



FUNDAÇÃO OSWALDO CRUZ
CENTRO DE PESQUISA GONÇALO MONIZ

FIOCRUZ

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

DISSERTAÇÃO DE MESTRADO

**LASP-HIV1-REStool: DESENVOLVIMENTO DE UMA FERRAMENTA DE
BIOINFORMÁTICA PARA ANÁLISE DE RESISTÊNCIA DO HIV-1 AOS
ANTIRRETROVIRAIS**

DOMINGOS RAMON MOREAU DA CUNHA

Salvador – Bahia – Brasil

2010

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

**LASP-HIV1-RESTool: DESENVOLVIMENTO DE UMA FERRAMENTA DE
BIOINFORMÁTICA PARA ANÁLISE DE RESISTÊNCIA DO HIV-1 AOS
ANTIRRETROVIRAIS**

DOMINGOS RAMON MOREAU DA CUNHA

Dissertação apresentada ao programa de Pós-Graduação em Biotecnologia em Saúde e Medicina Investigativa do Centro de Pesquisa Gonçalo Moniz (CPqGM) como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Orientador: Prof. Dr. Luiz Carlos Júnior Alcântara
Co-orientador: Prof. Dr. José Carlos Couto Fernandez

**Salvador – Bahia – Brasil
2010**

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

C9721 Cunha, Domingos Ramon Moreau da
Lasp-HIV1-restool: desenvolvimento de uma ferramenta de bioinformática
para análise de resistência do HIV-1 aos antirretrovirais [manuscrito] /
Domingos Ramon Moreau da Cunha. - 2010.

99 f.: il. ; 30 cm.

Datilografado (fotocópia).

Dissertação (mestrado) – Fundação Oswaldo Cruz, Centro de Pesquisas
Gonçalo Moniz, 2010.

Orientador: Prof. Dr. Prof. Dr. Luiz Carlos Júnior Alcântara, Laboratório
Avançado de Saúde Pública.

1. Bioinformática. 2. HIV. 3. Retrovirus. 4. Mutação. I. Título.

CDU 004:616.98

“LASP – HIV1 – RESToll: DESENVOLVIMENTO DE UMA FERRAMENTA DE BIOINFORMÁTICA
PARA ANÁLISE DE RESISTÊNCIA DE HIV-1 AOS ANTIRRETROVIRAIS”

DOMINGOS RAMON MOREAU DA CUNHA

FOLHA DE APROVAÇÃO

COMISSÃO EXAMINADORA



Dr. Walter de Araújo Eyer Silva
Professor Adjunto
UNI-RIO



Dr.ª Maria de Lourdes Farre Valle
Assistente de Pesquisa em Saúde Pública
CPqGM / FIOCRUZ

AGRADECIMENTOS

A Deus, pela minha existência e por sua presença constante em minha vida.

As minhas queridas e amadas esposa e filhas por serem a minha maior razão de viver e por terem me incentivado, auxiliado e suportado os meus momentos de ausência e impaciência.

Ao meu pai e a minha mãe que sempre estiveram ao meu lado mesmo que em pensamentos, me apoiando e incentivando.

Ao meu irmão e a minha irmã pelo apoio e amizade nos momentos mais difíceis.

Ao Dr. Luiz Carlos Junior Alcântara por ter me iniciado na bioinformática, me orientado neste trabalho com seu grande conhecimento e pela sua amizade.

Ao Dr. José Carlos Couto Fernandez por concordar em ser meu co-orientador e com sua vasta experiência e conhecimento ter dado uma contribuição de fundamental importância para a realização deste trabalho, além da amizade.

Ao Dr. Bernardo Galvão Castro-Filho pela amizade, pelos momentos de conversa que, mesmo quando informais, são bons ensinamentos e por ter me aceito no LASP.

A Dra. Mariza Gonçalves Morgado, por ter me aceito no Laboratório de AIDS e Imunologia Molecular e pela gentileza com a qual me recebeu.

A todos os amigos do Laboratório Avançado de Saúde Pública, dos quais, mesmo ausente, sempre me lembro com muita alegria e saudades.

Aos novos colegas do Laboratório de AIDS e Imunologia Molecular do Instituto Oswaldo Cruz.

“A mente que se abre a uma nova idéia
jamais voltará ao seu tamanho original.”

(Albert Einstein)

RESUMO

As terapias antirretrovirais (TARV) trouxeram um grande benefício para os pacientes infectados pelo HIV, porém não conseguem impedir totalmente o surgimento de formas virais resistentes, causadas principalmente pela elevada taxa mutacional do vírus. O desenvolvimento de resistência do HIV aos antirretrovirais é um fator limitante para o sucesso da TARV. Os portadores de vírus resistentes, além de não responderem adequadamente ao tratamento, podem também transmitir estes vírus mutantes, representando grave problema de saúde pública. O acúmulo de mutações de resistência aos antirretrovirais representa um desafio importante na melhoria do tratamento de pacientes com vírus multiresistentes.

Atualmente, dois tipos de métodos estão estabelecidos para avaliação da resistência ou sensibilidade do HIV aos antirretrovirais: os testes fenotípicos e genotípicos de resistência. As análises de resistência do HIV aos antirretrovirais, avançaram muito nos últimos anos com o desenvolvimento dos algoritmos para avaliação de resistência. Atualmente, uma série de sistemas de interpretação de resistência fenotípica e genotípica do HIV aos antirretrovirais estão disponíveis na Internet. Estes sistemas identificam respectivamente os níveis de sensibilidade *in vitro* aos antirretrovirais e mutações associadas à resistência em sequências virais, e retornam o perfil de resistência do HIV, funcionando portanto, como importantes ferramentas para utilização em análises de resistência.

No presente trabalho, foi desenvolvida uma ferramenta de bioinformática para análise de resistência do HIV-1 aos antirretrovirais, chamada de LASP-HIV1-ResTool. A ferramenta LASP-HIV1-ResTool é capaz de otimizar a análise de dados genômicos do HIV-1, é de fácil manuseio e está disponível no site da unidade de bioinformática do LASP/CPqGM/FIOCRUZ, Bioáfrica e IOC/FIOCRUZ para domínio público. Esta ferramenta pode acessar um banco de dados com sequências gênicas previamente armazenadas, extrair informações relativas à resistência aos antirretrovirais, comparar sequências de HIV-1 dos bancos de dados públicos, identificando mutações nos genes que codificam a transcriptase reversa e a protease e associar esses dados a resistência a cada antirretroviral utilizado nas terapias anti-HIV/AIDS. Adicionalmente, a ferramenta possibilita localizar sítios pós-traducionais e traçar um perfil de distribuição das sequências armazenadas no banco de dados local, tanto por resistência, quanto por sítios pós-traducionais, para servirem de parâmetro de comparação com as sequências

submetidas pelos usuários da ferramenta.

A ferramenta LASP-HIV1-ResTool apresenta uma tabela de quantidade e percentual de sequências brasileiras do HIV-1, dispostas por subtipo e outra tabela apresentando a quantidade e o percentual de seqüências das regiões *Protease* e *Transcriptase Reversa* do HIV-1 armazenadas no Banco de Dados. Os usuários da ferramenta podem analisar sequências do banco de dados que compõe a ferramenta ou submeter suas próprias sequências de HIV-1 no formato FASTA.

Durante o processamento das sequências, o sistema identifica a sequência através do número de acesso e localiza a posição inicial da sequência dentro da região gênica escolhida. Em seguida identifica a fase de leitura correta e realiza a conversão da sequência de nucleotídeos em sequência de aminoácidos. Inicia-se então o processo de localização de todas as mutações ocorridas, tanto na sequência de ácido nucléico como na sequência de aminoácido. O sistema então apresenta uma tabela apenas com as mutações que conferem a resistência do HIV-1 contra o antirretroviral. Além disso, exibe o grau de resistência de cada mutação e indica, ao final da tabela, se a amostra de HIV-1 submetida na análise é susceptível, apresenta resistência intermediária ou se é resistente ao antirretroviral. O sistema exibe dois gráficos que relacionam: a quantidade de sítios pós-traducionais em função do total de sequências submetidas e os padrões de resistência aos antirretrovirais (susceptível, intermediária e resistente) em função do total de sequências. Além de apresentar um gráfico de barras mostrando a quantidade de ocorrências de cada mutação de resistência no conjunto de sequências analisadas.

A ferramenta LASP-HIV1-ResTool também gera uma página em formato XML contendo os dados das análises do usuário, que podem ser salvas em um arquivo. Os arquivos XML podem ser importados de outras ferramentas tais como: planilhas eletrônicas, programas de análise estatística e sistemas gerenciadores de bancos de dados.

Para validação da aplicabilidade da ferramenta LASP-HIV1-ResTool foram utilizadas 111 amostras correspondentes a região gênica da protease e da transcriptase reversa derivadas da genotipagem da resistência do HIV-1 provenientes de pacientes infectados, que apresentavam falha virológica a TARV. Quando comparados os dados obtidos na ferramenta LASP-HIV1-ResTool com os dados obtidos através das análises realizadas com a ferramenta de Stanford, pode-se observar a similaridade entre os resultados. Por outro lado, pode-se observar diferenças entre os resultados das análises realizadas com a ferramenta LASP-HIV1-ResTool quando comparados com a ferramenta o sistema RENAGENO.

ABSTRACT

The antiretroviral therapy (HAART) has brought a great benefit to HIV-infected patients, but can not completely prevent the emergence of resistant viral forms, mainly caused by the high mutation rate of the virus. The development of HIV resistance to antiretroviral drugs is a limiting factor for the success of HAART. The patients with resistant virus, that do not respond adequately to treatment, may also transmit this virus mutants, representing a serious public health problem. The accumulation of resistance mutations to antiretroviral drugs represents a major challenge in improving the treatment of multiresistant virus. Currently, two types of methods are established to assess resistance or susceptibility of HIV to antiretroviral drugs: the phenotypic and genotypic resistance. However, the analysis of HIV resistance to antiretrovirals, have advanced greatly in recent years with the development of algorithms for evaluation of resistance. Currently, a number of systems for interpretation of HIV resistance to antiretrovirals, using algorithms, are available on the Internet. These systems identify the mutations associated with resistance, in amino acids sequences and return the viral resistance profile of HIV, acting therefore as, important tools for use in resistance analysis. In this study, we developed a bioinformatics tool for analysis of HIV-1 resistance to antiretrovirals. This tool is easy to use, will be available at the bioinformatics unit of LASP / CPqGM / FIOCRUZ, Bioafrica and IOC / FIOCRUZ in public domain and is able to optimize the analysis of genomic data of HIV-1. This tool can access a database of gene sequences previously stored, extract information, compare HIV-1 sequences from public databases, identifying mutations in genes that encode reverse transcriptase and protease, and associate that data with resistance to the individual anti -retroviral therapies used in anti-HIV/AIDS. Additionally, the tool allows locate post-translational sites and establish a distribution profile of the sequences stored in local databases, both for resistance and by post-translational sites to serve as parameter for comparison with the sequences submitted by users of the tool.

LISTA DE FIGURAS, TABELAS E GRÁFICOS

Figura		Página
01	Epidemiologia do HIV.	6
02	Estrutura morfológica do HIV-1.	10
03	Organização genômica do HIV-1.	12
04	Ciclo de replicação do HIV-1.	14
05	Sistema HIV Sequence Database.	36
06	Página do HIV Sequence Data Base contento as sequências selecionadas para download.	37
07	Página do HIV Sequence Data Base contendo as sequências e o quadro com as opções de downloads.	39
08	Esquematização da organização do sistema de subtipagem do HIV-1 REGA subtyping tool.	41
09	Página inicial do sistema de subtipagem do HIV-1 REGA subtyping tool.	43
10	Página d sistema de subtipagem do HIV-1 REGA subtyping tool com as sequências subtipadas do HIV-1 do Brasil	44
11	Organização das sequências e inclusão no HIVbase.	45
12	Sistema HIVbase.	46
13	Esquematização do banco de ados relacional.	54
14	Página de apresentação da ferramenta LASP-HIV1-ResTool.	60
15	Página utilizada para a escolha dos parâmetros de análise.	61

16	Página utilizada para a submissão das sequências de HIV-1 em arquivos no formato FASTA.	62
17	Topo da tela de resultados da análise de resistência aos antirretrovirais e sítios pós-traducionais.	65
18	Base da tela de resultados da análise de resistência aos antirretrovirais e sítios pós-traducionais.	66
19	Página contendo a sequência de nucleotídeo submetida e sua respectiva tradução em aminoácidos.	67
20	Página contendo as análises no formato XML para serem armazenadas.	68

Tabela

1	Antirretrovirais inibidores nucleosídeos da transcriptase reversa aprovados pelo FDA.	18
2	Antirretrovirais inibidores não-nucleosídeos da transcriptase reversa aprovados pelo FDA.	18
3	Antirretrovirais inibidores da protease aprovados pelo FDA.	19
4	Antirretrovirais inibidores de integrase aprovados pelo FDA.	20
5	Regras para a identificação de sítios pós-traducionais.	50

Gráfico

1	Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta LASP-HIV1-ResTool.	70
2	Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta Stanford.	70

3	Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta RENAGENO.	71
4	Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta LASP-HIV1-ResTool	71

LISTA DE ABREVIATURAS E SIGLAS

	Acquired Immune Deficiency Syndrome ou Síndrome da Imunodeficiência
AIDS	Adquirida
cDNA	Ácido Desoxirribonucléico complementar
CD4	cluster of differentiation 4 ou agrupamento de diferenciação tipo 4
CPqGM	Centro de Pesquisa Gonçalo Moniz
CRFs	Circulating Recombinant Forms ou Formas Recombinantes Circulantes
DNA	Ácido Desoxirribonucléico
DST	Doença Sexualmente Transmissível
FASTA	Formato de representação de sequência nucleotídica
FDA	Food and Drug Administration ou Administração de Alimentos e Drogas
FIOCRUZ	Fundação Oswaldo Cruz
	Highly Active Anti-Retroviral Therapy ou Terapia Antirretroviral Altamente
HAART	Ativa
HIV	Human Immunodeficiency Vírus ou Vírus da Imunodeficiência Humana
IFs	Inibidores de Fusão
in vitro	Experimentos realizados em meio artificial (“em vidro”)
INIs	Inibidores De Integrase
IPs	Inibidores da Protease
LASP	Laboratório Avançado em Saúde Pública
<i>Nef</i>	Negative factor ou fator negativo
NIAID	National Institute of Allergy and Infectious Diseases
NIH	National Institutes of Health
NNRTIs	Inibidores Não Nucleosídeos da Transcriptase Reversa

NRTIs	Inibidores Nucleosídeos da Transcriptase Reversa
PCR	Reação de Polimerase em Cadeia
PN-DST/AIDS	Programa Nacional de Doenças Sexualmente Transmissíveis e da AIDS
PHP	PHP: Hypertext Preprocessor (sigla em inglês)
PR	Protease
RENAGENO	Rede Nacional de Genotipagem
rev	Regulatory factor for expression ou Fator regulador da expressão
RNA	Ácido Ribonucléico
mRNA	Ácido Ribonucléico mensageiro
SGBD	Sistema Gerenciador de Banco de Dados
TARV	Terapia Antirretroviral
Tat	Transactivator ou Transativador
TR	Transcriptase Reversa
	Joint United Nations Programme on HIV/AIDS) ou Programa Conjunto das
UNAIDS	Nações Unidas sobre a HIV/AIDS
<i>Vif</i>	Viral infectivity factor ou Fator da infectividade viral
<i>Vpr</i>	viral protein R ou proteína R Viral
<i>Vpu</i>	viral protein U ou Proteína U viral
<i>Vpx</i>	Gene homólogo ao <i>vpr</i> encontrado no HIV-2
XML	Extensible Markup Language ou Linguagem de Marcação Estendida

SUMÁRIO

	Página
1. INTRODUÇÃO	4
1.1. A EPIDEMIA DA AIDS	5
1.2. O VÍRUS DA IMUNODEFICIÊNCIA HUMANA (HIV)	8
1.2.1. CICLO DE REPLICAÇÃO DO HIV	13
1.2.2. CLASSIFICAÇÃO DO HIV	15
1.3. TERAPIA ANTIRRETROVIRAL	17
1.4. RESISTÊNCIA AOS ANTIRRETROVIRAIS	21
1.5. SISTEMAS DE INTERPRETAÇÃO DE ALGORÍTMOS DE RESISTÊNCIA	24
2.0. RACIONAL	28
3.0. OBJETIVOS	32
3.1. GERAL	33
3.2. ESPECÍFICOS	33
4.0. MATERIAL E MÉTODOS	34
4.1. OBTENÇÃO DAS SEQUÊNCIAS	35
4.2. ORGANIZAÇÃO DAS SEQUÊNCIAS	40
4.2.1. SUBTIPAGEM DAS SEQUÊNCIAS	40
4.2.2. ESTRATIFICAÇÃO DAS SEQUÊNCIAS	44
4.2.3. HIVBASE	45
4.3. ARMAZENAMENTO DAS SEQUÊNCIAS	47
4.4. ALGORITMO DE RESISTÊNCIA	48

4.5. ALGORITIMO DE RESISTÊNCIA	49
4.6. CRIAÇÃO DA FERRAMENTA DE ANÁLISE DE RESISTÊNCIA	50
5.0. RESULTADOS	53
5.1. 1ª PARTE: DESCRIÇÃO DA CONSTRUÇÃO DA FERRAMENTA	53
5.1.1. BANCO DE DADOS	53
5.1.1.2 TABELA SEQUENCIA	54
5.1.1.3. TABELA REGIAO	55
5.1.1.4. TABELA GENOMA	56
5.1.1.5. TABELA RESISTENCIA	56
5.1.1.6. TABELA SUBTIPO	57
5.1.1.7. TABELA SUBTIPOREGIÃO	57
5.1.1.8. TABELA DROGA	57
5.1.2. PROGRAMAÇÃO DE SCRIPT PARA MANIPULAÇÃO DE SEQUÊNCIAS	58
5.1.2.1 - ORGANIZAÇÃO DAS SEQUÊNCIAS	58
5.1.2.3. FERRAMENTA DE ANÁLISE DE RESISTÊNCIA	58
5.2. 2ª PARTE: DESCRIÇÃO DA UTILIZAÇÃO DA FERRAMENTA	59
5.2.1. SUBMISSÃO DOS DADOS PARA ANÁLISE	59
5.2.2. RESULTADO DA ANÁLISE DOS DADOS SUBMETIDOS A FERRAMENTA	64
5.3. PARTE 3: DESCRIÇÃO DA UTILIZAÇÃO DA FERRAMENTA	69
5.3.1. AVALIAÇÃO COMPARATIVA DOS RESULTADOS DAS ANÁLISES GERADAS PELA FERRAMENTA	69
6.0. DISCUSSÃO	72

6.1. SÍTIOS PÓS-TRADUCIONAIS	74
6.2. AVALIAÇÃO COMPARATIVA DOS RESULTADOS DAS ANÁLISES GERADAS PELA FERRAMENTA	77
7. CONCLUSÕES	79
8. REFERÊNCIAS BIBLIOGRÁFICAS	81
9. ANEXOS	94
9.1. TRABALHOS APRESENTADOS EM CONGRESSOS	95
9.2. CURSOS RELACIONADOS A BIOINFORMÁTICA EM HIV/AIDS	95
9.3. PUBLICAÇÕES CIENTÍFICAS RELACIONADAS A BIOINFORMÁTICA EM HIV/AIDS	96
9.4. MOTIVAÇÃO PARA DESENVOLVIMENTO DA FERRAMENTA LASP-HIV1-ResTool	97

INTRODUÇÃO

1. INTRODUÇÃO:

1.1. A EPIDEMIA DE AIDS

A síndrome de imunodeficiência adquirida (do inglês AIDS) foi descrita inicialmente nos Estados Unidos no início dos anos 80, após a observação de um surto de pneumonia provocado pelo *Pneumocystis carini* e da ocorrência de um tipo raro de câncer vascular, conhecido como Sarcoma de Kaposi, entre homossexuais masculinos previamente saudáveis (GOTTLIEB *et al*, 1981). Estima-se, porém, através de estudos retrospectivos, que os primeiros casos de AIDS tenham provavelmente surgido bem antes dos casos descritos nos Estados Unidos. Estudos realizados em amostras retrospectivas de tecidos e sangue, congelados de pacientes que morreram com sintomas relacionados à AIDS, sugerem que provavelmente os primeiros casos de AIDS datam dos anos 60 (FROLAND *et al.*, 1988).

Atualmente, estima-se que cerca de 33,4 milhões de pessoas no mundo estejam vivendo com o HIV/AIDS e que destas, cerca 31,3 milhões sejam adultos e cerca de 2,1 milhões sejam crianças com menos de 15 anos. Em 2008, o número estimado de novas infecções foi de 2,7 milhões e o número de mortes chegou a 2 milhões em todo o mundo (Figura 1). Neste mesmo período, mais da metade dos casos de novas infecções no mundo, corresponderam a pessoas jovens entre 15 e 24 anos, com mais de 7.000 novos casos de infecção por dia. A situação se encontra bem pior na África Subsaariana, onde, no final de 2007, 12 milhões de crianças se encontravam órfãs da AIDS. (BOLETIM EPIDEMIOLÓGICO ONUSIDA/UNAIDS, 2007 disponível no site www.who.ch). A África é o continente mais afetado pela AIDS, especificamente a região Subsaariana é a região onde existe a maior prevalência. Nesta região, encontram-se cerca de 67% de todas as pessoas vivendo com AIDS e, em 2007, foi onde ocorreram cerca de 72% de todas as

mortes associadas à AIDS no mundo. Na América Latina estima-se que cerca de 140.000 pessoas tenham sido infectadas em 2007, elevando o número de pessoas vivendo com HIV para cerca de 1.8 milhões (BOLETIM EPIDEMIOLÓGICO ONUSIDA/UNAIDS, 2008).

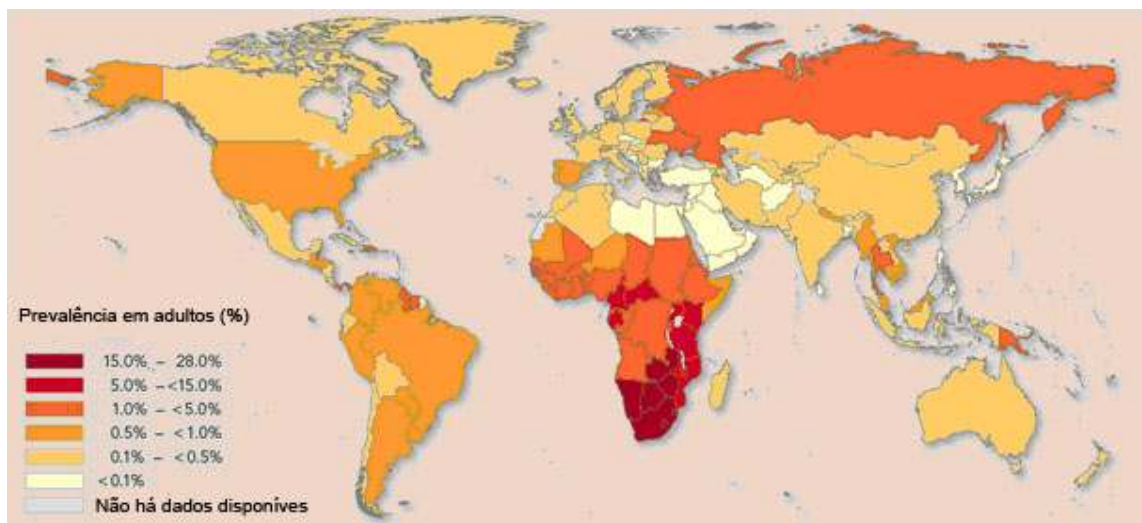


Figura 1. Epidemiologia do HIV (Adaptado de http://data.unaids.org/pub/GlobalReport/2008/GR08_2007_HIVPrevWallMap_GR08_en.jpg).

Durante duas décadas, a epidemia de AIDS se espalhou por todo o mundo, principalmente entre os países em desenvolvimento. Este fato também se observou no Brasil, onde as grandes metrópoles da região sudeste do Brasil posicionam-se entre as áreas mais afetadas pela epidemia da AIDS nas Américas (UNAIDS, 2008). As taxas de incidência de casos de AIDS no Brasil são diretamente proporcionais à distribuição populacional dos municípios, ou seja, quanto menor o tamanho das cidades, maior o aclave das taxas de incidência ao longo do tempo, mostrando que a epidemia teve início nos grandes centros urbanos, mas que são estes mesmos centros que detêm o menor aumento relativo (BOLETIM EPIDEMIOLÓGICO - DST/AIDS, 2008).

