



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE

CURSO DE GRADUAÇÃO EM BIOMEDICINA

Hegger Machado Fritsch

**ANÁLISE DA FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DA FORMA
RECOMBINANTE CRF31_BC DO HIV-1 NO BRASIL**

Porto Alegre

2016

Hegger Machado Fritsch

**ANÁLISE DA FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DA FORMA
RECOMBINANTE CRF31_BC DO HIV-1 NO BRASIL**

Trabalho de conclusão de curso de graduação em Biomedicina apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para obtenção do título de Bacharel(a) em Biomedicina.

Área de Habilitação: Biologia Molecular

Orientador: Dr^a. Sabrina Esteves de Matos Almeida

Co-orientador: Dr. Tiago Gräf

Porto Alegre

2016

CIP - Catalogação na Publicação

Machado Fritsch, Hegger

Análise da filogeografia e história demográfica da forma recombinante CRF31_BC do HIV-1 no Brasil / Hegger Machado Fritsch. -- 2016.
70 f.

Orientadora: Sabrina Esteves de Matos Almeida.
Coorientador: Tiago Gräf.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto de Ciências Básicas da Saúde, Curso de Biomedicina, Porto Alegre, BR-RS, 2016.

1. Epidemiologia Molecular. 2. Dispersão Viral.
3. HIV. 4. Forma Recombinante CRF31_BC. I. Esteves de Matos Almeida, Sabrina, orient. II. Gräf, Tiago, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Hegger Machado Fritsch

**ANÁLISE DA FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DA FORMA
RECOMBINANTE CRF31_BC DO HIV-1 NO BRASIL**

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel(a) em Biomedicina.

Aprovado em: ____ de _____ de _____.

BANCA EXAMINADORA

Dr. Dennis Maletich Junqueira
Centro Universitário Ritter dos Reis

Dr. Gustavo Fioravanti Vieira
Universidade Federal do Rio Grande do Sul

Dr^a. Sabrina Esteves de Matos Almeida
Universidade Feevale

“ A persistência é o caminho do êxito.”

Charlie Chaplin

Agradecimentos

Agradeço primeiramente a Deus pela vida e força para chegar ao final desta caminhada.

Agradeço aos meus pais Ana Lúcia e Paulo Roberto pelo apoio incondicional em todos os momentos da minha vida, por serem meu exemplo de força, dedicação, amor e, principalmente, por batalharem meus sonhos comigo e compreenderem meus momentos de ausência. Este trabalho é pra vocês!!

Agradeço a minha família e amigos que, nesses anos, me ensinaram que para estar junto não é preciso estar perto fisicamente. Em cada momento dessa caminhada levei vocês no meu coração.

Aos Doutores Tiago Gräf e Sabrina Esteves de Matos Almeida, profissionais incríveis com quem pude trabalhar e foram de suma importância para a concretização deste trabalho e sonho de trabalhar com HIV. Obrigada por todo suporte teórico, prático e discussões ao longo desta jornada.

Aos colegas da Turma 9 – Biomedicina UFRGS, em especial a Thainá Garbino, Guilherme Negrini, Maurício Barth, Livia Fratini e Laura Lago por terem compartilhado comigo os melhores e piores momentos do curso.

Agradeço especialmente a Janaína Pereira Pinto, minha prima, por ser meu primeiro exemplo na ciência, a Dr^a Ludmila Entiauspe e futura Dr^a Lisiane Ortiz por terem dividido meus primeiros momentos dentro de um laboratório. Jamais esquecerei aqueles momentos no HU-FURG. Obrigada por serem meus exemplos de ética, dedicação e paixão pela profissão.

Agradeço a todos os professores que, nesses cinco anos, semearam a semente do questionamento e busca de conhecimento.

E agradeço, por fim, ao Álvaro da Rosa Cunha por todo amor, companheirismo, amizade, sonhos compartilhados e força nesses momentos finais de curso.

RESUMO

A situação epidemiológica da epidemia de HIV-1 da região Sul do Brasil é peculiar: além de apresentar as maiores taxas de incidência e mortalidade, no sul do país, há alta frequência de subtipo C co-circulando com o subtipo B, o que por sua vez levou a origem da forma recombinante CRF31_BC. Recentemente, essa forma recombinante, antes descrita apenas em Porto Alegre, foi identificada em cidades do interior e em outros estados brasileiros, sugerindo um processo de expansão. Em vista destes novos dados, investigações que visem descrever cidade de origem, rotas de dispersão e história demográfica da epidemia da CRF31_BC fazem-se necessários. O dataset analisado aqui foi composto de sequências disponíveis em bancos de dados públicos e de sequências novas geradas a partir de uma amostragem em cidades do interior do RS e SC. O padrão de recombinação analisado por Bootscanning e confirmado por inferência filogenética. Árvores em escala de tempo foram inferidas por análise Bayesiana, no programa BEAST, utilizando-se os modelos Bayesian Skyline, GTR+G4+I e relógio molecular relaxado não correlacionado. Para realizar as análises demográficas foi testado o melhor modelo de crescimento populacional e as análises filogeográficas foram executadas utilizando-se o modelo de transição assimétrico com BSSVS. Nossas análises reconstruíram a origem da CRF31_BC para o ano de 1985 (1978-1989 95%HPD), muito provavelmente na cidade de Porto Alegre. As análises filogeográficas indicaram Porto Alegre como centro de dispersão da CRF31_BC para o interior do estado, bem como para outros estados brasileiros. Foi possível observar também que algumas cidades já apresentam epidemias autossustentadas. Esta forma recombinante tem seu crescimento descrito pelo modelo logístico, com um aumento exponencial nos primeiros 7,5 anos de epidemia e estabilização do crescimento a partir de 1992. O presente estudo destaca o papel de Porto Alegre como local de origem da forma recombinante CRF31_BC e no processo de disseminação dessa variante do HIV-1 pelo Brasil. Estudos anteriores já haviam reportado Porto Alegre como centro de dispersão do subtipo C pelo Brasil, o que juntamente com nossos resultados destaca a importância do monitoramento da epidemia do HIV-1 nesta cidade e do desenvolvimento de políticas de saúde pública para diminuir a incidência de Aids.

Lista de Figuras

Figura 1: Figura esquemática dos principais eventos do ciclo replicativo do HIV. Retirado de Gräf (2015), adaptado de Rambaut et al, 2004.	12
Figura 2: Representação gráfica da história natural da infecção pelo HIV. Adaptado de Goulder et al, 2004.	13
Figura 3: Distribuição mundial dos subtipos do HIV-1 e formas recombinantes circulantes (CRFs). Adaptado de Ariën et al, 2007.	16
Figura 4: (a) Representação de estrutura esquemática do genoma de forma recombinante CRF31_BC. (b) Análise de bootscanning mostrando o padrão de recombinação da CRF31. Adaptado de Santos et al, 2006.	19
Figura 5: Representação da diversidade molecular do HIV-1 em cidades do interior do Rio Grande do Sul e Santa Catarina. Adaptado de Gräf, 2015.	20
Figura 6: Análise de bootscan de uma sequência recombinante. Adaptado de Gräf, 2015.	27

Lista de Abreviaturas e Siglas

CCR5: receptor C-C de quimiocina tipo 5

CRF: forma recombinante circulante

CXCR4: receptor C-X-C de quimiocina tipo 4

DNA: ácido desoxirribonucleico

Env: gene do envelope

gp120: glicoproteína de superfície 120

gp41: glicoproteína de superfície 41

HIV-1: vírus da imunodeficiência humana tipo 1

HIV-2: vírus da imunodeficiência humana tipo 2

Pol: gene da polimerase

PR: Paraná

RNA: ácido ribonucleico

RS: Rio Grande do Sul

RT: transcriptase reversa

SC: Santa Catarina

URF: forma recombinante única

SUMÁRIO

1. INTRODUÇÃO	11
1.1 O vírus da Imunodeficiência Humana	11
1.2 O vírus e a Aids.....	12
1.3 A dinâmica evolutiva do HIV	14
1.3.1 Eventos de recombinação e o surgimento de partículas virais recombinantes.....	14
1.4 Epidemiologia molecular do HIV-1 no mundo	15
1.5 A epidemiologia molecular do HIV-1 no Brasil e a situação epidemiológica particular da região Sul.....	16
1.6 A CRF31_BC	18
1.7 O uso da bioinformática como ferramenta para análises evolutivas.	20
1.7.1 O uso da filogenia em estudos de epidemias virais.....	20
1.7.2 A construção da árvore filogenética.....	21
1.7.3 Modelos evolutivos	21
1.7.4 Os métodos de Máxima Verossimilhança e Inferência Bayesiana	23
1.7.5 Inferências filogeográficas	24
1.7.6 Detecção de eventos de recombinação.....	25
2. JUSTIFICATIVA	28
3. OBJETIVOS	29
4. ARTIGO CIENTÍFICO	30
5. CONCLUSÃO E PERSPECTIVAS FUTURAS	56
REFERÊNCIAS BIBLIOGRÁFICAS	57
ANEXOS	66
Anexo 1 – Normas de publicação do periódico Memórias do Instituto Oswaldo Cruz.....	66

1. INTRODUÇÃO

1.1 O vírus da Imunodeficiência Humana

O HIV é um vírus de RNA capaz de sintetizar, durante seu ciclo replicativo, um genoma intermediário de DNA, sendo alocado então na família *Retroviridae*. Pertencente ao gênero *Lentivirus*, apresenta um longo período de incubação e lento desenvolvimento da sintomatologia (Hoffmann et al, 2007). Existem dois tipos de HIV: o HIV-1 e HIV-2. Ambos apresentam morfologia semelhante, porém diferem quanto à patogenicidade e dispersão mundial. O HIV-1 é mundialmente conhecido por ser responsável pela pandemia mundial de Aids, enquanto o HIV-2 mostra-se menos patogênico, causando infecções geograficamente restritas ao oeste da África e a algumas localidades da Europa (Rambaut et al, 2004).

O genoma viral do HIV-1 é composto por duas fitas simples e similares de RNA, contendo aproximadamente 10 mil nucleotídeos cada, totalizando nove genes. Destes genes, três são ditos principais e presentes em todos os retrovírus: *env*, *gag* e *pol*. O gene *env* é responsável pela codificação de glicoproteínas de superfície (gp120 e gp41) compondo o envelope viral, enquanto o gene *gag* determina proteínas estruturais que constituem o capsídeo. Já o gene *pol* codifica as enzimas virais transcriptase reversa (RT), protease e integrase (Hoffmann et al, 2007). Os outros seis genes são acessórios e denominados *vif*, *vpu*, *vpr*, *tat*, *rev* e *nef*. Suas funções compreendem promoção da transcrição viral, transporte de RNA mensageiro ao citoplasma e evasão de mecanismos de resposta imune inata e adaptativa (Miller e Sarver, 1997; Cullen, 1998; Petter, 1998).

O ciclo replicativo viral inicia com a interação das glicoproteínas do envelope com a molécula de CD4 encontrada na superfície de células do sistema imunológico tais como linfócitos T, monócitos, macrófagos, células dendríticas e micróglia. Esta primeira interação necessita ser estável para promover a fusão do envelope viral com a membrana plasmática da célula hospedeira, com isso, ocorre uma segunda interação via co-receptores. Os co-receptores mais utilizados pelo vírus são CCR5 e/ou CXCR4. Após a entrada do vírus na célula, ocorre a desintegração do nucleocapsídeo e, conseqüentemente, liberação do material genético viral no citoplasma da célula infectada. Desta forma, as enzimas virais promovem a

transcrição reversa da fita de RNA, transporte até o núcleo e integração deste DNA proviral ao genoma celular. Ao tornar-se parte constituinte do genoma celular, o DNA recém-integrado é transcrito, permitindo a montagem de uma nova partícula viral. Esse novo vírion é formado por duas cópias do RNA viral e as enzimas virais transcriptase reversa, integrase e protease. A protease, por sua vez, tem como função principal agir na clivagem da poliproteína traduzida e na maturação do vírion após o brotamento, processo necessário para tornar essa partícula infecciosa (Hoffmann et al, 2007). As principais etapas do processo de replicação citadas acima estão representadas no desenho esquemático abaixo (Figura 1).

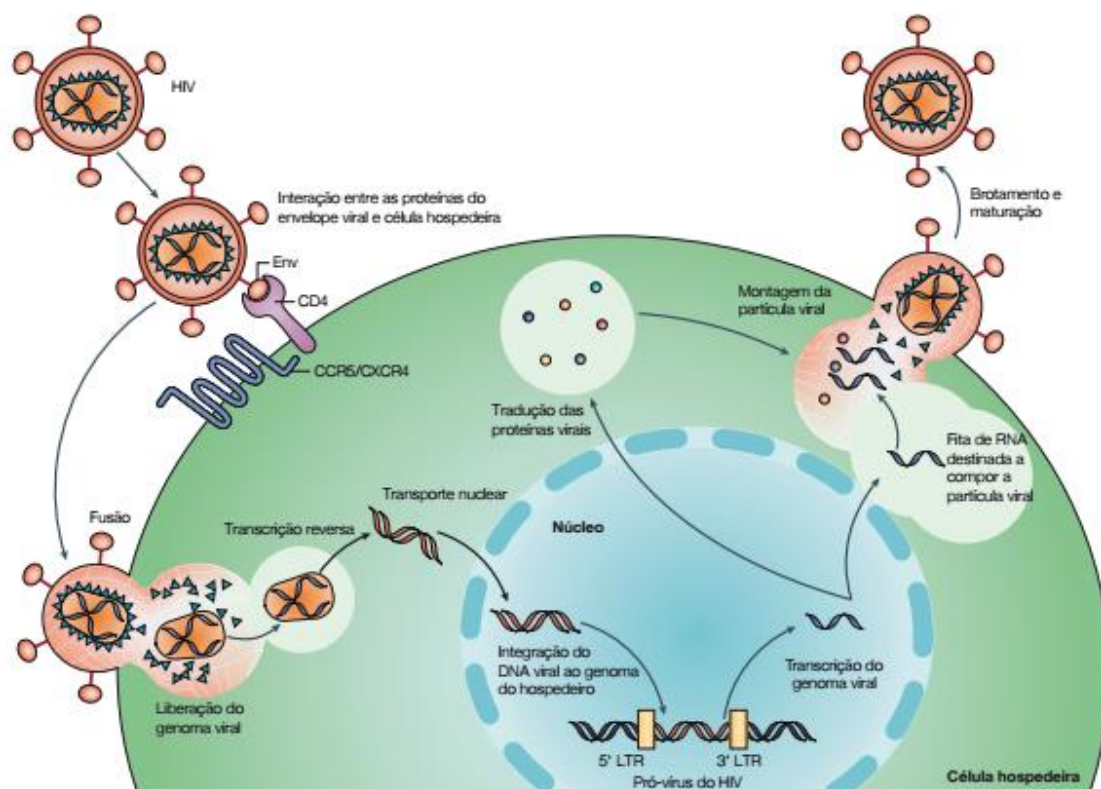


Figura 1: Figura esquemática dos principais eventos do ciclo replicativo do HIV. Retirado de Gräf (2015), adaptado de Rambaut et al, 2004.

1.2 O vírus e a Aids

As principais células-alvo de infecção pelo HIV são linfócitos T-CD4+. Este tipo celular é infectado e sofre depleção principalmente nos linfonodos, onde sua concentração é elevada e o ambiente pró-inflamatório local intensifica a replicação do vírus (Peterlin e Trono, 2003).

A história natural da doença causada pela infecção do HIV pode ser dividida em três fases: aguda (logo após a infecção), fase crônica e fase de Aids. Durante a fase aguda, o indivíduo apresenta elevada viremia e os sintomas são inespecíficos, semelhantes a um quadro gripal. Após algumas semanas, a ação de linfócitos T citotóxicos HIV-específicos conseguem controlar a carga viral (Mellors et al, 1995). A soroconversão, momento em que anticorpos anti-HIV circulantes tornam-se detectáveis, determina o final da fase aguda. O período que a precede é popularmente conhecido como janela imunológica.

