Title

HIV transmission networks among transgender women in Los Angeles County, CA, USA: a phylogenetic analysis of surveillance data

Journal

The Lancet HIV, 6(3)

ISSN

2352-3018

Authors

Ragonnet-Cronin, Manon

Hu, Yunyin W

Morris, Sheldon R

et al.

Publication Date

2019-03-01

DOI

10.1016/s2352-3018(18)30359-x

Peer reviewed

HHS Public Access

Author manuscript

Lancet HIV. Author manuscript; available in PMC 2019 December 02.

Published in final edited form as:

Lancet HIV. 2019 March ; 6(3): e164–e172. doi:10.1016/S2352-3018(18)30359-X.

HIV Transmission Networks Among Transgender Women in Los Angeles County: network analysis of surveillance data

Manon Ragonnet-Cronin[1,*], **Yunyin W. Hu**[2], **Sheldon R. Morris**[1], **Zhijuan Sheng**[2], **Kathleen Poortinga**[2], **Joel O. Wertheim**[1]

[1]Department of Medicine, University of California, San Diego, California, USA [MRC, PhD; SRM, MD; JOW, PhD]

[2]Division of HIV and STD Programs, Department of Public Health, Los Angeles County, California, USA [YWH, PhD; ZS, MSc, KP, MSc]

SUMMARY

Background—Transgender women (TGW) are among the groups at highest risk for HIV infection, with a prevalence of 27.7% in the United States, but despite this high risk, TGW have documented high rates of undiagnosed HIV infection. We propose that this disparity can be addressed by characterizing TGW in a molecular transmission network to prioritize public health activities.

Methods—Since 2006, HIV *pol* sequences from drug resistance testing have been reported to Los Angeles County (LAC) Department of Public Health and linked to demographic data, gender, and HIV transmission risk factor data for each case in the enhanced HIV/AIDS Reporting System (eHARS). We reconstructed a molecular transmission network using HIV-TRACE (pairwise genetic distance threshold of 0.015 substitutions/site) from the earliest *pol* sequences from 22,398 unique individuals, including 412 (2%) self-identified TGW. We examined the possible predictors of clustering using multivariate logistic regression. We characterized the genetically-linked partners of TGW and calculated assortativity—the tendency for persons to link to other persons with the same attributes—for each transmission risk group.

Findings—We found that 36% of individuals (8,133/22,398) clustered in the network across 1,722 molecular transmission clusters. TGW who indicated a sexual risk factor clustered at the highest frequency in the network: 147/345 (42.6%) linked to at least one other person (p<0.001). TGW were assortative in the network (0.06; p<0.001), indicating that they tended to link to other TGW. TGW were more likely than expected to link to other TGW and cisgender men who did not identify as men who have sex with men (MSM). TGW were less likely than expected to link to

*Corresponding author: manonragonnet@ucsd.edu (MRC).

AUTHOR CONTRIBUTIONS

YWH, ZS and KP collected and prepared data. JOW and MRC conceived of the study and performed all analyses. SM and JOW provided support and advice. YWH reviewed the statistical analyses. MRC wrote the manuscript and all authors revised the final draft.

COMPETING INTERESTS

JOW receives funding from the Centers for Disease Control and Prevention (CDC) and Gilead Sciences through grants and contracts awarded to UCSD.

DATA AVAILABILITY

California Code and Regulations (CCR) Title 17 prevents the submission of genetic sequences used in this study to GenBank.

MSM, despite the high prevalence of HIV among MSM. TGW were distributed across 126 clusters, and cis-individuals linked to one TGW were 9.2 times more likely to link to a second TGW than other individuals in the surveillance database. Reconstruction of the transmission network is limited by sample availability, but sequences were available for >40% of diagnoses.

Interpretation—TGW in LAC were more likely to cluster than any other risk group, suggesting high transmission rates—despite the small number of TGW in the network. TGW tended to cluster with other TGW, indicating shared risk activities (i.e., linked directly or through shared partners). This assortativity, and the observed tendency for linkage with cisgender men who did not identify as MSM, demonstrates the potential to use molecular epidemiology to both identify clusters likely to include undiagnosed HIV-infected TGW and improve the targeting of public health prevention and treatment services to TGW.

INTRODUCTION

The global HIV-1 pandemic is driven by geographical, gender, and socio-economic disparities (1). In North America and Europe, HIV burden is concentrated among marginalized and stigmatized populations, including sexual minorities and communities of color. In the United States (U.S.), men who have sex with men (MSM) are disproportionately affected, comprising 62% of HIV-1 diagnoses each year (2), and African Americans account for 44% of prevalent infections despite making up only 14% of the population (3). Transgender women (TGW; i.e., individuals assigned male at birth but who identify as female) are estimated to have an HIV prevalence of 27.7% (4), even higher than the 25% prevalence estimated for MSM (5). Of concern, African American TGW had even higher prevalence, averaging 56.3% (4). In parallel, cross-sectional HIV testing in Miami, San Francisco, and Los Angeles found an HIV prevalence rate of 12% among TGW with no previous test result indicating a high frequency of undiagnosed infection (4, 6).

Analysis of viral genetic sequences provides a route to uncovering transmission dynamics (7). HIV is particularly amenable to phylogenetic analysis, because of its rapid rate of evolution, making it possible to identify genetic networks of densely connected subpopulations (i.e., molecular transmission clusters) (8). These transmission clusters are presumed to comprise people at increased risk of HIV transmission or evidence of recent transmission events; however, densely sampled/sequenced sub-populations can also form clusters, even in the absence of increased transmission rates (9, 10). These clustering approaches have the capability to reveal patterns hidden from traditional epidemiological approaches [e.g., obscured transmission risk behaviors, like self-reported heterosexual males whose viruses cluster only with those from MSM (11, 12)].

California shoulders the second largest number of people living with HIV-1 in the U.S. as of 2014 (with 119,589 cases), following New York State (130,753 cases) (3). Within California, Los Angeles County (LAC) had the greatest number of HIV diagnoses in 2015 and has the largest burden of persons living with HIV infection in the State: 60,000. The LAC HIV epidemic is dominated by MSM, who account for 83% of recent diagnoses (2), Since 2006, HIV-1 genetic sequences, generated for routine antiretroviral resistance genotyping, have been reported to the LAC Department of Public Health.

