

The Central Role of Nondisclosed Men Who Have Sex With Men in Human Immunodeficiency Virus-1 Transmission Networks in Guangzhou, China

Huanchang Yan,^{1,a} Weiyun He,^{2,a} Liping Huang,¹ Hao Wu,² Yuanhao Liang,¹ Qingmei Li,² Jingwei Shui,¹ Cheng Wang,³ Emmanuel E. Dzakah,^{3,4} Zhigang Han,² and Shixing Tang^{1,3,6}

¹Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou, China, ²Guangzhou Center for Disease Control and Prevention, Guangzhou, China, ³Dermatology Hospital, Southern Medical University, Guangzhou, China, ⁴Department of Molecular Biology and Biotechnology, School of Biological Sciences, University of Cape Coast, Cape Coast, Ghana

Background. Men who have sex with men (MSM) are vulnerable risk group for human immunodeficiency virus (HIV)-1 infection. However, some MSM do not disclose their same-sex behavior and could impact the transmission and prevention of HIV-1 infection. Here, we evaluated the role of nondisclosed MSM in HIV-1 transmission in Guangzhou, China.

Methods. The HIV-1 *pol* sequences were obtained from HIV-infected subjects from 2008 to 2015. A transmission network was constructed using HIV TRANSMISSION Cluster Engine (HIV-TRACE) at a pairwise genetic distance of 0.5%. The position of nondisclosed MSM in the network was determined by centrality analysis.

Results. Nondisclosed MSM were inferred in 9.92% (61 of 615) of slightly older, self-reported non-MSM ($P = .006$). They were more likely to be married ($P = .002$) and less educated ($P < .001$) than the MSM with whom they clustered. Closeness centrality was bigger for nondisclosed MSM than for MSM ($P < .001$), indicating the central position of nondisclosed MSM in the networks. The average shortest path length was smaller for nondisclosed MSM than for MSM ($P < .001$), whereas radiality was bigger for nondisclosed MSM than for MSM, suggesting a relatively greater contribution of nondisclosed MSM in transmitting HIV-1 than MSM. Assortativity analysis indicated that nondisclosed MSM were more likely to link each other with coefficient of 0.025.

Conclusions. Nondisclosed MSM are a specific group, and they play an important role in HIV-1 transmission. They could be bisexual and might increase the risk of HIV-1 infection to their sex partners. Therefore, specific prevention and intervention targeting nondisclosed MSM are urgently needed.

Keywords. cluster analysis; HIV-1; men who have sex with men (MSM); nondisclosed MSM; transmission network.

Men who have sex with men (MSM) are at increasing risk for human immunodeficiency virus type 1 (HIV-1) infection in many countries and regions, although both HIV-1 prevalence and incidence are declining in the general population worldwide [1]. Human immunodeficiency virus-infected MSM show high HIV-1 burdens and substantial clustering of infections within transmission networks [2]. The MSM population has become one of the major risk groups for HIV-1 infection in China where there are approximately 20 million MSM. For Chinese MSM, the prevalence of HIV-1 infection has rapidly increased from 0.9% in 2003 to 7.8% in 2016, and, notably, 27.6% of the

124 000 newly diagnosed HIV-1 infections in 2016 were attributed to MSM transmission [3].

Due to social isolation, discrimination, and verbal or physical abuse, some MSM may not disclose their sexual orientation [4], particularly in low- and middle-income countries (LMICs), which in turn may put them at particularly high risk for HIV-1 transmission because of their identity of nondisclosed MSM and the lack of HIV-1 prevention services that are usually available to MSM. Furthermore, it is difficult to characterize HIV-1 transmission in nondisclosers through traditional “shoe-leather” epidemiologic reviews. Hence, it is critical to identify the specific group of MSM who do not disclose their homosexual orientation and better understand their needs for HIV-1 prevention.

In 2018, Ragonnet-Cronin et al [5] reported self-reported heterosexual men who clustered only with MSM and with no women as potential nondisclosed MSM. They adapted time-resolved phylogenetic analysis of HIV-1 sequences to create transmission networks and found that 18.6% of all clustered heterosexual men were actually nondisclosed MSM in the United Kingdom (UK). In addition, they found that nondisclosed MSM were more likely to link MSM and

Received 30 October 2019; editorial decision 27 April 2020; accepted 30 April 2020.

^aH. Y. and W. H. contributed equally to this manuscript.

Correspondence: Shixing Tang, MD, PhD, Department of Epidemiology, School of Public Health, Southern Medical University, 1838 Guangzhou Avenue, N., Guangzhou 510515, China (tangshixing@smu.edu.cn).

