



OPEN Characterization of the HIV-1 molecular network in a middle-aged population aged 50 years and older in a City in Southern Sichuan, China

Tingchun Peng^{1,3,7}, Jiayang Tang^{1,4,7}, Miaomiao Qiu¹, Zhen Lai¹, Junguo Xin^{1,6}, Shu Liang², Chang Zhou², Jianping Deng⁵, Ying Zhang⁵, Yali Zeng², Ling Su²✉ & Xiaohong Yang^{1,6}✉

This study aimed to investigate the characteristics of the HIV-1 molecular network among newly diagnosed HIV-infected patients in southern Sichuan City. Plasma samples will be collected from eligible study subjects ($n=1249$) during a cross-sectional survey conducted between 2016 and 2020. The HIV-1 polymerase (pol) gene sequences obtained from the collected samples will be used to perform phylogenetic analysis and characterize the genetic subtypes' molecular transmission networks. HIV-1 pol region sequences were successfully amplified in 898 cases, and seven genotypes were obtained, with CRF01_AE (331, 36.86%) and CRF07_BC (368, 40.98%) subtypes as the predominant prevalent strains in the region. 601 sequences entered the molecular transmission network. There were 302 highly connected individuals. Further multivariate analysis showed that the older the age (60–69 years, OR = 1.595, 95% CI: 1.026–2.479; ≥ 70 years, OR = 2.189, 95% CI: 1.295–3.699), RX and GJ counties (OR = 4.654, 95% CI: 2.776–7.803; OR = 6.847, 95% CI: 3.464–13.533) and CRF08_BC subtype (OR = 2.031, 95% CI: 1.225–3.367) were both more likely to be highly connected individuals. To effectively combat this local HIV-1 epidemic, HIV prevention and intervention programs should target older adults at least 60 years of age and registered residents in districts and counties within RX and GJ.

Keywords HIV-1, Genotype, Molecular network, China, Risk factors, Seniors

The worldwide spread and epidemic of the human immunodeficiency virus (HIV-1) has become a pressing public issue for all nations¹ and remains a significant public health concern in China. According to the latest data from China's comprehensive AIDS prevention and treatment system, as of December 2021, there were 1,147,410 individuals living with HIV/AIDS in China². Over the past decade, there has been a steady increase in the number of newly reported cases of middle-aged and older individuals living with HIV and AIDS (referred to as HIV/AIDS hereafter), rising from 22% in 2011 to 44% in 2020². In other regions of China, the proportion of HIV/AIDS cases reported annually among individuals over 50 has also increased yearly^{3,4}. The aging population with HIV/AIDS increases the socio-economic burden and poses challenges in treatment and prognosis due to their weakened immunity and the onset of more complications⁵. As such, the "13th Five-Year Plan of Action for Curbing and Preventing HIV/AIDS in China" identified middle-aged and elderly individuals as one of the priority groups for HIV/AIDS prevention and treatment in 2017⁶.

Epidemiological surveys are the most direct and effective method for tracking transmission and identifying transmission clusters that represent recent and ongoing HIV transmission and can identify key characteristics of potential risk based on transmission clusters, which can help guide intervention efforts for ongoing transmission⁷.

¹School of Public Health, Chengdu Medical College, Chengdu, China. ²Center for AIDS/STD Control and Prevention, Sichuan Center for Disease Control and Prevention, Chengdu, China. ³Yibing Center for Disease Control and Prevention, Yibing, China. ⁴Hospital-Acquired Infection Control Department, Sichuan Cancer Hospital, Chengdu, China. ⁵Inspection department, Zigong Center for Disease Control and Prevention, Zigong, China. ⁶Sichuan Provincial Key Laboratory of Philosophy and Social Sciences for Intelligent Medical Care and Elderly Health Management, Chengdu, China. ⁷Tingchun Peng and Jiayang Tang have contributed equally to this work. ✉email: sl7512@163.com; 21376686@qq.com

However, traditional epidemiological surveys face many obstacles due to limited resources, privacy concerns, and discrimination, so exploring new models of HIV prevention interventions and tapping into the search for potential high-risk transmitters and high-risk infections in the population is essential to curb the spread of AIDS and reduce the rate of new HIV infections⁸. HIV molecular transmission networks are emerging methods for molecular epidemiological analysis of HIV in recent years. It combines bioinformatics and phylogenetics to reveal HIV transmission characteristics and epidemiological patterns⁹. HIV molecular network studies is an effective way to track transmission and identify transmission clusters¹⁰. Based on the transmission clustering, critical characteristics of potentially high-risk can be identified, guiding future interventions¹¹. In recent years domestic and international techniques for constructing HIV-1 molecular transmission networks using the viral gene sequences of infected individuals are gradually being developed, which can more accurately determine potential transmission and identify active transmission networks, thus contributing to targeted interventions^{9–12}.

The incidence and mortality rates of HIV/AIDS infection in Sichuan Province have remained high, making it the most painful area for the AIDS epidemic in China, consistently reporting higher numbers of surviving cases than other provinces and cities⁴. The epidemic among individuals aged 50 years and above in a southern Sichuan city is particularly severe, posing a more significant challenge than in other parts of the province¹³. This region has been a critical area for HIV prevention and control in recent years, with molecular epidemiological studies and precision-based prevention and control measures still in their early stages. Thus, the most pressing task for HIV prevention and control in the region is identifying high-risk populations, analyzing possible HIV-1 transmission events, and examining the epidemiological transmission characteristics of HIV/AIDS subtypes. With this in mind, this study focused on exploring the HIV-1 transmission characteristics and potential transmission risks in this area by analyzing the molecular network of HIV-1. The findings of this study can provide a scientific basis and guiding suggestions for the implementation of precise HIV prevention and control measures and preventive interventions for high-risk populations to reduce the rate of new HIV infections.

Methods

Object of the study

Basic demographic information and blood samples were collected from 2017 to 2020 from middle-aged and older HIV-1-infected patients aged 50 years or older who were newly reported and previously reported not on ART in 2016 in six districts and counties within a city in southern Sichuan Province. Inclusion criteria: Study subjects were required to meet the following criteria: household registration in a city in southern Sichuan, positive HIV-1 antibody test result by HIV supplementation test, not yet receiving HIV antiviral treatment, no mental illness and intellectual disability, and informed consent.

Research methodology

Sequence acquisition

Plasma samples collected from eligible study subjects from 2016 to 2020 were subjected to RNA extraction and purification, reverse transcription PCR, and nested PCR for two rounds of amplification, and the amplified products were submitted to a sequencing company for sequencing. In this study, cDNA was synthesized by one-step RT-PCR transcription and the HIV pol gene region was amplified using PCR premix reagents. The amplification product was 1.3 kb in length and the reference strain was the pol gene region of the international standard strain HXB2. Reverse transcription was performed using AMV Reverse Transcriptase enzyme and template RNA to synthesize cDNA, combined with MAW 26 and RT21 for the first round of RT-PCR amplification, and PRO-1 and RT20 as primers as well as the products of the first round of PCR as templates for the second round of PCR amplification.