Here, we reconstructed the HIV-1 genetic transmission network from the LAC surveillance database, with a focus on transmission risk among TGW. We found that despite documented low rates of diagnosis (4), TGW clustered at higher rates than other risk groups. Furthermore, TGW were more likely than expected to link to each other and to cisgender males with a sexual transmission risk, rather than to MSM. We discuss how these results reveal a novel molecular epidemiological strategy that could be used to improve HIV-1 diagnosis rates in TGW and potentially reduce new HIV infections.

METHODS

Data sources

Since 2006, HIV-1 *protease* and *reverse transcriptase* (*pol*) genetic sequences generated during routine antiretroviral drug resistance testing have been reported to the LAC Department of Public Health. As of 2016, LAC HIV surveillance had received HIV-1 genetic sequences from 22,398 individuals residing or receiving care in LAC. Of 60,000 people estimated to be living with HIV, 49,976 had been diagnosed by 2015, thus 44.8% of diagnosed persons had a sequence available. We used the first genotype available for each individual. Information on treatment has been collected since 2006, and 69% of new cases since 2006 were treatment naïve at the time of their first genotype. Deduplication of cases is performed within the LAC database via a comprehensive procedure based on name, date of birth, address, and social security number.

For each case reported to the local HIV surveillance system, additional clinical and demographic data are available in the enhanced HIV/AIDS Reporting System (eHARS). We define TGW as people who were assigned male sex at birth but identify as women. In the LAC HIV surveillance database, transgender information was initially collected in a combined sex/gender field (male, female, male-to-female transgender, and female-to-male transgender) starting in the late 1990s. From 2009 onwards, a two-step method was implemented in the HIV/AIDS adult case-report form to identify transgender individuals, recording sex at birth alongside current gender identity. There are a variety of data sources for sex and gender information which may include provider reports (as abstracted from medical charts, physician's notes, and self-administered patient intake sheets), laboratory test reports, the Ryan White Program client registry, and public health investigation by surveillance and partner services staff. Other data available in eHARS include race/ethnicity (American Indian/Alaska Native, Asian/Pacific Islander, Black/African-American, Latino, White or multi-racial), transmission risk factor (MSM, people who inject drugs [PWID], MSM/PWID, heterosexual, perinatal, other, unknown), age at diagnosis, date of diagnosis, CD4 count at diagnosis, and date of last negative test. Age at diagnosis was treated as a categorical variable (0–12, 13–19, 20–29, 30–39, 40–49, 50–59, 60+) and date of diagnosis was analyzed as a continuous variable using month and year only. Where date of last negative test was available and less than 6 months before a positive HIV test, we classified individuals as "early" diagnoses. As a proxy for time since infection for other cases, we used CD4 count: >500, 200–500, <200 cells per μL.

A TGW who reports sex with cis-men may be classified as heterosexual (corresponding to their gender identity) or MSM (corresponding to their birth sex but disregarding their gender

identity). Therefore, for TGW, we collapsed transmission risk factor into two categories: TGW who reported injection drug use (TGW-PWID) and the remaining TGW who did not report injecting drugs and were likely to have been infected through sex. We classified this group as having a sexual risk factor (TGW-Sex). To permit meaningful comparison with cisgender males and females (individuals who identify with the sex they were assigned at birth), we categorized all cisgender individuals who reported injection drug use as PWID and those who reported perinatal exposure or other transmission risk factor as “other”. Individuals who reported heterosexual risk or no risk were classified as having sexual transmission risk. As such, the final risk categories differed from those assigned by HIV surveillance (see below).

The study was approved by both the University of California, San Diego and LAC Department of Public Health Institutional Review Boards.

Phylogenetic analyses

A molecular transmission network was constructed from genetic sequences using HIV-TRACE (13). In brief, HIV *pol* sequences were aligned to an HXB2 reference sequence and pairwise genetic distances were calculated under the Tamura-Nei 93 model. We did not remove codons associated with drug resistance as their removal has been demonstrated not to affect clustering using HIV-TRACE in similar datasets (14, 15). Each individual in the network is represented by a node, and nodes were linked to each other if their pairwise genetic distance was ≤0.015 substitutions/site. This threshold is in line with the expected divergence between sequences within an individual (16) and is in accordance with the genetic distance seen between named HIV risk partners (15). We further tested the sensitivity of our epidemiological inference at distance thresholds of ≤0.01 and ≤0.02 substitutions/site. Nodes linked to at least one other node are considered “clustered” in the transmission network. Ninety-seven percent of sequences were subtype B, but HIV-TRACE can create a single network regardless of subtype.

Statistical analyses

Clustered sequences are closely related genetically, indicating that they are likely to be part of the same transmission chain, and high rates of clustering within a population suggest increased rates of transmission. Therefore, we assessed the correlates of clustering using multivariate and univariate logistic regression. Date of HIV diagnosis, transmission risk group, age at diagnosis, race/ethnicity, CD4 count/early infection, and country of birth (U.S./U.S.-territories vs. foreign-born) were included as covariates in the multivariate regression models. Individuals for whom information was missing for one or more of these categories were categorized as “unknown”. For the purpose of the logistic regression, gender and transmission risk category were combined into a single variable. As such, our final transmission risk groups were: sexual risk cisgender females (F-Sex), cisgender female PWID (F-PWID), TGW-PWID, TGW-Sex, sexual risk cisgender males (M-Sex), MSM, MSM/PWID, cisgender male PWID (M-PWID), transgender men, and other (Supplementary Figure 1, Appendix p5). The “other” category comprised perinatal cases, blood product recipients and individuals for whom transmission risk could not be

ascertained. Using the same method and covariates, we then assessed the correlates of non-TGW cases clustering with TGW in the transmission network.

Assortativity is a network metric which describes, for a given characteristic (e.g., transmission risk factor) the tendency for nodes to link to other nodes with the same trait (i.e., do PWIDs link to PWIDs?) (17). Assortativity varies between -1 (completely disassortative) and 1 (completely assortative) and was calculated using the function available in the R igraph package v1.2.1 (18). F-Sex and M-Sex, and F-PWID and M-PWID, were combined into cisgender males and females with sexual risk and cisgender males and females who report injecting drugs, respectively, for this analysis, as we would expect them to mix with each other.

