HHS Public Access

Author manuscript

J Acquir Immune Defic Syndr. Author manuscript; available in PMC 2020 August 15.

Published in final edited form as:

J Acquir Immune Defic Syndr. 2019 August 15; 81(5): 578–584. doi:10.1097/QAI.0000000000002076.

New Subtype B Containing HIV-1 Circulating Recombinant of sub-Saharan Africa Origin in Nigerian Men who have Sex with Men

Erik Billings, PhD[*,1,2], **Gustavo H. Kijak, PharmD, PhD**[**,1,2], **Eric Sanders-Buell, BS**[1,2], **Nicaise Ndembi, PhD**[3], **Anne Marie O'Sullivan, BS**[1,2], **Sylvia Adebajo, MD, PhD**[4], **Afoke Kokogho, MBBS**[1,2], **Mark Milazzo, MS**[1,2], **Kara Lombardi, MS**[1,2], **Stefan Baral, MD, MPH**[5], **Rebecca Nowak, PhD**[6], **Habib Ramadhani, MD, PhD**[6], **Robert Gramzinski, PhD**[1,7], **Merlin L. Robb, MD**[1,2], **Nelson L. Michael, MD, PhD**[1,7], **Manhattan E. Charurat, PhD**[6], **Julie Ake, MD**[1,7], **Trevor A. Crowell, MD, PhD**[1,2], **Sodsai Tovanabutra, PhD**[1,2], and **MHRP Viral Sequencing Core and the TRUST/RV368 Study Group**

[1]U.S. Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD, USA

[2]Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA

[3]Institute of Human Virology-Nigeria, Abuja, Nigeria

[4]Population Council Nigeria, Abuja, Nigeria

[5]Johns Hopkins University, Baltimore, MD, USA

[6]Institute of Human Virology, University of Maryland, Baltimore, MD, USA

[7]Walter Reed Army Institute of Research, Silver Spring, MD, USA

[*]Current affiliation: Children's National Hospital, Washington, DC, USA

[**]Current affiliation: GSK Vaccines, Rockville, MD, USA

Abstract

Background—HIV-1 circulating recombinant forms (CRF) containing subtype B are uncommon in sub-Saharan Africa. Prevalent infections observed during enrollment of a prospective study of men who have sex with men (MSM) from Lagos, Nigeria revealed the presence of a family of subtype B and CRF02_AG recombinants. This report describes the HIV-1 genetic diversity within a high-risk, high-prevalence, and previously undersampled cohort of Nigerian MSM.

Methods—Between 2013 and 2016, 672 MSM were enrolled at the Lagos site of the TRUST/RV368 study. Prevalent HIV-1 infections were initially characterized via *pol* sequencing and

Corresponding Author: Sodsai Tovanabutra, PhD, 503 Robert Grant Avenue, Room 2N25, Silver Spring, MD 20910, USA, Tel: 301-319-9993, stovanabutra@hivresearch.org.

phylogenetic subtyping analysis. Samples demonstrating the presence of subtype B were further characterized by near full-length sequencing, phylogenetic, and Bayesian analyses.

Results—Within this cohort, HIV-1 prevalence was 59%. The major subtype was CRF02_AG (57%), followed by CRF02/B recombinants (15%), subtype G (13%), and smaller amounts of A1, B, and other recombinants. Nine clusters of closely related *pol* sequences indicate ongoing transmission events within this cohort. Among the CRF02_AG/B, a new CRF was identified and termed CRF95_02/B. Shared risk factors and Bayesian phylogenetic inference of the new CRF95_02B and the similarly structured CRF56_cpx, indicate a Nigerian or West African origin of CRF56_cpx prior to its observation in France.

Conclusion—With high HIV-1 prevalence, new strains, and multiple transmission networks, this cohort of Nigerian MSM represents a previously hidden reservoir of HIV-1 strains, including the newly-identified CRF95_02B and closely-related CRF56_cpx. These strains will need to be considered during vaccine selection and development in order to optimize the design of a globally effective HIV-1 vaccine.

