scientific reports

OPEN

Unraveling HIV-1 transmission patterns through molecular surveillance in a hotspot in Yunnan Province, China

Hui Yang[1,5], Huichao Chen[2,5], Enlong Yao[3,5], Jie Dai[2], Yu Han[2], Xiaomei Jin[2], Min Yang[2], Zhijun Zeng[2], Pengyan Sun[2], Yuhua Shi[2], Yanling Ma[2✉], Manhong Jia[2✉] & Min Chen[4✉]

Honghe Prefecture, Yunnan Province, China, is a key area affected by HIV/AIDS. Given the high number of infected individuals, understanding HIV transmission in this region is crucial to controlling the epidemic in Yunnan Province. A molecular epidemiology study with 1004 newly diagnosed HIV-1 cases in 2022 was conducted to characterize transmission and guide targeted interventions. Of the 833 samples genotyped, nine HIV-1 genotypes were identified, with CRF08_BC (56.7%), URFs (17.8%), CRF07_BC (13.9%) and CRF01_AE (7.1%) being the most prevalent. These genotypes were widely distributed but concentrated in the center and east of the prefecture. Molecular network analysis showed varying clustering rates among counties, with Luxi County having the highest and Honghe County the lowest. Low education levels were associated with higher clustering. Cross-county transmission patterns were revealed, with strong links between specific counties. Of the three spatial subdivisions identified by the cohesive subgroup analysis, subgroup I included eight counties running from northeast to southwest. Persons aged 50–59 years and ethnic minorities were more likely to be identified in cross-county transmission. Notably, the overall prevalent level of HIV-1 pre-treatment drug resistance was 6.2% (50/809), with higher levels in Luxi (18.2%, 8/44), Hekou (14.3%, 3/21) and Yuanyang (10.9%, 13/119). This study investigated HIV-1 transmission patterns using molecular network analysis in a hotspot of HIV transmission, and informed the application of HIV molecular surveillance in practice.

Keywords HIV-1, Genotypes, Molecular network, Spatial analysis, Pre-treatment drug resistance

Epidemic caused by human immunodeficiency virus type 1 (HIV-1) remains a serious global public health threat[1]. According to the report of the joint United Nations Program on HIV/AIDS (UNAIDS), by the end of 2022, 39 million [33.1–45.7 million] people worldwide were living with HIV, with the estimated 1.3 million [1.0–1.7 million] new HIV infections in 2022[2]. In 2023, China reported 58,903 cases of Acquired Immune Deficiency Syndrome (AIDS) and 22,137 deaths related to the disease[3]. The persistence of HIV transmission in different countries and regions makes achieving the goal of ending the global AIDS epidemic by 2030 a major challenge[4].

Yunnan Province is located on the south-western border of China and shares borders with Vietnam, Laos and Myanmar. Historically, frequent drug trafficking led to the spread of HIV among injecting drug users (IDUs). Subsequently, the intersection of intravenous drug use and commercial sexual behavior had facilitated the introduction of HIV into the general population[5]. Yunnan Province became one of the provinces most affected by HIV/AIDS in China[6]. In recent years, Yunnan Province has made positive progress in HIV prevention and treatment. By 2020, Yunnan Province achieved the "90–90-90" target for AIDS prevention and treatment[7],

[1]Yunnan Provincial Key Laboratory of Public Health and Biosafety & School of Public Health, Kunming Medical University, Kunming, Yunnan Province, China. [2]Yunnan Provincial Key Laboratory of Public Health and Biosafety & Institute for AIDS/STD Control and Prevention, Yunnan Center for Disease Control and Prevention, No 1177, Xianghe Street, Chenggong District, Kunming 650500, Yunnan Province, China. [3]Division for AIDS/STD Control and Prevention, Honghe Center for Disease Control and Prevention, Honghe, Yunnan Province, China. [4]Yunnan Provincial Key Laboratory of Public Health and Biosafety & Health Laboratory Center, Yunnan Center for Disease Control and Prevention, No 1177, Xianghe Street, Chenggong District, Kunming 650500, Yunnan Province, China. [5]Hui Yang, Huichao Chen and Enlong Yao contributed equally to this work. ✉email: mayanling68@aliyun.com; jiamanhong@hotmail.com; chenminyx@hotmail.com

namely, 90% of all people living with HIV knew their HIV status, 90% of all people with diagnosed HIV infection received sustained antiretroviral therapy, and 90% of all people receiving antiretroviral therapy had viral suppression. However, Yunnan Province still leads the country in the number of people living with HIV/AIDS. In order to achieve the goal of ending the AIDS epidemic by 2030, with the advancement of treatment-centered responses, it is also necessary to further interrupt transmission networks by detecting and controlling transmission risks.

HIV molecular surveillance is important for accurate prevention and control of AIDS[8,9]. Based on the relevant influencing factors, effective prevention and control measures can be taken in a timely manner to improve the effectiveness of prevention and control[9]. In addition, by constructing molecular transmission networks among HIV-infected individuals, HIV transmission relationships and patterns can be revealed in molecular dimensions, which can be used to identify and track HIV transmission pathways, helping to identify potential transmission chains and infection risks[10,11]. Further, based on the molecular network, combined with information such as case surveillance and social behavior, extended surveys can identify high-risk populations associated with the molecular network, providing an idea for effective precision HIV prevention and control[10,12]. Currently, HIV molecular networks are used in different regions and populations. Rapid response to transmission clusters has been used as a primary strategy for ending the AIDS epidemic in the United States[13]. In China, HIV molecular networks have been actively explored to understand HIV-1 transmission trends, identify the population for priority interventions and evaluate the effectiveness of interventions[14].

Honghe Prefecture is located in the southern part of Yunnan Province, bordering Vietnam to the south, with three border counties, namely Hekou, Jinping and Luchun. Due to its large population and the long duration of the AIDS epidemic, Honghe Prefecture has become a key area for AIDS prevention and treatment in Yunnan Province. Based on historical surveillance data, Honghe Prefecture's people living HIV/AIDS approximately accounted for 20% of those in Yunnan Province, ranking first in the province. The risk of local HIV transmission remains high and the epidemiological situation is complex. To reveal the characteristics of the local HIV epidemic and provide a basis for the development of effective prevention and control strategies, a HIV molecular epidemiological study was carried out in Honghe Prefecture.

Results

Demographic characteristics of the participants

From 1,004 HIV-1 positive plasma samples, 627 *gag*, 809 *pol* and 723 *env* sequences were successfully amplified and sequenced. Based on the genotypes of at least two genetic regions per sample, 833 samples were assigned with final genotyping results. The demographic characteristics of the 833 participants showed no statistical differences with those of the 171 failing to obtain genotypes (Supplementary Table 1).

