



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS – ICB  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

RAIMUNDO CRISTOVÃO FERREIRA PEREIRA

**RELAÇÃO ENTRE NÍVEIS DE CÉLULAS T CD4+ E PRESSÃO SELETIVA NOS  
GENES *env* E *vif* DO HIV-1 SUBTIPOS B E C**

BELÉM

2015

RAIMUNDO CRISTOVÃO FERREIRA PEREIRA

**RELAÇÃO ENTRE NÍVEIS DE CÉLULAS T CD4+ E PRESSÃO SELETIVA NOS  
GENES *env* E *vif* DO HIV-1 SUBTIPOS B E C**

Dissertação submetida ao Programa de Pós-Graduação em Biotecnologia da UFPA como requisito para obtenção do grau de Mestre em Biotecnologia.

*Orientador:* Prof. Dr ÉLCIO DE SOUZA LEAL

BELÉM

2015

Dados Internacionais de Catalogação-na-Publicação (CIP)  
Sistema de Bibliotecas da UFPA

---

Pereira, Raimundo Cristovão Ferreira, 1990-  
Relação entre níveis de células T CD4+ e pressão  
seletiva nos genes env e vif do HIV-1 subtipos B e C /  
Raimundo Cristovão Ferreira Pereira. - 2015.

Orientador: Élcio de Souza Leal.  
Dissertação (Mestrado) - Universidade  
Federal do Pará, Instituto de Ciências  
Biológicas, Programa de Pós-Graduação em  
Biotecnologia, Belém, 2015.

1. Infecções por HIV. 2. Linfócitos. 3.  
Genes. 4. HIV (Vírus). I. Título.

CDD 22. ed. 616.9792

---

RAIMUNDO CRISTOVÃO FERREIRA PEREIRA

**RELAÇÃO ENTRE NÍVEIS DE CÉLULAS T CD4+ E PRESSÃO SELETIVA NOS  
GENES *env* E *vif* DO HIV-1 SUBTIPOS B E C**

Data da avaliação: \_\_\_/\_\_\_/\_\_\_

Conceito: \_\_\_\_\_.

BANCA EXAMINADORA

---

Prof. PhD. Élcio de Souza Leal  
(UFPA– Presidente da banca)

---

Prof. Dr Chubert Bernardo Castro de Sena  
(UFPA– Membro)

---

Prof<sup>a</sup>. Dr<sup>a</sup> Edna Cristina Santos Franco  
(IEC/UFPA– Membro)

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Fabiola Elizabeth Villanova  
(UFPA– Membro)

## AGRADECIMENTOS

Á DEUS pelo dom da existência e sabedoria, e a minha família, a qual dedico a minha vitória.

Minha querida mãe, amiga e amparo de todas as horas: Raimunda Marques Ferreira.

Meu pai e herói: Antonio Pereira.

Minhas irmãs: Antonia Pereira, Aparecida Pereira, Maria de Jesus Pereira e meus sobrinhos.

Meus sinceros agradecimentos à família Mesquita Muniz, a qual me deu um lar e me permitiu a realização de meu mestrado em nome de meu amigo Anderson Muniz.

Meu orientador: Élcio Leal, por ter me aceitado e conduzido meu trabalho.

Meus amigos: Victor Hugo e Saymon Teixeira, pelo apoio e motivação nas horas difíceis.

Minha amiga e parceira de trabalhos: Marta Barreiros.

Minha amiga-irmã: Andreza Muniz simplesmente por existir em minha vida.

A todos que contribuíram de alguma forma para meu crescimento, formação e em meu mestrado, minha gratidão também é de vocês.

Até aqui, me ajudou o Senhor!

Jesus Cristo, “pão da vida e luz das Nações”.

## RESUMO

Estudos anteriores mostraram haver uma relação direta entre níveis de linfócitos T CD4+ e taxas evolutivas do HIV-1 (DIAZ *et al.*, 2008; LEMEY *et al.*, 2007; NOSTROM *et al.*, 2014). Além disso, outros fatores também afetam a variabilidade do no gene *env*: por exemplo ação de anticorpos neutralizantes - Nabs (FROST *et al.*, 2005), a variação nos sítios de glicosilação (LEAL *et al.*, 2012; LEAL *et al.*, 2008), a ligação nos receptores da célula alvo (*i.e.*, CD4, CXCR4, CCR5), e ainda, a escolha dos receptores CCR5 para CXCR4 (MILD *et al.*, 2013). Com isso a relação entre diversificação viral e níveis células T CD4+ no gene *env* pode ser circunstancial. O gene *vif*, por outro lado, é mais conservado e não sujeito ao viés presente no gene *env* (*i.e.*, sítios de glicosilação, etc.). Assim, para estudar a influência dos níveis de células T CD4+ na variabilidade do HIV, foi usado a estimativa do regime seletivo (dN e dS) através do modelo de códons e análise filogenética de indivíduos não relacionados (inter-hospedeiro). Foram utilizadas sequências do HIV-1 de indivíduos não correlacionados, obtidas a partir do banco de dados de Los Alamos. As sequências foram separadas em categorias de níveis de linfócitos T CD4, e posteriormente analisadas. A análise mostrou não haver relação direta entre os níveis de células T CD4+ e as taxas evolutivas em gp120 e no gene *vif* do HIV-1 dos subtipos B e C em uma abordagem populacional.

**Palavras-chave:** Linfócitos T CD4+; *vif*; *env*; Diversidade Viral; Pressão Seletiva.

## ABSTRACT

Previous studies have shown a direct relationship between levels of CD4+ T-lymphocytes and evolutionary rates of HIV-1 (DIAZ *et al*, 2008; LEMEY *et al*, 2007; NOSTROM *et al*, 2014.). Other factors also affect the variability of the *env* gene: for example function of neutralizing antibodies - Nabs (FROST *et al.*, 2005), the variation in glycosylation sites; (LEAL *et al*, 2012 LEAL *et al.*, 2008), binding to target cell receptors (i.e, CD4, CXCR4, CCR5), and also the switch from CCR5 to CXCR4 receptor (MILD *et al.*, 2013). Thus, the relationship between viral levels diversification and CD4 + T cells in the *env* gene can be circumstantial. The *vif* gene, on the other hand, it is retained and not subject to bias present in the *env* gene (i.e, glycosylation sites, etc.). Thus, to study the influence of CD4 + T cells levels in HIV variability, the estimate of the selective regimes was used (Non-synonymous Substitutions - dN and Synonymous Substitutions - dS) by a codon-based model and phylogenetic analysis of unrelated individuals (inter-host). HIV-1 sequences were used of the not-correlated individuals, obtained from the Los Alamos Database. The sequences were separated into CD4+ levels of categories and then analyzed. The analysis revealed no direct correlation between CD4+ T cell levels and evolutionary rates in gp120 and *vif* gene from HIV-1, subtypes B and C in a population approach.

**Keywords:** CD4 + T-lymphocytes; *vif*; *env*; Viral diversity; Selective pressure.

## LISTA DE FIGURAS E TABELAS

---

- Figura 1** Níveis de Células T CD4 + e Carga Viral (RNA) em Indivíduos infectados por HIV-1.
- Figura 2** Organização do Genoma do HIV.
- Figura 3** Representação estrutural do Vírus da Imunodeficiência Humana – HIV.
- Figura 4** Entrada do HIV na célula hospedeira.
- Figura 5** Ciclo Replicativo do HIV-1.
- Figura 6** Montagem do Complexo Vif-EloBC-A3G.
- Figura 7** Categorias de CD4+ para o gene *env* (gp120) subtipo B.
- Figura 8** Categorias de CD4+ para o gene *vif* subtipo B.
- Figura 9** Categorias de CD4+ para o gene *env* (gp120) subtipo C.
- Figura 10** Categorias de CD4+ para o gene *vif* subtipo C.
- Figura 11** Substituições Sinônimas e Não-Sinônimas no gene *env* (gp120) Subtipos B e C.
- Figura 12** Substituições Sinônimas e Não-Sinônimas no gene *vif* Subtipos B e C.
-

## LISTA DE SIGLAS

---

AIDS	Síndrome da Imunodeficiência Humana Adquirida
APOBEC	<i>Apolipoprotein B mRNA-editing catalytic polypeptide.</i>
CD4	Receptor celular (Grupamento de diferenciação 4)
CCR5	Receptor de quimiocina ligado à proteína G
CRFs	Formas recombinantes circulantes
CXCR4	Receptor C-X-C de Quimiocina tipo 4
Cul5	Culina 5, proteína componente do complexo ubiquitina ligase
dN	Número de substituições não sinônimas por sitio não sinônimo
dS	Número de substituições sinônimas por sítios sinônimos (reversões)
DNA	Ácido Desoxirribonucléico
cDNA	DNA complementar
Env	Proteína do Envelope Viral
EloB	Elonguina B, proteína componente do complexo ubiquitina ligase
EloC	Elonguina C, proteína componente do complexo ubiquitina ligase
EloBC	Complexo formado pela Elonguina B e Elonguina C
Gag	Proteína viral envolvida na montagem das partículas virais
G2	Fase do ciclo celular “Gap 2”
gRNA	RNA Genômico Viral
HIV	Vírus da Imunodeficiência Humana
HLA-B	Antígeno leucocitário humano - B
LTR	Repetições terminais longas
LANL	Laboratório Nacional de Los Alamos
ML	Máxima Verossimilhança
mRNA	RNA Mensageiro

RNA	Ácido Ribonucléico
tRNA <sup>3Lys</sup>	RNA transportador de Lisina
Vif	Fator de Infectividade Viral

1.	INTRODUÇÃO.....	11
2.	REFERENCIAL TEÓRICO.....	13
2.1.	O Vírus da Imunodeficiência Humana tipo 1 (HIV-1).....	13
2.2.	Ciclo Replicativo do HIV-1 .....	14
2.2.1.	Envelope viral (Env) .....	15
2.2.2	Fator de Infecção Viral (Vif).....	19
2.3	Diversidade do HIV-1 .....	20
2.4.	Pressão seletiva.....	21
3.	OBJETIVOS.....	24
3.1.	Objetivo geral.....	24
3.2.	Objetivos específicos.....	24
4.	MATERIAIS E MÉTODOS .....	25
4.1.	Obtenção das Sequências .....	25
4.2.	Categorização das sequências de acordo com os níveis de células T CD4+ .....	25
4.3.	Alinhamento e Inferência filogenética .....	28
4.4.	Análise da Pressão Seletiva .....	28
4.5.	Grupos controles.....	29
5.	RESULTADOS .....	30
5.1	Taxas de Substituições Sinônimas (dS) e Não-Sinônimas (dN) em gp120 do HIV-1 .....	30
5.2.	Taxas de Substituições Sinônimas (dS) e Não-Sinônimas (dN) do gene vif do HIV-1 .....	33
7.	CONCLUSÃO.....	37
8.	REFERÊNCIAS .....	38

## 1. INTRODUÇÃO

O curso clínico da infecção pelo HIV é marcado por uma considerável variabilidade na taxa de progressão para a AIDS, em função dos diversos fatores virais e do hospedeiro que são determinantes da progressão para a doença (MIGUELES & CONNORS, 2010). Dentre eles, a taxa de replicação do vírus expressa pela carga viral, é dependente dos níveis de linfócitos T CD4+ (MELLORS *et al.*, 1997).

A infecção pelo vírus causa a depleção progressiva dessas células, componentes importantes do sistema imune (CHITRA *et al.*, 2011; KNIPE & HOWLEY, 2013), como consequência direta da morte celular (GARG *et al.*, 2012), levando à um comprometimento das defesas do hospedeiro e a AIDS (MIGUEL, 2009; ACHESON, 2011; ABBAS *et al.*, 2012), como mostrado na figura 1 abaixo.