Open Forum Infectious Diseases®

© The Author(s) 2020. Published by Oxford University Press on behalf of Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
DOI: 10.1093/ofid/ofaa154

heterosexual women, and they might increase the risk of HIV-1 infection in their female partners. This group of nondisclosed MSM was only revealed in the UK since 2014 and has not yet been fully investigated [6].

In LMICs, few research studies on MSM sexual orientation disclosure have been reported. A cross-sectional survey in China found that more than 80% MSM did not disclose their sexual orientation to healthcare professionals [7]. In the UK, nondisclosed MSM are more likely to be in peripheral position in the transmission networks. They usually chose nondisclosed MSM as sexual partners and might exhibit lower-risk sexual behavior than MSM [5]. However, the characteristics of nondisclosed MSM and their role in HIV-1 transmission remain to be elucidated in other countries.

In this study, we adapted HIV TRANSMISSION Cluster Engine (HIV-TRACE), a recently developed tool for rapid and automated characterization of transmission networks from large sets of HIV-1 genetic sequences [8], and used HIV-1 sequence data from HIV-positive subjects diagnosed in Guangzhou city between 2008 and 2015, to locate nondisclosed MSM in HIV-1 transmission network and to explore their characteristics and potential role in HIV-1 epidemics.

METHODS

Study Design

The study included 3079 HIV-1 *pol* sequences (HXB2 position 2253–3821) obtained from 3684 HIV-positive subjects who were diagnosed and claimed to be infected through heterosexual transmission (HET), MSM, or intravenous drug users (IDUs) from 2008 to 2015 in Guangzhou, China. Demographic characteristics (age, gender, marital status, and education) and risk groups (HET, MSM, or IDUs) of HIV-1 infection were extracted from the electronic follow-up records at Guangzhou Center for Disease Control and Prevention (Guangzhou CDC). Sequences were aligned with the representative sequences of HIV-1 group M subtypes and circulating recombinant forms (CRFs) obtained from the Los Alamos database (<http://www.hiv.lanl.gov>) using MAFFT [9]. The HIV-1 subtypes were determined by the phylogenetic tree constructed by IQ-TREE 1.6.9 [10] using an ultrafast bootstrap with 1000 replicates, as well as the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT) with 1000 replicates. Human immunodeficiency virus-1 CRF genotyping was done by using the RIP online subtyping tool [11]. The HIV-1 *pol* sequences described here are available under the following GenBank accession numbers: MN424584–427369.

In addition, a background sequence dataset of HIV-1 subtype B, CRF01_AE, CRF07_BC, and CRF55_01B was compiled using HIV BLAST from the Los Alamos National Laboratory (LANL) database (June 2019). To limit the size of alignments, the 10 most similar reference sequences for each Guangzhou

sequence were selected. Sequences with no information of gender or risk groups of HIV-1 infection were excluded.

This study was approved by the Institutional Review Board of the Guangzhou CDC (No. 2017030). Written informed consent was obtained from the participants.

Transmission Network Analysis

The Tamura-Nei 93 nucleotide substitution model was used to calculate pairwise genetic distance for all the sequences analyzed. The Transmission network was constructed using HIV-TRACE with a threshold of pairwise genetic distance of 0.5%, which is more appropriate for identifying rapidly growing clusters [12]. In brief, each sequence is represented by a node in the network, and nodes are linked together if their pairwise genetic distance is 0.5% or less. Network centrality was measured by using NetworkAnalyzer 2.7 implemented in Cytoscape 3.7.0 [13], and network assortativity was calculated by using NetworkX 1.9.1. package in Python 3.7.3.

The position within the transmission networks was determined based on centrality measurements including average shortest path length, closeness centrality, radiality, and degree to quantify how important a node is in viral transmission. Among them, the average shortest path length represents the distance between 2 connected nodes [14]. Closeness centrality usually measures the speed that the virus spreads from a given node to other reachable nodes in the network and is defined as the reciprocal of the average shortest path length of a node [15]. Radiality, a node centrality index, is obtained by subtracting the average shortest path length of a node from the diameter of the connected component plus 1 [16]. Another measurement of the contribution of transmission within a network is the number of edges linked to a node, also called the node degree [17]. Network assortativity was adapted to determine the transmission links, ie, whether nondisclosed MSM link to each other rather than MSM in the clusters. The value of assortativity ranges from -1 to 1 , which means completely disassortative and definitely assortative, respectively [18].

To reduce the possible bias due to the size of the clusters, we also calculated the centrality indicators after removing the clusters of $n = 2$ and conducted centrality analysis among the group with an equal number of inferred nondisclosed MSM and MSM, in which MSM were randomly chosen to match the number of nondisclosed MSM.