(1) First round of PCR primers:

Outer upstream primer MAW 26 : 5'-TTGGAAATGTGGAAAGGAAGGAC-3'; HXB2 2028–2050;

Outer downstream primer RT21 : 5'-CTGTATTTCTGCTATTAAGTCTTTTGATGGG-3'; HXB2 3509–3539.

(2) Second round of PCR primers:

Medial upstream primer PRO-1 : 5'-CAGAGCCAACAGCCCCACCA-3'; HXB2 2147–2166.

Medial downstream primer RT20: 5'-CTTCTGTATATCCATTGACAGTCCAGCT-3'; HXB2 3299–3327.

Translated with DeepL.com (free version) The obtained sequences were cut and spliced using Sequencer 5.6. The spliced sequences were compared, clipped, and saved using Aliview 1.26. The resulting sequences were quality controlled using the WHO HIVDR QC Tool online tool (https://pssm.cfenet.ubc.ca/who_qc/) to determine whether the resulting sequences had inter-sample contamination, contamination between samples and laboratory strains, and assess the proportion of mixed bases. To proceed with subtyping and molecular network analysis, remove sequences with a length < 900 bp and a proportion of mixed bases greater than or equal to 15%.

Sequence typing

The compared sequences were submitted to the website HIV database (<https://www.hiv.lanl.gov/content/index>), and preliminary typing was performed using the HIV BLAST and RIP functions on the website tools page. Fasttree was applied to construct ML evolutionary trees to confirm the typing further. The phylogenetic tree was visualized and edited using Figtree 1.4.4, and the typing status of the samples was counted. Those clustered with

the reference strain and the support value of the node in the cluster $\geq 70\%$ were determined as the same subtype as the reference strain^{11,12} and compared with the typing results on the HIV database website.

Molecular communication network construction

The molecular transmission network was constructed using the gene distance threshold method, and the sequences that were collated and thoroughly compared were used to calculate the gene distances between sequences using the TN93 model in HyPhy software¹². In order to analyze the transmission association of more infected individuals and to include as many infected individuals as possible, the combined effects of gene distance threshold selection on molecular cluster resolution and entry rate, as well as subtypes, were fully considered, and the threshold with the most significant number of large molecular clusters and nodes was selected as the threshold for constructing the molecular network^{10,11}. Visualize and edit the molecular propagation network using Cytoscape 3.7.2. A node represents a patient. Degree value is the most critical parameter in transmission network analysis, defining highly connected individuals according to the results of degree value analysis; the higher the degree value means that the node corresponds to the more potential contacts of the infected person in the network, the higher the possibility of being the central figure of the network and the higher the possibility of transmitting HIV⁹, in this study, the degree value is skewed. The median six is used to define highly connected individuals¹⁰. Clusters are two or more nodes connected to form a cluster.

Statistical methods

Data were entered using EXCEL, and SPSS 26.0 was used for statistical analysis. Categorical data were analyzed descriptively using frequency and composition ratios, single factor analysis was performed using the chi-square test or Fisher's exact probability method, an unconditional logistic regression model was used to analyze the factors associated with HIV and the test level was set at $\alpha = 0.05$.

Ethical matters

All data in this study were kept under strict confidentiality measures, and particular researchers organized and analyzed samples and data information. The computers where the data were saved were password protected, special encrypted USB flash drives were used for data copying, and files delivered through the network were encrypted to ensure the safety of the data from leakage.

All experiments involving human participants were conducted in accordance with relevant guidelines and regulations. The study was approved by the Biomedical Ethics Committee of Chengdu Medical College (approval number: [(2021NO.04)]). Informed consent was obtained from all participants or their legal guardians prior to participation in the study.

Results

Basic characteristics and subtype distribution

The number of people newly reported in the city in 2017–2020 and previously reported before 2017 who did not start antiretroviral therapy in middle-aged and older adults over 50 years of age totaled 1253, and the final inclusion was consistent with 1249, with 898 cases successfully amplified to obtain HIV-1 pol region sequences, accounting for 71.90%. Genetic subtypes of different genotypes were clustered in the phylogenetic tree (Fig. 1), and seven were identified.

The study population was mainly aged 60–69 years (359,40.00%), male (673,74.90%), Han Chinese (895,99.70%), married (474,52.80%), elementary school educated (506,56.30%), with heterosexual transmission as the main route of infection (880,98.00%), and domicile in FS (343,38.10%) and RX (178,19.80%) counties. The molecular type was dominated by CRF01_AE (331,36.86%), CRF07_BC (368,40.98%), CRF08_BC (173,19.27%), followed by CRF85_BC (21,2.34%), CRF57_BC (3,0.33%), CRF64_BC (1,0.11%), and CRF55_01B (1,0.11%). The differences between the distribution of HIV-1 subtypes were statistically significant ($P < 0.05$) for age and domicile location but not for other characteristics. Age 60–69 years (48.90%) was higher for the CRF01_AE subtype, and age 50–59 years was predominant for the other subtypes. Several subtypes were distributed in all districts and counties, with the largest proportion in FS County. For details, see (Table 1) and (Fig. 2).

Characteristics of HIV-1 molecular transmission network

In order to construct the molecular propagation network diagram, the comprehensive influence of the selection of gene distance threshold on the resolution and access rate of molecular clusters was fully considered, and the optimal gene distance of 1.1% was selected to construct the molecular propagation network (Fig. 3A–B). There were 601 sequences entered the molecular propagation network, with a total entry rate of 66.90% (601/898), and the network consisted of 95 propagation clusters (39 CRF01_AE, 40 CRF07_BC, 14 CRF08_BC, and two others), 601 nodes, and 3144 edges. There were 48 macromolecular clusters (≥ 3 nodes) (16 CRF01_AE, 22 CRF07_BC, 9 CRF08_BC), with the number of individuals within the clusters ranging from 3 to 63; the range of each node degree was 1–57, with a median of 6. The most significant molecular cluster was subtype CRF01_AE cluster 1, consisting of 63 sequences, 50 males and 13 females, all of which were heterosexual contact infection transmission, mainly concentrated in RX (80.90%), with a degree range of 2–57 and a median of 43. Infected individuals entering the molecular network were mainly 60–69 years old, male, Han Chinese, married, heterosexual transmission, elementary school education, and FS county. The primary concentration areas of macromolecular clusters were in RX and FS, followed by GJ. The differences in the entry rate of each genotype were statistically significant, with the highest entry rate of CRF01_AE (Table 1).

