UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

"HIV infection in Angola: Molecular epidemiology, diagnosis and antibody neutralization"

Francisco Neves Dos Santos Martin

Orientadores: Prof. Doutor Nuno Eduardo Moura dos Santos da Costa Taveira

Prof. Doutor João Manuel Braz Gonçalves

Tese especialmente elaborada para a obtenção do grau de Doutor em Farmácia, especialidade Microbiologia.

2022

UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

“HIV infection in Angola: Molecular epidemiology, diagnosis and antibody neutralization”

Francisco Neves Dos Santos Martin

Orientadores: Prof. Doutor Nuno Eduardo Moura dos Santos da Costa Taveira

Prof. Doutor João Manuel Braz Gonçalves

Tese especialmente elaborada para obtenção do grau de Doutor em Farmácia (Microbiologia)

Júri:

Presidente: Doutora Maria Alexandra de Oliveira Silva Braga Pedreira de Brito, Professora Associada com Agregação e membro do Conselho Científico da Faculdade de Farmácia da Universidade de Lisboa
Vogais: Doutora Ana Barroso Abecasis, Professora Auxiliar, Instituto de Higiene e Medicina Tropical da Universidade Nova de Lisboa
Doutor Nuno Eduardo Moura dos Santos da Costa Taveira, Professor Catedrático, Instituto Universitário Egas Moniz, orientador
Doutora Perpétua da Conceição Rodrigues Gomes Cavaco Silva, Professora Associada, Instituto Universitário Egas Moniz
Doutor Fernando Manuel Tavares Maltez, Professor Auxiliar Convidado, Faculdade de Medicina da Universidade de Lisboa
Doutor José Miguel Azevedo Pereira, Professor Auxiliar com Agregação, Faculdade de Farmácia da Universidade de Lisboa

Francisco Martin teve o apoio financeiro da Fundação para a Ciência e Tecnologia através de uma bolsa de doutoramento (SFRH/BD/87488/2012).

2022

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Farmácia de Lisboa pelos conteúdos nele apresentados.

Francisco Martin teve o apoio financeiro da Fundação para a Ciência e Tecnologia através de uma bolsa de doutoramento (SFRH/BD/87488/2012).

“Faz-te um Homem”

-Amaro Jesus Neves

Acknowledgements

Foi uma longa, feliz e fascinante etapa que irei recordar sempre com bastante carinho e que me permitiu abrir horizontes e ser o ser humano que sou hoje.

Em primeiro lugar, gostaria de agradecer ao Professor Doutor Nuno Taveira, na qualidade de orientador desta tese, por me ter envolvido e dado a conhecer desde cedo a área da investigação. Posso afirmar que a investigação científica supera tudo o que fiz até hoje, em termos de conhecimento, superação, persistência, cooperação, envolvência e principalmente resiliência. É fascinante! "Ficou o bichinho".

Obrigado pela oportunidade de participar em diferentes projetos em que estivemos juntos e por toda a partilha que levou ao crescimento do meu conhecimento científico. Fui muito feliz nesta equipa.

Obrigado ainda por toda a compreensão, supervisão, orientação e apoio nesta reta final.

Em segundo lugar gostaria de agradecer ao Professor Doutor José Moniz Pereira, como ex-diretor do Departamento de Microbiologia da Faculdade de Farmácia, por ter permitido desenvolver o meu trabalho nesta unidade e também por ter iniciado a minha coorientação.

O meu muito obrigado ao Professor Doutor João Gonçalves por ter abraçado a coorientação do meu projeto de doutoramento já na fase final da redação da dissertação.

Obrigado também a todos os docentes que trabalham no departamento, em particular à Prof. Quirina, ao Prof. José Miguel, à Prof. Isabel Portugal e ao Prof. Jorge Vítor.

Um especial agradecimento à Dr.ª Sofia Clemente e restantes colaboradores do Hospital da Divina Providência em Luanda, Angola. Sem a vossa colaboração este trabalho não teria sido possível. Muito obrigado!

Aos doentes Angolanos infetados por VIH que participaram nos estudos desenvolvidos no decurso deste projeto, o meu muito obrigado!

À Fundação para a Ciência e Tecnologia o meu obrigado pelo apoio financeiro que permitiu desenvolver este trabalho.

Obrigado a todos os meus colegas de laboratório, Joana Duarte, Andreia Martins, Cláudia Paladino, Cheila Rocha, Inês Bártolo, Pedro Borrego, Carla Silva, João Perdigão, Marta Calado, Marta Gíria, Joana Vital, Ana Rita, Inês Figueiredo, Inês Moranguinho, Rita Calado, Rita Mateus, José Marcelino, Rute Marcelino e Jaciara pela amizade, conselhos, esclarecimentos, apoio e momentos tão bem passados que deixam saudades.

À minha família, um muito obrigado por todo o apoio, compreensão e por sempre me motivarem a fazer o melhor, e ser melhor.

Em especial aos meus pais, ao meus sogros, tios, primos, irmão, avó, à minha esposa e aos meus filhos por nunca desacreditarem em mim e sentirem-se sempre orgulhosos do meu esforço e dedicação.

A estabilidade, educação e formação que me têm dado e continuam a dar torna tudo mais fácil. Vocês conhecem-me como ninguém e sabem o que penso em relação a este assunto de querer estar sempre a melhorar e de ser um eterno insatisfeito. Contudo, deixo aqui por escrito, eu não vejo o meu sucesso como algo extraordinário que deva ser celebrado, mas sim como uma obrigação por tudo o que fizeram e continuam a fazer por mim.

PREFACE

The research described in this work was performed under the supervision of Prof. Doutor Nuno Taveira and the cosupervision of Prof. Doutor João Gonçalves. All experimental work was performed at HIV Evolution, Epidemiology and Prevention laboratory, Research institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa.

The results obtained in this thesis were described in the following publications:

- Bártolo I, Zakovic S, **Martin F**, Palladino C, Carvalho P, Camacho R, Thamm S, Clemente S, Taveira N. HIV-1 diversity, transmission dynamics and primary drug resistance in Angola. PLoS ONE. 2014; 9(12): e113626

- **Martin F**, Palladino C, Mateus R, Bolzan A, Clemente S, Gomes P, Taveira N. Early infant diagnosis of HIV-1 infection in Luanda, Angola, using a new DNA PCR assay and dried blood spots. PLoS ONE. 2017;12(7): e0181352.

- **Martin F,** Marcelino J, Palladino C, Bártolo I, Tracana S, Moranguinho I, Gonçalves P, Mateus R, Calado R, Borrego P, Leitner T, Clemente S, Taveira N. Long-term and low-level envelope C2V3 stimulation from highly diverse virus isolates leads to frequent development of broad and elite antibody neutralization in HIV-1 infected individuals. Available at medRxiv https://doi.org/10.1101/2022.01.27.22269759 (Manuscript submitted to eBioMedicine)

Peer-reviewed abstracts published in international journals:

Martin F, Palladino C, Mateus R, Clemente S, Gomes P, Taveira N. Evaluation of an in-house molecular HIV-1 Test to assess Mother-to-Child HIV-1 transmission in angola (the APEHC cohort) JAIDS Journal of Acquired Immune Deficiency Syndromes. 2016; (71): Abstract_P83

Martin F, Palladino C, Mateus R, Diniz AR, Calado R, Figueiredo I, Bártolo I, Borrego P, Clemente S, Taveira. Broad and potent neutralizing antibody responses in HIV-1 infected Angolan patients: implications for vaccine design and efficacy. HIV Medicine, 20 (Suppl. 9), 3–316. Abstract_PE38/9

Oral communications:

Martin F, Palladino C, Mateus R, Diniz A.R, Calado R, Figueiredo I, Bártolo I, Borrego P, Clemente S, Taveira N. Broad And Potent Neutralizing Antibody Responses In HIV-1 Infected Angolan Patients: Implications For Vaccine Design And Efficacy. 1st Annual Meeting in Pharmaceutical Sciences, Coimbra, Portugal, November, 2019. (Invited communication)

Martin F, Palladino C, Mateus R, Diniz A.R, Calado R, Figueiredo I, Bártolo I, Borrego P, Clemente S, Taveira N. Broad And Potent Neutralizing Antibody Responses In HIV-1 Infected Angolan Patients: Implications For Vaccine Design And Efficacy. 11th iMed UL. Postgraduate Students Meeting, Lisbon, Portugal, 15th of July, 2019. (awarded best oral communication)

Martin F, Palladino C, Mateus R, Diniz AR, Calado R, Clemente S, Taveira N. Characterization of the neutralizing antibody responses in HIV-1 infected patients from Angola and their impact on disease progression. 4o Congresso Nacional de Medicina Tropical. Instituto de Higiene e Medicina Tropical. 1o encontro lusófono de SIDA, Tuberculose e Doenças oportunistas. Instituto de Higiene e Medicina Tropical. 19-21 April 2017.

Poster Communications:

Martin F, Palladino C, Mateus R, Diniz AR, Calado R, Clemente S, Taveira N. Neutralizing antibody response in HIV-1 infected patients from Angola. Keystone Symposia, Molecular and Celular Biology, March, 2018.

Martin F, Palladino C., Mateus R, Diniz AR, Calado R, Clemente S, Taveira N. Neutralizing antibody response in HIV-1 infected patients from Angola. 9th iMed.ULisboa Postgraduate Students Meeting and 2nd i3DU Meeting. Poster number 92. 13-14 July 2017. Faculty of pharmacy. Lisbon.

Martin F, Palladino C, Mateus R, Diniz AR, Calado R, Clemente S, Taveira N. Neutralizing antibody response in HIV-1 infected patients from Angola. Ciência 2017. 3-5 July 2017. Centro de Congressos de Lisboa.

Martin F, Palladino C, Mateus R, Clemente S, Gomes P, Taveira N. Evaluation of an In-house Molecular HIV-1 Test to Assess Mother-to-Child HIV-1 Transmission in Angola (the APEHC Cohort). 17th Annual International Meeting of the Institute of Human Virology, Baltimore, USA, 26th September 2015.

Martin F, Mateus R, Calado R, Diniz AR, Palladino C, Clemente S, Taveira N. Characterization of the neutralizing antibody responses in HIV-1 infected patients in Angola. 7th iMed.UL Postgraduate Students Meeting, Lisbon, Portugal, 16th of July, 2015.

Martin F, Palladino C, Taveira N. Development of a new in-house real-time DNA PCR assay for the early diagnosis of perinatal HIV-1 infection in dried blood spots. 6th iMed.UL Postgraduate Students Meeting, Lisbon, Portugal, 2nd July, 2014.

Martin F, Bártolo I, Carvalho P, Camacho R, Clemente S, Taveira N. HIV-1 genetic diversity and transmitted drug resistance for the recent introduction of CRF01_AE. 18th International Bioinformatics workshop on virus evolution and molecular epidemiology, Gainesville, Florida, USA, 25th August 2013.

Martin F, Palladino C, Taveira N. Development of a new in-house real-time DNA PCR assay for the early diagnosis of perinatal HIV-1 infection in dried blood spots. 5th iMed.UL Postgraduate Students Meeting, Lisbon, Portugal, 2nd June, 2013.

Other publications

The research described below was produced during the PhD, however is not reflected in the present thesis. The studies were presented in the form of research papers, scientific book chapters, oral and poster communications in international and national meetings. The following publications are listed by order of publication date, the latest being the first listed.

Research papers:

Marcelino R, Gramacho F, **Martin F**, Brogueira P, Janeiro N, Afonso C, Badura R, Valadas E, Mansinho K, Caldeira L, Taveira N, Marcelino J. Antibody response against selected epitopes in the HIV-1 envelope gp41 ectodomain contributes to reduce viral burden in HIV-1 infected patients. Scientific Reports, 2021, 11:8993.

Calado R, Duarte J, Borrego P, Marcelino JM, Bártolo I, **Martin F**, Figueiredo I, Almeida S, Graça L, Vítor J, Aires da Silva F, Dias I, Carrapiço B, Taveira N. A Prime-Boost Immunization Strategy with Vaccinia Virus Expressing Novel gp120 Envelope Glycoprotein from a CRF02_AG Isolate Elicits Cross-Clade Tier 2 HIV-1 Neutralizing Antibodies. Vaccines (Basel). 2020 Apr 7;8(2):171.

Cunha-Santos C, Ricardo P, Perdigao L, **Martin F**, Gomes Oliveira J, Cardoso M, Manuel A, Taveira N, Goncalves J. Inhibition of HIV replication through siRNA carried by CXCR4-targeted chimeric nanobody. Cellular and Molecular Life Sciences. 2019.

Ezeonwumelu I, Bártolo, **Martin F**, Abecasis A, Campos T, Romero-Severson E O, Leitner T, Taveira N. Accidental Father-to-Son HIV-1 Transmission During the Seroconversion Period. Aids Research and Human Retroviruses, 2018, 34 (10): 857 - 862.

Book Chapters:

Taveira N, Moranguinho I, **Martin F. O agente.** Manual sobre SIDA – 6ª Edição. Editor Francisco Antunes e Fernando Maltez. Permanyer Portugal, 2021.

Taveira N and **Martin F**. **O agente.** Manual sobre SIDA – 5ª Edição. Editor Francisco Antunes, Fernando Maltez. Permanyer Portugal, 2018.

Peer-reviewed abstracts published in international journals:

Martin F, Martins A, Maia F, Rocha C, Borrego P, Antunes F, Caldeira L, Valadas E, Taveira N. Resistance mutations to protease inhibitors in proviral DNA of HIV-2 infected patients predict response to treatment. Abstract Supplement HIV Glasgow 2018. Journal of the International AIDS Society 2018, 21(S8):e25187. Abstract_P305.

Martins A, Maia F, **Martin F**, Rocha C, Valadas E, Antunes F, Borrego P, Taveira N. Protease diversity and resistance to protease inhibitors of viruses archived in peripheral blood mononuclear cells of HIV-2 infected patients. Reviews in Antiviral Theraphy & Infectious Diseases. 2014;(2): Abstract_P35

Gonçalves P, **Martin F**, Borges P, Espirito Santo M, Taveira N, Marcelino J. Neutralising and non-neutralising antibodies response in HIV-1-infected individuals from Mozambique. BMJ global health. 2019 EDC.161: Abstract_A61.

Oral communications:

Taveira N, Diniz AR, Borrego P, **Martin F**, Gomes P, Gonçalves F, Caixas U, Vaz Pinto I, Barahona, I, Bártolo I. Dolutegravir has a highly potent activity against most primary HIV-2 isolates includ- ing those that have acquired resistance to Raltegravir. – June 1-5, 2017 – ASM Microbe – New Orleans, USA.

Diniz AR, Borrego P, **Martin F**, Gomes P, Gonçalves F, Caixas U, Vaz Pinto I, Barahona, I, Bártolo I, Taveira N. Dolutegravir has a highly potent activity against most primary HIV-2 isolates includ- ing those that are resistant to raltegravir. – 3-9 May, 2017 – rede SAÚDE 4th Annual Conference, ULisboa Innovation Week – Lisbon, Portugal.

Calado R, Duarte J, Borrego P, Marcelino JM, Wilton J, Bártolo I, **Martin F**, Barroso H, Almeida SCP, Graça L, Taveira N. Novel HIV-1 gp120 glycoproteins derived from non-B subtypes induces tier 2 cross-clade neutralizing antibodies in mice. 4º Congresso Nacional de Medicina Tropical. 1o encontro lusófono de SIDA, Tuberculose e Doenças oportunistas. Instituto de Higiene e Medicina Tropical. 19-21 April 2017.

Poster Communications:

Calado R, Duarte J, Borrego P, Marcelino JM, Wilton J, Bártolo I, **Martin F**, Almeida SCP4, Barroso H, Graça L, Taveira N. A prime-boost immunization strategy with Vaccinia virus expressing novel HIV-1 gp120 glycoproteins induces tier 2 neutralizing antibodies in mice. 9th iMed.ULisboa Postgraduate Students Meeting and 2nd i3DU Meeting.13-14 July 2017. Faculty of pharmacy. Lisbon.

Calado R, Duarte J, Borrego P, Marcelino JM, Wilton J, Bártolo I, **Martin F**, Almeida SCP4, Barroso H, Graça L, Taveira N. Induction of tier 2 cross-clade neutralizing antibodies in mice by a prime-boost immunization strategy with Vaccinia virus expressing novel HIV-1 gp120 glycoproteins from non-B subtypes. Ciência 2017. 3-5 July 2017. Centro de Congressos de Lisboa.

Figueiredo IB, Calado R, **Martin F**, Borrego P, Cardoso F, Taveira N, Barroso H. Antibody response to an HIV-1/HIV-2 chimeric envelope glycoprotein in mice. 3a Semana da inovação. 3-9 May 2017. Reitoria da Universidade de Lisboa (ULisboa). Poster presentation.

Figueiredo IB, Calado R, **Martin F**, Borrego P, Cardoso F, Taveira N, Barroso H. Antibody response to an HIV-1/HIV-2 chimeric envelope glycoprotein in mice. Keystone Symposia, HIV Vaccines, Keystone, March, 2017.

Calado R, Duarte J, Borrego P, Marcelino JM, Wilton J, Bártolo I, **Martin F**, Clemente S, Almeida SCP, Barroso H, Graça L, Taveira N. A novel prime-boost immunization strategy with Vaccinia virus expressing HIV-1 gp120 derived from non-B subtypes induces autologous and heterologous Tier 2 HIV-1 cross-clade neutralizing antibodies in mice. Keystone Symposia, HIV Vaccines, Keystone, March, 2017.

Martins A, Maia F, **Martin F**, Rocha C, Valadas E, Antunes F, Borrego P, Taveira N. Protease diversity and resistance to protease inhibitors of viruses archived in peripheral blood mononuclear cells of HIV-2 infected patients. 12th European Workshop on HIV &

Hepatitis treatment strategies & antiviral drug resistance 26 – 28 March 2014, Barcelona, Spain.

Resumo

As metas definidas pela UNAIDS para o controlo da infeção por HIV até 2030, definem objectivos de reduzir 75% dos novos casos e mortes por HIV até 2020 e 90% até 2030, em comparação com 2010 [1]. Contudo, estima-se que nenhum país da África Subsariana atingiu a meta dos 75% de redução dos novos casos de infeção [2]. Inclusive, em Angola entre 2010 e 2018 a incidência da infeção por HIV-1 aumentou 61,2%. Contudo, as mais recentes estimativas sugerem que a prevalência parece estar agora estabilizada com 340.000 pessoas a viverem com HIV [2, 3]. As mulheres e crianças estão entre as populações mais afetadas [2, 3]. A terapêutica antirretrovírica está disponível em Angola desde 2000 para os doentes infetados por HIV-1, contudo só a partir de 2004 passou a ser dispensada de forma gratuita. No final de 2020 o número de pessoas a viver com HIV em tratamento eram 111.168, aproximadamente 33% das pessoas a necessitar de tratamento [3].

Os objetivos desta tese foram: 1) obter um melhor conhecimento da diversidade da infeção por HIV-1, da dinâmica de transmissão e resistência aos antirretrovíricos em Angola; 2) melhorar o diagnóstico da infeção por transmissão vertical em Angola; e 3) caracterizar a resposta neutralizante contra o vírus, e determinar os factores víricos e do hospedeiro a ela associada. A epidemia da infeção por HIV é peculiar em Angola, uma vez que circulam na população formas altamente divergentes do vírus. Dados já publicados sobre a epidemiologia molecular da infeção por HIV-1 em Angola demonstram que circulam na população praticamente todos os subtipos puros bem como vários subsubtipos, várias formas recombinantes circulantes, bem como formas recombinantes únicas [4, 5]. Esta elevada diversidade genética dos vírus circulantes espelha uma epidemia antiga, que data da primeira metade do século 20 [6].

Neste contexto, o primeiro objetivo específico desta tese foi caracterizar a diversidade genética do HIV-1 em Angola, e a dinâmica de transmissão dos subtipos e formas recombinantes do HIV-1 presentes no país, bem como conhecer a prevalência de transmissão de mutações de resistência aos antirretrovíricos, 5 anos após a massificação da distribuição da terapêutica antirretrovírica (capítulo 2). Para tal, sequenciou-se e analisou-se filogeneticamente sequências víricas, de 139 amostras de plasma colhidas em 2009, de doentes Angolanos HIV-1 positivos sem experiência prévia de tratamento, maioritariamente residentes em Luanda. Por forma a determinar as tendências evolutivas comparam-se estas sequências de 2009 do gene *pol* com sequências de 2001 anteriormente analisadas[7].

Verificou-se uma diminuição significativa da prevalência do subtipo A de 2001 para 2009 (40,0% para 10,8%, $p = 0,0019$), enquanto a prevalência de formas recombinantes únicas aumentou 2 vezes (40,0% para 83,1%, $p < 0,0001$). Em 2009, 47,1% dos vírus eram subtipos puros (foram identificados todos os subtipos exceto o B), 47,1% eram vírus recombinantes e 5,8% eram não tipáveis. Os recombinantes únicos mais frequentemente identificados (U/H) formaram um grupo monofilético altamente suportado, sugerindo uma origem local e comum destas formas recombinantes. No que diz respeito à transmissão de mutações de resistência aos antirretrovíricos, neste estudo em 2009, verificámos uma prevalência muito baixa (0,7%), somente num doente foi identificada a K103N o que compara com 1,6% de prevalência observada em 2001 [7]. O elevado número de recentes e pequenos grupos de transmissão que identificámos neste estudo, bem como a identificação de novas formas recombinantes únicas, são consistentes com uma epidemia de HIV-1 em crescimento, impulsionada principalmente pela transmissão heterossexual.

Uma epidemia de HIV em crescimento com elevada diversidade genética coloca desafios a diferentes níveis, principalmente no diagnóstico, tratamento e prevenção da infeção. Este facto aliado à elevada taxa de transmissão vertical registada em Angola, 25,0% em 2014 [8], estiveram na base do estudo apresentado no capítulo 3, que teve como principal objetivo o desenvolvimento e validação de um teste molecular qualitativo, sensível e barato de deteção de DNA pró-viral de HIV-1, em amostras de sangue total periférico colhidas em papel de filtro, por forma a permitir o diagnóstico da infeção perinatal em recém-nascidos expostos ao HIV-1. O gene da integrase foi usado como alvo dado ser um dos genes mais conservados do genoma do vírus. O limite de deteção determinado por regressão *Probit* foi avaliado usando diluições limite de plasmídeos recombinantes contendo o gene da integrase de todos os subtipos de HIV-1 relevantes, a forma recombinante circulante CRF02_AG e células ACH-2, que têm a particularidade de conter uma única cópia de HIV-1 integrada no genoma. O teste teve a capacidade de detetar todos os subtipos de HIV-1 com um limite de deteção de 14 cópias. A sensibilidade e especificidade clínicas foram avaliadas em 100 amostras de adultos infetados por HIV-1, 5 amostras de crianças também infetadas, 50 amostras de voluntários saudáveis e 139 amostras de recém-nascidos de mães Angolanas infetadas por HIV-1, com dados serológicos aos 18 meses de vida, que foram usados como referência de infeção. O teste teve também a capacidade de detetar a infeção por HIV-1, 4 semanas após o nascimento. A sensibilidade e especificidade clínicas do método foram de 100,0% avaliadas nas 139 amostras dos recém-nascidos expostos ao HIV-1. A taxa de transmissão vertical registada do nosso coorte foi de 2,2% entre Janeiro de 2012 e Outubro de 2014, que contrasta com a taxa registada a nível nacional em período semelhante de 25,0%, e espelha os excelentes cuidados materno-infantis prestados no Hospital da Divina Providência em Luanda, Angola. O baixo custo e a simplicidade do

teste tornam-no adequado para a implementação em Angola e outros países com recursos limitados.

O quarto capítulo desta tese teve como principal objetivo a caracterização da resposta neutralizante contra o HIV-1 em 322 doentes Angolanos infetados e perceber quais os principais determinantes virológicos e do hospedeiro dessa resposta imunológica. A indução de anticorpos neutralizantes de largo espectro contra o HIV é considerada de extrema importância para o controlo da infeção, contudo, até à data nenhuma vacina testada demonstrou eficácia em induzir este tipo de anticorpos [9-20]. Para caracterizar a resposta neutralizante usámos um painel de 12 pseudovirus de Env de difícil neutralização (*tier* 2) representativos das estirpes que circulam a nível mundial e usámos amostras de plasma de doentes Angolanos infetados por HIV-1 colhidas em 2009 (n=178) e 2014 (n=58). Amplificámos e determinámos por análise filogenética o subtipo e tropismo do vírus no gene *env*, especificamente na região C2V3C3 nas amostras colhidas em 2009 (n= 110 sequências) e comparámos com resultados anteriores determinados em amostras colhidas em 2001 (n= 96 sequências). Nas amostras de 2014 não foi possível a amplificação dado que a larga maioria dos doentes estava sob terapêutica antirretrovírica. Para além da caracterização da resposta neutralizante e tipagem das amostras, determinámos os possíveis epitopos neutralizantes por comparação com anticorpos neutralizantes monoclonais de largo espectro anteriormente caracterizados. O título de ligação contra polipéptidos recombinantes, compreendendo as regiões C2, V3, C3 de diferentes subtipos de HIV-1, numa subpopulação de doentes com resposta neutralizante conhecida (n= 48 em 2009 e n= 16 em 2014) foi também avaliado. De seguida explorámos possíveis associações e correlações entre a resposta neutralizante, o subtipo do vírus, o título de ligação e as características demográficas e da infeção, da população Angolana infetada por HIV-1. A análise filogenética permitiu identificar diversos subtipos de HIV-

1 incluindo o A1, A2, B, C, D, F1, G, H, J, estirpes não tipáveis e formas recombinantes circulantes. O subtipo A, foi o subtipo puro predominante em ambos os anos de colheita, 2001 e 2009, contudo, o subtipo C aumentou significativamente (2,2 vezes, $p = 0,0095$) em 2009. Dos 176 isolados em que estavam disponíveis ambas as sequências da protease e da região C2V3C3, 74 (42,0%) eram não-recombinantes e 102 (58,0%) eram recombinantes, sendo que os recombinantes prevaleceram relativamente aos subtipos puros em ambos os anos. Notavelmente, considerando as 236 amostras em que caracterizámos o perfil de neutralização, aproximadamente 56,0% dos doentes Angolanos exibiram respostas neutralizantes de largo espectro. A frequência de doentes com capacidade de neutralização elite foi elevadada em 2014, quando os doentes estavam sob terapêutica antirretrovírica e com virémias de baixo nível comparativamente com 2009, onde a larga maioria dos doentes não tinha experiência prévia de terapêutica antirretrovírica. Esta resposta neutralizante de largo espectro foi associada ao subtipo C, idade mais avançada e contagens de linfócitos T CD4+ mais baixas. Verificámos também, uma forte associação entre a resposta neutralizante e o título de ligação contra os polipéptidos recombinantes C2V3C3 dos diferentes subtipos. O título de ligação de anticorpos contra a região C2V3C3 no presente estudo foi um bom indicador do espectro e potência da neutralização. A resposta neutralizante teve como alvo na maioria dos casos o super epitopo baseado em N-glicanos na V3, mas anticorpos específicos para o vértice da V2, o local de ligação do CD4, a região proximal da membrana na gp41 e contra epitopos desconhecidos, também foram identificados nalguns doentes. As regiões V3 e C3 foram significativamente menos variáveis e menos sujeitas a seleção positiva nos doentes com respostas neutralizantes elite em comparação com os doentes com respostas neutralizantes mais fracas ou ausentes, o que sugere um papel ativo dos anticorpos neutralizantes de largo espectro, que têm como alvo estas regiões, no controlo da

replicação e diversificação do vírus. Concluindo, o desenvolvimento de anticorpos de largo espectro de neutralização contra o HIV-1 requer estimulação de longo prazo e de baixo nível da região V3C3 do invólucro por parte de isolados de subtipo C altamente diversos. Estes resultados têm implicações diretas para o desenho de uma nova geração de vacinas contra o HIV-1.

Palavras-chave: diversidade genética do HIV, epidemia de HIV em Angola, resposta neutralizante, diagnóstico da infeção perinatal por HIV.

Abstract

The UNAIDS fast-track goals established the need to reduce 75% of new HIV cases and deaths by 2020 and 90% by 2030, compared to 2010 [1]. However, it is estimated that no country in sub-Saharan Africa has reached the target of 75% reduction in new cases of HIV infection [2]. HIV-1 incidence in Angola increased 61.2% from 2010 to 2018 and the prevalence is now stable with 340,000 people living with HIV [2, 3]. Women and children are the most affected populations [3]. The aims of this thesis were: 1) to get a better understanding of the HIV-1 diversity, transmission dynamics and transmitted drug resistance (TDR) in Angola, 2) improve the early infant diagnosis (EID) in Angola and 3) characterize the neutralizing antibody responses against HIV-1 and assess the possible viral and host factors associated with it.

The HIV epidemics in Angola is peculiar, since highly divergent forms of the virus circulates in the population. In this context our first objective was to assess HIV-1 diversity, transmission dynamics and prevalence of transmitted drug resistance (TDR) in Angola in 2009, in 139 drug naïve HIV-1 infected individuals and compare the results before ART scale-up (chapter 2). We saw an increase in genetic diversity between 2001 and 2009, with the prevalence of subtype A decreasing significantly while the prevalence of unique recombinant forms (URFs) increased 2-fold. Also, local U/H recombinants were newly identified. TDR mutation K103N was found in one (0.7%) patient. Overall, transmission of drug resistant strains was still negligible in Luanda in 2009 and the emergence of new URFs are consistent with a rising HIV-1 epidemics. Our second objective was to develop and validate a sensitive, simple and cheap qualitative proviral DNA PCR-based assay for early infant diagnosis (EID) in HIV-1-exposed infants (n=139) using dried blood spots (DBS) (chapter 3). We were able to successfully validate the assay, the limit of detection (LOD) using several integrase recombinant plasmids was

14 copies and clinical sensitivity and specificity were high. The percentage of HIV-1 mother-to-child-transmission (MTCT) between January 2012 and October 2014 was only 2.2%. In chapter 4 we performed the first detailed characterization of the neutralizing antibody response in 322 Angolan HIV-1 infected patients and identified its determinants. Remarkably, 56% of the individuals had broad cross-neutralizing activity. Also, cross-clade neutralization was positively associated with subtype C infection and negatively associated with CD4 counts and antibody binding titers against envelope C2V3C3 region was a good indicator of neutralization breadth and potency. In chapter 4 we concluded that development of broad and elite antibody neutralization against HIV-1 requires long-term and low-level envelope V3C3 stimulation from highly diverse subtype C isolates.

Keywords: HIV genetic diversity, HIV epidemic in Angola, neutralizing antibody response, early infant diagnosis (EID) of HIV

Abbreviations

Ab Antibody

Ad Adenovirus

ADCC Antibody-dependent cell-mediated cytotoxicity

ADCVI Antibody-dependent cell-mediated virus inhibition

aLRT Approximate likelihood-ratio test

AIDS Acquired Immunodeficiency Syndrom**e**

APC Antigen presenting cells

ART Antiretroviral therapy

ARV Antiretroviral

BSA Bovine serum albumin

BSL-2 Biosafety level 2

bNAb Broadly neutralizing antibody

CA Conic shaped viral capsid

CD4bs CD4 binding site

CCR5 C-C chemokine receptor type 5

CDC Center for disease control and prevention

CMV Cytomegalovirus

CO2 Carbon dioxide

COVID-19 Coronavirus disease 2019

CRF Circulating recombinant form

CTL T cytotoxic lymphocytes

DC Dendritic cell

DMEM Dulbecco´s minimal essential medium

DNA Deoxyribonucleic acid

dsDNA Double stranded DNA

DRC Democratic Republic of Congo

DRM Drug resistance mutations

EID Early infant diagnosis

ELISA Enzyme-Linked Immunosorbent Assay

Fab Antigen-binding fragment

FBS Fetal bovine serum

Fc Fragment crystallizable region

FcRs Fc receptors

FP Fusion peptide

GALT Gut associated lymphoid tissue

GC Germinal Centers

GPCR G-protein-coupled receptor

HIV-1 Human immunodeficiency virus type 1

HIV-2 Human immunodeficiency virus type 2

HLA Human leukocyte antigen

HR1 Heptad Region 1

HR2 Heptad region 2

HTLV Human T-cell leukaemia viruses

Hu-BLT Humanized mice

HuMAbs Human Monoclonal antibodies

ICOS Inducible T cell co-stimulator

IDUs Injecting Drug users

IFN Type I interferon

IgG Immunoglobulin

IN Integrase

INIs Integrase inhibitors

IQR Interquartile range

KIR Killer immunoglobulin receptor

LTR Long terminal repeat

MA Matrix protein

ML Maximum likelihood

MPER Membrane proximal external region

mRNA Messenger RNA

MRCA Most recent common ancestor

MSM Men who have sex with men

MTCT Mother-to-child Transmission

NAb Neutralizing antibody

NC Nucleocapsid protein

NHPs Non-human primates

NNRTIs non-nucleoside reverse transcriptase inhibitor

NK Natural killer cells

NVP Nevirapine

OD Optical density

PAMPs Pathogen-associated molecular patterns

PBMC Peripheral blood mononuclear cell

PBS Phosphate buffered saline

PCR Polymerase chain reaction

PEP Post-exposure prophylaxis

PI Protease Inhibitors

PIC Pre-integration complex

PMTCT prevention of Mother-to-child transmission

PNGS potential-N-linked glycosylation sites

PrEP Pre-exposure prophylaxis

PRRs Pathogen-recognition receptors

PR Protease

RC Republic of Congo

RIP Recombinant Identification Program

RLU Relative light units

RNA Ribonucleic acid

RRE Rev Response element

RT Reverse transcriptase

RTC Reverse transcriptase complexes

SD Standard deviation

SHM Somatic Hypermutation

SHIV Simian-Human Immunodeficiency Virus

SIV Simian immunodeficiency virus

SPF Specific pathogen free

SU Surface glycoprotein

TAR Trans-acting response

TB Tuberculosis

TD Transmembrane domain

TDR transmitted drug resistance

Tfh T Follicular helper cells

Tfr T Follicular regulatory cells

TLR Toll-like receptor

TM Transmembrane glycoprotein

TNF- α Tumor necrosis factor α

URF Unique recombinant form

UNAIDS Joint United Nations Program on HIV/AIDS

VLPs Virus-like particles

VSV Vesicular stomatitis virus

VV Vaccinia virus

WHO World Health Organization

Units

ºC Celsius degrees

kb Kilobase

kDa kilodalton

ml milliliter

nm nanometers

µg micrograms

µl microliter

Table of contents

Chapter 1

Introduction

1. General introduction

1.1 HIV Discovery

The first cases of patients with the acquired immunodeficiency syndrome (AIDS) were reported in 1981 in the United States following the observation of opportunistic infections in young homosexual men [21]. Soon after the first reports in 1983, HIV the causative agent of AIDS was identified by Luc Montagnier and Françoise Barré-Sinoussi at the Pasteur Institute (France) [22, 23].

In 1986 HIV-2 was isolated in patients from Guinea-Bissau and Cape Verde Islands (West Africa) interned at Hospital Egas Moniz in Lisbon (Portugal) [24, 25]. Isolation of this new virus was possible due to the insightful work of Maria Odette Santos Ferreira at the Faculty of Pharmacy Universidade de Lisboa [25].

The identification and characterization of HIV was a major achievement, duely acknowledged by The Nobel Foundation in 2008, with the award of the Nobel Prize for Medicine to Luc Montagnier and Françoise Barré-Sinoussi. The knowledge built over the years trying to reduce the burden of HIV-1 infection globally also played an important role in minimizing the impact of other infectious diseases such as COVID-19 in terms of diagnosis, transmission, prevention and treatment [26, 27].

1.2 HIV in Sub-Saharan Africa and Angola

Despite recent progress in reducing the number of new infections, 1.5 million [1.0 million - 2.0 million] people became newly infected with HIV in 2020 globally and HIV/AIDS is still among the leading causes of disease burden and mortality in sub-Saharan Africa [28]. The UNAIDS fast-track goals for the control of HIV infection by 2030, define clear

and measurable goals for the implementation of public health policies and establish the need to reduce 75% of new HIV cases and deaths by 2020 and 90% by 2030, compared to 2010 [1]. However, none of the 44 Sub-Saharan African countries has reached the target of 75% reduction in new cases of HIV infection in 2020 [2]. Sub-Saharan Africa remains severely affected by the epidemic accounting for 67% of the people living with HIV in the world and for 60% of the new infections [28].

Angola is a South-western African country bordered by Republic of Congo, Democratic Republic of Congo, Zambia and Namibia. According to the UNAIDS report on the global AIDS epidemic 2013 [29] the estimated HIV prevalence and new infections in adults have decreased between 2001 and 2012 in all the bordering countries of Angola. For example, in the Republic of Congo HIV prevalence decreased from 4.7% to 2.8% and the number of new infections decreased from 6,600 to 3,400. In contrast, the estimated number of adults living with HIV in Angola has increased in the same period from 110,000 to 220,000 (1.8% vs 2.3% prevalence) and the estimated number of new infections rose from 16,000 to 23,000 [29]. This increasing tendency continued from 2010 onwards [2]. HIV-1 incidence in Angola increased 61.2% from 2010 to 2018 and the prevalence is now stable with 340,000 people living with HIV [2, 3]. Women and children are the most vulnerable populations [3]. Despite the recent advances in prevention of mother to child transmission (MTCT), Angola reported one of the highest rates of MTCT (25%) among the 22 priority countries included in the UNAIDS global plan in 2015 [8, 30]. Since then, the incidence of HIV-1 infection in children has been decreasing, however final vertical transmission rate including breastfeeding was 18.6% in 2020 [8] and is estimated that 5,200 children (aged 0-14 years) acquired HIV in 2020 [3]. In 2020, 11,000 women aged 15 and over became newly infected by HIV in Angola, more than 2 times the number of man in the same age group [3].

The development and availability of antiretroviral drugs helped to significantly improve the conditions of life of the HIV infected individuals, and especially to increase their life expectancy [31]. However, due to poverty related barriers, access to antiretroviral therapy remains limited in Africa and an effective vaccine is hoped to be a practical and cost-effective intervention to control the spread of the HIV/AIDS pandemic. ART has been available in Angola since 2000 for those infected with HIV who could afford buying ARV drugs. Since 2004, a national plan has been implemented to provide free ARV drugs to HIV-1 infected individuals using the WHO public health approach to ARV delivery [32]. At the end of 2012 the number of people on ART was 39,704 [29], 48% of the adults in need of treatment based on WHO 2010 guidelines [33]. More recent data shows that the access to antiretroviral therapy (ART) in Angola is still reason for concern, since at the end of 2020 the number of people living with HIV undergoing treatment was 111,168, approximately 33% of those in need of treatment [3]. The rising trend in the number of people living with HIV in Angola associated with the decrease in mortality due to ART [3], will inevitably increase the demand for ART in the near future. Increasing ART coverage of adults and children will inevitably lead to increasing drug resistance and raise the need for newer antiretroviral (ARV) drugs, such as 2nd generation Integrase Inhibitors (INIs) and Protease Inhibitors (PIs) that are not commonly available in Angola [34].

The frequency of transmitted drug resistance (TDR) in Angolan patients has risen from 1.6% in 2001 [7] to 16.3% in 2008–2010 [35, 36] suggesting that TDR may be an important public health problem in Angola. A more recent observational study registered a prevalence of 18% of drug resistant mutations (DRM) in 42 HIV-1 infected pregnant women naive to antiretroviral therapy [37]. Further work is required to characterize TDR level in Luanda as only a few patients living in this province have been included in previous surveys.

1.3 HIV-1 structure, genome organisation and replication cycle

1.3.1 Viral structure and genome organisation

HIV-1 is a lentivirus belonging to the retroviridae family. A common feature of all retroviruses is that they synthesize DNA from their RNA genome via the reverse transcriptase enzyme. The mature HIV virion, is spherical in form with a diameter of around 120 nanometers (nm). The virion is composed of two identical copies of positive single-strand RNA together with the reverse transcriptase (RT), protease (PR), integrase (IN) and RNase H enzymes encapsulated within the viral core (CA; p24), together with the accessory proteins Nef, Vif, Vpr and Vpu, which is enclosed within the matrix (MA; p17) [13, 38]. A lipid bilayer that is derived from the membrane of the host cell, as a result of the process of budding, envelopes the HIV virion, into which the gp41 is embedded, which in turn anchors the gp120 in a trimer of heterodimers (Figure 1).

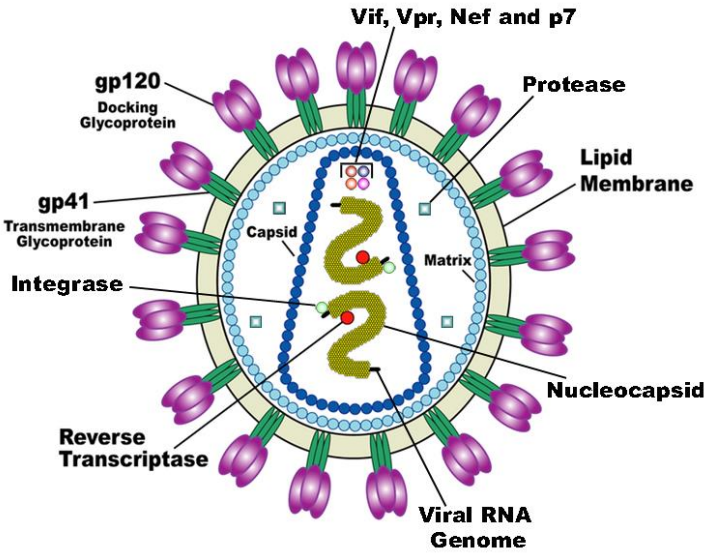


Figure 1- Schematic representation of a mature HIV-1 virion illustrating major viral components. Figure from http://commons.wikimedia.org/wiki/File:HIV_Virion.en.png (Consulted in July 2021).

HIV-1 virus consists of three structural genes: *gag*, *pol*, and *env* with flanking long terminal repeat (LTR) sequences at each end of the genome. In addition, HIV-1 possesses regulatory genes (*tat*, *rev*), and accessory genes (*vif*, *vpr*, *vpu*, and *nef*) (Figure 2).

The *gag* codes for the internal structural proteins of the virus: matrix (MA, p17), capsid (CA, p24), and nucleocapsid (NC, p7). The *pol* codes for the viral enzymes: reverse transcriptase (RT), which contains both DNA polymerase and associated ribonuclease H (RNase H) activity, integrase (IN) and protease (PR) [38, 39]. The *env* codes for viral envelope glycoproteins as a precursor (gp160), which is then processed to a surface glycoprotein, gp120 and a trans-membrane glycoprotein, gp41. The mature gp120.gp41 proteins are bound by non-covalent interactions and are associated as a trimer on the surface of virions. The envelope (Env) protein is responsible for recognition of cellular receptors and viral entry into cells.

The *tat* and *rev* regulatory genes modulate transcriptional and post-transcriptional steps of virus gene expression and are essential for virus replication and propagation. Tat acts by binding to the trans-activation response (TAR) RNA element and activates transcription initiation and elongation from the LTR promoter [40]. Rev acts by binding to Rev response element (RRE) and promotes the nuclear export, stabilization, and utilization of the viral mRNAs containing RRE [41, 42]. The genes *vif*, *vpr*, *vpu*, and *nef* are termed acessory genes however they play an important role modulating diferent steps of viral replication. Vif promotes the infectivity but not the production of viral particles. In the absence of Vif, the produced viral particles are defective, but the cell to cell transmission of virus is not affected significantly [43]. In addition, Vif prevents the action of the cellular APOBEC.3G (Apo-lipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G) a potent antiretroviral cytidine deaminase [44]. APOBEC.3G exerts its antiviral effect during reverse transcription to trigger G-to-A hyper-mutation in the

nascent retroviral DNA [45]. Vpr that is incorporated into the virion facilitates the nuclear localization of the preintegration complexes, and modulate cell division in order to accelerate the production of HIV proteins by arresting infected cells at the G2 phase of the cell cycle [46]. Vpx is found in HIV-2, but not in HIV-1. This accessory gene is a homolog of HIV-1 *vpr*, however the function in relation to Vpr is not fully elucidated; both are incorporated into virions at levels comparable to Gag proteins through interactions with Gag p6 [47]. Vpu is unique to HIV-1 and some SIV (e.g. SIVcpz). It has two essential biological functions; it degrades CD4 in the endoplasmic reticulum [48] and promotes extracellular release of viral particles [49]. Nef is important in pathogenesis, it down-regulates cell surface CD4 molecules, and increases the infectivity of viral particles [50].

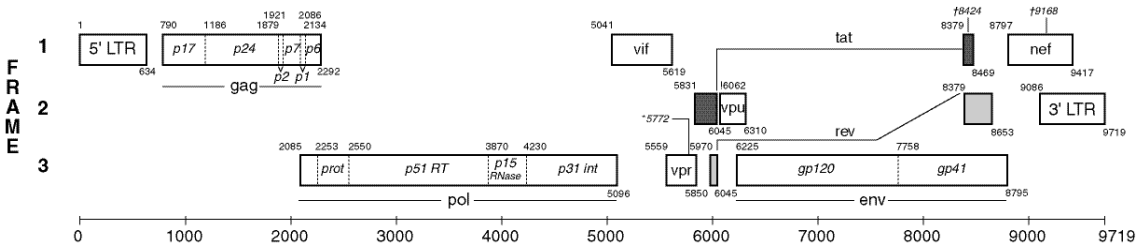


Figure 2- Organization and landmarks of the HIV-1 DNA genome**.** Open reading frames are shown as rectangles. The gene start is indicated by number in the upper left corner of each rectangle (ATG start codon). The number in the lower right indicates the position of the stop codon. The *tat* and *rev* spliced exons are shown in dark grey and grey, respectively. The numbering positions in the bottom of the figure are relative to HXB2 strain. Figure from Los Alamos National Library, http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html (consulted in July 2021)

1.3.2 Replication cycle

The replication cycle of HIV takes place in different phases (Figure 3). Entry is the first step in the process of HIV infection and requires binding of HIV gp120 to CD4 in host

cells [51]. This results in conformational changes in gp120, which then exposes the co-receptor binding sites. The co-receptors required for entry of HIV-1 are mainly CCR5 and CXCR4, and define viral tropism. Immediately following gp120 and co-receptor binding, further conformational changes take place in gp41, which allows it to expose the fusion peptide. The fusion peptide is then inserted into the target cell membrane enabling further conformation changes in gp41 that drive virus to cell fusion and subsequent entry into the cells [51, 52]. Cell to cell transmission of HIV is another effective mechanism of virus proliferation that do not necesseraly involve binding of HIV Env to the host cell receptor [53].

Following membrane fusion the virus core enters and uncoats into the cytoplasm of the target cell. The virus RNA genomes are converted into double-stranded DNA. Reverse transcriptase (RT) has two distinct enzymatic activities: it is a DNA polymerase capable of copying either RNA or a DNA template into a complementary DNA sequence; and it is an RNase H, capable of degrading the RNA strand of an RNA.DNA duplex into small pieces once it has been used as a template for the first DNA strand [54]. The reverse transcription is initiated, and viral RNA is converted into a DNA/RNA hybrid. The template RNA is degraded by the RNase H activity and RT polymerase synthetises the complementary DNA strand to form the double helix DNA molecule. The viral DNA was thought to be translocated from the cytoplasm to the nucleus as part of the pre-integration complex (PIC) (viral DNA associated with MA, RT, and IN). However, very recently new insights about HIV-1 replication cycle have been published showing that HIV-1 CA might travel intact to the nucleous of the host cell (Figure 4), apparently the diameter of the nucleous pore is sufficient to allow the passage of undisrupted CA. This disruptive finding is particularly important since a new drug class of ARVs, capsid inhibitors, will soon be available [55-57]. The viral capside containing the PIC is transported by the

microtubules to the vinicity of the nucleus, in a process that involves the binding to different nucleoporins (NUP62, NUP153, NUP358) and cleavage polyadenylation specificity factor 6 (CPSF6). In the nucleus reverse transcription is completed and the CA and PIC desagregates probably by the mechanic pressure of the newly synthesized viral DNA. The integration of the viral DNA in the genome of the host cell is mediated by the viral IN and the cellular cofactors NUP153 and LEDGF/p75. The integration process is completed when cellular repair enzymes fill in the gaps between the integrated viral DNA and the host target DNA, and is then designated proviral DNA [58].

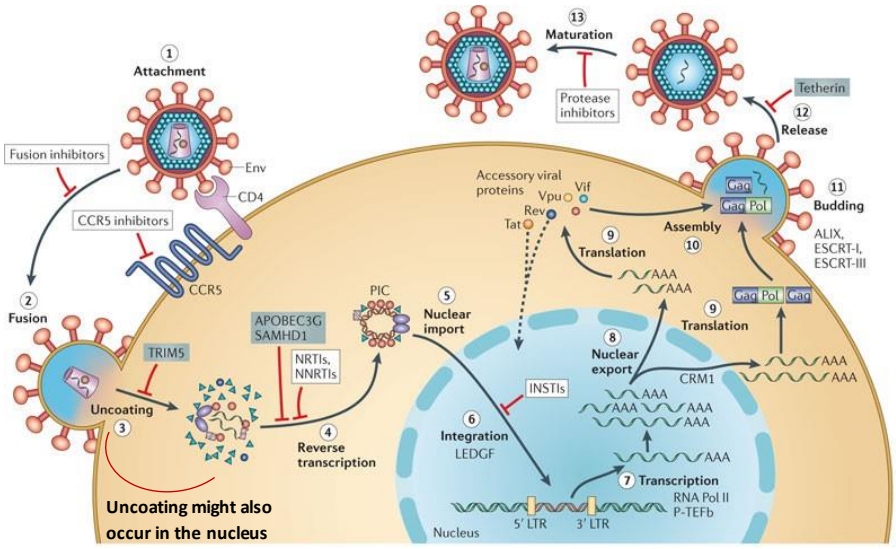


Figure 3 - Schematic representation of the life cycle of HIV. (Figure adapted from [58])

After integration, transcription is mediated by the host cell RNA polymerase II using integrated provirus as the template and generates full length viral RNAs that serve as mRNA. The HIV-1 LTR serves as the site of transcriptional initiation and harbors cis-acting elements required for RNA synthesis. The transcription initiates at the U3/R junction that contains binding sites for numerous proteins that participate in the regulation of the expression of the viral genome (e.g. Sp.1 and two NF.κB binding sites). Successful transcription leads to the generation of HIV viral transcripts that are derived from a single full-length transcript by alternative splicing, generating messenger RNA (mRNA) with

common 5´and 3´ends [59]. HIV-1 transcripts can be grouped into three different classes: double spliced mRNA or early transcripts encode early regulatory proteins such as Tat, Nef and Rev**;** single spliced mRNA or late transcripts encoding Env, Vif, Vpr and Vpu; and unspliced and complete mRNA that encode for the polyprotein precursors Gag and Gag-Pol and are incorporated in the viral particles as genomic RNA [59, 60]. To complete the expression of the late transcript proteins from the early transcripts, Tat and Rev are necessary. Tat binds to a short secondary structure located in the R region of the 5'LTR sequence known as the transactivation response region (TAR). Tat plays a critical role in up-regulating transcription from the LTR by more than 100-fold. There are two possible ways that Tat can increase HIV-1 RNA synthesis, one is to augment transcription initiation and the other one is to improve the activity of RNA polymerase [61, 62]. The transport of unspliced and single spliced mRNA outside the nucleus to the cytoplasm to be translated is performed by Rev. This process is mediated by the binding of Rev to the Rev response element (RRE), a 240 base region of complex RNA secondary structure, present in the midle of the *env* gene. Via the nuclear export factor (NES), Rev binds to the chromosome region maintenance 1 (Crm1) (nuclear export factor) in the presence of RAs-related Nuclear protein (Ran GTPase) connected to guanisine triphosphate (GTP). Posteriorly, DEAD-box polypeptide (DDX1) and DEAD-box polypeptide 3 (DDX3) bind to the N-terminal domain of Rev. The translocation process through the nucleus pore seems to be mediated by the DDX1 and DDX3 that bind to the nucleoporins. In the cytoplasm the hydrolyses of GTP induces the release of Rev from the viral mRNA. The translation of the viral mRNA is determined by the internal ribosome entry segment (IRES) present in the beginning of the mRNA [62, 63].