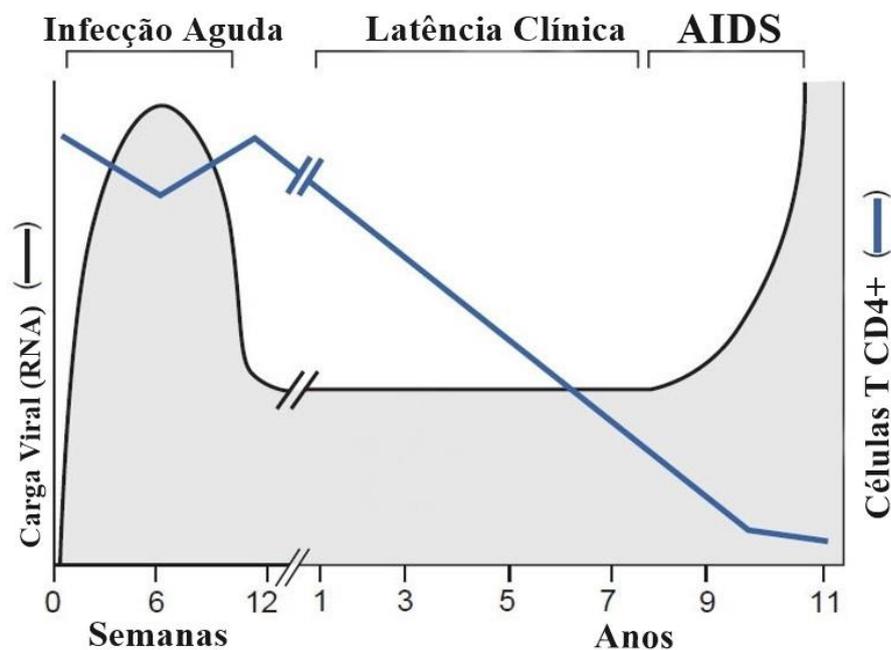


Fig.1 Níveis de Células T CD4 + e Carga Viral (RNA) em Indivíduos infectados por HIV-1. A linha azul representa os níveis de células T CD4+ (células/mm<sup>3</sup> de sangue) e a linha preta, os níveis de carga viral expressa pela quantidade de RNA viral presente no plasma sanguíneo (cópias/mL de sangue). Adaptado de: (ANCHESON, 2011).

Atualmente, para melhor compreensão do processo evolutivo do HIV, modelos evolucionários têm mostrado serem poderosas ferramentas nesses estudos (SHARP, 1997), dentre eles, o modelo da razão de substituições sinônimas/não sinônimas ( $dN/dS=\omega$ ), é responsável por indicar o tipo de pressão de seleção atuante em determinados códons de um

gene (YANG & NIELSEN, 2002). Logo, estudos das frequências absolutas de dS e dN separadamente podem promover maiores conhecimentos (SEO *et al.*, 2004).

O presente estudo analisou a intensidade e o tipo de pressão de seleção atuante na subunidade gp120 do gene *env* e no gene *vif* do HIV-1, subtipos B e C, com base nos valores de dN e dS, em relação aos níveis de células T CD4+ de indivíduos não correlacionados, em uma abordagem populacional.

## 2. REFERENCIAL TEÓRICO

### 2.1. O Vírus da Imunodeficiência Humana tipo 1 (HIV-1)

O HIV é um vírus envelopado, com cerca de 100 à 120nm e cada partícula viral contém o genoma composto por duas cópias de RNA de polaridade positiva, cada uma com cerca de nove mil e setecentos pares de bases (9,7kb) e as enzimas virais Transcriptase Reversa (RT), Protease (PR) e Integrase (IN), importantes nos primeiros passos da infecção pelo vírus, antes que novas proteínas sejam produzidas (KNIPE & HOWLEY, 2013).

O vírus pertence à família *Retroviridae*, e assim como os demais membros desta família, o genoma contém três genes estruturais: *gag* (*group-specific antigen*) responsável por codificar as proteínas do capsídeo viral; *pol* (*polimerases*) que codifica as enzimas RT, Proteases e Integrases; e o gene *env* (*envelope*), a Glicoproteína gp160, que após a clivagem por uma protease da célula hospedeira, dá origem às glicoproteínas do envelope: gp120 (*Surface Unit* - SU) e gp41 (*Trans Membrane* - TM) (MADIGAN *et al.*, 2010; MURPHY *et al.* 2010).

Além dos genes estruturais, o HIV-1 contém ainda, três genes regulatórios tais como: *tat*, *rev*, e *nef* e três auxiliares que incluem *vif* (*Virion Infectivity Factor*), *vpr* e *vpu* (FRANKEL & YOUNG, 1998), estes últimos dispensáveis para o crescimento viral em células *in vitro*, no entanto são considerados essenciais para a patogênese viral e processo adaptativo à célula hospedeira (LI *et al.*, 2005).

As figuras abaixo esquematizam o genoma do vírus (Fig. 2) e a estrutura da partícula viral do HIV-1 na fase mais precoce e na fase mais tardia do final do ciclo replicativo (Fig. 3).

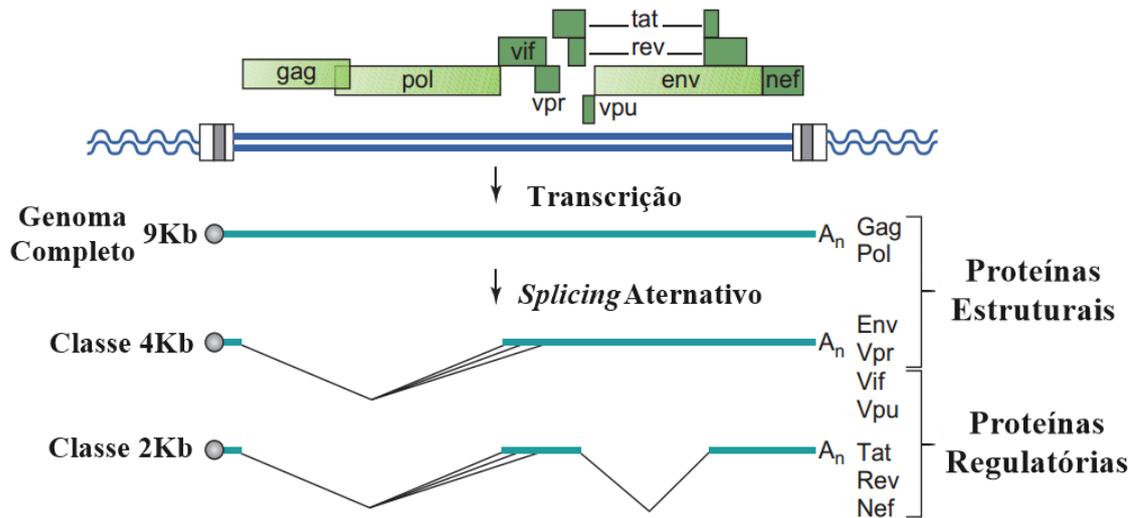


Fig. 2 Organização do Genoma do HIV, genes estruturais em verde claro e genes acessórios em verde escuro e classe de mRNAs. O mapa do DNA províral integrado, franqueado por duas LTRs “Long Terminal Repeats”. Abaixo, três classes de mRNAs de diferentes tamanhos por “splicing” alternativo e proteínas virais produzidas. Adaptado de (ACHESON, 2011).

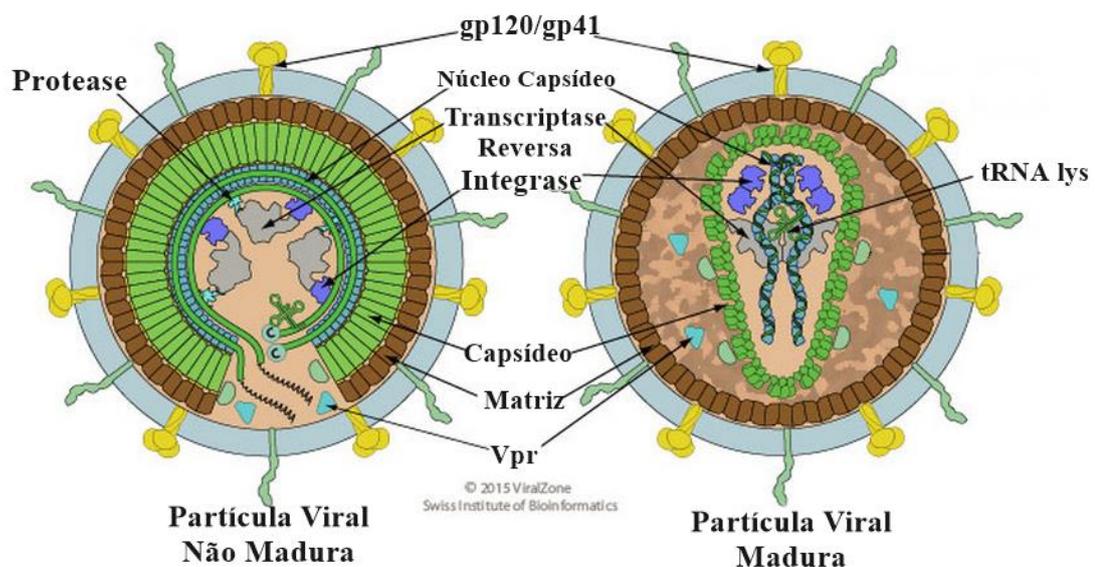


Fig. 3 Representação estrutural do Vírus da Imunodeficiência Humana – HIV. Na esquerda a representação da partícula viral em fase de maturação (não infecciosa) e na direita a partícula viral madura (infecciosa). Adaptado de ([http://viralzone.expasy.org/all\\_by\\_species/7.html](http://viralzone.expasy.org/all_by_species/7.html))

## 2.2. Ciclo Replicativo do HIV-1

O principal determinante na patogênese causada pelo HIV é o tropismo viral por células que expressam a molécula CD4 em sua superfície, tais como linfócitos T auxiliares e macrófagos, dentre outras (PANTOPHLET, *et al.*, 2006; ABBAS *et al.*, 2012), e como

consequência direta desse processo, a redução destas células que atuam nas respostas imunológicas do hospedeiro, causando um quadro de imunossupressão induzida pelo vírus (GUANGULI *et al.*, 2012).

O ciclo replicativo do vírus é considerado complexo e sua duração e evolução dependem do tipo de célula alvo e da ativação celular (SIMON *et al.*, 2006), e principalmente de duas glicoproteínas presentes no envelope necessárias para o início da infecção: gp120 e gp41. Ambas se ligam aos receptores de superfície (CD4) presente nos linfócitos T do hospedeiro (ACHESON, 2011).

### **2.2.1. Envelope viral (Env)**

A glicoproteína do envelope viral tem por finalidade primária facilitar a fusão das membranas virais e celulares permitindo a entrada do vírus na célula hospedeira. Logo, Env é encontrada disposta sobre a superfície do vírus e das células infectadas pelo HIV, como hetero-trímeros, e cada trímero contém três monômeros, que são constituídos por uma unidade de ligação ao receptor de superfície (gp120) e uma unidade de transmembrana fusogênico (gp41) que medeia a fusão de membranas. A subunidade gp120 liga-se a CD4 e ao co-receptor CXCR4 (X4) sobre os linfócitos T auxiliares (CHAN *et al.*, 1997; DUMAS *et al.*, 2014) ou CCR5 (R5) de macrófagos (O'CONNELL *et al.*, 2013), como ilustrado abaixo na (Fig. 4).