A fase crônica tem como principal característica um período de equilíbrio entre a replicação viral e resposta imune do hospedeiro. Esta fase é assintomática e pode perdurar anos, contudo, há progressiva morte de linfócitos T-CD4+. Com a redução destas células, o indivíduo adquire um caráter imunossuprimido. A fase de Aids é atingida quando os níveis de linfócitos T-CD4+ encontram-se em limites inferiores a 350 células/ μ L de sangue, tornando o sistema imunológico do hospedeiro ineficiente em protegê-lo de infecções por patógenos oportunistas (Hoffmann et al, 2007).

Na Figura 2 é possível observar a história natural da doença e a representação dos níveis de linfócitos T-CD4+ circulantes e partículas virais ao longo das fases descritas acima.

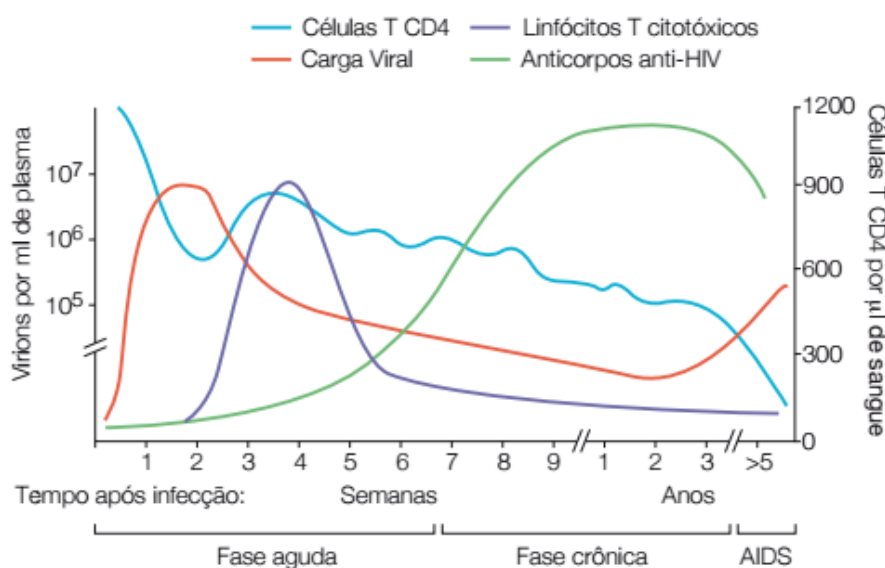


Figura 2: Representação gráfica da história natural da infecção pelo HIV. Adaptado de Goulder et al, 2004.

1.3 A dinâmica evolutiva do HIV

A principal marca do HIV é a extraordinária taxa evolutiva que resulta em alta diversidade molecular e grande dinamismo na epidemia de Aids (Tebit e Arts, 2011). A diversidade de subtipos é decorrente de uma combinação de fatores que tornam esse patógeno um dos organismos que apresentam evolução mais rápida, sendo possível mensurar essa evolução em uma escala de tempo curta (Rambaut et al, 2004; Grenfell et al, 2004). Um dos fatores associados a essa rápida evolução é a inexistência de maquinaria e processos de correção de erros da enzima RT, culminando com alta taxa mutacional. A etapa de transcrição reversa, necessária durante a replicação viral, gera em torno de 0,2 mutações por genoma por ciclo replicativo (Preston et al, 1988). Essa elevada taxa mutacional torna-se ainda mais acentuada devido ao número de 10^{10} a 10^{12} novos vírions por dia (Perelson et al, 1996), propiciando que o HIV possua uma taxa evolutiva estimada em 2 a 4×10^{-3} substituições por sítio a cada ano (Abecasis et al, 2009).

Além dos mecanismos capazes de gerar variação entre partículas virais do mesmo subtipo (variação intrasubtipo), altas taxas de recombinação entre variantes distintas do HIV também contribuem para o aumento da diversidade deste patógeno (Vuilleumier et al, 2015).

Devido à rápida taxa evolutiva do HIV, intra ou intersubtipo, a resposta imunológica frente ao vírus torna-se ineficaz. Além disso, o desenvolvimento de vacinas é inviabilizado ou com eficácia reduzida (Ramos et al, 1999), bem como permite que o vírus adquira resistência aos tratamentos antirretrovirais disponíveis, podendo ocasionar falha terapêutica (Grenfell et al, 2004).

1.3.1 Eventos de recombinação e o surgimento de partículas virais recombinantes

A recombinação viral é decorrente de eventos de múltipla infecção. A múltipla infecção pode ocorrer devido a uma co-infecção, onde há aquisição de dois subtipos virais diferentes de forma simultânea ou superinfecção, quando há aquisição sequencial de subtipos (Smith et al, 2005). Esses eventos de recombinação podem ocorrer com vírus do mesmo subtipo ou subtipos distintos (Liu et al, 2002; Yang et al, 2005), porém, no primeiro caso, a detecção é dificultada, visto que as sequências

parentais que originaram aquele híbrido precisam ser divergentes o suficiente para identificar a qual ancestral ela pertence (Lemey et al, 2009).

Os eventos de recombinação entre variantes distintas iniciam quando uma célula é co-infectada por duas partículas virais e, no momento do empacotamento do material genético, fitas de RNA de origem distinta são inseridas no mesmo capsídeo, formando uma partícula viral híbrida. A recombinação então ocorre no momento da infecção de uma nova célula, quando a enzima RT inicia a síntese de DNA e pode ocorrer a troca entre as diferentes fitas de RNA molde (Vuilleumier et al, 2015).

Localidades geográficas onde há co-circulação de mais de um subtipo do HIV-1 em uma mesma população são propícios ao surgimento de recombinantes (Tebit e Arts, 2011). Quando esta forma é restrita a um indivíduo ela é denominada forma recombinante única (URF), mas se é encontrada em três ou mais indivíduos epidemiologicamente não relacionado é denominada forma recombinante circulante (CRF) (Lau e Wong, 2013).

1.4 Epidemiologia molecular do HIV-1 no mundo

Em função dos mecanismos evolutivos e sucesso em disseminar por um grande número de indivíduos, o HIV-1 pode adquirir vasta diversidade genética, sendo classificado em três grupos: M, N e O. Os grupos N e O encontram-se distribuídos de forma mais discreta, sendo endêmicos em regiões do oeste africano, enquanto o grupo M apresenta maior relevância epidemiológica e clínica, visto que é o responsável pela pandemia atual. Esse grupo é subdividido em subtipos puros (A-D, F-H, J e K) (Robertson et al, 2000; Rambaut et al, 2004), além de alguns sub-subtipos (A1-A6 e F1-F2) (Lau e Wong, 2013).

Além dos subtipos já descritos, a alta diversidade viral do grupo M é também observada quando consideramos as formas recombinantes intersubtipos. Segundo o banco de sequências Los Alamos HIV Database (<http://www.hiv.lanl.gov>), atualmente, já são descritas 72 CRFs, correspondendo a 20% das novas infecções mundiais (Lau e Wong, 2013), além de infecções causadas por URFs.

Segundo a UNAIDS há, atualmente, mais de 34 milhões de pessoas vivendo com HIV-1. Os subtipos virais associados e os números envolvendo a infecção

apresentam distribuição não uniforme no território global (WHO/UNAIDS, 2013) conforme é mostrado na Figura 3, sendo a África Central a região que apresenta a epidemia mais diversificada (Hemelaar et al, 2006; Osmanov et al, 2002). O subtipo C é o mais prevalente mundialmente, sendo encontrado em países populosos como a Índia e região sul da África, enquanto o subtipo B, a forma mais dispersa no mundo, é responsável pela maioria de casos no continente americano, Europa e Oceania (Ariën et al, 2007).

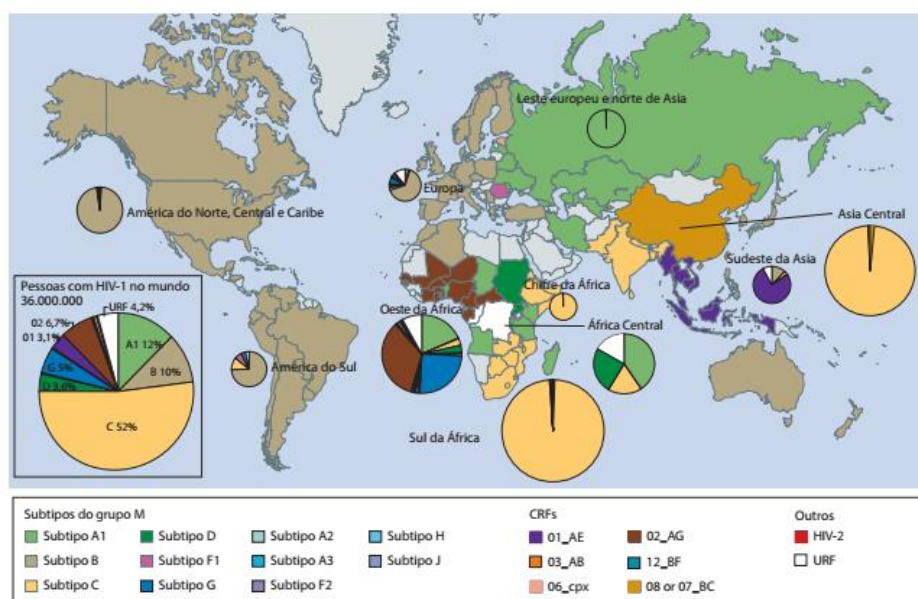


Figura 3: Distribuição mundial dos subtipos do HIV-1 e formas recombinantes circulantes (CRFs). Adaptado de Ariën et al, 2007.

1.5 A epidemiologia molecular do HIV-1 no Brasil e a situação epidemiológica particular da região Sul.

O Brasil é um país de dimensões continentais extensas, característica que propicia variação nos números acerca da epidemia de HIV/Aids conforme a região geográfica analisada. Segundo o Ministério da Saúde existem no Brasil 734 mil indivíduos vivendo com HIV (Coordenação de DST e Aids; Ministério da Saúde, 2014) . A prevalência de Aids entre indivíduos de 15 a 49 anos permanece estável em torno de 0,6% desde 2004. Da mesma forma, a incidência anual de novos casos de Aids é estável desde 2002, sendo em torno de 21 a cada 100 mil habitantes. Tais índices são similares em todas as regiões do país exceto no Sul. Na região sul, a incidência atual é de aproximadamente 30 novos casos anuais a cada 100 mil

habitantes, propiciando que a taxa de mortalidade nessa região seja superior à encontrada no restante do país. Entre os estados do Sul, o Rio Grande do Sul (RS) apresenta o pior cenário, sendo que em 2014, a taxa de detecção de Aids foi de 41 casos a cada 100 mil habitantes. Além disso, Porto Alegre é a cidade que apresenta a maior incidência entre as capitais do país, com taxas de detecção de 96,2 casos para cada 100 mil habitantes, número quase cinco vezes maior que a taxa brasileira (Coordenação de DST e Aids; Ministério da Saúde, 2014).

Diferenças consideráveis são observadas também quando analisamos as taxa de mortalidade e incidência encontrada nos estados do sul. A mortalidade por Aids no Sul do país encontra-se por volta de 8,5 mortes a cada 100 mil habitantes, enquanto a média brasileira permanece em torno de 6. Além disso, nos últimos 10 anos, Santa Catarina (SC) e Rio Grande do Sul (RS) vêm revelando incidência constante acima de 30 novos casos por 100 mil habitantes, enquanto para o estado do Paraná (PR) esse índice mantém-se abaixo de 20.

Outro fator epidemiológico que sofreu modificações ao longo da epidemia foi a razão de sexo. A razão de sexo é um índice que estima número de homens infectados por mulheres infectadas. De forma geral, existe, atualmente uma tendência de equilíbrio, visto que a razão de casos de Aids entre homens e mulheres no país passou de 15:1, em 1986 para 1,8:1 em 2013. O Sul, por sua vez, é a região com a maior proporção de mulheres infectadas, em torno de 10 para 14 homens (Coordenação de DST e Aids; Ministério da Saúde, 2014).

Os fatores que desencadeiam as contrastantes diferenças observadas na epidemiologia de HIV/Aids no país e no Rio Grande do Sul ainda são desconhecidos, porém podem estar relacionados a um fato peculiar: a circulação de HIV-1 subtipo C. Enquanto no país como um todo o subtipo B é responsável por 70-90% das infecções, (Teixeira et al, 2004; Sa Filho et al, 2005; Gadelha et al, 2003; Cerqueira et al, 2004; Couto-Fernandez et al, 2006; Pereira et al, 2006), a circulação do subtipo C não ultrapassa 3% (Pedroso et al, 2007; Cardoso et al, 2009, Sa Filho et al, 2005; Gadelha et al, 2003; Machado et al, 2009). Contudo, no RS o HIV-1C é encontrado com prevalência variando entre 39 (de Medeiros et al, 2011) e 44%

(Simon et al, 2010) em Porto Alegre e Canoas, respectivamente e 56% relatado por Silveira et al (2012) na cidade de Rio Grande.

Além da predominância do subtipo C, a epidemia de HIV no RS também é caracterizada pela circulação de formas recombinantes BC, dentre elas a CRF31_BC (Brígido et al, 2007; Brindeiro et al, 2003; Guimaraes et al, 2002; Santos et al, 2006; Soares et al, 2003a, 2003b).

1.6 A CRF31_BC

A CRF31_BC é uma forma recombinante que apresenta uma estrutura mosaica devido à inserção de um pequeno fragmento de subtipo B (em torno de 200 nucleotídeos) no gene da transcriptase reversa, enquanto todas as outras regiões permanecem classificadas como subtipo C (Figura 4) (Santos et al, 2006). Essa CRF está presente na região metropolitana de Porto Alegre, com prevalência variável de 19% (de Medeiros et al, 2011), 26% (Dias et al, 2009) e 35% (Simon et al, 2010) das infecções por HIV-1 naquela área.

Embora a prevalência da CRF31_BC no RS seja elevada, a frequência estimada de formas recombinantes únicas entre o B e o C permanecem baixos, atingindo taxas de 2% (Santos et al, 2007) a 8% na região (Brígido et al, 2007). Essa baixa taxa de infecções por URFs BC é surpreendente, considerando a alta prevalência de subtipo B e C concomitantemente nessa região por um longo período.

O estudo que caracterizou pela primeira vez a CRF31_BC sugere Porto Alegre como local de origem dessa forma recombinante (Santos et al, 2006), porém esta hipótese não foi testada em nenhum outro estudo. Além disso, outros estudos sugerem que a emergência dessa cepa híbrida date por volta de 1987 (IC95%: 1957-1998) (Santos et al, 2007) ou 1988 (IC95%: 1982-1992) (Passaes et al, 2009), contudo, a data precisa de início da epidemia de CRF31_BC no Sul do Brasil ainda permanece incerta, bem como seu local de origem.