The HIV Env glycoprotein is synthesized in the rough endoplasmic reticulum (RER) to generate the Env precursor protein, gp160 and then transported to the Golgi apparatus, where it is cleaved by a host protease (furin) into gp120 and gp41. The Env protein is incorporated in the plasma membrane of the cell, while Gag and Gag-Pol are assembled into viral capsids in the cytoplasm [60, 64]. Two identical copies of the viral genomic RNA bound by p7 Gag product form a nucleoprotein complex. The newly formed nucleoprotein particle migrates to the plasma membrane at the site of insertion of Env, and Gag and Gag-Pol precursors [60]. Immature virus particles are assembled at the plasma membrane and are released by budding through the plasma membrane, acquiring a portion of the plasma membrane that contains gp41 and gp120 [62]. Pol, the protease, and the Gag proteins are generated by proteolytic cleavage mediated by the protease domain of the Gag-Pol precursor polypeptides upon release of the particles from the cell, thereby producing mature virus particles [65].

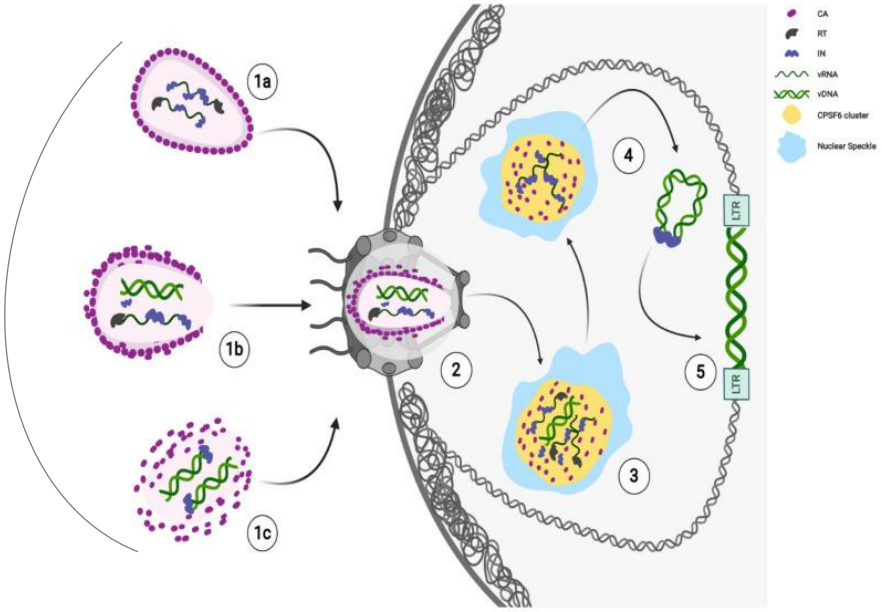


Figure 4- New mechanism of CA nuclear import. Different potential core states can reach the nucleus: nearly intact (1a), remodeled (1b), or partially uncoated (1c). Figure adapted from [56]

1.4 HIV envelope glycoproteins

HIV-1 Env glycoproteins are assembled as Env spikes. They are responsible for interacting with cellular receptors and initiating the fusion of the viral and cell membranes. They are the only viral target of neutralizing antibodies. The functional envelope spike consists of a trimer of heterodimers formed by two glycoproteins, gp120 (the exterior envelope glycoprotein) and gp41 (the transmembrane glycoprotein)[13]. Three gp120 molecules interact non-covalently with three gp41 units forming an oligomer, where the trimeric structure is maintained by the interactions between the gp41 domains. The Env trimer oscillates between open and closed conformations, most of the time the trimer is closed except when it interacts with the host cell receptors[13].

1.4.1 The gp120 molecule

The gp120 molecule consists of five relatively conserved regions (C1-C5), interspersed between five variable regions (V1-V5), which are delimited by cysteine residues forming disulfide bonds [66]. Gp120 is a highly glycosylated protein, N-linked glycans account for half of its mass, with a small proportion being O-linked sugars [13, 67]. The gp120 core is composed of three general areas: the inner domain, the outer domain, and the bridging sheet (Figure 5). The inner domain is formed mainly by the C1 and C5 regions which interact with the gp41 trans-membrane unit [68]. The inner domain is devoid of glycosylation [69]. The outer domain is heavily glycosylated to shield its antigenic surface, protecting it from antibody recognition. The proximal end of the outer domain includes V4 and V5 variable loops, whereas the distal end includes the base of the V3 loop, which interacts through hydrogen-bonds with the V1/V2 stem emanating from the inner domain. In between the outer and inner domains is the bridging sheet region, formed by four antiparallel β-sheets: β2 and β3, which constitute the stem of the V1/V2 loop; and

the β20 and β21 of the C4 region [70]. Besides these structural sites, gp120 has two main functional sites: the CD4 binding site (CD4bs) and co-receptor binding site. The CD4bs is a conformational region that is only apparent in the context of the liganded structure of gp120. The CD4 binding loop projects away from the centre of the outer domain and the β20-β21 segment of bridging sheet [62]. The α-helices of the inner domain, the CD4 binding loop and the β20-β21 segment of bridging sheet create a long, narrow cavity, lined principally with hydrophobic side chains, in which many of the residues that are presumed to contact CD4 are located near or within this long cavity [71]. The binding of CD4 induces large conformational changes in the inner domain, which leads to the formation of the bridging sheet and co-receptor binding site [72]. Neither the receptor (CD4) nor the co-receptor (CXCR4 or CCR5) site is properly formed in the unliganded conformation of the gp120 core. In the unliganded conformation, the bridging sheet can close up to create the co-receptor binding surface, which is flanked by the V1–V2 and V3 loops [73]. The V3 loop has been shown as the major determinant of HIV tropism, which demonstrates its involvement in the co-receptor-binding site [74].

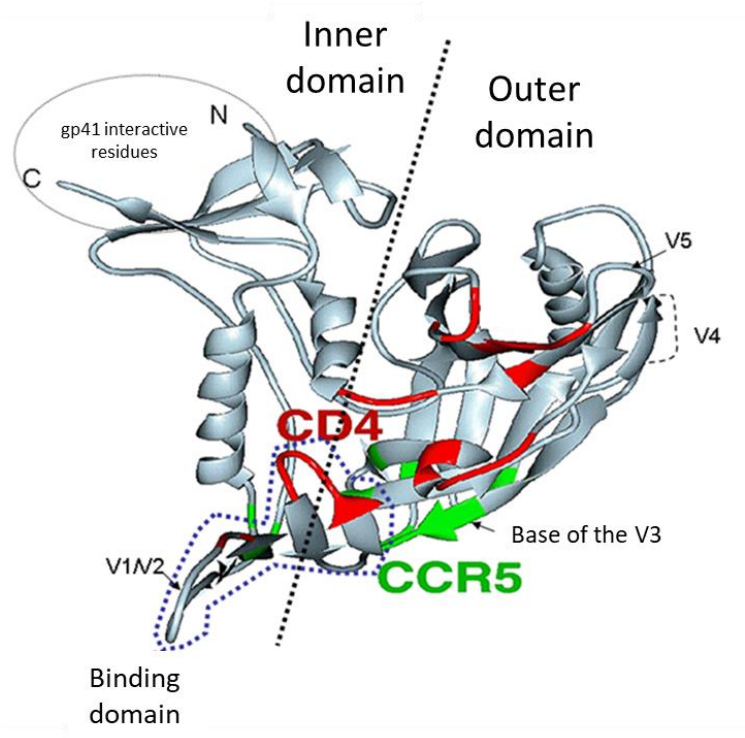


Figure 5 - gp120 molecular structure representing the outer domain, the inner domain and the V1/V2 small domain. CD4 binding site and CCR5 in the base of the V3 are coloured red and green respectively. Adapted from [62]

1.4.2 The gp41 molecule

The gp41 molecule is a transmembrane glycoprotein that is less variable and less glycosylated than gp120. It interacts non-covalently with gp120 and is responsible for maintaining the trimeric structure of the envelope glycoprotein, although its structure in the native conformation is unknown. The HIV-1 envelope glycoprotein gp41 is a homotrimeric structure formed by three gp41 monomers, with each monomer non-covalently associated with gp120. Each gp41 molecule consists of three domains: an extracellular domain (ectodomain), a transmembrane domain (TMD), and a C terminal cytoplasmic tail [13, 72].

1.5 HIV co-receptor usage and tropism

Not long after the discovery of the CD4 molecule as the major receptor for HIV [75], new evidence started to accumulate indicating that CD4 alone was not sufficient for HIV to enter the target cells. In 1986, Maddon et al. showed that CD4 expressed on mouse cells allowed the virus to bind but did not confer virus entry [76]. These results supported the conclusion that this restriction was due to the requirement for a cofactor of unknown identity that was specific to human cells [77, 78]. Subsequently, several in vitro studies using CD4+ T cell lines and macrophages showed different levels of HIV-1 infectivity depending on the cell lines used, which led to the identification of two main cell co-receptors that HIV-1 requires to enter the host cell CCR5 and CXCR4. The first HIV-1 co-receptor was identified in 1996 and was named fusin, because it mediated HIV-1 fusion [79] and was thereafter renamed CXCR4 [80]. The second HIV-1 co-receptor was

identified based on the finding that the CC-chemokines (i.e. RANTES, MIP.1α, and MIP.1β) that are the natural ligands of CCR5 could block the infection of certain HIV-1 isolates in primary macrophages [77, 78, 80, 81]. HIV-1 strains able to use both co-receptors were termed dual-tropic (D-tropic). Both CCR5 and CXCR4 belong to the superfamily of seven transmembrane (7TM) G protein-coupled receptors. More than fourteen other 7TM receptors or structural-related molecules have been suggested to act as co-receptors for entry of HIV-1 *in vitro*. Currently, there is little evidence to suggest that co-receptors other than CCR5 and CXCR4 are used significantly in vivo, although there is a body of literature that demonstrates that HIV-1 employs a number of methods to get into target cells [82-84].

HIV tropism is now commonly defined based on the co-receptor usage which is defined as the ability of a particular HIV-1 virus to infect a target cell using a specific co-receptor, either CCR5, or CXCR4 or both. The major genotypic determinant for HIV-1 co-receptor usage is the V3 variable loops of the gp120 envelope glycoprotein [85]. Different bioinformatic tools have been developed to predict HIV-1 co-receptor usage from the amino acid sequence of V3, taking into account the key amino acids at positions 11 and 25, plus other sites in V3 that differ between CCR5 and CXCR4. The most commonly known and used is geno2pheno from the Max Planck Institute Informatik and was used to determine coreceptor usage in chapter 4 of this thesis [86].

2. HIV genetic diversity

2.1 HIV classification

HIV is a member of the lentivirus genus of the *Retroviridae* family. The name lentivirus refers to slowly replicating viruses, because these viruses take a long time before it

induces the full-blown disease. Two types of HIV have been characterized: HIV-1 and HIV-2 and both share 40% to 50% genetic homology, with the greatest sequence divergence localized in the envelope gene [87]. HIV-1 is thought to be the result of cross-species transmission of simian immunodeficiency viruses (SIVs) isolated from chimpanzees (SIVcpz) or gorilas (SIVgor)[88], while HIV-2 is most closely related to a virus found in sooty mangabeys (SIVsm) [89].

HIV-1 is divided into group M (main), the group O (outlier), the group N (non-M/ non-O), and the more recently identified group P. Group M is responsible for the worldwide HIV-1 epidemic, the other groups represent a minority of HIV-1 strains and don't have global expression. HIV group M is classified into 10 subtypes: A, B, C, D, F, G, H, J, K and more recently identified subtype L[90] and 8 subsubtyes (A1-A6 e F1-F2). [91, 92].

Studies have shown that intra-subtype genetic distance can differ by up to 20% in the *env* gene and inter-subtype genetic distances can reach up to 35% [92] [13]. Sequencing full-length genomes have led to the identification of inter-subtype recombinants know as circulating recombinant forms (CRFs) [93]. These are presumed to be the result of recombination between different subtypes within an individual patient concurrently infected with HIV-1 of two or more subtypes. The inter-subtype recombinant genomes become designated as CRFs if; i) the identical recombinant viruses are identified in at least three epidemiologically unlinked people, ii) are characterized by full-length genome sequencing that share the same recombinant structure, and iii) form a monophyletic cluster in all regions of the genome; and as URFs, if only one or two sequences are available [94]. Recombinants are currently estimated to be responsible for at approximately 22% of HIV-1 infections worldwide [93]. The CRFs are named with a number sequential in the order in which they are reported in the literature and followed by the letters of the subtype involved, starting with CRF01_AE. If the recombinants

consist of more than two subtypes involved, they are therefore replaced by designation "cpx", meaning complex, e.g. CRF04_cpx (A, G, H, K, and U). Taxonomically, the CRFs are at the same level as the subtype [93]. Currently, more than 100 CRFs for HIV-1 (CRF01 to CRF102) and only one for HIV- 2 (HIV2.CRF01_AB) are found in the HIV database at Los Alamos National Laboratory (https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html), most of them having been described in Africa. Some CRFs are major strains circulating in certain regions and responsible of for more recent epidemics. For example CRF0l_AE and CRF02_AG are dominant in Thailand, Asia and West Africa respectively [93].

2.2 Global distribution of HIV subtypes

HIV-1 genetic subtypes are unevenly distributed in different geographical locations. According to recent studies, the most prevalent HIV-1 subtypes are subtypes A, B and C [93]. Subtype C accounts for nearly half (48%) of all HIV-1 infections in 2010-2015, while subtypes B, A, G and D account for 12.1%, 10.3%, 4.6% and 2.7%, respectively [92]. The subtypes F, H, J, and K all together account for approximately 1% of infections. The circulating recombinant forms CRF02_AG and CRF01_AE are responsible for 7.7% and 5.3% of cases respectively. Other CRFs account for the remaining 3.7% of infections. All recombinant forms (all CRFs and URFs) are responsible for over 20% of infections worldwide [93]. The global distribution and prevalence of HIV-1 worldwide are shown in Figure 6. This distribution reflects the present situation and might be susceptible to modifications in the next years. Subtype A viruses are predominant in Central and Eastern Africa (Kenya, Rwanda, Uganda, and Tanzania) and in Eastern European countries formerly constituting the Soviet Union. Subtype B, which is the most widely disseminated subtype, is predominant in North and Latin America, the Caribbean,

Europe, and Australia. It is also common in several countries of Southeast Asia, North Africa, Middle East (Israel), and among South-African and Russian homosexual men. Subtype C is predominant in southern Africa, Ethiopia and India. Subtype D viruses are found principally in East Africa and to a lesser extent in West Africa. CRF01_AE and subtype B co-circulate in South-East Asia, whereas CRF02_AG, along with other recombinants, dominates in West and West Central Africa. In South America, the epidemic is a mixture of subtype B and BF recombinants, with a small proportion of subtype C infections. In East Asia subtypes B, C and BC recombinant strains dominate. Central Africa harbors a complex mixture of rare subtypes (F, G, H, J and K) and recombinants, without any predominant strain. [91-93].

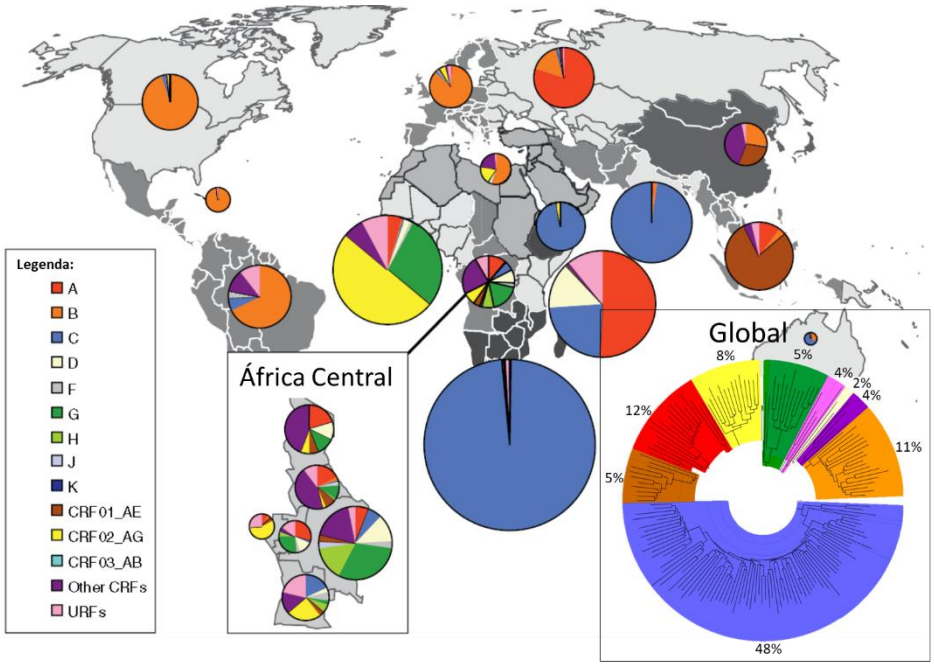


Figure 6 – Global distribution of HIV-1 pure subtypes and CRFs. The pie charts in the map represent the distribution of HIV-1 subtypes and CRFs from 2014 to 2017 for each region. The dimension of the pie charts are proportional to the number of people living with HIV in that region. The colours representing the different subtypes and CRFs are indicated in the legent on the left. In the bottom right is represented a phylogenetic tree showing the more prevalent subtypes and CRFs in the world based in the HIV-1 *pol* gene sequences obtained from (https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html) adapted from [95].

Angola is one of the countries with the most diverse HIV epidemic in the world [91]. In Angola the HIV-1 epidemic is highly complex with all HIV-1 subtypes, several circulant recombinant forms (CRFs), unique recombinant forms (URFs) and untypable (U) strains reported [4, 5, 7, 34-37, 96-99]. This genetic complexity poses significant challenges to laboratory diagnosis, antiretroviral treatment (ART) and effectiveness of prevention strategies [15, 34, 37, 100].

2.3 Consequences of HIV-1 genetic diversity

2.3.1 Impacts of HIV-1 genetic diversity on transmission and disease progression

Earlier cohort studies found an association between CRF01_AE and heterosexual transmission as well as between subtype B and intravenous drug use [101, 102]. However later on, contrary results were published from a longitudinal study performed in Thailand that found an increased probability of CRF01_AE transmission among IDUs compared with subtype B [103]. A more recent study performed in HIV-discordant couples in Uganda found that subtype A was associated with a significant higher rate of heterosexual transmission than subtype D [104]. The rate of transmission may be related to differences in subtype-specific co-receptor tropism. The R5 tropic HIV strains are more frequently transmitted than strains that use CXCR4, however X4 tropic viruses emerge later in infected patients and are associated with more rapid disease progression [105]. It was found that HIV-1 subtype D was more prone to use CXCR4 in early infection which may in part explain their reduced heterosexual transmissibility when compared to other genetic forms [104, 106, 107] whereas subtype A mostly used CCR5 even in late infection. This may explain why HIV-1 subtype D infected patients had more rapid progression that those infected with subtype A in Uganda, Kenya, and Tanzania. The percentage of

CXCR4 viruses appears lower in subtype C than in subtype B, even when the viruses are obtained from patients with advanced AIDS [108]. A previous study in Tanzania suggested that subtypes A and C are more likely to be perinatally transmitted than subtype D [109], and that pregnant women infected with subtype C were more frequently susceptible to transmit HIV to their children than those infected with subtype B [110]. Also a recent publication from the Swiss HIV cohort study showed a substancial increase in transmission of non B-subtypes between 1990 and 2019 among men who have sex with men (MSM) [111].

A number of studies have suggested differences in the rate of disease progression associated with infection by different HIV-1 subtypes [112]. The best studied of these are subtypes A, D and C, which co-circulate in several African countries [15, 113-115]. Large cohort studies addressed the issue of the impact of HIV-1 subtype on disease progression [15, 115, 116]. Most of these studies point to higher virulence of subtype C strains in comparison to other HIV-1 subtypes [15, 115-117]. One of those studies an European seroconverter cohort showed that HIV-1 subtype significantly influenced CD4 count at seroconversion and rate of decline, subtype C-infected participants had the lowest CD4 levels at seroconversion [116]. A multinational clinical trial PEARLS (ACTG A5175) has shown that HIV-1 subtype C infection was associated with higher virological risk failure compared to subtype B and other non-subtype B infections [117]. A large sub-Saharan African cohort also found that subtype C and subtype D-infected participants had faster disease progression in comparison to subtype A [15]. Moreover, a retrospective cohort study (1996-2007) reported that African patients infected with HIV-1 non-B subtypes had slower rates of disease progression compared to Haitians and Canadians infected with subtype B viruses [118]. Another study showed that Senegalese woman infected with non-A subtypes were 8 times more likely to develop AIDS than those

infected with subtype A [119]. Also, a study of a Kenyan cohort showed that patients infected with subtype D had a higher mortality rate and a faster decline in CD4+ count than those infected with subtype A or C [120]. The propensity of subtype D to exhibit a greater degree of using dual co-receptor than other subtypes [106] may help to explain the observation that subtype D appears to be associated with a more rapid rate of disease progression than other subtypes.

A very recent study, conduct in sub-saharan Africa comparing the viral replication capacity of viruses isolated from HIV-1 infected patients from East and West Africa, found that the viruses isolated from West African patients had higher replication capacity in a subtype dependent manner and that CRF02_AG was associated with lower replication capacity due to the presence of a specific polymorfism in Gag [121].

An important unanswered question is the biological basis for these differences. A possible clue comes from data suggesting that emergence of X4 variants is less common in HIV-1 subtype A infection, compared with HIV-1subtype D infection, but the differences were not statistically significant, and the cross-sectional design precluded analysing the effect of coreceptor usage on the risk of disease progression [113]. It is important to explore the relationship between emergence of X4 virus and subsequent disease progression in other cohorts with different viral subtypes in circulation. A recent study analysing the determinants of HIV-1 subtype C pathogenesis found that R5 tropic subtype C virus strains in the absence of co-receptor switch have the ability to use alternative coreceptors *in vitro* during the course of infection, in particular FPRL1 [122], which may account for the apparent higher virulence of subtype C strains. However, finding a cause effect in all of these cohort studies is very difficult to measure since there are several other factors that can impact transmission and disease progression besides viral diversity, e.g. behavioral, epidemiological and immunological factors [123].

2.3.2 Impacts of HIV-1 genetic diversity on diagnosis

HIV-1 fourth-generation immunoassays are able to detect all known HIV-1 group M subtypes, group O and HIV-2 positive samples with very high sensitivity and specificity. In contrast, the results obtained from antigen-antibody rapid tests are far from satisfactory with more diverse HIV strains [124]. These fourth-generation immunoassays and the ultra sensitive 5th generation assays detecting Ag p24 provide an advantage for detection of infection during the window period prior to seroconversion since the diagnostic window may be reduced by an average of 5 days relative to an IgM-sensitive EIA [124, 125]. However, such advanced assays are often not available in resource limited countries where most new infections occur. In field situations with a high diversity of circulating HIV strains, such as in Angola, the performance of rapid diagnostic tests is much less satisfactory with sensitivities ranging from 94.1 to 100% and specificities ranging from 88.0% to 98.8% [126]. Also, early antiretroviral therapy diminishes the production of antibodies which compromises the performance of serological tests [127, 128]. PCR-based assays for viral load measurements, also have difficulty detecting and reliably quantifying HIV-1 RNA when testing diverse genetic variants of HIV-1 from Africa, and different assays frequently yield discordant viral load results [129-131]. Furthermore, relative to mother-to-child-transmission (MTCT), serological assays do not allow the early diagnosis of HIV-1 infection because of the persistence of maternal HIV-1 antibodies in infants during the first 12-18 months of life. The WHO recommends the use of molecular-based virological testing to determine the infection status for HIV-1-exposed infants during the first 4-6 weeks of life or at the earliest opportunity thereafter [132]. Despite the high accuracy of fourth-generation and HIV RNA tests, their sensitivity could potentially be affected in settings of expanded ART for prevention of MTCT (option B and B+), which reduce circulating HIV-1 RNA and viral particles [133].

Qualitative DNA PCR test which detect proviral DNA in peripheral blood mononuclear cells (PBMCs) is therefore recommended for early infant diagnosis (EID) of HIV-1 and is the most widely implemented test in resource-limited settings [134, 135].

2.3.3 Impact of HIV-1 genetic diversity on antiretroviral therapy

The development of resistance to antiretroviral drugs continues to be an important problem in the treatment of HIV-infected individuals. Studies around the world have demonstrated that different HIV-1 subtypes have similar susceptibilities to currently used antiretroviral drugs, which were originally developed based on subtype B viruses [136, 137]. However, viral subtype can be associated with treatment virological failure [117]. Different HIV genetic forms carry in their genomes genetic signatures and polymorphisms that could alter the structure of viral proteins which are targeted by drugs, thus impairing ART efficacy. Several mutations are generally required for the virus to become resistant to protease inhibitors (PI) and integrase inhibitors (INIs), whereas a single amino acid substitution can induce resistance to the non-nucleoside reverse transcriptase inhibitor (NNRTIs) [138]. The NNRTI resistance mutation V106M occurs in subtype C and CRF01_AE, but not in subtype B, and the protease inhibitor (PI) mutation L89I/V occurs in subtypes C, F and G, but not in B [139]. Among non-B subtypes differences are also notable, as nevirapine resistance mutations developed more frequently in subtype D than A in a mother-to-child transmission prevention study using single-dose nevirapine [140]. The probability of selecting resistance varies between HIV-1 subtypes. There are specific drug resistance mutations pathways characteristic of certain HIV subtypes which might pose a major challenge in epidemics driven by highly divergent HIV strains, such as in Angola. In fact, a small surveillance study conducted in Angola recently showed a high prevalence (18%) of drug resistance mutations in HIV-1 infected pregnant women [34, 37].

3. Neutralizing antibodies against HIV infection

3.1 Antibody response to HIV

Following acute HIV infection an abundance of antibodies are elicited. Antibodies have the ability to inhibit HIV-1 infection through multiple pathways: they can bind cell-free virus and prevent the infection or they can complex with Fcδ receptor to block HIV-1 through effector cell mechanisms [13]. The progression of these antibodies include: i) binding antibodies that first develop within 8 days after plasma virus detection and initially exist as antigen–antibody complexes, followed a few days later by circulating anti-gp41 antibodies, and further few weeks later with anti-gp120 antibodies targeting mostly the V3 loop [141]; ii) non-neutralizing antibodies that act together with innate immune cells to kill virus infected cells known as antibody dependent cell-mediated cytotoxicity (ADCC) and antibody dependent cell-mediated viral inhibition (ADCVI) [142]; iii) neutralizing antibodies (Nabs). Neutralizing antibodies aim to block viral entry and subsequent infection by binding to exposed regions on the envelope. It is generally recognised that a successful HIV-1 vaccine should elicit potent and broadly neutralizing antibodies similar to those found in some HIV-1 controllers [13, 143, 144]. Such antibodies may protect patients from disease progression and can neutralize a wide range of genetically diverse HIV-1 subtypes [13].

3.1.1 Non-neutralizing antibodies (ADCC and ADCVI)

The HIV-1 Env glycoprotein is highly immunogenic but the antibodies elicited by it during infection are generally either non-neutralizing (nnAbs) or lack neutralizing breadth or potency against primary HIV-1 strains thus failing to inhibit viral replication in infected individuals [145, 146]. Nonetheless, non-neutralizing antibodies can clear the virus by binding to the infected cells and initiate the recruitment of activated effectors cells, which in turn induces cytolysis or apoptosis of infected cells. ADCC is the result of the formation of a complex between the IgG Fab portion of the antibody with the viral protein on the cell surface and binding of the Fc portion to the Fc receptors of the antibody [142]. Fc receptors are expressed on natural killer cells, monocytes, macrophages, dendritic cells and neutrophils [142]. The binding to the Fc receptors can lead to release of antiviral cytokines [147] resulting in the killing of the infected target cell [148]. ADCVI also involves the interaction between a target cell, ADCVI antibody and an effector cell. However, rather than causing cell death, ADCVI antibodies aim to reduce the viral output from infected target cells [149].

Vaccine studies in both humans and non-human primate model systems have brought some evidence that non-neutralizing antibodies may provide protection from infection. In fact, in human vaccine studies the few correlates of protection identified until now, in the only HIV-1 vaccine clinical trial (RV144) that demonstrated some level of protection against HIV-1 infection, were related to non-neutralizing antibodies to V1/V2, high levels of antibody-dependent cellular cytotoxicity (ADCC) after controlling for IgA, and HIV-1-specific IgG3 responses [9, 10, 12, 13, 117, 145, 150-154].

3.1.2 Neutralizing antibodies (nAbs)

In HIV-1 infection, nAbs can block the virus-cell interaction by inhibiting the binding of the virion to CD4 and co-receptors on the cell surface, therefore preventing

conformational changes of the virus envelope that are required for subsequent steps in the virus life cycle. The earliest neutralizing antibodies can be detected within months of infection in most HIV-1 infected individuals [155, 156]. Most HIV-1 infected individuals produce antibodies capable of inhibiting their own virus (autologous virus) and are known as autologous neutralizing antibodies [155-158]. These antibodies target immunogenic exposed regions of the HIV-1 virion; however their neutralization capacity is transient and have little impact in the control of the infection already established due to the continuous capacity of HIV-1 to diversify and escape recognition of these antibodies [13, 159-163]. The appearance of neutralization escape variants soon after the autologous response supports the notion that these antibodies exert immunological pressure on the virus [155, 164]. Autologous antibody response primarily target variable regions rather than the conserved regions of the HIV-1 envelope, explaining the strain specificity of these antibodies [165, 166]. The V1V2 region was therefore shown to be a frequent target of autologous nAbs in HIV-1 and Simian–Human Immunodeficiency Virus (SHIV) [167] [168]. A study reported on the detection of autologous nAbs as early as 52 days after detection of HIV-specific antibodies in acutely infected patients [155]. Gray et al, have evaluated autologous and heterologous neutralizing antibody responses in 14 HIV-1 subtype C acutely infected individuals [156]. Their results revealed that potent autologous neutralizing antibodies are produced within 3 to 12 months post-infection with an increase in autologous antibody production observed within the first 6 months [156]. Another study identified the C3V4 region, specifically the C3 α.2-helix as a major target of autologous neutralization antibodies in subtype C infections [168]. V4 does not appear to be a significant autologous nAb target, although changes in this region may mediate neutralization escape [169] while the role of V5 is less clear. This continuous interdependent cycle of neutralizing antibodies driving the selection of escape mutants,

that in turn drive the maturation of different B-cell percursors selecting new antibodies, might increase the probability of eliciting broad neutralizing antibodies (bNAbs).

3.1.3 Broad neutralizing antibodies (bNAbs)

In contrast to the early autologous Nab responses, antibodies capable of neutralizing heterologous viruses (heterologous neutralization) develop later in infection and can be remarkably potent and cross reactive [13, 155, 156]. Only a small percentage of chronically infected patients can develop broadly cross-reactive nAbs against multiple HIV-1 viruses [146, 170-172]. The reasons why some individuals develop broadly cross-reactive nAbs is still not fully understood, however some determinants of development of bNAbs were already identified such as viral diversity, duration of infection and viral load levels [146, 173]. Broadly cross-reactive nAbs response was initially thought to be quite uncommon [174], but recent studies have described the presence of broadly cross-reactive nAbs in different cohorts [173, 175, 176]. These studies indicate that such antibodies are not rare [173, 175, 176], providing evidence that the natural B cell response can generate broadly cross-reactive nAbs against HIV-1. This type of antibody response could be the key to stop or at least contain HIV-1 transmission if they could be induce by vaccines prior to virus exposure [159-161], since some bNAbs have showed the ability to prevent and suppress HIV-1 infection in Humans [177-182]. However, the effectiveness of antibodies elicited by candidate immunogens and vaccines to neutralize heterologous primary HIV-1 strains *in vitro* and *in vivo* tested so far has been very limited [9-14]. Broadly neutralizing monoclonal antibodies against HIV-1 have been identified and studied comprehensively over the last years [10, 15, 136, 150, 151, 159-161, 172, 183-212]. To date more than one hundred bNAbs against HIV-1 have been produced (https://www.hiv.lanl.gov/content/immunology/ab_search.html). bNAbs target six

indentified epitopes in the HIV-1 envelope: the CD4 binding site (CD4bs); the V1/V2; V3 glycan; gp41 MPER, gp41 and gp120 interface residues and the fusion peptide [13, 213, 214] (Figure 7).

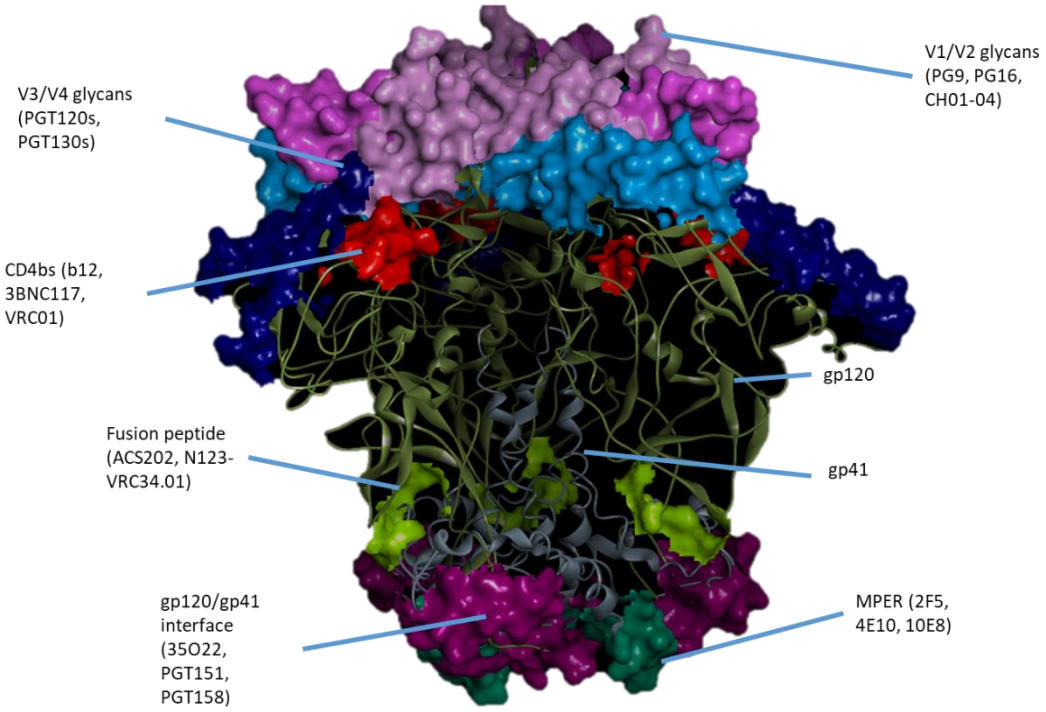


Figure 7- Molecular structure of the HIV-1 envelope gp160 showing the functional domains and neutralizing epitopes. Model based on the glycosylated BG505SOSIP.664 trimer. PDB ID 4ZMJ. gp120 is coloured in olive green and gp41 in gray. The six bNAbs target epitope regions are highlighted in different colours. V1/V2 in pink (V1 dark pink, V2 light pink); V3/V4 in blue (V3 dark blue, V4 light blue); CD4bs in red; fusion peptide in light green; gp120 and gp41 interface residues in magenta; MPER region in dark green. Example of bNAbs targeting the specific regions are given. Adapted from [62]

These new bNAbs were derived from donors infected with different HIV-1 subtypes and the success of this effort was based on the combination of three strategies: a) the selection of chronically infected individuals with potent and cross-subtype reactive serum antibodies; b) the use of novel selection of screening approaches; and c) the development of efficient methods to isolate human monoclonal antibodies [13].

However, the mechanisms underlying the elicitation of such antibodies by B cell populations has just started to be unveiled [2, 14, 26, 55, 56, 111, 121, 215-228]. Guiding the immune system to elicit such bNAbs remains a major challenge due to extremely complex antibody maturation pathways and high levels of somatic hypermutations (SHM) required by bNAbs to acquire neutralization breadth [215].

HIV-1 has developed multiple escape mechanisms to avoid neutralization. Such features include the inaccessibility of relevant epitopes due to the trimeric structure of envelope, shielded envelope glycoproteins by glycans which makes them recognised as host-derived; conformational masking of receptor-binding sites; and hypervariable loops that are exposed but capable to change Nab epitopes, nucleotide substitutions, insertions and deletions, modification of envelope [9, 10, 155, 203, 229], temporary epitope exposure and non-functional envelope spikes which are not expressed by mature functional spikes that may deviate the immune response from functional targets [230]. HIV envelope is heavily glycosylated, with almost 50% of the total mass consisting of poorly or non-immunogenic glycans [66, 231] shielding antibody access to the epitope. Also, the unfavourable stoichiometry with very limited number of gp160 glycoproteins per virion (between 21-42 SU molecules or 7-14 trimers per particle) likely reduce the ability of antibodies to bind simultaneously to two Env molecules (bivalent antibody binding) [232, 233].

Passive immunization of primates challenged with chimeric SHIV strains has shown that human bNAbs can protect against infection and are effective against intravenous [234, 235], oral [236] or intravaginal challenges [235, 237]. Passive administration of bNAbs has been shown to protect humanized mice and macaques against high-dose challenge with HIV or SHIVs viruses [234, 238, 239]. In addition, bNAbs have been shown to have an active role in therapy contributing to the decline of plasma viremia to undetectable

levels in HIV-1 infected humanized mice and SHIV-infected macaques, especially when combined with other bNAbs and/or ART [240, 241]. Recently, Nussenzweig et al. have demonstrated that the passive administration of HuMAb 3BNC117 was safe and effective in reducing viral load in HIV-1 infected individuals [177, 242]. Also, the potential use of bNAbs as long acting agents for the prevention and treatment of HIV-1 infection when used together with other ARVs is being explored [224]. Thus, bNAbs can be used to prevent HIV-1 infection but also disease progression.

3.2 Vaccine development

Eliminating HIV from the human population will require a successful vaccine. However, after 40 years of HIV-1 pandemic, no vaccine exists for clinical use to prevent HIV-1 infection. Candidate vaccines evaluated to date have either failed or have shown very modest efficacy. The reasons are multiple and include the remarkable high HIV-1 diversity; and the host's inability to mount antibodies to targets within conserved envelope regions that confer broad neutralization [13]. Recent insights of ways that vaccines can potentially stimulate protective T- and B-cell immunity, the identification of new targets for bNAbs, and the discovery of new mechanisms of host control of HIV-bNAbs induction offer renewed hope for the development of a well-tolerated and effective preventive HIV-1 vaccine [13, 144].

A new generation of optimized immunogens that mimic the native Env trimer such as BG505 SOSIP.664 gp140 [9, 10, 15-20], capable of expressing multiple epitopes for bNAbs including quaternary epitopes, may lead to the development of an effective HIV-1 vaccine. However to date such trimers failed to consistently elicit bNAbs capable of heterologous neutralization of tier 2 viruses (the tier of the majority of the primary HIV-1 strains) [9, 10, 12, 15-20].

Traditional vaccine strategies have focused on live attenuated viruses, whole killed viruses and protein subunits [243, 244]. These approaches have proven successful against other viruses such as the influenza virus but raise great safety concerns with regard to HIV-1 [245, 246]. More recent vaccine strategies have made use of gene delivery technologies such as plasmid DNA vaccines, recombinant viral vectors (attenuated or replication-incompetent viruses) such as adenoviruses, poxviruses and Vaccinia virus [14, 212, 247-249].

To date more than 100 clinical trials have been performed in last 30 years to evaluate HIV vaccine candidates. So far, only seven HIV-1 vaccine candidates have completed efficacy trials and none has succeeded in inducing bNAbs, as noted in the Thai Phase III clinical trial (RV144) conducted in Thailand in 2009 [13, 144]. The STEP study conducted in America, Caribbean and Australia, which comprised of a recombinant adenovirus vector that expressed HIV-1 subtype B *gag*, *pol* and *nef* genes unexpectedly was brought to an early end due to safety concerns [247] and more recently the clinical trial (HVTN 702) was stoped due to lack of efficacy [14].

In summary there are several obstacles for the development of an effective HIV vaccine, i) the high genetic diversity of the virus due to the lack of proofreading ativity of the RT, ii) the early and rapid destruction of the cellular immune system, iii) the early establishment of latent viral reservoirs in different cells and compartments, iv) the unclear immune correlates of protection, v) safety concerns regarding the use of attenuated viruses in humans, vi) the conformational changes of the Env trimer (open and closed), vii) the occlusion of neutralizing epitopes in the outer surface of the virus and also the difficulty in exposing these epitopes in vaccine candidates [13, 144].

Nevertheless, the efforts for finding an effective HIV vaccine continue, the major strategies that are being pursued include different approaches, passive immunization with

bNAbs and induction of broad nAbs with vaccine candidates, trying to augment the quality and the quantity of non-neutralizing V1V2 antibodies as seen in the RV144 immune correlates study, the development of vaccine vectors that better represent critical T-cell epitopes and the viral diversity of circulating strains [13] and mRNA vaccines that proved efficacious against certain variants of SARS-CoV-2 [250].

Aims and work plan

For more than 30 years our group and other investigators from the Faculty of Pharmacy, Universiade de Lisboa, have been studying the HIV epidemics in Angola, in terms of seroprevalence, genetic diversity, antiretroviral drug resistance surveillance and characterization of new recombinant forms and full genome sequences [4, 5, 7, 99, 251, 252]. However, in an epidemic constantly being fueled by new infections, where highly divergent forms of the virus circulate, it is important to monitor the evolution of the epidemics and continue to further characterize the impact of this rising HIV epidemics on transmission, TDR, diagnosis and disease progression.

In the 2nd chapter of this thesis the main goal was to further characterize the HIV-1 epidemic in Angola by assessing the HIV-1 diversity and prevalence of transmitted drug resistance (TDR) in Angola in 2009, five years after ART scale-up and determine the prevalent transmission dynamics.

First we extracted, amplified, sequenced and subtyped by phylogenetic analysis HIV-1 *pol* sequences from plasma samples collected in 2009, from 139 HIV-1 infected Angolan patients naïve to antiretroviral therapy residing in Luanda. To check for differences in viral subtype and diversity and in order to analyse the evolutionary trends we compared our reults with a similar survey that our group performed in 2001[7]. Resistance mutation

analysis was performed using the Stanford genotypic resistance interpretation algorithm. Mutations specifically associated with transmitted HIV-1 drug resistance were analyzed with the Calibrated Population Resistance Tool (CPR). For performing the transmission network analysis we extended the study population to all other Angolan patients for which *pol* sequences were available in the Los Alamos HIV Sequence Database (n=364), in order to avoid overestimation of relatedness between sequences due to the use of scarce data.

Women and children are the most affected populations by HIV/AIDS in Angola [3]. HIV-1 MTCT is the main route of infection among the pediatric population. In 2014 alone more than 4,500 children acquired HIV-1 through mother-to-child transmission (MTCT) [8] and is estimated that 5,200 children acquired HIV in 2020 [3]. Early infant diagnosis (EID) of HIV-1 infection and early antiretroviral treatment (ART) initiation reduces HIV-1-related mortality and long-term morbidity as well as the size of the HIV-1 reservoirs [253, 254]. Molecular-based virological testing enables determination of the infection status of HIV-1-exposed infants in the first days or weeks of life. However, the high cost of the commercially available tests and the relatively poor performance with dried blood spots (DBS) and highly diverse isolates have hampered its implementation in Angola.

In the 3rd Chapter of this thesis the main goal was to develop and validate a sensitive, simple and low-cost qualitative DBS-based HIV-1 DNA PCR for early infant diagnosis of HIV-1 infection in Angola and potentially other less resourced countries. We started by selecting a specific pair of primers targeting a highly conserved region in the integrase of HIV-1 with the capability of paring with all HIV-1 genotypes prevailing in Angola. We then constructed control plamids containing the integrase gene of all prevailing HIV-1 subtypes in Angola in order to determine the analytical sensitivity of our method. We performed serious dilutions of these control plasmids and ACH-2 cells, which contain a

single copy of HIV-1B provirus per cell, in seronegative blood and spiked filter papers. Clinical sensitivity was determined on dried blood spots (DBS) samples from 100 HIV-1-infected adult patients, 5 local samples of HIV-1 infected infants, 50 healthy volunteers and 139 HIV-1-exposed infants of the Angolan Pediatric HIV Cohort (APEHC) with serology at 18 months of life. For extracting the viral DNA from all DBS we used a simple and inexpensive method based on the cationic resin Chelex. For the detection of the viral DNA we used a nested PCR to amplify a small (194 bp) fragment in the integrase region and the human gene C-C chemokine receptor 5 (CCR5) was used as an internal control to confirm the presence and quality of genomic DNA. All samples were screened in triplicate, using HIV-1 serology results at 12 months as diagnostic reference.

Considering that vaccine effectiveness will depend on the extent to which induced antibodies will be capable to neutralize the global diversity of circulating HIV-1 variants, studying HIV-1 epidemics like the Angolan makes most sense. Angola HIV-1 epidemics is peculiar, as it is a very old epidemic driven by highly divergent forms of the virus including all M subtypes, circulating recombinant forms (CRFs), unique recombinant forms (URFs) and many untypable (U) sequences [5, 7, 96-99]. However, host immunity to these complex HIV-1 strains in Angola has never been studied.

In the 4th Chapter of this thesis the aim was to make the first detailed characterization of the neutralizing antibody response and determine viral and host factors associated with this response in Angolan patients infected with HIV-1.

We performed a cross–sectional population study with a total of 375 HIV-1 infected Angolan plasma samples, collected in 2001, 2009 and 2014. To characterize the neutralizing antibody responses we used a panel of 12 tier 2 Env-pseudotyped viruses representative of the strains circulating worldwide and used plasma samples from HIV-1 infected Angolan patients collected in 2009 (n=178) and 2014 (n=58). Briefly, Env-

pseudotyped viruses were produced by co-transfection of Env-expressing plasmids in 293T cells with PSG3.1Δenv plasmid as backbone and titrated in TZM-bl cells. A plasma dilution of 1:40 was used on an initial neutralizing screening assay and a neutralization score was determined in order to be reflective of neutralization potency and breadth. Neutralization of Env-pseudotyped viruses was measured using Tat-regulated firefly luciferase (luc) reporter gene expression to quantify reductions in virus infection in TZM-bl cells. VSV-G-pseudotyped particles were used to test neutralization specificity. A subset of plasma samples showing broad cross-neutralizing activity in the initial neutralizing screening assay were selected for determinining the 50% inhibitory dilution (ID50) (n=28 in 2009 and n=10 in 2014). We also amplified and determined by phylogenetic analysis the viral subtype and tropism in the *env* gene, specifically in the C2V3C3 region in samples collected in 2009 (n=110 sequences) and compared with previous results determined in samples collected in 2001 (n=96 sequences). In the 2014 samples, amplification was not possible as the vast majority of patients were under antiretroviral therapy. In addition to characterizing the neutralizing response and subtyping the samples, we determined the possible neutralizing epitopes of the polyclonal plasma samples by comparison with neutralizing epitopes from previously characterized bNAbs. Binding titer against constructed recombinant polypeptides, comprising the C2, V3, C3 regions of different HIV-1 subtypes, was determined in a subpopulation of patients with known neutralizing response (n=48 in 2009 and n=16 in 2014). We then explored possible associations and correlations between neutralizing response, viral subtype, binding titer to the C2V3C3 recombinant polypeptides, and demographic and infection characteristics of the HIV-1 infected Angolan population.

Chapter 2

HIV-1 Diversity, Transmission Dynamics and Primary Drug Resistance in Angola

Inês Bártolo [1,2], Suzana Zakovic [2], Francisco Martin [1,2], Claudia Palladino [1], Patrícia Carvalho [3], Ricardo Camacho [3,4], Sven Thamm [5], Sofia Clemente [6], Nuno Taveira [1,2*]

[1] Unidade dos Retrovírus e Infecções Associadas, Centro de Patogénese Molecular e Instituto de Investigacão do Medicamento (iMed.ULisboa), Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal. [2] Centro de Investigação Interdisciplinar Egas Moniz, Instituto Superior de Ciências da Saúde Egas Moniz, Monte de Caparica, Portugal. [3] Laboratório de Biologia Molecular, Centro Hospitalar Lisboa Ocidental, Hospital Egas Moniz, Lisboa, Portugal. [4] Rega Institute for Medical Research, KU Leuven, Leuven, Belgium. [5] Abbott GmbH & Co. KG, Wiesbaden, Germany. [6] Hospital da Divina Providência, Serviço de Doenças Infecciosas, Luanda, Angola

PLoS ONE December 2014 9(12): e113626.

https://doi.org/10.1371/journal.pone.0113626

Keywords: HIV-1 epidemics; Drug resistance; Transmission dynamics; Angola

Abstract

Objectives: To assess HIV-1 diversity, transmission dynamics and prevalence of transmitted drug resistance (TDR) in Angola, five years after ART scale-up.