Statistical Analysis

We used a multivariate logistic regression model with the outcome being either clustered in the network (1) or not (0). Age, gender, marital status, education level, risk group of HIV-1 infection, and HIV-1 subtype or genotype were included as predictor variables. The difference between nondisclosed MSM and MSM was determined by using 2-sample *t* test and Mann-Whitney *U* test in IBM SPSS Statistics 25. In addition, a logistic

regression model was adapted to predict the variables associated with the outcome being inferred as nondisclosed MSM. The independent role of individual factor was evaluated by using univariate logistic regression followed by the evaluation of the interaction of these factors in the multivariate logistic regression. The factors with $P > .1$ were removed from the multivariate model by using backward stepwise analysis.

RESULTS

Characteristics of Human Immunodeficiency Virus-1 Sequences Analyzed

A total of 3079 HIV-1 *pol* sequences were obtained from 3684 HIV-positive people who were diagnosed in Guangzhou between 2008 and 2015. We selected 2786 sequences, each representing a distinct HIV-positive individual from the 4 most common HIV-1 genotypes in China: subtype B (174, 5.65%), CRF01_AE (1093, 35.50%), CRF07_BC (1155, 37.48%), and CRF55_01B (364, 11.82%) in this study (Figure 1, Table 1). They were infected through MSM (1977, 71.0%), HET (501, 18.0%) or IDU (308, 11.1%), respectively (Table 1). Significant difference was observed among persons infected through MSM, HET, or IDU with regard to age, gender, marital status, education level, and the proportion of HIV-1 genotypes (Supplementary Table S1).

In our study, the proportion of the total HET sequences for men and women was 68.1% (341 of 501) and 31.9% (160 of 501), respectively, and was similar to that of the unclustered sequences for men (64.9%, 270 of 416) and women (35.1%, 146

of 416). The proportion of the total IDU sequences for men and women was 88.96% (274 of 308) and 11.04% (34 of 308), respectively, and was similar to that of the unclustered sequences for men (88.51%, 262 of 296) and women (11.49%, 34 of 296). These results indicated no significant sampling bias for the unclustered self-disclosed HET and IDU sequences attributed to men or women. It is interesting to note that 17.06% (85 of 501) of the HET sequences and 3.90% (12 of 308) of the IDU sequences clustered, whereas 35.05% (693 of 1977) of the MSM sequences clustered (Table 1). Further analysis also indicated that MSM were 1.95-fold more likely to cluster than those of HET (odds ratio[OR], 1.95; 95% confidence interval [CI], 1.48–2.57; $P < .001$) (Table 1). In addition, HIV-1 sequences of CRF01_AE (OR, 2.10; 95% CI, 1.36–3.23; $P = .001$), CRF07_BC (OR, 2.44; 95% CI, 1.59–3.74; $P < .001$), and CRF55_01B (OR, 3.71; 95% CI, 2.35–5.87; $P < .001$) were more likely to cluster than those of HIV-1 subtype B, whereas the sequences of HIV-infected men were more likely to cluster than those of women (OR, 2.80; 95% CI, 1.53–5.11; $P < .001$) (Table 1).

In our study, we retrieved 323 *pol* sequences of HIV-1 subtype B, CRF01_AE, CRF07_BC, and CRF55_01B from the LANL database as background sequences. All the background sequences have the information of gender and risk groups of HIV-1 infection. Of 3109 sequences analyzed, 28.36% (790 of 2786) of the Guangzhou sequences and 22 background sequences had a putative linkage with at least 1 other sequence at a pairwise genetic distance of 0.5%.