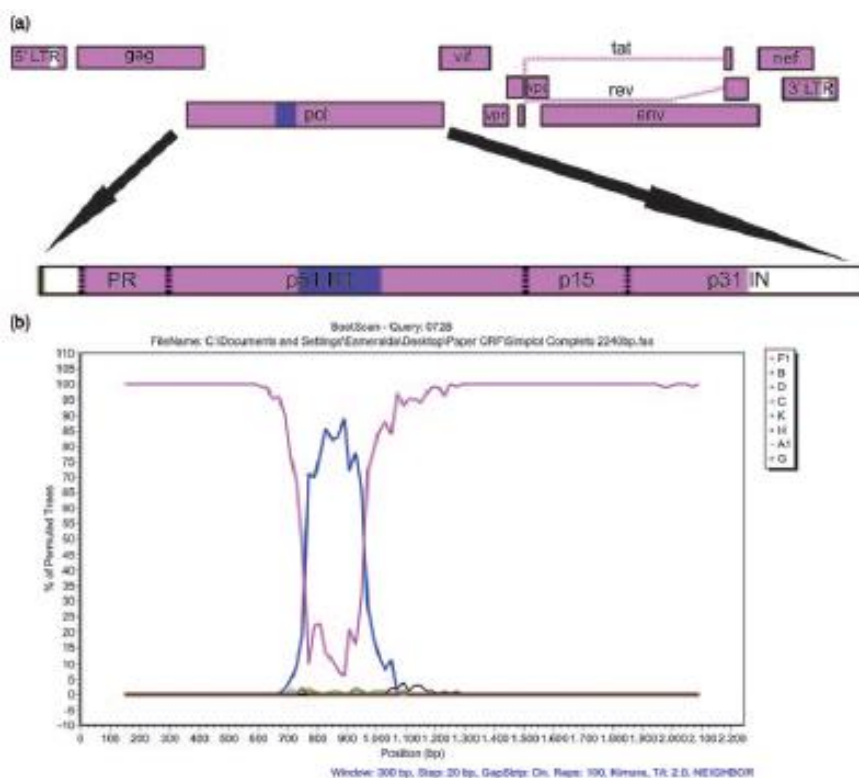


Figura 4: (a) Representação de estrutura esquemática do genoma de forma recombinante CRF31_BC. (b) Análise de bootscanning mostrando o padrão de recombinação da CRF31. Adaptado de Santos et al, 2006.

Além disso, a CRF31_BC, inicialmente descrita apenas na cidade de Porto Alegre, já foi relatada em estudos conduzidos por Gräf (2015) em cidades do interior do Rio Grande do Sul como Caxias do Sul, Santiago e Lajeado e no estado de Santa Catarina, podendo indicar um processo de expansão pelo território (Figura 5).

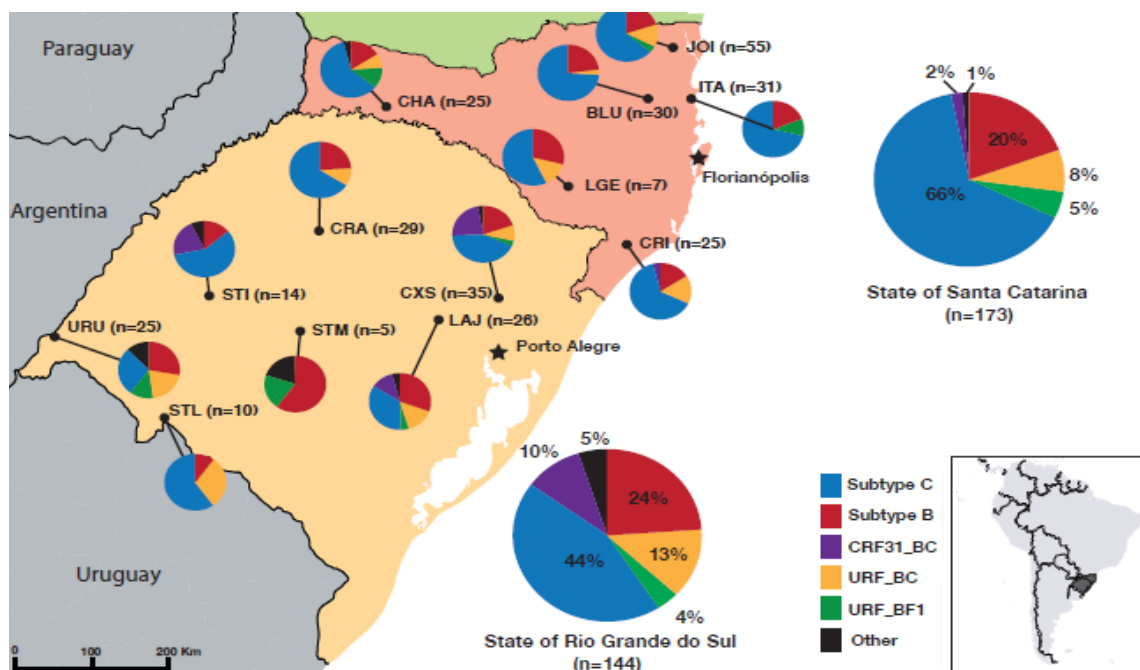


Figura 5: Representação da diversidade molecular do HIV-1 em cidades do interior do Rio Grande do Sul e Santa Catarina. Adaptado de Gräf, 2015.

1.7 O uso da bioinformática como ferramenta para análises evolutivas.

1.7.1 O uso da filogenia em estudos de epidemias virais

Filogenia é o estudo das relações evolutivas entre genes e organismos. A partir da filogenia clássica, já era possível estimar relações entre as espécies de interesse através da comparação de estruturas morfológicas, utilizando-se para isso o registro fóssil (Freeman e Herron, 2009). Além da filogenia clássica, há também a filogenia molecular. A filogenia molecular, por sua vez, utiliza-se de sequências nucleotídicas, aminoácidos ou outros dados moleculares para estabelecer relações de ancestralidade e relação entre os organismos. Devido à ausência de impressões fósseis deixadas por vírus, o estudo da filogenia de epidemias virais somente é possível através da análise de dados moleculares. Desta forma, a única maneira de descrever a história evolutiva de uma população viral é através de estudos filogenéticos de populações existentes ou pela observação de genomas virais antigos inseridos em espécies atuais (Lemey et al, 2009^a).

A filogenia molecular tornou-se uma ferramenta fundamental para investigação de novos vírus, bem como para descrever relações filogenéticas com outras populações, sua origem e história de disseminação. Esta abordagem filogenética que busca traçar a história evolutiva viral, descrevendo padrões e taxas de

transmissão, também é conhecida como epidemiologia molecular (Holmes, 1998). Essas ferramentas também são úteis no campo da virologia e epidemiologia, já que permitem aplicação no monitoramento do surgimento de cepas resistentes, diversidade genética e combate de epidemias por programas de vacinas (McCormack e Clewley, 2002).

1.7.2 A construção da árvore filogenética

Como resultado de uma análise filogenética é reconstruído um gráfico que descreve a relação descendente-ancestral dos organismos: a árvore filogenética. Como pressuposto, o método filogenético considera que as sequências analisadas são homólogas, ou seja, compartilham um ancestral comum. Ao longo do tempo de geração de um organismo, vão se acumulando modificações como deleções, inserções e trocas de nucleotídeos em seu genoma até o ponto em que, a partir de uma comparação, não é mais possível determinar sua relação filogenética com outra sequência. Em geral, para que uma árvore filogenética seja considerada útil para inferências de ancestralidade é preciso que as sequências nucleotídicas utilizadas na sua construção tenha, ao menos, 60% de similaridade entre elas (Lemey et al, 2009^a).

Para a construção desse gráfico é necessário, primeiramente, proceder ao alinhamento múltiplo das sequências. O alinhamento, realizado por um algoritmo, é etapa crucial no processo de inferência filogenética, pois é, neste momento, que os caracteres homólogos são posicionados na mesma coluna para serem analisados. Entretanto, devido a possível presença de deleções e inserções, torna-se importante que o examinador faça uma inspeção manual do arquivo gerado (Lemey et al, 2009). Os algoritmos de alinhamento mais utilizados são o Clustal W e o Muscle (Thompson et al, 1994; Edgar, 2004).

1.7.3 Modelos evolutivos

Os fatores que influenciam a forma com que os organismos evoluem podem ser diversos. Para a inferência filogenética ser acurada, é preciso que ela reflita a ação dessas forças evolutivas, contudo, com os métodos disponíveis atualmente, é possível modelar apenas o impacto de alguns desses elementos. Atualmente um

modelo evolutivo pode ser composto por um modelo de substituição nucleotídica, modelo de relógio molecular, modelo de coalescência e modelo filogeográfico de difusão espacial.

Os modelos de substituição nucleotídica descrevem o processo de troca de nucleotídeos de uma sequência genômica ao longo do tempo (Whelan, 2001). Devido a ocorrência de mutações reversas e múltiplas trocas na mesma posição nucleotídica o número real de substituições que possa ter ocorrido entre duas sequências pode ser maior que a distância genética observada entre elas. Os modelos de substituição nucleotídica tem por objetivo corrigir ou minimizar o efeito das mudanças que não podem ser mensuradas. Os diferentes modelos de substituição fazem diversas suposições acerca da probabilidade de um nucleotídeo ser trocado por outro e sobre a frequência em que isso ocorre. Para análise do HIV-1, o modelo de substituição nucleotídica mais utilizado é o GTR (Tavaré, 1986), considerando que cada nucleotídeo pode ter uma frequência e probabilidade de mudança próprios. Além disso, podem considerar que a taxa de mudança não é constante ao longo dos sítios e existência de sítios invariáveis (Yang, 1993).

A hipótese de relógio molecular indica que a distância genética entre genes é proporcional ao tempo de divergência entre as espécies a que esses genes pertencem (Zuckerandl e Pauling, 1962 e 1965). A partir de modelos de relógio molecular pode-se calcular a taxa evolutiva de um gene que, por sua vez, é utilizada para datar o ancestral comum mais recente de dois táxons. Uma maneira de auxiliar a datação de ancestralidade é a utilização de registros fósseis, desta forma o relógio molecular pode ser ajustado. Contudo, para populações virais, a inferência da taxa evolutiva torna-se mais precisa quando utilizamos sequências isoladas em diferentes momentos no tempo, criando um banco heterocrônico. Quando as modificações genômicas em uma dada espécie acumulam-se significativamente ao longo de um período, tal população é dita como de evolução mensurável no tempo (MEP – *Mensurable evolving population*) (Drummond et al, 2003). Vírus de RNA, dentre eles o HIV, estão entre as principais MEPs estudadas devido a sua elevada taxa mutacional. Há dois modelos de relógio molecular que podem ser utilizados em estudos filogenéticos: o modelo restrito e o relaxado. A principal diferença entre os dois é que o modelo relaxado assume que as taxas evolutivas podem mudar ao

longo da história evolutiva do organismo, enquanto no modelo restrito as taxas permanecem constantes (Drummond et al, 2006).

Outro fator importante na análise de uma população viral é a descrição de sua história demográfica. Os modelos de coalescência têm sido a metodologia utilizada para descrever como os parâmetros demográficos de uma dada população comportam-se ao longo do tempo, trazendo informações sobre seu tamanho e se este sofreu alterações (Griffiths e Tavaré, 1994). Atualmente, os modelos de coalescência em utilização são aplicáveis à MEPS (Rodrigo e Felsenstein, 1999) e utilizam parâmetros de crescimento demográfico (constante, exponencial, logístico e expansão) ou modelos não paramétricos que aceitam mudanças no padrão de crescimento através do tempo (skyline, skyride ou skygrid) (Ho e Shapiro, 2011; Gill et al, 2013). Os modelos de coalescência são de suma importância para análises de filodinâmica, pois a partir deles é possível descrever a história demográfica de epidemias, estimando número de infectados e taxas de expansão, tornando possível o desenvolvimento de medidas de contenção para episódios subsequentes (Rambaut et al, 2008; Pybus et al, 2001; Robbins et al, 2003).

1.7.4 Os métodos de Máxima Verossimilhança e Inferência Bayesiana

Verossimilhança é a probabilidade de uma hipótese gerada ser verdadeira. Quando aplicada à filogenia, a verossimilhança mostra a concordância dos dados com a hipótese em estudo, a partir de uma árvore construída sob um modelo evolutivo. O método de Máxima Verossimilhança (ML) é uma forma de encontrar qual topologia, tamanho de ramos e parâmetros evolutivos são capazes de descrever com maior fidedignidade os eventos pelos quais as sequências do alinhamento passaram (Lemey et al, 2009a).

Na busca pela árvore de ML são otimizados os tamanhos dos ramos e topologia. Esse processo de otimização gera, a cada etapa, um novo valor de verossimilhança. Se o novo valor for superior ao obtido pela árvore antiga, a nova é mantida e a de menor valor é descartada. Esta busca estende-se até que nenhuma modificação possível seja capaz de melhorar a verossimilhança da árvore. Uma das vantagens do método de ML sob métodos mais antigos é a capacidade de se integrar alguns modelos evolutivos (principalmente de substituição nucleotídica) à

busca da melhor árvore filogenética. Contudo, devido à análise de diversos parâmetros, o peso computacional desse método é elevado (Guindon et al, 2010).

Já a Inferência Bayesiana trata-se de um método para inferir a probabilidade de um evento ocorrer, incorporando o conhecimento de evidências prévias (também denominada probabilidades a priori ou *prior*). Na estatística Bayesiana, a possibilidade do evento ocorrer é chamada de probabilidade posterior (PP), sendo esta determinada pela análise dos dados e consideração dos *priors*. Uma das vantagens da análise Bayesiana em relação à Máxima Verossimilhança é que, no processo de reconstrução das árvores filogenéticas, árvores de menor probabilidade também podem ser “aceitas”. Este é um passo importante para atingir regiões de alta probabilidade que se localizam após um vale de baixa probabilidade no espaço paramétrico. Esta característica torna-se importante na análise, visto que o modelo de ML assume apenas parâmetros de maior verossimilhança, impossibilitando a exploração completa do espaço de possibilidades. Outra grande vantagem da inferência Bayesiana em relação a outros métodos é a acomodação da incerteza em suas análises. Isso é possível já que o objetivo final é descrever a distribuição das probabilidades posteriores no espaço paramétrico e não buscar valores que maximizem a verossimilhança (Brown e Kilmer, 2003).

O principal software de aplicação utilizado atualmente para análises bayesianas é o BEAST. Este programa é capaz de permitir a utilização de inúmeros modelos evolutivos tais como modelo de substituição nucleotídica, demográficos, relógio molecular e modelos de difusão filogeográfica (Drummond et al, 2012), tornando-se uma das principais ferramentas da atualidade para estudos de filodinâmica e filogeografia de patógenos virais.

1.7.5 Inferências filogeográficas

Filogeografia refere-se à análise filogenética de organismos, considerando o cenário geográfico da sua distribuição pelo território. Com isso, abre-se a possibilidade de testar hipóteses sobre o impacto dos eventos geográficos na separação e especiação dos organismos, explicando diferenças macroevolutivas observadas (Avise et al, 1987).

O início da análise filogeográfica provém da árvore filogenética que melhor reproduz as relações evolutivas entre as amostras. A estruturação geográfica das sequências pode ser analisada como se fosse um traço herdado do ancestral que permite estimar a localização geográfica de origem da epidemia. A dedução desses traços ancestrais pode ser realizada através de Parcimônia (Fitch, 1971), Máxima Verossimilhança (Cunningham et al, 1998) ou estatística Bayesiana (Lemey et al, 2009b).

Quando a estatística Bayesiana é escolhida para traçar as rotas de disseminação de uma epidemia viral, a localização geográfica é tratada como um caractere. Um modelo de difusão espacial serve como molde no qual as mudanças sofridas indicam eventos de migração entre uma localidade e outra. Os modelos de difusão utilizados para inferências filogeográficas podem especificar taxas de transição simétricas (a taxa de fluxo é igual entre as duas cidades: $A \rightarrow B = B \rightarrow A$) ou assimétricas (a taxa de fluxo entre as cidades é diferente: $A \rightarrow B \neq B \rightarrow A$). Além disso é possível utilizar recurso de escolha de variáveis (BSSVS) que procura manter no modelo apenas transições de maior relevância e identificar quais possuem maior suporte estatístico (Lemey et al, 2009b).

O aprimoramento de técnicas e desenvolvimento de novas metodologias de análise filogenética e filogeográfica têm contribuído muito para a caracterização de epidemias virais que acometem o homem na atualidade. São vários os exemplos recentes de estudos que aplicaram as técnicas e ajudaram a desvendar a dinâmica de expansão de epidemias no mundo, como no exemplo da introdução de Zika vírus e chikungunya nas Américas (Faria et al, 2016) ou do dispersão mundial do HIV a partir da África (Faria et al, 2014).

