

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

ANÁLISE DAS CARACTERÍSTICAS GENÉTICAS DA REGIÃO V3 E DO  
TROPISMO PELOS CO-RECEPTORES CCR5 E CXCR4 DO VÍRUS DA  
IMUNODEFICIÊNCIA HUMANA TIPO 1 SUBTIPOS B, C E RECOMBINANTES BC  
ATRAVÉS DE FERRAMENTAS GENOTÍPICAS

ANDRÉA CRISTINA VANNI

Dissertação submetida ao Programa de Pós-Graduação  
em Genética e Biologia Molecular da UFRGS como  
requisito parcial para obtenção do grau de Mestre.

Orientador

Prof. Dr. José Artur Bogo Chies

Porto Alegre

Março, 2011

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

DEPARTAMENTO DE GENÉTICA

INSTITUTO DE BIOCÊNCIAS

LABORATÓRIO DE IMUNOGENÉTICA

UNIVERSIDADE DE CAXIAS DO SUL

CENTRO DE CIÊNCIAS DA SAÚDE

LABORATÓRIO DE PESQUISA EM HIV/AIDS

UNIVERSIDADE FEDERAL DE SÃO PAULO

DEPARTAMENTO DE MEDICINA

DISCIPLINA DE INFECTOLOGIA

LABORATÓRIO DE RETROVIROLOGIA

Dedico este trabalho de conclusão de mestrado à minha mãe Sônia, à minha irmã Débora e ao meu namorado Tiago que me incentivaram e apoiaram durante todo o desenvolvimento deste trabalho.

## AGRADECIMENTOS

Agradeço:

Ao meu orientador, José Artur, por ter aceitado me orientar sem me conhecer previamente, pela atenção, orientação, contribuição científica e principalmente pela compreensão em diversas situações;

Ao pesquisador colaborador Ricardo Diaz por viabilizar a análise das amostras, pela orientação e pela importante contribuição científica;

Ao meu ex-chefe, Ricardo de Souza, onde estiver, por me permitir realizar este trabalho, disponibilizar as amostras, sugerir o tema do trabalho e viabilizar o contato com Ricardo Diaz;

Aos colegas do LPHA, Rosa Dea, Leonardo, Sérgio, Carla, Machline, Luiz Gustavo e Cristina pelo apoio, incentivo, compreensão durante minha ausência e amizade. Em especial à Dea pelos recursos humanos, ao Sérgio pela análise estatística e ao Leonardo pela assessoria científica;

À pesquisadora colaboradora Maria Cecilia por supervisionar o sequenciamento das amostras e pelo auxílio em diversas situações;

À Aline pela revisão da língua inglesa do artigo e aos demais bolsistas do LPHA, Gustavo, Sara e Gabriela, pelo incentivo e amizade;

À Leila, Ana Rachel e Michelle pela execução do sequenciamento das amostras, pela atenção e acolhimento no laboratório da UNIFESP;

Aos colegas de mestrado, em especial à Rúbia e ao Denis pelo auxílio com a análise das sequências e pela amizade;

À amiga Giovana por me hospedar em sua casa nas viagens para Porto Alegre;

À minha família pelo apoio, incentivo, dedicação, amizade, companheirismo, compreensão e repreensão nas horas necessárias.

## SUMÁRIO

|  |    |
|--|----|
| LISTA DE ABREVIATURAS E UNIDADES.....                                      | 6  |
| RESUMO .....   | 8  |
| ABSTRACT .....   | 9  |
| 1 INTRODUÇÃO.....  | 10 |
| 1.1 HISTÓRICO .....  | 10 |
| 1.2 EPIDEMIOLOGIA .....  | 12 |
| 1.3 HISTÓRIA NATURAL DA INFECÇÃO PELO HIV .....                            | 13 |
| 1.4 CLASSIFICAÇÃO DA INFECÇÃO PELO HIV .....                               | 15 |
| 1.5 BIOLOGIA DO HIV.....   | 18 |
| 1.5.1 Nomenclatura do HIV .....  | 18 |
| 1.5.2 Morfologia do HIV-1 .....  | 19 |
| 1.5.3 Organização do Genoma do HIV-1.....                                  | 21 |
| 1.6 CICLO DE REPLICAÇÃO DO HIV-1 .....                                     | 21 |
| 1.7 CO-RECEPTORES CCR5 E CXCR4 .....                                       | 25 |
| 1.7.1 CCR5 $\Delta$ 32 .....   | 27 |
| 1.8 TROPISMO DO HIV-1 POR CO-RECEPTORES.....                               | 28 |
| 1.9 INIBIDORES DA ENTRADA CELULAR DO HIV-1 .....                           | 33 |
| 1.10 ORIGEM E DISTRIBUIÇÃO DO HIV-1 SUBTIPOS B, C E RECOMBINANTES BC ..... | 35 |
| 1.11 TROPISMO DO HIV-1 SUBTIPOS B, C E RECOMBINANTES BC .....              | 39 |
| 1.12 DETERMINAÇÃO DO TROPISMO VIRAL POR CO-RECEPTORES CCR5 E CXCR4.....    | 41 |
| 1.12.1 Caracterização Genética da Região V3.....                           | 43 |
| 1.12.2 Predição Genotípica do Tropismo por Co-receptores.....              | 45 |
| 1.13 JUSTIFICATIVA .....   | 48 |
| 2 MANUSCRITO.....  | 50 |
| 3 CONCLUSÕES GERAIS E PERSPECTIVAS .....                                   | 77 |
| 4 REFERÊNCIAS BIBLIOGRÁFICAS .....   | 80 |

## LISTA DE ABREVIATURAS E UNIDADES

AIDS: Síndrome da Imunodeficiência Adquirida (do inglês *Acquired Immunodeficiency Syndrome*)

AZT: Azidotimidina

CCL: Ligante quimioquímico (C-C motivo) [do inglês *chemokine (C-C motif) ligand*]

CCR: Receptor quimioquímico (C-C motivo) [do inglês *chemokine (C-C motif) receptor*]

CDC: Sigla originada do inglês *Centers for Disease Control and Prevention*

cDNA: DNA complementar (do inglês *complementary DNA*)

CRF: Forma recombinante circulante (do inglês *circulating recombinant form*)

CXCR: Receptor quimioquímico (CXC motivo) [do inglês *chemokine (CXC motif) receptor*]

DNA: Ácido desoxirribonucléico (do inglês *deoxyribonucleic acid*)

EDTA: Ácido etilendiaminotetracético (do inglês *ethylenediaminetetraacetic acid*)

FIV: Vírus da Imunodeficiência Felina (do inglês *Feline Immunodeficiency Virus*)

gp: Glicoproteína

HIV: Vírus da Imunodeficiência Humana (do inglês *Human Immunodeficiency Virus*)

HTLV: Vírus T Linfotrófico Humano (do inglês *Human T-lymphotropic Virus*)

Kb: Quilobase (do inglês, *kilobase*)

LTR: Repetições terminais longas (do inglês *long terminal repeats*)

M: Molar

mL: Mililitro

μL: Microlitro

nm: Nanômetro

NSI: Não indutor de sincício (do inglês *non-syncytium-inducing*)

## LISTA DE ABREVIATURAS E UNIDADES

OMS: Organização Mundial da Saúde

pb: Pares de bases

PBMC: Célula mononuclear do sangue periférico (do inglês *peripheral blood mononuclear cell*)

PCR: Reação em cadeia da polimerase (do inglês *polymerase chain reaction*)

PSSM: Matriz de escore de posição específica (do inglês *position-specific scoring matrices*)

NIH: Sigla originada do inglês *National Institutes of Health*

RF: Sigla originada do inglês *random forest*

RNA: Ácido ribonucléico (do inglês *ribonucleic acid*)

SDF: Fator derivado de célula estromal (do inglês *stromal cell-derived factor*)

SI: Indutor de sincício (do inglês *syncytium-inducing*)

SIV: Vírus da Imunodeficiência Símia (do inglês *Simian Immunodeficiency Virus*)

SVM: Máquina de vetores de suporte (do inglês *support vector machine*)

TBE: Tris-borato-EDTA

URF: Forma recombinante única (do inglês *unique recombinant form*)

V: Volts

WHO: Sigla originada do inglês *World Health Organization*

## RESUMO

Este estudo analisou as características genéticas da região V3 e o uso de co-receptores dos subtipos de HIV-1 mais prevalentes no Rio Grande do Sul através de ferramentas genotípicas. Amostras de plasma de 105 pacientes infectados com os subtipos B (35), C (35) ou recombinantes BC (35) com base na sequência do gene *pol* foram selecionadas. Após *nested* PCR, a região V3 do gene *env* foi sequenciada e o tropismo foi determinado através das ferramentas genotípicas Geno2pheno\_coreceptor<sub>10%</sub>, WebPSSM e R5/X4-Pred. Um total de 90 amostras foram eficientemente amplificadas e sequenciadas. A análise do subtipo da região do gene *env* identificou 37 B e 53 C. A maior proporção de cepas X4 entre o subtipo B foi 16,2% (predita pela Geno2pheno<sub>10%</sub>) e entre o subtipo C foi 32,1% (predita pela WebPSSM<sub>SI/NSI(C)</sub>). A prevalência de cepas X4 em infecções recentes foi 18,7%, sendo identificadas exclusivamente no subtipo C. A análise da região V3 revelou diferenças e similaridades entre as sequências do subtipo B e C. Para ambos os subtipos o comprimento da alça V3 se mostrou conservado, sendo que a pequena variabilidade observada foi associada com uso de CXCR4. A perda do sítio de N-glicosilação foi raramente observada e não foi relacionada com o tropismo viral. Carga total aumentada e a presença de aminoácidos básicos nas posições 11/25 foram associadas ao uso de CXCR4 apenas no subtipo B. Entre as cepas do subtipo B, os tetrâmeros observados foram GPGR (43,2%), GWGR (16,2%), GFGR (16,2%) e APGR (13,5%). O tetrâmero GPGQ esteve presente em 84,9% dos isolados do subtipo C e o tetrâmero GPGR foi associado com o uso de CXCR4 nesse subtipo. Estes resultados reforçam prévias observações de diferenças subtipo-específicas no uso de co-receptores pelo HIV-1 e corrobora os relatos de evolução do HIV-1 subtipo C para uso de CXCR4.



## ABSTRACT

This study analyzed the genetic characteristics of the V3 region and the co-receptor usage of the main HIV-1 subtypes from Rio Grande do Sul through genotypic tools. Plasma samples from 105 patients known to be infected with subtypes B (35), C (35) or BC (35) recombinants based on *pol* sequences were selected. After nested PCR, the V3 region of *env* gene was sequenced and the tropism inferred using the genotypic tools Geno2pheno\_coreceptor<sub>10%</sub>, WebPSSM and R5/X4-Pred. A total of 90 samples were successfully amplified and sequenced. Subtyping analyses of *env* gene identified 37 B and 53 C. The higher X4 strains proportions among subtype B was 16.2% (predicted by Geno2pheno<sub>10%</sub>) and among subtype C was 32.1% (predicted by WebPSSM<sub>SI/NSI(C)</sub>). X4 prevalence in recent seroconverters was 18.7%, being exclusively identified in subtype C. Analysis of the V3 region revealed both differences and similarities among subtype B and C sequences. Both subtypes showed a conserved V3 length and the small variability observed was associated with CXCR4 usage. Loss of the *N*-glycosylation site was rarely observed and did not correlate to viral tropism. Increased net charge and basic amino acids at positions 11/25 were associated to CXCR4 usage only in subtype B isolates. Among subtype B strains, the tetramers observed were GPGR (43.2%), GWGR (16.2%), GFGR (16.2%) and APGR (13.5%). The tetramer GPGQ was present in 84.9% of subtype C sequences and the GPGR motif was associated with CXCR4 usage in subtype C. These finds reinforce previous observations of subtype-specific differences in HIV-1 co-receptor usage and corroborate reports of evolution of subtype C viruses for CXCR4 usage.

## 1 INTRODUÇÃO

Com mais de 30 milhões de óbitos desde sua descoberta, a Síndrome da Imunodeficiência Adquirida (AIDS, do inglês *Acquired Immunodeficiency Syndrome*) é uma das maiores pandemias da história (UNAIDS, 2008; UNAIDS, 2010). O agente etiológico da AIDS é o Vírus da Imunodeficiência Humana (HIV, do inglês *Human Immunodeficiency Virus*), cujo alvo principal é o linfócito T CD4<sup>+</sup>. Este linfócito coordena diversas funções imunológicas importantes e a perda dessas funções resulta no enfraquecimento progressivo da resposta imune, produzindo o quadro de AIDS (CDC, 1993).

### 1.1 Histórico

Em 1981 surgiram as primeiras descrições de imunodeficiência com causa desconhecida em homens homossexuais jovens, que se caracterizavam pelo aparecimento de doenças raras como o câncer Sarcoma de Kaposi e infecções oportunistas como, por exemplo, *Pneumocystis carinii* e *Candida albicans*. Essa nova doença se estendeu rapidamente para outras populações: usuários de drogas intravenosas, recém-nascidos, hemofílicos e receptores de transfusão sanguínea (Prusiner, 2002; Hoffmann *et al*, 2007). Finalmente em 1983, o grupo do pesquisador Dr. Luc Montagnier isolou um retrovírus a partir de uma cultura de linfócitos de um paciente com linfadenopatia e chamou-o de vírus associado à imunodeficiência (Montagnier, 2002). Em 1984 o grupo do Dr. Robert Gallo, dois anos e meio após a identificação da AIDS, evidenciou que aquele retrovírus é a causa da AIDS. O Dr. Robert Gallo também descreveu métodos de isolamento, produção

contínua e análise protéica do vírus que levaram ao desenvolvimento do primeiro teste de diagnóstico sanguíneo do, hoje chamado, HIV (Gallo, 2002).

A partir desse período, houve grandes avanços nas pesquisas. Em 1985 o genoma do HIV foi clonado e sequenciado. Em seguida, a variação antigênica do HIV foi descrita, variações de sequência gênica foram encontradas em populações virais do mesmo paciente, macrófagos foram identificados como alvos do vírus, vários modos de transmissão viral foram elucidados, todos os genes do HIV e a maioria de suas proteínas foram definidos, e o CD4 foi identificado como receptor do HIV. Em 1986 o HIV-2 foi isolado de pacientes do oeste da África. Em 1987 foi desenvolvido o primeiro medicamento anti-HIV, azidotimidina (AZT), e em 1995 surgiram os inibidores da protease e a terapia tripla de drogas (Gallo, 2002; Montagnier, 2002).

Duas descobertas independentes entre 1995 e 1996 abriram uma nova área de pesquisa em AIDS: uma delas foi a identificação de fatores supressores do HIV liberados por células T CD8<sup>+</sup>, identificados como quimiocinas inflamatórias CC, que poderiam inibir a replicação do HIV-1, e a outra foi a identificação de um segundo receptor (co-receptor) utilizado preferencialmente por algumas cepas, com características moleculares de receptor quimiocínico, chamado CXCR4 [do inglês *Chemokine (CXC motif) receptor*] (Moore *et al.*, 2004; Lusso, 2006). Essas descobertas propiciaram diversos avanços: a identificação do principal co-receptor do HIV, o CCR5 [do inglês *Chemokine (C-C motif) receptor*]; a evidência de resistência de base genética à infecção pelo HIV produzida por uma deleção de 32 pares de bases no gene do CCR5; e a identificação da quimiocina SDF-1 como um ligante específico do CXCR4. Esses e subseqüentes avanços abriram novas perspectivas para desenvolvimento de novas medidas para o controle da AIDS (Lusso, 2006).

## 1.2 Epidemiologia

A epidemia mundial do HIV está estabilizada desde 2000, porém em um nível muito elevado. Em 2009 estimou-se que 33,3 milhões de pessoas eram portadoras do HIV, correspondendo a um aumento de 27% da estimativa observada em 1999 (26,2 milhões). A África Subsaariana permanece sendo a região mais afetada pela epidemia, sendo responsável por 68% da população mundial infectada e 72% dos óbitos por AIDS em 2009 (UNAIDS, 2008; UNAIDS, 2010).

A incidência da infecção por HIV declinou 19% entre 1999 e 2009 globalmente. O número anual de novas infecções diminuiu gradualmente como resultado da implantação de programas de prevenção e tratamento em diversos países: em 1997 a estimativa foi de 3,2 milhões de novas infecções, em 1999 foi de 3,1 milhões, em 2001 diminuiu para 3 milhões, em 2007 declinou para 2,7 milhões e em 2009 estimou-se que 2,6 milhões de pessoas se infectaram pelo HIV no mundo. Após décadas de aumento da mortalidade por AIDS, o número de óbitos também está diminuindo gradualmente, decaindo de 2,1 milhões em 2004 para 1,8 milhões em 2009, como reflexo da rápida expansão da disponibilidade de tratamento antirretroviral, bem como, suporte e cuidado de indivíduos infectados. Em apenas seis anos o número de pessoas recebendo tratamento antirretroviral aumentou dez vezes em países de baixa e média renda, alcançando 3 milhões de indivíduos no final de 2007. Em 2009 esse número elevou-se para 5,2 milhões de pacientes recebendo tratamento, correspondendo a um aumento de 30% em relação a 2008 (UNAIDS, 2008; UNAIDS, 2010).

O Brasil possui cerca de um terço da população infectada da América do Sul e Central, cuja estimativa em 2009 foi de 1,4 milhões de indivíduos portadores. O país

possui um índice elevado de pacientes em tratamento antirretroviral (50 - 89%) e em 2007 estava entre os 20 países que possuíam cobertura de 50 - 75% de terapia para adultos e crianças com HIV na fase avançada. Como reflexo dessa cobertura, estima-se que entre 1996 e 2009 o Brasil ganhou mais de 1,2 milhões de anos em expectativa de vida (UNAIDS, 2008; UNAIDS, 2010).

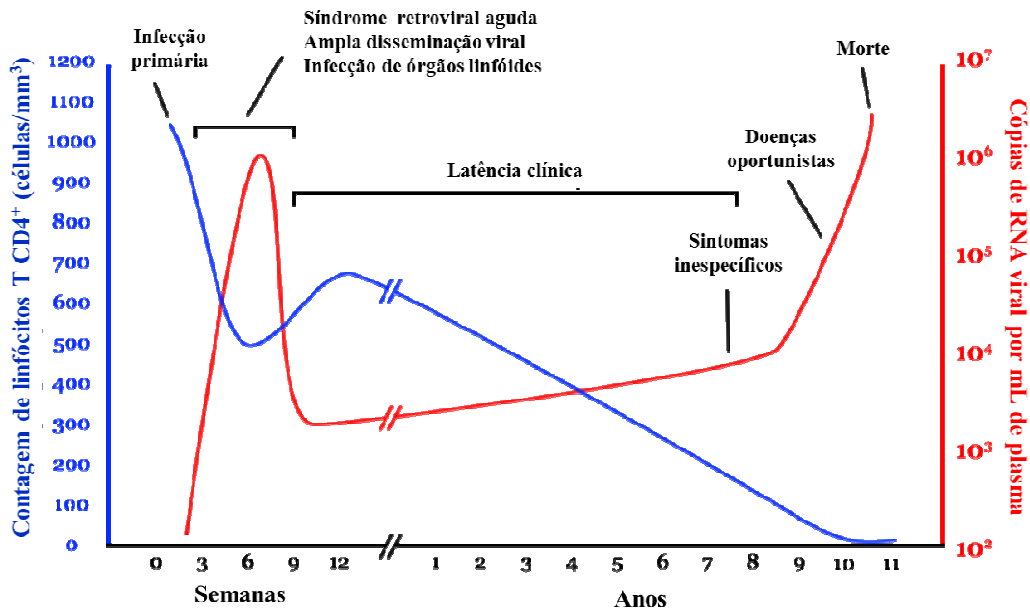
No período entre 1980 e metade de 2009 foram notificados 544.846 casos de AIDS no Brasil, com uma média de 34.838 casos/ano nos últimos cinco anos. Destes casos, aproximadamente 84,6% são indivíduos com faixa etária entre 20 e 49 anos e aproximadamente 65,4% pertencem ao sexo masculino. Até 2008 foram registrados 217.091 óbitos por AIDS no país, sendo que 19.490 ocorreram no Rio Grande do Sul. O número de notificações no Rio Grande do Sul correspondeu a 9,6% do registro nacional no mesmo período (Ministério da Saúde, 2009).

### 1.3 História Natural da Infecção pelo HIV

A história natural da infecção pelo HIV é definida como a progressão da doença, sem tratamento antirretroviral, desde a transmissão do vírus até a morte do indivíduo. O curso da infecção é caracterizado por três fases: fase aguda ou infecção primária, fase assintomática e AIDS. Durante a fase aguda, que corresponde ao período desde transmissão até o desenvolvimento de anticorpos anti-HIV, observam-se níveis extremamente elevados de viremia e decréscimo acentuado dos linfócitos T CD4<sup>+</sup> (Figura 1). Nessa fase os vírus se replicam extensivamente pela ausência de resposta imune adaptativa. A infecção aguda é uma fase bastante heterogênea e frequentemente sintomática, produzindo sintomas inespecíficos, similares a gripe, de uma a duas semanas

de duração, incluindo úlceras orais, febre, erupções maculopapulares, linfadenopatia, faringite, perda de peso, artralgia, indisposição, meningite asséptica e mialgia. Após duas a quatro semanas de infecção, o surgimento da resposta imune celular específica e subsequente produção de anticorpos anti-HIV leva ao declínio da carga viral. Os níveis de linfócitos T CD4<sup>+</sup> se expandem, porém não alcançam os níveis pré-infecção sem tratamento antirretroviral. Na fase assintomática, após três a quatro meses de infecção, o equilíbrio entre a replicação viral e a resposta do sistema imunológico é alcançado e as manifestações clínicas da doença podem não ocorrer por anos. Mesmo na ausência de terapia antirretroviral, o período de latência clínica pode se estender por mais de dez anos. Durante o curso da latência clínica, o número de linfócitos T CD4<sup>+</sup> decresce continuamente e sintomas não indicativos de AIDS podem aparecer, como manifestações imunológicas, hematológicas e dermatológicas leves. Quando os níveis de linfócitos T CD4<sup>+</sup> decaem para menos de 200 células/ $\mu$ L, emergem diversas doenças oportunistas indicativas de AIDS (Figura 1) (Sierra *et al.*, 2005; Hoffmann *et al.*, 2007; Qu *et al.*, 2008).

O curso da infecção pelo HIV pode variar drasticamente: indivíduos que apresentam rápido declínio de células T CD4<sup>+</sup> e progridem para AIDS em dois a cinco anos de infecção são chamados progressores rápidos; progressores típicos normalmente progridem para AIDS dentro de oito a dez anos; pacientes definidos como não progressores de longa data apresentam lenta diminuição de células T CD4<sup>+</sup>, mantendo níveis normais por mais de dez anos, viremia baixa e eventual progressão da doença; um pequeno grupo, controladores de elite (*elite controllers*), se caracteriza por apresentar carga viral indetectável persistente na ausência de terapia antirretroviral, em alguns casos por mais de 25 anos (Qu *et al.*, 2008).



**Figura 1.** Diagrama da história natural do HIV (Hoffmann *et al.*, 2007).

#### 1.4 Classificação da Infecção pelo HIV

Para monitoramento da epidemia do HIV, o CDC (do inglês, *Centers for Disease Control and Prevention*) e a Organização Mundial da Saúde (OMS) desenvolveram, independentemente, sistemas de definição de caso e classificação da infecção pelo HIV. Os sistemas visam gerar dados comparáveis e padronizados para vigilância em saúde pública, que considerem todos os estágios da doença, para direcionar investimentos e ações de prevenção e tratamento (WHO, 2007; CDC, 2008).