Keywords

Molecular epidemiology; CRF; MSM; Nigeria

INTRODUCTION

As of 2015, Nigeria has the largest population in Sub-Saharan Africa and second highest HIV-1 burden in the world.[1,2] National surveys have observed substantial differences in HIV-1 prevalence among all reproductive aged adults (3.4%)[3] and key affected populations (KAPs), including female sex workers (8.6–19.4% depending on brothel-based or mobile business type), their clients, men who have sex with men (MSM; 22.9%), and people who inject drugs (3.4%).[4] As a risk group, MSM in Nigeria face general social stigmatization and potential criminal prosecution.[5] These negative motivators have led MSM to be less willing to seek HIV testing and engage in care, which has resulted in general undersampling[6,7] of the MSM population and a lack of MSM-specific molecular epidemiology data. Over the past decade, some pioneering studies[8–10] have begun to quantify the extent of HIV-1 prevalence among Nigerian MSM. However, studies intended to molecularly characterize HIV-1 strains among MSM cohorts have only recently begun. As part of efforts to address the monitoring and treatment needs of MSM in Nigeria, the TRUST/RV368 study was initiated to monitor, characterize, and provide HIV-1 treatment for cohorts of MSM and transgender women (TGW) in the cities of Abuja and Lagos.[11]

Molecular characterization of an epidemic establishes the genetic background of HIV-1 subtypes or circulating recombinant forms (CRFs) that are typically found within that region. The genotypes of circulating HIV-1 strains guide vaccine development, testing kit reagent selection, and phylogenetic tracking of the epidemic's progress. Most HIV-1 infections belong to group M, which contains 9 subtypes (A, B, C, D, F, G, H, J, and K) and greater than 96 CRFs. Of those, subtypes A, C, D, G and CRF02_AG represent the majority (89%) of the publicly available sequences from sub-Saharan Africa (hiv.lanl.gov), with the remaining fraction composed of other recombinants and smaller amounts of subtypes B, F,

H, J, and K. Subtype C dominates in South Africa, the surrounding regions, and Ethiopia; subtypes A, C, and D compete for dominance in Kenya, Tanzania, and Uganda of east Africa; and CRF02_AG and subtype G comprise the major strains of Nigeria and its west African neighbors. In regions where multiple subtypes are found circulating in significant proportions, superinfections can lead to the development of unique recombinant forms (URFs) that can be transmitted and result in the establishment and growth of new CRFs. The HIV-1 epidemics of east and west Africa are particularly diverse and complex due to the presence of multiple co-circulating subtypes and a resulting plethora of URFs and CRFs.[12,13]

Subtypes A, B, C, D, G and recombinants CRF01_AE and CRF02_AG represent over 90% of prevalent global HIV-1 infections.[14] As a region, sub-Saharan Africa is home to approximately 70% of people living with HIV-1[15] and although it contains a genetically diverse collection of HIV-1 strains, it does not have significant amounts of subtype B or CRF01_AE (0.29% and 0.24%, respectively, of sub-Saharan sequences as of April, 2018: hiv.lanl.gov). The global MSM networks are known to be a risk factor for HIV-1 transmission. CRF56_cpx, was discovered in year 2010 among French MSM with close links to West Africa indicated the existing network of MSM between the continents. We hypothesized that a molecular epidemiology study of HIV-1 strains in MSM enrolled in a study conducted in Lagos, Nigeria would provide new insight information on circulating and recently introduced strains to the area. Unexpectedly, this study provides a population level example of a hidden reservoir of infectious CRF02_AG/B recombinant strains within a well-known, yet understudied group of high-risk individuals. Among these recombinant strains, 4 of them shared similar recombinant structure and have been designated as the newly discovered CRF95_02B. Additionally, phylodynamic analysis is presented that these previously unknown subtype B and CRF02_AG recombinants share a common origin with CRF56_cpx.

METHODS

Study participants

Between 2013 and 2016, the Lagos site of the TRUST/RV368 prospective observational cohort study enrolled 672 people assigned the male sex at birth and followed them at 3 month intervals for 18 months. Males aged 18 years and older who reported receptive or insertive anal intercourse with a male partner within the twelve months prior to enrollment were recruited using respondent driven sampling as previously described.[11] Upon enrollment, 614 out of 672 participants were tested for HIV and were administered a questionnaire to assess risk behaviors. HIV testing was performed on finger-stick blood samples using the Government of Nigeria HIV diagnostic algorithm, which employs parallel sero-reactive rapid device tests: Determine® (Alere, Waltham, MA, USA) and Uni-gold® (Trinity Biotech, Co-Wicklow, Ireland) with Stat-Pak HIV (Chembio, Medford, NY, USA) serving as a tie-breaker.

Sequencing

Plasma samples from 150 out of 362 HIV-1 positive participants with a viral load ≥ 1000 copies/mL underwent viral RNA extraction, cDNA generation, and near-endpoint dilution prior to PCR amplification, purification, and sequencing with an ABI 3730XL capillary sequencer as previously described.[17] *Pol* sequences were obtained using nested primers selected such that resulting PCR amplification products covered HXB2 referenced positions 2273–3869. Samples containing subtype B in *pol* were subjected to near-full genome sequencing via PCR amplification of complete (HXB2: 789–8713) or half-length (HXB2: 789–5852 and 4560–8713) genome products.