1.7.6 Detecção de eventos de recombinação

A evolução dos organismos não é apenas direcionada pelo acúmulo de mutações aleatórias ao longo das gerações (evolução vertical). A transferência horizontal de material genético, decorrente da mistura entre indivíduos de linhagens distintas possui um papel importante na diferenciação entre indivíduos, especialmente em populações virais. Embora as técnicas de inferência filogenética tenham evoluído, a representação em formato de árvore não é capaz de inferir sinais

de evolução horizontal, por isso torna-se necessária a utilização de métodos alternativos para identificar formas recombinantes.

Para que a detecção de eventos de recombinação ocorra é necessário que as sequências parentais que originaram a forma híbrida sejam suficientemente divergentes para identificar quem é o ancestral de cada fragmento (Lemey et al, 2009a). Quando buscamos eventos de recombinação envolvendo subtipos distintos do HIV-1 essa condição é facilmente cumprida, porém quando ocorre recombinação intrasubtipo, a divergência entre as sequências torna-se tão pequena a ponto de dificultar a detecção. Outro empecilho capaz de impossibilitar a detecção de recombinação são mutações de ponto que, por muitas vezes, acumulam-se na região recombinada. Caso esse evento tenha ocorrido há muito tempo, a referência da linhagem parental pode não ser mais encontrada.

Os diferentes métodos utilizados para promover o reconhecimento de eventos de recombinação apresentam um funcionamento similar. A partir de um alinhamento, ocorre a partição das sequências em porções sobrepostas chamadas de janela. A relação evolutiva entre as sequências é avaliada separadamente para cada janela, de forma progressiva, até cobrir todo o alinhamento. As principais diferenças entre esses métodos encontram-se na forma de medir a relação entre as sequências e o método estatístico que é aplicado a essa medida.

Um método muito utilizado para detecção de recombinação é o *bootscan*, atualmente implementado no Simplot (Lole et al, 1999) e RDP3 (Martin et al, 2010). Através dessa metodologia são construídas árvores para cada janela criada no alinhamento para avaliar as relações evolutivas da sequência de interesse com sequências referência, ou seja, os fragmentos da sequência são comparados às sequências de referência (de subtipos distintos) a fim de avaliar a similaridade entre elas.

O suporte dos clados formados é calculado por bootstrap e quando um evento é detectado, o padrão do bootstrap muda, significando que a sequência alvo pulou de um clado agrupado com uma sequência parental para outro, conforme mostra na Figura 6.

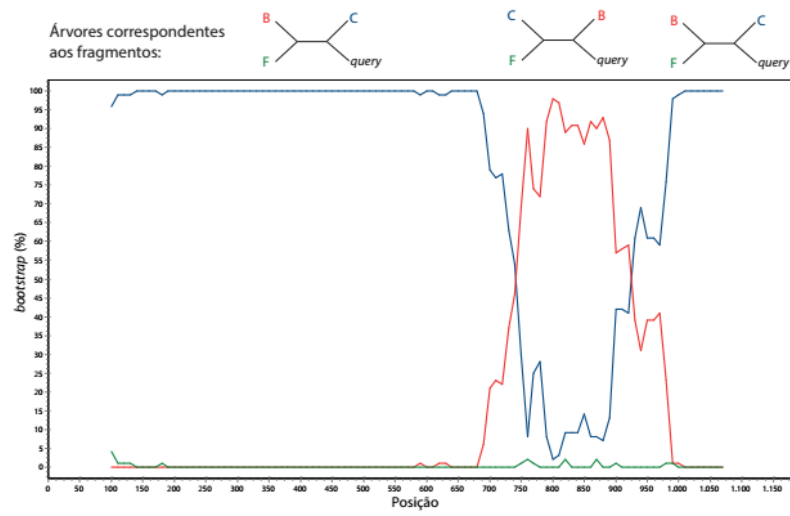


Figura 6: Análise de bootscan de uma sequência recombinante. Adaptado de Gräf, 2015.

2. JUSTIFICATIVA

A CRF31_BC é uma forma recombinante de suma importância para a epidemia local. Essa CRF, anteriormente encontrada apenas em Porto Alegre, já foi relatada em cidades do interior do estado e região Sul (Gräf, 2015), bem como em outras regiões do país e, até mesmo, na Argentina, indicando um possível processo de expansão da epidemia (Ferreira et al, 2011; Jones et al, 2009). Essa dispersão pode ainda revelar um processo de crescimento da população infectada pela CRF31_BC. Um estudo anterior já descreveu que a taxa de expansão populacional da CRF1_BC pode ser superior às taxas estimadas para o subtipo B e C (Bello et al., 2009). Em vista de um evidente processo de disseminação da CRF31_BC e da disponibilidade de novas sequências virais, torna-se possível descrever a história filogeográfica, assim como o provável local de surgimento desta cepa. Além disso, faz-se necessário a atualização das estimativas de crescimento populacional da CRF31_BC, pois tais análises podem trazer informações úteis para o desenvolvimento de novas campanhas de saúde destinadas ao público mais vulnerável, além do desenvolvimento de vacinas e estudos diferenciais que busquem mensurar a virulência dos subtipos virais e descrever as cadeias de transmissão associadas à forma recombinante, visando combater o HIV e possíveis novas infecções.

3. OBJETIVOS

3.1 Objetivos Gerais

O presente trabalho tem como objetivo principal descrever e analisar a filodinâmica da forma recombinante CRF31_BC do HIV-1 no Sul do Brasil.

3.2 Objetivos Específicos

- a) Estimar a provável data de surgimento do recombinante CRF31;
- b) Estimar o possível local de origem do recombinante CRF31;
- c) Caracterizar o padrão de dispersão filogeográfico da cepa CRF31 pelo Brasil;
- d) Reconstruir a história demográfica desse recombinante.

4. ARTIGO CIENTÍFICO

(Formatado de acordo com as regras do periódico Memórias do Instituto Oswaldo Cruz, disponível no Anexo 1).

ANÁLISE DA FILOGEOGRAFIA E HISTÓRIA DEMOGRÁFICA DA FORMA RECOMBINANTE CRF31_BC DO HIV-1 NO BRASIL

Hegger M Fritsch^{1,2}, Tiago Gräf³ e Sabrina EM Almeida^{2,4}

1. Graduação em Biomedicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
2. Centro de Desenvolvimento Científico Tecnológico, Fundação Estadual de Produção e Pesquisa em Saúde, Porto Alegre, RS, Brasil
3. Africa Centre for Health and Population Studies, University of Kwazulu-Natal, Durban, South Africa.
4. Universidade Feevale, Novo Hamburgo, RS, Brasil

Resumo

A situação epidemiológica da região Sul do Brasil é peculiar: apresenta a taxa de incidência e mortalidade mais elevadas do país e também predominância do subtipo C e circulação da CRF31_BC. Recentemente essa forma recombinante, antes descrita apenas em Porto Alegre, foi identificada no interior do estado do Rio Grande do Sul e em outros estados brasileiros, sugerindo um processo de expansão. Sabendo que o local de origem permanece incerto e frente a um aparente processo de dispersão, este estudo tem por objetivo descrever a cidade de origem, rotas de dispersão e história demográfica deste recombinante. Foram utilizadas 95 sequências e aplicados os modelos de coalescência Bayesian Skyline e testados quatro modelos

paramétricos de crescimento populacional, as análises filogeográficas foram executadas através do modelo de transição assimétrico com BSSVS. Porto Alegre mostrou-se a origem e centro da dispersão da CRF31_BC para a maioria das localidades analisadas, ademais foram observadas algumas transições independentes do Rio Grande do Sul. A taxa de crescimento estimada foi semelhante ao subtipo C. A partir dos resultados, conclui-se que a CRF31_BC não está concentrada apenas em Porto Alegre e encontra-se em um processo ativo de dispersão, onde, em algumas localidades, já há epidemias autossustentadas.

Palavras-chave: filogeografia, filogenia, recombinante CRF31_BC, HIV

Introdução

Uma das características mais marcantes do HIV-1 é sua grande variabilidade genética, sendo dividido em nove subtipos (A-D, F-H, J e K) (Lau & Wong. 2013). Essa característica é decorrente de uma combinação de fatores que promovem mutação e recombinação no genoma do HIV (Lau & Wong. 2013), tornando-o um dos organismos de mais rápida taxa evolutiva (Rambaut et al. 2004) (Grenfell et al. 2004). A grande diversidade molecular do HIV pode ser decorrente de mecanismos intrínsecos ao vírus como a elevada taxa de erro da enzima transcriptase reversa (RT) e ausência de mecanismos de reparo (Hu & Temin. 1990, Preston et al. 1988), porém também pode ser ocasionada por troca de segmentos gênicos entre partículas virais de subtipos distintos (recombinação inter-subtipo) ou similares (recombinação intra-subtipo) (Rousseau et al. 2007). Locais onde há co-circulação de mais de um subtipo do HIV-1 em uma mesma população são propícios ao surgimento de

formas recombinantes (Faria et al. 2014). Quando essa forma é restrita a menos de três indivíduos não relacionados ela é denominada forma recombinante única (URF), mas se é encontrada em três ou mais indivíduos epidemiologicamente não relacionados é denominada forma recombinante circulante (CRF) (Hemelaar. 2013). Atualmente, segundo o banco de dados Los Alamos HIV Database já foram descritas por volta de 70 CRFs, responsáveis por mais de 20% das novas infecções no mundo (Lau & Wong. 2013).

No Brasil, há, atualmente, 734 mil indivíduos vivendo com HIV (Coordenação de DST e AIDS. 2014), sendo a maioria das infecções causadas pelo subtipo B. As taxas de prevalência de Aids e incidência anual na população permanecem estáveis, contudo o cenário epidemiológico da Região Sul apresenta contrastantes diferenças. No Sul, além das taxas de incidência e mortalidade serem os mais elevados do país (Coordenação de DST e AIDS. 2014), essa região apresenta alta prevalência de subtipo C do HIV-1. Outra característica da região Sul, mais especificamente do estado do Rio Grande do Sul (RS), é a presença da forma recombinante CRF31_BC, resultado da co-circulação dos subtipos C e B (Brígido et al. 2007). Esta variante apresenta estrutura genômica mosaica, com um fragmento de subtipo B de aproximadamente 200 nucleotídeos inserido no gene da transcriptase reversa, enquanto todas as outras regiões gênicas são classificadas como subtipo C (Santos et al. 2006). A CRF31_BC é particularmente prevalente na região metropolitana de Porto Alegre, com prevalência variando entre 19% (Medeiros et al. 2010), 26% (Dias et al. 2009) e 35% (Simon et al. 2010).

Estudos anteriores relataram casos de infecção pela CRF31_BC principalmente em Porto Alegre, sendo sua data de origem estimada em 1988 (Passaes et al. 2009). Recentemente, este recombinante foi também identificada em cidades do interior do RS, sugerindo sua expansão pelo estado (Gräf 2015). Além disso, casos esporádicos já haviam sido anteriormente descritos em outros estados brasileiros e até na Argentina (Ferreira et al. 2011, Jones et al. 2009). Em vista de o local de origem da CRF31_BC ainda permanecer incerto e frente a uma aparente expansão pelo país, este estudo tem por objetivo descrever a cidade de origem, rotas de dispersão e história demográfica da epidemia da CRF31_BC.

Materiais e Métodos

Este estudo avaliou sequências geradas a partir de amostras de sangue coletadas de pacientes soropositivos para o desenvolvimento da tese de doutorado intitulada “Caracterização molecular da epidemia do HIV-1 em cidades do interior de Santa Catarina e Rio Grande do Sul e filogeográfica do subtipo C no Brasil”. Tal projeto foi submetido e aprovado pelo Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre (projeto de número 09-236) e pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Santa Catarina (UFSC), parecer número 34560. As amostras foram coletadas em 13 cidades do interior de Santa Catarina e Rio Grande do Sul a partir de cartões FTA (Whatman Inc.) e secos à temperatura ambiente. A limpeza dos confetes seguiu protocolo do fabricante utilizando de 3 a 5 confetes de sangue com 0.5mm de diâmetro, sendo limpos e secos. Posteriormente de 3 a 5 confetes foram inseridos no mix de PCR para

amplificação. O DNA pro-viral foi amplificado por Nested-PCR usando primers para o gene da polimerase (pol) (HXB2 2274-3545 pb). Os produtos de PCR foram purificados e sequenciados utilizando BigDye Terminator v3.1 Cycle Sequencing (Applied Biosystems) e analisados por ABI 3100 Genetic Analyzer (Applied Biosystems).

Também foram utilizadas sequências nucleotídicas disponíveis no banco público de sequências específicas de HIV Los Alamos HIV Sequence Database (<http://www.hiv.lanl.gov/>). Foram selecionadas todas as sequências classificadas como CRF31_BC e todas as sequências brasileiras classificadas como recombinante BC, desde que possuíssem data e o segmento disponível correspondesse à região da polimerase. Além disso, para posterior análise de filogenia, foram selecionadas randomicamente sequências gênicas da pol do subtipo A, utilizado como grupo externo, F1, B e C brasileiros.

As sequências selecionadas foram submetidas às ferramentas de subtipagem online REGA (de Oliveira et al. 2005) e SCUEAL (Pond et al. 2009). A classificação atribuída a cada sequência foi confirmada pela construção de árvores filogenéticas. As sequências de referências utilizadas para a inferência filogenética foram obtidas do banco de dados Los Alamos HIV Sequence Database (www.hiv.lanl.gov) e alinhadas utilizando o algoritmo Muscle (Edgar 2004). Árvores de Máxima Verossimilhança foram construídas usando o programa Seaview (Gouy et al. 2010) e PhyML com modelo de substituição nucleotídica GTR+G4+I. O padrão de recombinação das sequências agrupadas no cluster das CRF31_BC foi confirmado pelo método

de Bootscanning utilizando o software Simplot v3.5.1 (Lole et al. 1999) com janela de 200pb e passo de 20pb.

Árvores filogenéticas com escala de tempo foram estimadas utilizando-se estatística Bayesiana, para todas as sequências classificadas como CRF31_BC. Uma primeira análise exploratória foi realizada utilizando os seguintes modelos evolutivos: modelo de substituição nucleotídica GTR+G4+I, modelo de relógio molecular relaxado, não correlacionado com distribuição lognormal (*lognormal uncorrelated relaxed clock*) e modelo de coalescência não-paramétrico *Bayesian Skyline*. Para inferir a taxa de crescimento da epidemia foram testados quatro modelos demográficos paramétricos distintos para o mesmo dataset: modelo de crescimento exponencial, logístico, expansão e de população constante. Os demais parâmetros foram mantidos iguais ao da análise exploratória. A comparação dos modelos demográficos foi feito pelo método de Estimativa da Verossimilhança Marginal (MLE). As análises filogenéticas Bayesianas foram realizadas utilizando-se o pacote BEAST v. 1.8.2 (Drummond et al. 2012). Os resultados gerados foram analisados usando o software TRACER v. 1.6 e a convergência dos parâmetros foi considerada satisfatória quando o Tamanho Efeitoivo da Amostra (ESS) era superior a 200. O dataset correu por 1×10^8 gerações e *burn-in* utilizado foi de 10%.