Methods: Population sequencing of the pol gene was performed on 139 plasma samples collected in 2009 from drug-naive HIV-1 infected individuals living in Luanda. HIV-1 subtypes were determined using phylogenetic analysis. Drug resistance mutations were identified using the Calibrated Population Resistance Tool (CPR). Transmission networks were determined using phylogenetic analysis of all Angolan sequences present in the databases. Evolutionary trends were determined by comparison with a similar survey performed in 2001.

Results: 47.1% of the viruses were pure subtypes (all except B), 47.1% were recombinants and 5.8% were untypable. The prevalence of subtype A decreased significantly from 2001 to 2009 (40.0% to 10.8%, P=0.0019) while the prevalence of unique recombinant forms (URFs) increased 2-fold (40.0% to 83.1%, P<0.0001). The most frequent URFs comprised untypable sequences with subtypes H (U/H, n=7, 10.8%), A (U/A, n=6, 9.2%) and G (G/U, n=4, 6.2%). Newly identified U/H recombinants formed a highly supported monophyletic cluster suggesting a local and common origin. TDR mutation K103N was found in one (0.7%) patient (1.6% in 2001). Out of the 364 sequences sampled for transmission network analysis, 130 (35.7%) were part of a transmission network. Forty eight transmission clusters were identified; the majority (56.3%) comprised sequences sampled in 2008–2010 in Luanda which is consistent with a locally fueled epidemic. Very low genetic distance was found in 27 transmission pairs sampled in the same year, suggesting recent transmission events.

Conclusions: Transmission of drug resistant strains was still negligible in Luanda in 2009, five years after the scale-up of ART. The dominance of small and recent transmission clusters and the emergence of new URFs are consistent with a rising HIV-1 epidemics mainly driven by heterosexual transmission.

Introduction

Despite the recent decline in the number of people newly infected with HIV, around 35.3 million people were still living with HIV at the end of 2012 [29]. Sub-Saharan Africa remains severely affected by the epidemic accounting for 71% of the people living with HIV in the world and for 69.5% of the new infections [29]. Angola is a South-western African country bordered by Republic of Congo, Democratic Republic of Congo, Zambia and Namibia. According to the UNAIDS report on the global AIDS epidemic 2013 [29] the estimated HIV prevalence and new infections in adults have decreased between 2001 and 2012 in all the bordering countries of Angola. For example, in the Republic of Congo HIV prevalence decreased from 4.7% to 2.8% and the number of new infections decreased from 6,600 to 3,400. In contrast, the estimated number of adults living with HIV in Angola has increased in the same period from 110,000 to 220,000 (1.8% vs 2.3% prevalence) and the estimated number of new infections rose from 16,000 to 23,000 [29]. However a recent HIV seroprevalence survey performed on pregnant women in 36 sentinel sites in 18 provinces of Angola has found that on aggregate HIV prevalence did not vary significantly from 2004 up to 2011 (median 2.8%, range 2.7%–3.2%) although there was considerable variation across provinces [255]. Additional studies are clearly needed to better characterize the dynamics of the HIV epidemic in Angola.

HIV-1 epidemic in Angola is highly complex with all HIV-1 group M subtypes (except B), several circulant recombinant forms (CRFs), unique recombinant forms (URFs) and untypable (U) strains reported [4, 5, 7, 35, 36, 96-99]. This genetic complexity may pose a significant challenge to laboratory diagnosis and antiretroviral treatment (ART) effectiveness [15, 100], underscoring the importance of implementing regular surveys of HIV-1 diversity and its impact in this country.

Transmitted drug resistance (TDR) is a major public health problem, especially in resource-limited settings as it can determine rapid loss of effectiveness of firstline antiretroviral (ARV) regimens [206, 210]. Drug-naive individuals that acquire a virus with drug resistance mutations (DRMs) begin ART with a higher risk of virologic failure and of developing resistance [206, 256]. The absence of proper patient monitoring may lead to increased emergence and transmission of resistant strains [257]. ART has been available in Angola since 2000 for those infected with HIV who could buy ARV drugs. Since 2004, a national plan has been implemented to provide free ARV drugs to HIV-1 infected individuals using the WHO public health approach to ARV delivery [32]. At the end of 2012 the number of people on ART was 39,704 [29], 48% of the adults in need of treatment based on WHO 2010 guidelines [33]. The frequency of TDR in Angolan patients has risen from 1.6% in 2001 [7] to 16.3% in 2008–2010 [35, 36] suggesting that TDR may be an important public health problem in Angola. However, further work is required to characterize TDR level in Luanda as only a few patients living in this province have been included in previous surveys. In this study we aimed to better characterize the genetic diversity of HIV-1 and determine the prevalence of TDR in drug-naive patients in Luanda five years after ART scale-up in 2009. Additionally, to better understand the dynamics of the HIV-1 epidemics we performed the first investigation of HIV-1 transmission networks in Angola.

Materials and Methods

Study population

One hundred and thirty nine plasma samples were collected during 2009 from drug-naive HIV-1 positive individuals attending the Hospital da Divina Providência (HDP) in Luanda, Angola. This hospital is located in the Kilamba-Quiaxe district serving an

estimated population of 990,892 inhabitants, 13.4% of Luanda's population (7,395,977 habitants) [195]. Besides the patients attended at the main building, the hospital works with patients attending four health centers located in different regions of Luanda. The main criteria for patient inclusion in the study were those recommended by the WHO for this type of study [257]: confirmed diagnosis of HIV-1 infection, no pregnancy or first pregnancy (to exclude previous use of ARV for the prevention of mother-to-child transmission during delivery), no clinical diagnosis of AIDS (stage 1 and 2 WHO classification system for HIV infection) and no ART exposure. Epidemiological, clinical, and virological characterization of the patients is given in Table 1. Serological diagnosis of HIV-1 infection was done using the rapid tests Determine HIV-1/2 (Abbott) and Uni-Gold Recombigen (Trinity Biotech). The number of CD4+ T cells was determined using the ABACUS 5 Junior Hematology analyzer. Plasma viral load was determined in a subset of patients using the Abbott Real Time HIV-1 assay (Abbott Laboratories). The study was conducted according to the Declaration of Helsinki and was reviewed and approved by the Board of Directors of Hospital da Divina Providência (Luanda, Angola) and the National Ethics Committee of Angola. Written informed consent was obtained from all participants. The study was verbally explained to the patients before they signed the written consent. For the transmission network study, to avoid overestimation of relatedness between sequences due to the use of scarce data [258] we extended the study population to all other Angolan patients for which *pol* sequences were available in the Los Alamos HIV Sequence Database [259]. Hence, in addition to our present sequences we used 226 Angolan *pol* sequences collected from the Los Alamos HIV Sequence Database, counting in total 364 sequences. These sequences were derived from samples collected in 1993 and 2001 (n=86) [5, 7, 98, 99], 2009 (n=39) [259] and 2008-2010 (n=101) [35, 36]. Most sequences (n=64, 28.3%) were obtained from patients attending

different medical facilities in and near Luanda (including Hospital Sanatório de Luanda, Laboratório da Força Aérea Nacional Angolana, Hospital Militar Principal, Clínica Sagrada Esperança, Centro Nacional de Sangue and São Lucas Medical Center in the village Kifangondo). Remaining sequences were obtained from patients attending Hospital services in Cabinda (n==20, 8.8%), Namibe (n=4, 1.8%), Benguela (n=4, 1.8%), Zaire (n=3, 1.3%), Cuanza Norte, Bengo and Huila (n=1, 0.4%, each), and from patients living in Central (n=7, 3.1%), North (n=3, 1.3%) and South (n=2, 0.9%) of Angola. Origin of 116 (51.3%) patients was not available. Fourteen patients were on ART. Because HIV transmission networks are mainly confined to a country [260] no sequences outside Angola were included in the present study.

Table 1. Epidemiological, clinical, and virological characteristics of HIV-1 Angolan patients analyzed in this study.

Variables	**Samples**
Patients [n (%)]	139 (100)
Age [mean (SD), years]	36 (14) (n=139)
Gender [n (%)]	
Male	50 (36.0)
Female	87 (62.6)
Unknown	2 (1.4)
Transmission route [n (%)]	
Heterosexual	120 (86.3)
Vertical	14 (10.1)
Blood transfusion	2 (1.4)
Unknown	3 (2.2)
CD4 [mean (range), cells/ml]	240.5 (1-1914) (n=106)
HIV RNA [mean, (SD) log10, copies/ml]	5.1 (1.0) (n=86)
Pure subtype [n (%)]	65 (47.1) (n=138)
Untypable	8 (5.8) (n=138)
Recombinants [n (%)]	65 (47.1) (n=138)

doi:10.1371/journal.pone.0113626.t001

Viral RNA extraction, PCR amplification and sequencing

Viral RNA was extracted from 140 ml plasma using QIAmp Viral RNA Mini Kit (Qiagen). RT-PCR was performed with Titan One Tube RT-PCR System (Roche). Nested PCR was done using an in-house method described elsewhere [5, 7, 98, 99]. Thermal cycling conditions for PCR and primers sequence and position were previously described [5, 7, 98, 99]. DNA sequences were obtained with Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems) and an automated sequencer (3100-Avant Genetic Analyzer, Applied Biosystems).

Phylogenetic and recombination analysis

Sequences were aligned with reference strains collected from the Los Alamos HIV Sequence Database [259] using ClustalX [261]. Maximum-likelihood (ML) phylogenetic analyses [262] were performed using the best-fit model of molecular evolution estimated by Modeltest v3.7 under the Akaike information criterion [263]. ML trees were inferred, with program PhyML using Seaview software [264]. To find the ML tree, an iterative heuristic method combining two different tree rearrangement methods was used: nearest neighbor interchange (NNI) and subtree pruning and regrafting (SPR). The reliability of the obtained topology was estimated with the approximate likelihood-ratio test (aLRT) [264]. Recombination analysis was performed by bootscanning using SimPlot [265]. For transmission network analysis, protease (PR) and reverse transcriptase (RT) sequences were concatenated in SeaView [264]. Sequences were aligned with ClustalX [261] and manually edited in MEGA [211]. Codons associated to drug resistance were stripped from the alignment to exclude convergent evolution [206]. Phylogenetic analysis was performed using a single alignment with all subtypes and CRFs included as previously described [260, 266]. Best-fit model was chosen with Modeltest v3.7 under the Akaike information criterion [263]. ML tree was constructed in PhyML incorporated in Bioportal server [267]. Reliability of the tree was assessed using bootstrap replication (1000

replicates). Genetic distance for clusters with bootstrap support ≥90% was measured in MEGA [211]. Sequences with genetic distance, 0.05 (range 0.000–0.049) substitutions per site were considered genetically related and patients were assumed to belong to the same transmission cluster [268]. Automatic cluster detection was used to confirm the transmission clusters that were initially detected. This was performed with PhyloPart program based on Approximate ML tree obtained with FastTree program [269]. Clusters were detected through the depth-first search of reliable nodes with patristic distance under 1st percentile threshold of the whole tree distance [199]. According to this threshold, transmission clusters were recognized with patristic distance, 0.07 substitutions per site and node reliability ≥90%.

Resistance mutation analysis

Resistance mutation analysis was performed using the Stanford genotypic resistance interpretation algorithm [270]. Mutations specifically associated with transmitted HIV-1 drug resistance were analyzed with the Calibrated Population Resistance Tool (CPR) (http://cpr.stanford.edu) [271, 272].

Statistical analysis

Statistical analysis was performed with GraphPad Prism version 5.00 for Windows, (GraphPad Software). The Spearman rank test and linear regression analysis were used to quantify the magnitude and direction of the correlation between viral load and CD4+ T cells. The Mann-Whitney U test was used to compare independent groups. The frequencies of drug resistance mutations of Angolan viruses were compared with those available at the Stanford HIV Drug Resistance Database [270] for the same subtypes using Fisher's exact test. P-values <0.05 were considered significant.

GenBank accession numbers

Sequences have been assigned the following GenBank accession numbers: KF853612–KF853892.

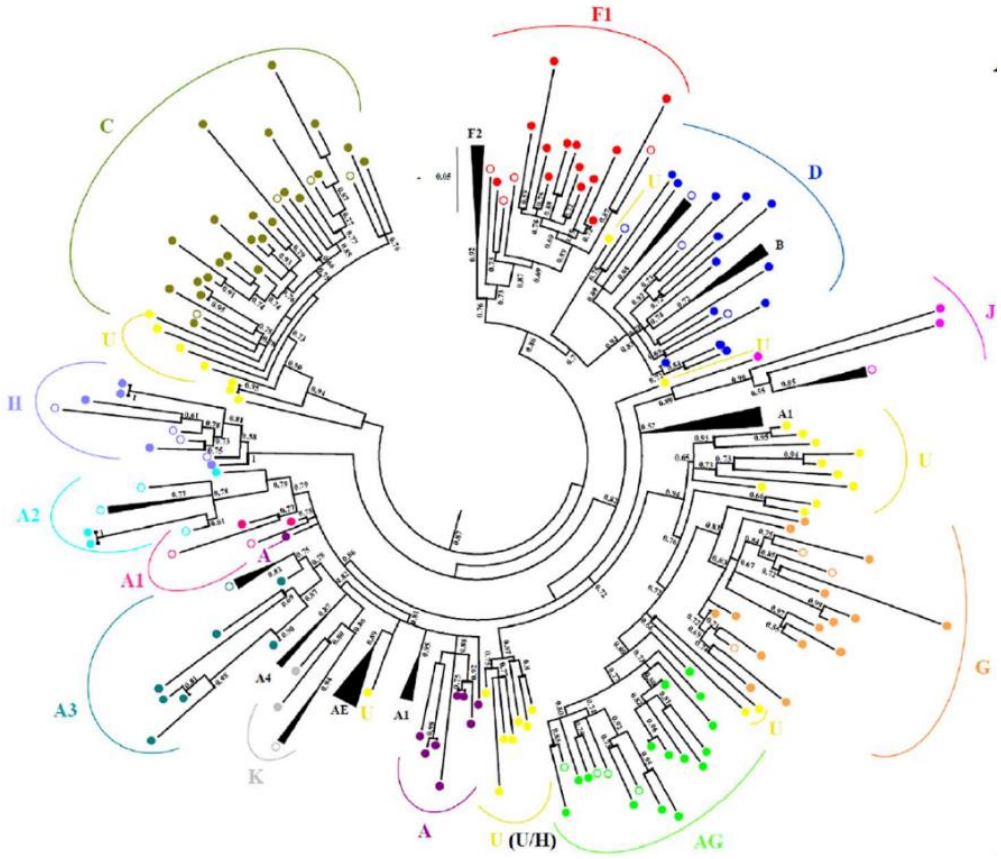
Results

HIV-1 genetic diversity

Plasma samples were obtained from 139 HIV-1 infected individuals. The mean age of the patients was 36 years (SD, 14) and most (62.6%) were women (Table 1). The main route of transmission was heterosexual contact (86.3%). As expected, plasma viral load was high in most patients (mean 5.1 $\log_{10}$ copies/ml) and the number of CD4+ T cells was low (mean, 240.5 cells/ml). Viral load and CD4+ T cells were negatively correlated (n=73, Spearman r=-0.3319, P=0.0041). Sequencing and phylogenetic analysis of the PR region was completed successfully for 139 (100%) patients; RT sequences were also obtained for all but one of these patients (n=138, 99.3%). Phylogenetic analysis showed that all viruses belonged to HIV-1 group M (Figure 1 A and B). Out of the 138 isolates for which there was PR and RT sequences, 65 (47.1%) sequences were non-recombinant and 65 (47.1%) were recombinant, of which 11 (16.9%) were CRF02_AG, and 8 (5.8%) were untypable (U). The following pure subtypes and sub-subtypes were identified: A (n=3, 4.6%), A1 (n=1, 1.5%), A2 (n=3, 4.6%), C (n=24, 36.9%), D (n=9, 13.8%), F1 (n=13, 20.0%), G (n=7, 10.8%), H (n=3, 4.6%) and J (n=2, 3.2%). Thirty different patterns of recombination were found. Subtypes A1, A2, A3, C, D, F1, G, H, J and K, and CRF02_AG and U sequences were involved in recombination events. Most of the recombinants (n=54, 83.1%) were URFs; in almost half of the recombinants (n=31, 47.7%) one of the regions was untypable. The most frequent URFs comprised untypable sequences with subtypes H (U/H, n=7, 10.8%), A (U/A, n=6, 9.2%) and G (G/U, n=4, 6.2%). The U/H recombinants had a mean genetic distance of 0.062 substitutions per site

and formed a highly supported monophyletic cluster in both genomic regions indicating that they share the same origin (Figure 1A and B). A similar U/H cluster has been described recently in Angola but the Province of origin of the patients in this cluster has not been disclosed [35, 36]. Phylogenetic analyses revealed a close evolutionary relatedness of all U/H sequences suggesting that the origin of this emerging URF is Luanda (Figure 2). The analyses of the evolution of HIV-1 genetic diversity in Luanda from 2001 to 2009, showed that there was a significant decrease in the prevalence of subtype A (3.7 fold difference, P=0.0019) which was replaced by subtype C as the dominant subtype (Table 2). Moreover, the percentage of URFs increased more than twice during the same period (P<0.0001).

A)



B)

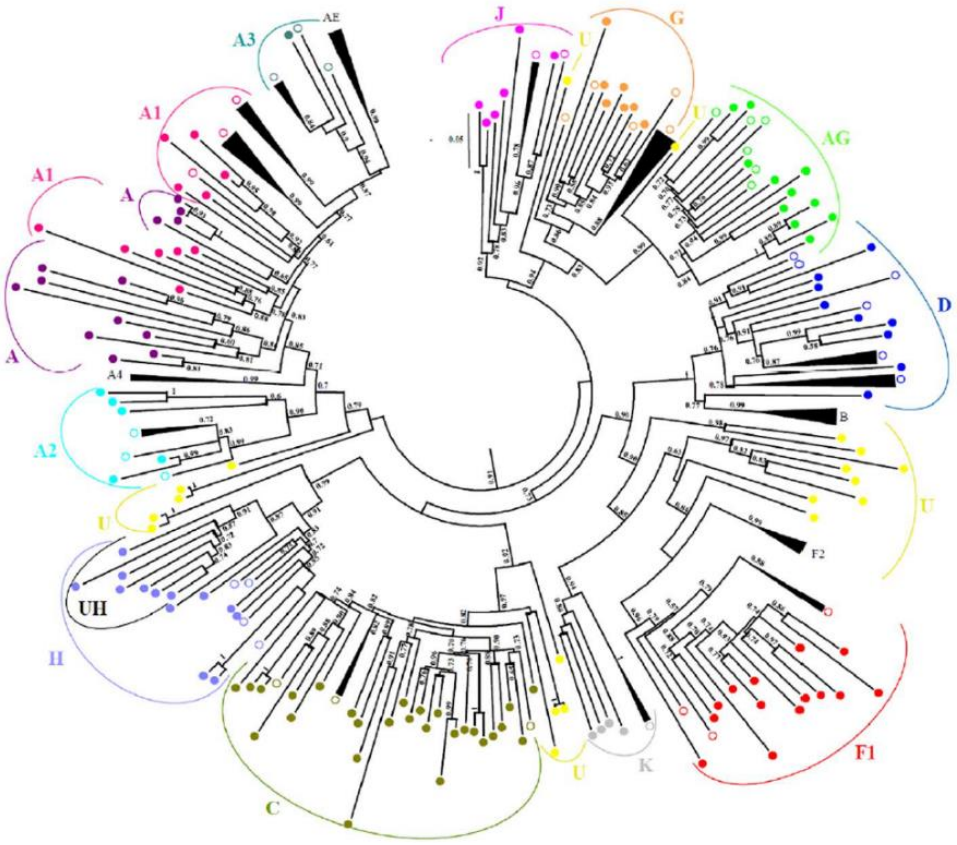


Figure 1. Genetic subtypes and evolutionary relationships of the viruses sequenced in this study. Maximum likelihood phylogenetic trees of PR (A) and RT (B) regions were constructed with reference sequences from all HIV-1 subtypes and sub-subtypes (empty circles) and with the Angolan sequences (filled circles). In each tree, the aLRT values supporting the internal branches defining a subtype or a sub-subtype are shown. The scale represents number of base substitutions per site.

Drug resistance mutations and other polymorphisms

There were no major mutations associated with resistance to protease inhibitors (PIs). The minor resistance mutations L10I and L10V, associated with resistance to most of the PIs when present with other mutations [273, 274], were found in 15.7% and 17.6% of the isolates, respectively (Table S1). This is higher than the frequencies previously described for untreated patients (6.8% and 8.2%) [270]. V11I, associated with resistance to darunavir [274, 275], was detected in 7.7% of subtype F isolates and 13.3% of

CRF02_AG isolates. This frequency is significantly higher than that found in sequences of the same subtypes available in the Stanford Database [270] (Table S1). K20I was found in almost all G and CRF02_AG isolates and is a natural polymorphism of both genetic forms [276]. K20V was found in one patient harboring a CRF02_AG virus. K20I/V codons are non-polymorphic in most subtypes [136, 270]. They appear to be selected most commonly by nelfinavir and to reduce its susceptibility [277, 278]. A71T was found in one patient infected with a subtype C virus and A71V was found in two patients infect with subtype D. The latter mutation has never been described for subtype D [270]. A71T/V are polymorphisms that occur in 2–3% of untreated individuals but the frequency increases in patients receiving PIs [279-281]. In subtype D isolates the frequency of polymorphism I13V was significantly lower than that found in sequences of the same subtype available in the Stanford database [270] (Table S1). Similar findings were obtained for K14R, E35D and R57K in subtype A and for L89M in subtype G. For all the other polymorphisms the frequencies found in the Angolan isolates were significantly higher when compared with the worldwide sequences available from untreated patients [270]. Subtype F isolates were the most polymorphic followed by subtype A and CRF02_AG.

In the RT, we detected the K103N mutation in one patient (1/138, 0.7%) that was infected with a subtype G virus (Table S2). This mutation confers high-level resistance to nevirapine, delavirdine and efavirenz [270]. In subtype F the frequency of polymorphisms A272P and I326V was significantly lower than that found in sequences of the same subtype available in the Stanford Database [270]. Similar findings were obtained for K11T, D123AS, K173S, Q174K, V179I, Q207E, R211S, T286A, E312D and G335DE in subtype A, I293V, I329L and G335D in subtype C and T200A, V292I and G335D in CRF02_AG. For all the other polymorphisms the frequencies found in the Angolan

isolates were significantly higher when compared with the sequences available from untreated patients worldwide. Compared with the 2001 survey there was a 2.3-fold decrease in the prevalence of TDR (1.6% vs 0.7%).

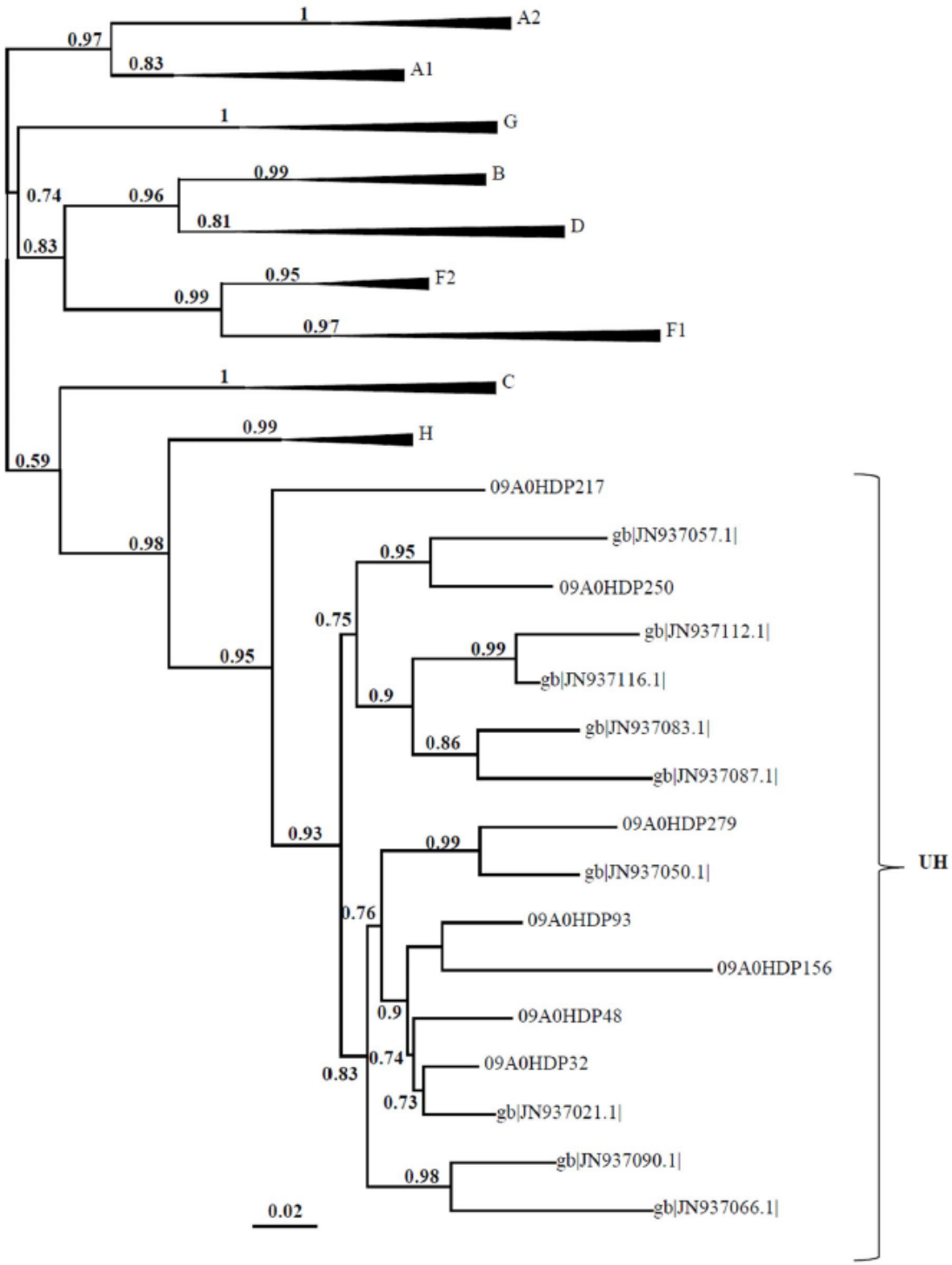


Figure 2. Evolutionary relationships of the U/H recombinants. Sequences of U/H recombinants (named with 09AOHDP) were aligned with those of a previous study (named gb, GenBank) [35, 36]. The aLRT values supporting the internal branches defining a subtype or a sub-subtype are shown. The scale represents number of base substitutions per site.

Table 2. Evolution of HIV-1 genetic diversity in Luanda from 2001 to 2009

HIV-1 genetic forms	**2001** n (%)	**2009** n (%)	**P value[a]**
Pure subtypes	46/86 (53.5)	65/138 (47.1)	0.6771
Recombinants	40/86 (46.5)	65/138 (47.1)	
First generation CRFs	5/40 (16)	11/65 (16.9)	0.5902
URFs	16/40 (40)	54/65 (83.1)	<0.0001
Prevalence of subtype A	17/46 (40)	7/65 (10.8)	0.0019
Prevalence of subtype C	15/46 (33)	24/65 (36.9)	0.6899
Prevalence of subtype F	6/46 (13)	13/65 (20)	0.4453

a Fisher's exact test.

Transmission network analysis

To better characterize the dynamics of the current HIV-1 epidemics in Angola and assist in the implementation of more focused prevention strategies we performed a transmission network analysis. The majority of the 364 sequences included in this sub-study were from patients residing in Luanda (n=202, 55.5%); the remaining sequences were derived from patients from seven other provinces of Angola (n=46, 12.6%) or their origin was unknown (n=116, 31.9%). Forty eight transmission clusters were identified comprising 130 patient sequences (35.7% of the sampled patients) (Figure 3); more than half of these (52.3%) reported being heterosexual (Table S3). Consistent with this, small clusters comprising two closely related strains were dominant (n=33, 68%). Only three large transmission chains, each comprising seven individuals, were found. As expected, most sequences (n=98, 75.4%) in the transmission clusters were sampled in 2008, 2009 and 2010 and most clusters (N=27, 56.3%) comprised only sequences sampled in these dates (Table

S3). All but one sequence (from Namibe) with available information on the origin were from Luanda indicating that the current epidemic is mostly sustained by local transmission. Notably, 14 (29.2%) transmission clusters comprised 2001 and 2008–2010 sequences suggesting a long standing presence of some of the more current viruses circulating in Luanda. In 3 (21.4%) of these clusters 2001 sequences originated from Cabinda, Benguela and Lunda Norte Provinces. Six clusters (12.5%) comprised only sequences sampled in 2001 and one cluster (2.1%) comprised only sequences sampled in Cabinda in 1993 (Table S3). These clusters were considered uninformative for the current study.

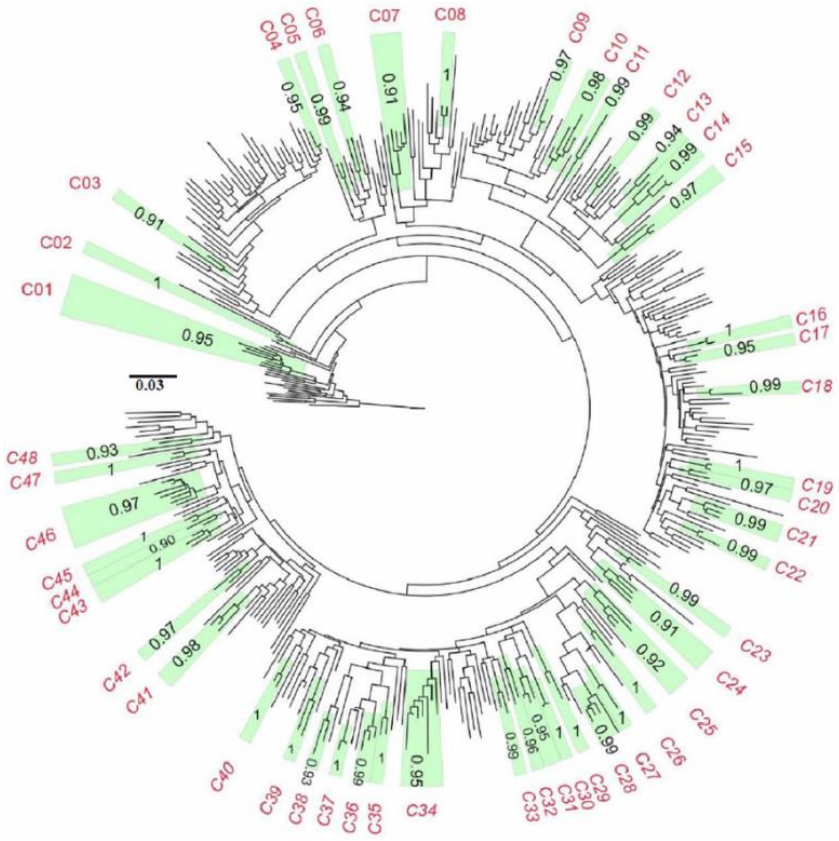


Figure 3. Transmission cluster analysis. Maximum likelihood tree with the 48 transmission clusters colored in green. Maximum likelihood tree was constructed in PhyML. Reliability of the tree was assessed using bootstrap resampling (1000 replicates). Bootstrap values and cluster number are indicated in each cluster. A bootstrap value of 0.7 (70%) or greater indicate significant support for the clusters. The scale represents number of base substitutions per site.

Finally, in 27 transmission pairs sampled in the same year (including pairwise clusters within larger transmission chains) very low genetic distance was found (median 0.005 nucleotide substitutions per site; range 0.000–0.007) suggesting recent transmission events [282]. Potential sample mix-up for pairwise clusters with 0.000 genetic distances was excluded based on visual inspection of pairwise alignments and origin of the samples. The main features of the individuals included in the transmission networks were no different compared to individuals outside the transmission networks (Table 3).

Table 3. Demographic, immunologic and virologic characteristics of HIV-1 patients included in the transmission networks and patients outside of the networks.

Variables	**Total**	**Transmission networks**	**Remaining patients**	**P value**
Patients [n (%)]	364 (100)	130 (35.71)	234 (64.28)	-
Gender [n (%)]				
Male	100 (27.47)	44 (33.85)	56 (23.93)	
Female	142 (39)	47 (36.15)	95 (40.6)	0.1250[a]
Unknown	122 (33.52)	39 (30)	83 (35.47)	
Transmission route [n (%)]				
Heterosexual	194 (53.3)	68 (52.3)	126 (53.85)	
Vertical	15 (4.12)	7 (5.38)	8 (3.42)	0.8208[a]
Others	10 (2.74)	4 (3.08)	6 (2.56)	
Unknown	145 (39.83)	51 (39.23)	94 (40.17)	
Origin [n (%)]				
Luanda	200 (54.94)	78 (60)	122 (52.14)	
Cabinda	20 (5.49)	4 (3.08)	16 (6.84)	0.2850[a]
Others	28 (7.29)	8 (6.15)	20 (8.55)	
Unknown	116 (31.87)	40 (30.77)	76 (32.48)	
CD4 [mean (range), cells/µL		335.39 (1-1914) (n=41)	300.12 (12-790) (n=78)	0.7309[b]
HIV RNA [mean, (range) $log_{10,}$ copies/mL]		5.15 (1.94-7) (n=33)	5.06 (1.6-6.68) (n=52)	0.6391[b]

a – Chi-square test; b – Mann-Whitney test;

Discussion

We assessed HIV-1 diversity, transmission dynamics and prevalence of TDR in Luanda in 2009, five years after the scale-up of ART and compared these data with our previous survey performed in 2001 [5, 7, 98, 99]. Individuals included in this study had a low CD4 count which was directly related with high viral load. These features are consistent with the reported absence of ART [283, 284]. Like in 2001, no major PIs resistance mutations were found in the study population which is consistent with the fact that first-line regimens used in Angola do not include PIs [285]. However, some minor mutations in the PR and many unusual polymorphisms were detected suggesting that some Angolan isolates might have a low genetic barrier for resistance to some PIs [270]. The K103N mutation, which confers high-level resistance to nevirapine, delavirdine and efavirenz [270], was found in one patient accounting for a 0.7% prevalence rate of TDR which is 2.3 fold lower compared to the 2001 survey [7]. This residual TDR prevalence is similar to that of several African countries that also use the public health approach to ART [200, 201, 286-288] and suggests that the most common first-line ARV regimens will be effective in this population. Similar to previous studies, the HIV-1 epidemic in Luanda in 2009 was highly

complex being characterized by the presence of almost all subtypes (A, C, D, F, G, H and J; 47.1%), complex recombinants viruses (47.1%) and untypable (5.8%) strains [4, 5, 7, 35, 36, 96-99, 289]. A high number of our sequences fall at basal positions on the phylogenetic trees (pre-subtype branches) which is consistent with the long standing presence of HIV-1 in Angola [5, 7, 98, 99]. In addition, some strains from Angola have little organized substructure and form weaker clusters within phylogenetic trees than the global reference sequences, not allowing a clear distinction between subtypes. As a consequence, the current global subtype classification may not reflect the extent of diversity in this region [290]. The prevailing subtype in 2009 in Luanda was subtype C

(36.9%) followed by sub-subtype F1 (20.0%) whereas in 2001 it was subtype A followed by subtype C [5, 7, 98, 99]. The significant decrease in the prevalence of subtype A and increase in subtype C observed in 2009 could be explained by the increasing predominance of subtype C in the bordering countries, namely in the south region of Democratic Republic of Congo [202] and in Zambia [35, 36, 200, 207]. As in 2001, almost half of our sequences were recombinant comprising all group M subtypes as well as CRF02_AG and U sequences. The frequency of CRFs did not change between 2001 and 2009, but the frequency of URFs more than doubled in the same period.

This is on contrast to the global and regional distribution of HIV-1 genetic forms between 2000 and 2007, where there was a notable increase in the proportion of CRFs and a decrease in URFs [203]. Importantly, the results indicate that the Angolan HIV-1 epidemic is still increasing in genetic complexity and suggest high rates of co-infection and/or superinfection [208] which is consistent with an increasing HIV-1 incidence and prevalence [29, 255, 291]. The most common URF was U/H found in seven strains (10.8% of the recombinants and 5.1% of the total population). This new recombinant strain was found in unrelated patients and its sequences clustered in a highly supported monophyletic group suggesting that it was originally produced in Luanda. The close relationship with U/H sequences recently reported elsewhere in Angola [35, 36], indicates that this new recombinant is already established in Angola. Sequencing the full-length genome of this recombinant strain will be needed to determine if this is a new CRF.

A large number of transmission clusters were identified in this study which included 35.7% of the analyzed samples. This is not uncommon in HIV epidemics as within a smaller population or even globally HIV infected individuals are often part of wide transmission networks [195, 260, 292]. Small clusters mostly comprising two sequences were dominant over large clusters which is consistent with heterosexual contact being the

main route of transmission reported in most patients [293]. While most transmission clusters comprised only sequences from Luanda and were therefore consistent with a locally propelled epidemic, some clusters contained sequences from Luanda and from other locations in Angola consistent with a more complex origin and transmission dynamics going well beyond the borders of the capital city. Finally, based on high sequence homology between patients in transmission clusters, a large number of potential recent infections were inferred. Overall, the results are consistent with a rising HIV-1 epidemic in Luanda [29, 255, 291]. Further surveys are required to obtain a clearer picture of the dynamics of the current HIV-1 epidemics at the national level. In conclusion, transmission of drug resistant strains was still negligible in Luanda in 2009, five years after the scale-up of ART. The dominance of small and recent transmission clusters and the emergence of new URFs are consistent with a rising HIV-1 epidemics mainly driven by heterosexual transmission.

Supporting Information

Table S1. Minor mutations and natural polymorphisms detected in the PR of drug-naive patients from Luanda.

Codons	Patients (n)												P value (Angola vs World)[#]					
	Angola						World [*]											
	A (n =21)	C (n = 25)	D (n = 13)	F1 (n = 13)	G (n = 15)	AG (n = 15)	A (n = 3736)	C (n = 7638)	D (n = 1373)	F (n = 817)	G (n = 1208)	AG (n = 3183)	A	C	D	F	G	AG
L10I	6	1	3	3	2	1	411	122	47	77	212	166	**0.0132**	0.875	**0.0098**	0.0992	1	0.7381
L10V	4	-	-	9	2	2	250	-	-	250	35	267	**0.0009**	-	-	**< 0.0001**	0.073	0.8242
V11I	-	-	-	1		2	-	-	-	0	-	73	-	-	-	**0.0133**	-	**0.0496**
T12A	-	1	-	-	2	-	-	107	-	-	37	-	-	0.8019	-	-	0.0801	-
T12I	-	-	-	1		-	-	-	-	33		-	-	-	-	0.4218		-
T12K	-	-	-	1	2	-	-	-	-	25	36	-	-	-	-	0.3409	0.0806	-
T12M	-	-	-	1		-	-	-	-	8	-	-	-	-	-	0.1331	**0.0001**	-
T12N	-	-	-	-	2	-	-	-	-	-	0	-	-	-	-	-	-	-
T12P	-	-	1	-	-	1	-	-	21	-	-	0	-	-	0.1885	-	-	**< 0.0001**
T12S	-	18	2	-	-	1	-	5041	32	-	-	0	-	0.6738	**0.0385**	-	-	**< 0.0001**
I13A	-	-	-	-	1	1	-	-	-	-	0	225	-	-	-	-	**0.0123**	0.6568
I13V	17	4	2	11	14	14	3250	298	919	123	1184	2928	0.6209	**0.0096**	**0.0002**	**< 0.0001**	0.2677	0.7753
K14E	-	2	-	-	-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-	-
K14R	2	2	1	1	12	9	1569	840	84	98	773	1942	**0.0053**	0.8743	0.5624	1	0.2805	0.9809
I15L	-	-	-	-	-	1	-	-	-	-	-	64	-	-	-	-	-	0.7204
I15V	7	19	-	8	4	3	1137	6492	-	621	181	226	0.96	0.3289	-	0.3232	0.2642	0.1524
G16A	1	-	-	-	-	-	60	-	-	-	-	-	0.7830	-	-	-	-	-
G16E	9	4	-	2	-	6	523	458	-	131	-	700	**0.0005**	0.0935	-	1	-	0.1721
L19I	1	11	3	1	1	-	112	4736	84	47	64		0.8661	0.1	**0.0432**	0.5136	0.5613	-
L19P	-	-	-	1	-	3	-	-	-	0	-	477	-	-	-	**0.0145**	-	0.857
L19T	-	6	-	-	-	-	-	1069	-	-	-	-	-	-	-	-	-	-
L19V	2	8	2	-	-		0	840	0	-	-	-	**< 0.0001**	**0.0025**	**< 0.0001**	-	-	-
K20I	1	-	-	-	15	14	164	-	-	-	1184	3024	0.6928	-	-	-	1	0.7662

K20M	1	-	-	1	-	-	0	-	-	15	-	-	**< 0.0001**	-	-	0.225	-	-
K20R	6	5	2	5	-	-	523	1375	136	245	-	-	0.1096	0.9991	0.6698	0.5463	-	-
K20V	-	-	-	-	-	1	-	-		-	-	35	-	-	-	-	-	0.4166
L33F	-	-	1	-	-		-	-	33	-	-		-	-	**0.0094**	-	-	-
E35D	13	4	5	12	6	3	3437	1991	247	735	628	732	**< 0.0001**	0.3592	0.0697	1	0.4393	0.9742
E35G	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0123**	-
E35N	1	-	-	-	-	1	60	-	-	-	-	80	0.8806	-	-	-	-	0.8432
M36I	21	21	10	13	15	14	3661	6492	851	784	1196	3119	0.8994	0.8877	0.3914	1	1	0.7204
M36L	-	3	-	-	-	-	-	244	-	-	-	-	-	0.0547	-	-	-	-
N37A	-	1	-	-	-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-	-
N37D	2	-	4	1	-	-	560	-	107	60	-	-	0.694	-	**0.0158**	1	-	-
N37E	2	-	-	-	-	-	45	-	-	-	-	-	**0.0148**	-	-	-	-	-
N37S	-	2	-	1	3	6	-	649	-	0	29	207	-	0.787	-	**0.0157**	**0.006**	**< 0.0001**
P39A	1	-	-	-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-	-	-
P39Q	-	1	-	-	-	-	-	0	-	-	-	-	-	-	**< 0.0001**	-	-	-
P39S	-	1	1	-	-	2	-	-	44	-	-	70	-	0.7191	0.3501	-	-	**0.0426**
P39T	-	-	1	-	-	-	-	-	15	-	-	-	-		0.1407	-	-	-
R41K	15	21	13	13	15	15	3587	6416	1318	744	1160	2960	**< 0.0001**	0.7848	1	0.6186	1	0.5790
R57K	3	-	-	11	-	-	1718	-	-	686	-	-	**0.0071**	-	-	1	-	-
Q61H	-	-	-	2	1	-	-	-	-	12	0	-	-	-	-	**0.0185**	**0.0123**	-
Q61N	-	-	1	-	-	-	-	-	29	-	-	-	-	-	0.2441		-	-
I62V	2	-	5	3	-	-	168	-	206	98	-	-	0.7208	-	**0.0484**	0.2032	-	-
L63N		-	-	-	2	-	-	-	-	-	0	-	-	-	-	-	**0.0001**	-
L63P	2	-	3	-	2	3	411	-	426	-	230	509	0.8749	0.6529	0.7648	-	0.7495	0.9446
L63Q	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**< 0.0001**
L63S	-	2	5	-	-	-	-	359	80	-	-	-	-	0.7606	**0.0007**	-	-	-
L63T	4	7	2	-	-	-	163	2521	58	-	-	-	**0.0041**	0.2658	0.1058	-	-	-
L63V	4	5	-	12	2	1	135	556	-	147	0	0	**0.0016**	**0.04**	-	**< 0.0001**	**0.0001**	**< 0.0001**
I64L	-	-	-	-	1	2	-	-	-	-	41	194	-	-	-	-	0.4098	0.531
I64M	-	1	1	-	-	-	-	84	0	-	-	-	-	**< 0.0001**	**0.0094**	-	-	-
E65D	-	-	1	11	-	1	-	-	165	196	-	0	-	-	1	**< 0.0001**		**< 0.0001**
C67E	-	-	-	-	4	-	-	-	-	-	411	-	-	-	-	-	0.7846	-
C67G	-	-	-	-	2	-	-	-	-	-	18	-	-	-	-	-	**0.0235**	-
C67N	-	-	1	-		-	-	-	0	-	-	-	-	-	**0.0094**	-	-	-
C67S	-	-	-	-	2	-	-	-	-	-	145	-	-	-	-	-	0.6994	-
C67Y	-	3	-	1	-	-	-	0	-	0	-	-	-	**< 0.0001**	-	**0.0157**	-	-
H69K	19	-	1	-	1148	15	3624	-	47	-	15	3088	0.271	-	0.3689	-	1	0.9339

H69Q	1	-	2	1	-	-	71	-	77	74	-	-	0.8763	-	0.1668	1	-	-
H69R	1	-		-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-	-	-
H69Y	-	-	1	3	-	-	-	-	220	35	-	-	-	-	0.7052	**0.0184**	-	-
K70Q	-	-	-	1	-	-	-	-		0	-	-	-	-	-	**0.0157**	-	-
K70R	4	-	1	2	1	3	273	-	84	27	110	350	0.1022	-	0.5624	0.0725	1	0.5743
A71P	-	1		-	-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-	-
A71T	-	1		-	-	-	-	497	-	-	-	-	-	0.9193	-	-	-	-
A71V	-	-	2	-	-	-	-	-	0	-	-	-	-	-	**< 0.0001**	-	-	-
I72M	-	-		-	-	1	-	-	-	-	-	73	-	-	-	-	-	0.7924
I72T	-	-	1	4	-	-	-	-	62	188	-	-	-	-	0.4553	0.5114	-	-
I72V	-	-	-	-	1	-	-	-	-	-	19	-	-	-	-	-	0.2202	-
V77L	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0123**
V77I	-	-	-	1	-	-	-	-	-	61	-	-	-	-	-	1	-	-
V82I	-	-	-	-	14	2	-	-	-		6798	102	-	-	-	-	0.9024	0.1397
L89F	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0157**	-	-
L89I	2	-	-	-	2	1	0	-	-	-	22	53	**< 0.0001**	-	-	-	**0.0332**	0.4941
L89M	19	23	-	12	13	14	3661	6340	-	400	1184	1148	0.0985	0.3527	-	**0.0016**	**0.0385**	0.5379
T91V	2	-	-	-	-	-	64	-	-	-	-	-	0.0596	-	-	-	-	-
I93L	2	24	9	-	-	-	897	7333	47	-	-	-	0.1953	0.6102	**< 0.0001**	-	-	-

*According to the Stanford HIV Drug Resistance Database; #Fisher exact test; Common accessory mutations are in bold letters; Common highly polymorphic compensatory mutations are in bold underline letters.

doi:10.1371/journal.pone.0113626.s001

Table S2. Drug resistance mutations and natural polymorphisms detected in the RT of drug-naive patients from Luanda

Codons	Patients (%)																	
	Angola						World						P value (Angola vs World)[#+]					
	A	C	D	F1	G	AG	A	C	D	F	G	AG						
	(n = 28)	(n = 26)	(n = 9)	(n = 14)	(n = 7)	(n = 11)	(n = 3108)	(n = 6794)	(n = 1095)	(n = 413)	(n = 1197)	(n = 2269)	A	C	D	F	G	AG
K103N	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-
P4S	-	2	-	-	-	-	-	272	-	-	-	-	-	0.6486	-		-	-
P4T	-	1	-	-	1	-	-	0	-	-	0	-	-	**<0.0001**	-	**0.0058**	-	-
P4QK	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
I5V	2	-	-	1	-	-	59	-	-	0	-	-	0.1891	-	-	**0.0328**	-	-
E6D		-	1	1	-	-	-	-	19	31	-	-	-	-	0.1366	1	-	-
E6K	1	1	-	-	-	1	106	150	-	-	-	30	0.634	0.8065	-	-	-	0.3604
E6N	1	-	-	-	-	-	53	-	-	-	-	-	0.9792		-	-	-	-
K11D	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
K11Q	-	-	-	-	1	-	-	-	-	-	12	-	-	-	-	-	0.0734	
K11R	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
K11S	4	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
K11T	5	1	-	-	-	-	1274	0	-	-	-	-	0.0222	**<0.0001**	-	-	-	-
K20R	4	1	3	5	1	1	591	462	76	50	168	132	0.694	0.836	**0.0152**	**0.0238**	1	0.855
V21A	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
V21I	-	-	-	1	4	-	-	-	-	5	65	-		-	-	0.1823	**0.0003**	-
T27S	1	2	-	-	1	-	31	197	-		0	-	0.6857	0.4144	-	-	**0.0058**	-
V35K	-	4	-	2	-	-	-	211	-	10	-	-	-	0.0026	-	0.0546	-	-
V35I	-	-	1	1	-	-	-		59	1	-	-	-	-	0.3617	0.0674	-	-
V35T	28	22	7	11	7	11	2890	6318	986	347	1125	2042	0.2803	0.1995	0.5697	0.4815	-	-
E36A	1	22	-	-	-	-	93	4824	-	-	-	-	0.7056	0.1899	-	-	1	0.548
E36V	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-