Phylogenetic analysis

Viral sequence subtypes were determined using a combination of the HIV-1 Genotyping Tool at NCBI,[18] a modified local installation of the jumping profile Hidden Markov Model (jpHMM) tool[19] which incorporated additional reference sequences from West Africa (see Reference List 1, Supplemental Digital Content), BLAST analysis (HIV BLAST, hiv.lanl.gov), and phylogenetic tree analysis. Recombinant genome structures were determined by first approximating the genome breakpoints using bootscanning with Simplot[20] v3.5.1 and the previously described jpHMM tool, followed by refinement of the breakpoints by visual inspection (SeqPublish, hiv.lanl.gov) using six reference sequences each for subtype B and CRF02_AG (see Reference List 2, Supplemental Digital Content). Confirmation of the parent subtype identity was provided via BLAST and bootstrap analysis of maximum likelihood trees constructed from the resulting sub-genome fragments. Recombinant genome structures were visualized using RecDraw.[21] Times to most recent common ancestor (tMRCA) were calculated using the Bayesian Markov chain Monte Carlo algorithm within BEAST[22] v1.8.4 with the GTR substitution model, strict clock, gamma site (5) heterogeneity model, exponential growth tree prior, and year of collection (study samples) or year of publication (reference sequences) as listed by the Los Alamos National Laboratory HIV sequence database (hiv.lanl.gov). Model parameters for tMRCA calculations were selected following comparisons of results which probed substitution models (GTR and HKY), site heterogeneity models (gamma and gamma+invariant), codon partitions (none and three), clock types (strict and relaxed with lognormal and gamma distributions), and tree priors (constant size and exponential growth). Chain lengths were selected such that the effective sample size of the posterior was greater than 200. The mean tMRCA and 95% highest posterior density values for each clade were reported by Tracer v1.6.0.[23] Multiple sequence alignments were generated with HIVAlign[24] using the Markov model[25] option. Maximum likelihood phylogenetic trees were calculated with DIVEIN[26] using the GTR model (with estimated invariant sites and gamma distribution), 100 replicates, and masked hypervariable regions in the variable loops of *env*. Transmission network calculations were performed with the TN93[27] algorithm (github.com/veg/tn93) using the average ambiguous base option, a genetic distance cut-off of 1.5%, and the *pol* sequences described above after masking of Drug-Resistance Surveillance sites.[28] The HIV-1 sequences described here are available under GenBank accession numbers: MH654824 - MH654973 for *pol*, and MH666153 - MH666162 for full-length genomes.

RESULTS

Study participant demographics, social environment, and prevalence of HIV-1, chlamydia, and gonorrhea for the combined locations of Abuja and Lagos have been previously published.[5,11,29,30] During enrollment, 614 participants were screened for HIV-1 and 362 prevalent infections were observed within the Lagos cohort yielding an HIV-1 prevalence rate of 59%. Of these, 205 had HIV-1 RNA ≥1000 copies/mL within 120 days of diagnosis and 150 underwent analysis of *pol* sequences. The participants who underwent viral sequencing had median age 23 (interquartile range [IQR] 20–27) years, median CD4 437.5 (IQR 344–547) cells/mm^3, and median HIV-1 RNA 4.51 (IQR 4.15–4.92) log_{10}copies/ml. Among them, self-reported gender identity included 76.7% men, 14.7% women, and 8.7% other/unknown. Fifty-seven percent reported their sexual orientation as gay/homosexual while the others were bisexual. Most of the participants (70%) had senior secondary level and 26% had higher education. Their occupations ranged from not working (23%), student (25%), professional/self-employed (15%), entertainment (9%), labor (5%), and other/unknown (23%). The phylogenetic analysis results revealed an HIV-1 subtype distribution of 57% CRF02_AG, 15% CRF02_AG/B recombinants, 13% subtype G, 3% subtype A1, 2% subtype B, and smaller amounts of recombinants containing CRF02_AG, A1, or G.

Application of a genetic distance cut-off of 1.5% in *pol* identified nine transmission network clusters containing two to five participants each and composed of pure CRF02_AG (five clusters) or subtype G (four clusters) strains. A maximum likelihood tree showing the branching patterns of subtype and transmission network clusters within the *pol* sequences is shown in Figure 1.