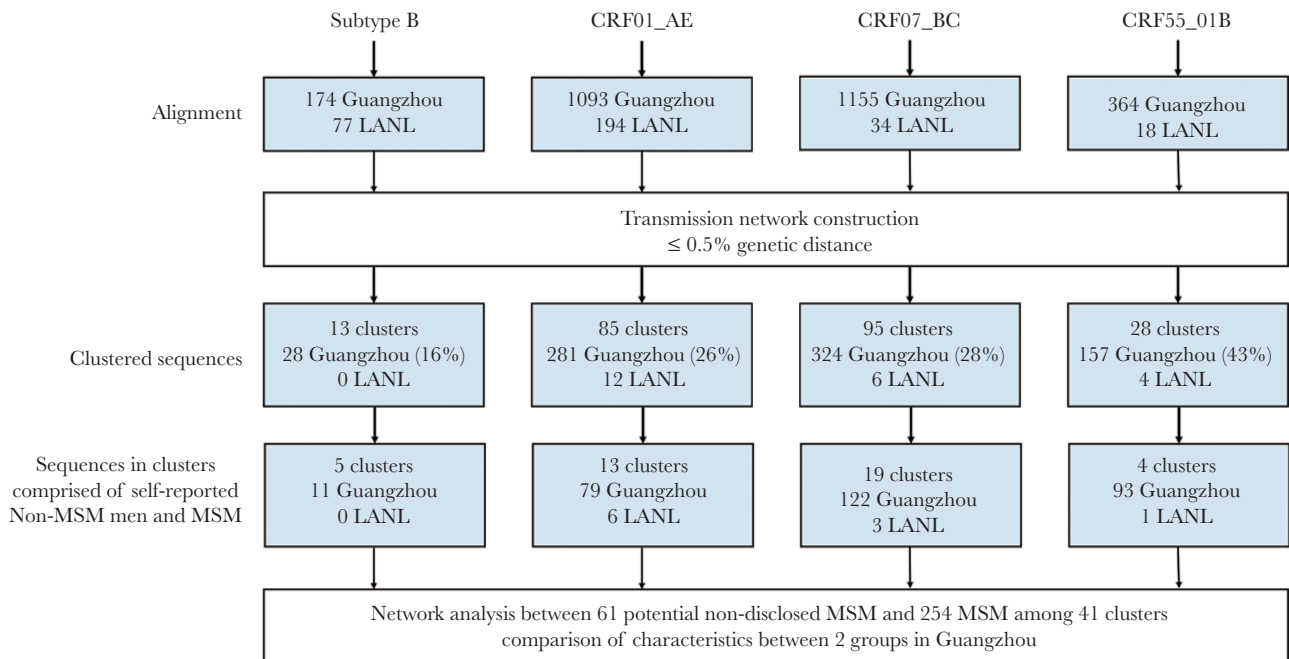


Figure 1. Data processing flow chart. The number of sequences at each stage of analysis were shown. Only clusters with at least 1 Guangzhou sequence were counted in the transmission network. Percentages were calculated with the total number of sequences from that human immunodeficiency virus-1 subtype or genotype as the denominator. LANL, Los Alamos National Laboratory; MSM, men who have sex with men.

Table 1. Comparison Between the Clustered HIV-1 Sequences in the Transmission Network and the Rest of the HIV-1 Sequences Obtained in Guangzhou, China^a

| Characteristics | Sequences Clustered (n = 790) | Sequences Unclustered (n = 1996) | Odds Ratio (95% CI) | PValue |
|-----------------------|-------------------------------|----------------------------------|---------------------|--------|
| HIV-1 Genotype | | | | |
| B | 28 (3.5%) | 146 (7.3%) | Reference | |
| CRF01_AE | 281 (35.6%) | 812 (40.7%) | 2.10 (1.36–3.23) | .001 |
| CRF07_BC | 324 (41.0%) | 831 (41.6%) | 2.44 (1.59–3.74) | <.001 |
| CRF55_01B | 157 (19.9%) | 207 (10.4%) | 3.71 (2.35–5.87) | <.001 |
| Risk Group | | | | |
| HET | 85 (10.8%) | 416 (20.8%) | Reference | |
| IDU | 12 (1.5%) | 296 (14.8%) | 0.16 (0.09–0.30) | <.001 |
| MSM | 693 (87.7%) | 1284 (64.3%) | 1.95 (1.48–2.57) | <.001 |
| Gender | | | | |
| Female | 14 (1.8%) | 180 (9.0%) | Reference | |
| Male | 776 (98.2%) | 1816 (91.0%) | 2.80 (1.53–5.11) | <.001 |

Abbreviations: CI, confidence interval; HET, heterosexual individuals; HIV, human immunodeficiency virus; IDU, intravenous drug users; MSM, men who have sex with men.

^aStatistical significances were calculated using multivariate logistic regression analysis.

These sequences formed 221 clusters and their sizes ranged from 2 to 88 (Figures 1 and 2A). Among them, 60.63% (134 of 221) of these transmission cluster sizes are $n = 2$. The proportion of cluster sizes 3–5, 6–10, and >10 is 31.67%, 3.62%, and 4.07%, respectively (Supplementary Table S2). Furthermore, 72.85% (161 of 221) of these clusters were composed of MSM only (Supplementary Table S3). A total of 19.91% of the clusters contain MSM and the subjects infected through HET or IDU, whereas 7.23% of the cluster are non-MSM (Supplementary Table S3). These results indicated that HIV-1 transmission is more frequent among MSM, but majority of these HIV-1 infections were one-to-one transmission.

