



HHS Public Access

Author manuscript

AIDS. Author manuscript; available in PMC 2021 February 01.

Published in final edited form as:

AIDS. 2020 February 01; 34(2): 227–236. doi:10.1097/QAD.0000000000002409.

Genetic Clustering Analysis for HIV Infection among Men Who Have Sex with Men in Nigeria: Implications for Intervention

Yuruo LI¹, Hongjie LIU¹, Habib O. RAMADHANI², Nicaise NDEMBI^{2,3}, Trevor A. CROWELL^{4,5}, Gustavo KIJAK^{4,5,a}, Merlin L. ROBB^{4,5}, Julie A. AKE⁴, Afoke KOKOGHO^{9,10}, Rebecca G. NOWAK², Charlotte GAYDOS⁶, Stefan D. BARAL⁷, Erik VOLZ⁸, Sodsai TOVANABUTRA^{4,5}, Man CHARURAT², TRUST/RV368 Study Group

¹Department of Epidemiology, University of Maryland School of Public Health, College Park, MD, USA

²Institute of Human Virology, University of Maryland School of Medicine, Baltimore, MD, USA

³Institute of Human Virology Nigeria, Abuja, Nigeria

⁴U.S. Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD, USA

⁵Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA

⁶Johns Hopkins University School of Medicine, Baltimore, MD, USA

⁷Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

⁸Department of Infectious Disease Epidemiology, Imperial College, London, UK

⁹Henry M. Jackson Foundation Medical Research International, Abuja, Nigeria

¹⁰U.S. Army Medical Research Directorate-Africa/Nigeria, Abuja, Nigeria

Abstract

Background: The HIV epidemic continues to grow among men who have sex with men (MSM) in countries across sub-Saharan Africa including Nigeria. To inform prevention efforts, we used a phylogenetic cluster method to characterize HIV genetic clusters and factors associated with cluster formation among MSM living with HIV in Nigeria.

Corresponding author contact information: Yuruo Li, yuruo.li12@gmail.com, Phone number: 404-539-6249, Mailing Address: 4200 Valley Dr, College Park, MD 20742.

^acurrent affiliation: GSK Vaccines, Rockville, MD, USA

Accession numbers:

The HIV-1 sequences described in this study are available under GenBank accession numbers: [MK784335](#) to [MK787541](#) and [MH654824–MH654973](#).

Disclaimer

The views expressed are those of the authors and should not be construed to represent the positions of the U.S. Army, the Department of Defense, or the Department of Health and Human Services. The investigators have adhered to the policies for protection of human subjects as prescribed in AR-70. Gustavo Kijak is currently employed by GSK Vaccines but his participation in this study precedes the current employment.

This manuscript was presented at the poster section in International AIDS Conference(IAS) 2018.

Methods: We analyzed HIV-1 *pol* sequences from 417 MSM living with HIV enrolled in the TRUST/RV368 cohort between 2013 and 2017 in Abuja and Lagos, Nigeria. A genetically-linked cluster was defined among participants whose sequences had pairwise genetic distance $\leq 1.5\%$. Binary and multinomial logistic regressions were used to estimate adjusted odds ratios (AORs) and 95% confidence intervals (CIs) for factors associated with HIV genetic cluster membership and size.

Results: Among 417 MSM living with HIV, 153 (36.7%) were genetically linked. Participants with higher viral load (AOR=1.72 95% CI: 1.04–2.86), no female partners (AOR=3.66; 95% CI: 1.97–6.08), and self-identified as male gender (compared with self-identified as bigender) (AOR=3.42; 95% CI: 1.08–10.78) had higher odds of being in a genetic cluster. Compared to unlinked participants, MSM who had high school education (AOR=23.84; 95% CI: 2.66–213.49), were employed (AOR=3.41; 95% CI: 1.89–10.70), had bacterial STIs (AOR=3.98; 95% CI: 0.89–17.22) and were not taking antiretroviral therapy (AOR=6.61; 95% CI: 2.25–19.37) had higher odds of being in a large cluster (size>4).

Conclusions: Comprehensive HIV prevention packages should include behavioral and biological components, including early diagnosis and treatment of both HIV and bacterial STIs to optimally reduce the risk of HIV transmission and acquisition.

Keywords

phylogenetic cluster; HIV; MSM; Nigeria; Intervention

Introduction

Genetic clustering methods have been widely used to identify characteristics of specific populations who are at heightened risk of transmitting HIV to uninfected individuals [1–11]. Examination of similarities between genetically-linked sequences [12–15] has shown that highly similar virus strains can be inferred as being connected by a short chain of transmission [16] and these clustered sequences may be linked as transmission pairs or belong to a larger local transmission network [7]. However, factors such as the recency of infection, sampling fractions (i.e. the intensity of sampling of local transmissions), survivor bias (i.e. younger individuals are more likely to cluster), and migration status (i.e. immigrants are more likely to be infected while abroad) may also contribute to any observed phylogenetic linkages [7]. Phylogenetic studies conducted among key populations [3, 11, 13, 17–23] in high-income countries have combined epidemiologic data and genetic clustering patterns to examine factors associated with cluster size and membership, including demographic characteristics, geographic variables, and risk behaviors. For example, transmission related to same-sex sexual practices among men [2, 14, 17, 21, 23], younger age [13, 17, 21], being an immigrant [13], urban residence [17] and substance use [23] were found to be positively associated with cluster membership.