A total of 23 *pol* sequences were identified as recombinants between CRF02_AG and subtype B. The participants who harbored these recombinants did not have different demographic characteristics and geographic location compared to others infected with non-CRF02_AG/B recombinants (see Table 1, Supplemental Digital Content). From the cluster of 15 CRF02_AG/B recombinants circled in Figure 1, eight samples forming mini-clusters with similar recombination patterns in *pol* were selected for near full-length sequencing. Recombinant breakpoint analysis of the full-length genomes revealed parental subtype B contributions to the *gag* and *pol* genes of most of the selected samples amidst a background of CRF02_AG and a small amount of subtype G (in one sample). Five of the full-length genomes were identified as representatives of a new CRF, termed CRF95_02B, as shown in Figure 2a. This new CRF has a background of CRF02_AG with contributions from subtype B to the C-terminus of *gag p17,* thumb of *RT,*[31] αA-αD motifs of *RNase,*[32] and the catalytic core of *integrase*.[33] The remaining strains share various genomic breakpoints in common with the CRF95_02B strains, but have diverged through additional recombination or evolutionary events, Figure 2b.

Recombinant breakpoint analysis identified eight breakpoints within the HIV-1 genomes representing the new CRF95_02B. The breakpoints define nine fragments, which were aligned with subtype reference strains to build maximum-likelihood trees for subtype confirmation through bootstrap assessment (see Supplemental Figures 1–3, Supplemental Digital Content). Bootstrap values ≥ 85% support the assigned parent subtype for seven of

nine fragments. Supportive bootstrap values for the two remaining fragments could not be obtained owing to the length of Fragment II (104 nucleotides) and the divergence of Fragment V from CRF02_AG and subtype A, consequently the parental subtypes for those two fragments were confirmed as subtype B and CRF02_AG, respectively, using BLAST similarity searching.

BLAST similarity searching of full-length and individual fragments of the CRF95_02B strains (as described above) indicated the presence of a potentially related recombinant: CRF56_cpx[34]. Multiple sequence alignments from the *gag-pol, int,* and *gp120-gp41* regions containing only a single subtype in common to both CRF95_02B and CRF56_cpx strains were used to construct maximum likelihood trees, which supported the genetic similarity implied by the BLAST results and genomic structures (see Supplemental Figure 4, Supplemental Digital Content). The results of tMRCA calculations using the same genomic regions described above in an alignment containing 108 CRF02_AG and 58 subtype B reference sequences (see Reference List 3, Supplemental Digital Content) displayed large differences in the predicted tMRCA for the CRF95_02B and CRF56_cpx clades. The mean tMRCA for the *gag-pol, int,* and *gp120-gp41* regions ranged from 26.5 to 31.9 years for CRF95_02B and 7.3 to 8.8 years for CRF56_cpx. As a single clade, the mean tMRCA for those same regions ranged from 27.0 to 34.9 years, which largely overlaps the predicted tMRCA for the CRF95_02B clade. The mean and 95% highest posterior density interval for each genomic region per clade is shown in Supplemental Figure 5b, Supplemental Digital Content.

DISCUSSION

The TRUST/RV368 study was implemented to address the HIV-1 treatment needs of MSM in Nigeria and obtain the molecular information necessary to guide HIV-1 vaccine development efforts. The results presented in this report describe the molecular genetics of HIV-1 within a cohort of Nigerian MSM at the time of study enrollment. The large proportions of CRF02_AG (57%) and subtype G (13%) observed within this cohort are consistent with the predominant strains observed among all Nigerian people living with HIV[12] as well as other specific populations in Nigeria.[35] However, the presence of a significant proportion of CRF02_AG/B recombinants (15%) is a new finding within a West African cohort. Detailed examination of these strains revealed the presence of a new CRF termed CRF95_02B. Analysis of the CRF95_02B strains suggested a relationship with the CRF56_cpx[16] strains observed in France in 2010. Calculations of tMRCA for shared regions across the genome show large differences between the mean tMRCA of CRF95_02B (26.5–31.9 years) and CRF56_cpx (7.3–8.8 years). The tMRCA of the CRF56_cpx strains likely represents sampling of closely related transmissions of CRF56_cpx (as evidenced by the similarity between the CRF56_cpx strains shown in Supplemental Figure 4, Supplemental Digital Content). However, the tMRCA of a combined clade of CFR95_02B and CRF56_cpx strains is nearly identical to the tMRCA for the CRF95_02B strains alone. When viewed in the context of shared risk factors,[16] similar recombinant breakpoints, and similar phylogenetic clustering, the combined tMRCA calculations are supportive of a common origin for both CRF95_02B and CRF56_cpx. If correct, then the current structure of the CRF56_cpx genome is the result of further recombination events since that time. In

the past, there has been strong correlation of MSM interaction between HIV-1 infected individuals of West Africa origin and the French MSM populations as supported by the existence of Unique Recombinant Forms (URFs) B/CRF02_AG, URF5_B/02/G and CRF56_cpx identified among MSM in France. The CRF02_AG/B recombinants identified in this cohort might be recombinant strains generated previously among MSM who traveled between West Africa and France.

