1253-65, 2002.

REQUEJO HI. Worldwide molecular epidemiology of HIV. *Rev Saude Publica*. 40(2):331-45, 2006.

RIBEIRO RM, BONHOEFFER S. A stochastic model for primary HIV infection: optimal timing of therapy. *AIDS*. 13 (3): 351-7, 1999.

RIBEIRO, R. M.; BONHOEFFER, S. Production of resistant HIV mutants during antiretroviral therapy. **Proc Natl Acad Sci USA**, v. 99, p. 15572-15577, 2002.

RIBEIRO, R. M.; MOHRI, H.; HO, D. D.; PERELSON, A.L. S. 2002. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? **Proc Natl Acad Sci USA**, v. 97, p. 7681-7686, 2000.

RINALDO C, HUANG XL, FAN ZF, DING M, BELTZ L, LOGAR A, PANICALI D, MAZZARA G, LIEBMANN J, COTTRILL M, ET AL. High levels of anti-human immunodeficiency virus type 1 (HIV-1) memory cytotoxic T-lymphocyte activity and low viral load are associated with lack of disease in HIV-1-infected long-term nonprogressors. **J Virol**. 69 (9): 5838-42, 1995.

RIVIERE Y, MCCHESENEY MB, PORROT F, TANNEAU-SALVADORI F, SANSONETTI P, LOPEZ O, PIALOUX G, FEUILLIE V, MOLLEREAU M, CHAMARET S, ET AL. Gag-specific cytotoxic responses to HIV type 1 are associated with a decreased risk of progression to AIDS-related complex or AIDS. **AIDS Res Hum Retroviruses**. 11 (8): 903-7, 1995.

RIVIERE, Y. et al. Gag-specific cytotoxic responses to HIV type 1 are associated with a decreased risk of progression to AIDS-related complex or AIDS. **AIDS Res Hum Retrov**, v. 11, p. 903-907, 1997.

RITS-VOLLOCH S, FREY G, HARRISON SC, CHEN B. Restraining the conformation of HIV-1 gp120 by removing a flexible loop. **EMBO J**. 25 (20): 5026-35, 2006.

ROBERTSON, D. L. et al. Recombination in HIV-1. **Nature**, v. 374, p. 124–126, 1995.

ROETH JF, COLLINS KL. Human immunodeficiency virus type 1 Nef: adapting to intracellular trafficking pathways. **Microbiol Mol Biol Rev**. 70 (2): 548-63, 2006.

ROGERS JS. Maximum likelihood estimation of phylogenetic trees is consistent when substitution rates vary according to the invariable sites plus gamma distribution. *Syst Biol.* 50 (5): 713-22, 2001.

ROSENBERG, E. S. et al. Vigorous HIV-1-specific CD4+ T cell responses associated with control of viremia. *Science*, v. 278, p. 1447-50, 1997.

ROSENBERG ES, LAROSA L, FLYNN T, ROBBINS G, WALKER BD. Characterization of HIV-1-specific T-helper cells in acute and chronic infection. *Immunol Lett.* 66 (1-3): 89-93, 1999.

ROSS, H. A.; RODRIGO, A. Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration. *J Virol*, v. 76, p. 11715-11719, 2002.

ROUX KH, TAYLOR KA. AIDS virus envelope spike structure. *Curr Opin Struct Biol.* 17 (2): 244-52, 2007.

ROUZINE, I.; COFFIN, J. M. Interplay between experimental and theory in development of a working model for HIV-1 population dynamics. In: DOMINGO, E.; WEBSTER, R.; HOLLAND, J. (Ed.). *Origin and evolution of viruses*. 225-262, 1999a.

ROUZINE, I.; COFFIN, J. M. Linkage disequilibrium test implies a large effective population number for HIV in vivo. *Proc Natl Acad Sci USA*, v. 96, p. 10758-10763, 1999b.

SANABANI S, KLEINE NETO W, KALMAR EM, DIAZ RS, JANINI LM, SABINO EC. Analysis of the near full length genomes of HIV-1 subtypes B, F and BF recombinant from a cohort of 14 patients in Sao Paulo, Brazil. *Infect Genet Evol.* 6 (5): 368-77, 2006.

SANCHEZ-MERINO V, NIE S, LUZURIAGA K. HIV-1-specific CD8+ T cell responses and viral evolution in women and infants. *J Immunol.* 175 (10): 6976-86, 2005.

SANDBERG JK, FAST NM, JORDAN KA, FURLAN SN, BARBOUR JD, FENNELLY G, DOBROSZYCKI J, SPIEGEL HM, WIZNIA A, ROSENBERG MG, NIXON DF. HIV-specific CD8+ T cell function in children with vertically acquired HIV-1 infection is critically influenced by age and the state of the CD4+ T cell compartment. *J Immunol.* 170 (8): 4403-10, 2003.