As análises filogeográficas foram executadas em uma abordagem integrada à reconstrução das árvores filogenéticas no pacote BEAST. Para estimar o local de origem e a dispersão da CRF31_BC foi informado ao programa o local de origem de cada sequência, quando a informação da cidade não estava

disponível utilizamos o estado. Para checarmos a influência da heterogeneidade do número de amostras por localidade nas análises filogeográficas, três novos datasets foram criados onde o número de sequências da cidade de Porto Alegre foi randomicamente reduzido a 12. As análises filogenéticas foram executadas utilizando-se do modelo de transição assimétrico, com inferência social de rede com BSSVS. Todos os demais parâmetros evolutivos permaneceram iguais à análise exploratória. A análise por modelo assimétrico correu com comprimento de cadeia de 5×10^8 e burn-in de 10%. A árvore com máxima credibilidade (MCC) foi obtida a partir do software TreeAnnotator v1.8.3 e visualizada no FigTree v1.4.2. As taxas de transição de sequências virais entre as localidades foi analisado pelo programa SPREAD v1.0.6 (Bielejec et al. 2011) e a intensidade das transições entre as cidades foi inferida pelo Bayes Factor (BF) com cut-off de 3.0.

Resultados

O dataset gerado foi formado por 95 sequências classificadas como CRF31_BC (Tabela 1). Para 15 sequências, a informação de local de origem não estava disponível, sendo que para estas a análise filogeográfica foi responsável por estimar a provável localidade de amostragem.

Nossas análises reconstruíram que o ano provável de origem da CRF31_BC é 1985 (95% HPD = 1978 -1989) e que seu surgimento ocorreu muito provavelmente em Porto Alegre (PP=0,998) (Fig. 1). Tais resultados foram confirmados pelas análises dos datasets reduzidos. Através da reconstrução das localidades nos nodos ancestrais das árvores filogenéticas é

possível presumir que, inicialmente, essa forma recombinante circulou localmente e seu processo de disseminação para outras cidades se iniciou cerca de cinco anos após seu surgimento. Os primeiros eventos de transmissão para outras cidades datam do final da década de 1980 e início da década de 1990. As primeiras cidades onde essa introdução ocorreu são Florianópolis (1989, 1986-1993 95%HPD), Caxias do Sul (1989, 1987-1990 95%HPD), Lajeado (1990, 1986-1994 95%HPD) e o estado de Mato Grosso (1990, 1987-1990 95%HPD).

Na análise, foi possível observar que a CRF31_BC está circulando em algumas localidades independentemente de novas introduções provenientes de Porto Alegre, ou seja, podemos assumir que a forma recombinante está circulando internamente na população. Essa circulação autossustentada pode ser observada nas cidades de Caxias do Sul, Florianópolis, Lajeado, Criciúma e no estado de Goiás. Além disso, a partir da Fig.1 podemos inferir que as sequências da CRF31_BC dos estados de São Paulo e Rio de Janeiro são relacionadas, visto que as amostras permaneceram em um mesmo clado da árvore.

A intensidade e significância das transições virais entre as cidades foi mensurada pela análise de Bayes Factor com cut-off de 3.0 (Tabela 2). As análises filogeográficas estimaram que Porto Alegre é uma cidade central para a dispersão da CRF31_BC dentro do Rio Grande do Sul e para outros estados do país, sendo origem de 14 das 22 transições significativas (63,63%) (Tabela 2 e Fig. 2). Além de Porto Alegre, Caxias do Sul também foi estimada como

desempenhando um papel importante na disseminação da CRF31_BC, mais precisamente para Itajaí e para o Piauí. Interessante ainda notar que o estado do Rio de Janeiro foi estimado como origem em transições para Pernambuco, São Paulo e inclusive para a cidade gaúcha de Rio Grande. Entre os resultados, também chama a atenção a circulação da CRF31_BC entre os estados de Goiás e Mato Grosso, no Centro-Oeste do Brasil (Fig. 2).

A partir do gráfico de crescimento populacional, gerado pelo modelo Bayesian SkyLine (Fig. 3), podemos observar que, logo após seu surgimento, a epidemia da CRF31_BC apresentou um crescimento de forma exponencial nos primeiros 7,5 anos, seguido por uma estabilização do crescimento, a partir de 1992, que perdura até os dias atuais. Esses resultados sugerem que o modelo paramétrico que melhor descreve o comportamento demográfico desta população é o logístico. Entretanto, neste trabalho testamos qual o melhor modelo demográfico para a CRF31_BC e de fato crescimento logístico apresentou-se como o melhor modelo para os dados analisados. A análise de crescimento populacional da CRF31_BC indicou uma taxa de crescimento média estimada em 0.66/ano (0,42-0,88 95%HPD), além disso o tempo de duplicação da população estimado foi de 1,04 (0,78-1,65 95%HPD), indicando que, no momento exponencial da curva logística de crescimento, a população viral da CRF31_BC dobrava de tamanho efetivo a cada ano.

Discussão

A data de surgimento da CRF31_BC estimada neste estudo (1985, 1978 - 1989 95%HPD) revelou-se levemente mais precoce do que já havia sido

estimado por Passaes et al (2009) (1988,1979-1993 95%HPD) e Santo et al. (2007) (1987, 1957-1998 95%HPD), apesar da sobreposição dos intervalos de confiança. Esta variação de resultados pode ser explicada pelo número reduzido de sequências analisadas em estudos anteriores conduzidos por Santos et al. (2007) (n=21 sequências) e por Passaes et al. (2009) (n= 47 sequências) contra um tamanho amostral de 95 sequências no presente trabalho. Além disso o fato de obtermos intervalos de confiança mais restritos é evidência de uma estimativa mais precisa. Os resultados encontrados aqui estão em concordância com a estimativa de origem dos subtipos B e C no Brasil. Segundo Bello et al. (2006, 2007) a introdução do subtipo B no Brasil ocorreu em meados de 1960, enquanto a introdução do C foi estimada em 1976 (1966-1983, 95%HPD) (Delatorre et al. 2013). O fato de estudos reportarem infecção por CRF31_BC em indivíduos diagnosticados no início dos anos 1990 (Santos et al. 2007) (Passaes et al. 2009) restringe o surgimento da CRF31_BC em algum ponto anterior a 1990, mas posterior a 1976.

Nesse estudo, foi possível inferir que provavelmente o local de origem da CRF31_BC seja a cidade de Porto Alegre. Embora nenhum outro estudo tenha procurado comprovar Porto Alegre como cidade de origem desta forma recombinante, esta hipótese já havia sido sugerida por Passaes et al. (2009). A presença de uma população de aproximadamente 4 milhões de habitantes, co-circulação do subtipo C e B e a maior taxa de incidência de Aids no país possibilitou a criação do cenário perfeito para que ocorresse a recombinação inter-subtipo e emergência dessa forma recombinante. Outro fator que pode ter ajudado no surgimento da CRF31_BC foi a circulação do subtipo C na

população de usuários de drogas injetáveis em Porto Alegre antes da década de 1990 (Almeida et al. 2012). Esta via de transmissão é conhecida por facilitar eventos de recombinação devido a grande carga viral transmitida no momento do compartilhamento das seringas (Zhao et al. 2014).

Além do ponto de origem, Porto Alegre foi também estimada como sendo o centro principal de disseminação da CRF31_BC para outras localidades dentro do próprio estado para outras unidades federativas e, até mesmo, para fora do país (Fig. 2 e Tabela 2). Um estudo recente também descreveu Porto Alegre com sendo o ponto central de dispersão do subtipo C do HIV-1 no Brasil, mostrando que a alta prevalência de infecções por HIV na cidade é um dos preditores deste processo (Gräf et al. 2015). A alta prevalência da CRF31_BC em cidades como Lajeado e Caxias do Sul, próximas ao foco de origem, pode ser explicada pela proximidade geográfica com Porto Alegre, facilitando os processos migratórios. Já para as cidades de Santiago e Rio Grande, localizadas a 443 e 318 quilômetros da capital respectivamente, o fluxo de pessoas deve ser menor, restringido pela distância geográfica. Rio Grande, por sua vez, apresenta intensa circulação e rotatividade de pessoas possibilitada pela presença de zona portuária, motivo pelo qual poderíamos justificar a transição com o estado do Rio de Janeiro, embora não tenhamos encontrado referências que suportem essa hipótese.

Além de ser um núcleo urbano para a disseminação dentro do próprio estado, Porto Alegre atua também como um centro de dispersão para outros estados do Brasil. Uma das transições interestaduais mais intensas reveladas

nesse estudo é com a capital de Santa Catarina: Florianópolis. Um dos motivos pelo qual a presença da CRF31_BC pode ser explicada é devido à relativa proximidade geográfica entre as capitais (478 quilômetros) e ao turismo. Contudo, vale ressaltar também que no estado de Santa Catarina a epidemia causada pelo subtipo C é muito forte, podendo indicar uma rede de transmissão e troca de variantes virais com o RS.

A disseminação da CRF31_BC envolvendo o Rio Grande do Sul e outras regiões do país pode ser explicada, principalmente, pelos processos migratórios da população. Podemos inferir que há fluxo populacional entre os estados do Sul, de forma mais intensa envolvendo cidades do RS, com outros centros urbanos do país pelo aumento da frequência de HIV-1C. A presença desse subtipo, que é considerado raro ou pouco frequente em outras regiões do país, indica que há transição viral de zonas onde o HIV-1C é mais frequente para zonas de menor prevalência. Um exemplo disso é o aumento da prevalência de subtipo C no estado de São Paulo. Em estudo realizado por Brígido et al. (2011) a prevalência encontrada já superou os 8%. Segundo Souza et al. (2009), responsável por relatar uma criança infectada pela CRF31_BC, a presença do recombinante pode indicar a emergência dessa forma no estado e sugerir que possa haver impacto de outras epidemias onde esta forma é mais presente.

A presença da CRF31_BC em estados do Centro-Oeste pode estar intimamente relacionada ao processo de deslocamento de populações do Sul para a região em busca de emprego na agricultura e pecuária. Como

observamos neste trabalho, não apenas Porto Alegre, mas a cidade de Santiago também atua de forma importante na rota de disseminação para a região Centro-Oeste, fortalecendo a hipótese que a migração de trabalhadores do setor agropecuário possa estar agindo nesse processo. Possivelmente a migração dessas populações também tenha sido responsável também pelo aumento do subtipo C no Centro-Oeste onde já são relatadas 17,7% de HIV-1 subtipo C ou que contenham segmentos C em Cuiabá (Ferreira et al. 2011) e aproximadamente 20% na população carcerária do Centro-Oeste (Cardoso et al. 2011).

Em nosso estudo foi possível observar também que há transição viral de CRF31 ocorrendo de forma independente, não vinculada ao Rio Grande do Sul. Observamos que há interligação entre a epidemia de São Paulo e Rio de Janeiro, visto que nesses estados estão alocados os centros financeiros, comerciais e industriários mais importantes do país, gerando um grande fluxo populacional entre estes dois estados. Além disso, foi possível identificar transições entre o estado do Rio de Janeiro e Pernambuco, possivelmente associadas ao turismo. Outra transição independente observada em nosso estudo envolve os estados de Goiás e Mato Grosso, justificada pelo comércio agrícola e pecuarista da região, possibilitando que a epidemia em ambas as localidades sejam mantidas.

Os resultados que descrevem a história demográfica mostram-se consistentes com um processo de crescimento logístico, indicando que a epidemia causada por essa forma recombinante não continua a expandir-se de

forma exponencial nos últimos anos. Segundo estudo realizado por Bello et al. (2009), a CRF31_BC teve um período de crescimento exponencial entre os anos de 1980 e 1990, seguidos por um período de estabilidade. Os resultados obtidos neste estudo são parcialmente condizentes com o observado por Bello et al. (2009), porém o período de crescimento exponencial encontrado aqui compreende os anos de 1985 a 1992, aproximadamente. A dinâmica de crescimento da CRF31_BC aparenta ter sofrido modificações ao longo dos anos, apresentando uma lentificação na taxa de crescimento quando comparada a estudos anteriores. A taxa de crescimento descrita em Bello et al. (2009) estima que a epidemia cresça aproximadamente 1,3/ano (0,44-2,26, 95%HPD) com *doubling time* de 0,55 (0,31-1,57, 95%HPD), contudo nossos resultados indicam que a taxa de crescimento da CRF31_BC está mais lenta, visto que encontramos taxa de crescimento média de 0,66/ano (0,42-0,88 95%HPD) e *doubling time* de 1,04 (0,78-1,65, 95%HPD). Além disso, o intervalo de confiança obtido é menor, indicando que nossas estimativas apresentam maior precisão, descrevendo melhor o comportamento da população ao longo do tempo. Estas diferenças devem ser devido ao dataset analisado, sendo que no presente estudo sequências mais recentes também estão presentes. Um maior número de sequências também deve ser o motivo de estimativas mais precisas reportadas pelo nosso estudo, tendo em vista que os intervalos de confiança estimados por Bello et al. (2009) para os dois parâmetros são maiores que os agora estabelecidos. Quando comparamos a taxa de crescimento da CRF31_BC com as taxas dos subtipos B e C, é possível inferir que a forma recombinante cresceu de maneira similar ao

subtipo C no Rio Grande do Sul (0,81, 0,40-1,26 95%HPD) e de forma mais acentuada que o subtipo B no Brasil (0,56, 0,35-0,80) (Bello et al.2009, Bello et al. 2007). A diferença entre taxas de propagação entre os subtipos B, C e recombinante CRF31_BC podem estar mais atreladas às variações nas redes de transmissão do que na infectividade e transmissibilidade viral, visto que não há estabelecida relação entre os diferentes subtipos do HIV-1 e capacidade de se disseminar na população humana (Bello et al. 2009).

A partir dos resultados obtidos é possível concluir que Porto Alegre, além de ser local de origem da forma recombinante CRF31_BC, desempenha um papel importante no processo de disseminação de novas variantes do HIV-1 pelo Brasil. O entendimento das dinâmicas de crescimento e dispersão viral do HIV-1 é de grande valia para o desenvolvimento de políticas públicas que visem combater a epidemia de Aids e reduzir o número de novas infecções nesse núcleo urbano, principalmente pela compreensão das cadeias de transmissão associadas a esses fenômenos e determinação de populações vulneráveis.

Estudos adicionais que busquem avaliar os demais recombinantes presentes no cenário do Sul do país, principalmente as URFs, também são importantes, visto que são indicativos de múltiplas infecções. O melhor entendimento dos fatores associados à repetida exposição poderá esclarecer o comportamento de risco da população, bem como melhorar os programas de prevenção e ampliar as campanhas de testagem. Por isso, torna-se necessário a manutenção de um programa continuado em vigilância epidemiológica e

monitoramento do impacto das novas formas virias no prognóstico dessas novas infecções.

Referências Bibliográficas

Almeida SE, de Medeiros RM, Junqueira DM, Gräf T, Passaes CP, Bello G, et al. Temporal dynamics of HIV-1 circulating subtypes in distinct exposure categories in southern Brazil. *Virology*. 2012;9:306.

Bello G, Eyer-Silva WA, Couto-Fernandez JC, Guimarães ML, Chequer-Fernandez SL, Teixeira SL, et al. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect Genet Evol*. 2007;7(2):263-70.

Bello G, Guimarães ML, Morgado MG. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS*. 2006;20(5):763-8.

Bello G, Guimarães ML, Passaes CP, Matos Almeida SE, Veloso VG, Morgado MG. Short communication: Evidences of recent decline in the expansion rate of the HIV type 1 subtype C and CRF31_BC epidemics in southern Brazil. *AIDS Res Hum Retroviruses*. 2009;25(11):1065-9.

Bielejec F, Rambaut A, Suchard MA, Lemey P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics*. 2011;27(20):2910-2.

Brígido LF, Ferreira JL, Almeida VC, Rocha SQ, Ragazzo TG, Estevam DL, et al. Southern Brazil HIV type 1 C expansion into the state of São Paulo, Brazil. *AIDS Res Hum Retroviruses*. 2011;27(3):339-44.

Brígido LF, Nunes CC, Oliveira CM, Knoll RK, Ferreira JL, Freitas CA, et al. HIV type 1 subtype C and CB Pol recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. *AIDS Res Hum Retroviruses*. 2007;23(12):1579-86.

Cardoso LP, da Silveira AA, Francisco RB, da Guarda Reis MN, Stefani MM. Molecular characteristics of HIV type 1 infection among prisoners from Central Western Brazil. *AIDS Res Hum Retroviruses*. 2011;27(12):1349-53.