The newly described CRF95_02B and its cluster of related strains was readily detected in this cohort. The diversity evident in the phylogenetic cluster patterns and tMRCA results shown here, indicate it has been in circulation within this cohort for decades. CRF95_02B is only newly identified due to under-surveillance of this group. Although only prevalent strains are presented here, the presence of transmission networks in this cohort indicates a possible exposure to this CRF in the community (manuscript in preparation). With high HIV-1 prevalence, new strains, and multiple transmission networks, this cohort of Nigerian MSM represents a previously hidden reservoir of HIV-1 strains that will need to be considered during vaccine immunogen selection and development. This study demonstrates the need to continue epidemiological monitoring of all groups engaged in high-risk activities, identifying their risks, and finding solutions for prevention. The cross-reactivity of both subtype C and mosaic HIV-1 vaccines currently in efficacy studies in southern Africa[36,37] to this new CRF, and related variants, will be critical to optimize the design of globally effective HIV-1 vaccines in key affected populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors are grateful to the participants, TRUST/RV368 Study Group, MHRP personnel, and MHRP Viral Sequencing Core team for their involvement and assistance in this study. The TRUST/RV368 Study Group includes Principal Investigators: Manhattan Charurat (IHV, University of Maryland, Baltimore, MD, USA), Julie Ake (MHRP, Walter Reed Army Institute of Research, Silver Spring, MD, USA); Co-Investigators: Sylvia Adebajo, Stefan Baral, Erik Billings, Trevor Crowell, George Eluwa, Abiola Fasina, Charlotte Gaydos, Sosthenes Ketende, Afoke Kokogho, Hongjie Liu, Jennifer Malia, Olumide Makanjuola, Nelson Michael, Nicaise Ndembi, Jean Njab, Rebecca Nowak, Oluwasolape Olawore, Zahra Parker, Sheila Peel, Habib Ramadhani, Merlin Robb, Cristina Rodriguez-Hart, Eric Sanders-Buell, Sodsai Tovanabutra, Erik Volz; Institutions: Institute of Human Virology at the University of Maryland School of Medicine (IHV-UMB), University of Maryland School of Public Health (UMD SPH), Johns Hopkins Bloomberg School of Public Health (JHSPH), Johns Hopkins University School of Medicine (JHUSOM), U.S. Military HIV Research Program (MHRP), Walter Reed Army Institute of Research (WRAIR), Henry M. Jackson Foundation for the Advancement of Military Medicine (HJF), Henry M. Jackson Foundation Medical Research International (HJFMRI), Institute of Human Virology Nigeria (IHVN), International Centre for Advocacy for the Right to Health (ICARH), The Initiative for Equal Rights (TIERS), Population Council (Pop Council) Nigeria, Imperial College London.

Funding Statement

This work was supported by a cooperative agreement between the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., and the U.S. Department of Defense [W81XWH-11-2-0174]; the National Institutes of Health [R01 MH099001, R01 AI120913, R01 MH110358]; Fogarty Epidemiology Research Training for Public Health Impact in Nigeria program [D43TW010051]; and the President's Emergency Plan for AIDS Relief through a cooperative agreement between the Department of Health and Human Services/Centers for Disease Control and Prevention, Global AIDS Program, and the Institute for Human Virology-Nigeria [NU2GGH002099].