In parallel, we counted links in the network between each pair of transmission risk groups to estimate mixing patterns between TGW and other groups (17). In order to adjust for degree (i.e., the total number of links connecting a given node), the number of links for each individual was divided by that individual's degree. This correction was performed because some individuals have far more links than others but we do not wish to over count those individuals' contribution to mixing between transmission risk groups. Assortativity is influenced by the ratio of node labels (PWID, MSM etc.). To assess the statistical significance of observed patterns of mixing and assortativity given the relative representation of each transmission risk group in these clusters, we generated expected distributions for parameters by randomly permuting transmission risk group labels on the static network 1,000 times in R v3.4.1.

RESULTS

In the LAC transmission network, 8,133 of 22,398 (36.3%) unique individuals were clustered at 0.015 substitutions/site. The network was composed of 1,722 clusters comprising between 2 and 116 nodes (Figure 1, Supplementary Figures 2 & 3, Appendix p6). The majority of the sequences in the surveillance database (14,932/22,398; 67%) were from MSM, and this proportion was even greater among clustered sequences (5993/8133; 73.7%). The LAC dataset contained sequences and demographic data from 412 TGW, including 67 TGW-PWID. The mean age of TGW at diagnosis was 29 years and their average current age was 50. TGW were less likely to be White than other cases in the dataset (Fisher's exact test; $p<0.001$; Figure 2). The number of sequences collected and the proportion of sequences clustering each year have increased overall, and for TGW-Sex specifically, but for not TGW-PWID, as diagnoses among PWID have decreased over time in LAC (2) (Supplementary Figure 4, Appendix p8).

We sought to determine which demographic/risk characteristics were associated with clustering in order to identify subpopulations with higher rates of transmission. TGW-Sex clustered at the highest frequency in the network (42.6%, compared with 40.1% for MSM) (Figure 3) and had the highest odds of clustering in the univariate analyses ($p<0.001$; Supplementary Table 1). In the multivariate analysis, the adjusted odds ratio (AOR) for clustering was even higher for TGW-PWID than for TGW-Sex and MSM (Figure 3). However, the AOR for clustering of TGW-PWID were affected by the date of the HIV

diagnosis, with 90% of TGW-PWID diagnoses having taken place before 2007 (Supplementary Figure 4, Appendix p8), and consequently their odds of clustering were lower than TGW-Sex in the univariate analysis (Supplementary Table 1, Appendix p1). Individuals diagnosed with a higher CD4 count, likely to have been diagnosed closer to the time of infection, were more likely to cluster but the effect was modest. Individuals with a documented negative HIV-test within 6 months prior to diagnosis, classified as “early”, were more likely to cluster in the univariate analysis, but this effect was not significant in the multivariate model (Supplementary Table 1, Appendix p1, Figure 3). An age trend was apparent, with younger individuals significantly more likely to cluster and older individuals significantly less likely to cluster (Figure 3). Individuals of Latino ethnicity were the largest racial/ethnic group (44.2% of population) and the group most likely to cluster. Individuals born outside the U.S. were less likely to cluster than those born in the U.S./US-Territories. Variables associated with clustering were consistent across genetic distance thresholds (Supplementary Table 2, Appendix p2).

We estimated assortativity, the tendency of nodes sharing attributes to link together, by transmission risk group, across the network. The 167 TGW were distributed across 126 clusters, with 21 clusters containing >1 TGW. Whereas only 503/21,986 (2.3%) of non-TGW individuals linked to at least one TGW in the network, 106/503 (21.1%) of those linked to a second TGW (Figure 1). Therefore, individuals linked to one TGW were 9.2 times more likely to link to two TGW than other individuals in the surveillance database. MSM, MSM/PWID, cisgenders with a sexual risk, TGW-Sex and TGW-PWID were all significantly assortative in the network (Figure 4). MSM were most likely to link to each other (Assortativity coefficient = 0.17; $p<0.001$). The assortativity coefficient for TGW with sexual transmission risk was 0.06 ($p<0.001$; i.e., an assortativity coefficient this extreme was not observed in any of the 1000 network permutations); however, absolute assortativity of TGW was low relative to cisgenders with a sexual risk and MSM, because the total number of TGW in the network is small. In contrast, cisgender PWID did not link assortatively, indicating that they were dispersed among other risk groups in the network. At genetic distance thresholds of ≤0.01 and ≤0.02 substitutions per site, MSM, cisgenders with a sexual risk and TGW-Sex remained highly significantly assortative (Supplementary Figure 5, Appendix p9).

We characterized the subpopulations clustering with TGW by constructing a linear regression model distinguishing between non-TGW clustering with TGW and those that did not. M-Sex, MSM/PWID, and M-PWID were all more likely than MSM to be clustered with TGW; and foreign-born individuals were less likely to cluster with TGW than U.S. born individuals (Table 1). There were no significant differences by age or race/ethnicity.

Finally, we explored the connectivity between each pair of transmission risk groups, adjusting for node degree, to determine who TGW linked to. The network reconstruction method creates a network in which far more links are present than in the true transmission network. Because the vast majority of the nodes in the network represent MSM, we expect high linkage to MSM for all transmission risk groups, and that is indeed what we observed (Supplementary Table 4). Nonetheless, we also observed trends towards TGW-Sex and TGW-PWID linking to each other as well as with M-Sex. To assess the statistical

significance of this observation, we estimated the expected proportion of links between TGW-Sex and TGW-PWID and each of the other risk groups using the randomly permuted networks (Table 2). For both TGW-Sex and TGW-PWID, the proportion of links to TGW-Sex and TGW-PWID was higher than expected, whereas the proportion of links to MSM was 25–30% lower than expected. For TGW-Sex, the proportion of links with M-Sex was higher than expected. Nonetheless the majority of TGW links (75%) were with MSM. Identical mixing patterns were seen across genetic distance thresholds (See Supplementary Tables 3, 4 and 5, Appendix p3, which include an extended version of Table 2, displaying mixing between every pair of transmission risk groups).