Definição de caso: ambos os órgãos consideram como critério para definir um caso como infecção por HIV a confirmação laboratorial, não sendo suficiente o diagnóstico baseado em critérios clínicos, incluindo condições indicativas de AIDS, e imunológicos (contagem de linfócitos T CD4<sup>+</sup>), como era preconizado há algum tempo (CDC, 1993;

WHO, 2007; CDC 2008). O CDC recomenda como critério laboratorial para adultos e adolescentes ( $\geq 13$  anos) dois resultados positivos por dois testes diferentes de detecção de anticorpos anti-HIV ou um resultado positivo por um teste virológico (detecção de ácido nucléico, antígeno p24 ou isolamento viral) (CDC, 2008). A OMS recomenda como critério para adultos e crianças ( $\geq 18$  meses) dois resultados positivos por dois testes diferentes de detecção de anticorpos anti-HIV e/ou dois resultados positivos por dois testes virológicos diferentes (WHO, 2007).

O CDC classifica a infecção pelo HIV em estágios de acordo com a sintomatologia clínica e o estado imunológico do paciente. A classificação para adultos e adolescentes ( $\geq 13$  anos) contém quatro estágios: 1, 2, 3 e desconhecido, que são definidos de acordo com a contagem de linfócitos T CD4<sup>+</sup> e ocorrência de condições clínicas indicativas de AIDS. A severidade da doença aumenta do estágio 1 até o estágio 3, correspondente ao quadro de AIDS. Para fins de vigilância epidemiológica, a progressão da doença é classificada de menos severa para a mais severa, sendo que uma vez classificado em um estágio severo, um caso não pode ser reclassificado em um estágio menos severo.

Estágio 1: nenhuma condição clínica indicativa de AIDS e contagem de linfócitos T CD4<sup>+</sup>  $\geq 500$  células/ $\mu$ L ou porcentagem  $\geq 29\%$  dos linfócitos totais.

Estágio 2: nenhuma condição clínica indicativa de AIDS e contagem de linfócitos T CD4<sup>+</sup> entre 200 e 499 células/ $\mu$ L ou porcentagem entre 14 e 28% dos linfócitos totais.

Estágio 3 (AIDS): uma condição clínica indicativa de AIDS ou contagem de linfócitos T CD4<sup>+</sup>  $< 200$  células/ $\mu$ L ou porcentagem  $< 14\%$  dos linfócitos totais.

Estágio desconhecido: condições clínicas do paciente e contagem de linfócitos T CD4<sup>+</sup> indisponíveis (CDC, 2008).



A OMS classifica a infecção pelo HIV em quatro estágios clínicos e quatro imunológicos. O estágio clínico em adultos ( $\geq 15$  anos) é classificado de acordo com a ocorrência de sintomas relacionados ao HIV, progredindo do estágio 1 até o 4 conforme a severidade da doença. Um caso pode ser reclassificado para um estágio menos severo, uma vez que pacientes que iniciam terapia antirretroviral podem reverter e melhorar sua condição clínica.

Estágio clínico 1: assintomático.

Estágio clínico 2: sintomas moderados.

Estágio clínico 3: sintomas avançados.

Estágio clínico 4: sintomas severos.

A classificação imunológica é baseada na contagem de linfócitos T CD4<sup>+</sup>. Para indivíduos com idade superior a 5 anos, os valores de contagem celular para os quatro estágios são:

Ausência de imunodeficiência ou não significativa:  $> 500$  células/ $\mu\text{L}$ .

Imunodeficiência moderada: 350 – 499 células/ $\mu\text{L}$ .

Imunodeficiência avançada: 200 – 349 células/ $\mu\text{L}$ .

Imunodeficiência severa:  $< 200$  células/ $\mu\text{L}$ .

Além do propósito de vigilância epidemiológica, a OMS recomenda a utilização das classificações clínica e imunológica no manejo clínico do paciente como guia para intervenções, incluindo início, monitoramento e interrupção de terapia antirretroviral (WHO, 2007).

## 1.5 Biologia do HIV

O HIV é membro do gênero lentivírus e pertence à família dos retrovírus, os quais se caracterizam por possuir um genoma de RNA (do inglês *Ribonucleic acid*) e um capsídeo em forma de cone (Turner e Summers, 1999; ICTVdB, 2006). Infecções por lentivírus caracteristicamente levam a doença crônica, apresentam longos períodos de latência clínica, replicação viral persistente e envolvimento do sistema nervoso central. Outros exemplos de lentivírus incluem o Vírus da Imunodeficiência Símia (SIV, do inglês *Simian Immunodeficiency Virus*), o Vírus da Imunodeficiência Felina (FIV, do inglês *Feline Immunodeficiency Virus*), o visna vírus e o vírus da anemia infecciosa equina (Turner e Summers, 1999; Hoffmann *et al.*, 2007).

### 1.5.1 Nomenclatura do HIV

Similar a outros retrovírus, o HIV apresenta um alto grau de diversidade gênica e é classificado em tipos, grupos, subtipos, subsubtipos, CRFs (do inglês, *Circulating recombinant form*) e URFs (do inglês, *Unique recombinant form*). Esta diversidade é reflexo da alta taxa de mutação, rápida renovação viral e eventos frequentes de recombinação (Robertson *et al.*, 2000). A grande maioria das infecções mundiais é produzida pelo HIV tipo 1, enquanto que o HIV tipo 2 se restringe a países do oeste da África, algumas regiões da Europa e sudoeste da Índia, sendo que Portugal tem a maior prevalência de infecções por HIV-2 na Europa (10-13% do total de infecções). As principais diferenças do HIV-2 incluem menor patogenicidade, virulência e taxa de transmissão e maior controle imunológico da infecção (Reeves e Doms, 2002).

O HIV tipo 1 é classificado em 3 grupos: M, N e O. Recentemente foi proposta a inclusão do grupo P (Los Alamos National Laboratory, 2010). O grupo M, responsável pela epidemia global da AIDS, é classificado em 9 subtipos (A-D, F-H, J e K), 4 subsubtipos (A1, A2, F1 e F2), 49 CRFs e inúmeras URFs. Os subtipos representam diferentes linhagens de HIV e possuem alguma associação geográfica. Os sub-subtipos correspondem a distintas linhagens que não são geneticamente distantes o suficiente para serem consideradas um novo subtipo. Recombinações entre subtipos são eventos comuns e quando cepas recombinantes são transmitidas e se tornam circulantes são classificadas como CRF (Robertson *et al.*, 2000; Los Alamos National Laboratory, 2010). O grupo N, referido como grupo novo, foi identificado em apenas alguns indivíduos de Camarões. O grupo O, assim como o grupo M, apresenta grande diversidade, porém raramente é encontrado (Los Alamos National Laboratory, 2010). O HIV-2 é dividido em sete grupos (anteriormente denominados subtipos): A até G (Reeves e Doms, 2002).

O HIV na espécie humana provavelmente originou-se da transmissão zoonótica do Vírus da Imunodeficiência Símia de primatas. O HIV-1 dos grupos M e N estão mais relacionados com SIV de chimpanzés, o HIV-1 do grupo O com SIV de gorilas e o HIV-2 com SIV de macacos da espécie *sooty mangabey* (Reeves e Doms, 2002; Los Alamos National Laboratory, 2010).

### 1.5.2 Morfologia do HIV-1

As partículas virais do HIV-1 tem um diâmetro de 80 a 100 nm e são constituídas por um envelope, uma matriz protéica, um capsídeo e um nucleocapsídeo (ICTVdB, 2006). O envelope é composto por uma bicamada fosfolipídica que é derivada da membrana de

células hospedeiras. Ancoradas ao envelope encontram-se aproximadamente 72 unidades do complexo protéico gp160, que se projetam na superfície viral. A gp160 é constituída pela glicoproteína externa gp120 e pela glicoproteína transmembrana gp41, a qual ancora gp120 ao envelope. A matrix protéica é constituída pela proteína p17 e reveste a superfície interna da membrana viral, circundando o capsídeo, o que assegura a integridade da partícula viral. O capsídeo é constituído pela proteína p24, está localizado no centro do vírus e circunda o genoma viral e três enzimas essenciais: protease, transcriptase reversa e integrase. O material genético é estabilizado em um complexo ribonucleoprotéico que contém a proteína p7. A proteína p6 é necessária para os últimos estágios da montagem da partícula viral e liberação da célula hospedeira. A partícula viral também contém as proteínas acessórias Nef, Vif, Vpr, Rev, Tat e Vpu (Turner e Summers, 1999; Kuiken *et al.*, 2010).

A gp120 é o principal alvo de anticorpos neutralizantes e contém sítios de ligação para o receptor CD4 e para co-receptores quimiocínicos (Sierra *et al.*, 2005). Ela é extremamente glicosilada e contém cinco regiões variáveis (V1-V5), das quais quatro (V1-V4) formam alças que são expostas na superfície. Essa variabilidade da superfície provavelmente é importante para o escape do sistema imune (Turner e Summers, 1999). A estrutura do centro da gp120 consiste de dois domínios, um interno e um externo, e quatro pontes de folhetos- $\beta$  antiparalelos. O domínio interno contém as regiões N e C-terminal e V1-V2. O domínio externo contém V3, V4 e V5 e é ligado ao domínio interno pelos folhetos- $\beta$ . As regiões N e C-terminais são direcionadas para a membrana viral, enquanto que os sítios de ligação a receptores V1/V2, folhetos- $\beta$  e alça V3, são direcionados para a célula hospedeira (Turner e Summers, 1999; Sierra *et al.*, 2005; Lusso, 2006).

### 1.5.3 Organização do Genoma do HIV-1

O genoma viral é constituído por duas cópias de RNA de fita simples que codificam nove genes, denominados *gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, e *vpu*. Três desses genes são essenciais para a replicação viral, pois codificam as proteínas estruturais da partícula do retrovírus: *gag*, que significa “group-antigen”, codifica as proteínas do nucleocapsídeo, *pol*, que significa polimerase, codifica as enzimas protease, transcriptase reversa, integrase e RNase H, e *env*, que significa envelope, codifica a gp160, a qual é processada em gp120 e gp41. Os demais genes codificam proteínas reguladoras da expressão gênica viral e proteínas acessórias (Hoffmann *et al.*, 2007; Kuiken *et al.*, 2010).

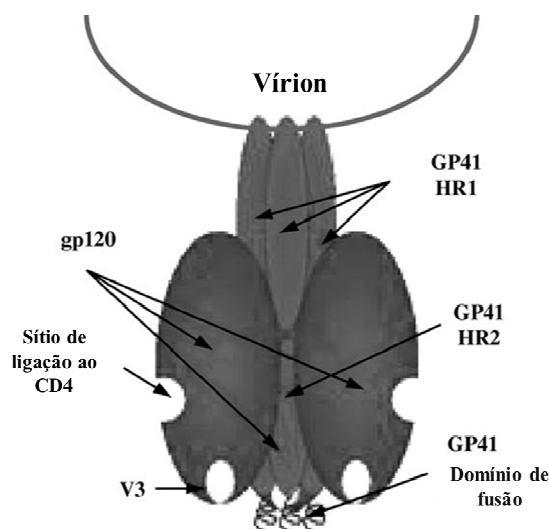
A estrutura clássica do genoma de um retrovírus é: 5’LTR-gag-pol-env-LTR3’. As regiões de LTR (do inglês *Long terminal repeat*) representam a parte inicial e final do genoma viral que são conectadas ao DNA (do inglês *Deoxyribonucleic acid*) da célula hospedeira após integração e não codificam proteína viral (Hoffmann *et al.*, 2007; Kuiken *et al.*, 2010). Cada LTR possui três regiões: U3, R e U5, que contêm sítios de ligação de fatores de transcrição celular e elementos de resposta a trans-ativação mediada por Tat (Sierra *et al.*, 2005).

### 1.6 Ciclo de Replicação do HIV-1

A infecção da célula hospedeira inicia com a ligação do vírus à superfície celular mediada pela interação da gp120 primeiramente com o receptor celular CD4, principal receptor para HIV-1 e 2 (Turner e Summers, 1999; Sierra *et al.*, 2005). O CD4 é uma glicoproteína presente na subpopulação de linfócitos T auxiliares, nos precursores de

células T localizados na medula óssea e no timo, nos monócitos, macrófagos, eosinófilos, células dendríticas e microglíocitos do sistema nervoso central (Hoffmann *et al.*, 2007). A interação CD4 - gp120 é necessária, porém não é suficiente para a entrada do vírus na célula. Os lentivírus de primatas necessitam de outras proteínas na superfície celular para que ocorra a fusão das membranas viral e celular (Turner e Summers, 1999). No caso do HIV-1 esta fusão ocorre principalmente através dos co-receptores quimiocínicos CCR5 e CXCR4, porém algumas raras cepas utilizam co-receptores secundários, como CCR2, CCR3, CCR8, CCR9, CX<sub>3</sub>CR1, D6, CXCR7/RDC-1, STRL33, Gpr15, Gpr1, APJ, ChemR23. Ao contrário do CCR5 e do CXCR4, que são expressos nos dois maiores alvos do HIV-1 (linfócitos T CD4<sup>+</sup> e fagócitos mononucleares), esses co-receptores secundários apresentam um padrão de expressão baixo e tecido-específico (Lusso, 2006; Alkhatib; 2009).

O sítio de ligação aos co-receptores da gp120 não está totalmente exposto antes da ligação ao CD4. O HIV desenvolveu uma estratégia única de interação com seus receptores celulares, um mecanismo eficiente de esconder seus epítomos altamente conservados do ataque dos anticorpos do hospedeiro. A ligação ao CD4 induz uma mudança conformacional na gp120, que expõe o sítio de ligação ao co-receptor quimiocínico, promovendo interações gp120/co-receptor. Os principais determinantes da especificidade do HIV pelo co-receptor localizam-se dentro do domínio V3 da gp120, o qual faz contato direto com o co-receptor durante o processo de entrada do vírus (Figura 2). No entanto, a maior superfície de contato do vírus com o co-receptor são os folhetos- $\beta$  e estruturas adjacentes. Sendo altamente conservados, estes filamentos determinam o reconhecimento do co-receptor, enquanto que a estrutura e a carga da alça V3 determinam a especificidade para CCR5 ou CXCR4 (Turner e Summers, 1999; Lusso, 2006, Alkhatib, 2009).

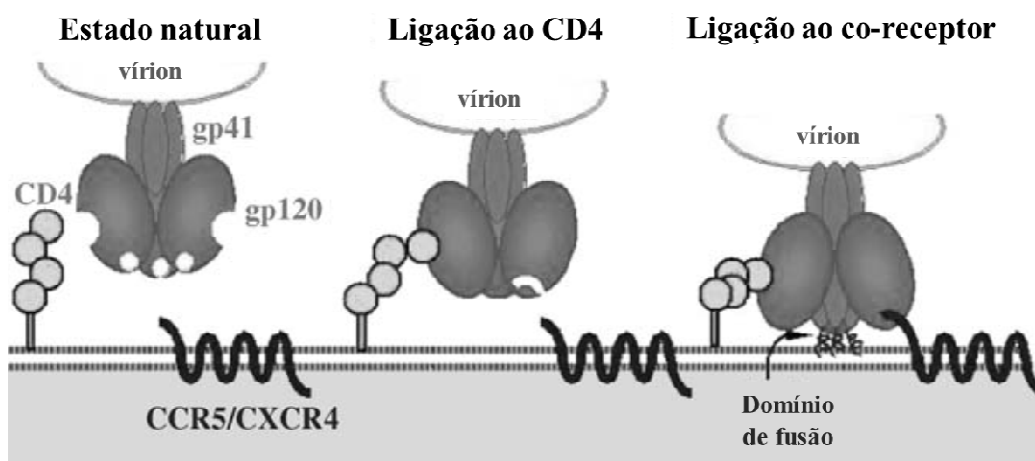


**Figura 2.** Glicoproteína 160 do HIV: a glicoproteína externa gp120 contém os sítios de ligação ao receptor CD4 e aos co-receptores. A alça V3 determina a especificidade pelo co-receptor (Sierra *et al.*, 2005).

A ligação ao co-receptor leva a uma nova mudança conformacional da gp120 que é transmitida para a gp41, resultando na sua mudança conformacional. A gp41 é crucial para a fusão das membranas viral e celular e sua nova conformação promove a inserção de sua terminação amino na membrana da célula alvo. A gp41 forma poros na membrana plasmática e ancora gp120 na célula alvo e no vírion, engatilhando a fusão das membranas (Figura 3) (Turner e Summers, 1999; Sierra *et al.*, 2005; Hoffmann *et al.*, 2007).

A fusão das membranas é seguida da liberação do core viral no citoplasma celular. O genoma é transcrito em uma fita dupla de DNA pela transcriptase reversa. A fidelidade da transcrição reversa é influenciada pela presença da proteína celular APOBEC3G que induz numerosas mutações na fita de cDNA recém-sintetizada, através da catálise da desaminação de citidina para uridina. Durante a replicação, essa mutação resulta na incorporação da adenosina ao invés da guanina na fita complementar (Minton K, 2005). A proteína viral Vif neutraliza a ação da APOBEC3G através da degradação e da redução de sua expressão. A seguir o DNA viral é transportado para a membrana nuclear pela proteína

Vpr e entra no núcleo através do poro nuclear, onde é integrado ao cromossomo do hospedeiro pela integrase. Nos primeiros ciclos de transcrição do DNA proviral pela RNA polimerase II celular são produzidos níveis basais das proteínas Nef, Rev e Tat. Após produção de quantidade suficiente de Tat, essa proteína controla a transcrição dos demais genes do HIV, através da ligação a LTR. A proteína Rev se liga às moléculas de RNA viral e facilita seu transporte através da membrana nuclear. No citoplasma os RNAs mensageiros são traduzidos pelos polissomos celulares. As proteínas virais e o RNA genômico são transportados para a membrana celular para a montagem do vírus. Para que isto ocorra, é essencial que haja um decréscimo do número de moléculas CD4 na membrana plasmática, evitando interações com a gp120 recém-sintetizada. A proteína Nef acelera a endocitose e a subsequente degradação das moléculas CD4, a gp160 prende as moléculas CD4 recém-sintetizadas dentro do retículo endoplasmático e Vpu degrada essas moléculas dentro do retículo. Adicionalmente, a Vpu facilita a liberação do vírion da membrana plasmática através da formação de poros. O vírus imaturo é liberado e seus polipeptídeos são processados pela protease viral, produzindo partículas virais maduras (Sierra *et al.*, 2005).



**Figura 3.** Modelo de ligação e entrada do vírus na célula alvo (Sierra *et al.*, 2005).



## 1.7 Co-receptores CCR5 e CXCR4

Os receptores quimiocínicos, como o CCR5 e o CXCR4, são proteínas de 68-120 aminoácidos e podem ser classificados de acordo com a presença de quatro cisteínas conservadas: as primeiras duas podem ser adjacentes (CC –  $\beta$ -quimiocinas), podem ser separadas por um resíduo de aminoácido (CXC –  $\alpha$ -quimiocinas) ou por três resíduos de aminoácidos (CX<sub>3</sub>C), ou uma das cisteínas pode estar ausente (C-quimiocinas). Eles apresentam alto grau de homologia estrutural entre si e podem compartilhar ligantes. Os receptores quimiocínicos pertencem à família de receptores com sete regiões transmembrana e são intracelularmente ligados a proteínas G (Figura 4) (Rubial-Ares *et al.*, 2004; Hoffmann *et al.*, 2007).

O sítio de ligação de quimiocinas desses receptores é localizado próximo à face extracelular, assim o sinal desencadeado pela quimiocina ligante deve ser transferido através da membrana plasmática para a proteína G. A ligação da quimiocina provoca uma mudança conformacional no receptor que é transmitida para os domínios citoplasmáticos da proteína, habilitando esta a se ligar à proteína G, o que desencadeia uma cascata de reações, produzindo uma resposta celular. Receptores acoplados à proteína G representam aproximadamente 3% dos genes do genoma humano, medeiam diversas funções celulares, como angiogênese e resposta imune, e sua habilidade de conduzir sinais através da membrana plasmática os torna alvos ideais para o desenvolvimento de drogas (Alkhatib, 2009; Lodowski e Palczewski, 2009). Embora receptores quimiocínicos tenham a função fisiológica de mediar sinais pró-inflamatórios, no caso da infecção por HIV eles foram subvertidos em sítios de acoplamento viral durante a entrada na célula (Lodowski e Palczewski, 2009).



O co-receptor CCR5 é expresso em macrófagos, monócitos, células *natural killers*, células T de memória, células dendríticas imaturas, incluindo as células de Langerhans da mucosa genital (Ruibal-Ares *et al.*, 2004; Kawamura *et al.*, 2005). O CXCR4 é expresso em uma gama maior de células, incluindo neutrófilos, monócitos, linfócitos T primários, linfócitos B, macrófagos, células dendríticas, células hematopoiéticas progenitoras, tímócitos, células endoteliais vasculares e neurônios (Ruibal-Ares *et al.*, 2004; van't Wout and Jensen, 2008).

### 1.7.1 CCR5 $\Delta$ 32

O alelo mutante CCR5 $\Delta$ 32 possui uma deleção de 32pb na região correspondente à segunda alça extracelular do co-receptor. Essa mutação provoca mudança na fase de leitura do gene e o surgimento de um códon de parada precoce, resultando na produção de uma proteína truncada que não é transportada para a superfície celular (Liu *et al.*, 1996). O CCR5 $\Delta$ 32 é frequente em populações de origem europeia, raro no Oriente Médio e na Ásia, e praticamente ausente na África. Indivíduos heterozigotos para esse alelo apresentam maior resistência à infecção pelo HIV-1 e progressão mais lenta para AIDS quando comparados com homozigotos para o alelo normal. Indivíduos homozigotos para o alelo mutante apresentam forte resistência à infecção, porém não absoluta (Salem *et al.*, 2009). Infecções por cepas de HIV-1 que utilizam exclusivamente o CXCR4 ou ambos co-receptores (CCR5 *in vitro*) foram relatadas entre homossexuais e pacientes hemofílicos homozigotos para CCR5 $\Delta$ 32. O efeito protetor conferido pela homozigose do mutante contra a infecção pelo HIV-1 pode ser resultado tanto da ausência de expressão do CCR5

na superfície celular quanto da regulação negativa da expressão do CXCR4 pela proteína mutante CCR5 $\Delta$ 32 (Moore *et al.*, 2004; Alkhatib *et al.*, 2009).

Apesar de promover resistência relativa à infecção por HIV, este alelo já foi associado a diferentes doenças e/ou a aumento de morbidade em determinadas condições patológicas. Por exemplo, a alta frequência de CCR5 $\Delta$ 32 foi associada com severidade da doença em artrite idiopática juvenil e com menor sobrevida em pacientes com esclerose múltipla. Em relação às doenças infecciosas, maiores frequências desse alelo foram observadas em pacientes Gregos com brucelose (revisado por Vargas *et al.*, 2009). Adicionalmente, essa variante é um forte fator de risco para desenvolvimento de infecção sintomática pelo Vírus do Oeste do Nilo, estando associado com casos letais da doença (Glass *et al.*, 2006).

Já a perda funcional de CXCR4 em camundongos resulta em fenótipo letal, defeitos de desenvolvimento cerebelar e vascular, hematopoiese e cardiogênese (van't Wout and Jensen, 2008; Alkhatib *et al.*, 2009).