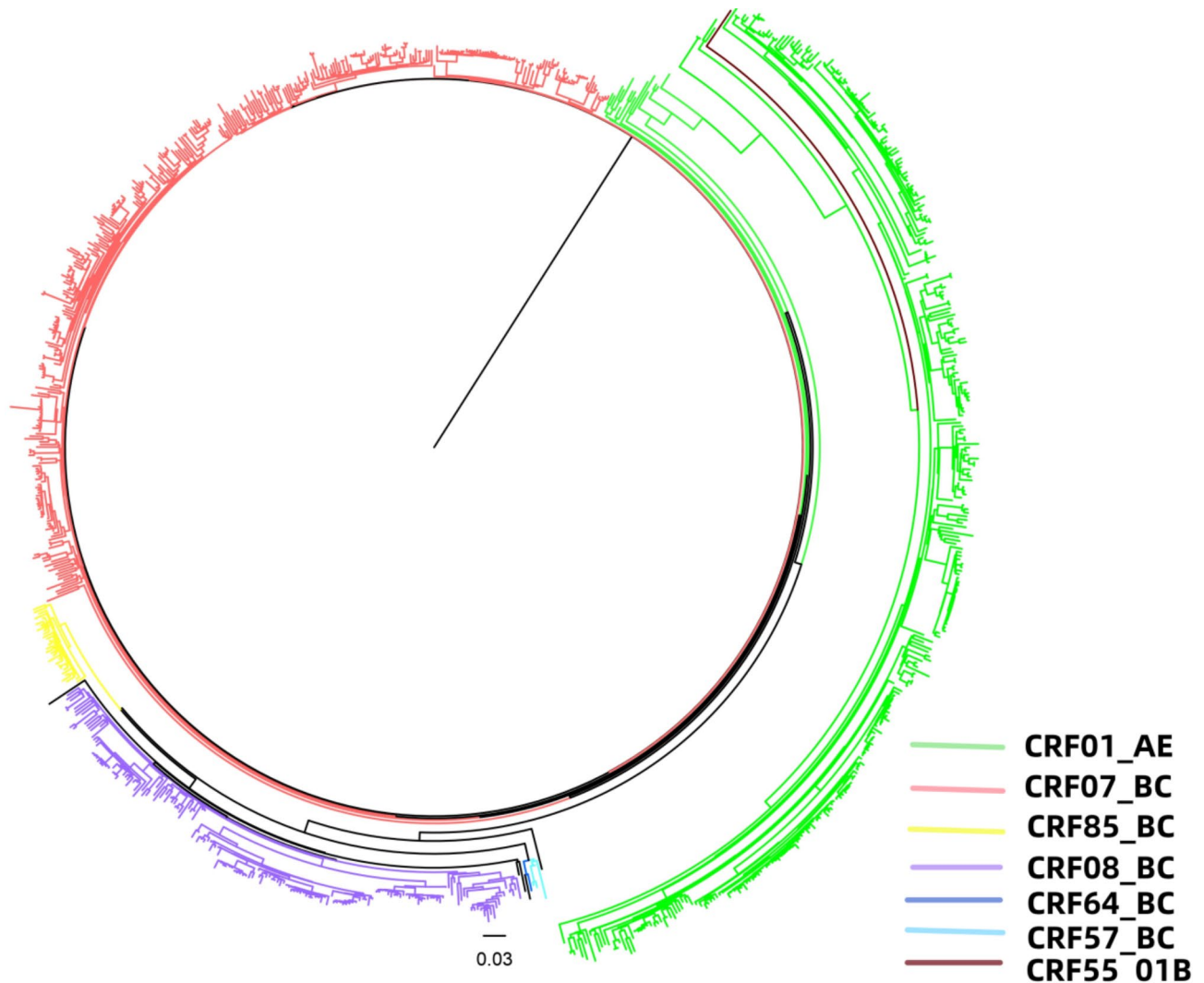


Fig. 1. Maximum likelihood phylogenetic tree of 898 *pol* sequences of HIV-1 in a city in southern Sichuan. 331 HIV-1 sequences were branched with the CRF01_AE reference sequence (depicted in green), 368 HIV-1 sequences were branched with the CRF07_BC reference sequence (depicted in pink), 173 HIV-1 sequences were clustered with the CRF08_BC reference sequence (depicted in purple), and 21 HIV-1 CRF85_BC sequences were identified (depicted in yellow), 3 CRF57_BC sequences (depicted in light blue), 1 CRF64_BC sequence (depicted in dark blue), and 1 CRF55_01B (depicted in dark brown) were also identified—the reference sequences from the Los Alamos HIV sequence database. On the upper edge of the CRF01_AE group, a number of seemingly ambiguous sequences were typed as well as verified and confirmed by the HIV BLAST and RIP functions in the HIV database page, all of which were CRF01_AE.

Analysis of factors influencing the molecular transmission network of HIV-1

Univariate and multifactor analyses were conducted with whether or not they were enrolled in the network as the dependent variable and age, gender, ethnicity, marriage, and route of infection as the independent variables. The results of the univariate analysis were combined with professional significance, and the influencing factors with $P < 0.10$ were finally selected as independent variables for inclusion in the multifactor analysis. The results of multifactorial logistic regression analysis showed that older age was more likely to be in the network (60–69, OR = 1.552, 95% CI: 1.103–2.184; ≥ 70 , OR = 2.259, 95% CI: 1.466–3.481), and CRF01_AE was more likely to be in the network compared to CRF07_BC and other subtypes. For details, see (Table 2).

Analysis of factors affecting individuals with high connectivity

Of the 601 subjects who entered the molecular network, 302 had a degree ≥ 6 and were defined as highly connected individuals in the molecular network. All highly connected individuals were infected through heterosexual transmission and were Han Chinese. Univariate analysis showed statistically significant differences in the distribution of highly connected individuals by age, sex, marriage, education, location, and genotype; further multivariate analysis showed that the older the age (compared with 50–59 years, 60–69 years, OR = 1.595, 95% CI: 1.026–2.479; ≥ 70 years, OR = 2.189, 95% CI: 1.295–3.699), two counties with household registration

