

Migration interacts with the local transmission of HIV in developed trade areas: A molecular transmission network analysis in China

Jiafeng Zhang^a, Jiaming Yao^a, Jun Jiang^a, Xiaohong Pan^{a,*}, Mingyu Luo^a, Yan Xia^a, Qin Fan^a, Xiaobei Ding^a, Jianjun Ruan^b, Andreas Handel^c, Justin Bahl^{d,e}, Wanjun Chen^a, Lianqi Zha^b, Tao Fu^{b,**}

^a Department of HIV/AIDS & STD control and prevention, Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou 310051, China

^b Yiwu Municipal Center for Disease Control and Prevention, Yiwu 322000, China

^c Department of Epidemiology and Biostatistics, Health Informatics Institute, Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA, USA

^d Center for the Ecology of Infectious Diseases, Department of Infectious Disease, Department of Epidemiology and Biostatistics, Institute of Bioinformatics, University of Georgia, Athens, GA, USA

^e Duke-NUS Graduate Medical School, Singapore

ARTICLE INFO

Keywords:

HIV-1
Molecular epidemiology
Transmission network
Phylogenetic

ABSTRACT

The HIV-1 epidemic is a remarkable public health concern in China, especially in developed trade areas. We aimed to investigate the interaction of migration with the local transmission network in a typical trade area, Yiwu City, the world's largest commodity distribution center. Based on 390 pol sequences from 413 participants diagnosed between 2014 and 2016, putative transmission clusters and the underlying demographic and behavioral characteristics were analyzed. Recent infection status was determined by HIV-1 limiting antigen avidity enzyme immunoassay to identify active clusters. Multiple subtypes were identified, with a predominance of CRF01_AE (47.4%) and CRF07_BC (40.8%), followed by 9 other subtypes and 8 URFs. Multivariable analyses revealed that individuals in clusters were more likely to be local residents, infected through heterosexual behaviors, and infected with CRF01_AE ($P < .05$). Of men who have sex with men (MSM), 81% were linked to other MSM, and only 3% were linked to heterosexual women. Of heterosexual women, 67% were linked to heterosexual men, and 11% to MSM. Yiwu residents were more likely to link to locals than that of migrants (43% vs 20%, $P < .001$). By contrast, local MSM and migrant MSM all had high percentages of linkage to migrant MSM (57% vs 69%, $P = .069$). Our findings reveal that migration promotes the dissemination and dynamic change of HIV, which are interwoven between locals and migrants. The results highlight the far-reaching influence of migrant MSM on the local HIV transmission network.

1. Introduction

Despite substantial efforts to control human immunodeficiency virus-1 (HIV-1), the HIV epidemic remains disproportionately severe across the world. By the end of 2018, approximately 37.9 million people were living with HIV (UNAIDS, 2018), of which 1.25 million were located in China (Lyu and Chen, 2019). Sexual transmission is the main mode of transmission in China, and its hidden nature presents a challenge for intervention. Another challenge for the expanding HIV epidemic is frequent population flow, especially in trade areas (Liu et al., 2014; Mendelsohn et al., 2015). Population mobility is of great

significance in HIV dissemination and circulation (Aldrich and Hemelaar, 2012). In China, numerous migrants frequently move between economically developed areas to seek better employment opportunities and living conditions. By the end of 2017, the size of the migrant population had reached 244 million nationwide (National Health Commission, 2018). Several studies have indicated that migrants account for a large proportion of the HIV epidemic in China (Chan et al., 2015; Dennis et al., 2014).

With this frequent population flow, developed trade areas (e.g., Shenzhen (Zhao et al., 2016), Shanghai (Li et al., 2016), and Beijing (Zhao et al., 2018)) in China have been increasingly characterized by

* Corresponding author at: Department of HIV/AIDS & STD Control and Prevention, Zhejiang Provincial Center for Disease Control and Prevention, No.3399, Binsheng Road, Hangzhou 310051, China.

** Corresponding author at: Yiwu Municipal Center for Disease Control and Prevention, No. 48, Xuefeng East Road, Yiwu 322000, China.

E-mail addresses: xhpan@cdc.zj.cn (X. Pan), 775883023@qq.com (T. Fu).

growth of the AIDS epidemic. The notion that migrant populations are the source of HIV import remains controversial. This hypothesis is supported by a previous study reporting that the vast majority (90.3%) of HIV-1-infected men who have sex with men (MSM) in Shenzhen were migrants and that these migrants had significantly different HIV-1 subtype distributions than local MSM (Zhao et al., 2016). Others have documented that the spread of HIV among local residents is the primary cause of the increase in HIV. The tools of HIV molecular epidemiology can be used to understand transmission dynamics and assist in the design of prevention trials. New developments in recent years, especially molecular transmission network analysis (Aldrich and Hemelaar, 2012; Chan et al., 2015; Chen et al., 2018; Dennis et al., 2014; Li et al., 2018; Little et al., 2014; Oster et al., 2015; Wertheim et al., 2014), have allowed a new understanding of transmission between subpopulations (Chan et al., 2015; Li et al., 2018).

In this study, we applied molecular transmission network analysis to investigate the transmission dynamics of HIV in a typical trade area. We inferred the local molecular transmission network based on HIV-1 pol sequences, and the results enhance the understanding of the impact of population mobility on HIV-1 transmission in trade areas and provide a reference for designing effective intervention trials.

2. Materials and methods

2.1. Study participants and sample collection

We conducted a retrospective molecular epidemiology study in a world-renowned trade city (Yiwu) in China. Yiwu City is a city of approximately 2,000,000 people in central Zhejiang province in East China, including a registered local population of 0.74 million and a floating population of 1.43 million, and is the world's largest commodity distribution center. In addition, an estimated 500,000 overseas merchants come to the city for procurement each year.