Of the 833 individuals, 59.7% (497/833) were reported from four out of thirteen counties (Mengzi, Yuanyang, Kaiyuan and Gejiu). And 88.1% (734/833) were Honghe natives, 5.4% (45/833), 5.8% (48/833) and 0.7% (6/833) were from other cities in Yunnan Province, other provinces and foreign countries, respectively. The male to female ratio was 2.18:1. Age ranged from 16 to 89 years, with a median age of 47 years. The dominant ethnic groups were Han (40.5%, 337/833), Hani (23.2%, 193/833) and Yi (20.6%, 172/833). Unmarried, married and divorced/widowed persons accounted for 24.1% (201/833), 43.7% (364/833) and 32.2% (268/833), respectively. Those with a primary school education or less accounted for 66.0% (550/833). The main occupation was farmer (80.9%, 674/833). Of the participants, 93.3% (777/833) were infected through heterosexual contact.

Prevalence characteristics of HIV-1 genotypes

Of the samples genotyped, the main HIV-1 genotypes were CRF08_BC (56.7%, 472/833), URFs (17.8%, 148/833), CRF07_BC (13.9%, 116/833) and CRF01_AE (7.1%, 59/833). The remaining HIV-1 genotypes included CRF85_BC (2.2%, 18/833), CRF101_01B (1.1%, 9/833), Subtype C (0.7%, 6/833), CRF55_01B (0.5%, 4/833) and CRF64_BC (0.1%, 1/833). Among URFs, the most common recombinant form was BC recombinants (72.3%, 107/833); others included CRF01_AE/C (7.4%, 11/833), CRF07_BC/CRF08_BC (6.8%, 10/833), CRF01_AE/CRF07_BC (6.1%, 9/833), CRF01_AE/CRF08_BC (3.4%, 5/833), CRF01_AE/BC (2.0%, 3/833), CRF01_AE/B (0.7%,1/833), CRF101_01B/C (0.7%, 1/833), and CRF55_01B/C (0.7%, 1/833).

Spatial distribution characteristics of HIV-1 genotypes

Overall, the four main HIV-1 genotypes had a wide distribution, but primarily in eastern and central regions (Fig. 1). Specifically, CRF08_BC was concentrated in Mengzi City, Gejiu City and Yuanyang County, CRF07_BC was concentrated in Mengzi City, Kaiyuan City and Gejiu City, and CRF01_AE was concentrated in Mengzi City, while URFs were mostly found in Mengzi City, Kaiyuan City, Gejiu City and Yuanyang County. The other genotypes had a more restricted spatial distribution, with CRF85_BC mainly in Kaiyuan County and CRF101_01B mainly in Yuanyang County.

Global and local spatial autocorrelation analyses showed that CRF07_BC formed a high-high cluster area in Mengzi City and Kaiyuan City, URFs formed a high-high cluster area in Kaiyuan City.

HIV-1 molecular network analysis

The HIV-1 molecular network was constructed with *pol* sequences using a combination of phylogenetic tree analysis and genetic distance. Of 809 *pol* sequences, 344 segregated into 124 clusters. The number of sequences per cluster ranged from 2 to 19.

The multivariable logistic analysis showed that reporting area and education level were associated with the clustering rate (Table 1). Compared with the clustering rate in Kaiyuan County (42.5%, 45/106), which was equivalent to the general clustering rate in the whole prefecture (42.5%, 344/809), the clustering rate in Honghe County (19.0%, 4/21) was significantly lower (OR (95% CI): 0.271 (0.084–0.877)), the clustering rate in Luxi
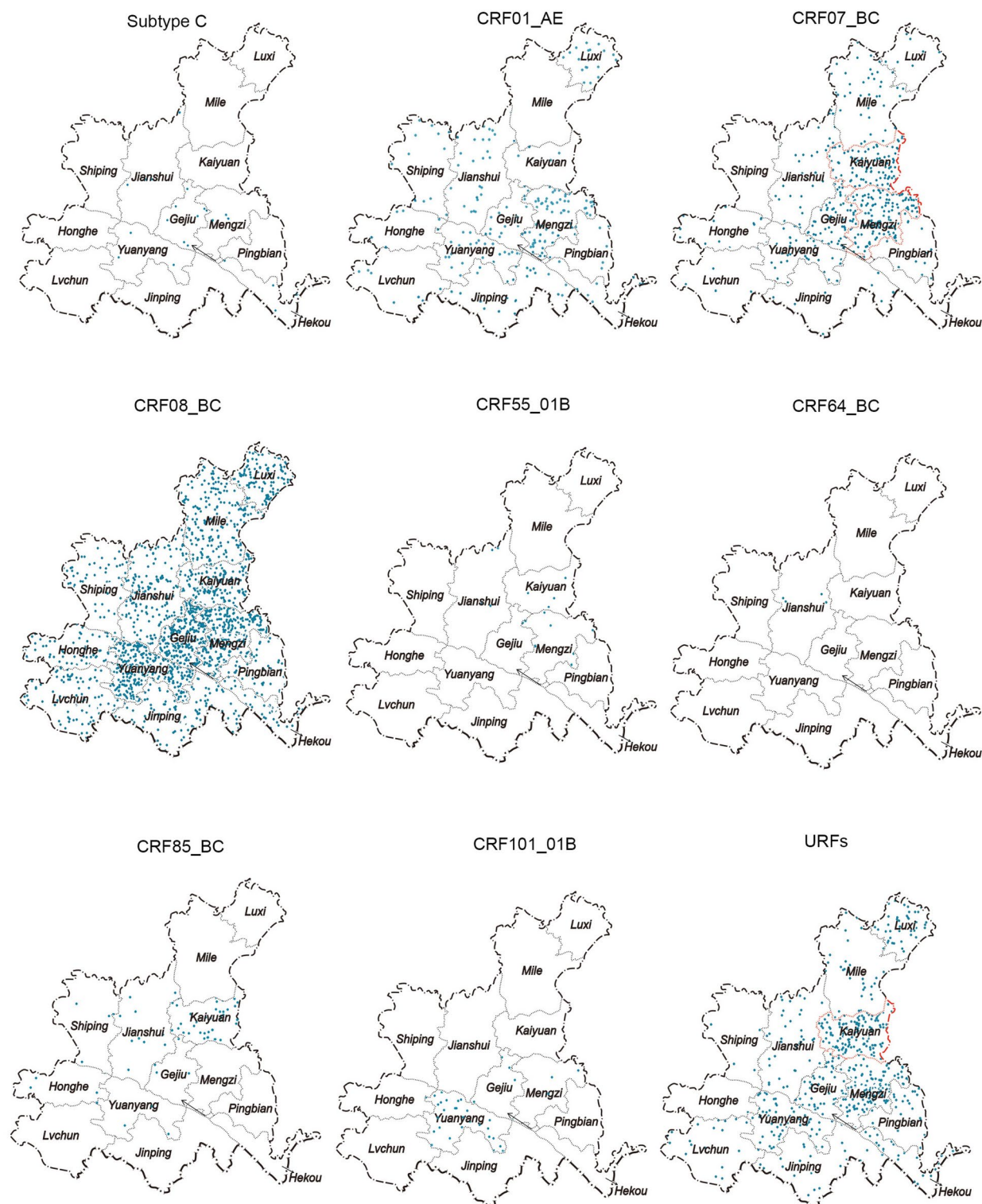
Fig. 1. Spatial distribution of HIV-1 genotypes in Honghe Prefecture. The dot density maps for Subtype C, CRF01_AE, CRF07_BC, CRF08_BC, CRF55_01B, CRF64_BC, CRF85_BC, CRF101_01B and URFs. One dot represented 0.03% of the total cases subtyped. The areas outlined in red were the high-high aggregation areas analyzed by the local spatial autocorrelation analysis. Administrative boundary data on were downloaded from the National Catalogue Service for Geographic Information (https://www.webmap.cn). The map content approval number: Yun-S-(2024)14.

