Instituto de Medicina Tropical de São Paulo
Universidade de São Paulo

Liã Bárbara Arruda

Determinação do tropismo do HIV-1 pelos correceptores CCR5 e CXCR4 pelo uso de ferramentas de bioinformática

Dissertação apresentada ao Instituto de Medicina Tropical de São Paulo, Universidade de São Paulo, para obtenção do título de Mestre em Ciências, no programa de Medicina Tropical

Orientador: Prof. Dr. Jorge Simão do Rosário Casseb

São Paulo
2010

Liã Bárbara Arruda

Determinação do tropismo de HIV-1 pelos correceptores CCR5 e CXCR4 pelo uso de ferramentas de bioinformática

Dissertação apresentada ao Instituto de Medicina Tropical de São Paulo, Universidade de São Paulo, para a obtenção do título de Mestre em Ciências

Área de concentração: Doenças Tropicais e Saúde Internacional
Orientador: Prof. Dr. Jorge Simão do Rosário Casseb

São Paulo
2010

Ficha catalográfica
Preparada pela Biblioteca do Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo
© Reprodução autorizada pelo autor

Arruda, Liã Bárbara.

Determinação do tropismo do HIV-1 pelos correceptores CCR5 e CXCR4 pelo uso de ferramentas de bioinformática / Liã Bárbara Arruda. – São Paulo, 2010.

Dissertação (Mestrado) – Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Área de concentração: Doenças Tropicais e Saúde Internacional

Orientador: Jorge Simão do Rosário Casseb

Descritores: 1. HIV. 2. TROPISMO. 3. BIOINFORMÁTICA.

Universidade de São Paulo

Insitituto de Medicina Tropical de São Paulo

Candidato: Liã Bárbara Arruda

Titulo da Dissertação: “Determinação do tropismo do HIV-1 pelos correceptores CCR5 e CXCR4 pelo uso de ferramentas de bioinformática”

Orientador: Prof. Dr. Jorge Simão do Rosário Casseb

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**, em sessão pública realizada a ...../...../....., considerou

( ) APROVADO(A) **( ) REPROVADO(A)**

Examinador(a) Assinatura.....

Nome.....

Instituição.....

Examinador(a) Assinatura.....

Nome.....

Instituição.....

Presidente Assinatura.....

Nome.....

Instituição.....

A Margrit e Hildegard, que são o meu alicerce.

AGRADECIMENTOS

Aos pacientes que foram extremamente compreensivos e, tão generosamente, aceitaram participar deste estudo.

Ao Programa de Pós-graduação em Medicina Tropical do Instituto de Medicina Tropical de São Paulo pela oportunidade de cursar o Mestrado.

Ao Dr. Alberto Duarte que abriu as portas do laboratório para que eu pudesse além desenvolver este projeto, integrar à família LIM56. Ao Dr. Gil Benard e à Dra. Maria Sato pelas oportunidades e por manterem estas portas abertas.

Ao Dr. Jorge Casseb por incontáveis razões, mas principalmente pela suprema paciência, essencial para a orientação deste projeto (e desta aluna inquieta).

Ao Dr. Cláudio Gonsalez, Dr. Lucas Medeiros, Dr. Eduardo Lagonegro e à Dra. Ana Paula Veiga e Enf. Daniela pela colaboração imprescindível no recrutamento de pacientes do ambulatório ADEE3002.

Ao Dr. José Eduardo Levi e a Dra. Ester Sabino pelas críticas e contribuições no exame de qualificação.

À Fundação de Amparo à Pesquisa do Estado de São Paulo por prover a bolsa de mestrado.

À minha pequena e inestimável família: minha mãe e minha *Oma*, que são meu verdadeiro alicerce.

Precedendo a lista de todos os outros colaboradores diretos e indiretos, fico na esperança de não esquecer ninguém. Entretanto, antecipo o pedido de perdão para o caso de lapsos de memória.

Aos amigos do grupo Imunovirologia (os Cassebetes): Adriana de Brito e Fernando Costa que me apoiam desde os primeiros passos; Ingrid Olah e Rosana Alcade que me cederam os primeiros protocolos; Maira Martinez e Paulo Quatroqui por me ajudarem (e aturarem) nos experimentos; Fábio Cabral e Karen Brunialti pela constante troca de idéias e Karen Oliveira, que também leva o espírito do grupo. Um forte abraço a Fábio

Cabral e Fernando Costa pela paciência na revisão do texto. A todos, um imenso agradecimento pela amizade, colaboração e incentivo.

Ao Dr. Heitor de Andrade e à Solange Ferreira por resolverem todas as dúvidas sobre a pós-graduação.

À Rosangêla Araújo pelo companheirismo e colaboração essencial durante a coleta das amostras.

A todos que compõem a secretaria, especialmente Edna Reis e Juliana Pisok, mestres na arte de desenrolar burocracias.

A Noemia Orii e José Eduardo Martins pela realização dos exames de CD4+/CD8+ e carga viral, e Fábio Silva pela grande ajuda no sequenciador.

Ao Demetrius Vignati pelas soluções em informática e ao Rodrigo Gomes pela edição das figuras.

À Cláudia Finazzo pela amizade e por me apresentar ao Prof. Raymundo Soares de Azevedo Neto quem desmistificou a curva ROC para mim.

Ao Rafael Martins e Guilherme Silveira pelas intermináveis discussões e contribuições.

À Vanessa Batista e Viviane Bressani pelas confidências e amizade e a muitos outros amigos e colegas do LIM56: Léia, Juliana, Soraya, Paula, Luli, Luciana, Camila, Nilson, Shinai, Cyro, Mayce, Nátalli, Suzana, Adriana L., Andressa, Dr. Alexandre, Dra. Telma, Dra. Alessandra, Dr. Dewton...

E mesmo que a lista fosse infindável não contemplaria a todos a quem sou grata.

“Para chegar à realidade, uma ideia começa por se apoderar de espíritos fervorosos e escraviza-os. A partir desse momento, eles pertencem-lhe e não veem diante de si se não o objetivo a atingir. Por vezes, esse objetivo parece inatingível: quanto mais nos adiantamos, mais ele nos parece distante. Mas que importa? Os escravos de uma ideia são incapazes de desanimar.”

Marie Curie

RESUMO

Arruda LB. *Determinação do tropismo do HIV-1 pelos correcpetores CCR5 e CXCR4 pelo uso de ferramentas de bioinformática* (dissertação). São Paulo: Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo; 2010.

A sequência de 35 aminoácidos da alça V3 da gp120 do gene *env* do HIV-1 é o principal determinante do tropismo viral pelos correceptores CCR5 ou CXCR4, utilizados pelo HIV-1 para a entrada na célula. O desenvolvimento de estratégias antirretrovirais baseadas no uso dos correceptores representa um avanço importante para o controle da progressão da infecção. Entretando, o uso clínico dos antagonistas de CCR5 implica na determinação do tropismo das cepas virais do indivíduo infectado e os programas preditores de bioinformática para a determinação do tropismo poderiam ser uma alternativa mais acessível para a triagem dos candidatos ao uso dos antagonistas de CCR5. Este estudo teve como objetivo utilizar ferramentas de bioinformática para a predição de tropismo e avaliar sua aplicabilidade na prática clínica. Foram coletadas amostras de sangue periférico de 101 indivíduos infectados pelo HIV-1 e sob acompanhamento clínico, dos quais foram extraídas amostras de DNA proveniente de PBMCs. As amostras de DNA foram amplificadas por PCR para a região da alça V3, das quais foram obtidas 94 sequências. Os sistemas preditivos foram avaliados utilizando 185 sequências com tropismo conhecido provenientes de banco de dados. Com base nesta análise foi possível elaborar um algoritmo para a predição do tropismo com 94% de confiabilidade. Assim, a predição das 94 amostras demonstrou uma prevalência de 80% (n=75) de cepas R5 e 20% (n=19) de cepas X4. Os sistemas preditivos de tropismo podem representar uma importante estratégia para a triagem dos candidatos ao uso dos antagonistas de coreceptor, porém, não são capazes de substituir completamente os ensaios padrão-ouro para a determinação do tropismo.

DESCRITORES: HIV-1. Tropismo. Bioinformática.

ABSTRACT

Arruda LB. *Determination of HIV-1 coreceptor usage by CCR5 and CXCR4 coreceptors using bioinformatic tools* (dissertation). Sao Paulo: Institute of Tropical Medicine of Sao Paulo of University of Sao Paulo; 2010.

The 35 amino acids of the V3-gp120 of HIV-1 *env* gene is the main determinant of viral tropism by the coreceptors CCR5 and CXCR4 used for HIV-1 cell entry. The development of antiretroviral strategies based on the coreceptor usage represents an important step to control the infection progression. However, the clinical application of CCR5 antagonists involves the coreceptor usage determination of viral strains in the infected individual. The bioinformatics predictive programs for coreceptor usage determination could be a more available alternative for screening candidates to receive CCR5 antagonists. This study aimed to employ bioinformatics tools to predict tropism and assess its applicability in clinical practice. Peripheral blood samples were collected from 101 individuals infected with HIV-1 and under clinical follow-up. DNA samples were extracted from PBMCs. The DNA samples were amplified by PCR and 94 V3 sequences were obtained. The predictive systems were evaluated using 185 sequences of known tropism from a database. This analysis provides the construction of an algorithm showing 94% of reliability. Thus, the 94 sample prediction showed a prevalence of 80% (n=75) of R5 strain and 20% (n=19) of X4 strain. The predictive systems could be an important strategy in the screening of the tropism. Nonetheless, they are not able to fully replace the coreceptor usage gold-standard assays.

DESCRIPTORS: HIV-1. Tropism. Bioinformatics.

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação da estrutura morfológica do HIV. 20
Figura 2 – Mapa gênico do HIV-1, sequência referência HXB2. 20
Figura 3 – Mecanismo de entrada do HIV. 23
Figura 4 – Tropismo celular.. 26
Figura 5 – Evolução do HIV-1 em um indivíduo infectado. 27
Figura 6 – Ilustração do método utilizado no ensaio *Trofile.* 33
Figura 7 – Fluxograma das amostras coletadas e sequenciadas. 46
Figura 8 – Distribuição dos subtipos de HIV-1 encontrados na população estudada 48
Figura 9 – Carga líquida das sequências obtidas no banco de dados de Los Alamos 48
Figura 10 – Curva ROC do conjunto de testes geno2pheno$_{[coreceptor]}$, sinsi, x4r5 e SVM 51
Figura 11 – Algoritmo para determinação do tropismo utilizando sistemas preditivos. 53
Figura 12 – Carga elétrica líquida das amostras obtidas neste estudo..... 55

LISTA DE TABELAS

Tabela 1 – Testes de predição de tropismo oferecidos pelas plataformas livres 35
Tabela 2 – Sequências dos *primers* utilizados para amplificar e sequenciar a região V3 do gene *env* 40
Tabela 3 – Testes de predição de tropismo utilizados neste estudo..... 43
Tabela 4 – Número de sequências com tropismo conhecido obtidas no banco de dados de Los Alamos 44
Tabela 5 – Dados demográficos e laboratoriais 47
Tabela 6 – Número de sequências obtidas de Los Alamos com tropismo conhecido e sua predição pelas ferramentas de bioinformática 49
Tabela 7 – Sensibilidade e especificidade dos testes de bioinformática para a predição de cepas X4 50
Tabela 8 – Área abaixo da curva ROC dos sistemas preditivos 51
Tabela 9 – Relação entre os dados laboratoriais e o tropismo inferido 54

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ADEE 3002	Ambulatório de Imunodeficiências Secundárias do Departamento de Dermatologia
ART	terapia antirretroviral (*antiretroviral therapy*)
CAPPesq	Comissão de Ética para Análise de Projetos de Pesquisa
cDNA	DNA complementar
cel/mm³	Células por milímetro cúbico
CEP	Comissão de Ética em Pesquisa
cp/mL	Cópias por mililitro
CRFs	Formas recombinantes circulantes (*circulating recombinant forms*)
DNA	Ácido desoxirribonucléico (*Deoxyribonucleic Acid*)
dNTPs	Deoxiribonucleosídeos 5'-trifosfatados (dATP, dTTP, dCTP, dGTP) (*deoxynucleoside triphosphates*)
EDTA	Ácido etilenodiaminotetracético (*ethylenediaminetetraacetic acid*)
EUA	Estados Unidos da América
gp	Glicoproteína
HAART	Terapia Antirretroviral Altamente Ativa (*highly active antiretroviral therapy)*
HCFMUSP	Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
HIV-1	Vírus da Imunodeficiência Humana tipo 1 (*Human Immunodeficiency Virus type 1*)
LTR	Região de longa repetição terminal (*long terminal repeat*)
Kb	Quilobases (10^3 pares de base)
$MgCl_2$	Cloreto de magnésio
µL	Microlitro (10^{-6} litro)
mL	Mililitro (10^{-3} litro)
mM	Milimolar (10^{-3} molar)
min	Minuto

ng	Nanograma (10^{-9} grama)
nm	Nanômetro (10^{-9} metro)
NNRTI	Inibidores da transcriptase reversa não-nucleosídeos (*non-nucleoside reverse transcriptase inhibitors*)
NSI	Não indutores de sincício (*non-syncytium-inducing*)
pb	Pares de bases
PBMCs	Células mononucleares do sangue periférico (*Peripheral Blood Mononuclear Cells*)
PCR	Reação em cadeia da polimerase (*Polymerase Chain Reaction*)
pmol	Picomol (10^{-12} mol)
PSSM	Matriz de escore de posição específica (*position specific score matrice*)
R5	Tropismo pelo correceptor CCR5
R5X4	Tropismo pelos correceptores CCR5 e CXCR4
RANTES	Regulado sobre a ativação de célula T expressa e secretada (*regulated upon activation T cell expressed and secreted*)
RNA	ácido ribonucléico (*ribonucleic acid*)
RT	transcriptase reversa (*reverse transcriptase*)
RVA	ensaio de vírus recombinantes (*recombinant vírus assay*)
SDF	Fator derivado de estroma celular (*stromal cell-derived factor*)
s	Segundo
SI	Indutores de sincício (*syncytium-inducing*)
SIV	Vírus da Imunodeficiência Símia (*Simian Immunodeficiency Virus*)
SMV	Ferramenta de suporte de vetores (*support vector machine*)
TCLE	Termo de Consentimento Livre e Esclarecido
U	Unidade

UV	Ultravioleta
X4	Tropismo pelo correceptor CXCR4
xg	Gravidade
~	Aproximadamente
°C	Graus Celsius
%	Porcento

Aminoácidos

A (Ala)	Alanina
C (Cys)	Cisteína
D (Asp)	Aspartato ou ácido aspártico
E (Glu)	Glutamato ou ácido glutâmico
F (Phe)	Fenilalanina
G (Gly)	Glicina
H (His)	Histidina
I (Ile)	Isoleucina
K (Lys)	Lisina
L (Leu)	Leucina
M (Met)	Metionina
N (Asn)	Asparagina
P (Pro)	Prolina
Q (Gln)	Glutamina
R (Arg)	Arginina
S (Ser)	Serina
T (Thr)	Treonina
V (Val)	Valina
W (Trp)	Triptofano
Y (Tyr)	Tirosina

Nucleotídeos

A	Adenina
C	Citosina
G	Guanina
T	Timina

SUMÁRIO

1 INTRODUÇÃO 18
1.1 O Vírus 19
1.2 A variabilidade genética do HIV-1 20
1.3 A entrada do HIV-1 na célula hospedeira 22
1.4 O tropismo do HIV 23
1.4.1 Os correceptores 24
1.4.2 As cepas do HIV-1 25
1.4.3 A mudança de tropismo 27
1.5 Os antagonistas de correceptor 29
1.6 Determinação do tropismo 31
1.6.1 Ensaios fenotípicos 32
1.6.2 Ensaios genotípicos e programas preditores 34
2 OBJETIVOS 37
2.1 Objetivo geral 37
2.2 Objetivos específicos 37
3 MÉTODOS 38
3.1 Casuística 38
3.1.1 Critérios de inclusão 38
3.1.2 Critérios de exclusão 38
3.1.3 Aspectos éticos 38
3.2 Coleta das amostras 39
3.3 Separação de células mononucleares do sangue periférico 39
3.4 Extração do DNA pró-viral 40
3.5 Amplificação da região V3 do gene *env* 40
3.6 Detecção dos produtos de PCR 41
3.7 Purificação dos produtos amplificados 41
3.8 Reação de sequenciamento 41
3.9 Análise e interpretação do sequenciamento 42
3.9.1 Determinação dos subtipos 42
3.9.2 Determinação da carga elétrica líquida 42

3.10 Predição do tropismo 43
3.11 Avaliação da predição do tropismo 43
3.12 Análise estatística 44
4 RESULTADOS 46
4.1 Dados demográficos e laboratoriais 46
4.2 Prevalência dos subtipos de HIV-1 47
4.3 Avaliação da carga elétrica líquida 48
4.4 Predição do tropismo 49
4.4.1 Avaliação da predição do tropismo 49
4.4.2 Elaboração de um algoritmo para a predição do tropismo 51
4.4.3 Predição do tropismo das amostras obtidas neste estudo 53
5 DISCUSSÃO 56
5.1 Considerações metodológicas 58
5.2 Prevalência dos subtipos de HIV-1 59
5.3 Avaliação dos preditores de tropismo 59
5.4 Considerações finais 64
6 CONCLUSÕES 66
REFERÊNCIAS 67
ANEXO 72
Anexo A – Certificado emitido pela Comissão de Ética 72
APÊNDICE 73
Apêndice A – Termo de Consentimento Livre e Esclarecido 73

1 INTRODUÇÃO

A falha no tratamento, em pacientes portadores do Vírus da Imunodeficiência Humana (HIV), devido à emergência de cepas virais com resistência às drogas dirigiu o desenvolvimento de novas classes de antirretrovirais capazes de manter a carga viral a níveis indetectáveis. Entretanto, para pacientes com resistência a diversas drogas as opções de tratamento permanecem bastante limitadas (Hughes e Nelson, 2009).