Variables	Total(% ^a) n = 898	The frequency and percentage of different HIV-1 genotypes(% ^b)				Z/ χ^2	P-value ^f
		01_AE	07_BC	08_BC	Others ^c		
Age (years)						9.348	0.025
50-	334(37.20)	93(28.10)	158(42.90)	71(41.00)	12(46.20)		
60-	359(40.00)	162(48.90)	128(34.80)	64(37.00)	5(19.20)		
≥ 70	205(22.80)	76(23.00)	82(22.30)	38(22.00)	9(34.60)		
Gender						0.677	0.879
Male	673(74.90)	245(74.00)	276(75.00)	131(75.70)	21(80.80)		
Female	225(25.10)	86(26.00)	92(25.00)	42(24.30)	5(19.20)		
Ethnicity						1.436	0.697
Han	895(99.70)	329(99.40)	367(99.70)	173(100.00)	26(100.00)		
Others	3(0.30)	2(0.60)	1(0.30)	0(0.00)	0(0.00)		
Marital Status						2.767	0.837
Unmarried	110(12.20)	42(12.70)	44(12.00)	23(13.30)	1(3.80)		
Married	474(52.80)	168(50.80)	199(54.10)	92(53.20)	15(57.70)		
Divorced /widowed	314(35.00)	121(36.50)	125(34.00)	58(33.50)	10(38.50)		
Route of infection ^d						13.365	0.147
HST	880(98.00)	325(98.20)	357(97.00)	173(100.00)	25(96.20)		
MSM	9(1.00)	3(0.90)	6(1.60)	0(0.00)	0(0.00)		
IDU	3(0.30)	0(0.00)	3(0.80)	0(0.00)	0(0.00)		
Others/unknown	6(0.70)	3(0.90)	2(0.60)	0(0.00)	1(3.80)		
Education Level						4.494	0.213
Illiteracy	152(16.90)	62(18.70)	59(16.00)	27(15.60)	4(15.40)		
Primary School	506(56.30)	195(58.90)	200(54.40)	97(56.10)	14(53.80)		
Junior High School	202(22.50)	60(18.10)	94(25.50)	41(23.70)	7(26.90)		
High School	33(3.70)	13(3.90)	11(3.00)	8(4.60)	1(3.90)		
College and above	5(0.60)	1(0.30)	4(1.10)	0(0.00)	0(0.00)		
Domicile Location						97.791	<0.001
FS	343(38.20)	152(45.90)	119(32.30)	62(35.90)	10(38.50)		
RX	178(19.80)	88(26.60)	80(21.70)	7(4.00)	3(11.50)		
DA	91(10.10)	20(6.00)	38(10.30)	27(15.60)	6(23.10)		
ZLJ	107(11.90)	30(9.10)	41(11.10)	33(19.10)	3(11.50)		
GJ	91(10.10)	17(5.10)	58(15.80)	14(8.10)	2(7.70)		
YT	88(9.90)	24(7.30)	32(8.70)	30(17.30)	2(7.70)		
First CD4 (pcs/ul)						2.118	0.548
≤ 200	384(42.80)	148(44.70)	152(41.30)	72(41.60)	12(46.20)		
200	448(49.90)	164(49.50)	189(51.40)	82(47.40)	13(50.00)		
≥ 500	57(6.30)	17(5.10)	25(6.80)	14(8.10)	1(3.80)		
Unknown ^e	9(1.00)	2(0.60)	2(0.50)	5(2.90)	0(0.00)		
Year of diagnosis						20.601	0.057
≤ 2016	35(3.90)	10(3.00)	21(5.70)	4(2.30)	0(0.00)		
2017	162(17.70)	50(15.10)	61(16.50)	42(24.80)	8(30.80)		
2018	260(28.80)	106(32.00)	103(28.00)	46(27.20)	4(15.40)		
2019	246(27.20)	87(26.30)	104(28.30)	45(26.00)	10(38.40)		
2020	195(21.70)	78(23.60)	79(21.50)	34(19.70)	4(15.40)		
Whether to enter the network						37.766	<0.001
Yes	601(66.90)	253(76.40)	207(56.30)	127(73.40)	14(53.80)		
No	297(33.10)	78(23.60)	161(43.70)	46(26.60)	12(46.20)		

Table 1. Analysis of the general demographic characteristics of each genetic subtype of the study population.

^aNumbers in parentheses indicate cases as a percentage of the total 898 study subjects. ^bThe numbers in parentheses indicate the percentage of subtypes to each variable. ^cOther subtypes include 21 cases of CRF85_BC, 3 cases of CRF57_BC, 1 case of CRF64_BC, 1 case of CRF55_01B. ^dHST: heterosexual, MSM: men who have sex with men, IDU: injection drug use. ^eNo CD4 test was done. ^fP-value < 0.05 was considered statistically significant.

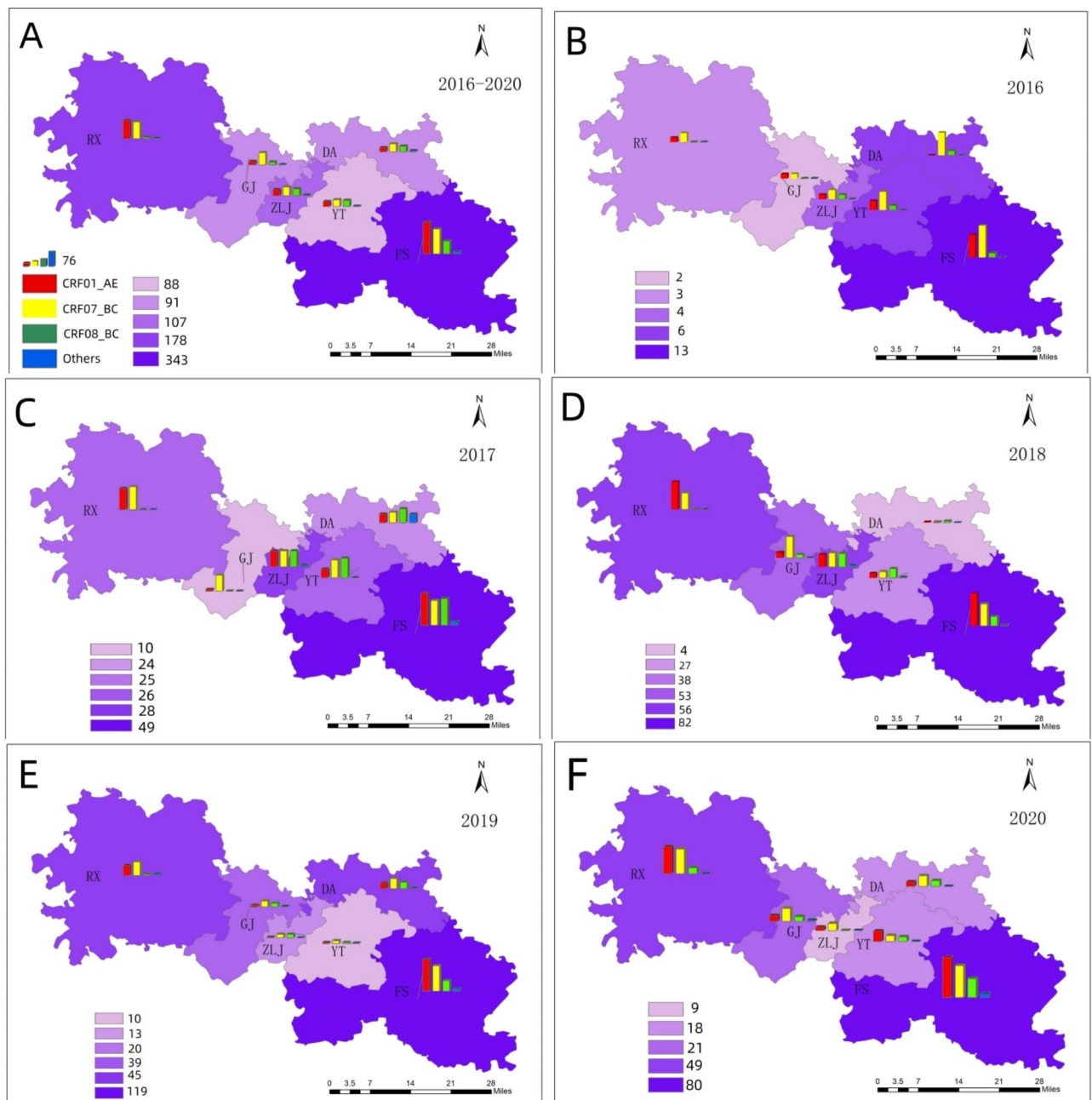


Fig. 2. Distribution of HIV-1 genotypes in different areas of a city in southern Sichuan: A, B, C, D, E, and F. The six maps represent the distribution of the overall and the number of genetic subtypes in each year from 2016 to 2020, respectively. Different colors of the bar graph represent different subtypes, and the height of the bar graph represents the number of each subtype, with the highest value being 76; the darker the background color of the map, the higher the number of infected people in the region.

of RX and GJ (compared to FS county, OR=4.654, 95% CI: 2.776–7.803; OR=6.847, 95% CI: 3.464–13.533) and CRF08_BC subtype (compared to CRF01_AE, OR=2.031, 95% CI: 1.225–3.367) were both more likely to be highly connected individuals, i.e., more inclined to connect six and more transmitters. High school literacy was less likely to be a highly connected individual compared to illiterate (OR=0.262, 95% CI: 0.086–0.799), CRF07_BC subtype than CRF01_AE (OR=0.415, 95% CI: 0.267–0.647). See Fig. 4 for details.