Variables	Total	Cases in networks n (%)	Univariate analysis		Multivariate analysis	
			P	OR (95% CI)	P	OR (95% CI)
Report area			0.003		0.016	
Kaiyuan	106	45 (42.5)	–	1.000	–	1.000
Gejiu	101	33 (32.7)	0.148	0.658 (0.373–1.160)	0.148	0.652 (0.365–1.163)
Mengzi	159	66 (41.5)	0.879	0.962 (0.585–1.583)	0.928	0.977 (0.586–1.627)
Mile	64	30 (46.9)	0.574	1.196 (0.641–2.232)	0.561	1.211 (0.635–2.308)
Pingbian	21	7 (33.3)	0.439	0.678 (0.253–1.816)	0.388	0.642 (0.235–1.754)
Jianshui	65	26 (40.0)	0.752	0.904 (0.482–1.694)	0.850	0.939 (0.49–1.801)
Shiping	16	7 (43.8)	0.922	1.054 (0.365–3.044)	0.837	1.119 (0.383–3.268)
Luxi	44	29 (65.9)	0.010	2.621 (1.260–5.453)	0.032	2.279 (1.075–4.83)
Yuanyang	119	45 (37.8)	0.479	0.824 (0.483–1.407)	0.284	0.736 (0.420–1.29)
Honghe	21	4 (19.0)	0.053	0.319 (0.100–1.013)	0.029	0.271 (0.084–0.877)
Jinping	42	22 (52.4)	0.275	1.491 (0.728–3.056)	0.544	1.261 (0.596–2.666)
Lvchun	30	21 (70.0)	0.010	3.163 (1.324–7.555)	0.051	2.441 (0.998–5.972)
Hekou	21	9 (42.9)	0.973	1.017 (0.395–2.619)	0.908	0.945 (0.36–2.477)
Registered residence			0.141			
Other cities in Yunnan	45	13 (28.9)	–	1.000		
Honghe Prefecture	711	311 (43.7)	0.054	1.914 (0.988–3.708)		
Other provinces	46	16 (34.8)	0.547	1.313 (0.542–3.182)		
Foreinger	7	4 (57.1)	0.153	3.282 (0.643–16.749)		
Gender						
Male	557	231 (41.5)	–	1.000		
Female	252	113 (44.8)	0.370	1.147 (0.850–1.549)		
Age			0.566			
30–39	171	64 (37.4)	–	1.000		
< 30	76	34 (44.7)	0.279	1.353 (0.782–2.341)		
40–49	210	92 (43.8)	0.208	1.303 (0.863–1.969)		
50–59	232	105 (45.3)	0.116	1.382 (0.923–2.069)		
≥ 60	120	49 (40.8)	0.557	1.154 (0.715–1.861)		
Race/ethnicity			0.331			
Hani	193	77 (39.9)	–	1.000		
Han	322	133 (41.3)	0.753	1.060 (0.737–1.525)		
Yi	163	69 (42.3)	0.642	1.106 (0.724–1.689)		
Others	131	65 (49.6)	0.084	1.484 (0.948–2.321)		
Marital Status			0.007		0.050	
Unmarried	198	74 (37.4)	–	1.000	–	1.000
Married	355	173 (48.7)	0.010	1.593 (1.117–2.272)	0.127	1.337 (0.920–1.941)
Divorced/Widowed	256	97 (37.9)	0.910	1.022 (0.697–1.499)	0.554	0.886 (0.594–1.322)
Education			0.004		0.015	
Senior middle school and above	79	22 (27.8)	–	1.000	–	1.000
Junior middle school	194	75 (38.7)	0.092	1.633 (0.923–2.889)	0.185	1.505 (0.822–2.755)
Primary school	365	159 (43.6)	0.011	2.000 (1.173–3.410)	0.038	1.868 (1.034–3.375)
Illiteracy	171	88 (51.5)	0.001	2.747 (1.544–4.887)	0.003	2.660 (1.391–5.089)
Occupation						
Others	152	53 (34.9)	–	1.000	–	1.000
Farmers	657	291 (44.3)	0.035	1.485 (1.029–2.145)	0.863	1.040 (0.664–1.629)
Infection Routes			0.441			
Introvenous drug use	18	5 (27.8)	–	1.000		
Heterosexual contact	753	322 (42.8)	0.211	1.942 (0.686–5.504)		
Homosexual contact	38	17 (44.7)	0.229	2.105 (0.625–7.083)		