DISCUSSION

We found that TGW had the highest odds of clustering in the LAC network, indicating that their risk of being in a molecular transmission cluster exceeds that of even MSM. Our findings also reveal that TGW occupy a distinct position in the LAC transmission network. TGW were significantly more likely to be the genetically-linked partners of cisgender males not reporting injection drug use or MSM contact (M-Sex) than expected. Furthermore, TGW tended to cluster assortatively with other TGW in the network (i.e., having one TGW in a cluster increased the odds of finding another TGW in that same cluster).

The patterns of clustering among TGW observed here suggest a potentially powerful strategy for using the molecular transmission network to improve public health outcomes. Assortativity among TGW indicates that non-TGW who are genetically linked to a TGW are nine times more likely to be clustered with a second TGW. Based on this finding, we propose that non-TGW with a genetic link to a TGW may be more likely to identify additional HIV-infected or at-risk TGW via partner services, than non-TGW who are not genetically linked to TGW. At present, in LAC and in much of the U.S., partner elicitation services are not universally offered. Further, typically less than half of interviews result in the identification of a partner (19). Molecular epidemiology could be used to prioritize these genetically linked non-TGW cases for partner elicitation interviews by public health investigators, with the expectation of identifying more HIV-infected, undiagnosed TGW, high-risk HIV-uninfected TGW, or HIV-infected TGW who are not in care. This targeted approach could lead to improved HIV diagnosis, linkage to HIV care, and pre-exposure prophylaxis (PrEP) access among TGW.

Clearly, the named sexual partners of TGW should also continue to be considered a high priority group for HIV research and interventions. In interviews, 20% of the male partners of TGW have reported being HIV-positive (20, 21). Although TGW think of themselves as a distinct community, their non-TGW partners may not, which makes them more difficult to identify (20). Molecular epidemiology represents a tool to identify this high-risk population. Importantly, if validated, this type of network-targeted approach would be applicable to any group that clusters assortatively in a molecular transmission network. Nonetheless, we acknowledge that there is a difference between the individuals named during a partner services interview and those individuals with a genetic link (15).

Genetic clustering methods have been rightly criticized for potential bias towards identifying subpopulations with higher sampling rates rather than higher transmission rates (22–24). Consistently high frequencies of clustering among MSM (14, 25, 26) and individuals diagnosed with acute/early stage infection (23, 27) may reflect elevated diagnosis rates rather than exceptional transmission rates; thus, clustering analyses could potentially divert public health focus from where it is most needed (22). That being said, high clustering has been consistent with shorter transmission intervals in time-resolved analyses (8), and the algorithm used here has demonstrated ability to detect subpopulations with higher transmission rates in simulations (22). Strikingly, in our analysis, the highest clustering rates were seen among TGW, a group documented to have low diagnosis rates (4, 6, 28), suggesting that in this instance, a genetic clustering approach works well as a tool to identify a hidden high risk population in the absence of increased sampling rates. However, the proportion of people recently diagnosed with HIV-1 in LAC who have a reported *pol* sequence is only between 40–50%, suggesting the potential for sampling bias. Further, TGW are more likely to engage in care after HIV diagnosis (29), increasing the likelihood of having an HIV sequence in the surveillance database. To address this potential bias, we used CD4 count at diagnosis as a proxy for time since infection in our multivariate regression. Although a higher CD4 count (suggesting a shorter time between infection and diagnosis) was indeed associated with clustering, the effect was weak and our main finding was robust including this covariate. Furthermore, all genetic network analyses, such as this one, are limited because they geographically constricted and affected by sampling, and we cannot account for migration or transmission events occurring outside of LAC. Nonetheless, as we found that individuals from outside the U.S. were less likely to cluster than those from within the U.S., this migration should not bias our results.

Importantly, a limitation of our clustering analysis is that HIV-TRACE does not infer directionality, and we cannot distinguish between transmitters and recipients in our clusters. However, our inference is not unduly influenced by this limitation, because identifying genetically linked partners is sufficient for deciding whether to prioritize individuals for public health interventions. We find that TGW are more likely to be involved in HIV transmission events, but we cannot state whether they more frequently the transmitter or recipient. This finding highlights the importance of allocating public health and other services towards the HIV-infected and at-risk transgender community.

Our finding that TGW link preferentially to M-Sex (who will be composed mainly of male heterosexuals) is particularly meaningful given that MSM have far higher HIV prevalence than male heterosexuals and are expected to be the source of the majority of infections. This finding is in agreement with interviews of TGW (4) and their partners (20). In a study of male partners of TGW in San Francisco, half the TGW described themselves as straight, and only 10% identified as gay (20). Although the genetic transmission network alone does not conclusively reveal source of infection for TGW (sex with M-Sex, sex with MSM, or shared needles), traits-based phylogenetic analysis on these clusters may further elucidate transmission risk for TGW. Nonetheless, reliance on self-reporting of transmission risk can be influenced by MSM who do not disclose their risk factors (11, 12). Reliable estimates of TGW and diagnosis rates in U.S. populations are unfortunately lacking, but would be helpful for assessing the impact of public health services provided to TGW and their partners.

In conclusion, we report that TGW in LAC were more likely to cluster in a molecular transmission network than other risk groups, suggesting high transmission rates—despite low representation of TGW in the database. TGW were genetically linked to M-Sex more than expected and to MSM less than expected. TGW tended to be part of the same clusters, indicating linkage either directly or through shared partners. This assortativity highlights the potential to use molecular epidemiology to both identify transmission clusters likely to include undiagnosed or undisclosed HIV-infected TGW and improve public health prevention and treatment activities towards TGW.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We wish to thank Brooke Sullivan for providing feedback on the manuscript, Ying Ou for his assistance in collecting, processing, and managing HIV-1 sequence data, and Reilly Hostager for help in preparing tables and figures.

