UC San Diego
UC San Diego Previously Published Works

Title
Characterization of HIV Risk Behaviors and Clusters Using HIV-Transmission Cluster Engine Among a Cohort of Persons Living with HIV in Washington, DC

Permalink
https://escholarship.org/uc/item/0tj895q6

Journal
AIDS Research and Human Retroviruses, 37(9)

ISSN
0889-2229

Authors
Wilbourn, Brittany
Saafir-Callaway, Brittani
Jair, Kamwing
et al.

Publication Date
2021-09-01

DOI
10.1089/aid.2021.0031

Peer reviewed

Characterization of HIV Risk Behaviors and Clusters Using HIV-Transmission Cluster Engine Among a Cohort of Persons Living with HIV in Washington, DC

Brittany Wilbourn,[1,i] Brittani Saafir-Callaway,[2] Kamwing Jair,[1] Joel O. Wertheim,[3]
Oliver Laeyendeker,[4,5] Jeanne A. Jordan,[1] Michael Kharfen,[2]
and Amanda Castel,[1] on behalf of the DC Cohort Executive Committee

Abstract

Molecular epidemiology (ME) is one tool used to end the HIV epidemic in the United States. We combined clinical and behavioral data with HIV sequence data to identify any overlap in clusters generated from different sequence datasets; to characterize HIV transmission clusters; and to identify correlates of clustering among people living with HIV (PLWH) in Washington, District of Columbia (DC). First, Sanger sequences from DC Cohort participants, a longitudinal HIV study, were combined with next-generation sequences (NGS) from participants in a ME substudy to identify clusters. Next, demographic and self-reported behavioral data from ME substudy participants were used to identify risks of secondary transmission. Finally, we combined NGS from ME substudy participants with Sanger sequences in the DC Molecular HIV Surveillance database to identify clusters. Cluster analyses used HIV-Transmission Cluster Engine to identify linked pairs of sequences (defined as distance ≤1.5%). Twenty-eight clusters of ≥3 sequences (size range: 3–12) representing 108 (3%) participants were identified. None of the five largest clusters (size range: 5–12) included newly diagnosed PLWH. Thirty-four percent of ME substudy participants ($n = 213$) reported condomless sex during their last sexual encounter and 14% reported a Syphilis diagnosis in the past year. Seven transmission clusters (size range: 2–19) were identified in the final analysis, each containing at least one ME substudy participant. Substudy participants in clusters from the third analysis were present in clusters from the first analysis. Combining HIV sequence, clinical and behavioral data provided insights into HIV transmission that may not be identified using traditional epidemiological methods alone. Specifically, the sexual risk behaviors and STI diagnoses reported in the substudy survey may not have been disclosed during Partner Services activities and the survey data complemented clinical data to fully characterize transmission clusters. These findings can be used to enhance local efforts to interrupt transmission and avert new infections.

Keywords: HIV, molecular epidemiology, HIV clusters, HIV-TRACE, District of Columbia

Introduction

THE HIV/AIDS EPIDEMIC continues to present domestic and international public health challenges. The U.S. Centers for Disease Control and Prevention (CDC) estimates that >1.1 million people are currently living with HIV in the United States,[1] with close to 40,000 new infections each year.[2] The Ending the HIV Epidemic (EHE) initiative aims to reduce the number of new HIV infections in the United States by 90% by 2030.[3,4] The initiative has four pillars: (1)

---

[1]Department of Epidemiology, Milken Institute School of Public Health, The George Washington University, Washington, District of Columbia, USA.
[2]HIV/AIDS, Hepatitis, STD and TB Administration, DC Health, Washington, District of Columbia, USA.
[3]Department of Medicine, University of California San Diego, LA Jolla, California, USA.
[4]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.
[5]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Baltimore, Maryland, USA.
[i]ORCID ID (https://orcid.org/0000-0003-3428-1392).

diagnose all people living with HIV (PLWH) as early as possible; (2) *treat* PLWH quickly after diagnosis to achieve sustained viral suppression; (3) *prevent* new HIV infections using proven methods; (4) and rapidly detect and *respond* to growing HIV clusters to further prevent new infections.

The fourth pillar of the EHE initiative can be addressed through recent applications of molecular epidemiology (ME) that allow for better understanding of HIV transmission dynamics among and between high-risk groups.[5–9] HIV sequence data can be used to identify transmission clusters comprising individuals with similar HIV sequences and who may represent potential transmission partners.[5,10,11] Clusters can be further characterized by combining sequence data and epidemiologic data.[12–14] Clusters identified using sequence data can also complement the sexual networks constructed based on contacts elicited during contact tracing and partner notification investigations.[14–17] Although many studies have analyzed HIV sequence data retrospectively to generate transmission clusters,[16,18,19] identifying and characterizing HIV networks prospectively can inform interventions to rapidly interrupt disease transmission [11,20,21] and the HIV-Transmission Cluster Engine (HIV-TRACE) is one tool allowing for the identification of HIV clusters.[22]

The District of Columbia (DC) has a severe HIV epidemic: ~2% of the population is currently living with HIV and the leading transmission risk among both incident and prevalent cases are male-to-male sexual contact and heterosexual sexual contact.[23] One initiative to combat the epidemic specific to DC is the 90/90/90/50 Plan, which aims to have 90% of all District residents with HIV know their HIV status, 90% of DC residents living with HIV engaged in treatment, and 90% of those in treatment reach viral suppression, culminating in a 50% reduction of new HIV diagnoses by 2020.[24] The plan acknowledges the implications that cluster identification and characterization will have for its success and DC Health has been funded by the CDC to conduct Molecular HIV Surveillance (MHS) since 2016.[6,25] As of December 31, 2018, 47% of PLWH in DC have sequences in MHS. Given the potential for molecular HIV sequencing to contribute to interrupting HIV transmission in DC, we sought to identify any overlap in clusters generated from analyzing different sequence datasets; to characterize HIV transmission clusters; and to identify correlates of clustering among a large cohort of PLWH in DC.