## 1.8 Tropismo do HIV-1 por Co-receptores

A alta taxa de mutação do HIV-1 permite a geração de vírus que escapam do sistema imune do hospedeiro e origina propriedades biológicas variantes, como o tropismo por diferentes co-receptores celulares. Antes da descoberta dos co-receptores virais, os diferentes fenótipos do HIV-1 eram reconhecidos através do tropismo por células T ou por macrófagos e pela indução ou não de formação de sincício, sendo suas prevalências correlacionadas com o estágio clínico (Lusso, 2006). Isolados de HIV-1 eram classificados como monocitotrópicos (M-trópicos) quando o vírus se propagava em culturas de

macrófagos, sendo incapaz de infectar linhagens de células T. Quando o inverso ocorria, os isolados eram classificados como trópicos por células T (T-trópicos), porém em ambos os isolados os vírus eram capazes de se replicar em células T primárias (Hoffmann *et al.*, 2007). Adicionalmente, os vírus eram classificados de acordo com a habilidade de infectar e induzir sincício (células gigantes multinucleadas) em células MT-2, que expressam CXCR4 e não expressam CCR5. Vírus que não infectam essas células são chamados não indutores de sincício (NSI, do inglês *non-syncytium-inducing*), enquanto que vírus que infectam são chamados indutores de sincício (SI, do inglês *syncytium-inducing*) (Moore *et al.*, 2004; van't Wout and Jensen, 2008).

O surgimento da área das quimiocinas revelou a base fisiológica da variabilidade do HIV-1: o uso diferencial dos dois principais co-receptores, CXCR4 e CCR5 (Lusso, 2006). Cepas que utilizam CCR5 são chamadas R5 e correspondem aos vírus M-trópicos e NSI, e cepas que usam CXCR4 são chamadas X4 e correspondem aos vírus T-trópicos e SI. Vírus capazes de utilizar ambos co-receptores são chamados R5X4 (Berger *et al.*, 1998). O uso de ambos co-receptores por um isolado viral se deve à presença de: a) clones com duplo tropismo R5X4 ou b) mistura de clones puros R5 e X4 ou c) mistura desses três tipos virais (van't Wout and Jensen, 2008). A correspondência entre vírus R5 e vírus M-trópicos não é perfeita, pois macrófagos expressam CXCR4 em menor proporção, portanto podem ser infectados por cepas X4 (Ruibal-Ares *et al.*, 2004). Já a capacidade de induzir formação de sincício em células MT-2 indica a presença de vírus que usam CXCR4, mas não exclusivamente, cepas SI podem corresponder a vírus X4 ou R5X4 (Moore *et al.*, 2004).

O início da infecção por HIV-1 é caracterizado pela predominância de cepas virais R5, as quais são responsáveis pela maioria das transmissões e predominam na fase assintomática da doença. As cepas X4 surgem em estágios tardios da infecção em 50-60%

dos pacientes e estão correlacionadas com a progressão da doença, observando-se queda na quantidade de células T CD4<sup>+</sup> e aumento da carga viral (Sierra *et al.*, 2005). Foi demonstrado que as variantes X4 evoluem de ancestrais R5, sugerindo que não ocorre a propagação de antigas cepas X4 transmitidas, que foram sequestradas em reservatórios celulares por longo período (Salemi *et al.*, 2007). A transmissão predominante de cepas R5 pode ser explicada por uma combinação de mecanismos imunológicos celulares e humoral, como ligação e inativação preferencial de vírus X4 por mucinas e proteínas antivirais inatas, preferencial transcitose (transporte de macromoléculas através do interior da célula por meio de vesículas) de vírus R5, ligação preferencial de vírus R5 por macrófagos, células dendríticas e células de Langerhans, neutralização preferencial de vírus X4 (van't Wout and Jensen, 2008) e expressão constitutiva de altos níveis de SDF-1 por células epiteliais da mucosa genital e intestinal (Lusso, 2006). Outra explicação seria a ocorrência de infecção abortiva na transmissão de vírus X4 devido à rápida citólise que este provoca em células T CD4<sup>+</sup> de memória. Essa citólise rápida não é observada em células T CD4<sup>+</sup> ativadas, fenótipo prevalente na infecção tardia, quando os vírus X4 emergem (Mosier, 2009).

Durante a infecção aguda, as cepas R5 têm vantagem seletiva sobre as cepas X4, devido ao seu tropismo preferencial por células dendríticas que transportam patógenos do local da infecção para os linfonodos e os apresentam para células T e B para inicialização da resposta imune adaptativa. Após internalização, o HIV-1 permanece infeccioso dentro da célula dendrítica por mais de cinco dias. Quando a célula dendrítica interage com células T CD4<sup>+</sup> nos linfonodos, o vírus encontra um ambiente favorável para rápida replicação. Assim, a replicação seletiva de cepas R5 dentro do linfonodo ocorre independentemente da rota de transmissão (vaginal, retal, intravenosa ou vertical),

predominando no início da infecção. Essa vantagem replicativa explica o fato de que em casos de transmissão de cepas X4, os vírus R5 se tornam predominantes após um período inicial de infecção, enquanto que X4 praticamente desaparecem. Outra explicação para esses casos seria a evolução das cepas X4 para R5 por pressão do ambiente, favorável ao uso de CCR5 (Moore *et al.*, 2004).

Os primeiros anos de infecção pelo HIV-1, em que cepas R5 são abundantes, podem ser considerados como uma infecção predominantemente do tecido linfóide do trato digestivo. Esse tecido contém mais da metade dos linfócitos T do corpo humano e é o sítio primário de replicação do vírus. A maioria das células T CD4<sup>+</sup> desse tecido expressam CCR5, talvez pelos altos níveis de expressão de quimiocinas ligantes a CCR5 neste local, que direcionam o tráfego de linfócitos T. Adicionalmente, a expressão de CXCR4 no tecido linfóide do trato digestivo é baixa e a produção localizada de SDF-1 impede o uso desse co-receptor pelo HIV-1. Ativação imune provocada por infecções, incluindo o HIV, aumenta a expressão de CCR5 nas células T ativadas, o que facilita a replicação de cepas R5 continuamente, até a exaustão do sistema imune (Moore *et al.*, 2004).

A heterogeneidade da expressão de co-receptores nos diferentes alvos celulares permite a evolução contínua dos vírus, tanto R5 como X4, no indivíduo infectado. Mutações graduais dentro da região da alça V3 produzem o aumento da afinidade entre o vírus e o co-receptor e perda da sensibilidade viral a quimiocinas ligantes aos co-receptores. Essa evolução adaptativa provoca o escape viral do controle das quimiocinas antivirais, bem como, de medicamentos antirretrovirais inibidores da entrada viral na célula-alvo (Scarlati *et al.*, 1997; van't Wout and Jensen, 2008). Durante o curso da infecção, a evolução da região do envelope viral pode aprimorar continuamente o uso de CCR5 ou o diminuir com surgimento do uso de CXCR4 (Mosier, 2009). Após a

emergência de cepas X4, os vírus R5 permanecem presentes. O uso de ambos co-receptores com a mesma eficiência é uma fase transitória breve do curso natural da doença. Populações puras de cepas X4 são raras e frequentemente estão associadas a estágios tardios da doença, onde ocorre transmissão de cepas X4 (van't Wout and Jensen, 2008). Não se conhece os fatores que determinam o surgimento dos vírus X4, sugere-se que seja resultado de alterações nas pressões seletivas de quimiocinas endógenas que se ligam ao CCR5 e ao CXCR4, facilitando a propagação de cepas que usam CXCR4. A presença de infecções oportunistas também pode induzir ou reprimir a produção de quimiocinas, selecionando uma ou outra cepa (Lusso, 2006). A disponibilidade de células-alvo também pode estar relacionada com surgimento de variantes X4: pacientes infectados apresentam ao longo do tempo uma taxa aumentada de renovação de células T primárias, como consequência da progressiva depleção de células T de memória na periferia. Adicionalmente, foi sugerido que o timo tem um importante papel na evolução e disseminação de vírus X4: a análise de cepas X4 de origem recombinante demonstrou que a recombinação ocorreu entre sequências das regiões V1/V2 de vírus R5 periférico e sequências V3 de vírus X4 do timo. Os tímócitos expressam altos níveis dos receptores CD4 e CXCR4, sendo ambiente favorável para propagação de cepas X4, e o tráfego de linfócitos do timo para os tecidos linfóides disseminaria essas cepas sistemicamente. Evidências mostram que, ao contrário dos relatos antigos de que a função tímica declina na adolescência e quase desaparece na fase adulta, o timo do adulto retém parte da função timopoiética e continua a produzir e liberar células T primárias para a periferia (Salemi *et al.*, 2007). Além de infectar e destruir tímócitos imaturos mais eficientemente que vírus R5, as cepas X4 infectam e destroem a células T em desenvolvimento no timo, prejudicando sua função de produção dessas células. Se uma grande quantidade de células



T imaturas é afetada, ocorre um decréscimo na produção das linhagens celulares CD4<sup>+</sup> e CD8<sup>+</sup> primárias e a resposta imune contra o HIV e outras infecções pode diminuir (Moore *et al.*, 2004).

Também não há um consenso sobre o significado biológico da mudança fenotípica, se ela é responsável pela progressão da doença ou se é consequência da progressiva imunodeficiência. A maior patogenicidade observada nas cepas X4, reflexo do amplo espectro de células-alvo, é um indicativo de que o seu surgimento é a causa da aceleração da doença. Porém, o fato de que uma considerável fração de pacientes em fase terminal da AIDS nunca apresentou mudança do tropismo para CXCR4 e que a alteração fenotípica é reversível, podendo ser apenas transitória, prova que o uso desse co-receptor não é essencial para o desenvolvimento da AIDS (Lusso, 2006). Adicionalmente, cepas X4 podem surgir em qualquer estágio da doença, tanto após o diagnóstico de AIDS quanto em pacientes com contagens de CD4 estáveis (van't Wout and Jensen, 2008).

### 1.9 Inibidores da Entrada Celular do HIV-1

A relevância biológica dos co-receptores CCR5 e CXCR4 levou ao desenvolvimento de uma nova classe terapêutica, os inibidores da entrada do HIV na célula. A predominância do uso de CCR5 durante todo o curso da infecção, seu papel essencial na transmissão e a resistência quase total a infecção observada em indivíduos homocigotos para a mutação CCR5 $\Delta$ 32, tornaram o CCR5 o alvo primário dessas novas drogas (Lodowski e Palczewski, 2009; Lusso, 2006). O primeiro composto antagonista de CCR5 aprovado para uso clínico foi o Maraviroc (Selzentry, Pfizer Incorporation), um segundo composto está na etapa de ensaios clínicos de fase três, o Vicriviroc (*Schering-*

*Plough Research Institute*) e diversos outros estão em desenvolvimento pré-clínico ou clínico. O mecanismo de ação dos antagonistas de CCR5 é a inibição da entrada do vírus na célula através da ligação à cavidade localizada entre as hélices, expostas na membrana, do CCR5, estabilizando o co-receptor em uma conformação que não é reconhecida eficientemente pelo HIV-1. Esses medicamentos prendem o HIV fora da célula-alvo, limitando sua habilidade de reabastecer os reservatórios celulares latentes. Inibidores de CCR5 não são eficientes em reduzir a carga viral de pacientes portadores de cepas X4, sendo recomendados somente para tratamento de infecções por cepas R5 puras. Portanto, o tropismo da população viral do paciente deve ser determinado antes da administração desses medicamentos (Moore and Kuritzkes, 2009).

A resistência aos inibidores de CCR5 pode ocorrer de duas formas: emergência de vírus que usam CXCR4, por expansão de cepas minoritárias pré-existentes ao tratamento, ou evolução das cepas R5. A resistência de variantes R5 pode ocorrer através do uso mais eficiente dos co-receptores livres, por exemplo, por aumento da afinidade da região do envelope viral pelo CCR5 ou por aumento da velocidade de fusão das membranas após a ligação com CCR5. Esse mecanismo confere resistência parcial aos inibidores de CCR5. A resistência completa ocorre pelo desenvolvimento da capacidade de utilizar o CCR5 ligado ao inibidor para entrar na célula através de mudanças na sequência da alça V3, além de continuar utilizando os CCR5 livres. Essas variantes são capazes de se ligar a essa nova configuração de CCR5, mantendo a capacidade de se ligar ao CCR5 original (Moore and Kuritzkes, 2009).

Em pacientes cuja presença de cepas X4 foi provocada pela utilização de inibidores de CCR5 observa-se reversão do fenótipo viral para o uso de CCR5 após a interrupção do tratamento, indicando que a pressão seletiva contra cepas X4 foi preservada.

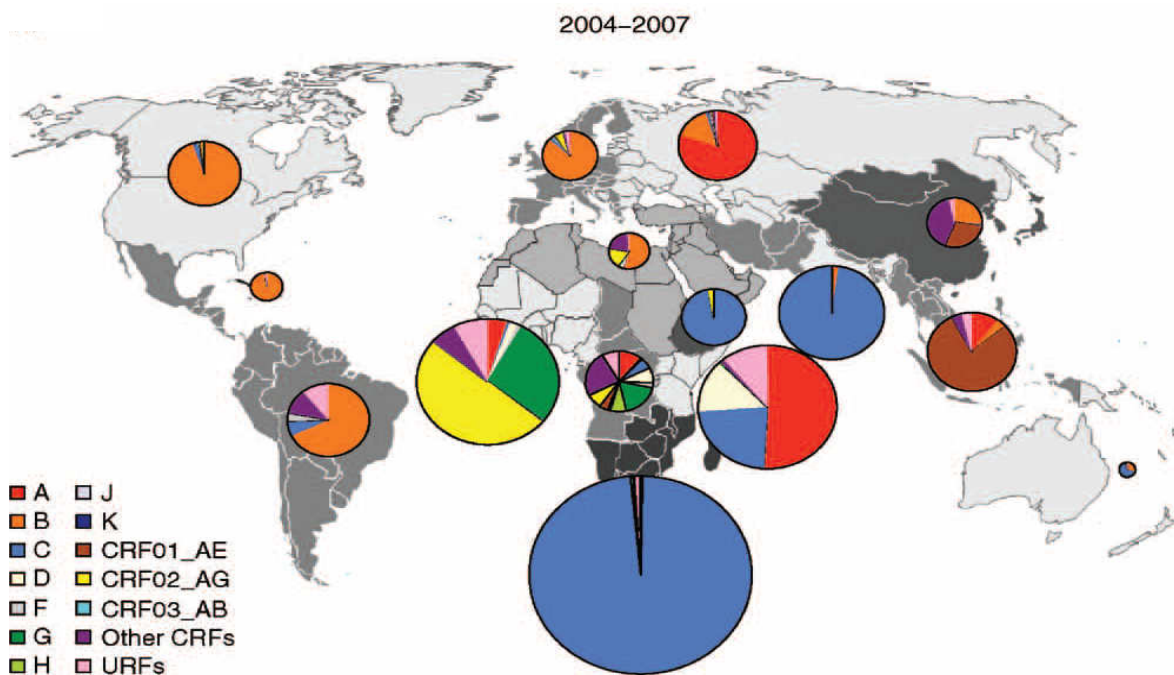
Adicionalmente, estudos *in vitro* mostram que vírus R5 resistentes aos inibidores de CCR5 apresentam resistência cruzada a outros inibidores de CCR5, porém continuam sensíveis às demais classes de medicamentos antirretrovirais (van't Wout and Jensen, 2008; Moore and Kuritzkes, 2009).

#### 1.10 Origem e Distribuição do HIV-1 Subtipos B, C e Recombinantes BC

A distribuição dos subtipos do HIV-1 varia amplamente no mundo, sendo que o subtipo C é o mais prevalente, responsável por quase a metade (48%) das infecções. Os subtipos B e A causam 11 e 12% das infecções, respectivamente, seguidos por CRF02\_AG (8%), CRF01\_AE (5%), subtipo G (5%) and D (2%). Enquanto na Europa e na América do Norte predomina o subtipo B, na África e na América do Sul há uma maior diversidade, predominando respectivamente os subtipos (A, B, C, D e G) e (B, C e F) (Figura 5) (Hemelaar *et al.*, 2011).

O Brasil apresenta uma complexa distribuição de subtipos, devido a sua ampla extensão territorial, sendo que a epidemia é causada principalmente pelo subtipo B, circulando também os subtipos F1, C e D, e as formas recombinantes B/F1 e B/C, variando conforme a região geográfica. As regiões sudeste, nordeste e central apresentam uma distribuição similar, onde o subtipo mais prevalente é o B (70-90%), seguido do F1 e recombinantes B/F1 (5-30%) e uma pequena proporção do subtipo C (<8%). Já na região sul, principalmente no Rio Grande do Sul, observa-se uma alta frequência do subtipo C e recombinantes B/C (20-70%) seguido do subtipo B (25-60%), e uma menor fração do F1 e recombinantes B/F1 (<10%) (Bello *et al.*, 2008; Ferreira *et al.*, 2008; Araujo *et al.*, 2010; Franca *et al.*, 2010; Toledo *et al.*, 2010). Dias *et al.* (2009) analisaram 128 amostras de

pacientes virgens de tratamento da cidade de Porto Alegre, coletadas entre julho de 2004 e dezembro de 2005, e detectaram uma prevalência de 29% do subtipo C, 23,4% da CRF31\_BC, 22,6% do subtipo B, 10% do subtipo F e 15% de formas mosaicas.



**Figura 5.** Distribuição regional dos subtipos e formas recombinantes do HIV-1 em 2004-2007. O mundo foi dividido em 15 grupos de países. Países de um mesmo grupo estão sombreados com a mesma cor. As cores dos gráficos em pizza representam os diferentes subtipos especificados na legenda. As áreas dos gráficos em pizza correspondem ao número relativo de pessoas infectadas com HIV-1 nas regiões (Hemelaar *et al.*, 2011).

A epidemia do subtipo B no Brasil é mais antiga que a do subtipo C. O subtipo B parece ter sido introduzido no país entre metade da década de 1960 e início da década de 1970, provavelmente vindo dos Estados Unidos (Bello *et al.*, 2006; Bello *et al.*, 2007). Já a epidemia do subtipo C parece ter sido resultado da introdução de uma única cepa originada de Burundi, localizado na região central-leste da África. Essa introdução provavelmente ocorreu na região sul do país na década de 80, 10 a 15 anos após a introdução do subtipo B, época em que inúmeros refugiados da guerra civil de Burundi se espalharam por

diversos países (Bello *et al.*, 2008). Análises filogenéticas demonstram que essa introdução foi provavelmente um evento único, pois os vírus do subtipo C brasileiros se agrupam em um clado bem definido (Santos *et al.*, 2006). Adicionalmente, sugere-se que o subtipo C brasileiro se espalhou para Argentina e Uruguai, já que esse subtipo foi detectado nesses países pela primeira vez em 1999 e 2001, respectivamente, e se agrupam dentro do mesmo clado das sequências do Brasil (Monteiro *et al.*, 2007).

Análises filogenéticas do genoma do HIV-1 demonstram que alguns isolados se agrupam em dois ou mais subtipos, dependendo da região genômica analisada. Essas formas mosaicas surgem através de recombinação durante a transcrição reversa e algumas delas podem alcançar níveis epidêmicos de disseminação (Robertson *et al.*, 1995). Vírus recombinantes são designados como CRFs se o genoma viral for totalmente sequenciado e detectado em pelo menos três indivíduos sem relação epidemiológica (Robertson *et al.*, 2000). O surgimento de CRFs é favorecido pela co-circulação de distintos subtipos em uma população. Entre 2004 e 2007, CRFs eram responsáveis por 16,5% das infecções no mundo, embora essa proporção seja maior em algumas regiões, como o sul e o sudeste da Ásia (82,3%) e o oeste da África (56,3%) (Hemelaar *et al.*, 2011). Na América Latina, a CRF12\_BF é a mais prevalente, sendo encontrada na Argentina, Peru, Bolívia, Brasil e Uruguai (Hierholzer J *et al.*, 2002; Brígido *et al.*, 2005; Butler *et al.*, 2007). No Brasil, além dessa CRF, já foram identificadas outras cinco: CRF28\_BF, CRF29\_BF, CRF31\_BC, CRF39\_BF e CRF40\_BF, indicando que os vírus estão evoluindo dinamicamente no país (De Sá Filho *et al.*, 2006; Santos *et al.*, 2006; Guimarães *et al.*, 2008). A co-prevalência dos subtipos B e C no Rio Grande do Sul propiciou a origem da CFR31\_BC, que se caracteriza por possuir um fragmento de aproximadamente 240 nucleotídeos do gene *pol* do subtipo B, compreendendo os códons 138 – 217 do gene da transcriptase reversa,

incorporado em um genoma característico do subtipo C (Santos *et al.*, 2006). A CRF31\_BC circula na região sul do país a mais de 20 anos, sugerindo que o evento de recombinação ocorreu entre 5 a 10 anos após a introdução do subtipo C no Brasil. Provavelmente esta variante se originou na capital do estado ou proximidades, onde foi detectada em indivíduos diagnosticados no início da década de 1990 (Santos *et al.*, 2006; Santos *et al.*, 2007; Passaes *et al.*, 2009).

A prevalência do subtipo C e recombinantes BC está aumentando no sul do país. Santos *et al.* (2007) analisaram amostras das cidades de Porto Alegre e Rio Grande, coletadas entre julho de 2002 e janeiro de 2003, e observaram que 45,4% das infecções eram causadas pelo subtipo B, seguido do subtipo C (34,9%), CRF31\_BC (7,4%) e os subtipos F, D e formas mosaicas em menor frequência. Monteiro *et al.* (2007) observaram 27,3% de cepas do subtipo C, 27,3% de recombinantes BC, 36,4% do subtipo B, 4,5% do subtipo F e 4,5% de recombinante B/F entre 22 amostras coletadas em 2003 de indivíduos de Porto Alegre. Já Brígido *et al.* (2007), analisando amostras provenientes de Porto Alegre, Itajaí (SC) e Balneário Camboriú (SC), coletadas entre julho de 2004 e março de 2006, encontraram uma prevalência de aproximadamente 50% e 20% para o subtipo C e CRF31\_BC, respectivamente, entre amostras provenientes de Porto Alegre e também demonstraram a presença dessa CRF no estado de Santa Catarina, em menor proporção (3,6%). Adicionalmente, Souza *et al.* (2009) reportaram a ocorrência de um paciente com 10 anos de idade portador da CRF31\_BC em São Paulo, tendo adquirido a infecção através de transmissão materna. Esses dados sugerem uma expansão dessa variante na última década, o que enfatiza sua relevância epidemiológica.

### 1.11 Tropismo do HIV-1 Subtipos B, C e Recombinantes BC

O uso de co-receptores e sua mudança fenotípica tem sido intensamente estudados em isolados do subtipo B. Nestes, mais de 50% dos pacientes com AIDS (Koot *et al.*, 1996; Scarlatti *et al.*, 1997) e cerca de 16% dos pacientes virgens de tratamento recém diagnosticados (Frange *et al.*, 2009) apresentam múltiplo uso de co-receptores. Já o tropismo dos demais subtipos virais tem sido menos caracterizado. Alguns estudos demonstram que há diferenças no uso de co-receptores entre o subtipo B e C, devido a diferenças de sequência gênica e conformação da região V3 da gp120. Essa região viral no subtipo B é menos constricta que no subtipo C, sendo capaz de acomodar mutações que conferem tropismo por CXCR4. Contrariamente, a alça V3 do subtipo C possui uma conformação mais restrita, limitando seu potencial de evoluir para X4 (Coetzer *et al.*, 2006; Patel *et al.*, 2008). Os vírus do subtipo C de diferentes regiões do mundo e de diferentes estágios da doença, incluindo o estágio da AIDS, utilizam quase exclusivamente o CCR5, apresentando mínima conversão para uso do CXCR4. Coetzer *et al.* (2006) citam ter observado apenas 10% de cepas com tropismo por CXCR4 em um estudo de 231 isolados de HIV-1 subtipo C, as quais se correlacionaram significativamente com contagens de CD4 inferiores a 200 células/ $\mu$ L. Zhang *et al.* (2009) observaram 100% de cepas R5 entre 28 sequências desse subtipo provenientes da China retiradas de *Los Alamos HIV Sequence Database*. A manutenção do fenótipo R5 durante todo o curso da doença em combinação com carga viral alta observada em estágios avançados pode explicar a rápida disseminação desse subtipo no mundo, já que cepas trópicas por CCR5 são mais eficientemente transmitidas (Pollakis *et al.*, 2001). No entanto, recentes estudos sugerem que o subtipo C do sul da África está evoluindo para o uso do CXCR4. Connell *et al.*

(2008) encontraram 30% de cepas X4 entre isolados de pacientes com AIDS e virgens de tratamento antirretroviral de Johannesburg. Johnston *et al.* (2003) observaram 35% de cepas trópicas por CXCR4 em pacientes em tratamento antirretroviral de Zimbábue.