Devido às descobertas fundamentais sobre o mecanismo de fusão/entrada, fatores que afetam a história natural da transmissão do HIV e o desenvolvimento de novos conceitos de tratamento e prevenção da infecção pelo HIV, foi desenvolvida uma série de estratégias terapêuticas capazes de inibir a progressão da infecção (Alkhatib e Berger, 2007). Uma contribuição importante foi a descoberta dos receptores de quimiocinas CCR5 e CXCR4 e seu papel como correceptores essenciais para a entrada viral, que provocou uma década de pesquisas sobre os mecanismos de entrada e possíveis alvos para tratamento e prevenção do HIV (Hughes e Nelson, 2009).

A afinidade ou tropismo *in vitro* de diferentes cepas do HIV pelos correceptores CCR5 e CXCR4 incitou novas perspectivas e trouxe importantes informações sobre a história natural da infecção. Porém, explicações sobre o aparecimento de cepas com tropismo para o correceptor CXCR4 e com tropismo duplo durante a progressão da doença e sobre a diferença na patogenicidade, ainda representam uma incógnita (Alkhatib e Berger, 2007). Assim, entender a base molecular do tropismo do HIV-1 pode prover informações genéticas capazes de definir cofatores celulares que são críticos para passos específicos do ciclo de vida viral. A caracterização molecular destes cofatores poderá identificar alvos potenciais para o desenvolvimento de novos agentes antivirais (Cullen, 2001).

O desenvolvimento de estratégias terapêuticas baseadas no uso dos correceptores representa um avanço fundamental para o controle da progressão da infecção, sendo utilizadas principalmente como adjuvantes

em terapias aplicadas em pacientes infectados com cepas multirresistentes à terapia convencional (Emmelkamp e Rockstroh, 2007). Porém, a utilização dos antagonistas de CCR5 implica na determinação do tropismo das cepas de HIV prevalentes no indivíduo infectado. Entretanto, os ensaios fenotípicos para seleção dos candidatos ao uso destas novas estratégias antirretrovirais ainda não constituem uma rotina clínica, devido aos seus altos custos e baixa disponibilidade, restringindo o seu acesso (Braun e Wiesmann, 2007). Assim, a difusão dos *softwares* preditores de tropismo a partir de sequências de nucleotídeos ou aminoácidos da região V3 pode representar uma alternativa mais viável para a triagem dos candidatos ao uso dos antagonistas de CCR5 (Sierra et al., 2007).

A utilização das ferramentas de bioinformática propicia estudos que possam caracterizar alterações genotípicas específicas do tropismo viral, e a identificação da mudança de tropismo ao longo da infecção, contribuindo com a adoção de estratégias terapêuticas que visam a inibição da progressão da infecção pelo HIV.

1.1 O Vírus

Os agentes etiológicos da Síndrome da Imunodeficiência Adquirida (Aids), HIV tipos 1 e 2, pertencem à família *Retroviridae*, que são vírus de RNA envelopados, contendo a enzima transcriptase reversa (RT – *reverse transcriptase*), responsável pela capacidade deste grupo de vírus se replicar através de um DNA pró-viral (Karpas, 2004).

O genoma viral é constituído por duas moléculas idênticas de RNA fita simples (~ 9,7 Kb) com polaridade positiva e compreende nove genes flanqueados por regiões de longa repetição terminal (LTR – *long terminal repeat*) (Figuras 1 e 2). A replicação eficiente dos retrovírus depende de três genes: *gag*, *pol* e *env*. Os genes que codificam a expressão de proteínas regulatórias são *tat* e *rev*, enquanto *vif*, *nef*, *vpu* e *vpr* são classificados como genes acessórios por não serem essenciais na replicação *in vitro* (Rubbert et al., 2007; Turner e Summers, 1999).

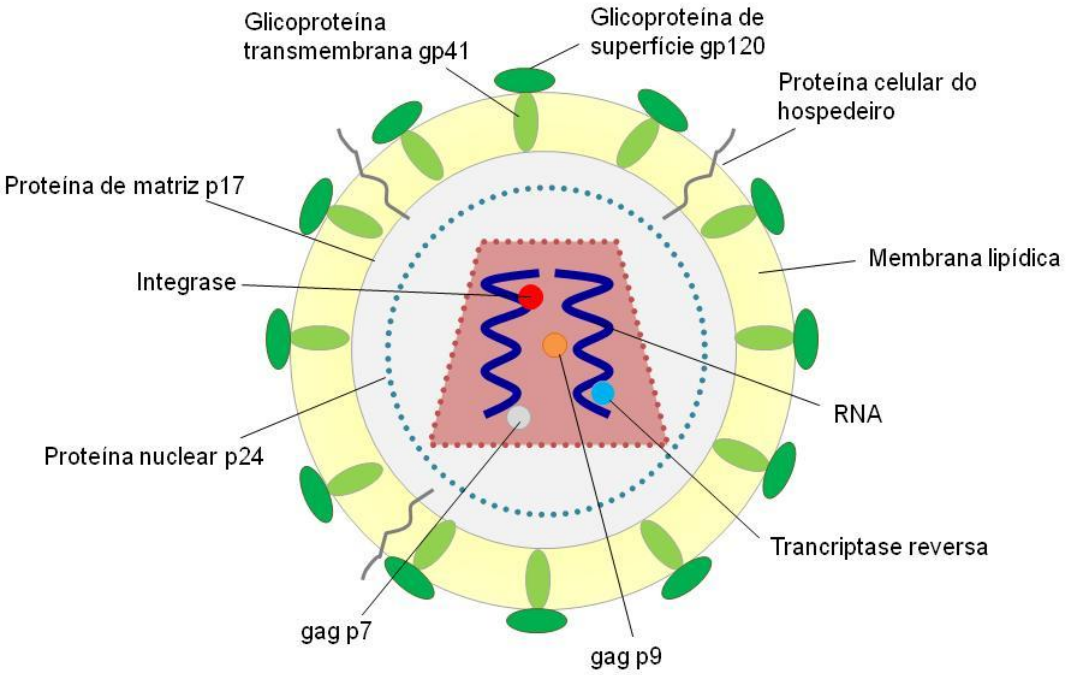

Figura 1 – Representação da estrutura morfológica do HIV. Baseado em Rubbert et al., (2007).

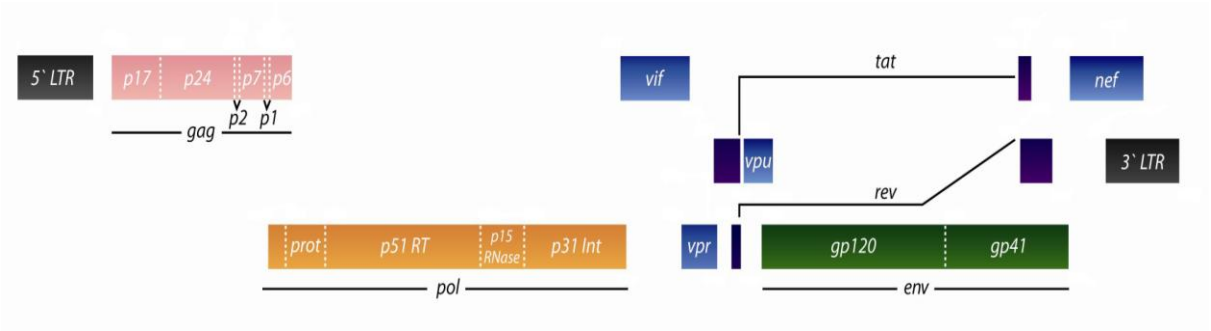

Figura 2 – Mapa gênico do HIV-1, sequência referência HXB2. Adaptado do banco de dados de sequências do HIV do Laboratório Nacional de Los Alamos (http://hiv-web.lanl.gov).

1.2 A variabilidade genética do HIV-1

Patógenos e vírus, em particular, estão sujeitos a uma forte pressão seletiva durante a infecção, desenvolvendo um elevado grau de variabilidade genética, o que, consequentemente, é uma das principais características do HIV. Esta variabilidade é decorrente dos frequentes erros durante a replicação pela transcriptase reversa e da rápida troca de gerações de

vírions (*turnover*) nos indivíduos infectados (Peeters, 2000; Simon-Loriere et al., 2009).

A variabilidade do genoma do HIV envolve um lento acúmulo de mutações pontuais (inserções e deleções) produzidas pela ausência de atividade de exonuclease da transcriptase reversa e da alta capacidade recombinogênica desta polimerase. A recombinação é um importante mecanismo evolutivo que contribui para a diversificação genética permitindo grandes alterações genômicas. Uma vez que a recombinação geralmente envolve genes que codificam produtos funcionais, a probabilidade de produzir uma progênie viável é grande comparada com a inserção de um número equivalente de mutações pontuais aleatórias (Simon-Loriere et al., 2009).

Para ocorrer recombinação, uma mesma célula precisa ser infectada por vírions geneticamente distintos. Quando duas variantes virais distintas infectam uma mesma célula, ocorre um complexo processo envolvendo múltiplos *crossing-overs*, originando um genoma único e infectivo. A recombinação, portanto, tem origem em uma dupla-infecção ou superinfecção, onde o portador de um determinado subtipo se reinfecta em contato com outro subtipo. Apesar de ser mais evidente entre cepas de subtipos diferentes, a recombinação ocorre igualmente em todos os níveis filogenéticos (Peeters, 2000; Simon-Loriere et al., 2009).

A recombinação do HIV implica em aspectos chave da patogênese viral como escape do sistema imunológico do hospedeiro, transmissibilidade, evolução para resistência aos antirretrovirais e transmissão entre espécies (Simon-Loriere et al., 2009). Em virtude desta flexibilidade genética, o HIV-1 pode ser dividido em grupos, subtipos e sub-subtipos além de formas recombinantes circulantes (CRFs – *circulating recombinant forms*). Os grupos M (*Main*), O (*Outlier*) e N (*Non-M/Non-O*) resultam de transmissões interespecíficas independes do SIVcpz*Ptt* (*Simian Immunodeficiency Virus* – vírus da imunodeficiência símia) dos chimpanzés *Pan troglodytes troglodytes* para os humanos, enquanto os grupos O e N são raramente encontrados e têm alta diversidade genética (Keele et al., 2006; Peeters, 2000). A maior

parte dos isolados de HIV-1 pertence ao grupo M que é subdividido em nove subtipos: A, B, C, D, F, G, H, J e K. Existem ainda sub-subtipos (A1, A2; F1, F2, etc.) e mais de 40 CRFs descritas atualmente (Leitner et al., 2005; Peeters, 2000; Robertson et al., 2000).

1.3 A entrada do HIV-1 na célula hospedeira

O gene *env* do HIV produz a poliproteína gp160 que é clivada em duas subunidades: a proteína de superfície gp120 e a proteína transmembrana gp41. A proteína de superfície viral gp120 tem um alto grau de variabilidade genética apresentando, em sua sequência de aminoácidos, cinco regiões variáveis (V1-V5) interespaçadas por quatro regiões constantes (C1-C4), sendo que três gp120-gp41 associadas formam uma estrutura trimérica em forma de pico na superfície do envelope glicoprotéico (Este e Telenti, 2007; Hughes e Nelson, 2009).

As proteínas de superfície do HIV ligam-se à superfície celular através do receptor primário CD4 (membro da superfamília das imunoglobulinas), ancorando o vírus na superfície da célula hospedeira e promovendo uma interação adicional à proteína correceptora (usualmente um membro da família de receptores de quimiocinas). Apesar de uma dúzia de correceptores de HIV terem sido identificados *in vitro*, todas as cepas de HIV-1 utilizam CCR5 e CXCR4 para a entrada na célula *in vivo* (Este e Telenti, 2007; Hughes e Nelson, 2009).

Após a ligação CD4-vírus e o reconhecimento do correceptor ocorre uma mudança conformacional na gp120 expondo a alça V3 (Figura 3). A interação entre a alça V3 e o correceptor implica no melhor ancoramento entre o CD4 celular e a gp120. Esse ancoramento leva a exposição da, até então inacessível, gp41 promovendo a fusão das membranas. Acredita-se que a gp41 assuma uma conformação de seis hélices aproximando as membranas o suficiente para formar o poro de fusão no intuito de internalizar o vírus (Este e Telenti, 2007; Hughes e Nelson, 2009).

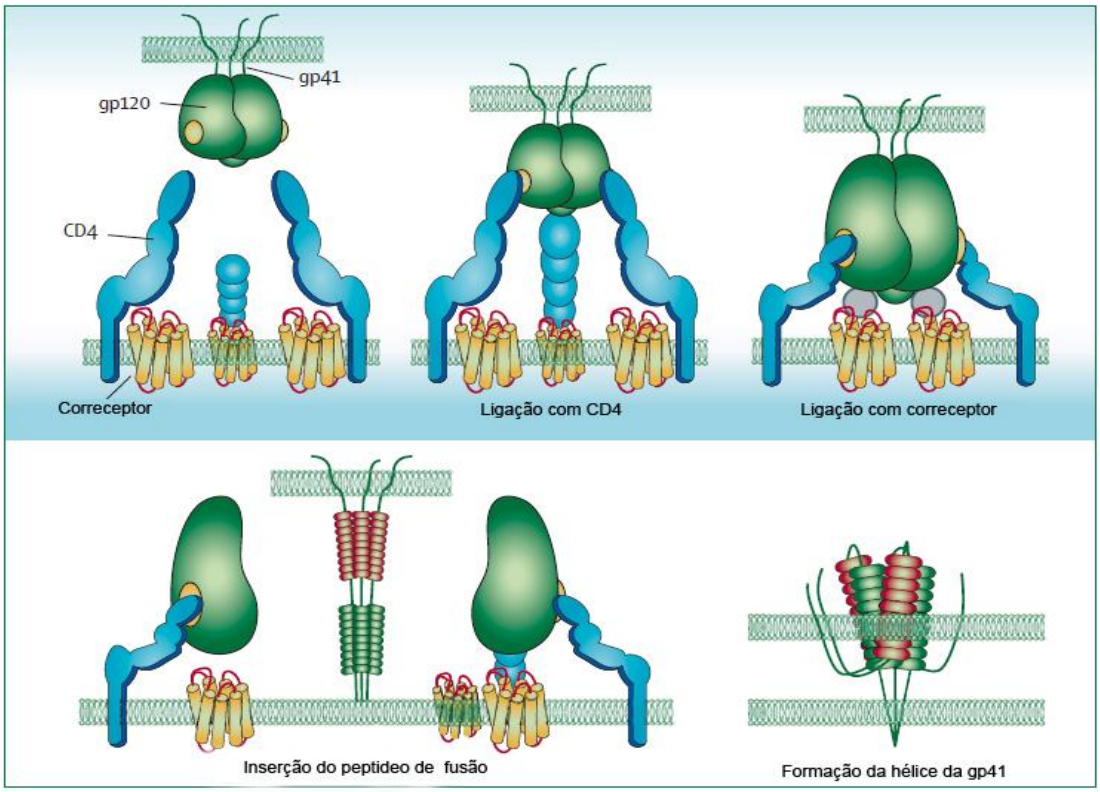

Figura 3 – Mecanismo de entrada do HIV. Durante a ligação com o CD4 (sítio de ligação em amarelo) a gp120 muda de conformação, expondo os epítopos que se ligam aos receptores de quimiocinas. Posteriormente, a gp41 adota a conformação fusiogênica e as hélices do N-terminal (verde) e C-terminal (vermelho) formam uma estrutura de *hairpin*, levando à aproximação das membranas viral e celular e resultando na fusão das membranas. Adaptado de Este e Telenti (2007).

1.4 O tropismo do HIV

O tropismo de um vírus é definido pelas células, dentro de uma determinada espécie, que são capazes de suportar uma infecção produtiva por aquele vírus. Os vírus são extremamente dependentes de fatores da célula hospedeira para o sucesso completo do seu ciclo de vida, incluindo não apenas receptores na superfície celular, mas também uma série de cofatores intracelulares. Quanto mais dependente de um cofator particular é o vírus, mais limitada será sua capacidade de infectar diferentes tipos de células. O HIV-1 apresenta um tropismo muito específico em termos não somente de número de espécies, mas também de número de células em uma espécie particular que são permissivas a este vírus patogênico (Cullen, 2001).

1.4.1 Os correceptores

Uma década após a descoberta de que o CD4 é o principal receptor do HIV, dois receptores de quimiocinas com sete domínios transmembrana acoplados a proteína G, CCR5 e CXCR4, foram descobertos como correceptores determinantes para a entrada do HIV na célula hospedeira (Cullen, 2001; Hughes e Nelson, 2009). Apesar do CCR5 e CXCR4 representarem os correceptores mais relevantes para o HIV-1 *in vivo*, foram identificados *in vitro* outros receptores de quimiocinas, como o CCR2, CCR3, CCR8, CCR9, STRL33, Gpr15, Gpr1, APJ e ChemR23 que podem ser utilizados para a entrada de certos isolados de HIV (Munerato et al., 2003; Rubbert et al., 2007).