No Brasil, de acordo com dados do Ministério da Saúde, o número total de casos da doença no nosso país entre o ano de 1980 e junho de 2008 é de 506.499, distribuídos entre todas as regiões. Sendo 305.725 o número de casos na região Sudeste, 95.552 na região Sul, 58.348 na região Nordeste, 28.719 na região Centro Oeste e 18.155 na região

Norte (BOLETIM EPIDEMIOLÓGICO - DST/AIDS, 2008). Estes dados apontam que a epidemia da AIDS em algumas regiões esteja em um processo de estabilização, embora mantenha patamares elevados. O número de casos entre os homens demonstra esta tendência à estabilização, no entanto entre as mulheres, o número de casos é o maior desde o início da epidemia. A taxa de mortalidade também aponta para um quadro de estabilidade. Entre os usuários de drogas injetáveis, o número de casos de AIDS vem mantendo a tendência de queda observada nos últimos anos. (BOLETIM EPIDEMIOLÓGICO – DST/AIDS, 2004). Regionalmente, considerando a taxa de incidência da AIDS (número de casos/ 100 mil habitantes), o crescimento da epidemia é observado em todas as regiões do país, exceto no Sudeste, onde há estabilização com tendência de queda. (BOLETIM EPIDEMIOLÓGICO – DST/AIDS, 2008).

A análise recente dos números da epidemia de AIDS no Brasil, também chama a atenção pelo aumento progressivo da incidência na região sul (BOLETIM EPIDEMIOLÓGICO - DST/AIDS, 2009) assim como a feminilização, municipalização e pauperização da epidemia na maior parte das regiões (PARKER *et al.*, 1996; BASTOS *et al.*, 2000; CASTILHO *et al.*, 2000; EYER-SILVA *et al.*, 2006; SÁ *et al.*, 2007).

Em 1996, pela primeira vez em 16 anos de história da epidemia mundial, o número de novos casos de AIDS caiu devido ao sucesso nos esforços na prevenção e aos novos tratamentos antirretrovirais (BOLETIM EPIDEMIOLÓGICO -DST/AIDS 1996; MARINS *et al.*, 2003). A partir da introdução de novos inibidores da transcriptase reversa e principalmente dos inibidores da protease, culminando no desenvolvimento da Terapia Antirretroviral Altamente Ativa (HAART), as mortes pela doença foram reduzidas em até 20% no Brasil e no mundo (MARINS *et al.*, 2003).

1.2. O VÍRUS DA IMUNODEFICIÊNCIA HUMANA (HIV)

O vírus da imunodeficiência humana (HIV do inglês, Human Immunodeficiency vírus)), agente etiológico da AIDS, é um retrovírus da família *Retroviridae* do gênero *Lentivirus* (BARRÉ-SINOUSSE *et al.*, 1983) que utiliza como principal alvo para infecção e replicação, células que apresentam o receptor CD4 na superfície da membrana plasmática, principalmente linfócitos T CD4⁺, monócitos e macrófagos (COHEN *et al.*, 1998). A principal consequência da infecção por este vírus no indivíduo infectado é a depleção progressiva desta população linfocitária, fundamental para a resposta imunológica celular e humoral efetora, resultando no surgimento da AIDS e o desenvolvimento de doenças oportunistas e neoplasias (COHEN *et al.*, 2002).

Atualmente, o HIV é classificado em dois tipos: HIV tipo1 (HIV-1) e HIV tipo 2 (HIV-2) (PEETERS *et al.*, 1989; GAO *et al.*, 1999; JOHNSON *et al.*, 1992; LEITNER *et al.*, 2005). Essa definição se baseia principalmente em diferenças na sua estrutura genômica, que podem ser superiores a 60%, bem como, pela sua relação com demais lentivírus de primatas (LEITNER *et al.*, 2005). O HIV-1 foi identificado em 1983, dois anos depois da descrição dos primeiros casos de AIDS nos EUA e Europa (GOTTLIEB *et al.*, 1981), a partir da cultura de células de linfonodo de um paciente apresentando síndrome de linfadenopatia persistente (BARRÉ-SINOUSSE *et al.*, 1983). O HIV-1 está distribuído por todo o mundo e atualmente é o principal responsável pela pandemia de AIDS (BOLETIM EPIDEMIOLÓGICO ONUSIDA/UNAIDS). O HIV-2 foi identificado em 1986 (CLAVEL *et al.*, 1987) em países do oeste africano onde se concentram a maior parte dos indivíduos infectados, além de casos esporádicos na Europa (Portugal, França) e Ásia (Índia e China) (HU *et al.*, 1996). Os dois retrovírus apresentam as mesmas vias de infecção, contudo estudos mostram potenciais de patogenicidade diferenciados (ADAM MACNEIL *et al.*, 2007), onde o HIV-1 apresenta altas taxas de viremia e transmissão, enquanto o

HIV-2 tem uma evolução mais lenta para a AIDS (MARLINK *et al.*, 1994).

A estrutura da partícula viral madura (vírion) apresenta formato esférico, mede aproximadamente 100 nanômetros de diâmetro, contendo um nucleocapsídeo excêntrico sob a forma de bastão (Figura 2). Conforme os demais retrovírus, o HIV é um vírus envelopado, seu capsídeo é constituído de proteínas internas e nele está contido seu material genético composto de duas fitas de RNA idênticas de polaridade positiva. Também, as enzimas virais necessárias a replicação do vírus: a transcriptase reversa (TR), a protease (PR) e a integrase (INT). Associadas ao material genômico do HIV-1 organizam-se monômeros de duas proteínas do nucleocapsídeo, as proteínas p7 e p9, necessárias à estabilidade da partícula infecciosa.

O envelope viral é constituído por uma bicamada fosfolipídica de origem celular e pelas glicoproteínas gp120 e gp41, responsáveis pela entrada do HIV na célula hospedeira (FREED & MARTIN, 1994). Logo abaixo do envelope viral está a matriz, formada pela proteína p17 que possui uma alta afinidade com a membrana plasmática e o capsídeo da partícula viral formado pela proteína p24, um importante antígeno viral (GOTLINGER *et al.*, 2001). Envolvendo este complexo ribonucléico encontramos a proteína p24 responsável pela formação do capsídeo e que constitui o core cilíndrico da partícula viral madura (Figura 2).

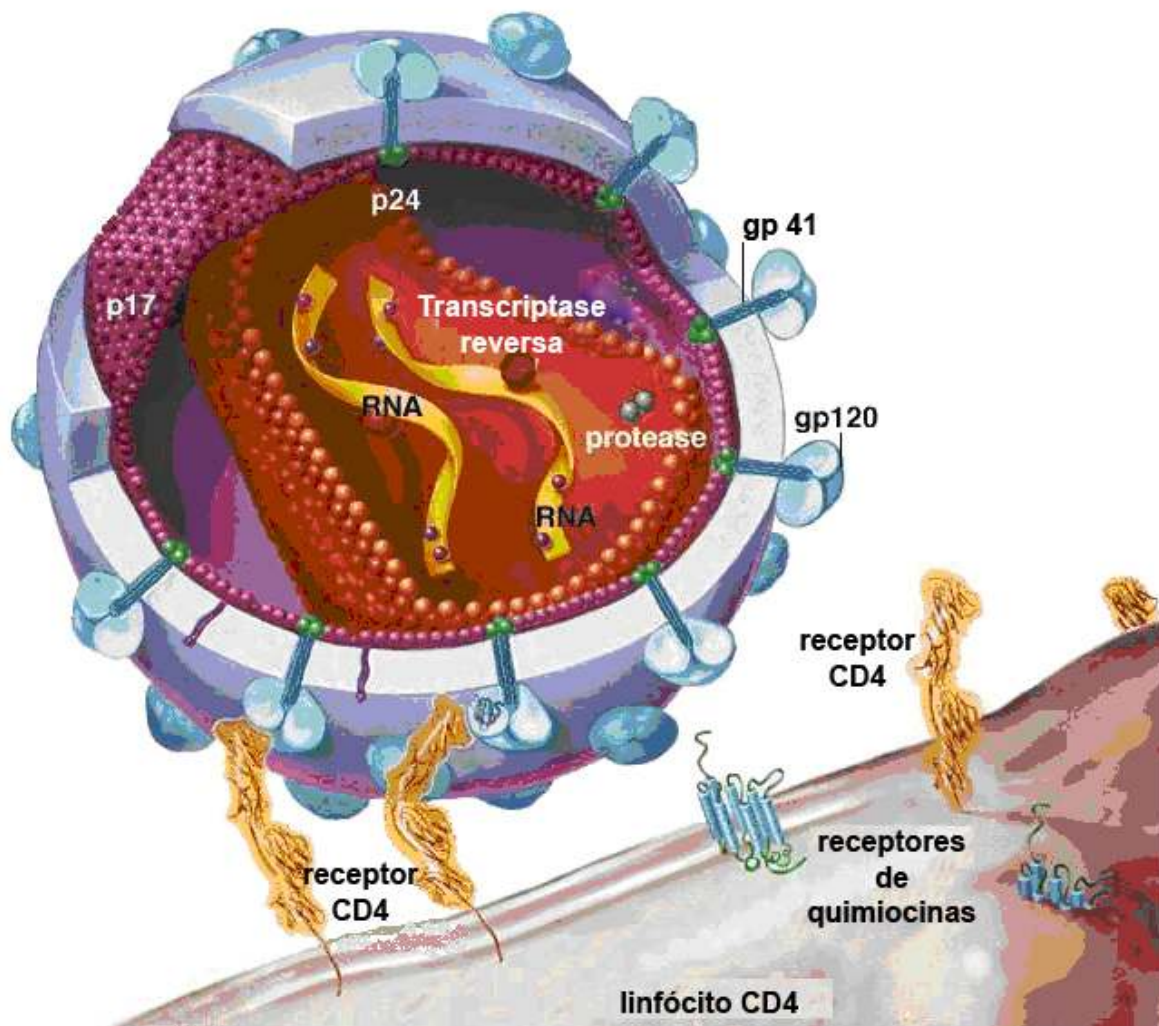


Figura 2. Estrutura morfológica do HIV-1 (Adaptado de <http://www.critpath.org/aric/library/img005.htm>).

O genoma do HIV-1 (**Figura 3**) é constituído por duas fitas de RNA, cadeia simples, com polaridade positiva. Contém aproximadamente 9.800 nucleotídeos (9,8 Kb), e é formado por três genes estruturais (*gag*, *env*, *pol*), além de seis genes acessórios e regulatórios da replicação viral (*rev*, *tat*, *vif*, *vpx*, *vpr*, *nef*). O genoma viral é flanqueado pelas seqüências LTR (long terminal repeat), localizadas nas extremidades (3', 5') do genoma pro-viral, que têm um papel importante na integração do provírus ao DNA da célula hospedeira.

O gene *env* codifica uma glicoproteína precursora (160 kDa) que gera, após clivada pelas proteases celulares, a gp120 e a gp41, que são glicoproteínas de superfície,

responsáveis respectivamente, pela ligação ao receptor CD4, aos co-receptores CCR5 e/ou CXCR4 e posterior fusão de membranas viral e celular para entrada do vírus na célula hospedeira (FREED & MARTIN, 1994).

O gene *gag* codifica uma proteína de 55 kDa (p55) que é clivada pela protease viral nas proteínas p24 (capsídeo), p17 (matriz), p7 (nucleocapsídeo) e p6 (núcleo) (WELKER *et al.*, 2000; STRACK *et al.*, 2003).

O gene *pol* é responsável pela codificação de uma cadeia protéica precursora que após ser clivada pela protease, dá origem as seguintes enzimas: transcriptase reversa (66 Kb) que possui atividade de polimerase e RNase H, realiza a retrotranscrição do RNA do genoma viral em DNA complementar ou provírus; protease (11 Kb), que media a clivagem do precursor *gag/pol* e outros precursores que levarão a maturação viral e a integrase (34 Kb), que promove a integração do DNA proviral ao genoma da célula hospedeira.

Existem também os genes estruturais *tat* e *rev*, que codificam proteínas regulatórias e os genes *vif*, *vpx*, *vpr*, *nef*, e *vpu* responsáveis pela codificação de proteínas acessórias que desempenham um papel importante nas diferentes etapas da replicação do HIV (**Figura 3**). O HIV-2 possui o gene *vpx* no lugar ao que corresponderia ao gene *vpu* do HIV-1). (FIELDS VIROLOGY, 2007)

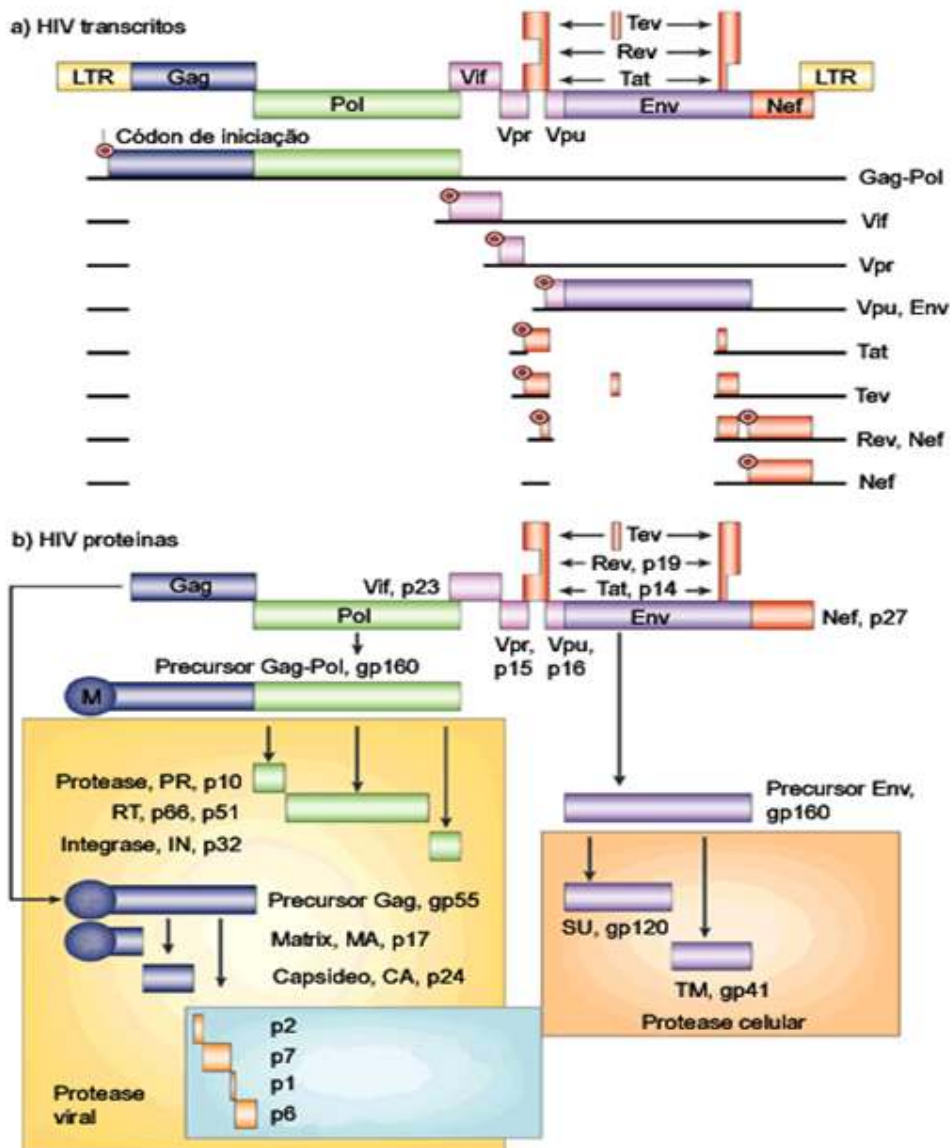


Figura 3. Organização genômica do HIV-1: genes estruturais (*gag*, *env*, *pol*), genes acessórios (*rev*, *tat*, *vif*, *vpx*, *vpu*, *nef*), seqüências LTR (long terminal repeat) que têm um papel importante na integração do provírus ao DNA da célula hospedeira, o gene *env* codifica as glicoproteínas de superfície gp120 e a gp41, *gag* codifica as proteínas p24 (capsídeo), p17 (matriz), p7 (nucleocapsídeo) e p6 (núcleo) e o gene *pol* codificador das enzimas transcriptase reversa, protease e a integrase (Retirado de Uhaweb.hartford.edu/alosisi.WEBPAGES/auct_140/HIV.HTM).

1.2.1. CICLO DE REPLICAÇÃO DO HIV

O primeiro passo para o início da infecção é a ligação da partícula viral a receptores específicos na superfície da célula alvo. O HIV começa sua infecção ligando a glicoproteína de superfície gp120 ao receptor CD4 da célula hospedeira. A molécula de CD4 faz parte da superfamília de imunoglobulinas, sendo um receptor de superfície celular expresso em uma das subpopulações de linfócitos T, em monócitos, e em outras células como macrófagos, células da microglia e células dendríticas. Dois co-receptores principais foram também identificados, como os principais componentes necessários para que o HIV se incorpore à célula: o primeiro foi o CXCR4, que é utilizado pelos vírus linfotrópicos e o segundo, o CCR5 utilizado pelos vírus macrofagotrópicos (Revisto por FRANKEL *et al.*, 1998). Este reconhecimento de co-receptores e de como o HIV se funde à célula abriu novas possibilidades para o estudo com drogas antivirais. Uma variedade de novos agentes terapêuticos está sendo desenvolvida para impedir a infecção obstruindo a fusão do HIV com a célula hospedeira (HIV MEDICINE 2007).

Posteriormente, mudanças estruturais ocorrem nesta interação e a glicoproteína do envelope gp41 leva a fusão das membranas viral e celular possibilitando a introdução do capsídeo na célula CD4. Uma vez que o HIV penetrou na célula, seu material genético (RNA) é liberado no citoplasma da célula hospedeira. O RNA viral é convertido em cDNA, processo chamado de transcrição reversa, realizado pela enzima transcriptase reversa. Seguindo-se a transcrição reversa, o cDNA é rapidamente transportado para o núcleo da célula hospedeira. A ação da integrase, resulta na integração estável do cDNA do genoma viral no DNA cromossômico sendo denominado pró-vírus. O conjunto de RNAs transcritos são então transportados para o citoplasma, onde serão então traduzidos, e constituirão novas partículas virais (HOPE, 1997).

Os vírions são inicialmente montados próximos à membrana celular na forma de partículas imaturas, compostas de um envelope glicoprotéico, RNA genômico e poliproteínas virais. Após ou durante o brotamento, as partículas virais passam por uma modificação morfológica conhecida como maturação, que consiste na clivagem de poliproteínas gag e gag-pol pela protease viral, produzindo as enzimas de replicação e proteínas estruturais do capsídeo (Figura 4). O processamento das poliproteínas no vírion completa o ciclo de replicação do HIV, os vírions maduros são então capazes de infectar outras células (FIELDS VIROLOGY 2007)

A história natural da infecção pelo HIV-1 começou a ser alterada a partir da segunda metade da década de 90 com o surgimento da terapia antirretroviral altamente ativa HAART, revertendo quadros clínicos sintomáticos e o estado de imunossupressão observados em pacientes com AIDS. Nos indivíduos assintomáticos, a HAART pode impedir a progressão da doença, aumentando com isso a sobrevida destes pacientes e reduzindo as taxas de morbidade e mortalidade associadas a esta infecção no mundo (OSMANOV *et al.*, 2002) e no Brasil (MARINS *et al.*, 2003).

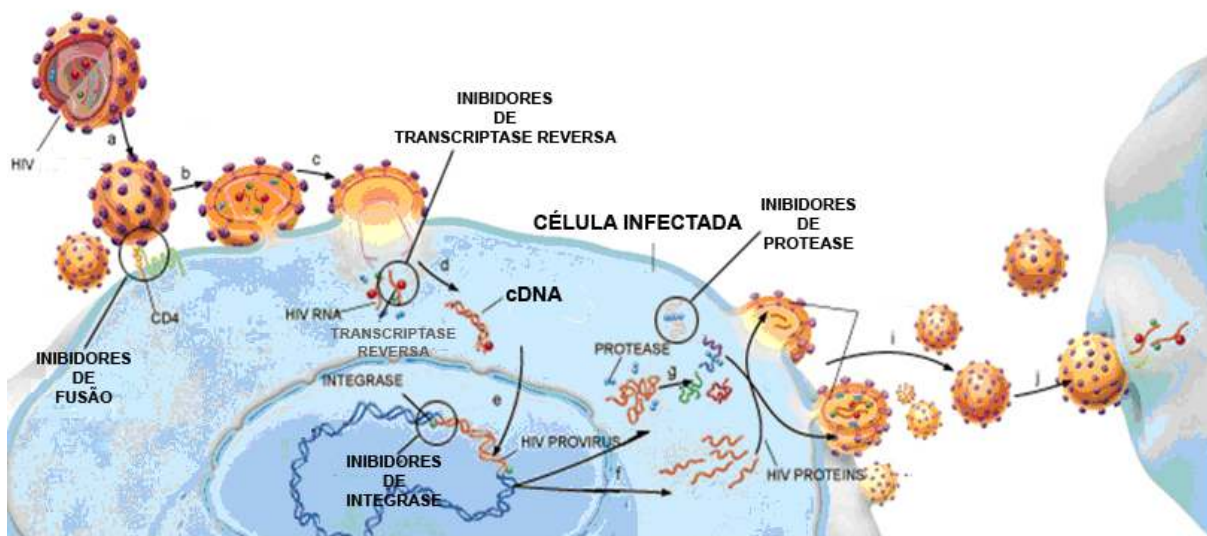


Figura 4: Representação das principais etapas do ciclo de replicação do HIV-1, indicando os principais alvos dos antirretrovirais. Ciclo biológico do HIV-1 (Adaptado de [http://www.lia.ufsc.br/JAF1a.ppt#283,15,Ciclo do HIV](http://www.lia.ufsc.br/JAF1a.ppt#283,15,Ciclo%20do%20HIV)).

1.2.2. CLASSIFICAÇÃO DO HIV

As análises filogenéticas de numerosas amostras de HIV-1 provenientes de diversas localidades geográficas permitiram estabelecer a classificação do HIV em tipos, grupos, subtipos, sub-subtipos e formas recombinantes circulantes (CRFs- circulating recombinant forms) [LEITNER *et al.*, LOS ALAMOS, 2005).

Os dois tipos de vírus da AIDS, o HIV-1 e o HIV-2 são diferenciados com base na sua organização genômica e nas relações filogenéticas entre si, e entre os outros retrovírus de primatas. Para o HIV-1 foram descritos três grupos distintos: M (major), O (outlier) (JAFFE *et al.*, 1998) e N (new) (SIMON *et al.*, 1998). Destes, o mais prevalente é o grupo M, responsável pela atual pandemia, que ainda é subdividido em pelo menos nove subtipos classificados de A a D, F a H, J e K, além de 46 CRFs. Atualmente, os subtipos F e A encontram-se subdivididos em dois e 4 sub-subtipos, respectivamente F1 e F2 (TRIQUES *et al.*, 2000) e A1 a A4 (GAO *et al.*, 2001).

A variabilidade genética que distancia as diferentes amostras do grupo M entre si pode chegar à cerca de 30% para as seqüências do gene *env*, cerca de 15% para as seqüências do gene *gag* (TAKEBE, 2004), e 5% para as seqüências do gene *pol* (LOS ALAMOS, 2005). Já para os grupos de HIV-1 esta diferença genética pode chegar a 35% na sua estrutura.

O surgimento de formas recombinantes pode ser considerado uma propriedade fundamental dos retrovírus em razão da natureza diplóide de seu genoma de RNA e da possibilidade da transcriptase reversa atuar numa ou noutra fita durante a síntese do cDNA. Conseqüentemente, de uma célula duplamente infectada pode sair um vírus que possua uma fita de RNA de cada vírus de subtipos diferentes gerando um recombinante. Até o momento foram descritas 43 CRFs no mundo (LEITNER, LOS ALAMOS, 2005) e recentemente, foram identificadas no Brasil duas novas formas recombinantes, a

CRF39_BF, e CRF40_BF, apresentando seqüências dos subtipos B e F (GUIMARÃES *et al.*, 2008). A recombinação genética do HIV tem sido descrita em todos os níveis filogenéticos, entre lentivirus de primatas, entre tipos, grupos, subtipos e subsubtipos do HIV. Esta recombinação tem implicações importantes para o entendimento da epidemia do HIV, na patogenia da doença, na resistência às drogas e no desenvolvimento de vacinas (PEETERS *et al.*, 2000; LOS ALAMOS 2007).

O HIV-2 possui aproximadamente 50% de identidade, entre as seqüências de aminoácidos, com o HIV-1. O HIV-2 pode ser dividido em 8 subtipos de A – H (Damond e cols, 2004).