Embora o polimorfismo genético das cepas de HIV-1 brasileiras seja bem conhecido, o uso de co-receptores até então foi pouco investigado. Ferraro *et al.* (2001) caracterizaram o uso de co-receptores de 18 isolados virais provenientes das cidades de Santos, Rio de Janeiro, Belo Horizonte e Salvador, sendo 17 do subtipo B e 1 do subtipo F, e encontraram uma predominância do uso de CCR5 (11 R5, 3 X4 e 4 R5X4). Araujo *et al.* (2010) analisaram 46 sequências virais dos subtipos B, F1 e recombinantes B/F1 de indivíduos de Salvador, encontrando 24% de cepas trópicas por CXCR4. Franca *et al.* (2011) observaram uma prevalência de 31% de cepas X4 entre 45 sequências dos subtipos B e F1 de Ribeirão Preto. O tropismo do subtipo C no Brasil foi caracterizado somente em dois estudos, cujo número amostral foi pequeno. Monteiro *et al.* (2007) investigaram o tropismo de 7 isolados virais de Porto Alegre e encontraram 4 cepas R5 do subtipo C, 1 cepa X4 do subtipo C e duas cepas R5X4 recombinantes BC. Ferreira *et al.* (2008) analisaram 34 sequências dos subtipos B e C de Curitiba e encontraram uma prevalência de 18% de cepas trópicas por CXCR4, sendo todas do subtipo B. Na China, onde também há uma alta prevalência de recombinantes BC e onde circulam as outras duas CRFs BC descritas (CRF07\_BC e CRF08\_BC), o tropismo viral se caracteriza por predominância do uso de CCR5: 100% das amostras estudadas por Guo *et al.* (2007), Ma *et al.* (2009) e Zhang *et al.* (2009) apresentam esse fenótipo.



## 1.12 Determinação do Tropismo Viral por Co-receptores CCR5 e CXCR4

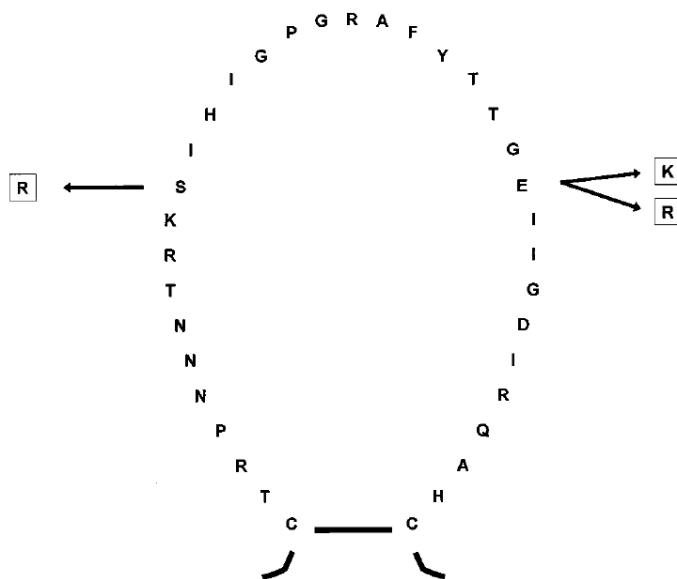
A necessidade do conhecimento do tropismo do HIV-1 para o emprego da terapêutica de antagonistas do CCR5, para o acompanhamento da progressão da doença durante terapia antirretroviral combinada e para auxiliar na compreensão da história natural da epidemia (Daar *et al.*, 2007) tem levado ao desenvolvimento de diversas metodologias para determinação do uso de co-receptores. O método padrão-ouro para caracterização do tropismo do HIV-1 é o ensaio fenotípico, realizado a partir de isolados primários de HIV ou a partir de vírus recombinantes (Sánchez *et al.*, 2010). Atualmente estão disponíveis quatro ensaios fenotípicos recombinantes: Trofile<sup>®</sup> (Monogram Biosciences), Phenoscript<sup>®</sup> (Eurofins-VIRalliance), XtrackC/PhenX-R<sup>®</sup> (inPheno) e VircoType<sup>®</sup> (Virco) (van't Wout and Jensen, 2008). O Trofile é o único ensaio aprovado para uso clínico e é o mais amplamente utilizado mundialmente (Sánchez *et al.*, 2010). As metodologias fenotípicas recombinantes consistem basicamente na geração de vírus recombinantes infecciosos, a partir do plasma do paciente, contendo o gene do envelope, total ou parcial, da população viral do paciente. Subsequentemente, esses vírus são testados em linhagens de células indicadoras que expressam o receptor CD4 e o co-receptor CCR5 ou CXCR4. Portanto, esses ensaios permitem a distinção de populações puras R5 ou X4 e populações com duplo tropismo (R5X4). Os ensaios fenotípicos recombinantes são capazes de detectar variantes minoritárias quando estas representam 5 a 10% da população. Em 2008 foi desenvolvida uma nova versão do Trofile<sup>®</sup> com maior sensibilidade (ES Trofile<sup>®</sup>), capaz de detectar variantes minoritárias em nível de 0,3% da população viral. No entanto, testes fenotípicos geralmente requerem amostras de plasma com carga viral mínima de 1.000 cópias/mL. Assim, pacientes em regimes de tratamento supressivo, que são candidatos para tratamento

com inibidores de CCR5 por razões que não sejam falha terapêutica, não podem ser testados (van't Wout and Jensen, 2008).

Apesar de sua acurácia, a metodologia fenotípica é complexa, onerosa, demorada e sua implantação é limitada por dificuldades técnicas, requerendo conhecimento especializado e estrutura laboratorial sofisticada, tornando-a impraticável na rotina clínica. Adicionalmente, os ensaios recombinantes disponíveis são financeiramente inacessíveis para a maioria das instituições, principalmente para as que necessitam transportar as amostras internacionalmente. Metodologias genotípicas, baseadas na sequência da alça V3 da gp120, tem sido desenvolvidas como alternativa para predição do tropismo viral. Alguns estudos tem demonstrado boa correlação entre métodos genotípicos e fenotípicos (Delobel *et al.*, 2007; Raymond *et al.*, 2008; Raymond *et al.*, 2010; Sánchez *et al.*, 2010; Seclén *et al.*, 2010). No entanto, metodologias genotípicas não são capazes de diferenciar cepas X4 e R5X4, sendo ambas classificadas como vírus X4, e são capazes de detectar variantes minoritárias somente quando estas estão em uma proporção de no mínimo 25% da população viral (van't Wout and Jensen, 2008; Seclén *et al.*, 2010). Contudo, esses métodos possuem maior potencial de aplicação na prática clínica por serem mais simples, rápidos, de fácil execução e com custo menor. Adicionalmente, eles podem ser realizados a partir de amostras de plasma, através da amplificação do RNA viral, ou de células mononucleares do sangue periférico (PBMC, do inglês *peripheral blood mononuclear cell*), através da amplificação do DNA pró-viral. Assim, as metodologias genotípicas são uma opção para determinação do tropismo em pacientes com viremia indetectável (Seclén *et al.*, 2010).

### 1.12.1 Caracterização Genética da Região V3

O tropismo viral por co-receptores é determinado principalmente pela estrutura molecular da alça V3 da gp120 do envelope viral. Essa região geralmente é constituída por 35 aminoácidos, podendo variar de 32 a 37 aminoácidos; possui uma cisteína em cada uma das extremidades conectadas por uma ligação dissulfídica; se caracteriza por ser catiônica, glicosilada e altamente variável, tolerando um número considerável de variações de aminoácidos sem perder suas propriedades funcionais (Figura 6) (De Jong *et al.*, 1992; Hartley *et al.*, 2005). O tropismo por co-receptores é influenciado pela variação de aminoácidos específicos, inserções e deleções, carga total dos aminoácidos, sequência de amino ácidos do tetrâmero e glicosilação dentro da região V3.

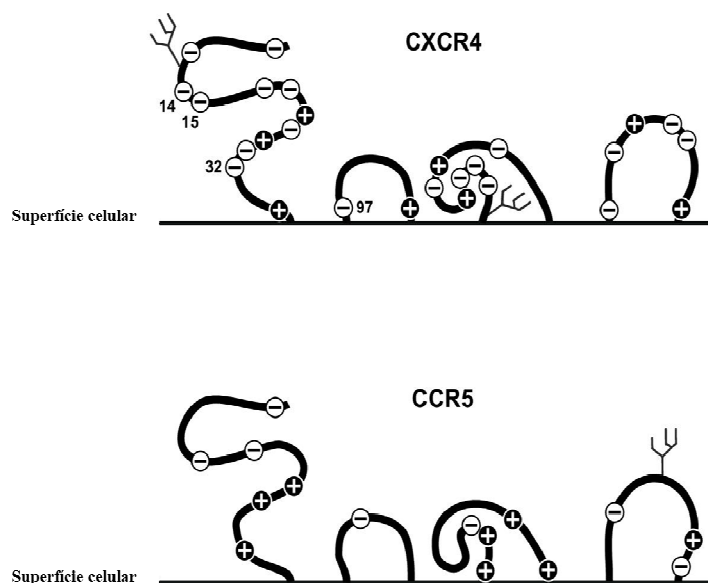


**Figura 6.** Representação esquemática da sequência de aminoácidos da região V3 e a variação de aminoácidos nas posições 11 (direita) e 25 (esquerda). A presença dos aminoácidos ácidos arginina (R) e lisina (K) nessas posições são indicativas de uso de CXCR4 (Fouchier *et al.*, 1995).

Em geral, a presença de aminoácidos básicos nas posições 11 e 25, maior positividade da carga total dos aminoácidos, variabilidade no comprimento da região V3 e

dos resíduos do tetrâmero e ausência do sítio de *N*-glicosilação são características indicativas do uso de CXCR4 (De Jong *et al.*, 1992; Fouchier *et al.*, 1992; Fouchier *et al.*, 1995; Hartley *et al.*, 2005).

A tendência dos vírus com maior número de aminoácidos carregados positivamente se ligarem ao CXCR4 pode ser explicada pela atração eletrostática entre V3 e a superfície extracelular desse co-receptor que possui mais cargas negativas que CCR5 (naturalmente mais neutro) (Figura 7).



**Figura 7.** Representação dos aminoácidos carregados nos domínios extracelulares de CXCR4 e CCR5 (Thorsen *et al.*, 2002).

O padrão de glicosilação do envelope do HIV-1 pode influenciar o uso de co-receptores celulares. A perda do sítio de *N*-glicosilação dentro da região V3, posição 6-8 (sítio NXT/S, em que X representa qualquer aminoácido), em associação com carga positiva total alta, pode provocar a mudança do fenótipo do vírus de R5 para X4 (Pollakis *et al.*, 2001; Nabatov *et al.*, 2004). Essa mudança pode ser provocada pelo desmascaramento de aminoácidos carregados positivamente, aumentando a carga total de

V3, ou pela ação indireta da perda do glicano sobre a estrutura secundária da alça V3 (Hartley *et al.*, 2005). A região V3 de vírus do subtipo C em geral apresenta carga total menor e alta frequência de glicosilação, corroborando com a baixa frequência de cepas X4 entre isolados desse subtipo (Pollakis *et al.*, 2001).

A alça V3 possui uma região conservada que forma um giro na estrutura da alça e alterações de aminoácidos nesse tetrâmero podem alterar o uso de co-receptores (De Jong *et al.*, 1992; Hartley *et al.*, 2005). A epidemia brasileira do HIV-1 subtipo B se caracteriza por apresentar alta prevalência dos tetrâmeros GPGR, frequente na América do Norte e Europa, e GWGR, também chamado B'-GWGR, que é característico das cepas do Brasil. Cada um desses tetrâmeros está presente em aproximadamente 50% dos isolados do subtipo B no Brasil, variando de região para região (Bongertz *et al.*, 2000; de Brito *et al.*, 2006; Diaz *et al.*, 2008). O subtipo C do HIV-1 apresenta o tetrâmero GPGQ que se acredita que tenha um papel limitador do uso de CXCR4 (Coetzer *et al.*, 2006; Raymond *et al.*, 2010).

#### 1.12.2 Predição Genotípica do Tropismo por Co-receptores

A predição genotípica do uso de co-receptores pode ser realizada através das regras 11/25 e valor da carga total dos aminoácidos, pela combinação dessas duas regras e por ferramentas de bioinformática disponíveis gratuitamente na internet.

**Regra 11/25:** a presença de aminoácidos básicos (carga positiva), arginina (R) e lisina (K), na posição 11 e/ou 25 da alça V3 é associada com uso do CXCR4, enquanto que a presença de aminoácidos neutros ou ácidos (carga negativa), ácido aspártico (D) e ácido

glutâmico (E), é associada com uso do CCR5 (De Jong *et al.*, 1992; Fouchier *et al.*, 1992; Fouchier *et al.*, 1995).

**Regra da carga total dos aminoácidos:** a carga total é calculada subtraindo-se o número de aminoácidos com cargas negativas (D e E) do número de aminoácidos com cargas positivas (R e K) da região V3. Carga total maior ou igual a 5 é associada com o uso do CXCR4, enquanto que carga menor que 5 é associada com o uso do CCR5.

A combinação dessas duas regras aumenta a sensibilidade e especificidade para a predição do uso de co-receptores (Delobel *et al.*, 2007; Raymond *et al.*, 2008; Séclen *et al.*, 2010).

Até o momento foram descritas três regras combinadas para predição do tropismo viral:

**Regra de Garrido:** a presença de aminoácidos básicos nas posições 11/25 **ou** carga total  $\geq 5$  é preditiva de cepa X4, caso contrário é R5 (Garrido *et al.*, 2008).

**Regra de Delobel:** a presença de aminoácidos básicos nas posições 11/25 **e** carga total  $\geq 5$  é preditiva de cepa X4, caso contrário é R5 (Delobel *et al.*, 2007).

**Regra de Raymond:** a presença de 11 R/K ou 25K ou ambos; **ou** 25R e carga total  $\geq 5$ ; **ou** carga total  $\geq 6$  é preditiva de cepa X4, caso contrário é R5 (Raymond *et al.*, 2008).

As ferramentas de bioinformática disponíveis online mais utilizadas para predição do uso do co-receptor viral são: WebPSSM, Geno2pheno\_coreceptor, WetCat e R5/X4-Pred e NSI/SI-Pred.

**WebPSSM:** essa ferramenta prediz o uso do co-receptor a partir da sequência de aminoácidos da região V3 no formato fasta. A sequência é alinhada com uma sequência consenso e a predição é realizada pelo método estatístico matriz de escore de posição específica (PSSM, do inglês *position-specific scoring matrices*). Esse método pontua cada resíduo de aminoácido, gerando uma pontuação geral para a sequência, e a classifica como R5 ou X4 de acordo com um ponto de corte estabelecido. Pontuações altas estão

associadas com variantes X4. Essa ferramenta permite selecionar três matrizes diferentes: X4/R5 subtipo B, SI/NSI subtipo B e SI/NSI subtipo C. A matriz X4/R5 utiliza sequências cujo fenótipo para co-receptor é conhecido, enquanto que as matrizes SI/NSI utilizam sequências cujo fenótipo para indução ou não de sincício em células MT-2 é conhecido. A predição pode ser realizada somente para sequências dos subtipos B e C (Jensen *et al.*, 2003). Endereço eletrônico: <http://indra.mullins.microbiol.washington.edu/webpssm/>.

**Geno2pheno\_coreceptor:** essa ferramenta prediz o uso do co-receptor a partir da sequência de nucleotídeos da região V3 no formato fasta ou no formato plano (somente a sequência). A sequência é traduzida, alinhada com a sequência referência HXB2 e a predição é realizada pelo método estatístico denominado “máquina de vetores de suporte” (SVM, do inglês *support vector machine*). Esse método reconhece padrões em grande volume de dados e os classifica. Para determinação do tropismo por essa ferramenta é necessário escolher o nível de significância da predição (taxa de falso positivo, ou seja, probabilidade de classificar um vírus R5 erroneamente como X4). Os níveis de significância disponíveis são 1%, 2,5%, 5%, 10%, 15% e 20%, sendo que quanto maior o nível de significância, maior é a sensibilidade de detecção de variantes X4 e menor é a especificidade. Essa ferramenta permite a adição de parâmetros clínicos para melhorar a predição, como: carga viral, contagem de células CD4<sup>+</sup> e CD8<sup>+</sup>, porcentagem de células CD4<sup>+</sup> e genótipo do CCR5 do paciente. A predição pode ser realizada para sequências do subtipo B e subtipos não-B (Sing *et al.*, 2007). Endereço eletrônico: <http://coreceptor.bioinf.mpi-inf.mpg.de/index.php>.

**Wetcat:** essa ferramenta disponibiliza diversos métodos estatísticos para predição do uso do co-receptor. Três métodos são baseados em árvore de decisão (Wetcat C4.5, C4.5p8p12 e PART), um método é baseado em SVM (Wetcat SVM) e um método é similar a regra

11/25 (Wecat chargerule). A predição é realizada a partir de sequências de aminoácidos da região V3 que devem ser previamente alinhadas com a sequência consenso, descrita da página da ferramenta. O banco de dados Wetcat inclui amostras do subtipo B e de subtipos não-B (Séclen *et al.*, 2010). Endereço eletrônico: <http://genomiac2.ucsd.edu:8080/wecat/v3.html>.

**R5/X4-Pred e NSI/SI-Pred:** essa ferramenta utiliza um método estatístico robusto para predição do tropismo viral, *random forest* (RF). Esse método consiste na construção de árvores de regressão através de seleção de características aleatórias e algoritmo de aprendizagem de máquina. A predição é realizada através da agregação das predições do conjunto de dados. Essa ferramenta inclui na análise, além da carga total e do perfil de aminoácidos, a polaridade da alça V3. A predição é realizada a partir de sequências de aminoácidos da região V3 no formato fasta. Esse método apresenta bom desempenho para predição de sequências dos subtipos B, C e não-B e não-C. Assim como a WebPSSM, os métodos R5/X4 e NSI/SI utilizam como banco de dados sequências com o fenótipo conhecido para uso de co-receptor ou indução de sincício, respectivamente (Xu *et al.*, 2007). Endereço eletrônico: <http://bioinfo.gnway.net/HIV-1/PhenoPred.php>.

### 1.13 Justificativa

O perfil do tropismo por co-receptores e das características genéticas da região V3 do gene *env* da população viral do Rio Grande do Sul necessitam ser investigados, pois auxiliarão na compreensão da patogênese viral e da dinâmica da epidemia nessa região. Além disso, servirão como base para implementação de intervenções nos serviços de saúde, especialmente relacionadas a estratégias terapêuticas. A prevalência do subtipo C



em nossa região e sua origem africana intensificam a necessidade de pesquisas do uso de co-receptores neste subtipo, visto que a emergência do tropismo por CXCR4 em vírus do subtipo C pode ter implicações importantes para a transmissão viral, patogênese, progressão da doença e resposta à nova classe terapêutica de inibidores da entrada viral na célula.

A disponibilização de novas linhas medicamentosas antirretrovirais é indispensável para a continuidade do tratamento de pacientes que apresentam sérios sintomas de toxicidade devido ao uso prolongado das drogas usuais e de pacientes portadores de vírus resistentes a estas drogas. Para estes pacientes, as opções de tratamento são atualmente limitadas, ocasionando ineficiente inibição da replicação viral e contribuindo para o progresso da doença. Além disso, como a incidência de variantes virais resistentes aumentou na população infectada em tratamento, está ocorrendo a transmissão destas variantes para indivíduos virgens de tratamento antirretroviral, provocando maior incidência de falha terapêutica (Little *et al.*, 2002). O conhecimento do perfil do tropismo por co-receptores da população viral mais prevalente no Rio Grande do Sul auxiliará no manejo do regime terapêutico antirretroviral do Brasil.

Com base nos dados expostos, este estudo se propõe a analisar as características da região da alça V3 e o tropismo por co-receptores CCR5 e CXCR4 de HIV-1 subtipos B, C e recombinantes BC circulantes no Rio Grande do Sul, isolados de pacientes virgens de tratamento antirretroviral através de ferramentas genotípicas.

2 MANUSCRITO

MANUSCRITO EM PREPARAÇÃO PARA SUBMISSÃO À REVISTA

*AIDS RESEARCH AND HUMAN RETROVIRUSES*

CO-RECEPTOR USAGE PROFILE AND V3 GENETIC CHARACTERISTICS OF  
HIV-1 SUBTYPES B, C AND BC RECOMBINANTS FROM DRUG NAÏVE PATIENTS  
IN SOUTHERN BRAZIL

Andréa Cristina Vanni<sup>1,2,\*</sup>, Rosa Dea Sperhackle<sup>1</sup>, Sérgio Kakuta Kato<sup>1,4</sup>, Maria Cecilia  
Araripe Sucupira<sup>3</sup>, Ricardo Sobhie Diaz<sup>3</sup>, José Artur Bogo Chies<sup>2</sup>

<sup>1</sup>HIV/AIDS Research Laboratory, University of Caxias do Sul, Rio Grande do Sul, Brazil.

<sup>2</sup>Immunogenetics Laboratory, Federal University of Rio Grande do Sul, Rio Grande do Sul, Brazil.

<sup>3</sup>Retrovirology Laboratory, Federal University of Sao Paulo, Sao Paulo, Brazil.

<sup>4</sup>Federal University of Health Sciences of Porto Alegre, Rio Grande do Sul, Brazil.

\* Corresponding author:

University of Caxias do Sul, Building S, Room 315, HIV/AIDS Research Laboratory. Street address: Francisco Getúlio Vargas, 1130, Caxias do Sul, Rio Grande do Sul, Brazil. ZIP code: 95070-560. Telephone: +55 54 3218-2737. Fax number: +55 54 3218-2056. E-mail: acvanni@gmail.com and acvanni@ucs.br

## Abstract

Changes on genetic sequence and conformation of the V3 loop of HIV-1 gp120 have been reported to cause subtype-specific differences in co-receptor usage. This study analyzed the V3 genetic characteristics and the co-receptor tropism of HIV-1 subtypes B, C and BC recombinants of drug naïve individuals from Rio Grande do Sul through genotypic tools. Stored plasma specimens from 105 patients known to be infected with subtypes B (35), C (35) or BC (35) recombinants based on RT and PR sequences along the *pol* gene were selected. After nested PCR, the V3 region of *env* gene was sequenced and the tropism inferred using the genotypic tools Geno2pheno\_coreceptor<sub>10%</sub>, WebPSSM and R5/X4-Pred. A total of 90 samples were successfully amplified and sequenced. Subtypes found at *env* region were 37 B and 53 C. The higher CXCR4-tropic strains proportions among subtype B was 16.2% (predicted by Geno2pheno<sub>10%</sub>) and among subtype C was 32.1% (predicted by WebPSSM<sub>SI/NSI(C)</sub>). X4 prevalence in recent seroconverters was 18.7%, being exclusively identified in subtype C. There was no association between tropism and viral load. V3 characteristics associated to CXCR4 usage were length variability, mean net charge and basic amino acids at positions 11/25 for subtype B sequences and for subtype C sequences was just length variability. Loss of N-glycosylation was rarely observed and did not correlate to viral tropism. Among subtype B viruses, the tetramers observed were GPGR (43.2%), GWGR (16.2%), GFGR (16.2%) and the uncommon APGR (13.5%). The tetramer GPGQ was present in 84.9% of subtype C sequences and the GPGR motif was associated with CXCR4 usage in subtype C. These finds reinforce previous observations of subtype-specific differences in HIV-1 co-receptor usage and corroborate reports of evolution of subtype C viruses for CXCR4 usage.