As quimiocinas são proteínas pró-inflamatórias de 68 a 150 aminoácidos que, fisiologicamente, são responsáveis pela mediação da quimiotaxia das células T e fagócitos para as zonas inflamatórias e que na infecção por HIV-1 podem bloquear a replicação viral, competindo pela ligação aos seus receptores naturais. De acordo com seu motivo comum de cisteínas, as quimiocinas são classificadas em dois grandes grupos: CXC (α-quimiocinas) e CC (β-quimiocinas) (Pollakis et al., 2004; Rubbert et al., 2007). Ao receptor CCR5 ligam-se as β-quimiocinas RANTES (*regulated upon activation T cell expressed and secreted* – regulado sobre a ativação de célula T expressa e secretada) MIP-1α (*macrophage inhibitory protein* - proteína inibitória de macrófago) e MIP-1β, que são capazes de inibir a entrada de isolados virais com tropismo para CCR5 em células T. O SDF-1 (*stromal cell-derived factor* - fator derivado de estroma celular) é o ligante natural do receptor CXCR4 e, portanto, capaz de inibir a entrada de isolados com tropismo para CXCR4 em células T (Munerato et al., 2003; Pollakis et al., 2004; Rubbert et al., 2007).

Polimorfismos genéticos ocorrem naturalmente em receptores de quimiocinas e podem estar associados com a progressão mais rápida ou mais favorável da infecção. O polimorfismo mais significativo para infecção por HIV-1 é a deleção de 32 pb no gene CCR5, o que inibe a expressão

deste receptor na superfície celular. A homozigose da mutação CCR5Δ32 revelou a resistência completa à infecção pelo HIV em indivíduos expostos ao vírus com tropismo para este correceptor, havendo relatos de casos isolados de infecção por cepas X4 nestes indivíduos. Em heterozigose (Δ32/wt) esta mutação representa uma expressão diminuída de CCR5 na superfície celular e os indivíduos infectados tendem a uma progressão mais lenta da infecção. Esta descoberta reforça a importância do correceptor CCR5 na história natural da infecção (Alkhatib e Berger, 2007; Rubbert et al., 2007)

O polimorfismo CCR5Δ32 ocorre em homozigose em 1% da população caucasiana e a frequência da heterozigose ocorre entre cinco e 15%. Apesar desta mutação não ser prejudicial ao funcionamento fisiológico normal do sistema imune, sua presença reduz a progressão da infecção pelo HIV e em receptores CCR5 não-funcionais leva a completa resistência da infecção pelo HIV, sugerindo a superioridade dos vírus com tropismo para CCR5 na transmissão do HIV, embora, desta vez, direcionada pelo hospedeiro em vez de fatores virais (Emmelkamp e Rockstroh, 2007; Hughes e Nelson, 2009; Jones e Nelson, 2007).

1.4.2 As cepas do HIV-1

O HIV-1 pode ser caracterizado de acordo com o correceptor utilizado na infecção das células T CD4+ (Figura 4). Os previamente descritos “macrófagotrópicos”, vírus não-indutores de sincício (NSI), correspondem aos vírus CCR5-trópicos (R5), enquanto os “células T-trópicos”, vírus indutores de sincício (SI), os quais são associados à progressão da infecção, correspondem aos vírus CXCR4-trópicos (X4) (Simon-Loriere et al., 2009). Isolados de HIV classificados com R5X4 são cepas capazes de infectar tanto células com CCR5 quanto células com CXCR4 na superfície, e desta forma são denominadas cepas duplo-trópicas (D) ou ainda podem representar uma mistura (M) de cepas R5 e X4 que coinfectam um mesmo indivíduo, uma vez que ainda não existem métodos que sejam capazes de discernir uma

mistura de um conjunto de cepas com tropismo duplo (Cullen, 2001; Este e Telenti, 2007; Hughes e Nelson, 2009; Koning et al., 2002).

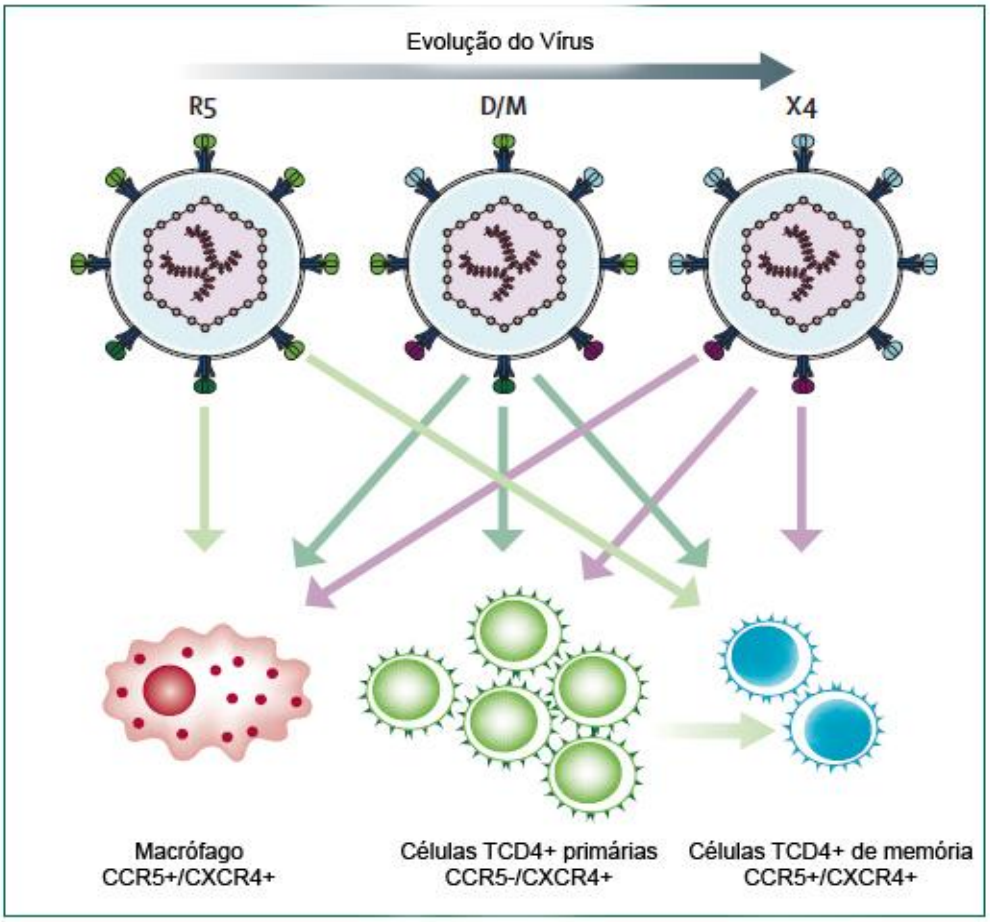

Figura 4 – Tropismo celular. Células T CD4+ primárias são CXCR4+ e CCR5-, e raramente infectadas por vírus R5. Cepas R5 são comumente transmitidas e persistem ao longo da infecção, mas podem envolver populações virais duplo-trópicas (D) ou misturas (M), e o fenótipo X4 aparece em estágios avançados da infecção. Adaptado de Esté e Telenti (2007).

Com relação a história natural, as cepas R5 são relatadas como envolvidas na infecção primária e predominando na fase assintomática, enquanto as cepas X4 emergiriam ao longo da infecção e seriam responsáveis pela depleção acelerada de células T e rápida progressão da doença (Figura 5). As cepas R5X4 surgem durante a infecção, na transição de R5 para X4, mas o papel evolutivo desta transição ainda não foi decifrado (Fouchier et al., 1995; Jensen et al., 2003; Mild et al., 2009).

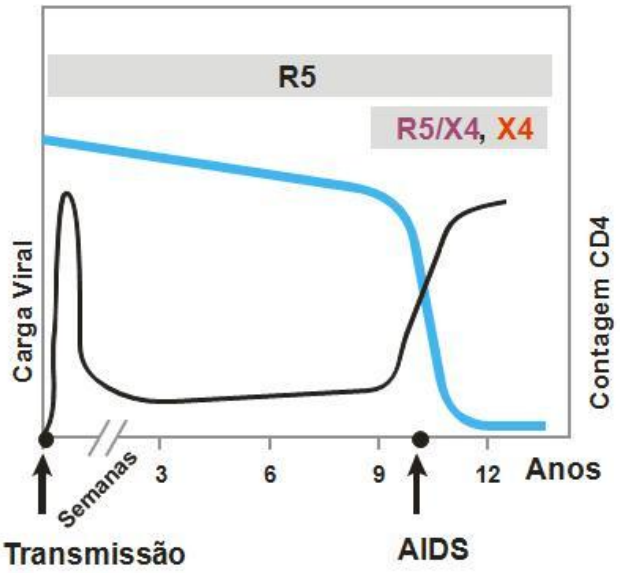

Figura 5 – Evolução do HIV-1 em um indivíduo infectado, desde a transmissão, a fase assintomática e a progressão da doença. Adaptado de Alkatib e Berger (2007).

O tropismo prevalente em um indivíduo infectado é independente da rota de transmissão ou do tropismo que predomina na fonte. Os vírus R5 parecem melhor adaptados virologicamente em relação aos X4. Estudos sugerem que 81-88% dos infectados que nunca utilizaram antirretrovirais sejam R5, 12-19% sejam R5X4 e menos de 1% seja X4. Entre os pacientes não-expostos a antirretrovirais, vírus X4 estão associados a baixas contagens de células T CD4 e carga viral alta. Se os X4 fossem mais virulentos, a progressão da infecção seria esperada. Entretanto, 50% dos pacientes que morreram por AIDS mantinham-se R5 e a possibilidade de dois fenótipos R5 tem sido sugerida como vírus R5 iniciais e tardios com diferentes efeitos imunológicos (Hughes e Nelson, 2009).

1.4.3 A mudança de tropismo

Os mecanismos utilizados pelo HIV para a mudança do tropismo ainda não estão claros. Uma possibilidade é que a população viral seja regulada por fatores da célula hospedeira e do sistema imune no início da infecção, resultando em uma população precoce de R5 que é mais propensa a manter o uso exclusivo de R5 ou evoluir para X4 ao longo da infecção.

Isso explica porque apenas alguns indivíduos desenvolvem cepas X4 (Mild et al., 2009).

Os esforços para identificar o gene do HIV-1 que controlava os fenótipos levaram a análises mutacionais mais detalhadas que revelaram que o determinante primário do tropismo tecidual do HIV-1 está alocado no terceiro domínio variável (V3) da subunidade gp120 do envelope, mas que regiões do primeiro e segundo domínios também podem modular o tropismo tecidual em um menor grau. A alça V3, portanto, é o principal determinante do tropismo, mas mutações genéticas fora desta área também podem afetar o tropismo, como a extensão da região V2 ou o número de sítios de glicosilação na região V1-V3 (Cullen, 2001; Hughes e Nelson, 2009; Mild et al., 2009).

Observou-se que isolados R5, usualmente, apresentam carga elétrica mais baixa na região V3, quando comparados a isolados X4, enquanto que o correceptor CCR5 apresenta carga elétrica mais alta que o CXCR4. Portanto, mutações que conferem trocas de aminoácidos podem alterar a carga elétrica e, consequentemente, o tropismo viral (Pollakis et al., 2004; Sierra et al., 2007). A ligação entre V3 e o correceptor é carreada pela presença de aminoácidos básicos (K ou R), aminoácidos ácidos (D ou E) ou modificações pós-transcricionais (geralmente N ou O-glicosilações ou sulfatação da tirosina), consequentemente, mutações que afetam a carga elétrica da região V3 estão intrinsecamente correlacionadas com a seletividade do correceptor (Rosen et al., 2006; Sierra et al., 2007).

A sequência dos 35 aminoácidos da região V3 é o principal determinante do tropismo do vírus pelo correceptor. Na progressão da infecção, alterações na região V3 podem acarretar na mudança do tropismo, levando a substituição de populações R5 por cepas X4 (Rosen et al., 2006; Sierra et al., 2007). A presença de resíduos básicos nas posições 11 e 25 (posições 306 e 322 na sequência consenso do subtipo B) está associada a vírus X4 ou duplo-trópicos, enquanto a presença de um resíduo de carga negativa e um neutro nas posições 25 e 11, respectivamente, está correlacionada com vírus R5. Numerosas investigações confirmam que a

presença de um resíduo carregado positivamente na posição 25 é capaz de converter o tropismo de uma cepa R5 para X4 (Poon et al., 2007; Rosen et al., 2006; Sander et al., 2007). Ainda não está claro se o aparecimento de aminoácidos básicos nas posições 11 e 25 é suficiente ou mesmo necessário para a emergência de cepas X4, mas diferenças intrínsecas na alça V3 entre os subtipos podem levar a diferenças na emergência das cepas X4 e é provável que a troca do tropismo seja um processo gradual de acúmulo de mutações (Hughes e Nelson, 2009; Jensen et al., 2003).

1.5 Os antagonistas de correceptor

O conhecimento da patogênese da infecção pelo HIV-1 permitiu o desenvolvimento de novas estratégias terapêuticas. A introdução dos inibidores de protease e dos inibidores da transcriptase reversa não nucleosídeos (NNRTI) nos regimes terapêuticos antirretrovirais em 1995, iniciou a era da terapia antirretroviral altamente ativa (HAART), resultando na melhoria do prognóstico dos pacientes infectados (Rubbert et al., 2007). Ainda assim, surgiram novos problemas relacionados à toxicidade dos fármacos a curto e longo prazo e a ocorrência de mutações de resistência viral. Recentemente, novas classes de drogas antirretrovirais baseadas na variabilidade genotípica e fenotípica viral foram desenvolvidas com foco no processo de entrada do vírus na célula como os inibidores de ligação, inibidores de fusão e os antagonistas de correceptores. Concomitantemente, o desenvolvimento de novas técnicas de triagem para a avaliação do tropismo viral, vem auxiliando a escolha da melhor classe de antirretrovirais ou adjuvantes a serem usados na clínica médica, de acordo com a necessidade do paciente (Braun e Wiesmann, 2007; Rubbert et al., 2007; Whitcomb et al., 2007).

Os inibidores de ligação pertencem a uma classe de moléculas que inibem a ligação inicial da gp120 com o receptor CD4. Os inibidores de fusão atuam na gp41, impedindo sua mudança conformacional final, necessária

para a fusão do envelope viral à membrana celular, e consequente entrada do HIV-1 na célula alvo (Whitcomb et al., 2007).

Os avanços nos estudos acerca dos correceptores propiciou o surgimento de novas estratégias de tratamento da infecção pelo HIV que podem agir diretamente sobre os correceptores por meio de ligantes aos correceptores (drogas de baixo peso molecular, anticorpos, derivados de quimiocinas) ou tratamentos de modulação do correceptor (modalidades de terapia gênica), tendo como principal alvo os sítios de ligação da gp120 (Alkhatib e Berger, 2007).

Os antagonistas de correceptores são uma nova classe de drogas, que bloqueiam a entrada do vírus na célula por interagirem com os receptores de quimiocinas CCR5 e CXCR4, usados como correceptores pelo HIV-1. Deste modo, atuam diretamente nos correceptores presentes nas células-alvo, propiciando a redução do aparecimento de mutações que conferem resistência aos antirretrovirais (Sander et al., 2007; Whitcomb et al., 2007). Por esta razão, a falência ao tratamento anti-HIV de uso corrente, ocasionada por cepas multirresistentes aos inibidores de protease e transcriptase reversa, poderá ser contornada com o uso de antagonistas de correceptores (Lorenzen et al., 2007).

Uma vez que o polimorfismo CCR5Δ32 não afeta o funcionamento normal do sistema imunológico, mas intefere significativamente na infecção e progressão da doença, o bloqueio da função do receptor CCR5 representa uma estratégia eficiente para a terapia antirretroviral (Emmelkamp e Rockstroh, 2007; Jones e Nelson, 2007; Lorenzen et al., 2007). Dentre as diferentes moléculas antagonistas de CCR5 em desenvolvimento, destacam-se três: aplaviroc (APL, AK-602, GlaxoSmithKline, Reino Unido), vicriviroc (VVC, SCH-D, Schering-Ploug, EUA) e maraviroque (MVC, UK-427,857, Pfizer, EUA). O desenvolvimento do aplaviroc foi encerrado, em 2005, nas fases II/III dos testes clínicos devido a ocorrência de hepatoxicidade em vários pacientes. Houve um atraso nos testes com o vicriviroc durante a fase II devido a ocorrência de linfomas malignos e adenomas, porém tais suspeitas não foram confirmadas e atualmente os testes clínicos encontram-

se na fase III (Emmelkamp e Rockstroh, 2007; Jones e Nelson, 2007; Kondru et al., 2008).

Em 2007, o maraviroque (*Celsentry*, Pfizer) foi a primeira molécula antagonista a ser aprovada para uso clínico. Este é o agente antagonista do CCR5 mais avançado, testado em várias doses, e apresentou efeitos adversos leves como hipotensão leve, cefaléia e náuseas e pode ser usado concomitantemente com outros antirretrovirais. Os ensaios clínicos comprovaram a eficácia do maraviroque, demonstrando uma queda na carga viral para níveis indetectáveis e aumento na contagem de linfócitos T CD4+ (Alkhatib e Berger, 2007; Bredeek e Harbour, 2007; Jones e Nelson, 2007).