These methods are starting to be applied in countries across sub-Saharan Africa to give additional insights into the dynamics of local HIV epidemics and to mobilize resources towards achieving epidemic control. For example, De Oliveira et al. found through clustered sequences that men aged 25–40 years old were the primary source of HIV transmissions to

adolescent girls in KwaZulu-Natal, South Africa [6]. This finding informed the Determined, Resilient, Empowered, AIDS-free, Mentored and Safe (DREAMS) initiative, a global partnership to reduce HIV infections among adolescent girls and young women in 10 sub-Saharan African countries[24]. In a mixed HIV epidemic like Nigeria, with multiple risk groups requiring intervention to halt onward transmission, there is a need to document the characteristics of genetic clusters and examine factors that determine the formation of clusters to inform the development of interventions that decrease HIV transmission in high-risk populations. Some of the highest HIV incidence rates among key populations in Nigeria are seen in men who have sex with men (MSM) [25]. Using a dynamic infectious disease model fitted to time-scaled phylogenies, we previously estimated transmission patterns of MSM and a representative sample of newly enrolled treatment-naïve people living with HIV [26].

The purpose of this study was to describe HIV genetic clusters among HIV-infected MSM and to examine associations between cluster membership, risk behaviors, and preventive practices in two Nigerian communities with a high burden of HIV.

Methods

Study Protocol

TRUST/RV368 is a multicenter prospective observational cohort that recruited MSM in Abuja and Lagos, Nigeria, using respondent-driven sampling (RDS) between 2013 and 2017 as previously described [27, 28]. Eligible participants included those who were assigned male sex at birth, aged 16 and older in Abuja or 18 and older in Lagos, and reported receptive or insertive anal intercourse in the previous 12 months. Demographic and behavioral data were obtained from structured interviews. Upon enrollment, participants were screened for HIV using fingerstick blood samples in a parallel testing algorithm with Determine (Alere, Waltham, MA, USA) and Uni-gold (Trinity Biotech, Co-Wicklow, Ireland) rapid tests with HIV-1/2 Stat-Pak (Chembio Diagnostics, Medford, NY) as a tie-breaker for discrepant results.

All participant plasma samples with HIV RNA $\geq 1,000$ copies/mL at enrollment were genetically sequenced for clustering analysis. Samples with HIV RNA below this cut-off were excluded because of poor amplification. Participants who were HIV uninfected at enrollment were followed every three months for up to 18 months and incident infections were sequenced at the time of seroconversion.

Ethical Considerations

The institutional review boards at the Nigerian Federal Capital Territory Health Research Ethics Committee, the Nigerian Ministry of Defense, the University of Maryland Baltimore, and the Walter Reed Army Institute of Research reviewed and approved the research protocol. All participants provided written informed consent and data were de-identified prior to analysis.

Genetic Sequencing and Pairwise Calculation

Sequences of the HIV-1 *pol* gene (corresponding to HXB2 positions 2273–3869 or 2108–3308) were generated using previously-described methods [18, 26, 29]. Pairwise genetic distances, corresponding to HXB2 positions 2317–3249, were calculated using the TN93 model[30]. Genetic clusters were defined among individuals whose sequences had a genetic distance of $\leq 1.5\%$ [15, 31]. The cut-off at 1.5% was used for consistency with prior studies of MSM populations and to conservatively estimate any direct linkages between transmission pairs as shown by Kroon et al [32].

The HIV-1 Genotyping Tool at the National Center for Biotechnology Information, Jumping Profile HMM Tool at GLOBICS, REGA HIV Subtyping Tool at BIOAFRICA, NCBI BLASTn tool, and the HIV BLAST tool at the Los Alamos HIV sequence database website (https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST.html) were used to determine the subtype(s) of each sequence. If the subtype(s) of a sequence from all tools were in agreement, a final subtype result was assigned. If the results were different, neighbor joining trees of the sequences along with relevant HIV-1 reference subtypes or CRFs were made at various breakpoints and over the span of the whole sequence to determine the genetic relatedness of the sample to reference sequences[33].

Sequence Quality Control

Obtained sequencing electropherograms were visually inspected using Sequencher 5.4 (Gene Codes Corp., Ann Arbor, Michigan, USA) at two independent laboratories (the Institute of Human Virology Nigeria, Abuja, Nigeria and the U.S. Military HIV Research Program in Bethesda, Maryland, USA) to verify that each nucleotide base was covered by at least three reads, one of which had to be in the opposite direction as the other two. Sequences were first aligned using HIV Align (<http://www.hiv.lanl.gov/content/sequence/VIRALIGN/viralalign.html>) and the alignments were manually edited using Geneious (<http://www.geneious.com>). Sequence genetic relatedness was assessed in MEGA version 5.2.2. Samples with sequences that were $< 1.0\%$ different and had been processed on the same day were re-processed and re-sequenced to rule out cross-sample contamination.

Dependent Variables:

Cluster Membership: if a participant had a genetic distance less than 1.5% from another participant, both participants were classified as part of a genetic cluster. For analysis, we evaluated cluster membership in two ways: as a dichotomized outcome based on whether or not a given participant was a member of any genetic cluster and as a categorical outcome based on whether a given participant was not in any genetic cluster, in a small genetic cluster (cluster size of 2–3 participants), or in a genetic cluster of 4 or more participants.

Independent Variables:

Antiretroviral Therapy (ART) Use: A comprehensive chart review was performed to determine ART exposure among all participants with sequencing data. A participant was considered ART experienced if he initiated three-drug ART more than 28 days before the sample for sequencing was drawn.

Consistent Condom Use: Participants self-reported their frequency of condom use during vaginal sex, receptive anal sex and insertive anal sex in the previous 12 months. The frequencies were measured on a 5-point Likert scale: never, almost never, about half of the time, almost always, and always. Participants were classified as using condoms consistently if they reported almost always or always using condoms for all three types of sexual intercourse.

Bacterial Sexually Transmitted Infection (STI) Status: Participants provided blood, urine, oropharyngeal swabs and rectal swabs for bacterial STI testing. If they were diagnosed with syphilis, gonorrhea, chlamydia, or presented with symptoms consistent with an otherwise undiagnosed syndrome consistent with a bacterial STI, they were classified as having a bacterial STI.

Number of Male Partners: Participants reported the average number of male partners for anal sex in the past 12 months. We dichotomized responses as either less than or equal to 1 male partner or more than 1 male partner to minimize measurement error.