## Introduction

The HIV-1 entry into target cells involves the interaction between the viral envelope complex and cellular receptors. It is mediated by a sequential binding of the envelope gp120 with the receptor CD4 and a chemokine co-receptor, usually CCR5 or CXCR4<sup>1,2</sup>. Accordingly to the co-receptor usage, virus strains are classified as R5, X4 or R5X4<sup>3</sup>. Virus populations showing the later phenotype may contain only R5X4 viruses or, most commonly, different mixtures of these three types<sup>4</sup>. This classification has replaced the previous identification method of viral isolates based on their ability to induce syncytia in MT-2 cells, which express only CXCR4. Thus, R5 viruses do not infect MT-2 cells and are called non-syncytium-inducing (NSI), while X4 and R5X4 viruses infect these cells and are called syncytium-inducing (SI)<sup>3,4</sup>. Due to a combination of multiple selective mechanisms, R5 strains are most commonly transmitted and are predominant at early stages of infection, being detected throughout the infection<sup>4</sup>. X4 variants may emerge during the course of the infection, generally in advanced disease stages, through evolution from R5 strains<sup>5</sup>. The emergence of CXCR4-tropic viruses is associated with a faster CD4<sup>+</sup>T cells counts decline, a viral load increase and progression to AIDS<sup>2,6</sup>.

The specificity of HIV-1 co-receptor usage is believed to be dictated mainly by the structure and charge of the third hypervariable loop (V3 loop) of the gp120<sup>7</sup>. Changes on its genetic sequence and conformation have been reported to cause subtype-specific differences in co-receptor tropism<sup>8,9</sup>. Some V3 characteristics such as sequence length variability, increased positively amino acid charge, loss of glycosylation sites, variation of specific amino acids and residues at crown motif have been identified as influencing co-receptor usage in HIV subtypes B and C<sup>8,10-16</sup>. Multiple co-receptor usage is found in about 50% of AIDS patients infected with subtype B viruses<sup>17,18</sup>. HIV-1 subtype C infections, in

which a predominance of CCR5-tropic strains at various stages of diseases is observed<sup>8,19</sup>, are evolving to use CXCR4 for cell entry in some regions of Africa<sup>10,20,21</sup>. Both subtypes B and C, as well as BC recombinants, dominate HIV-1 epidemic in the southernmost state of Brazil, Rio Grande do Sul, accounting for approximately 35-45%, 25-35% and 7-25% of infections in this region, respectively<sup>22,23</sup>. The co-receptor tropism of these main circulating strains remains to be elucidated, especially when considering the African origin of Brazilian subtype C<sup>24,25</sup>.

Phenotypic methodologies are considered the gold standard for HIV-1 tropism assignment. Despite their accuracy, these assays are very complex, time-consuming, expensive, and require special facilities and expertise. As an alternative method for tropism determination, several genotypic approaches, based on analysis of V3 loop sequence, were developed. These genotypic methods are easier, faster, cheaper and have demonstrated good performance when compared with phenotypic tests, representing a suitable alternative methodology for co-receptor prediction in places where phenotypic assays are not available or are financially unfeasible<sup>10,14,26-32</sup>.

Considering the importance of tropism assessment for the therapeutic use of CCR5 antagonists, disease natural history comprehension and prediction of disease progression<sup>6</sup>, this study proposes to characterize the V3 region and analyze the co-receptor usage of HIV-1 subtypes B, C and BC recombinants of antiretroviral-naïve individuals from Rio Grande do Sul by genotypic tools.

## Materials and Methods

### Study population

HIV-1 infected antiretroviral-naïve individuals enrolled on Project AMPLIAR (research consortium developed to study laboratory methods and risk factors for HIV infection in south Brazil) between 2006 and 2008 attended at four sites located in three different cities in Rio Grande do Sul: Municipal Ambulatory of Infectology of Caxias do Sul (Caxias do Sul), Hospital Sanatório Partenon (Porto Alegre), Health Center Vila dos Comerciantes (Porto Alegre) and Specialized Service in DST/HIV/AIDS Herbert de Souza (Viamão) were studied. A total of 105 plasma samples, stored at  $-80^{\circ}\text{C}$ , from patients known to be infected with subtypes B (35), C (35) or BC (35) recombinants based on reverse transcriptase and protease sequences along the *pol* gene were recovered. At the time of recruitment, informed consent was obtained from all the patients and demographic information was recorded.

Viral load was quantified using Versant HIV-1 RNA 3.0 Assay (Siemens Healthcare Diagnostics, USA) which quantifies the viral RNA from 1 mL of plasma through bDNA technique, having a quantification ranging from 50 to 500,000 HIV-1 RNA copies/mL. Recent seroconversion was detected using the Aware<sup>TM</sup> BED<sup>TM</sup> EIA HIV-1 Incidence Test (Calypte Biomedical Corporation, USA), which classifies individuals as recent seroconverters (infection duration  $\leq 155$  days) or with long-term infections (infection duration  $> 155$  days) based on the proportion of HIV-1 specific IgG with respect to total IgG.

### **V3 amplification and sequencing**

V3 amplification and sequencing were performed at the Retrovirology Laboratory of the Federal University of Sao Paulo. Viral RNA was extracted from 200 µL plasma samples using the QIAamp<sup>®</sup> Viral RNA Mini Kit (Qiagen, Germany) and the cDNA was obtained through reverse transcription using SuperScript III Reverse Transcriptase (Invitrogen, USA), according to the manufacturer's instructions. The V3 region was amplified by nested PCR using three different set of primers:

1) Outer primers: V3 outer 5' (5' ATA AGC TTC AAT GTA CAC ATG GAA TT 3') and V3 outer 3' (5' ATG AAT TCA TTA CAG TAG AAA AAT TCC C 3'). Inner primers: V3 inner 5' (5' ATA AGC TTG CAG TCT AGC AGA AGA AGA 3') and V3 inner 3' (5' ATG AAT TCT GGG TCC CTC CTG AGG A 3')<sup>33</sup>.

2) Outer primers: ED5 (5' ATG GGA TCA AAG CCT AAA GCC ATG TG 3') and ED12 (5' AGT GCT TCC TGC TGC TCC CAA GAA CCC AAG 3'). Inner primers: ES7 (5' TGT AAA ACG ACG GCC AGT CTG TTA AAT GGC AGT CTA GC 3') and ES8 (5' CAG GAA ACA GCT ATG ACC CAC TTC TCC AAT TGT CCC TCA 3')<sup>34</sup>.

3) Outer primers: EX1 (5' AAT GTC AGC ACA GTA CAA TGT ACA 3') and EX10 (5' CGC CCA TAG TGC TTC CTG CTG CTC 3'). Inner primers: IN1 (5' GTT AAA TGG CAG TCT AGC AGA AGA 3') and IN10 (5' TAT AAT TCA CTT CTC CAA TTG TCC3')<sup>35</sup>.

The set of primers 1, 2 and 3 produce a final fragment of 335, 655 and 671pb, respectively. The amplification conditions for both first and second rounds for set 1 were: 94°C for 5 min, 35 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 1 min and 72°C for 7 min. For set 2, the conditions were: 94°C for 5 min, 35 cycles at 94°C for 30 s, 55°C for

45 s, 72°C for 1 min and 72°C for 5 min. Conditions for set 3 were: 94°C for 10min, 35 cycles at 94°C for 45 s, 55°C for 30 s, 72°C for 45 s and 72°C for 5min.

PCR amplicons were purified using Montage PCR Centrifugal Filter Devices (Millipore Corporation, USA) and directly sequenced in both directions in the ABI 3130 Xl Genetic Analyzer using Big Dye Terminator V.3.1 Cycle Sequencing kit (Applied Biosystems, USA). The pair of primers used for V3 sequencing was V3 inner 5' and 3'. Nucleotide sequences were analyzed with Sequencher program (Gene Codes Corporation Ann Arbor, MI). The sequences quality was analyzed using Quality Control - HIV-1 Sequence Quality Analysis program from Los Alamos National Laboratory (LANL) (<http://www.hiv.lanl.gov/content/sequence/QC/index.html>) for identification of stop codons, frameshifts and hypermutation. Nucleotide mixtures were considered when the second highest peak in the electropherogram was > 25% of the main peak. Codons containing mixtures were translated into every possible amino acid and all different sequences generated through each amino acid permutation were analyzed, except the sequences containing stop codons. Nucleotide sequences showing an unsuitable electropherogram were excluded, as well as, sequences with eight or more nucleotide mixtures within the V3 region due to the elevated number of amino acid sequences generated. Nucleotide sequences were translated using Codon Alignment program from LANL (<http://www.hiv.lanl.gov/content/sequence/CodonAlign/codonalign.html>). Alignments were performed and manually edited using BioEdit Sequence Alignment Editor Software. Phylogenetic analysis of the nucleotide sequences excluded any possibility of sample contamination (data not shown).



## **HIV-1 subtype determination**

Subtype analysis of the envelope region was performed using the following online subtyping methods: Recombinant Identification Program (RIP) from LANL (confidence threshold = 95%, window size = 200) available at <http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html> and REGA HIV-1 & 2 Automated Subtyping Tool Version 2.0 available at <http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html>.

## **Genotypic prediction of HIV-1 co-receptor usage**

The prediction of co-receptor usage was performed using three freely available online genotypic tools: Geno2pheno\_coreceptor, WebPSSM and R5/X4-Pred. Genotypic predictions do not allow R5X4 and X4 strains differentiation, being both classified as X4 viruses.

The Geno2pheno predictor is based on the support vector machine method and its dataset includes B and non-B subtype sequences. A rate of 10% of false positive was used in this study<sup>10,27</sup>. This tool is available at <http://coreceptor.bioinf.mpi-inf.mpg.de/>.

The WebPSSM tool uses the statistical method of position-specific scoring matrices for predicting the co-receptor usage and allows the selection among different matrices. The matrix X4/R5 (B) was used for subtype B and the matrix SI/NSI (C) for subtype C analysis<sup>31,36</sup>. This predictor is available at <http://indra.mullins.microbiol.washington.edu/webpssm/>.

The R5/X4-Pred is based on the random forest method and the prediction includes the V3 loop polarity, in addition to net charge and amino acid profiles. It has a good performance for tropism prediction of subtypes B, C, non-B and non-C sequences<sup>32</sup>. The

website (<http://bioinfo.gnway.net/HIV-1/PhenoPred.php>) is no longer available and the analyses were performed using the software graciously provided by Dr. Chiyu Zhang.

### **Statistical analyses**

Demographic and baseline characteristics of the study population were recorded as absolute numbers, percentage and mean  $\pm$  standard deviation. Genotypic results were expressed as absolute and relative frequencies. Associations between categorical parameters and genotypic results were assessed using Fisher's exact test. Quantitative variables were compared using Student's *t*-test. Differences were considered as significant if *p* values were below 0.05. All reported *p* values were two-sided. All statistical analyses were performed using the SPSS v.18.0 software.

### **Results**

The V3 region was successfully amplified from 101 out of 105 plasma samples of the HIV-1 infected patients previously selected. The first set of primers amplified the V3 fragment of 34 samples. Of the 71 remaining samples submitted to amplification with the second set of primers, 61 were successfully amplified. Finally, it was possible to amplify 6 of the 10 remaining samples using the third set of primers. Among the 101 V3 sequences analyzed, 11 showed an inadequate electropherogram or showed more than eight nucleotide mixtures within the V3 region and were excluded from this study. Large genetic heterogeneity in the viral envelope gene most likely explains it. A total of 90 patients were evaluated and the translation of all V3 sequences generated by amino acid permutations resulted in 1237 different sequences. The baseline characteristics of the study population are summarized in Table 1. Overall, 58.9% of the individuals were male, 17.8% were

recent seroconverters and 96.7% reported sexual exposure. The mean age was  $35.28 \pm 10.45$  years. The subtyping analysis of *env* region identified 37 (41.1%) patients with subtype B and 53 (58.9%) with subtype C infection. The mean viral load was  $4.34 \pm 0.86$  log HIV-1 RNA copies/mL and there was no significant difference between subtype B and C infections. The subtype classification considering both regions (*pol/env*) were: 29 B/B (32.2%), 27 C/C (30.0%), 24 BC/C (26.7%), 7 BC/B (7.8%), 2 B/C (2.2%) and 1 C/B (1.1%). The prevalence of X4 viruses predicted by each genotypic tool among the different subtypes is shown in Table 2. As the *env* region determines the viral tropism, the genotypic analyses were carried out considering just the *env* subtype.

### **V3 Sequence Characterization and Co-receptor Usage Prediction of HIV-1 Subtype B**

Among the 37 subtype B infections, CXCR4-tropic strains were detected in 6 individuals using Geno2pheno<sub>10%</sub>, 4 individuals using WebPSSM<sub>X4/R5(B)</sub> and 4 individuals using R5/X4-Pred (Table 2). X4 viruses were detected by all genotypic tools in 3 patients and in 5 patients they were detected just by one predictor (Figure 1). The mean viral load was not significantly different according to CCR5- or CXCR4- using variants classified by each genotypic tool (Table 3).

The V3 sequences from R5 and X4 strains were compared for length, net charge, *N*-glycosylation site, amino acid at positions 11/25 and crown motif (Figure 1). The V3 region typically consists of 35 amino acids and length variability has been associated with CXCR4 co-receptor usage in subtype B and C infections<sup>8,10,13</sup>. The length characteristic showed to be conserved; only one sequence had 34 amino acids and displayed the X4 phenotype predicted by all the genotypic tools.

CXCR4 tropism has been associated with an increased number of positively charged amino acids at V3 loop<sup>8,10,11</sup>. Among the 29 viruses predicted as R5 by the three genotypic tools, 25 (86,2%) had a net charge of  $< +5$ , with the majority of samples having +3 and +4; 2 had a net charge of +5; and only 2 had a net charge of  $> +5$ . The 3 X4 strains classified as such by all predictors had a net charge of +4 or +5. The X4 strain identified exclusively by WebPSSM<sub>X4/R5(B)</sub> showed a net charge of +6; the X4 strain predicted exclusively by R5/X4-Pred showed +4; and the 3 X4 strains predicted by Geno2pheno<sub>10%</sub> showed  $< +5$  (Figure 1). The mean net charge was increased in X4 strains when compared to R5 strains predicted by each genotypic tool, being statistically significant only in the WebPSSM<sub>X4/R5(B)</sub> prediction (Table 3).

The presence of the basic amino acids arginine (R) or lysine (K) at the position 11 and/or 25 and the loss of a potential *N*-glycosylation site, at positions 6-8 (motif NXT/S) of the V3 region, have also been associated with CXCR4 co-receptor usage<sup>8,10,12,13,15,16</sup>. Among the 37 sequences, two had an R at position 11 of V3 and both were predicted as X4 viruses by the three tools; three had a K at position 25, of which 1 was classified as CXCR4-tropic strain by all the predictors, one was classified as CXCR4-tropic strain just by WebPSSM<sub>X4/R5(B)</sub> and 1 was predicted as an R5 virus. These results were statistically significant (Table 3). All the subtype B viruses, regardless of the co-receptor usage, had the potential *N*-glycosylation site conserved (Figure 1).

In Brazilian HIV-1 subtype B epidemic, two distinct strains cocirculates: the United States/Europe-like strain which contains the GPGR crown motif at the tip of V3 loop and the Brazilian variant named B' - GWGR, each of them accounting for approximately half of all subtype B infections in Brazil<sup>37,38,39</sup>. The GPGR motif was detected in 16 (43.2%) individuals, of which six had amino acid ambiguities at different positions. The GWGR

variants were presented in six (16.2%) individuals, but one had an ambiguity W/G and another had an ambiguity G/R at the second glycine position. One individual showed an AWGR motif. Additionally, six individuals had a GFGR motif (one of them containing R/G ambiguity) and five individuals had an APGR motif (one of them containing A/G ambiguity) (Table 3).

### **V3 Sequence Characterization and Co-receptor Usage Prediction of HIV-1 Subtype C**

The prevalence of X4 viruses among the subtype C infections obtained by the different predictors was: 13.2% using Geno2pheno<sub>10%</sub>, 7.5% using R5/X4-Pred and 32.1% using WebPSSM<sub>SI/NSI(C)</sub> (Table 2). Despite WebPSSM<sub>SI/NSI(C)</sub> has been developed for subtype C analysis, it has relatively low specificity (81.8 - 86%) for X4 detection<sup>10,20</sup>. In this study, the CXCR4 usage was predicted in 10 individuals solely by this tool. None of the isolates were exclusively classified as X4 strains by Geno2pheno<sub>10%</sub> or R5/X4-Pred tools (Figure 2). As in subtype B infections, there was no association between mean viral load and tropism determined by the different genotypic tools (Table 3).

V3 amino acid sequence alignments of the subtype C strains are shown in Figure 2, together with genotypic results. The V3 length was also preserved in this subtype; one sequence had 34 amino acids and was predicted as X4 strain by all the genotypic tools. The net charge among the 36 CCR5-tropic isolates, predicted as such by all genotypic tools, was predominantly < +5 (91.7%), with the majority of samples having +3; two strains had a net charge of +5; and just one strain had +6. Most of X4 viruses, independently of the predictor, had a net charge of +3 or +4 (Figure 2). The mean net charge was increased in X4 strains as compared to R5 strains, but it was not statistically significant using any of the predictors (Table 3).

The evaluation of positions 11 and 25 and *N*-glycosylation site of the V3 loop revealed the absence of basic amino acids at position 11, the presence of the amino acid arginine at position 25 in just one sequence which was predicted as X4 viruses by the three genotypic tools, and the loss of the *N*-glycosylation site in two sequences, one of them predicted as X4 and another predicted as R5 by all genotypic tools (Figure 1).

The GPGQ crown motif at the V3 tip is conserved in subtype C viruses<sup>8,10,40</sup>. The GPGQ motif was present in 84.9% of the subtype C isolates, of which seven had amino acid ambiguities at different positions. This motif was mostly preserved in CCR5-tropic strains (presenting in 97.2% of them) when compared to CXCR4-tropic strains. The substitution of glutamine residue for arginine residue at GPGQ crown was responsible for most of X4 motif variation. The association between GPGR and CXCR4 usage was statistically significant using WebPSSM<sub>SI/NSI(C)</sub> ( $p=0.001$ ) and Geno2pheno<sub>10%</sub> ( $p=0.006$ ). The B'-GWGR signature was not detected in subtype C (Figure 2).

### **V3 Genotypic Characterization of HIV-1 Recent Seroconverters**

The CXCR4-tropic strains proportion among the patients with recent infection was 18.7%. All recent seroconverters with subtype B infection harbored only R5 viruses, but one isolate showed features indicative of CXCR4 usage: net charge of +6 and a lysine at position 25. Two out of 6 sequences presented GPGR replacements (Figure 1). Three of the 10 recent seroconverters (30%) with subtype C infection harbored X4 viruses. One of them was classified as X4 by all the predictors and lost the *N*-glycosylation site. The other two isolates were predicted as CXCR4-using variants by WebPSSM<sub>SI/NSI(C)</sub> (Figure 2).

## Discussion

The genetic inference of co-receptor usage profile and the V3 genotypic characteristics of HIV-1 viral population from Rio Grande do Sul, southernmost state of Brazil, are firstly reported here. Three different genotypic tools widely used were applied for the characterization of co-receptor usage from subtypes B, C and BC recombinants strains which are mainly responsible for HIV-1 epidemic in Rio Grande do Sul<sup>22,23</sup>. The higher prevalence of CXCR4-tropic strains found in infections caused by subtypes B and C (thus classified according to *env* sequences) were, respectively, 16.2% predicted by Geno2pheno<sub>10%</sub> and 32.1% predicted by WebPSSM<sub>SI/NSI(C)</sub>. The X4 proportion among subtype B viruses is in agreement with other studies conducted on antiretroviral-naïve individuals (15.9% - 16.9%)<sup>29,41</sup>. The high X4 prevalence found among subtype C infections is similar to recent reports from subtype C epidemic in eastern and southern Africa: antiretroviral-naïve AIDS patients from Johannesburg showed 30% of CXCR4-using viruses<sup>20</sup> and treatment-experienced individuals from Malawi<sup>10</sup> and Zimbabwe<sup>21</sup> exhibited 29% and 35% of X4 viruses, respectively. These studies suggest a subtype C epidemic evolution for CXCR4 usage in that area, while in other world's regions this subtype viruses remain predominantly CCR5-tropic, as observed in China<sup>19</sup>, France<sup>10</sup> and India<sup>40</sup>. The central-east African origin of the south Brazilian subtype C epidemic, which was introduced in Rio Grande do Sul around the early 1980s<sup>24</sup> led us to speculate that the features of the strains triggering the viral evolution for CXCR4 usage in African infections could be present in isolates from Rio Grande do Sul and be responsible for the high prevalence of X4 viruses among subtype C in our country.

Several recent studies have reported a good performance of the current genotypic tools to predict the HIV-1 co-receptor usage compared to phenotypic assays, showing

concordances of 84.5 – 91.4% for WebPSSM<sub>X4/R5(B)</sub>, 69 – 88.6% for Geno2pheno<sub>10%</sub> and 84.3% for WebPSSM<sub>SI/NSI(C)</sub><sup>10,26,28,30</sup>. R5/X4-Pred demonstrated to have an accuracy of 94% in subtype B tropism predictions and accuracy of 96.7% in subtype C tropism predictions<sup>32</sup>. The WebPSSM<sub>SI/NSI(C)</sub> tool was developed for subtype C analysis<sup>36</sup>, but two studies assessing genotypic performance for viral tropism determination compared to phenotypic assays reported a high sensibility (93.3% and 100%) yet a relatively low specificity (81.8% and 86%) of this tool for detecting CXCR4-using viruses<sup>10,20</sup>. Therefore, as the X4 proportion among subtype C viruses was determined exclusively by the WebPSSM<sub>SI/NSI(C)</sub> tool, the actual prevalence of CXCR4-tropic strains among HIV-1 subtype C in south Brazil could be lower than 32.1%. In this study, no significantly higher plasma HIV-RNA levels were observed among individuals infected with X4 viruses as compared to individuals infected with R5 viruses, regardless of viral subtype. This same result was reported in studies involving seroconverters, drug-naïve and antiretroviral-experienced individuals<sup>35,42,43</sup>.

Interestingly, sequence analysis of the V3 region pointed to both similarities and differences among the features of subtype C and B strains. The V3 length was heavily conserved in both subtype B and C viruses from southern Brazil when compared to V3 sequences from different locations where a length range of 32-37 amino acids has been described<sup>8,10</sup>. A small variability observed among Brazilian V3 sequences was a result of deletion of the G amino acid at position 24 in two sequences and this was related to CXCR4 usage, in agreement with previous V3 characterization studies developed in both subtype B and C populations<sup>8,10</sup>. In this study, the potential *N*-glycosylation site was not associated with co-receptor usage in both subtypes, since the large majority of strains showed the integrity of this motif.