Coordenação de DST e AIDS, Ministério da Saúde. Boletim Epidemiológico AIDS. Ano III ed2014.

de Medeiros RM, Junqueira DM, Matte MC, Barcellos NT, Chies JA, Matos Almeida SE. Co-circulation HIV-1 subtypes B, C, and CRF31_BC in a drug-naïve population from Southernmost Brazil: analysis of primary resistance mutations. *J Med Virol*. 2011;83(10):1682-8.

de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics*. 2005;21(19):3797-800.

Delatorre E, Couto-Fernandez JC, Guimarães ML, Vaz Cardoso LP, de Alcantara KC, Stefani MM, et al. Tracing the origin and northward dissemination dynamics of HIV-1 subtype C in Brazil. *PLoS One*. 2013;8(9):e74072.

Dias CF, Nunes CC, Freitas IO, Lamego IS, Oliveira IM, Gilli S, et al. High prevalence and association of HIV-1 non-B subtype with specific sexual

transmission risk among antiretroviral naïve patients in Porto Alegre, RS, Brazil. *Rev Inst Med Trop Sao Paulo*. 2009;51(4):191-6.

Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29(8):1969-73.

Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004;5:113.

Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014;346(6205):56-61.

Ferreira AS, Cardoso LP, Stefani MM. Moderate prevalence of transmitted drug resistance and high HIV-1 genetic diversity in patients from Mato Grosso State, Central Western Brazil. *J Med Virol*. 2011;83(8):1301-7.

Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol*. 2010;27(2):221-4.

Gräf T. Caracterização molecular da epidemia do HIV-1 em cidades do interior de Santa Catarina e Rio Grande do Sul e filogeografia do subtipo C no Brasil. Tese de Doutorado. Florianópolis, Santa Catarina - Brasil.: Universidade Federal de Santa Catarina; 2015. p. 151.

Gräf T, Vrancken B, Maletich Junqueira D, de Medeiros RM, Suchard MA, Lemey P, et al. Contribution of Epidemiological Predictors in Unraveling the Phylogeographic History of HIV-1 Subtype C in Brazil. *J Virol.* 2015;89(24):12341-8.

Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science.* 2004;303(5656):327-32.

Hemelaar J. Implications of HIV diversity for the HIV-1 pandemic. *J Infect.* 2013;66(5):391-400.

Hu WS, Temin HM. Retroviral recombination and reverse transcription. *Science.* 1990;250(4985):1227-33.

Jones LR, Dilemnia DA, Manrique JM, Moretti F, Salomón H, Gomez-Carrillo M. In-depth analysis of the origins of HIV type 1 subtype C in South America. *AIDS Res Hum Retroviruses.* 2009;25(10):951-9.

Lau KA, Wong JJ. Current Trends of HIV Recombination Worldwide. *Infect Dis Rep.* 2013;5(Suppl 1):e4.

Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol.* 1999;73(1):152-60.

Passaes CP, Bello G, Lorete RS, Matos Almeida SE, Junqueira DM, Veloso VG, et al. Genetic characterization of HIV-1 BC recombinants and evolutionary history of the CRF31_BC in Southern Brazil. *Infect Genet Evol.* 2009;9(4):474-82.

Pond SLK, Posada D, Stawiski E, Chappey C, Poon A, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput Biol.* 2009.

Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science.* 1988;242(4882):1168-71.

Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. *Nat Rev Genet.* 2004;5(1):52-61.

Rousseau CM, Learn GH, Bhattacharya T, Nickle DC, Heckerman D, Chetty S, et al. Extensive intrasubtype recombination in South African human immunodeficiency virus type 1 subtype C infections. *J Virol.* 2007;81(9):4492-500.

Santos AF, Schrago CG, Martinez AM, Mendoza-Sassi R, Silveira J, Sousa TM, et al. Epidemiologic and evolutionary trends of HIV-1 CRF31_BC-related strains in southern Brazil. *J Acquir Immune Defic Syndr.* 2007;45(3):328-33.

Santos AF, Sousa TM, Soares EA, Sanabani S, Martinez AM, Sprinz E, et al. Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS*. 2006;20(16):2011-9.

Simon D, Béria JU, Tietzmann DC, Carli R, Stein AT, Lunge VR. Prevalence of HIV-1 subtypes in patients of an urban center in Southern Brazil. *Rev Saude Publica*. 2010;44(6):1094-101.

Souza AC, Oliveira CM, Marques HH, Levi JE. Description of an HIV-1 BC recombinant virus identified in a pediatric patient in the city of São Paulo. *Braz J Infect Dis*. 2009;13(1):67-9.

Zhao K, Du J, Zheng W, Yu XF. HIV-1 transmission among injection drug users leads to novel recombinants circulating in southern China. *AIDS Res Hum Retroviruses*. 2014;30(10):1006-9.

Figuras e Tabelas

Tabela 1: Distribuição geográfica das sequências CRF31_BC constituintes do dataset analisado neste estudo.

Estado	Cidade	N	%
RS	Porto Alegre	48	50,53%
	Caxias do Sul	8	8,42%
	Lajeado	4	4,21%
	Santiago	3	3,16%
	Rio Grande	1	1,05%
SC	Itajaí	1	1,05%
	Criciúma	2	2,11%
	Florianópolis	3	3,16%
SP	São Paulo	1	1,05%
RJ	Rio de Janeiro	2	2,11%
MS	Campo Grande	1	1,05%
MT	Cuiabá	1	1,05%
GO	Goiânia	2	2,11%
PI	Teresina	1	1,05%
PE	Recife	1	1,05%
Exterior	Madri	1	1,05%
	NA	15	15,78%
Total		95	100%

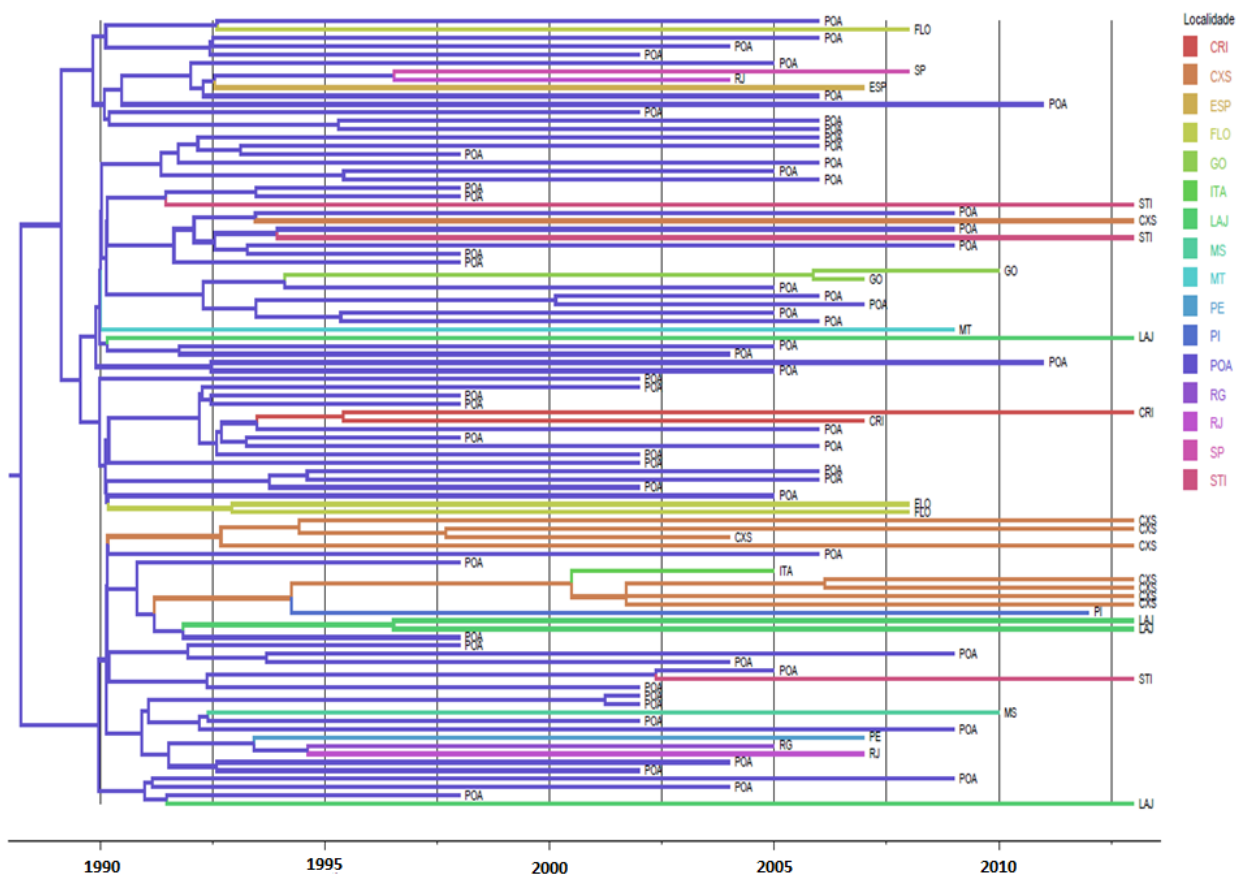


Fig. 1: Árvore filogenética Bayesiana em escala de tempo das sequências *pol* da CRF31_BC. Reconstrução das localidades ancestrais foi realizada por modelo assimétrico de transição usando BSSVS. Ramos da árvore estão coloridos conforme legenda. Siglas: CRI (Criciúma), CXS (Caxias do Sul), ESP (Espanha), FLO (Florianópolis), GO (Goiás), ITA (Vale do Itajaí), LAJ (Lajeado), MS (Mato Grosso do Sul), MT (Mato Grosso), PE (Pernambuco), PI (Piauí), POA (Porto Alegre), RG (Rio Grande), RJ (Rio de Janeiro), SP (São Paulo) e STI (Santiago).

Tabela 2: Transições significativas da forma recombinante CRF31_BC entre as localidades analisadas e Bayes Factor de suporte.

Bayes Factor (BF)	Cidade de Origem	Cidade de Destino
BF=1738	Porto Alegre	Caxias do Sul
BF=1406	Porto Alegre	Lajeado
BF=1223	Porto Alegre	Santiago
BF=793	Porto Alegre	Florianópolis
BF=146	Caxias do Sul	Itajaí
BF=91	Porto Alegre	Rio de Janeiro
BF=85	Porto Alegre	Criciúma
BF=12	Porto Alegre	Espanha
BF=11	Porto Alegre	Goiás
BF=10	Rio de Janeiro	São Paulo
BF=10	Porto Alegre	Mato Grosso
BF=9	Porto Alegre	Pernambuco
BF=9	Porto Alegre	Mato Grosso do Sul
BF=7	Caxias do Sul	Piauí
BF=7	Rio de Janeiro	Rio Grande
BF=7	Porto Alegre	Rio Grande
BF=6	Porto Alegre	São Paulo
BF=4	Porto Alegre	Pernambuco
BF=4	Rio de Janeiro	Pernambuco
BF=3	Mato Grosso	Goiás
BF=3	Santiago	Mato Grosso do Sul
BF=3	Goiás	Mato Grosso

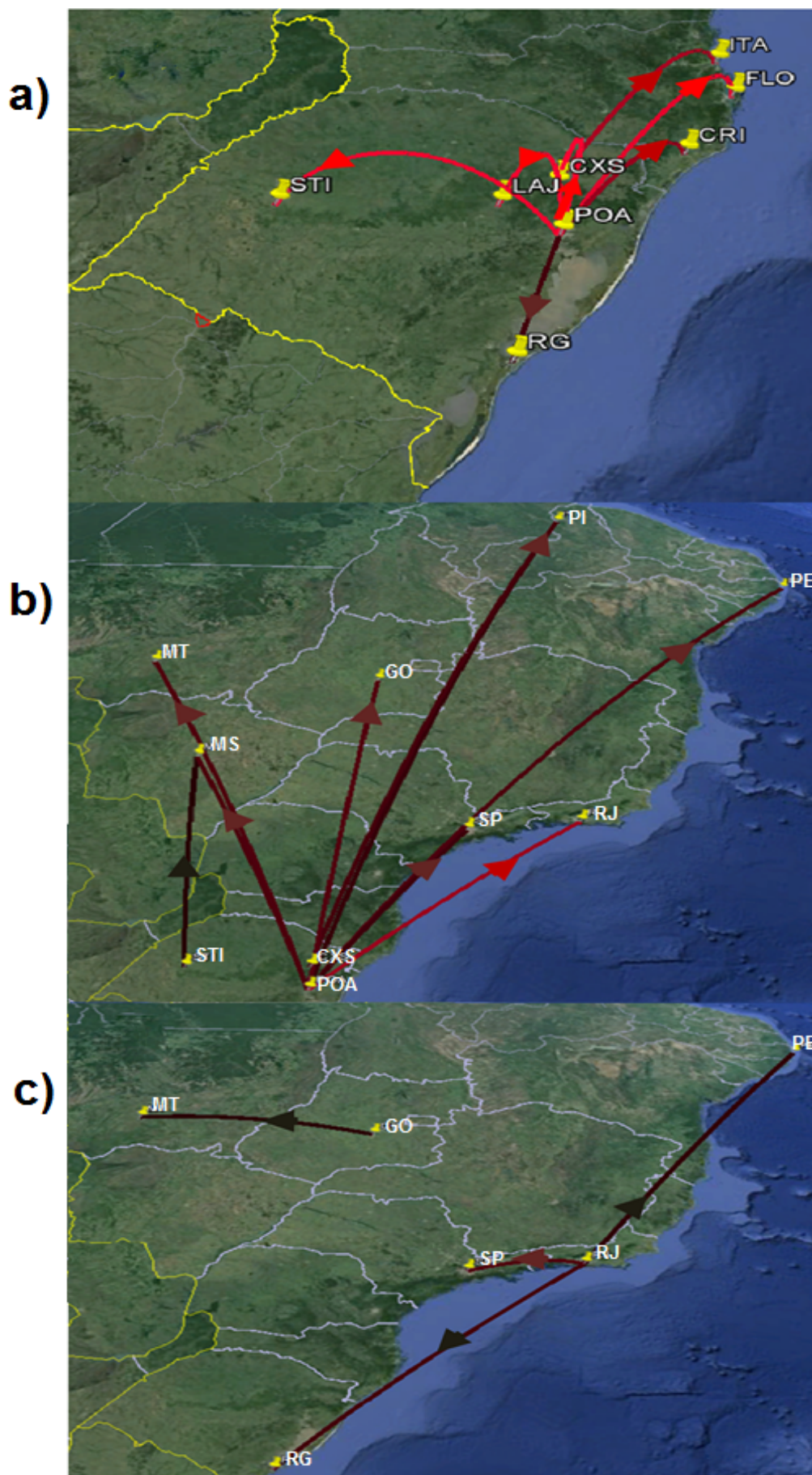


Fig. 2: Transições significativas da CRF31_BC entre localidades amostradas no Brasil. Setas indicam a direção da transição e estão coloridas conforme o BF de suporte. Vermelho mais claro indica suporte maior e cores mais escuras suporte menor. a) Transições significativas da CRF31_BC intra-região Sul. b) Transições significativas da CRF31_BC com origem na região Sul. c) Transições significativas da CRF31_BC independentes do RS.