REFERENCES

1. National Agency for the Control of AIDS. Nigeria Global AIDS Response Country Progress Report. 2015; http://www.unaids.org/sites/default/files/country/documents/NGA_narrative_report_2015.pdf. Accessed April 11, 2018.
2. UNAIDS. Country Factsheets Nigeria. 2016; http://www.unaids.org/en/regionscountries/countries/nigeria. Accessed June 12, 2018.
3. Federal Ministry of Health. National HIV & AIDS and Reproductive Health Survey, 2012 (NARHS Plus). Abuja, Nigeria: Federal Ministry of Health; 2013.
4. Federal Ministry of Health. Nigeria Integrated Biological and Behavioural Surveillance Survey 2014. Abuja, Nigeria: Federal Ministry of Health; 2015.
5. Schwartz SR, Nowak RG, Orazulike I, et al. The immediate eff ect of the Same-Sex Marriage Prohibition Act on stigma, discrimination, and engagement on HIV prevention and treatment services in men who have sex with men in Nigeria: analysis of prospective data from the TRUST cohort. Lancet HIV. 2015;2(7):e299–306. [PubMed: 26125047]
6. Dutta A, Maiga M. An Assessment of Policy toward Most-at-Risk Populations for HIV/AIDS in West Africa. Accra, Ghana: Action for West Africa (AWARE-II) Project;2011.
7. Vu L, Andrinopoulos K, Tun W, Adebajo S. High levels of unprotected anal intercourse and never testing for HIV among men who have sex with men in Nigeria: evidence from a cross-sectional survey for the need for innovative approaches to HIV prevention. Sex Transm Infect. 2013;89(8):659–665. [PubMed: 23851190]
8. Merrigan M, Azeez A, Afolabi B, et al. HIV prevalence and risk behaviours among men having sex with men in Nigeria. Sex Transm Infect. 2011;87(1):65–70. [PubMed: 20820061]
9. Adebajo SB, Eluwa GI, Allman D, et al. Prevalence of Internalized Homophobia and HIV Associated Risks among Men who have Sex with Men in Nigeria. African Journal of Reproductive Health. 2012;16(4):21–28. [PubMed: 23444540]
10. Vu L, Adebajo S, Tun W, et al. High HIV prevalence among men who have sex with men in Nigeria: implications for combination prevention. J Acquir Immune Defic Syndr. 2013;63(2):221–227. [PubMed: 23406978]
11. Keshinro B, Crowell TA, Nowak RG, et al. High prevalence of HIV, chlamydia and gonorrhoea among men who have sex with men and transgender women attending trusted community centres in Abuja and Lagos, Nigeria. J Int AIDS Soc. 2016;19(1):21270. [PubMed: 27931519]
12. Heipertz RA Jr., Ayemoba O, Sanders-Buell E, et al. Significant contribution of subtype G to HIV-1 genetic complexity in Nigeria identified by a newly developed subtyping assay specific for subtype G and CRF02_AG. Medicine (Baltimore). 2016;95(32):e4346. [PubMed: 27512845]
13. Billings E, Sanders-Buell E, Bose M, et al. HIV-1 Genetic Diversity Among Incident Infections in Mbeya, Tanzania. AIDS Res Hum Retroviruses. 2017;33(4):373–381. [PubMed: 27841669]
14. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH, Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000–2007. Aids. 2011;25(5):679–689. [PubMed: 21297424]
15. Joint United Nations Programme on HIV/AIDS (UNAIDS) Gap report: 2014. Geneva, Switzerland 2014.
16. Leoz M, Chaix ML, Delaugerre C, et al. Circulation of multiple patterns of unique recombinant forms B/CRF02_AG in France: precursor signs of the emergence of an upcoming CRF B/02. Aids. 2011;25(11):1371–1377. [PubMed: 21522007]
17. Heipertz RA Jr., Sanders-Buell E, Kijak G, et al. Molecular epidemiology of early and acute HIV type 1 infections in the United States Navy and Marine Corps, 2005–2010. AIDS Res Hum Retroviruses. 2013;29(10):1310–1320. [PubMed: 23972100]
18. Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. Nucleic Acids Res. 2004;32(Web Server issue):W654–659. [PubMed: 15215470]
19. Schultz AK, Zhang M, Bulla I, et al. jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic Acids Res. 2009;37(Web Server issue):W647–651. [PubMed: 19443440]

20. Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol. 1999;73(1):152–160. [PubMed: 9847317]
21. Kijak GH, Tovanabutra S, Beyrer C, et al. RecDraw: a software package for the representation of HIV-1 recombinant structures. AIDS Res Hum Retroviruses. 2010;26(12):1317–1321. [PubMed: 20961275]
22. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29(8):1969–1973. [PubMed: 22367748]
23. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 2018.
24. Gaschen B, Kuiken C, Korber B, Foley B. Retrieval and on-the-fly alignment of sequence fragments from the HIV database. Bioinformatics. 2001;17(5):415–418. [PubMed: 11331235]
25. Eddy SR. Profile hidden Markov models. Bioinformatics. 1998;14(9):755–763. [PubMed: 9918945]
26. Deng W, Maust BS, Nickle DC, et al. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. BioTechniques. 2010;48(5):405–408. [PubMed: 20569214]
27. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10(3):512–526. [PubMed: 8336541]
28. Bennett DE, Camacho RJ, Otelea D, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724. [PubMed: 19266092]
29. Crowell TA, Keshinro B, Baral SD, et al. Stigma, access to healthcare, and HIV risks among men who sell sex to men in Nigeria. J Int AIDS Soc. 2017;20(1):21489. [PubMed: 28453241]
30. Rodriguez-Hart C, Musci R, Nowak RG, et al. Sexual Stigma Patterns Among Nigerian Men Who Have Sex with Men and Their Link to HIV and Sexually Transmitted Infection Prevalence. AIDS Behav. 2018;22(5):1662–1670. [PubMed: 29168069]
31. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science. 1992;256(5065):1783–1790. [PubMed: 1377403]
32. Davies JF 2nd, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. Science. 1991;252(5002):88–95. [PubMed: 1707186]
33. Chen JC, Krucinski J, Miercke LJ, et al. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proc Natl Acad Sci U S A. 2000;97(15):8233–8238. [PubMed: 10890912]
34. Leoz M, Feyertag F, Charpentier C, et al. Characterization of CRF56_cpx, a new circulating B/CRF02/G recombinant form identified in MSM in France. Aids. 2013;27(14):2309–2312. [PubMed: 24157908]
35. Charurat M, Nasidi A, Delaney K, et al. Characterization of acute HIV-1 infection in high-risk Nigerian populations. The Journal of infectious diseases. 2012;205(8):1239–1247. [PubMed: 22357663]
36. HVTN 705: A Study to Assess the Efficacy of a Heterologous Prime/Boost Vaccine Regimen of Ad26.Mos4.HIV and Aluminum Phosphate-Adjuvanted Clade C gp140 in Preventing Human Immunodeficiency Virus (HIV) −1 Infection in Women in Sub-Saharan Africa. https://clinicaltrials.gov/ct2/show/NCT03060629. Accessed September 7, 2018.
37. HVTN 702: Pivotal Phase 2b/3 ALVAC/Bivalent gp120/MF59 HIV Vaccine Prevention Safety and Efficacy Study in South Africa (HVTN702). https://clinicaltrials.gov/ct2/show/NCT02968849. Accessed September 7, 2018.