Characteristics of Nondisclosed Men Who Have Sex With Men

Based on the transmission networks constructed by HIV-TRACE, 18.55% (41 of 221) of the clusters contained 1 or more than 1 nondisclosed MSM and 1 or more than 1 MSM but not women (Figure 2A). These clusters included 61 nondisclosed MSM and 254 MSM (Figure 1), indicating that 9.92% (61 of 615) of the self-reported non-MSM may actually be nondisclosed MSM. Of 41 clusters containing nondisclosed MSM, 18 (43.9%) had only 2 sequences, and the maximum cluster size of Cluster168 was 88 (Figure 2B). The median of links for each node was 3.0 (interquartile range [IQR], 2.0–7.0). The average genetic distance among these links was $0.33\% \pm 0.11\%$. Most of these clusters were infected with CRF07_BC (46.3%) or CRF01_AE (31.7%), but the largest cluster (Cluster168) was infected with CRF55_01B. Compared with MSM, nondisclosed MSM were slightly older ($P = .006$), more likely to be married ($P = .002$), and less likely to be educated ($P < .001$) than MSM (Table 2). The proportion of HIV-1 genotypes was not significantly different between the two groups ($P = .219$).

Position of Nondisclosed Men Who Have Sex With Men in the Transmission Network

Clusters created in this study indicated that nondisclosed MSM could be in the central (Figure 2C) or peripheral (Figure 2D) positions within a transmission network. Nondisclosed MSM could also link both MSM and themselves (Figure 2E) in the transmission network. For all of the transmission networks containing nondisclosed MSM and MSM, the average shortest path length was significantly smaller for nondisclosed MSM (1.0; IQR, 1.0–2.0) than for MSM (2.0; IQR, 1.2–3.2; $P < .001$),

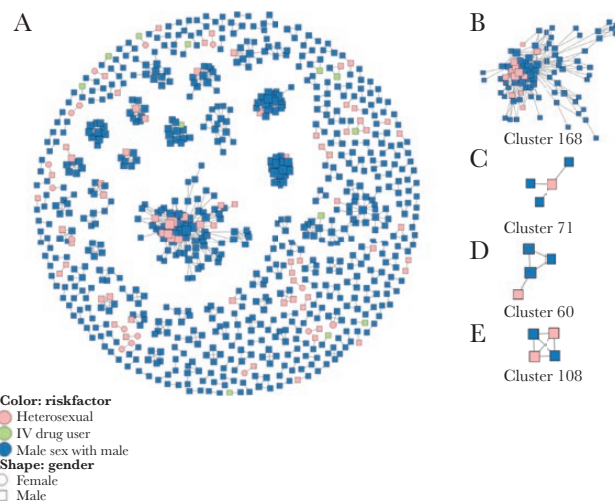


Figure 2. Clusters containing men who have sex with men (MSM) and heterosexual men and positions of nondisclosed MSM within MSM clusters. (A) All human immunodeficiency virus-1 transmission clusters identified in this study. (B) Clusters containing multiple nondisclosed MSM and MSM. (C) Nondisclosed MSM tended to be on the center of clusters. (D) Nondisclosed MSM tended to be on the periphery of clusters. (E) Nondisclosed MSM linked to other nondisclosed MSM in MSM clusters. Potential nondisclosed MSM were inferred as self-reported heterosexual men who clustered with MSM and with no women.

Table 2. Comparison of Demographic Characteristics of HIV-Infected Nondisclosed MSM and MSM in Guangzhou, China^a

| Characteristic | Nondisclosed MSM (n = 61) | MSM (n = 244) ^b | P Value |
|------------------------------|---------------------------|----------------------------|---------|
| Age (mean ± SD) | 35.61 ± 10.90 | 31.36 ± 8.29 | .006 |
| Marital status, n (%) | | | .002 |
| Unmarried | 27 (44.3%) | 159 (65.2%) | |
| Married or cohabiting | 26 (42.6%) | 54 (22.1%) | |
| Divorced or separated | 5 (8.2%) | 18 (7.4%) | |
| Unknown | 3 (4.9%) | 13 (5.3%) | |
| Education Levels, n (%) | | | <.001 |
| Primary school or less | 8 (13.1%) | 7 (2.87%) | |
| Junior or senior high school | 37 (60.7%) | 132 (54.1%) | |
| University or academy | 12 (19.7%) | 97 (39.8%) | |
| Unknown | 4 (6.6%) | 8 (3.3%) | |
| HIV-1 Genotype, n (%) | | | .219 |
| B | 5 (8.2%) | 6 (2.5%) | |
| CRF01_AE | 15 (24.6%) | 64 (26.2%) | |
| CRF07_BC | 26 (42.6%) | 96 (39.3%) | |
| CRF55_01B | 15 (24.6%) | 78 (32.0%) | |

Abbreviations: CI, confidence interval; HIV-1, human immunodeficiency virus-1; LANL, Los Alamos National Laboratory; MSM, men who have sex with men; SD, standard deviation.
^aStatistical significances were calculated using 2-sample *t* test, χ^2 test, and Mann-Whitney *U* test.