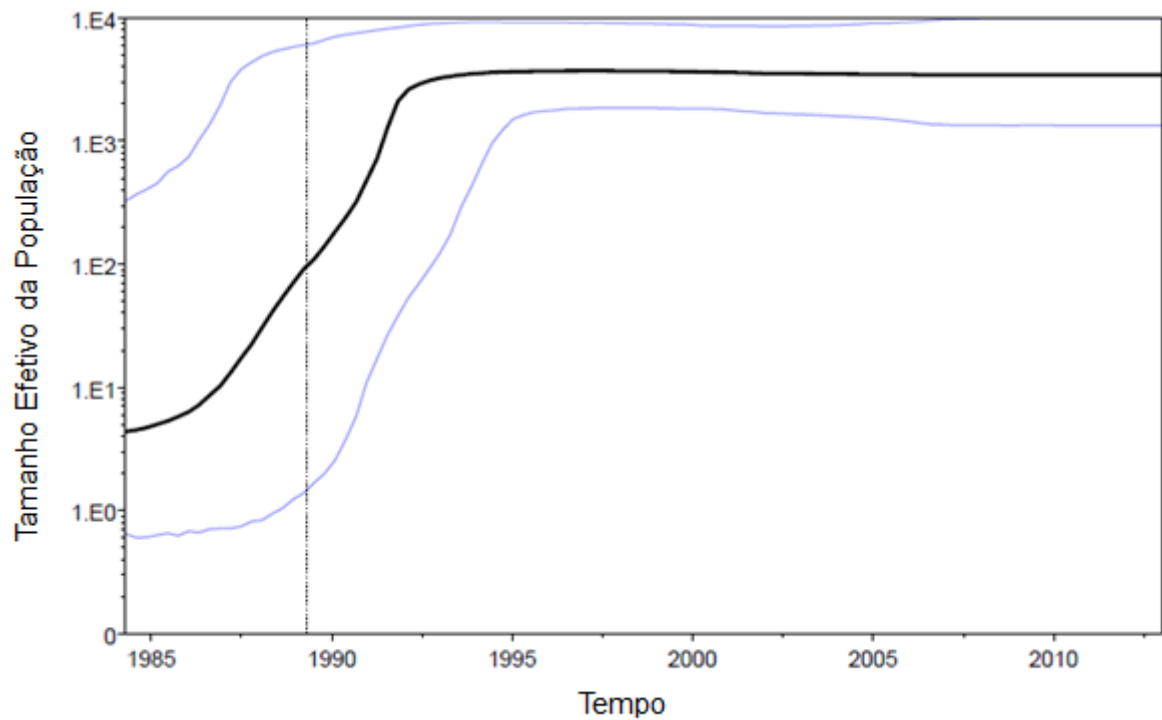


Fig. 3: Gráfico de crescimento populacional da CRF31_BC obtido pelo método de análise Bayesiana por Skyline. Valores médios da taxa de crescimento são representados pela linha preta e os intervalos de 95% HPD são representados pelas linhas azuis. A barra pontilhada representa o limite inferior da altura da raiz estimada.

5. CONCLUSÃO E PERSPECTIVAS FUTURAS

A partir dos resultados obtidos é possível concluir que Porto Alegre, além de ser local de origem da forma recombinante CRF31_BC, desempenha um papel importante no processo de disseminação de novas variantes do HIV-1 pelo Brasil. O entendimento das dinâmicas de crescimento e dispersão viral do HIV-1 é de grande valia para o desenvolvimento de políticas públicas que visem combater a epidemia de Aids e reduzir o número de novas infecções nesse núcleo urbano, principalmente pela compreensão das cadeias de transmissão associadas a esses fenômenos e determinação de populações vulneráveis. Estudos adicionais que busquem avaliar os demais recombinantes presentes no cenário do Sul do país, principalmente as URFs, também são importantes, visto que são indicativos de múltiplas infecções. O melhor entendimento dos fatores associados à repetida exposição poderá esclarecer o comportamento de risco da população, bem como melhorar os programas de prevenção e ampliar as campanhas de testagem. Por isso, torna-se necessário a manutenção de um programa continuado em vigilância epidemiológica e monitoramento do impacto das novas formas virais no prognóstico dessas novas infecções.

REFERÊNCIAS BIBLIOGRÁFICAS

ABECASIS, A. B.; VANDAMME, A. M.; LEMEY, P. Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. **J Virol**, v. 83, n. 24, p. 12917-24, Dec 2009. ISSN 1098-5514. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19793809> >.

ARIËN, K. K.; VANHAM, G.; ARTS, E. J. Is HIV-1 evolving to a less virulent form in humans? **Nat Rev Microbiol**, v. 5, n. 2, p. 141-51, Feb 2007. ISSN 1740-1534. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17203103> >.

AVISE, J. C. et al. **Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics**. *Ann Rev Ecol Syst.* 18: 489-522 p. 1987.

BELLO, G. et al. Short communication: Evidences of recent decline in the expansion rate of the HIV type 1 subtype C and CRF31_BC epidemics in southern Brazil. **AIDS Res Hum Retroviruses**, v. 25, n. 11, p. 1065-9, Nov 2009. ISSN 1931-8405. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19895209> >.

BRINDEIRO, R. M. et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. **AIDS**, v. 17, n. 7, p. 1063-9, May 2003. ISSN 0269-9370. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12700457> >.

BROWN, J. W.; KILMER, A. J. **The state of bayesian phylogenetics: Bayes for the uninitiated**.: Kingston: Queen's University 2003.

BRÍGIDO, L. F. et al. HIV type 1 subtype C and CB Pol recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. **AIDS Res Hum Retroviruses**, v. 23, n. 12, p. 1579-86, Dec 2007. ISSN 0889-2229. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18160017> >.

CARDOSO, L. P.; QUEIROZ, B. B.; STEFANI, M. M. HIV-1 pol phylogenetic diversity and antiretroviral resistance mutations in treatment naïve patients from Central West Brazil. **J Clin Virol**, v. 46, n. 2, p. 134-9, Oct 2009. ISSN 1873-5967. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19682948> >.

CERQUEIRA, D. M. et al. Antiretroviral resistance and genetic diversity of human immunodeficiency virus type 1 isolates from the Federal District, Central Brazil. **Mem Inst Oswaldo Cruz**, v. 99, n. 8, p. 877-82, Dec 2004. ISSN 0074-0276. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15761606> >.

COORDENAÇÃO DE DST E AIDS, Ministério da Saúde. **Boletim Epidemiológico AIDS** 2014.

COUTO-FERNANDEZ, J. C. et al. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: tracing the origin of this subtype in Brazil. **AIDS Res Hum Retroviruses**, v. 22, n. 2, p. 207-11, Feb 2006. ISSN 0889-2229. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16478405> >.

CULLEN, B. R. HIV-1 auxiliary proteins: making connections in a dying cell. **Cell**, v. 93, n. 5, p. 685-92, May 1998. ISSN 0092-8674. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9630214> >.

CUNNINGHAM, C. W.; OMLAND, K. E.; OAKLEY, T. H. Reconstructing ancestral character states: a critical reappraisal. **Trends Ecol Evol**, v. 13, n. 9, p. 361-6, Sep 1998. ISSN 0169-5347. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21238344> >.

DE MEDEIROS, R. M. et al. Co-circulation HIV-1 subtypes B, C, and CRF31_BC in a drug-naïve population from Southernmost Brazil: analysis of primary resistance mutations. **J Med Virol**, v. 83, n. 10, p. 1682-8, Oct 2011. ISSN 1096-9071. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21837783> >.

DIAS, C. F. et al. High prevalence and association of HIV-1 non-B subtype with specific sexual transmission risk among antiretroviral naïve patients in Porto Alegre, RS, Brazil. **Rev Inst Med Trop Sao Paulo**, v. 51, n. 4, p. 191-6, 2009 Jul-Aug 2009. ISSN 1678-9946. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19738998> >.

DRUMMOND, A. J. et al. **Measurably evolving populations**. *Trends Ecol Evol*. 9: 481-488 p. 2003.

DRUMMOND, A. J. et al. Relaxed phylogenetics and dating with confidence. **PLoS Biol**, v. 4, n. 5, p. e88, May 2006. ISSN 1545-7885. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16683862> >.

DRUMMOND, A. J. et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7. **Mol Biol Evol**, v. 29, n. 8, p. 1969-73, Aug 2012. ISSN 1537-1719. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22367748> >.

ZUCKERKANDL, E; PAULING, L. **Molecular disease, evolution and genetic heterogeneity**. Kasha M, Pullman B. Horizons of biochemistry.: New York: Academic press.: 185-225 p. 1962.

ZUCKERKANDL, E; PAULING, L. **Evolutionary divergence and convergence in proteins**. Bryson V, Vogel HJ. *Envolving genes and proteins*.: New York: Academic press: 97-166 p. 1965.

EDGAR, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, v. 5, p. 113, Aug 2004. ISSN 1471-2105. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15318951> >.

FARIA, N. R. et al. Zika virus in the Americas: Early epidemiological and genetic findings. **Science**, v. 352, n. 6283, p. 345-9, Apr 2016. ISSN 1095-9203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/27013429> >.

FARIA, N. R. et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. **Science**, v. 346, n. 6205, p. 56-61, Oct 2014. ISSN 1095-9203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25278604> >.

FERREIRA, A. S.; CARDOSO, L. P.; STEFANI, M. M. Moderate prevalence of transmitted drug resistance and high HIV-1 genetic diversity in patients from Mato Grosso State, Central Western Brazil. **J Med Virol**, v. 83, n. 8, p. 1301-7, Aug 2011. ISSN 1096-9071. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21678433> >.

FITCH, W. M. **Toward defining the course of evolution: minimum change for a specific tree topology**. *Syst Zool*. 20: 406-416 p. 1971.

FREEMAN, S.; HERRON, J. **Análise Evolutiva**. 4 ed. Porto Alegre: Artmed, 2009.

GADELHA, S. R. et al. Molecular epidemiology of human immunodeficiency virus-1 in the state of Ceará, Northeast, Brazil. **Mem Inst Oswaldo Cruz**, v. 98, n. 4, p. 461-3, Jun 2003. ISSN 0074-0276. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12937754> >.

GILL, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. **Mol Biol Evol**, v. 30, n. 3, p. 713-24, Mar 2013. ISSN 1537-1719. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23180580> >.

GRENFELL, B. T. et al. Unifying the epidemiological and evolutionary dynamics of pathogens. **Science**, v. 303, n. 5656, p. 327-32, Jan 2004. ISSN 1095-9203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14726583> >.

GRIFFITHS, R. C.; TAVARÉ, S. Sampling theory for neutral alleles in a varying environment. **Philos Trans R Soc Lond B Biol Sci**, v. 344, n. 1310, p. 403-10, Jun 1994. ISSN 0962-8436. Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/7800710> >.

GRÄF, T. **Caracterização molecular da epidemia do HIV-1 em cidades do interior de Santa Catarina e Rio Grande do Sul e filogeografia do subtipo C no Brasil.** Tese de Doutorado. Florianópolis, Santa Catarina - Brasil.: Universidade Federal de Santa Catarina: 151 p. 2015.

GUIMARÃES, M. L. et al. High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. **AIDS Res Hum Retroviruses**, v. 18, n. 17, p. 1261-9, Nov 2002. ISSN 0889-2229. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12487814> >.

GUINDON, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. **Syst Biol**, v. 59, n. 3, p. 307-21, May 2010. ISSN 1076-836X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20525638> >.

HEMELAAR, J. et al. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. **AIDS**, v. 20, n. 16, p. W13-23, Oct 2006. ISSN 0269-9370. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17053344> >.

HO, S. Y.; SHAPIRO, B. Skyline-plot methods for estimating demographic history from nucleotide sequences. **Mol Ecol Resour**, v. 11, n. 3, p. 423-34, May 2011. ISSN 1755-0998. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21481200> >.

HOFFMANN, C.; ROCKSTROH, J.; KAMPS, B. **HIV Medicine**. 15. Paris: Flying Publisher, 2007.

HOLMES, E. C. Molecular epidemiology and evolution of emerging infectious diseases. **Br Med Bull**, v. 54, n. 3, p. 533-43, 1998. ISSN 0007-1420. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10326282> >.

JONES, L. R. et al. In-depth analysis of the origins of HIV type 1 subtype C in South America. **AIDS Res Hum Retroviruses**, v. 25, n. 10, p. 951-9, Oct 2009. ISSN 1931-8405. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19842791> >.

LAU, K. A.; WONG, J. J. Current Trends of HIV Recombination Worldwide. **Infect Dis Rep**, v. 5, n. Suppl 1, p. e4, Jun 2013. ISSN 2036-7430. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24470968> >.

LEMEY, P. et al. Bayesian phylogeography finds its roots. **PLoS Comput Biol**, v. 5, n. 9, p. e1000520, Sep 2009. ISSN 1553-7358. Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/19779555> >.

LEMEY, P. et al. Phylogeography takes a relaxed random walk in continuous space and time. **Mol Biol Evol**, v. 27, n. 8, p. 1877-85, Aug 2010. ISSN 1537-1719. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20203288> >.

LEMEY, P.; SALEMI, M.; VANDAMME, A. **The Phylogenetic Handbook: a practical approach to phylogenetic analysis and hypothesis testing**. Cambridge: Cambridge University Press, 2009a.

LIU, S. L. et al. Selection for human immunodeficiency virus type 1 recombinants in a patient with rapid progression to AIDS. **J Virol**, v. 76, n. 21, p. 10674-84, Nov 2002. ISSN 0022-538X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12368309> >.

LOLE, K. S. et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. **J Virol**, v. 73, n. 1, p. 152-60, Jan 1999. ISSN 0022-538X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9847317> >.

MACHADO, L. F. et al. Molecular epidemiology of HIV type 1 in northern Brazil: identification of subtypes C and D and the introduction of CRF02_AG in the Amazon region of Brazil. **AIDS Res Hum Retroviruses**, v. 25, n. 10, p. 961-6, Oct 2009. ISSN 1931-8405. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19795985> >.

MARTIN, D. P. et al. RDP3: a flexible and fast computer program for analyzing recombination. **Bioinformatics**, v. 26, n. 19, p. 2462-3, Oct 2010. ISSN 1367-4811. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20798170> >.

MCCORMACK, G. P.; CLEWLEY, J. P. The application of molecular phylogenetics to the analysis of viral genome diversity and evolution. **Rev Med Virol**, v. 12, n. 4, p. 221-38, 2002 Jul-Aug 2002. ISSN 1052-9276. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12125014> >.

MELLORS, J. W. et al. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. **Ann Intern Med**, v. 122, n. 8, p. 573-9, Apr 1995. ISSN 0003-4819. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/7887550> >.

MILLER, R. H.; SARVER, N. HIV accessory proteins as therapeutic targets. **Nat Med**, v. 3, n. 4, p. 389-94, Apr 1997. ISSN 1078-8956. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9095171> >.

OSMANOV, S. et al. Estimated global distribution and regional spread of HIV-1

genetic subtypes in the year 2000. **J Acquir Immune Defic Syndr**, v. 29, n. 2, p. 184-90, Feb 2002. ISSN 1525-4135. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11832690> >.

PASSAES, C. P. et al. Genetic characterization of HIV-1 BC recombinants and evolutionary history of the CRF31_BC in Southern Brazil. **Infect Genet Evol**, v. 9, n. 4, p. 474-82, Jul 2009. ISSN 1567-7257. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19460312> >.

PEDROSO, C. et al. High prevalence of primary antiretroviral resistance among HIV-1-infected adults and children in Bahia, a northeast state of Brazil. **J Acquir Immune Defic Syndr**, v. 45, n. 2, p. 251-3, Jun 2007. ISSN 1525-4135. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17527096> >.

PEREIRA, G. A. et al. Human immunodeficiency virus type 1 and hepatitis C virus co-infection and viral subtypes at an HIV testing center in Brazil. **J Med Virol**, v. 78, n. 6, p. 719-23, Jun 2006. ISSN 0146-6615. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16628586> >.

PERELSON, A. S. et al. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. **Science**, v. 271, n. 5255, p. 1582-6, Mar 1996. ISSN 0036-8075. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/8599114> >.

PETER, F. HIV nef: the mother of all evil? **Immunity**, v. 9, n. 4, p. 433-7, Oct 1998. ISSN 1074-7613. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9806629> >.