Sexual Positions: The sexual positions of the participants were inferred from the self-reported number of partners for insertive or receptive anal sex in the past 12 months. Participants were categorized as insertive or receptive if they only reported insertive or receptive sex, respectively. If participants reported both, they were classified as versatile.

All the behavioral questionnaire and clinical evaluations were completed at the enrollment or seroconversion visit concurrent with the collection of blood for sequencing.

Statistical Analyses

We used multivariable logistic regression to evaluate characteristics associated with genetic clustering and multinomial logistic regression to assess factors associated with cluster size. The primary dependent variables were a binary category of any clustering and a multinomial category for none, small (size 2–3) and large (size 4 and above) genetic clusters. The independent demographic variables were age, education, gender identity, employment status, religion, and sexual orientation. We also assessed the associations between cluster membership and ART status, consistent condom use, STI status, sexual position and number of male sexual partners. Viral load and CD4 count at the time of sequencing and the stage of HIV infection (prevalent or incident) were included as potential confounding variables[7]. Because viral load was positively skewed, \log_{10} transformations were used in the final models. Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) were reported. The regression analyses were adjusted for non-random sampling method by including RDS weights. Sensitivity analyses were conducted to evaluate the odds of clustering at various genetic distance threshold: 1%, 2%, and 4.5%.

A correlation permutation test was also performed to evaluate the likelihood of observing any clustering for certain behaviors. We calculated the Pearson's correlation of ART use, STI status and the number of male partners between every pair of genetically linked participants and compared these with the null distribution of correlation obtained from randomly permuting observed data. The achieved significance level (ASL) and null

distribution were reported. Although the TRUST/RV368 cohort is a longitudinal study, these analyses were conducted cross-sectionally because all independent variables were collected at the same time as sequencing. R package, igraph, and SAS 9.4 were used in data management and analysis.

Results

Descriptive analysis

A total of 417 HIV viremic MSM were eligible for virus sequencing, including 153 (36.7%) who were members of genetic clusters. The demographic characteristics of the sample stratified by cluster status are presented in Table 1. More than half of participants reported inconsistent condom use in the past 12 months and having a bacterial STI regardless of their clustering status. MSM who were not members of an HIV genetic cluster were more likely to have female sexual partners in the past 12 months (50.4%) than MSM in an HIV cluster (37.3%). ART use was low, only 6.5% among those with clustering and 12.1% among those without clustering.

Sequences were obtained from 376 (90.2%) prevalent and 41 (9.8%) incident HIV infections. The average genetic distance between sequences was 8.3% with a range of 0.00%–15.37%. A total of 207 genetically-linked ties presented in 46 genetically-linked clusters (Figure 1). Among the genetically-linked clusters, 30 (65.2%) were comprised of pairs and 10 (21.7%) included 3–5 participants. The most common HIV subtype was CRF02_AG and its associated recombinants, which was found in 319 (76.5%) sequenced participants. The largest HIV genetic cluster included 15 participants whose HIV was a mix of subtype A1 and A1/U recombinant.

As depicted in Figure 1, 8 (50%) of the incident cases among clustered individuals were concentrated in the two largest clusters and not evenly distributed throughout the clusters. Participants within the same cluster were more likely to have the same STI and report similar frequencies of condom use, although condom use was low among all participants.

Model Results:

Tables 2 and 3 show the results of logistic regression models that evaluated factors associated with membership in an HIV genetic cluster. Compared to participants who self-identified as bigender, male participants had 2.42 times higher odds of being in a cluster (95% CI: 1.08–10.78). Compared to MSM who reported sex with female partners in the past 12 months, MSM with no female partners had 2.66 times higher odds of being in a cluster (95% CI: 1.97–6.80). Viral load was also significantly higher (AOR=1.72; 95% CI: 1.04–2.86) in clustered participants.

In multinomial logistic regressions, participants with more than a high school education had higher odds of being in a large cluster (size 4 and above vs. un-clustered: AOR = 23.84; 95% CI: 2.66–213.49), but lower odds of being in a small cluster (size 2–3 vs. un-clustered: AOR=0.34; 95% CI: 0.08–1.51) as compared to not being in an HIV genetic cluster. Participants who self-identified as male had higher odds of being in a small cluster (size 2–3 vs. un-clustered: AOR = 17.28; 95% CI: 2.93–102.12) than participants who self-identified

as bigender. Compared to MSM who reported sex with female partners in the past 12 months, MSM with no female partner had 3.67 (95% CI: 2.05–10.64) times higher odds of being in a large cluster (size 4 and above) and 2.17(95% CI: 1.89–5.32) times higher odds of being in a small cluster (size 2–3), as opposed to un-clustered MSM. As shown in Table 3, ART-inexperienced participants had higher odds of being in a large cluster (4 cluster size vs no cluster: AOR = 6.61; 95% CI: 2.25–19.37), but there was no statistically significant difference between the probability of clustering between participants who were not in small clusters(size 2–3) and un-clustering participants. Compared to participants that were not members of an HIV genetic cluster, those within a large cluster (4) had slightly higher odds of having bacterial STIs (AOR = 3.98; 95% CI: 0.89–17.22), and those within a small cluster (2–3 cluster size) had higher odds of self-reporting receptive sex (AOR=3.48; 95% CI: 2.03–5.92) and versatile sexual position(AOR=2.87; 95% CI: 2.41–3.41), as opposed to insertive sexual position.

Inferences drawn about risk behaviors associated with genetic clustering were relatively robust across sensitivity analyses using different genetic cut-offs to determine clusters (Supplemental Tables 1 and 2). Although not all results were statistically significant, the direction of most of the results was consistent. Younger age (20–25 vs. 25 and above), no female partners, high viral load, ART inexperience, and having bacterial STIs were associated with increased odds of cluster membership across different cut-offs. Male gender identity was positively associated with clustering when lower cut-offs were used (1%, 2%) but the association was reversed with a higher cutoff of 4.5%. Higher CD4 count was associated with higher odds of clustering probability at the 4.5% cutoff but not significant for the lower cutoffs.