^bData are not available for 10 background LANL sequences.

whereas the closeness centrality (1.0, IQR = 0.5–1.0 vs 0.5, IQR = 0.3–0.8; $P < .001$) and the radiality (1.0, IQR = 0.8–1.0 vs 0.8, IQR = 0.6–0.9; $P < .001$) were significantly greater in nondisclosed MSM than in MSM (Table 3), indicating that nondisclosed MSM were in a central position of the transmission networks and played an important role in HIV-1 transmission. Only the degree was slightly smaller for nondisclosed MSM (2.0; IQR, 1.0–5.5) than for MSM (3.0; IQR, 2.0–7.0; $P = .008$), suggesting less connects in nondisclosed MSM than in the MSM with whom they clustered.

Table 3. Analysis of Centrality Indicators for Nondisclosed MSM and MSM in Guangzhou, China

| Indicators ^a | All clusters | | | Restricted to clusters of size $n \geq 3$ | | | Restricted to clusters of size $n \geq 3$ and equal number of cases ^b | | |
|------------------------------|---------------------------|---------------|---------|---|------------------|---------|--|----------------|---------|
| | Nondisclosed MSM (n = 61) | MSM (n = 254) | P Value | Nondisclosed MSM (n = 43) | MSM (n = 236) | P Value | Nondisclosed MSM (n = 43) | MSM (n = 43) | P Value |
| Average shortest path length | 1.0 (1.0–2.0) | 2.0 (1.2–3.2) | <.001 | 1.6 (1.0–2.3) | 2.2 (1.3–3.3) | .001 | 1.6 (1.0–2.3) | 1.8 (1.2–2.5) | .046 |
| Closeness centrality | 1.0 (0.5–1.0) | 0.5 (0.3–0.8) | <.001 | 0.6 (0.4–1.0) | 0.5 (0.3–0.8) | .001 | 0.64 (0.4–1.0) | 0.56 (0.4–0.8) | .035 |
| Degree | 2.0 (1.0–5.5) | 3.0 (2.0–7.0) | .008 | 3.0 (2.0–9.0) | 4.0 (2.0–7.0) | .715 | 3.0 (2.0–9.0) | 2.0 (2.0–5.0) | .069 |
| Radiality | 1.0 (0.8–1.0) | 0.8 (0.6–0.9) | <.001 | 0.79 (0.72–1.00) | 0.73 (0.62–0.88) | .002 | 0.79 (0.72–1.00) | 0.75 (0.6–0.9) | .020 |

Abbreviation: MSM, men who have sex with men.

^aData are median (IQR) for the abnormal distribution of indicators. Statistical significance was calculated using a Mann-Whitney test.

^bAnalysis for equal number of nondisclosed MSM and MSM using Wilcoxon signed rank test.

Of note, for the transmission cluster size $n = 2$, the contribution of each member to the centrality is equal. Therefore, we recalculated the average shortest path length and other centrality indicators after removing the clusters $n = 2$ and got similar results (Table 3). To evaluate the potential bias due to the larger MSM-dominant components, we further analyzed the centrality indicators using equal number of nondisclosed MSM and MSM, and we found no significant difference (Table 3). These results indicate that cluster size did not significantly affect the results of centrality analysis.

Furthermore, we also conducted centrality analysis in the absence of additional LANL sequences. The results indicated that 782 Guangzhou sequences formed 217 clusters with the cluster size of 2 to 87. A total of 60 potential nondisclosed MSM were inferred. The sensitivity analysis showed very similar results (Supplementary Table S4) to those described in Table 3, indicating no significant influence of additional LANL sequences when determining the place of nondisclosed MSM in the transmission networks.

To investigate whether nondisclosed MSM link to each other more closely than to MSM in the clusters, we chose 8 clusters containing at least 2 MSM and at least 2 self-reported heterosexual men. There were 152 MSM and 26 heterosexual men among the 8 clusters in the true networks. The assortativity coefficient was 0.025, indicating that nondisclosed MSM were more likely to link to each other.

Factors Associated With Nondisclosed Men Who Have Sex With Men

We used both univariate and multivariate logistic regression analysis to identify factors that could predict nondisclosed MSM among the self-reported non-MSM (Supplementary Tables S5 and S6). We found that nondisclosed MSM were more likely to be younger ($P = .003$), less likely to be infected with CRF01_AE ($P < .001$) or CRF07_BC ($P = .003$) than CRF55_01B, and more likely to be infected through HET than IDU (OR, 4.56; 95% CI, 2.20–9.46; $P < .001$) (Supplementary Table S7). However, no significant difference was found regarding marital status and

education levels between nondisclosed MSM and self-reported non-MSM.