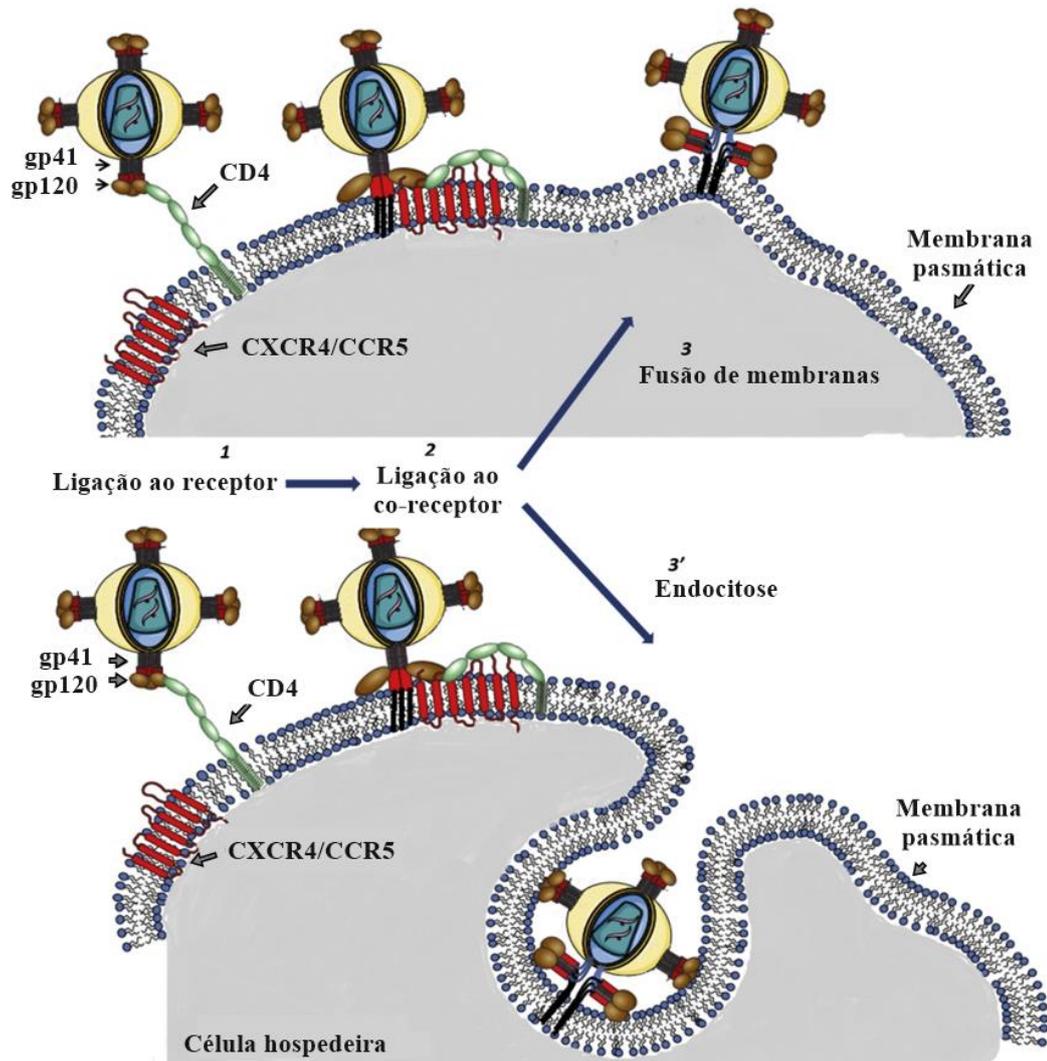


Fig. 4 Entrada do HIV na célula hospedeira. O esquema representa a ligação Env – CD4/CXCR4-CCR5 responsável por mediar a entrada do vírus por meio de fusão de membranas ou por endocitose. Adaptado de (DUMAS *et al.*, 2014).

Vírus que usam o receptor CCR5 (R5) são predominantes nos estágios mais precoces da infecção e mais frequentemente se mantêm durante o curso da infecção (KEELE *et al.*, 2008), já nos estágios finais da infecção, em cerca de 70% dos indivíduos infectados pelo subtipo B os vírus utilizaram o receptor CXCR4 (X4) em vez de R5 (ESBJORNSSON *et al.*, 2010), evidenciando serem mais comuns na fase tardia da infecção, e associados com a acelerada perda de células T CD4<sup>+</sup> (MILD *et al.*, 2013).

Essas diferenças são em função das variações na glicoproteína gp120 (ACHESON, 2011), pois o gene *env* é o que está mais sobre constantes taxas de evolução dentre os outros que compõem o genoma viral (ROSS & RODRIGO, 2002).

A ligação de gp120 ao receptor CD4 da células hospedeiras, desencadeia uma sequência complexa de eventos que envolvem várias alterações conformacionais na gp120 que resultam na exposição de locais de ligação ao co-receptor (GUTTMAN *et al.*, 2014).

Uma vez que o vírus entra na célula hospedeira por endocitose ou por fusão de membranas (DUMAS *et al.*, 2014), é então liberado no citoplasma o complexo nucleoproteico que estruturalmente é composto pela transcriptase reversa (RT), duas cópias de RNA genômico (gRNA), um RNA transportador de lisina (tRNA<sup>3Lys</sup>), e as proteínas Integrase, Vpr, Nef e NC (LEVIN *et al.*, 2010). A transcriptase reversa (RT) utilizando o tRNA<sup>3Lys</sup> do vírus como *primer*, inicia a síntese reversa do DNA complementar (cDNA) de fita negativa (KLEIMAN *et al.*, 2004).

O sítio de ligação da tRNA<sup>3Lys</sup> ao gRNA corresponde a região de ligação ao *primer* (PBS). Depois de finalizado a síntese de cDNA(-), a transcriptase atua também como ribonuclease H, e degrada o gRNA expondo a fita de cDNA negativa que servirá de molde para a síntese do cDNA de fita positiva. Depois de sintetizada, o cDNA dupla fita é transportado até o núcleo e inserido ao cromossomo da célula hospedeira por ação da enzima integrase (KLEIMAN *et al.*, 2004).

Uma vez integrado, o cDNA (denominado provirus) é então transcrito pela RNA polimerase II do hospedeiro em uma fita de mRNA que, no citoplasma, será traduzido dando origem às proteínas virais (SIMON *et al.*, 2006; CHAN & KIM, 1998), como ilustrado abaixo na (Fig. 5).

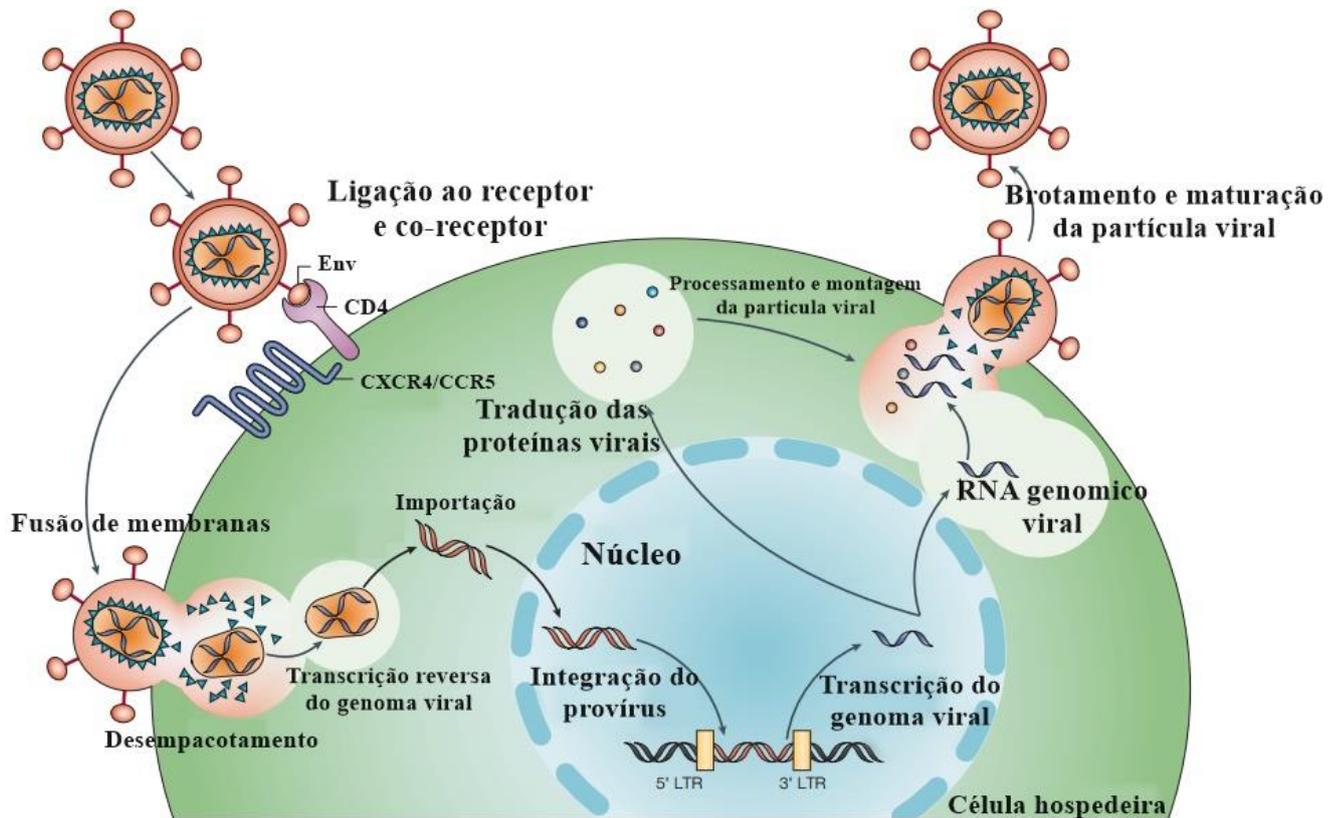


Fig. 5 Ciclo Replicativo do HIV-1. Representação da partícula viral, desde sua interação com o receptor e co-receptor celular; fusão de membranas; desempacotamento; transcrição reversa; importação; integração; transcrição do RNA viral; tradução e o brotamento de uma nova partícula viral da célula hospedeira.

Adaptado de (RAMBAUT *et al.*, 2004).

A montagem do vírus é direcionada pela proteína viral denominada Gag, que é uma proteína multi-domínio formada por três sítios: matriz, região capsídica e a nucleocapsídica (NC) e três pequenos peptídeos denominados SP1, SP2 e p6 (BRIGGS; KRÄUSSLICH; 2011).

Logo, em resposta à infecção pelo HIV, as células hospedeiras respondem de diferentes modos, ativando mecanismos da resposta inata e humoral contra o vírus (LI *et al.*, 2005). Um dos mecanismos intracelulares da resposta inata envolvidos na resistência a infecção viral é a família das proteínas APOBEC, as quais são alvos da proteína viral Vif (*Viral Infectivity Factor*).

### 2.2.2 Fator de Infecção Viral (Vif)

O Fator de Infectividade viral (Vif) é uma proteína acessória, essencial à replicação do HIV em linfócitos, macrófagos e em algumas linhagens de células conhecidas pelo termo “não permissivas”, ou seja, os vírus que eram produzidos a partir de células “não permissivas”, não eram mais infecciosos (STREBEL *et al.*, 1987).

Hoje, sabe-se que este mecanismo de resposta da célula hospedeira é uma proteína chamada APOBEC (Apolipoproteína B Enzima Catalítica Polipeptídeo *like* 3G de Edição de mRNA), um potente inibidor da infecção do HIV em células “não permissivas” (HARRIS *et al.*, 2002).

A proteína Vif desempenha diferentes funções no ciclo de vida do HIV, tais como: supressão da síntese de proteínas A3G (APOBEC3G) e empacotamento das mesmas nas partículas virais, bloqueio do ciclo celular na fase G2, e também possui atividade de chaperona de RNA; no entanto, o mecanismo de ação molecular mais abordado refere-se ao bloqueio das proteínas citidina desaminases APOBECs, responsáveis pela resposta imune inata do hospedeiro contra infecções virais (STOPAK *et al.*, 2003; MARIANI *et al.*, 2003; KAO *et al.*, 2003; IZUMI *et al.*, 2010).

A Vif, neste caso, age através do recrutamento do complexo ubiquitina E3 ligase composto pelas proteínas celulares Elonguina B e C (EloBC), Nedd8, Culina-5 (Cul5) e RING-box 2 (Rbx2), que é responsável pela poliubiquitinação das APOBEC3 e consequente degradação mediada pelo complexo proteassomal celular (SHEEHY *et al.*, 2003; ZHANG *et al.*, 2008; ZHOU *et al.*, 2012; XIAO *et al.*, 2007).

Vif consiste em vários domínios funcionais envolvidos no reconhecimento da APOBEC e complexo ubiquitina E3 ligase, mas hoje são melhores caracterizados por apenas três domínios (BARRAUD *et al.*, 2008). O primeiro domínio é conhecido como SOCS-box o qual inclui o BC-box e os motivos Cullin-box (STANLEY *et al.*, 2008). O BC-box tem uma elevada afinidade para o complexo EloBC (STANLEY *et al.*, 2008; WOLFE *et al.*, 2010).

O segundo domínio é referido como HCCH, que está localizado a jusante de BC-box, e corresponde aos resíduos 100-142 (XIAO *et al.*, 2006; MEHLE *et al.*, 2006; XIAO *et al.*, 2007), sendo altamente conservado entre Vif de diferentes lentivírus (PAUL *et al.*, 2006) e é responsável por mediar a interação Cul5 (XIAO *et al.*, 2007; WOLFE *et al.*, 2010; MEHLE *et al.*, 2006) e contém dois resíduos de cisteína (C114, C133) e dois de histidina (H108, H139), que são componentes essenciais da sequência de polipeptídeos para formar um motivo dedo de zinco “*zinc-finger*” (XIAO *et al.*, 2007; WOLFE *et al.*, 2010).