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.* 20 (16): 2011-9, 2006.

SCHLIEKELMAN P, GARNER C, SLATKIN M. Natural selection and resistance to HIV. *Nature.* 411 (6837): 545-6, 2001.

SCHRAGO C. G. An empirical examination of the standard errors of maximum likelihood phylogenetic parameters under the molecular clock via bootstrapping. *Genet. Mol. Res.* 5 (1): 233-241, 2006.

SHANKARAPPA R, MARGOLICK JB, GANGE SJ, RODRIGO AG, UPCHURCH D, FARZADEGAN H, GUPTA P, RINALDO CR, LEARN GH, HE X, HUANG XL, MULLINS JI. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol*, v. 73, p. 10489-10502, 1999.

SHANKARAPPA R, CHATTERJEE R, LEARN GH, NEOGI D, DING M, ROY P, GHOSH A, KINGSLEY L, HARRISON L, MULLINS JI, GUPTA P. Human immunodeficiency virus type 1 env sequences from Calcutta in eastern India:

identification of features that distinguish subtype C sequences in India from other subtype C sequences. *J Virol.* 75 (21): 10479-87, 2001.

SHARMA PL, CRUMPACKER CS. Attenuated replication of human immunodeficiency virus type 1 with a didanosine-selected reverse transcriptase mutation. *J Virol.* 71 (11): 8846-51, 1997.

SHELDON, J. G.; Condra, J. H. Genotypic analysis methods for detection of drug resistance mutations in the HIV-1 proteinase and reverse transcriptase genes. *Antivir Ther*, v. 4, p. 135-142, 1999.

SHERIDAN P.L., SHELINE C., MILOCO S.H., JONES K.A. Tat and the HIV promoter: a model for RNA-mediated regulation of transcription. *Semin Virol.* 4: 69-80, 1991.

SIERRA S, KUPFER B, KAISER R. Basics of the virology of HIV-1 and its replication. *J Clin Virol.* 34 (4): 233-44, 2005.

SMITH, M.W. et al. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science*, v. 277, p. 959-965, 1997.

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.* 34 (5): 520-6, 2003.

SRINIVASAN A, ANAND R, YORK D, RANGANATHAN P, FEORINO P, SCHOCHETMAN G, CURRAN J, KALYANARAMAN VS, LUCIW PA, SANCHEZ-PESCADOR R. Molecular characterization of human immunodeficiency virus from

Zaire: nucleotide sequence analysis identifies conserved and variable domains in the envelope gene. *Gene*.52 (1): 71-82, 1987.

STAMATATOS L, CHENG-MAYER C. Evidence that the structural conformation of envelope gp120 affects human immunodeficiency virus type 1 infectivity, host range, and syncytium-forming ability. *J Virol*. 67 (9): 5635-9, 1993.

STIEGLER G, KUNERT R, PURTSCHER M, WOLBANK S, VOGLAUER R, STEINDL F, KATINGER H. A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res Hum Retroviruses*. 17 (18): 1757-65, 2001.

SUZUKI Y, AND GOJOBORI T. A method for detecting positive selection at single amino acid sites. *Mol. Biol. Evol*. 16:1315–1328, 1999.

SWOFFORD, D. L. **PAUP***. **Phylogenetic analysis using parsimony (*and other methods)**. Version 4. Massachusetts: Sinauer Associates,1998.

SWOFFORD DL, WADDELL PJ, HUELSENBECK JP, FOSTER PG, LEWIS PO, ROGERS JS. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst Biol*. 50 (4): 525-39, 2001.

TANTILLO C, DING J, JACOBO-MOLINA A, NANNI RG, BOYER PL, HUGHES SH, PAUWELS R, ANDRIES K, JANSSEN PA, ARNOLD E. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance.

TAYLOR L. H. et al. Risk factor for human disease emergence. **Philos Trans R Soc Lond**, v. 356, p. 983-990, 2001.

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

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**. 25 (24): 4876-82, 1997.

THOMSON MM, VILLAHERMOSA ML, VAZQUEZ-DE-PARGA E, CUEVAS MT, DELGADO E, MANJON N, MEDRANO L, PEREZ-ALVAREZ L, CONTRERAS G, CARRILLO MG, SALOMON H, NAJERA R. Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina. **AIDS**. 14 (7): 897-9, 2000.

THOMSON MM, DELGADO E, HERRERO I, VILLAHERMOSA ML, VAZQUEZ-DE PARGA E, CUEVAS MT, CARMONA R, MEDRANO L, PEREZ-ALVAREZ L, CUEVAS L, NAJERA R. Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. **J Gen Virol**. 83 (Pt 1): 107-19, 2002.