T39A	1	-	2	1	2	5	162	-	80	83	60	60	0.9695	-	0.1141	0.3208	0.0464	**<0.0001**
T39D	1	3	-	-	-	-	0	1495	-	-	-	-	**<0.0001**	**0.0004**	-	-	-	-
T39E	-	22	-	-	-	-	-	4756	-	-	-	-	-	0.1588	-	-	-	-
T39K	9	-	-	2	1	-	870	-	-	16	83	-	0.7829	-	-	0.1131	0.398	-
T39L	1	-	3	3	1	-	47	-	22	28	49	-	0.9121	-	**0.0005**	0.073	0.2574	-
T39M	3	-	-	-	-	-	112	-	-	-	-	-	0.1368	-	-	-	-	-
T39N	4	-	-	-	-	-	177	-	-	-	-	-	0.1521	-	-	-	-	-
T39P	-	-	-	-	1	-	-	-	-	0	-	-	-	-	-	0.0058	-	-
T39R	-	-	-	1	-	-	-	-	-	7	-	-	-	-	-	0.2358	-	-
T39S	1	-	-	1	-	-	34	-	-	5	-	-	0.7348	-	-	0.1823	-	-
E40A	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
E40D	7	4	5	1	4	5	373	482	285	83	239	227	0.0707	0.084	0.0332	0.3208	**0.034**	**<0.0001**
K43E	3	-	-	1	-	-	183	-	-	0	-	-	0.5	-	-	**0.035**	-	-
K43Q	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0058**	-
K43R	-	1	-	-	2	-	-	50	-	-	204	-	-	0.4858	-	-	0.3427	-
S48T	1	24	-	-	-	-	78	5911	-	-	-	-	0.8035	0.6094	-	-	-	-
K49R	4	-	7	2	1	-	342	-	722	79	132	-	0.8035	-	0.278	1	0.5603	-
K64R	5	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
V69I	11	5	8	1	6	3	1492	2718	1051	74	922	771	0.4657	0.1313	1	0.4803	1	0.8812
S68G	-	-	-	-	1	2	-	-	-	-	24	79	-	-	-	-	0.1369	0.0701
S68R	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
A98S	-	-	-	-	4	-	-	-	-	-	407	-	-	-	-	-	0.2385	-
K102Q	3	-	-	-	-	-	53	-	-	-	-	-	**0.0041**	-	-	-	-	-
K103R	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
K104R	1	1	-	2	-	2	0	197	-	10	-	129	**<0.0001**	0.7655	-	0.0546	-	0.3387
K104T	-	1	-	-	-	-	-	0	-	-	-	-		**<0.0001**	-	-	-	-
V118I	32	1	-	1	-	-	56	170	-	13	-	-	**<0.0001**	**0.0024**	-	0.3776	-	-
V118M	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0073**	-	-	-
D121C	-	-	2	-	-	-	-	-	48	-	-	-	-	0.8175	**0.0473**	-	-	-
D121H	4	1	1	-	1	-	373	285	107	-	65	-	0.9377	0.688	0.5627	-	0.3268	-
D121Y	-	3	3	-	-	-	-	815	372	-	-	-	-	0.8175	1	-	-	-

K122A	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
K122E	25	20	7	3	5	3	2797	6115	898	66	778	817	0.8478	0.059	1	0.4815	1	0.7739
K122V	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	
D123G	-	3	-	-	1	1	-	1427	-	-	40	75	-	0.2703	-	-	0.2158	0.8224
D123N	11	3	-	1	1	-	777	951	-	0	132	-	0.1295	0.9382	-	0.0328	0.5603	-
D123S	7	5	-	-	3	1	1461	1563	-	-	443	123	**0.0329**	0.6737	-	-	0.7144	0.8958
I135K	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
I135L	-	-	-	1	-	-	-	-	-	95	-	-	-	-	-	0.208	-	-
I135R	-	-	1	-	-	-	-	-	26	-	-	-	-	-	0.1804	-	-	-
I135T	8	-	1	2	1		1492	-	252	40	323	-	0.063	-	0.6903	0.6378	0.6818	-
I135V	-	5	1	9	3	5	-	435	3	120	335	1498	-	**0.024**	**0.0287**	**0.0137**	0.4083	0.2641
I135Y	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
E138A	-	2	-	-	-	-	-	387	-	-	-	-	-	0.9371	-	-	-	-
E138D	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**		-	-	-	-
E138V	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
I142T	-	1	-	-	-	-	-	37	-	-	-	-	-	0.2453	-	-	-	-
I142V	5	5	1	-	-	-	404	285	105	-	-	-	0.6326	**0.0019**	0.5556	-	-	-
K154R	1	1	-	-	-	-	0	0	-	-	-	-	**<0.0001**	**<0.0001**	-	-	-	-
P157L	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
P157R	-	-	-	2	-	-	-	-	-	0	-	-	-	-	-	**0.001**	-	-
A158S	5	4	1	-	-	-	496	747	12	-	-	-	0.9889	**0.6893**	0.0908	-	-	-
S162A	1	-	-	-	1	9	137	-	-	-	395	2133	0.8042	-	-	-	0.4368	0.2904
S162C	1	7	1	11	2	-	177	747	0	219	49	-	0.9416	**0.0397**	**0.0073**	0.0988	**0.0322**	-
S162E	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
S162N	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
S162T	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
S162W	-	-	-	1	-	-	-	-	-	5	-	-	-	-	-	0.1823	-	-
T165I	1	3	2	-	1	-	146	414	81	-	47.0	-	0.8663	0.4554	0.1165	-	0.2484	-
T165P	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
K166R	2	3	1	-	-	-	81	951	22		-	-	0.3694	0.9382	0.1556	-	-	-
E169D	1	-	2	2	-	1	109	-	120	153	-	27	0.6188	-	0.2186	0.7159	-	0.3166

E169G	1	-	-	-	-	-	0	-	-		-	-	**<0.0001**	-	-	-	-	-
K173A	15	14	-	2	2	-	435	4212	-	215	168	-	**<0.0001**	0.2671	-	**0.0058**	0.2587	-
K173D	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
K173I	-	-	-	7	-	3	-	-	-	35	-	193	-	-	-	**0.0001**	-	0.0937
K173L	3	-	-	1	-	-	839	-	-	0	-	-	0.0852	-	-	**0.0328**	-	-
K173M	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
K173N	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
K173R	-	1	-	1	1	-	-	0	-	7	65	-		**<0.0001**	-	0.2358	0.3268	-
K173S	4	-	-	-	-	1	1554	-	-	-	-	43	**0.0004**	-	-	-	-	0.5273
K173T	1	10	-	-	3	3	96	1699	-	-	515	1520	0.6881	0.1759	-	-	1	0.3303
Q174K	14	11	-	14	5	-	2300	2174	-	372	910	-	**0.0078**	0.5341	-	0.3797	0.6760	-
Q174N	-	1	1	-	-	-	-	11	0	-	-	-	-	**0.0332**	**0.0073**	-	-	-
Q174R	-	1	1	-	1	7	-	401	11	-	63	1566	-	0.9783	0.0840	-	0.3184	0.2381
D177E	21	20	7	4	5	7	2642	5299	996	149	1101	2133	0.2272	0.9160	0.5337	0.7782	0.1052	**0.0004**
D177G	2	-	-	-	-	-	87	-	-	-	-	-	1	-	-	-	-	-
D177N	-	-	-	1	-	-	-	-	-	10	-	-	-	-	-	0.3101	-	-
I178L	3	2	3	3	-	-	106	482	131	50	-	-	0.1136	0.7917	0.0620	0.3973	-	-
I178M	3	-	2	1	-	3	174	-	471	19	-	635	0.4494	-	0.4779	0.4945	-	0.7764
I178V	-	3	-	-	-	-	-	115	-	-	-	-	-	**0.0020**	-	-	-	-
V179D	-	2		1	-	-	-	0	-	1	-	-	-	**<0.0001**	-	0.1823	-	-
V179G	-	-	-	1	-	-	-	-	-	5	-	-	-	-	-	0.1823	-	-
V179I	1	-	-	1	-	-	1678	-	-	20	-	-	**<0.0001**	-	-	0.1585	-	-
G196E	8	2	1	-	2	-	193	299	39	-	71	-	**<0.0001**	0.7360	0.2566	-	0.0625	-
G196K	-	1	-	-	-	-	-	122	-	-	-	-	-	0.9634	-	-	-	-
G196S	1	-	-	-	-	-	27	-	-	-	-	-	**<0.0001**	-	-	-	-	-
T200A	19	26	-	14	7	7	1150	6250	-	182	1101	2065	**0.0015**	0.2537	-	-	1	**0.0088**
T200E	-	-	5	-	-	-	-	-	164	-	-	-	-	-	**0.0030**	-	-	-
T200I	2	-	-	-	-	-	118	-	-	-	-	-	0.6715	-	-	-	-	-
T200M	-	-	1	-	-	-	-	-	34	-	-	-	-	-	0.2280	-	-	-
T200R	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
T200V	-	-	-	-	-	4	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**

I202V	2	2	-	2	-	-	622	510	-	26	-	-	0.1442	0.7361	-	0.2320	-	-
E204D	1	1	1	-	-	-	0	0	0	-	-	-	**<0.0001**	**<0.0001**	**0.0073**	-	-	-
E204K	-	-	1	-	-	1	-	-	81	-	-	52	-	-	0.4621	-	-	0.6241
E204L	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0073**	-	-	-
E204Q	1	-	2	-	-	-	0	-	0	-	-	-	**<0.0001**	-	**<0.0001**	-	-	-
Q207A	18	5	-	-	1	3	2704	584	-	-	55	172	**0.0011**	0.1148	-	-	0.2841	0.0601
Q207D	4	2	-	1	-	-	87	428	-	30	-	-	**0.0024**	0.9103	-	1	-	-
Q207E	1	17	7	11	1	7	180	4552	756	223	515	1679	0.9247	0.9728	0.4470	0.0998	0.2495	-
Q207G	-	-	-	-	-	1	-	-	-	-	-	95	-	-	-	-	-	0.9558
Q207K	-	-	-	-	3	-	-	-	-	-	359	-	-	-	-	-	0.4354	-
Q207N	2	1		1			44	211	-	16	-	-	0.0854	0.7271	-	0.4389	-	-
Q207S	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
R211A	-	-	-	1	-	-	-	-	-	12	-	-	-	-	-	0.3558	-	-
R211K	15	14	6	7	1	9	466	4484	879	314	443	1203	**<0.0001**	0.2722	0.6609	0.0516	0.4338	0.1081
R211N	-	-	-	-	1	1	-	-	-	-	37	23	-	-	-	-	0.2015	0.2552
R211S	10	-	-	1	5	-	2238	-	-	5	383	-	**<0.0001**	-	-	0.1823	0.0388	-
F214L	6	5	3	-	1	-	404	883	38	-	144	-	0.3003	0.5152	**0.0129**	-	0.5937	-
P243A	2	2	1	1	-	-	0	0	14	66	-	-	**<0.0001**	**<0.0001**	0.1041	0.7067	-	-
P243S	-	1	1	1	-	-	-	0	35	0	-	-	-	**<0.0001**	0.2338	0.0328	-	-
P243T	1	3		1	-	-	0	163	-	10	-	-	**<0.0001**	**0.0173**	-	0.3101	-	-
I244V	-	-	5	-	-	-	-		44	-	-	-		-	**<0.0001**	-	-	-
V245E	7	1	1	-	-	-	528	211	164	-	-	-	0.3845	0.7271	1	-	-	-
V245H	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-
V245K	1	2	4	1	-	1	292	482	482	5	-	168	0.4666	0.7971	0.7368	0.1823	-	0.7160
V245L	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
V245M	3	-	-	-	-	-	466	-	-	-	-	-	0.7144	-	-	-	-	-
V245Q	13	21	1	11	17	10	1243	5571	175	380	1065	1072	0.6184	0.9260	1	0.1050	0.3887	**0.0096**
V245S	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-
V245T	1	-	1	-	-	-	402	-	153	-	-	-	0.0015	-	1	-	-	-
E248D	10	8	-	7	4	4	995	679	-	244	814	386	0.8303	0.0014	-	0.5844	0.6867	0.1939
E248N	2	-	-	4	-	-	205	-	-	58	-	-	0.7901	-	-	0.1306	-	-

D250T	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
S251A	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
S251D	-	1	-	-	1	4	-	14	-	-	36	0	-	0.0763	-	-	0.1967	**<0.0001**
S251H	3	-	-	-	-	-	40	-	-	-	-	-	**0.0006**		-	-	-	-
S251N	1	1	-	1	-	-	71	88	-	7	-	-	0.8365	0.8221	-	0.2358	-	-
S251T	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**		-	-	-	-
K275Q	3	1	-	1	2		121	279	-	4	77	-	0.1748	0.6684	-	0.1542	0.1263	-
K275R	1	2	-	1	2	1	53	394	-	78	0	116	0.9792	0.9935	-	0.4825	**<0.0001**	0.9296
A272K	-	2	-	-	-	-	-	82	-	-	-	-	-	**0.0356**	-	-	-	-
A272P	11	18	7	11	6	-	1108	5435	788	392	1053	-	0.8402	0.2613	0.4550	**0.0378**	0.5937	-
A272S	6	1		2	1		106	214	-	10	57	-	**<0.0001**	0.7193	-	0.0546	0.2928	-
A272T	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
V276I	-	1	1	-	1	2	-	88	49	-	62	170	-	0.7808	0.3109	-	0.3142	0.4431
V276T	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
K277Q	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
K277R	8	12	6	3	1	-	528	4076	624	120	227	-	0.1711	0.2161	0.4779	0.7656	1	-
K277S	-	-	-	-	1	1	-	-	-	-	0	408	-	-	-	-	**0.0058**	0.7093
Q278H	-	2	3	1	3	-	-	299	64	66	91	-	-	0.7360	**0.0096**	0.7067	**0.0128**	-
Q278N	-	1	-	1	-	-	-	183		6	-	-	-	0.8070	-	0.2095	-	-
K281R	6	4	1	4	2	1	653	598	105	27	109	522	0.8580	0.4039	0.5556	**0.0137**	0.1302	0.4620
L283I	3	2	-	-	-	-	127	197	-	-	-	-	0.2022	0.3868	-	-	-	-
R284K	-	1	-	-	-	4	-	95	-	-	-	34	-	0.8231	-	-	-	**<0.0001**
T286A	18	20	6	13	7	11	2673	4416	712	355	1053	2156	**0.0026**	0.2861	0.7206	0.7033	1	0.9498
T286V	-	1	-	-	-	-	-	183	-	-	-	-	-	0.8070	-	-	-	-
E291D	23	24	-	11	6	11	2766	6386	-	384	1089	2019	0.3963	0.9584	-	0.079	0.4862	0.4946
E291I	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
V292I	20	23	1	13	7	6	1958	6318	110	326	1137	1883	0.4694	-	0.5732	0.3182	1	**0.0361**
I293V	27	20	6	13	7	10	3046	6386	843	405.0	1149	2224	0.9326	**0.0013**	1	0.2613	1	0.5500
P194A	1	-	-	-	-	-	127	-	-	-	-	-	0.7319	-	-	-	-	-
P294S	-	-	-	1	-	-	-	-	-	7	-	-	-	-	-	0.1434	-	-
P294T	16	2	1	-	2	9	2020	347	41	-	68	1566	0.5043	0.8799	0.2677	-	0.0579	0.5556

L295M	2	-	-	-	-	-	106	-	-	-	-	-	0.5771	-	-	-	-	-
L295W	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
E297A	1	-	-	1	5	-	0	-	-	0	802	-	**<0.0001**	-	-	**0.0328**	-	-
E297D	2	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
E297K	4	2	6	-	-	-	199	102	142	-	-	-	0.1929	0.0768	**0.0001**	-	-	-
E297R	2	-	-	8	-	-	53	-	-	244	-	-	0.1445	-	-	1	-	-
E297T	-	-	-	3	1	-	-	-	-	95	103	-	-	-	-	1	0.4695	-
L310I	-	1	-	1	-	4	-	82	-	5	-	41	-	0.7422	-	0.1823	-	**<0.0001**
K311R	3	1	-	-	-	-	118	401	-	-	-	-	0.1617	0.9783	-	-	-	-
E312D	8	-	-	-	-	-	1865	-	-	-	-	-	**0.0015**	-	-	-	-	-
E312N	1	-	-	-	-	-	342	-	-	-	-	-	0.3419	-	-	-	-	-
E312P	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
E312Q	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
E312T	3	-	-	-	-	-	37	-	-	-	-	-	**0.0003**	-	-	-	-	-
V317A	3	5	-	1	-	-	246	747	-	58	-	-	0.8459	0.3056		0.7033	-	-
V317E	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-
S322A	1	-	-	-	-	7	106	-	-	-	-	163	0.6340	-	-	-	-	**<0.0001**
S322E	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
S322T	-	-	-	-	-	1	-	-	-	-	-	102	-	-	-	-	-	0.9964
K323E	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
K323I	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**	-	-	-	-	-
K323R	1	-	-	-	-	1	0	-	-	-	-	0	**<0.0001**	-	-	-	-	**<0.0001**
D324E	7	6	1	3	4	-	224	747	45	50	850	-	**0.0013**	0.0993	0.2896	0.3973	0.4214	-
D324F	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
L325I	-	-	-	-	2	-	-	-	-	-	22	-	-	-	-	-	**0.0075**	-
I326K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I326L	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
I326V	11	-	-	1	1	8	497	-	-	245	40	1339	**0.0021**	-	-	**<0.0001**	0.2158	0.5382
E328D	1	1	2	-	-	-	0	82	0	-	-	-	**<0.0001**	0.7422	**<0.0001**	-	-	-
E328K	-	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**<0.0001**
I329L	-	1	1	1	-	-	-	353	57	27	-	-	-	**0.0438**	0.1546	1	-	-

I329V	1	1	-	1	6	-	106	129	-	50	1041	-	0.6340	**0.0004**	-	1	1	-
G333D	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0073**	-	-	-
G333E	-	-	1	1	-	-	-	-	0	0	-	-	-	-	**0.0073**	**0.0328**	-	-
G333L	-	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0058**	-
Q334A	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
Q334C	-	1	-	-	-	-	-	102	-	-	-	-	-	0.8627	-	-	-	-
Q334D	-	1	-	-	-	-	-	747	-	-	-	-	-	0.3954	-	-	-	-
Q334E	-	-	-	-	1	-	-	-	-	-	14	-	-	-	-	-	0.0842	-
Q334H	2	-	1	-	-	1	56	-	41	-	-	0	0.1664	-	0.2677	-	-	<**0.0001**
Q334K	-	1	-	-	-	-	-	0	-	-	-	-	-	<**0.0001**	-	-	-	-
Q334N	2	4	-	1	-	-	0	747	-	14	-	-	<**0.0001**	0.6893	-	0.3987	-	-
Q334R	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	**0.0328**	-	-
Q334Y	-	-	1	1	-	-	-	-	11	0	-	-	-	-	0.0840	**0.0328**	-	-
G335A	1	-	-	-	-	-	0	-	-	-	-	-	<**0.0001**	-	-	-	-	-
G335D	4	1	-	-	-	2	2704	5435	-	-	-	1951	<**0.0001**	<**0.0001**		-	-	<**0.0001**
G335L	-	1	-	-	-	-	-	0	-	-	-	-	-	<**0.0001**	-	-	-	-
G335E	1	-	-	-	-	-	2705	-	-	-	-	-	<**0.0001**	-	-	-	-	-
G335R	1	-	-	-	-	-	0	-	-	-	-	-	<**0.0001**	-	-	-	-	-
G335S	-	-	1	-	-	-	-	-	50	-	-	-	-	-	0.3161	-	-	-

*According to the Stanford HIV Drug Resistance Database; #Fisher exact test; Drug resistance mutations are in bold letters. Additional NRTI-selected mutations are in underline letters. Additional NNRTI polymorphic accessory mutations are in italic letters. doi:10.1371/journal.pone.0113626.s002

Table S3. Epidemiological characteristics of the patients included in the transmission clusters.

Cluster number	Sequence	Sampling date	Year of birth	Gender	Reported transmission route	Region
1	09AGHDP119	2009	1980	F	Heterosexual	Luanda
	09AGHDP226	2009	n.a.	F	n.a.	Luanda
	09AGHDP208	2009	1970	F	Heterosexual	Luanda
	JN937038	2008	1975	F	n.a.	n.a.
	01AOHM176	2001	1963	M	Heterosexual	Luanda
	JQ616884	2009	n.a.	n.a.	n.a.	n.a.
	JN937034	2008	n.a.	n.a.	n.a.	n.a.
2	JQ616880	2009	n.a.	n.a.	n.a.	n.a.
	JN937047	2009	n.a.	n.a.	n.a.	n.a.
3	09AGHDP237	2009	1957	M	Heterosexual	Luanda
	01AOCSE126	2001	1976	F	n.a.	Lunda Norte
4	01AOLFA13	2001	1971	M	Heterosexual	Luanda
	01AOHAB86	2001	1963	F	n.a.	Luanda
5	JN937098	2010	n.a.	n.a.	n.a.	n.a.
	JN937104	2010	n.a.	n.a.	n.a.	n.a.
6	09AGHDP186	2009	1982	F	Heterosexual	Luanda
	JQ616882	2009	n.a.	n.a.	n.a.	n.a.
7	09AGHDP62	2009	1973	F	Heterosexual	Luanda
	JN937097	2010	n.a.	n.a.	n.a.	n.a.
	01AOSNS09	2001	1981	F	Homosexual	Luanda
	09AGHDP42	2009	1967	M	Heterosexual	Luanda
	09AGHDP289	2009	1985	F	Heterosexual	Luanda
8	93AOHDC247	1993	n.a.	F	n.a.	Cabinda
	93AOHDC253	1993	n.a.	F	n.a.	Cabinda
9	JN937116	2010	n.a.	n.a.	n.a.	n.a.
	JN937112	2010	1976	M	Heterosexual	Central
10	JN937050	2009	n.a.	n.a.	n.a.	n.a.
	JQ616883	2009	n.a.	n.a.	n.a.	n.a.
	09AGHDP279	2009	1960	M	Heterosexual	Luanda
	01AOHJM06	2001	1968	F	Heterosexual	Luanda
11	01AOSNS01	2001	1959	M	Heterosexual	Luanda
	01AOHDP73	2001	1958	M	Heterosexual	Luanda
12	JN937046	2009	n.a.	n.a.	n.a.	n.a.
	JN937037	2008	n.a.	n.a.	n.a.	n.a.

13	JN937061	2009	n.a.	n.a.	n.a.	n.a.
	JN937101	2010	n.a.	n.a.	n.a.	n.a.
14	09AGHDP233	2009	1979	M	Heterosexual	Luanda
	09AGHDP280	2009	2004	M	Vertical	Luanda
	JN937054	2009	1964	F	Heterosexual	Central
	01AOSNS56	2001	1965	M	n.a.	Luanda
15	09AGHDP231	2009	1987	F	Heterosexual	Luanda
	09AGHDP164	2009	1992	F	Heterosexual	Luanda
	01AOHDC229	2001	1964	M	Heterosexual	Cabinda
16	09AGHDP44	2009	2007	M	Vertical	Luanda
	09AGHDP68	2009	1985	F	Heterosexual	Luanda
17	09AGHDP50	2009	1963	M	Heterosexual	Luanda
	09AGHDP118	2009	1978	F	Heterosexual	Luanda
18	09AGHDP242	2009	1970	M	Heterosexual	Luanda
	09AGHDP111	2009	1965	F	Heterosexual	Luanda
19	JN937017	2008	1987	F	Heterosexual	Namibe
	JN937040	2008	1967	M	MSM	Luanda
20	JN937095	2010	1965	M	Heterosexual	Benguela
	01AOHDP71	2001	1949	F	Heterosexual	Luanda
21	01AOHJM64	2001	1975	F	n.a.	Luanda
	01AOSNS55	2001	1966	F	n.a.	Luanda
	09AGHDP240	2009	1982	F	Heterosexual	Luanda
22	09AGHDP74	2009	1939	M	Heterosexual	Luanda
	JQ616899	2009	n.a.	n.a.	n.a.	n.a.
23	09AGHDP94	2009	1973	M	Heterosexual	Luanda
	09AGHDP30	2009	2007	M	Vertical	Luanda
24	JN937059	2009	n.a.	n.a.	n.a.	n.a.
	JN937060	2009	n.a.	n.a.	n.a.	n.a.
	JN937115	2010	1981	F	n.a.	Central
	09AGHDP82	2009	2006	M	Vertical	Luanda
25	09AGHDP212	2009	1981	F	Heterosexual	Luanda
	01AOSNS04	2001	1969	M	Bisexual	Luanda
	JQ616913	2009	n.a.	n.a.	n.a.	n.a.
	09AGHDP263	2009	1978	F	Heterosexual	Luanda
26	09AGHDP267	2009	1971	M	Heterosexual	Luanda
	01AOSNS24	2001	1973	F	Heterosexual	Luanda
27	01AOLFA90	2001	1977	M	n.a.	Luanda
	01AOSNS49	2001	1957	M	Bisexual	Cabinda
28	JN937081	2010	n.a.	n.a.	n.a.	n.a.
	JN937070	2010	n.a.	n.a.	n.a.	n.a.
	JN937073	2010	n.a.	n.a.	n.a.	n.a.
	JN937079	2010	n.a.	n.a.	n.a.	n.a.
29	09AGHDP169	2009	1986	M	Heterosexual	Luanda

	09AGHDP167	2009	1980	M	Heterosexual	Luanda
30	09AGHDP130	2009	1984	F	Heterosexual	Luanda
	09AGHDP145	2009	1976	M	Heterosexual	Luanda
31	01AOLFA19	2001	1963	M	Heterosexual	Luanda
	01AOLFA17	2001	n.a.	F	Heterosexual	Luanda
	JQ616890	2009	n.a.	n.a.	n.a.	n.a.
32	09AGHDP49	2009	1973	M	Heterosexual	Luanda
	01AOCSE136	2001	1972	M	Heterosexual	Luanda
33	JQ616894	2009	n.a.	n.a.	n.a.	n.a.
	JQ616891	2009	n.a.	n.a.	n.a.	n.a.
34	09AGHDP64	2009	1963	F	Heterosexual	Luanda
	09AGHDP200	2009	1969	M	Heterosexual	Luanda
	JN937085	2010	n.a.	n.a.	n.a.	n.a.
	09AGHDP296	2009	1976	F	Heterosexual	n.a.
	01AOLFA94	2001	1969	M	Heterosexual	Luanda
	01AOSNS36	2001	n.a.	F	Heterosexual	Luanda
	JN937075	2010	n.a.	n.a.	n.a.	n.a.
35	09AGHDP281	2009	1978	F	Heterosexual	Luanda
	JQ616886	2009	n.a.	n.a.	n.a.	n.a.
36	01AOSNS03	2001	1964	M	Heterosexual	Luanda
	01AOLFA14	2001	1956	M	Heterosexual	Luanda
	01AOLFA18	2001	1958	F	Heterosexual	Luanda
37	09AGHDP274	2009	1976	M	Heterosexual	Luanda
	09AGHDP204 (a2)	2009	2006	F	Vertical	Luanda
38	09AGHDP106	2009	2006	M	Vertical	Luanda
	09AGHDP57	2009	1966	F	Heterosexual	Luanda
39	09AGHDP290	2009	1980	M	Heterosexual	Luanda
	JN937058	2009	n.a.	n.a.	n.a.	n.a.
40	09AGHDP100	2009	1969	F	Heterosexual	Luanda
	JN937033	2008	n.a.	n.a.	n.a.	n.a.
41	JN937035	2008	n.a.	n.a.	n.a.	n.a.
	JN937072	2010	n.a.	n.a.	n.a.	n.a.
	09AGHDP37	2009	1972	F	Heterosexual	Luanda
42	09AGHDP157	2009	1958	M	Heterosexual	Luanda
	JQ616905	2009	n.a.	n.a.	n.a.	n.a.
43	JQ616885	2009	n.a.	n.a.	n.a.	n.a.
	09AGHDP20	2009	1982	F	Heterosexual	Luanda
	09AGHDP245	2009	1989	F	Heterosexual	Luanda
44	01AOCSE125	2001	1969	M	Heterosexual	Luanda
	01AOHDP75	2001	1973	F	Heterosexual	Luanda
45	JN937031	2008	n.a.	n.a.	n.a.	n.a.
	JN937106	2010	1991	F	Heterosexual	Central

46	09AGHDP258	2009	2006	M	Heterosexual	Luanda
	09AGHDP266	2009	1986	F	Heterosexual	Luanda
	01AOSNS40	2001	1976	F	Heterosexual	Luanda
	09AGHDP86	2009	2006	M	Vertical	Luanda
	JN937028	2008	n.a.	n.a.	n.a.	Central
	JQ616898	2009	n.a.	n.a.	n.a.	n.a.
	09AGHDP105	2009	1976	M	Heterosexual	Luanda
47	09AGHDP201	2009	1951	M	Heterosexual	Luanda
	JQ616897	2009	n.a.	n.a.	n.a.	n.a.
48	01AOLFA97	2001	1958	F	Heterosexual	Luanda
	01AOSNS37	2001	1973	F	n.a.	Luanda

n.a.- not available. doi:10.1371/journal.pone.0113626.s003

Author Contributions

Conceived and designed the experiments: NT IB SC RC SZ. Performed the experiments: IB SC PC FM SZ. Analyzed the data: IB SZ CP ST RC NT. Contributed reagents/materials/analysis tools: IB SZ FM ST SC. Wrote the paper: IB SZ NT FM.

Chapter 3

Early Infant Diagnosis of HIV-1 Infection in Luanda, Angola, Using a New DNA PCR Assay and Dried Blood Spots

Short title: Early Infant Diagnosis of HIV-1 in Luanda

Francisco Martin[1,¶], Claudia Palladino[1,¶], Rita Mateus[1], Anna Bolzan[2], Perpétua Gomes[3,4], José Brito[4], Ana Patrícia Carvalho[3], Yolanda Cardoso[5], Cristovão Domingos[6], Vanda Sofia Lôa Clemente[2], Nuno Taveira[1,4]

[1]Research Institute for Medicines (iMed.Ulisboa), Faculdade de Farmácia, Universidade de Lisboa, Lisbon, Portugal. [2] Hospital da Divina Providência, Luanda, Angola. [3]Laboratório de Microbiologia Clínica e Biologia Molecular, Serviço de Patologia Clínica, Centro Hospitalar de Lisboa Ocidental, E.P.E., Lisbon, Portugal. [4] Centro de Investigação Interdisciplinar Egas Moniz, Instituto Superior de Ciências da Saúde Egas Moniz, Caparica, Portugal. [5] Ministério da Saúde de Angola - Instituto Nacional de Saúde Pública (INSP), Luanda, Angola. [6] Ministério da Saúde de Angola - Instituto Nacional de Luta contra a Sida (INLS), Luanda, Angola.

¶These authors contributed equally to this work.

PLoS ONE 2017 July; 12(7): e0181352. https://doi.org/10.1371/journal.pone.0181352

Keywords: HIV-1 early infant diagnosis; proviral DNA; dried blood spots; Luanda, Angola

Abstract

Background: Early diagnosis and treatment reduces HIV-1-related mortality, morbidity and size of viral reservoirs in infants infected perinatally. Commercial molecular tests enable the early diagnosis of infection in infants but the high cost and low sensitivity with dried blood spots (DBS) limit their use in sub-Saharan Africa.

Objectives: To develop and validate a sensitive and cheap qualitative proviral DNA PCR-based assay for early infant diagnosis (EID) in HIV-1-exposed infants using DBS samples.

Study design: Chelex-based method was used to extract DNA from DBS samples followed by a nested PCR assay using primers for the HIV-1 integrase gene. Limit of detection (LoD) was determined by Probit regression using limiting dilutions of newly produced recombinant plasmids with the integrase gene of all HIV-1 subtypes and ACH-2 cells. Clinical sensitivity and specificity were evaluated on 100 HIV-1 infected adults; 5 infected infants; 50 healthy volunteers; 139 HIV-1-exposed infants of the Angolan Pediatric HIV Cohort (APEHC) with serology at 18 months of life.

Results: All subtypes and CRF02_AG were amplified with a LoD of 14 copies. HIV-1 infection in infants was detected at month 1 of life. Sensitivity rate in adults varied with viral load, while diagnostic specificity was 100%. The percentage of HIV-1 MTCT cases between January 2012 and October 2014 was 2.2%. The cost per test was 8-10 USD which is 2- to 4-fold lower in comparison to commercial assays.

Conclusions: The new PCR assay enables early and accurate EID. The simplicity and low-cost of the assay make it suitable for generalized implementation in Angola and other resource-constrained countries.

Introduction

HIV-1 mother-to-child-transmission (MTCT) is the main mode of infection among the pediatric population and is disproportionately affecting children in impoverished countries. Despite the decline in MTCT rate in recent years in most of the sub-Saharan Africa, it is estimated that 150,000 children became newly infected with HIV in 2015 [30]. Children infected perinatally are at high risk of rapid disease progression and death during the first year of life without antiretroviral therapy (ART) [197]. Given the reported benefits of early ART initiation in reducing HIV-1-related mortality and long-term morbidity [253] and reducing the size of the HIV-1 reservoirs [254], early HIV-1 diagnosis in newborns represents the critical gateway to timely initiation of life-saving ART. Serological assays do not permit the early diagnosis of HIV-1 infection because of the persistence of maternal HIV-1 antibodies in infants during the first 12-18 months of life. The WHO recommends the use of molecular-based virological testing to determine the infection status for HIV-1-exposed infants during the first 4-6 weeks of life or at the earliest opportunity thereafter [132]. Despite the high accuracy of tests which detect HIV RNA or p24, their sensitivity could potentially be affected in settings of expanded ART for prevention of MTCT (option B and B+), which reduce circulating HIV-1 RNA and viral particles [133]. Qualitative DNA PCR test which detect proviral DNA in peripheral blood mononuclear cells (PBMCs) is recommended for early infant diagnosis (EID) of HIV-1 and is the most widely implemented test in resource-limited settings [134, 135]. The considerable uptake of HIV-1 DNA molecular tests is driven by the lower costs compared with quantitative assays along with their good sensitivity when performed on blood microsamples or dried blood spots (DBS) [134]. The use of DBS presents several advantages such as reduced costs for collection, storage and shipping. Thus, DBS samples are convenient for increasing access to testing in settings with poor healthcare provision

and referral laboratories [294]. Currently, two HIV-1 DNA assays are commercially available for EID using DBS: the COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Qualitative Test (Roche Molecular System Inc., Branchburg, NJ), which recently replaced the Roche Amplicor® DNA test v1.5, and the RealTi*m*e HIV-1 Qualitative Test (Abbott Molecular, Des Plaines, IL) [135]. These assays are highly sensitive but require sophisticated instrument platforms whose cost with related equipment (e.g. centrifuge; biosafety cabinet; freezer) can range from about US$ 100,000 to more than US$ 200,000. Recently, innovative technologies designed for use at or near the point-of-care (such as the Cepheid Xpert® HIV-1 Qual assay and the Alere™ q HIV-1/2 Detect) have been developed and the Xpert® HIV-1 Qual assay has been validated for both whole blood and DBS specimens [135].

Despite the recent advances in prevention of MTCT, Angola reported one of the highest rates of MTCT (25%) among the 22 priority countries included in the UNAIDS global plan [30]. The EID national program implemented in 2007 based on the Roche Amplicor® DNA test v1.5 was interrupted in 2012; consequently the coverage of virological testing for infants is currently very low and only 14% of HIV-1-exposed newborns received virological HIV-1 testing within the first 2 months of life in 2014 [30]. Several challenges may prevent the implementation of molecular diagnostic tests such as the high cost of available commercial PCR assays and their limited sensitivity with highly divergent HIV-1 subtypes which co-circulate in the country [5-7, 98, 99, 212, 252].

In this article, we developed and validated a new qualitative HIV-1 DNA PCR assay for the early infant diagnosis of HIV-1 infection in Angola and other countries with similar complex epidemics. Using this highly sensitive and specific assay we identified the new cases of MTCT of HIV-1 in the APEHC pediatric cohort recently established in a major hospital in Luanda.

Participants and Methods

Study Design

This assay validation study describes the use of DBS to diagnose HIV-1 in infants born to HIV-infected mothers in Angola between January 2012 and October 2014. The STROBE checklist was used to help design and conduct the study [295].

Ethics Statement

This study was conducted according to the Declaration of Helsinki with the approval of the National Ethical Committee of Angola and the Ethic Committee of the *Centro Hospitalar de Lisboa Ocidental, E.P.E* in Portugal. Written informed consent was obtained from all participants or from parents/guardians for their infants and from HIV-1 seronegative healthy volunteers.

The Angolan Pediatric HIV Cohort and clinical samples

The Angolan Pediatric HIV Cohort (APEHC) has prospectively collected data on HIV-1-infected pregnant women and their infants attending the municipal *Hospital da Divina Providência* (HDP) since March 2012. HDP is located in the Luanda district, which is the most populated district of Luanda city, Angola. Epidemiological, clinical and laboratory data were collected at study entry and every 6 months thereafter for women. Infants were followed according to the perinatal care service offered at the HDP, which includes clinical and biological examination at months 1, 3, 6, and 18 of life. No specific recommendations for HIV treatment was made for women enrolled in the cohort and physicians followed the WHO guidelines for prevention of MTCT [134]. Infants received NVP once daily from birth through age 4-6 weeks in accordance with option B [134] . HIV-1 testing for both mothers and newborns was free of charge and was performed using two rapid tests for detection of antibodies against HIV-1/2 (Determine HIV ½ and Uni-

Gold HIV) as recommended by WHO [296]. The diagnosis required consistent results of the two different tests. An infant was considered as infected if anti-HIV-1 antibodies were detected on two separate samples collected at least three months apart and persisted after 18 months of age; it was considered non-infected if serological testing was negative on two separate samples before or after 18 months. Undetermined cases were those with discordant results between the two rapid tests and were further tested using an ELISA assay (Vironostika HIV Uni-form II Ag/Ab ELISA test; bioMérieux, France). DBS specimens were obtained from all infants born between January 2012 and October 2014 from HIV-1 infected women enrolled in the APEHC. Additionally, five DBS samples from HIV-1 infected infants aged 2- to 12-days old were obtained from the *Instituto Nacional de Saúde Pública* (INSP). DBS samples were also collected from six women attending the HDP with indeterminate HIV serologic testing results. DBS samples were prepared by spotting 125 µL of whole blood, collected by heel prick for infants and by finger-prick for adults, onto filter paper cards (Whatman® Human ID Blood Stain Cards BFC 180). DBS were dried over night at room temperature, individually enveloped in Glassine paper, inserted in a zip-lock polyethylene bag (Deltalab S.L., Spain) with silica desiccant and stored at -20ºC. DBS were subsequently shipped at room temperature to our laboratory at the University of Lisbon for testing.

Two sets of samples were used as clinical controls. One-hundred DBS specimens were collected during 2014 from adults attending a central hospital in Lisbon (*Centro Hospitalar de Lisboa Ocidental, E.P.E.*), who had a confirmed HIV-1 diagnosis done by a 4th generation assay followed by an immunoblot assay on two separate samples. Plasma HIV-1 RNA level was determined in these patients by COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Test, v2.0 (Roche Diagnostic Systems) with a lower limit of detection

of 20 copies/mL. $CD4^+$ T-lymphocytes were quantified by flow cytometry. Fifty DBS collected from HIV-1 seronegative healthy volunteers were used as negative controls.

To further test the diagnostic specificity of the assay, samples obtained from patients infected with HBV, HCV and CMV were also analyzed. These samples were collected from HIV-1 seronegative adults attending the *Centro Hospitalar de Lisboa Ocidental, E.P.E.* who had a confirmed diagnosis of hepatitis B (HBV) done by detection of hepatitis B surface antigen, HBsAg, (N=10), or hepatitis C (HCV) done by detection of antibodies to HCV and viral RNA (N=10), or cytomegalovirus (CMV) done by detection of antibodies to CMV (N=10).

Production of DBS samples with control plasmids

A 1553 bp *pol* gene product comprising the highly conserved IN gene region was amplified from the reference subtype B HIV-1 SG3.1 [297] and from nine primary viral isolates belonging to the most common HIV-1 clades circulating in Angola S1 Table by RT-PCR as described previously [4, 5, 7, 98, 99]. Primers for this PCR were designed using PerlPrimer® v1.1.21 software and reference sequences present in the Los Alamos HIV sequence database (http://www.hiv.lanl.gov). Their sequence and location in the HIV-1 genome are shown in S2 Table. PCR products were cloned into pcDNA 3.1D/V5-His-TOPO plasmid using the protocol indicated by the manufacturer (Invitrogene Corp., Carlsbad, CA) and sequenced by the Sanger method. Phylogenetic analysis was used to confirm the subtype of the isolates (data not shown). Purified control plasmids were quantified by spectroscopy at 260 nm using a calibration curve and then serially diluted in HIV-1 seronegative blood and spotted (125 µL) onto Human ID blood stain cards to obtain a concentration of 50 to 5,000 copies/DBS.

Production of DBS samples with ACH-2 cells

ACH-2 cells were used as analytical control for DNA PCR assays as they contain a single, integrated HIV-1 subtype B proviral DNA per cell [298, 299]. ACH-2 cells were diluted in 125 µL of HIV seronegative blood (5 log serial dilutions) and spotted on Human ID blood stain cards to obtain a DBS control panel containing 50-5,000 cells.

Chelex DNA extraction method

DNA was extracted from DBS samples using the polyvalent cationic resin Chelex 100® (Bio-Rad Laboratories, Hercules, CA, USA). Briefly, six circles of 5 mm diameter were punched from each DBS spot into a 1.5 mL tube. After 30 min washing with sterile water and 3 min centrifugation (15,400 g), 250 µL of 5% Chelex solution was added to the pellet and samples were incubated at 56ºC for 15 minutes. The samples were vortexed and centrifuged (15,400 g) for 3 min, prior to a final incubation at 100ºC for 8 minutes. Finally, samples were centrifuged and stored at -20ºC until the PCR reaction was performed.

PCR amplification of proviral HIV-1 DNA

A nested PCR was used to amplify a 194 bp fragment of the IN gene. First- and second-round amplifications were performed using the same reaction and cycling conditions. The 25 µL reaction volume contained 1X NH4 buffer, 3 mmol/L MgCl2, 0.5 µmol/L of each dNTP, 0.3 µmol/L forward and reverse primers S2 Table, 1U of Taq DNA polymerase (Bioline® Reagents Ltd, London, UK) plus 2.5 µL of DNA solution from the clinical or control samples (first-round PCR). For the second-round PCR reaction we used 2.5ul of the first-round PCR reaction. Amplification cycling conditions were as follows: denaturation step of 94°C/3 min followed by 35 cycles of denaturation at 94°C/ 45 sec, annealing at 56°C/35 sec, extension at 72°C/ 1 min followed by a single final extension step at 72ºC/ 15 min. Amplified products were visualized with green safe staining after electrophoresis in 2% agarose gel. In all cases the human gene C-C chemokine receptor

5 (CCR5) was used as an internal control to confirm the presence and quality of genomic DNA. CCR5 was amplified using primers CCR5c and CCR5d [107] and the cycling conditions used for the amplification of the IN gene region.

Limit of detection (LoD)

DNA extracted from DBS samples spotted with serial dilutions of control plasmids or ACH-2 cells was subjected to the nested PCR protocol described above. Each template was amplified ≥10 times and the LoD for each subtype and its 95% fiducial confidence interval were estimated by probit regression analysis.

Results

Analytical sensitivity

The LoD of the assay was determined by probit regression analysis with DBS spiked with serial dilutions of control plasmids containing the IN gene regions from HIV-1 subtypes A, C, D, F, G, H, J and CRF02_AG. For HIV-1 subtype B, DBS spiked with increasing number of ACH-2 cells, which contain one integrated proviral DNA copy per cell, were used. When the subtype was included in the probit model as an independent factor, the parallelism test chi-square was significant ($\chi 2 = 64.7$; $df = 8$; $p < 0.001$) which rejects the assumption of equal slopes across subtypes. Therefore, the probit analysis was implemented separately for each subtype. Under these conditions the LoD of the assay varied between 4.3 and 14.4 copies depending on subtype Table 1.

Table 1. Limit of detection of the assay for different HIV-1 subtypes as determined using probit regression analysis.

Subtype	Probit results*		
	Significance of Pearson χ^2 goodness of fit test	**LoD (copies/PCR)**	**95% Confidence Interval**
A1	0.330	5.0	4.1 – 8.4
H	0.300	10.1	7.9 – 19.7
B	0.885	9.3	5.0 – 33.1
F	0.914	10.4	7.7 – 27.2
G	0.997	11.8	9.6 – 22.6
J	0.267	4.3	3.3 – 6.3
CRF02_AG	0.252	4.4	3.4 – 6.4
C	0.077	5.5	4.7 – 31.0
D	0.624	14.4	10.4 – 25.2

* Determined in DBS spiked with serial dilutions of control plasmids except for subtype B that was determined in DBS spiked with ACH-2 cells (S1 Fig. and S3 Table).

One hundred and twenty five µl of infant's blood, which is the amount present in each DBS, has 0.4-0.6x10^6 PBMCs [300]. HIV-1 infected infants harbor an estimate of 13,000 to 75,400 HIV proviral copies per 10^6 PBMCs [301]. In each PCR reaction, we used 1/100 of the extracted DNA solution (2.5 out of 250 µl); considering only the lower limit of PBMCs that the infants may have in this amount of blood this corresponds to 0.4x10^4 PBMCs per reaction. Considering the lower number of proviral copies that the HIV-1 infected infants may have in this number of cells this corresponds to a minimum of 52 HIV-1 proviral copies. This is more than 11-fold higher than the lower LoD of our PCR assay assuring that it has high enough sensitivity to detect HIV-1 infection in all infected infants.

Diagnostic sensitivity and specificity

Diagnostic specificity of the EID assay was 100% since all the HIV-1 seronegative samples tested (n=186, 50 adults from Portugal and 136 infant DBS samples from the APEHC cohort) were negative for the presence of HIV-1 proviral DNA. This was further confirmed using 30 samples from patients infected with HBV, HCV or CMV as all gave negative results using the new PCR assay S2 Fig.

The clinical sensitivity of the assay was determined with DBS samples collected from 100 HIV-1 infected adult patients from Portugal (patients with chronic infection), from 5 confirmed HIV-1 positive infants obtained from the *Instituto Nacional de Luta contra a Sida* (INLS), Luanda, Angola (infection in these patients was confirmed by serology at month 18 of life and by detection of HIV-1 DNA using the Nuclisens EasyMag/EasyQ, Biomérieux) and from the infants of the APEHC cohort. Regarding the Portuguese adult patients, the median CD4 count was 608 cells/mm^3 (min-max:83-2,075). Nine subjects were severely immunosuppressed (<200 cells/mm^3), 35 had CD4 count of 200-499 cells/mm^3, and 52 had ≥500 cells/mm^3. HIV-1 proviral DNA was detected in 14.3%, 56.3% and 85.7% of the patients with plasma viral load of <20 copies/mL, 20-1,000 copies/mL and >1,000 copies/mL, respectively Table 2. All five HIV-1 infected infants from the INLS were HIV-1 DNA positive by our assay.

Table 2. Performance of the new PCR assay in HIV-1-infected adults from Portugal.

EID assay	**Adult HIV-1-infected patients (N=100)**			**Total**
	Undetectable viral load (<20 copies/mL)	**Viral load of 20-1,000 copies/mL**	**Viral load of >1,000 copies/mL**	
Positive	11	9	6	26
Negative	66	7	1	74
Total	77	16	7	100
Percentage of detection	14.3	56.3	85.7	--

A total of 154 HIV-1-exposed infants were enrolled in the APEHC cohort and one DBS card containing 4 blood spots per infant was available for the lab tests. The median age was 1 month: 83% (129/154) of infants were 1 month of age, 7% (11/154) were 2-5

months of age, and 9% (14/154) were 6-12 months of age; 50% were girls (n=77). For the specificity and sensitivity analyses, 15 patients were excluded as follows: 11 infants dropped-out from routine clinical care and 4 infants died before the serology results were confirmed. Those patients were all negative according to our assay. Three out of the 139 samples that were analyzed by our assay were HIV-1 DNA positive; infection with HIV-1 in these infants was confirmed by serology at month 18 (Table 3 and S3 Fig.).

Table 3. Sensitivity and specificity of the new assay for early infant diagnosis of HIV-1 infection in Angola.

	Infants with HIV-1 serology at month 18 and/or HIV-1 DNA test (N=144)*	
Our HIV-1 DNA PCR assay	**Positive**	**Negative**
Positive	8	0
Negative	0	136
Total	8	136
Sensitivity	100.0 %	
Specificity	100.0 %	

* Infants from the APEHC cohort (N=139) plus infants (N=5) with HIV-1 infection confirmed at the INLS in Luanda, Angola.

All 136 infants with negative results with our assay were HIV-1 seronegative at month 18. Therefore, 3 out of 139 (2.2%) infants from the APEHC cohort were infected with HIV-1 between January 2012 and October 2014 acquiring the virus through MTCT. Among the HIV-1-infected infants, one was an 8-month-old girl born in healthcare facilities at the end of 2012 who received oral zidovudine and formula feeding; her mother initiated ART (lamivudine/zidovudine/nevirapine) during the second trimester of

pregnancy and received intrapartum zidovudine. The second infant was a one month old boy born at home in July 2014 who received formula feeding and his mother initiated ART (tenofovir/lamivudine/efavirenz) during the third trimester of pregnancy; no prophylaxis with zidovudine was administered. The third HIV-1-infected newbom was a disabled 7–month-old girl who was transferred to another hospital soon after delivery. No information on prophylaxis was available.

Finally, we calculated the cost per test using our assay, including cost of filter papers, reagents, equipment maintenance, and human resources to be 8-10 USD which is about 1/2 to 1/4 of the cost of commercialized tests in Angola. Hands-on time required to perform the assay was comparable based on information taken from company websites and references [302-304] (Table 4).

Table 4. Comparison of cost and operational features of our in-house assay with commercial assays.

Features	**In-house EID assay**	**AMPLICOR™ HIV-1 DNA Test v1.5**	**Abbott RealTi*me* HIV-1 Qualitative**
Type of assay	PCR/qualitative	PCR/qualitative	Real Time PCR/Qualitative
Specimen volume	100-125 µl	200-500 µl	100-200 µl
Target of amplification	Pol (IN)	Gag	Pol (IN)
Genotypes detected	All subtypes, CRF02	Subtypes A-H	Subtypes A-H, CRF01, CRF02, Groups O and N
Analytical sensitivity (for DBS specimens)	112 copies/ml	Not available	839 copies/ml
Time for result	6-7 hours	7-8 hours	8 hours
Cost/test (USD)*	8-10	15-30	37
Number of samples/run	1-30 samples	9-21 samples	21-96 samples
Equipment required	Thermocycler, microcentrifuge, heat block, gel electrophoresis	Thermocycler, ELISA, reader/washer, microcentrifuge	M2000sp, M2000rt equipment
Equipment cost	14,000$	25,000$	150,000$

*Prices of tests performed with commercial assays vary considerably depending on quantities, infrastructure, support required and country of implementation.

Discussion

Children are among the most vulnerable to be at risk for HIV infection but in spite of this the AIDS response in sub-Saharan Africa has largely left them behind [30]. In this regard, Angola has only registered a 25% reduction of new infections among infants since 2009 [30] which led the Assistant Secretary-General of the United Nations to declare in 2015 that the epidemic in Angola might worsen if an effective AIDS response is not reinvigorated. EID of HIV-1 infection in infants at risk enables early treatment and care of the infected infants. Significant progress has been made in many sub-Saharan countries in implementing EID services following the introduction of HIV DNA testing on DBS [198]. However, the high cost of commercially available HIV-1 DNA tests and the perceived sensitivity problems related to the very diverse and complex viral strains circulating in Angola have prevented their implementation in this country. At the HDP where our cohort is based and in most other hospitals in Angola, pediatric diagnosis of HIV-1 infection is still done by serology at month 12 of life which significantly delays the initiation of treatment [305]. To support EID service expansion in Luanda we developed and validated a new HIV-1 DNA assay to be used on DBS samples. To account for the very diverse HIV-1 strains present in Luanda, we had to produce a new set of control plasmids containing the IN gene, our target for amplification, from local HIV-1 subtypes. Phylogenetic analysis showed that most of the new IN sequences fall at basal positions on the phylogenetic trees (pre-subtype branches) which is consistent with the complexity of the HIV-1 strains circulating in Luanda and with Angola being one of the epicenters of the HIV-1 epidemic [4-7, 98, 99, 212, 252].