No Brasil a distribuição dos subtipos é complexa quando comparada com outros países da América do Sul. O principal subtipo circulante é o subtipo B, embora resultados de vários estudos realizados no país mostram a presença dos subtipos C, principalmente no Sul do país (WHO, 1994; LOUREIRO, 1998; STEFANI *et al.*, 2000), subtipo F e o segundo mais prevalente (MORGADO *et al.*, 1994; COUTO-FERNANDEZ *et al.*, 1999; VICENTE *et al.*, 2000; STEFANI *et al.*, 2000), seguido pelo número crescente de genomas recombinantes B/F (SABINO *et al.*, 1994, BRINDEIRO *et al.*, 2003; COUTO-FERNANDEZ *et al.*, 2005). Também foram, identificado casos esporádicos da circulação do subtipo D (MORGADO *et al.*, 1998; COUTO-FERNANDEZ *et al.*, 2006), subtipo A (CARIDE *et al.*, 2000) e formas recombinantes B/C no sul do Brasil (CORNELISSEN *et al.*, 1996) (Revisto por MORGADO *et al.*, 2002). Recentemente, houve a primeira descrição da circulação no Brasil da forma recombinante prevalente na África Ocidental CRF2_AG ocorrida no Rio de Janeiro (COUTO-FERNANDEZ *et al.*, 2005).

1.3. TERAPIA ANTIRRETROVIRAL

A terapia antirretroviral (TARV) tem por objetivo principal a melhoria da qualidade de vida dos infectados pelo HIV através da supressão da replicação viral, pela atuação de drogas inibidoras específicas em diferentes etapas do ciclo replicativo do HIV-1. A supressão da progressão viral deve ser duradoura, objetivando a não progressão da imunodeficiência do paciente, para reequilíbrio de seu sistema imunológico. Atualmente, vinte e cinco compostos estão aprovados pelo FDA (*US Food and Drug Administration*) para uso em humanos, e se dividem em seis categorias, de acordo com a sua composição química e etapa de intervenção no ciclo de replicação do HIV-1. Os inibidores da transcriptase reversa (TR) são divididos em nucleosídeos/nucleotídeos (INTRs) e inibidores não-nucleosídeos (INNTRs). Os inibidores da protease (IPs), inibidores de fusão (IFs), inibidores de entrada ou de correceptor e inibidor de integrase (DE CLERCQ, 2009). Os inibidores nucleosídeos da TR (INTRs) (Tabela 1) foram os primeiros antirretrovirais desenvolvidos para o tratamento contra o HIV/AIDS, sendo o AZT (azidotimidina) o primeiro a ser aprovado em 1987 (WARNKE *et al.*, 2007). Ainda hoje o uso de INTRs tem posição de destaque na TARV, pela ação que exerce no bloqueio da síntese de DNA proviral, através da incorporação de análogos nucleosídeos e/ou nucleotídicos durante o processo de transcrição reversa do RNA viral (YARCHOAN *et al.*, 1986). Os inibidores não-nucleosídicos da TR (INNTR) (Tabela 2) atuam ligando-se a uma região próxima do sítio ativo da TR, levando a alterações estruturais no sítio ativo da enzima e a consequente inibição da atividade de polimerase.

Inibidores Nucleosídeos da transcriptase reversa (NRTIs)	
Nome genérico	Abreviatura
Emtricitabina	FTC
Lamivudina	3TC
Zidovudina	AZT
Didanosina	DDI
Tenofovir	TDF
Estavudina	d4T
Abacavir	ABC

Tabela 1. Antirretrovirais inibidores nucleosídeos da transcriptase reversa aprovados pelo FDA para o uso na terapia anti-HIV/AIDS. Adaptado de (<http://hivdb.stanford.edu/pages/asi/releaseNotes/>).

Inibidores não-nucleosídeos da transcriptase reversa (NNRTIs)	
Nome genérico	Abreviatura
Delavirdina	DLV
Efavirenz	EFV
Etravirina	ETR
Nevirapina	NVP

Tabela 2. Antirretrovirais inibidores não-nucleosídeos da transcriptase reversa aprovados pelo FDA para o uso na terapia anti-HIV/AIDS. Adaptado de (<http://hivdb.stanford.edu/pages/asi/releaseNotes/>).

Os inibidores de protease atuam nas últimas etapas do ciclo replicativo do HIV, se ligando ao sítio ativo da PR viral e inibindo a atividade proteolítica da enzima. Com isso, não haverá a clivagem das poliproteínas precursoras gag e gag-pol e formação de novas partículas virais maduras com capacidade infectante (DE CLERCQ, 2002). Atualmente, existem pelo menos 8 inibidores de protease sintetizados a partir de epítomos que compõem o sitio ativo da enzima (Tabela 1).

Inibidores da protease (IPs)	
Nome genérico	Abreviatura
tipranavir/r	TPV/r
Indinavir/r	IDV/r
saquinavir/r	SQV/r
lopinavir/r	LPV/r
fosamprenavir/r	FPV/r
atazanavir/r	ATV/r
Nelfinavir	NFV
darunavir/r	DRV/r

Tabela 3. Antirretrovirais inibidores da protease aprovados pelo FDA para o uso na terapia anti-HIV/AIDS. Adaptado de (<http://hivdb.stanford.edu/pages/asi/releaseNotes/>).

Os inibidores de entrada e fusão atuam nas etapas iniciais do ciclo de replicação viral, agindo respectivamente, na ligação da glicoproteína gp120 ao coreceptor celular CCR5 e na ligação ao domínio HR1 da glicoproteína gp41. Os inibidores de correceptores (Maraviroc e Vicriviroc) são os únicos compostos disponíveis nesta classe de medicamentos e atuam como antagonistas aos receptores de quimiocina CCR5 presente em macrófagos, impedindo a perfeita interação vírus-célula. A enfuvirtida (T20) é a única droga disponível desta classe e atua se ligando ao domínio HR1 da gp41, bloqueando a formação da estrutura de hélices necessária à aproximação e fusão das membranas viral e celular (WARNKE *et al.*, 2007).

Os inibidores da integrase (Tabela 4) representam um novo alvo para o bloqueio do ciclo replicativo viral e atuam impedindo o transporte e integração do genoma pró-viral ao da célula eucariota. Atualmente, o raltegravir é a única droga licenciada no tratamento anti-HIV/AIDS e normalmente é utilizado em resgate terapêutico de pacientes que

apresentam resistência aos demais antirretrovirais utilizados em TARV de primeira e segunda linhas (CONSENSO BRASILEIRO PARA TERAPIA ANTIRRETROVIRAL, 2009).

Os inibidores de integrase (Inis)	
Nome genérico	Abreviatura
Raltegravir	RAL
Elvitegravir	EVG

Tabela 4. Antirretrovirais inibidores da integrase aprovados pelo FDA para o uso na terapia anti-HIV/AIDS. Adaptado de (<http://hivdb.stanford.edu/pages/asi/releaseNotes/>).

Buscando a supressão máxima dos níveis de replicação do HIV-1 no indivíduo infectado, em 1996 se estabeleceu a terapia antirretroviral altamente ativa (HAART). A HAART consiste em uma combinação de pelo menos duas drogas da classe dos INTRs, um da classe dos INNTR e/ou um inibidor de protease. A introdução da HAART no atual tratamento antiHIV/AIDS trouxe um expressivo aumento de sobrevida e melhora na qualidade de vida de pacientes infectados pelo HIV (PERELSON *et al.*, 1996; SEPKOWITZ, 2001; MARINS *et al.*, 2003).

No Brasil, a aTARV teve início em 1991 com a introdução do AZT através do Sistema Único de Saúde (SUS). Em 1996, foi introduzida a HAART combinada atendendo gratuitamente a grande maioria dos pacientes com HIV/AIDS, sendo administrada segundo recomendações técnicas definidas pelo Comitê Técnico Assessor para TARV (CONSENSO BRASILEIRO PARA TERAPIA ANTIRRETROVIRAL, 2009). A política de acesso universal ao tratamento anti-HIV/AIDS promovido pelo Ministério da Saúde através do Departamento de DST, AIDS e Hepatites Virais se consolidou ao longo dos anos e atualmente, é reconhecida internacionalmente como uma das mais abrangentes e

completas do mundo. Aproximadamente, 200.000 pacientes estão em TARV no Brasil, o que tem contribuído para a melhoria de qualidade de vida destes pacientes, além da redução significativa dos níveis de morbidade e mortalidade (MARINS, 2003).

1.4. RESISTÊNCIA AOS ANTIRRETROVIRAIS

Apesar dos benefícios obtidos pela TARV nos últimos anos, o tratamento não impede totalmente o surgimento de formas virais resistentes às drogas (WINTERS, 1998). Isto deve, entre outros fatores, à falta de atividade de correção de incorporação de nucleotídeos pela TR. Associado a uma elevada taxa mutacional – cerca de $3,4 \times 10^{-5}$ mutações por sítio nucleotídico (MANSKY & TEMIN, 1995), uma alta taxa de replicação do vírus (MACNEIL *et al.*, 2007) e uma grande quantidade de partículas geradas diariamente no indivíduo infectado – cerca de 10^9 a 10^{10} partículas. Portanto, praticamente uma mutação pode ser formada no genoma viral a cada ciclo replicativo (PELERSON *et al.*, 1996).

O desenvolvimento de resistência do HIV aos antirretrovirais é um fator limitante para o sucesso da TARV. Os portadores de vírus resistentes, além de não responderem adequadamente ao tratamento, podem também transmitir estes vírus mutantes, representando grave problema de saúde pública (INOCENCIO *et al.*, 2009). A falha terapêutica é um evento multifatorial, que além do surgimento de vírus com mutações de resistência, pode ocorrer também por disfunção do sistema imune, falta de adesão do paciente ao tratamento, concentrações inibitórias das drogas sub-ótimas e deficiência na absorção do medicamento pelo paciente (DURANT *et al.*, 1999; SCHAFFER, 2002). Além disso, a presença do HIV em santuários no organismo dificulta o acesso do antirretroviral

fazendo com que a droga atue precariamente contra o HIV.

O surgimento de mutações de resistência no genoma viral leva a uma queda na susceptibilidade aos antirretrovirais, reduzindo progressivamente a potência do esquema terapêutico. Com isso, o HIV continua a replicar na presença da TARV, selecionando vírus com mutações de resistência as diferentes drogas. Este acúmulo progressivo de mutações de resistência, pode acarretar resistência cruzada à outras drogas da mesma classe ao longo do tempo (SCHAFER, 2002).

As mutações de resistência do HIV são classificadas em mutações primárias (ou principal) e secundárias (ou acessórias, ou compensatórias). As mutações primárias reduzem significativamente a susceptibilidade às drogas pela qual foi selecionada. Embora contribuam para a adaptabilidade viral quando na presença da TARV, geralmente levam uma redução na capacidade replicativa (“fitness” viral). As mutações secundárias são selecionadas normalmente de modo a compensar a capacidade replicativa prejudicada. Contudo, podem contribuir para o aumento da resistência, quando em combinação com mutações primárias (DIAZ *et al.*, 2008). Normalmente, estas mutações têm pouca ou nenhuma influência sobre o vírus na ausência de mutações primárias (RIDKY & LEIS, 1995).

O acúmulo de mutações de resistência aos antirretrovirais representa um desafio importante na melhoria do tratamento de pacientes com vírus multirresistentes. O desenvolvimento de resistência à várias classes de antirretrovirais, limita o número de novos regimes terapêuticos e o sucesso virológico de terapias de resgate subsequentes.

Mais recentemente, testes avaliação de resistência ou de susceptibilidade do HIV às drogas foram incorporados ao monitoramento e tratamento de pacientes com HIV/AIDS (DURANT *et al.*, 1999). Estes testes podem ajudar a determinar qual droga poderá ser eficaz, ou que apresente alguma atividade residual, facilitando a adoção de regimes terapêuticos mais eficazes, diminuindo também custos e toxicidade da terapia.

Atualmente, dois tipos de métodos estão estabelecidos para avaliação da resistência ou sensibilidade do HIV aos antirretrovirais: os testes fenotípicos e genotípicos de resistência (WILSON, 2003).

Os testes fenotípicos de resistência quantificam *in vitro* a susceptibilidade do vírus aos antirretrovirais, ou seja, determinam a quantidade de droga necessária para inibir a replicação do HIV-1. A replicação viral é medida em culturas celulares submetidas a pressão selectiva de concentrações crescentes da droga, comparativamente à replicação de vírus selvagens. As concentrações inibitórias das drogas são expressas em valores de CI_{50} (concentração inibitória de 50%), ou seja, a concentração de droga necessária para inibir a replicação viral em 50%.

Os testes de genotipagem se baseiam na análise de mutações associadas à resistência presentes no genoma viral. Basicamente, as mutações são identificadas através do sequenciamento molecular direto de determinadas regiões do genoma do HIV, codificadoras de proteínas e enzimas alvo dos diferentes antirretrovirais. Os fragmentos genômicos analisados são gerados a partir da retrotranscrição do RNA viral purificado do plasma do paciente e amplificado molecularmente pela técnica de reação de polimerase em cadeia (PCR), específico para os genes da TR e protease do HIV-1. Os testes de genotipagem fornecem uma medida indireta da resistência aos antirretrovirais e detectam mutantes que compreendam pelo menos 20 a 30% do total da população viral no indivíduo infectado. Atualmente, um grande número de mutações de resistência associadas à queda de susceptibilidade da maioria das drogas foram descritas (Stanford). A análise dos padrões genotípicos de resistência é baseada na correlação entre o genótipo e o fenótipo. São analisados dados disponíveis de estudos *in vitro*, observações clínicas e testes combinados, onde mutações genotípicas são analisadas conjuntamente com dados de resistência fenotípica. Contudo, o elevado número de mutações e de diferentes padrões de resistência, que podem também conter mutações principais e

secundárias, torna difícil a determinação do grau de resistência para alguns antirretrovirais. As análises de resistência do HIV aos antirretrovirais, avançaram muito nos últimos anos com o desenvolvimento dos algoritmos para avaliação de resistência. Um algoritmo é um conjunto de procedimentos pré-definidos, com início, meio e fim que devem ser adotados para expressar a execução de uma determinada tarefa (TAVARES, 1998). Os algoritmos de resistência são dinâmicos e suas regras podem ser atualizadas à medida que novas pesquisas venham a indicar mudanças do perfil de resistência do HIV. Porém, os sistemas computadorizados que interpretam os algoritmos são estruturados para realizar a interpretação do algoritmo independente de suas variações, deixando desta forma o pesquisador livre para realizar as mudanças nos algoritmos sem se preocupar com a programação do sistema.

1.5. SISTEMAS DE INTERPRETAÇÃO DE ALGORITMOS DE RESISTÊNCIA

Um algoritmo é um conjunto de procedimentos pré-definidos, com início, meio e fim que devem ser adotados para expressar a execução de uma determinada tarefa. Dentro do contexto da resistência do HIV aos antirretrovirais, os algoritmos funcionam como um conjunto de procedimentos que, baseados na literatura e em ensaios clínicos, atribuem valores matemáticos a cada mutação observada em sequências de aminoácidos do HIV e a partir do conjunto das mutações, atribui um valor final para a resistência a cada droga antirretroviral.

Atualmente, uma série de sistemas de interpretação de resistência do HIV aos antirretrovirais, utilizando algoritmos, estão disponíveis gratuitamente na Internet (HIV medicine, 2008). Estes sistemas identificam as mutações associadas à resistência viral e, baseados nos algoritmos fundamentados em diferentes lógicas de interpretação,

analisam as sequências virais normalmente submetidas em formato fasta e retornam o perfil de resistência e sensibilidade as drogas de cada sequência genômica do HIV. Estes sistemas são de grande importância para a realização desse tipo de análise e funcionam como ferramentas de bioinformática para utilização, por médicos e virologistas capacitados a interpretar os complexos perfis de resistência as drogas e tomar as decisões necessárias à estruturação do tratamento do paciente. Entre os sistemas disponíveis podemos citar como exemplo o HIV Drug Resistance Database (<http://hivdb.stanford.edu/index.html>) da Universidade de Stanford que utiliza um modelo matemático para atribuir pontuação e/ou penalidades as diferentes mutações na sequência de aminoácidos. Este processo avalia o efeito que cada mutação confere a sequência analisada, em relação ao tipo selvagem do HIV. O sistema de pontuação é dividido em cinco faixas de resistência, estabelecendo a faixa entre 0 e 9 como suscetíveis, entre 10 e 14 como baixo nível de resistência, entre 15 e 29 com um baixo nível de resistência porém com indicativo de uma exposição anterior a classe da droga antirretroviral, entre 30 e 59 um grau de resistência intermediário entre a baixa e a alta resistência e na faixa igual ou maior que 60, um alto grau de resistência, em geral, com pouca ou nenhuma resposta virológica ao tratamento com a droga (RHEE *et al.*, 2003; SHAFER *et al.*, 2000; SHAFER, 2006; LIU, 2006; RHEE *et al.*, 2006; GIFFORD *et al.*, 2009).

Como exemplo de sistema de análise de resistência, podemos citar também o RegaDB (<http://www.rega.kuleuven.be/cev/regadb/index.php?id=2>), que é um banco de dados com diversas ferramentas de bioinformática, que podem ser adquiridas e instaladas em um computador local. Com esta ferramenta é possível armazenar dados clínicos relacionados ao HIV. Após a instalação, diversos serviços são oferecidos pelo RegaDB. Dentre estes serviços está Drug Resistance interpretation, que é uma ferramenta baseada em vários algoritmos e possui suporte para o formato de arquivo XML. O RegaDB é uma ferramenta desenvolvida e uma colaboração conjunta entre o Instituto de Rega,

Katholieke Universiteit Leuven e MyBioData Biomédica IT Solutions (LIBIN *et al.*, 2007; LIBIN *et al.*, 2008; IMBRECHTS *et al.*, 2009).

Outro exemplo de sistema de avaliação de resistência do HIV é o algoritmo Brasileiro de interpretação disponibilizado aos Laboratórios da Rede Nacional de Genotipagem (RENAGENO) (http://forrest.ime.usp.br:3001/public_analyses) e também através da digitação das diferentes mutações pelo médico. Este sistema utiliza uma regra de combinação das diferentes mutações para avaliação dos níveis de resistência. A regra foi estabelecida e é permanentemente atualizada por grupo técnico assessor que incorpora achados relevantes de descrições de mutações associadas as drogas). O sistema RENAGENO considera três faixas de interpretação da resistência: suscetível, intermediária e resistente. O software é integrado por toda Rede Pública de Saúde em nível federal e é apoiado pelo PNDS/AIDS, que requereu o desenvolvimento do programa em 2004 (www.aids.gov.br/algoritmo). As regras para a composição do algoritmo brasileiro de interpretação de reistencia do HIV-1 para para rede nacional de genotipagem (RENAGENO), são apresentadas na forma em uma tabela e disponível no site do Depto. De DST, AIDS e Hepatites Virais do Ministério da Saúde (www.aids.gov.br/algoritmo) (. Na primeira coluna da tabela encontram-se listadas as siglas dos antirretrovirais:ABC (Abacavir), DDI (Didanosine), 3TC (Lamivudine), FTC(Emtricitabine), d4T (Stavudine), TDF (Tenofovir), AZT (Zidovudine), EFV (Efavirenz), NVP (Nevirapine), FPV (Fosamprenavir), IDV (Indinavir), LPV (Lopinavir), NFV (Nelfinavir), RTV (Ritonavir), SQV (Saquinavir), ATV (Atazanavir), DRV (Darunavir), ETR (Etravirine), TPV (Tipranavir); na segunda coluna o número de combinações da regra correspondente a cada antirretroviral; na terceira coluna encontram-se as combinações das diferentes mutações de resistencia frente a cada droga constituindo a lógica de cada regra. ; na quarta coluna estão listadas as letra que representam os perfis de susceptibilidade (S) aos antirretrovirais, resistência intermediária (I) e resistência completa (R), representando a interpretação das regras da terceira coluna.

Atualmente, cerca de 200 mil pacientes recebem medicamentos antirretrovirais

pelo Ministério da Saúde. (BOLETIM EPIDEMIOLÓGICO DST e AIDS/MS 2009).

Outro importante sistema de interpretação de resistência aos antirretrovirais é o algoritmo francês HIV Drug Resistance Genotypic Interpretation (<http://pugliese.p.free.fr/index.htm>), desenvolvido pelo grupo francês “AC11: Grupo de Resistência” da Agência Nacional para Pesquisa da AIDS. Este sistema utiliza também um sistema de pontuação que considera três níveis de resistência: “nenhuma prova de resistência”, “possível resistência” e “resistência”

(<http://www.hivfrenchresistance.org/2009/Algo-2009.pdf>) (PRADIER *et al.*, 1998).

Existem outros importantes sites de avaliação de resistência que fornecem a possibilidade de interface entre outros algoritmos (http://resdb.lanl.gov/Resist_DB). Mais recentemente, alguns sites também disponibilizam informações de avaliação fenotípica de sensibilidade às drogas, através de informações pareadas genotípicas e fenotípicas disponíveis nestes bancos de dados (<http://www.geno2pheno.org>, <http://hivdb.stanford.edu>)

RACIONAL

2. RACIONAL

As terapias com antirretrovirais são amplamente utilizadas na supressão da replicação do HIV-1 no indivíduo infectado e têm como principal consequência a diminuição das taxas de mortalidade, morbidade e hospitalização, proporcionando grandes benefícios aos pacientes infectados pelo HIV-1 (SEPKOWITZ *et al.*, 2001; MARINS *et al.*, 2003). Porém, tais terapias trazem, também, problemas como desenvolvimento de resistência, toxicidade, complexidade terapêutica e alto custo (POMERANTZ *et al.*, 1995). Para minimizar os problemas causados pelas terapias, é de suma importância uma escolha cautelosa da terapia inicial que deve ser baseada tanto na avaliação clínica, como nos parâmetros laboratoriais (carga viral e contagem de células T-CD4⁺), que vão determinar o grau de imunodeficiência já existente e o risco de progressão (NIES-KRASKE *et al.*, 2009).

A identificação prévia das mutações e do grau de resistência que elas conferem a determinado antirretroviral é de fundamental importância no tratamento do paciente infectado com o HIV-1 (DURANT *et al.*, 1999). A detecção de mutações específicas encontradas em cepas resistentes, pode favorecer à compreensão dos mecanismos associados à resistência, permitindo a escolha de drogas eficazes, desenvolvimento de novas drogas e conseqüentemente o aumento da eficácia do tratamento. A bioinformática sendo uma ciência que possibilita a análise das sequências armazenadas em bancos de dados públicos, é uma grande ferramenta para auxiliar na identificação das mutações e suas possíveis relações com a formação de sítios pós-traducionais, através de análise computacional. A utilização de programas de bioinformática desenvolvidos com o objetivo de associar as mutações ao grau de resistência que elas conferem, são de grande valia para identificar de forma sistemática, o grau de resistência viral as drogas que compoem a TARV. Esta identificação tem importância tanto para o início da terapia como para

verificação e estruturação de falha terapêutica. Assim, a detecção de mutantes contendo em seu genoma mutações de resistência pode constituir um marcador precoce da progressão da doença, e, portanto indicar alterações na terapia antes do surgimento de manifestações clínicas e laboratoriais e ainda auxiliar no desenho racional de utilização das drogas.

No presente trabalho, foi desenvolvida uma ferramenta de bioinformática para análise de resistência do HIV-1 aos antirretrovirais. Esta ferramenta de fácil manuseio, estará disponível no site da unidade de bioinformática do LASP/CPqGM/FIOCRUZ, Bioafrica, IOC/FIOCRUZ para domínio público e é capaz de otimizar a análise de dados genômicos do HIV-1. Esta ferramenta pode acessar um banco de dados com seqüências gênicas previamente armazenadas, extrair informações relativas, comparar seqüências de HIV-1 dos bancos de dados públicos, identificando mutações nos genes que codificam a transcriptase reversa e a protease e associar esses dados a resistência a cada anti-retroviral e suas combinações, utilizados nas terapias anti-HIV/AIDS. Adicionalmente, a ferramenta possibilita localizar sítios pós-traducionais e traçar um perfil de distribuição das seqüências armazenadas no banco de dados local, tanto por resistência, quanto por sítios pós-traducionais, para servirem de parâmetro de comparação com as seqüências submetidas pelos usuários da ferramenta.

A ferramenta LASP-HIV1-ResTool apresenta uma tabela de quantidade e percentual de sequências brasileiras do HIV-1, dispostas por subtipo e outra tabela apresentando a quantidade e o percentual de seqüências das regiões *Protease* e *Trasncriptase Reversa* do HIV-1 armazenadas no Banco de Dados. Os usuários da ferramenta podem analisar sequências do banco de dados que compõe a ferramenta ou submeter suas próprias sequências de HIV-1 no formato FASTA.