Methods

Study population

The DC Cohort study is an ongoing longitudinal cohort study of PLWH receiving care at 15 clinical sites in Washington, DC.[26] In brief, enrollment began in January 2011 and persons living with HIV are eligible to participate if they receive care at 1 of the 15 participating clinics sites. Upon enrollment, clinic patients are assigned a participant ID (PID) that is not associated with any identifiable information and clinical data from the HIV care visits of consenting patients are abstracted from each site's electronic health record (EHR) into the study database. Periodically, DC Cohort data are linked to the DC Health HIV/AIDS, Hepatitis, STD, and TB Administration (HAHSTA) services and surveillance databases, inclusive of HIV molecular sequence data.[26,27] The linkage process is explained to participants during the consent process. As of 2016, 40% of PLWH in DC were enrolled in the DC Cohort.[28] In 2016, the DC Cohort launched an ME substudy, which focused on triangulating clinical, molecular sequence, and behavioral data to construct molecular clusters and describe behaviors and clinical features that might be associated with clustering. DC Cohort participants with molecular HIV sequences generated using commercial Sanger sequencing generated during routine HIV care and reported to the DC Health from 2011 to 2017 were included in this analysis. To identify PLWH at risk for transmitting infection, DC Cohort participants who were either (1) diagnosed in the 12 months before their study enrollment or (2) diagnosed more than 12 months before their study enrollment and viremic [viral load (VL) >1,500 copies/mL] as of their most recent VL test in the past 12 months were eligible to participate in the ME substudy. Participants were recruited by Research Assistants from 13 clinical sites to provide consent, complete a behavioral survey, and provide a blood sample for molecular sequencing. Upon enrollment in the substudy, a separate PID was assigned that is linked to their Cohort PID, but not to any identifying information. Substudy participants who consented to the DC Cohort by March 1, 2020 were included in this analysis, and the molecular sequences of these participants were generated using next-generation sequencing (NGS).[29] Prospectively collected NGS data were also linked with DC Health HAHSTA in an attempt to add missing sequences to the MHS database, and a subset of blood samples ($n = 22$) collected from participants was shared with researchers from Johns Hopkins University to determine the recency of HIV infection.[30] For recency testing, participant samples needed to be collected within 3 months of diagnosis and participants had to be naive to antiretroviral therapy (ART) with detectable virus. Any data shared with researchers external to the George Washington University are assigned a third PID to maintain data confidentiality. Figure 1 provides the flow of participants, samples, and sequences for the study. The study protocol, informed consent documents, and survey instrument were approved by the George Washington University Institutional Review Board (IRB), the DC Health IRB, and the IRB's of the clinical sites.

Measures

Data routinely extracted from the EHRs of DC Cohort participants include demographics (*i.e.*, sex at birth, race/ethnicity, and age), transmission risk, history of sexually transmitted infections (STIs), chronic hepatitis B virus (HBV) or hepatitis C virus (HCV), VL, history of ARV treatment, and ARV resistance. Participants enrolled in the ME substudy completed an extensive self-administered cross-sectional survey of risk behaviors and risk reduction strategies through REDCap, a secure, web-based application designed to support data capture for research studies.[31] The survey included measures of sociodemographic characteristics, sexual activity and risk behaviors with recent sex partners, risk reduction strategies, and ARV medication use. A limited set of DC Cohort clinical data were also matched to the self-reported survey responses.

Statistical and sequence analysis

We conducted three separate statistical analyses using these databases (Fig. 1). Our first analysis aimed to identify HIV clusters and the characteristics of persons in those
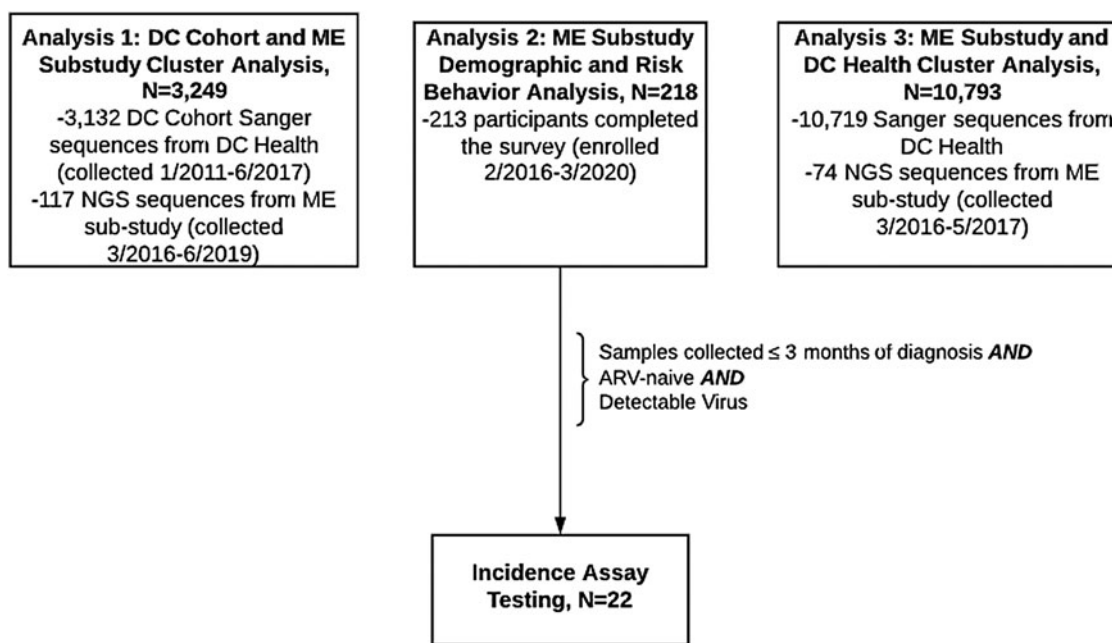
FIG. 1. Flow of participants, samples, and sequences. ARV, antiretroviral; ME, molecular epidemiology; NGS, next-generation sequences.

clusters among 3,249 DC Cohort participants using a combination of Sanger sequences ($n = 3,132$) reported to DC Health between January 2011 and June 2017 and available NGS ($n = 117$) from the ME substudy collected from March 2016 to June 2019. The second analysis described the demographic and risk behaviors of 218 participants enrolled in the ME substudy between February 2016 and March 2020. This analysis also included incidence assay testing on 22 samples. Our third analysis linked available ME substudy NGS ($n = 74$) collected from March 2016 to May 2017 with all available Sanger sequences in the DC Health Molecular HIV Surveillance (DC MHS) database ($n = 10,719$) to determine whether ME substudy participants contributed to DC MHS cluster identification.