Com relação aos antagonistas de CXCR4, apenas a droga AMD3100 apresentou uma redução significativa da carga viral em um pequeno número de pacientes com prevalência de cepas X4. Esta droga não será desenvolvida para uso clínico devido a sua toxicidade, porém, outros compostos desta mesma série estão sendo avaliados, como o AMD070 e AMD887, e ambos exibiram atividade frente a cepas multiresistentes a drogas *in vitro* (Jones e Nelson, 2007).

1.6 Determinação do tropismo

O uso clínico dos antagonistas de CCR5 implica na triagem dos candidatos a administração desta nova estratégia terapêutica. Visto que indivíduos com prevalência de cepas R5X4 ou X4 não têm indicação para o uso dos antagonistas de CCR5, a determinação do tropismo viral é essencial para a triagem de candidatos ao uso clínico desta classe de drogas (Sander et al., 2007; Sierra et al., 2007; Skrabal et al., 2007). Por esta razão, foram desenvolvidos ensaios moleculares e programas preditores capazes de inferir o tropismo das cepas de HIV-1. Estas metodologias ainda não são uma rotina e os ensaios padrão-ouro permanecem restritos a alguns laboratórios no exterior (Braun e Wiesmann, 2007; Whitcomb et al., 2007).

1.6.1 Ensaios fenotípicos

Os primeiros ensaios para a avaliação do fenótipo viral foram desenvolvidos na década de 80, com objetivo de classificar as cepas em indutoras de sincício (SI) ou não indutoras de sincício (NSI). A habilidade de replicação do HIV em linhagens celulares específicas, associada à fusão celular (sincício) indica a presença de vírus X4. Entretanto, estes ensaios requerem isolados virais variados e é um procedimento bastante trabalhoso (Braun e Wiesmann, 2007).

Os ensaios fenotípicos com vírus recombinantes (RVA – *recombinant viral assays*) são baseados em populações de pseudovírus marcados, capazes de recombinarem com as populações virais testadas, identificando o tropismo pelas populações celulares infectadas (Whitcomb et al., 2007). Atualmente existem quatro ensaios comerciais: *Trofile* (Monogram Biosciences, EUA), *Phenoscript* (VIRalliance, França), *Xtrack^{c}/PhenX-R* (inPheno AG, Suíça) e *Virco* (Virco BVBA, Bélgica). Estes ensaios geram pseudovírus que apresentam o gene do envelope, por completo ou determinados fragmentos, provenientes da população viral dos pacientes (Braun e Wiesmann, 2007).

O *Trofile* é um ensaio de vírus recombinantes em ciclo único de replicação e, atualmente, é o ensaio comercial mais comumente utilizado e o RVA melhor avaliado para a determinação do tropismo de HIV-1 em estudos clínicos (Braun e Wiesmann, 2007). Este ensaio usa uma região do gene *env* com tamanho aproximado de 2,5 Kb para a determinação do tropismo, que é amplificado por PCR (*Polymerase Chain Reaction* – reação em cadeia da polimerase) e inserido em um vetor de expressão do envelope. Neste ensaio, uma linhagem de células HEK293 (linhagem de células embrionárias renais) é utilizada para a transfecção do vetor de expressão *env* e o vetor de HIV genômico transportando o gene *reporter* luciferase. O processo de infecção é realizado em células U87 (linhagem de células de glioma humano) que expressam CD4/CXCR4 ou CD4/CCR5 na superfície celular (Figura 6). A quantificação da emissão de luz pela expressão do gene

reporter de luciferase é controlada pela presença de antagonistas de correceptor (Braun e Wiesmann, 2007; Whitcomb et al., 2007).

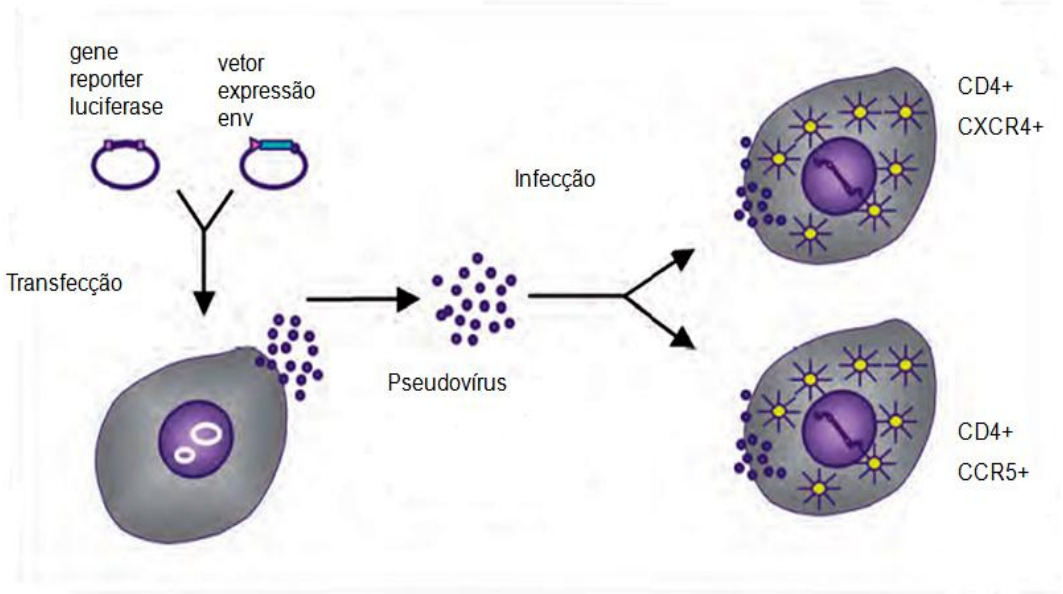

Figura 6 – Ilustração do método utilizado no ensaio *Trofile.* Células HEK-293 são cotransfectadas com um vetor de expressão contendo parte do gene *env* proveniente de uma amostra de plasma do paciente e um vetor de HIV genômico contendo um gene *reporter* luciferase na região do *env*. São produzidas partículas de pseudovírus usadas para infectar células U87 expressando os receptores CD4/CCR5 ou CD4/CXCR4. O tropismo é detectado pela atividade da luciferase das células infectadas. Adaptado de Braun e Wiesmann (2007).

O ensaio *Trofile* é bastante preciso e apresenta uma adequada reprodutibilidade. Entretanto, este ensaio apresenta algumas limitações como os altos custos, a demora para a confirmação do resultado, a difícil disponibilidade (as amostras precisam ser enviadas para o exterior), os critérios para a realização do ensaio e o limitado acesso ao países em desenvolvimento (Braun e Wiesmann, 2007; Genebat et al., 2009). De um modo geral, por serem baseados em experimentos de cultura celular, os ensaios fenotípicos implicam em maior custo, maior tempo para a padronização e realização dos ensaios quando comparados aos testes genotípicos (Braun e Wiesmann, 2007; Sander et al., 2007; Sierra et al., 2007).

1.6.2 Ensaios genotípicos e programas preditores

Muitos dos determinantes gênicos para a utilização do correceptor estão presentes no envelope do HIV-1, em especial na alça V3, influenciando na especificidade do correceptor utilizado pelas variantes virais, onde poucas trocas de aminoácidos são suficientes para a alteração do tropismo viral. Consequentemente, esta região é um importante alvo para avaliações baseadas em bioinformática capazes de inferir o fenótipo do tropismo a partir de predições genotípicas utilizando dados de sequenciamento (Clevestig et al., 2006; De Jong et al., 1992; Skrabal et al., 2007).

Foram publicados diferentes protocolos baseados em sequências de aminoácidos da região V3. Um teste simples e bastante popular é a regra 11/25: se aminoácidos básicos (arginina ou lisina) encontram-se nas posições 11 ou 25 da região V3 o vírus é X4, logo, se não há aminoácidos básicos nesta posição, considera-se R5. Este teste é preciso para a predição de isolados R5, mas não é considerado muito confiável para isolados X4, podendo conduzir a conclusões incompletas ou ambíguas (Jensen et al., 2003; Sierra et al., 2007). Para contornar esta situação, foram publicados diferentes protocolos para a predição do tropismo baseados em sequências de aminoácidos da região V3 utilizando ferramentas de bioinformática capazes de ampliar os parâmetros de análise, garantindo uma predição mais eficiente. Apenas alguns destes métodos de predição estão disponíveis em servidores livres na *internet*: *WetCat*, *WebPSSM* e *geno2pheno*$_{[coreceptor]}$ (Tabela 1). Estas ferramentas utilizam sequências de nucleotídeos ou aminoácidos da região V3, que são enviadas via *internet* e realizam a predição em tempo real através de matrizes, algoritmos e bases de dados. É possível inserir dados adicionais como carga viral e contagem de linfócitos T que melhoram a especificidade do teste (Sierra et al., 2007).

Tabela 1 - Testes de predição de tropismo oferecidos pelas plataformas livres

Plataforma	Testes	Descrição dos testes
WetCat[1]	*C4.5*	Árvore de decisão
	C4.5 only p8 and p12	Árvore de decisão
	PART	Regra a partir de árvore de decisão
	SVM	Matriz estatística
	Charge Rule	Regra de cargas 11/25
WebPSSM[2]	*x4r5*	Matriz de escores a partir de dados genotípicos
	sinsi	Matriz de escores a partir de dados fenotípicos
Geno2pheno *[coreceptor]*[3]	*Clonal data*	SVM a partir das sequências fornecidas
	Clinical data	SVM a partir das sequências e dados clínicos fornecidos

[1] http://genomiac2.ucsd.edu:8080/wetcat/v3.html
[2] http://indra.mullins.microbiol.washington.edu/webpssm/
[3] http://coreceptor.bioinf.mpi-inf.mpg.de/index.php

O *WetCat* é um serviço desenvolvido e mantido pela Universidade da Califórnia (São Francisco, EUA) que consiste em uma regra de cargas implementada por três diferentes árvores de decisão e uma ferramenta de suporte vetorial (SVM – *support vector machine*). As sequências de nucleotídeos precisam ser traduzidas para aminoácidos e alinhadas manualmente com a sequência consenso disponibilizada pelo *website*, um processo trabalhoso e que consome bastante tempo. A vantagem é a possibilidade de inserir várias sequências e obter o resultado da predição do tropismo em uma mesma corrida (Sierra et al., 2007).

O servidor *WebPSSM* é mantido pela Universidade de Washington (Washington, EUA) e prediz o correceptor a partir de uma sequência de aminoácidos desenvolvida através de uma matriz de escore de posição específica (PSSM – *position specific score matrice*). O alinhamento das amostras com a sequência consenso é automático e pode ser realizado com sequências do subtipo B ou C. O resultado, além de indicar o tropismo, traz

um valor quantitativo, o escore de predição e o intervalo de confiança, entre outros dados (Sierra et al., 2007).

A predição pelo *geno2pheno*$_{[coreceptor]}$, mantido pelo Instituto de Informática Max Planck (Saarbrücken, Alemanha) é realizada através de um SVM implementado com outros algoritmos preditivos podendo inserir sequências de nucleotídeos ou aminoácidos que contenham a região V3. É possível ainda, inserir parâmetros clínicos obtendo-se dois resultados: a análise comparando a sequência de aminoácidos ou nucleotídeos com um padrão de dados obtidos por clones e comparando com dados obtidos de marcadores clínicos de 1000 pacientes que nunca fizeram uso de terapia antirretroviral (*naïves*) (Sierra et al., 2007).

A utilização de ferramentas de bioinformática para a predição do tropismo quando comparado com a aplicação de ensaios com vírus recombinantes, torna-se uma alternativa mais simples e financeiramente mais viável para a triagem dos candidatos ao uso de antagonistas de CCR5. A inclusão desta estratégia mais específica, capaz de selecionar os candidatos ao uso dos antagonistas de CCR5, poderia atuar como adjuvante na prática clínica, uma vez que a inclusão desta nova classe de drogas na terapia antirretroviral pode desacelerar a progressão da infecção, beneficiando, desta forma, o prognóstico e a qualidade de vida do paciente infectado.

2 OBJETIVOS

2.1 Objetivo geral

Identificar o tropismo do HIV-1 pelos correceptores CCR5 e CXCR4 a partir das sequências da região V3 da gp120 do envelope do HIV-1 pelo uso de bioinformática.

2.2 Objetivos específicos

- Utilizar instrumentos de bioinformática para predizer o tropismo viral aos correceptores;
- Determinar a prevalência do tropismo do HIV-1 aos correceptores CCR5 e CXCR4;
- Avaliar a aplicabilidade das ferramentas de bioinformática para predição do tropismo viral.

3 MÉTODOS

3.1 Casuística

Foram convidados a participar deste estudo 101 pacientes infectados pelo HIV-1 e em acompanhamento clínico no Ambulatório de Imunodeficiências Secundárias ADEE 3002 – Departamento de Dermatologia – Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP).

3.1.1 Critérios de inclusão

Foram incluídos no estudo somente indivíduos adultos, homens ou mulheres, com sorologia positiva para HIV-1 e idade superior a 18 anos.

3.1.2 Critérios de exclusão

Foram excluídos do estudo pacientes com dados clínicos e laboratoriais incompletos.

3.1.3 Aspectos éticos

Os indivíduos que aceitaram participar deste estudo assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (Apêndice A). Todos os dados individuais são confidenciais e em nenhum momento serão divulgados. O processo foi aprovado pela Comissão de Ética em Pesquisa do HCFMUSP, CAPPesq nº **0108/08** (Anexo A).

3.2 Coleta das amostras

Foram coletadas amostras de sangue periférico através de punção venosa em dois tubos (5 mL) contendo EDTA dos indivíduos que aceitaram participar do estudo. Após a coleta, as amostras foram centrifugadas a 450 xg, por 5 min, à temperatura ambiente. O plasma foi separado em duas alíquotas de 500 µL e estocado a -70°C para possível uso.

3.3 Separação de células mononucleares do sangue periférico

As amostras de sangue total foram diluídas em solução salina na proporção 1:1 e transferidas para tubos cônicos de 50 mL contendo Ficoll-Paque Plus (GE Healthcare, Suécia) na proporção 1:3 e posteriormente foram centrifugadas a 860 xg, por 20 min, a 16°C.

As células mononucleares do sangue periférico (PBMCs – *Peripheral Blood Mononuclear Cells*) separadas pelo gradiente de densidade foram removidas e transferidas para um tubo cônico de 15 mL e lavadas com solução salina sob centrifugação de 450 xg, por 10 min, a 16°C. Em seguida o sobrenadante foi removido e foi adicionado 1 mL de solução de lise de hemácias (*GFX Genomic Blood DNA Purification Kit*, GE Healthcare, Reino Unido) seguido por uma incubação de 5 min, à temperatura ambiente. Uma segunda lavagem com solução salina foi realizada nas mesmas condições da lavagem anterior.

O sobrenadante foi removido e o *pellet* (precipitado) de células foi ressuspendido em 1 mL de solução salina para contagem de leucócitos no contador hematológico *Cell-Dyn 1400* (Abbott, EUA). Foram separadas duas alíquotas contendo 2.10^6 células/mL que foram centrifugadas a 6000 xg, por 3 min, à temperatura ambiente. Após a remoção do sobrenadante uma alíquota foi estocada a -70°C e outra armazenada a -20°C até a sua utilização.

3.4 Extração do DNA pró-viral

O DNA pró-viral foi extraído a partir das PBMCs das 101 amostras, utilizando o *GFX Genomic Blood DNA Purification Kit* (GE Healthcare, Reino Unido), seguindo as orientações fornecidas pelo fabricante e estocado à temperatura de -20°C até sua utilização.

3.5 Amplificação da região V3 do gene *env*

O DNA extraído foi submetido a amplificação por PCR *nested* (aninhada), utilizando os *primers* (iniciadores) *outer* (externos) senso ED5 e anti-senso ED12 e os *primers inner* (internos) senso ED31 e antissenso ED33 (Delwart et al., 1995) (Tabela 2).

Tabela 2 – Sequências dos *primers* utilizados para amplificar e sequenciar a região V3 do gene *env*

Nome	**Sequência (5' – 3')**	**Posição[1]**
ED5 (*outer*)	ATGGGATCAAAGCCTAAAGCCATGTG	6556 – 6581
ED12 (*outer*)	AGTGCTTCCTGCTGCTCCCAAGAACCCAAG	7822 – 7792
ED31 (*inner*)	CCTCAGCCATTACACAGGCCTGTCCAAAG	6816 – 6844
ED33 (*inner*)	TTACAGTAGAAAAATTCCCCTC	7359 – 7380

Nota: [1] Posição sobre a sequência referência HXB2, acesso no GenBank nº K03455

As amplificações foram realizadas em termociclador convencional, seguindo as etapas para a PCR *outer:* uma denaturação inicial de 94°C durante 1 min; seguida de 35 ciclos de 94°C por 45 s, 55°C por 45 s e 72°C por 2 min; e uma extensão final a 72°C por 10 min. Para a PCR *inner* foram seguidas as seguintes condições: uma denaturação de 94°C durante 1 min; seguida de 35 ciclos de 94°C por 1 min, 55°C por 1 min e 72°C por 1:30 min; com uma extensão final de 72°C por 10 min. As condições para uma reação final de PCR de 50 μL foram de 1x tampão de PCR (10x), 2 mM de $MgCl_2$, 2 μM de cada primer, 0,2 mM de cada dNTPs, 1 U de Taq DNA polimerase e 5

μL de DNA extraído (20 – 50 ng/μL) para a PCR *outer* e 5 μL (20 – 100 ng/μL) de produto de PCR para a PCR *inner*.