FUNDING

This work was supported by a California HIV/AIDS Research Program IDEA Award (ID15-SD-052), an NIH-NIAID K01 Career Development Award (K01AI110181), and an NIH-NIAID R01 (AI135992). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Richardson ET, Collins SE, Kung T, Jones JH, Hoan Tram K, Boggiano VL, et al. Gender inequality and HIV transmission: a global analysis. J Int AIDS Soc. 2014;17:19035. [PubMed: 24976436]
2. Division of HIV and STD Programs Los Angeles County Department of Public Health. 2015 Annual HIV/STD Surveillance Report. 2018.
3. Centers for Disease Control and Prevention. Diagnoses of HIV Infection in the United States and Dependent Areas, 2016. 2017.
4. Herbst JH, Jacobs ED, Finlayson TJ, McKleroy VS, Neumann MS, Crepaz N, et al. Estimating HIV prevalence and risk behaviors of transgender persons in the United States: a systematic review. AIDS Behav. 2008;12(1):1–17. [PubMed: 17694429]
5. Centers for Disease Control and Prevention. HIV prevalence, unrecognized infection, and HIV testing among men who have sex with men--five U.S. cities, June 2004-April 2005. MMWR Morb Mortal Wkly Rep. 2005;54(24):597–601. [PubMed: 15973239]
6. Schulden JD, Song B, Barros A, Mares-DelGrasso A, Martin CW, Ramirez R, et al. Rapid HIV testing in transgender communities by community-based organizations in three cities. Public Health Rep. 2008; 123 Suppl 3:101–14. [PubMed: 19166094]
7. Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303(5656):327–32. [PubMed: 14726583]
8. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ. Episodic sexual transmission of HIV revealed by molecular phylodynamics. PLoS Med. 2008;5(3):e50. [PubMed: 18351795]
9. Poon AF, Gustafson R, Daly P, Zerr L, Demlow SE, Wong J, et al. Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: an implementation case study. Lancet HIV. 2016;3(5):e231–8. [PubMed: 27126490]
10. Smith DM, May SJ, Tweeten S, Drumright L, Pacold ME, Kosakovsky Pond SL, et al. A public health model for the molecular surveillance of HIV transmission in San Diego, California. AIDS. 2009;23(2):225–32. [PubMed: 19098493]

11. Hue S, Brown AE, Ragonnet-Cronin M, Lycett SJ, Dunn DT, Fearnhill E, et al. Phylogenetic analyses reveal HIV-1 infections between men misclassified as heterosexual transmissions. Aids. 2014;28(13):1967–75. [PubMed: 24991999]
12. Ragonnet-Cronin M, Hue S, Hodcroft EB, Tostevin A, Dunn D, Fawcett T, et al. Nondisclosed MSM in UK HIV transmission networks: phylogenetic analysis of surveillance data. Lancet HIV. 2018.
13. 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.
14. 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(4):444–51. [PubMed: 26302431]
15. Wertheim JO, Kosakovsky Pond SL, Forgione LA, Mehta SR, Murrell B, Shah S, et al. Social and Genetic Networks of HIV-1 Transmission in New York City. PLoS Pathog. 2017;13(1):e1006000. [PubMed: 28068413]
16. Hue 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):719–28. [PubMed: 15075506]
17. Newman ME. Mixing patterns in networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2003;67(2 Pt 2):026126.
18. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006;Complex Systems:1695.
19. Katz DA, Hogben M, Dooley SW Jr., Golden MR. Increasing public health partner services for human immunodeficiency virus: results of a second national survey. Sex Transm Dis. 2010;37(8): 469–75. [PubMed: 20661113]
20. Operario D, Burton J, Underhill K, Sevelius J. Men who have sex with transgender women: challenges to category-based HIV prevention. AIDS Behav. 2008;12(1):18–26. [PubMed: 17705095]
21. Operario D, Nemoto T, Iwamoto M, Moore T. Risk for HIV and unprotected sexual behavior in male primary partners of transgender women. Arch Sex Behav. 2011;40(6):1255–61. [PubMed: 21604064]
22. Poon AF. Impacts and shortcomings of genetic clustering methods for infectious disease outbreaks. Virus Evol. 2016;2(2):vew031.
23. 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]
24. Dearlove BL, Xiang F, Frost SDW. Biased phylodynamic inferences from analysing clusters of viral sequences. Virus Evol. 2017;3(2):vex020.
25. Ragonnet-Cronin M, Lycett SJ, Hodcroft EB, Hue S, Fearnhill E, Brown AE, et al. Transmission of Non-B HIV Subtypes in the United Kingdom Is Increasingly Driven by Large Non-Heterosexual Transmission Clusters. Journal of Infectious Diseases. 2016;213(9):1410–8. [PubMed: 26704616]
26. Kouyos RD, von Wyl V, Yerly S, Boni J, Taffe P, Shah C, et al. Molecular epidemiology reveals long-term changes in HIV type 1 subtype B transmission in Switzerland. J Infect Dis. 2010;201(10):1488–97. [PubMed: 20384495]
27. Pao D, Fisher M, Hue S, Dean G, Murphy G, Cane PA, et al. Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. AIDS. 2005;19(1):85–90. [PubMed: 15627037]
28. Garofalo R, Kuhns LM, Reisner SL, Mimiaga MJ. Behavioral Interventions to Prevent HIV Transmission and Acquisition for Transgender Women: A Critical Review. J Acquir Immune Defic Syndr. 2016;72 Suppl 3:S220–5. [PubMed: 27429186]
29. Division of HIV and STD Programs Los Angeles County Department of Public Health. HIV Surveillance Reports: 2014 Los Angeles County HIV Cascades and PLWH Estimate 2016.

Evidence before this study

We searched Google scholar for the terms “HIV” “genetic”, “transmission” “networks” on 4th June 2018, then added the term “transgender” to the search, with no date limits. HIV is spread through contacts within a sexual network. The virus accumulates genetic mutations within the same timeframe as transmission events. The transmission history of the virus is a subset of the network that can be reconstructed from HIV genetic sequences (although some transmission events may be missed for various reasons). The structure of the reconstructed transmission network can be informative in terms of risk factors associated with transmission and to inform interventions. Furthermore, several studies have demonstrated that it is possible to gain insights into transmission of HIV among groups difficult to investigate using traditional epidemiological tools such as contact tracing. We found no molecular epidemiological analyses specific to transgender women, despite them being one of the groups with the highest prevalence of HIV in the United States.

Added value of this study

We reconstructed the HIV transmission network using all HIV sequences available from Los Angeles County. We identified transgender women within these networks and looked at how they were connected to other risk groups in the network. We found that transgender women were more connected to each other and to heterosexual men, and less connected to men who have sex with men, than expected.