O local de ligação A3G é o terceiro domínio da proteína Vif. Esta região é essencial para dirigir a ubiquitinação de A3G pelo complexo ubiquitina E3 ligase, e corresponde a uma região não linear localizado na extremidade N-terminal da Vif (HE *et al.*, 2008), assim, formando o complexo Vif-EloBC-A3G (da COSTA *et al.*, 2014), mostrado na figura abaixo (Fig. 6).

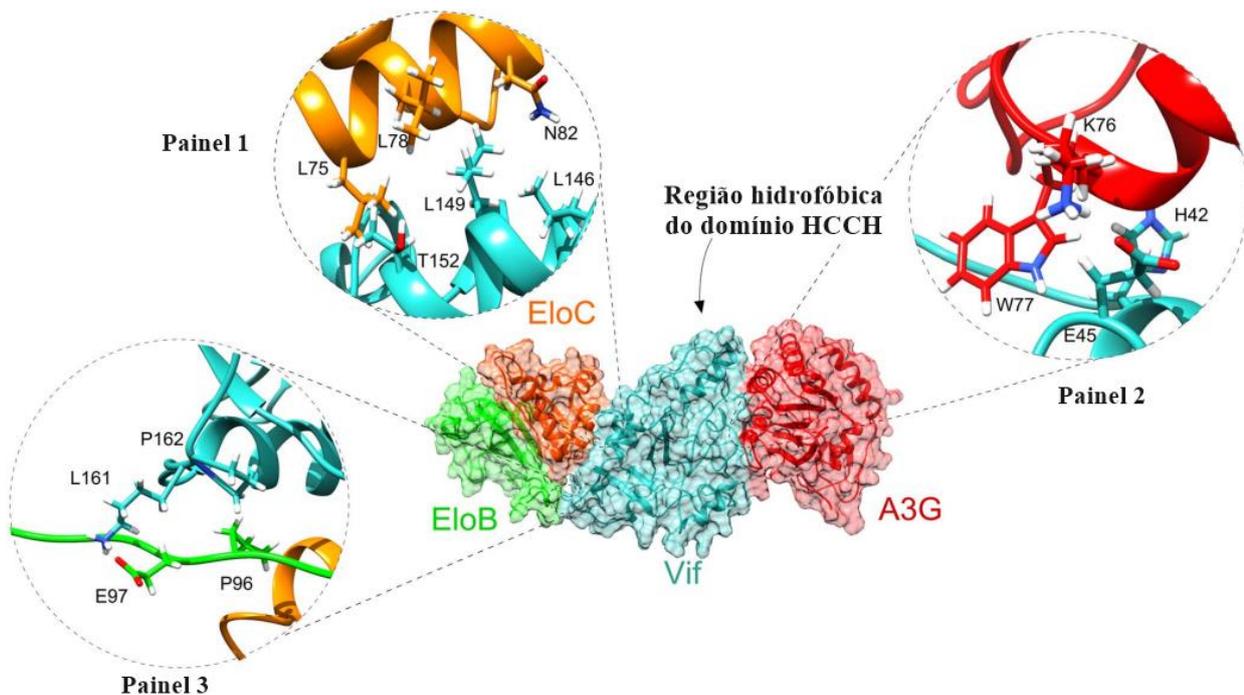


Fig. 6 Montagem do Complexo Vif-EloBC-A3G. O primeiro painel mostra a interação dos resíduos de aminoácidos do motivo BC-Box da Vif (em ciano) com EloC (em laranja); o segundo, na região N-terminal, mostra a interação de Vif, via ligação tipo ponte de hidrogênio com A3G (em vermelho); o último painel mostra a interação hidrofóbica dos resíduos de aminoácidos do motivo Cullin-Box da Vif, com os resíduos de aminoácidos da região C-terminal da EloB (em verde).

Adaptado de (da COSTA *et al.*, 2014).

### 2.3 Diversidade do HIV-1

Uma das características marcantes do HIV-1 que desafia pesquisadores do mundo todo é a rápida taxa de evolução e a ampla variabilidade genética do vírus que resulta em populações mais dinâmicas na epidemia da AIDS (TEBIT & ARTS, 2011), que já afeta cerca de 35,3 milhões de pessoas em todo o mundo (UNAIDS/WHO, 2013) e estima-se aproximadamente 28 milhões de novos casos até 2030 (UNAIDS/WHO, 2015).

Atualmente o HIV-1 é classificado em quatro grupos (M, N, O e P) que fazem parte da diversidade genética do vírus (SHARP, HAHN, 2011), sendo que, o grupo M “Major”, possui

nove subtipos com características divergentes (A, B, C, D, F, G, H, J e K) (TEBIT & ARTS, 2011; LIHANA *et al.*, 2012) e 72 formas recombinantes circulantes “CRFs” (PESSÔA *et al.*, 2014).

O subtipo B é considerado o mais prevalente no Brasil, seguido pelo F e C (BRINDEIRO *et al.*, 2003; CARDOSO *et al.*, 2009), no entanto, o subtipo C tem a maior frequência no sul do país e está com expansão em relação aos subtipos B e F, indicando maior taxa de incidência nos últimos anos (SANTOS *et al.*, 2007; ALMEIDA *et al.*, 2012).

A rápida evolução viral do HIV-1 é devida as altas taxas de mutação ( $3.10^{-5}$  mutações/sítio/cada ciclo) e pela rápida adaptação ao hospedeiro, que está associada à eventos de recombinação, a quantidade de partículas virais produzidas ( $\sim 10^{10} - 10^{12}$  novos vírus por dia) (RAMBAUT *et al.*, 2004) e rapidez em adaptar-se as pressões evolutivas impostas pelo sistema imune ( $\sim 1$  substituição/2,5 meses) (WILLIAMSON, 2003).

Logo uma das consequências diretas desse processo evolutivo do HIV, está no desenvolvimento de resistência contra drogas, ou ainda, o surgimento de subtipos duplamente resistentes, fixados em determinada população (RAMBAUT *et al.*, 2004).

#### **2.4. Pressão seletiva**

Dentre outros fatores determinantes do processo evolutivo de diversos microrganismos, a seleção positiva e a recombinação homóloga são os principais, ambos contribuem para a variabilidade genética entre eles, que acaba determinando a diversidade populacional e a adaptação dos microrganismos patogênicos a diferentes grupos de hospedeiros (LEWIS-ROGERS *et al.*, 2008; PETERSEN *et al.*, 2007).

A comparação de dados genômicos de diferentes espécies pode ser utilizada para a identificação de genes afetados pela seleção natural (PETERSEN *et al.*, 2007), o que hoje é bem mais explorado em virtude dos variados métodos de análises filogenéticas utilizados para comparar a relação entre os genes, inferindo a história comum entre eles.

Conhecer genes que apresentam taxas de mutação maior, é importante para aplicação prática, ajudando a desenvolver vacinas e drogas, uma vez que os alvos serão regiões conservadas do genoma do patógeno (POLLEY *et al.*, 2003).

A redundância do código genético permite que as sequências de nucleotídeos evoluam mais rapidamente do que as sequências de proteínas. Uma mudança na última base de um códon resulta na alteração do amino ácido correspondente em apenas 30% dos casos (SALEMI & VANDAMME, 2003).

Logo, a frequência de mutações sinônimas é bem maior quando comparadas às mutações não sinônimas (YANG & BIELAWSKI, 2000). Isto ocorre porque mutações não sinônimas podem alterar a funcionalidade da proteína codificada em comparação com o alelo selvagem, o que resulta na exclusão do alelo mutante da população, devido à ação da seleção purificadora ou negativa (AGUILETA *et al.*, 2010).

No entanto, em outras situações, o alinhamento de códons pode mostrar que podem existir posições em que a frequência de mutações não sinônimas é mais elevada do que a frequência desse tipo de mutação no restante do alinhamento em questão, indicando uma pressão de seleção positiva para a variação ao invés da conservação dessa proteína (AGUILETA *et al.*, 2009; 2010).

Os modelos de substituição de códons demonstram como a seleção natural atua sobre uma determinada proteína, através de mudanças na sequência nucleotídica. Em função disso, a análise de seleção positiva é realizada diretamente nas sequências codificadoras (SCHERER, 2002).

A taxa de substituições não sinônimas/sinônimas ( $\omega = d_N/d_S$ ) é uma medida que indica que tipo de pressão de seleção está atuando em uma determinada proteína, sendo que  $\omega = 1$  indica evolução neutra,  $\omega < 1$  indica seleção purificadora e  $\omega > 1$  indica seleção positiva (YANG & NIELSEN, 2002).

Na evolução viral intra-hospedeiro dentre outros fatores, é importante mencionar a diferença nas substituições sinônimas, as quais podem refletir diferenças na taxa de mutação ou tempo de geração, enquanto que as não sinônimas, podem estar ligadas a mudança na pressão seletiva e tamanho populacional efetivo (SEO, 2004).

A análise de diversos genes por meio da taxa de substituições sinônimas e não sinônimas revelou a ação da seleção positiva sobre eles (YANG & BIELAWSKI, 2000), e desde a primeira identificação da seleção positiva atuante em sequências de proteínas (HILL & HASTIE, 1987), a estimativa das taxas de substituições sinônimas (dS) e não-sinônimas (dN), tem se tornado uma ferramenta padrão nos estudos de evolução molecular (SHARP, 1997).

As análises de estimativa das frequências relativas de (dS/dN) tiveram poder substancialmente aumentado através do desenvolvimento de modelos de verossimilhança baseada em códons que permitem que a seleção varie entre os sítios (NIELSEN & YANG, 1998). Embora as frequências absolutas de dS e dN separadamente possam promover maiores conhecimentos (SEO *et al.*, 2004).

Um exemplo disso foram as diferenças entre as taxas de substituições sinônimas (dS) na evolução do HIV-1 intra-hospedeiro, que podem sugerir diferenças na taxa de mutação do

vírus ou taxas de replicação viral, enquanto que as diferenças nas taxas de (dN), sugerem as mudanças de pressão seletiva ou tamanho populacional efetivo (SEO *et al.*, 2004; LEE *et al.*, 2008).

Lemey *et al.*, (2007), mostrou que a progressão do HIV-1 para AIDS parece ser predita pelas taxas de (dS), as quais são indicativas da dinâmica replicativa do vírus. Já no trabalho de Lee *et al.*, (2008) foi mostrado que a taxa de evolução viral intra-hospedeiro não é constante, mas sim em conformidade com a taxa de declínio das células T CD4+.

Neste contexto, Diaz *et al.*, (2008) reportou uma correlação entre os níveis de células T CD4+ ( $>200/\text{mm}^3$ ) com valores maiores de (dS/dN) no gene *env* em variantes do subtipo B do HIV-1 na epidemia brasileira, como indicativo de seleção positiva, abrindo caminhos para realização deste estudo no gene *env*, com a abordagem da região codificante da unidade de superfície da glicoproteína do envelope viral (gp120) e no gene *vif* dos subtipos B e C do HIV-1.

### **3. OBJETIVOS**

#### **3.1. Objetivo geral**

Analisar os regimes seletivos atantes nos genes *env* (gp120) e *vif* do HIV-1 em indivíduos não correlacionados.

#### **3.2. Objetivos específicos**

- 1 Correlacionar os níveis de células T CD4+ com a intensidade da pressão seletiva nos genes *env* e *vif*.
- 2 Analisar a diversidade adaptativa nos genes *env* e *vif* de isolados de HIV-1 subtipos B e C.
- 3 Avaliar o efeito dos níveis de células T CD4+ na intensidade de seleção no HIV-1 em populações de indivíduos não correlacionados.

## 4. MATERIAIS E MÉTODOS

### 4.1. Obtenção das Sequências

As sequências codificantes dos genes *vif* e *env* (gp120) do HIV-1, subtipos B e C, foram obtidas a partir da base de dados do Laboratório Nacional de Los Alamos (LANL) - HIV DATABANK (<http://hiv-web.lanl.gov/>) e agrupados por categorias de acordo com os valores de células T CD4+. Para gp120, foi utilizado a subunidade gp120, com 1410 pares de bases, composta por 300 sequências para o subtipo B, separadas em 3 categorias (D1, D2 e D3), com 100 sequências em cada, e 120 sequências para o subtipo C, com 1365 pares de base (pbs), separadas em 2 categorias (D1 e D2) com 60 sequências em cada.