Table 1. Associated demographic factors of the individuals clustering in the HIV-1 molecular network.
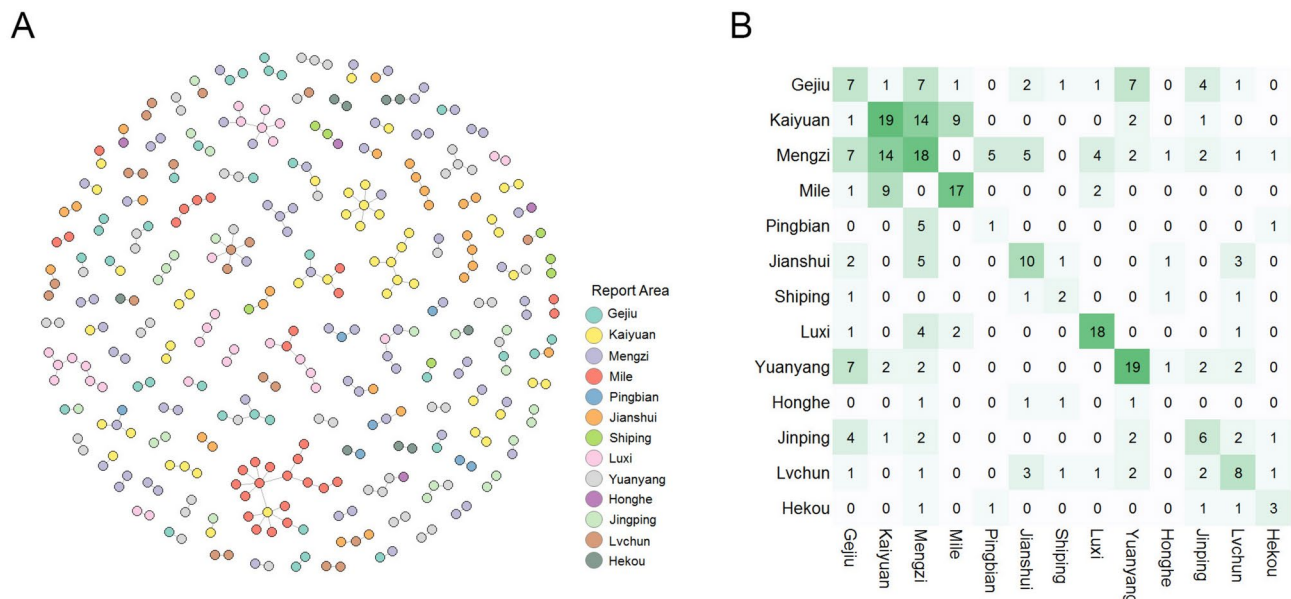
A

B

Fig. 2. Cross-country HIV-1 genetic transmission networks. (**A**) The HIV-1 molecular network coded by the reporting area of the individuals in the network. (**B**) The matrix of HIV-1 genetic links between counties. The cells on the diagonal represented the number of links in the same region. The other cells represented the number of links between different regions.
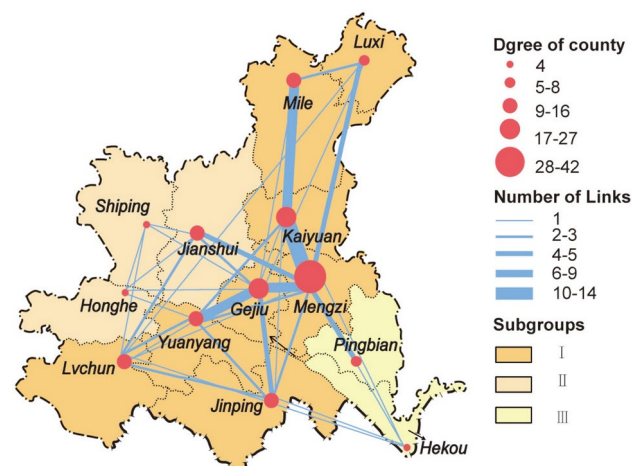

Fig. 3. The visualization of cross-county transmission on the map of Honghe prefecture. The size of the dot represented the degree of cross-county transmission for the corresponding county, indicating its centrality. The thickness of the line represented the number of links between counties and indicated the strength of the link. The different colored areas represented the cohesive subgroups obtained using the faction analysis. The map content approval number: Yun-S-(2024)14.

County was significantly higher (OR (95% CI): 2.279 (1.075–4.83)). And the clustering rates in the illiterate individuals (51.5%, 88/171) and those with primary school education (43.6%, 159/365) were significantly higher than that in those with a senior high school education or above (27.8%, 22/79).

Cross-regional HIV-1 transmission network

Among the 220 links found in the HIV-1 molecular network, 58.1% (128/220) of links connected nodes within the same county, while 41.8% (92/220) connected nodes from different counties (Fig. 2), which suggested the presence of cross-county transmission. With the number of links as the strength of connection between counties, Mengzi-Kaiyuan, Kaiyuan-Mile, Mengzi-Gejiu, Gejiu-Yuanyang formed strong connections between each other (Fig. 3).

To explore the internal associations and grouping of counties in the cross-regional transmission network, the cohesive subgroups were further analyzed. The result showed that there were three subgroups (Fig. 3). The

subgroups I included Mengzi, Kaiyuan, Mile, Luxi, Gejiu, Yuanyang, Jinping and Lvchun. The subgroup II included Jianshui, Honghe and Shiping. The subgroup III included Hekou and Pingbian.

Among the 344 individuals in the network, 154 were involved in cross-county transmission. Multivariate logistic regression analysis suggested that age and ethnicity were associated with cross-county transmission (Table 2). Those aged 50–59 years (OR (95% CI): 2.103 (1.015–4.358)) and of ethnic minorities (OR (95% CI): 1.904 (1.026–3.532)) were more likely associated with cross-county transmission.

Prevalence of pretreatment HIV-1 drug resistance

Among the individuals with *pol* sequences, the overall prevalence of HIV-1 pre-treatment drug resistance was 6.2% (50/809). Further, the prevalence of PDR to NNRTIs, NRTIs and PIs were 5.2% (42/809), 0.9% (7/809) and 0.4% (3/809), respectively. The overall prevalence of PDR among all counties showed statistical difference ($\chi^2 = 23.789$, $P = 0.022$), with those in Luxi (18.2%, 8/44), Hekou (14.3%, 3/21), and Yuanyang (10.9%, 13/119) ranking among the top three (Table 3). The prevalence of PDR to NNRTIs among all counties showed statistical difference ($\chi^2 = 22.920$, $P = 0.028$), with those in Luxi (15.9%, 7/44) and Hekou (14.3%, 3/21) exceeding 10% (Table 3).