All Sanger sequences were ordered for genotyping as part of routine HIV care at the clinical sites and reported to HAHSTA per DC Health guidelines. NGS laboratory sequencing for this study has been described elsewhere.[32] Consensus Sanger and NGS were treated the same in alignment and distance analysis, given that RNA NGS consensus sequences behave like Sanger sequences in HIV-TRACE networks, assuming that ambiguous nucleotides are called when their mixture is between 20% and 80%.[33] HIV-TRACE was used to identify molecular transmission clusters using a pairwise genetic distance threshold ≤0.015 substitutions/site from the reverse transcriptase (*RT*) or protease *(PR/RT)* region (HXB2 coordinates: 2,253–3,749) and an ambiguity threshold of 1.5%.[22] If multiple sequences (Sanger or NGS) were available for a single participant, all sequences were included. A genetic link was established between pairs of participants if at least one of their sequences was below the genetic distance threshold. All participants included in the cluster analyses were stratified by cluster status (in a cluster of three or more vs. a pair). Univariate analyses using chi-square and Wilcoxon rank sum tests were conducted to describe and examine differences between participants. All analysis was conducted in SAS 9.4 (Cary, NC).

Results

Analysis 1: Clusters and characteristics of DC Cohort participants by cluster size

People in clusters were mixed with respect to possible HIV risk behaviors (54% men who have sex with men and 34% heterosexual sex), the majority of participants (60%) were on ARVs at the time of sequencing, and the median VL was 10,538 copies/mL. Three of these five largest clusters included only PLWH with Sanger-generated HIV sequences. Participants in clusters of three or more PLWH were significantly younger (median age 32 vs. 44, $p < .0001$) and had been living with HIV for less time (median 9 vs. 14 years, $p < .0001$) compared with those in clusters of ≤2 (Table 1). No significant differences were observed with respect to race/ethnicity, sex, transmission risk, history of STIs, HBV or HCV, nor in VL or CD4 count among those PLWH in clusters of ≥3 versus those in clusters of ≤2 persons. Sequence data were available for a total of 3,249 DC Cohort participants, 97% ($n = 3,132$) of whom had Sanger-generated HIV sequences and 3% ($n = 117$) of whom had NGS-generated HIV sequences. Among the 3,249 sequences analyzed, HIV-TRACE found 208 genetic links connecting 268 individuals. A total of 28 clusters of three or more sequences (range: 3–12) representing a total of 108 (3%) participants were identified (Fig. 2). The five largest clusters ($n = 35$ participants) ranged in size from 5 to 12 PLWH, and none of the clusters included newly diagnosed PLWH (Table 2).

Analysis 2: Behavioral characteristics of ME substudy participants

Of the 218 participants enrolled in this substudy, 213 completed the behavioral survey and 196 participants had available EHR data (Table 3). Participants were a median age of 44 [interquartile range (IQR) = 32–54], and the majority were men (69%) and identified as Black (80%). The median

TABLE 1. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF PARTICIPANTS BY CLUSTER STATUS ($N = 3{,}249$)

Characteristic	In a cluster of ≥3[a] n (%)	In a cluster of ≤2[a] n (%)	Total n	χ^2 p-Value[b]
	108 (3)	3,141 (97)	3,249	
Age at substudy consent, median (IQR)	32 (24–44)	44 (32–53)	44 (32–53)	**<.0001**
Sex at birth				
Male	81 (76)	2,123 (68)	2,204 (68)	.0810
Female	26 (24)	1,013 (32)	1,039 (32)	
Race/ethnicity				
NH Black	91 (85)	2,699 (86)	2,790 (86)	.4332
NH White	9 (8)	247 (8)	256 (8)	
Hispanic	7 (6)	140 (4)	147 (4)	
Other[c]	0 (0)	50 (2)	50 (2)	
State of residence				
DC	84 (91)	2,006 (90)	2,090 (90)	.6422
MD	8 (9)	189 (8)	197 (8)	
VA	0 (0)	35 (1)	35 (1)	
Other	0 (0)	5 (0.2)	5 (0.2)	
Mode of transmission				
Men who have sex with men	61 (58)	1,491 (48)	1,552 (48)	.0888
Heterosexual sex	30 (28)	1,010 (32)	1,040 (32)	
Persons who inject drugs	13 (12)	417 (13)	430 (13)	
Other[d]	2 (2)	216 (7)	218 (7)	
Hx of STIs[e]	17 (18)	328 (15)	345 (15)	.3145
Hx HBV	3 (3)	190 (8)	193 (8)	.0741
Hx HCV	6 (6)	254 (11)	260 (11)	.1484
Length of HIV infection (years)				
≤1	0 (0)	18 (1)	18 (1)	<.0001
2–5	11 (10)	221 (7)	232 (7)	
6–10	57 (54)	893 (28)	950 (29)	
11–15	20 (19)	637 (20)	657 (20)	
16–19	4 (4)	385 (12)	389 (12)	
≥20	14 (13)	976 (31)	99 (31)	
CD4 (cells/μL)[f]	419 (258, 598)	369 (186, 559)	371 (187, 560)	.0885
VL (copies/mL)[f]	13,100 (1340, 47750)	10,460 (670, 51200)	10,510 (680, 50780)	.3878
ARV history				
Experienced	69 (75)	1,766 (79)	1,832 (79)	.4720
Naïve	8 (9)	129 (6)	137 (6)	
Unknown	15 (16)	347 (15)	362 (15)	