Durante o processamento das sequências, o sistema identifica a seqüência através do número de acesso e localiza a posição inicial da seqüência dentro da região gênica

escolhida. Em seguida identifica a fase de leitura correta e realiza a conversão da sequência de nucleotídeos em sequência de aminoácidos. Inicia-se então o processo de localização de todas as mutações ocorridas, tanto na sequência de ácido nucléico como na sequência de aminoácido. O sistema então apresenta uma tabela apenas com as mutações que conferem a resistência do HIV-1 contra o antirretroviral. Além disso, exibe o grau de resistência de cada mutação e indica, ao final da tabela, se a amostra de HIV-1 submetida na análise é susceptível, apresenta resistência intermediária ou se é resistente ao antirretroviral. O sistema exibe dois gráficos que relacionam: a quantidade de sítios pós-traducionais em função do total de sequências submetidas e os padrões de resistência aos antirretrovirais (susceptível, intermediária e resistente) em função do total de sequências. Além de apresentar um gráfico de barras mostrando a quantidade de ocorrências de cada mutação de resistência no conjunto de sequências analisadas.

A ferramenta LASP-HIV1-ResTool também gera uma página em formato XML contendo os dados das análises do usuário, que podem ser salvas em um arquivo. Os arquivos XML podem ser importados de outras ferramentas tais como: planilhas eletrônicas, programas de análise estatística e sistemas gerenciadores de bancos de dados.

O desenvolvimento da ferramenta LASP-HIV1-ResTool, além de oferecer um mecanismo eficiente para contribuir com as pesquisas relacionadas a resistência do HIV-1 aos antirretrovirais, possibilita a aquisição de conhecimento e experiência no campo da bioinformática, fator este muito importante para desenvolvimentos de novos sistemas informatizados voltados para as áreas biológicas.

OBJETIVOS

3. OBJETIVOS

3.1. GERAL:

Desenvolvimento de uma ferramenta de bioinformática capaz de otimizar a análise de seqüências de HIV-1 , identificando mutações de resistência nos genes da transcriptase reversa e a protease e associando esses dados à resistência a cada antirretroviral utilizado no tratamento dos indivíduos infectados pelo HIV-1.

3.2. ESPECÍFICOS:

1 – Desenvolver um Banco de Dados local com seqüências do HIV-1 do Brasil, obtidas de Bancos de Dados Públicos.

2 – Desenvolver uma ferramenta on-line, de domínio público, para análise de resistência a drogas anti-retrovirais, e de sítios pós-traducionais.

3 – Analisar seqüências do HIV-1, quanto sua distribuição pela ocorrência de mutações e o grau de resistência por ela aferidas a cada droga anti-retroviral e por sítios pós-traducionais, utilizando a ferramenta desenvolvida.

4 – Disponibilizar a ferramenta de análise de resistência no site do Laboratório Avançado em saúde Pública-Lasp/CPqGM/FIOCRUZ-Ba para utilização pública.

MATERIAL E MÉTODOS

4. MATERIAL E MÉTODOS:

4.1. OBTENÇÃO DAS SEQUÊNCIAS

O trabalho foi iniciado com a obtenção das seqüências de HIV-1 do Brasil depositados no banco de dados HIV Sequence Database de Los Alamos (<http://www.hiv.lanl.gov/>).

O “HIV Sequence Database” é um sistema de acesso a um banco de dados que contém seqüências genômicas, mutações relacionadas à resistência, epitopos de Imunologia e ensaios de vacinas relacionados ao HIV. Somado ao sistema de banco de dados, o site possui diversas outras ferramentas que podem ser utilizadas para analisar esses dados. Este projeto é Mantido pela Divisão de AIDS do National Institute of Allergy and Infectious Diseases (NIAID), parte do National Institutes of Health (NIH).

Ao acessar o “HIV Sequence Database”, é exibida a primeira tela do sistema, onde são inseridos os parâmetros que especificam as seqüências a serem obtidas, de acordo com o interesse do usuário. As seqüências podem ser selecionadas de acordo com parâmetros pré-definidos como: vírus, país, subtipo, região genômica, número de acesso.

Para a estruturação do banco de seqüências necessário ao presente projeto, na variável “Virus” foi selecionado HIV-1, na variável “Sampling country” foi inserido BR (sigla que representa o Brasil), na variável “Subtype” foi escolhido o valor “Any subtype” (utilizado para representar a escolha de todos os subtipos) e na variável “Genomic region” (campo que representa a região genômica) foi escolhida a região codificadora da *protease* (FIGURA 5).

Sequence Search Interface

Tips

- Click or mouse over the field name for specific tips
- The *italicized fields* are listed in output by default
- To list fields that are not listed by default or included in the search, put an asterisk (*) in the input box.
- Use the + and - to see more or fewer search fields
- For other details about each field, see [Help](#)

Last GenBank update: 2010-02-02

[Advanced Search](#)

Sequence Information

Accession number <input type="text"/>	Subtype	<div style="border: 1px solid black; padding: 2px;"> Any subtype No subtype A A1 A2 B </div>
Sequence name <input type="text"/>		
Sequence length <input type="text"/>		
exact <input checked="" type="checkbox"/> Sampling year <input type="text"/>	<input type="checkbox"/> Include recombinants	
Virus <input type="text" value="HIV-1"/>	<input type="checkbox"/> Include problematic sequences	

More sequence information

Find all sequences for a specific gene or region (HIV-1 and SIVcpz)

<div style="border: 1px solid black; padding: 2px;"> p6 Pol CDS p51 (RT) p15 (RNase H) p31 (Integrase) orf97 VIF CDS </div>	Or define start <input type="text"/> and end <input type="text"/> <input type="checkbox"/> Include fragments of minimum length <input type="text" value="100"/>
--	--

Combine database sequences with your own sequence alignment (HIV-1 and SIVcpz)

Publication Information

Patient Information

Geographical Information

Sampling country <input type="text" value="BR"/>	Sampling site <input type="text"/>	
Infection country <input type="text"/>	Infection site <input type="text"/>	
Geographic region <input type="text" value="Any"/>		

Output

List records per page Show results selected Show SQL

[Advanced Search](#)

last modified: Tue Feb 9 08:06 2010

Questions or comments? Contact us at seq-info@lanl.gov.

FIGURA 5. Sistema HIV Sequence Database para aquisição das sequencias e HIV-1 do Brasil.

Após a escolha dos parâmetros de busca das seqüências Brasileiras de HIV-1 e o clique no botão “Search”, o sistema nos direciona para a tela seguinte, onde estão listadas todas as seqüências selecionadas conforme os parâmetros inseridos. (FIGURA 6).

Please note that if a genomic region is specified, only HIV-1 sequences can be found

Displaying 1 - 100 of 2675 sequences found:

Note: 9 [problematic](#) sequences were removed from this result. [Click here to repeat search to include problematic sequences.](#)

 record to
 records per page

Click on field name to sort in ascending or descending order

#	Select	Patient Code (id)	Accession Name	Subtype	Country	Sampling Year	Genomic Region	Sequence Length	Organism
1	<input type="checkbox"/> Blast		AB485641 BZ167	B	BR			9644	HIV-1
2	<input type="checkbox"/> Blast		AB485642 BZ167	B	BR			9662	HIV-1
3	<input type="checkbox"/> Blast		AB485656 BZ163	F1	BR			9602	HIV-1
4	<input type="checkbox"/> Blast		AB485657 BZ163	F1	BR			9602	HIV-1
5	<input type="checkbox"/> Blast	BR020(143)	AF005494 93BR020_1	F1	BR	1993		8968	HIV-1
6	<input type="checkbox"/> Blast	BR029(58)	AF005495 93BR029_4	BF1	BR	1993		8954	HIV-1
7	<input type="checkbox"/> Blast		AF079981 BR31	B	BR			297	HIV-1
8	<input type="checkbox"/> Blast		AF079982 BR32	B	BR			297	HIV-1
9	<input type="checkbox"/> Blast		AF079983 BR33	B	BR			297	HIV-1
10	<input type="checkbox"/> Blast		AF079984 BR40	B	BR			297	HIV-1
11	<input type="checkbox"/> Blast		AF079985 BR41	F1	BR			297	HIV-1
12	<input type="checkbox"/> Blast	BR57(10018)	AF079986 BR57	F1	BR			297	HIV-1
13	<input type="checkbox"/> Blast	BR58(10883)	AF079987 BR58	F1	BR			297	HIV-1
14	<input type="checkbox"/> Blast	BR59(10884)	AF079988 BR59	F1	BR			297	HIV-1
15	<input type="checkbox"/> Blast		AF079989 BR60	F1	BR			297	HIV-1
16	<input type="checkbox"/> Blast		AF079990 BR70	B	BR			297	HIV-1
17	<input type="checkbox"/> Blast		AF079991 BR71	B	BR			297	HIV-1
18	<input type="checkbox"/> Blast		AF079992 BR72	B	BR			297	HIV-1
19	<input type="checkbox"/> Blast		AF079993 BR73	B	BR			297	HIV-1
20	<input type="checkbox"/> Blast		AF079994 BR75	B	BR			297	HIV-1
21	<input type="checkbox"/> Blast	BR97(10885)	AF079995 BR97	F1	BR			297	HIV-1
22	<input type="checkbox"/> Blast		AF079996 BR112	F1	BR			297	HIV-1
23	<input type="checkbox"/> Blast		AF099155 BR34	B	BR			297	HIV-1
24	<input type="checkbox"/> Blast		AF099156 BR43	B	BR			297	HIV-1
25	<input type="checkbox"/> Blast		AF099157 BR45B	B	BR			297	HIV-1
26	<input type="checkbox"/> Blast		AF099158 BR45F	F1	BR			297	HIV-1
27	<input type="checkbox"/> Blast		AF099159 BR46	F1	BR			297	HIV-1

Figura 6. Página do HIV Sequence Data Base contendo as seqüências selecionadas para download.

As sequências brasileiras da protease do HIV-1 de interesse foram selecionadas através do link “Select all” e obtidas através do link “Download sequences”. O sistema oferece ainda a possibilidade de processamento preliminar das sequências de HIV-1, como a retirada de gaps e possibilidade de armazenamento em diferentes formatos de arquivo, como fasta, philip, mega etc. para utilização nos diferentes programas bioinformática aplicados à filogenia do HIV-1 (FIGURA 7). Na seleção das nossas sequências retiramos os intervalos (“gaps”), que são espessamentos introduzidos nas sequências para efeito de alinhamento múltiplo, selecionando no campo “Gap handling” a opção “strip”. Na opção “Show names as” incluímos a variável “Accession” que representa o código identificador da sequência no Genbank, mantivemos o restante das variáveis com seus respectivos valores padrão.

Download options

Align (only HIV-1 and SIVcpz) Format: Gap handling:
 Sequence type:

Clip to selected region Include HXB2 Reference Sequence (K03455)
 Show names as: or [compose a label](#)

Displaying 1 - 100 of 2675 sequences found:

[Exclude problematic sequences](#)

[Select all](#) [Unselect all](#) [Invert selection](#) [Show all](#) [One sequence/patient](#) record records per pa

Click on field name to sort in ascending or descending order

#	Select	Patient (id)	Code	Accession Name	Subtype	Country	Sampling Year	Genomic Region	Sequence Length	Organism
1	<input checked="" type="checkbox"/>	Blast		AB485641 BZ167	B	BR			9644	HIV-1
2	<input checked="" type="checkbox"/>	Blast		AB485642 BZ167	B	BR			9662	HIV-1
3	<input checked="" type="checkbox"/>	Blast		AB485656 BZ163	F1	BR			9602	HIV-1
4	<input checked="" type="checkbox"/>	Blast		AB485657 BZ163	F1	BR			9602	HIV-1
5	<input checked="" type="checkbox"/>	Blast	BR020(143)	AF005494 93BR020_1	F1	BR	1993		8968	HIV-1
6	<input checked="" type="checkbox"/>	Blast	BR029(58)	AF005495 93BR029_4	BF1	BR	1993		8954	HIV-1
7	<input checked="" type="checkbox"/>	Blast		AF079981 BR31	B	BR			297	HIV-1
8	<input checked="" type="checkbox"/>	Blast		AF079982 BR32	B	BR			297	HIV-1
9	<input checked="" type="checkbox"/>	Blast		AF079983 BR33	B	BR			297	HIV-1
10	<input checked="" type="checkbox"/>	Blast		AF079984 BR40	B	BR			297	HIV-1
11	<input checked="" type="checkbox"/>	Blast		AF079985 BR41	F1	BR			297	HIV-1
12	<input checked="" type="checkbox"/>	Blast	BR57(10018)	AF079986 BR57	F1	BR			297	HIV-1
13	<input checked="" type="checkbox"/>	Blast	BR58(10883)	AF079987 BR58	F1	BR			297	HIV-1
14	<input checked="" type="checkbox"/>	Blast	BR59(10884)	AF079988 BR59	F1	BR			297	HIV-1
15	<input checked="" type="checkbox"/>	Blast		AF079989 BR60	F1	BR			297	HIV-1
16	<input checked="" type="checkbox"/>	Blast		AF079990 BR70	B	BR			297	HIV-1
17	<input checked="" type="checkbox"/>	Blast		AF079991 BR71	B	BR			297	HIV-1
18	<input checked="" type="checkbox"/>	Blast		AF079992 BR72	B	BR			297	HIV-1
19	<input checked="" type="checkbox"/>	Blast		AF079993 BR73	B	BR			297	HIV-1
20	<input checked="" type="checkbox"/>	Blast		AF079994 BR75	B	BR			297	HIV-1
21	<input checked="" type="checkbox"/>	Blast	BR97(10885)	AF079995 BR97	F1	BR			297	HIV-1
22	<input checked="" type="checkbox"/>	Blast		AF079996 BR112	F1	BR			297	HIV-1

Figura 7. Página do HIV Sequence Data Base contendo as sequências e o quadro com as opções de downloads.

Como o número de sequências selecionadas foi maior do que 1000, o sistema nos enviou para uma nova tela para inserirmos um e-mail pelo qual recebemos as sequências posteriormente num arquivo no formato FASTA. Em seguida repetimos os mesmos procedimentos para a obtenção das sequências da Transcriptase Reversa.

4.2. ORGANIZAÇÃO DAS SEQUÊNCIAS

4.2.1. SUBTIPAGEM DA SEQUÊNCIAS

O programa de bioinformática “REGA subtyping tool” é um sistema automatizado de análise genotípica de fácil manipulação, aplicado a estudos filogenéticos de sequências do HIV-1 e outros patógenos virais. O programa foi desenvolvido com as linguagens de programação Java e Perl e encontra-se disponível na Internet no endereço eletrônico <http://bioafrica.net/subtypetool> (T. DE OLIVEIRA *et al.*, 2005). O “REGA subtyping tool” combina análises filogenéticas com métodos de “bootscanning” para realizar a subtipagem gênica de sequências completas ou fragmentos de sequências do HIV-1. Ao ser comparado com os métodos tradicionais de investigação de subtipos por filogenia, sobre dados previamente publicados, o “REGA subtyping tool”, apresentou entre 97,5% e 100% de precisão na subtipagem de não-recombinantes e formas recombinantes circulantes (CRFs) do HIV-1, incluindo 108 full-length , 108 e 221 seqüências gag env obtidas a partir do banco de dados de Los Alamos. O sistema, de subtipagem pode ser facilmente adquirido através de “download” e instalado em sistemas UNIX ou Linux.

O sistema inicia o processamento com a construção de uma árvore filogenética pelo método a distância HKY no software PAUP. As sequências sob consulta são confrontadas com as de referência do grupo M para os subtipos AD, FH, J e KA de K. A

etapa seguinte envolve a construção de uma segunda árvore usando a seqüência sob consulta, o HIV-1 e as seqüências de CRFs. Na terceira etapa, a seqüência de consulta é dividida em pequenos segmentos e é analisada por recombinação utilizando “bootscanning” com uma janela deslizante de 400 pares de bases movendo-se em etapas de 50 pares de bases. Finalmente, na quarta etapa, o alinhamento é examinado para verificar se contém sinais filogenéticos suficiente para a determinação do subtipo, utilizando então o software Treepuzzle (FIGURA 8).

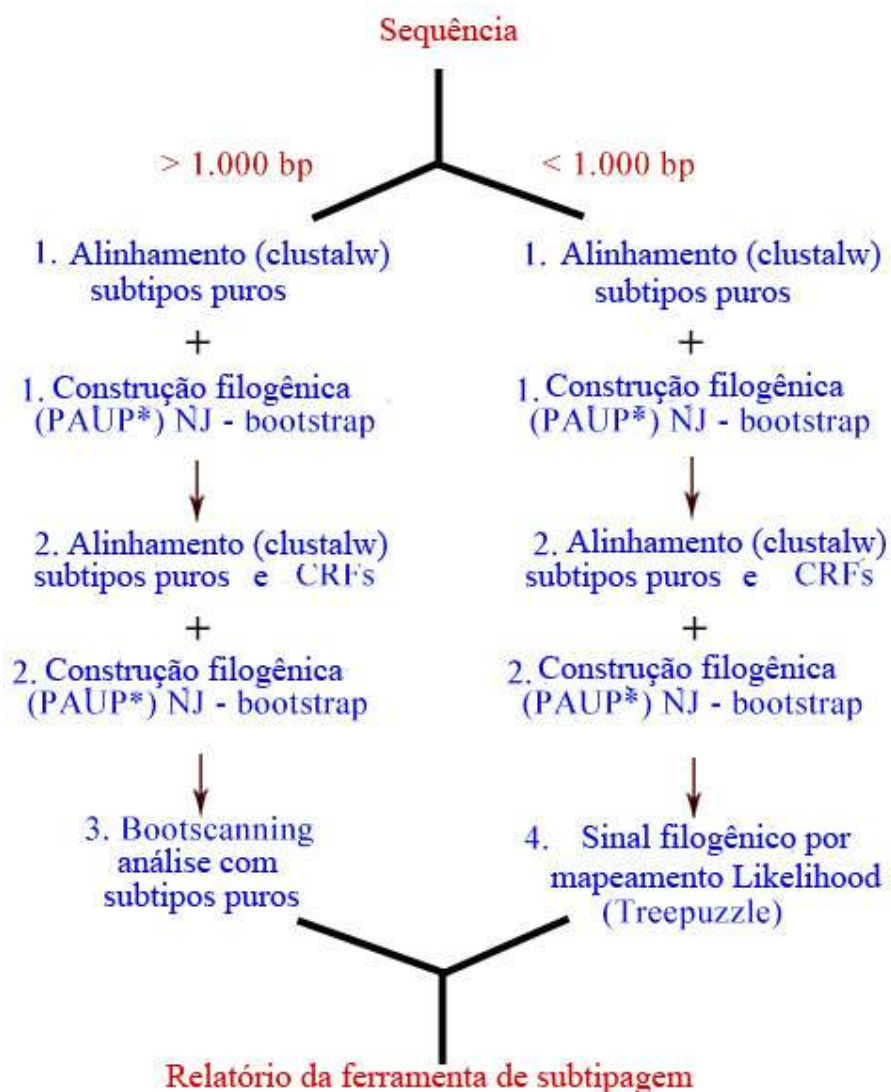


FIGURA 8. Esquematização da organização do sistema de subtipagem do HIV-1 REGA subtyping tool.

Todas as seqüências que não estavam caracterizadas genotipicamente quanto ao subtipo viral, foram submetidas ao REGA HIV-1 Subtyping Tool (<http://www.bioafrica.net/rega-genotype/html/indexhiv.html>) para que fosse feita a identificação do subtipo do HIV-1 através de análise filogenética. Ao acessar o REGA HIV-1 Subtyping Tool, a primeira tela é apresentada, com uma caixa de texto para a inclusão das seqüências, em formato FASTA, a serem caracterizadas(FIGURA 9). Clicando no botão “Run” o sistema envia as seqüências para serem analisadas e na próxima tela obtemos o output da subtipagem do HIV-1 tendo a possibilidade de obter os arquivos na forma xml, pdf e exl.



REGA HIV-1 & 2 Automated Subtyping Tool (Version 2.0)

This tool is designed to use phylogenetic methods in order to identify the subtype of a specific sequence. The sequence is analysed for recombination using bootscanning methods.

Note for batch analysis: The REGA subtype tool accepts up to 1000 sequences at a time.

Enter here your input data as FASTA format.

[Choose a mirror to subtype your sequences](#)

```
>GQ401300
CCTCAAATCACTCTTTGGCAGCGACCCCTTGTCAACAATAAGAGTAGGGGG
ACAGCTAAAGGAGGCTCTCTTAGACACAGGAGCAGATGATACAGTATTAG
AAGAAATGAAATGCCAGGAAATGGAAACCAAAAATGATAGGAGGAATT
GGGGTTTTATCAAAGTAAGACAGTATGATCAAATACCTATAGAAATTTG
TGGAAAAAGGCTATAGGTACAGTACTAGTGGGACCTACACCTGTCAACA
TAATTGGAAGAAACCTGTTGACTCAGTTGGGGTGCACACTAAATTTT
>GQ401301
CCTCAGATCACTCTTTGGCAGCGACCCCTTGTCTCAATAAAAGTAGGGGG
ACAGACAAAGGAGGCTCTGTTAGACACAGGAGCAGATGATACAGTATTAG
AAGAAATAAAATGCCAGGAAATGGAAGCCAAAATGATAGGGGGAATT
GGAGGTTTTATCAAAGTAAGACAGTATGATCAAATACCTATAGAAATTTG
```

[Submit sequences](#) [How to cite](#) [HIV Tutorials](#) [HIV Decision Trees](#) [HIV Subtyping Process](#) [HIV Example Sequences](#) [Contact us](#)

Developed by: Tulio de Oliveira, Koen Deforche, Sharon Cassol, Andrew Rambaut and Anne-Mieke Vandamme.

Developed in cooperation with the [Evolutionary Biology Group](#) at University of Oxford, UK, the [HIV-1 Pathogenesis and Immunotherapeutics Program](#) at University of Pretoria, South Africa, and the [REGA Institute](#) at the Katholieke Universiteit Leuven, Belgium.

Funded by the Marie Curie Fellowship, Flanders Bilateral Cooperation Program and the Wellcome Trust (grant # 061238).

For questions, suggestions or problems please contact: [Dr. Tulio de Oliveira](#).

FIGURA 9. Página inicial do sistema de subtipagem do HIV-1 REGA subtyping tool.

A tela seguinte do REGA HIV-1 Subtyping Tool realiza o processamento das amostras e retorna o subtipo de cada sequência submetida (FIGURA 10).

HIV Genotyping Tool Results











Name	Length	Report	Assignment	Support	Genome
GQ48130	297bp	Report	HIV-1 Subtype C	95.8	
GQ48131	297bp	Report	HIV-1 Subtype C	100.0	
GQ48132	297bp	Report	HIV-1 Subtype C	98.8	
GQ48133	297bp	Report	HIV-1 Subtype C	97.8	
GQ48134	297bp	Report	HIV-1 Subtype C	94.8	
GQ48135	297bp	Report	HIV-1 Subtype C	100.0	
GQ48137	297bp	Report	HIV-1 Subtype C	99.8	
GQ48138	297bp	Report	HIV-1 Subtype C	98.8	
GQ48139	297bp	Report	HIV-1 Subtype C	96.8	
GQ48130	297bp	Report	HIV-1 Subtype B	78.8	

Figura 10. Página d sistema de subtipagem do HIV-1 REGA subtyping tool com as sequências subtipadas do HIV-1 do Brasil.

4.2.2. ESTRATIFICAÇÃO DAS SEQUÊNCIAS

As sequências foram organizadas com o auxílio de um script, desenvolvido na linguagem de programação PHP, que processa o arquivo FASTA com as seqüências obtidas no HIV Sequence Database, cria diretórios e arquivos FASTA nomeados e organizados de acordo com as informações obtidas em cada seqüência (número de acesso, país de origem, sub-tipo, ano de amostragem)(FIGURA 11).