DISCUSSION

We inferred approximately 10% of self-reported non-MSM as nondisclosed MSM in Guangzhou, China and their central position in HIV-1 transmission networks. Unlike the peripheral position of nondisclosed MSM that have been reported in the UK [5], we found that nondisclosed MSM in China were more likely to be in the central position, indicating their important role and relatively great contribution of HIV-1 transmission. However, the position within the transmission clusters could also be related to the timing of HIV-1 infection. Unfortunately, the influence of the timing of infection was not evaluated due to the lack of the information. Our preliminary results and previous study indicate that nondisclosed MSM should be considered as a specific target group with high risk in HIV-1 transmission and epidemics.

In our study, 28.36% of the Guangzhou sequences linked with at least 1 other sequence at a pairwise genetic distance of $\leq 0.5\%$ and were included in the transmission clusters. Our results are in line with the findings in the UK where 28.8% of the patient sequences were included in the resulting networks by using a cutoff value of 5 years to most recent common ancestor sequences [5]. We must emphasize that 9.92% of self-reported non-MSM were inferred to be nondisclosed MSM in Guangzhou. This tends to be a relatively conservative estimation because we chose a relatively small pairwise genetic distance of 0.5%, given the rapidly spreading HIV-1 among MSM. Ragonnet-Cronin et al [5] identified 18.6% of all clustered heterosexual men in the networks to be nondisclosed MSM in the UK. In another study, Hué et al [6] reported that up to 21% of black African men who report only heterosexual behavior became infected through sex with men in the UK. These results suggested that the actual proportion of nondisclosed MSM may be underestimated and should be further investigated in China.

Compared with MSM with whom they clustered, nondisclosed MSM were more likely to be older, married, and less educated. Our results are consistent with earlier findings that low rates of MSM sexual orientation disclosure in China may be associated with limited access to MSM services, insufficient social and psychological support, and social discrimination [7, 19]. We also found that among the self-reported non-MSM, infection with HIV-1 CRF55_01B or through HET rather than IDU showed a greater likelihood of being inferred as nondisclosed MSM. These results suggest that nondisclosed MSM in Guangzhou are actually bisexual and have both male and female sex partners, which in turn may increase the risk of their sex partners to infect HIV-1. The similar study conducted in the UK also supports that nondisclosed MSM link MSM and heterosexual epidemics [5].

Human immunodeficiency virus-1 CRF55_01B was first described in 2006 in China and is a recombinant of CRF01_AE and subtype Thailand B' [20]. A study in Shenzhen city, which is very close to Guangzhou city, showed that among HIV-infected MSM, the proportion of CRF55_01B increased from 0 in 2006 to 17% in 2012, indicating a rapid spread of HIV-1 CRF55_01B infection in this population [21]. These results also suggest that HIV-1 CRF55_01B is in an stage of active transmission and rapid expansion, which may affect HIV-1 epidemics in China. It remains to be investigated whether the increased prevalence of CRF55_01B among HIV-infected MSM is due to the wide spread of CRF55_01B infection in nondisclosed MSM or better virus fitness for CRF55_01B.

Compared with previous HIV-1 transmission network studies, which were limited to a single HIV-1 subtype at a time, HIV-TRACE is substantially powerful in analyzing multiple HIV-1 subtypes or genotypes simultaneously [8]. A limitation of our study was that the transmission network created by HIV-TRACE cannot determine the direction and order of transmission events. Ragonnet-Cronin et al [5] have pointed out that by using such a reconstruction method, the degree, ie, the number of links within a transmission network, may be overestimated and does not represent the true events of onward transmissions. In our study, we used 3 robust centrality indicators, including average shortest path length, closeness centrality, and radiality to investigate the position of nondisclosed MSM within a transmission network to reduce possible bias due to the size of clusters or the lack of the timing of HIV-1 infection. The results of the 3 centrality indicators were significantly different between nondisclosed MSM and MSM with whom they clustered, indicating that nondisclosed MSM may play an important role in HIV-1 transmission in Guangzhou. Our results are obviously different from those reported in the UK [5], and they may reveal differences in sexual or transmission behavior of nondisclosed MSM in different countries or regions. However, more detailed analysis including the information of transmission directions and timing of HIV-1 infection may gain crucial insight into the behavior of this population. Ratmann et al [22] recently specified to use viral deep-sequencing to reconstruct HIV-1 transmission networks and to determine the direction of transmission in these networks. This new technique provides us with a novel tool for further characterization of the role of nondisclosed MSM in HIV-1 transmission. In addition, we did not test the role of nondisclosed MSM in transmitting HIV-1 between MSM and heterosexual epidemics through broker analysis because there were only 16 heterosexual women in our clusters and 14 of them were in the clusters with only 2 members. A nationwide study is necessary to include a sufficient number of HIV-1 sequences and HIV-infected subjects from different regions. We want to emphasize that the inference of nondisclosed MSM is based on their linkage to MSM. Therefore, any comparison with MSM is inherently biased.