Implications of all the available evidence

The way in which people are connected through the genetic transmission network provides information on transmission patterns within the population. Transmission clusters comprising at least one transgender woman are attractive targets for interventions aimed at finding additional undiagnosed and at-risk transgender women, because individuals within that cluster are more likely to have other transgender women among their sexual/ social contact networks. This study highlights the potential for molecular epidemiology to guide interventions towards subpopulations with high HIV prevalence but low diagnosis rates.

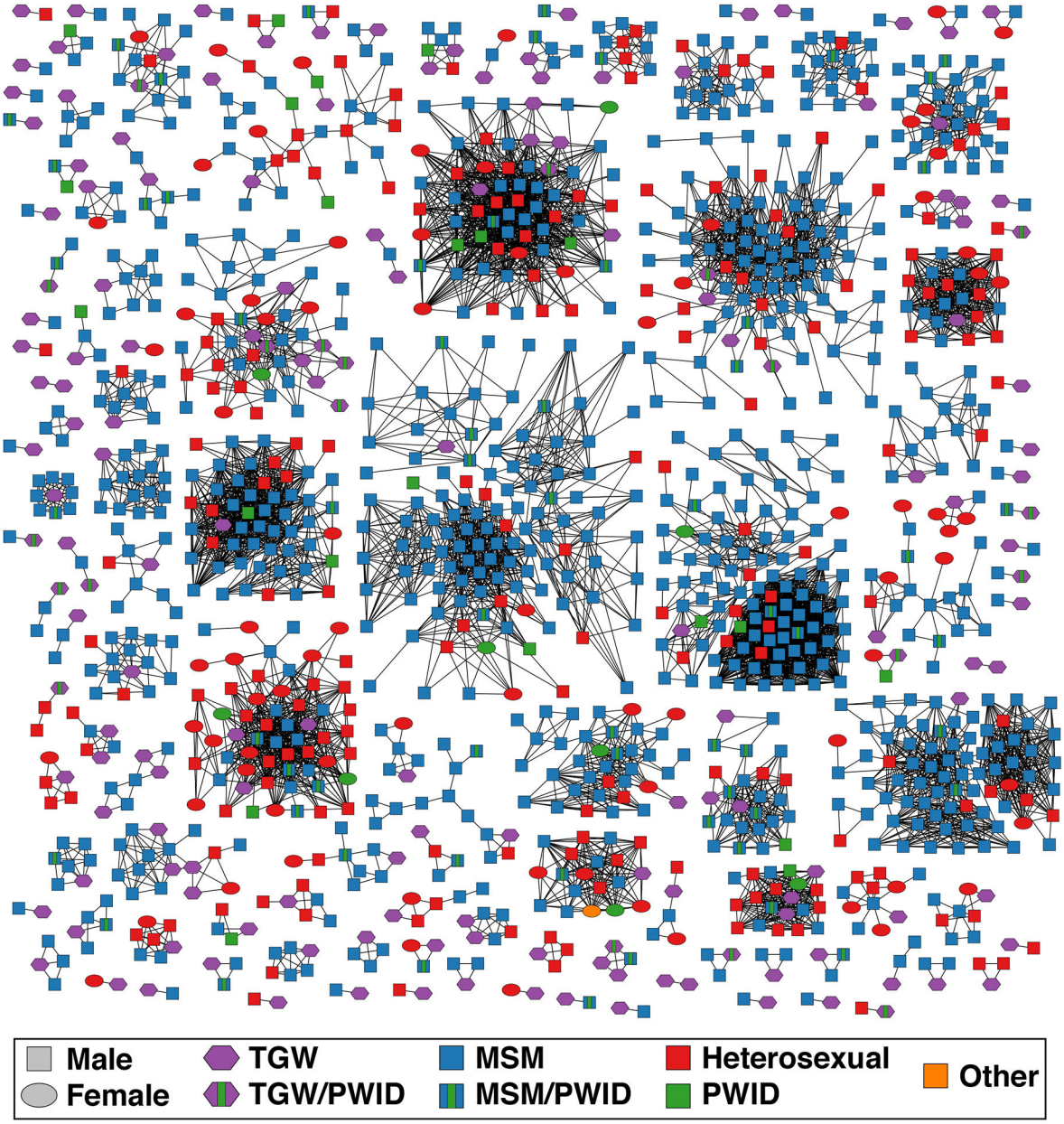

Figure 1:

Molecular transmission clusters in Los Angeles County with at least one transgender woman (TGW). Node shape denotes gender and color denotes transmission risk factor. Edges represent genetic distance of ≤0.015 substitutions/site. Sex, sexual risk; PWID, people who inject drugs; TGW, transgender women; MSM, men who have sex with men.