To lower the costs, we used the Chelex method of DNA extraction which is also quick and easy to perform. This method had been previously applied to diagnostic screening in

HIV-1-exposed infants in Rwanda showing reliable results when used in combination with either an in-house nested PCR or the Roche Amplicor HIV-1 DNA assay version 1.5 [306]. For the purpose of the EID, the Chelex DNA extraction method represents a low-cost alternative to commercial kits such as the QIAamp™ DNA Investigator Kit which costs 6.8 USD per reaction in Portugal and at least twice that in Angola. The Chelex method that we have used costs less than 3 USD per reaction, is very quick to perform, and does not use hazardous solvents. Moreover, a recent study that compared the yield of DNA extracted from blood samples applied to Whatman™ FTA™ cards using different methods showed that the amount of DNA recovered with the Chelex procedure (2.5 ng from a blood spot of 6.0 mm) is similar to or larger than the amount of DNA recovered with the other methods [307].

Our nested PCR assay showed a very low LoD for all the complex HIV-1 genotypes that we used as controls, suggesting that it was appropriate for early diagnosis of HIV-1 infection in infants in Angola. Indeed, using this assay we could detect all HIV-1 infected infants at month 1 of life. The good performance of the assay was also demonstrated in HIV-1 infected adults where we could detect HIV-1 DNA in 14.3% of patients with undetectable viral load (plasma viral load of <20 copies/mL).

The low percentage of HIV-1 infected infants (2.2%) in the APEHC cohort between 2012 and 2014 contrasts with the rate reported in 2014 at a national level of 25% [30] and confirms the effectiveness of the WHO-based prevention program implemented since 2007 at the HDP [305]. The detection of HIV-1 infection in infants as early as 1 month after birth makes this new assay suitable to health care centers following option B+ of WHO guidelines that recommend EID at 4-6 weeks of life. Moreover, our test might be useful to determine HIV-1 infection status when serology results are indeterminate after 12-18 months of age.

One possible shortcoming of our study is that we could not make a head-to-head comparison of our assay with a commercial test because currently there is no adequate platform for EID testing from DBS in Angola. In fact, the EID national program implemented at the *Instituto Nacional de Saúde Pública* (INSP) with the support of the Clinton Foundation was discontinued in 2012. NucliSens® HIV-1 QT test (bioMérieux, Inc., Durham, NC) is still used at the INLS but its performance is severely affected by the genetic heterogeneity of HIV-1. As reported by several studies, the test has low accuracy in detecting or quantifying specific group M subtypes (A, C, F, and G), recombinants (CRF01_AE, CRF02_AG), and group O [308-311]. Another possible shortcoming is the limited size of the prospectively enrolled cohort and consequently the clinical evaluation of the assay. Thus, further study in the clinical setting is likely warranted.

Conclusions

The high analytical and clinical sensitivity of our EID assay have enabled accurate, early and low cost diagnosis of HIV-1 infection in exposed infants in Angola. The low percentage of HIV-1 MTCT case observed within the APEHC cohort is consistent with the current high standard of pediatric care provided at HDP. The simplicity and low-cost of the assay make it suitable for generalized implementation in Angola and other resource-constrained countries.

Competing interests

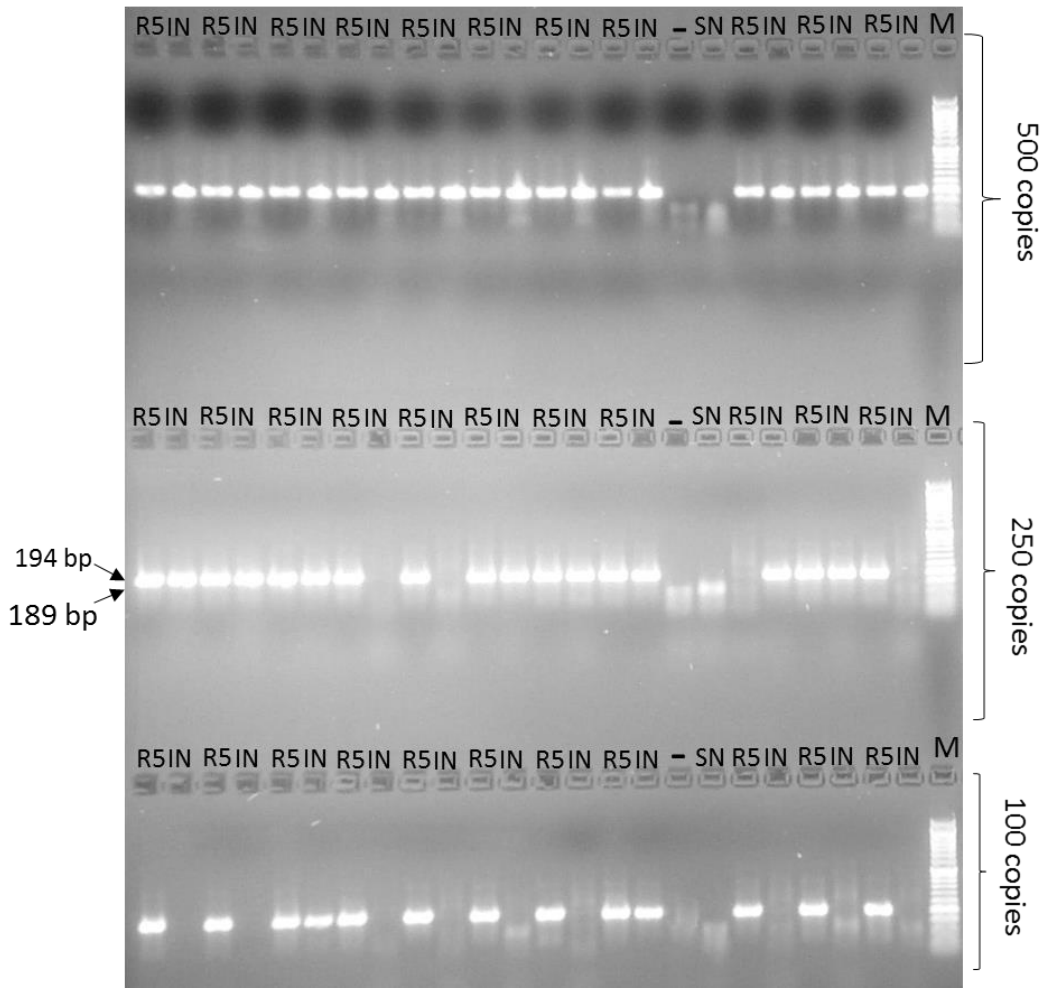
The authors declare that they have no conflicts of interest.

Acknowledgements

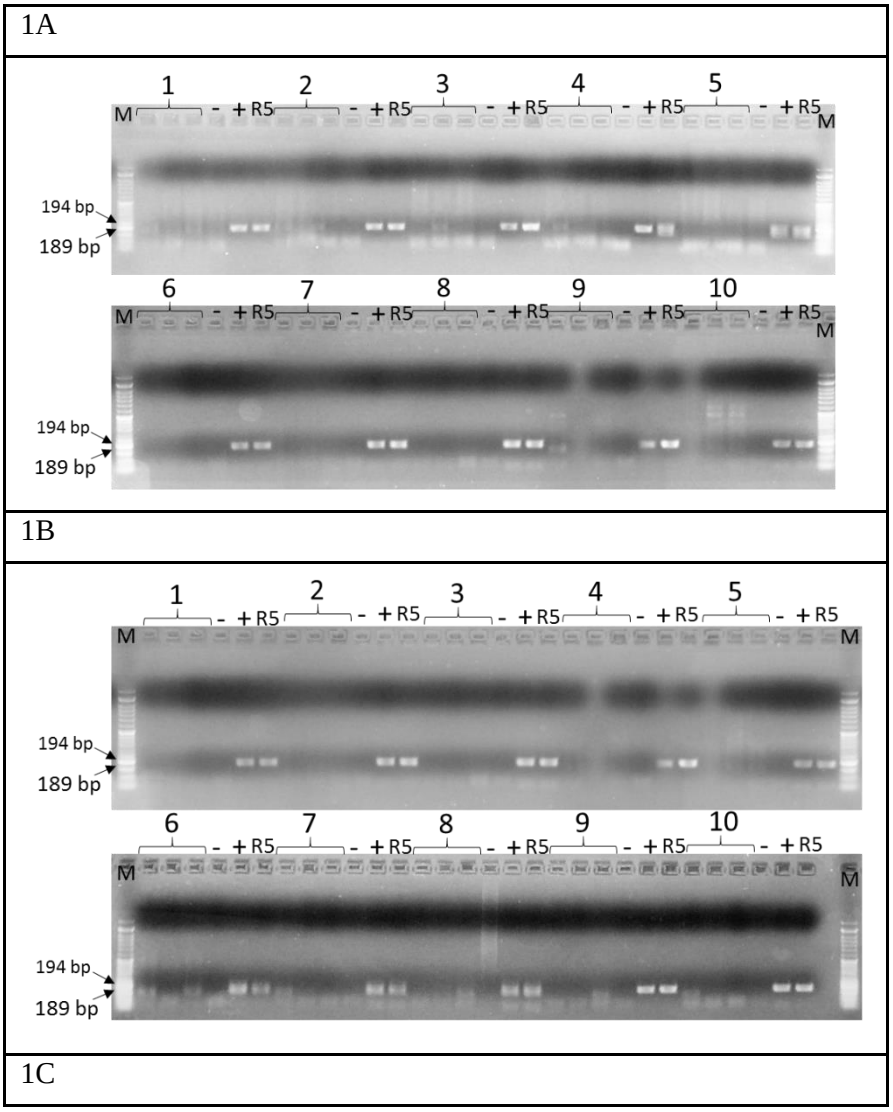
We greatly acknowledge the contribution and efforts of the staff involved at the *Hospital da Divina Providência* in Luanda for helping us to conduct the study. ACH-2 cells were

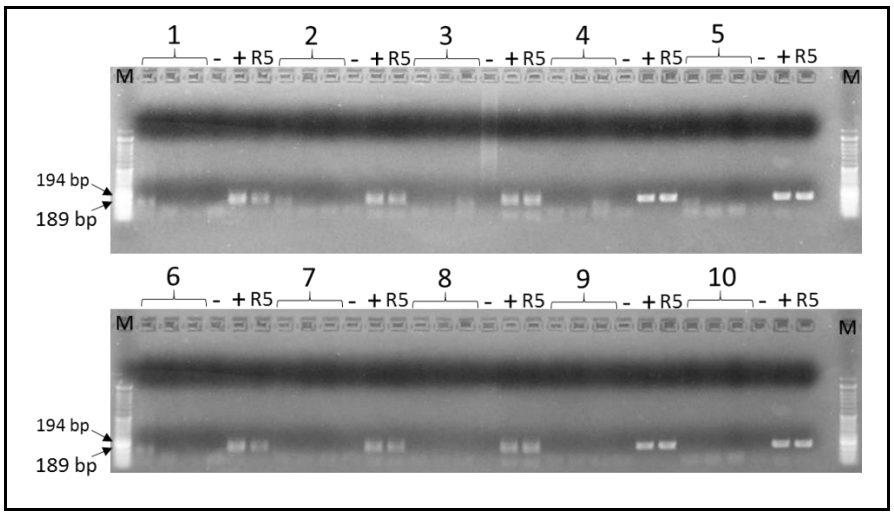
obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH. We appreciate the participation of all the patients without whom this study would not have been possible. Financial support was provided by the *Fundação para a Ciência e a Tecnologia* (FCT) Portugal (PTDC/SAU-EPI/122400/2010), part of the EDCTP2 programme supported by the European Union. Francisco Martin is supported by the Portuguese *Fundação para a Ciência e Tecnologia* (FCT) (grant number SFRH/BD/87488/2012). Claudia Palladino is supported by the Portuguese *Fundação para a Ciência e Tecnologia* (FCT) (grant number SFRH/BPD/77448/2011).

Supporting Information

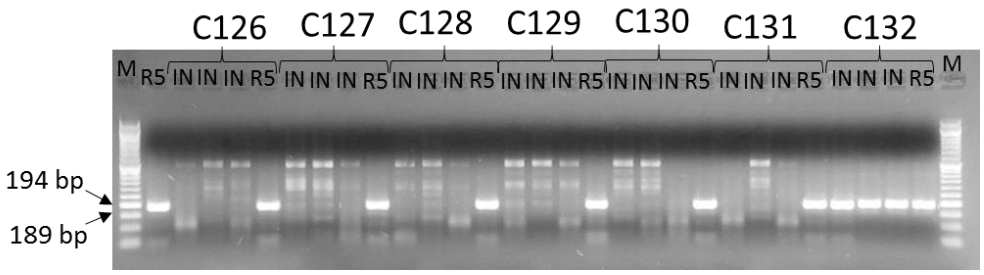


S1 Fig. Representative example of the results of the limit of detection (LoD) of the new PCR assay for subtype J control plasmid**.** 500, 250 and 100 copies of control plasmid were added to 125 ul of seronegative HIV blood and spotted in Human ID bloodstain cards. Extracted DNA was amplified by nested-PCR and amplified products were run on a 2% agarose gel with green safe staining. Each samples was amplified >10 times. (M) Molecular weight marker (NZY Leader VI); (IN) HIV-1 integrase fragment (194 bp); (R5) CCR5 gene fragment (189 bp); (SN) HIV-1 seronegative control; (-): ddH2O. The LoD for the subtype J was 4.3 copies/PCR (95% confidence interval: 3.3-6.3).





S2 Fig. Results of the diagnostic specificity experiments using the new PCR assay with samples collected from adult patients infected with HBV, HCV or CMV. Amplification of samples from patients infected with HBV, HCV or CMV using the new PCR assay conditions. A) Samples of HBV-infected patients (n=10); B) Samples of HCV-infected patients (n=10); C) Samples of CMV-infected patients (n=10). Samples were tested in triplicate. PCR products were run on 2% agarose gel and stained with green safe. (M) Molecular weight marker (NZY Leader VI); (-) HIV-1 seronegative sample; (+) HIV-1 seropositive sample (194 bp); (R5) CCR5 gene (189 bp).



S3 Fig. Representative example of the results obtained using the new PCR assay on samples collected from infants enrolled in the APHEC cohort. Each infant was assigned an anonymized code (C126-C132). C126-C131 are uninfected infants whereas C132 is an HIV-1 infected infant. Samples were tested in triplicate. Amplified products were run on a 2% agarose gel with green safe staining. (M) Molecular weight marker (NZY leader VI); (IN) HIV-1 integrase fragment (194 bp); (R5) CCR5 gene fragment (189 bp).

Table S1. Origin and genotype of the virus isolates used to produce the control plasmids.

Isolate	Accession number	Sampling date	Country of Origin (Province)	Patient	Genotype (IN gene)
93AOHDC249	KU296949	1993	Angola (Cabinda)	adult	A1
93AOHDC251	KU296950	1993	Angola (Cabinda)	adult	H
93AOHDC253	KU296951	1993	Angola (Cabinda)	adult	J
09AOHDP34	KU296952	2009	Angola (Luanda)	adult	C
09AOHDP110	KU296953	2009	Angola (Luanda)	adult	D
09AOHDP157	KU296954	2009	Angola (Luanda)	adult	G
09AOHDP237	KU296955	2009	Angola (Luanda)	adult	F1
01PTHDECJN	KU296956	1998	Portugal* (Lisbon)	infant	CRF02_AG
00PTHDEEBB	KU296957	2000	Portugal* (Lisbon)	infant	G

*Patients infected in Luanda

Table S2. Sequence and location of PCR primers used in this study and size of amplified.

This table reports the sequence and location of primers used for the amplification of the IN gene in clinical specimens, reference plasmids and ACH-2 cells.

PCR type	Primers	Sequence (5'-3')	Position in the HIV-1 HXB2	Band size (bp)
First-round PCR	F_IN_out	AACATAGTAACAGAYTCACARTATGC	4,029-4,055	1,553
	R_IN_out	TGGTCTTCTGGGGCTTGTTCCAT	5,582-5,559	
	APEHC_IN_F	AATTGGAGAGCAATGGCTAGTGA	4,281-4,303	194

Second-round PCR	APEHC_IN_R	CACTGGCTACATGGACTGCTAC	4,473-4,452	

Table S3. Limit of detection (LoD) of HIV-1 subtype B DNA in ACH-2 cells using probit regression analysis**.** This table relates to the determination of the LoD of the in-house EID molecular test in ACH-2 cells using probit regression analysis. The same principle was applied to the control plasmids in order to determine the LoD for the different subtypes tested.

Cells and provirus per DBS	**Cells and provirus per PCR**	**No. Detected (%)**	**Probit value**
5,000	50	10 (100)	NA *
1,000	10	10 (100)	NA *
500	5	8 (80)	5.84
250	2.5	6 (60)	5.25
100	1	3 (30)	4.48
50	0.5	2 (20)	4.16

* NA, not applicable.

Chapter 4

Long-term and low-level envelope C2V3 stimulation from highly diverse virus isolates leads to frequent development of broad and elite antibody neutralization in HIV-1 infected individuals

Francisco Martin[1], José Marcelino[1,2], Claudia Palladino[1], Inês Bártolo[1], Susana Tracana[1], Inês Moranguinho[1], Rita Mateus[1], Rita Calado[1], Pedro Borrego[1], Thomas Leitner[3], Sofia Clemente[4], Nuno Taveira[1,2#]

[1]Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisbon, Portugal; [2]Centro de Investigação Interdisciplinar Egas Moniz CIIEM, Instituto Superior de Ciências da Saúde Egas Moniz, Caparica, Portugal; [3]Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA; [4]Hospital da Divina Providência, Luanda, Angola;

Manuscript submitted to eBioMedicine (Available at medRxiv https://doi.org/10.1101/2022.01.27.22269759)

Keywords: HIV-1 evolution; genetic diversity; antibody neutralization; Angola

Abstract

Elicitation of potent neutralizing antibodies against genetically diverse HIV-1 isolates is important for an effective HIV-1 vaccine. Some HIV-1 infected patients produce such broadly neutralizing antibodies (bNAbs). Identification of host and viral correlates of bNAb production may help develop the next generation of HIV-1 vaccines. We carried out the first detailed characterization of the neutralizing antibody response and identify viral and host factors associated with the development of bNAbs in HIV-1 infected patients from Angola, one of the oldest, more dynamic, and diverse HIV-1 epidemics in the world. Plasma samples from 322 HIV-1 infected patients were collected in 2001, 2009 and 2014. Phylogenetic analysis of C2V3C3 envelope sequences identified a diverse array of subtypes including A1, A2, B, C, D, F1, G, H, J, untypable strains, and recombinant forms which prevailed over pure subtypes. Notably, 56% of the patients developed cross, broad, or elite neutralizing responses against a reference panel of tier 2 Env-pseudoviruses far exceeding results obtained elsewhere in the world. The frequency of elite neutralizers was higher in 2014, when patients were on ART and had low viremia, than in 2009 when patients were drug naive. In drug naïve patients, broad neutralization was associated with subtype C infection, lower CD4+ T cell counts, higher age, or higher titer of C2V3C3-specific antibodies relative to patients that did not develop bNAbs. Neutralizing antibodies targeted the V3-glycan supersite in most patients but antibodies specific for the V2 apex, the CD4 binding site, the gp41 membrane-proximal external region (MPER) and unknown epitopes were also found in some patients. V3 and C3 regions were significantly less variable and less subject to positive selection in elite neutralizers compared to weak or no neutralizers suggesting an active role of bNabs directed against these regions in controlling HIV-1 replication and diversification. Hence, development of broad and elite antibody neutralization against HIV-1 requires long-term

and low-level envelope V3C3 stimulation from highly diverse subtype C isolates. These results have direct implications for the design of a new generation of HIV-1 vaccines.

Introduction

The HIV-1 Env glycoprotein is highly immunogenic but, in general, the antibodies elicited by it during infection lack neutralizing breadth or potency against primary HIV-1 strains thus failing to inhibit viral replication in infected individuals [145, 146]. In natural infection only 5 to 30% of adult individuals develop broadly neutralizing antibodies (bNAbs) after several years of infection [92, 145, 146, 159, 161-163, 173, 181, 182, 189, 194, 196, 212, 312-319] and these bNAbs have little impact in the control of the infection due to the continuous capacity of HIV-1 to diversify and escape these antibodies [13, 159-163]. However, some recombinant human bNAbs supress viral replication in HIV-1 infected individuals [177-182], prevent human infection by some HIV-1 strains [14], and passive immunization in animal models can protect from infection and/or disease progression (reviewed in [320]). Therefore, bNAbs are promising tools to restrict HIV-1 transmission and control disease progression if they could be induced by vaccination. Unfortunately, so far, antibodies elicited by candidate immunogens and vaccines have shown a very limited ability to neutralize heterologous primary HIV-1 strains [9-20].

bNAbs target five highly conserved epitopes in the HIV-1 envelope: the CD4 binding site (CD4bs); the V2 apex; V3 glycan supersite; gp41 MPER, and gp41/gp120 interface [213, 214, 321]. However, the mechanisms underlying the elicitation of such antibodies by B cell populations are still largely unknown [215, 216, 223, 322]. Guiding the immune system to elicit such bNAbs remains a major challenge due to extremely complex antibody maturation pathways and high levels of somatic hypermutation (SHM) required by bNabs to acquire neutralization breadth (reviewed in [322]). An exception to this rule is the V3-glycan supersite bNAb lineage that does not require extensive antibody-affinity

maturation [314, 323] allowing their development in early stages of infection [181, 314, 315], and explaining their high prevalence in recently infected individuals [173]. Furthermore, V3-specific IgG binding and neutralizing responses in pregnant woman living with HIV-1 predict low risk of mother-to-child-transmission of HIV-1 [316]. In HIV-2 infected individuals the V3 loop is a dominant target of bNAbs such that V3 undergoes extensive sequence, conformational and functional alterations to escape antibody neutralization [212, 217, 324-326]. Such findings, together with the proved therapeutic value of V3-glycan supersite bNAbs [181], highlight this epitope as a key target for HIV vaccine design.

Understanding the mechanisms underlying the production of bNabs against HIV-1 in some individuals during natural infection is of crucial importance for the development of improved immunogens and immunization strategies. Gray et al. [165] showed that patients infected with HIV-1 clade C rarely produce antibodies binding to the 2F5 neutralizing epitope in gp41 suggesting a correlation between HIV-1 subtype and neutralizing response. However, other studies found limited to no impact of HIV-1 subtype in plasma neutralization, suggesting that HIV-1 group M subtypes and neutralization response evolved independently [157, 162, 163, 327-329]. More recently, in a large longitudinal Sub-Saharan HIV primary infection cohort, cross-clade plasma neutralization was strongly correlated with subtype C infection [173]. Also, Rusert et al. [146] found a strong association between plasma neutralization specificity and HIV-1 subtype, with subtype B viruses being more vulnerable to CD4-binding-site specific antibodies and non-B subtype viruses being more vulnerable to V2-glycan specific neutralizing antibodies. In this study, V3-glycan and MPER-specific neutralizing responses were independent of viral subtype. The differences observed between studies might be related with the different assay conditions used to assess neutralizing activity,

in particular with the selected indicator virus panel that should represent the global HIV-1 diversity and be standardized to allow inter-study comparison [162].

Considering that vaccine effectiveness will depend on the extent to which induced antibodies will neutralize the global diversity of circulating HIV-1 variants, it is important to characterize HIV-1 antibody responses in different epidemics and geographies. For example, we have shown recently that the frequency and level of binding antibody response to selected epitopes in the envelope transmembrane gp41 glycoprotein differ between HIV-1 infected patients from Germany, France, and Portugal that have different subtype distribution [330]. In Switzerland, data analysis of a large Swiss HIV Cohort showed that ethnicity was associated with bnAb induction being black participants more prone to develop bNAb responses [146].

The neutralizing antibody response of HIV-1 infected patients from Angolan has never been evaluated. Angolan HIV-1 epidemic is peculiar, as it is driven by all subtypes and multiple CRFs and URFs [5-7, 96-99, 212, 252]. In addition, because it is a very old epidemic, highly divergent and ancestral forms of the different subtypes are often present [5-7, 96-99, 331]. It has been suggested that the genetic complexity of the virus quaisespecies present in HIV-1 individuals is directly related to the development of neutralization breadth regardless of infection duration (reviewed in [171]). This should be particularly evident in old epidemics such as the one of Angola. Hence, characterizing the antibody responses and HIV-1 evolution in this population may provide new insights into the development and evolution of the neutralizing antibody response against HIV-1 and into vaccine design. Here, we carried out the first detailed characterization of the neutralizing antibody response against HIV-1 in Angola and identified viral and host factors associated with the neutralizing response.

Materials and Methods

Study population and ethics statement

This cross-sectional observational retrospective study included 322 HIV-1 infected adults. Plasma samples were collected in 2001, 2009 and 2014 at the Hospital da Divina Providência (HDP), a referral hospital in Luanda, the capital city of Angola. Eligible participants had ≥19 years of age, were not pregnant and had a serological diagnosis of HIV-1 [Determine HIV-1/2 (Abbott) and Uni-Gold Recombigen (Trinity Biotech) rapid tests]. Plasma viral load and number of CD4+ T cells were determined in a subset of patients using the Abbott Real Time HIV-1 assay (Abbott Laboratories) and the ABACUS 5 Junior Hematology analyser, respectively. The study was conducted according to the Declaration of Helsinki and was reviewed and approved by the National Ethics Committee of Angola. The study was verbally explained to all the patients before obtaining their written consent.

Cell lines

TZM-bl and HEK-293T cell lines were obtained from the NIH AIDS Reagent Program (https://www.niaid.nih.gov/research/nih-aids-reagent-program). TZM-bl cells were engineered from HeLa cells that constitutively express CXCR4 to express large amounts of CD4, CCR5 and a firefly luciferase reporter gene under the control of the HIV-1 LTR [332]. Cells were cultured at 37oC, 5% CO2 using Dulbecco minimal essential medium (DMEM) supplemented with 10 % heat-inactivated fetal bovine serum and with 100 units/ml of penicillin and 100 µg/ml of streptomycin.

Viral RNA extraction, PCR amplification, sequencing, and phylogenetic analysis

Before viral RNA extraction, 1 ml of plasma was centrifuged at 35.000rpm (61,793g) for 1h at 4ºC to concentrate viral particles. Supernatant was stored at -80oC for other applications and pelleted material was resuspended with 560µl of Buffer AVL+RNA carrier from the QIAmp® Viral RNA Mini Kit (Qiagen) and the manufacturer's protocol was followed. Reverse transcription was performed with NZY First-Strand cDNA Synthesis Kit (NZYtech, Portugal) and a 534bp fragment comprising the C2V3C3 env region was amplified by PCR using an in-house method described elsewhere [5, 7, 98, 99]. Sequencing of the C2V3C3 amplicons was performed with BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). Sequences were aligned using Muscle in MEGA version 6 software [211, 333] with reference strains collected from the Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov/). Maximum-likelihood (ML) phylogenetic analyses were performed using the best-fit model of nucleotide substitution as estimated by Modeltest v3.7 under the Akaike information criterion [263]. Maximum Likelihood (ML) trees were inferred with PhyML 3.0 [264, 334]. Tree searching was done with nearest neighbor interchange (NNI) and subtree pruning and regrafting (SPR). The reliability of the obtained topology was estimated with bootstrap with 500 replicates [264, 334]. Determination of coreceptor usage was made based on the V3 loop sequence using geno2pheno [coreceptor] webtool (https://coreceptor.geno2pheno.org/) [86]. False positive rates (FPR) used were 10% as recommended [191]. Selective pressure was examined with the DATAMONKEY web-server [335], after removing all positions containing gaps and missing data from the dataset. All estimations were performed using the MG94 codon substitution model crossed with the nucleotide substitution model GTR previously selected with Modeltest. Four different approaches were used to identify codons under selection: single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), internal fixed effects likelihood (IFEL) and relaxed-effects likelihood

(REL) methods [336]. While SLAC, FEL and REL detect sites under selection at the external branches of the phylogenetic three, IFEL identifies such sites only along the internal branches. To classify a site as positively or negatively selected the cut-off P-value was 10% for SLAC, FEL and IFEL. For REL, codons under selection were detected with a cut-off value for the Bayes factor of 50.

Entropy and N-linked glycosylation analysis

Potential N-linked glycosylation sites were identified using the N-Glycosite software [337], and the entropy at each amino acid position was measured with Shannon's entropy-one and Shannon's entropy-two online tools, all available at the Los Alamos National Laboratory HIV sequence database (http://www.hiv.lanl.gov/).

Production of C2V3C3 polypeptides and analysis of antibody reactivity

Six 178 amino acids polypeptides comprising the part of C2, V3 and part of C3 envelope regions (position 212–390 in gp120 in HIV-1 HXB2) of HIV-1 isolates circulating in Angola (subtypes C, G, H, J, and CRF02_AG) and Portugal (subtype B) were expressed in Escherichia coli and purified as described previously [212]. Briefly, a DNA fragment of 534 nucleotides comprising the C2, V3 and C3 coding regions (position 6858-7392 in HIV-1 HXB2) was amplified from plasmids containing the full-length envelope gene using the primers described elsewhere and cloned into the bacterial expression vector pTrcHis (Invitrogen) [212]. Expression of C2V3C3 polypeptides in Escherichia coli strain TOP10 was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) according to the manufacturer´s instructions, and protein purification was performed using Dynabeads® His-tag Isolation & Pulldown (Life Technologies). Bradford assay (Bio-Rad) was performed to determine protein concentration. Purified recombinant polypeptides were analysed by SDS-12% PAGE.

Antibody reactivity against these polypeptides was determined using an ELISA assay as described [212]. In brief, 96- wells ELISA plates were coated overnight at 4^0C with 100 µl of 0.05M bicarbonate-coating buffer (pH=9.4) containing 100 ng of the different C2V2C3 polypeptides. Plates were blocked with 2% gelatine in Tris-buffered saline (TBS) 1X for 1h at room temperature. Serial dilutions of plasma (1:100 to 1:3200) in primary antibody buffer (TBS 1X + 0.05 % Tween 20 + 5 % blocking solution) were added to the wells and the plates were incubated for 2h at room temperature. Goat anti-human IgG conjugated to alkaline phosphatase, diluted 1:2000 in the primary antibody buffer and 5% goat serum was added to the wells. Plates were washed between steps with TBSt (TBS 1X + 0.05 % Tween 20). Plates were developed by adding Sigma-fast p-nitrophenol phosphate diluted in deionized water. The plates were incubated for 20 minutes at room temperature away from light. Optical density was read at 405nm on a microplate reader. The clinical cut-off value of the assay was calculated as the mean OD value of HIV-seronegative samples plus 2 times the standard deviation (SD). Binding antibody titers were calculated as the highest plasma dilutions giving a positive reaction (OD / cut-off> 1).

Production of Env-pseudotyped viruses

A reference panel of 12 tier 2 HIV-1 Env-pseudotyped viruses of subtypes C (n=3), A (n=1), CRF07_BC (n=2), CRF01_AE (n=2), B (n=2), G (n=1) and AC recombinant (n=1) were produced using the Global panel of HIV-1 Env clones [162], obtained through the NIH AIDS reagent program. Env-pseudotyped viruses were produced by transfection of Env-expressing plasmids in 293T cells using PSG3.1Δenv as backbone, in a 1:3 ratio using JetPRIME® DNA transfection reagent. Viral stocks were filtered through 0.45 µm pore size filters after 48 hours and stored at -80°C until use.

Plasma neutralization assay

Neutralization of the Env-pseudotyped viruses was assessed in TZM-bl cells using Tat-induced luciferase (Luc) reporter gene expression to quantify the reduction in virus infection as described previously [338]. Briefly, TZM-BL cells (10,000 cells/well) were seeded the day before the neutralization assay to allow adherence of the cells to the bottom of the wells. Heat inactivated plasma samples (56^0C for 30 min) were incubated at 1:40 dilution in triplicate with the respective Env-pseudotyped virus for 1 hr at 37^0C before transfer to TZM-bl cells. Following 48 hours, percent neutralization was determined by calculating the difference in average RLU between test wells containing plasma samples and the wells containing the Env-pseudotyped virus from the indicator panel after the normalization of the results using the average RLU of cell control wells. Results were considered valid if the average RLU of virus wells was >10 times the average RLU of cell control wells. A virus pseudotyped with the envelope glycoprotein of vesicular stomatitis virus (VSV-G) was used as neutralization specificity control.

Neutralizing antibody titers were determined for a subset of plasma samples showing broad cross-neutralizing activity (n=38). In this case, 100 µL of 2-fold serial dilutions beginning at 1:40 were mixed with 100 µL of each virus (200 TCID50/well) and incubated for 1 h before adding to the cells. After 48 h, culture medium was removed from each well, and plates were analyzed for luciferase activity as described above. Wells with medium were used as background control, and virus-cell wells were included as infection control. Neutralizing titer (ID50) was defined as the highest dilution for which 50% neutralization was achieved.

Neutralization score and plasma categorization

To categorize the neutralizing activity of the Angolan samples in terms of potency and breadth we used a previously described scoring system [146, 167, 330]. A score of 0 was attributed when neutralizing activity against a given virus of the panel was less than 20%, a score of 1 when neutralization ranged between 20-<50%, a score of 2 for 50-<80% neutralization and a score of 3 for ≥80% neutralization. The overall neutralization score (NS) for a given plasma was obtained by adding the scores against the 12 Env-pseudoviruses of the panel and reflects neutralization potency and breadth. As a validated and worldwide accepted classification system to define neutralizing activity is lacking, for the purpose of the present study we classified plasmas with scores 25-36 as elite neutralizers, 18-24 as broad neutralizers, 6-17 as cross neutralizers and <6 as weak or no neutralizers. According to this classification an elite neutralizer sample must neutralize ≥9 viruses of the panel with a neutralization potency ≥80%.

Prediction of bnAb epitope specificities by clustering analysis

The neutralizing antibody specificities were determined for a subset of patients exhibiting broad and elite neutralization capacity using cluster analysis with human bNAbs targeting the main neutralizing epitopes on the viral envelope and capable to neutralize at least half of the 12 Env-pseudotyped virus panel as described previously [162]. Neutralization heatmaps and clusters were computed via the online tool ClustVis using a predefined correlation clustering distance method (Pearson correlation subtracted from 1) based on the average distance of all possible pairs. ClustVis is a web tool for visualizing clustering of multivariate data (available at https://biit.cs.ut.ee/clustvis/) [339].

Statistical analysis

The statistical analysis was performed with GraphPad Prism version 5.01 or 9.0 (GraphPad Software Incorporated, San Diego, California, USA). The Mann-Whitney,

Kruskal-Wallis or Fisher's exact tests were used to compare differences between groups. The Spearman rank test was used to quantify the magnitude and direction of the correlation between antibody neutralization activity and plasma binding titers against C2V3C3 polypeptides, CD4+ T cell counts, viral subtype and age of patients. Hypothesis tests were two-tailed and P values <0.05 were considered significant.

To test the potential correlation between neutralization score and genetic distance of the clinical samples to the neutralization panel viruses, we used amino acid sequences and Hamming distances that included gaps as characters because: 1) neutralization occurs on the amino acid level, 2) does not depend on the evolutionary path to a state-combination, and 3) indels may have significant effects on antibody binding. Genetic distances were calculated using DECIPHER [340], regression analysis was performed using R version R-4.0.3 [341], and visualization using ggplot2 [342].

Results

Characterization of the study population and infecting HIV-1 isolates

Overall, 375 plasma samples from 322 adult HIV-1 infected patients from three sampling years, 2001 (n=106), 2009 (n=210) and 2014 (n=59) were included in the analysis. Epidemiological, clinical, demographic, and virological characterization of the patients is given in Table S1. The median age of the patients was 34 years and most (n=242, 64.5%) were women. The main route of transmission was heterosexual contact (n=304, 81.1%). There were no significant differences in age and gender between sampling years. The median plasma viral load (VL) at the time of sampling was significantly higher in 2001 relative to 2009 (4.2-fold higher) and 2014 (33.5-fold higher). The median number of CD4+ T cells in 2014 was 1.8-fold higher when compared to 2009 (p=0.0015). The

significantly lower VL and higher CD4+ T cell number in 2014 is consistent with most patients being on cART which was not the case in 2001 and 2009.

Sequencing and phylogenetic analysis of the C2V3C3 Env region was completed successfully for 206 patients from 2001 (n=96/106, 90.6%) and 2009 (n=110/210, 52.4%). The following subtypes were identified: A1 (2001, n=33, 34.4%; 2009, n=32, 29.1%), A2 (2001, n=6, 6.3%; 2009, n=3, 2.7%), B (2001, n=2, 2.1%; 2009, n=2, 1.8%), C (2001, n=12, 12.5%; 2009, n=30, 27.3%), D (2001, n=2, 2.1%; 2009, n=8, 7.3%), F1 (2001, n=5, 5.2%; 2009, n=6, 5.5%), G (2001, n=8, 8.3%; 2009, n=11, 10%), H (2001, n=19, 19.8%; 2009, n=15, 13.6%), and J (2001, n=3, 3.1%; 2009, n=0, 0.0%). Untypable U strains were 4.2% (n=6) in 2001 and 2.7% (n=3) in 2009 (Figure S1). Subtype A prevailed in 2001 and 2009, but subtype C increased significantly (2.2-fold, P= 0.0095) in 2009. Out of the 176 isolates for which there were protease (PR) and C2V3C3 sequences available, 74 (42.0%) were non-recombinant and 102 (58.0%) were recombinant. Recombinant strains prevailed over pure subtypes in 2001 and 2009 (Table S2).

The genotypic analysis of tropism showed that most viruses were R5 in 2001 (82.3%, N=79) and in 2009 (85.5%, N=94), without significant differences between sampling years (Figure S2). Unfortunately, we could not sequence the C2V3C3 region from most of the 2014 samples due to their low or undetectable viral load (Table S1). Moreover, the lack of plasma prevented further analysis in samples collected in 2001.

Characterization of the antibody response

In total, 236 Angolan plasma samples were screened for neutralization breadth and potency against the 12 ENV-pseudotyped indicator panel, 178 samples from 2009 and 58 from 2014, amounting to 2832 plasma/virus combinations (Figure S3). In 2009, 80.9%

(144/178) of Angolan patients had the capacity to neutralize at least one virus from the indicator panel; this increased to 93.1% (54/58) in 2014 (Figure 1). Likewise, the mean percent neutralization (27.43%, 95%CI: 25.94, 28.92 in 2009 vs 60.52%, 95%CI: 57.37, 63.66 in 2014, p<0.0001) and the mean neutralization breadth (4.39, 95%CI: 3.83, 4.95 in 2009 vs 8.40, 95%CI: 7.32, 9.48 in 2014, p<0.0001) were higher in 2014 relative to 2009.

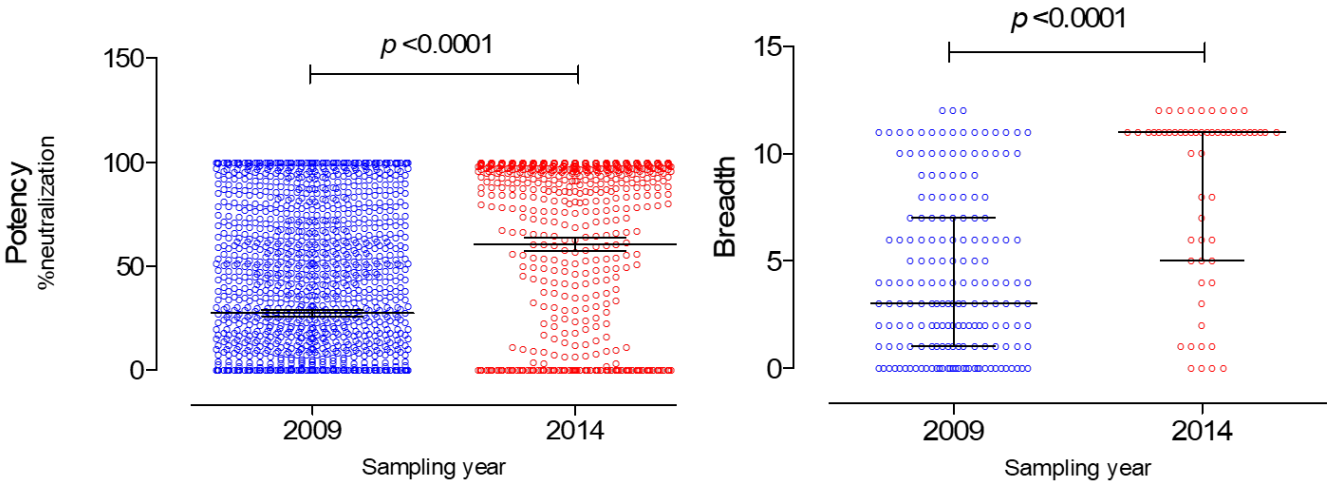


Figure 1– Neutralization potency and breadth per sampling year. A) Potency of neutralization (% neutralization at 1:40 plasma dilution) of samples collected in 2009 and 2014 as assessed against the 12-Env-pseudotyped virus indicator panel. Mean and 95% confidence intervals are shown. B) Neutralization breadth (number of Env-pseudotyped virus neutralized with a neutralization value >20%) in samples collected in 2009 and 2014. Median and interquartile range are shown. P values were obtained using the Mann Whitney U test.

Percent neutralization for each plasma-virus combination was recorded as a breadth-potency matrix: ≥80% neutralization received a score of 3, 50% to <80% a score of 2, 20% to <50% a score of 1, and <20% received a score of 0. Plasma samples were then ranked by the sum of scores in order to reflect their potency and breadth [146, 167]. Breadth, potency and neutralization score were directly correlated as expected (Figure S4).

Mean NS was 11.71, 95% CI [10.22, 13.19] ranging from 0 to 36 and median was 7, IQR [2.0, 21.0]. Remarkably, approximately 30% (n= 68/236) of the patients developed antibody responses with the capacity to potently neutralize at least half the viruses from the panel (Figure 2A). Overall, considering both sampling years, 18.6% (n= 44/236) of study participants were elite neutralizers (NS $\geq$ 25), 10.2% (n= 24/236) were broad neutralizers (18 $\leq$ NS < 25), 27.1% (n= 64/236) were cross neutralizers (6 $\leq$ NS < 17), and 44.1% (n= 104/236) were weak neutralizers or did not neutralize any virus of the panel (Figure 2B).

Correlates of the neutralizing response

The neutralizing antibody response has been previously associated with viral load, CD4+ T cell count, viral diversity and infection time [146, 173]. We first analyzed the impact of sample collection time on the neutralizing antibody responses of the HIV infected Angolan patients. Strikingly, median NS was 6.2-fold higher in 2014 relative to 2009 (31.00, IQR [10.50, 33.00] vs 5.00, IQR [1.00, 13.25], p<0.0001) (Figure 2A). Consistent with this, the frequency of elite neutralizers was 9.5-fold higher in 2014 than in 2009 [57% (n=33/58) vs 6% (n= 11/178), p<0.0001], and weak or no neutralizers were 2.7-fold more frequent in 2009 than in 2014 [52% (n= 93/178) vs 19.0% (n= 11/58), p<0.0001] (Figure 2B). Broad neutralizers were 2.4-fold more frequent in 2009 than in 2014 [12% (n= 21/178) vs 5% (n= 3/58), p=0.2107] and a similar trend was observed for cross neutralizers [30% (n= 53/179) vs 19.0% (n= 11/58), p=0.1270]. We also analysed matching plasma pairs from 2009 and 2014 to determine the evolution of neutralizing antibody response as a function of infection time. In line with the previous results, neutralizing score increased in 2014 relative to 2009 in 31 out of the 38 matched plasma pairs analysed (81.6%). (Figure 2C). The NS was unrelated with the sex of the patients (Figure 2D).

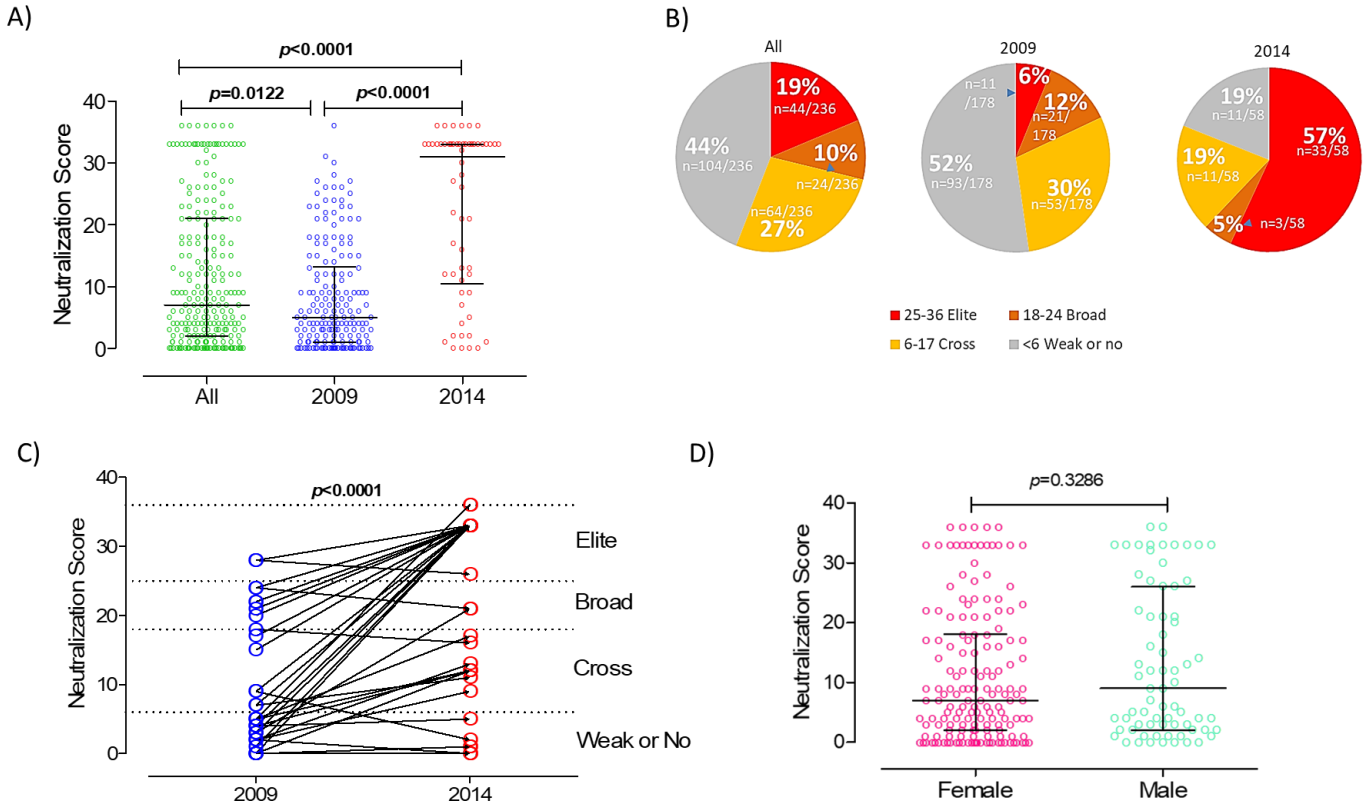


Figure 2- Neutralization score (NS) in HIV-1 infected patients from Angola as a function of year of sampling and sex. A) NS in 2009 is represented in blue and in 2014 in red; NS in all patients is in green. B) Angolan patients were categorized into 4 groups according to the NS as follows: no or weak neutralizers, <6 (grey); Cross neutralizers, 6-17 (yellow); Broad neutralizers, 18-24 (orange); Elite neutralizers, 25-36 (red). C) Neutralization score in matched samples collected in 2009 and 2014, showing the NS categories. D) NS in males and females. Median and interquartile range are shown. P values were obtained using the Mann Whitney U test.