Permutation tests found no association between HIV genetic links and ART use ($\rho=0.12$, ASL=0.20), STI status ($\rho=0.09$, ASL = 0.31), number of male partners ($\rho=0.02$, ASL = 0.78) or consistent condom use ($\rho=-0.04$, ASL = 0.65).

Discussion

In the setting of a mixed epidemic, we found that one-third of Nigerian MSM living with HIV in the TRUST/RV368 cohort were genetically clustered. MSM who were ART-inexperienced and only engaged in same-sex sexual practices were more likely to belong to any genetic cluster and particularly a large cluster (4). MSM who were employed, had a higher education level, and had bacterial STIs were also more likely to be in a large genetic cluster (4). Although clustered individuals were more likely to engage in high-risk behaviors, permutation tests suggest that this pattern was not significant at the dyadic level. The HIV risk behaviors of one linked individual were not directly correlated with the risk behaviors of his linked partner.

Previous studies on genetic clustering in countries across sub-Saharan Africa found smaller phylogenetic clusters (size <5) than observed in this study [1, 8, 9]. Our largest reported cluster was composed of 15 individuals and 5 out of 46 clusters were composed of more than 6 individuals. If we increased the genetic distance threshold for defining a cluster to 4.5%, as used in de Oliverira et al.[6], the largest cluster would be composed of 255

individuals. Fifty percent of the MSM in this study reported having at least one female partner, consistent with our prior work showing that a substantial portion of female HIV infections was attributable to transmissions from MSM reporting unmet HIV prevention needs [26]. However, since we only included male participants in our study, further research is needed to verify any linkages between Nigerian MSM and heterosexual female populations. Recruitment into the cohort used RDS, which tends to recruit participants who are alike and connected, thereby potentially contributing to the clustering pattern observed. However, among 207 genetically-linked pairs, only 7 were directly linked by RDS. The geodesic distance from RDS recruitment chains was not correlated with genetic distance (Pearson's correlation: $\rho = 0.05$). Previous studies in El Salvador[34] and Croatia[20] also found no correlation between RDS recruitment links and phylogenetic distances. Since the RDS recruitment strategy did not restrict social contacts between participants and limited the number of recruitments, the degree to which the transmission network overlapped with the RDS network is unclear.

Recent infections were more likely to cluster than prevalent infections with 50% of the incident cases concentrating in the two largest genetically-linked clusters among clustering participants. The clustering of incident cases may be due to an increased risk of transmission among larger clusters. Another possible explanation is that new infections are more contagious and more likely to produce other incident cases because of their higher plasma HIV levels [35, 36]. The high transmission rate during acute HIV infection, combined with our finding that participants in large clusters were less likely to be on ART, underscore the urgency of ART initiation as soon as possible after seroconversion. HIV prevention interventions that focus on sexual networks of recently diagnosed individuals and network-based HIV screening among key populations may facilitate early diagnosis and treatment.

Our data also highlight the burden of sexually transmitted co-infections among HIV genetic clusters. For the majority of participants, STIs were presumably acquired after the formation of phylogenetic linkages, suggesting that high-risk behaviors persisted after infection with HIV. Since bacterial STIs could facilitate HIV transmission, comprehensive HIV treatment packages should include condom education, STI screening and treatment, as well as ART to optimally curb further transmission for both HIV and STIs [28, 37].

There were several strengths in our study, including an RDS recruitment strategy that allowed us to sample a highly marginalized population of MSM in Nigeria, a prospective design that allowed us to characterize newly-diagnosed incident infections, and detailed behavioral data that allowed us to inform targeted intervention strategies. However, this study had some limitations. First, we could only sequence HIV from individuals who had a viral load greater than 1000 copies/ml, which may have introduced bias to our analyses. Among 968 HIV-infected participants in this cohort, only 43% had a high enough viral load to be sequenced and individuals on suppressive ART could not be included in these analyses. Beyond the TRUST cohort, this sequenced sample only represented approximately less than 10% of the local MSM population (The local population size in Abuja was estimated through RDS scheme and self-reported network size by RDS Analyst[38–40]. Mean: 2470; 95% CI: 2330–2920). The small sampling fraction limited our ability to infer the underlying transmission network. Second, we could only assess the association between HIV risk

behaviors and cluster membership at the aggregated level, which did not account for variations of each cluster. Third, like all genetic clustering studies, we were unable to determine whether the results of genetic linkage were due to direct or indirect transmissions.

Conclusion

We documented HIV genetic clusters among MSM in Nigeria, where a mixed HIV epidemic is disproportionately impacting key populations such as MSM. We found that MSM who were in HIV genetic clusters were less likely to report ART use, had high viral loads often related to recent HIV acquisition, and did not consistently use condoms, all of which are important targets for combination interventions to prevent onward HIV transmission. The high prevalence of bacterial STIs among MSM in HIV genetic clusters underscores the need for behavioral interventions that promote safer sex with early diagnosis and treatment of bacterial STIs. Ultimately, without evidence-based and human rights affirming implementation strategies to prevent the acquisition and transmission of HIV for all of those at risk, an HIV-free future remains improbable in Nigeria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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, Oluvide Makanjuola, Nelson Michael, Nicaise Ndembu, Jean Njab, Rebecca Nowak, Oluwasolape Olowore, 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 (HJM), 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, and the Institute for Human Virology-Nigeria [NU2GGH002099].