[a]Totals may not sum to N owing to missing data.
[b]Chi-square or Wilcoxon test; significant p values <.05 are given in bold.
[c]Other race includes mixed race individuals, Asians, Alaska Natives, American Indians, Native Hawaiian, Pacific Islanders, and unknown race.
[d]Other mode of transmission includes perinatal transmission, hemophilia, blood transfusion, and occupational exposure (health care workers).
[e]Chlamydia, gonorrhea, syphilis, trichomoniasis.
[f]CD4 and VL are at the time of specimen collection.
ARV, antiretroviral; DC, District of Columbia; Dx, diagnosis; HBV, hepatitis B virus; HCV, hepatitis C virus; Hx, history; IQR, interquartile range; NH, non-Hispanic; MD, Maryland; STIs, sexually transmitted infections; VA, Virginia; VL, viral load.

length of HIV diagnosis was 12 years and more than a third of participants (36%) acquired HIV through male-to-male sexual contact. The median VL upon enrollment into the substudy was 6,975 copies/mL (IQR = 40–66900) and most participants (93%) were ARV experienced upon enrollment in the substudy. Most participants (65%) had engaged in sexual activity in the prior 12 months, and the mean number of sex partners in the prior 12 months was 11 (standard deviation = 42). Almost half of participants (49%) described their most recent sexual partner as a ''primary'' partner and almost half of participants (45%) reported that their most recent partner was HIV negative. Twenty-nine percent of HIV-negative partners were reported to be taking pre-

exposure prophylaxis. More than one third of participants (34%) reported condomless sex during their last sexual encounter and 14% of participants reported having been diagnosed with Syphilis in the past 12 months. The majority of participants self-reported current ARV use (75%) and 63% reported taking all their ARVs in the past 30 days.

Analysis 3: Linkage of ME substudy and DC Health molecular HIV surveillance sequences

NGS HIV sequence data from 74 participants enrolled in the ME substudy from March 2016 to May 2017 were shared with DC Health HAHSTA, of which 15 participants did not
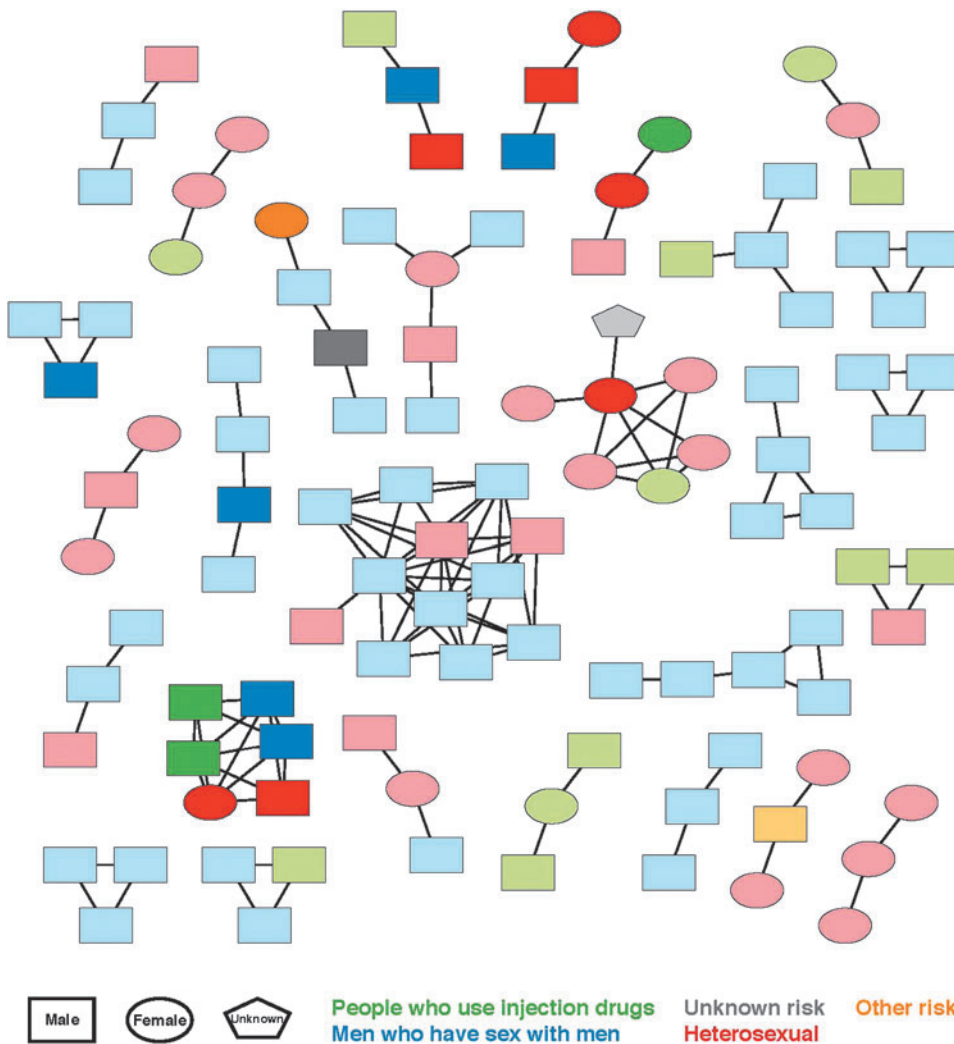
FIG. 2. Retrospective and prospective clusters of three or more participants in the DC Cohort ME substudy ($n = 28$ clusters from 108 participants). *Edges* denote genetic distance ≤0.015 substitutions/site. *Light shading* denotes Sanger-generated sequences. *Dark shading* denotes NGS-generated sequences. *Color* denotes HIV transmission risk. *Shape* denotes sex assigned at birth. DC, District of Columbia. Color images are available online.

have a sequence previously reported to the DC MHS system. Of the 59 participants whose HIV sequences were already in MHS, the ME substudy NGS data had an earlier sequence date for 11 (18.6%) of these individuals. When the 74 NGS-generated HIV sequences from the ME substudy and 10,719 Sanger HIV sequences from all PLWH in DC from the DC MHS database were analyzed, HIV-TRACE found 350 genetic links connecting 49 individuals, 12 of whom were enrolled in the substudy (Fig. 3). A total of seven transmission clusters (range in size: 2–19) were identified, each containing at least one participant from the ME substudy; one cluster of six contained five participants from the substudy (Table 4). Of the five substudy participants in this cluster, two were also identified in the third DC Cohort cluster described in Table 1. The largest cluster comprised mostly men (90%), with the majority of these participants infected through male-to-male sexual contact (67%). Furthermore, two substudy participants identified in ME substudy/DC MHS clusters were identified in DC Cohort dyads from the first analysis. Of the 22 participants whose samples were sent for incidence testing, 5 were determined to be recently infected, 3 of whom were Black or African American and 3 of whom had an HIV mode of transmission of MSM (data not shown).