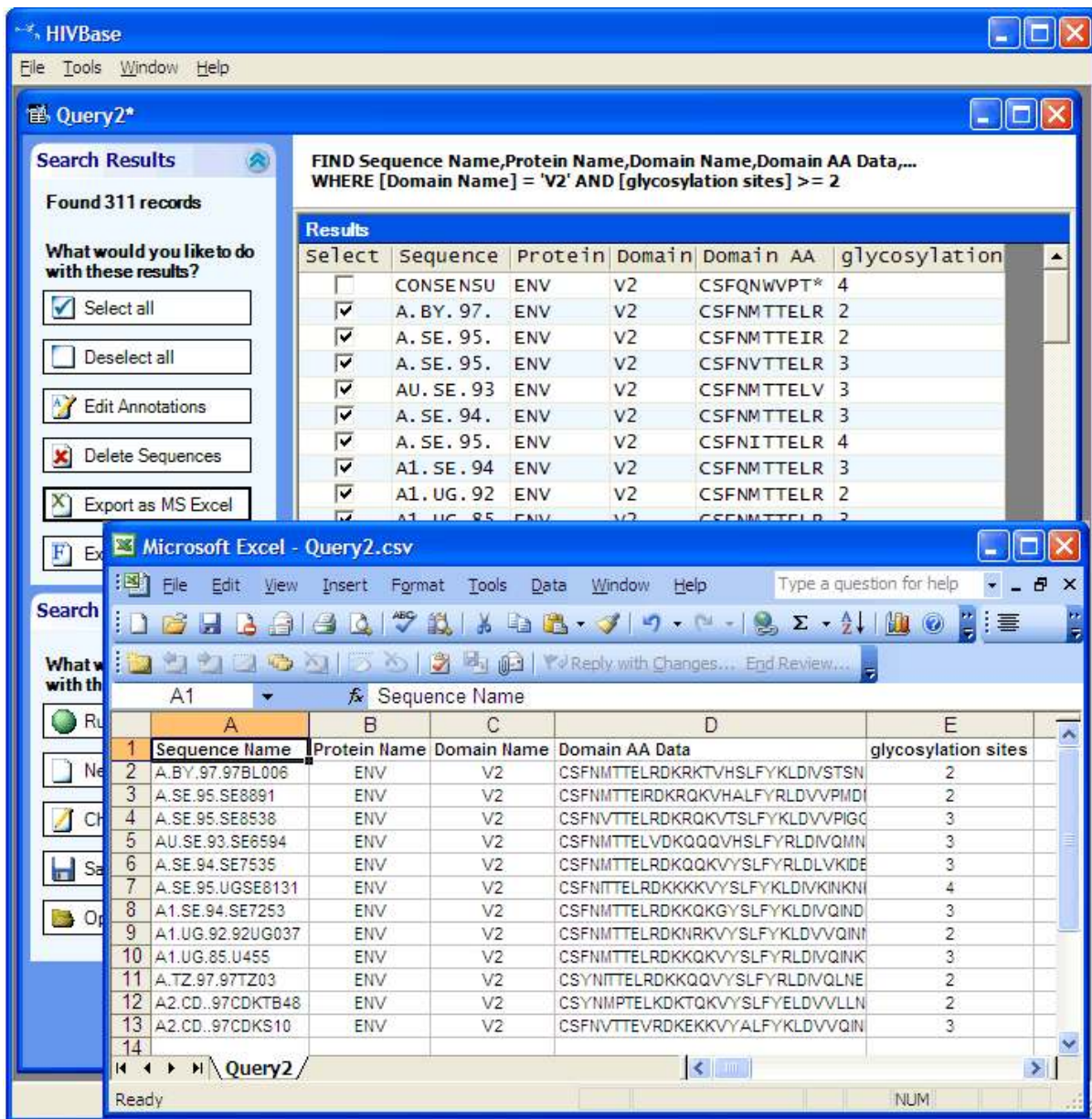


FIGURA 12. Sistema HIVBASE para manipulação e organização das sequencias HIVBASE (adaptado de: <http://www.genejohnson.net/Export.aspx>).

Para estruturação e organização do banco de dados genômicos utilizamos as ferramentas “Data Import Tools” que importa sequências de ácidos nucleicos ou aminoácidos, individualmente ou em arquivos no formato FASTA e identifica as proteínas e os motivos que elas contêm. Com o recurso “Query Data” é possível filtrar as característica específicas ao interesse da pesquisa associadas a sequência. E para finalizar, o HIVbase possui a ferramenta “Export” que permite a exportação dos dados,

selecionados através do recurso “Query Data”, em diversos formatos que podem ser usados em outros softwares como: Microsoft Word, Excel, SAS, ou qualquer aplicativo de alimento de sequências que reconheça o formato FASTA.

4.3. ARMAZENAMENTO DAS SEQUÊNCIAS

Para o armazenamento das sequências, foi criado um banco de dados relacional utilizando o SGBD (Sistema Gerenciador de Bancos de Dados) Mysql, em um servidor Linux, no LASP (Laboratório Avançado de Saúde Pública/CPqGM/FIOCRUZ). O Mysql é um sistema de gerenciamento de banco de dados Open Source (código fonte disponível para ser estudado e adaptado às necessidades específicas de cada usuário), desenvolvido e disponibilizado gratuitamente no web site do MySQL (<http://www.mysql.com/downloads/mysql/>). As informações mais recentes do MYSQL são disponibilizadas constantemente em seu web site possibilitando, desta forma, uma melhor compreensão e um gerenciamento mais eficiente de seus recursos. O MYSQL tem capacidade para armazenar uma grande quantidade de dados. Este é um atributo muito importante para favorecer análises relacionadas de sequências genômicas, que exigem grande quantidade de espaço e excelência em desempenho. O MySQL usa a GPL (GNU General Public License - Licença Pública Geral GNU), disponível no web site <http://www.fsf.org/licenses>, que define as diferentes regras para as diferentes maneiras com que o sistema pode ser utilizado. O banco de dados relacional é dividido em diversas tabelas que armazenam dados necessários ao funcionamento do sistema. Cada tabela armazena os dados necessários a realização de uma determinada função, além de uma chave que funciona como índice único e possibilita a interação com as outras tabelas do sistema. O banco de dados possui uma tabela que armazena as sequências e os dados

a elas relacionados, uma tabela que armazena dados relacionados às drogas antirretrovirais utilizadas nas terapias anti HIV, além de uma tabela para armazenar dados relacionados às mutações e seus escores de resistência. O banco de dados, também é composto por uma tabela que armazena a sequência de referência (HXB2) dividida em regiões gênicas. A sequência HXB2 é utilizada como base para a comparação com as sequências que deverão ser analisadas.

4.4. ALGORITMO DE RESISTÊNCIA

O grau de resistência mensurado pela ferramenta de análise aos antirretrovirais, baseia-se no sistema de pontuação utilizado na lógica do algoritmo do HIV Drug Resistance Database da Universidade de Stanford, disponível no web site <http://hivdb.stanford.edu/pages/asi/index.html#references> Hirsch MS, (RHEE *et al.*, 2003; SHAFER *et al.*, 2000; SHAFER, 2006; LIU, 2006; RHEE *et al.*, 2006; GIFFORD *et al.*, 2009), com algumas adaptações relacionadas à faixa intermediária de resistência. O sistema de pontuação do HIV Drug Resistance Database dividi-se em cinco faixas de sensibilidade. A primeira faixa, que varia entre 0 e 9 classifica a sequência de aminoácidos do HIV-1 como susceptível, a segunda faixa que varia de 10 a 14 como baixo nível de resistência, entre 15 e 29 com um baixo nível de resistência porém com indicativo de uma exposição anterior a classe da droga antirretroviral, entre 30 e 59 um grau de resistência intermediário entre a baixa e a alta resistência e na faixa igual ou maior que 60, um alto grau de resistência, em geral, com pouca ou nenhuma resposta virológica ao tratamento com a droga. As mutações que conferem tais resistências vêm sendo descritas na literatura na medida em que são identificadas e comprovadas por ensaios clínicos. A partir daí passam a ser considerados pelos algoritmos. A posição de

cada mutação e seu respectivo grau de resistência são então cadastrados no banco de dados para serem acessadas pela ferramenta.

4.5. SÍTIOS PÓS-TRADUCIONAIS

As glicoproteínas gp120 e gp41 do envelope do HIV-1 são estruturas determinantes do processo de entrada na célula do hospedeiro, através da ligação ao receptor CD4 e co-receptores e da fusão das membranas viral e celular respectivamente.

Os níveis de glicosilação das proteínas do envelope viral tem um papel crucial na patogênese viral, impactando diretamente na conformação estrutural e funcional do envelope, influenciando a antigenicidade e imunogenicidade da proteína. Portanto a identificação e conhecimento dos níveis e padrões de glicosilação de isolados virais provenientes de pacientes brasileiros são fundamentais para o desenvolvimento de futuros imunógenos destinados a indução de uma resposta imune humoral protetora contra o HIV-1 no Brasil.

A tabela 5 apresenta algumas regras para identificação de sítios pos-traducionais (sítios de glicosilação e forforilação), já incorporadas à ferramenta HIV-1ResTool, além de outras duas regras que servem de exemplo de regras que também podem ser incorporadas à ferramenta. As letras em negrito da primeira coluna identificam cada regra, sendo: **A.** a presença de asparagina na primeira posição e serina ou treonina na terceira posição, Além da ausência de prolina na segunda e na quarta posições, sugerem a presença de um sitio de glicosilação. **B.** a presença de serina ou treonina na primeira posição e um aminoácido qualquer na segunda posição, além da presença de arginina ou lisina na terceira posição, sugerem a presença de um sitio de fosforilação PKC. **C.** a presença de serina ou treonina na primeira posição mais dois aminoácido quaisquer na segunda e terceira posições, além da presença de ácido aspártico ou ácido glutâmico na quarta posição, sugerem a presença de um sitio de fosforilação CK2. **D.** a presença de glicina na primeira posição mais dois aminoácido quaisquer na terceira e quarta posições e a presença de serina, treonina, alanina, glicina, cisteina ou asparagina na quinta posição, além da ausência de ácido glutâmico, ácido aspártico, arginina, lisina, histidina, prolina, fenilamina, tirosina e triptofano na segunda posição e prolina na sexta posição, sugerem a presença de um sitio de miristitação.

	Sítios Pos-traducionais	Regras
A.	Glicosilação	N-{P}-[ST]-{P}
B.	PKC_fosforilação	[ST]-x-[RK]
C.	CK2_fosforilação	[ST]-x(2)-[DE]
D.	Miristilação	G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}

Tabela 5. Regras para a identificação de sítios pós-traducionais.

4.6. CRIAÇÃO DA FERRAMENTA DE ANÁLISE DE RESISTÊNCIA

Após serem coletadas e a organizadas, as sequências foram armazenadas no banco de dados para que se desse início ao processo de desenvolvimento da ferramenta para análise de resistência as drogas antirretrovirais e sítios pós-traducionais de glicosilação e fosforilação. A ferramenta para análise de resistência e sítios pós-traducionais foi desenvolvida com a linguagem de programação PHP, e está hospedada, também no servidor do LASP para funcionar como uma das diversas ferramentas do site <http://www.bahiana.edu.br/lasp/>. Adicionalmente, a ferramenta localiza os domínios potenciais protéicos de glicosilação e fosforilação utilizando regras estabelecidas no tutorial da ferramenta de bioinformática PROSITE (<http://ca.expasy.org/cgi-bin/prosite-list.pl#ptm>) (Sigrist C.J.A. *et al.*, 2002; Hulo N. *et al.*, 2006; Bologna G. *et al.*, 2004), além de criar e disponibilizar um link para acesso aos resultados das análises no formato XML. O formato XML (Extensible Markup Language) é uma poderosa ferramenta utilizada principalmente para a transmissão de dados entre diferentes tipos de sistemas. Esta característica é muito importante para sanar os problemas de incompatibilidade no momento em que os dados devem ser transportado de uma plataforma computacional para outra. Os sistemas utilizados para a criação, armazenamento e manutenção da

ferramenta desenvolvida no presente trabalho serão de domínio público seguindo as normas da GPL (General Public License).

RESULTADOS

5. RESULTADOS

5.1. PARTE 1: DESCRIÇÃO DA CONSTRUÇÃO DA FERRAMENTA:

5.1.1. BANCO DE DADOS

Foi desenvolvido um banco de dados relacional onde foram armazenadas seqüências das regiões gênicas da *protease* e da transcriptase reversa de isolados de HIV-1 obtidos de pacientes brasileiros e cadastrados no HIV Sequence Database (<http://www.hiv.lanl.gov/>).

O banco de dados e os elementos que compõem sua estrutura (tabelas, campos) possuem uma nomenclatura independente de regras de acentuação para manter a consistência técnica necessária para o melhor funcionamento do sistema (S. Pompilho).

O banco de dados intitulado “resistencia” possui sete tabelas conectadas entre si por campos chave. As tabelas são estruturas lógicas, organizadas em linhas e colunas, utilizadas para armazenar dados de forma sistematizada possibilitando o acesso e a manipulação por sistemas computadorizados. Os dados das tabelas são ocultos aos usuários do sistema, que têm acesso apenas a informação gerada pelo processamento dos dados. Os campos chave são constituídos por conjuntos de um ou mais atributos e são encontrados em duas formas a chave primária e a chave estrangeira. A chave primária (PK - Primary Key) que nunca se repete, para manter uma importante característica de unicidade ao registro, além de funcionar como índice para a localização dos registros dentro do banco de dados. A chave estrangeira (FK - Foreign Key) que pode se repetir e dar origem ao relacionamento entre as tabelas. A chave estrangeira em uma tabela corresponde a chave primária na tabela a qual se pretende realizar o relacionamento (FIGURA 13).

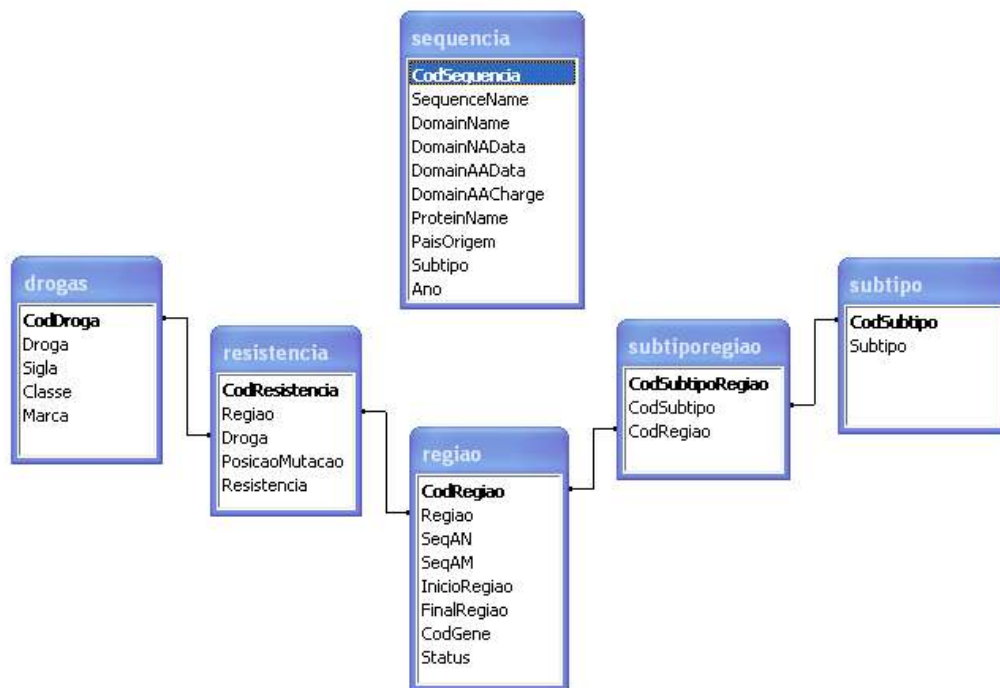


FIGURA 13. Esquematização do banco de dados relacional acessado pela ferramenta LASP-HIV-1ResTool.

5.1.1.2 TABELA SEQUENCIA

A tabela denominada “sequencia” é composta pelo campo “codSequencia” que armazena a chave de identificação de cada registro dentro da tabela (chave primária), o campo “SequenceName”, armazena o número de acesso de cada sequência, o campo “DomainName” armazena a chave de conexão (chave estrangeira) da tabela sequência com a tabela região, o campo “DomainNADData” armazena as sequências de nucleotídeos, o campo “DomainAAData” armazena a sequência de aminoácidos, os campos restantes desta tabela são PaisOrigem, Subtipo e Ano, e armazenam respectivamente o país de

origem da sequência, o subtipo ao qual a sequência pertence e o ano de obtenção da amostra.

5.1.1.3. TABELA REGIAO

A tabela “regiao” é composta por campos que armazenam dados relacionados às regiões gênicas do HIV-1. O primeiro campo, “codRegiao”, armazena a chave primária, o segundo campo, denominado “Regiao” armazena o nome das regiões gênicas e o terceiro e quarto campos armazenam respectivamente as sequências de nucleotídeos e aminoácidos da amostra de referência do HIV-1 (HXB2). A amostra de referência é a sequência considerada original, sem as mutações que conferem resistência aos antirretrovirais. A tabela “regiao” possui também o campo codGene (chave estrangeira), responsável pela conexão com a tabela “hivgenoma”. A tabela “regiao” possui também o campo “codGene” (chave estrangeira), responsável pela conexão com a tabela “hivgenoma”. Adicionalmente, a tabela “Regiao” possui o campo Status que recebe valores booleanos (positivo ou negativo). Quando preenchido, o campo “Status” é interpretado como verdadeiro e a região gênica passa a ser considerada nas análises de resistência. Quando, porém o campo Status está vazio, a interpretado para a região gênica é falsa e o sistema deixa de considerá-la nas análises de resistência.

5.1.1.4. TABELA GENOMA

A tabela “hivgenoma” armazena o genoma do HIV-1 (HXB2), dividido por regiões gênicas. O campo “codGenoma” possui a função de chave primária. Os campos “DomainNA” e “DomainAA” armazenam respectivamente as sequências de nucleotídeos e aminoácidos de cada região gênica.

5.1.1.5. TABELA RESISTENCIA

A tabela “resistencia” é responsável por armazenar os dados utilizados para inferir o impacto que cada mutação, na sequência de aminoácido, exerce sobre os diversos tipos de drogas utilizadas nas terapias antirretroviral. O primeiro campo da tabela, intitulado “codResistencia”, armazena a chave primária. O campo “Regiao” armazena a chave estrangeira que possibilita a conexão da tabela “resistencia” com a tabela “regiao”, o campo “Droga” conecta a tabela “regiao” com a tabela “droga” com a tabela. Entre os campos da tabela resistencia encontram-se os campos “PosicaoMutacao” e “Resistencia”, de fundamental importância para a análise de resistência do HIV aos antirretrovirais. O campo “PosicaoMutacao” armazena os dados que indicam a ocorrência de cada mutação, informando o aminoácido original e o aminoácido mutante, além de indicar a posição onde ocorre a mutação. Enquanto o campo “Resistencia” armazena o dado que mensura o impacto da mutação na resistência do HIV à droga antirretroviral.

5.1.1.6. TABELA SUBTIPO

A tabela “subtipo” é uma tabela auxiliar, onde são armazenados os subtipos do HIV-1, atualmente identificados. O campo “codSubtipo” armazena a chave primária e o campo “Subtipo” armazena os subtipos de HIV-1, representados por letras (A,B,C,D,E,F,G,H,J,K).

5.1.1.7. TABELA SUBTIPOREGIAO

A tabela “subtiporegiao” implementa a associação entre a tabela “subtipo” e a tabela “regiao”. Esta associação é feita com a utilização dos campos “codSubtipo” e “codRegiao”, que funcionam como chave estrangeira na tabela “subtiporegiao”.

5.1.1.8. TABELA DROGA

Os dados relativos as drogas utilizadas nas terapias antirretroviral encontram-se armazenados na tabela “drogas”. Esta tabela possui o campo “codDroga” como chave primária, o campo “Droga” contendo o nome de cada antirretroviral e o campo “Sigla” para armazenar a sigla da droga. O antirretroviral, muitas vezes é conhecido por sua sigla.

5.1.2. PROGRAMAÇÃO DE SCRIPT PARA MANIPULAÇÃO DE SEQUÊNCIAS

5.1.2.1 - ORGANIZAÇÃO DAS SEQUÊNCIAS

Após a criação do banco de dados, iniciou-se o processo de obtenção e organização das seqüências de HIV-1. Para armazenar as sequências no banco de dados, foi necessária a estratificação de cada seqüência por subtipo, país de origem, região gênica e ano de obtenção. Nesta etapa do projeto, foi desenvolvido um script com a linguagem de programação PHP, a mesma linguagem utilizada na criação da ferramenta de análise de resistência aos antirretrovirais. O Script processa e estratifica seqüências de arquivos no formato FASTA, além de organizá-las em diretórios específicos às características de cada seqüência. Após sua execução, o Script processa o arquivo FASTA que contém todas as seqüências, lê o cabeçalho de cada seqüência (país de origem, subtipo, ano de obtenção) e cria arquivos e diretórios para armazenar cada seqüência de acordo com as características descritas. Após a submissão ao "Script", as seqüências foram importadas pelo HIVbase onde foi finalizada a estratificação e organização. Após organizadas, as seqüências foram inseridas no banco de dados.

5.1.2.3. FERRAMENTA DE ANÁLISE DE RESISTÊNCIA

A ferramenta de análise de resistência aos antirretrovirais e de sítios pós-traducionais foi desenvolvida e hospedada no servidor Linux, localizado no Laboratório Avançado de Saúde Pública no Centro de Pesquisas Gonçalo Moniz (LASP/CPqGM/FIOCRUZ), e pode ser acessada através do link <http://lasp.bahia.fiocruz.br>.

5.2. PARTE 2: DESCRIÇÃO DA UTILIZAÇÃO DA FERRAMENTA

5.2.1. SUBMISSÃO DAS SEQUÊNCIAS PARA ANÁLISE

Em sua tela inicial (FIGURA 14), a ferramenta apresenta duas tabelas. A primeira tabela exibe números que representam a quantidade e o percentual de seqüências brasileiras do HIV-1, dispostas por subtipo. A segunda tabela representa a quantidade e o percentual de seqüências das regiões *Protease* e *Trasncriptase Reversado* armazenadas no Banco de Dados. A tela inicial apresenta também uma caixa de texto onde os usuários podem incluir suas próprias seqüências para serem submetidas à ferramenta (FIGURA 15).

A ferramenta também possibilita ao usuário a submissão de seqüências armazenadas em arquivo FASTA, bastando para isso utilizar o botão "Procurar". Em seguida o usuário deve selecionar o arquivo no formato FASTA para a realização das análises (FIGURA 16).



[Translate English](#)

Observe que : Esta ferramenta foi projetada para a análise de Resistência do HIV-1 aos antirretrovirais nas regiões específicas da transcriptase reversa e da protease.

A Ferramenta LASP-HIV1-ResTool possui um bom desempenho com até 1.000 sequências de protease e 500 sequências de transcriptase reversa.

subtipo	quantidade	%
A	1	0.01%
B	4142	46.06%
C	501	5.57%
D	70	0.78%
F	57	0.63%
Y	4220	46.93%
Total	8991	100%

Adicione abaixo as sequências no formato Fasta

região	quantidade	%
PS1	57	0.63%
PROTEASE	2626	29.2%
Total	2683	29.84%

Selecione a região genômica

[Submeter sequências](#)

[Sequências de exemplo](#)

Desenvolvido por: Domingos Ramon Moreau da Cunha, José Carlos Couto Fernandez, Túlio de Oliveira, Bernardo Galvão Castro Filho, Luiz Carlos Júnior Alcântara

Desenvolvido em cooperação com o Laboratório Avançado de Saúde Pública (LASP), CPqGM, FIOCRUZ, Brazil, o Laboratório de Aids e Imunologia Molecular, IOC, FOCRUZ, Brazil e a Universidade de Western Cape - Cape Town - África do Sul.

Questões sobre a Ferramenta de Análise de Resistência do HIV-1: [Ramon Moreau](#).

Sugestões ou problemas: [Dr Luiz Alcântara](#).

Figura 14. Página de apresentação da ferramenta: exibe a quantidade de sequências por cada subtipo e por regiões gênicas (PR e TR); Os links que compõem a ferramenta e as caixas de inserção de sequências para submissão.



Translate English

Observe que : Esta ferramenta foi projetada para a análise de Resistência do HIV-1 aos antirretrovirais nas regiões específicas da transcriptase reversa e da protease.

A Ferramenta LASP-HIV1-ResTool possui um bom desempenho com até 1.000 seqüências de protease e 500 seqüências de transcriptase reversa.