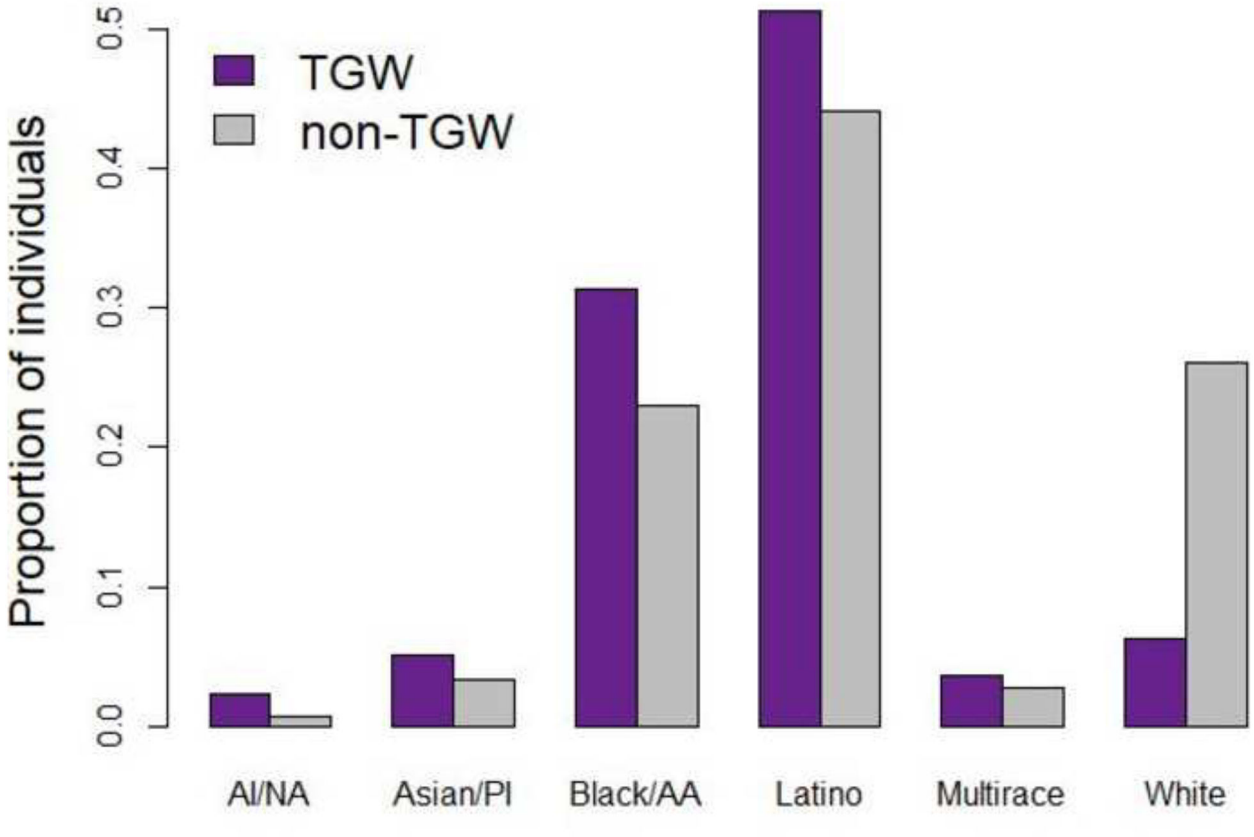

Figure 2:

Race/ethnicity of transgender women (TGW) and other individuals with sequence data available in the Los Angeles County dataset. There were 412 individuals in the TGW group compared to 21,986 non-TGW. AI/NA, American Indian/Native Alaskan; PI, Pacific Islander; AA, African American. TGW were less likely to be white than other cases in the dataset (Fisher's test, $p<0.001$).

	Total	Clustered	AOR	*p*
Diagnosis date	22,398	8,133 (36.3%)	1.18***	<0.0001
Stage/ CD4 count				
Early	226	165 (73%)	1.34	0.11
CD4 >500	6,317	2,753 (43.6%)	1.12*	0.01
CD4 200-500	8,520	3,546 (41.6%)	1	REF
CD4 <200	7,200	1,598 (22.2%)	0.57***	<0.0001
Transmission risk				
F-PWID	328	61 (18.6%)	1.07	0.7184
F-Sex	1979	480 (24.3%)	1	REF
M-PWID	453	96 (21.2%)	1.16	0.3468
M-Sex	2647	949 (35.9%)	1.37***	0.0002
MSM	14932	5993 (40.1%)	1.99***	<0.0001
MSM/PWID	1448	96 (6.6%)	1.81***	<0.0001
Other	182	16 (8.8%)	2.42	0.2912
TGM	17	3 (17.6%)	0.62	0.5
TGW-PWID	67	20 (29.9%)	2.61**	0.0053
TGW-Sex	345	147 (42.6%)	2.31***	<0.0001
Age category				
<13	169	15 (8.9%)	0.51	0.44
13-19	963	444 (46.1%)	2.06***	<0.0001
20-24	3,763	1,774 (47.1%)	1.68***	<0.0001
25-29	4,486	1,807 (40.3%)	1.39***	<0.0001
30-39	7,442	2,312 (31.1%)	1	REF
40-49	3,945	1,233 (31.3%)	0.77***	<0.0001
50-59	1,350	454 (33.6%)	0.66***	<0.0001
60+	280	94 (33.6%)	0.64**	0.01
Race/ethnicity				
Asian/Pacific Islander	749	308 (41.1%)	0.75**	0.004
Black/African American	5,187	1,607 (31%)	0.62***	<0.0001
Latino	9,911	4,074 (41.1%)	1	REF
Mixed race	628	217 (34.6%)	0.66***	0.0007
Native American	166	53 (31.9%)	0.70	0.1234
White	5,748	1,868 (32.5%)	0.77***	<0.0001
Country of birth				
USA/Territories	12,666	4,470 (35.3%)	1	REF
Foreign-born	5,644	1,955 (34.6%)	0.81***	<0.0001

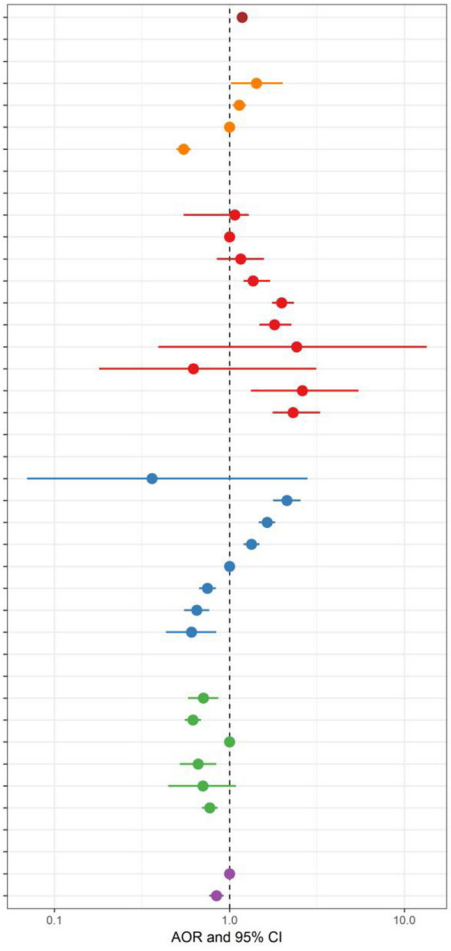

Figure 3:

Demographic breakdown of the persons reported with HIV-1 sequence data in LAC with adjusted odds ratio (AOR) for clustering. The “total” column indicates the number of individuals in the LAC surveillance population in that category and the “clustered” column indicates the number and percentage of individuals in that category who were clustered. The AOR for diagnosis date indicates that individuals diagnosed in each year were 1.18 times more likely to be clustered than individuals sampled in the previous year. Individuals classified as “Early” are those who tested negative for HIV within 6 months before diagnosis. F-PWID, cisgender female person who injects drugs; F-Sex, female with sexual risk; M-PWID, cisgender male person who injects drugs; M-Sex, cisgender male with sexual risk; MSM men who have sex with men; MSM/PWID men who have sex with men and inject drugs; TGM, transgender men; TGW-PWID, transgender women who inject drugs; TGW-Sex, transgender women with sexual risk. * indicates $p<0.05$, ** $p<0.01$, and *** $p<0.001$.