References

1. Bezemer D, Faria NR, Hassan A, Hamers RL, Mutua G, Anzala O, et al. HIV Type 1 transmission networks among men having sex with men and heterosexuals in Kenya. *AIDS Res Hum Retroviruses* 2014; 30(2):118–126. [PubMed: 23947948]
2. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, 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; 10:262. [PubMed: 20822507]

3. Chan PA, Hogan JW, Huang A, DeLong A, Salemi M, Mayer KH, et al. Phylogenetic Investigation of a Statewide HIV-1 Epidemic Reveals Ongoing and Active Transmission Networks Among Men Who Have Sex With Men. *J Acquir Immune Defic Syndr* 2015; 70(4):428–435. [PubMed: 26258569]
4. Chan PA, Reitsma MB, DeLong A, Boucek B, Nunn A, Salemi M, et al. Phylogenetic and geospatial evaluation of HIV-1 subtype diversity at the largest HIV center in Rhode Island. *Infect Genet Evol* 2014; 28:358–366. [PubMed: 24721515]
5. Chehadeh W, Albaksami O, Altawalah H, Ahmad S, Madi N, John SE, et al. Phylogenetic analysis of HIV-1 subtypes and drug resistance profile among treatment-naive people in Kuwait. *J Med Virol* 2015; 87(9):1521–1526. [PubMed: 25976289]
6. de Oliveira T, Kharsany AB, Graf T, Cawood C, Khanyile D, Grobler A, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV* 2017; 4(1):e41–e50. [PubMed: 27914874]
7. Frost SD, Pillay D. Understanding drivers of phylogenetic clustering in molecular epidemiological studies of HIV. *J Infect Dis* 2015; 211(6):856–858. [PubMed: 25312038]
8. Grabowski MK, Herbeck JT, Poon AFY. Genetic Cluster Analysis for HIV Prevention. *Curr HIV/AIDS Rep* 2018; 15(2):182–189. [PubMed: 29460226]
9. Grabowski MK, Lessler J, Redd AD, Kagaayi J, Laeyendecker O, Ndyanabo A, et al. The role of viral introductions in sustaining community-based HIV epidemics in rural Uganda: evidence from spatial clustering, phylogenetics, and egocentric transmission models. *PLoS Med* 2014; 11(3):e1001610. [PubMed: 24595023]
10. Kiwuwa-Muyingo S, Nazziwa J, Ssemwanga D, Ilmonen P, Njai H, Ndemi N, et al. HIV-1 transmission networks in high risk fishing communities on the shores of Lake Victoria in Uganda: A phylogenetic and epidemiological approach. *PLoS One* 2017; 12(10):e0185818. [PubMed: 29023474]
11. Wertheim JO, Leigh Brown AJ, Hepler NL, Mehta SR, Richman DD, Smith DM, et al. The global transmission network of HIV-1. *J Infect Dis* 2014; 209(2):304–313. [PubMed: 24151309]
12. Little SJ, Kosakovsky Pond SL, Anderson CM, Young JA, Wertheim JO, Mehta SR, et al. Using HIV networks to inform real time prevention interventions. *PLoS One* 2014; 9(6):e98443. [PubMed: 24901437]
13. Dennis AM, Hue S, Pasquale D, Napravnik S, Sebastian J, Miller WC, et al. HIV Transmission Patterns Among Immigrant Latinos Illuminated by the Integration of Phylogenetic and Migration Data. *AIDS Res Hum Retroviruses* 2015; 31(10):973–980. [PubMed: 26214548]
14. Fearnhill E, Gourlay A, Malyuta R, Simmons R, Ferns RB, Grant P, et al. A Phylogenetic Analysis of Human Immunodeficiency Virus Type 1 Sequences in Kiev: Findings Among Key Populations. *Clin Infect Dis* 2017; 65(7):1127–1135. [PubMed: 28575385]
15. Poon AF. Impacts and shortcomings of genetic clustering methods for infectious disease outbreaks. *Virus Evol* 2016; 2(2):vew031.
16. Volz EM, Koopman JS, Ward MJ, Brown AL, Frost SD. Simple epidemiological dynamics explain phylogenetic clustering of HIV from patients with recent infection. *PLoS Comput Biol* 2012; 8(6):e1002552. [PubMed: 22761556]
17. Dennis AM, Volz E, Frost A, Hossain M, Poon AFY, Rebeiro PF, et al. HIV-1 Transmission Clustering and Phylodynamics Highlight the Important Role of Young Men Who Have Sex with Men. *AIDS Res Hum Retroviruses* 2018; 34(10):879–888. [PubMed: 30027754]
18. Heipertz RA Jr, Sanders-Buell E, Kijak G, Howell S, Lazzaro M, Jagodzinski LL, et al. Molecular epidemiology of early and acute HIV type 1 infections in the United States Navy and Marine Corps, 2005–2010. *AIDS research and human retroviruses* 2013; 29(10):1310–1320. [PubMed: 23972100]
19. Kusejko K, Kadelka C, Marzel A, Battgay M, Bernasconi E, Calmy A, et al. Inferring the age difference in HIV transmission pairs by applying phylogenetic methods on the HIV transmission network of the Swiss HIV Cohort Study. *Virus Evol* 2018; 4(2):vey024.
20. Lepej SZ, Vrakela IB, Poljak M, Bozicevic I, Begovac J. Phylogenetic analysis of HIV sequences obtained in a respondent-driven sampling study of men who have sex with men. *AIDS Res Hum Retroviruses* 2009; 25(12):1335–1338. [PubMed: 20001315]