Discussion

Principal findings

This study utilized a combination of Sanger and NGS, clinical and behavioral data among a cohort of PLWH to identify any overlap in clusters generated from analyzing different sequence datasets, characterize HIV clusters, and identify clinical and behavioral correlates of clustering. Based on clusters identified by analyzing Sanger and NGS using HIV-TRACE (Figs. 2 and 3), cluster composition by sex and mode of transmission revealed a high level of genetic clustering among men who had sex with other men, consistent with several other studies.[5–7,12] Also consistent with other studies, factors associated with clustering included younger age[7,11,12] and less time living with HIV (Table 1).[11] However, in our study no significant differences in cluster size were observed with respect to race/ethnicity, sex, history of STIs, HBV or HCV, or HIV clinical parameters (Table 1), despite our hypothesis that these might be associated with secondary transmission risks. By characterizing clusters and identifying correlates of clustering, HIV prevention and treatment efforts can be further targeted.

Demographically, the median age of participants in this study was 44 years and the majority of participants were men and Black (Table 1). These characteristics are similar to those

TABLE 2. CHARACTERISTICS OF PARTICIPANTS (N=35) IN FIVE LARGEST CLUSTERS AMONG DC COHORT PARTICIPANTS WITH AVAILABLE MOLECULAR SEQUENCES

Participant characteristics	Cluster no.					
	1	2	3	4	5	Total
Cluster size	12	7	6	5	5	35
Retrospective Sanger vs. Prospective NGS	12 Retro	6 Retro, 1 Pro	6 Pro	5 Retro	5 Retro	35
No. of new diagnoses in cluster	0	0	0	0	0	0
Years since HIV dx, median (IQR)	9 (8, 10)	7 (7, 8)	22 (21, 24)	8 (7, 9)	12 (11, 12)	9 (8, 12)
Age at time of sequencing, median (IQR)	25 (23, 31)	28 (24, 31)	49 (47, 55)	32 (26, 34)	28 (28, 29)	30 (24, 33)
Mode of transmission						
Men who have sex with men	9 (75)	0 (0)	2 (33)	3 (60)	5 (100)	19 (54)
Heterosexual sex	3 (25)	5 (71)	2 (33)	2 (40)	0 (0)	12 (34)
Persons who inject drugs	0 (0)	1 (14)	2 (33)	0 (0)	0 (0)	3 (9)
On ARVs at time of sequencing	6 (50)	4 (57)	6 (100)	2 (40)	3 (60)	21 (60)
VL at time of sequencing (copies/mL), median (IQR)	28,356 (5567, 78902)	20,260 (8004, 30872)	11,448 (4322, 20315)	3,630 (2560, 17320)	590 (411, 2600)	10,538 (1042, 37378)

TABLE 3. BEHAVIORAL CHARACTERISTICS OF PROSPECTIVE PARTICIPANTS (N=213)

Demographics	Total[a]
	n (%)
Age, median (IQR)	44 (32, 54)
Sex at birth	
Male	135/196 (69)
Race	
NH Black	157/196 (80)
NH White	15/196 (8)
Mode of transmission	
Men who have sex with men	71/196 (36)
Heterosexual sex	57/196 (29)
Persons who inject drugs	11/196 (6)
VL at enrollment, median (IQR)	6,975 (40–66900)
ARV exposure at sub-study enrollment	
Experienced	183/196 (93)
Naive	7/196 (4)
Months since diagnosis, median (IQR)	140 (8–1433)

Risk behaviors	
Sex in the past 12 months	138 (65)
Number of partners in the past 12 months, median (IQR)	2 (1–5)
Most recent partner type	
Primary partner	65/134 (49)
Casual	41/134 (31)
Steady, nonprimary partner	14/134 (10)
HIV status of most recent partner	
HIV negative	60/133 (45)
HIV positive	39/133 (29)
Most recent HIV-negative partner on PrEP	17/59 (29)
Condomless sex at last sexual encounter	72 (34)
Self-reported STI diagnosis in the past 12 months	
Syphilis	29 (14)
Chlamydia	19 (9)
Gonorrhea	14 (7)
Self-reported currently taking ARVs	159 (75)
Self-reported 30-day adherence[b]	
All	96/153 (63)
Most	36/153 (23)

[a]Totals may not sum to N owing to missing data.
[b]Among those who self-reported taking ARVs.
PrEP, preexposure prophylaxis.

of the larger population of PLWH in DC.[28] Participants in our study also reported risk behaviors that have previously been associated with the potential for secondary HIV transmission including a high number of sex partners,[34,35] condomless sex,[34–37] and recent STI diagnoses.[38] However, the majority of participants in our study also self-reported high levels of ARV adherence despite having detectable VLs.