The 50% neutralization titers (ID50) against the 12 Env-pseudotyped virus indicator panel were determined in a subset of plasma samples from 2009 (n=28) and 2014 (n=10) showing broad and elite neutralizing activity (Figure 3A). When comparing unmatched samples, neutralization titers were significantly higher in 2014 than in 2009 [median log10 ID50 in 2009= 1.903, IQ: 1.602-2.505 (n=336 plasma-virus pairs) vs median log10 ID50 in 2014= 2.204, IQ: 1.903- 2.806 (n= 120 plasma-virus pairs), p=0.0013] (Figure3 A/B)

A)

	Year	ID	Clade	Gender	Age	Number of CD4+ T cells/µl	HIV-1 Clade												VSV	Neutralization score (NS)
							C			A	G	AC	AE		B		BC			
							25710	CE0217	CE1176	398F1	1632	246F3	CNE8	CNE55	TRO11	X2278	X2000	CH119		
2009 (n=28)		33	F1	M	40	208	160	<40	40	80	40	40	80	<40	40	320	<40	40	<40	21
		43	G	F	24	NA	80	<40	80	80	<40	80	160	<40	40	<40	<40	40	<40	18
		53	C	F	45	NA	640	<40	1280	160	80	320	<40	320	<40	<40	160	320	<40	24
		54	F1	F	29	239	160	80	640	80	80	80	<40	<40	<40	40	160	320	<40	27
		61	C	M	42	109	640	640	640	2560	640	1280	<40	1280	640	320	2560	1280	<40	33
		72	A1	F	47	270	160	<40	<40	2560	<40	640	<40	<40	80	<40	40	<40	<40	19
		77	C	F	40	86	160	40	40	80	80	40	40	<40	40	320	40	40	<40	25
		103	C	M	38	181	80	40	40	40	<40	320	1280	80	40	40	40	160	<40	30
		104	ND	F	28	NA	40	<40	80	160	80	80	<40	40	<40	<40	40	160	<40	24
		122	H	M	52	NA	640	40	1280	40	40	<40	<40	<40	80	80	<40	160	<40	21
		125	G	F	26	306	640	40	640	40	80	80	640	<40	<40	80	<40	160	<40	17
		131	F1	M	43	272	320	40	640	<40	40	40	160	<40	<40	40	80	640	<40	24
		137	C	F	39	NA	80	80	40	80	40	80	320	80	40	80	<40	160	<40	29
		143	C	F	39	359	40	40	80	1280	40	80	80	<40	80	160	40	80	<40	28
		153	H	F	20	585	40	40	<40	640	80	<40	40	80	40	320	<40	<40	<40	18
		169	C	M	22	106	80	40	80	640	160	320	320	160	80	160	40	1280	<40	36
		175	C	F	21	488	40	<40	<40	40	80	1280	640	320	80	<40	<40	640	<40	22
		178	H	F	31	125	40	40	<40	<40	40	<40	320	160	160	<40	80	160	<40	25
		183	D	M	56	NA	80	80	40	640	40	1280	640	160	<40	160	<40	80	<40	26
		184	H	M	45	58	80	160	40	<40	80	640	320	160	320	320	80	80	<40	21
		192	H	F	42	248	80	<40	40	160	40	<40	40	<40	<40	320	40	<40	<40	18
		198	C	F	27	737	160	80	160	80	40	80	40	<40	40	80	40	320	<40	23
		199	G	F	27	402	320	40	40	640	<40	<40	40	40	40	80	40	80	<40	20
		205	C	F	30	85	320	40	<40	40	40	40	40	<40	<40	40	40	40	<40	23
		209	ND	M	31	ND	640	40	80	640	40	40	160	40	80	80	<40	80	<40	24
		278	G	F	30	27	80	160	<40	640	160	640	160	640	ND	<40	160	160	<40	27
		279	H	M	48	NA	1280	640	1280	640	40	1280	80	640	ND	640	320	640	<40	32
		293	G	F	39	234	40	<40	<40	40	40	<40	<40	1280	ND	40	<40	640	<40	21
2014 (n=10)		1	ND	F	39	555	40	320	160	1280	40	40	40	80	40	40	40	40	<40	36
		21	ND	F	32	NA	160	80	80	160	80	<40	80	160	80	160	80	<40	<40	30
		31	A*	F	38	209+	1280	640	1280	2560	1280	1280	1280	640	1280	2560	1280	640	<40	36
		37	G*	F	42	649	320	320	320	1280	40	40	40	640	40	40	40	40	<40	36
		42	G*	M	47	183+	160	40	160	320	160	<40	80	40	80	1280	320	40	<40	27
		50	C	M	51	326	640	160	640	640	640	160	320	160	320	640	640	320	<40	36
		57	A2*	F	48	NA	320	160	160	640	160	160	320	160	320	320	320	160	<40	36
		143	C	F	39	359+	160	<40	80	1280	<40	40	<40	<40	160	160	320	160	<40	26
		165	ND	F	44	228+	640	320	320	640	<40	<40	80	<40	40	<40	<40	<40	<40	36
		171	ND	F	51	NA	40	<40	<40	40	40	320	160	320	640	<40	320	640	<40	21

B)

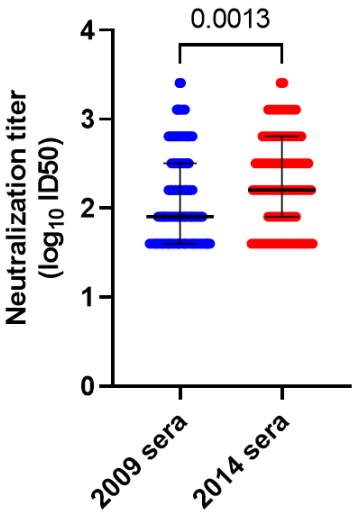


Figure 3- Antibody neutralization titers in a subset of unmatched plasma samples from elite and broad neutralizers from 2009 (n=28) and 2014 (n=10). A) Heatmap of the neutralization titers (ID50) and neutralization score against the 12 Env-pseudotyped virus indicator panel. ID50 values are color-coded, with darker colors implying higher ID50 values. *HIV subtype determined in the pol gene. + Number of CD4+ T cells determined in 2009; ND- not done due to lack of sample. B) Comparison of antibody neutralization

titers in 2009 and 2014. Log10 ID50 values obtained by each patient sample against the 12 Env-pseudotyped virus indicator panel are plotted. Lines indicate the median with interquartile range. P value was obtained using the Mann Whitney U test.

Overall, these results suggest that duration of infection is an important correlate of the potency and breadth of the neutralizing antibody response in Angolan patients infected with HIV-1.

To analyse the impact of HIV-1 subtype on neutralization by Angolan samples we compared the neutralization score (NS) in patients infected with subtypes C (n=27) and A1 (n=26), the two prevailing subtypes in Angola, and in patients infected with the other subtypes and recombinant forms (n=56). Only samples collected in 2009 were included in this analysis due to the limited number of samples genotyped in 2014. NS varied significantly with infecting virus subtype (p=0.014), with subtype C leading to significantly higher NS than subtype A1 [median NS= 17.00, IQR (6.00, 25.00) vs 6.00 IQR (3.50, 15.00), p=0.0103] or other subtypes [median NS= 17.00, IQR (6.00, 25.00) vs 8.00, IQR (1.00, 17.75), p=0.0087] (Figure 4A). These results indicate that virus subtype is a major determinant of the neutralizing antibody response in our patients.

The indicator virus panel used in the neutralization experiments contains three subtype C strains (25710, CE1176, and CE0217) that could be more closely related to the subtype C isolates from the Angolan patients and explain the higher NS observed in patients infected with subtype C viruses. To examine this issue, we compared the susceptibility of the reference panel isolates to neutralization and found a significant variation related to virus subtype (Figure 4B). The easiest viruses to neutralize were isolate 25710, a subtype C from India, and 398F1, a subtype A from Tanzania. On the other hand, viruses most resistant to neutralization were 2278, a subtype B from Spain and CNE8, a CRF01_AE

from China. Interestingly, subtype C isolates CE1176 and CE0217 from Malawi were significantly more resistant to neutralization than 25710 suggesting a closer relationship of subtype C isolates from Angola to this Indian subtype C isolate and showing that subtype per is not the main determinant of susceptibility to antibody neutralization. To investigate the impact of the evolutionary distance between the HIV-1 Angolan isolates and the indicator virus panel on the neutralizing antibody responses, we aligned the C2V3C3 amino acid sequences from the Angolan isolates (year 2009) with those from the indicator virus panel. As expected, subtype C viruses from the patients were more closely related with subtype C viruses of the indicator panel relative to other subtypes (Figure S5A). There was a significant negative correlation of amino acid distance of the indicator panel to NS (Spearman r= -0.2319, p = 0.019) (Figure S5B). Hence, the closer the isolate from the indicator panel was to the patient's C2V3C3 amino acid sequence, the easier it was neutralized. On average, clade C reference strain 25710 from the indicator panel was the closest indicator virus panel member to the Angolan isolates and, not surprisingly, it was the easiest virus to neutralize. At the other end, clade B reference strain 2278 was the furthest away from the C2V3C3 Angolan sequences and was the most difficult virus to neutralize along with the CRF01_AE virus (CNE8). Nevertheless, many patients infected with all subtypes developed potent bNAb responses despite the high genetic distance to the viruses of the indicator panel, indicating that other factors besides the relatedness with the indicator panel contribute to the potency of the neutralizing response.

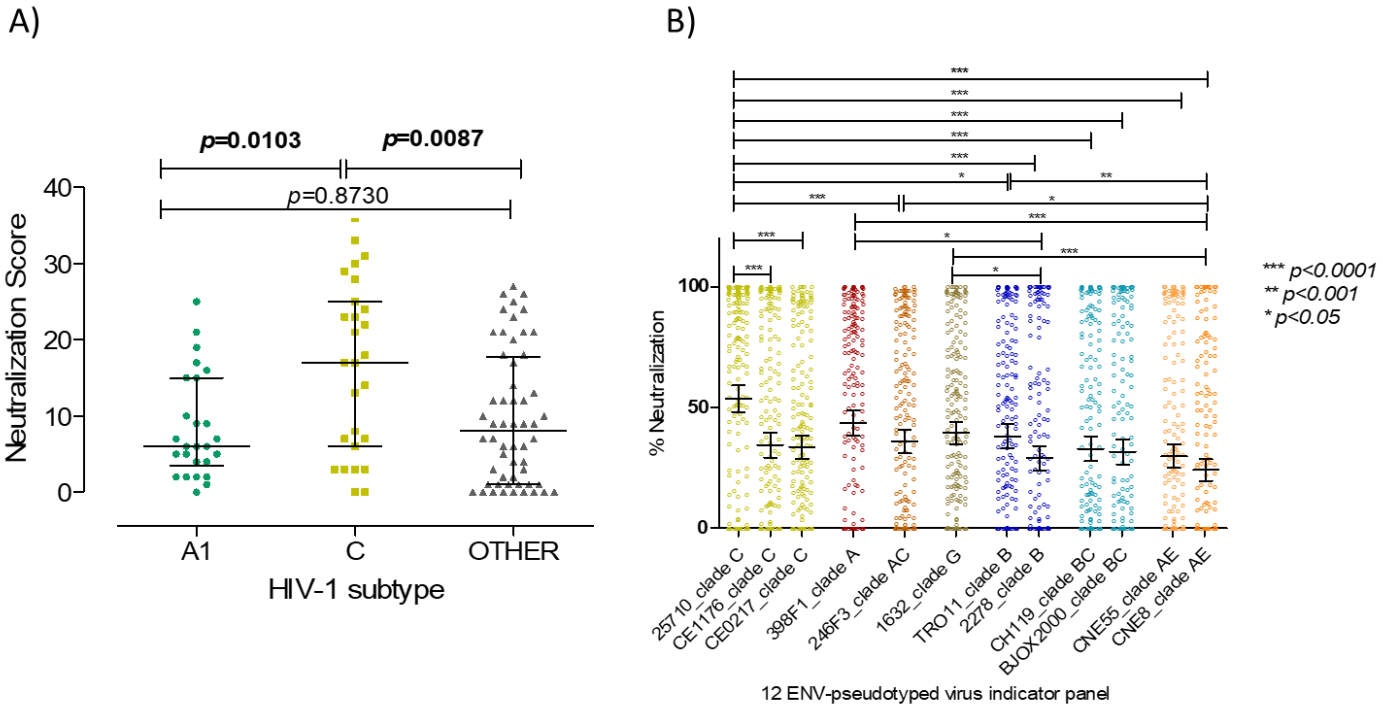


Figure 4– Impact of HIV-1 subtype on antibody neutralization. A) Neutralization score in patients infected with the two most common subtypes in Angola (year 2009), C (n=27) and A1 (n=26), and in patients infected with other subtypes and recombinant forms (n=56). The Kruskal-Wallis nonparametric test was used to analyse the difference in median NS for all subtypes (p=0.014). Dunns multiple comparison test was used to analyse differences in NS between subtypes. Median and interquartile range are shown. B) Percent neutralization of each of the 12 Env-pseudotyped virus indicator panel by the plasma samples (at 1:40 dilution) from the Angolan patients (N=236). Mean percent neutralization and 95% confidence interval bars against a given virus from the indicator panel are shown. Statistically significant differences are represented by the P values obtained with Dunns multiple comparison test. *** p<0.0001, ** p<0.001, *p<0.05.

Drug naïve patients (year 2009) with ≤200 CD4+ T cells/µl at study entry had significantly higher NS values than patients with >200 CD4+ T cells/µl [median NS in patients ≤200 CD4+ T cell counts was 7.00 (IQR, 3.50, 21.00) vs 4.00 (1.00, 12.00) in patients with > 200 CD4+ T cell counts, p = 0.0193] (Figure 5A). Moreover, NS values

were inversely associated with CD4+ T cell counts (Spearman r=-0.3043, p=0.0005) (Figure 5B) and directly associated with age (Spearman r=0.1644, p =0.0302) in these patients (Figure 5C). These results suggest that elicitation of high levels of broadly neutralizing antibodies in these patients is directly correlated with prolonged antigenic stimulation [343].

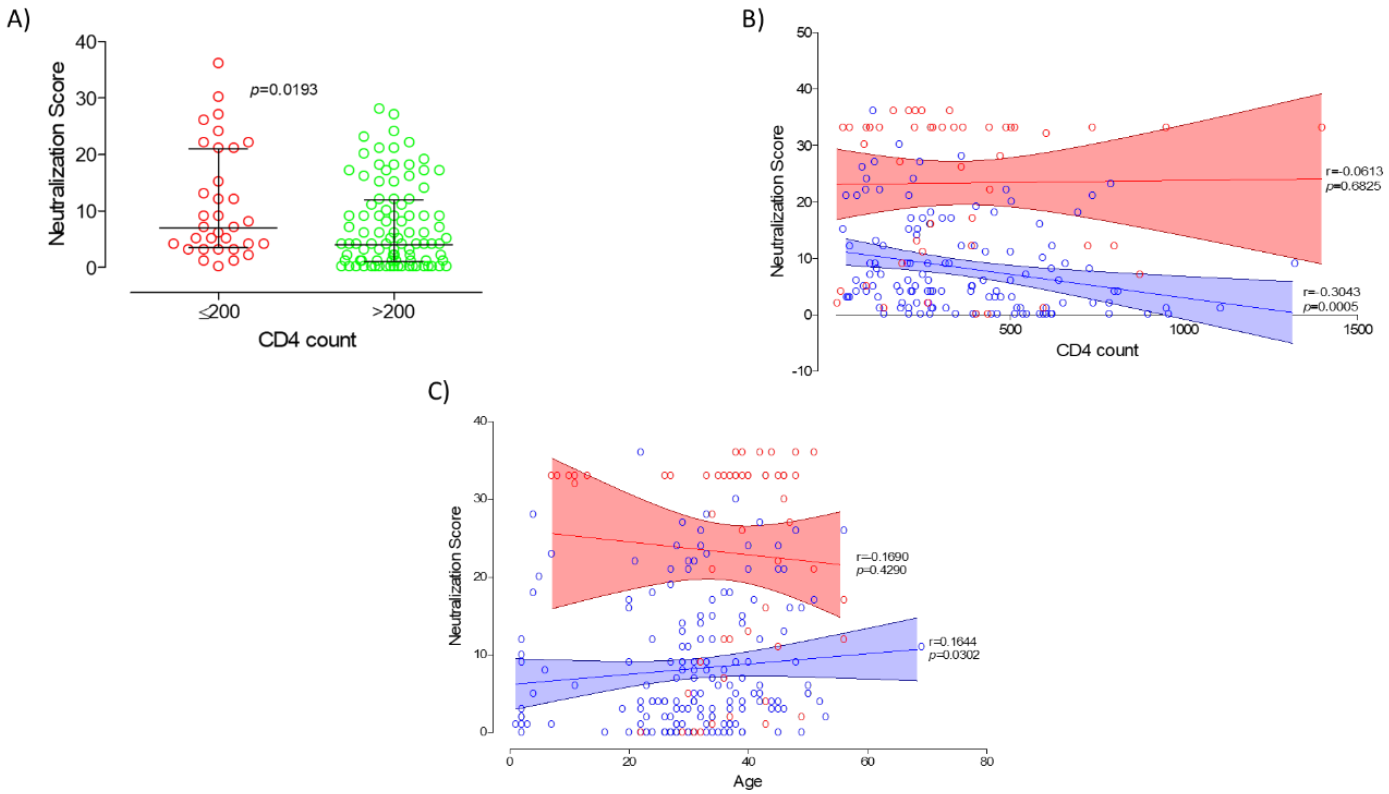


Figure 5- Correlation between neutralization score, CD4+ T cell counts and patient's age. A) Neutralization score differences between 2009 patient's with ≤ 200 CD4+ T cell counts at study entry and patients with >200 CD4 T cell counts. Median and interquartile range are shown. P values were obtained using the Mann Whitney U test; B) Correlation of neutralization score with CD4+ T cell counts in 2009 and 2014; C) Correlation of neutralization score with patient's age in 2009 and 2014. Samples collected in 2009 are shown in blue and samples collected in 2014 in red. Linear trend is shown with mean and 95% CI bands; Spearman r and P values are indicated.

Epitope specificities of the plasma neutralizing antibodies

In the same subset of 38 plasma samples (n=28 from 2009 and n=10 from 2014) from broad and elite neutralizers, the epitope specificities were mapped using a computational clustering tool based on the epitope specificities of a panel of human bNAbs [339]. Six (15.8%) samples did not cluster with any of the bNAbs. Thirty-two (84.2%) samples clustered with one of the bNAbs. Of these most samples (68.8%, 22/32) clustered with PGT128 and 2G12, two bnAbs that target the V3 glycan supersite with important contact residues in V3 and V4 (Figure 6) [188, 204, 328]. Five (15.6%) samples clustered with bnAb 4E10 that targets the gp41 membrane-proximal external region (MPER) [344]. Four (12.5%) samples clustered with VRC01 and VRC-CH31 bnAbs that target the CD4 binding site. Finally, one (3.1%) sample clustered with PG16 and PG9 that target the V1V2 glycans. These results indicate that the V3 glycan supersite is the dominant broadly neutralizing epitope in Angolan patients.

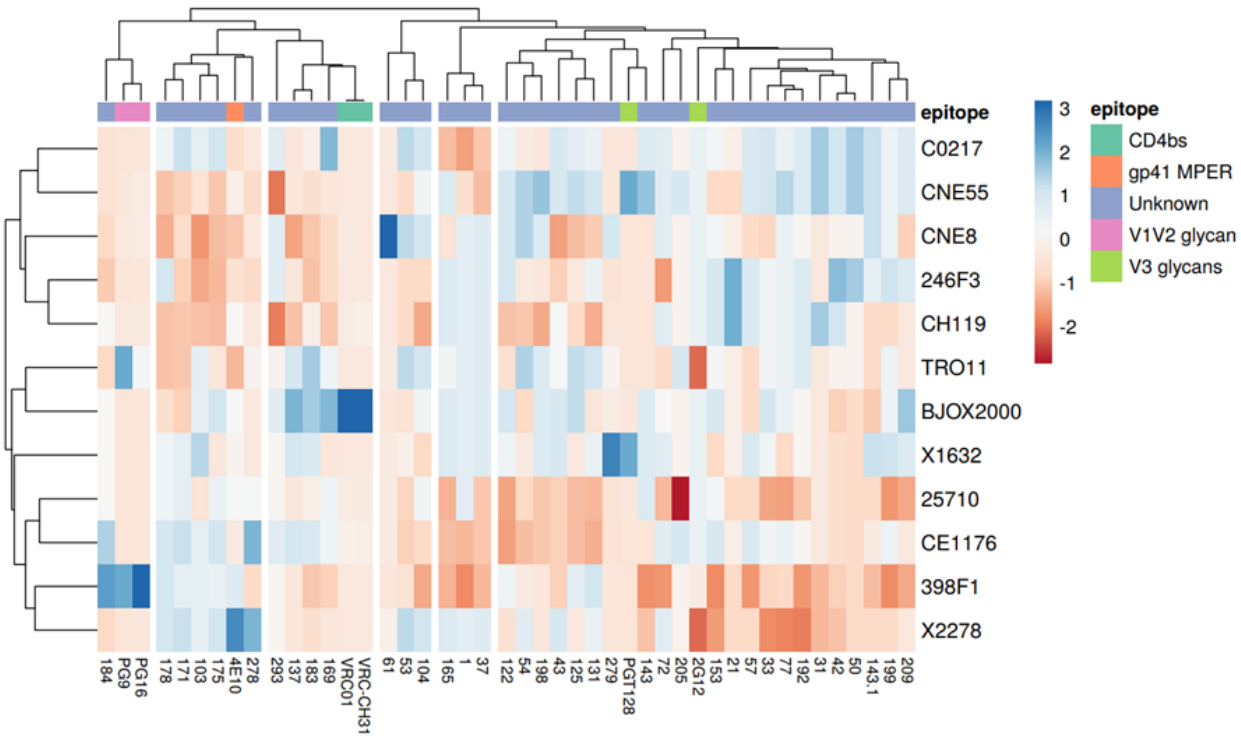


Figure 6- Cluster analysis and heatmap of the predicted epitope specificity in the top neutralizing patients from Angola. In the top of the columns, known bnAb epitopes are coloured according to the respective epitope specificities as shown by the legend. The

identification of the plasma samples and bnAbs is shown in the bottom of the columns. Cluster analysis for both rows and columns were computed according to the Pearson correlation [339]. Blue colours in the heatmap represent lower neutralization activity and red colours higher neutralization activity. Each column represents the neutralization values of a given plasma sample or a bnAb of known specificities against the 12 Env-pseudotyped virus panel whose names are indicated to the left.

Neutralization score is directly related with titer of C2V3C3-binding antibodies in all subtypes

The antibody binding reactivity against a panel of recombinant polypeptides comprising the C2, V3, and C3 envelope regions of subtypes B, C, G, H, J, and CRF02_AG was characterized in a subset of samples from 2009 (n=48) and 2014 (n=16) with known antibody neutralization profile. All but the B polypeptide were derived from Angolan isolates. All but six samples from five patients reacted with all C2V3C3 polypeptides demonstrating the high antigenicity of this envelope region (Figure S6). In 2009, patients had significantly higher median antibody binding titers against subtype C than against subtypes G (p=0.0007), H (p=0.0282), J (p=0.0052), and CRF02_AG (p=0.0149). Of note, median antibody binding titers were always higher in 2014 relative to 2009 regardless of the C2V3C3 polypeptide subtype but this was not significant except for CRF02_AG.

The higher antibody reactivity against subtype C antigen could be related with the higher neutralizing responses observed in subtype C infected patients. We therefore investigated possible associations between neutralization score, C2V3C3 antibody binding titer and subtype. Remarkably, C2V3C3 antibody binding titer was positively associated with NS values, i.e., patients with higher antibody binding titers to C2V3C3 polypeptides had higher neutralizing antibody responses (Figure 7). This was significant for all subtypes

of the C2V3C3 recombinant polypeptides tested confirming this epitope as an important neutralizing domain in these patients independent of subtype.

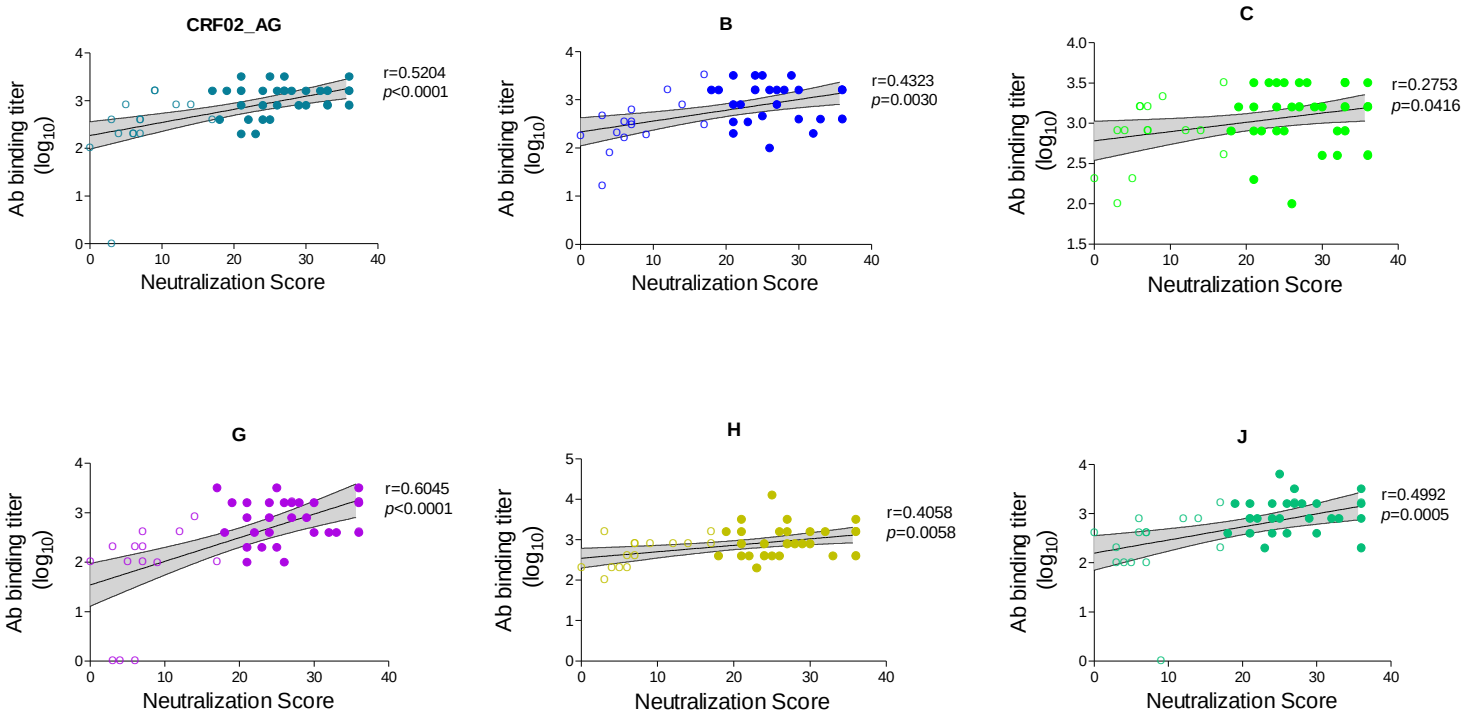


Figure 7– Association between antibody binding titer to C2V3C3 recombinant polypeptides of different subtypes and neutralization score (year 2009). Filled symbols are the antibody binding titers of the broad/elite neutralizers to a given C2V3C3 subtype. Unfilled symbols are the antibody binding titers of the no/weak and cross neutralizers to a given C2V3C3 subtype. Associations were assessed by Spearman analyses. P-values and Spearman r values are indicated. Linear trend is shown with mean and 95% CI bands.

Impact of the neutralizing antibodies in the diversity and evolution of C2V3C3

Neutralizing antibodies targeting the C2, V3 and C3 envelope regions are common in HIV-1 infected individuals [173] and escape from these antibodies leads to higher diversity in these regions as well as to higher positive selection and convergent evolution

[345-347]. To investigate the impact of the neutralizing antibodies in the diversity and evolution of the envelope glycoproteins of the viruses infecting our patients, we analysed amino acid entropy and the sites under selective pressure in the C2, V3 and C3 regions in the different neutralization categories for samples collected in 2009. Considering the three regions together, mean overall entropy values were similar in all neutralization categories: weak/no neutralizers= 0.5727 [95% confidence interval (CI): 0.4753, 0.6241]; cross neutralizers= 0.5931 (0.5012, 0.6849); broad neutralizers= 0.5381 (0.4472, 0.6291); and elite neutralizers= 0.4106 (0.3313, 0.4899). Regardless of neutralization category, the region with higher mean entropy was C3 [0.8528 (0.7556, 0.9500)] followed by V3 [0.4659 (0.3903, 0.5414)] and C2 [0.3635 (0.3092, 0.4178)] ($p<0.0001$). We then plotted Shannon's entropy differences in C2V3C3 between no/weak neutralizers and cross, broad and elite neutralizers. This analysis revealed that viruses from elite neutralizers were far less variable than viruses from weak/no neutralizers as seen by the number of amino acids with positive entropy differences relative to amino acids with negative entropy differences (37 vs 16 sites, $p=0.0023$) (Figure 8). On the other hand, viruses from broad and cross neutralizers did not vary significantly from viruses from no/weak neutralizers. Relative to no/weak neutralizers, the most variable amino acid residues in the broad and elite neutralizers were found in V3 and/or C3 (broad neutralizers: 1 site in C2 vs 7 sites in C3, $p<0.0001$; elite neutralizers: 3 sites in C2 vs 13 sites in V3C3, $p<0.0001$), two regions that contain broadly neutralizing epitopes (Figure 8).

Diversifying selection in C2V3C3 varied according to the different neutralization categories ($p<0.001$) (Table S3). Considering only sites that were selected by at least two methods, weak/no neutralizers had a total of 9 positively selected sites, cross neutralizers 6, broad neutralizers 3 and elite neutralizers 1. Regardless of neutralization category most sites under selective pressure were present in the C3 region.

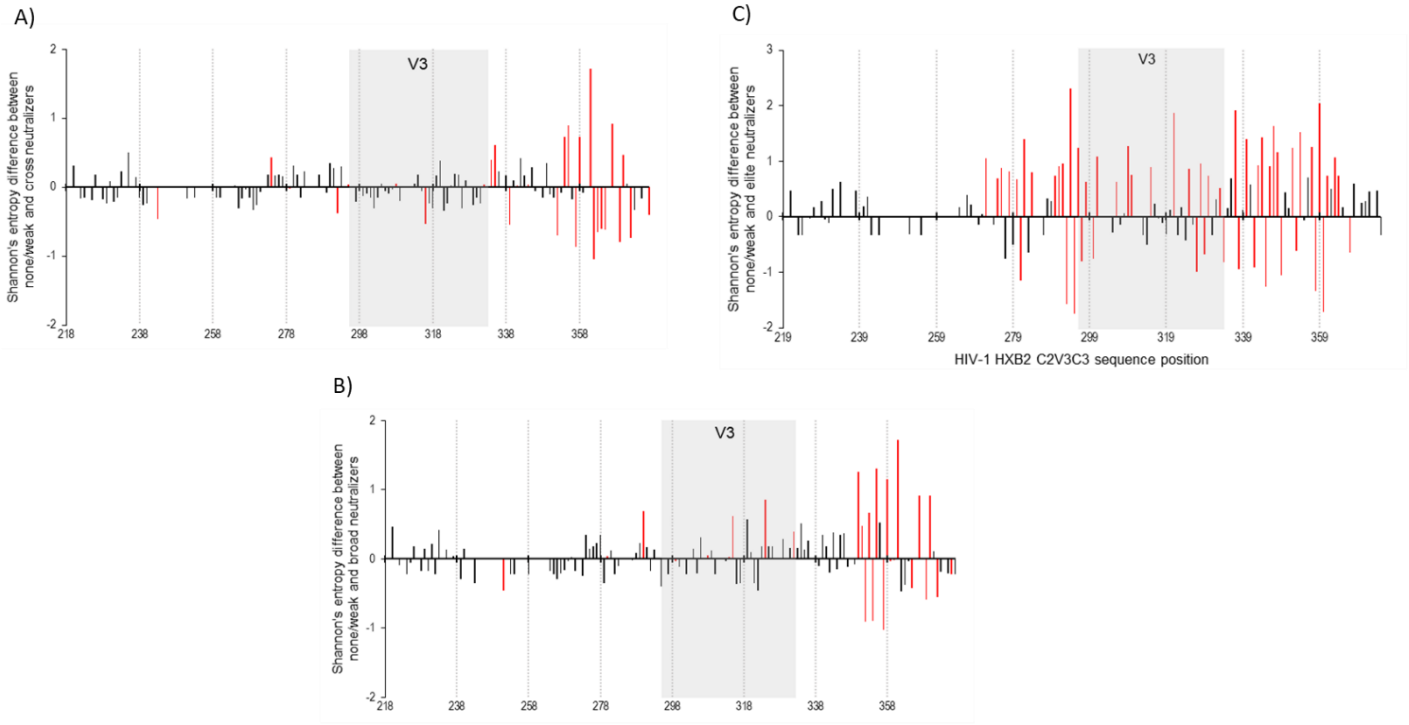


Figure 8 – Amino acid entropy difference in the C2V3C3 region between neutralization categories. A) Shannon's entropy difference between No/Weak and Cross neutralizers. B) Shannon's entropy difference between No/Weak and Broad neutralizers. C) Shannon's entropy difference between No/Weak and Elite neutralizers. Sites with significant entropy difference ($p \leq 0.05$) are shown in red. Gray boxes delimitate the V3 region. Numbers in the x axes indicate the amino acid position in HIV-1 HXB2.

The mean number of N-glycosylation sites in C2V3C3 was similar in all neutralization categories [Non-neutralizers: 9.2 (range: 8-12); cross-neutralizers: 9.2 (range: 7-11), broad-neutralizers: 9.4 (range: 7-11); elite-neutralizers: 9.8 (range: 8-11)] (Table S4). C3 had more potential N-glycosylation sites than C2 or V3 but sites in V3 and C2 were more conserved. For example, sites 241, 262, 276 and 289 in C3 were present in ≥70% of strains and site 301 in the V3 crown was present in all but two strains (97%). In C3, site 332, which together with site 301 in V3 and other elements in V1, V3 and V4 is part of

the V3-glycan supersite [170, 348], was highly conserved (70%) in all neutralization categories.

Discussion

HIV-1 was introduced in Angola from Kinshasa, the capital city of the Democratic Republic of Congo (DRC), likely in 1910-1940 making the Angolan epidemic the second oldest in the world [6]. Like in the DRC, the Angolan epidemic has been driven by all subtypes but B, untypable and highly divergent strains, and multiple CRFs and URFs [5, 96, 99, 252]. In this study we confirmed the extremely high diversity and evolving complexity of HIV-1 strains present in Angola. Subtypes A and C dominated over other subtypes but all other Env subtypes were present along with untypable basal strains and recombinant strains that prevailed over pure subtypes. The remarkable diversity and evolution of HIV-1 in Angola is driven by the increasing number of new infections [2], the limited access to antiretroviral therapy, and high levels of drug resistance [99, 349]. The high diversity and rapid evolution of HIV-1 in this country can pose a serious challenge to vaccination and other preventive efforts. At the individual level, the long-term B cell stimulation by this highly diversified ensemble of viruses may have promoted the development of exceptional neutralization breadth [146, 166, 170, 172, 173, 184, 350, 351]. We found that the majority (56%) of the patients in our cohort developed cross, broad, or elite neutralizing responses. These results far exceed those from previous cohort studies in sub-Saharan Africa [167, 172, 173, 176, 352]. For example, Beirnaert et al. found 10.6% broad neutralizers in Cameroon [352] and Landais et al. found about 15% broad neutralizers in a cohort of HIV-1 infected patients from Eastern and South Africa [173]. When compared to cohort studies from other geographies where subtype B dominates, the frequency of patients with bNAb responses reported in our study was also

much higher[146, 163, 353]. For example, Rusert et al. [146] in Switzerland found that most patients (79.1%) showed weak or no neutralization breadth, which compares to 44% in our cohort, and that only 1.3% were elite neutralizers which compares to 19% in our cohort. This divergence may be related with many factors besides the diversity of infecting viruses, such as the HLA genotype and ethnicity of the patients, viral load, CD4+ T cell counts, and duration of infection [146, 166, 172, 173, 354].

In agreement with other studies, neutralization score was inversely correlated with CD4+ T cell counts in 2009 when the patients were naïve to ART and had high viral loads [146, 166, 170, 173]. This is generally associated with high envelope stimulation of B cells and inevitably leads to B-cell exhaustion in chronic viraemic HIV-1 infection [355]. Remarkably, however, the frequency of elite neutralizers and the mean neutralization score in matched and unmatched patients increased significantly in 2014, when patients were already undergoing ART, relative to 2009. The boost in the quality of the neutralizing response in these patients suggest good restoration of the B cell compartment with ART which is uncommon in chronic HIV-1 infection [355-357]. Moreover, the moderate level of plasma viremia (median 11,660 HIV-1 RNA copies /ml, IQR, 380-30,060) found in these patients may have provided the low-level antigenic stimulation needed for the full maturation of memory B cells and bNAb production [355, 358, 359]. This has precedent in HIV-2 infection where most patients are infected for long periods and produce potent and broadly neutralizing responses in a setting of low plasma viremia [324, 360, 361]. In this model, B cell exposure to low-level envelope antigens, likely in lymphoid tissues, during prolonged infection periods leads to the generation of highly specific envelope C2V3C3- specific antibodies as well as broad and potent neutralizing antibodies [317].

Viral type and subtype as well as the nature of the epitope target on the viral envelope impact the antibody maturation process as seen by the frequency of elicitation [212, 361] and epitope specificity of bnAbs [146, 171, 348]. Differences in envelope structure and epitope exposure, length of variable loops, type of V3 motifs, N-glycosylation patterns, and conservation of key sites have helped to explain why certain HIV-1 subtypes like subtype C are better at promoting the elicitation of neutralizing antibodies [146, 171, 173, 176, 212, 348, 362]. In line with these studies, we found that infection with subtype C viruses was associated with enhanced neutralization breadth and potency. In general, subtype C infected individuals have shown a bias to V2-glycan directed antibody responses, and subtype C envelope from transmitted viruses have been less prone to neutralization by V3-directed antibodies due to the absence of the N332-glycan in the C3 region [146, 173, 176, 348, 363-365]. This was not the case in our study as most top neutralizers had V3-directed antibodies that were able to neutralize the subtype C isolates from the virus panel, and the N301 and N332 glycans defining the V3-glycan supersite were highly conserved in the patient's isolates. Supporting the major role of the V3 and C3 envelope regions in the development of bNAbs in our cohort, we found a strong direct correlation between the titer of antibodies binding to C2V3C3 envelope polypeptides from all subtypes and neutralization score. Nonetheless, antibodies specific for the V2 apex, the CD4 binding site, the gp41 MPER and/or unknown epitopes were also found in some patients revealing the complexity of the neutralizing antibody responses in these patients.

We also looked at the variability of patient's sequences in the envelope C2V3C3 region to assess the impact of escape from neutralizing antibodies on viral evolution and diversity. V3 and C3 were the most variable regions which is consistent with the dominant role of neutralizing antibodies targeting these regions in these patients [173]. V3 bNAb

recognition sites and sites associated with resistance to neutralization such as N295 [166, 170, 173] were under positive selection in broad and elite neutralizers. However, elite neutralizers exhibited far less variability and lower number of sites under selective pressure in V3 and C3 relative to weak or low neutralizers. The convergence of the viral swarm to a few resistant strains provides convincing evidence for the crucial role of V3- and C3-directed bNAbs in controlling HIV-1 replication and diversification in these patients [176, 212, 316, 366].

In conclusion, an exceptionally high number of Angolan patients infected with HIV-1 elicits broad and elite neutralizing antibodies mostly targeting the V3-glycan supersite. This is associated with long-term and low-level V3- and C3- antigenic stimulation by the highly diverse isolates circulating in this country especially subtype C. These results have direct implications for the design of a new generation of HIV-1 vaccines.

Acknowledgments

We greatly acknowledge the contribution and efforts of the staff and patients from the Hospital da Divina Providência in Luanda for this study. TZM-bl cells were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH. This work was supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal, under project grants UIDB/04138/2020 and UIDP/04138/2020. This study was in part supported by the NIH/NIAID under grant R01AI087520. Francisco Martin was supported by FCT under PhD grant number SFRH/BD/87488/2012. CP is funded by FCT under a contract-program as defined by DL No. 57/2016 and Law No. 57/2017.

Supporting Information

Table S1- Characteristics of the HIV-1-infected patients.

Characteristics	2001	2009	2014
Total number of patients, n (%)	106 (28.3)	210 (56.0)	59* (15.7)
Age (years) median (IQR)	32 (26-40)	32 (28-39)	39 (36-46)
Sex, n (%)			
Female	53 (50.0)	145 (69.0)	44 (74.6)
Male	43 (40.6)	65 (31.0)	15 (25.4)
Unknown	10 (9.4)	--	--
Geographic origin, n (%):			
Angola	94 (88.7)	207 (98.6)	57 (96.6)
RDC	--	3 (1.4)	1 (1.7)
Unknown	12 (11.3)	--	1 (1.7)
HIV-1 mode of transmission, n (%):			
Heterosexual	38 (35.8)	210 (100.0)	56 (94.9)
Bisexual	4 (3.8)	--	--
IDU	1 (0.9)	--	--
Transfusion	1 (0.9)	--	--
Unknown	62 (58.5)	--	3 (5.1)
CD4+ T cell count	--	N=162	N=21
CD4+ T cell count/mm^3, median (IQR)	N/A	265 (133-448)	475 (343-569)
Plasma viral load	N=16	N=71	N=13
VL (copies/ml), median (IQR)	390,877 (209,172-704,286)	93,391 (28,222-510,579)	11,660 (380-30,060)
Undetectable, n (%)	--	--	N=9 (69.2)
Unknown, n (%)	90 (84.9)	139 (66.2)	46 (78.0)
Co-morbidities, n (%):			
Tuberculosis	--	33 (15.7)	--
HBV	--	17 (8.1)	--
TB+HBV co-infections	--	3 (1.4)	--
Other	--	48 (22.9)	--
WHO Clinical stage, n (%):			
Asymptomatic	13 (12.3)	--	--

Symptomatic intermediate	20 (18.9)	--	1 (1.7)
AIDS	11 (10.4)	--	1 (1.7)
Unknown	62 (58.5)	210 (100.0)	57 (96.6)
cART, n (%):			
cART-naïve	102 (96.2)	202 (96.2)	1 (1.7)
cART-exposed	4 (3.8)	1 (0.5)	21 (35.6)
Unknown	--	7 (3.3)	37 (62.7)

Legend: N/A, not available; RDC, Republic Democratic of Congo; IDU, intravenous drug user; VL, viral load; cART, combined antiretroviral therapy; IQR, interquartile range; HBV, Hepatitis B virus. *53/59 HIV-1 infected patients were followed longitudinally from 2009.

Table S2- Main C2V3C3 subtypes in Angola in 2001 and 2009

Genetic forms	**2001** **N (%)**	**2009** **N (%)**	**P value** [a]
Pure subtypes	35/88 (39.8)	39/88 (44.3)	0.6470
Recombinant forms	53/88 (60.2)	49/88 (55.7)	
Subtype A	33/96 (34.4)	32/110 (29.1)	0.4542
Subtype C	12/96 (12.5)	30/110 (27.3)	**0.0095**
Subtype H	19/96 (19.8)	15/110 (13.6)	0.2625

[a]Fisher's exact test

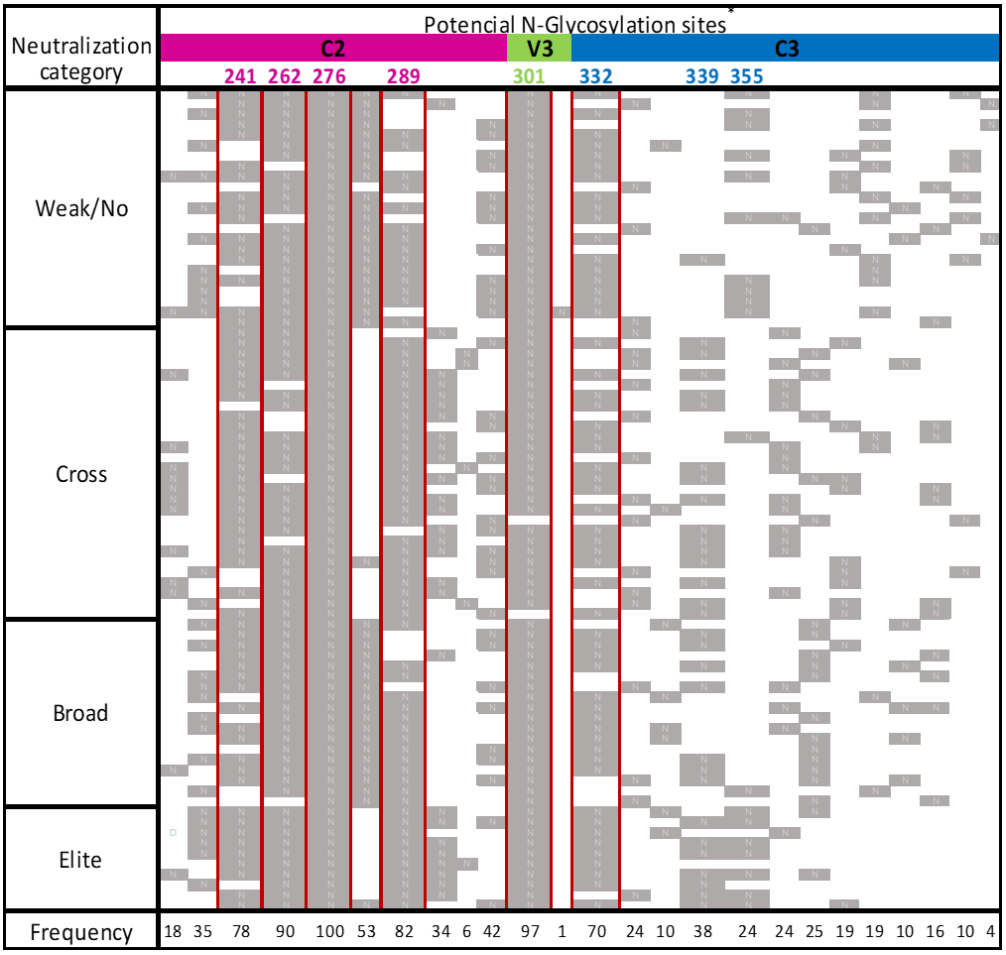
Table S3- Positively selected sites in the C2, V3 and C3 regions in the four neutralization categories selected at least by two methods

Neutralization category	Codon*	SLAC	*p-value*	REL	PP	FEL	*p-value*	IFEL	*p-value*
no/Weak	293	3.673	0.101	**1.191**	1.000	1.028	0.254	**4.219**	0.059
	335	**3.816**	0.048	**1.612**	0.993	**0.809**	0.036	**1.376**	0.016
	336	2.793	0.139	**1.307**	1.000	3.023	0.105	**8.145**	0.020
	343	**3.820**	0.092	**1.237**	1.000	0.418	0.650	0.316	0.747
	344	1.933	0.277	**1.457**	0.997	**0.599**	0.053	0.067	0.853
	346	2.924	0.134	**1.386**	1.000	**1.587**	0.046	**0.781**	0.089
	347	**4.361**	0.039	**1.414**	1.000	**0.980**	0.100	0.432	0.440
	361	**3.785**	0.097	**1.267**	1.000	0.540	0.715	0.420	0.754
	362	**4.384**	0.034	**1.307**	1.000	1.015	0.140	0.344	0.562
Cross	318	**2.646**	0.085	0.496	0.647	**0.317**	0.053	0.000	1.000
	336	**2.582**	0.096	0.719	0.889	**0.735**	0.045	0.522	0.152
	337	**3.791**	0.041	0.721	0.887	**0.520**	0.060	0.132	0.469
	344	2.185	0.188	**0.852**	0.989	**0.599**	0.012	**1.184**	0.095
	346	**4.272**	0.009	**0.862**	0.991	**0.587**	0.002	**0.366**	0.066
	365	**2.034**	0.053	-0.040	0.068	**0.210**	0.038	0.000	1.000
Broad	335	**2.928**	0.055	0.935	0.835	**0.999**	0.047	0.000	1.000
	347	**3.383**	0.034	0.927	0.830	**1.282**	0.031	0.603	0.409
	363	**2.908**	0.068	0.583	0.609	0.978	0.274	**4.188**	0.027
Elite	295	2.067	0.138	**4.302**	1.000	**10.915**	0.020	8.826	0.724

*Codons selected with 10% level of significance (SLAC, FEL and IFEL) or above a Bayes Factor of 50 (REL) selected by at least 2 methods and numbered according to codon position of HIV-1 HXB2. PP, posterior probabilities. Codons selected

simultaneously by SLAC, FEL and REL are bold and underlined. Bold dN-dS differences correspond to significant P-values or posterior probabilities.

Table S4- Frequency and distribution of potential N-glycosylation sites in the C2, V3 and C3 regions across neutralization categories.



*Relevant N-glycosylation sites are highlighted and coloured according the position in C2V3C3. Higher frequency glycosylation sites are boxed in red. Sites were numbered according to the reference strain HIV-1 HXB2.

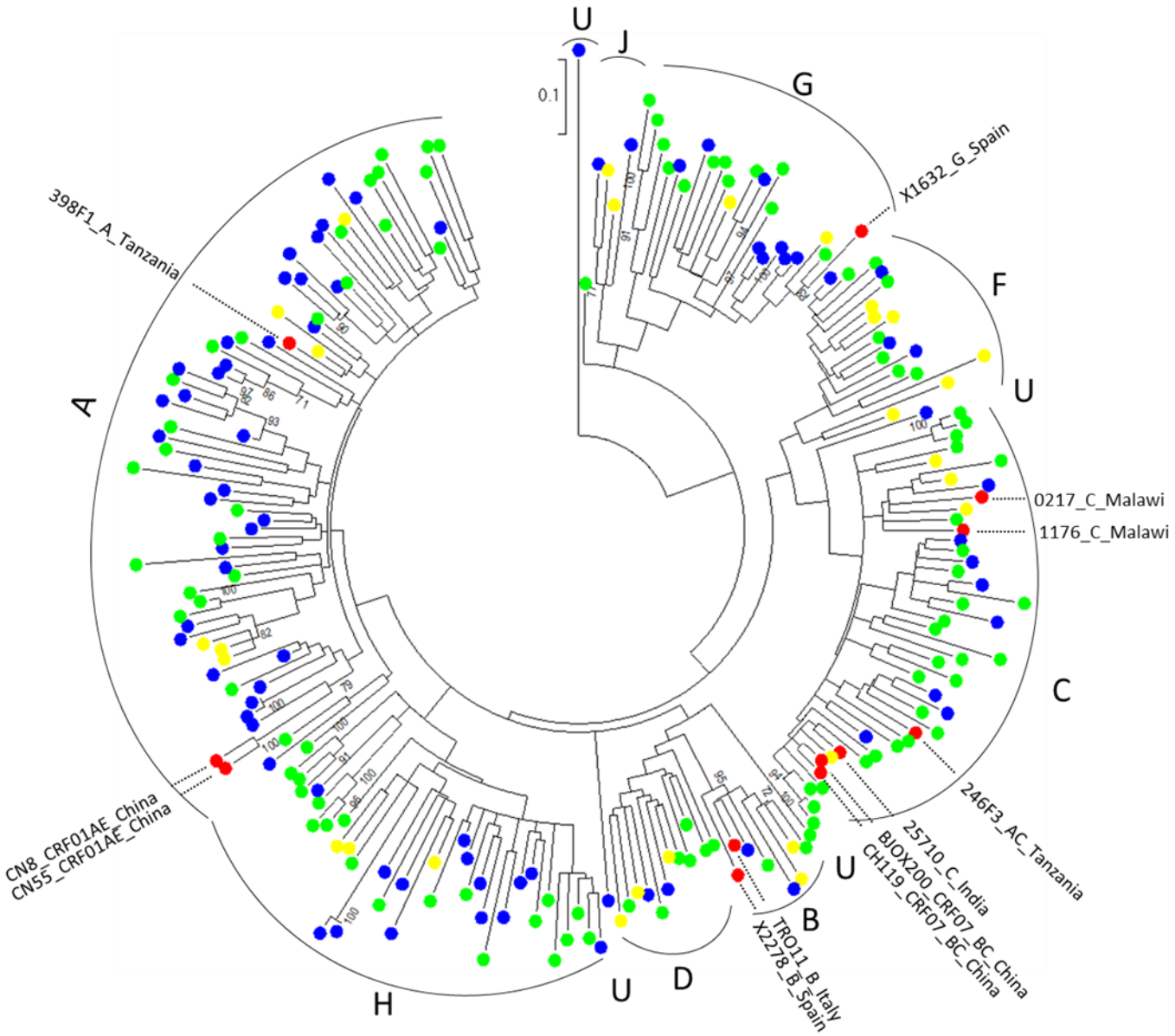


Figure S1- Phylogenetic relationship between the Angolan HIV-1 C2V3C3 sequences. Maximum likelihood phylogenetic tree of C2V3C3 region was constructed with reference sequences from all HIV-1 subtypes (yellow dots) with the 2001 (blue dots) and 2009 (green dots) Angolan sequences and the virus sequences from the indicator panel (red dots).