Discussion
Statement of principal findings
In this study, a comprehensive molecular epidemiological study was conducted on newly diagnosed HIV-1 infected individuals in Honghe Prefecture in 2022. CRF08_BC, URFs, CRF07_BC, and CRF01_AE were the

Variables	Total	Individuals involved in cross-county transmission	Univariate analysis		Multivariate analysis	
			P	OR (95% CI)	P	OR (95% CI)
Registered residence			0.746			
Honghe Prefecture	311	140	–	1.000		
Other cities in Yunnan	13	4	0.318	0.543 (0.164–1.800)		
Other provinces	16	8	0.697	1.221 (0.447–3.337)		
Foreinger	4	2	0.842	1.221 (0.170–8.782)		
Gender						
Male	231	102	–	1.000		
Female	113	52	0.744	1.078 (0.686–1.694)		
Age			0.013		0.077	
≥60	49	15	–	1.000	–	1.000
50–59	105	52	0.029	2.224 (1.085–4.559)	0.045	2.103 (1.015–4.358)
40–49	92	33	0.531	1.268 (0.604–2.663)	0.859	1.072 (0.498–2.306)
30–39	64	33	0.027	2.413 (1.105–5.267)	0.135	1.872 (0.823–4.260)
< 30	34	21	0.006	3.662 (1.458–9.193)	0.138	2.212 (0.775–6.308)
Race/ethnicity			0.079		0.135	
Han	133	48	–	1.000	–	1.000
Hani	77	40	0.026	1.914 (1.082–3.386)	0.074	1.791 (0.944–3.396)
Yi	69	33	0.108	1.623 (0.900–2.929)	0.167	1.547 (0.833–2.872)
Others	65	33	0.050	1.826 (1.001–3.332)	0.041	1.904 (1.026–3.532)
Marital Status			0.267			
Married	173	70	–	1.000		
Unmarried	74	37	0.167	1.471 (0.851–2.544)		
Divorced/Widowed	97	47	0.204	1.383 (0.838–2.282)		
Education			0.455			
Illiteracy	88	37	–	1.000		
Primary school	159	68	0.912	1.030 (0.608–1.745)		
Junior middle school	75	36	0.446	1.272 (0.684–2.365)		
Senior middle school and above	22	13	0.155	1.991 (0.770–5.145)		
Occupation						
Farmers	291	129	–	1.000		
Others	53	25	0.702	1.121 (0.624–2.016)		
Infection Routes			0.084		0.265	
Heterosexual contact	322	139	–	1.000	–	1.000
Homosexual contact	17	12	0.034	3.160 (1.088–9.178)	0.115	2.627 (0.789–8.742)
Introvenous drug use	5	3	0.459	1.975 (0.326–11.980)	0.648	1.538 (0.242–9.754)

Table 2. Associated demographic factors of the individuals involved in cross-county transmission.

County	Total	Drug resistance n (%)	NNRTIs resistance n (%)	NRTIs resistance n (%)	PIs resistance n (%)
Jianshui	65	3 (4.6)	3 (4.6)	0	0
Kaiyuan	106	4 (3.8)	3 (2.8)	1 (0.9)	0
Shiping	16	0	0	0	0
Lvchun	30	0	0	0	0
Pingbian	21	2 (9.5)	2 (9.5)	0	0
Mengzi	159	8 (5.0)	6 (3.8)	2 (1.3)	2 (1.3)
Mile	64	3 (4.7)	2 (3.1)	1 (1.6)	0
Gejiu	101	4 (4.0)	3 (3.0)	1 (1.0)	0
Hekou	21	3 (14.3)	3 (14.3)	0	0
Yuanyang	119	13 (10.9)	11 (9.2)	2 (1.7)	0
Luxi	44	8 (18.2)	7 (15.9)	0	1 (2.3)
Honghe	21	1 (4.8)	1 (4.8)	0	0
Jinping	42	1 (2.4)	1 (2.4)	0	0
Total	809	50 (6.2)	42 (5.2)	7 (0.9)	3 (0.4)

Table 3. Pre-treatment drug resistance by county.

predominant HIV-1 genotypes, accounting for 95.4% of the total infections, of which CRF07_BC and URFs formed spatial clusters. A molecular network was constructed to investigate the local transmission patterns and identify transmission hotspots. A moderate level of HIV-1 pre-treatment drug resistance was identified, with higher levels in several counties. These findings provide a basis for further formulating prevention and control measures.

Comparison with other studies
As reported in the previous study, CRF08_BC, CRF07_BC, CRF01_AE, and URFs were identified as the most prevalent genotypes, although their ranking varied by region[15]. For example, in Zhaotong City located in northern Yunnan Province, these four genotypes were ordered as CRF08_BC, CRF01_AE, CRF07_BC and URFs[16]; in Baoshan City located in western Yunnan Province, they were ordered as URFs, CRF08_BC, CRF01_AE and CRF07_BC[17]. This study showed that the order of these four HIV-1 genotypes in Honghe Prefecture was CRF08_BC, URFs, CRF07_BC and CRF01_AE, which accounted for 95.4% of the total. These regional variations could be associated with distinct epidemiological contexts (e.g., timeline of epidemic establishment, population mobility patterns) as suggested by previous comparative studies[18].

The study uniquely identified specific spatial clustering patterns and transmission hotspots, which have not been as extensively explored in other regional studies. The proportion of CRF08_BC was more than half. Spatially, CRF08_BC showed a broad distribution, being detected in all 13 counties. CRF07_BC formed a high-high cluster area in Mengzi City and Kaiyuan City, which are the economic and cultural center of Honghe Prefecture. Both CRF08_BC and CRF07_BC originated from subtype B' and C recombination among intravenous drug users in the western Yunnan Province in the 1990s, with CRF08_BC spreading mainly east and CRF07_BC north[19–21], which could partially explain the observed predominance of CRF08_BC in eastern Yunnan, where Honghe Prefecture is geographically situated. CRF01_AE showed no obvious clustering but was centrally distributed in Mengzi City, the capital of Honghe Prefecture, and sporadically distributed elsewhere. These findings indicate spatial heterogeneity in HIV-1 transmission patterns across regions.

The increase in viral genetic variation and diversity, including the introduction of new subtypes/CRFs and a high proportion of URFs, was a common trend observed in other regions. CRF101_01B, subtype C, CRF55_01B and CRF64_BC were reported for the first time in Honghe Prefecture. CRF101_01B was first identified in Zhaotong City and Kunming City, Yunnan Province[22], CRF64_BC was in Dehong Prefecture, Yunnan Province[23,24]. CRF55_01B was first identified in MSM in eastern China and has spread nationwide[25,26]. All CRF55_01B cases in this study were detected among MSM, highlighting the potential for interregional spread of the virus. URFs, predominantly BC recombinants (79.1%), were the second most common genotype in Honghe Prefecture. Spatially, URFs formed a high-high cluster in Kaiyuan City, the epicenter of the HIV-1 epidemic in Honghe Prefecture.

Meaning of the study
The HIV-1 molecular network is commonly used to study transmission characteristics and the population at risk[14]. As found in this study, Luxi County, with the highest clustering rate in Honghe Prefecture, emerged as a potential hotspot of HIV transmission, requiring timely interventions, such as epidemiological investigation, targeted testing and prompt treatment services[27]. Conversely, Honghe County's lowest clustering rate highlighted the need for improve testing strategies and contact tracing to better detect infected individuals. Low literacy levels were found to be associated with higher clustering rates in the molecular network, highlighting the need for health education programs tailored to populations with limited access to HIV prevention information.