Additional pertinent information was gleaned by combining or sharing sequence data from participants in the ME substudy with that from other entities. Fifteen of the participants whose NGS data were shared with DC Health did not
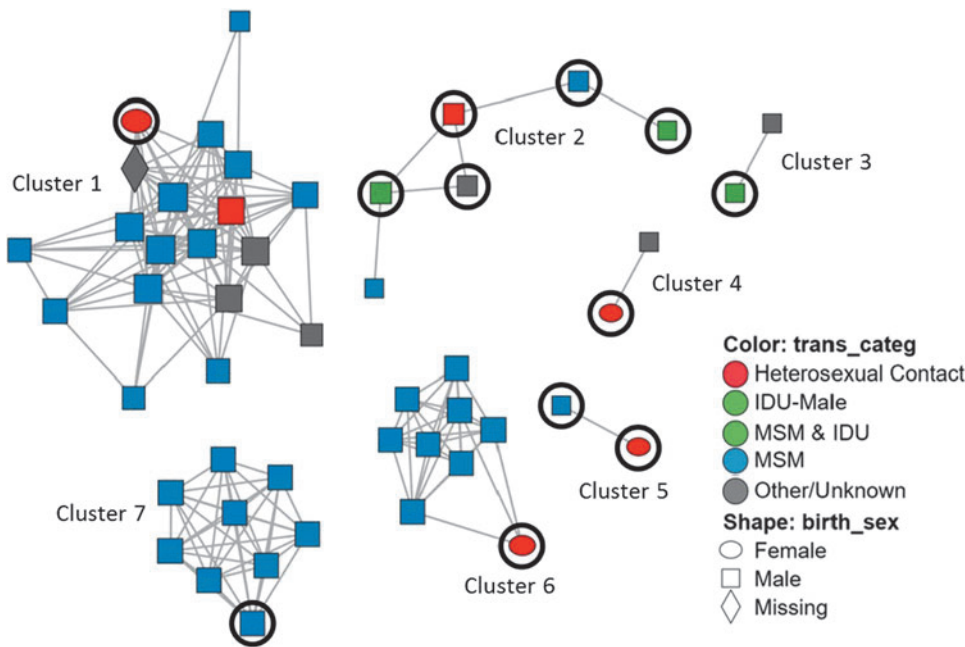
FIG. 3. ME substudy and DC Health molecular surveillance clusters ($n = 7$ clusters). *Edges* denote genetic distance ≤0.015 substitutions/site. *Color* denotes transmission risk. *Shape* denotes sex assigned at birth. *Circles* denote ME sub-study participants. IDU, injection drug user; MSM, men who have sex with men. Color images are available online.

have a Sanger sequence in the DC MHS database. Thus, sequence data collected as part of ME substudy might assist DC Health with achieving a more complete MHS system, an important issue reported nationally by other health departments conducting molecular HIV surveillance.[39] Furthermore, 11 participants who already had Sanger-generated HIV sequences in the MHS had NGS-generated HIV sequence data with an earlier sample date. This suggests that HIV sequence data generated as part of the ME substudy may be timelier than that in the DC MHS database. Analyzing both NGS from participants in the substudy with Sanger sequences in the MHS resulted in the identification of phylogenetic clusters that differed from the clusters identified when NGS from participants in the substudy were analyzed with Sanger sequences from Cohort participants. The addition of NGS data from the substudy could assist DC Health in identifying clusters that should be prioritized for intervention.[40] Finally, incidence testing revealed that the majority of participants in the substudy who had been recently diagnosed and whose samples had been tested were not, in fact, recently infected.

Limitations

The findings of this analysis should be interpreted in light of several limitations. Our sample only represents a subset of PLWH receiving care in DC between 2011 and 2020 who also had molecular genotyping conducted as part of their routine HIV care or enrolled in the ME substudy, which limits the generalizability of our findings. Moreover, as is the case in many geographic areas, the DC MHS does not contain molecular sequence data for all PLWH in DC. Linkage between the DC Cohort and DC Health HAHSTA did not result in molecular sequence data for all Cohort participants; only 37% of Cohort participants had sequence data available through the linkage. To address this issues, DC Health is working with local providers to encourage genotype testing per U.S. Department of Health and Human Services Guide-

lines.[41] Behavioral data collected from participants in the ME substudy were self-reported; thus, the data may be subject to selection and social desirability bias. However, it should be noted that the surveys were largely self-administered likely minimizing social desirability. Furthermore, although we used a combination of different sequencing techniques, because HIV-TRACE uses a 15% cutoff regardless of whether the HIV sequence data were generated by Sanger or NGS methods, data can be compared across these two different methods. Of importance, the role of MHS in interrupting HIV transmission is being questioned by community members, which may further limit peoples' willingness to have genotyping performed or to participate in Partner Services efforts that are informed by MHS data. Finally, the molecular clusters presented represent potential transmission partners; however, we cannot definitively assign a transmission link between any two persons or assess directionality of transmission, as third parties not captured in our data collection efforts may also be involved in transmission clusters. Despite this limitation inherent in phylogenetic analyses, we are still able to make inferences from ambiguous genetic linkage. Chief among the concerns expressed by researchers and community activists alike is the potential for prosecution of PLWH because of HIV criminalization laws.[42–47] In addition to the use of PIDs to protect participant data, DC Health has also drafted changes to existing legislation to better protect HIV surveillance data from being released by court order.[48] Community buy-in is needed to fully harness MHS data for the prevention and treatment of HIV and local community engagement efforts are underway to better understand and address concerns around public health interventions that utilize surveillance data.[49,50]

Conclusion

The success of the respond pillar of the EHE initiative hinges on the ability to identify and interrupt HIV transmission quickly. Toward this goal, we combined Sanger

TABLE 4. CHARACTERISTICS OF PARTICIPANTS IN MOLECULAR EPIDEMIOLOGY AND DC HEALTH CLUSTERS (N=49)