TUNGATURTHI PK, SAWAYA BE, SINGH SP, TOMKOWICZ B, AYYAVOO V, KHALILI K, COLLMAN RG, AMINI S, SRINIVASAN A. Role of HIV-1 Vpr in AIDS pathogenesis: relevance and implications of intravirion, intracellular and free Vpr. **Biomed Pharmacother**. 57 (1): 20-4, 2003.

TURNER BG, SUMMERS MF. Structural biology of HIV. **J Mol Biol**. 285 (1): 1-32, 1999.

VAISHNAV YN, WONG-STAAAL F. The biochemistry of AIDS. *Annu Rev Biochem.*60: 577-630, 1991.

VAN OPIJNEN T, DE RONDE A, BOERLIJST MC, BERKHOUT B. Adaptation of HIV-1 Depends on the Host-Cell Environment. *PLoS ONE.* 2: e271, 2007.

VANCOTT TC, POLONIS VR, LOOMIS LD, MICHAEL NL, NARA PL, BIRX DL. Differential role of V3-specific antibodies in neutralization assays involving primary and laboratory-adapted isolates of HIV type 1. *AIDS Res Hum Retroviruses.* 11 (11): 1379-91, 1995.

VERONESE FD, COPELAND TD, OROSZLAN S, GALLO RC, SARNGADHARAN MG. Biochemical and immunological analysis of human immunodeficiency virus gag gene products p17 and p24. *J Virol.* 62 (3): 795-801, 1988.

ZANOTTO PM, KALLAS EG, DE SOUZA RF, HOLMES EC. Genealogical evidence for positive selection in the nef gene of HIV-1. *Genetics.* 153 (3): 1077-89, 1999.

ZHANG WH, HOCKLEY DJ, NERMUT MV, JONES IM. Functional consequences of mutations in HIV-1 Gag p55 selected by CTL pressure. *Virology.* 203 (1): 101-5, 1994.

ZHANG Z, ZHAO QX, FU JL, YAO JX, HE Y, JIN L, WANG FS. Characteristics of HIV-1-specific CD8 T-cell responses and their role in loss of viremia in children chronically infected with HIV-1 undergoing highly active antiretroviral therapy. *Chin Med J (Engl).* 119 (23): 1949-57, 2006.

WANG WK, CHEN MY, CHUANG CY, JEANG KT, HUANG LM. Molecular biology of human immunodeficiency virus type 1. *J Microbiol Immunol Infect.* 33 (3): 131-40, 2000.

WELLENSIEK BP, SUNDARAVARADAN V, RAMAKRISHNAN R, AHMAD N. Molecular characterization of the HIV-1 gag nucleocapsid gene associated with vertical transmission. *Retrovirology*. 3: 21, 2006.

WILLIANSO, S. Adaptation in the env gene of HIV-1 and evolutionary theories of disease progression. *Mol Biol Evol*, v. 20, p. 1318-1325, 2003.

WLODAWER A, VONDRASEK J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct*. 27: 249-84, 1998.

WOOLHOUSE, M. E. J. et al. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Gen*, v. 32, p. 569-577, 2002.

WU, H., LATHEY, J., RUAN, P., et al. (2004), Relationship of Plasma HIV-1 RNA Dynamics to Baseline Factors and Virological Responses to Highly Active Antiretroviral Therapy in Adolescents (Aged 12-22 Years) Infected through High-Risk Behavior, *Journal of Infectious Diseases*, 189, 593-601.

YANG Z. Statistical properties of the maximum likelihood method of phylogenetic estimation and comparison with distance matrix methods. *Syst. Biol*. 43: 329-342, 1994a.

YANG Z. Estimating the pattern of nucleotide substitution. *J. Mol. Evol*. 39:105-111, 1994b.

YANG OO, CHURCH J, KITCHEN CM, KILPATRICK R, ALI A, GENG Y, KILLIAN MS, SABADO RL, NG H, SUEN J, BRYSON Y, JAMIESON BD, KROGSTAD P. Genetic and stochastic influences on the interaction of human immunodeficiency virus type 1 and cytotoxic T lymphocytes in identical twins. *J Virol*. 79 (24): 15368-75, 2005.

YANG W, BIELAWSKI JP, YANG Z. Widespread adaptive evolution in the human immunodeficiency virus type 1 genome. *J Mol Evol*. 57 (2): 212-21, 2003.

YANG Z., NIELSEN R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* 19 (6): 908 – 917, 2002.

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.*66 (8): 4966-71, 1992.

YUSIM K, KESMIR C, GASCHEN B, ADDO MM, ALTFELD M, BRUNAK S, CHIGAEV A, DETOURS V, KORBER BT. Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. *J Virol.* 76 (17): 8757-68, 2002.