A total of 413 newly confirmed HIV-1-positive blood samples, without antiviral therapy, were continuously collected between January 2014 and December 2016 through routine follow-up. The sampling density for newly confirmed individuals was estimated to reach 80.7% (413/512) in this study. The corresponding demographic data were also collected by the staff working at the local center for disease control (CDC). HIV-1 infection status was first screened by an enzyme-linked immunosorbent assay (ELISA) (GenScreen ULTRA HIV Ag-Ab, BIO-RAD, USA) and confirmed by Western blot assay (HIV BLOT 2.2, MP Diagnostics, Singapore). The study was approved by the Medical Ethics Committee of the National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention. All methods in this study were carried out in accordance with approved guidelines and regulations.

2.2. Recent infection detection

The HIV-1 limiting antigen avidity enzyme immunoassay (HIV-1 LAg-Avidity EIA, Sedia, Portland, OR) was used to test the serum of the first HIV-1 positive sample at diagnosis to identify if it was a new infection or a long-term infection. When the study subjects' CD4 lymphocyte counts went below 200 cells/ μ l or if they were diagnosed with AIDS in a 6-month follow-up period, the LAg-Avidity EIA "recent infection" results were negated.

2.3. Amplification of HIV-1 gene fragments

Viral RNA was extracted from 140 μ l of plasma using the QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. RNA samples were directly subjected to reverse transcription PCR and nested PCR to generate pol fragments (HXB2: 2147–3462). The pol fragment covering the entire protease (PR) and the first 300 codons of the reverse transcriptase (RT) gene was

amplified using the One Step RT-PCR Kit (Takara, Dalian, China). The details of amplification and sequencing were as previously described (Zhang et al., 2015). Amplified products were analyzed using 1.0% agarose gel electrophoresis. Target PCR products were sent to Hangzhou TsingKe Bio Tech Co. for purification and sequencing by using five overlapping primers.

2.4. Sequence analysis

The trimming and assembly of sequences were performed using DNA sequence analysis software Sequencher v5.0 (Gene Codes, Ann Arbor, MI). ClustalW multiple alignments were performed using BioEdit v7.0. Reference sequences were obtained from the Los Alamos National Laboratory (LANL) HIV sequence database (<https://www.hiv.lanl.gov>), which covers the major HIV-1 subtypes and circulating recombinant forms (CRFs). By using MEGA v6.0, phylogenetic tree analyses were performed using the neighbor-joining method based on the Kimura 2-parameter model with 1000 bootstrap replicates. The viral sequences were determined as the subtypes of reference strains by phylogenetic tree nodes with bootstrap support values of more than 75%. They were considered as URFs if the query sequences did not match any known subtypes or CRFs. To demonstrate possible inter-subtype recombinations, candidate sequences were analyzed using the Recombination Identification Program (RIP) v3.0 (Siepel et al., 1995) (<https://www.hiv.lanl.gov/content/sequence/RIP/RIP.html>).

The nucleotide sequence sets of the pol gene in FASTA format were submitted to the online Calibrated Population Resistance (CPR) Tool v6.0 (<http://cpr.stanford.edu/cpr.cgi>). Sequences with surveillance drug resistance mutations (SDRMs) were selected using the Genotypic Resistance Interpretation Algorithm based on the list recommended by the WHO in 2009 (Bennett et al., 2009).

2.5. Construction and analysis of molecular transmission networks

TN93 substitution model (Aldous et al., 2012; Chan et al., 2015; Oster et al., 2015; Wertheim et al., 2017; Wertheim et al., 2014; Zhao et al., 2018) was used in the calculation of the pairwise genetic distances (GDs) among viral sequences by MEGA v6.0. The optimal GD threshold was selected across the range of GDs from 0.0025 to 0.020 expected substitutions per site when the number of transmission clusters was maximized to provide the highest resolution (Wertheim et al., 2017; Zhao et al., 2018). A putative link between two individuals was considered when their GD was less than or equal to the threshold. The network was constructed by identifying pairs of sequences (nodes) and their potential transmission relationships (edge) (Aldous et al., 2012; Chan et al., 2015; Little et al., 2014; Wertheim et al., 2014; Zhao et al., 2018) utilizing the visualization software Cytoscape v3.6.2. Clusters with ≥ 4 nodes were defined as larger clusters (Oster et al., 2015) in this study. Clusters were regarded as active when they included ≥ 3 subjects with new infections. To understand the extent to which HIV transmission occurs between subjects with different categories, we analyzed the characteristics of potential transmission partners based on the number of links (Oster et al., 2015). For the designated transmission category, the proportions of links associated with different categories were calculated.

2.6. Statistical analyses

Statistical analyses were conducted using SPSS v19.0 (IBM, Armonk, NY). Categorical variables were compared using Person's chi-square test or Fisher's exact test. We constructed univariate and multivariate logistic regression analysis to test the association between HIV clustering and participants' demographic and behavioral characteristics. Variables determined from a univariate analysis ($P < .10$) were inserted in a multiple logistic regression model in order to analyze the factors that influenced HIV clustering. P values less than 0.05 were considered

Table 1
Univariate and multivariate analysis of clustered and non-clustered subjects in the molecular network.