Discussion

The current study found that the newly diagnosed infections among middle-aged and older people aged 50 years and above in this region in 2017–2020 were predominantly male, heterosexual transmission, with low literacy (73.20% were in elementary school and below). Their domicile was mainly in two districts, RX and FS¹³, which is consistent with the results of previous survey studies, indicating that in these two districts, more emphasis

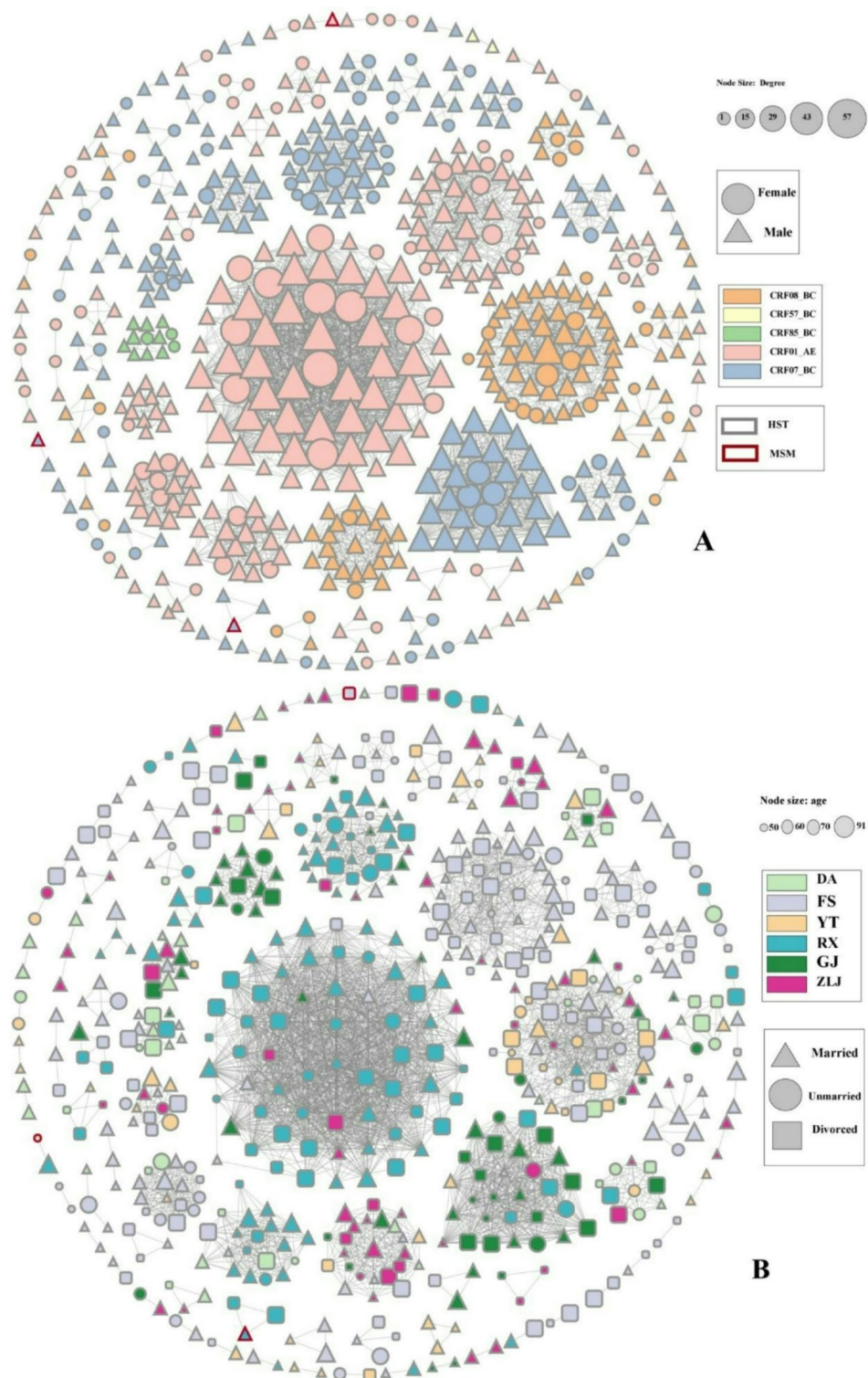


Fig. 3. HIV-1 molecular transmission networks. (A) HIV-1 subtype and gender, route of infection molecular transmission network, nodes of different colors represent different subtypes, different shapes represent different genders, different border colors represent different routes of infection, larger nodes represent higher degree values, values range 1–57. (B) molecular transmission network diagram of different domicile locations and marital status, age, nodes of different colors represent different regions, different shapes represent different marital status, the larger the node represents the older the age, value range 50–91. divorced: divorced or widowed.

Variables	Total number of incoming networks(%) <i>N</i> = 601	Single factor analysis		Multi-factor logistic analysis	
		<i>Z</i> / χ^2	<i>P</i>	OR(95%CI)	<i>P</i> -values
Age (years)		-5.321	0.000		
50-	188(56.30)			1	
60-	255(71.00)			1.552(1.103–2.184)	0.012
≥ 70	158(77.10)			2.259(1.466–3.481)	< 0.001
Gender		3.596	0.058		
Male	462(68.60)			1	
Female	139(61.80)			0.831(0.592–1.168)	0.286
Ethnicity		1.535	0.215		
Han	600(67.00)				
Others	1(33.30)				
Marital Status		1.399	0.497		
Unmarried	75(68.20)				
Married	309(65.20)				
Divorced /widowed	217(69.20)				
Route of infection		10.788	0.005		
HST	598(67.50)			1	
MSM	3(33.30)			0.268(0.061–1.173)	0.080
IDU	0(0.00)			0.000(0.000)	0.999
Education Level		-2.701	0.007		
Illiteracy	106(69.70)			1	
Primary School	349(69.00)			1.158(0.761–1.762)	0.494
Junior High School	121(59.90)			0.952(0.586–1.548)	0.843
High School	23(69.70)			1.323(0.544–3.218)	0.537
College and above	2(40.00)			0.364(0.053–2.479)	0.302
Domicile Location		7.111	0.213		
FS	221(64.40)				
RX	126(70.80)				
DA	54(59.30)				
ZLJ	73(68.20)				
GJ	68(74.70)				
YT	59(67.00)				
First CD4 (pcs/ul)		-1.250	0.211		
≤ 200	247(64.30)				
200	310(69.20)				
≥ 500	39(68.40)				
Unknown	5(55.60)				
Year of diagnosis		21.715	0.000		
≤ 2016	17(48.60)			1	
2017	119(73.50)			2.034(0.884–4.680)	0.095
2018	190(73.10)			1.958(0.881–4.351)	0.099
2019	164(66.70)			1.474(0.665–3.268)	0.339
2020	111(56.90)			0.990(0.444–2.210)	0.981
HIV-1 Genotypes		37.766	0.000		
01_AE	253(76.40)			1	
07_BC	207(56.30)			0.427(0.304–0.600)	< 0.001
08_BC	127(73.40)			0.855(0.553–1.323)	0.482
Others	14(53.80)			0.334(0.143–0.780)	0.011