Participant characteristics	Cluster no.							Total
	1	2	3	4	5	6	7	
Cluster size	19	9	9	6	2	2	2	49
DC health vs. molecular Epi	18 DOH, 1 ME	8 DOH, 1 ME	8 DOH, 1 ME	1 DOH, 5 ME	1 DOH, 1 ME	2 ME	1 DOH, 1 ME	49
Sex at birth								
Male	17 (89)	9 (100)	8 (89)	6 (100)	2 (100)	1 (50)	1 (50)	44 (90)
Female	1 (5)	0	1 (11)	0	0	1 (50)	1 (50)	4 (8)
Unknown	1 (5)	0	0	0	0	0	0	1 (2)
Mode of transmission								
Men who have sex with men	13 (68)	9 (100)	8 (89)	2 (33)	0	0	1 (50)	33 (67)
Heterosexual sex	2 (10)	0	1 (11)	1 (17)	0	1 (50)	1 (50)	6 (12)
Persons who inject drugs	0	0	0	2 (33)	1 (50)	0	0	3 (6)
Unknown	4 (21)	0	0	1 (17)	1 (50)	1 (50)	0	7 (14)

and targeted NGS HIV sequence data with clinical and behavioral data that was shared with collaborating partners to better identify and characterize transmission clusters among PLWH in the DC metro area. These complementary data sources have the potential to provide important insight into HIV transmission that may not be uncovered through traditional epidemiological methods (*e.g.*, contact tracing and partner notification) alone. Our results indicate a high level of clustering among MSM, many of whom may be engaging in sexual risk behaviors that could result in secondary HIV transmission. The insight acquired from this analysis can be used to focus local public health efforts on the rapid identification of new HIV infections and related clusters to help interrupt disease transmission and avert new infections.

Authors' Contributions

Acknowledgments

Ethical approval: Written informed consent was obtained from all participants before enrollment in the DC Cohort and the ME substudy. The DC Cohort and ME studies were approved by the Institutional Review Board at The George Washington University, which serves as the IRB of record for Whitman-Walker Health, La Clinica del Pueblo, Family and Medical Counseling Service, Unity Health Care, The

GW Medical Faculty Associates, MetroHealth, and Children's National Health System (pediatric and adolescent clinics). The study was independently approved by the IRBs of record at Howard University Hospital (adult and pediatric clinics), MedStar Washington Hospital Center, Georgetown University, and the Veterans Affairs Medical Center.

Author Disclosure Statement

Funding Information

References

1. Centers for Disease Control and Prevention. PrEP 101. 2020. Available at https://www.cdc.gov/hiv/pdf/library/consumer-info-sheets/cdc-hiv-consumer-info-sheet-prep-101.pdf, accessed June 3, 2021.

2. Centers for Disease Control and Prevention. Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 (Updated), Atlanta, GA. 31;2020.

3. US Department of Health and Human Services. What is Ending the HIV Epidemic: A plan for America? 2020. Available at https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview, accessed June 3, 2021.

4. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP: Ending the HIV Epidemic: A plan for the United States. JAMA 2019;321:844–845.

5. Oster AM, Wertheim JO, Hernandez AL, Ocfemia MC, Saduvala N, Hall HI: Using molecular HIV surveillance data to understand transmission between subpopulations in the United States. J Acquir Immune Defic Syndr 2015;70: 444–451.

6. Dennis AM, Hue S, Pasquale D, et al.: HIV transmission patterns among immigrant Latinos illuminated by the integration of phylogenetic and migration data. AIDS Res Hum Retroviruses 2015;31:973–980.

7. 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; 68:46–54.

8. Whiteside YO, Song R, Wertheim JO, Oster AM: Molecular analysis allows inference into HIV transmission among young men who have sex with men in the United States. AIDS 2015;29:2517–2522.

9. Aldous JL, Pond SK, Poon A, et al.: Characterizing HIV transmission networks across the United States. Clin Infect Dis 2012;55:1135–1143.

10. Little SJ, Kosakovsky Pond SL, Anderson CM, et al.: Using HIV networks to inform real time prevention interventions. PLoS One 2014;9:e98443.

11. Dennis AM, Hue S, Hurt CB, et al.: Phylogenetic insights into regional HIV transmission. AIDS 2012;26:1813–1822.

12. Chalmet K, Staelens D, Blot S, 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.

13. Dennis AM, Herbeck JT, Brown AL, et al.: Phylogenetic studies of transmission dynamics in generalized HIV epidemics: An essential tool where the burden is greatest? J Acquir Immune Defic Syndr 2014;67:181–195.

14. Pasquale DK, Doherty IA, Sampson LA, et al.: Leveraging phylogenetics to understand HIV transmission and partner notification networks. J Acquir Immune Defic Syndr 2018; 78:367–375.

15. Delva W, Leventhal GE, Helleringer S: Connecting the dots: Network data and models in HIV epidemiology. AIDS 2016;30:2009–2020.

16. Smith DM, May SJ, Tweeten S, et al.: A public health model for the molecular surveillance of HIV transmission in San Diego, California. AIDS 2009;23:225–232.

17. Wertheim JO, Kosakovsky Pond SL, Forgione LA, et al.: Social and genetic networks of HIV-1 transmission in New York City. PLoS Pathog 2017;13:e1006000.

18. Ahn MY, Wertheim JO, Kim WJ, et al.: Short communication: HIV-1 transmission networks across South Korea. AIDS Res Hum Retroviruses 2017;33:827–831.

19. Oster AM, Pieniazek D, Zhang X, et al.: Demographic but not geographic insularity in HIV transmission among young black MSM. AIDS 2011;25:2157–2165.

20. Poon AFY, Gustafson R, Daly P, et al.: Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: An implementation case study. Lancet HIV 2016;3:e231–e238.

21. Wertheim JO, Murrell B, Mehta SR, et al.: Growth of HIV-1 molecular transmission clusters in New York City. J Infect Dis 2018;218:1943–1953.

22. Kosakovsky Pond SL, Weaver S, Leigh Brown AJ, Wertheim JO: HIV-TRACE (TRAnsmission Cluster Engine): A tool for large scale molecular epidemiology of HIV-1 and other rapidly evolving pathogens. Mol Biol Evol 2018;35: 1812–1819.

23. District of Columbia Department of Health HIV/AIDS, Hepatitis, STD, and TB Administration. Annual Epidemiology & Surveillance Report: Data Through December 2018. 2019. Available at: https://dchealth.dc.gov/sites/default/files/dc/sites/doh/publication/attachments/2019%20Annual%20Surveillance%20Report_09062019.pdf, accessed June 3, 2021.