Variables	Subjects	Subtypes				Sequences		Univariate analysis	Multivariate analysis
	(N = 413)	Total (N = 390)	CRF01_AE	CRF07_BC	Others	Non-clustered	Clustered	OR(95% CI)	AOR(95% CI)
Diagnosis time									
2014	98(23.7)	97(24.9)	48	36	13	39	58	1	
2015	147(35.6)	136(34.9)	59	58	19	66	70	0.713(0.421–1.208)	
2016	168(40.7)	157(40.3)	78	65	14	70	87	0.836(0.500–1.397)	
Gender									
Male	348(84.3)	329(84.4)	161	134	34	149	180	1	
Female	65(15.7)	61(15.6)	24	25	12	26	35	0.784(0.452–1.360)	
Age group									
≤ 25 yrs	94(22.8)	88(22.6)	47	34	7	42	46	1	
26–34 yrs	118(28.6)	110(28.2)	59	36	15	61	49	0.733(0.418–1.287)	0.802(0.446–1.441)
≥ 35 yrs	201(48.7)	192(49.2)	79	89	24	72	120	1.522(0.914–2.535)	1.382(0.798–2.395)
Marital status									
Single	176(42.6)	164(42.1)	91	57	16	77	87	1	
Married	156(37.8)	148(38.0)	61	69	18	58	90	1.373(0.875–2.155)	
Divorced/widowed	81(19.6)	78(20.0)	33	33	12	40	38	0.841(0.049–1.442)	
Education									
Primary school or illiterate	101(24.5)	98(25.1)	38	45	15	40	58	1	
Junior high school	161(39.0)	154(39.5)	73	68	13	72	82	0.785(0.471–1.311)	
Senior high school	89(21.5)	83(21.3)	43	31	9	35	48	0.946(0.523–1.712)	
College or above	62(15.0)	55(14.1)	31	15	9	28	27	0.665(0.342–1.293)	
Registered household									
Yiwu County	93(22.5)	87(22.3)	48	32	7	22	65	1	
Other areas in Zhejiang province	44(10.7)	43(11.0)	16	19	8	21	22	0.355(0.164–0.765)	0.463(0.208–1.031)
Other provinces	276(66.8)	260(66.7)	121	108	31	132	128	0.328(0.191–0.564)	0.387(0.221–0.676)
Occupation									
Workers and peasants	164(39.7)	156(40.0)	66	67	23	71	85	1	
Commercial service workers	144(34.9)	134(34.4)	75	49	10	64	70	0.914(0.575–1.451)	
Service workers in public places	40(9.7)	37(9.5)	10	20	7	16	21	1.096(0.532–2.258)	
Domestic workers and unemployed individuals	30(7.3)	29(7.4)	17	10	2	8	21	2.193(0.916–5.250)	
Cadres, retirees and students	12(2.9)	11(2.8)	6	4	1	5	6	1.002(0.294–3.422)	
Other and unknown	23(5.6)	23(5.9)	11	9	3	11	12	0.911(0.379–2.190)	
Transmission route									
Homosexual behavior	230(55.7)	215(55.1)	106	85	24	110	105	1	
Heterosexual behavior	181(43.8)	173(44.4)	79	73	21	64	109	1.784(1.186–2.684)	1.619(1.046–2.505)
Mother-to-child	1(0.2)	1(0.3)	0	1	0	0	1	–	
Intravenous drug use	1(0.2)	1(0.3)	0	0	1	1	0	–	
STD history									
Yes	99(24.0)	93(23.9)	44	39	10	45	48	1	
No	309(74.8)	293(75.1)	139	118	36	129	164	1.192(0.747–1.902)	
Unknown	5(1.2)	4(1.0)	2	2	0	1	3	2.807(0.292–27.954)	
First CD4 count after HIV confirmation									
< 200 cells/μl	96(23.2)	90(23.1)	43	33	14	41	49	1	
200–349 cells/μl	162(39.3)	153(39.2)	65	68	20	65	88	1.133(0.671–1.914)	
> 350 cells/μl	155(37.5)	147(37.7)	77	58	12	69	78	0.946(0.559–1.601)	
Infection status									
Recent infection	99(24.0)	94(24.1)	44	45	5	41	53	1	
Latent infection	314(76.0)	296(75.9)	141	114	41	134	162	0.935(0.586–1.493)	
Subtypes									
CRF01_AE	185(44.8)	185(47.4)	/	/	/	70	115	1	
CRF07_BC	159(38.5)	159(40.8)	/	/	/	80	79	0.511(0.266–0.981)	0.549(0.348–0.866)
Other types	46(11.1)	46(11.8)	/	/	/	25	21	0.601(0.391–0.924)	0.518(0.261–1.029)
ND	23(5.6)	/	/	/	/	/	/	/	

ND: not determined (due to failed amplification).

statistically significant.

3. Results

3.1. Demographic characteristics of the study subjects

A total of 413 newly confirmed HIV-positive samples were collected from 2014 to 2016 (Table 1). Of all subjects, 84.3% (348/413) were

male, and 15.7% (65/413) were female. The median age of the subjects was 34.0 years (range: 3–79 years). A total of 22.5% (93/413) were local residents, 10.7% (44/413) were from other regions in Zhejiang province, and 66.8% (276/413) were from other provinces. Of the participants, 42.6% (176/413) were single, 37.8% (156/413) were married, and 19.6% (81/413) were divorced or widowed. A total of 63.4% (262/413) had junior high school education or below. For occupation, workers and peasants accounted for 39.7%, followed by

commercial service workers (34.9%) and other occupations (25.4%). According to the results of the LAg-Avidity EIA, 99 specimens (24.0%) were classified as recent HIV-1 infection, and 314 specimens (76.0%) were classified as being associated with chronic HIV-1 infection. Of the 99 subjects with recent infection, the majority were male (86.7%), 35 years old and above (51.5%), had junior high school education or below (63.6%), migrated from other provinces (63.6%), infected by homosexual behaviors (64.6%) and infected with CRF01_AE (46.8%, 44/94) or CRF07_BC (47.9%, 45/94).

The distribution of genotypes by sampling time and by the participants' gender, age, marital status, education, household registered, occupation, transmission route, sexually transmitted disease (STD) history, first CD4 count after HIV confirmation and infection status revealed no significant differences when tested by Pearson's χ^2 test ($P > .05$). Among the subjects with new infection, the proportions of CRF01_AE and CRF07_BC were close (46.8% vs 47.9%). Among the subjects with latent infection, the proportions of CRF01_AE and CRF07_BC were 47.6% (141/296) and 38.5% (114/296), respectively, though there was no statistical difference ($\chi^2 = 5.946$, $P = .051$).

3.2. HIV-1 subtype distribution

A total of 23 of 413 samples failed to amplify, while 390 (94.4%) were successfully amplified and sequenced and included in the phylogenetic analysis. CRF01_AE was the dominant subtype (47.4%, 185/390), followed by CRF07_BC (40.8%, 159/390), CRF08_BC (2.8%, 11/390), CRF55_01B (2.6%, 10/390), subtype B (2.1%, 8/390), CRF64_BC (0.5%, 2/390), CRF68_01B (0.5%, 2/390), CRF02_AG (0.3%, 1/390), CRF45_cpx (0.3%), CRF59_01B (0.3%), CRF85_BC (0.3%), URF(01_AE/07_BC) (1.8%, 7/390), and URF(B/C) (0.5%, 2/390).