Table 2. Analysis of factors influencing HIV-1 subtype entry into the molecular network. ^aThe numbers in parentheses indicate the percentage of cases to their respective variables of the same type.

should be placed on the middle-aged and older male population with low education. The HIV prevalent strains among HIV-infected patients in this region were multiple recombinant genetic subtypes, with CRF01_AE (331, 36.86%) and CRF07_BC (368, 40.98%) being the predominant prevalent strains, followed by CRF08_BC (173, 19.27%), CRF85_BC (21, 2.34%), CRF57_BC (3, 0.33%), CRF64_BC (1, 0.11%), CRF55_01B (1, 0.11%), which is

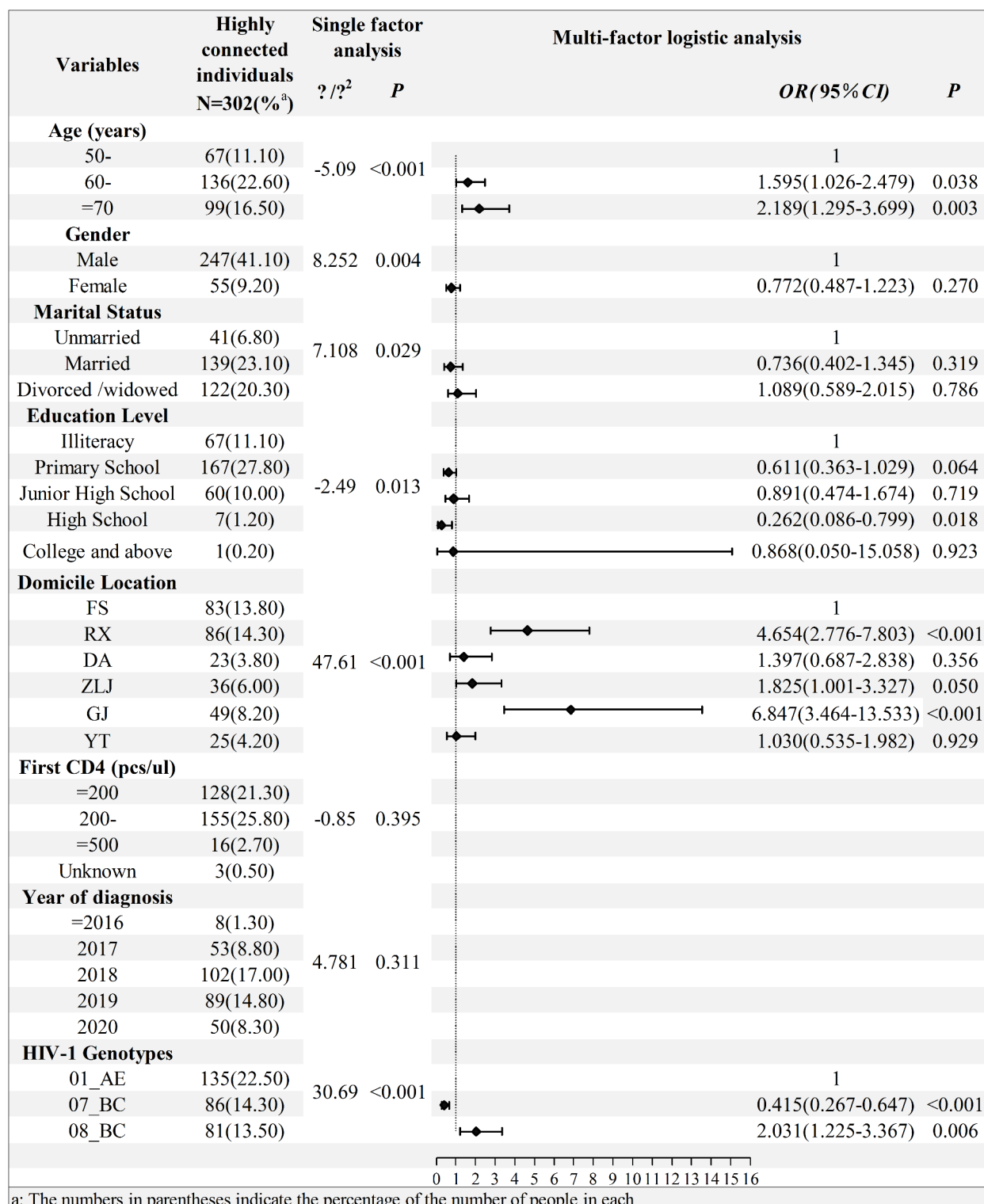


Fig. 4. Analysis of influencing factors of highly connected individuals.

consistent with previous national and provincial HIV molecular epidemiological findings^{14,15}, which is different from other regions where subtypes B (Henan) and CRF08_BC (Yunnan) are the dominant strains¹⁶. CRF01_AE was identified in the 1990s from Thailand and imported to southwestern China in the 1990s, with predominant heterosexual population transmission¹⁷, CRF07_BC with the predominant heterosexual and homosexual transmission, and both subtypes have since become prevalent among MSM, and CRF08_BC with predominantly heterosexual transmission^{10,18}. The results of this study are generally consistent with the routes of infection described in published articles, which also suggest that there may be a specific size of local MSM groups and

HIV infections that have not yet been detected. They may contribute to the local transmission and prevalence of CRF01_AE and CRF07_BC strains together with heterosexuals. In addition, a case of CRF55_01B subtype newly emerged in China in recent years was found. CRF55_01B was first reported to be found in the MSM population¹⁹. The CRF55_01B subtype found there was heterosexual transmission, suggesting the emergence of a new strain among local heterosexuals, which may also be caused by MSM concealing the infection route and other behaviors due to social stigma and shame²⁰ and the specific reasons for this need to be further investigated.