This study found that trans-county transmission was common, accounting for 41.8% of all links. Spatial visualization revealed higher intensity associations primarily between neighboring counties, forming a connected

spreading belt (Mile—Kaiyuan—Mengzi—Gejiu—Yuanyang). Mengzi, Kaiyuan, Gejiu, and Yuanyang were identified as probable hotspots. Thus, it is important for cross-regional HIV prevention and control.

This study identified three cohesive subgroups, with the largest coinciding with the aforementioned spreading belt, extending from the northeast to the southwest and including 8 of the 13 counties. The other two subgroups were smaller, located in the southeast (Pingbian and Hekou) and northwest (Jianshui, Shiping and Honghe). The structure of the transmission network can be influenced by geography, economy, transport and culture. The largest subgroup included Honghe Prefecture's political, economic, and tourist hubs, well-connected by highways and high-speed rail. Therefore, in addition to traditional strategies, interregional cooperation and information-sharing mechanisms can help control the risk of transmission.

At the individual level, people aged 50–59 and from ethnic minorities showed higher representation in interregional transmission. In Yunnan, the proportion of HIV-positive individuals aged ≥ 50 years increased from 8.9% in 2011 to 21.3% in 2017. Among newly reported HIV cases ≥ 50 years of age in China from 2015 to 2022, 46.8% involved commercial heterosexual sex, yet only 51.0% of the elderly knew the core AIDS prevention knowledge[28]. As this study focused on newly diagnosed cases, the possibility that they had been infected several years previously could not be excluded. However, older adults should be a key focus for HIV prevention, taking into account their specific needs and circumstances[29]. Ethnic minorities, with high mobility and low education, were also vulnerable. Effective interventions targeting these key populations are crucial for controlling transmission.

China's prevalence of transmitted drug resistance has risen to a moderate level (7.8%), with NNRTI resistance at 6.7%[30]. Pretreatment drug resistance is related to treatment effectiveness. According to WHO guidelines, public health measures are needed when NNRTI PDR prevalence exceeds 10%[31]. The overall prevalence of PDR in Honghe Prefecture was moderate (6.2%), but varied widely between counties. Luxi and Hekou had high levels of PDR, exceeding the 10% threshold for public health response, and Luxi had the highest rate of infected individuals entering the network, risking further spread of drug-resistant strains. Improved surveillance and epidemiological investigations are needed to assess the effectiveness of treatment, detect failures and prevent the spread of drug-resistance strains[10,27,32].

Unanswered questions and future research
Despite the comprehensive nature of the study, there are still unanswered questions. Future research should include longitudinal studies to track the evolution of HIV-1 genotypes and transmission patterns over time. Studies should be conducted to develop and validate risk assessment tools based on the results of this study. Further research is needed to understand the behavioral and social factors that contribute to high-risk behavior among older people and ethnic minorities. Further studies should focus on the effectiveness of interregional cooperation and information-sharing mechanisms in controlling the spread of HIV-1.

Strengths and weaknesses of the study
The study has several strengths. It used a comprehensive molecular epidemiological approach, providing detailed insights into the genetic diversity and transmission patterns of HIV-1. The use of molecular network analysis allowed the identification of transmission hotspots and cross-regional transmission links. The large sample size and detailed demographic and geographic data allowed for robust statistical analysis. However, the study has some weaknesses. It is limited to a single year and may not capture long-term trends in HIV-1 transmission. The results of the study are specific to one region and may not be generalizable to other regions.

Conclusion
In Honghe Prefecture, which currently has the largest number of reported infections in Yunnan Province, HIV genetics was even more diverse and complex. CRF08_BC, URFs, CRF07_BC and CRF01_AE were the predominant HIV-1 strains, and some strains not previously reported locally were also detected in this study. The four predominant strains were widely distributed, but mainly in the eastern and central regions, with URFs and CRF07_BC forming clustering areas. HIV-1 molecular network analyses showed that reporting region and literacy correlated with clustering rates, suggesting hotspot regions and populations for transmission. Cross-regional transmission showed different strengths of links between counties and formed specific spatial structural subdivisions. The age-specific group and ethnic minorities on the move were the main drivers of local inter-regional transmission. The overall level of pre-treatment drug resistance was moderate, but there were large regional differences that would challenge the effectiveness of control measures. These findings provided epidemiological evidence to optimize HIV control strategies in the local area.

Methods
Study participants and sample collection
A total of 1004 newly reported HIV/AIDS cases were recruited from medical and health institutions in Honghe Prefecture from January to December 2022. The inclusion criteria included: (1) age ≥ 15 years; (2) voluntary participation in this study and signing of informed consent. The adults' written informed consents were provided by themselves, the juveniles' were provided by their legal guardians. This study was in accordance with the Declaration of Helsinki, and was approved by the Ethical Review Committee of Yunnan Provincial Centre for Disease Control and Prevention (YNCDCER 2021-14).

Application and sequencing of HIV-1 gene
HIV-1 RNA was extracted from 200 μl of plasma using a fully automated nucleic acid extractor and accompanying kit (Tianlong, Xi'an, China). The partial *gag* (HXB2: 781–1861), *pol* (HXB2: 2147–3462) and *env* (HXB2: 7002–

7541) genes were amplified using nested polymerase chain reaction (PCR). The first round of PCR amplification was performed with 5 μl of viral RNA using the One Step RNA PCR Kit (Takara, Dalian, China), and the second round of amplification was performed with 5 μl of the production from the first round of PCR using 2× Taq PCR MasterMix (Tiangen, Beijing, China). The specific primers and amplification procedures are detailed in a previous study[17]. The amplified products were sent to Beijing Biomed Gene Technology Co. (Beijing, China) for Sanger sequencing using ABI 3730XL.

HIV-1 genotyping

Sequence assembly was performed using Sequencher 5.1 (Gene Codes, Ann Arbor, MI). Sequence alignment and cleaning were performed using BioEdit 7.0 and AliView. HIV-1 genotyping was performed jointly using HIV sequence typing tools on the HIV Gene Sequences Database of China (https://nmdc.cn/hiv/tool/sequence)[33], COntext-based Modeling for Expeditious Typing (COMET HIV-1, http://comet.retrovirology.lu)[34] and REGA HIV-1 Subtyping Tool-Version 3.0 (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/). The final genotype of each sample was determined jointly by at least two of the *gag*, *pol* and *env* regions.