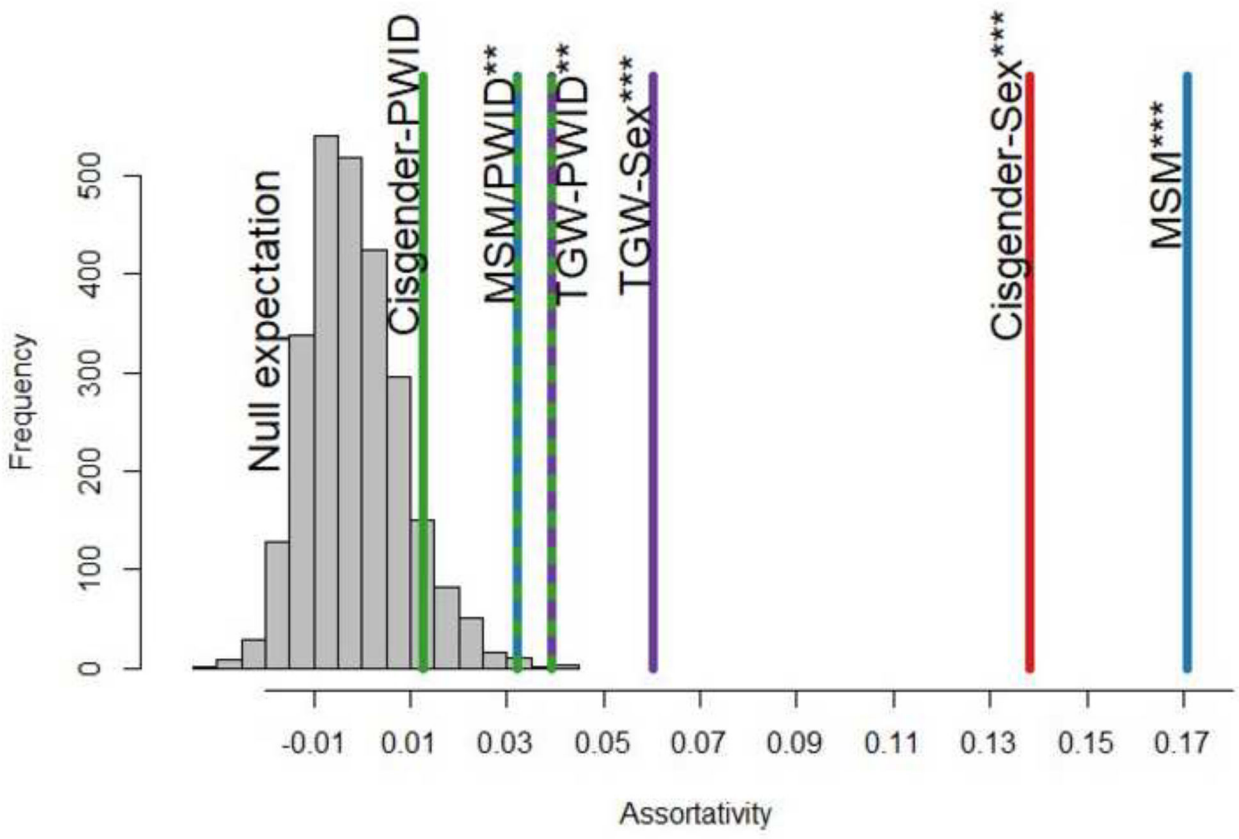

Figure 4:

Assortativity broken down by self-reported risk group. The null distribution of expected assortativity is shown in grey and the observed assortativity for each risk group is displayed in a different color. MSM, men who have sex with men; Sex, sexual risk; TGW-Sex, transgender women with sexual risk; PWID, people who inject drugs; TGW-PWID, transgender women who inject drugs. Significant assortativity is denoted by **p<0.01, ***p<0.001.

Table 1:

Correlates of clustering with transgender women

	AOR	95% CI
Diagnosis year	1.03***	1.01 – 1.04
RISK		
MSM	-	
Cisgender male Sex	1.74***	1.45 – 2.07
MSM/PWID	1.39*	1.03 – 1.85
Cisgender male PWID	1.79*	1.07 – 2.86
BIRTH COUNTRY		
U.S./ U.S. Territories	-	
Foreign-born	0.81*	0.67 – 0.97

Only variables with significant association in the multivariate regression model are shown. MSM, men who have sex with men; Sex, sexual risk; PWID, people who inject drugs; AOR, adjusted odds ratio; CI, confidence interval;

* p<0.05,

*** p<0.001.

Table 2:

Ratio of the observed proportion of pairwise links compared to the mean of the simulated proportion of pairwise links (and 95% confidence intervals). Ratios above 1 indicate an overrepresentation of those relationships in the true network compared to random expectation and ratios below 1 indicate an underrepresentation of those relationships.

	F-Sex	F-PWID	TGW-PWID	TGW-Sex	M-Sex	MSM	MSM/PWID	M-PWID
TGW-Sex	1.09 (0.81–1.62)	0.73 (0.25-Inf)	6.55*** (1.5-Inf)	4.65*** (2.08-Inf)	1.53** (1.21–2.01)	0.75** (0.71–0.8)	1.68** (1.11–2.6)	1.82* (1.05-Inf)
TGW-PWID	0.72 (0.37–4.4)	0 (0-NaN)	11.99*** (2.2-Inf)	6.9*** (1.97-Inf)	1.38 (0.78–5.2)	0.69** (0.59–0.85)	2.43* (1.07–9.6)	2.65 (0.48-Inf)

F, cisgender female; PWID, people who inject drugs; TGW, transgender women; Sex, sexual risk; M, cisgender male; MSM, men who have sex with men.

* p<0.05,

** p<0.01,

*** p<0.001.