601 HIV-infected patients entered the molecular network, with an overall entry rate of 66.90%, higher than some similar studies¹⁵, probably due to the different genetic distance thresholds selected or the different methods used to construct the molecular network. The transmission chains in the HIV-1 molecular network do not represent direct or indirect transmission relationships. However, the formation of transmission networks is associated with the sociological and behavioral characteristics of infected individuals within the transmission clusters¹⁰. Identifying local HIV-1 transmission clusters may provide a scientific basis for precise interventions. The results of this study showed that older age was more likely to enter the network compared to 50–59 years, and the risk level increased with age (60–69 years, OR = 1.552, 95% CI: 1.103–2.184; ≥70 years, OR = 2.259, 95% CI: 1.466–3.481). The CRF01_AE subtype was also another contributing factor to the formation of the network at this site. Regarding the most extensive transmission cluster, CRF01_AE-infected individuals formed the largest transmission cluster (63 nodes) and were most likely to have large clusters of aggregation. Hence, the risk of transmission was higher for those infected in the large CRF01_AE cluster. The vast majority of transmission modes entering the transmission network in this study were heterosexual infections. However, the study in the Liaoning region showed that it was mainly clustered in homosexual transmission²¹, while the study by LIU et al.²² showed that factors such as gender, transmission mode, and education level were not significantly correlated with entry into the molecular transmission network, suggesting that the transmission network situation varies in different regions. Targeted prevention and control measures need to be proposed for different regions.

The results of logistic regression analysis of highly connected individuals showed that age ≥ 60 years, domicile in RX and GJ, and CRF08_BC subtype were risk factors for being highly connected individuals, and high school education and CRF07_BC subtype were less likely to be highly connected individuals. The findings that older infected individuals are more likely to be highly connected are inconsistent with previous studies in which younger MSM was more likely to be high-risk transmitters²³, which may be related to the local demographic and economic environment and the lack of a scaled-up MSM organization. The problem of aging is more prominent in the local area, especially the older people over 60 years old who are less mobile and generally stay in the local area. Hence, the corresponding scope of sexual activities and objects is also in the local area^{13,24}, with the loss of their labor force, lack of economic resources, and lower economic level. At the same time, this population mainly lives in rural areas; their lives are more monotonous and empty²⁵, influenced by the local “Banbancha” culture²⁴, which makes it easier to consume in low-grade sex trade places²⁶, increases the risk of HIV infection in middle-aged and elderly people, making them more likely to be highly connected individuals, which may also be one of the reasons for the higher rate of local molecular network clustering in this age group than in other studies. It may also be related to the generally low literacy level of this local middle-aged and elderly population, their insufficient knowledge of HIV-related knowledge²⁷, the difficulties in prevention intervention efforts for them, and the lack of awareness of condom use during sexual intercourse in this population²⁸. Compared to those with low literacy, those with high literacy have a higher awareness of HIV-related protection and testing²⁹, making them less likely to be highly connected. Therefore, it is suggested that special interventions should be carried out for the elderly with low literacy and aged ≥ 60 years, especially for HIV-infected people. The intervention models such as “social self-worth intervention” or “health education based on empowerment theory” for middle-aged and elderly HIV-infected patients by Bhatta³⁰ and Peng³¹ can be used to improve knowledge about HIV, condom use, and treatment compliance, and have specific reference value for reducing the possibility of second-generation transmission among local middle-aged and elderly infected patients. At the same time, by combining the analysis of infection route and subtype strains, and because there is no formal and influential MSM organization in the local area, we cannot rely on the corresponding social organizations to conduct mapping and targeted interventions for the MSM population. Because of social stigma and psychological pressure, we cannot exclude the existence of some hidden MSM in the local area. Therefore, the local community should better intervene, prevent, and control the MSM population. The infected individuals in RX and GJ are more likely to be highly connected, probably because there are more surviving infected individuals in RX, the two counties are adjacent to each other, and GJ is a central urban area with a developed economy. There may be more sex service establishments compared to other areas²³, which increases the risk of HIV transmission.

There are some shortcomings in this study, some infected samples did not obtain qualified sequences due to storage time and sequencing quality, and we can only analyze samples that have been diagnosed; those who have been infected but not diagnosed were not included in the analysis, resulting in a possible deviation of the inclusion rate of this study from the actual situation and limited representativeness.

Conclusion

The present study provided distribution of HIV-1 genotypes in middle-aged and older people aged 50 years and above in a city in southern Sichuan, which were more complex and diverse than before. The dominant local strains were CRF01_AE and CRF07_BC, followed by CRF08_BC and other genotypes. The overall molecular entry rate was 66.90%, and the factors affecting entry were age and genetic subtypes. Age ≥ 60 years, domicile in RX and GJ, and CRF08_BC subtype were more likely to be highly connected individuals. The MSM population may be underestimated, suggesting that intervention behaviors in this particular group should be expanded in addition to interventions in the above critical heterosexually transmitted elderly population.

In summary, this study revealed the characteristics of HIV-1 transmission in a city in southern Sichuan from the population level by constructing a molecular transmission network and assessing the transmission

risk and early warning based on the molecular network situation, suggesting the combination of molecular network analysis methods and public health methods to achieve the prevention and control effect of the precise intervention.

Data availability

Availability of data and materials The datasets generated and/or analyzed during the current study are not publicly available due to privacy concerns related to AIDS patients. In accordance with ethical approval requirements, the data are subject to strict privacy protections and cannot be publicly stored. However, the data can be made available from the corresponding author upon reasonable request. For further information or access to the data, readers may contact the corresponding author at [21376686@qq.com]. Data is deposited in National Microbiology Data Center (NMDC) with accession numbers NMDCN0006LMD (<https://nmhc.cn/resource/genomics/sequence/detail/NMDCN0006LMD>).