Para o gene *vif* foram utilizadas 64 sequências do subtipo B, com 579 bps, separadas em 2 categorias (D1 e D2) com 32 sequências em cada, e 124 sequências do subtipo C, com 579 bps, separadas também em 2 categorias (D1 e D2), com 62 sequências em cada.

Todas as sequências utilizadas nesse trabalho, obtidas a partir da base de dados de Los Alamos (LANL), atendiam o critério de inclusão de somente sequências de indivíduos não tratados com antirretrovirais “*drug-naive*”, afim de eliminar o viés da pressão de seleção induzida pelo tratamento antirretroviral.

### 4.2. Categorização das sequências de acordo com os níveis de células T CD4+

**Subtipo B.** Para o gene *env* (gp120), com mais sequências disponíveis no banco de dados (LANL) HIV DATABANK (<http://hiv-web.lanl.gov/>), foram agrupados em três categorias: D1, composta por 100 sequências com níveis de células T entre 100 e 300 CD4/mm<sup>3</sup>; D2, composta por 100 sequências com níveis de células T entre 400 e 600 CD4/mm<sup>3</sup> e D3, com 100 sequências com níveis de células T acima de 700 CD4/mm<sup>3</sup>, como mostrado abaixo (Fig. 7).

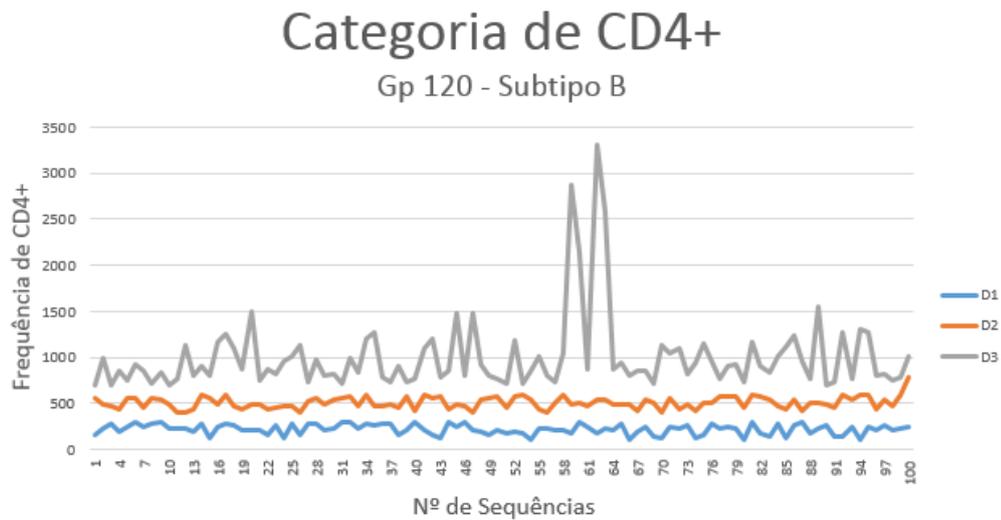


Fig.7 Categorias de CD4+ para o gene *env* (gp120) subtipo B. A linha superior (cinza) representa os níveis mais altos de CD4+ (>700CD4+), a linha em laranja, os valores entre (400-600CD4+) e a linha inferior (azul), níveis mais baixos de CD4+ (100-300CD4+). O eixo Y representa a frequência de CD4+ e o eixo X, o número de sequências.

Para o gene *vif*, foram agrupados em duas categorias, com 32 sequências em cada: D1, com valores de células T entre 100 e 300 CD4/mm<sup>3</sup> e D2, com valores de células T acima de 400 CD4/mm<sup>3</sup>, mostrado abaixo (Fig. 10).

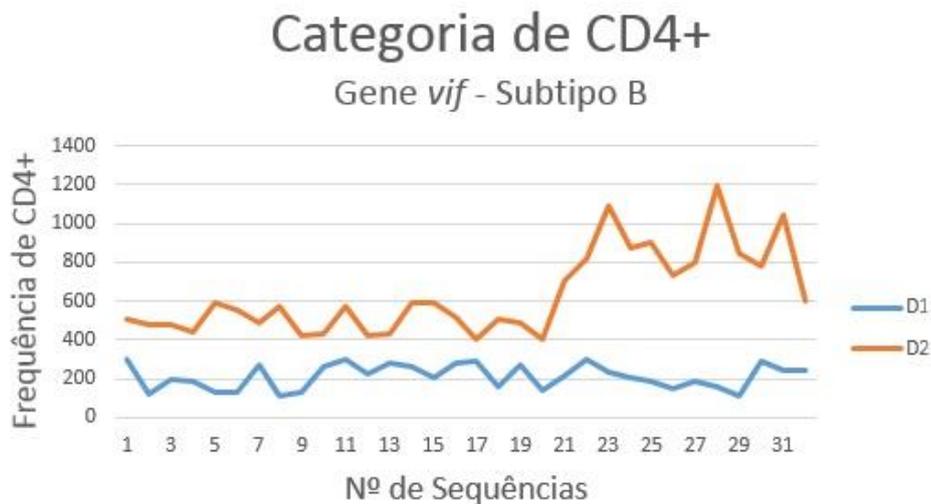


Fig.8 Categorias de CD4+ para o gene *vif* subtipo B. A linha superior (laranja) representa os níveis mais altos de CD4+ (>400 CD4+), e a linha inferior (azul), níveis mais baixos de CD4+ (100-300 CD4+). O eixo Y representa a frequência de CD4+ e o eixo X, o número de sequências.

**Subtipo C.** No gene *env* (gp120), em função da pouca quantidade de sequências disponíveis, foram agrupadas em 2 categorias: D1, composta com 60 sequências com valores de células T entre 100 e 300 CD4/mm<sup>3</sup> e D2, composta com 60 sequências com valores de células T acima de 400 CD4/mm<sup>3</sup>, mostrado abaixo (Fig. 11).

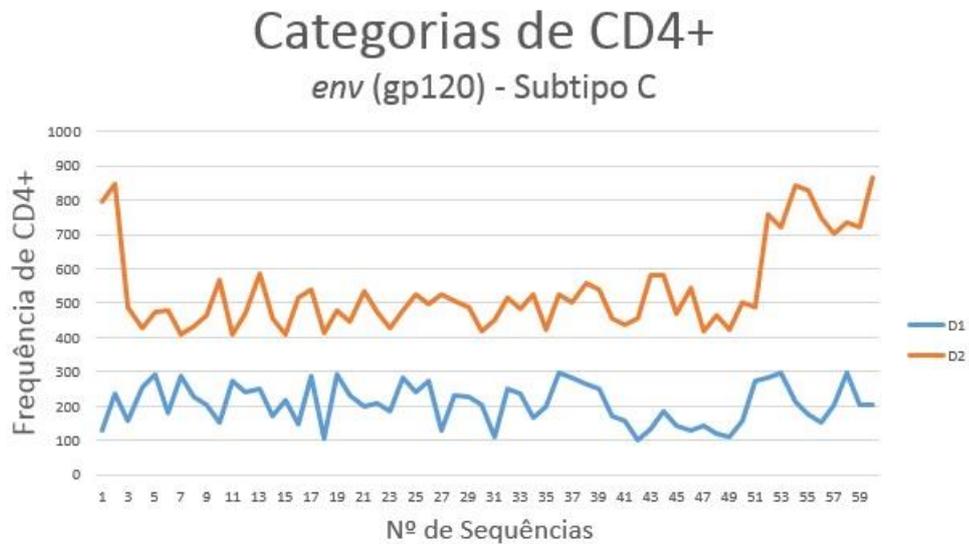


Fig.9 Categorias de CD4+ para o gene *env* (gp120) subtipo C. A linha superior (laranja) representa os níveis mais altos de CD4+ (>400CD4+), e a linha inferior (azul), níveis mais baixos de CD4+ (100-300CD4+). O eixo Y representa a frequência de CD4+ e o eixo X, o número de sequências.

Para o gene *vif*, 2 categorias respectivamente com 62 sequências em cada: D1, com valores de células T entre 100 e 300CD4/mm<sup>3</sup> e D2 com valores de células T acima de 400CD4/mm<sup>3</sup>, como mostrado abaixo (Fig. 12).

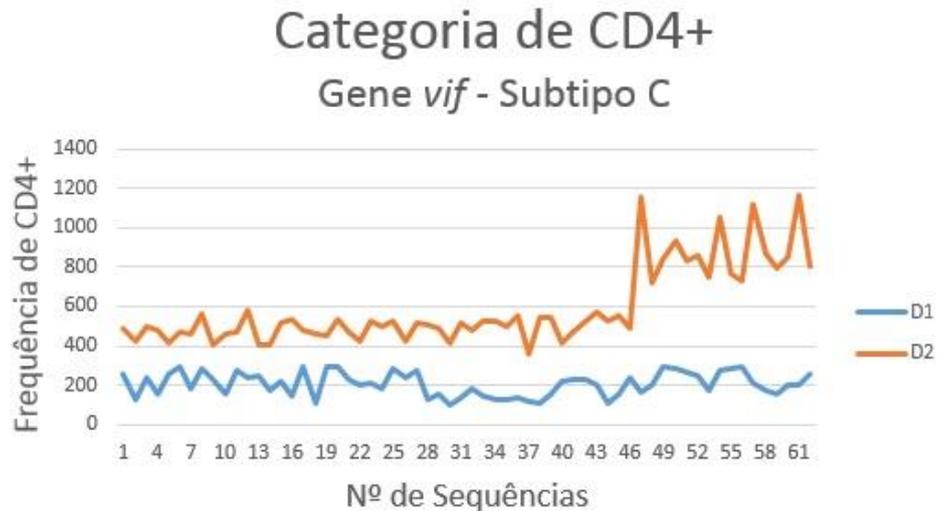


Fig.10 Categorias de CD4+ para o gene *vif* subtipo C. A linha superior (laranja) representa os níveis mais altos de CD4+ (>400CD4+), e a linha inferior (azul), níveis mais baixos de CD4+ (100-300CD4+). O eixo Y representa a frequência de CD4+ e o eixo X, o número de sequências.

### 4.3. Alinhamento e Inferência filogenética

As sequências foram alinhadas manualmente usando o programa ClustalX (2.1) (THOMPSON *et al.*, 1997; CHENNA *et al.*, 2003; LAKIN *et al.*, 2007). Sequências com inserções/exclusões de códons de parada foram excluídas da análise. Conjuntos de sequências dos subtipos do HIV-1 de referência obtidas a partir de GenBank foram incluídos nos alinhamentos. Os alinhamentos foram editados manualmente afim de manter a fase de leitura, esta etapa foi realizada através do programa BioEdit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>).

As árvores de máxima verossimilhança foram geradas pelo programa MEGA, versão 6.1 (KUMAR, TAMURA E NEI, 2013), utilizando o método de *Neighbor-joining* de busca de árvores. Inferências filogenéticas foram feitas utilizando o modelo de substituição de nucleotídeos geral tempo reversível (GTR) (RODRIGUEZ *et al.*, 1990).

### 4.4. Análise da Pressão Seletiva

Foram utilizados os métodos de par-a-par e máxima verossimilhança (ML) códon-específico para estimar as pressões seletivas. Inicialmente, foram compilados os conjuntos de dados, como descrito anteriormente e determinou-se as médias dS e dN utilizando o método de códons, apresentado no programa Hy-Phy (SERGEI *et al.*, 2005), implementado no

MEGA versão 6.1 (KUMAR, TAMURA E NEI, 2013). Esta abordagem usa códons, para calcular a distribuição de dN/dS ( $\omega$ ) entre os sítios e leva em conta as relações filogenéticas entre as sequências.

#### 4.5. Grupos controles

Para gp120, os grupos controles foram feitos juntando todas as sequências de gp120 em um mesmo alinhamento e depois foram selecionadas aleatoriamente para compor os grupos controles 1, 2 e 3, para o subtipo B e controle 1 e 2 para o subtipo C. Nos gráficos, o eixo Y representa as frequências absolutas de dS no quadro à esquerda, e as frequências de dN no quadro à direita, para cada painel respectivamente. As categorias de valores de CD4+ dos genes analisados são representados por D1, D2 e D3 (subtipo B) e D1 e D2 (subtipo C) no eixo X dos gráficos.