Adicione abaixo as seqüências no formato Fasta

subtipo	quantidade	%
A	1	0.01%
B	4142	46.06%
C	501	5.57%
D	70	0.78%
F	57	0.63%
Y	4220	46.93%
Total	8991	100%

```
>DQ190986
CCTCAGATCACTCTTTGGCAGCGACCCCTAGTCCCAATAAAAAGTAGGAGGGCAACTAAA
GGAAGCTCTATTAGACACAGGAGCAGATGATACAGTATTAGAAGAAATAGATTTGCCAG
GAAAATGGAARACCAAAAATAATAGGGGGAATTGGAGTTTTGTAAAAGTAAGACAGTAT
GATCAGATACCCATAGAAATTTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGGCC
TACACCTGCCAACATAATTGGAAGAAATTTGATGACTCAGATTGGCTGTACTTTAAAT
```

Procurar...

região	quantidade	%
PS1	57	0.63%
PROTEASE	2626	29.2%
Total	2683	29.84%

Selecione a região genômica **PROTEASE**

Subtipo

B C D F G J

Droga

Tipranavir/r

Indinavir/r

Saquinavir/r

Lopinavir/r

Fosamprenavir/r

Atazanavir/r

Nelfinavir

Darunavir/r

Enviar

[Submeter seqüências](#)

[Seqüências de exemplo](#)

Desenvolvido por: Domingos Ramon Moreau da Cunha, José Carlos Couto Fernandez, Túlio de Oliveira, Bernardo Galvão Castro Filho, Luiz Carlos Júnior Alcântara

Desenvolvido em cooperação com o Laboratório Avançado de Saúde Pública (LASP), CPqGM, FIOCRUZ, Brazil, o Laboratório de Aids e Imunologia Molecular, IOC, FIOCRUZ, Brazil e a Universidade de Western Cape - Cape Town - África do Sul.

Questões sobre a Ferramenta de Análise de Resistência do HIV-1: [Ramon Moreau](#).

Sugestões ou problemas: [Dr Luiz Alcântara](#).

FIGURA 15. Página utilizada para a escolha dos parâmetros de análise (subtipo, região gênica e droga antirretroviral).

Escolher arquivo a carregar

Examinar: Sequências

hiv-db

Nome do arquivo: hiv-db

Arquivos do tipo: Todos os Arquivos (*.*)

Botões: Abrir, Cancelar

subtipo quantidade %

A	1	0.0
B	4142	46.0
C	501	5.5
D	70	0.7
F	57	0.6
Y	4220	46.9
Total	8991	100

região quantidade %

P51	57	0.6
PROTEASE	2626	29.2
Total	2683	29.84%

[Submeter sequências](#)

[Sequências de exemplo](#)

Desenvolvido por: Domingos Ramon Moreau da Cunha, José Carlos Couto Fernandez, Túlio de Oliveira, Bernardo Galvão Castro Filho, Luiz Carlos Júnior Alcântara

Desenvolvido em cooperação com o Laboratório Avançado de Saúde Pública (LASP), CPqGM, FIOCRUZ, Brazil, o Laboratório de Aids e Imunologia Molecular, IOC, FIOCRUZ, Brazil e a Universidade de Western Cape - Cape Town - África do Sul.

Questões sobre a Ferramenta de Análise de Resistência do HIV-1: [Ramon Moreau](#).
 Sugestões ou problemas: [Dr Luiz Alcântara](#).

FIGURA 16. Página utilizada para a submissão das sequências de HIV-1 em arquivos no formato FASTA.

Em sua tela inicial, a Ferramenta de Análise de Resistência aos antirretrovirais, também possui os seguintes **Links**: **Submeter sequências**, **Como citar**, **Tutorial**, **Sequências de exemplo** e **Contatos**.

- 1) **Submeter sequências** retorna para a página inicial da ferramenta.
- 2) **Como citar** exibe uma tela que orienta como o usuário deve fazer a citação da ferramenta para análise de resistência do HIV-1 aos antirretrovirais LASP-HIV1-ResTool.
- 3) **Tutorial**: exibe um tutorial completo da ferramenta, com um passo a passo simulando, através de imagens, cada tela da LASP-HIV1_ResTool.
- 4) **Sequências de exemplo**: exibe uma tela com oito sequências de HIV-1, sendo quatro correspondentes a Transcriptase Reversa e as outras quatro correspondentes a protease.
- 5) **Contatos**: exibe uma tela com o e-mail dos pesquisadores envolvidos no projeto, para esclarecimento de possíveis dúvidas sobre a utilização da ferramenta. Adicionalmente pode ser usado para registro e envio de sugestões ou de aviso de falhas no sistema.

Caso o usuário não adicione nenhuma seqüência, a ferramenta irá considerar apenas as sequências do Banco de dados local para a análise. Posteriormente, o usuário deve selecionar a região gênica correspondente a seqüência que deseja submeter ou as do banco de dados local, as quais ele tenha interesse. Após a seleção da região gênica, a ferramenta exibirá os subtipos de HIV-1 que estão armazenados no banco de dados e as drogas antirretrovirais que atuam sobre a região gênica escolhida. Depois de seguir os passos anteriores, o usuário deve selecionar o subtipo e a droga de seu interesse e somente após esses procedimentos submeter seus dados para análise através do botão 'Enviar'.

Após a submissão das sequências, o sistema exibe os resultados na tela seguinte (FIGURA 17).

5.2.2. RESULTADO DA ANÁLISE DAS SEQUÊNCIAS SUBMETIDAS À FERRAMENTA

Como resultado do processamento, o sistema identifica a sequência através do número de acesso e localiza a posição em que a mesma se inicia dentro da região gênica escolhida. Em seguida, o sistema identifica a fase de leitura correta da sequência e realiza a conversão da sequência de nucleotídeos em sequência de aminoácidos. Inicia-se então o processo de localização de todas as mutações ocorridas, tanto na sequência de ácido nucléico como na sequência de aminoácido. O sistema então apresenta uma tabela apenas com as mutações que conferem a resistência do HIV-1 contra o antirretroviral. Além disso, exibe o grau de resistência de cada mutação e indica, ao final da tabela, se a amostra de HIV-1 submetida na análise é susceptível, apresenta resistência intermediária ou se é resistente ao antirretroviral. Adicionalmente, a ferramenta identifica a presença de sítios pós-traducionais de glicosilação e fosforilação, suas respectivas posições na sequência de aminoácidos, e exibe o gráfico que relaciona quantidade de sítios pós-traducionais por sequência.

Se o usuário não inserir nenhuma sequência, o sistema exibe na tela seguinte os parâmetros relativos à análise (número de acesso da sequência, subtipo, região gênica, antirretroviral, as mutações e suas posições, o grau de resistência de cada mutação e o grau de resistência total (Figura 17).

Ao final da página, o sistema exibe dois gráficos que relacionam: a quantidade de sítios pós-traducionais em função do total de sequências submetidas e os padrões de resistência aos antirretrovirais (susceptível, intermediária e resistente) em função do total de sequências. Além de apresentar um gráfico de barras mostrando a quantidade de ocorrências de cada mutação de resistência no conjunto de sequências analisadas (Figura 18).



Análise de resistência a TIPRANA VIR em amostras de PROTEASE do subtipo B do Brasil.

Número de Amoças: 2021811

Primeira base de letters

Posição Inicial: 1

Mutações na Sequência de Nucleotídeos:

G4A G7A G41A A189G G110A T189C G258A A277C T282C

Mutações na Sequência de Aminoácidos:

VH S47D L459 Y771 I94L

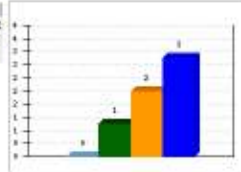
Tabela de Resistência

Mutação/Posição Resistência Resistência Total

0 0 0 0

Sites Pós-Tradução

PKC_Resistência	CR1_Resistência	Mutação
12TR14	26110128	2704207132
	31791214	4600306910
		840071396



Número de Amoças: 2021811

Primeira base de letters

Posição Inicial: 1

Mutações na Sequência de Nucleotídeos:

G7A A110 C16A A181C G41A A48D G48A T18C Q188A A189G Q191A G121A Q181A A189G C182A T189C C187T T281C A214Q T240C

Mutações na Sequência de Aminoácidos:

VH T12P I15V I168I S17D E41K S44E S44V L45F V524 L90L

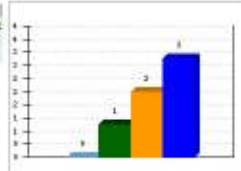
Tabela de Resistência

Mutação/Posição Resistência Resistência Total

S44E	2	
S44V	12	
V524	8	
L90L	2	28

Sites Pós-Tradução

PKC_Resistência	CR1_Resistência	Mutação
12TR14	26110128	2704207132
	31791214	4600306910
		840071396



Número de Amoças: 2021811

Primeira base de letters

Posição Inicial: 1

Mutações na Sequência de Nucleotídeos:

G7A A20T Q198A Q199A A120Q A184Q T189C A190C A214Q A218Q T281C C187T T189C

Mutações na Sequência de Aminoácidos:

VH S48E E37H S21V L459 I94L T12V

Tabela de Resistência

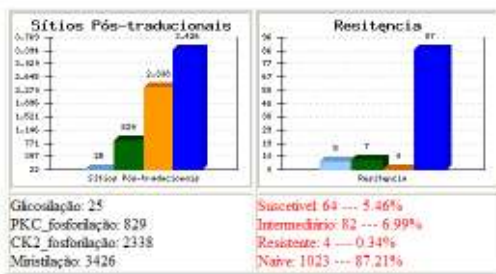
Mutação/Posição Resistência Resistência Total

0 0 0 0

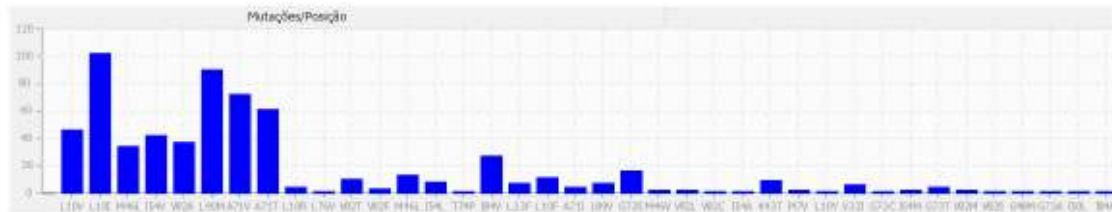
Sites Pós-Tradução

PKC_Resistência	CR1_Resistência	Mutação
12TR14	26110128	2704207132

Figura 17. Topo da tela de resultados da análise de resistência aos antirretrovirais e sítios pós-traducionais.



Sequências Submetidas --> 1174



[Acessar o arquivo XML](#)

FIGURA 18. Base da tela de resultados da análise de resistência aos antirretrovirais e sítios pós-traducionais.

O número de acesso de cada sequência é apresentado como um link que permite o acesso do usuário a uma nova página contendo a seqüência de nucleotídeos e sua respectiva seqüência de aminoácidos (FIGURA 19).

Seqüência de nucleotídeos

```
>DQ358810
CCTCAAATCACTCTTTGGCAACGACCCCTCGTCAACAATAAGATAGGAGGGCAACTAAAGGAAGCTCTAT
TAGATACAGGACAGATGATACAGTATTAGAAGAAATGGATTGCCAGGAAGATGGAAACCAGAAATGATA
GGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACCCATAGAAATCTGTGGACATAAAGC
TATAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGCTTGGCT
GCACITTAAT
```

Seqüência de aminoácidos

```
>DQ358810
PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMDLPGRWKPKMIGGIGGFIKVRQYDQPIEICGHK
AIGTVLIGTPVNIIGRNLLTQLGCTLN
```

[Verar](#)

FIGURA 19. Página contendo a seqüência de nucleotídeo submetida e sua respectiva tradução em aminoácidos.

O usuário pode ter acesso às informações geradas pela ferramenta, através de um link exposto no final da página. Este link direciona o usuário a uma página em formato XML contendo todas as informações de suas análises (FIGURA 20). A página XML pode ser salva para um arquivo utilizado para transportar os dados. Após salvo, o arquivo.xml pode ser importado de outras ferramentas tais como: planilhas eletrônicas (excel), programas de análise estatística (SPSS), programas de bancos de dados (access).

```

<?xml version="1.0" encoding="iso-8859-1" ?>
- <!--
  Ferramenta de análise de resistência aos antirretrovirais no formato XML

  Resultado da análise de amostras de HIV-1 (26-02-2010 - 15:37)

  Laboratório Avançado de Saúde Pública (LABSP/CPqSM/FIOCRUZ)

  Laboratório de AIDS e Imunologia Molecular (LABAIDS/IOC/FIOCRUZ)
-->
- <ResistenciaAnalise>
- <DQ358810>

  <Sequencia>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCACAATAAAGATAGGAGGGCCAACTAAAGGAAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGGATTGC
</Sequencia>
  <FaseLeitura>Primeira fase de leitura</FaseLeitura>
  <Inicio>1</Inicio>
  <mutacaoAN>G6A</mutacaoAN>
  <mutacaoAN>G7A</mutacaoAN>
  <mutacaoAN>G48A</mutacaoAN>
  <mutacaoAN>A109G</mutacaoAN>
  <mutacaoAN>G110A</mutacaoAN>
  <mutacaoAN>T188C</mutacaoAN>
  <mutacaoAN>G229A</mutacaoAN>
  <mutacaoAN>A277C</mutacaoAN>
  <mutacaoAN>T282C</mutacaoAN>
  <mutacaoAA>V3I</mutacaoAA>
  <mutacaoAA>S37D</mutacaoAA>
  <mutacaoAA>L63P</mutacaoAA>
  <mutacaoAA>V77I</mutacaoAA>
  <mutacaoAA>I93L</mutacaoAA>
- <Resistencia>
  <posicao />
</Resistencia>
- <SitosPostraducionais>
  <fosforilacao>12TIK14</fosforilacao>
  <CK2fosforilacao>26TGAD29</CK2fosforilacao>
  <CK2fosforilacao>31TVLE34</CK2fosforilacao>
  <miristilacao>27GADDTV32</miristilacao>
  <miristilacao>48GGIGGF53</miristilacao>
  <miristilacao>94GCTLN98</miristilacao>
</SitosPostraducionais>
</DQ358810>
+ <DQ358809>
+ <DQ358808>
+ <DQ358805>
+ <DQ190987>
+ <DQ190988>
+ <DQ190986>
+ <DQ190982>
+ <DQ190981>
+ <DQ190980>
+ <DQ190979>
+ <DQ190974>
+ <DQ190978>
+ <DQ190969>
+ <DQ190973>
+ <DQ190968>
+ <DQ190967>
+ <DQ190965>
+ <DQ190966>
+ <DQ190960>
+ <DQ190961>
+ <DQ190964>
+ <DQ190953>
+ <DQ190954>
+ <AY569880>
+ <AY569881>
+ <AY569883>
+ <AY569884>
+ <AY569885>
+ <AY569886>
+ <AY569887>
+ <AY569888>
+ <AY569890>
+ <AY569891>
+ <AY569892>

```

FIGURA 20. Página contendo as análises no formato XML para serem armazenadas.

5.3. PARTE 3: DESCRIÇÃO DA UTILIZAÇÃO DA FERRAMENTA

5.3.1. AVALIAÇÃO COMPARATIVA DOS RESULTADOS DAS ANÁLISES GERADAS PELA FERRAMENTA.

Com o objetivo de confrontar os dados gerados após a submissão das sequências às ferramentas de análise de resistência LASP-HIV1-ResTool (desenvolvida nesse trabalho) e Hiv Drug Resistance Database (http://hivdb.stanford.edu/pages/algs/sierra_sequence.html) (ferramenta de Stanford), criamos um script, também na linguagem de programação PHP. O Script ao ser executado processa os arquivos XML, gerado pelo LASP-HIV1-ResTool, e HTML gerado pelo Hiv Drug Resistance Database e compara os dados de resistência obtidos nas sequência de cada arquivo, nas duas ferramentas. Em seguida o script calcula o percentual de similaridade entre os resultados. O mesmo procedimento é utilizado para a análise comparativa entre o ferramenta LASP-HIV1-ResTool e a ferramenta de interpretação do algoritmo brasileiro RENAGENO (http://forrest.ime.usp.br:3001/public_analyses), neste caso porém, o script lê o arquivo XML gerado pelo LASP-HIV1-ResTool e o arquivo HTML gerado pela Sistema RENAGENO. Este procedimento é de grande relevância para avaliar a eficácia da ferramenta LASP-HIV1-ResTool, desenvolvida no presente trabalho.

Para validação da aplicabilidade da ferramenta LASP-HIV1-Restool foram utilizadas 111 amostras correspondentes a região gênica da protease e da transcriptase reversa derivadas da genotipagem da resistência do HIV-1 provenientes de pacientes infectados, que apresentavam falha virológica a TARV.

Os gráficos abaixo mostram o padrão de resistência das sequências submetidas a três diferentes classes de antiretrovirais. O gráfico 1 foi gerado utilizando dados obtidos na ferramenta desenvolvida nesse trabalho (LASP-HIV1-ResTool), o gráfico 2 a partir de análises utilizando dados gerados pela ferramenta de análise de resistência de Stanford (Drug Resistance Database de) e o gráfico 3 foi gerado pelos dados resultantes das análises obtidas através do sistema brasileiro (RENAGENO). Pode-se observar a similaridade entre os resultados obtidos utilizando a nossa ferramenta e a de Stanford (comparação entre os gráficos 1 e 2). Por outro lado, pode-se observar diferenças entre os resultados das análises realizadas com a utilização o sistema RENAGENO, tanto em relação a nossa ferramenta (comparação entre os gráficos 3 e 1) quanto em relação à de Stanford (comparação entre os gráficos 2 e 3). Houve diferença estatística, especificamente da faixa de sensibilidade intermediária aos IP, NRTI e NNRTI. De

maneira global as Análises de variância apontaram $p > 0,05$ para a comparação entre o gráfico 2 e o gráfico 3 e $p < 0,05$ para as comparações entre 3 e 1 e entre 3 e 2, através do teste ANOVA fator único.

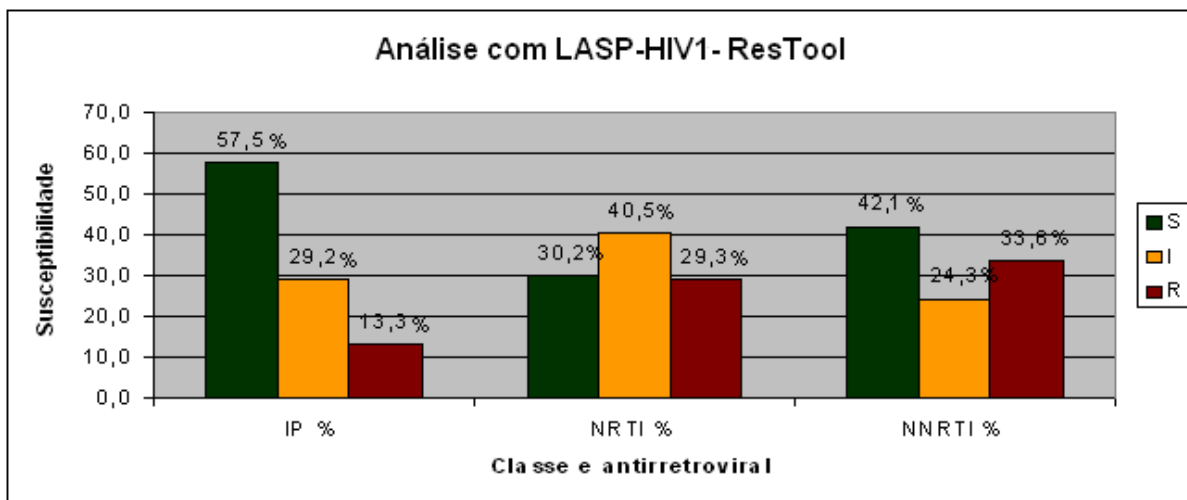


Gráfico 1: Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta LASP-HIV1-ResTool. Os níveis de sensibilidade aos ARV estão representados em verde para susceptível (S), em amarelo para resistência intermediária (I) e em vermelho para resistência completa (R). Os valores estão expressos em níveis percentuais.

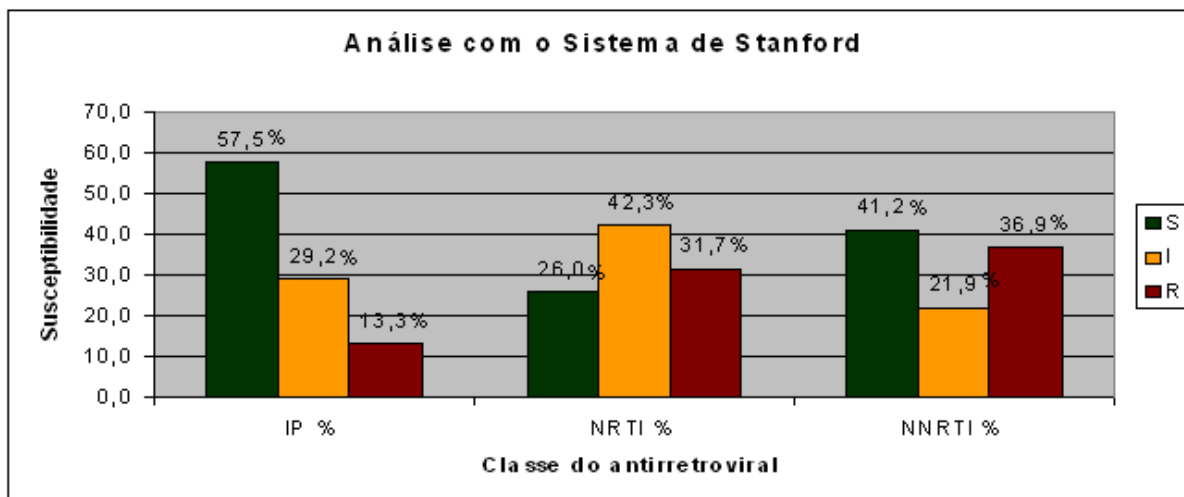


Gráfico 2: Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta de Stanford. Os níveis de sensibilidade aos ARV estão representados em verde para susceptível (S), em amarelo para resistência intermediária (I) e em vermelho para resistência completa (R).

(I) e em vermelho para resistência completa (R). Os valores estão expressos em níveis percentuais.

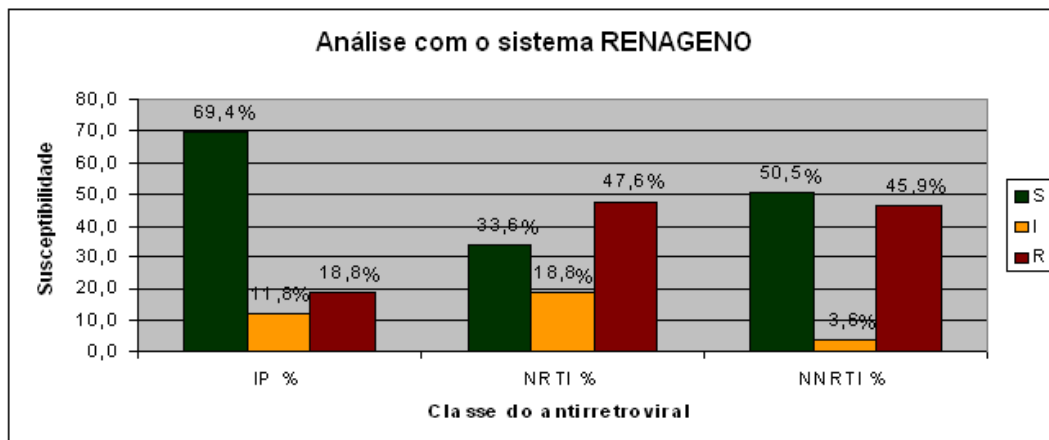


Gráfico 3: Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados na ferramenta da RENAGENO. Os níveis de sensibilidade aos ARV estão representados em verde para susceptível (S), em amarelo para resistência intermediária (I) e em vermelho para resistência completa (R). Os valores estão expressos em níveis percentuais.