3.6 Detecção dos produtos de PCR

Os produtos de PCR foram submetidos à eletroforese em gel de agarose 1,5% a 100 V por 60 min. A visualização foi realizada com a adição de tampão carregador azul de bromofenol 10x à amostra, na proporção 1:4 e sob incidência de luz UV.

A quantificação e a determinação do tamanho dos fragmentos foi realizada utilizando o padrão de peso molecular de 100 pb como parâmetro de concentração. Foram considerados adequados os fragmentos com aproximadamente 500 pb e concentração entre 20 e 100 ng/μL.

3.7 Purificação dos produtos amplificados

Os produtos das PCR *inner* foram purificados com o *QIAquick Purification Kit* (Qiagen, Alemanha), conforme o protocolo fornecido pelo fabricante.

3.8 Reação de sequenciamento

Para a reação de sequenciamento foi utilizado o *kit ABI PRISM Big Dye Terminator v. 3.1 Cycle Sequencing Ready Reaction* (PE Applied Biosystems, EUA), adicionando 20 – 100 ng/μL do produto amplificado purificado e 1μM dos respectivos *primers inner*. As amplificações foram realizadas em termociclador convencional seguindo 25 ciclos de 96°C por 10 s, 55°C por 5 s e 60°C por 4 min. As amostras foram precipitadas com isopropanol 80% e ao *pellet* seco foram adicionados 10 μL de formamida. A corrida foi executada em sequenciador automático *ABI Prism 3100* (PE Applied Biosystems, EUA).

3.9 Análise e interpretação do sequenciamento

As sequências senso e antissenso obtidas foram sobrepostas, avaliadas e editadas através do programa *SeqScape* v. 2.5.0 (Applied Biosystems, EUA). As sequências consenso obtidas foram analisadas utilizando o programa *BioEdit* v.7.0.8 (Ibis Biosciences, EUA) e alinhadas por *ClustalW* com a sequência de DNA da alça V3 proveniente da cepa HXB2 obtida no banco de dados do Laboratório Nacional de Los Alamos (número de acesso no *GenBank*: K03455). A região V3 foi isolada e os nucleotídeos foram traduzidos para aminoácidos.

3.9.1 Determinação dos subtipos

Os subtipos das sequências foram determinados utilizando o programa *jpHMM* (*jumping profile Hidden Markov Model*) mantido pela Universidade de Göttingen, Alemanha (Disponível em: http://jphmm.gobics.de/, acesso em: 04 out. 2009 e 16 fev. 2010). Este programa é uma aproximação probabilística que compara uma sequência com alinhamentos múltiplos de uma família de sequências (Schultz et al., 2009).

3.9.2 Determinação da carga elétrica líquida

A carga elétrica líquida das sequências de aminoácidos foi outro parâmetro considerado como indicativo para o tropismo, uma vez que cepas com carga líquida ≤ 4 na região V3 tendem a apresentar tropismo pelos correceptores CCR5 (Xu et al., 2007). A carga líquida de uma proteína é a função da soma de suas cargas positivas e negativas. Os resíduos com caráter básico arginina (R), histidina (H) e lisina (K) apresentam a carga elétrica positiva (+1), enquanto os resíduos de caráter ácido aspartato (D) e glutamato (E) apresentam carga negativa (-1), sendo os demais resíduos, considerados neutros, ou com carga elétrica nula.

3.10 Predição do tropismo

As sequências de 35 aminoácidos da região V3 foram editadas de acordo com as exigências das plataformas de predição e submetidas a oito testes disponíveis (Tabela 3). O tropismo foi denominado R5 ou X4 quando houve concordância de resultados entre os oito testes e quando houve discordância na predição em pelo menos um dos testes, o resultado foi denominado inconclusivo.

Tabela 3 - Testes de predição de tropismo utilizados neste estudo

Plataforma	**Testes**
WetCat	*C4.5*
	C4.5 only p8 and p12
	PART
	SVM
	Charge Rule
WebPSSM	*x4r5*
	sinsi
Geno2pheno $_{[coreceptor]}$	*Clonal data*

3.11 Avaliação da predição do tropismo

Adicionalmente, com a finalidade de avaliar a sensibilidade e especificidade, bem como a aplicabilidade das ferramentas preditivas, foram analisadas sequências da região V3 obtidas a partir do banco de dados do Laboratório Nacional de Los Alamos (Disponível em: http://www.hiv.lanl.gov, acesso em 21 fev. 2010). O banco de dados informa o tropismo empírico, determinado através de ensaios fenotípicos, portanto, na ausência de um ensaio padrão-ouro, estas sequências com tropismo conhecido foram imprescindíveis para a avaliação dos sistemas preditivos.

Este banco de dados contém 3231 sequências V3 com tropismo para CCR5, 457 com tropismo para X4 e 555 duplo-trópicas. Destas 4243

sequências foram selecionadas 3122 sequências dos subtipos B, C e F1, uma vez que são os principais subtipos encontrados no Brasil. Destas, foram excluídas aquelas contendo 34 ou 36 aminoácidos na alça V3, restando apenas as sequências contendo 35 aminoácidos.

Devido ao grande número de sequencias disponíveis, nos conjuntos que continham mais de 100 sequencias, foi selecionada uma representatividade de aproximadamente 5% para serem incluídas no estudo, perfazendo um total de 185 sequências com tropismo conhecido que foram submetidas aos testes preditivos de tropismo (exceto geno2pheno$_{[coreceptor]}$ – *clinical data*, por não haver dados clínicos disponíveis destas sequências). A Tabela 4 sumariza o número de sequências disponíveis e as que foram incluídas no estudo.

Tabela 4 – Número de sequências com tropismo conhecido obtidas no banco de dados de Los Alamos

	Subtipo			
Tropismo	**B**	**C**	**F1**	**Total**
R5	115 (2214)	23 (349)	2 (2)	**140**
X4	13 (186)	5 (19)	1 (1)	**19**
R5X4	18 (325)	7 (25)	1 (1)	**26**
Total	**146**	**35**	**4**	**185**

Nota: os números entre parênteses indicam o total de sequências disponíveis no banco de dados.

3.12 Análise estatística

Os dados demográficos e laboratoriais (contagem de linfócitos T CD4+ e T CD8 + e carga viral) dos pacientes sob terapia antirretroviral e daqueles que não faziam uso de medicamentos foram analisados através de estatística descritiva com intervalo de confiança de 95% para a determinação das médias e desvios–padrões, enquanto a diferença estatística foi determinada pelo teste de Mann-Whitney, considerando significativo $p<0,05$.

Para comparação de grupos de dados foi utilizado o teste exato de Fisher, considerando $p<0,05$.

Os sistemas preditivos foram avaliados através do cálculo de sensibilidade, especificidade, valores preditivos e curva ROC (*Receiver Operating Characteristic*).

As médias e desvios-padrões das cargas elétricas líquidas das sequências obtidas neste estudo e daquelas obtidas em Los Alamos foram calculadas por estatística preditiva com intervalo de confiança de 95%.

As análises estatísticas foram realizadas utilizando os programas *GraphPad Prism v. 5.00* (GraphPad Software, EUA) e *PASW Statistics 17* (SPSS Inc., EUA).

4 RESULTADOS

Foram incluídos no estudo 101 indivíduos infectados pelo HIV, em acompanhamento clínico, sendo possível a amplificação por PCR da região V3 do gene *env* de 100 pacientes, dos quais o sequenciamento foi possível em 94 amostras (Figura 7).

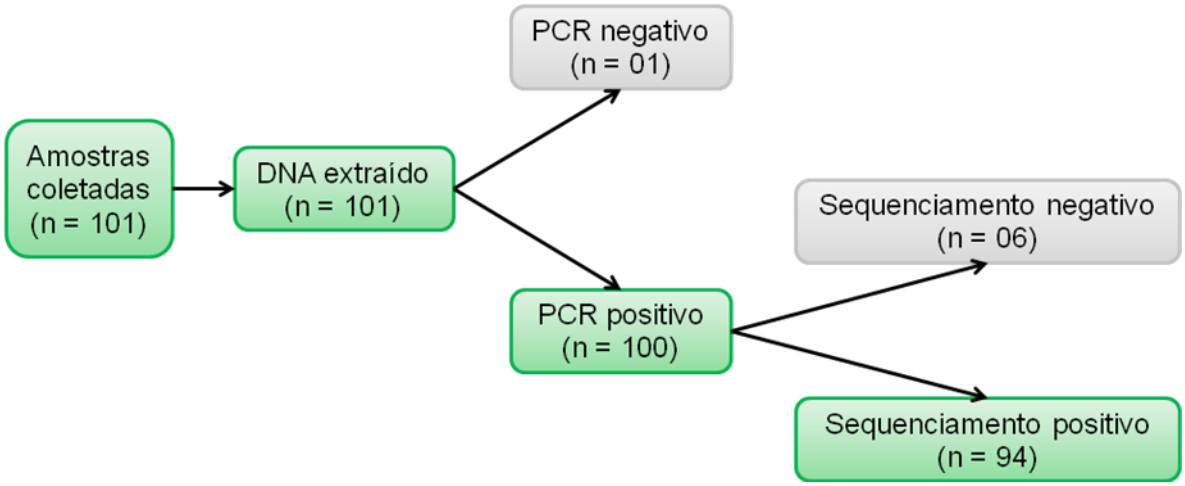

Figura 7 – Fluxograma das amostras coletadas e sequenciadas. Foram coletadas 101 amostras, porém, apenas 94 apresentaram sequências legíveis da região V3-gp120 após o sequenciamento.

4.1 Dados demográficos e laboratoriais

Os dados demográficos e laboratoriais dos 94 pacientes encontram-se relacionados na Tabela 5. Foram incluídos no estudo 67 homens e 27 mulheres, com média de idade de 40 (± 10) anos. Com relação à via de transmissão, em 70,2% dos casos foi por contato sexual (homossexual e heterossexual). A maior parte dos pacientes está fazendo uso de antirretrovirais (Sob HAART), enquanto 27,5% permanecem sem indicação para medicação (*naïves*).

Tabela 5 – Dados demográficos e laboratoriais

		Frequência (%)
Média de idade		39,8
Desvio-padrão		± 9,9
Sexo	Homens	67 (71,3)
	Mulheres	27 (28,7)
Via de transmissão	Heterossexual	38 (40,4)
	Homossexual	32 (34)
	Contato sanguíneo	5 (5,3)
	Usuários de drogas injetáveis	4 (4,3)
	Desconhecida	15 (16)
Tratamento (uso de HAART[1])	Sim	68 (72,3)
	Não	26 (27,7)
Dados laboratoriais (Média, desvio-padrão)	T CD4+ (cel/mm^3)	495 (± 280)
	Média T CD8+ (cel/mm^3)	1029 (± 505)
	Média carga viral (cp/mL)	9992 (± 33514)

[1] *Highly Active Antiretroviral Therapy* (terapia antirretroviral altamente ativa)

Com relação aos resultados de exames laboratoriais, a média da contagem de linfócitos T CD4+ em pacientes *naïves* e sob HAART foi de 476 (± 288) e 502 (± 280) respectivamente, e a contagem de linfócitos T CD8+ foi de 1056 (± 494) e 1019 (± 512) respectivamente, não havendo diferença estatística entre os dois grupos. A média da carga viral plasmática entre os pacientes *naïve* foi de 25741 e entre os pacientes sob terapia foi de 4027, com p>0,0001. Um total de 50 pacientes (53%) apresentou carga viral plasmática em níveis indetectáveis.

4.2 Prevalência dos subtipos de HIV-1

O subtipo B predomina nesta população, representando cerca de 83% da população (Figura 8). A variante brasileira B', que apresenta o motivo GWGR na coroa da alça V3, foi encontrada em 22,3% das sequências de subtipo B.

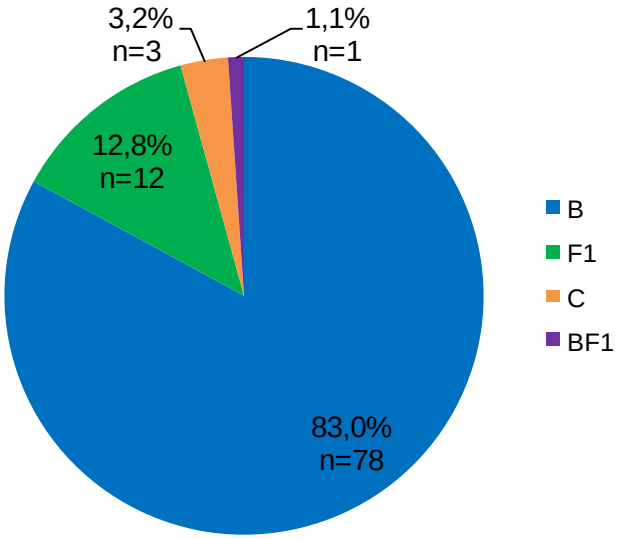

Figura 8 – Distribuição dos subtipos de HIV-1 encontrados na população estudada

4.3 Avaliação da carga elétrica líquida

A avaliação da carga elétrica das sequências do banco de dados revelou que as sequências R5 apresentaram carga líquida variando entre 1 e 8, para R5X4 houve variação entre 4 e 8 e as sequências X4 apresentaram carga líquida entre 4 e 10. A carga elétrica líquida média foi de 4,8 (± 1), 6,3 (± 0,88) e 7,1 (± 1,41), com medianas de 5,0, 6,0 e 7,0 para sequências R5, R5X4 e X4 respectivamente (Figura 9).

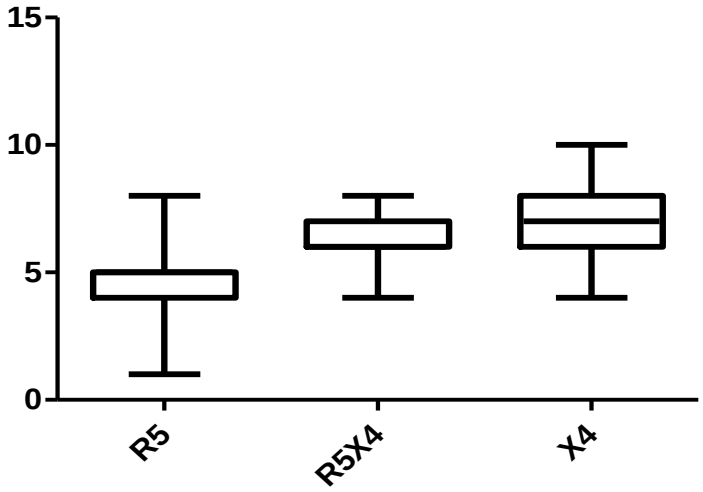

Figura 9 – Carga líquida das sequências obtidas no banco de dados de Los Alamos

4.4 Predição do tropismo

As 185 sequências obtidas no banco de dados de Los Alamos e as 94 sequências obtidas neste estudo foram submetidas aos oito testes preditivos alocados nas três plataformas de bioinformática, previamente descritas. O teste *geno2pheno*$_{[coreceptor]}$ – *clinical data* não foi avaliado neste estudo devido à ausência de dados clínicos das sequências do banco de dados.

As plataformas de predição não são capazes de diferenciar cepas R5X4 de cepas X4, considerando ambas como X4 e quando houve discordância na predição em pelo menos um dos oito testes o resultado foi denominado inconclusivo até a realização de avaliações mais específicas.

4.4.1 Avaliação da predição do tropismo

Devido à presença de casos inconclusivos e da ausência de um ensaio padrão-ouro (*Trofile*) para a confirmação dos resultados, foi necessária uma melhor compreensão dos sistemas preditivos para avaliar sua aplicabilidade na prática clínica.

Assim, 185 sequências da região V3, com tropismo determinado empiricamente, foram selecionadas a partir do banco de dados do Laboratório Nacional de Los Alamos e submetidas aos oito testes preditivos. Em 77,8% (n=144) dos casos houve concordância entre a predição e o tropismo empírico, quando os oito testes apresentaram o mesmo resultado, havendo erro na predição de 3,2% (n=6) dos casos, enquanto em 19% (n=35) dos casos a predição foi considerada inconclusiva (Tabela 6).

Tabela 6 – Número de sequências obtidas de Los Alamos com tropismo conhecido e sua predição pelas ferramentas de bioinformática

Predição	R5 (n=140)	R5X4 (n=26)	X4 (n=19)
R5	119	3	2
X4	1	13	12
Inconclusivo	20	10	5

Os testes foram avaliados individualmente e demonstraram sensibilidade entre 87% para o teste *geno2pheno*$_{[coreceptor]}$ e 56% para *C4.5 p8p12*, enquanto a especificidade variou entre 99% no teste *x4r5* e 90% para *geno2pheno*$_{[coreceptor]}$ (Tabela 7).