Fourteen (3.6%, 14/390) were identified as harboring SDRMs. All individuals with SDRMs harbored one mutation each, and the predominant subtype was CRF01_AE (11/14), followed by CRF07_BC ($n = 1$), CRF08_BC ($n = 1$) and CRF68_01B ($n = 1$). Six individuals harboring the M46L mutation were found in a cluster (CRF01_AE-C5, Fig. 2A).

3.3. Characterization of genetic transmission networks

A total of 390 pol sequences were subsequently used for genetic transmission network analysis. The optimal GD thresholds were selected as 0.015, 0.010 and 0.0175 substitutions/site for CRF01_AE, CRF07_BC and other subtypes, respectively. A total of 215 (55.1%) sequences were linked to at least one other sequence, and they segregated into 64 distinct clusters, with the number of sequences per cluster ranging from 2 to 16 (Fig. 1A). Forty-one clusters (64.1%) were composed of 2 sequences, and 15 (23.4%) were composed of 4 or more

sequences. All nodes in the network had the number of links ranging from 1 to 11, of which 94 (43.7%) had one link (one potential transmission partner), 174 (80.9%) had no more than 5 links with other nodes, and only 5 nodes had more than 10 links (Fig. 1B). The GD between the pairwise sequences in each cluster ranged from 0.000 to 0.0166 substitutions per site (Fig. 1C). Pairs with a genetic distance of ≤ 0.010 substitutions per site accounted for 76.1% (264/347) of all sequence pairs.

Of the 64 clusters, 28 belonged to CRF01_AE, 27 to CRF07_BC, three to CRF55_01B, three to CRF08_BC, and one each to CRF68_01B, CRF64_BC and URF(01_AE/07_BC), respectively (Fig. 2). We found that the majority of clusters (64.1%, 41/64) had 2 individuals ("dyads"), of which 10 clusters had individuals who were married.

3.4. Analysis of persons with potential transmission links

Of the individuals in clusters, 83.7% were male, 55.8% were ≥ 35 years old, 59.5% had household registered in other provinces, 50.7% were infected by heterosexual behaviors, 48.8% were infected by homosexual behaviors, 75.3% were associated with chronic HIV-1 infection, and 53.5% were infected with CRF01_AE (Table 1). Age group ($P = .009$), registered household ($P < .001$), transmission route ($P = .005$) and subtype ($P = .026$) were associated with clustering (Table 1). According to multivariable logistic regression analysis, individuals in clusters were more likely to be infected from heterosexual behaviors [OR(95% confidence interval): 1.619(1.046–2.505), $P = .017$], less likely to be migrants from other provinces [OR: 0.387(0.221–0.676), $P = .001$] and less likely to be infected with CRF07_BC [OR: 0.549(0.348–0.866), $P = .026$].

As shown in Fig. 3A, of MSM, 81% were linked to other MSM, and only 3% were linked to heterosexual women. Male heterosexuals were commonly linked to male heterosexuals (54%), female heterosexuals (24%) and MSM (21%). Female heterosexuals were most commonly linked to heterosexual men (67%), and 11% were linked to MSM.

As shown in Fig. 3B, of the local residents, 49% were linked to migrants from other provinces, while 43% were linked to local residents. Of the migrants from other areas in Zhejiang province and other provinces, 19% and 20% were linked to local residents, respectively. The local residents had a higher percentage of linkage to locals than that of migrants from other provinces (43% vs 20%, $P < .001$).

MSM with households registered in Yiwu and migrant MSM all had high percentages of linkage to migrant MSM (57% vs 69%, $P = .069$) (Fig. 3C). Heterosexuals with households registered in Yiwu had a higher percentage of linkage to local heterosexuals than to migrants (42% vs 22%, $P < .001$) (Fig. 3C). Migrant heterosexuals were most commonly linked to migrant heterosexuals (55%), and 22% were linked to local heterosexuals.

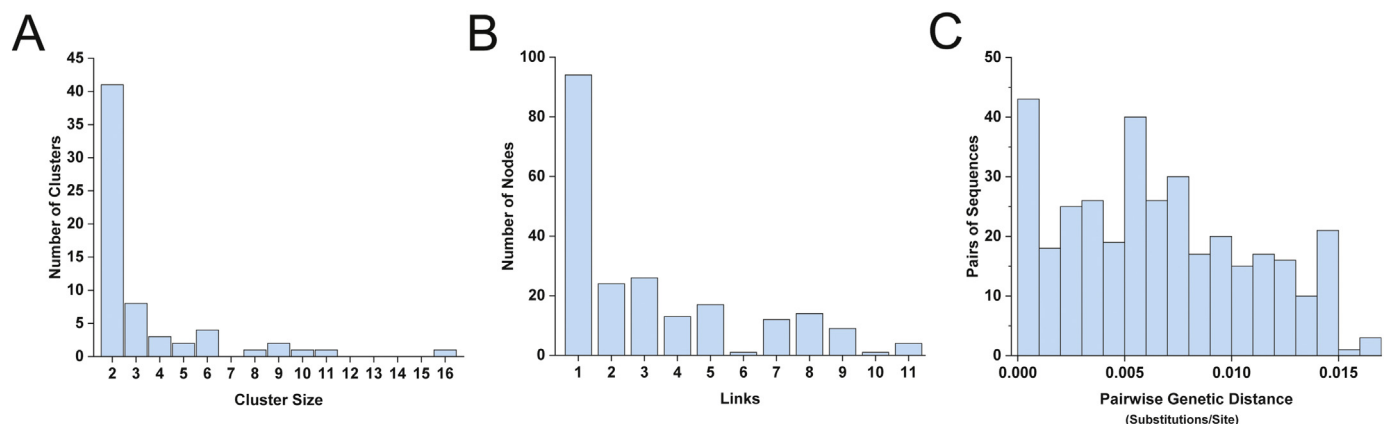


Fig. 1. The characteristics of the molecular transmission networks. (A) Distribution of molecular transmission clusters by cluster size; (B) Distribution of nodes in clusters by links; (C) Distribution of pairs of sequences by pairwise genetic distance.