21. Lubelchek RJ, Hoehnen SC, Hotton AL, Kincaid SL, Barker DE, French AL. Transmission clustering among newly diagnosed HIV patients in Chicago, 2008 to 2011: using phylogenetics to expand knowledge of regional HIV transmission patterns. *J Acquir Immune Defic Syndr* 2015; 68(1):46–54.
22. Mehta SR, Murrell B, Anderson CM, Kosakovsky Pond SL, Wertheim JO, Young JA, et al. Using HIV Sequence and Epidemiologic Data to Assess the Effect of Self-referral Testing for Acute HIV Infection on Incident Diagnoses in San Diego, California. *Clin Infect Dis* 2016; 63(1):101–107. [PubMed: 27174704]
23. Morgan E, Nyaku AN, D'Aquila RT, Schneider JA. Determinants of HIV Phylogenetic Clustering in Chicago Among Young Black Men Who Have Sex With Men From the uConnect Cohort. *J Acquir Immune Defic Syndr* 2017; 75(3):265–270. [PubMed: 28328553]
24. Fleischman J, Peck K. Addressing HIV risk in adolescent girls and young women. CSIS Global Health Policy Center 2015.
25. Nowak RG, Mitchell A, Crowell TA, Liu H, Ketende S, Ramadhani HO, et al. Individual and sexual network predictors of HIV incidence among men who have sex with men in Nigeria. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2019.
26. Volz EM, Ndemi N, Nowak R, Kijak GH, Idoko J, Dakum P, et al. Phylodynamic analysis to inform prevention efforts in mixed HIV epidemics. *Virus Evol* 2017; 3(2):vex014.
27. Keshinro B, Crowell TA, Nowak RG, Adebajo S, Peel S, Gaydos CA, 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]
28. Ramadhani HO, Liu H, Nowak RG, Crowell TA, Ndomb T, Gaydos C, et al. Sexual partner characteristics and incident rectal *Neisseria gonorrhoeae* and *Chlamydia trachomatis* infections among gay men and other men who have sex with men (MSM): a prospective cohort in Abuja and Lagos, Nigeria. *Sex Transm Infect* 2017; 93(5):348–355. [PubMed: 28235839]
29. Etiebet MA, Shepherd J, Nowak RG, Charurat M, Chang H, Ajayi S, et al. Tenofovir-based regimens associated with less drug resistance in HIV-1-infected Nigerians failing first-line antiretroviral therapy. *AIDS* 2013; 27(4):553–561. [PubMed: 23079810]
30. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution* 1993; 10(3):512–526. [PubMed: 8336541]
31. Hué S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *Aids* 2004; 18(5):719–728. [PubMed: 15075506]
32. Kroon E, Pham PT, Sirivichayakul S, Trichavaroj R, Colby DJ, Pinyakorn S, et al. Transmission dynamics among participants initiating antiretroviral therapy upon diagnosis of early acute HIV-1 infection in Thailand. *AIDS* 2018; 32(16):2373–2381. [PubMed: 30096068]
33. Delatorre E, Mir D, Bello G. Spatiotemporal dynamics of the HIV-1 subtype G epidemic in West and Central Africa. *PLoS One* 2014; 9(2):e98908. [PubMed: 24918930]
34. Dennis AM, Murillo W, de Maria Hernandez F, Guardado ME, Nieto AI, Lorenzana de Rivera I, et al. Social network-based recruitment successfully reveals HIV-1 transmission networks among high-risk individuals in El Salvador. *J Acquir Immune Defic Syndr* 2013; 63(1):135–141. [PubMed: 23364512]
35. Cohen MS. Amplified transmission of HIV-1: missing link in the HIV pandemic. *Transactions of the American Clinical and Climatological Association* 2006; 117:213–224; discussion 225. [PubMed: 18528475]
36. Mayer KH, Skeer MR, O'Cleirigh C, Goshe BM, Safren SA. Factors associated with amplified HIV transmission behavior among American men who have sex with men engaged in care: implications for clinical providers. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine* 2014; 47(2):165–171. [PubMed: 23873338]
37. Fox J, White PJ, Macdonald N, Weber J, McClure M, Fidler S, et al. Reductions in HIV transmission risk behaviour following diagnosis of primary HIV infection: a cohort of high-risk men who have sex with men. *HIV Med* 2009; 10(7):432–438. [PubMed: 19459996]

38. Handcock MSF, Ian E.; Gile Krista J. RDS Analyst: Software for the Analysis of Respondent-Driven Sampling Data. In. 0.42 ed; 2014.
39. Sulaberidze L, Mirzazadeh A, Chikovani I, Shengelia N, Tsereteli N, Gotsadze G. Population Size Estimation of Men Who Have Sex with Men in Tbilisi, Georgia; Multiple Methods and Triangulation of Findings. Plos One 2016; 11(2).
40. Handcock MS, Gilo KJ, Mar CM. Estimating hidden population size using Respondent-Driven Sampling data. Electron J Stat 2014; 8:1491–1521. [PubMed: 26180577]