Tabela 7 – Sensibilidade e especificidade dos testes de bioinformática para a predição de cepas X4 em sequências do banco de dados

	Sensibilidade (IC=95%)	**Especificidade (IC=95%)**	**VPP**	**VPN**
geno2pheno$_{[coreceptor]}$	86,7 (0,73-0,95)	89,3 (0,83-0,94)	0,72	0,95
sinsi	77,8 (0,63-0,88)	97,1 (0,93-0,99)	0,89	0,93
x4r5	75,6 (0,60-0,87)	98,6 (0,95-0,99)	0,94	0,93
SVM	75,6 (0,60-0,87)	95 (0,89-0,98)	0,83	0,92
Charge rule	75,6 (0,60-0,87)	92,9 (0,87-0,97)	0,77	0,92
PART	68,9 (0,53-0,82)	92,7 (0,87-0,97)	0,76	0,9
C4.5	57,8 (0,42-0,72)	94,3 (0,89-0,98)	0,76	0,87
C4.5 p8p12	55,6 (0,4-0,7)	94,3 (0,89-0,98)	0,76	0,87

Nota: IC: intervalo de confiança; VPP: valor preditivo positivo; VPN: valor preditivo negativo

A área abaixo da curva ROC foi calculada para avaliar a variação da sensibilidade e especificidade para os diferentes testes, como uma tentativa de inferir qual o sistema preditivo mais adequado para a determinação do tropismo do HIV. Foram realizadas uma análise individual e uma agrupando os testes que apresentaram resultados de sensibilidade e especificidade mais satisfatórios. A Tabela 8 sumariza a área abaixo da curva ROC e a Figura 10 indica a curva com a melhor área encontrada neste estudo.

Tabela 8 – Área abaixo da curva ROC dos sistemas preditivos

Teste preditivo	Área abaixo da curva
geno2pheno$_{[coreceptor]}$, *sinsi*, *x4r5* e *SVM*	0,923
Todos	0,914
geno2pheno$_{[coreceptor]}$	0,880
sinsi	0,875
x4r5	0,871
SVM	0,853
Charge rule	0,842
PART	0,809
C4.5	0,760
C4.5 p8p12	0,749

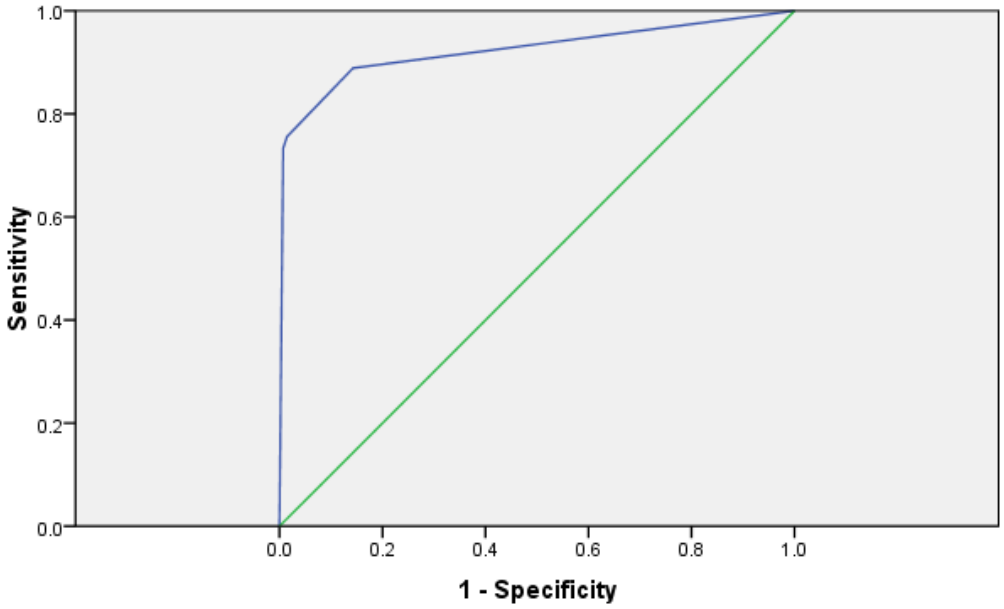

Figura 10 – Curva ROC do conjunto de testes geno2pheno$_{[coreceptor]}$, sinsi, x4r5 e SVM. Área abaixo da curva = 0,923

4.4.2 Elaboração de um algoritmo para a predição do tropismo

Diante dos resultados obtidos com a avaliação das ferramentas preditivas e da presença de resultados inconclusivos, foi necessária a

elaboração de um algoritmo capaz de melhorar a eficiência da predição, bem como, seu papel na prática clínica.

A avaliação da área da curva ROC indicou que o uso combinado dos testes preditivos: *geno2pheno*$_{[coreceptor]}$, *sinsi*, *x4r5* e *SVM* apresentou um índice de precisão maior, quando comparado com os índices individuais. Desta forma, foi possível inferir que o uso em conjunto destes quatro testes seria a melhor estratégia para a predição do tropismo. Entretanto, estes quatro testes apresentaram resultados discordantes em 14% (n=26) dos casos. A permanência de casos inconclusivos levou à avaliação de outros parâmetros, como o número de acertos na predição, o tamanho do banco de dados no qual se baseia cada teste e a carga elétrica da sequência de aminoácidos.

Uma matriz de escores foi construída atribuindo valores arbitrários e ordenados considerando especificidade, sensibilidade, número de acertos, área abaixo da curva ROC e tamanho do banco de dados de cada teste preditivo. Os escores resultaram da somatória dos valores atribuídos a cada teste preditivo, dos quais, foi possível inferir a seguinte ordem de confiabilidade no resultado da predição: *sinsi* > *x4r5* > geno2pheno$_{[coreceptor]}$ > *SVM*.

Assim, o algoritmo sugerido foi utilizar os resultados provenientes do uso em conjunto dos quatro testes e para os casos inconclusivos, considerar o resultado de sinsi. Seguindo esta regra houve erro de predição em 7,6% (n=14) dos casos.

Com o objetivo de definir uma estratégia com eficiência mínima de 95%, a carga elétrica líquida foi utilizada como discriminador nos casos inconclusivos. Desta forma, foi possível elaborar uma estratégia com eficiência de 94,5%, cujo algoritmo está ilustrado na Figura 11.

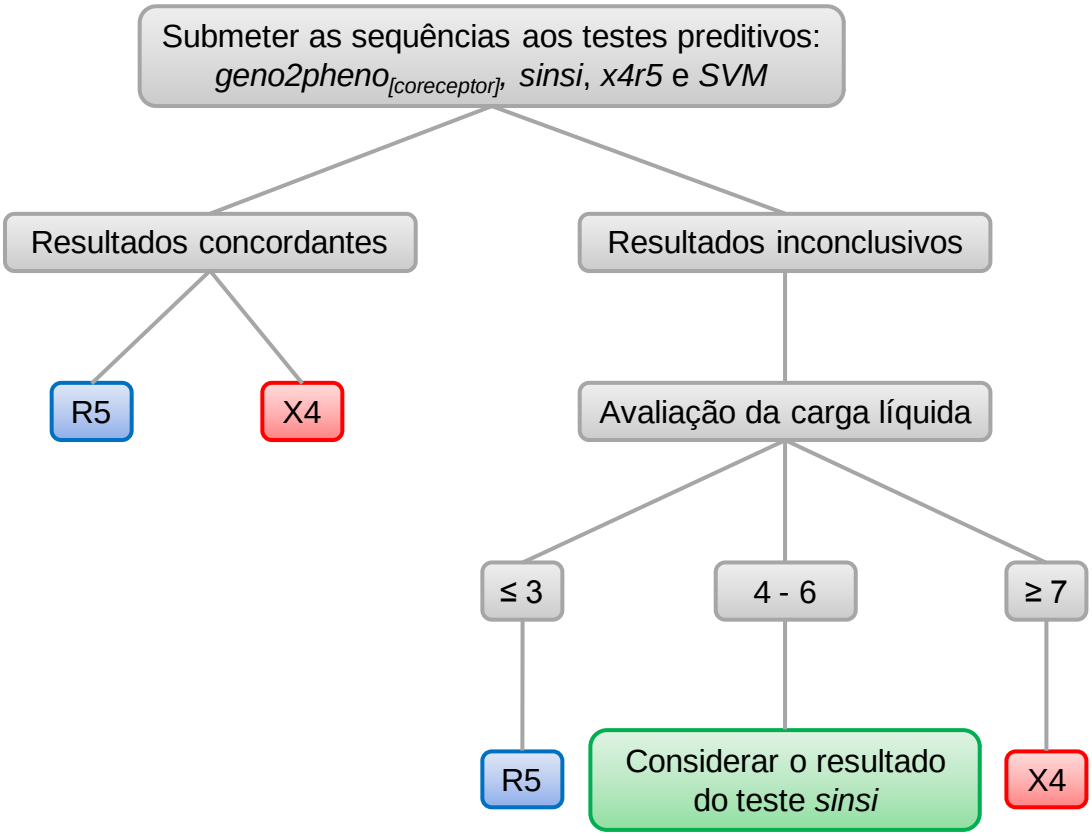

Figura 11 – Algoritmo para determinação do tropismo utilizando sistemas preditivos.

4.4.3 Predição do tropismo das amostras obtidas neste estudo

As 94 sequências obtidas neste estudo foram submetidas aos oito testes preditivos, das quais, 38,3% (n=36) sequências foram preditas como cepas R5 e 5,3% (n=5) foram preditas X4, somando 43,6% de concordância entre os nove testes, enquanto 56,4% (n=53) dos casos foram considerados inconclusivos.

Quando submetidas ao algoritmo de predição, a concordância de resultados entre os testes *geno2pheno*$_{[coreceptor]}$, *sinsi*, *x4r5* e *SVM*, demonstrou a predição de 38 sequências R5, oito X4 e a presença de 48 casos inconclusivos. Dos casos inconclusivos, nove puderam ser discriminados utilizando a carga líquida. O resultado final indicou a presença de 79,8% (n=75) de cepas com tropismo pelo correceptor CCR5 e 20,2% (n=19) de cepas capazes de utilizar o correceptor CXCR4 para a entrada na célula.

A avaliação dos dados laboratoriais entre os portadores de cepas R5 e cepas X4 não demonstrou diferença significativa entre as médias de contagem de linfócitos T CD4+ e T CD8+ e carga viral plasmática. Quando os grupos são divididos com relação ao tratamento, observou-se uma diferença significativa nos portadores de cepas R5, os pacientes naïves apresentaram carga viral significativamente mais alta que os pacientes e sob HAART (*p*<0,0001). Este mesmo grupo apresentou carga viral significativamente mais alta em relação aos portadores *naïve* de cepas X4 (Tabela 9).

Tabela 9 – Relação entre os dados laboratoriais e o tropismo inferido

		R5 (n=75)	**X4 (n=19)**	***p***
		n=21	**n=5**	
Naïves	Média T CD4+ (células/mm^3)	518 (± 286)	308 (± 252)	NS
	Média T CD8+ (células/mm^3)	1089 (± 463)	926 (± 646)	NS
	Média carga viral (cópias/mL)	30910 (± 65620)	2637 (± 5033)	0,011
		n=54	**n=14**	
Sob HAART	Média T CD4+ (células /mm^3)	511 (± 289)	468 (± 245)	NS
	Média T CD8+ (células /mm^3)	1046 (± 520)	919 (± 483)	NS
	Média carga viral (cópias/mL)	3876 (± 10697)	4586 (± 12569)	NS

Nota: NS: não significativo; valor de p >0,05

Com referência aos subtipos, o grupo de cepas R5 apresentou todos os subtipos encontrados na coorte estuda (n=60 B; 11 F1; 3 C; 1 BF1), enquanto no grupo de cepas X4, houve apenas um representante do subtipo F1 e os demais foram do subtipo B (n=18), dos quais três foram

caracterizados como a variante B'. A presença da variante B' não apresentou diferença significativa entre grupos R5 e X4.

A análise da carga elétrica líquida demonstrou média de 5,1 (± 0,96) e 7 (± 1,2) e medianas de 5,0 e 7,0 para os grupos R5 e X4 respectivamente (Figura 12).

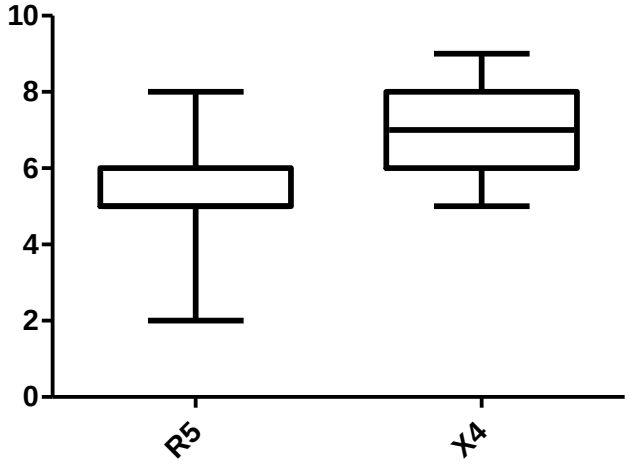

Figura 12 – Carga elétrica líquida das amostras obtidas neste estudo

Em síntese, a avaliação dos sistemas preditivos de bioinformática utilizados neste estudo e a avaliação adicional da carga elétrica líquida propiciaram a elaboração de uma estratégia de análise capaz de aumentar a confiabilidade da predição, possibilitando o seu uso na prática clínica.

5 DISCUSSÃO

Uma das principais características do HIV-1 é sua alta taxa de variabilidade genética, demonstrada pelo grande número de variantes que circulam em um mesmo indivíduo. Esta marcante diversidade genética é resultante não só de fatores virais, como as altas taxas de mutação e replicação e a atividade recombinogênica, mas também à pressão exercida pelo sistema imune do hospedeiro. Uma vez que esta heterogeneidade atua como uma eficiente estratégia de escape às defesas do hospedeiro, outros fatores externos são necessários para combater a infecção (Peeters, 2000; Simon-Loriere et al., 2009).

Neste contexto, a terapia antirretroviral representa um importante papel para aumentar a sobrevida e diminuir a morbidade dos pacientes. O intuito da introdução de medicamentos antirretrovirais é diminuir o número de partículas virais circulantes, capazes de infectar novas células e manter o ciclo replicativo, reduzindo a viremia a níveis indetectáveis. Entretanto, nenhum tratamento é capaz de erradicar completamente o vírus, a pressão da droga apenas suprime a replicação viral e a aderência inadequada do paciente ao tratamento pode levar à emergência de vírus resistentes, ocasionando a falha terapêutica. Logo, os indivíduos infectados com cepas multirresistentes, têm opções de tratamento cada vez mais limitadas (Jones e Nelson, 2007; Lieberman-Blum et al., 2008).

Novas moléculas vêm sendo desenvolvidas com o objetivo de se tornarem novas drogas para suprir a necessidade de antirretrovirais eficientes em pacientes com falha terapêutica, e por sua vez, novos pontos do ciclo replicativo do HIV estão surgindo como alvo destas novas estratégias terapêuticas. Um exemplo é o maraviroque, um antagonista de correceptor CCR5 aprovado pra uso clínico em 2007, mas que ainda não foi incorporado ao consenso brasileiro de terapia antirretroviral devido à necessidade de um ensaio prévio para determinar o tropismo viral em candidatos ao uso desse medicamento (Jones e Nelson, 2007).

Existem diferentes protocolos de ensaios capazes de determinar o tropismo viral. Os ensaios fenotípicos evoluíram de culturas de linhagens de células MT2 onde se observa a formação ou não de sincício na presença do HIV e, atualmente, envolvem linhagens celulares que originalmente não são permissivas ao HIV, mas são modificadas para expressar CD4 e receptores de quimiocinas em sua superfície. Linhagens celulares derivadas de um glioma humano (U87) e de um osteosarcoma humano (HOS ou GHOST) são as mais comumente utilizadas para este propósito (Blaak et al., 2005; Braun e Wiesmann, 2007; Jensen et al., 2003). O desenvolvimento biotecnológico originou o painel atual de ensaios para a determinação de tropismo de HIV, conhecidos como ensaios de vírus recombinantes, nos quais culturas celulares são infectadas por pseudovírus recombinantes, como o que acontece, por exemplo, no ensaio *Trofile* (Monogram Biosciences, EUA) considerado o padrão-ouro entre os ensaios fenotípicos atuais (Braun e Wiesmann, 2007).

Apesar da relativa popularidade e da posição de ensaios padrão-ouro, os ensaios de vírus recombinantes permanecem pouco acessíveis devido aos altos custos e a necessidade de envio das amostras para o exterior. Estes fatores implicaram no desenvolvimento de estratégias alternativas e mais acessíveis aos pacientes que necessitam introduzir o antagonista de CCR5 em seu tratamento (Sierra et al., 2007).

O primeiro indicativo genotípico para a mudança de tropismo foi a presença de resíduos positivos nas posições 11 e 25 da alça V3 da gp120 (De Jong et al., 1992). A partir desta observação foram desenvolvidos instrumentos de bioinformática que incorporaram algoritmos e bancos de dados para estimar o tropismo a partir de sequências de aminoácidos da alça V3. Estes sistemas preditivos estão disponíveis na *internet*, sendo as plataformas *WetCat*, *WebPSSM* e *geno2pheno*$_{[coreceptor]}$ as mais difundidas. Contudo, até o momento, nenhum destes testes foi capaz de substituir completamente os ensaios fenotípicos (Chueca et al., 2009; Garrido et al., 2008; Sierra et al., 2007).

5.1 Considerações metodológicas

Primeiramente, é importante ressaltar que o emprego dos preditores de bioinformática é dependente do sequenciamento da alça V3 e que o material genético do HIV é constituído de duas cópias de RNA que são convertidas em cDNA pela transcriptase reversa e inseridas no genoma hospedeiro pela ação da integrase. Por conseguinte, partindo do pressuposto de que o sangue periférico seria o espécime mais prático de ser obtido, é possível obter cópias do genoma viral a partir do RNA proveniente de partículas virais circulantes no plasma ou a partir do DNA pró-viral inserido no genoma das PBMCs. Considerando que 53% dos pacientes incluídos neste estudo apresentaram carga viral em níveis indetectáveis, impossibilitando a amplificação a partir do RNA plasmático devido ao baixo número de cópias, a extração de DNA a partir de PBMCs foi considerada a opção mais viável para a obtenção de material genético viral.