24. DC Health, DC Appleseed. 90/90/90/50 Plan: Ending the HIV Epidemic in the District of Columbia by 2020. 2016. https://doh.dc.gov/sites/default/files/dc/sites/doh/page_content/attachments/DC%2090-90-90-50%20Plan%20-%20FINAL.pdf, accessed June 3, 2021.

25. DC Appleseed. Ending the HIV Epidemic in DC: 2018 Progress Report. 2018. http://www.dcappleseed.com/wp-content/uploads/2018/12/Ending-the-HIV-Epidemic-in-DC-2018-Progress-Report.pdf, accessed June 3, 2021.

26. Greenberg AE, Hays H, Castel AD, et al.: Development of a large urban longitudinal HIV clinical cohort using a web-based platform to merge electronically and manually abstracted data from disparate medical record systems: Technical challenges and innovative solutions. J Am Med Inform Assoc 2016;23:635–643.

27. Castel A, Terzian A, Opoku J, et al.: Defining care patterns and outcomes among persons living with HIV in Washington, DC: Linkage of clinical cohort and surveillance data. JMIR Public Health Surveillance 2018;4:e33.

28. Opoku J, Doshi R, Castel A, et al.: Comparison of clinical outcomes of persons living with HIV by enrollment status in Washington, DC: Evaluation of a large longitudinal HIV cohort study. JMIR Public Health Surveillance 2020;6:e16061.

29. Metzker ML: Sequencing technologies—The next generation. Nat Rev Genet 2010;11:31–46.

30. Konikoff J, Brookmeyer R, Longosz AF, et al.: Performance of a limiting-antigen avidity enzyme immunoassay for cross-sectional estimation of HIV incidence in the United States. PLoS One 2013;8:e82772.

31. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–381.

32. Jair K, McCann CD, Reed H, et al.: Validation of publicly-available software used in analyzing NGS data for HIV-1 drug resistance mutations and transmission networks in a Washington, DC, Cohort. PLoS One 2019;14:e0214820.

33. Collura RV, Patterson W, Swain C-A, et al.: Implications of next-generation sequencing for drug resistance and cluster detection. Conference on Retroviruses and Opportunistic Infections; 2020; Boston, MA.

34. Weinhardt L, Kelly J, Brondino M, et al.: HIV Transmission risk behavior among men and women living with HIV in 4 cities in the United States. J Acquir Immune Defic Syndr 2004;36:1057–1066.

35. Sullivan KM, Dawson Rose C, Phillips JC, et al.: Sexual transmission-risk behaviour among HIV-positive persons: A multisite study using social action theory. J Adv Nurs 2017;73:162–176.

36. Du P, Crook T, Whitener C, Albright P, Greenawalt D, Zurlo J: HIV transmission risk behaviors among people living with HIV/AIDS: The need to integrate HIV prevention interventions and public health strategies into HIV care. J Public Health Manag Pract 2015;21:E1–E10.

37. Golin C, Marks G, Wright J, et al.: Psychosocial characteristics and sexual behaviors of people in care for HIV infection: An examination of men who have sex with men, heterosexual men and women. AIDS Behav 2009;13:1129–1142.

38. Secco AA, Akselrod H, Czeresnia J, et al.: Sexually transmitted infections in persons living with HIV infection and estimated HIV transmission risk: Trends over time from the DC Cohort. Sex Transm Infect 2020;96:89–95.

39. Dasgupta S, Hall HI, Hernandez AL, Ocfemia MCB, Saduvala N, Oster AM: Receipt and timing of HIV drug resistance testing in six U.S. jurisdictions. AIDS Care 2017;29:1567–1575.

40. Oster A, France A, Panneer N, et al.: Identifying clusters of recent and rapid HIV transmission through analysis of molecular surveillance data. J Acquir Immune Defic Syndr 2018;79:543–553.

41. Panel of Antiretroviral Guidelines for Adults and Adolescents. Guidelines for Use of Antiretroviral Agents in Adults and Adolescents with HIV. Department of Health and Human Services, Washington, DC, 2019.

42. Dawson L, Benbow N, Fletcher FE, et al.: Addressing ethical challenges in US-based HIV phylogenetic research. J Infect Dis 2020;222:1997–2006.

43. Nelson B: Questioning the Benefits of Molecular Surveillance. In: POZ. New York, NY, 2020.

44. Smith JC: Fourth Pillar or ''Third Rail?:'' Towards a community-centered understanding of the role of molecular HIV surveillance in Ending the HIV Epidemic in the United States. Am J Bioeth 2020;20:5–6.

45. Centers for Disease Control and Prevention. HIV and STD Criminalization Laws. 2020. Available at https://www.cdc.gov/hiv/policies/law/states/exposure.html, accessed June 3, 2021.

46. The Center for HIV Law and Policy. Is Molecular HIV Surveillance Worth the Risk? Brooklyn, NY, 2019.

47. Ryan B: Will the Genetic Analysis–Based HIV Surveillance Safeguard Privacy? In: POZ. New York, NY, 2018.

48. Council of the District of Columbia. HIV/AIDS Data Privacy Protection Amendment Act of 2021, Washington, DC, 2021.

49. DC Center for AIDS Research. New NIH CFAR Ending the HIV Epidemic supplement recipients: Adam Allston, PhD, MPH, MSW and Brittani Saafir-Callaway, PhD, MPH. 2019. Available at https://dccfar.gwu.edu/new-nih-cfar-ending-hiv-epidemic-supplement-recipients-adam-allston-phd-mph-msw-and-brittani-saafir, accessed June 3, 2021.

50. National Institute of Allergy and Infectious Diseases. CFAR/ARC Ending the HIV Epidemic Supplement Awards. 2020. Available at https://www.niaid.nih.gov/research/cfar-arc-ending-hiv-epidemic-supplement-awards, accessed June 3, 2021.

Address correspondence to:
Brittany Wilbourn
Department of Epidemiology
Milken Institute School of Public Health
The George Washington University
950 New Hampshire Avenue Northwest, Suite 500
Washington, DC 20037
USA

E-mail: blewis1@gwu.edu