HIV-1 molecular network analysis

HIV-1 molecular clusters were constructed using HIV transmission network tool on the HIV Gene Sequences Database of China (https://nmdc.cn/hiv/molecularnetwork), using the method of evolutionary analysis combined with gene distance. The web-based HIV transmission network tool uses FastTree software to construct maximum likelihood (ML) phylogenetic trees[35], and Cluster Picker software to extract molecular clusters based on selected genetic distance and bootstrap support[36]. These two underlying analysis software programs are also recommended by China's technical guidelines[27]. When using FastTree to construct ML phylogenetic trees, the generalized time-reversible (GTR) model of nucleotide evolution and the Shimodaira-Hasegawa test for local support values are used. According to the previous studies[37], to select the optimum genetic distance threshold, the effect of different genetic distances on cluster identification were evaluated. The genetic distance was varied from 0.5% to 4.5% with 90% node support value. When the genetic distance was greater than 3.5%, the growth of the number of clusters no decreased significantly (Supplementary Fig. 1). Thus, a genetic distance of 3.5% was chosen as the optimum genetic distance to construct the network.

Spatial analysis of HIV-1 strains and cross-county transmission

The distribution of HIV-1 genotypes was visualized using dot density maps, where one dot represents 0.03% of the total number of cases. The clustering characteristics of specific genotypes among different counties were assessed based on global and local spatial autocorrelation Moran's I values, with 5% as the significance test level. A high-high cluster represents a high distribution of a given genotype in both the corresponding and neighboring counties. The data on administrative boundaries were obtained from the National Catalogue Service for Geographic Information (https://www.webmap.cn, accessed on 11 November 2022). The spatial analysis was performed with GeoDa 1.20 and QGIS 3.28.

The cross-county transmission in the molecular network was shown by a 13×13 matrix plot, where the corresponding number indicated the number of links between two counties. The cross-county transmission was further visualized by connecting the centers of the corresponding counties on the map. The thickness of the line represented the number of links, which also corresponded to the strength of the connection.

Cohesive subgroups refer to clusters of actors in a social network who are more connected to each other than to other actors outside the subgroup. Faction analysis is a method of cohesive subgroup analysis based on group reciprocity relationships[38]. In this study, the method of faction analysis was used to analyze the spatial partitioning in the overall network formed by cross-county transmission. The cohesive subgroup analysis was carried out using UCINET software.

Pretreatment HIV-1 drug resistance analysis

Pretreatment drug resistance (PDR) was analyzed using the Stanford HIVdb programme (https://hivdb.stanford.edu/hivdb/by-sequences/). In accordance with the WHO pre-treatment HIV drug resistance guidelines, sequences were considered drug resistant if they were low, intermediate or high resistance to the 12 monitoring drugs, including two NNRTIs (nevirapine (NVP) and efavirenz (EFV)), seven nucleotide reverse transcriptase inhibitors (NRTIs, abacavir (ABC), zidovudine (AZT), stavudine (D4T), didanosine (DDI), emtricitabine (FTC), lamivudine (3TC) and tenofovir (TDF)) and three proteinase inhibitors (PIs, atazanavir/r (ATV/r), darunavir/r (DRV/r) and lopinavir/r (LPV/r))[31].

Statistical analysis

Statistical analysis was performed using SPSS 19.0 software (SPSS Inc. Chicago, IL). Categorical variables were described by numbers and percentages, which were compared using Chi-square test. Associated demographic factors of individuals clustered in the HIV-1 molecular network and those of the individuals involved in cross-county transmission were analyzed by logistic regression. The logistic regression analysis was conducted in two stages. Initially, univariate logistic regression analysis was performed to identify potential factors associated with the outcome of interest. Variables with p-values less than 10% were considered for further analysis. Subsequently, multivariate logistic regression analysis was conducted. This analysis included all variables that showed a significant association ($p < 0.10$) in the univariate analysis. All tests were two-tailed and a p-value < 0.05 was considered statistically significant.

Data availability

The sequences obtained in this study were submitted to NCBI GenBank under accession numbers PP983259-PP985426. The datasets are available from the corresponding author on reasonable request.