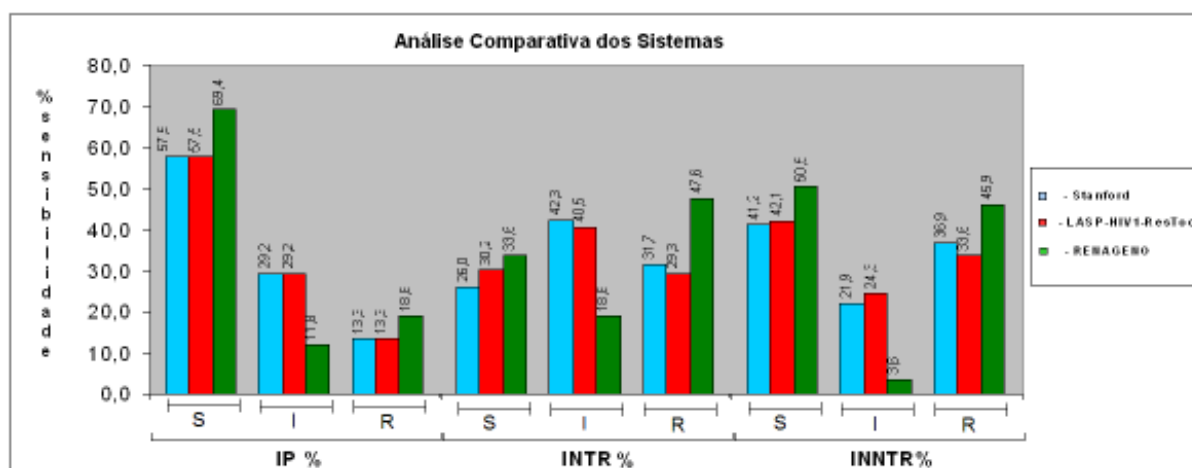


Gráfico 4: Avaliação dos perfis de resistência aos inibidores de protease (IP), inibidores nucleosídeos da transcriptase reversa (INTR) e inibidores não nucleosídeos (INNTR) realizados nas ferramentas LASP-HIV1-ResTool, RENAGENO e HIV Drug Resistance Database. As barras azuis apresentam os resultados das análises de resistência realizadas no HIV Drug Resistance Database, as barras vermelhas apresentam os resultados das análises de resistência realizadas no LASP-HIV1-ResTool e as barras verdes representam os resultados de resistência realizadas no sistema RENAGENO.

DISCUSSÃO

6.0. DISCUSSÃO

Apesar dos benefícios alcançados nos últimos anos pela terapia antirretroviral combinada de alta potência, o desenvolvimento de resistência do HIV-1 aos antirretrovirais é um fator limitante para o sucesso da TARV. A seleção de populações virais contendo no seu genoma mutações de resistência aos antirretrovirais, representa um desafio importante na reestruturação do tratamento de pacientes com vírus multirresistentes. O desenvolvimento de resistência a várias classes de antirretrovirais, limita o número de novos regimes terapêuticos e o sucesso virológico de terapias de resgate subsequentes.

Destinados a auxiliar a estruturação terapêutica destes pacientes surgiram os testes de avaliação de resistência ou de susceptibilidade do HIV às drogas, sendo rapidamente incorporados ao monitoramento e tratamento de pacientes com HIV/AIDS (DURANT *et al.*, 1999). No entanto, o benefício dos testes genotípicos de caracterização da resistência viral é limitado pela necessidade de interpretação do perfil mutacional para se obter uma previsão fidedigna da susceptibilidade à cada droga antirretroviral.

Existem atualmente, diversos sistemas de interpretação de resistência do HIV aos antirretrovirais disponíveis. Alguns destes sistemas são comercializados enquanto outros podem ser adquiridos ou acessados gratuitamente através dos web sites de seus desenvolvedores. Poucos estudos, porém, foram feitos para avaliar a similaridade entre os resultados das análises de tais sistemas. No presente trabalho, desenvolvemos uma ferramenta de bioinformática para análise de mutações de resistência do HIV-1 aos antirretrovirais entre outras funcionalidades incluindo avaliação de sítios pós-traducionais. Trata-se de uma ferramenta de fácil manuseio, que ficará disponível via web nos sites da unidade de bioinformática do LASP/CPqGM/FIOCRUZ e Bioafrica, para utilização em nível de domínio público, sendo capaz de otimizar a análise de dados genômicos de amostras de HIV-1. Esta ferramenta pode acessar um banco de dados com seqüências gênicas previamente armazenadas ou através de deposição de seqüências FASTA via internet, extrair informações relativas, comparar seqüências de HIV-1 dos bancos de dados públicos, identificando mutações de resistência aos diferentes antirretrovirais e sítios pós-traducionais. Além disso, associa esses dados à resistência e susceptibilidade a cada antirretroviral utilizado nas terapias anti-HIV/AIDS.

6.1. SÍTIOS PÓS-TRADUCIONAIS

Os níveis de glicosilação das proteínas do envelope viral tem um papel crucial na patogênese viral, impactando diretamente na conformação estrutural e funcional do envelope, influenciando a antigenicidade e imunogenicidade da proteína. Consequentemente, está envolvida nos mecanismos de evasão imune do HIV, através de mudanças conformacionais e ocultação de epítopos imunorelevantes (REITTER *et al.*, 1998; WEI *et al.*, 2003).

A estrutura trimérica do envelope é extensivamente glicosilada, com pelo menos 24 sítios potenciais de glicosilação espalhados pelos domínios conservados e variáveis da gp120 e entre 4 a 5 sítios potenciais localizados na porção extra membranar de gp41 (MIZUOCHI *et al.*, 1990; IRUNGU *et al.*, 2008).

As variações no padrão global de glicosilação do envelope avaliando sequências recuperadas do banco genômico de Los Alamos identificou pelo menos 51 padrões de glicosilação da proteína do envelope (KORBER *et al.*, 2001) contribuindo para o conhecimento de variações nos níveis de imunogenicidade da proteína env.

Portanto a identificação e conhecimento dos níveis e padrões de glicosilação de isolados virais provenientes de pacientes brasileiros são fundamentais para o desenvolvimento de futuros imunógenos destinados a indução de uma resposta imune humoral protetora contra o HIV-1 no Brasil.

Nossa ferramenta (LASP-HIV1-ResTool) é capaz de realizar este tipo de análise, pois apresenta em seu código a funcionalidade necessária para interpretar as regras utilizadas para a identificação dos sítios de glicosilação e fosforilação. Além da identificação dos sítios de glicosilação e fosforilação, a ferramenta de análise HIV-1ResTool foi estruturada para incorporar novas regras que sejam de interesse do usuário.

A ferramenta de bioinformática LASP-HIV1-Restool pode ser usada para identificar diferentes perfis de glicosilação do envelope de diferentes subtipos de HIV-1 e formas recombinantes intersubtipos, bem como identificar epítopos específicos e carboidratos em cada local de glicosilação. Essas informações podem contribuir para avaliação dos perfis de infectividade associados a cada subtipos do HIV-1 e em estudo para modulação da imunogenicidade por anticorpos neutralizantes.

No presente estudo, mostramos que existem diferenças nos níveis de interpretações genótípicas entre os algoritmos testados. Isto está em concordância com estudos anteriores, que encontraram diferenças entre algoritmos de interpretação para um determinado genótipo [DI GIAMBENEDETTO *et al.*, 2002; STURMER *et al.*, 2003]. A taxa

de discordância em nosso estudo é significativamente maior para a previsão de resistência aos inibidores nucleosídeos da TR e inibidores da protease, especificamente relacionada à faixa de sensibilidade intermediária. Isto se deve basicamente ao sistema de pontuação adotado pela nossa ferramenta que se assemelha mais ao modelo de Stanford que ao algoritmo RENAGENO que utiliza fundamentalmente uma regra de interpretação. Outro aspecto que possivelmente pode estar contribuindo com essa diferença será a junção dos sistemas de pontuação de Stanford onde agregamos as três faixas de sensibilidade intermediária para uma faixa na nossa ferramenta. Isso possivelmente irá super estimar os níveis de resistência do HIV-1 aos antirretrovirais, mas sem um impacto significativo nos níveis de sensibilidade e/ou resistência do vírus.

Atualmente, diferentes sistemas de interpretação de resistência do HIV-1 estão disponíveis comercialmente ou gratuitamente, algumas destas ferramentas apresentam interface de dados fenotípicos de resistência às drogas em combinação com o sistema de genotipagem.

Neste estudo, além do desenvolvimento da ferramenta LASP-HIV1-ResTool, realizamos a comparação dos resultados obtidos com amostras de pacientes apresentando falha a terapias antirretrovirais prévias, utilizando outros dois sistemas de diferente interpretação de resistência do HIV, visando a identificação de padrões mutacionais responsáveis por possíveis discordâncias entre estes sistemas.

Durante as etapas iniciais do desenvolvimento da ferramenta de Bioinformática percebemos que a organização das seqüências do HIV-1 é uma etapa fundamental do processo relacionado a de banco de dados. Neste sentido e visando o desenvolvimento de um sistema eficaz através de softwares livres (licenças públicas), utilizamos o sistema gerenciador de bancos de dados MYSQL(<http://www.mysql.com>). Este sistema já consagrado no mundo e bastante utilizado em diferentes seguimentos da informática, se mostrou eficaz e útil para o armazenamento e manipulação de sequências de HIV-1 obtidas através do GeneBank ou depositadas diretamente no sistema. O banco de dados, assim como a ferramenta LASP-HIV1-ResTool ficam hospedados em um servidor Linux, no LASP (Laboratório Avançado de Saúde Pública/CPqGM/FIOCRUZ).

Outro importante Sistema que se mostrou útil e ágil para a organização e manipulação das sequências do HIV-1 foi o programa HIVbase (<http://www.genejohnson.net>). Através deste sistema foi possível importar sequências de ácidos nucléicos ou aminoácidos no formato FASTA. Foi possível agilizar o processo de estratificação das sequências por regiões gênicas, subtipo, país de origem, além de outra informações relacionadas as sequências. Entretanto o HIVbase é um sistema proprietário

e comercializado a um custo substancial. Isto onera significativamente na criação de banco de dados utilizando esse programa. Por outro lado, a agilidade e eficácia do sistema durante o desenvolvimento da ferramenta mostrou uma relação de custo/benefício que justificou sua aquisição e ampla utilização em diferentes seguimentos institucionais (LASP/CPqGM/Fiocruz e IOC/FIOCRUZ).

Os “scripts” utilizados para a organização e o armazenamento das sequências foram desenvolvidos durante o processo de programação da ferramenta de bioinformática. Os demais programas utilizados mostraram-se confiáveis e eficientes no suporte a ferramenta de bioinformática para análise de resistência. O PHP é uma linguagem bastante utilizada na informática e se mostrou uma linguagem poderosa e versátil para o criação dos “scripts” e no desenvolvimento da ferramenta LASP-HIV1-ResTool.

Para efeito de programação da ferramenta para avaliação das mutações de resistência do HIV aos antirretrovirais, utilizamos o sistema de pontuação desenvolvido pela Universidade de Stanford (HIV Drug Resistance Database), com alguns ajustes realizados na faixa intermediária de sensibilidade aos antirretrovirais. Fundamentalmente, consideramos, para efeito de análise, a faixa superior (“Potential High-level resistance”) e inferior (“Low-level resistance”) do nível intermediário de resistência em uma única faixa de sensibilidade. Este sistema de três níveis de avaliação para a terapia antirretroviral é similar ao sistema utilizado pelo algoritmo brasileiro de interpretação de resistência do HIV-1.

As ferramentas de bioinformática desenvolvidas para realizar análises de resistência do HIV-1 aos antirretrovirais, frente aos diferentes perfis mutacionais, juntamente com os algoritmos de resistência, são de grande utilidade para dar suporte as terapias antirretrovirais. Existem, porém, inúmeras variáveis que acarretam grande complexidade na escolha da terapia mais adequada para cada paciente, tais como multifalhas ou resistência ampla, adesão do paciente ao antirretroviral, as variações metabólicas diferenciadas para cada indivíduo. Para que seja feita uma análise ampla e consistente de todas as variáveis envolvidas na terapia de cada paciente infectado pelo HIV é de suma importância a participação de médicos com grande conhecimento e experiência, capazes de conduzir a terapia mais adequada a cada paciente, levando também em consideração o estado clínica, psicológico e social de cada paciente.

6.2. AVALIAÇÃO COMPARATIVA DOS RESULTADOS DAS ANÁLISES GERADAS PELA FERRAMENTA.

Para a análise comparativa realizada para validação da aplicabilidade da ferramenta LASP-HIV1-ResTool utilizamos sequências de HIV-1 obtidas da genotipagem da resistência do HIV realizada pela RENAGENO em pacientes do Rio de Janeiro apresentando falha virológica a terapia.

Não podemos observar diferenças significativas nos padrões de interpretação entre a nossa ferramenta (gráfico 1) e os padrões de interpretação produzidos por Stanford (gráfico 2). Aparentemente a incorporação dos valores de pontuação em uma única faixa de sensibilidade intermediária, não mostrou impacto significativo na estimação da resistência viral, fator este preocupante no sentido de incorporação de erros de interpretação realizados pela ferramenta.

Por outro lado, foi possível observar diferenças significativas entre os perfis de interpretação da nossa ferramenta e Stanford, comparativamente aos perfis de avaliação do algoritmo brasileiro da RENAGENO (gráfico 3). Houve diferença estatística significativa, especificamente na faixa de sensibilidade intermediária aos inibidores nucleosídeos e não nucleosídeos da TR e inibidores da protease. De maneira global as Análises de variância apontaram $p > 0,05$ para a comparação entre a ferramenta LASP-HIV1-ResTool (gráfico 1) e Stanford (gráfico 2). Por outro lado observamos uma variância de $p < 0,05$ para as comparações realizadas entre a nossa ferramenta (gráfico 1) e o algoritmo brasileiro (gráfico 3), através do teste ANOVA fator único.

Ao realizarmos a análise comparativa entre o nosso sistema e o sistema RENAGENO (http://forrest.ime.usp.br:3001/public_analyses), podemos observar no gráfico 4 uma discordância entre a interpretação do perfil mutacional de resistência realizado pelo o LASP-HIV1-ResTool e o sistema RENAGENO. Esta diferença fica bastante evidente, principalmente nas faixas intermediárias de resistência, para as três categorias de antirretrovirais. A pequena diferença na interpretação das mutações de resistência encontrada entre o LASP-HIV1-ResTool e o sistema de Stanford provavelmente é resultado das diferenças observadas na quantidade de perfis intermediários de classificação de resistência, consideradas pelos dois sistemas. No LASP-HIV1-ResTool são consideradas três faixas de classificação (susceptível, intermediária e resistente), enquanto que em Stanford são utilizadas cinco faixas (susceptível, baixa resistência, resistência potencial baixa, resistência intermediária e alta resistência). Contudo não podemos detectar na nossa avaliação uma subestimativa dos

níveis de resistência pela nossa ferramenta. Ou seja, aqueles indivíduos que apresentavam baixos níveis de resistência aos antirretrovirais permaneceram na faixa intermediária de avaliação, confirmando assim a fidelidade na avaliação do perfil mutacional de resistência do HIV-1. O mesmo aconteceu com relação aos indivíduos com elevado risco de resistência completa registrada pelo seu perfil mutacional. A grande maioria das sequências permaneceu na faixa intermediária de sensibilidade e poucas sequências mudaram para a faixa de resistência completa.

Por outro lado fica a necessidade de estudos estatísticos mais apurados, especificamente aplicados para comprovação de significância de diferenças de pontuação/avaliação frente a cada droga antirretroviral inter e intra-algoritmo de interpretação.

Concluindo, nossos resultados mostram a viabilidade e eficiência da ferramenta LASP-HIV1-ResTool para interpretação de mutações de resistências aos antirretrovirais. A ferramenta LASP-HIV1-ResTool mostrou ter um excelente custo/benefício, ser de fácil aplicação e de livre acesso.

Além disso, os resultados descritos no presente trabalho demonstram a necessidade de uma uniformização entre os sistemas de análise de resistência do HIV aos antirretrovirais. Para isso ser possível torna-se necessária uma padronização uniforme e integrada dos fundamentos e sistemas de avaliação dos diferentes algoritmos de interpretação de resistência do HIV-1.

CONCLUSÕES

7. CONCLUSÕES

Neste trabalho realizamos o desenvolvimento uma ferramenta de bioinformática capaz de realizar análise de mutações de resistência do HIV-1 aos antirretrovirais, identificação de sítios pós-translacionais entre outras funcionalidades.

Utilizando programas de livre acesso foi possível construir uma ferramenta de fácil utilização e baixo custo, capaz de identificar a ocorrência de mutações de resistência do HIV-1 nas sequências de nucleotídeos e aminoácidos e indicar o potencial de resistência para cada antirretroviral.

Foi possível desenvolver também um banco de dados local onde estão armazenadas sequências brasileiras do HIV-1, assim como informações a elas relacionadas. Este banco de dados estará aberto à deposição de novas sequências, tanto brasileiras quanto de quaisquer outros pais.

A ferramenta de bioinformática LASP-HIV1-ResTool, mostrou a capacidade de localizar sítios pós-traducionais de glicosilação, fosforilação e miristilação em diferentes genes do HIV-1 através de diferentes sequências genômicas submetidas ao sistema.

A análise comparativa da ferramenta LASP-HIV1-ResTool com outros algoritmos de interpretação de resistência mostrou estreita correlação com os perfis de sensibilidade de Standford.

Os resultados destas comparações inter-ferramentas de avaliação de resistência do HIV-1 foram satisfatórios, mas demonstraram a necessidade de um aprimoramento e padronização dos modelos e fundamentos utilizados para este tipo de análise.

A ferramenta de bioinformática LASP-HIV1-ResTool encontra-se disponível para o uso geral e irrestrito de sua funcionalidade, contribuindo assim com o avanço das pesquisas relacionadas a bioinformática.

REFERÊNCIAS BIBLIOGRÁFICAS

8. REFERÊNCIAS BIBLIOGRÁFICAS

Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). **Science** 1983; 220: 868-71.

Bastos FI, Szwarcwald CL. AIDS and pauperization: principal concepts and empirical evidence. **Cad. Saúde Pública**, Rio de Janeiro. 16(Sup. 1): 65-76. 2000.

Bologna G, Yvon C, Duvaud S, Veuthey AL. N-Terminal myristoylation predictions by ensembles of neural networks. **Proteomics**. 2004 Jun;4(6):1626-32.

Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, Dantas MC, Barreira D, Teixeira PR, Tanuri A; Brazilian Network for Drug Resistance Surveillance. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. **AIDS**. 2003 May 2;17(7):1063-9.

Caride E, Brindeiro R, Hertogs K, Larder B, Dehertogh P, Machado E, de Sá CA, Eyer-Silva WA, Sion FS, Passioni LF, Menezes JA, Calazans AR, Tanuri A. Drug-resistant reverse transcriptase genotyping and phenotyping of B and non-B subtypes (F and A) of human immunodeficiency virus type I found in Brazilian patients failing HAART. **Virology**. 2000 Sep 15;275(1):107-15.

Castilho E.A., Bastos F.I., Scwarcwald C.L., Fonseca M.G.M.. AIDS in Brazil: a changing epidemic. **Cad. Saúde Pública**. Rio de Janeiro. 16(Sup. 1): 4-5. 2000.

Clavel F. HIV-2, the West African AIDS virus. **AIDS**. 1987 Sep;1(3):135-40.

Cohen, D.E.; Walker, B.D. Human Immunodeficiency Virus pathogenesis and prospects for immune control in patients with established infection. **Clin. Infect. Dis**. 2001; 32:1756-1768.

Cohen, O.J.; Weissman, D.; Fauci, A.S. The Immunopathogenesis of HIV Infection. In: PAUL, W. E. **Fundamental Immunology**. 4.ed. Philadelphia: Lippincott Raven, 1998.

Cornelissen M, Kampinga G, Zorgdrager F, Goudsmit J. Human immunodeficiency virus type 1 subtypes defined by env show high frequency of recombinant gag genes. The UNAIDS Network for HIV Isolation and Characterization. **J Virol**. 1996 Nov;70(11):8209-12.

Couto-Fernandez JC, Janssens W, Heyndrickx L, Motte J, Fransen K, Peeters M, Delaporte E, Galvão-Castro B, Piot P, van der Groen G. Genetic and antigenic variability of HIV type 1 in Brazil. **AIDS Res Hum Retroviruses**. 1994 Sep;10(9):1157-63.

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

Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, Oliveira SM, Arakaki-Sanchez D, Chequer PJ, Morgado MG. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio de Janeiro, Brazil: assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active antiretroviral therapy. **Mem Inst Oswaldo Cruz**. 2005 Feb;100(1):73-8. Epub 2005 Apr 12.

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

Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, Brun-Vézinet F, Robertson DL, Simon F. Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. **AIDS Res Hum Retroviruses**. 2004 Jun;20(6):666-72.

David M. Knipe, PhD; Peter M. Howley, MD; Diane E. Griffin, MD, PhD; Robert A. Lamb, PhD, ScD; Malcolm A. Martin, MD; Bernard Roizman, ScD; Stephen E. Straus, MD. **Field's Virology**. 5th, 2007, ISBN-10:0781760607, ISBN-13:9780781760607, 3177, 697.

De Clercq E. A new drug combination therapy for treatment-naive patients with HIV-1 infection, consisting of raltegravir, emtricitabine and tenofovir disoproxil fumarate. **Expert Opin Pharmacother**. 2009 Dec;10(17):2935-7.

Diaz, Ricardo S. , MD; Accetturi, Conceição A. ; Sucupira, Maria C..
Emergence of Resistance Mutations Preceding Virologic Failure in Patients Receiving Antiretroviral Therapy. **JAIDS**. 2008 - Volume 49 - Issue 1 - pp 111-113.

Di Giambenedetto S, Baldini F, Cori C, Rizzo MG, Cingolani A, Liuzzi G, Perno CF, De Luca A & Antinori A.
Poor prediction of virological outcomes by the interpretation of HIV-1 resistance to didanosine and abacavir using 11 available rule-based algorithms. **Antiviral Therapy** 2002; **7**:S76.

Durant J, Clevenbergh P, Halfon P, Delgiudice P, Porsin S, Simonet P, Montagne N, Boucher CA, Schapiro JM & Dellamonica P. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. **Lancet**. 1999; 353:2195–2199.

Erickson, J.W.; Gulnik, S.V.; Markowitz, M. Protease inhibitors: resistance, cross-resistance, fitness and the choice of initial and salvage therapies. **AIDS**. 1999; 13: S189-S204. SUPPLEMENTO 3.

Eyer-Silva WA, Morgado MG. Molecular epidemiology of HIV-1 infection in a small Brazilian county: usefulness of envelope and polymerase sequences to epidemiologic studies. *J Acquir Immune Defic Syndr*. 41:664-70. 2006.

Fauci AS, Mavilio D, Kottlilil S. NK cells in HIV infection: paradigm for protection or targets for ambush. **Nat Rev Immunol**. 2005 Nov;5(11):835-43.

Freed E.O.; Martin M.A.. Evidence for a functional interaction between the V1/V2 and C4 domains of human immunodeficiency virus type 1 envelope glycoprotein gp120. **J Virol**. 1994 Apr;68(4):2503-12.

Froland S.S., Jenum P., Lindboe C.F., Wefring K.W., Linnestad Pj, Böhmer T. HIV-1 infection in Norwegian family before 1970. **Lancet**. 1988 Jun 11;1(8598):1344-5.

Gao, F.; Bailes, E.; Robertson, D.L.; Chen, Y.; Rodenburg, C.M.; Michael, S.F.; Cummins, L.B.; Larry, O.A.; Peeters, M.; Shaw, G.M.; Sharp, P.M.; Hahn, B. Origin of HIV-1 on chimpanzee *Pan troglodytes troglodytes*. **Nature**. 1999; 397:436-441.

Gao F., et al., Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. **Nature**. 1999. 397(6718): p. 436-41.

Gao F, Vidal N, Li Y, Trask SA, Chen Y, Kostrikis LG, Ho DD, Kim J, Oh MD, Choe K, Salminen M, Robertson DL, Shaw GM, Hahn BH, Peeters M. Evidence of two distinct subsubtypes within the HIV-1 subtype A radiation. **AIDS Res Hum Retroviruses**. 2001 May 20;17(8):675-88.

Gifford RJ, Liu TF, Rhee SY, Kiuchi M, Hue S, Pillay D, Shafer RW (2009). The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance. **Bioinformatics**, 2009 Mar 20.

Göttlinger HG (2001). HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. pp. 2-28 in HIV Sequence Compendium 2001. Edited by: Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877.

Gottlieb G. J., Ragaz A., Vogel J.V., Friedman-Kien A., Rywlin A.M., Weiner E.A., Ackerman A.B.. A preliminary communication on extensively disseminated Kaposi's sarcoma in young homosexual men. **Am J Dermatopathol**. 1981; 3(2):111-4.

Guimarães ML, dos Santos Moreira A, Loureiro R, Galvão-Castro B, Morgado MG. High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. **AIDS Res Hum Retroviruses**. 2002 Nov; 20;18(17):1261-9.

Hahn, B.H. Viral genes and their products. In: **Textbook on AIDS Medicine**. New York: Williams & Williams. 1994; cap.3, p.21-43.

Hope TJ. Viral RNA export. **Chem Biol**. 1997 May;4(5):335-44.

Hu DJ, Dondero TJ, Rayfield MA, George JR, Schochetman G, Jaffe HW, Luo CC, Kalish ML, Weniger BG, Pau CP, Schable CA, Curran JW. The emerging genetic diversity of HIV. The importance of global surveillance for diagnostics, research, and prevention. **JAMA**. 1996 Jan 17;275(3):210-6.