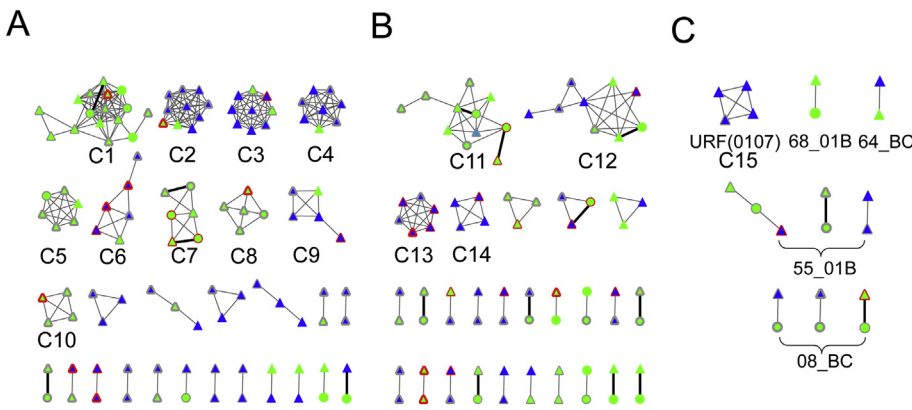


Fig. 2. The molecular transmission networks among newly diagnosed individuals. (A) Clusters with CRF01_AE; (B) Clusters with CRF07_BC; (C) Clusters with other subtypes. Males are labeled with solid upward-pointing triangles (▲), and females are labeled with solid circles (●). Subjects infected with HIV-1 through heterosexual contact are green; homosexual contact are blue. A red outline indicates new infection, and a gray outline indicates chronic infection determined by IAG. Thick, thin and no outline denote newly diagnosed individuals in 2016, 2015 and 2014, respectively. Each pair of spouses is indicated with a bold line to each other. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Characterization of large clusters

Fifteen large clusters (≥ 4 links), including 109 subjects, were found in our study. Ten belonged to CRF01_AE, 4 to CRF07_BC and 1 to URF (01_AE/07_BC), with the number of nodes ranging from 4 to 16 for a single cluster (Table 2). Of the 15 large clusters, 6 clusters (containing 47 subjects) were all infected through heterosexual contact, 4 clusters (containing 19 subjects) were all infected through homosexual contact, and 5 clusters were infected by mixed heterosexual or homosexual transmission. Most of the large clusters (73.3%, 11/15) contained both locals and migrants, with percentages of 32.1% and 67.9%, respectively. Five clusters were active according to the self-designed criteria, in which 4 clusters grew by new cases with recent infection both from locals (6 cases total) and migrants (7 cases total).

4. Discussion

We systematically depicted the molecular transmission network in a typical trade area and found that it was intertwined between local residents and migrant populations. Our analysis of mixing by populations from different household yields important new insights into the influence of migration on transmission networks in trade areas. We found that homosexual behavior in migrants had a greater influence than heterosexual behavior on the local transmission network. Thus tailored interventions that reduce transmission involving migrant MSM are likely to reduce HIV acquisition among local MSM as well.

Developed trade cities provide a great number of better employment opportunities, which attracting large migrant populations from rural areas. Migration contributes to the spread of different genotypes of HIV-1 (Aldrich and Hemelaar, 2012; Li et al., 2018; Li et al., 2016; Su et al., 2018; Zhao et al., 2016). Eleven genotypes were detected in this study, revealing high local HIV genetic diversity and implicating a wide variety of sources. Notably, some novel CRFs that were discovered in the past few years in China were circulating in the study trade area,

including CRF55_01B, CRF59_01B, CRF64_BC, CRF68_01B and CRF85_BC. CRF55_01B, originating from Shenzhen (Zhao et al., 2014), has disseminated widely among MSM in China (Han et al., 2015). Of the ten individuals infected with CRF55_01B, seven were infected through homosexual contact, and two women were considered to be infected from their spouses through heterosexual behavior in marriage. CRF55_01B spread from the homosexual to heterosexual populations through a “bisexual bridge”. CRF45_cpx was first identified in 2009 according to the genome sequences from Cameroon, Gabon, and the Democratic Republic of Congo (DRC) (de Silva et al., 2010). One woman was infected with CRF45_cpx, which was the first case reported in Zhejiang province. An additional epidemiological case survey verified that this individual had a heterosexual contact with an African individual. Thus, HIV diversity has increased due to interactions between infected individuals from other countries, which has complicated control efforts. All seven men were infected with URF(01_AE/07_BC) through homosexual behavior, indicating that the MSM populations play the role of an incubator for new recombinant strains. Previous studies revealed higher rates of dual-variant and multiple-variant HIV infection than in heterosexual people (Li et al., 2010), and URFs are prone to formation in MSM because of the specific behavioral features, including multiple sex partners, low rates of condom use and anal intercourse (Li et al., 2010; Li et al., 2015; Zhang et al., 2017). Given the co-circulation of the predominant subtypes CRF01_AE and CRF07_BC in the MSM population in Zhejiang (Zhang et al., 2017), it was not surprising that HIV-1 URFs were generated through recombination between CRF01_AE and CRF07_BC in MSM.