Os linfócitos T CD4+ ativados são responsáveis pela maior parte da produção de partículas virais (93-99%), enquanto 1-7% das partículas são provenientes principalmente de macrófagos e as células T de memória são responsáveis por uma produção de partículas virais inferior a 1% (Perelson et al., 1996). A duração de um linfócito infectado é de aproximadamente dois dias, enquanto a vida média de um vírion no plasma é de cerca de sete horas. Em média, a população de linfócitos T CD4+ periféricos é substituída a cada 15 dias em pacientes não expostos a antirretrovirais, sendo que em pacientes sob ART (*antiretroviral therapy* – terapia antirretroviral) a produção de linfócitos tende a ser superior à sua depleção (Ho et al., 1995; Perelson et al., 1996).

No contexto de tropismo viral, verificou-se que não há compartimentalização na produção de cepas R5 e X4 e as subpopulações virais encontradas no sangue periférico são produzidas por células com duração de vida semelhantes e não são geneticamente isoladas de um tipo particular de células (Ince et al., 2009). Um estudo, comparando o tropismo entre RNA plasmático e DNA proveniente de PBMCs demonstrou que em

75% dos pacientes foram observados genótipos semelhantes da sequência V3 em plasma e PBMCs (Saracino et al., 2009). Outro estudo semelhante mostrou que aproximadamente 20% das sequências provenientes de RNA plasmático foram preditas como cepas X4, enquanto mais de 50% das sequências provenientes de PBMCs foram preditas como X4, demonstrando o DNA pró-viral como uma fonte mais confiável de material genético viral para ensaios de tropismo tanto genotípicos, quanto fenotípicos. O uso do DNA pró-viral pode facilitar a execução de ensaios genotípicos de tropismo na rotina laboratorial, reduzindo a complexidade, o custo e o tempo do procedimento sem depender de cargas virais acima de 1000 cp/mL (Verhofstede et al., 2009).

5.2 Prevalência dos subtipos de HIV-1

As frequências dos subtipos de HIV-1 variam de acordo com a região geográfica estudada. A frequência de subtipos encontrada em nosso estudo B=83%, F1=12,8%, C=3,2%, e BF1=1,1%, confirmam os estudos brasileiros que revelaram a prevalência do subtipo B no país, podendo variar entre 70 e 90%, dependendo da área estudada; seguido pelo subtipo F, com três a 24%, sendo alguns recombinantes BF (cerca de 3%); e no sul do Brasil, o subtipo C representa aproximadamente 22% (de Martinez et al., 2002; Gadelha et al., 2003; Guimaraes et al., 2001; Sabino et al., 1996).

5.3 Avaliação dos preditores de tropismo

A determinação prévia do tropismo é indispensável para os candidatos ao uso dos antagonistas de CCR5, e a aprovação desta nova classe de antirretroviral para o uso clínico estimulou a realização de estudos que avaliam a eficiência destes ensaios, sejam fenotípicos ou genotípicos. O objetivo é instituir na rotina laboratorial metodologias pouco dispendiosas, porém, eficientes e que sejam capazes de fazer a triagem dos pacientes que tenham indicação para o uso dos antagonistas de CCR5. Assim, as

plataformas de bioinformática que predizem o tropismo a partir de sequências da alça V3 representam uma alternativa mais acessível que os ensaios fenotípicos baseados em vírus recombinantes.

Apesar das várias metodologias de bioinformática para a predição do tropismo propostas nos últimos anos, apenas três encontram-se disponíveis como ferramentas *online*. Neste estudo, foram avaliados os cinco testes disponíveis na plataforma *WetCat*, as duas matrizes do *WebPSSM* e a opção *clonal data* do *geno2pheno* $_{[coreceptor]}$, totalizando oito testes, com o objetivo de aumentar a confiabilidade da predição, visto que cada teste tem suas peculiaridades.

Até o momento, os três sistemas de predição estão restritos à análise da sequência da alça V3, mas apresentam diferenças em relação às metodologias aplicadas, aos conjuntos de dados nos quais são baseados e na maneira pela qual as sequências V3 devem ser fornecidas. O *WetCat*, ativo desde 2003, sem nenhuma atualização, é o sistema mais antigo e requer a inserção dos dados em uma formatação restrita, incluindo o alinhamento prévio da região V3 com uma sequência consenso específica. Em contraste, o *WebPSSM*, atualizado em 2006, constrói o alinhamento das sequências, enquanto o *geno2pheno* $_{[coreceptor]}$, iniciado em 2003 e atualizado quatro vezes, sendo a mais recente em novembro de 2009, é capaz de detectar e alinhar automaticamente a região V3 a partir da sequência inserida. O conjunto de dados destes dois últimos sistemas também é muito maior que do *WetCat*. Quanto à saída dos resultados, o *WetCat* classifica as sequências em X4 e não-X4, enquanto o *WebPSSM* demonstra um valor quantitativo que determina a probabilidade do vírus utilizar o correceptor CXCR4. Semelhantemente, o *geno2pheno* $_{[coreceptor]}$ permite a seleção de um nível de especificidade que mensura a probabilidade da predição estar correta, além de habilitar a inserção de marcadores clínicos como a contagem de linfócitos T CD4+ (Lengauer et al., 2007; Sierra et al., 2007).

O sistema *WetCat* disponibiliza cinco testes preditivos baseados em experimentos de validação cruzada de um conjunto de dados com 271 sequências. O teste *C4.5* realiza a predição a partir de uma árvore de

decisão, determinando o tropismo pela presença ou ausência de determinados aminoácidos em posições específicas, por exemplo: se existe uma glicina (G) na posição 12 e um glutamato (E) ou lisina (K) na posição 6 é predito um vírus X4 e assim por diante. O teste *C4.5* usando apenas as posições 8 e 12 (*C4.5 p8p12*), considera essas duas posições como as melhores para a determinação de tropismo por árvore de decisão, construindo uma árvore baseada somente nestas duas posições e, deste modo, o tropismo é determinado pela presença de certos resíduos nestas duas posições exclusivamente. O *PART* constrói uma regra a partir de uma árvore de decisão seguindo a descrição: se há uma S na posição 12 e uma I na posição 35 é um vírus R5, se há uma R na posição 12 é um vírus X4 e se há uma T na posição 8, N na p6, um *gap* na p16 e F na p23 é um vírus R5, em todos os outros casos é um vírus X4. O *SVM* (*support vector machine* – ferramenta de suporte de vetores) atribui escores para cada resíduo nas posições dentro de um gráfico de distribuição dos resíduos com três eixos, estimando predição do tropismo de acordo com a posição tridimensional do resíduo. A *Charge rule* (regra de cargas) segue a clássica 11/25, onde a presença dos resíduos positivos (H, K e R) nas posições 11 e 25 indica um vírus X4 (Pillai et al., 2003).

A plataforma *WebPSSM* utiliza matrizes de escore de posição específica (PSSM – *position-specific scoring matrices*) para analisar as sequências da região V3. Em síntese, o *PSSM* detecta a distribuição dos aminoácidos em posições específicas e associadas a agrupamentos de sequências determinados empiricamente. Este modelo é usado frequentemente para buscar sequências de DNA ou proteínas com um motivo particular, por exemplo, sítios regulatórios de transcrição, domínios *coiled-coil*, sítios de ligação ao complexo de histocompatibilidade principal de classe I (MHC-I), entre outros. O *PSSM* utiliza um conjunto de variações genéticas como base para comparações ou “modelo nulo” para facilitar a comparação de um resíduo em uma sequência com um grupo de alinhamento de sequências que apresenta a propriedade procurada. A comparação leva a um escore que pode ser interpretado como indicativo da

probabilidade de que a sequência apresente a propriedade de interesse. Para o desenvolvimento do *WebPSSM* foram utilizados agrupamentos empíricos de sequências V3 associadas com vírus X4 e R5 ou indutores e não indutores de sincício (SI/NSI), sendo do conjunto x4r5 composto por 213 sequências (17 X4, 168 R5 e 28 R5X4) e o conjunto sinsi por 257 sequências (70 SI e 187 NSI) (Jensen et al., 2003).

O *geno2pheno* $_{[coreceptor]}$ é baseado no modelo estatístico *SMV* e é um método de classificação dupla, mesmo que o tropismo apresente três classificações (R5, X4 E R5X4), este modelo foi desenvolvido para determinar se o vírus pode ou não utilizar CXCR4, portanto, vírus X4 e R5X4 são dispostos na classificação positiva. A escolha do *SMV* como modelo preditivo é proveniente da avaliação de outros métodos (incluindo a clássica regra de cargas, árvores de decisão, redes neurais, *PSSM* e misturas de regras de posições), onde a sensibilidade e a especificidade de cada modelo preditivo foram testados pelo método estatístico da curva ROC, determinando que *SMV* é o melhor modelo preditivo. O conjunto de dados é composto por 1100 sequências V3 (769 R5, 210 X4 e 131 R5X4) depositadas no banco de dados de Los Alamos (Sing et al., 2007).

Além das diferenças inerentes aos modelos preditivos, divergências nos resultados de cada teste, podem ser atribuídas à aplicação dos modelos preditivos, uma vez que em modelos que processam sequências provenientes de amostras de pacientes, a presença de mistura de nucleotídeos (posições onde dois ou mais nucleotídeos podem ser detectados entre as cepas prevalentes no paciente), contribui com uma perda substancial da sensibilidade e especificidade do modelo estatístico, quando comparado com a predição realizada a partir de sequências provenientes de clones, onde sequências não ambíguas são submetidas ao modelo estatístico (Lengauer et al., 2007).

A avaliação dos sistemas preditivos é imprescindível, quando se tem em vista a utilização desta estratégia na prática clínica. Devido à ausência de confirmação dos resultados por um ensaio padrão-ouro, foram utilizadas sequências da região V3 depositadas no banco de dados de Los Alamos e

que apresentam tropismo conhecido para avaliar a aplicabilidade dos testes preditivos. Os testes apresentaram especificidade acima de 90%, demonstrando que, em geral, são capazes de identificar as cepas R5. Entretanto, a sensibilidade esteve abaixo de 90%, o que indica uma ineficiência para a detecção de cepas X4.

Com o objetivo de garantir uma predição mais adequada, o estudo da curva ROC revelou a estratégia mais adequada para utilização dos sistemas preditivos, apresentando a melhor área abaixo da curva quando são utilizados em conjunto os testes *geno2pheno*$_{[coreceptor]}$, *sinsi*, *x4r5* e *SVM*. Em um estudo que também avaliou estes sistemas preditores, Chueca e colaboradores (2009) encontraram um resultado similar e concluíram que um algoritmo combinando as ferramentas de bioinformática *SMV*, *geno2pheno*$_{[coreceptor]}$, *x4r5* e *C4.5* poderia melhorar a predição do tropismo e mereceria avaliações complementares, uma vez que pode ser útil como uma estratégia de seleção na prática clínica.

De fato, quando são aplicados apenas esses quatro testes, os resultados inconclusivos persistem, porém, em menor proporção. Para minimizar este efeito, foram avaliadas as cargas elétricas líquidas das sequências da alça V3. Observou-se a tendência de cepas R5 apresentarem carga elétrica mais baixa, entre um e oito, enquanto cepas X4 tiveram variação de carga líquida entre quatro e 10. Esta mesma tendência foi observada em um estudo de Xu e colaboradores (2007), que avaliou as características da alça V3, e determinou que 90% sequencias R5 apresentaram carga líquida ≤ 4.

A partir destes resultados foi possível elaborar um algoritmo com 94% de confiabilidade na predição, associando quatro ferramentas preditivas e a carga elétrica líquida da alça V3. Com base neste algoritmo, foi realizada a inferência do tropismo das amostras obtidas neste estudo, demonstrando uma prevalência de cepas R5 (80%) nas 94 sequências estudadas, sendo que esta mesma proporção foi encontrada tanto entre os pacientes *naïve*, quanto em pacientes sob HAART. Alguns estudos demonstraram uma proporção semelhante, indicando que as cepas R5 são prevalentes no início

da infecção e as X4 emergem ao longo da doença, predominando em 10-20% dos pacientes não expostos a terapia antirretroviral e em 30-60% dos pacientes que desenvolvem resistência múltipla aos antirretrovirais (Jensen et al., 2003; Lengauer et al., 2007).

Os principais estudos a cerca do tropismo são realizados em coortes com prevalência do subtipo B. Alguns estudos indicam que a ocorrência de cepas R5X4 e X4 são raras no subtipo C (Dash et al., 2008; Tscherning et al., 1998) e os poucos estudos que relatam o tropismo no subtipo F indicam a ausência de cepas X4 e R5X4 (Ferraro et al., 2001; Tscherning et al., 1998), o que também foi verificado neste estudo. Os pacientes infectados com a variante B' apresentam uma progressão mais lenta da infecção, o que poderia indicar que estes vírus utilizam preferencialmente o correceptor CCR5 (Leal et al., 2008), entretanto, neste estudo foi possível observar que 16,7% das amostras X4 do subtipo B apresentaram o motivo GWGR.

5.4 Considerações finais

A elaboração de um algoritmo capaz de melhorar a confiabilidade da predição pode contribuir significativamente para a aplicação desta estratégia na triagem de candidatos ao uso dos antagonistas de correceptor. Entretanto, é importante ressaltar que esta estratégia também está sujeita a falhas, pois foi construída a partir de um conjunto específico de sequências e possíveis erros na anotação destas mesmas serão refletidos durante a predição de novas sequências. Esta interferência intrínseca pode ser observada pela diferença significativa ($p<0,0001$) quando são comparados o número de resultados inconclusivos em sequências do banco de dados (19%, n=35) e nas sequências obtidas neste estudo (56%, n=53), evidenciando um possível viés da análise, que no entanto, é inerente à metodologia aplicada. Portanto, não obstante à extensa avaliação dos sistemas preditores de tropismo disponíveis atualmente, estas ferramentas ainda apresentam algumas limitações e a realização de ensaios confirmatórios ainda é indispensável. Contudo, avanços na biotecnologia e

bioinformática poderão melhorar estes sistemas e ampliar seu uso na prática clínica e laboratorial.

Um fator importante para o aumento na qualidade da predição é em relação à metodologia para a obtenção das sequências, onde o uso de vetores de clonagem ou pirossequenciamento possibilitaria a detecção de cepas minoritárias. De fato, os ensaios fenotípicos detectam cepas minoritárias de até 1%, enquanto o sequenciamento direto de amostras dos pacientes é capaz de detectar a presença de cepas minoritárias superior a 20% da população viral total. Informações sobre a estrutura tridimensional da alça V3, dados clínicos e laboratoriais dos pacientes, como contagem de linfócitos T também são capazes de aumentar substancialmente a qualidade da predição (Lengauer et al., 2007; Sander et al., 2007).

Outro aspecto a ser salientado é que apesar da determinação do tropismo do HIV-1 estar centrada na alça V3, outros fatores encontrados na gp120, além destes 35 aminoácidos, podem interferir na afinidade do vírus pelo correceptor, sítios de N-glicosilação e variação na extensão da região V1/V2 (Hughes e Nelson, 2009; Mild et al., 2009).

Assim, o uso em conjunto das ferramentas de bioinformática disponíveis atualmente, apresenta uma predição mais confiável do que quando estas ferramentas são aplicadas isoladamente para determinar o tropismo. Diante das intensas avaliações deste estudo e, por consequência, da elaboração de um algoritmo capaz de melhorar a predição, a aplicação dos sistemas preditivos poderá representar uma alternativa viável, na prática clínica, atuando como estratégia de triagem para os candidatos ao uso dos antagonistas de CCR5, principalmente quando não for possível realizar um ensaio confirmatório para a determinação do tropismo.

6 CONCLUSÕES

- A avaliação dos sistemas preditivos utilizando sequências com tropismo conhecido possibilitou a elaboração de um algoritmo capaz de melhorar substancialmente a predição do tropismo;

- A determinação da carga elétrica líquida da alça V3 representou importante parâmetro adicional para complementar a predição do tropismo;

- Utilizando as ferramentas de bioinformática foi possível inferir uma prevalência de 80% de cepas R5 na população estudada;

- Até o momento, os testes preditivos disponíveis não devem ser utilizados isoladamente e testes confirmatórios são necessários para determinar o tropismo do HIV-1, sempre que possível.

REFERÊNCIAS

Alkhatib G, Berger EA. HIV coreceptors: from discovery and designation to new paradigms and promise. Eur J Med Res. 2007 Oct 15;12(9):375-84.

Blaak H, Boers PH, Gruters RA, Schuitemaker H, van der Ende ME, Osterhaus AD. CCR5, GPR15, and CXCR6 are major coreceptors of human immunodeficiency virus type 2 variants isolated from individuals with and without plasma viremia. J Virol. 2005 Feb;79(3):1686-700.

Braun P, Wiesmann F. Phenotypic assays for the determination of coreceptor tropism in HIV-1 infected individuals. Eur J Med Res. 2007 Oct 15;12(9):463-72.

Bredeek UF, Harbour MJ. CCR5 antagonists in the treatment of treatment-naive patients infected with CCR5 tropic HIV-1. Eur J Med Res. 2007 Oct 15;12(9):427-34.