The V3 loops of CXCR4-tropic strains have already been associated with an increased net charge in both B and C subtype<sup>8,10,11,12</sup>. The mean net charge was significantly higher among subtype B CXCR4-tropic strains where almost all X4 viruses had a V3 charge of  $\geq 4$ . However, among subtype C viruses the mean net charge showed to be higher, although not significantly, in CXCR4-tropic strains. Nearly 95% of X4 viruses had a V3 charge of  $\leq 4$ , suggesting that an increased number of basic amino acids at V3 loop was not the driving force for the transition from R5 to X4 viruses in this subtype C population.

The presence of R and K residues at positions 11 and 25 are indicative of SI viruses in subtype B. Substitutions with basic residues at these positions induce changes in secondary structure at the tip of the loop and were proposed to promote the transition from a NSI to a SI phenotype<sup>12,13</sup>. In accordance with these observations, subtype B showed an association between basic amino acids at these positions and CXCR4 usage. This association has also been described for subtype C viruses<sup>10</sup> however, in a study of V3 genetic characterization of subtype C isolates these positions were not necessarily associated with positively charged residues in CXCR4-using strains even though there was higher variation at these positions compared to R5 viruses<sup>8</sup>. In accordance to these findings, in this study, none subtype C sequence displayed a basic amino acid at the position 11, but X4 variants presented increased substitutions at this position when compared to R5 variants. The position 25 did not show differences according to co-receptor usage.

In the Brazilian subtype B epidemic prevails variants harboring GPGR or GWGR crown motifs. Four main motif variants were found in subtype B isolates: the GPGR tetramer, as expected, was the most prevalent (43.2%), being followed by the tetramers

GWGR (16.2%), GFGR (16.2%), which usually evolves from the GWGR sequences<sup>39</sup>, and the uncommon APGR (13.5%). The GWGR motif prevalence was intermediate as compared to the frequencies observed in central-east region (42-50%)<sup>37,38,39</sup> and central region (6.7%)<sup>44</sup>, and it is similar to northeast region (18.2%) from Brazil<sup>45</sup>, revealing the origin diversity of the circulating strains of the different regions. Recently it was suggested that the GWGR isolates might only have R5 variants due to constraints on the viral envelope to use the CXCR4 co-receptor<sup>46</sup>. Similarly, in this study all GWGR isolates were predicted as R5 viruses, except the isolate containing an amino acid ambiguity at tryptophan position (W/G). In this case, the X4 phenotype could be related to the variant containing the glycine residue at position 16, which was previously associated with the presence of basic amino acids at sites 11 and 25 of the V3 loop (feature of CXCR4-using viruses)<sup>46</sup>.

Point mutations in the highly conserved tip of the V3 loop can alter the virus phenotype<sup>11</sup>. Interestingly, the GPGR tetramer which is uncommon in non-subtypes B was related to CXCR4 use in subtype C. It is conceivable that replacement of GPGQ by other motifs represent the evolution acquired by HIV strains over time, which also relates to time for tropism change from CCR5 to CXCR4 using viruses. The association between GPGR and CXCR4 usage in subtype C was observed in other studies and it was speculated that the arginine at position 18 has a functional role for CXCR4 usage in subtype C viruses<sup>8,10,40</sup>. In fact, it has been reported that the arginine at position 18 may induce CXCR4 usage by increasing the charge or altering the conformation of the co-receptor-ligand site of the V3 loop, while the GPGQ conserved motif restricts the transition from R5 to X4 viruses in subtype C<sup>8,10,47</sup>.

A subtype C-specific change on V3 sequence which corresponds to the insertion of a glycine and an isoleucine (GI) right before the GPGQ motif position was previously described in about 30-40% of the X4 subtype C viruses and was associated with CXCR4 usage<sup>10,20</sup>. Notwithstanding, this feature was not observed among south Brazilian strains.

Although X4 genotype is normally associated with advanced disease<sup>2,6</sup>, it has been detected in recent seroconverters<sup>42,43,48</sup>. We found an X4 prevalence of 18.7% among individuals with recent infection as predicted by the genotyping analysis, this feature being exclusively identified in subtype C isolates in the present study. A similar proportion of CXCR4-using variants (17.2%) was described in Spain, but the study population was mainly composed by subtype B viruses (92%)<sup>42</sup>. A study conducted in 24 male seroconverters infected with subtype C viruses from Zimbabwe found a prevalence of 12.5%<sup>48</sup>. To our knowledge, this is the first report of a high X4 prevalence (30%) among recent seroconverters carrying subtype C viruses. The presence of X4 variants in recent infection may have implications for antiretroviral therapy and clinical prognosis, being associated with a greater decrease in CD4<sup>+</sup> T cell count over time and an increased risk of progression to AIDS<sup>31</sup>. It is important to consider that this result could be biased due to relatively small sample size, lack of confirmation using phenotypic assays which are considered gold standard for HIV-1 tropism assignment, and possible overestimation of recent infections by BED assay<sup>49</sup>.

South Brazilian HIV-1 strains from antiretroviral-naïve individuals exhibited a significant proportion of CXCR4 usage. The X4 viruses were detected in all subtypes analyzed (B, C and BC recombinants) and were most prevalent among subtype C infected patients, including recent seroconverters. Our findings are consistent with previous observations of subtype-specific differences in HIV-1 co-receptor usage. The subtypes B

and C of HIV-1 Brazilian epidemic displayed a different combination of features along the V3 sequence involved on viral evolution for CXCR4 usage. These observations suggest that the relation between X4 viruses and both disease progression and therapeutic answer for CCR5 inhibitors may be different among subtypes B and C from south Brazil.

## References

1. Dragic T, Litwin V, Allaway GP, Martin SR, *et al.*: HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 1996; 381(6584): 667-73.
2. Sierra S, Kupfer B and Kaiser R: Basics of the virology of HIV-1 and its replication. *J Clin Virol* 2005; 34: 233-44.
3. Berger EA, Doms RW, Fenyö EM, Korber BTM, *et al.*: A new classification for HIV-1. *Nature* 1998; 391: 240.
4. van't Wout AB and Jensen MA: HIV-1 tropism and its detection – impact on infection, transmission and treatment. *European Infectious Diseases* 2008; 2 (1): 60-64.
5. Salemi M, Burkhardt BR, Gray RR, Ghaffari G, Sleasman JW and Goodenow MM: Phylodynamics of HIV-1 in lymphoid and non-lymphoid tissues reveals a central role for the thymus in emergence of CXCR4-using quasispecies. *PLoS ONE* 2007; 2(9): e950.
6. Daar ES, Kesler KL, Petropoulos CJ, Huang W, *et al.*; Hemophilia Growth and Development Study: Baseline HIV type 1 coreceptor tropism predicts disease progression. *Clin Infect Dis* 2007; 45(5):643-9.
7. Lusso P: HIV and the chemokine system: 10 years later. *EMBO J* 2006; 25 (3): 447-56.
8. Coetzer M, Cilliers T, Ping LH, Swanstrom R and Morris L: Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology* 2006; 356: 95-105.
9. Patel MB, Hoffman NG and Swanstrom R: Subtype-specific conformational differences within the V3 region of subtype B and subtype C human immunodeficiency virus type 1 Env proteins. *J Virol* 2008; 82(2): 903-16.
10. Raymond S, Delobel P, Mavigner M, Ferradini L, *et al.*: Prediction of HIV type 1 subtype C tropism by genotypic algorithms built from subtype B viruses. *J Acquir Immune Defic Syndr* 2010; 53(2):167-75.

11. De Jong JJ, Goudsmit J, Keulen W, Klaver B, Krone W, Tersmette M and De Ronde A: Human Immunodeficiency Virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *J Virol* 1992; 66(2): 757-65.
12. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, *et al.*: Phenotype-associated sequence variation in the third variable domain of the Human Immunodeficiency Virus type 1 gp120 molecule. *J Virol* 1992; 66(5):3183-87.
13. Fouchier RA, Brouwer M, Broersen SM and Schuitemaker H: Simple determination of Human Immunodeficiency Virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol* 1995; 33(4):906-11.
14. Delobel P, Nugeyre MT, Cazabat M, Pasquier C, *et al.*: Population-based sequencing of the V3 region *env* for predicting the coreceptor usage of human immunodeficiency virus type 1 quasispecies. *J Clin Microbiol* 2007; 45(5): 1572-80.
15. Nabatov AA, Pollakis G, Linnemann T, Kliphuis A, Chalaby MIM and Paxton WA: Inpatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. *J Virol* 2004; 78(1): 524-30.
16. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J and Paxton WA: N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *J Biol Chem* 2001; 276(16): 13433-41.
17. Scarlatti G, Tresoldi E, Björndal A, Fredriksson R, *et al.*: In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nature Medicine* 1997; 3(11):1259-65.
18. Koot M, van 't Wout AB, Kootstra NA, de Goede RE, Tersmette M and Schuitemaker H: Relation between changes in cellular load, evolution of viral phenotype, and the clonal composition of virus populations in the course of human immunodeficiency virus type 1 infection. *J Infect Dis* 1996; 173(2): 349-54.
19. Zhang C, Xu S, Wei J and Guo H: Predicted co-receptor tropism and sequence characteristics of China HIV-1 V3 loops: implications for the future usage of CCR5 antagonists and AIDS vaccine development. *Int J Infect Dis* 2009; 13: e212-e216.
20. Connell BJ, Michler K, Capovilla A, Venter WDF, Stevens WS and Papathanasopoulos MA: Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. *AIDS* 2008; 22(7): 896-99.
21. Johnston ER, Zijenah LS, Mutetwa S, Kantor R, Kittinunvorakoon C, Katzenstein DA: High frequency of syncytium-inducing and CXCR4-tropic viruses among human immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral treatment. *J Virol* 2003; 77(13):7682-8.

22. Santos AF, Schrago CG, Martinez AMB, Mendonza-Sassi R, *et al.*: Epidemiologic and evolutionary trends of HIV-1 CRF31\_BC-related strains in southern Brazil. *J Acquir Immune Defic Syndr* 2007; 45(3): 328-33.
23. Monteiro JP, Ferraro GA, Oliveira T, Goldani LZ, *et al.*: Genetic and biologic characterization of HIV type 1 subtype C isolates from south Brazil. *AIDS Res Hum Retroviruses* 2007; 23(1):135-43.
24. Bello G, Passaes CPB, Guimarães ML, Lorete RS, *et al.*: Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS* 2008; 22: 1993-2000.
25. Sanabani S, Neto WK, de Sa Filho DJ, Diaz RS, *et al.*: Full-length genome analysis of human immunodeficiency virus type 1 subtype C in Brazil. *AIDS Res Hum Retroviruses* 2006; 22(2):171-6.
26. Poveda E, Briz V, Roulet V, González MM, Faudon JL, Skrabal K and Soriano V: Correlation between a phenotypic assay and three bioinformatic tools for determining HIV co-receptor use. *AIDS* 2007; 21(11): 1487-90.
27. Raymond S, Delobel P, Mavigner M, Cazabat M, *et al.*: Correlation between genotypic predictions based on V3 sequences and phenotypic determination of HIV-1 tropism. *AIDS* 2008; 22: F11-F16.
28. Seclén E, Garrido C, González Mdel M, González-Lahoz J, *et al.*: High sensitivity of specific genotypic tools for detection of X4 variants in antiretroviral-experienced patients suitable to be treated with CCR5 antagonists. *J Antimicrob Chemother* 2010; 65(7): 1486-92.
29. Sánchez V, Masiá M, Robledano C, Padilla S, Ramos JM and Gutiérrez F: Performance of genotypic algorithms for predicting HIV-1 tropism measured against the enhanced-sensitivity Trofile coreceptor tropism assay. *J Clin Microbiol* 2010; 48(11):4135-9.
30. Garrido C, Roulet V, Chueca N, Poveda E, Aguilera A, *et al.*: Evaluation of eight different bioinformatics tools to predict viral tropism in different Human Immunodeficiency Virus type 1 subtypes. *J Clin Microbiol* 2008; 46(3): 887-91.
31. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, *et al.*: Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of Human Immunodeficiency Virus type 1 env V3 loop sequences. *J Virol* 2003; 77(24):13376-88.
32. Xu S, Huang X, Xu H and Zhang C: Improved prediction of coreceptor usage and phenotype of HIV-1 based on combined features of V3 loop sequence using random forest. *J Microbiol* 2007; 45(5):441-46.

33. Wolfs TF, Zwart G, Bakker M, Valk M, Kuiken CL and Goudsmit J: Naturally occurring mutations within HIV-1 V3 genomic RNA lead to antigenic variation dependent on a single amino acid substitution. *Virology* 1991; 185(1):195-205.
34. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, *et al.*: Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* 1993; 262(5137):1257-61.
35. Saracino A, Monno L, Cibelli DC, Punzi G, *et al.*: Co-receptor switch during HAART is independent of virological success. *J Med Virol* 2009; 81(12):2036-44.
36. Jensen MA, Coetzer M, van 't Wout AB, Morris L, Mullins JI : A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences. *J Virol* 2006; 80(10):4698-704.
37. de Brito A, Komninakis SC, Novoa P, de Oliveira RM, *et al.*: Women infected with HIV type 1 Brazilian variant, subtype B (B'-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). *Clin Infect Dis* 2006; 43:1476–81.
38. Bongertz V, Bou-Habib DC, Brígido LF, Caseiro M, *et al.*; Brazilian Network for HIV Isolation and Characterization: HIV-1 diversity in Brazil: genetic, biologic, and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. *J Acquir Immune Defic Syndr* 2000; 23(2):184-93.
39. Diaz RS, Leal E, Sanabani S, Sucupira MC, Tanuri A, Sabino EC, Janini LM: Selective regimes and evolutionary rates of HIV-1 subtype B V3 variants in the Brazilian epidemic. *Virology* 2008; 381(2):184-93.
40. Neogi U, Prarthana SB, D'Souza G, Decosta A, Kuttiatt VS, Ranga U and Shet A: Co-receptor tropism prediction among 1045 Indian HIV-1 subtype C sequences: Therapeutic implications for India. *AIDS Res Ther* 2010; 7:24.
41. Frange P, Galimand J, Goujard C, Deveau C, *et al.*: High frequency of X4/DM-tropic viruses in PBMC samples from patients with primary HIV-1 subtype-B infection in 1996-2007: the French ANRS CO06 PRIMO Cohort Study. *J Antimicrob Chemother* 2009; 64(1):135-41.
42. de Mendoza C, Rodriguez C, García F, Eiros JM, Ruíz L, *et al.*; Spanish HIV Seroconverter Study Group: Prevalence of X4 tropic viruses in patients recently infected with HIV-1 and lack of association with transmission of drug resistance. *J Antimicrob Chemother* 2007; 59(4):698-704.
43. Poveda E, Briz V, de Mendoza C, Benito JM, *et al.*: Prevalence of X4 tropic HIV-1 variants in patients with differences in disease stage and exposure to antiretroviral therapy. *J Med Virol* 2007; 79(8):1040-6.

44. Véras NM, Véras VS, Ramalho ED, Kyaw C, *et al.*: HIV type 1 genetic variability in central Brazil. *AIDS Res Hum Retroviruses* 2007; 23(12):1481-90.
45. Araujo AF, Brites C, Monteiro-Cunha J, Santos LA, *et al.*: Lower prevalence of human immunodeficiency virus type 1 Brazilian subtype B found in northeastern Brazil with slower progression to AIDS. *AIDS Res Hum Retroviruses* 2010; 26(11):1249-54.
46. Leal E, Silva WP, Sucupira MC, Janini LM, Diaz RS: Molecular and structural characterization of HIV-1 subtype B Brazilian isolates with GWGR tetramer at the tip of the V3-loop. *Virology* 2008; 381(2):222-9.
47. Hartley O, Klasse PJ, Sattentau QJ and Moore JP (2005) V3: HIV's switch-hitter. *AIDS Res Hum Retroviruses* 21(2): 171-89.
48. Batra M, Tien PC, Shafer RW, Contag CH and Katzenstein DA: HIV type 1 envelope subtype C sequences from recent seroconverters in Zimbabwe. *AIDS Res Hum Retroviruses* 2000; 16(10):973-9.
49. Hargrove J, Humphrey J; ZVITAMBO Study Group: Short communication: Simplified estimation of the long-term specificity of the BED assay to improve estimates of HIV incidence. *AIDS Res Hum Retroviruses* 2010; 26(9):977-9.



## Figures and Tables

| Cons. B | V3 amino acid sequences alignment |       |   |             |       |   |   |     |     |     |     |     |     |     |   | V3 Length | Net charge | G2p 10% | WebPSSM X4R5(B) | R5/X4 Pred |     |   |       |         |   |         |     |     |       |     |     |     |    |     |     |     |     |     |     |    |    |    |
|---------|-----------------------------------|-------|---|-------------|-------|---|---|-----|-----|-----|-----|-----|-----|-----|---|-----------|------------|---------|-----------------|------------|-----|---|-------|---------|---|---------|-----|-----|-------|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|----|----|----|
|         | C                                 | T     | R | P           | N     | N | N | T   | R   | K   | S   | I   | H   | I   | G |           |            |         |                 |            | P   | G | R     | A       | F | Y       | A   | T   | G     | D   | I   | I   | G  | D   | I   | R   | Q   | A   | H   | C  |    |    |
| 100168  | .                                 | T/L/P | . | .           | N/S   | . | . | .   | R   | G   | .   | H/R | .   | .   | . | .         | .          | .       | F/V             | F          | .   | . | D/E   | K       | . | T       | .   | .   | .     | .   | .   | .   | .  | .   | .   | .   | .   | 35  | 4/5 | X4 | X4 | X4 |
| 100443  | .                                 | .     | . | .           | .     | . | . | .   | S/R | .   | .   | .   | .   | W/G | . | .         | .          | .       | L               | .          | T   | . | --    | .       | . | .       | D/N | .   | .     | .   | .   | Y   | .  | .   | .   | 34  | 3-5 | X4  | X4  | X4 |    |    |
| 120483  | .                                 | E     | . | .           | .     | . | . | .   | R   | .   | G   | .   | .   | .   | . | .         | .          | .       | S               | .          | .   | . | A     | E       | Q | .       | V   | .   | .     | .   | .   | .   | K  | .   | F   | .   | 35  | 4   | X4  | X4 | X4 |    |
| 100646  | .                                 | A     | . | .           | .     | . | . | .   | R   | D   | V   | .   | L   | .   | F | .         | R/G        | .       | T               | I          | Y/F | . | .     | E       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | .   | .   | 35  | 1/2 | X4  | R5 | R5 |    |
| 100802  | .                                 | .     | . | .           | G     | . | . | .   | .   | .   | .   | M   | .   | F   | . | .         | .          | .       | .               | .          | .   | . | K     | E       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | .   | 35  | 4   | X4  | R5  | R5 |    |    |
| 130095  | .                                 | I     | . | .           | S     | . | . | .   | .   | .   | V   | S   | M   | .   | F | .         | .          | .       | .               | I/T        | Y/F | . | .     | .       | . | .       | .   | .   | .     | .   | .   | K   | .  | .   | .   | 35  | 4   | X4  | R5  | R5 |    |    |
| 130129  | .                                 | .     | . | P/T/H/L/N/I | S/T/A | . | . | .   | .   | .   | .   | P/R | .   | .   | . | R/S/N/K   | .          | .       | .               | .          | Y/F | . | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | .   | 35  | 2-4 | R5  | R5  | X4 |    |    |
| 120200  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | .   | L   | .   | P/Q | . | .         | .          | .       | .               | F/L        | .   | . | .     | D/E/N/K | . | .       | .   | D/N | .     | .   | .   | .   | .  | .   | 35  | 3-6 | R5  | X4  | R5  |    |    |    |
| 100164  | .                                 | .     | . | P/L         | S     | . | . | .   | .   | .   | S   | .   | .   | .   | . | .         | .          | .       | V               | V          | .   | . | .     | .       | . | .       | .   | N   | .     | .   | .   | .   | .  | .   | 35  | 4   | R5  | R5  | R5  |    |    |    |
| 100197  | .                                 | .     | . | .           | G     | . | . | .   | .   | .   | .   | P   | I/L | .   | . | .         | K          | .       | F/I             | .          | .   | . | .     | E       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 3   | R5  | R5  | R5  |    |    |    |
| 100426  | .                                 | .     | . | .           | N/S   | . | . | .   | .   | .   | .   | .   | .   | A   | . | .         | .          | .       | .               | .          | .   | . | T     | .       | . | .       | .   | .   | .     | .   | .   | H/Y | .  | .   | 35  | 3   | R5  | R5  | R5  |    |    |    |
| 100620  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | .   | .   | .   | L   | . | .         | .          | .       | .               | .          | .   | . | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 3   | R5  | R5  | R5  |    |    |    |
| 100718  | .                                 | .     | . | .           | S     | . | . | .   | .   | .   | .   | .   | .   | .   | . | .         | .          | .       | .               | .          | .   | . | .     | E/K/Q   | . | .       | .   | D/N | .     | .   | .   | .   | .  | .   | 35  | 3-6 | R5  | R5  | R5  |    |    |    |
| 100827  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | H/P | I/M | .   | .   | . | R/G       | .          | .       | .               | .          | .   | . | .     | D/A/E   | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 2-4 | R5  | R5  | R5  |    |    |    |
| 110102  | .                                 | .     | . | .           | S     | . | . | .   | .   | .   | .   | T   | A   | W   | . | .         | .          | .       | L               | .          | .   | . | E     | .       | . | .       | .   | .   | .     | .   | K   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 110115  | .                                 | .     | . | .           | .     | . | . | S/G | V   | H/N | .   | G/A | .   | .   | . | .         | .          | .       | A/V             | F/I        | .   | . | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 3   | R5  | R5  | R5  |    |    |    |
| 110188  | .                                 | .     | . | .           | .     | . | . | G   | .   | .   | M   | .   | .   | .   | . | .         | .          | .       | T               | .          | .   | . | .     | .       | . | .       | N   | .   | .     | .   | .   | .   | .  | .   | 35  | 4   | R5  | R5  | R5  |    |    |    |
| 110225  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | .   | .   | .   | .   | . | .         | Q          | .       | .               | .          | .   | . | M     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 2   | R5  | R5  | R5  |    |    |    |
| 110263  | .                                 | I     | . | .           | N/H   | . | . | .   | .   | .   | .   | .   | .   | .   | . | .         | .          | .       | A/T             | .          | .   | . | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 3   | R5  | R5  | R5  |    |    |    |
| 110266  | .                                 | .     | . | .           | G     | . | . | .   | .   | .   | S   | F   | A   | .   | . | .         | .          | .       | A/T             | I          | .   | . | .     | Q       | . | I/T/V/A | .   | N   | .     | .   | I/L | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 110312  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | H/P | .   | .   | .   | . | R/S       | .          | .       | .               | .          | .   | . | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | .   | 35  | 2/3 | R5  | R5  | R5  |    |    |    |
| 120051  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | Q   | M   | .   | W   | . | .         | .          | .       | L               | .          | .   | . | E     | .       | V | .       | .   | .   | .     | .   | .   | .   | .  | 35  | 3   | R5  | R5  | R5  |     |    |    |    |
| 120084  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | .   | M   | .   | F   | . | .         | .          | .       | L               | .          | .   | V | .     | D/E     | . | .       | D/N | .   | .     | Q/K | .   | .   | .  | 35  | 3-5 | R5  | R5  | R5  |     |    |    |    |
| 120088  | .                                 | .     | . | .           | .     | . | . | .   | .   | H/Y | M   | .   | W   | .   | . | .         | .          | .       | F/L             | .          | .   | . | .     | D/N     | . | .       | .   | .   | .     | .   | .   | .   | .  | 35  | 3/4 | R5  | R5  | R5  |     |    |    |    |
| 120159  | .                                 | .     | . | .           | S     | . | . | .   | .   | .   | N/T | .   | .   | .   | . | .         | .          | .       | A/T             | F          | .   | . | .     | .       | . | .       | .   | .   | .     | .   | K   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 120202  | .                                 | .     | . | .           | .     | . | . | K/R | .   | .   | L   | .   | F   | .   | . | .         | .          | .       | T               | L          | .   | . | E     | A       | . | .       | .   | N   | .     | .   | .   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 120220  | .                                 | .     | . | .           | .     | . | . | R   | .   | .   | .   | A   | .   | .   | . | .         | .          | .       | .               | .          | .   | . | .     | A/E     | . | .       | .   | .   | .     | .   | .   | .   | .  | 35  | 3/4 | R5  | R5  | R5  |     |    |    |    |
| 120254  | .                                 | .     | . | .           | .     | . | . | R   | G/D | .   | .   | .   | .   | .   | . | .         | .          | .       | .               | Y/H        | A/T | . | .     | E       | . | .       | .   | .   | .     | .   | .   | .   | .  | 35  | 2/3 | R5  | R5  | R5  |     |    |    |    |
| 120280  | .                                 | .     | . | .           | .     | . | . | R   | .   | .   | .   | .   | .   | .   | . | .         | .          | .       | .               | F          | E   | . | .     | .       | V | .       | .   | .   | .     | K   | .   | .   | .  | 35  | 3   | R5  | R5  | R5  |     |    |    |    |
| 120288  | C/R                               | T/S   | . | .           | G     | . | . | .   | S/G | .   | M   | .   | W   | G/R | . | .         | .          | .       | T               | .          | .   | . | A     | .       | . | .       | .   | .   | Q/K/E | .   | .   | .   | .  | 35  | 3-7 | R5  | R5  | R5  |     |    |    |    |
| 120300  | .                                 | I     | . | .           | G     | . | . | .   | .   | .   | .   | A   | .   | .   | . | .         | .          | .       | .               | .          | .   | . | E     | .       | . | .       | .   | .   | .     | .   | K   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 120324  | .                                 | V     | . | .           | G     | . | . | R   | .   | .   | .   | .   | W   | .   | . | .         | .          | .       | S               | L          | .   | . | G/E/A | .       | . | .       | .   | .   | .     | .   | .   | .   | 35 | 2/3 | R5  | R5  | R5  |     |     |    |    |    |
| 120451  | .                                 | .     | . | .           | .     | . | . | .   | .   | S   | L   | .   | F   | .   | . | .         | .          | .       | L               | .          | .   | . | .     | .       | . | .       | .   | .   | .     | K   | Y   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 120579  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | L   | .   | .   | G/S | . | .         | .          | .       | T               | .          | .   | . | E     | I/M     | . | .       | .   | .   | .     | R/K | F   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 130006  | .                                 | .     | . | .           | .     | . | . | .   | .   | .   | .   | .   | .   | .   | . | .         | K          | .       | T               | .          | .   | H | .     | .       | . | .       | .   | .   | .     | .   | .   | .   | .  | 35  | 3   | R5  | R5  | R5  |     |    |    |    |
| 130035  | .                                 | I     | . | .           | .     | . | . | .   | G   | .   | .   | .   | .   | .   | . | .         | .          | .       | .               | .          | .   | . | .     | E       | . | .       | .   | .   | .     | .   | K   | .   | .  | 35  | 4   | R5  | R5  | R5  |     |    |    |    |
| 130159  | .                                 | .     | . | .           | S     | . | . | .   | .   | .   | M   | .   | W   | .   | . | .         | .          | .       | T               | .          | .   | . | .     | D/A     | . | .       | .   | D/N | .     | .   | .   | .   | .  | 35  | 3-5 | R5  | R5  | R5  |     |    |    |    |