Characterizing HIV subtypes and including transmission network analysis within epidemiological studies can shed light on the transmission patterns and inform prevention trials. In our previous cross-sectional molecular epidemiology study in Zhejiang province (Yang et al., 2017), we only found fragmental clusters, which may be due to the low sampling density. In this study, evaluation of putative transmission clusters using a comprehensive set of HIV-1 pol sequences in a

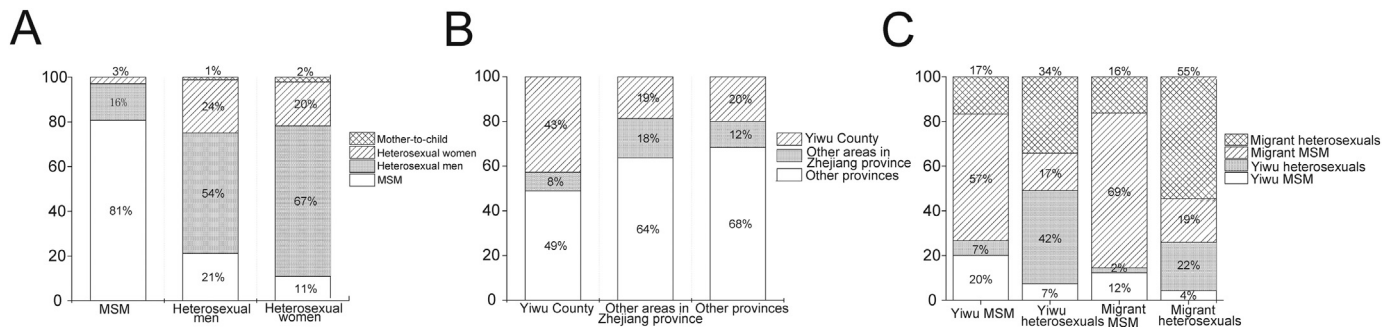


Fig. 3. The constitution of potential transmission partners of study subjects. (A) Subjects from different transmission categories; (B) Subjects from different locations of registration; (C) Subjects from different locations of registration with different transmission categories. MSM: men who have sex with men.

Table 2
Characteristics of the large molecular transmission clusters.

Subtypes	Cluster No.	Nodes	Genetic Distance ($\bar{x} \pm s$)	Age	Gender	Transmission route	Populations
CRF01_AE	C1	16	0.009 \pm 0.004	G1-2, G3-7, G4-7	M-12, F-4	Hetero	locals-6(1), migrants-10(1)
	C2	10	0.008 \pm 0.003	G1-2, G2-7, G3-1	M	Homo-8, Hetero-2	migrants-10(2)
	C3*	9	0.003 \pm 0.002	G1-6, G2-2, G3-1	M	Homo-7, Hetero-2	migrants-9(3)
	C4	8	0.008 \pm 0.004	G1-2, G2-3, G3-3	M	Homo-7, Hetero-1	locals-2, migrants-6
	C5	6	0.005 \pm 0.003	G2-3, G3-3	M-5, F-1	Hetero	locals-1, migrants-5
	C6*	6	0.007 \pm 0.004	G2-2, G3-3, G4-1	M	Homo-5, Hetero-1	locals-2(1), migrants-4(2)
	C7*	6	0.005 \pm 0.005	G3-5, G4-1	M-3, F-3	Hetero	locals-4(2), migrants-2(1)
	C8	5	0.001 \pm 0.001	G3-2, G4-3	M-4, F-1	Hetero	locals-5(1)
	C9	5	0.013 \pm 0.001	G1-2, G2-1, G3-2	M	Homo	locals-3(1), migrants-2
	C10	4	0.004 \pm 0.003	G3-2, G4-2	M	Hetero	locals-2, migrants-2(1)
CRF07_BC	C11	11	0.005 \pm 0.002	G1-2, G3-5, G4-4	M-8, F-3	Hetero-10, MTCT-1	locals-4(2), migrants-6
	C12	9	0.006 \pm 0.003	G2-2, G3-5, G4-2	M-8, F-1	Homo-5, Hetero-4	locals-2(1), migrants-7(1)
	C13*	6	0.004 \pm 0.002	G1-4, G2-2	M	Homo	locals-2(2), migrants-4(2)
	C14*	4	0.0008 \pm 0.001	G2-1, G3-3	M	Homo	locals-1(1), migrants-3(2)
URF(01_AE/07_BC)	C15	4	0.0013 \pm 0.001	G1-2, G2-1, G3-1	M	Homo	migrants-4(1)

G1: age group1(≤ 25 yrs), G2: age group2(26–34 yrs), G3: age group3(35–59 yrs), G4: age group4(≥ 60 yrs); M: male, F: female; Homo: homosexual contact, Hetero: heterosexual contact, MTCT: Mother-to-Child transmission. Active clusters are indicated by an asterisk. The number after the dash denotes the number of subjects. The number of subjects with recent infections is denoted in parentheses.

concentrated county identified significant local sexual transmission patterns among newly diagnosed individuals. An interweaved network of local residents and migrant populations was found in this study based on the analysis of potential transmission links. Moreover, most of the large clusters (73.3%, 11/15) in the network included both locals and migrants. Four active clusters underwent growth in new cases with recent infection of both locals and migrants. We noticed six individuals harboring the M46L mutation were found in a cluster (CRF01_AE-C5, Fig. 2A), hinting strongly their close transmission relationship. We also found high numbers of singleton lineages, which may indicate an important epidemiologic phenomenon driving HIV transmission in populations through migration or short-term travel. Our findings reveal that migration promotes the dissemination and dynamic change of HIV and leads to an interwoven network between locals and migrants.

A previous web-based study indicated that MSM migration in China may alter local MSM population sizes (Mi et al., 2016). In this study, we found that the influence of homosexual behavior in migrants on the local transmission network was greater than that of heterosexual behavior. Local heterosexuals had a higher percentage of linkage to local heterosexuals than to migrants (42% vs 22%, $P < .001$). However, migrant MSM had high percentages of linkage to both local MSM and migrant MSM (57% vs 69%, $P = .069$). Given the high-risk behavioral features of MSM, which include multiple casual sex partners (Zhang et al., 2012), and the dominance of migrants in the overall population, it is not surprising that a transmission network interconnecting the local MSM and migrant MSM tends to form. Previous studies reported the formation of a large social and sexual network (Li et al., 2018; Li et al., 2015; Zhang et al., 2017; Zhao et al., 2016), and the intertwined network of MSM in trade areas may be only a subset of a huge network with broader geographical scope. These analyses in our study offer unique insights into the heterogeneous impact of risk behavior on HIV-1 transmission, highlighting the far-reaching influence of migrant MSM on the local HIV epidemic.