Os grupos controles utilizados para o gene *vif*, foram feitos juntando todas as sequências de *vif* em um mesmo alinhamento e depois foram selecionadas aleatoriamente sequências para compor os grupos controles 1 e 2, para o subtipo B e controle 1 e 2 para o subtipo C. Nos gráficos da figura 12, o eixo Y representa as frequências absolutas de dS no quadro à esquerda, e as frequências de dN no quadro à direita. As categorias de valores de CD4+ dos genes analisados são representados por D1 e D2 (subtipo B) e D1 e D2 (subtipo C) no eixo X dos gráficos.

## 5. RESULTADOS

### 5.1 Taxas de Substituições Sinônimas (dS) e Não-Sinônimas (dN) em gp120 do HIV-1

A análise de seleção positiva foi feita utilizando o método de máxima verossimilhança (ML) baseada em códons “*codon-based*”, adicionalmente, como parte do pacote do software HyPhy (SERGEI *et al.*, 2005), implementado no MEGA 6.1 (TAMURA, *et al.*, 2013). Foram estimadas as taxas absolutas de dS e dN nos códons, levando em conta as relações filogenéticas entre as sequências.

Os resultados desta etapa, foram organizados em gráficos do tipo ‘BoxPlot’ para melhor visualização das médias, desvio padrão e os valores não comuns “*outliers*” entre as categorias analisadas. Na figura 11, foram analisados os regimes seletivos no gene *env* (gp120) dos subtipos B e C. No subtipo B, foram comparados em três categorias separadas por valores de células T CD4+ (D1 com valores de 100 a 300 células T CD4/mm<sup>3</sup>; D2 com valores de 400 a 600 CD4/mm<sup>3</sup> e D3 com valores acima de 700 CD4/mm<sup>3</sup>). Comparando os valores entre os grupos D1e D2; D1 e D3 e finalmente entre D2 e D3, no subtipo C foram comparados em duas categorias, também separadas de acordo com valores de células T CD4+ (D1 com valores de 100 a 300 CD4/mm<sup>3</sup>) e (D2 acima de 400 CD4/mm<sup>3</sup>), nas comparações foi usado o Teste *t-Student*, assumindo nível de significância de 5%, implementado no pacote do software Biostat 5.0 (AYRES *et al.*, 2007).

Para as substituições sinônimas (dS) no subtipo B, foi comparado o grupo com valores de CD4 entre 100 e 300 (D1) e o grupo com valores de CD4 entre 400 e 600 (D2), e obtidas, respectivamente, as seguintes médias 8.9544 e 9.1291 e o valor  $p = 0.5643$ , mostrando não haver diferença entres esses dois grupos. Para o grupo D1 e o grupo com níveis de CD4 acima de 700 (D3), as médias obtidas foram 8.5216 e 7.8244 e valor  $p = 0.0013^*$ . Esta comparação mostrou haver diferença entre os grupos D1 e D3. Na comparação os grupos D2 e D3 foram obtidas as seguintes médias: 8.7871 e 7.6807 e o teste estatístico indicou diferença significativa entre os valores de dS entre esses grupos ( $p\text{-value} < 0.0001^*$ ).

Para as substituições não-sinônimas (dN), foi comparado o grupo D1 e D2, que obteve as seguintes médias: D1= 8.0901 e D2= 8.7081, e o valor  $p \leq 0.0001^*$ , mostrando haver diferença entre os grupos. Para os grupos D1 e D3, as médias foram respectivamente 8.1950 e 7.5786 e o valor  $p \leq 0.0001^*$ , mostrando haver diferença entres os grupos. Na comparação dos

grupos D2 e D3, as médias obtidas foram respectivamente: 8.9489 e 7.6869 e o valor  $p = < 0.0001^*$ , também mostrando haver diferença entre os grupos.

No subtipo C, para as substituições sinônimas (dS), foi comparado o grupo com valores médios de CD4 de 100-300 (D1) e o grupo com valores médios de CD4 acima de 400 (D2), e obteve-se as seguintes médias: D1= 5.1536 e D2= 4.8464 e o valor  $p = 0.0843^*$ , mostrando não haver diferença significativa entre os grupos. Para as substituições não-sinônimas (dN), foi comparado o grupo D1 e D2, e obteve-se as seguintes médias: D1= 5.1301 e D2= 4.6992 e o valor  $p = 0.0012^*$  mostrando haver diferença entre os grupos.

Apesar desses resultados significativos para maioria dos grupos, as medidas centrais são muito semelhantes visualmente nos quadros do Box-plot, e apenas os valores não comuns “*outliers*” que diferem entre os grupos analisados, como mostrado abaixo (Fig. 11).

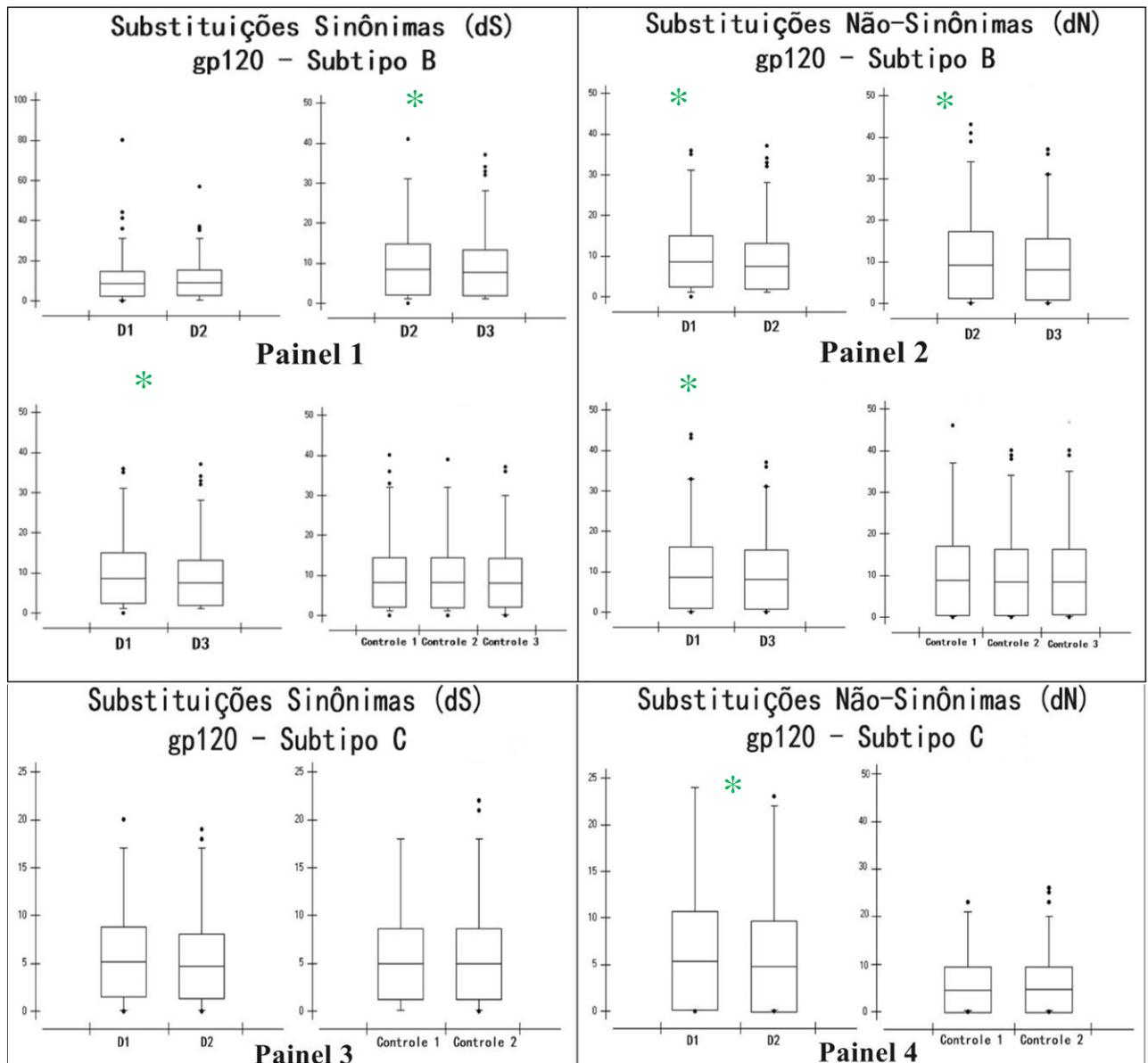


Fig.11 Substituições Sinônimas e Não-Sinônimas no gene *env* (gp120) Subtipo B (Painel 1 e 2) e Subtipo C (Painel 3 e 4). Os quadros representam os valores de limites mínimos e máximos separados pela mediana. Os pontos marcados em acima e abaixo da barra de erro padrão, representam os valores não comuns “*outliers*”. Os resultados significativos estão representados por asteriscos em verde. O eixo Y representa a frequência de Substituições dS e dN e o eixo X representa as categorias de sequência com de valores de CD4+.

## 5.2. Taxas de Substituições Sinônimas (dS) e Não-Sinônimas (dN) do gene *vif* do HIV-1

Para a análise do regime seletivo no gene *vif* do HIV-1, subtipo B as sequências foram igualmente organizadas em categorias de acordo com os níveis de células CD4. Assim o grupo D1 contém 32 sequências obtidas de pacientes com níveis de células T CD4 entre 100 e 300 células/mm<sup>3</sup>, como também o grupo D2, com 32 sequências de pacientes com níveis de células T CD4 acima de 400 células/mm<sup>3</sup>. Os resultados das análises de pressão seletiva para o subtipo B indicaram os valores médios das taxas de substituições sinônimas (dS) de 2.4353 e 2.4471, respectivamente para os grupos D1 e D2. A análise estatística indicou que não existe diferença ( $p= 0.9480$ ) nas taxas de substituições sinônimas entre esses grupos. Igualmente, as médias das taxas de substituições não-sinônimas (dN) foram 2.0125 e 2.1500 para os grupos D1 e D2 e o teste T indicou não haver diferença entre as taxas ( $p= 0.3609$ ), entre esses grupos.

As análises feitas no gene *vif* do HIV-1 do subtipo C as médias calculadas para as taxas de substituições sinônimas (dS) foram 4.3333 e 4.4167 respectivamente para os grupos D1 e D2, com 62 sequências em cada. O teste estatístico indicou não haver diferença ( $p= 0.7325$ ) nas taxas de substituições sinônimas entre os grupos. Da mesma forma o cálculo das taxas de substituições não-sinônimas (dN) obteve médias de 2.6022 e 2.4086 para os grupos D1 e D2 e o valor de P calculado foi de 0.2665, indicando que não existe diferença entre esses grupos. Os valores das taxas de substituições sinônimas (dS) e não-sinônimas significativas (dN) calculados individualmente para os códons do gene *vif* do HIV-1 do subtipo B e do subtipo C estão mostrados de forma gráfica na figura 12.