Hulo N., Bairoch A., Bulliard V., Cerutti L., De Castro E., Langendijk-Genevaux P.S., Pagni M., Sigrist C.J.A. The PROSITE database. **Nucleic Acids Res**. 34:D227-D230(2006).

Hulo N., Bairoch A., Bulliard V., Cerutti L., Cuche B., De Castro E., Lachaize C., Langendijk-Genevaux P.S., Sigrist C.J.A.. The 20 years of PROSITE. **Nucleic Acids Res**. 2007; 14.

Inocencio LA, Pereira AA, Sucupira MC, Fernandez JC, Jorge CP, Souza DF, Fink HT, Diaz RS, Becker IM, Suffert TA, Arruda MB, Macedo O, Simão MB, Tanuri A. Brazilian Network for HIV Drug Resistance Surveillance: a survey of individuals recently diagnosed with HIV. **J Int AIDS Soc**. 2009 Sep 18;12(1):20.

Irungu J, Go EP, Zhang Y, Dalpathado DS, Liao HX, Haynes BF, Desaire H. Comparison of HPLC/ESI-FTICR MS versus MALDI-TOF/TOF MS for glycopeptide analysis of a highly glycosylated HIV envelope glycoprotein. **J Am Soc Mass Spectrom**. 2008 Aug;19(8):1209-20. Epub 2008 May 24.

Jaffe HW, Schochetman G. Group O human immunodeficiency virus-1 infections. **Infect Dis Clin North Am**. 1998 Mar;12(1):39-46.

J. F. Zagury, G. Franchini, M. Reitz, E. Collalti, B. Starcich, L. Hall, K. Fagnoli, L. Jagodzinski, H.-G. Guo, F. Laureş, S. K. Arya, S. Josephs, D. Zagury, F. Wong-Staal, And R. C. Gallo. Genetic variability between isolates of human immunodeficiency virus (HIV) type 2 is comparable to the variability among HIV type 1. **Cell Biology**. 1988 August. Vol. 85, pp. 5941-5945.

Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. **Br Med Bull**. 2001;58:19-42.

Leitner T, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, Wolinsky S, and Korber B. HIV Sequence Compendium 2005 Eds. Published by Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR 06-0680.

Libin P, K Deforche, K Van Laethem, R Camacho and A.M. Vandamme. RegaDB: An open source, community-driven HIV data and analysis management environment. **Fifth European HIV Drug resistance workshop**, 28-30 maart 2007, Cascais – Portugal. Abstract 81. *Reviews in Antiviral Therapy* 2:82-83, 2007. Poster

Libin P, M Assel, A Loehden, K Theys, K Van Laethem, E Van Wijngaerden, G Lapadula, C Torti, A De Luca, L Ruiz, B Clotet, S Wesner, P Sloot and A-M Vandamme. Applying the RegaDB platform for data integration in the ViroLab project. **Sixth European HIV Drug Resistance Workshop**. From basic science to clinical implications. Budapest Hungary, 26-28 March 2008. Abstract 10. *Reviews in Antiviral Therapy*, 2: 55, 2008 Poster.

Ministério da Saúde - Boletim Epidemiológico - Aids e DST 2008. Ano V, Nº 01.

Ministério da Saúde - Boletim Epidemiológico - Aids e DST 2007. Ano IV, Nº 01.

Ministério da Saúde - Boletim Epidemiológico - Aids e DST 2004. Ano IV, Nº 01.

Regina Loureiro, Deborah Blank, Ana Rejane Wortmann, Themis Rosa, Morgado MG. Prevalence and risk factors associated to subtype C in HIV-1 positive patients being treated in the Ambulatório de Dermatologia Sanitária (Surgery of Sanitary Dermatology) of Secretaria da Saúde (Department of Health) of the state of Rio Grande do Sul. **Boletim da Saúde**, v. 16, n. 2, 2002.

Johnson, P. R. and V. M. Hirsch, SIV infection of macaques as a model for AIDS pathogenesis. **Int Rev Immunol**. 1992. 8(1): p. 55-63.

Mizuochi T, Matthews TJ, Kato M, Hamako J, Titani K, Solomon J, Feizi T. Diversity of oligosaccharide structures on the envelope glycoprotein gp 120 of human immunodeficiency virus 1 from the lymphoblastoid cell line H9. Presence of complex-type oligosaccharides with bisecting N-acetylglucosamine residues. **J Biol Chem**. 1990 May 25;265(15):8519-24.

Morgado MG, Sabino EC, Shpaer EG, Bongertz V, Brigido L, Guimaraes MD, Castilho EA, Galvão-Castro B, Mullins JI, Hendry RM, et al. V3 region polymorphisms in HIV-1 from Brazil: prevalence of subtype B strains divergent from North American/European prototype and detection of subtype F. **AIDS Res Hum Retroviruses**. 1994 May;10(5):569-76.

Morgado MG, Guimarães ML, Galvão-Castro B. HIV-1 polymorphism: a challenge for vaccine development - a review. **Mem Inst Oswaldo Cruz**. 2002 Mar;97(2):143-50.

Marlink R, Kanki P, Thior I, Travers K, Eisen G, Siby T, Traore I, Hsieh C.C., Dia M.C., Gueye Eh, et al. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. **Science**. 1994 Sep 9;265(5178):1587-90.

MacNeil A, Sarr AD, Sankalé JL, Meloni ST, Mboup S, Kanki P. Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection. **J Virol**. 2007 May;81(10):5325-30. Epub 2007 Feb 28.

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

Marins JR, Jamal LF, Chen SY, Barros MB, Hudes ES, Barbosa AA, Chequer P, Teixeira PR, Hearst N. Dramatic improvement in survival among adult Brazilian AIDS patients. **AIDS**. 2003 Jul 25;17(11):1675-82.

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

Nies-Kraske E, Schacker TW, Condoluci D, Orenstein J, Brenchley J, Fox C, Daucher M, Dewar R, Urban E, Hill B, Guenaga J, Hoover S, Maldarelli F, Hallahan CW, Horn J, Kottlil S, Chun TW, Folino M, Palmer S, Wiegand A, O'Shea MA, Metcalf JA, Douek DC, Coffin J, Haase A, Fauci AS, Dybul M. Evaluation of the pathogenesis of decreasing CD4(+) T cell counts in human immunodeficiency virus type 1-infected patients receiving successfully suppressive antiretroviral therapy. **J Infect Dis.** 2009 Jun 1;199(11):1648-56

Osmanov S, Pattou C, Walker N, Schwardländer B, Esparza J. Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. **J Acquir Immune Defic Syndr.** 2002 Feb 1;29(2):184-90.

Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. **N Engl J Med.** 1998 Mar 26;338(13):853-60.

Parker R, Galvão J. Quebrando o silêncio: mulheres e Aids no Brasil. Rio de Janeiro: Relume-Dumará: ABIA:IMS/UERJ; 1996; parte 1, cap. 2: 33.

Peeters, M., et al., Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *Aids*, 1989. 3(10): p. 625-30.

Pomerantz, R.J.; Trono, D. Genetic therapies for hiv infections: promise for the future. **AIDS.** 1995;9:985-993.

Peeters M, Sharp PM. Genetic diversity of HIV-1: the moving target. **AIDS.** 2000;14 Suppl 3:S129-40.

Peeters M (2000). Recombinant HIV sequences: Their role in the global epidemic. pp. 1-39-54 in HIV Sequence Compendium 2000. Edited by: Kuiken CL, Foley B, Hahn B, Korber B, McCutchan F, Marx PA, Mellors JW, Mullins JI, Sodroski J, and Wolinsky S. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

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**, páginas 1582 a 1586, Março, 1996.

Pradier C, Pugliese P, Caissotti C, S Wehrlen-Martini, Huynh-Van E, Pueyo B, Dellamonica P. Dossier médical informatisé pour les patients atteints d'infection par le VIH (ADDIS) : l'expérience du Service des Maladies Infectieuses du CHU de Nice. **Med Mal Inf.** 1998 ;28 :291-5.

Rhee SY Rhee, Matthew J. Gonzales, Rami Kantor, Bradley J. Betts, Jaideep Ravela, and Robert W. Shafer. Human immunodeficiency virus reverse transcriptase and protease sequence database. **Nucleic Acids Res.** Jan 2003; 31: 298 - 303.

Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris, Sirivichayakul S, Jorgensen L, Brigido LF, Schapiro JM, and Shafer RW for the international Non Subtype B HIV-1 Working Group (2006). HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. **AIDS.** 2006, 20(5):643-651, March 21

Reinhold Welker, Heinrich Hohenberg, Uwe Tessmer, Carola Huckhagel, and Hans-Georg Kräusslich. Biochemical and Structural Analysis of Isolated Mature Cores of Human Immunodeficiency Virus Type 1. **J Virology.** 2000 February; 74(3): 1168–1177.

Reitter, J. N.; Means, R. E.; Desrosiers, R. C. A role for carbohydrates in immune evasion in AIDS. **Nat. Med.** 1998, 4 (6), 679–684.

Ridky T, Leis J. Development of drug resistance to HIV-1 protease inhibitors. **J Biol Chem.** 1995 Dec 15;270(50):29621-3.

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, Korber B. HIV-1 nomenclature proposal. **Science.** 2000 Apr 7;288(5463):55-6.

Sepkowitz, K.A.; AIDS the first 20 years. **N. ENGL. J. MED.** 2001; 344:1764-1772.

Sá MS, Sampaio J, Haguihara T, Ventin FO, Brites C. Clinical and laboratory profile of HIV-positive patients at the moment of diagnosis in Bahia, Brazil. **Braz J Infect Dis**. Aug; 11(4): 395-8. 2007.

Sabino EC, Shpaer EG, Morgado MG, Korber BT, Diaz RS, Bongertz V, Cavalcante S, Galvão-Castro B, Mullins JI, Mayer A. Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. **J Virol**. 1994 Oct;68(10):6340-6.

Schubert U, Ott D.E., Chertova E.N., Welker R, Tessmer U, Princiotta Mf, Bennink J.R., Krausslich H.G., Yewdell J.W.. Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. **Proc Natl Acad Sci USA**. 2000 Nov 21;97(24):13057-62.

Sepkowitz KA. AIDS--the first 20 years. **N Engl J Med**. 2001 Jun 7;344(23):1764-72.

Sharp, J.M.; Robertson, DL.; Gao, F.; Harn, B.H. Origins and diversity of human immunodeficiency viruses. **AIDS**. 1994; 8: S27-S42. Suplementos 1.

Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. **Clin Microbiol Rev**. 2002 Apr;15(2):247-77.

Shafer RW, Duane R. Jung, and Bradley J. Betts. (2000) Human immunodeficiency virus type 1 reverse transcriptase and protease mutation search engine for queries. **Nature Medicine**, 6(11), 1290-1292.

Sigrist C.J.A., Cerutti L., Hulo N., Gattiker A., Falquet L., Pagni M., Bairoch A., Bucher P. PROSITE: a documented database using patterns and profiles as motif descriptors. **Brief Bioinform**. 3:265-274(2002).

Simon V, Vanderhoeven J, Hurley A, Ramratnam B, Louie M, Dawson K, Parkin N, Boden D, Markowitz M. Evolving patterns of HIV-1 resistance to antiretroviral agents in newly infected individuals. **AIDS**. 2002 Jul 26;16(11):1511-9.

S. Pompilho; Análise Essencial, **Editora Infobook**, 1995.

Stefani MM, Pereira GA, Martelli CM, Shindo N, Galvão-Castro B. Evidence of HIV-1 genetic diversity among pregnant women with AIDS or infected with HIV-1 in Central Brazil. **J Acquir Immune Defic Syndr**. 2000 Feb 1;23(2):205-7.

Stürmer M, Doerr HW, Staszewski S, Preiser W.
Comparison of nine resistance interpretation systems for HIV-1 genotyping.
Antivir Ther. 2003 Jun;8(3):239-44.

Strack B, Calistri A, Craig S, Popova E, Göttlinger H.G.. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. **Cell**. 2003 Sep 19;114(6):689-99.

Susanna Lamers, Scott Beason¹, Luke Dunlap,
Robert Compton And Marco Salemi.
Hivbase: a pc/Windows-based software offering storage and querying power for locally held hiv-1 genetic, experimental and clinical data. **Bioinformatics**. 2003; 19: 1-3.

Szwarcwald CL, Bastos FI, Esteves MA, De Andrade CL. The spread of the AIDS epidemic in Brazil from 1987 to 1996: a spatial analysis. **Cad Saude Publica**. 2000;16(## Suppl 1):7-19.

Takebe Y, Kusagawa S, Motomura K. Molecular epidemiology of HIV: tracking AIDS pandemic. **Pediatr Int**. 2004 Apr;46(2):236-44.

Tavares P. de Campos; Algoritmo, in "Enciclopédia Verbo Luso-Brasileira da Cultura, Edição Século XXI", Volume II, Editorial Verbo, Braga, Janeiro de 1998 ISBN 972-22-1864-6.

T. De Oliveira, K. Deforche, S. Cassol, M. Salminen, D. Paraskevis, C. Seebregts, J. Snoeck, E. J. Van Rensburg, A. M. J. Wensing, D.A. Van De Vijver, C. A. Boucher, R. Camacho, And A-M Vandamme. An Automated Genotyping System for Analysis of HIV-1 and other Microbial Sequences. **Bioinformatics**. 2005; 21 (19), 3797-3800.

Triques K, Bourgeois A, Vidal N, Mpoudi-Ngole E, Mulanga-Kabeya C, Nzilambi N, Torimiro N, Saman E, Delaporte E, Peeters M. Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. **AIDS Res Hum Retroviruses**. 2000 Jan 20;16(2):139-51.

Vicente AC, Otsuki K, Silva NB, Castilho MC, Barros FS, Pieniazek D, Hu D, Rayfield MA, Bretas G, Tanuri A. The HIV epidemic in the Amazon Basin is driven by prototypic and recombinant HIV-1 subtypes B and F. **J. Acquir Immune Defic Syndr**. 2000 Apr 1;23(4):327-31.

Voit, D.; Voit, Jg.; Pratt, C.W. Fundamentos de Bioquímica. Porto Alegre: **Artes Médicas Sul**, 2001.

Warnke D, Barreto J, and Temesgen Z. Antiretroviral drugs. **J Clin Pharmacol**. 2007, **47**: 1570–1579.

Wei, X.; Decker, J. M.; Wang, S.; Hui, H.; Kappes, J. C.; Wu, X.; Salazar-Gonzalez, J. F.; Salazar, M. G.; Kilby, J. M.; Saag, M. S.; Komarova, N. L.; Nowak, M. A.; Hahn, B. H.; Kwong, P. D.; Shaw, G. M. Antibody neutralization and escape by HIV-1. **Nature** 2003, 422 (6929), 307–312.

Who network for HIV isolation and characterization of HIV-1 variation in the WHO – sponsored vaccine evaluation sites: genetic screening, sequence analysis preliminary biological characterization of selected viral strains. **AIDS Res Hum Retroviruses**, v. 10, p. 1325-1341, 1994.

Wilson JW. Update on antiretroviral drug resistance testing: combining laboratory technology with patient care. **AIDS Read**. 2003 Jan;13(1):25-30, 35-8.

Winters, M.A., K.L. Coolley, Y.A. Girard, D.J. Levee, H. Hamdan, R.W. Shafer, D.A. Katzenstein, and T.C. Merigan. 1998. A 6-basepair insert in the reverse transcriptase gene of human immunodeficiency virus type 1 confers resistance to multiple nucleoside inhibitors. **J.Clin.Invest**. 102: 1769-1775.

Yarchoan R, Klecker RW, Weinhold KJ, Markham PD, Lyerly HK, Durack DT, Gelmann E, Lehrman SN, Blum RM, Barry DW, et al. Administration of 3'-azido-3'-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. **Lancet**. Mar 15; 1(8481): 575-80. 1986.

ANEXOS

9. ANEXOS

9.1. TRABALHOS APRESENTADOS EM CONGRESSOS

1 - Cunha DRM, Couto-Fernandez JC, Morgado MG, Oliveira T, Galvão-Castro B and Alcantara LC. Development Of A Bioinformatics Tool For Analysis Of Hiv-1 Drug Resistance And Postranslational Sites. VIII SIMPAIDS, 2009.

2 - Cunha DRM, Couto-Fernandez JC, Morgado MG, Oliveira T, Galvão-Castro B and Alcantara LC. Development of a bioinformatic tool for analysis of hiv-1 drug resistance mutations. XVIII International AIDS Conference (AIDS 2010).

9.2. CURSOS RELACIONADOS A BIOINFORMÁTICA EM HIV/AIDS

1 - Treinamento em Bioinformática e Análises Genéticas e Evolutivas Virais – LASP/CPqGM/FIOCRUZ - (29 de janeiro a 6 de fevereiro 2004).

2 - Estudos Avançados em Bioinformática e Análises Evolutivas Virais – LASP/CPqGM/FIOCRUZ - (de 14 a 19 de junho de 2004).

3 - Segundo Workshop Brasileiro em Evolução Viral e Epidemiologia Molecular – LASP/CPqGM/FIOCRUZ - (de 8 a 13 de novembro de 2004).

4 - Treinamento para a Utilização do Database HIV-Base –

LASP/CPqGM/FIOCRUZ - (de 15 a 18 de novembro de 2005).

5 - Capacitação em Bioinformática para pesquisa em HIV-AIDS: gestão e exploração de banco de dados (data management e data mining),(de 5 a 9 de novembro de 2007).

9.3. PUBLICAÇÕES CIENTÍFICAS RELACIONADAS A BIOINFORMÁTICA EM HIV/AIDS

1 - MIRANDA, A.C. ; DE OLIVEIRA, T. ; **CUNHA, D.R.M.** ; BOMFIM, C. ; GALVÃO-CASTRO, B. ; ALCANTARA, L.C.J. . Mapping the molecular characteristics of Brazilian human T-cell lymphotropic virus type 1 Env (gp46) and Pol amino acid sequences for vaccine design. Memórias do Instituto Oswaldo Cruz, v. 102, p. 000/000, 2007.

2 - QUEIROZ, A.T.L. ; MIRANDA, A.C. ; DE OLIVEIRA, T. ; **CUNHA, D.R.M.** ; URPIA, C. C. ; CARVALHO, C.M. ; GALVÃO-CASTRO, B. ; ALCANTARA, L.C.J. . Re-mapping the molecular features of the human immunodeficiency virus type 1 and human T-cell lymphotropic virus type 1 Brazilian sequences using a bioinformatics unit established in Salvador, Bahia, Brazil, to give support to the viral epidemiology studies. Memórias do Instituto Oswaldo Cruz (Online), v. 102, p. 133-139, 2007.

3 - QUEIROZ, A.T.L. ; SANTOS, L.A. ; **CUNHA, D.R.M.** ; DE OLIVEIRA, T. ; Watkins, D. ; GALVÃO-CASTRO, B. ; ALCANTARA, L.C.J. . Identification and characterization of previously described epitopes in HIV-1 subtypes B, C, F and BF in Brazil.. The Brazilian Journal of Infectious Diseases, v. 11, p. 27-30, 2007.

4 - ALCANTARA, L.C.J. ; MIRANDA, A.C. ; QUEIROZ, A.T.L. ; GALVÃO-CASTRO, B. ; CARVALHO, C.M. ; **CUNHA, D.R.M.** ; DE OLIVEIRA, T. . Bioinformática. In: Ministério da Saúde. (Org.). Pesquisa e Desenvolvimento Tecnológico em DST/HIV/Aids no Brasil.. Brasília: , 2005, v. Brasil, p. 263-266.

9.4. MOTIVAÇÃO PARA DESENVOLVIMENTO DA FERRAMENTA LASP-HIV1-ResTool

Avanços em biotecnologia ocorridos nos últimos anos vêm proporcionando a obtenção de grandes quantidades de seqüências genômicas de diversos organismos. Esses resultados são publicados e armazenados em bancos de dados públicos, e muitas vezes, por falta de ferramentas e pesquisadores treinados neste tipo de análise, acabam sendo subutilizados. Com o intuito de minimizar estas deficiências, o Dr. Luiz Carlos Alcântara, pesquisador do Laboratório Avançado de Saúde Pública do Centro de Pesquisas Gonçalo Moniz na Fundação Oswaldo Cruz (LASP/CPqGM/FIOCRUZ), propôs um projeto para a implantação de um núcleo de bioinformática no LASP, aprovado em 2003. Em dezembro de 2003, fui convidado, pelo Dr. Luiz Carlos Alcântara, a participar da implementação dos projetos de pesquisa **Implantação de um Núcleo de Bioinformática para dar Suporte ao Monitoramento de agentes Biológicos Emergentes, HTLV e HIV (Resistente e Incidente): Implicações no Controle Antirretroviral e Desenvolvimento de Vacinas e Desenvolvimento e Manutenção de um Núcleo de Referência em Bioinformática para dar Suporte e Treinamento aos Projetos Desenvolvidos pelo PN-DST/AIDS, Ministério da Saúde.**

O Núcleo de Bioinformática foi instalado em janeiro de 2004. Em sua infra-estrutura, o Núcleo de Bioinformática conta com um servidor Linux, onde está hospedado o web site do LASP (<http://lasp.cpqgm.fiocruz.br>), desenvolvido em conjunto com o Dr. Túlio de Oliveira (South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa), pelo Dr. Luiz Carlos Junior Alcântara e pelo Dr. Bernardo Galvão Castro-Filho. O Núcleo conta também com quinze computadores dispostos em bancadas e utilizados como estações de trabalho e estudos, pelos pesquisadores e participantes dos cursos e treinamentos sediados no LASP. As estações de trabalho operam com os sistemas Windows e Linux onde estão instalados os principais programas utilizados em análises de bioinformática.

O primeiro evento realizado no Núcleo de Bioinformática foi o **Treinamento em Bioinformática e Análises Genéticas e Evolutivas Virais** ocorrido entre os dias 29 de

janeiro e 6 de fevereiro 2004. Este treinamento teve como objetivos: treinar pesquisadores em análise genética e evolutiva de patógenos humanos, com ênfase no HIV e HTLV; coordenar e dar suporte a estudos de monitoramento do HIV, polimorfismo viral, infecção recente, resistência aos antirretrovirais e transmissão materno-fetal; dar suporte a outros estudos genômicos do LASP como HCV, Febre Amarela, Dengue, HTLV, HHV8, Sarampo, Rubéola; estimular o desenvolvimento de conhecimento na área de Bioinformática; promover colaboração técnica e científica com grupos de pesquisas nacionais e internacionais; elaborar um algoritmo de resgate terapêutico para subsidiar a terapia antirretroviral a ser estabelecida nos serviços clínicos associados; gerar seqüências nucleotídicas do genoma do HIV-1 de interesse estratégico para o desenvolvimento de vacinas.

O primeiro treinamento deu início a uma série de outros eventos nos anos seguintes. Entre os novos eventos, os principais foram: **Estudos Avançados em Bioinformática e Análises Evolutivas Virais**, entre 14 e 19 de junho de 2004; **Segundo Workshop Brasileiro em Evolução Viral e Epidemiologia Molecular**, entre 8 e 13 de novembro de 2004; **Treinamento para a Utilização do Database HIV-Base**, entre 15 e 18 de novembro de 2005; **Capacitação em Bioinformática para pesquisa em HIV-AIDS: gestão e exploração de banco de dados (data management e data mining)**, realizado entre 5 e 9 de novembro de 2007.

Os cursos e treinamentos no Laboratório de Bioinformática abriram novas perspectivas para o desenvolvimento de projetos voltados a bioinformática. No ano de 2007 iniciamos a criação de pequenos módulos de programação para interpretação de algoritmos para análise de resistência do HIV aos antirretrovirais. Utilizamos para isto seqüências disponíveis nos bancos de dados públicos. Dentre as muitas seqüências de DNA armazenadas em bancos de dados públicos, estão as do HIV (Vírus da Imunodeficiência Humana). Cerca de 279.946 seqüências genômicas do HIV já foram identificadas e catalogadas sendo aproximadamente 8.058 provenientes do Brasil (<http://www.ncbi.nlm.nih.gov/sites>).

Partindo do conhecimento e do material descritos acima foi possível dar início ao sistema que deu origem ao meu projeto de mestrado, iniciado no final do ano de 2007 intitulado: "Desenvolvimento de uma ferramenta de bioinformática para análise de resistência do HIV-1 a drogas anti-retrovirais.

Os treinamentos e cursos realizados no Núcleo de Bioinformática foram de grande importância para a aquisição de conhecimento das novas técnicas e métodos de análises moleculares e serviram de insumo para a publicação de diversos trabalhos.