CONCLUSIONS

Our results indicate the central role of nondisclosed MSM in HIV-1 transmission and support the urgent need for prevention and healthcare measures specifically targeting the nondisclosed MSM population because they are barely reached by traditional epidemiological investigation and may not receive appropriate healthcare programs designed for MSM. Their risk in HIV-1 transmission has not yet been fully studied. Further studies collaborated with social scientists should aim to better understand the characteristics of nondisclosed MSM and their needs.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Financial support. This work was funded by the National Key R&D Program of China (2018ZX10732401-003) and the Bureau of Science and Information Technology of Guangzhou Municipality (201604020011, 201704020219, 201707010184, and 201803040002).

Potential conflicts of interest. All authors: no reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References

1. van Griensven F, de Lind van Wijngaarden JW, Baral S, Grulich A. The global epidemic of HIV infection among men who have sex with men. *Curr Opin HIV AIDS* **2009**; 4:300–7.
2. Beyrer C, Baral SD, van Griensven F, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet* **2012**; 380:367–77.
3. Wang G. Current situation of HIV infection and AIDS in China. In: Information on AIDS prevention and control in China; **2017**:46–7.
4. Centers for Disease Control and Prevention (CDC). HIV/STD risks in young men who have sex with men who do not disclose their sexual orientation--six U.S. cities, 1994–2000. *MMWR Morb Mortal Wkly Rep* **2003**; 52:81–6.
5. Ragonnet-Cronin M, Hué S, Hodcroft EB, et al.; UK HIV Drug Resistance Database. Non-disclosed men who have sex with men in UK HIV transmission networks: phylogenetic analysis of surveillance data. *Lancet HIV* **2018**; 5:e309–16.
6. Hué S, Brown AE, Ragonnet-Cronin M, et al.; UK Collaboration on HIV Drug Resistance and the Collaborative HIV, Anti-HIV Drug Resistance Network (CHAIN). Phylogenetic analyses reveal HIV-1 infections between men misclassified as heterosexual transmissions. *AIDS* **2014**; 28:1967–75.
7. Tang W, Mao J, Tang S, et al.; SESH Study Group. Disclosure of sexual orientation to health professionals in China: results from an online cross-sectional study. *J Int AIDS Soc* **2017**; 20:21416.
8. Kosakovsky Pond SL, Weaver S, Leigh Brown AJ, Wertheim JO. HIV-TRACE (TRANsmision Cluster Engine): a tool for large scale molecular epidemiology of HIV-1 and other rapidly evolving pathogens. *Mol Biol Evol* **2018**; 35:1812–9.
9. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **2002**; 30:3059–66.
10. Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **2016**; 44:W232–5.
11. Siepel AC, Halpern AL, Macken C, Korber BT. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res Hum Retroviruses* **1995**; 11:1413–6.
12. Division of HIV/AIDS Prevention. Detecting and responding to HIV transmission clusters. Technical Report: Centers for Disease Control and Prevention **2018**. <https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-Clusters.pdf>. Accessed June 2018.
13. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003**; 13:2498–504.
14. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* **1998**; 393:440–2.
15. Newman ME. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys Rev E Stat Nonlin Soft Matter Phys* **2001**; 64:016132.
16. Brandes U. A faster algorithm for betweenness centrality. *J Math Sociol* **2001**; 25:163–77.
17. Broder AZ, Kumar R, Maghoul F, et al. Graph structure in the Web. *Computer Networks* **2000**; 33:309–20.
18. Newman ME. Mixing patterns in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* **2003**; 67:026126.
19. Koji U. Sexual orientation and psychological distress in adolescence: examining interpersonal stressors and social support processes. *Social Psychology Quarterly* **2005**; 68:258–277.
20. Han X, An M, Zhang W, et al. Genome sequences of a novel HIV-1 circulating recombinant form, CRF55_01B, identified in China. *Genome Announc* **2013**; 1:e00050–12.
21. Zhao J, Cai W, Zheng C, et al. Origin and outbreak of HIV-1 CRF55_01B among MSM in Shenzhen, China. *J Acquir Immune Defic Syndr* **2014**; 66:e65–7.
22. Ratmann O, Grabowski MK, Hall M, et al.; PANGEA Consortium and Rakai Health Sciences Program. Inferring HIV-1 transmission networks and sources of epidemic spread in Africa with deep-sequence phylogenetic analysis. *Nat Commun* **2019**; 10:1411.