Received: 26 November 2024; Accepted: 24 March 2025

Published online: 26 March 2025

References

- UNAIDS. 2018 GLOBAL HIV STATISTICS. 2018, [<http://aidsinfo.unaids.org>]
- China Center for Disease Control and Prevention. National annual AIDS statistics by region. 2022, [(2021). http://www.phscience.com.cn/Share/ky_sjml.jsp?id=c2ca694e-3995-4c7f-9078-3ed0aaf14556]
- Wu, G. et al. Higher risks of virologic failure and All-Cause deaths among older people living with HIV in Chongqing, China. *AIDS Res. Hum. Retroviruses*. **35** (11–12), 1095–1102 (2019).
- Yuan, F. et al. Epidemiological and Spatiotemporal analyses of HIV/AIDS prevalence among older adults in Sichuan, China between 2008 and 2019: A population-based study. *Int. J. Infect. Dis.* **105**, 769–775 (2021).
- Ham, L. et al. Four-Year trajectories of internal strengths and socioemotional support among Middle-Aged and older adults with HIV. *AIDS Behav.* **27** (2), 628–640 (2023).
- General Office of the State Council. Notice of the General Office of the State Council on the issuance of China's 13th Five-Year Plan of Action for Curbing and Preventing AIDS. 2023, [https://www.gov.cn/gongbao/content/2017/content_5171331.htm]
- Wei, L. et al. Epidemiological characteristics of HIV infection among college students in Nanjing, China: a cross-sectional survey. *BMJ Open*. **10** (5), e035889 (2020).
- Laura, P. et al. HIV general population surveys: Shedding light on the status of HIV epidemics and informing future Actions. *JAIDS J. Acquir. Immune Defic. Syndr.* **87** (Suppl 1), S2–S5 (2021).
- Chen, H. et al. Molecular transmission network of pretreatment drug resistance among human immunodeficiency virus-positive individuals and the impact of virological failure on those who received antiretroviral therapy in China. *Front Med (Lausanne)*, 9(965836). (2022).
- Li, K. et al. Using molecular transmission networks to understand the epidemic characteristics of HIV-1 CRF08_BC across China. *Emerg. Microbes Infect.* **10** (1), 497–506 (2021).
- Rose, R. et al. Identifying transmission clusters with cluster picker and HIV-TRACE. *AIDS Res. Hum. Retroviruses*. **33** (3), 211–218 (2017).
- Zhang, J. et al. Combining molecular network analysis and field epidemiology to quantify local HIV transmission and highlight ongoing epidemics. *Int. J. Infect. Dis.* **128**, 187–193 (2023).
- Zhang, Y. et al. Analysis of new infections among newly reported HIV/AIDS patients in Zigong City from 2018 to 2021. *Journal of Preventive Medicine Information*, 39(09)(1019–1024. [(2023). <https://link.cnki.net/urlid/51.1276.R.20230403.2313.004>].
- Zhou, C. et al. Characterization of HIV-1 molecular epidemiology and transmitted drug-resistance in newly diagnosed HIV-infected patients in Sichuan, China. *BMC Infect. Dis.* **22** (1), 602 (2022).
- Zhao, B. et al. Molecular Network Analysis Reveals Transmission of HIV-1 Drug-Resistant Strains Among Newly Diagnosed HIV-1 Infections in a Moderately HIV Endemic City in China. *Front Microbiol.*, 12(797771). (2021).
- Li, L. et al. Different distribution of HIV-1 subtype and drug resistance were found among treatment Naïve individuals in Henan, Guangxi, and Yunnan Province of China. *PLoS One*. **8** (10), e75777 (2013).
- Feng, Y. et al. The rapidly expanding CRF01_AE epidemic in China is driven by multiple lineages of HIV-1 viruses introduced in the 1990s. *AIDS* **27** (11), 1793–1802 (2013).
- Ge, Z. et al. HIV-1 CRF07_BC transmission dynamics in China: Two decades of National molecular surveillance. *Emerg. Microbes Infect.* **10** (1), 1919–1930 (2021).
- Xiaoxu, H. et al. Genome sequences of a novel HIV-1 Circulating Recombinant form, CRF5501B, identified in China. *Genome Announcements*. **1** (1), e00050–e00012 (2013).
- Ndione, A. G. et al. Sexuality-based stigma and access to care: Intersecting perspectives between healthcare providers and men who have sex with men in HIV care centres in Senegal. *Health Policy Plan.* **37** (5), 587–596 (2022).
- Ning, M. et al. HIV-1 molecular transmission network among sexually transmitted populations in Liaoning Province, China. *Medicine* **100** (28), e26640–e26640 (2021).
- Liu, L. et al. Survey of pretreatment HIV drug resistance and genetic transmission network analysis among HIV patients in a high drug-Use area of Southwest China. *Curr. HIV Res.* **17** (6), 441–451 (2019).
- Li, Q. H. et al. Young MSM changed Temporal HIV-1 epidemic pattern in Heilongjiang Province, China. *Front. Microbiol.*, **13** 131028383. (2022).
- Zhang, B. et al. 2019–2021 Zigong City MSM population characteristics and AIDS knowledge and behavior survey. *Applied Preventive Medicine*, 28(03)(275–278. [(2022). https://kns.cnki.net/kcms2/article/abstract?v=1i9K2Vab0pNZbB6-niGfwKp7UsSO4P2q2kTyYFvISqnl2PJxencMHXabsxsauiGyBrDF6NOPXf7s1d3O0qP_1GyFtnRpEvNRU2go3yXXR_7u5FkGOprfNoM3xOpIIRCjDrQ2hX2BmHBVuat6CMoP2YmnzGAndijy0HvvUGY-T764embKi_hX5E5J4uYkRByOkMTogcUE=&uniplatform=NZKPT&language=CHS].
- Yang, L. & Zhang, X. Research on the urban-rural differences and influencing factors of life satisfaction among elderly people in China. *J. Health Popul. Nutr.* **43** (1), 213 (2024).
- Gao, Y. & Lu, H. Research progress on the influencing factors of AIDS-related high-risk sexual behavior in the elderly. *China AIDS STD*. **26** (04), 452–454 (2020). <https://link.cnki.net/doi/10.13419/j.cnki.aids.2020.04.31>
- Aguiar, R., Leal, M. & Marques, A. Knowledge and attitudes about sexuality in the elderly with HIV. *Cien Saude Colet.* **25** (6), 2051–2062 (2020).
- Stephen, E. & Joseph, L. Age is not a condom: HIV and sexual health for older adults. *Curr. Sex. Health Rep.* **9** (3), 109–115 (2017).
- Lina, H. Survey and analysis of HIV/AIDS prevention and treatment knowledge and cognition among people with different education levels. *Henan Journal of Preventive Medicine*, 30(06)(450–452. [(2019). <https://link.cnki.net/doi/10.13515/j.cnki.hnjpm.1006-8414.2019.06.016>].

30. Bhatta, D. N. & Liabsuetrakul, T. Social self-value intervention for empowerment of HIV infected people using antiretroviral treatment: a randomized controlled trial. *Bmc Infect. Dis.* , **16**(1): 272–279 (2016).
31. Peng, S. et al. The Effect of Health Education Based on Empowerment Theory on AIDS KAP in Middle-aged and Older Adults in the Community. *Modern Clinical Medicine*, 48(01):1–5. [(2022). https://kns.cnki.net/kcms2/article/abstract?v=1i9K2Vab0pPL97RFwKyXYWiB4Hacyf5mGEjJNpn8Fjc79kuVd2EiLP0sLYY9uWOZb71nWJ7fSlvmqb0Xn1zlfCDmuLLtBGsXy7bhkc6RVOej-JP uSqG141GwdLIucrnOU2b0EXfwVpREjrj6zlE8PVHOTL1Pn8U2MDL_HNHp2IE9k5ZAo3-apj2p5kYOekzf8X5j221sxc=&uniplatform=NZKPT&language=CHS].

Acknowledgements

We acknowledge the work of staffs involved in collecting patient blood sample the staff of 21 municipal CDCs. We appreciate all the volunteer patients enrolled in this study.

Author contributions

LS and XHY contributed equally to this article as co-corresponding authors. JYT and LS designed the study. JYT, TCP, XHY, MMQ, ZL and JGX participated in the process of sequences editing and phylogenetic analyses. JYT and TCP performed statistical analysis. JYT, SL, CZ, YZ, JPD, YLZ and LS collected sample and demographic data. JYT and TCP participated in the writing process. All authors read and approved the final manuscript.

Funding

This study was supported by the Natural Science Foundation of Sichuan Province (2022NSFSC1547), the Joint Fund of Sichuan Collaborative Innovation Center for Aging and Elderly Health(YLKYYB2211), and the Chengdu Medical College-Yibin Center for Disease Control and Prevention Joint Research Fund(2021LHYB01 and 2023LHYB04).

Declarations

Ethics approval and consent to participate

This study was approved by the Biomedical Ethics Committee of Chengdu Medical College(2021NO.04).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.S. or X.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025