Figure 1. V3 amino acid sequences alignment and genotypic results of subtype B isolates. V3 amino acid sequences from 37 patients infected with HIV-1 subtype B (at *env* region) were aligned to the consensus sequence where dot identifies consensus amino acid, dash indicates a deletion and slash separates the possible amino acids at each position. The highlighted samples correspond to recent seroconverters. N-glycosylation site is boxed, positions 11 and 25 are shaded in light gray and the crown motif is shaded in dark gray. V3 length indicates the number of amino acids compounding this region. V3 net charge and genotype predicted by Geno2pheno<sub>10%</sub> (G2p 10%), WebPSSM<sub>X4/R5(B)</sub> and R5/X4-Pred are also showed. X4 predictions and characteristics indicative of X4 viruses are highlighted in red font.

| Cons. C | C   | T       | R   | P       | N   | N  | N   | T   | R   | K       | S   | I     | R   | I   | G   | P   | G     | Q   | T   | F   | Y | A  | T   | G   | D     | I  | I   | G   | D   | I   | R   | Q   | A     | H   | C  | V3 Length | Net charge | G2p 10% | WebPSSM <sub>SI/NSI(C)</sub> | R5/X4 Pred |    |
|---------|-----|---------|-----|---------|-----|----|-----|-----|-----|---------|-----|-------|-----|-----|-----|-----|-------|-----|-----|-----|---|----|-----|-----|-------|----|-----|-----|-----|-----|-----|-----|-------|-----|----|-----------|------------|---------|------------------------------|------------|----|
| 110285  | .   | .       | .   | .       | .   | K  | .   | K   | .   | T       | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | E     | .   | Q  | .         | 35         | 3       | X4                           | X4         | X4 |
| 120231  | .   | .       | .   | P/R     | S   | .  | N/Y | .   | .   | .       | .   | I/M/L | G   | M/V | .   | .   | .     | R   | A   | .   | . | .  | .   | .   | E     | V  | .   | .   | .   | .   | .   | Q/K | .     | .   | .  | 35        | 3-5        | X4      | X4                           | X4         |    |
| 120413  | .   | T/A     | .   | .       | G   | .  | .   | .   | R/S | .       | S/G | .     | A/G | .   | .   | .   | R     | .   | V   | .   | . | .  | .   | D/E | R     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3/4        | X4      | X4                           | X4         |    |
| 120571  | .   | .       | .   | .       | G   | .  | .   | .   | .   | T       | .   | V     | .   | .   | .   | .   | .     | .   | I   | .   | H | .  | .   | .   | .     | T  | .   | .   | .   | .   | .   | K   | .     | .   | .  | 34        | 3          | X4      | X4                           | X4         |    |
| 110164  | .   | .       | .   | .       | .   | .  | .   | .   | .   | T       | .   | M     | G   | .   | .   | .   | R     | .   | V   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 2          | X4      | X4                           | R5         |    |
| 120510  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | G   | .     | G   | .   | .   | .   | .     | .   | V   | .   | F | .  | .   | .   | A     | .  | .   | .   | .   | .   | .   | K   | .     | .   | .  | 35        | 4          | X4      | X4                           | R5         |    |
| 120523  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | G     | .   | .   | .   | .   | R     | A   | .   | .   | . | .  | .   | .   | E     | .  | .   | .   | .   | .   | .   | .   | .     | Y   | .  | .         | 35         | 3       | X4                           | X4         | R5 |
| 120190  | .   | .       | R/T | .       | S   | .  | .   | .   | .   | .       | .   | .     | .   | .   | G/R | .   | .     | .   | T/A | .   | . | .  | .   | .   | .     | .  | .   | .   | IT  | R/G | .   | .   | HL    | .   | .  | 35        | 1-4        | R5      | X4                           | R5         |    |
| 120228  | .   | .       | .   | .       | N/H | .  | .   | .   | .   | E       | .   | .     | .   | .   | G/R | G/R | Q/HL  | .   | .   | .   | . | .  | .   | .   | .     | T  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1-3        | R5      | X4                           | R5         |    |
| 100352  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | R/K | .   | .   | .   | .     | .   | A   | .   | . | .  | R   | .   | D/E   | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 4          | R5      | X4                           | R5         |    |
| 110113  | .   | T/I     | .   | .       | .   | .  | .   | .   | .   | .       | .   | G/S   | .   | .   | .   | R   | A     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | N   | .   | .   | .   | .   | .     | .   | .  | 35        | 4          | R5      | X4                           | R5         |    |
| 120258  | .   | .       | .   | .       | S/G | .  | .   | .   | .   | K/R     | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | AT | .   | .   | .     | IT | .   | .   | .   | .   | .   | .   | L/F   | C/R | .  | 35        | 3/4        | R5      | X4                           | R5         |    |
| 120453  | C/W | .       | .   | .       | .   | .  | T/S | .   | .   | K/E     | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | TP  | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1/3        | R5      | X4                           | R5         |    |
| 120524  | .   | .       | .   | .       | .   | .  | .   | .   | .   | E       | .   | .     | R/K | .   | .   | G/R | .     | .   | .   | .   | . | .  | .   | G/R | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1-3        | R5      | X4                           | R5         |    |
| 120532  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | Q   | .     | T   | IM  | .   | .   | R/K   | A   | .   | F/I | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 2          | R5      | X4                           | R5         |    |
| 130162  | .   | .       | .   | .       | N/H | .  | .   | .   | .   | .       | G   | .     | G   | .   | .   | .   | .     | .   | .   | Y/F | . | .  | D/N | A   | .     | .  | .   | .   | .   | .   | .   | .   | H/Y   | .   | .  | 35        | 2/3        | R5      | X4                           | R5         |    |
| 100339  | .   | .       | .   | .       | .   | .  | .   | .   | .   | K/E     | S/G | .     | G   | .   | .   | .   | Q/H/R | T/A | .   | .   | . | .  | .   | .   | .     | .  | D/N | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 0-4        | R5      | X4                           | R5         |    |
| 100184  | .   | .       | .   | .       | PH  | .  | .   | .   | .   | K/R/G/E | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | H/Q/Y | .   | .  | 35        | 1-3        | R5      | R5                           | R5         |    |
| 100224  | .   | I       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | D/G/E | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3/4        | R5      | R5                           | R5         |    |
| 100285  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 100590  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | A/E   | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3/4        | R5      | R5                           | R5         |    |
| 100696  | .   | M       | .   | .       | .   | .  | .   | .   | .   | D       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | .   | .   | K   | .   | Y   | .     | .   | 35 | 2         | R5         | R5      | R5                           |            |    |
| 100742  | .   | .       | .   | .       | N/S | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | .     | LV | .   | D/N | R/K | .   | H/Y | .   | .     | .   | 35 | 3/4       | R5         | R5      | R5                           |            |    |
| 100743  | .   | I       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | V     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 3         | R5         | R5      | R5                           |            |    |
| 100751  | .   | .       | R/K | .       | .   | .  | .   | .   | .   | .       | .   | .     | R/K | .   | .   | .   | .     | .   | .   | F   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 100810  | .   | .       | .   | .       | N/S | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | D/G   | IT | D/N | R/I | Q/K | .   | H/Y | .   | .     | .   | 35 | 3-5       | R5         | R5      | R5                           |            |    |
| 100822  | .   | .       | .   | .       | .   | .  | .   | .   | K/E | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | D/A   | .  | D/N | R/I | Q/K | .   | H/Y | .   | .     | .   | 35 | 0-6       | R5         | R5      | R5                           |            |    |
| 100826  | .   | .       | .   | .       | .   | .  | .   | .   | E   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | K   | .   | .   | .     | .   | .  | 35        | 2          | R5      | R5                           | R5         |    |
| 110224  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | E   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 110253  | .   | I       | .   | .       | .   | .  | .   | .   | E   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1          | R5      | R5                           | R5         |    |
| 120067  | .   | .       | .   | .       | .   | .  | .   | .   | E   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1          | R5      | R5                           | R5         |    |
| 120234  | .   | .       | .   | .       | .   | NT | .   | .   | E   | .       | .   | .     | .   | .   | .   | .   | .     | QL  | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | A   | .   | Y     | .   | .  | 35        | 1          | R5      | R5                           | R5         |    |
| 120257  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 120263  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | D/A/G | .  | D/N | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3-5        | R5      | R5                           | R5         |    |
| 120293  | .   | .       | .   | .       | .   | .  | .   | .   | R   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 120299  | .   | V/M     | .   | .       | .   | .  | .   | .   | K/E | .       | .   | .     | .   | .   | .   | .   | Q/H   | .   | .   | .   | . | .  | .   | .   | .     | .  | D/N | .   | .   | H/Y | .   | .   | .     | .   | 35 | 1-4       | R5         | R5      | R5                           |            |    |
| 120366  | .   | .       | .   | N/D/G/S | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | P/A | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | .   | .   | .   | H/Y | .   | .     | .   | 35 | 2/3       | R5         | R5      | R5                           |            |    |
| 120367  | .   | .       | .   | .       | .   | .  | .   | T   | .   | V       | .   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 2         | R5         | R5      | R5                           |            |    |
| 120372  | .   | .       | R/S | .       | .   | .  | .   | .   | .   | .       | V   | .     | .   | .   | .   | .   | .     | .   | A   | .   | . | .  | .   | .   | .     | .  | .   | N   | .   | .   | Y   | .   | .     | .   | 35 | 3/4       | R5         | R5      | R5                           |            |    |
| 120373  | .   | .       | .   | .       | .   | .  | .   | R   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 3         | R5         | R5      | R5                           |            |    |
| 120385  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | LV  | R/K   | .   | .   | G/R | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 3/4       | R5         | R5      | R5                           |            |    |
| 120418  | .   | T/IM    | .   | .       | .   | .  | .   | E   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | F/I | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 1         | R5         | R5      | R5                           |            |    |
| 120537  | .   | .       | .   | .       | .   | .  | .   | Q   | .   | T       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | K   | .     | Y   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 120573  | .   | .       | .   | .       | .   | .  | .   | .   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 3          | R5      | R5                           | R5         |    |
| 120631  | .   | T/K/E/A | .   | .       | .   | .  | .   | E   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | H/Y | .   | .     | .   | .  | 35        | 0-3        | R5      | R5                           | R5         |    |
| 120644  | .   | .       | .   | G       | .   | .  | .   | Q/E | .   | LV      | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1/2        | R5      | R5                           | R5         |    |
| 130170  | .   | .       | .   | .       | .   | .  | .   | K/E | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 1-4       | R5         | R5      | R5                           |            |    |
| 130234  | .   | .       | .   | .       | .   | .  | .   | T   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | E     | .  | .   | N   | .   | .   | .   | .   | .     | .   | 35 | 3         | R5         | R5      | R5                           |            |    |
| 130270  | .   | .       | .   | .       | .   | .  | .   | E   | .   | G       | .   | .     | .   | .   | .   | .   | R     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | K   | .   | .   | .   | .   | .     | .   | 35 | 3         | R5         | R5      | R5                           |            |    |
| 130278  | .   | .       | .   | .       | .   | .  | .   | E   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | D/N | .   | .   | .   | .   | .   | .     | .   | .  | 35        | 1/2        | R5      | R5                           | R5         |    |
| 130320  | .   | .       | R/T | .       | .   | .  | .   | K/E | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 0-3       | R5         | R5      | R5                           |            |    |
| 130333  | .   | .       | .   | .       | .   | .  | .   | T   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | V   | . | .  | .   | .   | .     | .  | .   | N   | .   | .   | Y   | .   | .     | .   | 35 | 3         | R5         | R5      | R5                           |            |    |
| 130348  | .   | .       | .   | .       | .   | .  | .   | Q   | .   | .       | .   | .     | .   | .   | .   | .   | .     | .   | .   | .   | . | .  | .   | .   | .     | .  | .   | .   | .   | .   | .   | .   | .     | .   | 35 | 2         | R5         | R5      | R5                           |            |    |

Figure 2. V3 amino acid sequences alignment and genotypic results of subtype C isolates. V3 amino acid sequences from 53 patients infected with HIV-1 subtype C (at *env* region) were aligned to the consensus sequence where dot identifies consensus amino acid, dash indicates a deletion and slash separates all possible amino acids at each position. The highlighted samples correspond to recent seroconverters. N-glycosylation site is boxed, positions 11 and 25 are shaded in light gray and the crown motif is shaded in dark gray. V3 length indicates the number of amino acids compounding this region. V3 net charge and genotype predicted by Geno2pheno<sub>10%</sub> (G2p 10%), WebPSSM<sub>SI/NSI(C)</sub> and R5/X4-Pred are also showed. X4 predictions and characteristics indicative of X4 viruses are highlighted in red font.

Table 1. Baseline characteristics of the study population

|  | Total (n=90)      | Subtype <i>env</i> |                   | <i>p</i> value |
|--|-------------------|--------------------|-------------------|----------------|
|  |                   | B (n=37)           | C (n=53)          |                |
| Male gender, no. (%)                         | 53 (58.9%)        | 22 (59.5%)         | 31 (58.5%)        | 1.000          |
| Mean age (years), $\pm$ SD                   | 35.28 $\pm$ 10.45 | 34.97 $\pm$ 9.35   | 35.49 $\pm$ 11.23 | 0.819          |
| Transmission route, no. (%)                  |                   |                    |                   |                |
| Sexual                                       | 87 (96.7%)        | 36 (97.3%)         | 51 (96.2%)        | 1.000          |
| Intravenous drug use                         | 5 (5.6%)          | 0 (0.0%)           | 5 (9.4%)          | 0.146          |
| Unknown                                      | 1 (1.1%)          | 1 (2.7%)           | 0 (0.0%)          | 0.856          |
| Recent seroconverters, no. (%)               | 16 (17.8%)        | 6 (16.2%)          | 10 (18.9%)        | 0.965          |
| Mean plasma HIV-RNA, log copies/mL, $\pm$ SD | 4.34 $\pm$ 0.86   | 4.44 $\pm$ 0.91    | 4.28 $\pm$ 0.83   | 0.388          |

SD, standard deviation

Table 2. Distribution of CXCR4-tropic strains among subtypes *pol/env* accordingly to different genotypic tools.

| Genotypic Tool            | Total <sup>+</sup><br>(n=90) | Subtype <i>env</i> <sup>+</sup> |    |   |                                 |    |   | <i>p</i> value |
|---------------------------|------------------------------|---------------------------------|----|---|---------------------------------|----|---|----------------|
|                           |                              | B (n=37)                        |    |   | C (n=53)                        |    |   |                |
|                           |                              | Subtype <i>pol</i> <sup>†</sup> |    |   | Subtype <i>pol</i> <sup>†</sup> |    |   |                |
|                           |                              | B                               | BC | C | C                               | BC | B |                |
| Geno2pheno <sub>10%</sub> | 13 (14.4%)                   | 6                               | 0  | 0 | 3                               | 3  | 1 | 0.924          |
|                           |                              | 6 (16.2%)                       |    |   | 7 (13.2%)                       |    |   |                |
| WebPSSM <sup>*</sup>      | 21 (23.3%)                   | 3                               | 1  | 0 | 8                               | 8  | 1 | 0.036          |
|                           |                              | 4 (10.8%)                       |    |   | 17 (32.1%)                      |    |   |                |
| R5X4-Pred                 | 8 (8.9%)                     | 3                               | 1  | 0 | 3                               | 1  | 0 | 0.874          |
|                           |                              | 4 (10.8%)                       |    |   | 4 (7.5%)                        |    |   |                |

<sup>\*</sup> HIV-1 subtype B co-receptor usage was assessed using WebPSSM<sub>X4/R5(B)</sub> and HIV-1 subtype C co-receptor usage was assessed using WebPSSM<sub>SI/NSI(C)</sub>.

<sup>+</sup> Results are expressed as absolute numbers and percentages.

<sup>†</sup> Results are expressed as absolute numbers.

Table 3. Profiles of R5 and X4 strains classified by the different predictors.

| Subtype | Characteristic              | Geno2pheno <sub>10%</sub> |           |          | WebPSSM*  |           |          | R5/X4-Pred |           |          |
|---------|-----------------------------|---------------------------|-----------|----------|-----------|-----------|----------|------------|-----------|----------|
|         |                             | R5 (n=31)                 | X4 (n=6)  | <i>p</i> | R5 (n=33) | X4 (n=4)  | <i>p</i> | R5 (n=33)  | X4 (n=4)  | <i>p</i> |
| B       | Mean viral load, ± SD       | 4.37±0.91                 | 4.79±0.85 | 0.305    | 4.45±0.94 | 4.29±0.70 | 0.732    | 4.39±0.94  | 4.77±0.55 | 0.436    |
|         | Mean net charge             | 3.87                      | 4.00      | 0.787    | 3.76      | 5.00      | 0.023    | 3.82       | 4.50      | 0.224    |
|         | Basic aa at positions 11/25 | 2                         | 3         | 0.022    | 1         | 4         | < 0.001  | 2          | 3         | 0.005    |
|         | Recent seroconverters       | 6                         | 0         | 0.561    | 6         | 0         | 1.000    | 6          | 0         | 1.000    |
|         |                             | R5 (n=46)                 | X4 (n=7)  | <i>p</i> | R5 (n=36) | X4(n=17)  | <i>p</i> | R5 (n=49)  | X4 (n=4)  | <i>p</i> |
| C       | Mean viral load, ± SD       | 4.30±0.85                 | 4.14±0.75 | 0.654    | 4.28±0.94 | 4.28±0.56 | 0.999    | 4.28±0.83  | 4.25±0.93 | 0.947    |
|         | Mean net charge             | 3.09                      | 3.43      | 0.431    | 3.00      | 3.41      | 0.188    | 3.08       | 3.75      | 0.227    |
|         | Basic aa at position 25     | 0                         | 1         | 0.132    | 0         | 1         | 0.321    | 0          | 1         | 0.075    |
|         | Recent seroconverters       | 9                         | 1         | 1.000    | 7         | 3         | 1.000    | 9          | 1         | 1.000    |

\* HIV-1 subtype B co-receptor usage was assessed using WebPSSM<sub>X4/R5(B)</sub> and HIV-1 subtype C co-receptor usage was assessed using WebPSSM<sub>sinsi(C)</sub>. SD: standard deviation. Basic aa at positions 11/25 (subtype B) or 25 (subtype C): indicates the number of sequences with R or K amino acids at these positions.

### 3 CONCLUSÕES GERAIS E PERSPECTIVAS

O conhecimento do tropismo do HIV-1 e das características genéticas da alça V3, principal região envolvida na determinação da especificidade por co-receptores, é importante para a compreensão da história natural da doença, uso terapêutico de antagonistas de CCR5 e predição da progressão da doença. Considerando que as infecções por HIV-1 no Rio Grande do Sul são causadas predominantemente pelos subtipos B, C e recombinantes BC, este estudo se propôs a analisar o uso de co-receptores e as características genéticas da região V3 dessas principais cepas circulantes no estado. Amostras de plasma de pacientes infectados com HIV-1 recrutados no Projeto Ampliar entre 2006 e 2008, cujo subtipo viral, baseado na região do gene *pol*, era conhecido, foram selecionadas. Todos os pacientes incluídos neste estudo são virgens de tratamento antirretroviral, representando assim a evolução natural do vírus em relação à progressão da doença. O subtipo da região *env* e o tropismo viral foram determinados através de ferramentas disponíveis na internet, após amplificação e sequenciamento da região V3 do gene *env*.

O perfil do tropismo por co-receptores das cepas de HIV-1 do subtipo B se mostrou semelhante ao descrito na literatura. Já os isolados do subtipo C apresentaram uma alta prevalência de cepas trópicas por CXCR4, incluindo isolados de pacientes com infecção recente, fato previamente observado somente entre vírus de algumas regiões da África. A origem africana do subtipo C brasileiro nos levou a sugerir que as mesmas características virais das cepas africanas que estão promovendo o uso de CXCR4 pelo HIV-1 são responsáveis pela alta prevalência de X4 observada nessa região. A presença de cepas trópicas por CXCR4 em indivíduos com infecção recente demonstra a rápida evolução do

vírus e está relacionada com mau prognóstico da doença (queda acelerada da contagem de células T CD4<sup>+</sup> e risco aumentado de progressão para AIDS). Adicionalmente, a detecção dessas variantes em indivíduos virgens de tratamento antirretroviral tem implicações terapêuticas importantes, diminuindo as opções de terapia medicamentosa, já que inibidores de CCR5 não podem ser administrados em portadores de cepas X4.