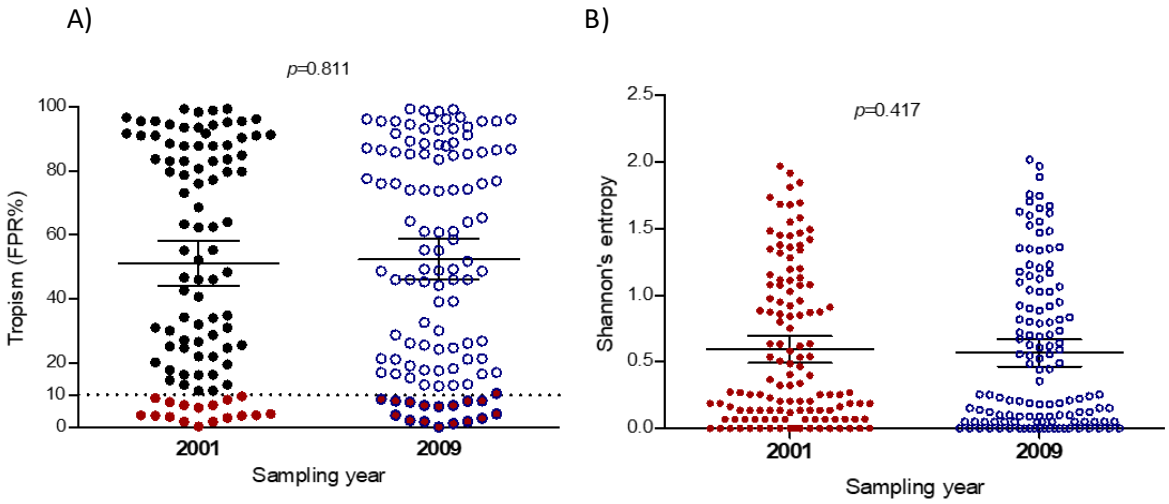
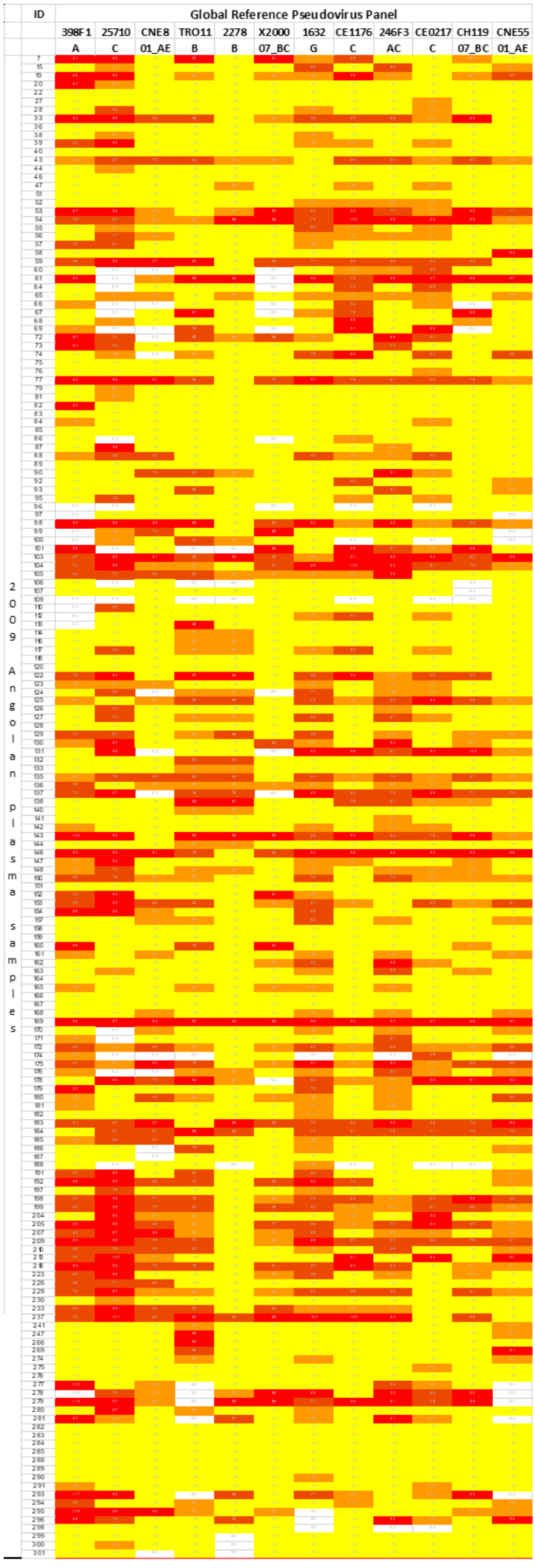


Figure S2- HIV tropism and amino acid diversity in 2001 and 2009. A) HIV V3-based tropism as determine in geno2pheno considering a false positive rate cut-off of 10%. The red dots are predicted X4 tropic virus. B) C2V3C3 amino acid diversity as assessed by Shannon's entropy. Mean entropy in the C2V3C3 region for each patient is shown, 2001 samples are represented in red filled dots and 2009 in blue unfilled dots. Variability at the amino acid level was calculated using Shannon's entropy-one online tool (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html). Mean and 95% confidence intervals are represented. P values were obtained using the Mann Whitney U test.


ID
Global Reference Pseudovirus Panel
398F1 A
25710 C
CNE8 01_AE
TRO11 B
2278 B
X2000 07_BC
1632 G
CE1176 C
246F3 AC
CE0217 C
CH119 07_BC
CNE55 01_AE
2009 Angolan plasma samples

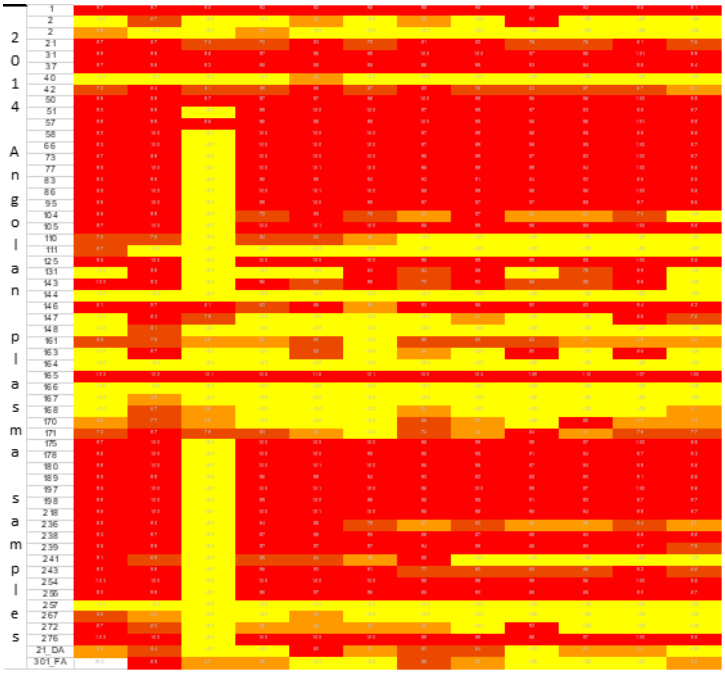


Figure S3- Heatmap showing the neutralizing activity of plasma samples from 2009 and 2014 against the 12-Env pseudotyped virus indicator panel. Percent neutralization was determined in TZM-bl cells with plasma samples diluted 1:40. White cells indicates non determined values; Yellow cells indicate <20% neutralization; orange highlighting indicates 20 to <50% neutralization; light brown highlighting indicates 50% to <80% neutralization; red highlighting indicates ≥80% neutralization. Virus subtype is indicated below the isolate common name of the Env-pseudotyped virus.

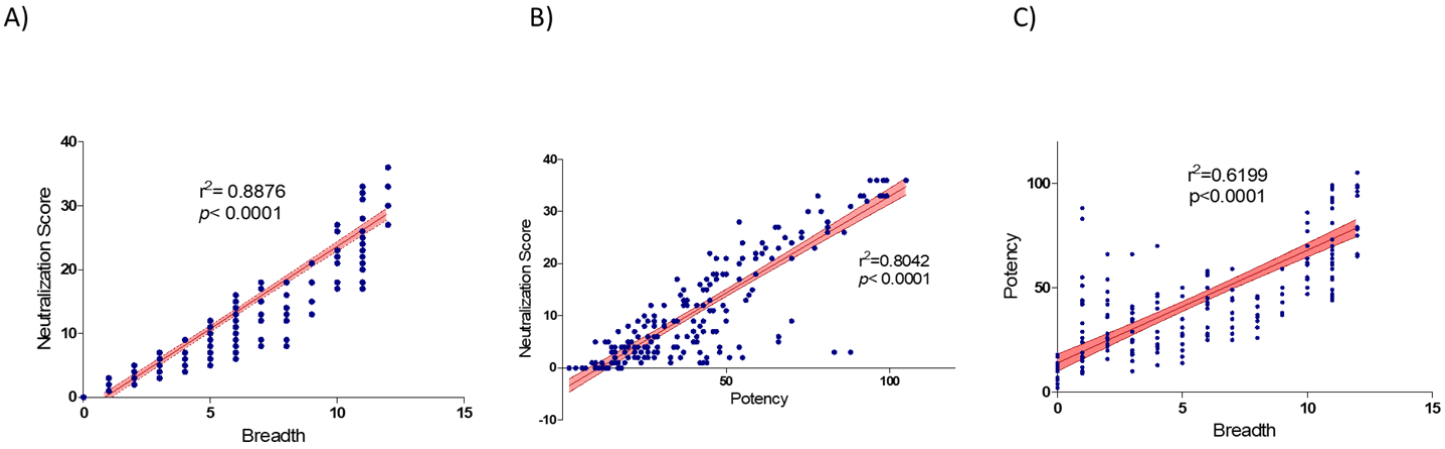
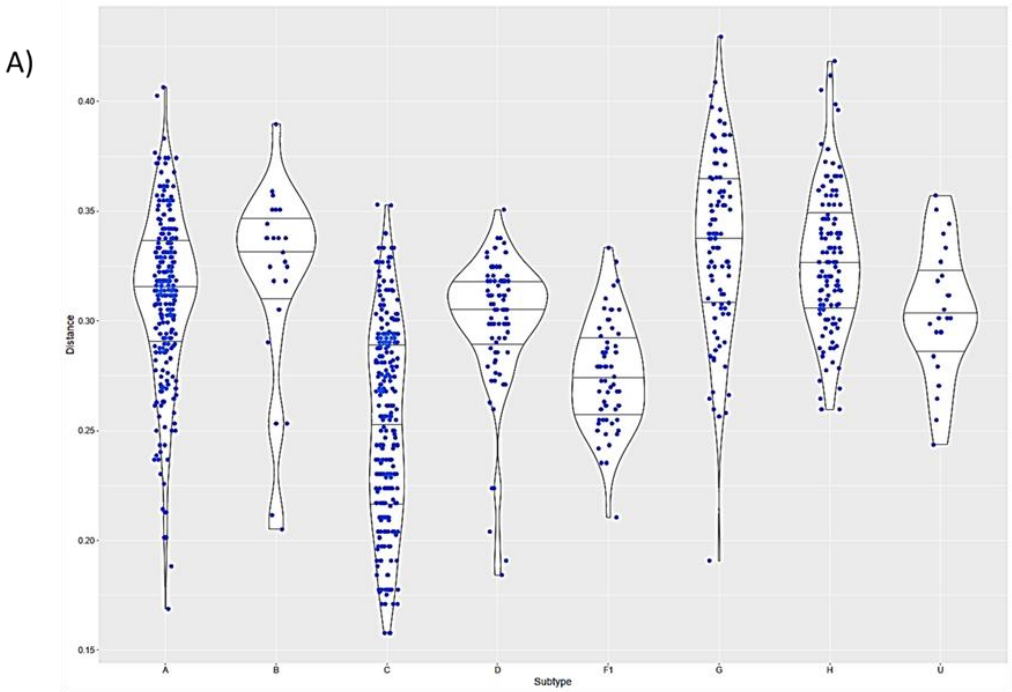


Figure S4- Neutralization breadth and potency predict neutralization score (NS). A) Correlation between neutralization breadth and potency in the 236 samples. B)

Correlation between NS and breadth. C) Correlation between NS and potency. Breadth was considered the number of pseudoviruses that were neutralized >20% and potency was the geometric mean of %neutralization against a given virus of the 12-virus indicator panel. Linear trend is shown with mean and 95% CI bands. The linear regression line is represented showing mean and 95% confidence interval error bands, goodness of fit r2 and P values are indicated.



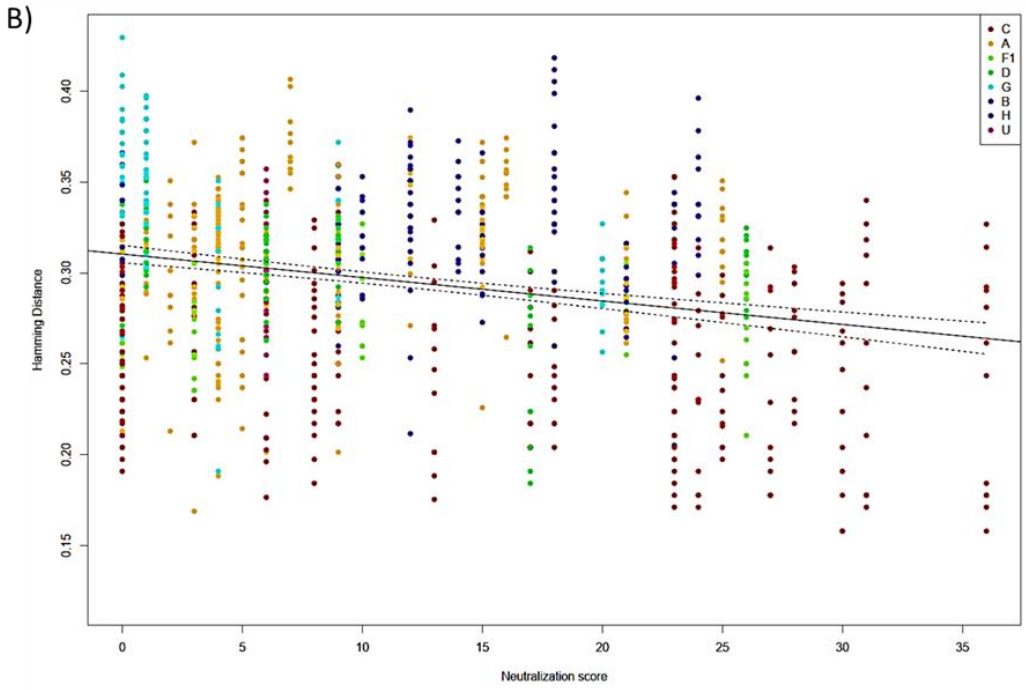


Figure S5- Impact of HIV-1 clade on antibody neutralization of the 12-virus panel. A) C2V3C3 amino acid distance of the viruses from the indicator panel to the viruses infecting the patients. B) Correlation between neutralization score and C2V3C3 amino acid distance of the viruses from the indicator panel to the viruses infecting the patients. Subtype of the patient's virus is indicated by color, linear trend is shown with mean and 95% CI bands.

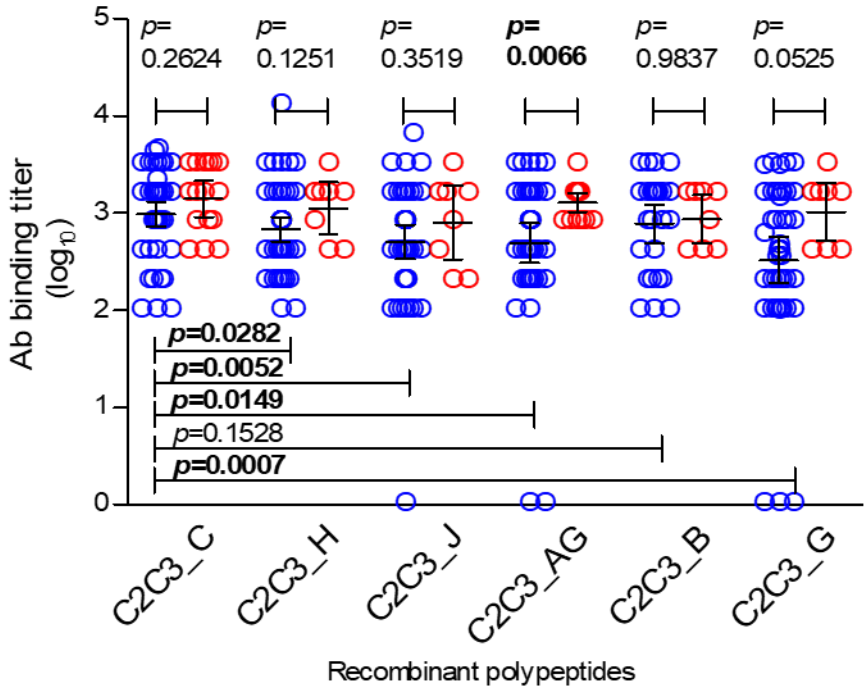


Figure S6- Antibody binding titers against the C2V3C3 recombinant polypeptides of different subtypes in patients from 2009 and 2014. Blue circles correspond to patients from 2009 and red circles to patients from 2014. Median and interquartile range are shown. P values were obtained using the Mann Whitney U test. P values <0.05 are shown in bold.

Chapter 5

General Discussion and Conclusions

In Chapter 2 of this thesis we characterized the HIV-1 diversity, transmission dynamics and prevalence of TDR in Luanda in 2009, five years after the scale-up of ART and compared these data with our previous survey performed in 2001 [5, 7, 98, 99]. Like in 2001, no major PIs resistance mutations were found in the study population which is consistent with the fact that first-line regimens used in Angola didn't include PIs in 2009 [285]. The K103N mutation, which confers high-level resistance to nevirapine, and efavirenz [270], was found in only one patient accounting for a 0.7% prevalence rate of TDR which was lower than the 2001 survey [7]. This residual TDR prevalence was similar to that of several African countries that also use the public health approach to ART [200, 201, 286-288]. However, more recent studies conducted in Luanda, by Sebastião *et al* showed an increase in HIV drug resistance, where more than 17% of drug naïve HIV-1 infected pregnant women presented DRMs mainly to the NNRTIs which is particular concerning due to the limited number of ART options available in Angola for the treatment of HIV-1 infection, that is still mainly based on NNRTIs with a backbone of two NRTIs [37, 349]. The unregulated and unmonitored use of ART obtained in the black market or abroad, the lack of alternative ART regimens as well as the displacement of HIV infected people to countries where ART is accessible for a long time are the most likely explanations for the increase of the HIV drug resistance seen in more recent years in Angola. Similar to previous studies and in line with more recent publications [34, 37, 349], the HIV-1 epidemic in Luanda in 2009 was highly complex being characterized by the presence of almost 50% of complex recombinant virus and all subtypes with the exception of subtypes B and K [4, 5, 35, 96, 289]. We saw that some strains from Angola had little organized substructure and formed weaker clusters within phylogenetic trees than the global reference sequences, not allowing a clear distinction between subtypes. As a consequence, the current global subtype classification may not reflect the extent of

diversity in this region [290] which might reflect the 5.8% of untypable sequences observed in our study. Moving forward, full genome sequences should be sequenced more frequently in order to get more detail about the the possible recombination events, as recently Bartolo *et al* identified a new H/U/CRF02_AG recombinant genome isolated from an HIV-1 infected Angolan patient [252]. The prevailing subtype in 2009 in Luanda was subtype C (36.9%) followed by subsubtype F1 (20.0%) whereas in 2001 it was subtype A followed by subtype C [7]. In 2018, in 34 HIV-1 infecteted pregnant women from Luanda, also analyzing the *pol* gene, Sebastião *et al.* identified the same prevailing subtypes as we did in 2009, 38% of subtype C the most frequent, followed by the subtype F1 18% [37]. The significant decrease in the prevalence of subtype A and increase in subtype C observed in 2009 and in more recent studies [34, 37] could be explained by the increasing predominance of subtype C in the bordering countries, namely in the south region of Democratic Republic of Congo [202] and in Zambia [35, 200, 207].

Importantly, the results indicate that the Angolan HIV-1 epidemic is still increasing in genetic complexity and suggest high rates of co-infection and/or superinfection [208] which is consistent with an increasing HIV-1 incidence and prevalence [29, 255, 291].

In chapter 2 we also attempted to perform the first assessement of HIV-1 transmission dynamics in Angola. We sucessfully identified a large number of transmission clusters in this study which included 35.7% of the analyzed samples. Overall, based on high sequence homology between patients in transmission clusters, a large number of potential recent infections were inferred which is consistent with a rising HIV-1 epidemic in Luanda driven by heterosexual transmission [29, 255, 291].

In chapter 3 we performed a much needed validation of a molecular test for the EID of HIV-1 infection with pratical and direct benefits for the children of the Angolan Perinatal HIV Cohort (APEHC). The high seroprevalence of HIV-1 infection in pregnant women

in Angola make children among the most at risk for HIV infection [367]. In spite of this the AIDS response in sub-Saharan Africa has largely left them behind [30]. In this regard, Angola has only registered a 25% reduction of new infections among infants since 2009 [30] which led to the implementation of several programs for prevention of mother to child transmission (PMTCT), due to the implementation of such programs 4,100 new HIV infections were averted in 2020 [3]. However, final vertical transmission rate including breastfeeding in 2020 in Angola was 18.6% [3]. Significant progress has been made in many sub-Saharan countries in implementing EID services following the introduction of HIV DNA testing on DBS [198]. Several recent studies shown the feasibility and advantages of using point of care (POC) platforms for the EID of HIV-1 infection in different African countries [226, 227, 368]. However, at the HDP where our cohort was based and in most other hospitals in Angola, pediatric diagnosis of HIV-1 infection was still done by serology at month 12 of life which significantly delays the initiation of treatment [305]. To support EID service expansion in Luanda we developed and validated a new HIV-1 DNA assay to be used on DBS samples.

To lower the costs, we used the Chelex method of DNA extraction which is also quick and easy to perform. This method had been previously applied to diagnostic screening in HIV-1-exposed infants in Rwanda showing reliable results when used in combination with either an in-house nested PCR or the Roche Amplicor HIV-1 DNA assay version 1.5 [306]. Also, the same extraction method was used before successfully to detect CMV co-infection in infants infected with HIV [369]. The Chelex method that we have used costs less than 3 USD per reaction, is very quick to perform, and does not use hazardous solvents. Our nested PCR assay showed a very low LoD for all the complex HIV-1 genotypes that we used as controls, suggesting that it was appropriate for early diagnosis of HIV-1 infection in infants in Angola. Indeed, using this assay we could detect all HIV-

1 infected infants at month 1 of life which makes this new assay suitable to health care centers following option B+ of WHO guidelines that recommend EID at 4-6 weeks of life. The low percentage of HIV-1 infected infants (2.2%) in the APEHC cohort between 2012 and 2014 confirms the effectiveness of the WHO-based prevention program implemented since 2007 at the HDP [305] and the high standard of care provided at HDP to the pregnant women infected with HIV. More recently, using a similar methodology, filter papers and nested PCR for the diagnosis of perinatal HIV-1[370] and other infectious disease [371] in rural and low income areas, has been sucessfuly used by others [370, 371]. Which shows the feasibility and adaptability of our method to be used not only to diagnose HIV infection but to detected other pathogens in diferent geographic areas [371]. In fact using an optimized version of the protocol that we developed and validated for the perinatal HIV-1 diagnosis, we as well as others [372] managed to detect and diagnose SARS-CoV-2 infection, which was especially important since the COVID-19 pandemic caused supply shortages of diagnostic tests. In comparison with current conventional methods for detecting SARS-CoV-2 (COVID-19) [373] the Chelex method presents important advantages in terms of safety, costs, and sensitivity [372].

In chaper 4 we performed the first detailed characterization of the antibody response and its viral and host determinants on a large cohort of HIV-1-infected Angolan patients. Overall, 29% (elite and broad neutralization categories) of the patients developed broad and potent neutralizing responses, roughly corresponding to more than 50% breadth on a diverse and representative 12 ENV-pseudotyped virus panel [162]. This result is in line with other studies suggesting that some degree of neutralization breadth develops in the majority of HIV infected individuals. However it is higher than that observed in previous cohort studies in sub-Saharan Africa [167, 172, 173, 176] and in Europe [146, 163]. This divergence is likely related with differences between HIV populations such as the HLA

genotype and ethnicity of the patients, diversity of infecting viruses, viral load, CD4 counts, and duration of infection [146, 166, 172, 173, 354]. Time of infection and viral diversity has been shown to influence the developing of neutralization breadth [146, 172, 173, 374]. In this sense, the age of the HIV-1 epidemic, the unusual set of HIV-1 clades and the overall high viral diversity found in Angola may have contributed to the high rate of broad neutralizers found in this study [2, 5, 6, 34, 37, 99, 252]. Indeed, in this thesis (Chapter 2 and 4) we observed an increasing viral variability that might impact neutralization capacity, both processes are mutually dependent, Env escape variants being selected in response to neutralizing responses and a greater diversity of escape variants selecting a greater variety of Abs creating a repetitive cycle resulting in the increased probability of eliciting bNAbs [166, 170].

Also, we saw a significant boost in the neutralizing response from 2009 to 2014. Although, we can argue that in 2009 some patients could still be in the process of mounting an effective neutralizing response as more than two thirds of the samples collected in 2014 were follow-ups of 2009 [146, 172, 173, 374]. Nevertheless, this result is no less remarkable since in 2014 the vast majority of the patients were under ART and it was shown before that restoration of the B cell function with ART is uncommon in chronic HIV-1 infection [356, 357]. However, we can hypothesize that similar to what happens in HIV-2, persistent stimulation of germinal centers in the lymph nodes in a seting of reduced viremia is linked to memory B-cell exhaustion resulting in increased HIV antibody affinity maturation [317]. In fact, moving forward we are going to further characterize and test the neutralizing activity of the Angolan HIV-1 infected patients against HIV-2 clinical isolates, preliminary results shows that cross-type neutralization is much more common than expected [375, 376], which raises the hypothesis that these

ancestral founder viruses that circulate in Angola have envelope features structurally more similar to HIV-2 that has important exposed epitopes, such as the V3 region [326].

Previous publications found that the development of neutralization breadth was strongly associated with viral load (VL) [173, 377]. Unfortunately VL determination was not standard of care at the time of our study in Angola, so CD4 count and clinic were the main markers used to measure disease progression. The finding that late presenters (<200 CD4+ T cell counts) had higher neutralization scores was in line with other studies and likely reflects higher viral replication in these patients. An intrinsic difference in CD4 levels between patients with or without neutralizing capacity cannot be totally excluded as CD4 levels prior to infection were not available. Even so, if there is indeed a direct impact from CD4+ T cell loss on bnAb development, this could potentially be due to the increase of survival time of auto-reactive B cells or by restraining the polyactivation of B cells in that way decreasing unspecific activation of precursor B cell populations [172].

In our study infection by subtype C viruses was associated with enhanced neutralization breadth and potency. Specific phenotypic and genotypic features of subtype C transmitted/founder viruses have been described that possibly support the development of neutralization breadth, such as less variable V3 region, shorter V1-V2 region and unique mutational patterns in the α2-helix in the C3 region [146, 168, 171, 176, 212, 348, 362, 364]. Highlighting the importance of the infecting virus in guiding the selection of bNAbs, a comprehensive analysis by Koyous *et al.*, where they prospectively followed selected transmission pairs identifying specific characteristics of the virus that make them more prone to develop bNAbs, emphasizing the importance that select HIV-1 viral variants have on the capability to initiate bnAb responses [353]. Furthermore, it has been shown that viral subtype impacts the elicitation [212] and epitope specificity of bNAbs [146, 171] which shows the importance of the subtype of the infecting virus in guiding

bnAb induction and maturation. Yet, in our study we must consider that subtype C viruses from the indicator panel more closely related to the infecting virus of the Angolan patients were more easily neutralized than subtypes with higher genetic differences such as CRF01_AE, evidencing a within clade neutralization tendency.

Despite not having assessed the neutralizing activity against the autologous virus, we looked at the variability of the Angolan HIV sequences in the C2V3C3 region to assess the impact of antibodies targeting this region on viral diversity. The V3 and C3 were the most variable regions across neutralizing categories which is in line with other studies that show that bNAbs targeting the V3 glycan are the most abundant [173] and first bNAbs [316] to be selected during the course of the natural infection, mainly because of the exposed nature of the V3 loop and the fact that contrary to other bNAbs, V3 targeting bNAbs do not require extensive somatic hypermutations [323]. Interestingly, patients exhibiting elite neutralizing capacity against the virus panel had far less variability in the C2V3C3 in comparison to patients with low neutralizing responses, indicating that the neutralizing response constrained the diversification having selected quasispecies more resistant to neutralization. Moreover, site-by-site analysis revealed that diversifying selection follow the same tendency as viral diversity and neutralizing responses, since the Angolan patients exhibiting low neutralizing responses were the ones showing more sites under selective pressure. Taken together these findings point to an increasing diversifying selection in Env evolving together with development of breadth to a certain extent and then a convergence of the resistant viral quasispecies with the acquisition of breadth.

Of note, important V3 bnAb recognition sites and sites associated with resistance to neutralization such as N295 were under positive selection only in patients with broad and elite neutralizing capacities, and N-glycans at positions 339 and 355 in C3 were only present simultaneously in elite neutralizers. These findings suggest that most Nab

epitopes in these patients are located in these V3 and C3 regions. In fact, despite N295 glycan not being a direct contact of bnAb PGT 135 it was shown by Kong *et. al.* [19] that N295 is an important recognition site for some HIV-1 strains. Also Seabright et al. shown that deletion of N295A resulted in a considerable decrease in 2G12 antibody binding [378]. Furthermore, as hypothesized by Gray et al. that clade C resistance to 2G12 was related to the lack of N295 glycan [379], in our study site N295 was under positive selective pressure in the elite neutralizers (60% of which harbouring clade C virus) despite lacking glycosylation at position 295 which may have prompted the high number of 2G12 and PGT 128 like bnAb responses seen in our study. Increasing the number of N-glycans in the envelope gp120 surface glycoprotein, or varying the position of glycosylation sites, has been associated with escape from IgG neutralizing antibody response in simian immunodeficiency virus (SIV) and HIV-1 infection [155, 380-382]. A recent study that followed HIV-1 subtype C infected individuals longitudinally from acute infection reported associations between the development of bNAbs and the presence of specific glycans in the C2V3C3 region on the gp120 [176], specifically on sites 301 and 332. In our study both these N-glycosylation sites were very conserved, being present in almost all sequences tested, irrespective of neutralization category and viral subtype, which also might have played a role in the high number of PGT128 and 2G12 bnAb like responses that we have seen. In contrast, N-glycosylation sites N339 and N355 were mainly present in the patients exhibiting elite neutralization capacity. These sites and particularly N339 were recently identified as an important epitope for CP506 lineage mAbs [366].

We found a strong association between antibody binding titers targeting the C2V3C3 region and neutralization capacity irrespective of viral subtype and patients' characteristics. These results confirm the important role of the C2V3C3 HIV-1 region as a target for antibody neutralization [176, 212, 316, 366]

Consistent with this we found that most of the neutralizing antibodies in Angolan patients had neutralization fingerprints similar to bNAbs targeting the V3 glycan, concordantly Landais *et al.* in a large cohort of Eastern and South African HIV infected patients found that the most prevalent target epitope was V3 glycan N332 supersite followed by the V2 apex [173]. The C2V3C3 region of HIV-1, as previously found, has an extended and highly accessible V3 loop [383]. Such conformation is entirely consistent with its immunodominant and neutralizing nature and with its crucial role in HIV-1 co-receptor binding and tropism. Concordantly, Calado et al. tested recently a vaccine HIV candidate in mice and showed that cross clade neutralization of tier 2 HIV-1 isolates was being driven essentially by antibodies targeting the V3 crown which provides further support for the V3 crown as a crucial vaccine target [212, 250]. However, we have to take into consideration that the envelope conformations of the global Env-pseudotyped virus panel might naturally favour the selection and identification of antibodies targeting V3, since it was shown by Han et al. [383] that Envs of distinct standardised virus panels are in an open conformation with the V3 loop exposed. Still, the high prevalence of Angolan patients with cross-clade neutralizing activity is consistent with long-term antigenic stimulation with highly divergent and diverse HIV-1 clades. Also, the antibody binding titer against envelope C2V3C3 region was a good indicator of neutralization breadth and potency and may also be a good indicator of vaccine efficacy.

Overall, the studies developed in these thesis show an increasing diversity of the HIV-1 epidemics in Angola which poses significant challenges to the diagnosis, treatment and management of people living with HIV. Lastly, the results obtained by characterizing the neutralizing antibody responses in this challenging setting confirms the importance of further characterize epidemics such as the Angolan with ancient HIV strains.

References

References

1. *UN General Assembly Political declaration on HIV and AIDS: on the fast track to accelerating the fight against HIV and to ending the AIDS epidemic by 2030. 2016.* https://www.unaids.org/sites/default/files/media_asset/2016-political-declaration-HIV-AIDS_en.pdf.
2. *Subnational mapping of HIV incidence and mortality among individuals aged 15-49 years in sub-Saharan Africa, 2000-18: a modelling study*. Lancet HIV, 2021. **8**(6): p. e363-e375.
3. AIDSinfo, *Global data on HIV epidemiology and response. Country factsheets: Angola 2020*

https://www.unaids.org/en/regionscountries/countries/angola.

4. Bartolo, I., et al., *High genetic diversity of human immunodeficiency virus type 1 in Angola*. AIDS Res Hum Retroviruses, 2005. **21**(4): p. 306-10.
5. Bartolo, I., et al., *Highly divergent subtypes and new recombinant forms prevail in the HIV/AIDS epidemic in Angola: new insights into the origins of the AIDS pandemic*. Infect Genet Evol, 2009. **9**(4): p. 672-82.
6. Pineda-Pena, A.C., et al., *On the contribution of Angola to the initial spread of HIV-1*. Infect Genet Evol, 2016. **46**: p. 219-222.
7. Bartolo, I., et al., *Antiretroviral drug resistance surveillance among treatment-naive human immunodeficiency virus type 1-infected individuals in Angola: evidence for low level of transmitted drug resistance*. Antimicrobial agents and chemotherapy, 2009. **53**(7): p. 3156-8.
8. UNAIDS, *2015 PROGRESS REPORT ON THE GLOBAL PLAN: towards the elimination of new HIV infections among children and keeping their mothers alive.*

https://www.unaids.org/en/resources/documents/2015/JC2774_2015ProgressReport_GlobalPlan.

9. Pugach, P., et al., *A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene*. J Virol, 2015. **89**(6): p. 3380-95.
10. Sanders, R.W., et al., *HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers.* Science, 2015. **349**(6244): p. aac4223.
11. Burton, D.R. and J.R. Mascola, *Antibody responses to envelope glycoproteins in HIV-1 infection*. Nat Immunol, 2015. **16**(6): p. 571-6.
12. Lee, J.H., et al., *Antibodies to a conformational epitope on gp41 neutralize HIV-1 by destabilizing the Env spike.* Nat Commun, 2015. **6**: p. 8167.
13. Stephenson, K.E., et al., *Vaccines and Broadly Neutralizing Antibodies for HIV-1 Prevention*. Annu Rev Immunol, 2020. **38**: p. 673-703.
14. Gray, G.E., et al., *Vaccine Efficacy of ALVAC-HIV and Bivalent Subtype C gp120-MF59 in Adults.* N Engl J Med, 2021. **384**(12): p. 1089-1100.
15. Amornkul, P.N., et al., *Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa*. AIDS, 2013. **27**(17): p. 2775-86.
16. van der Sluis, R.M., et al., *Dendritic cell-induced activation of latent HIV-1 provirus in actively proliferating primary T lymphocytes.* PLoS Pathog, 2013. **9**(3): p. e1003259.
17. Sanders, R.W., et al., *A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies.* PLoS Pathog, 2013. **9**(9): p. e1003618.
18. Kijak, G.H., et al., *Molecular evolution of the HIV-1 Thai epidemic between the time of RV144 immunogen selection to the execution of the vaccine efficacy trial.* J Virol, 2013. **87**(13): p. 7265-81.

19. Kong, L., et al., *Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120.* Nat Struct Mol Biol, 2013. **20**(7): p. 796-803.

20. Julien, J.P., et al., *Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens.* Proc Natl Acad Sci U S A, 2015. **112**(38): p. 11947-52.

21. Hymes, K.B., et al., *Kaposi's sarcoma in homosexual men-a report of eight cases.* Lancet, 1981. **2**(8247): p. 598-600.

22. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.

23. Chermann, J.C., et al., *Isolation of a new retrovirus in a patient at risk for acquired immunodeficiency syndrome.* Antibiot Chemother (1971), 1983. **32**: p. 48-53.

24. Clavel, F., et al., *Isolation of a new human retrovirus from West African patients with AIDS*. Science, 1986. **233**(4761): p. 343-6.

25. Clavel, F., et al., *Human immunodeficiency virus type 2 infection associated with AIDS in West Africa*. N Engl J Med, 1987. **316**(19): p. 1180-5.

26. Rhodes, S.D. and F.S. Sy, *Effectively Confronting the COVID-19 Pandemic: Critical Lessons From HIV Prevention, Care, and Treatment and Innovative Strategies to Conduct Community-Based and Community-Engaged Research Safely*. AIDS Educ Prev, 2020. **32**(6): p. 455-471.

27. Hans, L., et al., *Preparing for the next pandemic: Lessons from rapid scale-up of SARS-CoV-2 testing in a South African high-throughput automated HIV molecular laboratory*. Int J Infect Dis, 2021. **110**: p. 1-3.

28. *UNAIDS. Global HIV & AIDS statistics - 2020 fact sheet*

https://www.unaids.org/en/resources/fact-sheet. 2020.

29. UNAIDS, *UNAIDS report on the global AIDS epidemic 2013. Geneve. Avalaible http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf. Accessed 2014 Jan 30.* 2013.

30. UNAIDS, *AIDS by the numbers 2016* 2016, Joint United Nations Programme on HIV/AIDS (UNAIDS).

31. Reeves, J.D. and A.J. Piefer, *Emerging drug targets for antiretroviral therapy.* Drugs, 2005. **65**(13): p. 1747-66.

32. Gilks, C.F., et al., *The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings.* Lancet, 2006. **368**(9534): p. 505-10.

33. *WHO (2010) Antiretroviral therapy for HIV infection in adults and adolescents. Recommendations for a public health approach. 2010 revision Avalaible: http://whqlibdoc.who.int/publications/2010/9789241599764_eng.pdf. Accessed 15 January 2014.*

34. Sebastiao, C.S., J. Morais, and M. Brito, *Clinical and Public Health Implications of HIV- Genetic Diversity and Drug Resistance Mutations in Angola: A Systematic Review.* AIDS Rev, 2020. **23**(1): p. 48-56.

35. Afonso, J.M., et al., *HIV-1 genetic diversity and transmitted drug resistance mutations among patients from the North, Central and South regions of Angola.* PLoS One, 2012. **7**(8): p. e42996.

36. Afonso, J.M., M.G. Morgado, and G. Bello, *Evidence of multiple introductions of HIV-1 subtype C in Angola.* Infect Genet Evol, 2012. **12**(7): p. 1458-65.

37. Sebastiao, C.S., et al., *Genetic diversity and drug resistance of HIV-1 among infected pregnant women newly diagnosed in Luanda, Angola.* PLoS One, 2019. **14**(11): p. e0225251.

38. Coffin, J.M., *The virology of AIDS: 1990.* AIDS, 1990. **4 Suppl 1**: p. S1-8.

39. Bedwell, G.J. and A.N. Engelman, *Factors that mold the nuclear landscape of HIV-1 integration*. Nucleic Acids Res, 2021. **49**(2): p. 621-635.

40. Roy, S., et al., *A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation*. Genes Dev, 1990. **4**(8): p. 1365-73.

41. Sherpa, C., et al., *Evolution of the HIV-1 Rev Response Element during Natural Infection Reveals Nucleotide Changes That Correlate with Altered Structure and Increased Activity over Time*. J Virol, 2019. **93**(11).

42. Cullen, B.R. and M.H. Malim, *The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators.* Trends Biochem Sci, 1991. **16**(9): p. 346-50.

43. Dettenhofer, M., et al., *Association of human immunodeficiency virus type 1 Vif with RNA and its role in reverse transcription.* J Virol, 2000. **74**(19): p. 8938-45.

44. Sheehy, A.M., N.C. Gaddis, and M.H. Malim, *The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif.* Nat Med, 2003. **9**(11): p. 1404-7.

45. Stavrou, S. and S.R. Ross, *APOBEC3 Proteins in Viral Immunity*. J Immunol, 2015. **195**(10): p. 4565-70.

46. Le Rouzic, E. and S. Benichou, *The Vpr protein from HIV-1: distinct roles along the viral life cycle*. Retrovirology, 2005. **2**: p. 11.

47. Laguette, N., et al., *SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx.* Nature, 2011. **474**(7353): p. 654-7.

48. Willey, R.L., et al., *Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4*. J Virol, 1992. **66**(12): p. 7193-200.

49. Klimkait, T., et al., *The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release.* J Virol, 1990. **64**(2): p. 621-9.

50. Roeth, J.F. and K.L. Collins, *Human immunodeficiency virus type 1 Nef: adapting to intracellular trafficking pathways.* Microbiol Mol Biol Rev, 2006. **70**(2): p. 548-63.

51. Gomez, C. and T.J. Hope, *The ins and outs of HIV replication*. Cell Microbiol, 2005. **7**(5): p. 621-6.

52. Hladik, F. and M.J. McElrath, *Setting the stage: host invasion by HIV.* Nat Rev Immunol, 2008. **8**(6): p. 447-57.

53. Agosto, L.M., et al., *Highly active antiretroviral therapies are effective against HIV-1 cell-to-cell transmission*. PLoS Pathog, 2014. **10**(2): p. e1003982.

54. Arnold, E. and S.G. Sarafianos, *Molecular biology: an HIV secret uncovered.* Nature, 2008. **453**(7192): p. 169-70.

55. Zila, V., et al., *Cone-shaped HIV-1 capsids are transported through intact nuclear pores.* Cell, 2021. **184**(4): p. 1032-1046 e18.

56. Scoca, V. and F. Di Nunzio, *The HIV-1 Capsid: From Structural Component to Key Factor for Host Nuclear Invasion*. Viruses, 2021. **13**(2).

57. AlBurtamani, N., A. Paul, and A. Fassati, *The Role of Capsid in the Early Steps of HIV-1 Infection: New Insights into the Core of the Matter*. Viruses, 2021. **13**(6).

58. Engelman, A. and P. Cherepanov, *The structural biology of HIV-1: mechanistic and therapeutic insights.* Nat Rev Microbiol, 2012. **10**(4): p. 279-90.

59. Wu, Y., *HIV-1 gene expression: lessons from provirus and non-integrated DNA.* Retrovirology, 2004. **1**: p. 13.

60. Murphy, R.E. and J.S. Saad, *The Interplay between HIV-1 Gag Binding to the Plasma Membrane and Env Incorporation.* Viruses, 2020. **12**(5).

61. Moranguinho, I. and S.T. Valente, *Block-And-Lock: New Horizons for a Cure for HIV-1*. Viruses, 2020. **12**(12).

62. Taveira, N. and F. Martin, *O agente. Manual sobre SIDA – 5ª Edição. Editor Francisco Antunes, Fernando Maltez. 2018.* Permanyer Portugal., 2018.

63. Jackson, P.E.H., et al., *Sequence and Functional Variation in the HIV-1 Rev Regulatory Axis.* Curr HIV Res, 2020. **18**(2): p. 85-98.

64. Klingler, J., et al., *How HIV-1 Gag Manipulates Its Host Cell Proteins: A Focus on Interactors of the Nucleocapsid Domain.* Viruses, 2020. **12**(8).

65. Bird, S.W., et al., *The ins and outs of viral infection: keystone meeting review.* Viruses, 2014. **6**(9): p. 3652-62.

66. Leonard, C.K., et al., *Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.* J Biol Chem, 1990. **265**(18): p. 10373-82.

67. Bernstein, H.B., et al., *Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides.* J Virol, 1994. **68**(1): p. 463-8.

68. Helseth, E., et al., *Human immunodeficiency virus type 1 gp120 envelope glycoprotein regions important for association with the gp41 transmembrane glycoprotein.* J Virol, 1991. **65**(4): p. 2119-23.

69. Wyatt, R., et al., *The antigenic structure of the HIV gp120 envelope glycoprotein.* Nature, 1998. **393**(6686): p. 705-11.

70. Wyatt, R., et al., *Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding*. J Virol, 1995. **69**(9): p. 5723-33.

71. Kwong, P.D., et al., *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody*. Nature, 1998. **393**(6686): p. 648-59.

72. Pancera, M., et al., *Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility*. Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1166-71.

73. Huang, C.C., et al., *Structure of a V3-containing HIV-1 gp120 core*. Science, 2005. **310**(5750): p. 1025-8.

74. Chen, B., et al., *Structure of an unliganded simian immunodeficiency virus gp120 core*. Nature, 2005. **433**(7028): p. 834-41.

75. Dalgleish, A.G., et al., *The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus.* Nature, 1984. **312**(5996): p. 763-7.

76. Maddon, P.J., et al., *The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain*. Cell, 1986. **47**(3): p. 333-48.

77. Deng, H., et al., *Identification of a major co-receptor for primary isolates of HIV-1*. Nature, 1996. **381**(6584): p. 661-6.

78. Dragic, T., et al., *HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5*. Nature, 1996. **381**(6584): p. 667-73.

79. Alkhatib, G., et al., *CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1*. Science, 1996. **272**(5270): p. 1955-8.

80. Moriuchi, M., et al., *Cloning and analysis of the promoter region of CXCR4, a coreceptor for HIV-1 entry*. J Immunol, 1997. **159**(9): p. 4322-9.

81. Trkola, A., et al., *CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5*. Nature, 1996. **384**(6605): p. 184-7.

82. Allen, A.G., et al., *Gene Editing of HIV-1 Co-receptors to Prevent and/or Cure Virus Infection*. Front Microbiol, 2018. **9**: p. 2940.

83. Gombos, R.B., et al., *Inhibitory Effect of Individual or Combinations of Broadly Neutralizing Antibodies and Antiviral Reagents against Cell-Free and Cell-to-Cell HIV-1 Transmission*. J Virol, 2015. **89**(15): p. 7813-28.

84. Dufloo, J., T. Bruel, and O. Schwartz, *HIV-1 cell-to-cell transmission and broadly neutralizing antibodies*. Retrovirology, 2018. **15**(1): p. 51.

85. Hwang, S.S., et al., *Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1*. Science, 1991. **253**(5015): p. 71-4.

86. Lengauer, T., et al., *Bioinformatics prediction of HIV coreceptor usage*. Nat Biotechnol, 2007. **25**(12): p. 1407-10.

87. De Cock, K.M., F. Brun-Vezinet, and B. Soro, *HIV-1 and HIV-2 infections and AIDS in West Africa*. AIDS, 1991. **5 Suppl 1**: p. S21-8.

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

89. Chen, Z., et al., *Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range*

of simian immunodeficiency virus-infected sooty mangabeys. J Virol, 1997. **71**(5): p. 3953-60.

90. Yamaguchi, J., et al., *Brief Report: Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L.* J Acquir Immune Defic Syndr, 2020. **83**(3): p. 319-322.

91. Hemelaar, J., et al., *Country Level Diversity of the HIV-1 Pandemic between 1990 and 2015.* J Virol, 2020. **95**(2).

92. Hemelaar, J., et al., *Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis.* Lancet Infect Dis, 2019. **19**(2): p. 143-155.

93. Hemelaar, J., et al., *Global and regional epidemiology of HIV-1 recombinants in 1990-2015: a systematic review and global survey.* Lancet HIV, 2020. **7**(11): p. e772-e781.

94. Robertson, D.L., et al., *HIV-1 nomenclature proposal.* Science, 2000. **288**(5463): p. 55-6.

95. Hemelaar, J., *The origin and diversity of the HIV-1 pandemic.* Trends Mol Med, 2012. **18**(3): p. 182-92.

96. Abecasis, A., et al., *HIV-1 genetic variants circulation in the North of Angola.* Infect Genet Evol, 2005. **5**(3): p. 231-7.

97. Abecasis, A.B., et al., *HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics.* Retrovirology, 2013. **10**: p. 7.

98. Bartolo, I., et al., *HIV-1 genetic diversity and transmitted drug resistance in health care settings in Maputo, Mozambique.* J Acquir Immune Defic Syndr, 2009. **51**(3): p. 323-31.

99. Bartolo, I., et al., *HIV-1 diversity, transmission dynamics and primary drug resistance in Angola*. PLoS One, 2014. **9**(12): p. e113626.

100. Bartolo I, T.N., *HIV-1 Diversity and Its Implications in Diagnosis, Transmission, Disease Progression, and Antiretroviral Therapy*. In: Caliskan M, editor. Genetic Diversity in Microorganisms

Rijeka, Croatia: InTech - Open Access Publisher. pp. 171–214., 2012.

101. Gao, F., et al., *The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin.* J Virol, 1996. **70**(10): p. 7013-29.

102. Soto-Ramirez, L.E., et al., *HIV-1 Langerhans' cell tropism associated with heterosexual transmission of HIV.* Science, 1996. **271**(5253): p. 1291-3.

103. Hudgens, M.G., et al., *Subtype-specific transmission probabilities for human immunodeficiency virus type 1 among injecting drug users in Bangkok, Thailand.* Am J Epidemiol, 2002. **155**(2): p. 159-68.

104. Kiwanuka, N., et al., *HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda*. AIDS, 2009. **23**(18): p. 2479-84.

105. Berger, E.A., et al., *A new classification for HIV-1*. Nature, 1998. **391**(6664): p. 240.

106. Huang, W., et al., *Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations.* J Virol, 2007. **81**(15): p. 7885-93.

107. Huang, Y., et al., *The role of a mutant CCR5 allele in HIV-1 transmission and disease progression*. Nature medicine, 1996. **2**(11): p. 1240-3.

108. Cilliers, T., et al., *The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C.* J Virol, 2003. **77**(7): p. 4449-56.