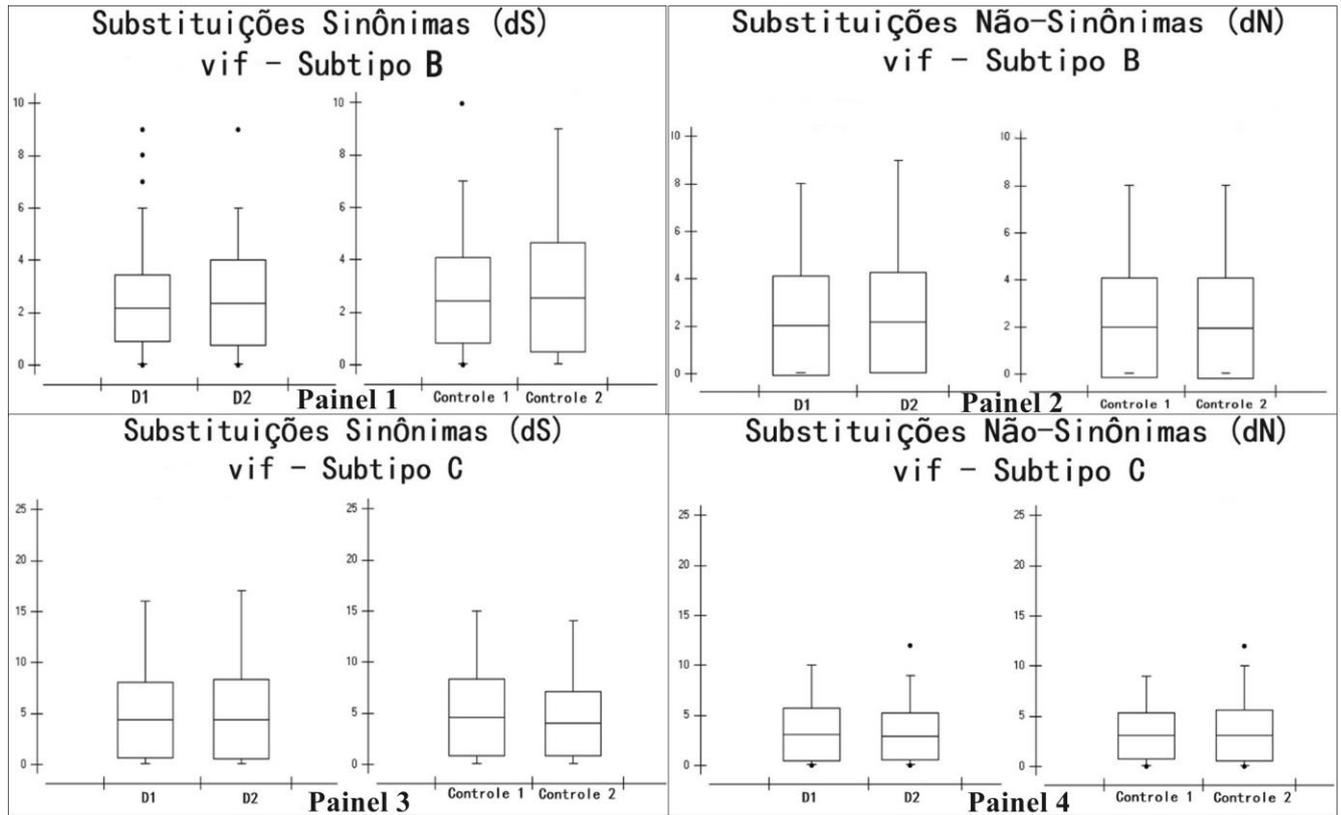


Fig.12 Substituições Sinônimas e Não-Sinônimas no gene *vif* Subtipo B (Painel 1 e 2) e subtipo C (Painel 3 e 4). Os quadros representam os valores de limites mínimos e máximos separados pela mediana. Os pontos marcados em acima e abaixo da barra de erro padrão, representam os valores não comuns “*outliers*”. O eixo Y representa a frequência de Substituições dS e dN e o eixo X representa as categorias de sequência com de valores de CD4+.

## 6. DISCUSSÃO

Neste trabalho foi realizado uma comparação entre os regimes seletivos atuantes nos genes *env* (gp120) e *vif* do HIV-1 e níveis de células T CD4+, através das estimativas de dS e dN, pelo método de máxima verossimilhança baseada em códons, uma vez que estudos anteriores mostraram haver relação proporcional entre maiores taxas evolutivas do vírus no gene do envelope (*env*) e níveis mais altos de células T CD4+ (>600 células/mL) (DIAZ *et al.*, 2008; LEMEY *et al.*, 2007; NOSTROM *et al.*, 2012).

Desta forma, os resultados deste trabalho mostraram não haver relação entre as taxas evolutivas do vírus e níveis de células T CD4+, ou seja, no gene do envelope (subunidade gp120) alguns resultados foram significativos mostrando diferença entre as categorias analisadas como observado no subtipo B (dS): D1 (100 a 300 CD4/mm<sup>3</sup>) e o grupo D3 (>700 CD4/mm<sup>3</sup>), com as médias respectivas de 8.5216 e 7.8244 e valor  $p = 0.0013^*$ , diferente do observado no gene *vif*, que não mostrou diferença para nenhuma das categorias analisadas, como por exemplo, médias de dS do subtipo B: 2.4353 e 2.4471, respectivamente para os grupos D1 e D2 e o valor  $p = 0.9480$ , não significativo.

Apesar dos valores significativos encontrados nas categorias de dS e dN estimados para o gene gp120, estes resultados apresentam códons com valores não comuns “outliers”, os quais talvez estejam interferindo no teste estatístico. Ainda, as médias, mostradas nas linhas centrais dos gráficos, tipo BoxPlot, entres as categorias analisadas variam pouco, logo esses resultados podem não ser reais representativos de diferenças nas taxas de substituições sinônimas e não-sinônimas na unidade gp120 do gene do envelope do HIV-1.

Estudos anteriores de adaptação molecular estimados pelas taxas absolutas de dS e dN têm provado serem ferramentas poderosas (SEO *et al.*, 2004; LEMEY *et al.*, 2007; LEE *et al.*, 2008) para entender a evolução do HIV-1. Lemey *et al* (2007) mostrou que as taxas altas de substituições sinônimas em sequências do gene *env* parecem prever a progressão para AIDS, e que por sua vez indicam aumento na dinâmica de replicação do vírus.

Por abordagem diferente Lee *et al.*, (2008) e também Leal *et al.*, (2012) mostraram que as taxas evolucionárias do HIV-1 intra-hospedeiro não são constantes, e sim menores comparadas ao declínio de células T CD4+, ambos trabalhos mostraram a relação direta entre taxas de evolução do HIV-1 e níveis de linfócitos T CD4+. Os trabalhos de Lemey *et al* (2007) e Diaz *et al* (2008) mostraram que a mesma relação direta entre altos níveis de níveis

de linfócitos T CD4<sup>+</sup> e altas taxas evolutivas do HIV-1 existem quando são analisadas amostras de indivíduos não correlacionados em populações distintas.

Outro achado importante foi mostrado por Yoshida *et al.*, (2011), no estudo entre indivíduos recém infectados e indivíduos infectados há algumas décadas, onde a pressão de seleção era maior em sequências do gene *env* mais antigas do que as mais atuais. No entanto, os autores não avaliaram se existia relação entre pressão seletiva do HIV-1 e os níveis de linfócitos T CD4<sup>+</sup> dos pacientes.

Recentemente Nostrom *et al.*, (2014) mostraram haver uma correlação entre níveis de linfócitos T CD4<sup>+</sup> e taxas de substituições sinônimas no gene *gag* do HIV-1. Nesse estudo foi utilizado sequências do gene *gag* em vez do gene *env*, e ainda, as sequências utilizadas na análise eram apenas de seis indivíduos HLA-B\*5701, os quais apresentam uma interação diferente do hospedeiro com o vírus durante a infecção, em função dessa interação, são associados à lenta progressão para doença (PEREYRA *et al.*, 2010; MIGUELES & CONNORS, 2010; NOSTROM *et al.*, 2012).

Para o gene *vif* nenhuma das categorias analisadas mostraram resultados significativos, semelhantes aos resultados de Bizinoto *et al.*, (2013), sugerindo que os regimes seletivos atuantes em pares de códons do gene *vif*, estejam mais associados com a interação e neutralização da proteína humana APOBEC3G, do que em relação aos níveis de células T CD4<sup>+</sup>.

## 7. CONCLUSÃO

O presente trabalho analisou a relação entre os níveis de células T CD4+ e a pressão seletiva na subunidade gp120 do gene do envelope e no gene *vif* dos subtipos B e C do HIV-1. Os resultados mostram não haver relação entre as taxas de substituições sinônimas e não-sinônimas do HIV-1 e níveis de linfócitos T CD4+ dos pacientes não-correlacionados em abordagem populacional.

Apesar do esperado ter sido uma relação direta entres esses regimes seletivos e níveis de células T CD4+ como mostrado em estudos anteriores (DIAZ *et al.*, 2008; LEMEY *et al.*, 2007; NOSTROM *et al.*, 2014), esses resultados necessitam de estudos posteriores, visando um melhor entendimento da dinâmica replicativa viral inter-hospedeiro.

Além disso, compreender os regimes seletivos atuantes nos genes do HIV-1, através das frequências absolutas de substituições sinônimas e não-sinônimas, só tornará cada vez mais perto a conquista do controle da infecção pelo vírus e progressão para AIDS, uma vez que como meta da UNAIDS/WHO, o controle total da infecção pelo HIV-1 até 2030.

## 8. REFERÊNCIAS

- ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. H.I.V. **Imunologia celular e molecular**. 7. ed. Rio de Janeiro: Elsevier, (2012).
- ACHESON, N. H. **Fundamentals of molecular virology**.—2nd ed. p.; cm. Includes bibliographical references and index. ISBN 978-0-470-90059-8. (2011).
- ALMEIDA SE, DE MEDEIROS RM, JUNQUEIRA DM, GRAF T, PASSAES CP, *et al.* **Temporal dynamics of HIV-1 circulating subtypes in distinct exposure categories in southern brazil**. *Virol J* 9: 306. doi: 10.1186/1743-422x-9-306. (2012).
- AYRES, M., AYRES JÚNIOR, M., AYRES, D.L. & SANTOS, A. A. **BIOESTAT – Aplicações estatísticas nas áreas das ciências bio-médicas**. Ong Mamiraua. Belém, PA. (2007).
- BARRAUD, P. *et al.* **Advances in the structural understanding of Vif proteins**. *Current HIV Research*, v. 6, p. 91–99. (2008).
- BIZINOTO. M, C.; YABE. S.; LEAL. É.; KISHINO, H.; MARTINS, L, DE OLIVEIRA,; DE LIMA, M, L.; MORAIS. R, E.; DIAZ, R, S.; JANINI. L, M,. **Codon pairs of the HIV-1 vif gene correlate with CD4+ T cell count**. *BMC Infectious Diseases*. 13:173 (2013).
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Coordenação Nacional de DST e AIDS. **Boletim Epidemiológico – AIDS**. Ano III, nº 1, 01 à 26ª semanas epidemiológicas, janeiro a junho, Brasília, (2013).
- BRIGGS J. A.; KRÄUSSLICH, H. G. **The molecular architecture of HIV**. *Journal of Molecular Biology*, v. 410 p.491-500. (2011).
- BRINDEIRO RM, *et al.* **Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals**. *AIDS*.;17 (7):1063-9. (2003).
- BUONAGURO L, TORNESELLO ML, BUONAGURO FM. **Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications**. *J Virol* 81: 10209–10219. (2007).
- CHAN DC, FASS D, BERGER JM, KIM PS. **Core structure of gp41 from the HIV envelope glycoprotein**. *Cell* 89: 263–273 (1997).
- CHAN D.C.; KIM, PS. **HIV entry and its inhibition**. *Cell*. v. 93; p. 681-4. (1998).

CHENNA, R., SUGAWARA, H., KOIKE, T., LOPEZ, R., GIBSON, T. J., HIGGINS, D. G. & THOMPSON, J. D. **"Multiple sequence alignment with the Clustal series of programs."** *Nucleic Acids Res* 31(13): 3497-500. (2003).

CHITRA P, BAKTHAVATSALAM B, PALVANNAN T. **Beta-2 microglobulin as an immunological marker to assess the progression of human immunodeficiency virus infected patients on highly active antiretroviral therapy.** *Clin Chim Acta.* 412:1151–1154. doi: 10.1016/j.cca.2011.01.037. (2011).

DA COSTA KS, LEAL E, DOS SANTOS AM, LIMA E LIMA AH, ALVES CN, ET AL. **Structural Analysis of Viral Infectivity Factor of HIV Type 1 and Its Interaction with A3G, EloC and EloB.** *PLoS ONE* 9(2): e89116. doi:10.1371/journal.pone.0089116. (2014).

DIAZ RS. *et al.*, **Selective regimes and evolutionary rates of HIV-1 subtype B V3 variants in the Brazilian epidemic.** *Virology.* 381: 184-193. (2008).

DUMAS F, PREIRA P, AND SALOME L, **"Membrane organization of virus and target cell plays a role in HIV entry,"** *Biochimie*, vol. 107, pp. 22–27, (2014).

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

GANGULI A, WANG J, GOURLEY DR. **Does combining antiretroviral agents in a single dosage form enhance quality of life of HIV/AIDS patients? A cost-utility study.** *Res Social Adm Pharm* 8(2):157–165. (2012).

GARG, H.; MOHL, J.; JOSHI, A. **HIV-1 Induced Bystander Apoptosis.** *Viruses*, 4, 3020-3043. (2012).

IZUMI T, IO K, MATSUI M, SHIRAKAWA K, SHINOHARA M, *et al.* **HIV-1 viral infectivity factor interacts with TP53 to induce G2 cell cycle arrest and positively regulate viral replication.** *Proc Natl Acad Sci U S A* 107: 20798–20803. (2010).