Our results suggested that HIV infection among heterosexual women likely originated from two main sources. First, the majority of this group's links were heterosexual men (67% of all links). Second, a proportion of the links were to MSM (11% of all links). A high prevalence of bisexual behavior among MSM in China and 26.3% (95% CI: 23.6–29.1%) of MSM had female sexual partners in the last six months (Chow et al., 2011), suggesting a substantial proportion of HIV acquisition among heterosexual women is from bisexual MSM. Our study also found that heterosexual men were much more likely to fall into transmission networks with other heterosexual men. In this study, we cannot determine specific epidemiologic relationships between these men, but we propose two hypotheses: (1) because most of the men self-reported

as heterosexual, this finding may point to a significant underdiagnosis of HIV infection in “bridge women” who have sex with men, and (2) some men concealed the truth of their MSM behavior. If so, this supports the underestimation of infection status in women and the proportion of MSM in men.

Although we assume high sampling density ($> 80\%$) in our study, the inference we used to assess phylogenetic relationships among pol sequences represents a subset of the overall transmission network, given that only those with available sequences from 2014 to 2016 were included. Unavailable sequences due to population outflow and loss to follow-up would affect the understanding of the full transmission network based on molecular analysis to some extent. Additionally, we cannot confirm the direction of HIV transmission between locals and migrants using phylogenetic methods in our study. A preferable approach would be to integrate the epidemiological survey into a phylogenetic inference framework to confirm the transmission network.

In conclusion, we used phylogenetic and statistical analyses to delineate HIV-1 transmission in a typical trade area and explored the interaction between locals and migrants in the context of frequent population flow. The findings may provide inspiration for studies in other trade areas around the world.

Disclosure statement

The authors declare that they have no conflict of interest.

Funding

This work was supported by the Mega-Projects of National Science Research for the 13th Five-Year Plan (grant number: 2017ZX10201101-002), Mega-Projects of National Science Research for the 12th Five-Year Plan (grant number: 2012ZX10001-002) and Zhejiang Provincial Medicine Science and Technology Plan (grant number: 2016RCB004, 2017KY287).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful to the staff of Yiwu Municipal CDC for

their efforts in subject recruitment and sample collection. The authors are especially grateful to Prof. Ping Zhong for thoughtful comments on this manuscript. Dr. Yi Shen, Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, is also thanked for his kind assistance.