109. Renjifo, B., et al., *Differences in perinatal transmission among human immunodeficiency virus type 1 genotypes.* J Hum Virol, 2001. **4**(1): p. 16-25.

110. Renjifo, B., et al., *In-utero transmission of quasispecies among human immunodeficiency virus type 1 genotypes.* Virology, 2003. **307**(2): p. 278-82.

111. Duran Ramirez, J.J., et al., *Increasing Frequency and Transmission of HIV-1 Non-B Subtypes among Men Who Have Sex with Men in the Swiss HIV Cohort Study.* J Infect Dis, 2021.

112. Kuritzkes, D.R., *HIV-1 subtype as a determinant of disease progression.* The Journal of infectious diseases, 2008. **197**(5): p. 638-9.

113. Kaleebu, P., et al., *Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1-positive persons in Uganda.* The Journal of infectious diseases, 2002. **185**(9): p. 1244-50.

114. Kiwanuka, N., et al., *Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection.* The Journal of infectious diseases, 2008. **197**(5): p. 707-13.

115. Mlisana, K., et al., *Challenges of diagnosing acute HIV-1 subtype C infection in African women: performance of a clinical algorithm and the need for point-of-care nucleic-acid based testing.* PLoS One, 2013. **8**(4): p. e62928.

116. Touloumi, G., et al., *Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts.* Clin Infect Dis, 2013. **56**(6): p. 888-97.

117. Kantor, R., et al., *Pretreatment HIV Drug Resistance and HIV-1 Subtype C Are Independently Associated With Virologic Failure: Results From the Multinational PEARLS (ACTG A5175) Clinical Trial.* Clin Infect Dis, 2015. **60**(10): p. 1541-9.

118. Keller, M., et al., *Impact of HIV-1 viral subtype on CD4+ T-cell decline and clinical outcomes in antiretroviral naive patients receiving universal healthcare.* AIDS, 2009. **23**(6): p. 731-7.

119. Kanki, P.J., et al., *Human immunodeficiency virus type 1 subtypes differ in disease progression.* J Infect Dis, 1999. **179**(1): p. 68-73.

120. Baeten, J.M., et al., *HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads.* J Infect Dis, 2007. **195**(8): p. 1177-80.

121. Farinre, O., et al., *Subtype-specific differences in Gag-protease replication capacity of HIV-1 isolates from East and West Africa.* Retrovirology, 2021. **18**(1): p. 11.

122. Cashin, K., et al., *Linkages between HIV-1 specificity for CCR5 or CXCR4 and in vitro usage of alternative coreceptors during progressive HIV-1 subtype C infection.* Retrovirology, 2013. **10**: p. 98.

123. Attia, S., et al., *Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis.* AIDS, 2009. **23**(11): p. 1397-404.

124. Stone, M., et al., *Comparison of Detection Limits of Fourth- and Fifth-Generation Combination HIV Antigen-Antibody, p24 Antigen, and Viral Load Assays on Diverse HIV Isolates.* J Clin Microbiol, 2018. **56**(8).

125. Fiebig, E.W., et al., *Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection.* AIDS, 2003. **17**(13): p. 1871-9.

126. Aghokeng, A.F., et al., *Inaccurate diagnosis of HIV-1 group M and O is a key challenge for ongoing universal access to antiretroviral treatment and HIV prevention in Cameroon.* PLoS One, 2009. **4**(11): p. e7702.

127. Hage-Sleiman, M., et al., *False-negative Results of Human Immunodeficiency Virus (HIV) Rapid Testing in HIV Controllers.* Clin Infect Dis, 2020. **70**(8): p. 1754-1757.

128. Fogel, J.M., et al., *Brief Report: Impact of Early Antiretroviral Therapy on the Performance of HIV Rapid Tests and HIV Incidence Assays.* J Acquir Immune Defic Syndr, 2017. **75**(4): p. 426-430.

129. Delaugerre, C., et al., *Clinical and resistance consequences of misquantification of plasma and cerebrospinal fluid human immunodeficiency virus type 1 (HIV-1) RNA in samples from an HIV-1 subtype G-infected patient.* J Clin Microbiol, 2009. **47**(11): p. 3763-4.

130. Korn, K., et al., *Single-point mutations causing more than 100-fold underestimation of human immunodeficiency virus type 1 (HIV-1) load with the Cobas TaqMan HIV-1 real-time PCR assay.* J Clin Microbiol, 2009. **47**(4): p. 1238-40.

131. Rouet, F., et al., *Comparison of the Generic HIV Viral Load assay with the Amplicor HIV-1 monitor v1.5 and Nuclisens HIV-1 EasyQ v1.2 techniques for plasma HIV-1 RNA quantitation of non-B subtypes: the Kesho Bora preparatory study.* J Virol Methods, 2010. **163**(2): p. 253-7.

132. WHO, *Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations. 2016 update*. . 2016, World Health Organization: Geneva.

133. Penazzato, M., et al., *Early infant diagnosis of HIV infection in low-income and middle-income countries: does one size fit all?* Lancet Infect Dis, 2014. **14**(7): p. 650-5.

134. WHO, *March 2014 Supplement to the 2013 Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection. Recommendations for a Public Health Approach*. . 2014, World Health Organization Geneva.

135. UNITAID, *HIV/AIDS diagnostics technology landscape - semi-annual update.* 2015, World Health Organization: Geneva.

136. Kantor, R., et al., *Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration.* PLoS Med, 2005. **2**(4): p. e112.

137. Kantor, R., *Impact of HIV-1 pol diversity on drug resistance and its clinical implications.* Curr Opin Infect Dis, 2006. **19**(6): p. 594-606.

138. Saag, M.S., et al., *Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults: 2020 Recommendations of the International Antiviral Society-USA Panel.* JAMA, 2020. **324**(16): p. 1651-1669.

139. Martinez-Cajas, J.L., et al., *Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008).* J Int AIDS Soc, 2009. **12**: p. 11.

140. Eshleman, S.H., et al., *Comparison of mother-to-child transmission rates in Ugandan women with subtype A versus D HIV-1 who received single-dose*

nevirapine prophylaxis: HIV Network For Prevention Trials 012. J Acquir Immune Defic Syndr, 2005. **39**(5): p. 593-7.

141. Tomaras, G.D., et al., *Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia*. J Virol, 2008. **82**(24): p. 12449-63.
142. Sarmay, G., et al., *Mapping and comparison of the interaction sites on the Fc region of IgG responsible for triggering antibody dependent cellular cytotoxicity (ADCC) through different types of human Fc gamma receptor*. Mol Immunol, 1992. **29**(5): p. 633-9.
143. Burton, D.R., *Advancing an HIV vaccine; advancing vaccinology.* Nat Rev Immunol, 2019. **19**(2): p. 77-78.
144. Dieffenbach, C.W. and A.S. Fauci, *The search for an HIV vaccine, the journey continues.* J Int AIDS Soc, 2020. **23**(5): p. e25506.
145. Horwitz, J.A., et al., *Non-neutralizing Antibodies Alter the Course of HIV-1 Infection In Vivo*. Cell, 2017. **170**(4): p. 637-648 e10.
146. Rusert, P., et al., *Determinants of HIV-1 broadly neutralizing antibody induction.* Nat Med, 2016. **22**(11): p. 1260-1267.
147. Russell, J.H. and T.J. Ley, *Lymphocyte-mediated cytotoxicity*. Annu Rev Immunol, 2002. **20**: p. 323-70.
148. Baum, L.L., et al., *HIV-1 gp120-specific antibody-dependent cell-mediated cytotoxicity correlates with rate of disease progression.* J Immunol, 1996. **157**(5): p. 2168-73.
149. Overbaugh, J. and L. Morris, *The Antibody Response against HIV-1*. Cold Spring Harb Perspect Med, 2012. **2**(1): p. a007039.

150. Huang, J., et al., *Broad and potent neutralization of HIV-1 by a gp41-specific human antibody*. Nature, 2012. **491**(7424): p. 406-12.

151. Haynes, B.F., et al., *Immune-correlates analysis of an HIV-1 vaccine efficacy trial.* N Engl J Med, 2012. **366**(14): p. 1275-86.

152. Sogaard, O.S., et al., *The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo*. PLoS Pathog, 2015. **11**(9): p. e1005142.

153. Baldwin, K.M., et al., *HLA class II diversity in HIV-1 uninfected individuals from the placebo arm of the RV144 Thai vaccine efficacy trial.* Tissue Antigens, 2015. **85**(2): p. 117-26.

154. Kim, J.H., J.L. Excler, and N.L. Michael, *Lessons from the RV144 Thai phase III HIV-1 vaccine trial and the search for correlates of protection*. Annu Rev Med, 2015. **66**: p. 423-37.

155. Wei, X., et al., *Antibody neutralization and escape by HIV-1*. Nature, 2003. **422**(6929): p. 307-12.

156. Gray, E.S., et al., *Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection*. J Virol, 2007. **81**(12): p. 6187-96.

157. Moore, J.P., et al., *Inter- and intraclade neutralization of human immunodeficiency virus type 1: genetic clades do not correspond to neutralization serotypes but partially correspond to gp120 antigenic serotypes*. J Virol, 1996. **70**(1): p. 427-44.

158. Arendrup, M., et al., *Autologous HIV-1 neutralizing antibodies: emergence of neutralization-resistant escape virus and subsequent development of escape virus neutralizing antibodies*. J Acquir Immune Defic Syndr, 1992. **5**(3): p. 303-7.

159. Walker, L.M., et al., *Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target.* Science, 2009. **326**(5950): p. 285-9.

160. Scheid, J.F., et al., *Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals.* Nature, 2009. **458**(7238): p. 636-40.

161. Simek, M.D., et al., *Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm.* J Virol, 2009. **83**(14): p. 7337-48.

162. deCamp, A., et al., *Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies.* J Virol, 2014. **88**(5): p. 2489-507.

163. Hraber, P., et al., *Impact of clade, geography, and age of the epidemic on HIV-1 neutralization by antibodies.* J Virol, 2014. **88**(21): p. 12623-43.

164. Deeks, S.G., et al., *Neutralizing antibody responses against autologous and heterologous viruses in acute versus chronic human immunodeficiency virus (HIV) infection: evidence for a constraint on the ability of HIV to completely evade neutralizing antibody responses.* J Virol, 2006. **80**(12): p. 6155-64.

165. Moore, P.L., et al., *Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection.* PLoS Pathog, 2009. **5**(9): p. e1000598.

166. Moore, P.L., *The Neutralizing Antibody Response to the HIV-1 Env Protein*. Curr HIV Res, 2018. **16**(1): p. 21-28.

167. Mishra, N., et al., *Broadly neutralizing plasma antibodies effective against autologous circulating viruses in infants with multivariant HIV-1 infection*. Nat Commun, 2020. **11**(1): p. 4409.

168. Moore, P.L., et al., *The c3-v4 region is a major target of autologous neutralizing antibodies in human immunodeficiency virus type 1 subtype C infection.* J Virol, 2008. **82**(4): p. 1860-9.

169. Sato, S., et al., *Potent antibody-mediated neutralization and evolution of antigenic escape variants of simian immunodeficiency virus strain SIVmac239 in vivo.* J Virol, 2008. **82**(19): p. 9739-52.

170. Landais, E. and P.L. Moore, *Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers.* Retrovirology, 2018. **15**(1): p. 61.

171. Stefic, K., et al., *Impact of HIV-1 Diversity on Its Sensitivity to Neutralization.* Vaccines (Basel), 2019. **7**(3).

172. Gray, E.S., et al., *The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection.* J Virol, 2011. **85**(10): p. 4828-40.

173. Landais, E., et al., *Broadly Neutralizing Antibody Responses in a Large Longitudinal Sub-Saharan HIV Primary Infection Cohort.* PLoS Pathog, 2016. **12**(1): p. e1005369.

174. Mascola, J.R. and D.C. Montefiori, *The role of antibodies in HIV vaccines.* Annu Rev Immunol, 2010. **28**: p. 413-44.

175. Stamatatos, L., et al., *Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine?* Nat Med, 2009. **15**(8): p. 866-70.

176. Ndlovu, B., et al., *Envelope characteristics in individuals who developed neutralizing antibodies targeting different epitopes in HIV-1 subtype C infection.* Virology, 2020. **546**: p. 1-12.

177. Caskey, M., et al., *Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117.* Nature, 2015. **522**(7557): p. 487-91.

178. Caskey, M., et al., *Corrigendum: Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117*. Nature, 2016.

179. Schoofs, T., et al., *HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1*. Science, 2016. **352**(6288): p. 997-1001.

180. Scheid, J.F., et al., *HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption.* Nature, 2016. **535**(7613): p. 556-60.

181. Caskey, M., et al., *Antibody 10-1074 suppresses viremia in HIV-1-infected individuals*. Nat Med, 2017. **23**(2): p. 185-191.

182. Nishimura, Y. and M.A. Martin, *Of Mice, Macaques, and Men: Broadly Neutralizing Antibody Immunotherapy for HIV-1*. Cell Host Microbe, 2017. **22**(2): p. 207-216.

183. Ng, O.T., et al., *HIV type 1 polymerase gene polymorphisms are associated with phenotypic differences in replication capacity and disease progression*. The Journal of infectious diseases, 2014. **209**(1): p. 66-73.

184. Huang, J., et al., *Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface.* Nature, 2014. **515**(7525): p. 138-42.

185. Falkowska, E., et al., *Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers.* Immunity, 2014. **40**(5): p. 657-68.

186. Bosque, A., et al., *Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells.* PLoS Pathog, 2011. **7**(10): p. e1002288.

187. Mehta, S.R., et al., *Using phylogeography to characterize the origins of the HIV-1 subtype F epidemic in Romania.* Infect Genet Evol, 2011. **11**(5): p. 975-9.

188. Li, Y., et al., *Mechanism of neutralization by the broadly neutralizing HIV-1 monoclonal antibody VRC01*. J Virol, 2011. **85**(17): p. 8954-67.

189. Tomaras, G.D., et al., *Polyclonal B cell responses to conserved neutralization epitopes in a subset of HIV-1-infected individuals.* J Virol, 2011. **85**(21): p. 11502-19.

190. Raymond, S., et al., *Genotypic prediction of HIV-1 subtype D tropism.* Retrovirology, 2011. **8**: p. 56.

191. Vandekerckhove, L.P., et al., *European guidelines on the clinical management of HIV-1 tropism testing.* Lancet Infect Dis, 2011. **11**(5): p. 394-407.

192. Lewis, D.J., et al., *Phase I randomised clinical trial of an HIV-1(CN54), clade C, trimeric envelope vaccine candidate delivered vaginally.* PLoS One, 2011. **6**(9): p. e25165.

193. Garcia, F., et al., *Safety and immunogenicity of a modified pox vector-based HIV/AIDS vaccine candidate expressing Env, Gag, Pol and Nef proteins of HIV-1 subtype B (MVA-B) in healthy HIV-1-uninfected volunteers: A phase I clinical trial (RISVAC02).* Vaccine, 2011. **29**(46): p. 8309-16.

194. Churchyard, G.J., et al., *A phase IIA randomized clinical trial of a multiclade HIV-1 DNA prime followed by a multiclade rAd5 HIV-1 vaccine boost in healthy adults (HVTN204).* PLoS One, 2011. **6**(8): p. e21225.

195. Leigh Brown, A.J., et al., *Transmission network parameters estimated from HIV sequences for a nationwide epidemic.* The Journal of infectious diseases, 2011. **204**(9): p. 1463-9.

196. Spearman, P., et al., *A trimeric, V2-deleted HIV-1 envelope glycoprotein vaccine elicits potent neutralizing antibodies but limited breadth of neutralization in human volunteers.* J Infect Dis, 2011. **203**(8): p. 1165-73.

197. Marston, M., et al., *Net survival of perinatally and postnatally HIV-infected children: a pooled analysis of individual data from sub-Saharan Africa*. Int J Epidemiol, 2011. **40**(2): p. 385-396.

198. Chatterjee, A., et al., *Implementing services for Early Infant Diagnosis (EID) of HIV: a comparative descriptive analysis of national programs in four countries*. BMC Public Health, 2011. **11**: p. 553.

199. Prosperi, M.C., et al., *A novel methodology for large-scale phylogeny partition*. Nat Commun, 2011. **2**: p. 321.

200. Price, M.A., et al., *Transmitted HIV type 1 drug resistance among individuals with recent HIV infection in East and Southern Africa*. AIDS Res Hum Retroviruses, 2011. **27**(1): p. 5-12.

201. Hamers, R.L., et al., *HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study*. Lancet Infect Dis, 2011. **11**(10): p. 750-9.

202. Djoko, C.F., et al., *High HIV type 1 group M pol diversity and low rate of antiretroviral resistance mutations among the uniformed services in Kinshasa, Democratic Republic of the Congo*. AIDS Res Hum Retroviruses, 2011. **27**(3): p. 323-9.

203. Hemelaar, J., et al., *Global trends in molecular epidemiology of HIV-1 during 2000-2007*. AIDS, 2011. **25**(5): p. 679-89.

204. Walker, L.M., et al., *Broad neutralization coverage of HIV by multiple highly potent antibodies*. Nature, 2011. **477**(7365): p. 466-70.

205. Santos, A., et al., *Evaluation of the diagnostic performance of the rapid test VIKIA HIV1/2 in a highly complex HIV-1 epidemic*. Diagn Microbiol Infect Dis, 2011. **71**(1): p. 90-2.

206. Gupta, R.K., et al., *Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis.* Lancet, 2012. **380**(9849): p. 1250-8.

207. Sullivan, P.S., et al., *Prevalence of seroconversion symptoms and relationship to set-point viral load: findings from a subtype C epidemic, 1995-2009.* AIDS, 2012. **26**(2): p. 175-84.

208. Yebra, G., et al., *Most HIV type 1 non-B infections in the Spanish cohort of antiretroviral treatment-naive HIV-infected patients (CoRIS) are due to recombinant viruses.* J Clin Microbiol, 2012. **50**(2): p. 407-13.

209. Delatorre, E. and G. Bello, *Phylodynamics of the HIV-1 epidemic in Cuba*. PLoS One, 2013. **8**(9): p. e72448.

210. Stadeli, K.M. and D.D. Richman, *Rates of emergence of HIV drug resistance in resource-limited settings: a systematic review.* Antivir Ther, 2013. **18**(1): p. 115-23.

211. Hall, B.G., *Building phylogenetic trees from molecular data with MEGA.* Mol Biol Evol, 2013. **30**(5): p. 1229-35.

212. Calado, R., et al., *A Prime-Boost Immunization Strategy with Vaccinia Virus Expressing Novel gp120 Envelope Glycoprotein from a CRF02_AG Isolate Elicits Cross-Clade Tier 2 HIV-1 Neutralizing Antibodies.* Vaccines (Basel), 2020. **8**(2).

213. Julg, B. and D.H. Barouch, *Neutralizing antibodies for HIV-1 prevention*. Curr Opin HIV AIDS, 2019. **14**(4): p. 318-324.

214. Cohen, Y.Z. and M. Caskey, *Broadly neutralizing antibodies for treatment and prevention of HIV-1 infection*. Curr Opin HIV AIDS, 2018. **13**(4): p. 366-373.

215. McGuire, A.T., et al., *Diverse recombinant HIV-1 Envs fail to activate B cells expressing the germline B cell receptors of the broadly neutralizing anti-HIV-1 antibodies PG9 and 447-52D.* J Virol, 2014. **88**(5): p. 2645-57.

216. McGuire, A.T., et al., *Specifically modified Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice.* Nat Commun, 2016. **7**: p. 10618.

217. Nityanandam, R. and R. Serra-Moreno, *BCA2/Rabring7 targets HIV-1 Gag for lysosomal degradation in a tetherin-independent manner.* PLoS Pathog, 2014. **10**(5): p. e1004151.

218. Troyano-Hernaez, P., et al., *Update in natural antiretroviral-associated mutations among HIV-2 variants and discrepancies across HIV-2 resistance interpretation tools.* AIDS Res Hum Retroviruses, 2021.

219. Laville, P., M. Petitjean, and L. Regad, *Structural Impacts of Drug-Resistance Mutations Appearing in HIV-2 Protease.* Molecules, 2021. **26**(3).

220. Zhang, Z., et al., *HIV-2 Vif and foamy virus Bet antagonize APOBEC3B by different mechanisms.* Virology, 2021. **554**: p. 17-27.

221. Umunnakwe, C.N., et al., *Specific Guanosines in the HIV-2 Leader RNA are Essential for Efficient Viral Genome Packaging.* J Mol Biol, 2021. **433**(2): p. 166718.

222. Berzow, D., et al., *Human Immunodeficiency Virus-2 (HIV-2): A Summary of the Present Standard of Care and Treatment Options for Individuals Living with HIV-2 in Western Europe.* Clin Infect Dis, 2021. **72**(3): p. 503-509.

223. Hu, Y., et al., *Virus Evolution and Neutralization Sensitivity in an HIV-1 Subtype B' Infected Plasma Donor with Broadly Neutralizing Activity.* Vaccines (Basel), 2021. **9**(4).

224. McFarland, E.J., et al., *Safety, Tolerability, and Pharmacokinetics of a Long-Acting Broadly Neutralizing HIV-1 Monoclonal Antibody VRC01LS in HIV-1-Exposed Newborn Infants.* J Infect Dis, 2021.

225. Sutcliffe, C.G., et al., *The NSEBA Demonstration Project: implementation of a point-of-care platform for early infant diagnosis of HIV in rural Zambia.* Trop Med Int Health, 2021.

226. De Broucker, G., et al., *The cost-effectiveness of scaling-up rapid point-of-care testing for early infant diagnosis of HIV in southern Zambia.* PLoS One, 2021. **16**(3): p. e0248217.

227. Sutcliffe, C.G., et al., *Point-of-care p24 antigen detection for early infant diagnosis of HIV infection: cross-sectional and longitudinal studies in Zambia.* BMC Infect Dis, 2021. **21**(1): p. 118.

228. Etoori, D., et al., *'If the results are negative, they motivate us'. Experiences of early infant diagnosis of HIV and engagement in Option B.* Glob Public Health, 2021. **16**(2): p. 186-200.

229. Krumm, S.A., et al., *Mechanisms of escape from the PGT128 family of anti-HIV broadly neutralizing antibodies.* Retrovirology, 2016. **13**(1): p. 8.

230. Mouquet, H., *Antibody B cell responses in HIV-1 infection.* Trends Immunol, 2014. **35**(11): p. 549-61.

231. Chen, B., et al., *Determining the structure of an unliganded and fully glycosylated SIV gp120 envelope glycoprotein.* Structure, 2005. **13**(2): p. 197-211.

232. Zhu, P., et al., *Distribution and three-dimensional structure of AIDS virus envelope spikes.* Nature, 2006. **441**(7095): p. 847-52.

233. Brandenberg, O.F., et al., *The HIV-1 Entry Process: A Stoichiometric View.* Trends Microbiol, 2015. **23**(12): p. 763-774.

234. Baba, T.W., et al., *Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection*. Nat Med, 2000. **6**(2): p. 200-6.

235. Mascola, J.R., S.S. Frankel, and K. Broliden, *HIV-1 entry at the mucosal surface: role of antibodies in protection*. AIDS, 2000. **14 Suppl 3**: p. S167-74.

236. Hofmann-Lehmann, R., et al., *Postnatal passive immunization of neonatal macaques with a triple combination of human monoclonal antibodies against oral simian-human immunodeficiency virus challenge*. J Virol, 2001. **75**(16): p. 7470-80.

237. Veazey, R.S., et al., *Protection of macaques from vaginal SHIV challenge by vaginally delivered inhibitors of virus-cell fusion*. Nature, 2005. **438**(7064): p. 99-102.

238. Sun, M., et al., *VRC01 antibody protects against vaginal and rectal transmission of human immunodeficiency virus 1 in hu-BLT mice*. Arch Virol, 2016. **161**(9): p. 2449-55.

239. Shibata, R., et al., *Neutralizing antibody directed against the HIV-1 envelope glycoprotein can completely block HIV-1/SIV chimeric virus infections of macaque monkeys*. Nat Med, 1999. **5**(2): p. 204-10.

240. Caskey, M., F. Klein, and M.C. Nussenzweig, *Broadly Neutralizing Antibodies for HIV-1 Prevention or Immunotherapy.* N Engl J Med, 2016. **375**(21): p. 2019-2021.

241. Horwitz, J.A., et al., *HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16538-43.

242. Caskey, M., et al., *Corrigendum: Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117*. Nature, 2016. **535**(7613): p. 580.

243. Learmont, J.C., et al., *Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. A report from the Sydney Blood Bank Cohort.* N Engl J Med, 1999. **340**(22): p. 1715-22.

244. Baba, T.W., et al., *Live attenuated, multiply deleted simian immunodeficiency virus causes AIDS in infant and adult macaques.* Nat Med, 1999. **5**(2): p. 194-203.

245. Giri, M., K.E. Ugen, and D.B. Weiner, *DNA vaccines against human immunodeficiency virus type 1 in the past decade*. Clin Microbiol Rev, 2004. **17**(2): p. 370-89.

246. Estcourt, M.J., A.J. McMichael, and T. Hanke, *DNA vaccines against human immunodeficiency virus type 1*. Immunol Rev, 2004. **199**: p. 144-55.

247. Priddy, F.H., et al., *Safety and immunogenicity of a replication-incompetent adenovirus type 5 HIV-1 clade B gag/pol/nef vaccine in healthy adults.* Clin Infect Dis, 2008. **46**(11): p. 1769-81.

248. Zhang, M., et al., *DNA prime-protein boost using subtype consensus Env was effective in eliciting neutralizing antibody responses against subtype BC HIV-1 viruses circulating in China*. Hum Vaccin Immunother, 2012. **8**(11): p. 1630-7.

249. Chapman, R., et al., *Heterologous prime-boost vaccination with DNA and MVA vaccines, expressing HIV-1 subtype C mosaic Gag virus-like particles, is highly immunogenic in mice*. PLoS One, 2017. **12**(3): p. e0173352.

250. Morris, L., *mRNA vaccines offer hope for HIV*. Nat Med, 2021. **27**(12): p. 2082-2084.

251. Santos-Ferreira, M.O., et al., *A study of seroprevalence of HIV-1 and HIV-2 in six provinces of People's Republic of Angola: clues to the spread of HIV infection.* J Acquir Immune Defic Syndr, 1990. **3**(8): p. 780-6.

252. Bartolo, I., et al., *Rare HIV-1 subtype J genomes and a new H/U/CRF02_AG recombinant genome suggests an ancient origin of HIV-1 in Angola*. AIDS Res Hum Retroviruses, 2016.

253. Violari, A., et al., *Early antiretroviral therapy and mortality among HIV-infected infants.* N Engl J Med, 2008. **359**(21): p. 2233-44.

254. Bitnun, A., et al., *Early initiation of combination antiretroviral therapy in HIV-1-infected newborns can achieve sustained virologic suppression with low frequency of CD4+ T cells carrying HIV in peripheral blood.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2014. **59**(7): p. 1012-9.

255. Catumbela E, F.A., Serrano D, Furtado ML, Gomes M, Shiraishi RW et al., *The Angola HIV epidemic, 2004-11: a case of change or stability?* . Lancet, 2013. **382**(suppl 2): p. s17.

256. Little, S.J., et al., *Antiretroviral-drug resistance among patients recently infected with HIV.* N Engl J Med, 2002. **347**(6): p. 385-94.

257. Bennett, D.E., et al., *Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment.* Antivir Ther, 2008. **13 Suppl 2**: p. 25-36.

258. Albert, T.L.a.J., *Reconstruction of HIV-1 Transmission Chains for Forensic Purposes.* AIDS Reviews, 2000(2): p. 241-251.

259. Los Alamos Sequence Database (2014) Los Alamos Sequence Database. Los Alamos National Laboratory, Los Alamos, New Mexico. Available:

http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html. Accessed 7 August 2014.

260. Frentz, D., et al., *Limited cross-border infections in patients newly diagnosed with HIV in Europe*. Retrovirology, 2013. **10**: p. 36.

261. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-8.

262. Felsenstein, J., *Evolutionary trees from DNA sequences: a maximum likelihood approach*. J Mol Evol, 1981. **17**(6): p. 368-76.

263. Posada, D. and K.A. Crandall, *MODELTEST: testing the model of DNA substitution*. Bioinformatics, 1998. **14**(9): p. 817-8.

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

265. 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, 1999. **73**(1): p. 152-60.

266. Chalmet, K., et al., *Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections*. BMC Infect Dis, 2010. **10**: p. 262.

267. Kumar, S., et al., *AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses*. BMC Bioinformatics, 2009. **10**: p. 357.

268. Lewis, F., et al., *Episodic sexual transmission of HIV revealed by molecular phylodynamics*. PLoS Med, 2008. **5**(3): p. e50.

269. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree 2--approximately maximum-likelihood trees for large alignments.* PLoS One, 2010. **5**(3): p. e9490.

270. *Stanford HIV Drug Resistance Database (2014) Stanford HIV Drug Resistance Database. Stanford University. Available: http://hivdb.stanford.edu/. Accessed 2014 Jan 22.*

271. Bennett, D.E., et al., *Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update*. PLoS One, 2009. **4**(3): p. e4724.

272. Johnson, V.A., et al., *Update of the drug resistance mutations in HIV-1: March 2013*. Top Antivir Med, 2013. **21**(1): p. 6-14.

273. Baxter, J.D., et al., *Genotypic changes in human immunodeficiency virus type 1 protease associated with reduced susceptibility and virologic response to the protease inhibitor tipranavir*. J Virol, 2006. **80**(21): p. 10794-801.

274. Poveda, E., et al., *Prevalence of darunavir resistance mutations in HIV-1-infected patients failing other protease inhibitors.* J Antimicrob Chemother, 2007. **60**(4): p. 885-8.

275. de Meyer, S., et al., *Resistance profile of darunavir: combined 24-week results from the POWER trials.* AIDS Res Hum Retroviruses, 2008. **24**(3): p. 379-88.

276. Holguin, A., et al., *Efficacy of antiretroviral therapy in individuals infected with HIV-1 non-B subtypes.* AIDS Rev, 2006. **8**(2): p. 98-107.

277. van Westen, G.J., et al., *Significantly improved HIV inhibitor efficacy prediction employing proteochemometric models generated from antivirogram data.* PLoS Comput Biol, 2013. **9**(2): p. e1002899.

278. Vermeiren, H., et al., *Prediction of HIV-1 drug susceptibility phenotype from the viral genotype using linear regression modeling*. J Virol Methods, 2007. **145**(1): p. 47-55.

279. Kempf, D.J., et al., *Analysis of the virological response with respect to baseline viral phenotype and genotype in protease inhibitor-experienced HIV-1-infected patients receiving lopinavir/ritonavir therapy.* Antivir Ther, 2002. **7**(3): p. 165-74.

280. Pellegrin, I., et al., *Interpretation of genotype and pharmacokinetics for resistance to fosamprenavir-ritonavir-based regimens in antiretroviral-experienced patients.* Antimicrob Agents Chemother, 2007. **51**(4): p. 1473-80.

281. Pellegrin, I., et al., *Virological responses to atazanavir-ritonavir-based regimens: resistance-substitutions score and pharmacokinetic parameters (Reyaphar study).* Antivir Ther, 2006. **11**(4): p. 421-9.

282. Brenner, B.G., et al., *High rates of forward transmission events after acute/early HIV-1 infection.* J Infect Dis, 2007. **195**(7): p. 951-9.

283. Langford, S.E., J. Ananworanich, and D.A. Cooper, *Predictors of disease progression in HIV infection: a review.* AIDS Res Ther, 2007. **4**: p. 11.

284. Mahnke, Y.D., et al., *Early immunologic and virologic predictors of clinical HIV-1 disease progression.* AIDS, 2013. **27**(5): p. 697-706.

285. *Instituto Nacional de Luta Contra a SIDA - Angola (2010) UNGASS 2010 - Relato´rio sobre o Progresso do Pa ı´s para dar Seguimento aos Compromissos da Sessa˜o Especial sobre VIH e SIDA da Assembleia Geral das Nac¸o˜es Unidas, 2008–2009.* *Avalaible: http://www.unaids.org/en/dataanalysis/knowyourresponse/countryprogressreports/2010countries/angola_2010_country_progress_report_es.pdf. Accessed 2014 Feb 15.*

286. Nwobegahay, J., et al., *Low prevalence of transmitted genetic drug resistance in a cohort of HIV infected naive patients entering antiretroviral treatment programs at two sites in northern South Africa*. J Med Virol, 2012. **84**(12): p. 1839-43.

287. Hunt, G.M., et al., *Surveillance of transmitted HIV-1 drug resistance in Gauteng and KwaZulu-Natal Provinces, South Africa, 2005-2009.* Clin Infect Dis, 2012. **54 Suppl 4**: p. S334-8.

288. Aghokeng, A.F., et al., *Evaluation of transmitted HIV drug resistance among recently-infected antenatal clinic attendees in four Central African countries.* Antivir Ther, 2009. **14**(3): p. 401-11.

289. Castelbranco, E.P., et al., *Frequency of primary resistance to antiretroviral drugs and genetic variability of HIV-1 among infected pregnant women recently diagnosed in Luanda-Angola*. AIDS Res Hum Retroviruses, 2010. **26**(12): p. 1313-6.

290. Rambaut, A., et al., *The causes and consequences of HIV evolution*. Nat Rev Genet, 2004. **5**(1): p. 52-61.

291. *Ministério da Saúde de Angola (2009) Boletim Epidemiológico VIH e SIDA. Instituto Nacional de Luta Contra a SIDA. Ano 0 N˚ 02 - Julho a Dezembro de 2009. Angola. Avalaible: http://sida.gov.ao/. Accessed 20 July 2011.*

292. Hue, S., et al., *Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups.* Proc Natl Acad Sci U S A, 2005. **102**(12): p. 4425-9.

293. Brenner, B., M.A. Wainberg, and M. Roger, *Phylogenetic inferences on HIV-1 transmission: implications for the design of prevention and treatment interventions.* AIDS, 2013. **27**(7): p. 1045-57.

294. Smit, P.W., et al., *Systematic review of the use of dried blood spots for monitoring HIV viral load and for early infant diagnosis.* PLoS One, 2014. **9**(3): p. e86461.

295. Field, N., et al., *Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement.* Lancet Infect Dis, 2014. **14**(4): p. 341-52.

296. WHO, *HIV Assays: Laboratory performance and other operational characteristics rapid diagnostic tests Combined detection of HIV-1/2 antibodies and discriminatory detection of HIV-1 and HIV-2 antibodies.* . 2015, World Health Organization: Geneva.

297. Ghosh, S.K., et al., *A molecular clone of HIV-1 tropic and cytopathic for human and chimpanzee lymphocytes.* Virology, 1993. **194**(2): p. 858-64.

298. Folks, T.M., et al., *Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a chronically infected T-cell clone.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(7): p. 2365-8.

299. Clouse, K.A., et al., *Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone.* Journal of immunology, 1989. **142**(2): p. 431-8.

300. Denny, T., et al., *Lymphocyte subsets in healthy children during the first 5 years of life.* JAMA, 1992. **267**(11): p. 1484-8.

301. Tetali, S., et al., *Virus load as a marker of disease progression in HIV-infected children.* AIDS Res Hum Retroviruses, 1996. **12**(8): p. 669-75.

302. Chang, J., et al., *Field evaluation of Abbott Real Time HIV-1 Qualitative test for early infant diagnosis using dried blood spots samples in comparison to Roche*

COBAS Ampliprep/COBAS TaqMan HIV-1 Qual test in Kenya. Journal of virological methods, 2014. **204**: p. 25-30.

303. Abbott and Molecular, *RealTime HIV-1 Qualitative. Product Description*, Abbott Molecular.

304. WHO, *Sources and prices of selected medicines and diagnostics for people living with HIV/AIDS/a joint UNICEF, UNAIDS, WHO, MSF project- 6th edition*. 2005, World Health Organization: Geneva.

305. Lussiana, C., et al., *Effectiveness of a prevention of mother-to-child HIV transmission programme in an urban hospital in Angola*. PLoS One, 2012. **7**(4): p. e36381.

306. Fischer, A., et al., *Simple DNA extraction method for dried blood spots and comparison of two PCR assays for diagnosis of vertical human immunodeficiency virus type 1 transmission in Rwanda*. J Clin Microbiol, 2004. **42**(1): p. 16-20.

307. Sciences, G.H.L., *Reliable extraction of DNA from Whatman™ FTA™ cards GE Healthcare Application note 28-9822-22 AA*. 2010, GE Healthcare Life Sciences.

308. Swanson, P., et al., *Impact of human immunodeficiency virus type 1 (HIV-1) genetic diversity on performance of four commercial viral load assays: LCx HIV RNA Quantitative, AMPLICOR HIV-1 MONITOR v1.5, VERSANT HIV-1 RNA 3.0, and NucliSens HIV-1 QT*. J Clin Microbiol, 2005. **43**(8): p. 3860-8.

309. Amendola, A., et al., *Underevaluation of HIV-1 plasma viral load by a commercially available assay in a cluster of patients infected with HIV-1 A/G circulating recombinant form (CRF02).* Journal of acquired immune deficiency syndromes, 2002. **31**(5): p. 488-94.

310. Antunes, R., et al., *Evaluation of the clinical sensitivities of three viral load assays with plasma samples from a pediatric population predominantly infected with*

human immunodeficiency virus type 1 subtype G and BG recombinant forms. J Clin Microbiol, 2003. **41**(7): p. 3361-7.

311. Nkengasong, J.N., et al., *Quantification of RNA in HIV type 1 subtypes D and G by NucliSens and Amplicor assays in Abidjan, Ivory Coast.* AIDS Res Hum Retroviruses, 1999. **15**(6): p. 495-8.
312. Hu, X., et al., *Profiling the neutralizing antibody response in chronically HIV-1 CRF07_BC-infected intravenous drug users naive to antiretroviral therapy.* Sci Rep, 2017. **7**: p. 46308.
313. Brandenberg, O.F., et al., *Predicting HIV-1 transmission and antibody neutralization efficacy in vivo from stoichiometric parameters.* PLoS Pathog, 2017. **13**(5): p. e1006313.
314. Bonsignori, M., et al., *Staged induction of HIV-1 glycan-dependent broadly neutralizing antibodies.* Sci Transl Med, 2017. **9**(381).
315. Anthony, C., et al., *Cooperation between Strain-Specific and Broadly Neutralizing Responses Limited Viral Escape and Prolonged the Exposure of the Broadly Neutralizing Epitope.* J Virol, 2017. **91**(18).
316. Martinez, D.R., et al., *Maternal Binding and Neutralizing IgG Responses Targeting the C-Terminal Region of the V3 Loop Are Predictive of Reduced Peripartum HIV-1 Transmission Risk.* J Virol, 2017. **91**(9).
317. Rocha, C., et al., *Potency of HIV-2-specific antibodies increase in direct association with loss of memory B cells.* AIDS, 2017. **31**(17): p. 2431-2433.
318. Doria-Rose, N.A., et al., *Mapping Polyclonal HIV-1 Antibody Responses via Next-Generation Neutralization Fingerprinting*. PLoS Pathog, 2017. **13**(1): p. e1006148.

319. Martin, F., et al., *Early infant diagnosis of HIV-1 infection in Luanda, Angola, using a new DNA PCR assay and dried blood spots.* PLoS One, 2017. **12**(7): p. e0181352.

320. Spencer, D.A., et al., *Polyfunctional Tier 2-Neutralizing Antibodies Cloned following HIV-1 Env Macaque Immunization Mirror Native Antibodies in a Human Donor*. J Immunol, 2021. **206**(5): p. 999-1012.

321. Spencer, D.A., et al., *Advancing HIV Broadly Neutralizing Antibodies: From Discovery to the Clinic*. Front Public Health, 2021. **9**: p. 690017.

322. Kwong, P.D. and J.R. Mascola, *HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure.* Immunity, 2018. **48**(5): p. 855-871.

323. Simonich, C.A., et al., *HIV-1 Neutralizing Antibodies with Limited Hypermutation from an Infant.* Cell, 2016. **166**(1): p. 77-87.

324. Marcelino, J.M., et al., *Resistance to antibody neutralization in HIV-2 infection occurs in late stage disease and is associated with X4 tropism.* AIDS, 2012. **26**(18): p. 2275-84.

325. Rocha, C., et al., *Evolution of the human immunodeficiency virus type 2 envelope in the first years of infection is associated with the dynamics of the neutralizing antibody response*. Retrovirology, 2013. **10**: p. 110.

326. Serra, P.A., N. Taveira, and R.C. Guedes, *Computational Modulation of the V3 Region of Glycoprotein gp125 of HIV-2*. Int J Mol Sci, 2021. **22**(4).

327. Kostrikis, L.G., et al., *Quantitative analysis of serum neutralization of human immunodeficiency virus type 1 from subtypes A, B, C, D, E, F, and I: lack of direct correlation between neutralization serotypes and genetic subtypes and evidence for prevalent serum-dependent infectivity enhancement.* J Virol, 1996. **70**(1): p. 445-58.

328. Trkola, A., et al., *Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1*. J Virol, 1996. **70**(2): p. 1100-8.

329. Ditse, Z., et al., *Effect of HIV Envelope Vaccination on the Subsequent Antibody Response to HIV Infection*. mSphere, 2020. **5**(1).

330. Marcelino, R., et al., *Antibody response against selected epitopes in the HIV-1 envelope gp41 ectodomain contributes to reduce viral burden in HIV-1 infected patients.* Sci Rep, 2021. **11**(1): p. 8993.

331. Faria, N.R., et al., *HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations.* Science, 2014. **346**(6205): p. 56-61.

332. Platt, E.J., et al., *Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1.* J Virol, 1998. **72**(4): p. 2855-64.

333. Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.* Mol Biol Evol, 2013. **30**(12): p. 2725-9.

334. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.* Syst Biol, 2010. **59**(3): p. 307-21.

335. Weaver, S., et al., *Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes.* Mol Biol Evol, 2018. **35**(3): p. 773-777.

336. Kosakovsky Pond, S.L. and S.D. Frost, *Not so different after all: a comparison of methods for detecting amino acid sites under selection.* Mol Biol Evol, 2005. **22**(5): p. 1208-22.

337. Zhang, M., et al., *Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin.* Glycobiology, 2004. **14**(12): p. 1229-46.

338. Mascola, J.R., et al., *Recommendations for the design and use of standard virus panels to assess neutralizing antibody responses elicited by candidate human immunodeficiency virus type 1 vaccines.* J Virol, 2005. **79**(16): p. 10103-7.

339. Metsalu, T. and J. Vilo, *ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap.* Nucleic Acids Res, 2015. **43**(W1): p. W566-70.

340. Firth, H.V., et al., *DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources.* Am J Hum Genet, 2009. **84**(4): p. 524-33.

341. *R Core Team, R: A language and environment for statistical computing, 2017. Disponível em https://www.R-project.org/.*

342. *Wickham, H., ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, 2016.*

343. Mellors, J.W., et al., *Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection.* Ann Intern Med, 1997. **126**(12): p. 946-54.

344. Brunel, F.M., et al., *Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody.* J Virol, 2006. **80**(4): p. 1680-7.

345. Barroso, H., et al., *Evolutionary and structural features of the C2, V3 and C3 envelope regions underlying the differences in HIV-1 and HIV-2 biology and infection.* PLoS One, 2011. **6**(1): p. e14548.

346. Mabvakure, B.M., et al., *Positive Selection at Key Residues in the HIV Envelope Distinguishes Broad and Strain-Specific Plasma Neutralizing Antibodies.* J Virol, 2019. **93**(6).

347. Dingens, A.S., et al., *An Antigenic Atlas of HIV-1 Escape from Broadly Neutralizing Antibodies Distinguishes Functional and Structural Epitopes.* Immunity, 2019. **50**(2): p. 520-532 e3.

348. Bricault, C.A., et al., *HIV-1 Neutralizing Antibody Signatures and Application to Epitope-Targeted Vaccine Design.* Cell Host Microbe, 2019. **26**(2): p. 296.

349. Sebastiao, C.S., J. Morais, and M. Brito, *Factors Influencing HIV Drug Resistance among Pregnant Women in Luanda, Angola: Findings from a Cross-Sectional Study.* Trop Med Infect Dis, 2021. **6**(1).

350. Bhiman, J.N., et al., *Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies.* Nat Med, 2015. **21**(11): p. 1332-6.

351. Liao, H.X., et al., *Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus.* Nature, 2013. **496**(7446): p. 469-76.

352. Beirnaert, E., et al., *Identification and characterization of sera from HIV-infected individuals with broad cross-neutralizing activity against group M (env clade A-H) and group O primary HIV-1 isolates.* J Med Virol, 2000. **62**(1): p. 14-24.

353. Kouyos, R.D., et al., *Tracing HIV-1 strains that imprint broadly neutralizing antibody responses.* Nature, 2018. **561**(7723): p. 406-410.

354. Sather, D.N., et al., *Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection.* J Virol, 2009. **83**(2): p. 757-69.

355. Moir, S. and A.S. Fauci, *B-cell responses to HIV infection.* Immunol Rev, 2017. **275**(1): p. 33-48.

356. Badura, R., et al., *Early ART in Acute HIV-1 Infection: Impact on the B-Cell Compartment.* Front Cell Infect Microbiol, 2020. **10**: p. 347.

357. Moir, S., et al., *B cells in early and chronic HIV infection: evidence for preservation of immune function associated with early initiation of antiretroviral therapy.* Blood, 2010. **116**(25): p. 5571-9.

358. Gach, J.S., et al., *HIV-1 specific antibody titers and neutralization among chronically infected patients on long-term suppressive antiretroviral therapy (ART): a cross-sectional study.* PLoS One, 2014. **9**(1): p. e85371.

359. Cizmeci, D., et al., *Distinct clonal evolution of B-cells in HIV controllers with neutralizing antibody breadth.* Elife, 2021. **10**.

360. Kong, R., et al., *Broad and potent neutralizing antibody responses elicited in natural HIV-2 infection.* J Virol, 2012. **86**(2): p. 947-60.

361. Rodriguez, S.K., et al., *Comparison of heterologous neutralizing antibody responses of human immunodeficiency virus type 1 (HIV-1)- and HIV-2-infected Senegalese patients: distinct patterns of breadth and magnitude distinguish HIV-1 and HIV-2 infections.* J Virol, 2007. **81**(10): p. 5331-8.

362. Bai, H., et al., *The breadth of HIV-1 neutralizing antibodies depends on the conservation of key sites in their epitopes.* PLoS Comput Biol, 2019. **15**(6): p. e1007056.

363. Rademeyer, C., et al., *Genetic characteristics of HIV-1 subtype C envelopes inducing cross-neutralizing antibodies.* Virology, 2007. **368**(1): p. 172-81.

364. Kumar, S., et al., *An HIV-1 Broadly Neutralizing Antibody from a Clade C-Infected Pediatric Elite Neutralizer Potently Neutralizes the Contemporaneous and Autologous Evolving Viruses.* J Virol, 2019. **93**(4).

365. Cheedarla, N., et al., *Evolution of Neutralization Response in HIV-1 Subtype C-Infected Individuals Exhibiting Broad Cross-Clade Neutralization of HIV-1 Strains.* Front Immunol, 2018. **9**: p. 618.

366. Lei, L., et al., *The HIV-1 Envelope Glycoprotein C3/V4 Region Defines a Prevalent Neutralization Epitope following Immunization.* Cell Rep, 2019. **27**(2): p. 586-598 e6.

367. Vueba, A.N., et al., *Prevalence of HIV and hepatitis B virus among pregnant women in Luanda (Angola): geospatial distribution and its association with socio-demographic and clinical-obstetric determinants.* Virol J, 2021. **18**(1): p. 239.

368. Matsinhe, M., et al., *Inpatient Point-of-Care HIV Early Infant Diagnosis in Mozambique to Improve Case Identification and Linkage to Antiretroviral Therapy.* Glob Health Sci Pract, 2021. **9**(1): p. 31-39.

369. Leruez-Ville, M., et al., *Detection of cytomegalovirus DNA on dried blood spots collected from infants infected with HIV: an in-house method adaptable in resource-limited settings.* J Virol Methods, 2013. **193**(2): p. 503-7.

370. Moragas, M., M.D. Golemba, and A. Mangano, *A new highly sensitive single-tube nested real-time PCR assay: Clinical utility in perinatal HIV-1 diagnosis.* J Virol Methods, 2021. **297**: p. 114273.

371. De Vitis, E., et al., *Real-time polymerase chain reaction on filter paper spotted samples: a gateway to molecular diagnosis of invasive bacterial diseases for rural areas in low-income countries.* Trans R Soc Trop Med Hyg, 2022. **116**(3): p. 233-241.

372. Guan, B., et al., *Sensitive extraction-free SARS-CoV-2 RNA virus detection using a chelating resin*. iScience, 2021. **24**(9): p. 102960.

373. *WHO, Diagnostic testing for SARS-CoV-2: interim guidance, 2020. Available: https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2. Accessed 10th December 2021*

374. Anthony, C., et al., *Correction for Anthony et al., "Cooperation between Strain-Specific and Broadly Neutralizing Responses Limited Viral Escape and Prolonged the Exposure of the Broadly Neutralizing Epitope".* J Virol, 2018. **92**(9).

375. Bottiger, B., et al., *Envelope cross-reactivity between human immunodeficiency virus types 1 and 2 detected by different serological methods: correlation between cross-neutralization and reactivity against the main neutralizing site.* J Virol, 1990. **64**(7): p. 3492-9.

376. Vidyavijayan, K.K., et al., *Cross Type Neutralizing Antibodies Detected in a Unique HIV-2 Infected Individual From India.* Front Immunol, 2018. **9**: p. 2841.

377. Doria-Rose, N.A., et al., *Breadth of human immunodeficiency virus-specific neutralizing activity in sera: clustering analysis and association with clinical variables.* J Virol, 2010. **84**(3): p. 1631-6.

378. Seabright, G.E., et al., *Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions.* Structure, 2020. **28**(8): p. 897-909 e6.

379. Gray, E.S., et al., *N-linked glycan modifications in gp120 of human immunodeficiency virus type 1 subtype C render partial sensitivity to 2G12 antibody neutralization.* J Virol, 2007. **81**(19): p. 10769-76.

380. Polzer, S., et al., *The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization.* Virology, 2002. **304**(1): p. 70-80.

381. Chackerian, B., L.M. Rudensey, and J. Overbaugh, *Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies.* J Virol, 1997. **71**(10): p. 7719-27.

382. Sagar, M., et al., *Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity*. J Virol, 2006. **80**(19): p. 9586-98.

383. Han, Q., et al., *Difficult-to-neutralize global HIV-1 isolates are neutralized by antibodies targeting open envelope conformations.* Nat Commun, 2019. **10**(1): p. 2898.