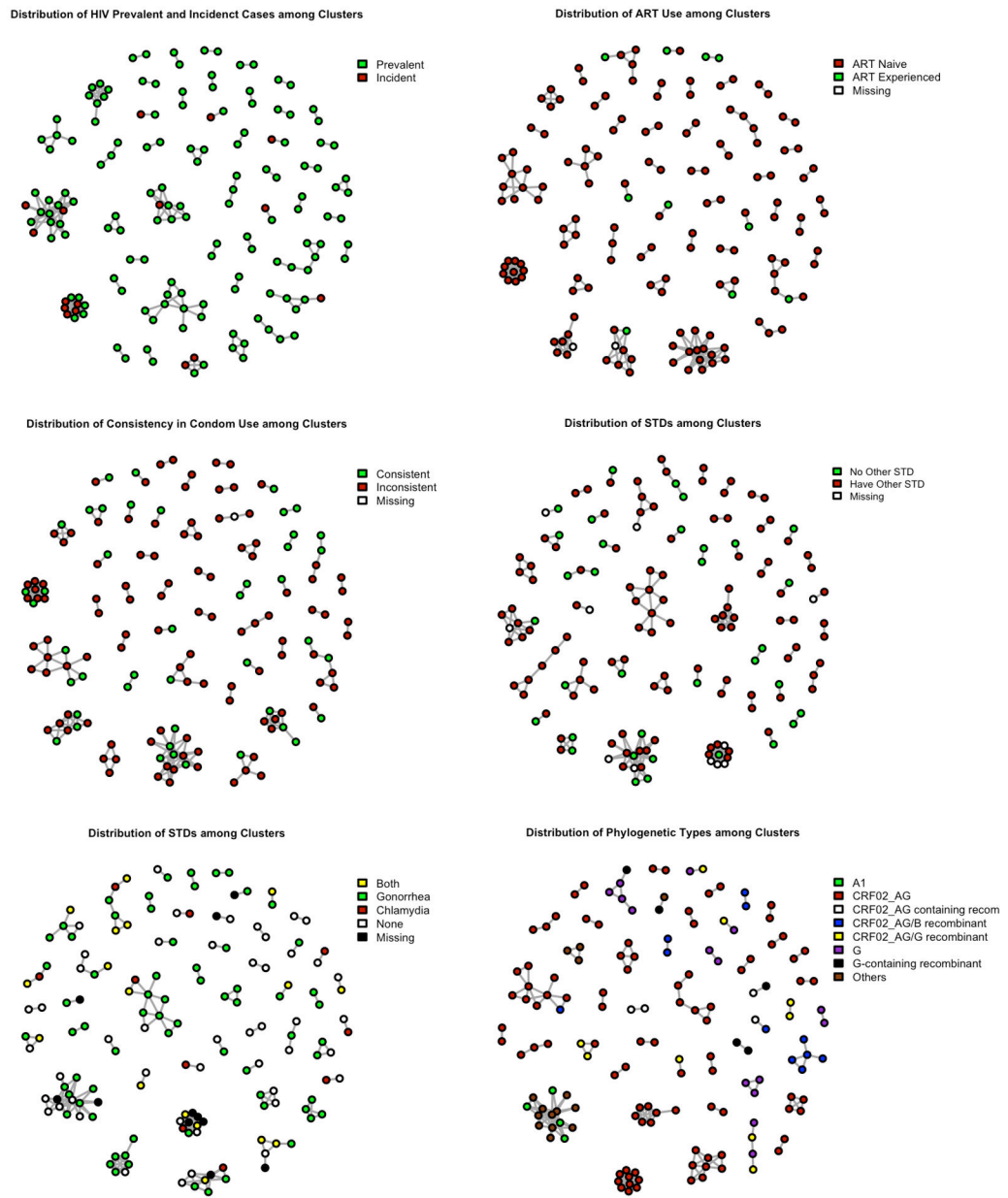


Figure 1. Distribution of risk behaviors in clusters.

Table 1:

Sample characteristics stratified by cluster among MSM in Nigeria

Characteristics	Not in a Cluster	In a cluster	Total	P-value ¹	RDS-Adjusted
	N=264(63.3%)	N=153(36.7%)	N=417(100%)		P-value ²
<i>Categorical</i>	<i>N (%)</i>	<i>N (%)</i>	<i>N (%)</i>		
Gender Identity					
Male	199(75.4%)	125(81.7%)	324(77.7%)	0.12	0.41
Transgender Women	35(13.3%)	19(12.4%)	54(13.0%)		
Both Male and Female	29(11.0%)	8(5.2%)	37(8.9%)		
Age					
16–19	27(10.2%)	24(15.7%)	51(12.2%)	0.07	0.60
20–24	109(41.3%)	71(46.4%)	180(43.2%)		
>=25	127(48.1%)	58(37.9%)	185(44.4%)		
Male Partners					
<=1	149(56.4%)	88(57.5%)	237(56.8%)	0.81	0.42
>1	114(43.2%)	64(41.8%)	178(42.7%)		
ART use					
ART Experienced	32(12.1%)	10(6.5%)	42(10.1%)	0.07	0.50
ART Inexperienced	231(87.5%)	141(92.3%)	372(89.2%)		
Education					
Above High School	235(89.0%)	141(92.2%)	376(90.2%)	0.71	0.56
Less than High School	23(8.7%)	12(7.8%)	35(8.4%)		
Missing	6(2.3%)	0(0.0%)	6(1.4%)		
Religion					
Christianity	192(72.7%)	123(80.4%)	315(75.5%)	0.09	0.71
Muslim or Atheist	71(26.9%)	30(19.6%)	101(24.2%)		
Occupation					
Employed	202(76.5%)	111(72.5%)	313(75.1%)	0.25	0.53
Unemployed	57(21.6%)	41(26.8%)	98(23.5%)		
Missing	5(1.9%)	1(0.7%)	6(1.4%)		
Consistent Condom Use					
Always	95(36.0%)	44(28.8%)	139(33.3%)	0.12	0.65
Not Always	165(62.5%)	108(70.6%)	273(65.5%)		
Missing	4(1.5%)	1(0.7%)	5(1.2%)		
STIs					
No other STIs	73(27.7%)	34(22.2%)	107(25.7%)	0.14	0.93
Have other STIs	163(61.7%)	108(70.6%)	271(65.0%)		
Missing	28(10.6%)	11(7.2%)	39(9.4%)		
Female Partner					
No Female Partner	130(49.2%)	95(62.1%)	277(66.4%)	0.01	<0.01
Have Female Partner	133(50.4%)	57(37.3%)	190(45.6%)		
Sexual Orientation					
				0.45	0.28