References

- Aldous, J.L., Pond, S.K., Poon, A., Jain, S., Qin, H., Kahn, J.S., Kitahata, M., Rodriguez, B., Dennis, A.M., Boswell, S.L., Haubrich, R., Smith, D.M., 2012. Characterizing HIV transmission networks across the United States. *Clin. Infect. Dis.* 55, 1135–1143.
- Aldrich, C., Hemelaar, J., 2012. Global HIV-1 diversity surveillance. *Trends Mol. Med.* 18, 691–694.
- Bennett, D.E., Camacho, R.J., Otelea, D., Kuritzkes, D.R., Fleury, H., Kiuchi, M., Heneine, W., Kantor, R., Jordan, M.R., Schapiro, J.M., Vandamme, A.M., Sandstrom, P., Boucher, C.A., van de Vijver, D., Rhee, S.Y., Liu, T.F., Pillay, D., Shafer, R.W., 2009. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* 4, e4724.
- Chan, P.A., Hogan, J.W., Huang, A., DeLong, A., Salemi, M., Mayer, K.H., Kantor, R., 2015. Phylogenetic investigation of a statewide HIV-1 epidemic reveals ongoing and active transmission networks among men who have sex with men. *J. Acquir. Immune Defic. Syndr.* 70, 428–435.
- Chen, M., Ma, Y., Chen, H., Dai, J., Dong, L., Yang, C., Li, Y., Luo, H., Zhang, R., Jin, X., Yang, L., Cheung, A.K.L., Jia, M., Song, Z., 2018. HIV-1 genetic transmission networks among men who have sex with men in Kunming, China. *PLoS One* 13, e0196548.
- Chow, E.P., Wilson, D.P., Zhang, L., 2011. What is the potential for bisexual men in China to act as a bridge of HIV transmission to the female population? Behavioural evidence from a systematic review and meta-analysis. *BMC Infect. Dis.* 11, 242.
- de Silva, T.I., Turner, R., Hue, S., Trikha, R., van Tienen, C., Onyango, C., Jaye, A., Foley, B., Whittle, H., Rowland-Jones, S.L., Cotten, M., 2010. HIV-1 subtype distribution in the Gambia and the significant presence of CRF49_cpx, a novel circulating recombinant form. *Retrovirology* 7, 82.
- Dennis, A.M., Herbeck, J.T., Brown, A.L., Kellam, P., de Oliveira, T., Pillay, D., Fraser, C., Cohen, M.S., 2014. Phylogenetic studies of transmission dynamics in generalized HIV epidemics: an essential tool where the burden is greatest? *J. Acquir. Immune Defic. Syndr.* 67, 181–195.
- Han, X., Takebe, Y., Zhang, W., An, M., Zhao, B., Hu, Q., Xu, J., Wu, H., Wu, J., Lu, L., Chen, X., Liang, S., Wang, Z., Yan, H., Fu, J., Cai, W., Zhuang, M., Liao, C., Shang, H., 2015. A large-scale survey of CRF55_01B from men-who-have-sex-with-men in China: implying the evolutionary history and public health impact. *Sci. Rep.* 5, 18147.
- Li, H., Bar, K.J., Wang, S., Decker, J.M., Chen, Y., Sun, C., Salazar-Gonzalez, J.F., Salazar, M.G., Learn, G.H., Morgan, C.J., Schumacher, J.E., Hraber, P., Giorgi, E.E., Bhattacharya, T., Korber, B.T., Perelson, A.S., Eron, J.J., Cohen, M.S., Hicks, C.B., Haynes, B.F., Markowitz, M., Keele, B.F., Hahn, B.H., Shaw, G.M., 2010. High multiplicity infection by HIV-1 in men who have sex with men. *PLoS Pathog.* 6, e1000890.
- Li, Z., Liao, L., Feng, Y., Zhang, J., Yan, J., He, C., Xu, W., Ruan, Y., Xing, H., Shao, Y., 2015. Trends of HIV subtypes and phylogenetic dynamics among young men who have sex with men in China, 2009–2014. *Sci. Rep.* 5, 16708.
- Li, X., Xue, Y., Lin, Y., Gai, J., Zhang, L., Cheng, H., Ning, Z., Zhou, L., Zhu, K., Vanham, G., Kang, L., Wang, Y., Zhuang, M., Pan, Q., Zhong, P., 2016. Evolutionary dynamics and complicated genetic transmission network patterns of HIV-1 CRF01_AE among MSM in Shanghai, China. *Sci. Rep.* 6, 34729.
- Li, X., Gao, R., Zhu, K., Wei, F., Fang, K., Li, W., Song, Y., Ge, Y., Ji, Y., Zhong, P., Wei, P., 2018. Genetic transmission networks reveal the transmission patterns of HIV-1 CRF01_AE in China. *Sex. Transm. Infect.* 94, 111–116.
- Little, S.J., Kosakovsky Pond, S.L., Anderson, C.M., Young, J.A., Wertheim, J.O., Mehta, S.R., May, S., Smith, D.M., 2014. Using HIV networks to inform real time prevention interventions. *PLoS One* 9, e98443.
- Liu, X., Erasmus, V., Wu, Q., Richardus, J.H., 2014. Behavioral and psychosocial interventions for HIV prevention in floating populations in China over the past decade: a systematic literature review and meta-analysis. *PLoS One* 9, e101006.
- Lyu, P., Chen, F.F., 2019. National HIV/AIDS epidemic estimation and interpretation in China. *Zhonghua Liu Xing Bing Xue Za Zhi* 40, 1191–1196.
- Mendelsohn, J.B., Calzavara, L., Light, L., Burchell, A.N., Ren, J., Kang, L., 2015. Design and implementation of a sexual health intervention for migrant construction workers situated in Shanghai, China. *Emerg. Themes Epidemiol.* 12, 16.
- Mi, G., Ma, B., Kleinman, N., Li, Z., Fuller, S., Bulterys, M., Hladik, W., Wu, Z., 2016. Hidden and mobile: a web-based study of migration patterns of men who have sex with men in China. *Clin. Infect. Dis.* 62, 1443–1447.
- National Health Commission, 2018. Report on China's Migrant Population Development. China Population Publishing House.
- Oster, A.M., Wertheim, J.O., Hernandez, A.L., Ocfemia, M.C., Saduvala, N., Hall, H.I., 2015. Using molecular HIV surveillance data to understand transmission between subpopulations in the United States. *J. Acquir. Immune Defic. Syndr.* 70, 444–451.
- Siepel, A.C., Halpern, A.L., Macken, C., Korber, B.T., 1995. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res. Hum. Retrovir.* 11, 1413–1416.
- Su, L., Liang, S., Hou, X., Zhong, P., Wei, D., Fu, Y., Ye, L., Xiong, L., Zeng, Y., Hu, Y., Yang, H., Wu, B., Zhang, L., Li, X., 2018. Impact of worker emigration on HIV epidemics in labour export areas: a molecular epidemiology investigation in Guangyuan, China. *Sci. Rep.* 8, 16046.
- UNAIDS, 2018. Global Factsheets 2018. <http://aidsinfo.unaids.org/>. https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf.
- Wertheim, J.O., Leigh Brown, A.J., Hepler, N.L., Mehta, S.R., Richman, D.D., Smith, D.M., Kosakovsky Pond, S.L., 2014. The global transmission network of HIV-1. *J. Infect. Dis.* 209, 304–313.
- Wertheim, J.O., Kosakovsky Pond, S.L., Forgiione, L.A., Mehta, S.R., Murrell, B., Shah, S., Smith, D.M., Scheffler, K., Torian, L.V., 2017. Social and genetic networks of HIV-1 transmission in New York City. *PLoS Pathog.* 13, e1006000.
- Yang, J.Z., Chen, W.J., Zhang, W.J., He, L., Zhang, J.F., Pan, X.H., 2017. Molecular epidemiology and transmission of HIV-1 infection in Zhejiang province, 2015. *Zhonghua Liu Xing Bing Xue Za Zhi* 38, 1551–1556.
- Zhang, L., Chow, E.P., Wilson, D.P., 2012. Distributions and trends in sexual behaviors and HIV incidence among men who have sex with men in China. *BMC Public Health* 12, 546.
- Zhang, J., Guo, Z., Yang, J., Pan, X., Jiang, J., Ding, X., Zhang, W., Xia, Y., Xu, Y., Huang, J., 2015. Genetic diversity of HIV-1 and transmitted drug resistance among newly diagnosed individuals with HIV infection in Hangzhou, China. *J. Med. Virol.* 87, 1668–1676.
- Zhang, J., Guo, Z., Pan, X., Zhang, W., Yang, J., Ding, X., Xu, Y., Xia, Y., Jiang, J., 2017. Highlighting the crucial role of Hangzhou in HIV-1 transmission among men who have sex with men in Zhejiang, China. *Sci. Rep.* 7, 13892.
- Zhao, J., Cai, W., Zheng, C., Yang, Z., Xin, R., Li, G., Wang, X., Chen, L., Zhong, P., Zhang, C., 2014. Origin and outbreak of HIV-1 CRF55_01B among MSM in Shenzhen, China. *J. Acquir. Immune Defic. Syndr.* 66, e65–e67.
- Zhao, J., Chen, L., Chaillon, A., Zheng, C., Cai, W., Yang, Z., Li, G., Gan, Y., Wang, X., Hu, Y., Zhong, P., Zhang, C., Smith, D.M., 2016. The dynamics of the HIV epidemic among men who have sex with men (MSM) from 2005 to 2012 in Shenzhen, China. *Sci. Rep.* 6, 28703.
- Zhao, S., Feng, Y., Xin, R., Li, Z., Hao, M., Liao, L., Ruan, Y., Lu, H., Shao, Y., Xin, H., 2018. Using molecular transmission network to explore the transmission characteristics of HIV-1 among men who have sex with men in Beijing. *Chin. J. AIDS STD* 24 (241–245), 306.