Chueca N, Garrido C, Alvarez M, Poveda E, de Dios Luna J, Zahonero N, et al. Improvement in the determination of HIV-1 tropism using the V3 gene sequence and a combination of bioinformatic tools. J Med Virol. 2009 May;81(5):763-7.

Clevestig P, Pramanik L, Leitner T, Ehrnst A. CCR5 use by human immunodeficiency virus type 1 is associated closely with the gp120 V3 loop N-linked glycosylation site. J Gen Virol. 2006 Mar;87(Pt 3):607-12.

Cullen BR. Species and Tissue Tropisms of HIV-1: Molecular Basis and Phenotypic Consequences. HIV Sequence Compendium. 2001:1-12.

Dash PK, Siddappa NB, Mangaiarkarasi A, Mahendarkar AV, Roshan P, Anand KK, et al. Exceptional molecular and coreceptor-requirement properties of molecular clones isolated from an Human Immunodeficiency Virus Type-1 subtype C infection. Retrovirology. 2008;5:25.

De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. J Virol. 1992 Nov;66(11):6777-80.

de Martinez AM, Barbosa EF, Ferreira PC, Cardoso FA, Silveira J, Sassi G, et al. Molecular epidemiology of HIV-1 in Rio Grande, RS, Brazil. Rev Soc Bras Med Trop. 2002 Sep-Oct;35(5):471-6.

Delwart EL, Herring B, Rodrigo AG, Mullins JI. Genetic subtyping of human immunodeficiency virus using a heteroduplex mobility assay. PCR Methods Appl. 1995 Apr;4(5):S202-16.

Emmelkamp JM, Rockstroh JK. CCR5 antagonists: comparison of efficacy, side effects, pharmacokinetics and interactions--review of the literature. Eur J Med Res. 2007 Oct 15;12(9):409-17.

Este JA, Telenti A. HIV entry inhibitors. Lancet. 2007 Jul 7;370(9581):81-8.

Ferraro GA, Mello MA, Sutmoller F, Van Weyenbergh J, Shindo N, Galvao-Castro B, et al. Biological characterization and chemokine receptor usage of HIV type 1 isolates prevalent in Brazil. AIDS Res Hum Retroviruses. 2001 Sep 1;17(13):1241-7.

Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. J Clin Microbiol. 1995 Apr;33(4):906-11.

Gadelha SR, Shindo N, Cruz JN, Morgado MG, Galvao-Castro B. Molecular epidemiology of human immunodeficiency virus-1 in the state of Ceara, Northeast, Brazil. Mem Inst Oswaldo Cruz. 2003 Jun;98(4):461-3.

Garrido C, Roulet V, Chueca N, Poveda E, Aguilera A, Skrabal K, et al. Evaluation of eight different bioinformatics tools to predict viral tropism in different human immunodeficiency virus type 1 subtypes. J Clin Microbiol. 2008 Mar;46(3):887-91.

Genebat M, Ruiz-Mateos E, Leon JA, Gonzalez-Serna A, Pulido I, Rivas I, et al. Correlation between the Trofile test and virological response to a short-term maraviroc exposure in HIV-infected patients. J Antimicrob Chemother. 2009 Oct;64(4):845-9.

Guimaraes ML, Bastos FI, Telles PR, Galvao-Castro B, Diaz RS, Bongertz V, et al. Retrovirus infections in a sample of injecting drug users in Rio de Janeiro City, Brazil: prevalence of HIV-1 subtypes, and co-infection with HTLV-I/II. J Clin Virol. 2001 May;21(2):143-51.

Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995 Jan 12;373(6510):123-6.

Hughes A, Nelson M. HIV entry: new insights and implications for patient management. Curr Opin Infect Dis. 2009 Feb;22(1):35-42.

Ince WL, Harrington PR, Schnell GL, Patel-Chhabra M, Burch CL, Menezes P, et al. Major coexisting human immunodeficiency virus type 1 env gene subpopulations in the peripheral blood are produced by cells with similar turnover rates and show little evidence of genetic compartmentalization. J Virol. 2009 May;83(9):4068-80.

Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX, 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 Dec;77(24):13376-88.

Jones R, Nelson M. The role of receptors in the HIV-1 entry process. Eur J Med Res. 2007 Oct 15;12(9):391-6.

Karpas A. Human retroviruses in leukaemia and AIDS: reflections on their discovery, biology and epidemiology. Biol Rev Camb Philos Soc. 2004 Nov;79(4):911-33.

Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science. 2006 Jul 28;313(5786):523-6.

Kondru R, Zhang J, Ji C, Mirzadegan T, Rotstein D, Sankuratri S, et al. Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. Mol Pharmacol. 2008 Mar;73(3):789-800.

Koning FA, van Rij RP, Schuitemaker H. Biological and Molecular Aspects of HIV-1 Coreceptor Usage. HIV Sequence Compendium. 2002:24-42.

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 Nov 25;381(2):222-9.

Leitner T, Korber B, Daniels M, Calef C, Foley B. HIV-1 Subtype and Circulating Recombinant Form (CRF) Reference Sequences, 2005. HIV Sequence Compendium. 2005:41-48.

Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. Nat Biotechnol. 2007 Dec;25(12):1407-10.

Lieberman-Blum SS, Fung HB, Bandres JC. Maraviroc: a CCR5-receptor antagonist for the treatment of HIV-1 infection. Clin Ther. 2008 Jul;30(7):1228-50.

Lorenzen T, Stoehr A, Walther I, Plettenberg A. CCR5 antagonists in the treatment of treatment-experienced patients infected with CCR5 tropic HIV-1. Eur J Med Res. 2007 Oct 15;12(9):419-25.

Mild M, Kvist A, Esbjornsson J, Karlsson I, Fenyo EM, Medstrand P. Differences in molecular evolution between switch (R5 to R5X4/X4-tropic) and non-switch (R5-tropic only) HIV-1 populations during infection. Infect Genet Evol. 2009 May 14.

Munerato P, Azevedo ML, Sucupira MC, Pardini R, Pinto GH, Catroxo M, et al. Frequency of polymorphisms of genes coding for HIV-1 co-receptors CCR5 and CCR2 in a Brazilian population. Braz J Infect Dis. 2003 Aug;7(4):236-40.

Peeters M. Recombinant HIV sequences: their role in the global epidemic. HIV Sequence Compendium. 2000:I-39-54.

Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science. 1996 Mar 15;271(5255):1582-6.

Pillai S, Good B, Richman D, Corbeil J. A new perspective on V3 phenotype prediction. AIDS Res Hum Retroviruses. 2003 Feb;19(2):145-9.

Pollakis G, Abebe A, Kliphuis A, Chalaby MI, Bakker M, Mengistu Y, et al. Phenotypic and genotypic comparisons of CCR5- and CXCR4-tropic human immunodeficiency virus type 1 biological clones isolated from subtype C-infected individuals. J Virol. 2004 Mar;78(6):2841-52.

Poon AF, Lewis FI, Pond SL, Frost SD. An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. PLoS Comput Biol. 2007 Nov;3(11):e231.

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 nomenclature proposal. Science. 2000 Apr 7;288(5463):55-6.

Rosen O, Sharon M, Quadt-Akabayov SR, Anglister J. Molecular switch for alternative conformations of the HIV-1 V3 region: implications for phenotype conversion. Proc Natl Acad Sci U S A. 2006 Sep 19;103(38):13950-5.

Sabino EC, Diaz RS, Brigido LF, Learn GH, Mullins JI, Reingold AL, et al. Distribution of HIV-1 subtypes seen in an AIDS clinic in Sao Paulo City, Brazil. AIDS. 1996 Nov;10(13):1579-84.

Sander O, Sing T, Sommer I, Low AJ, Cheung PK, Harrigan PR, et al. Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage. PLoS Comput Biol. 2007 Mar 30;3(3):e58.

Saracino A, Monno L, Cibelli DC, Punzi G, Brindicci G, Ladisa N, et al. Co-receptor switch during HAART is independent of virological success. J Med Virol. 2009 Dec;81(12):2036-44.

Schultz AK, Zhang M, Bulla I, Leitner T, Korber B, Morgenstern B, et al. jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic Acids Res. 2009 Jul 1;37(Web Server issue):W647-51.

Sierra S, Kaiser R, Thielen A, Lengauer T. Genotypic coreceptor analysis. Eur J Med Res. 2007 Oct 15;12(9):453-62.

Simon-Loriere E, Galetto R, Hamoudi M, Archer J, Lefeuvre P, Martin DP, et al. Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus. PLoS Pathog. 2009 May;5(5):e1000418.

Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, et al. Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. Antivir Ther. 2007;12(7):1097-106.

Skrabal K, Low AJ, Dong W, Sing T, Cheung PK, Mammano F, et al. Determining human immunodeficiency virus coreceptor use in a clinical setting: degree of correlation between two phenotypic assays and a bioinformatic model. J Clin Microbiol. 2007 Feb;45(2):279-84.

Tscherning C, Alaeus A, Fredriksson R, Bjorndal A, Deng H, Littman DR, et al. Differences in chemokine coreceptor usage between genetic subtypes of HIV-1. Virology. 1998 Feb 15;241(2):181-8.

Turner BG, Summers MF. Structural biology of HIV. J Mol Biol. 1999 Jan 8;285(1):1-32.

Verhofstede C, Vandekerckhove L, Eygen VV, Demecheleer E, Vandenbroucke I, Winters B, et al. CXCR4-using HIV type 1 variants are more commonly found in peripheral blood mononuclear cell DNA than in plasma RNA. J Acquir Immune Defic Syndr. 2009 Feb 1;50(2):126-36.

Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. Antimicrob Agents Chemother. 2007 Feb;51(2):566-75.

Xu S, Huang X, Xu H, 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 Oct;45(5):441-6.

ANEXO

Anexo A – Certificado emitido pela Comissão de Ética

HOSPITAL DAS CLÍNICAS
DA FACULDADE DE MEDICINA
DA UNIVERSIDADE DE SÃO PAULO

APROVAÇÃO

A Comissão de Ética para Análise de Projetos de Pesquisa - CAPPesq da Diretoria Clínica do Hospital das Clínicas e da Faculdade de Medicina da Universidade de São Paulo, em sessão de 04/06/2008, **APROVOU** o Protocolo de Pesquisa nº **0108/08**, intitulado: "**DETERMINAÇÃO DO TROPISMO DO HIV-1 PELOS CO-RECEPTORES CCR5 E CXCR4 PELO USO DE FERRAMENTAS DE BIOINFORMÁTICA**" apresentado pelo Departamento de **DERMATOLOGIA**, inclusive o Termo de Consentimento Livre e Esclarecido.

Cabe ao pesquisador elaborar e apresentar à CAPPesq, os relatórios parciais e final sobre a pesquisa (Resolução do Conselho Nacional de Saúde nº 196, de 10/10/1996, inciso IX.2, letra "c").

Pesquisador (a) Responsável: **Dr. Jorge Simão do Rosário Casseb**
Pesquisador (a) Executante: **Liã Bárbara Arruda**

CAPPesq, 06 de Junho de 2008

Prof. Dr. Eduardo Massad
Presidente da Comissão de Ética para Análise de Projetos de Pesquisa

Comissão de Ética para Análise de Projetos de Pesquisa do HCFMUSP e da FMUSP Diretoria Clínica do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo Rua Ovídio Pires de Campos, 255, 5º andar - CEP 05403 010 - São Paulo – SP Fone: 011 3069 6442 Fax: 011 3069 6492 e-mail: cappesq@hcnet.usp.br / secretariacappesq2@hcnet.usp.br

APÊNDICE

Apêndice A – Termo de Consentimento Livre e Esclarecido

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

CAPPesq 08/0108

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL

1.NOME:.:.....
DOCUMENTO DE IDENTIDADE Nº : ..... SEXO :.M □ F □
DATA NASCIMENTO: ...../...../.....
ENDEREÇO ..... Nº..... APTO: .....
BAIRRO:.....CIDADE.....
CEP:..... TELEFONE: DDD (.....) .....

2.RESPONSÁVEL LEGAL .....
NATUREZA (grau de parentesco, tutor, curador etc.) .....
DOCUMENTO DE IDENTIDADE :.....SEXO: M □ F □
DATA NASCIMENTO.: ...../...../.....
ENDEREÇO:.....Nº.....APTO: .....
BAIRRO:.....CIDADE:.....
CEP:.....TELEFONE:DDD(.....).....

DADOS SOBRE A PESQUISA

1. TÍTULO DO PROTOCOLO DE PESQUISA: DETERMINAÇÃO DO TROPISMO DO HIV-1 PELOS CO-RECEPTORES CCR5 E CXCR4 PELO USO DE FERRAMENTAS DE BIOINFORMÁTICA

PESQUISADOR : Dr. Jorge Simão do Rosário Casseb

CARGO/FUNÇÃO: Médico-Pesquisador INSCRIÇÃO CONSELHO REGIONAL Nº : CRM/SP 67217

2. UNIDADE DO HCFMUSP: Departamento de Dermatologia do Hospital das Clínicas da Faculdade de São Paulo

3. AVALIAÇÃO DO RISCO DA PESQUISA:

RISCO MÍNIMO □ RISCO MÉDIO □

RISCO BAIXO x RISCO MAIOR □

4.DURAÇÃO DA PESQUISA : 24 meses

O (a) senhor (a) está sendo convidado a participar de um estudo sobre o tropismo Vírus da Imunodeficiência Humana (HIV).

Para infectar uma pessoa, o HIV precisa se ligar a célula CD4, utilizando diferentes “portas de entrada”. A determinação de qual destas portas de entrada é utilizada pelo vírus é importante para a aplicação de novos tratamentos. O vírus apresenta afinidade (tropismo) por diferentes “portas de entrada”. Assim, o objetivo do nosso estudo é avaliar os testes de computador podem determinar o tropismo do HIV.

Para participar do estudo, serão solicitadas algumas informações de identificação pessoal (nome, sobrenome, data de nascimento, local de nascimento, endereço e outros) e o (a) senhor (a) deverá responder a um questionário. Estas informações serão mantidas em sigilo. Após esta etapa, o (a) senhor (a) coletará 5 mL (uma colher de sopa) através de punção venosa do sangue periférico.

Esta coleta será realizada com uma agulha esterilizada conectada em um tubo de vidro contendo anticoagulante (EDTA) e será introduzida em uma veia do seu antebraço, aspirando o sangue para dentro do tubo por um processo a vácuo.

A coleta não acarretará em risco a sua saúde e o desconforto gerado consiste em uma picada durante a introdução da agulha no antebraço, idêntica ao procedimento normal de coleta de sangue.

O (A) senhor (a) não receberá nenhum benefício direto por fazer parte deste estudo. As informações obtidas poderão ajudar a outras pessoas que tenham o HIV.

O (A) senhor (a) também tem o direito de optar por não participar do estudo. Caso o (a) senhor (a) decida não participar, isto em nada afetará o seu direito a receber tratamento na rede pública de saúde.

Em qualquer etapa do estudo, o (a) senhor (a) terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. O principal investigador é o Dr. Jorge Casseb que pode ser encontrados no Laboratório de Alergia e Imunologia Clínica e Experimental - LIM/56 Av. Enéas de Carvalho Aguiar, 470 / prédio IMT II / 3º andar ou pelos telefones: (11) 3061-7193 / 7457. Se o (a) senhor (a) tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) – Rua Ovídio Pires de Campos, 225 – 5º andar – tel.: 3069-6442 ramais 16, 17, 18 ou 20, FAX: 3069-6442 ramal 26 – E-mail: cappesq@hcnet.usp.br

Reforçamos que participar do estudo é um ato inteiramente voluntário. O (A) senhor (a) pode preferir não participar do estudo ou abandoná-lo a qualquer momento. O (A) senhor (a) será tratado da mesma forma, independentemente de sua decisão. Nós prestaremos todos os serviços clínicos e profissionais de diagnóstico e de laboratório que fizerem parte do estudo, sem que isso incorra em custo algum para o (a) senhor (a).

As informações obtidas serão analisadas em conjunto com as informações de outros pacientes, não sendo divulgado a identificação de nenhum paciente.

Nós lhe comunicaremos sobre novas informações deste ou de outros estudos que possam afetar sua saúde, bem-estar ou vontade de continuar participando deste estudo. Quando o estudo terminar, nós lhe informaremos os resultados.

Não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação. Se existir qualquer despesa adicional, ela será absorvida pelo orçamento da pesquisa.

Em caso de dano pessoal causado diretamente pelos procedimentos ou tratamentos propostos neste estudo (nexo causal comprovado), o participante tem direito a tratamento médico na Instituição, bem como às indenizações legalmente estabelecidas.

Todos os dados e o material coletado serão utilizados somente para esta pesquisa.

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo "DETERMINAÇÃO DO TROPISMO DO HIV-1 PELOS CO-RECEPTORES CCR5 E CXCR4 PELO USO DE FERRAMENTAS DE BIOINFORMÁTICA".

Eu discuti com o Dr. Jorge Casseb sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste Serviço.

Assinatura do paciente/representante legal Data ___/___/___

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste paciente ou representante legal para a participação neste estudo.

Assinatura do responsável pelo estudo Data ___/___/___

Instituto de Medicina Tropical de São Paulo
Universidade de São Paulo
Av. Dr. Enéas de Carvalho Aguiar, 470 CEP 05403-000
São Paulo – SP – Brasil
Fone/Fax (0xx11) 3061-7011
http://www.imtsp.usp.br