References

1. Bekker, L. G. The HIV epidemic 40 years on. *Nat. Rev. Microbiol.* **21**, 767–768. https://doi.org/10.1038/s41579-023-00979-y (2023).
2. UNAIDS. 2023 UNAIDS global AIDS update. Available online: https://thepath.unaids.org/. Accessed on 12 Jan 2025.
3. National Disease Control and Prevention Administration. Overview of the National Notifiable Infectious Disease Epidemic in 2023. Available online: https://www.ndcpa.gov.cn/jbkzzx/c100016/common/content/content_1836299733133275136.html. Accessed on January 12, 2025.
4. De Lay, P. R. et al. Ending AIDS as a public health threat by 2030: Time to reset targets for 2025. *PLoS Med.* **18**, e1003649. https://doi.org/10.1371/journal.pmed.1003649 (2021).
5. Lu, L. et al. The changing face of HIV in China. *Nature* **455**, 609–611. https://doi.org/10.1038/455609a (2008).
6. Jia, M. et al. The HIV epidemic in Yunnan Province, China, 1989–2007. *J. Acquir. Immune Defic. Syndr.* **53**(Suppl 1), S34-40. https://doi.org/10.1097/QAI.0b013e3181c7d6ff (2010).
7. Yanling Ma, Jincheng Lou, Yunfei Lao & Jia, M. Strategies and effectiveness in achieving the 90–90–90 target for AIDS prevention and treatment in Yunnan Province. *Chin J AIDS STD* **28** (2022).
8. Han, X., Zhao, B., An, M., Zhong, P. & Shang, H. Molecular network-based intervention brings us closer to ending the HIV pandemic. *Front. Med.* https://doi.org/10.1007/s11684-020-0756-y (2020).
9. Oster, A. M., France, A. M. & Mermin, J. Molecular epidemiology and the transformation of HIV prevention. *JAMA* **319**, 1657–1658. https://doi.org/10.1001/jama.2018.1513 (2018).
10. National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Division of HIV/AIDS. Prevention Detecting and Responding to HIV Transmission Clusters: A Guide for Health Department. Available from: 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 September 13, 2021.
11. Hassan, A. S., Pybus, O. G., Sanders, E. J., Albert, J. & Esbjornsson, J. Defining HIV-1 transmission clusters based on sequence data. *AIDS* **31**, 1211–1222. https://doi.org/10.1097/QAD.0000000000001470 (2017).
12. Arantes, I., Bello, G., Darcissac, E., Lacoste, V. & Nacher, M. Using phylogenetic surveillance and epidemiological data to understand the HIV-1 transmission dynamics in French Guiana. *AIDS* https://doi.org/10.1097/QAD.0000000000002817 (2021).
13. Fauci, A. S., Redfield, R. R., Sigounas, G., Weahkee, M. D. & Giroir, B. P. Ending the HIV epidemic: A plan for the United States. *JAMA* **321**, 844–845. https://doi.org/10.1001/jama.2019.1343 (2019).
14. Qiu, Y., Zhao, B., Han, X. & Zhong, P. HIV-1 molecular network-based studies in China. *Infect. Microbes Dis.* **4**, 20–25. https://doi.org/10.1097/im9.0000000000000086 (2022).
15. Chen, M. et al. Spatial clusters of HIV-1 genotypes in a recently infected population in Yunnan, China. *BMC Infect. Dis.* **19**, 669. https://doi.org/10.1186/s12879-019-4276-9 (2019).
16. Cao, R. et al. Using molecular network analysis to understand current HIV-1 transmission characteristics in an inland area of Yunnan, China. *Epidemiol. Infect.* **151**, e124. https://doi.org/10.1017/S0950268823001140 (2023).
17. Zhang, Y. et al. Using molecular network analysis to explore the characteristics of HIV-1 transmission in a China-Myanmar border area. *PLoS One* **17**, e0268143. https://doi.org/10.1371/journal.pone.0268143 (2022).
18. Paraskevis, D., Nikolopoulos, G. K., Magiorkinis, G., Hodges-Mameletzis, I. & Hatzakis, A. The application of HIV molecular epidemiology to public health. *Infect. Genet. Evol.* **46**, 159–168. https://doi.org/10.1016/j.meegid.2016.06.021 (2016).
19. Tee, K. K. et al. Temporal and spatial dynamics of human immunodeficiency virus type 1 circulating recombinant forms 08_BC and 07_BC in Asia. *J. Virol.* **82**, 9206–9215. https://doi.org/10.1128/JVI.00399-08 (2008).
20. Yang, R. et al. Identification and characterization of a new class of human immunodeficiency virus type 1 recombinants comprised of two circulating recombinant forms, CRF07_BC and CRF08_BC, China. *J. Virol.* **77**, 685–695 (2003).
21. Beyrer, C. et al. Overland heroin trafficking routes and HIV-1 spread in south and south-east Asia. *AIDS* **14**, 75–83 (2000).
22. Zhang, C. et al. Genetic characterization and recombinant history of a novel HIV-1 circulating recombinant form (CRF101_01B) identified in Yunnan, China. *Infect. Genet. Evol.* **73**, 109–112. https://doi.org/10.1016/j.meegid.2019.04.024 (2019).
23. Feng, Y. et al. Geographic origin and evolutionary history of China's two predominant HIV-1 circulating recombinant forms, CRF07_BC and CRF08_BC. *Sci. Rep.* **6**, 19279. https://doi.org/10.1038/srep19279 (2016).
24. Hsi, J. H. et al. Genome sequence of a novel HIV-1 circulating recombinant form (CRF64_BC) identified from Yunnan, China. *AIDS Res. Hum. Retrovir.* https://doi.org/10.1089/AID.2013.0234 (2013).
25. Han, X. et al. Genome Sequences of a Novel HIV-1 Circulating Recombinant Form CRF55_01B, Identified in China. *Genome Announc.*. https://doi.org/10.1128/genomeA.00050-12 (2013).
26. Gan, M. et al. The prevalence of CRF55_01B among HIV-1 strain and its connection with traffic development in China. *Emerg. Microbes Infect.* **10**, 256–265. https://doi.org/10.1080/22221751.2021.1884004 (2021).
27. Chian Center for Disease Control and Prevention. Technical Guide for HIV Transmission Networks Monitoring and Intervention. Accessed 2 Dec 2021 (2021).
28. Tang, H., Jin, Y. & Lyu, F. HIV/AIDS epidemic in the elderly and prevention and control challenges in China. *Chin. J. Epidemiol.* **44**, 1669–1672 (2023).
29. Tan, T. et al. Survey of AIDS knowledge awareness and prevalence of extramarital sex behavior in HIV-infected men aged 50 years and older in Chongqing. *Chin. J. Epidemiol.* **44**, 1692–1696 (2023).
30. Liu, X. et al. Changes in HIV-1 subtypes/sub-subtypes, and transmitted drug resistance among ART-naive HIV-infected individuals—China, 2004–2022. *China CDC Wkly* **5**, 664–671. https://doi.org/10.46234/ccdcw2023.129 (2023).
31. World Health Organization. Surveillance of HIV drug resistance in adults initiating antiretroviral therapy (pre-treatment HIV drug resistance) (http://apps.who.int/iris/bitstream/handle/10665/112802/9789241507196_eng.pdf?sequence=1). Accessed July 2014.
32. World Health Organization. Guidelines on the public health response to pretreatment HIV drug resistance (http://www.who.int/hiv/topics/drugresistance/en). Accessed 7 July 2017.
33. Feng, Y. et al. Research on HIV gene sequence data management and analysis system in China. *Chin. J. AIDS STD* **29**, 848–853 (2023).
34. Struck, D., Lawyer, G., Ternes, A. M., Schmit, J. C. & Bercoff, D. P. COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res* **42**, e144. https://doi.org/10.1093/nar/gku739 (2014).
35. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490. https://doi.org/10.1371/journal.pone.0009490 (2010).
36. Ragonnet-Cronin, M. et al. Automated analysis of phylogenetic clusters. *BMC Bioinform.* **14**, 317. https://doi.org/10.1186/1471-2105-14-317 (2013).

37. Junqueira, D. M., de Medeiros, R. M., Graf, T. & Almeida, S. E. Short-term dynamic and local epidemiological trends in the south American HIV-1B epidemic. *PLoS One* **11**, e0156712. https://doi.org/10.1371/journal.pone.0156712 (2016).
38. de Amorim, S. G., Barthélemy, J. P. & Ribeiro, C. C. Clustering and clique partitioning: Simulated annealing and tabu search approaches. *J. Classif.* **9**, 17–41 (1992).

Acknowledgements

Author contributions

M.C., M.J. and Y.M. conceived and designed the project. H.Y., H.C. and E.Y. collected the samples. H.Y., H.C., J.D., X.J., M.Y., Z.Z. and P.S. performed the experiments. E.Y., Y.H. and Y.S. collected the epidemiological data. H.Y., M.C. and Y.M. analyzed the data. H.Y., M.C. and M.J. prepared the manuscript. All authors reviewed the manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-94290-w.

Correspondence and requests for materials should be addressed to Y.M., M.J. or M.C.

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.