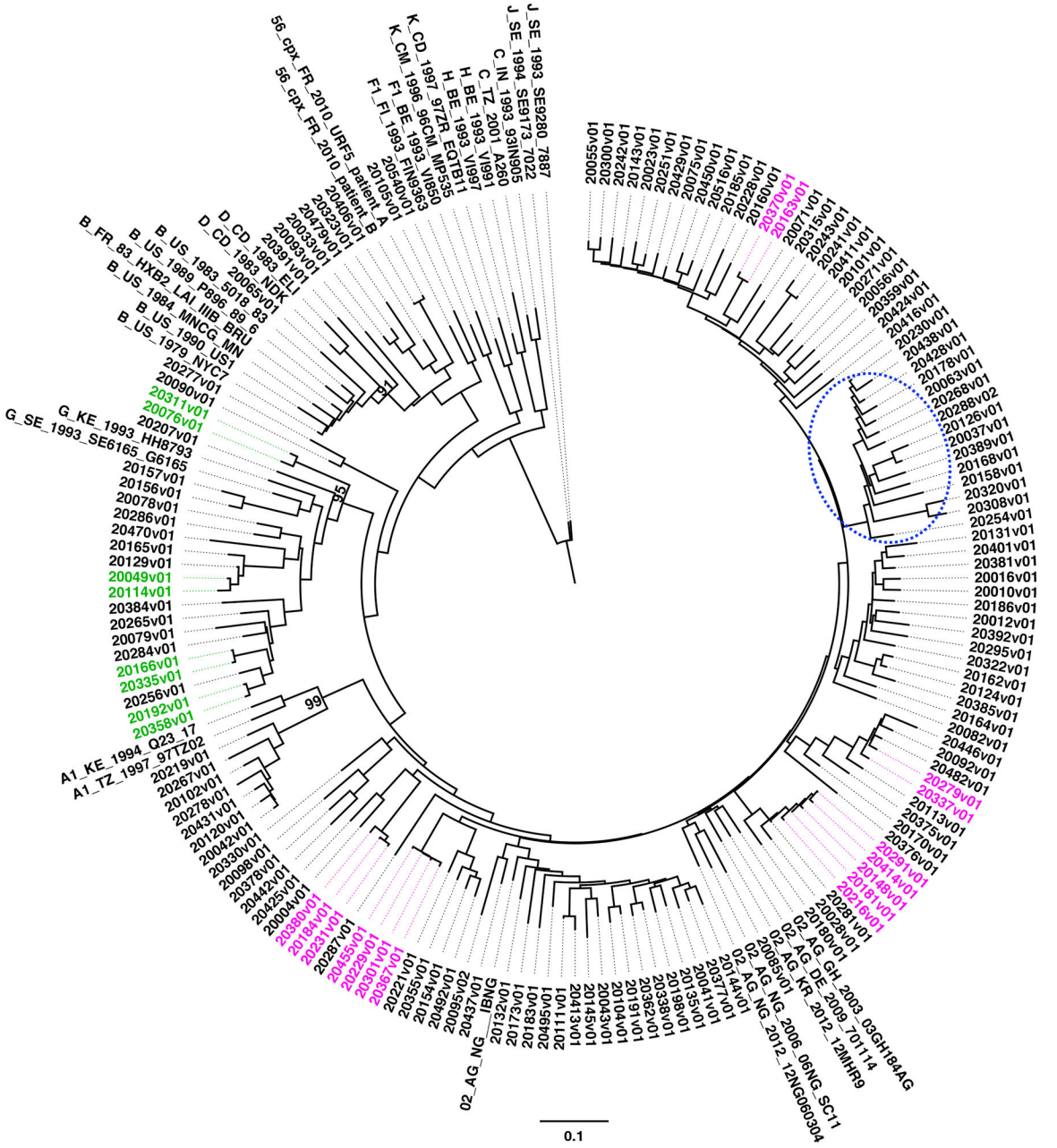

Figure 1.

Phylogenetic relationships and transmission network clusters among participant *pol* sequences. Maximum likelihood (ML) tree constructed using subtype reference strains and *pol* sequences (HXB2: 2273–3869) from prevalent infections observed at study enrollment. Colored sample numbers denote CRF02_AG (pink) or subtype G (green) sequences belonging to proposed transmission networks (defined as having a pairwise genetic distance ≤1.5%). The cluster circled in blue represents the CRF02_AG/B recombinants in which the CRF95_02B strains were observed. ML bootstrap values at relevant nodes are shown. The scale bar indicates a genetic distance of 10%.

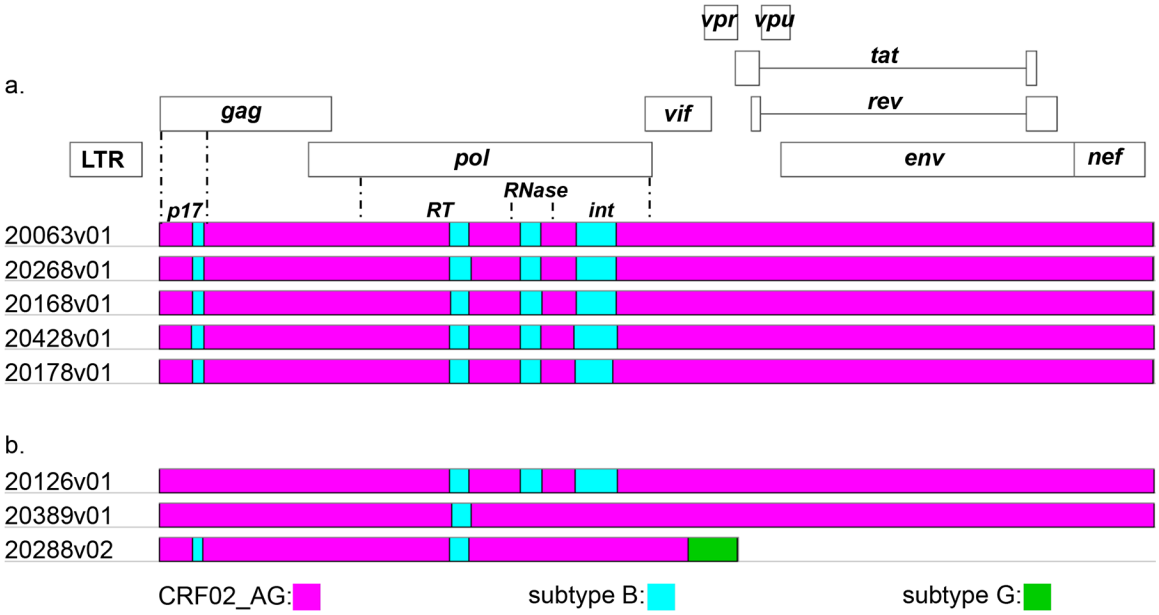

Figure 2.

Recombinant genome structures. Genome structures of the eight strains (seven full-length and one half-length) observed while characterizing CRF95_02B. (a) The HXB2 referenced positions of the CRF95 subtype B fragments are shown relative to genomic markers in *gag* and *pol* (dash-dot lines). (b) The similarly structured genomes of the related non-CRF95 strains. The colors of the fragments correspond to their predicted parent subtypes: CRF02_AG (pink), subtype B (blue), subtype G (green).