PETERLIN, B. M.; TRONO, D. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. **Nat Rev Immunol**, v. 3, n. 2, p. 97-107, Feb 2003. ISSN 1474-1733. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12563294> >.

PRESTON, B. D.; POIESZ, B. J.; LOEB, L. A. Fidelity of HIV-1 reverse transcriptase. **Science**, v. 242, n. 4882, p. 1168-71, Nov 1988. ISSN 0036-8075. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/2460924> >.

PYBUS, O. G. et al. The epidemic behavior of the hepatitis C virus. **Science**, v. 292, n. 5525, p. 2323-5, Jun 2001. ISSN 0036-8075. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11423661> >.

RAMBAUT, A. et al. The causes and consequences of HIV evolution. **Nat Rev Genet**, v. 5, n. 1, p. 52-61, Jan 2004. ISSN 1471-0056. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14708016> >.

RAMBAUT, A. et al. The genomic and epidemiological dynamics of human influenza A virus. **Nature**, v. 453, n. 7195, p. 615-9, May 2008. ISSN 1476-4687. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18418375> >.

RAMOS, A. et al. Dual and recombinant infections: an integral part of the HIV-1 epidemic in Brazil. **Emerg Infect Dis**, v. 5, n. 1, p. 65-74, 1999 Jan-Feb 1999. ISSN 1080-6040. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10081673> >.

ROBBINS, K. E. et al. U.S. Human immunodeficiency virus type 1 epidemic: date of origin, population history, and characterization of early strains. **J Virol**, v. 77, n. 11, p. 6359-66, Jun 2003. ISSN 0022-538X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12743293> >.

ROBERTSON, D. L. et al. HIV-1 nomenclature proposal. **Science**, v. 288, n. 5463, p. 55-6, Apr 2000. ISSN 0036-8075. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10766634> >.

RODRIGO, A.; FELSENSTEIN, J. **Coalescent approaches to HIV-1 population genetics**. Crandall K. The Evolution of HIV.: Baltimore: Johns Hopkins University Press: 233-272 p. 1999.

SA FILHO, D. J. et al. Analysis of full-length human immunodeficiency virus type 1 genome reveals a variable spectrum of subtypes B and f recombinants in São Paulo, Brazil. **AIDS Res Hum Retroviruses**, v. 21, n. 2, p. 145-51, Feb 2005. ISSN 0889-2229. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15725753> >.

SANTOS, A. F. et al. Epidemiologic and evolutionary trends of HIV-1 CRF31_BC-related strains in southern Brazil. **J Acquir Immune Defic Syndr**, v. 45, n. 3, p. 328-33, Jul 2007. ISSN 1525-4135. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17496564> >.

SANTOS, A. F. et al. Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. **AIDS**, v. 20, n. 16, p. 2011-9, Oct 2006. ISSN 0269-9370. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17053347> >.

SIMON, D. et al. [Prevalence of HIV-1 subtypes in patients of an urban center in Southern Brazil]. **Rev Saude Publica**, v. 44, n. 6, p. 1094-101, Dec 2010. ISSN 1518-8787. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20944890> >.

SMITH, D. M. et al. HIV drug resistance acquired through superinfection. **AIDS**, v. 19, n. 12, p. 1251-6, Aug 2005. ISSN 0269-9370. Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/16052079> >.

SOARES, E. A. et al. Epidemiologic and molecular characterization of human immunodeficiency virus type 1 in southern Brazil. **J Acquir Immune Defic Syndr**, v. 34, n. 5, p. 520-6, Dec 2003. ISSN 1525-4135. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14657764> >.

SOARES, M. A. et al. A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. **AIDS**, v. 17, n. 1, p. 11-21, Jan 2003. ISSN 0269-9370. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12478065> >.

TAVARÉ, S. **Some probabilistic and statistical problems in the analysis of DNA sequences**: Lectures on Mathematics in the Life Sciences. 17: 57-86 p. 1986.

TEBIT, D. M.; ARTS, E. J. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. **Lancet Infect Dis**, v. 11, n. 1, p. 45-56, Jan 2011. ISSN 1474-4457. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21126914> >.

TEIXEIRA, S. L. et al. HIV-1 infection among injection and ex-injection drug users from Rio de Janeiro, Brazil: prevalence, estimated incidence and genetic diversity. **J Clin Virol**, v. 31, n. 3, p. 221-6, Nov 2004. ISSN 1386-6532. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15465416> >.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res**, v. 22, n. 22, p. 4673-80, Nov 1994. ISSN 0305-1048. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/7984417> >.

VUILLEUMIER, S.; BONHOEFFER, S. Contribution of recombination to the evolutionary history of HIV. **Curr Opin HIV AIDS**, v. 10, n. 2, p. 84-9, Mar 2015. ISSN 1746-6318. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25565174> >.

WHELAN, S.; LIÒ, P.; GOLDMAN, N. Molecular phylogenetics: state-of-the-art methods for looking into the past. **Trends Genet**, v. 17, n. 5, p. 262-72, May 2001. ISSN 0168-9525. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11335036> >.

WHO/UNAIDS. **Global report: UNAIDS report on the global AIDS epidemic**: Joint United Nations Programme on HIV/AIDS 2013.

YANG, O. O. et al. Human immunodeficiency virus type 1 clade B superinfection:

evidence for differential immune containment of distinct clade B strains. **J Virol**, v. 79, n. 2, p. 860-8, Jan 2005. ISSN 0022-538X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15613314> >.

YANG, Z. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. **Mol Biol Evol**, v. 10, n. 6, p. 1396-401, Nov 1993. ISSN 0737-4038. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/8277861> >.

ANEXOS

Anexo 1 – Normas de publicação do periódico Memórias do Instituto Oswaldo Cruz.

Orientações para publicação disponível em:

<<http://memorias.ioc.fiocruz.br/instructions-to-authors>>.

The Memórias' content is freely accessible to readers and no publication fees are charged to authors. The Memórias do Instituto Oswaldo Cruz has decided to simplify the requirements regarding the format of submitted manuscripts. From now on, all manuscripts may be submitted in any text format as long as the common subdivision of scientific articles are followed, e.g. introduction, materials and methods, results, discussion and references. For Reviews, Perspectives and similar articles, authors may use the sections that best suit the structure and content of the proposed manuscript. All manuscripts should contain, besides the title and abstract, full details of authors and institutions, acknowledgements of any technical or financial assistance as well as state any conflicts of interest. This flexible text format will be used for the initial analysis and peer review. If the manuscript is accepted, authors will be requested to edit the text in accordance with the publication style of the Memórias."

Upon acceptance, the manuscript should be arranged in the following format:

The manuscript should be prepared using standard word processing software and should be printed (font size 12) double-spaced throughout the text, figure captions, and references, with margins of at least 3 cm. The figures should come in the extension tiff, with a minimum resolution of 300 dpi. Tables and legends to figures must be submitted all together in a single file. Figures, must be uploaded separately as supplementary file.

Running title: with up to 40 characters (letters and spaces)

Title: with up to 250 characters

Author's names: without titles or graduations

Institutional affiliations: full address of the corresponding author only

Summary: up to 200 words (100 words in case of short communications). It should emphasize new and important aspects of the study or observations.

Key words: 3-6 items must be provided. Terms from the Medical Subject Headings (Mesh) list of Index Medicus should be used.

Sponsorships: indicating the sources of financial support and change of address.

Introduction: should set the purpose of the study, give a brief summary (not a review) of previous relevant works, and state what new advance has been made in the investigation. It should not include data or conclusions from the work being reported.

Materials and Methods: should briefly give clear and sufficient information to permit the study to be repeated by others. Standard techniques need only be referenced.

Ethics: when reporting experiments on human subjects, indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 1983. When reporting experiments on animals, indicate whether the institution's or a national research council's guide for, or any national law on the care and use of laboratory animals was followed.

Results: should be a concise account of the new information discovered, with the least personal judgement. Do not repeat in text all the data in the tables and illustrations.

In case of describing New Species, should follow:

Name of the new species, authors (when it is the case), sp. nov., (Figs x-y)
[Ex: *An. (Nyssorhynchus) atacamensis* González and Sallum, sp. nov. (Figs 1-4)]

Previous reference to the new species (when it is the case)
[Ex: *An. pictipennis* of Rueda et al. (2008): 448.]

Diagnosis (or *Description*; all stages are described);

Type host (when it is the case);

Site of Infection (when it is the case);

Type-locality;

Type data and depository;

Other material examined (when it is the case);

Distribution;

Host-parasite data (such prevalence and other important data, when it is the same case);

Bionomics;

Etymology;

Taxonomic discussion (or simply DISCUSSION as internal title).

Discussion: should be limited to the significance of the new information and relate the new findings to existing knowledge. Only unavoidable citations should be included.

Acknowledgements: should be short and concise, and restricted to those absolutely necessary.

References

Must be accurate. Only citations that appear in the text should be referenced. Unpublished papers, unless accepted for publication, should not be cited. Work accepted for publication should be referred to as "in press" and a letter of acceptance of the journal must be provided. Unpublished data should only be cited in the text as "unpublished observations", and a letter of permission from the author must be provided. The references at the end of the paper should be arranged in alphabetic order according to the surname of the first author.

REFERENCES MUST BE CREATED IN WORD PROCESS

To assure your references is published as requested, please send archives without Mendeley Hyperlink or similar. References created with edition marks will cause improper corrections during editing process, causing delays.

THE TITLE OF JOURNALS

Should be abbreviated according to the style used in the Index Medicus.

Consult: <http://www2.bg.am.poznan.pl/czasopisma/medicus.php?lang=eng>

In the text use authors' surname and date: Lutz (1910) or (Lutz 1910)

With two authors it is: (Lutz & Neiva 1912) or Lutz and Neiva (1912)

When there are more than two authors, only the first is mentioned: Lutz et al. (1910) or (Lutz et al. 1910).

AT THE END OF THE PAPER USE THE FOLLOWING STYLES

Journal article

Chagas C, Villela E 1922. Forma cardiaca da tripanosomiase americana. Mem Inst Oswaldo Cruz 14: 15-61.

Book and Thesis

Forattini OP 1973. Entomologia Médica. Psychodidae, Phlebotominae, Leishmaniose, Bartonelose, Vol. IV, Edgard Blucher, São Paulo, 658 pp.

Morel CM 1983. Genes and Antigens of Parasites. A Laboratory Manual, 2nd ed., Fundação Oswaldo Cruz, Rio de Janeiro, xxii + 580 pp.

Mello-Silva CC 2005. Controle alternativo e alterações fisiológicas em *Biomphalaria glabrata* (Say, 1818), hospedeiro intermediário de *Schistosoma mansoni* Sambom, 1907 pela ação do látex de *Euphorbia splendens* var. *hislopii* N.E.B (Euphorbiaceae), PhD Thesis, Universidade Federal Rural do Rio de Janeiro, Seropédica, 85 pp.

Chapter in book

Cruz OG 1911. The prophylaxis of malaria in central and southern Brasil. In R Ross, The Prevention of Malaria, John Murray, London, p. 390-398.

Journal article on the Internet

Abood S. Quality improvement initiative in nursing homes: the ANA acts in an advisory role. Am J Nurs [serial on the Internet]. 2002 Jun [cited 2002 Aug 12];102(6):[about 3 p.]. Available from:

<http://www.nursingworld.org/AJN/2002/june/Wawatch.htm>

Monograph on the Internet

Foley KM, Gelband H, editors. Improving palliative care for cancer [monograph on the Internet]. Washington: National Academy Press; 2001 [cited 2002 Jul 9]. Available from: <http://www.nap.edu/books/0309074029/html/>.

Homepage/Web site

Cancer-Pain.org [homepage on the Internet]. New York: Association of Cancer Online Resources, Inc.; c2000-01 [updated 2002 May 16; cited 2002 Jul 9]. Available from: <http://www.cancer-pain.org/>.

Part of a homepage/Web site

American Medical Association [homepage on the Internet]. Chicago: The Association; c1995-2002 [updated 2001 Aug 23; cited 2002 Aug 12]. AMA Office of Group Practice Liaison; [about 2 screens]. Available from: <http://www.ama-assn.org/ama/pub/category/1736.html>

DATABASE ON THE INTERNET

Open database

Who's Certified [database on the Internet]. Evanston (IL): The American Board of Medical Specialists. c2000 - [cited 2001 Mar 8]. Available from: <http://www.abms.org/newsearch.asp>

Closed database

Jablonski S. Online Multiple Congenital Anomaly/Mental Retardation (MCA/MR) Syndromes [database on the Internet]. Bethesda (MD): National Library of Medicine (US). c1999 [updated 2001 Nov 20; cited 2002 Aug 12]. Available from: http://www.nlm.nih.gov/mesh/jablonski/syndrome_title.html

Part of a database on the Internet

MeSH Browser [database on the Internet]. Bethesda (MD): National Library of Medicine (US); 2002 - [cited 2003 Jun 10]. Meta-analysis; unique ID: D015201; [about 3 p.]. Available from:

<http://www.nlm.nih.gov/mesh/MBrowser.html> Files updated weekly. Updated June 15, 2005

Figures

FIGURES AND TABLES MUST BE UNDERSTANDABLE WITHOUT REFERENCE TO THE TEXT

Figures: presented in tiff format with a minimum of 300 dpi and photographs must be sharply focused, well contrasted, and if mounted onto a plate, the figures should be numbered consecutively with Arabic numbers. Magnification must be indicated by a line or bar in the figure, and referenced, if necessary in the caption (e.g., bar = 1 mm). Plates and line figures should either fit one column (8 cm) or the full width (16.5 cm) of the page and should be shorter than the page length to allow inclusion of the legend. Letters and numbers on figures should be of a legible size upon reduction or printing. A colour photograph illustrates the cover of each issue of the Journal and authors are invited to submit illustrations with legends from their manuscript for consideration for the cover.

Tables: should supplement, not duplicate, the text and should be numbered with Roman numerals. A short descriptive title should appear above each table, with any explanations or footnotes (identified with a, b, c, etc.) below.

Technical Notes: Technical Notes should communicate rapidly single novel techniques or original technical advances. The entire note should occupy no more than three printed pages including figures and/or tables (it means around 10 double-spaced typed Word file maximum). The text must not be not divided into sections. Therefore, the state of art must be very briefly presented; results must be rapidly presented and discussed at a time. Complementary tables and figures may be

published as supplementary data. References must be limited to few essential ones and cited at the end of the note, using the same format as in full papers. A brief summary and three key words must be provided.

Short communications: should communicate rapidly single results or techniques. They should occupy no more than three printed pages including figures and/or tables. They should not contain excessive references. References should be cited at the end of the paper using the same format as in full papers. A brief summary and three key words must be provided.

Genome Announcement and Highlights: this section is dedicated to publish new genome information from eukaryote parasites, virus, bacteria and their respective vectors. Authors who wants a fast peer review and publication cycle for their research results covering new genome sequences, re-sequencing and comparative genome analysis as well as the expression pattern of genomes are invited to submitted papers under the short communication format.

Alternative format: manuscripts may be submitted following the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" produced by the International Committee of Medical Journal Editors also known as the Vancouver Style. In this case, authors should follow the guidelines in the fifth edition (Annals of Internal Medicine 1997; 126: 36-47, or at the website <http://www.acponline.org/journals/resource/unifreqr/htm>) and will be responsible for modifying the manuscript where it differs from the instructions given here, if the manuscript is accepted for publication. Authors should also follow the Uniform Requirements for any guidelines that are omitted in these Instructions.

In case of clinical trials it's mandatory to inform the registration number of the REBEC platform.

A statement that the data/results of the manuscript are not plagiarism and have not been published elsewhere.

ONCE A PAPER IS ACCEPTED FOR PUBLICATION, THE AUTHORS MUST PROVIDE:

Page charges: there will be no page charges.

Proofs: one set of page proofs will be supplied for the author to check for typesetting accuracy, to be returned by the stipulated date. No changes to the original manuscript will be allowed at this stage.