A análise da sequência da alça V3 revelou características similares e divergentes entre os subtipos B e C. Em ambos os subtipos, a perda do sítio de *N*-glicosilação foi raramente observada, independentemente do tropismo viral, e o número de aminoácidos da alça V3 apresentou pouca variação. Essa pequena variação de comprimento foi associada com o uso de CXCR4, em concordância com os dados da literatura. A presença de aminoácidos básicos nas posições 11/25 e elevada carga total da alça V3 foram associadas com tropismo por CXCR4 somente no subtipo B. Esses resultados indicam que diferentes características virais estão envolvidas na evolução do HIV-1 para o uso de CXCR4 entre subtipos B e C dessa região.

A caracterização da região do tetrâmero do topo da alça V3 revelou a presença de quatro assinaturas mais prevalentes entre as cepas do subtipo B e duas assinaturas entre as cepas do subtipo C. Em concordância com as descrições para o subtipo B brasileiro: o tetrâmero GPGR foi o mais frequente, seguido dos tetrâmeros B'- GWGR, GFGR e APGR; e a presença do tetrâmero B'- GWGR foi relacionada com tropismo por CCR5. Entre os isolados do subtipo C, observou-se a prevalência do tetrâmero GPGQ, e o tetrâmero GPGR foi associado com uso de CXCR4, reforçando a hipótese de que a arginina na posição 18 facilita a utilização do CXCR4.

Este estudo encontrou uma prevalência significativa de cepas X4 entre indivíduos virgens de tratamento antirretroviral do Rio Grande do Sul, especialmente entre indivíduos

infectados com o subtipo C, o que indica que os inibidores de CCR5 devem ser utilizados cautelosamente nessa população. Esses dados devem ser confirmados através de estudos que analisem um número maior de indivíduos e que utilizem a metodologia fenotípica para determinação do tropismo viral, já que esta é considerada padrão-ouro e foi aprovada para uso na prática clínica.

#### 4 REFERÊNCIAS BIBLIOGRÁFICAS

Alkhatib G (2009) The biology of CCR5 and CXCR4. *Current Opinion in HIV and AIDS* 4: 96-103.

Araujo AF, Brites C, Monteiro-Cunha J, Santos LA, Galvao-Castro B, Alcantara LC (2010) Lower prevalence of human immunodeficiency virus type 1 Brazilian subtype B found in northeastern Brazil with slower progression to AIDS. *AIDS Research and Human Retroviruses* 26(11):1249-54.

Bello G, Guimarães ML and Morgado MG (2006) Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* 20(5): 763-68.

Bello G, Eyer-Silva WA, Couto-Fernandez JC, Guimarães ML, Chequer-Fernandez SL, Teixeira SL and Morgado MG (2007) Demographic history of HIV-1 subtypes B and F in Brazil. *Infection, Genetics and Evolution* 7(2): 263-70.

Bello G, Passaes CPB, Guimarães ML, Lorete RS, Almeida SEM, Medeiros RM, Alencastro PR and Morgado MG (2008) Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS* 22: 1993-2000.

Berger EA, Doms RW, Fenyö EM, Korber BTM, Littman DR, Moore JP, Sattentau QJ, Schuitemaker H, Sodroski J and Weiss RA (1998) A new classification for HIV-1. *Nature* 391: 240.

Bongertz V, Bou-Habib DC, Brígido LF, Caseiro M, Chequer PJ, Couto-Fernandez JC, Ferreira PC, Galvão-Castro B, Greco D, Guimarães ML, Linhares de Carvalho MI, Morgado MG, Oliveira CA, Osmanov S, Ramos CA, Rossini M, Sabino E, Tanuri A and Ueda M; Brazilian Network for HIV Isolation and Characterization (2000) HIV-1 diversity in Brazil: genetic, biologic, and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. *Journal of Acquired Immune Deficiency Syndromes* 23(2):184-93.

Brígido LF, Franco HM, Custódio RM, Oliveira CA, P Ferreira JL, Eira M, Bergel F, Araújo F, Carneiro JR, Rodrigues R (2005) Molecular characteristics of HIV type 1 circulating in São Paulo, Brazil. *AIDS Res Hum Retroviruses* 21(7):673-82.

Brígido LFM, Nunes CC, Oliveira CM, Knoll RK, Ferreira JLP, Freitas CA, Alves MA, Dias C and Rodrigues R (2007) HIV type 1 subtype C and CR1 recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. *AIDS Research and Human Retroviruses* 23(12): 1579-85.

Butler IF, Pandrea I, Marx PA and Apetrei C (2007) HIV genetic diversity: biological and public health consequences. *Current HIV Research* 5: 23-45.

CDC - Centers for Disease Control and Prevention (1993) Revised classification system for HIV infection and expanded surveillance case definition for AIDS among



adolescents and adults. United States. Disponível em: <http://www.cdc.gov/mmwr/preview/mmwrhtml/00018871.htm>. Acesso em: 16/05/2009.

CDC - Centers for Disease Control and Prevention (2008) Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged <18 months and for HIV infection and AIDS among children aged 18 months to < 13 years. United States. Disponível em: <http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5710a1.htm>. Acesso em: 01/08/2009.

Coetzer M, Cilliers T, Ping LH, Swanstrom R and Morris L (2006) Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology* 356: 95-105.

Connell BJ, Michler K, Capovilla A, Venter WDF, Stevens WS and Papathanasopoulos MA (2008) Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. *AIDS* 22(7): 896-99.

Daar ES, Kesler KL, Petropoulos CJ, Huang W, Bates M, Lail AE, Coakley EP, Gomperts ED and Donfield SM; Hemophilia Growth and Development Study (2007) Baseline HIV type 1 coreceptor tropism predicts disease progression. *Clinical Infectious Diseases* 45(5):643-9.

De Brito A, Komninakis SC, Novoa P, de Oliveira RM, Fonseca LA, Duarte AJ and Casseb J (2006) Women infected with HIV type 1 Brazilian variant, subtype B (B'-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). *Clinical Infectious Diseases* 43:1476-81.

De Sá Filho DJ, Sucupira MC, Caseiro MM, Sabino EC, Diaz RS, Janini LM (2006) Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS Research Human Retroviruses* 22(1): 1-13.

De Jong JJ, Goudsmit J, Keulen W, Klaver B, Krone W, Tersmette M and De Ronde A (1992) Human Immunodeficiency Virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *Journal of Virology* 66(2): 757-65.

Delobel P, Nugeyre MT, Cazabat M, Pasquier C, Marchou B, Massip P, Barre-Sinoussi F, Israel N and Izopet J (2007) Population-based sequencing of the V3 region *env* for predicting the coreceptor usage of human immunodeficiency virus type 1 quasispecies. *Journal of Clinical Microbiology* 45(5): 1572-80.

Dias CF, Nunes CC, Freitas IO, Lamego IS, Oliveira IM, Gilli S, Rodrigues R, Brigido LF (2009) High prevalence and association of HIV-1 non-B subtype with specific sexual transmission risk among antiretroviral naïve patients in Porto Alegre, RS, Brazil. *Revista do Instituto de Medicina Tropical de São Paulo* 51(4):191-96.

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

Ferraro GA, Mello MAG, Sutmoller F, Weyenbergh JV, Brigido LFM, Greco D, Morgado M, Shindo N, Galvão-Castro B and Bou-Habib DC (2001) Biological characterization and chemokine receptor usage of HIV type 1 isolates prevalent in Brazil. *AIDS Research and Human Retroviruses* 17(13): 1241-47.

Ferreira JLP, Thomaz M, Rodrigues R, Harrad D, Oliveira CM, Oliveira CAF, Batista JPG, Ito TS and Brigido LFM (2008) Molecular characterisation of newly identified HIV-1 infections in Curitiba, Brazil: preponderance of clade C among males with recent infections. *Memórias do Instituto Oswaldo Cruz* 103(8): 800-08.

Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F and Schuitemaker H (1992) Phenotype-associated sequence variation in the third variable domain of the Human Immunodeficiency Virus type 1 gp120 molecule. *Journal of Virology* 66(5):3183-87.

Fouchier RA, Brouwer M, Broersen SM and Schuitemaker H (1995) Simple determination of Human Immunodeficiency Virus type 1 syncytium-inducing V3 genotype by PCR. *Journal of Clinical Microbiology* 33(4):906-11.

Franca RFO, Castro-Jorge LA, Neto RJP, Jorge DMM, Lima DM, Colares JKB, Paula SO and Fonseca BAL (2011) Genotypic characteristics of HIV type 1 based on gp120 hypervariable region 3 of isolates from southern Brazil. *AIDS Research and Human Retroviruses* [Epub ahead of print].

Frange P, Galimand J, Goujard C, Deveau C, Ghosn J, Rouzioux C, Meyer L and Chaix ML (2009) High frequency of X4/DM-tropic viruses in PBMC samples from patients with primary HIV-1 subtype-B infection in 1996-2007: the French ANRS CO6 PRIMO Cohort Study. *Journal of Antimicrobial Chemotherapy* 64(1):135-41.

Gallo RC (2002) The early years of HIV/AIDS. *Science* 298: 1728-30.

Garrido C, Roulet V, Chueca N, Poveda E, Aguilera A, Skrabal K, Zahonero N, Carlos S, García F, Faudon JL, Soriano V and Mendoza C (2008) Evaluation of eight different bioinformatics tools to predict viral tropism in different Human Immunodeficiency Virus type 1 subtypes. *Journal of Clinical Microbiology* 46(3): 887-91.

Glass WG, McDermott DH, Lim JK, Lekhong S, Yu SF, Frank WA, Pape J, Cheshier RC and Murphy PM (2006) CCR5 deficiency increases risk of symptomatic West Nile virus infection. *Journal of Experimental Medicine* 203(1):35-40.

Guimarães ML, Eyer-Silva WA, Couto-Fernandez JC, Morgado MG (2008) Identification of two new CRF\_BF in Rio de Janeiro State, Brazil. *AIDS* 22(3): 433-35.

Guo Y, Ma L, Yuan L, Wang S, Sun J, Xu W, Xu J, Xing H, Hong K, Zhang X, Ruan Y, Zhang Y and Shao Y (2007) R5 to X4 coreceptor switch of human immunodeficiency virus type 1 B' and B'/C recombinant subtype isolates in China. *Chinese Medical Journal* 120(6): 522-25.

Hartley O, Klasse PJ, Sattentau QJ and Moore JP (2005) V3: HIV's switch-hitter. *AIDS Research and Human Retroviruses* 21(2): 171-89.

Hemelaar J, Gouws E, Ghys PD and Osmanov S WHO-UNAIDS Network for HIV Isolation and Characterisation (2011) Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* 25(5):679-89.

Hierholzer J, Montano S, Hoelscher M, Negrete M, Hierholzer M, Avila MM, Carrillo MG, Russi JC, Vinales J, Alava A, Acosta ME, Gianella A, Andrade R, Sanchez JL, Carrion G, Sanchez JL, Russell K, Robb M, Birx D, McCutchan F, Carr JK (2002) Molecular Epidemiology of HIV Type 1 in Ecuador, Peru, Bolivia, Uruguay, and Argentina. *AIDS Res Hum Retroviruses* 18(18):1339-50.

Hoffmann C, Rockstroh JK and Kamps BS (2007) HIV medicine 2007. 15 Ed. Flying Publisher. Paris. Disponível em: <http://www.hivmedicine.com/hivmedicine2007.pdf>. Acesso em: 27/01/2011.

ICTVdB Management - International Committee on Taxonomy of Viruses (2006) 00.061.1.06.009. Human immunodeficiency virus 1. In: *ICTVdB - The Universal Virus Database*, version 4. Büchen-Osmond, C. (Ed), Columbia University, New York, United States. Disponível em: <http://www.ictvdb.org/ICTVdB/00.061.1.06.009.htm>. Acesso em: 24/01/2011.

Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, McLaughlin S, Shankarappa R, Margolick JB and Mullins JI (2003) Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of Human Immunodeficiency Virus type 1 env V3 loop sequences. *Journal of Virology* 77(24):13376-88.

Johnston ER, Zijenah LS, Mutetwa S, Kantor R, Kittinunvorakoon C, Katzenstein DA (2003) High frequency of syncytium-inducing and CXCR4-tropic viruses among human immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral treatment. *Journal of Virology* 77(13):7682-8.

Kawamura T, Kurtz SE, Blauvelt A and Shimada S (2005) The role of Langerhans cells in the sexual transmission of HIV. *Journal of Dermatological Science* 40: 147-55.

Koot M, van 't Wout AB, Kootstra NA, de Goede RE, Tersmette M and Schuitemaker H (1996) Relation between changes in cellular load, evolution of viral phenotype, and the clonal composition of virus populations in the course of human immunodeficiency virus type 1 infection. *The Journal of Infectious Diseases* 173(2): 349-54.

Kuiken C, Leitner T, Foley B, Hahn B, Mullins J, Apetrei C, Mizrachi I, Rambaut A, Wolinsky S and Korber B (2010) HIV Sequence Compendium 2010. Los Alamos National Laboratory. Theoretical Biology and Biophysics, Los Alamos, New Mexico, p: 4-7.

Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA and Landau NR (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86: 367-77.

Little SJ, Holte S, Routy JP, Daar ES, Markowitz M, Collier AC, Koup RA, Mellors JW, Connick E, Conway B, Kilby M, Wang L, Whitcomb JM, Hellmann NS and Richman DD (2002) Antiretroviral-drug resistance among patients recently infected with HIV. *The New England Journal of Medicine* 347(6):385-94.

Lodowski DT and Palczewski K (2009) Chemokine receptors and other G protein-coupled receptors. *Current Opinion in HIV and AIDS* 4: 88-95.

Los Alamos National Laboratory (2010) HIV and SIV Nomenclature. United States. Disponível em: <http://www.hiv.lanl.gov/content/sequence/HelpDocs/subtypes-more.html>. Acesso em: 27/01/2011.

Lusso P (2006) HIV and the chemokine system: 10 years later. *The European Molecular Biology Organization Journal* 25 (3): 447-56.

Ma L, Guo Y, Yuan L, Huang Y, Sun J, Qu S, Yu X, Meng Z, He X, Jiang S and Shao Y (2009) Phenotypic and genotypic characterization of Human Immunodeficiency virus type 1 CRF07\_BC strains circulating in the Xinjiang province of China. *Retrovirology* 6:45.

Ministério da Saúde (2009) Boletim Epidemiológico AIDS e DST. Ano VI nº 01. Brasília.

Minton K (2005) New mechanism for APOBEC3G? *Nature Reviews Immunology* 5 (3): 183.

Montagnier L (2002) A history of HIV discovery. *Science* 298: 1727-1728.

Monteiro JP, Ferraro GA, Oliveira T, Goldani LZ, Kashima S, Alcantara LC, Morgado MG, Bou-Habib DC, Galvão-Castro B (2007) Genetic and biologic characterization of HIV type 1 subtype C isolates from south Brazil. *AIDS Research and Human Retroviruses* 23(1):135-43.

Moore JP, Kitchen SG, Pugach P and Zack JA (2004) The CCR5 and CXCR4 coreceptors – central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. *AIDS Research and Human Retroviruses* 20(1): 111-126.

Moore JP and Kuritzkes DR (2009) A pièce de resistance: how HIV-1 escapes small molecule CCR5 inhibitors. *Current Opinion in HIV and AIDS* 4: 118-24.

Mosier, DE (2009) How HIV changes its tropism: evolution and adaptation? *Current Opinion in HIV and AIDS* 4:125-30.

Nabatov AA, Pollakis G, Linnemann T, Kliphuis A, Chalaby MIM and Paxton WA (2004) Inpatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. *Journal of Virology* 78(1): 524-30.

Passaes CPB, Bello G, Lorete R, Almeida SEM, Junqueira DM, Veloso VG, Morgado MG, Guimarães ML (2009) Genetic characterization of HIV-1 BC recombinants and evolutionary history of the CFR31\_BC in southern Brazil. *Infection, Genetics and Evolution* 9: 474-82.

Patel MB, Hoffman NG and Swanstrom R (2008) Subtype-specific conformational differences within the V3 region of subtype B and subtype C human immunodeficiency virus type 1 Env proteins. *Journal of Virology* 82(2): 903-16.

Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J and Paxton WA (2001) N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *The Journal of Biological Chemistry* 276(16): 13433-41.

Poveda E, Briz V, Roulet V, González MM, Faudon JL, Skrabal K and Soriano V (2007) Correlation between a phenotypic assay and three bioinformatic tools for determining HIV co-receptor use. *AIDS* 21(11): 1487-90.

Prusiner SB (2002) Discovering the cause of AIDS. *Science* 298: 1726-27.

Qu W, Robinson M and Zhang FJ (2008) Factors influencing the natural history of HIV-1 infection. *Chinese Medical Journal* 121 (24): 2613-21.

Raymond S, Delobel P, Mavigner M, Cazabat M, Souyris C, Sandres-Sauné K, Cuzin L, Marchou B, Massip P and Izopet J (2008) Correlation between genotypic predictions based on V3 sequences and phenotypic determination of HIV-1 tropism. *AIDS* 22: F11-F16.

Raymond S, Delobel P, Mavigner M, Ferradini L, Cazabat M, Souyris C, Sandres-Sauné K, Pasquier C, Marchou B, Massip P and Izopet J (2010) Prediction of HIV type 1 subtype C tropism by genotypic algorithms built from subtype B viruses. *Journal of Acquired Immune Deficiency Syndromes* 53(2):167-75.

Reeves JD and Doms RW (2002) Human immunodeficiency virus type 2. *Journal of General Virology* 83:1253-65.

Robertson DL, Hahn BH, Sharp PM (1995) Recombination in AIDS viruses. *Journal of Molecular Evolution* 40(3):249-59.

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, *et al.* (2000) HIV-1 nomenclature proposal. *Science* 288: 55-58.

Rubial-Ares BH, Belmonte L, Baré PC, Parodi CM, Massud I and Bracco MME (2004) HIV-1 infection and chemokine receptor modulation. *Current HIV Research* 2: 39-50.

Salem AH, Farid E, Fadel R, Abu-Hijleh M, Almawi W, Han K and Batzer MA (2009) Distribution of four HIV type 1-resistance polymorphisms (CCR5 $\Delta$ 32, CCR5-m303, CCR2-64I and SDF1-3'A) in the Bahraini population. *AIDS Research and Human Retroviruses* 25 (10): 973-77.

Salemi M, Burkhardt BR, Gray RR, Ghaffari G, Sleasman JW and Goodenow MM (2007) Phylodynamics of HIV-1 in lymphoid and non-lymphoid tissues reveals a central role for the thymus in emergence of CXCR4-using quasispecies. *PLoS ONE* 2(9): e950.

Sánchez V, Masiá M, Robledano C, Padilla S, Ramos JM and Gutiérrez F (2010) Performance of genotypic algorithms for predicting HIV-1 tropism measured against the enhanced-sensitivity Trofile coreceptor tropism assay. *Journal of Clinical Microbiology* 48(11):4135-9.

Santos AF, Sousa TM, Soares EAJM, Sanabani S, Martinez AMB, Sprinz E, Silveira J, Sabino EC, Tanuri A and Soares MA (2006) Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS* 20: 2011-19.

Santos AF, Schrago CG, Martinez AMB, Mendonza-Sassi R, Silveira J, Sousa TM, Lengruher RB, Soares EAJM, Sprinz E and Soares MA (2007) Epidemiologic and evolutionary trends of HIV-1 CRF31\_BC-related strains in southern Brazil. *Journal of Acquired Immune Deficiency Syndromes* 45(3): 328-33.

Seclén E, Garrido C, González Mdel M, González-Lahoz J, de Mendoza C, Soriano V and Poveda E (2010) High sensitivity of specific genotypic tools for detection of X4 variants in antiretroviral-experienced patients suitable to be treated with CCR5 antagonists. *The Journal of Antimicrobial Chemotherapy* 65(7): 1486-92.

Sierra S, Kupfer B and Kaiser R (2005) Basics of the virology of HIV-1 and its replication. *Journal of Clinical Virology* 34: 233-44.

Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues F, Büch J, Däumer M, Kaiser R, Lengauer T and Harrigan PR (2007) Predicting HIV co-receptor usage based on genetic and clinical covariates. *Antiviral Therapy* 12:1097-106.

Souza AC, Oliveira CM, Marques HH and Levi JE (2009) Description of an HIV-1 BC Recombinant Virus Identified in a Pediatric Patient in the City of São Paulo. *The Brazilian Journal of Infectious Diseases* 13(1): 67-69.

Scarlatti G, Tresoldi E, Björndal A, Fredriksson R, Colognesi C, Deng HK, Malnati MS, Plebani A, Siccardi AG, Littman DR, Fenyö EM, Lusso P (1997) In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nature Medicine* 3(11):1259-65.

Thorsen I, Polzer S and Schreiber M (2002) Infection of cells expressing CXCR4 mutants lacking N-glycosylation at the N-terminal extracellular domain is enhanced for R5X4-dualtropic Human Immunodeficiency Virus type-1. *BMC Infectious Diseases* 2:31.

Toledo PVM, Carvalho DS, Rossi SG, Brindeiro R and Queiroz-Telles F (2010) Genetic diversity of Human Immunodeficiency virus-1 isolates in Paraná, Brazil. *Brazilian Journal of Infectious Diseases* 14(3): 230-236.

Turner BG and Summers MF (1999) Structural biology of HIV. *Journal of Molecular Biology* 285: 1-32.

UNAIDS – Joint United Nations Programme on HIV/AIDS (2008) Report on the global AIDS epidemic 2008. Geneva. Disponível em: [http://www.unaids.org/en/media/unaids/contentassets/dataimport/pub/globalreport/2008/jc1510\\_2008globalreport\\_en.pdf](http://www.unaids.org/en/media/unaids/contentassets/dataimport/pub/globalreport/2008/jc1510_2008globalreport_en.pdf). Acesso em: 22/01/2011.

UNAIDS – Joint United Nations Programme on HIV/AIDS (2010) Global report: UNAIDS report on the global AIDS epidemic 2010. Geneva. Disponível em: [http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2010/20101123\\_globalreport\\_en.pdf](http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2010/20101123_globalreport_en.pdf). Acesso em: 22/01/2011.

van't Wout AB and Jensen MA (2008) HIV-1 tropism and its detection – impact on infection, transmission and treatment. *European Infectious Diseases* 2 (1): 60-64.

Vargas AE, Cechim G, Correa JF, Gomes PA, Macedo Gde S, de Medeiros RM, Perotoni G, Rauber R, Villodre ES and Chies JA (2009) Pros and cons of a missing chemokine receptor--comments on "Is the European spatial distribution of the HIV-1-resistant CCR5-D32 allele formed by a breakdown of the pathocenosis due to the historical Roman expansion?" by Eric Faure and Manuela Royer-Carenzi (2008). *Infection, Genetics and Evolution* 9(4): 387-9.

Zhang C, Xu S, Wei J and Guo H (2009) Predicted co-receptor tropism and sequence characteristics of China HIV-1 V3 loops: implications for the future usage of CCR5 antagonists and AIDS vaccine development. *International Journal of Infectious Diseases* 13: e212-e216.

WHO – World Health Organization (2007) WHO Case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related

disease in adults and children. Geneva. Disponível em: <http://www.who.int/hiv/pub/guidelines/hivstaging/en/index.html>. Acesso em: 23/01/2011.

Xu S, Huang X, Xu H and Zhang C (2007) Improved prediction of coreceptor usage and phenotype of HIV-1 based on combined features of V3 loop sequence using random forest. *The Journal of Microbiology* 45(5):441-46.