Gay	95(36.0%)	61(39.9%)	156(37.4%)		
Bisexual	168(63.6%)	92(60.1%)	260(62.4%)		
WHO Stage				0.65	0.07
Stage 1	222(84.1%)	130(85.0%)	352(84.4%)		
Stage 3–4	25(9.5%)	17(11.1%)	42(10.1%)		
Missing	17(6.4%)	6(3.9%)	23(5.5%)		
Incidence				0.74	0.25
Prevalent Cases	239(90.5%)	137(89.5%)	376(90.2%)		
Incident Cases	25(9.5%)	16(10.5%)	41(9.8%)		
Sexual Position				0.01	0.43
Insertive	43(16.3%)	13(8.5%)	56(13.4%)		
Receptive	56(21.2%)	53(34.6%)	109(26.1%)		
Versatile	162(61.4%)	86(48.6%)	248(59.5%)		
Missing	3(1.1%)	1(0.7%)	5(1.0%)		
Subtypes				0.10	0.03
A1	2(0.8%)	4(2.6%)	6(1.4%)		
CRF02_AG	154(58.3%)	90(58.8%)	244(58.5%)		
CRF02_AG recombinant	13(4.9%)	4(2.6%)	17(4.1%)		
CRF02_AG/B recombinant	21(8.0%)	11(7.2%)	32(7.7%)		
CRF02_AG/G recombinant	17(6.4%)	9(5.9%)	26(6.2%)		
G	38(14.4%)	15(9.8%)	53(12.7%)		
G-containing recombinant	1(0.4%)	5(3.3%)	6(1.4%)		
Others	18(6.8%)	15(9.8%)	33(7.9%)		
CD4 Count(cells/mm³)				0.52	0.76
>=500	45(17.1%)	29(19.0%)	74(17.8%)		
350–449	63(23.5%)	44(28.8%)	106(25.4%)		
201–349	72(27.3%)	41(26.8%)	113(27.1%)		
<=200	82(31.1%)	39(25.5%)	121(29.0%)		
Missing	3(1.1%)	0(0.0%)	3(0.7%)		
Continuous	Mean (SD)	Mean (SD)	Mean (SD)		
Viral Load (Log₁₀copies/ml)	4.61(0.83)	4.68(0.74)	4.63(0.79)	0.38	<0.01
Missing	28(10.6%)	11(7.2%)	39(9.3%)		

¹: Chi-square p-value was reported excluding

²: P-value for Rao-Scott Chi-Square test was reported

Table 2:

Factors associated with cluster membership and size

	Membership in a Cluster		Cluster Size 2–3		Cluster Size>3	
	OR	95% CI	OR	95% CI	OR	95% CI
Age						
>25	<i>Reference</i>		<i>Reference</i>			
20–25	2.04	0.76–6.21	2.15	0.89–5.17 [^]	1.85	0.29–11.66
16–19	1.85	0.16–22.03	2.30	0.20–26.20	1.38	0.08–25.40
Gender Identity						
Both Male and Female	<i>Reference</i>		<i>Reference</i>			
Male	3.42	1.08–10.78*	17.28	2.93–102.12*	1.84	0.49–6.87
Transgender Women	1.71	0.36–6.80	5.22	0.17–159.89	1.48	0.50–4.38
Education						
Less than High School	<i>Reference</i>		<i>Reference</i>			
Above High School	1.24	0.32–4.83	0.34	0.08–1.51	23.84	2.66–213.49*
Religion						
Muslim or Atheist	<i>Reference</i>		<i>Reference</i>			
Christianity	1.46	0.41–6.89	2.14	0.61–7.46	1.00	0.17–5.84
Occupation						
Unemployed	<i>Reference</i>		<i>Reference</i>			
Employed	1.14	0.49–4.09	0.71	0.18–2.89	3.41	1.89–10.70*
Sexual Orientation						
Gay	<i>Reference</i>		<i>Reference</i>			
Bisexual	1.31	0.96–1.79 [^]	1.87	0.74–4.68	1.02	0.40–2.55
Female Partner						
Have Female Partner	<i>Reference</i>		<i>Reference</i>			
No Female Partner	3.66	1.97–6.80*	3.17	1.89–5.32*	4.67	2.05–10.64*
log₁₀ Viral Load	1.72	1.04–2.86*	2.08	1.08–4.02*	1.46	0.80–2.66
CD4						
>=500	<i>Reference</i>		<i>Reference</i>			
350–449	1.21	0.57–2.57	0.84	0.25–2.84	1.82	0.53–6.31
201–349	1.31	0.66–2.60	0.78	0.19–3.13	2.35	0.84–6.54
<=200	0.85	0.27–2.61	0.72	0.10–5.00	0.99	0.40–2.42
Site						
Abuja	<i>Reference</i>		<i>Reference</i>			
Lagos	0.59	0.29–1.18	0.58	0.25–1.34	0.59	0.26–1.35

* p < 0.05

[^] p < 0.1 for trend

Table 3:

Association between ART status, sexual behaviors, and cluster membership and size

	Membership in a cluster		Cluster Size			
	AOR ¹	95% CI	Cluster Size 2–3		Cluster Size>3	
	AOR ¹	95% CI	AOR ¹	95% CI	AOR ¹	95% CI
ART use¹						
ART experienced		<i>Reference</i>			<i>Reference</i>	
ART inexperienced	2.82	0.79–10.14	1.86	0.48–7.20	6.61	2.25–19.37*
Condom Use²						
Always		<i>Reference</i>			<i>Reference</i>	
Not Always	0.99	0.40–2.45	0.79	0.25–2.51	1.23	0.71–2.13
STI²						
No other STI		<i>Reference</i>			<i>Reference</i>	
Have other STI	1.16	0.31–4.35	0.55	0.12–2.47	3.98	0.89–17.22 [^]
Sexual Position²						
Insertive		<i>Reference</i>			<i>Reference</i>	
Receptive	1.97	0.44–8.43	3.48	2.03–5.92*	1.24	0.15–10.10
Versatile	1.90	0.69–5.23	2.87	2.41–3.41*	1.38	0.32–6.02
# of Male Partners²						
1		<i>Reference</i>			<i>Reference</i>	
>1	1.32	0.78–2.29	1.42	0.67–3.06	1.24	0.73–2.08

¹: Each model was adjusted for age, education, sexual orientation, employment status, HIV incidence and study site.

²: All the variables were fitted independently.

[^] p-value <0.10

* p-value <0.05