HARRIS RS, PETERSEN-MAHRT SK, NEUBERGER MS. **RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators.** *Mol Cell*; 10:1247–53. (2002)

HE Z, ZHANG W, CHEN G, XU R, YU XF. **Characterization of conserved motifs in HIV-1 Vif required for APOBEC3G and APOBEC3F interaction.** *J Mol Biol* 381: 1000–1011. (2008).

HILL RE, HASTIE ND. **Accelerated evolution in the reactive centre regions of serine protease inhibitors.** *Nature* 326: 96–99. (1987).

HUTCHINSON AB, FARNHAM PG, DEAN HD, EKWUEME DU, DEL RIO C, *et al.* **The economic burden of HIV in the United States in the era of highly active antiretroviral therapy: evidence of continuing racial and ethnic differences.** *Journal of acquired immune deficiency syndromes* 43: 451–457. PMID:16980906 doi: 10.1097/01.qai.0000243090.32866.4e. (2006).

KEELE, B.F., GIORGI, E.E., SALAZAR-GONZALEZ, J.F., DECKER, J.M., PHAM, K.T., SALAZAR, M.G., SUN, C., GRAYSON, T., WANG, S., LI, H., WEI, X., JIANG, C., KIRCHHERR, J.L., GAO, F., ANDERSON, J.A., PING, L.H., SWANSTROM, R., TOMARAS, G.D., BLATTNER, W.A., GOEPFERT, P.A., KILBY, J.M., SAAG, M.S., DELWART, E.L., BUSCH, M.P., COHEN, M.S., MONTEFIORI, D.C., HAYNES, B.F., GASCHEN, B., ATHREYA, G.S., LEE, H.Y., WOOD, N., SEOIGHE, C., PERELSON, A.S., BHATTACHARYA, T., KORBER, B.T., HAHN, B.H., SHAW, G.M. **Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection.** *Proc. Natl. Acad. Sci. USA* 105, 7552–7557. (2008).

KAO S, KHAN MA, MIYAGI E, PLISHKA R, BUCKLER-WHITE A, *et al.* **The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity.** *J Virol* 77: 11398–11407. (2003).

KLEIMAN, L; HALWANI, R; JAVANBAKHT, H. **The selective packaging and annealing of primer tRNA<sup>3</sup> Lys in HIV-1.** *Current HIV Research*, v.2, p. 163-75, (2004).

KNIPE, D. M.; HOWLEY, P. M. **Field's Virology.** 6th ed. Philadelphia, PA, USA: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2 v, 3091p. (2007).

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N.P., CHENNA, R., MCGETTIGAN, P.A., MCWILLIAM, H., VALENTIN, F., WALLACE, I.M., WILM, A., LOPEZ, R., THOMPSON, J.D., GIBSON, T.J., HIGGINS, D.G. **Clustal W and Clustal X version 2.0.** *Bioinformatics*, 23:2947-2948. (2007).

LEAL, E; SILVA, W P.; SUCUPIRA, M C.; JANINI, L. M; DIAZ, R S. **Molecular and structural characterization of HIV-1 subtype B Brazilian isolates with GWGR tetramer at the tip of the V3-loop.** *VIROLOGY*, v. 381, n. 2, p. 222-229, NOV 25. Citações Web of Science: 8. (2008).

LEAL, E; CASSEB, J; HENDRY, M; BUSCH, MICHAEL P.; DIAZ, R S. **Relaxation of Adaptive Evolution during the HIV-1 Infection Owing to Reduction of CD4+T Cell Counts.** *PLoS ONE*, v. 7, n. 6 JUN 29 .Citações Web of Science: 2. (2012).

LEE HY, PERELSON AS, PARK S-C, LEITNER T. **Dynamic Correlation between Intra-host HIV-1 Quasispecies Evolution and Disease Progression.** PLoS Comput Biol 4(12): e1000240. doi:10.1371/journal.pcbi.1000240. (2008).

LEMEY P, KOSAKOVSKY POND SL, DRUMMOND AJ, PYBUS OG, SHAPIRO B, BARROSO H, *et al.* **Synonymous Substitution Rates Predict HIV Disease Progression as a Result of Underlying Replication Dynamics.** PLoS Comput Biol 3(2): e29. doi:10.1371/journal.pcbi.0030029. (2007).

LEVIN, J. G. *et al.* **Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription.** RNA Biology, v.7, p.754-774. (2010).

LI, L. *et al.* **Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions.** Cell Research, v.15, p. 923-934. (2005).

LIHANA RW, SSEMWANGA D, ABIMIKU A, NDEMBI N. **Update on HIV-1 diversity in Africa: a decade in review.** AIDS Rev 14: 83–100. (2012)

LOS ALAMOS HIV DATABANK.

Disponível em: < <http://www.hiv.lanl.gov/content/index>>. Acesso: Junho (2014).

MADIGAN, M. T.; MARTINKO, J. M.; DUNLAP, P. V.; CLARK, D. P. **Microbiologia de Brock.** Porto Alegre: Artmed. 12 ed., 1160 p. (2010).

MARIANI R, CHEN D, SCHROFELBAUER B, NAVARRO F, KONIG R, *et al.* **Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif.** Cell 114: 21–31. (2003).

MEHLE A, THOMAS ER, RAJENDRAN KS, GABUZDA D. **A zinc-binding region in Vif binds Cul5 and determines cullin selection.** J Biol Chem 281: 17259–17265. (2006).

MIGUEL, A. **“VIH e o seu ciclo de vida”.** Revista mundo farmacêutico, setembro/outubro. 42:40-42. (2009).

MURPHY, K.; TRAVERS, P.; WALPORT, M. **Imunobiologia de Janeway.** Porto Alegre: ArtMed, 7ª edição. (2010).

NIELSEN R, YANG Z. **Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene.** Genetics 148: 929–936. (1998).

NORSTRÖM MM, VERAS NM, HUANG W, PROPER MCF, COOK J, *et al.* **Baseline CD4+ T Cell Counts Correlates with HIV-1 Synonymous Rate in HLA-B\*5701 Subjects with Different Risk of Disease Progression.** *PLoS Comput Biol* 10(9): e1003830. doi:10.1371/journal.pcbi.1003830, Set 4 (2014).

OLIVIA O'CONNELL, ALEXANDER REPIK, JACQUELINE D. REEVES, MARIA PAZ GONZALEZ-PEREZ, BRIANA QUITADAMO, ELIZABETH D. ANTON, MARIA DUENAS-DECAMP, PAUL PETERS, RONGHENG LIN, SUSAN ZOLLA-PAZNER, DAVIDE CORTI, AARON WALLACE, SHIXIA WANG, XIANG-PENG KONG, SHAN LU, AND PAUL R. CLAPHAM. **Efficiency of Bridging-Sheet Recruitment Explains HIV-1 R5 Envelope Glycoprotein Sensitivity to Soluble CD4 and Macrophage Tropism.** *J. Virol.* 87:1 187-198; Accepted manuscript posted online 10 October 2012, doi:10.1128/JVI.01834-12. January (2013).

PANTOPHLET R, BURTON DR. **GP120: target for neutralizing HIV-1 antibodies.** *Annu Rev Immunol* 24: 739–769. doi: 10.1146/annurev.immunol.24.021605.090557. (2006).

PAUL, I.; CUI, J.; MAYNARD, E. L. **Zinc binding to the HCCH motif of HIV-1 virion infectivity factor induces a conformational change that mediates protein-protein interactions.** *Proceedings of the National Academy of Sciences of the United States of America*, v. 103, p. 18475-80. (2006).

PESSÔA R, CARNEIRO PROIETTI AB DE F, BUSCH MP, SANABANI SS. **Identification of a Novel HIV-1 Circulating Recombinant Form (CRF72\_BF1) in Deep Sequencing Data from Blood Donors in Southeastern Brazil.** *Genome Announcements.* 2(3):e00386-14. doi:10.1128/genomeA.00386-14. (2014).

RAMBAUT, A. *et al.* **The causes and consequences of HIV evolution.** *Nature Reviews Genetics.* v. 5, p. 52–61. (2004).

RODRIGUEZ, F., OLIVER, J.L., MARIN, A., MEDINA, J.R. **The general stochastic model of nucleotide substitution.** *J. Theor. Biol.* 142 (4), 485–501. (1990).

ROSS HA, RODRIGO AG. **Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration.** *J Virol* 76: 11715–11720. (2002).

SANTOS AF, *et al.* **Epidemiologic and evolutionary trends of HIV-1 CRF31\_BC-related strains in southern Brazil.** *J Acquir Immune Defic Syndr.* 45(3):328-33. (2007).

SEO TK, KISHINO H, THORNE JL. **Estimating absolute rates of synonymous and nonsynonymous nucleotide substitution in order to characterize natural selection and date species divergences.** *Mol Biol Evol* 21: 1201–1213. (2004).

SERGEI L. KOSAKOVSKY POND, SIMON D. W. FROST, AND SPENCER V. MUSE. **HyPhy: hypothesis testing using phylogenies** *Bioinformatics*. 21 (5): 676-679 first published online October 27, 2004 doi:10.1093/bioinformatics/bti079. (2005).

SHARP PM, HAHN BH. **Origins of HIV and the AIDS pandemic**. Cold Spring Harb Perspect Med doi:10.1101/cshperspect.a006841. (2011).

SHEEHY AM, GADDIS NC, MALIM MH. **The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif**. *Nat Med* 9: 1404–1407. (2003).

SIMON, V., *et al.* **HIV/AIDS epidemiology, pathogenesis, prevention, and treatment**. *Lancet*.v. 5 p. 489-504. (2006).

STANLEY, B.J., *et al.* **Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly**. *Journal of Virology*, v. 82, p. 8656-63. (2008).

STOPAK K, DE NORONHA C, YONEMOTO W, GREENE WC. **HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability**. *Mol Cell* 12: 591–601. (2003).

STREBEL K, DAUGHERTY D, CLOUSE K, COHEN D, FOLKS T, MARTIN MA. **The HIV 'A' (sor) gene product is essential for virus infectivity**. *Nature* ; **328**:728–30. (1987).

TAMURA K, STECHER G, PETERSON D, FILIPSKI A, AND KUMAR S. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. *Molecular Biology and Evolution*: 30 2725-2729. (2013).

TAVARÉ S. **"Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences"**. *Lectures on Mathematics in the Life Sciences* (American Mathematical Society) 17: 57–86. (1986).

TEBIT DM, ARTS EJ: **Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease**. *Lancet Infect Dis*, **11**:45-56. (2011).

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):4876–4882. (1997).

UNAIDS. Global Report: **UNAIDS Report on the Global AIDS Epidemic 2013**. <http://www.unaids.org/en/resources/campaigns/globalreport2013/index.html> (2013).

WILEN, CRAIG B., JOHN C. TILTON, AND ROBERT W. DOMS. “**HIV: Cell Binding and Entry.**” *Cold Spring Harbor Perspectives in Medicine* 2.8 (2012): a006866.PMC. Web. 4 July (2015).

WILLIAMSON, S. **Adaptation in the env gene of HIV-1 and evolutionary theories of disease progression.** *Mol. Biol. Evol.* 20, 1318–1325. (2003).

WOLFE, L. S., *et al.* **Dissection of the HIV Vif interaction with human E3 ubiquitin ligase.** *Journal of Virology*, v. 84, p. 7135-7139. (2010).

XIAO, Z. *et al.* **Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-binding domain-stabilized hydrophobic interface in Vif.** *Virology*, v. 349, p. 290-9. (2006).

XIAO, Z., *et al.* **Characterization of a novel Cullin5 binding domain in HIV-1 Vif.** *Journal of Molecular Biology*, v. 373, p. 541-550. (2007).

YANG, Z. **Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution.** *Mol. Biol. Evol.* 15 (5), 568–573. (1998).

YOSHIDA I, SUGIURA W, SHIBATA J, REN F, YANG Z, TANAKA H. **Change of Positive Selection Pressure on HIV-1 Envelope Gene Inferred by Early and Recent Samples.** *PLoS ONE* 6(4): e18630. doi:10.1371/journal.pone.0018630. (2011).

ZHANG, W. *et al.* **Distinct determinants in HIV-1 Vif and human APOBEC3 proteins are required for the suppression of diverse host anti-viral proteins.** *PLoS ONE* 3, v. 3, p. 1-9. (2008).

ZHOU,X. *et al.* **Characterization of the Interaction of Full-Length HIV-1 Vif Protein with its Key Regulator CBFb and CRL5 E3 Ubiquitin Ligase Components.** *PLoS ONE*, v. 7, p. 1-10. (2012).