Letter to the Editor

Molecular Characterization of HIV-1-circulating Strains among Pre-Treatment Patients in Tibet*

Dorjiwangmo[1,&], GAN Meng Ze[2,&], DONG Ao Bo[2,&], Baimayangjin[1], Tashibazong[1], LIU Lei[2], LIAO Ling Jie[2], RUAN Yu Hua[2], SHAO Yi Ming[2], FENG Yi[2,#], and XING Hui[2,#]

The Tibet Autonomous Region (referred to as 'Tibet') is located in southwestern China, bordering the Yunnan Province, Sichuan Province, Xinjiang Uygur Autonomous Region, and other provinces with high HIV/AIDS prevalence, as well as India, Nepal, and other high-risk border countries. In 1994, the first patient with HIV-1 infection was detected at Zhangmu port, Namu County, Shigatse, Tibet. In 2005, the first AIDS cases were reported among local Tibetans. A total of 253 cases of HIV/AIDS were reported between 2004 and 2013. The number of HIV/AIDS cases in Tibet is increasing each year, and this epidemic encompasses seven cities throughout the region. Even though the overall epidemic has a low countrywide prevalence, the prevention and control of HIV-1 in Tibet cannot be ignored. Previous studies on patients with HIV-1 in Tibet have mainly focused on epidemiological studies, and to the best of our knowledge, no relevant study has reported the genotype of HIV-1. Therefore, this study aimed to analyze the genotype and source of HIV-1 infection strains in Tibet.

We considered the treatment sites in the city or county (district) as the basic sampling unit (including hospitals and Centers for Disease Control and Prevention, CDC), and the time of receiving antiviral treatment as the order, and patients with HIV infection were interviewed and blood samples collected upon providing their consent. The interviews and blood sample collections were conducted simultaneously. The blood samples of the subjects were collected from treatment sites in Lhasa, Changdu, Linzhi, Naqu, Shigatse, and Shannan. The inclusion criteria of the study subjects were as follows: (1) age ≥ 18 years; (2) patients with HIV infection who had not received any antiviral treatment from January 2018 to June 2019; (3) newly reported HIV-1 infections, and patients who restarted antiviral treatment after stopping the drug; and (4) enrollment depended on the time of receiving antiviral treatment at each treatment site. After we analyzed the blood samples, we conducted epidemiological surveys (including questionnaires and face-to-face interviews) with the subjects to supplement the analysis results. The investigation included a question regarding whether there were high-risk behaviors with foreign nationals. If yes, details such as country and year were recorded.

This study was approved by an ethics committee (Ethics Committee of National Center for AIDS/STD Control and Prevention, Chinese CDC, X140617334), and the permission of the health department was obtained to collect blood samples. All subjects in this study provided informed consent (including written and verbal). When a patient is diagnosed with HIV infection, their plasma samples can be used in research to control and prevent the associated disease.

The venous blood of the study subjects was collected using EDTA anticoagulation tubes. Plasma was separated according to standard procedures, frozen at −80 °C, and then transported to the National Center for AIDS/STD Control and Prevention, China CDC for testing. Viral RNA was extracted. Nested PCR was used to amplify the sequence of the HIV-1 *pol* region with reference to the gene region at position 2253–3553 of the international standard strain HXB2[1]. The gene sequences were spliced with sequencer 4.10.1 to obtain the complete sequences.

doi: 10.3967/bes2020.132

*This study was supported by the National Science and Technology Major Project 'Prevention and Control of Major Infectious Diseases such as AIDS and Viral Hepatitis' [2017ZX10201101], special funding for AIDS prevention.

1. Tibet Autonomous Region Center for Disease Control and Prevention, Lhasa 850000, Tibet, China; 2. State Key Laboratory of Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Beijing 102206, China

In the HIV databases (https://www.hiv.lanl.gov/content/sequence/VIRALIGN/viralign.html), we aligned the Tibet sequences with the reference sequences (the reference sequences included the major international epidemic strains A-D, F-H, and J-K, and the major epidemic recombinant strains in China) and used BioEdit 7.0.5.3 software to obtain clean sequence reads. Maximum-likelihood trees were generated using IQ-TREE version 2.0.6. Bootstrap trees were produced using the rapid bootstrapping algorithm with 1,000 bootstrap replicates. Clusters with a bootstrap value higher than 0.80 (80%) were defined as the same subtype. The maximum-likelihood trees were imported to FigTree v 1.4.3.

The major epidemic strains in Tibet were subjected to Bayesian analysis to investigate the region and time of their introduction to Tibet (more than 10% of HIV strains were identified by Bayesian analysis and the rest were identified by BLAST)[2]. We performed the Tibetan sequences using BLAST in HIV Databases (https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html). The Tibetan sequences and those obtained from BLAST were sorted by removing duplicates and sequences without years and regions, and the *pol* region fragments were intercepted for Bayesian analysis. We organized a total of 135 sequences of CRF07_BC, 143 sequences for CRF01_AE, 85 sequences for subtype C (to more accurately describe the spread of subtype C in Tibet, we screened 30 sequences of subtype C in the HIV Databases by region and year for Bayesian analysis), and 47 sequences of CRF08_BC for Bayesian analysis using BLAST.

To estimate the source of the strains, the Bayesian Markov Chain Monte Carlo (MCMC) model was used to perform the Bayesian discrete phylogeographic approach in BEAST v1.8.2 under a Bayesian skygrid demographic model[3]. Bayesian MCMC results were analyzed using Tracer v1.6. The generated MCC trees were visualized in FigTree v1.4.3 and were adjusted using Adobe Illustrator CS6.

From January 2018 to June 2019, 56 samples were collected from Lhasa, 15 from Changdu, 8 from Linzhi, 8 from Naqu, 3 from Shigatse, and 3 from Shannan. A total of 99 samples were collected, accounting for 38.8% (99/255) of the newly reported HIV-1 infection cases during the sampling period. Overall, 93 *pol* sequences were successfully amplified (the amplification rate was 93.9%). The samples mainly comprised patients who were 26–49 years old (72.0%; 67/93), male (63.4%; 59/93), Tibetan (75.3%; 70/93), married (52.7%; 49/93), heterosexual (61.3%; 57/93), and from Lhasa (60.2%; 56/93).

Phylogenetic analysis revealed that among 93 HIV-1 *pol* sequences in Tibet, CRF07_BC accounted for 38.7% (36/93), CRF01_AE for 25.8% (24/93), subtype C for 16.1% (15/93), CRF08_BC for 12.9% (12/93), CRF55_01B for 4.3% (4/93), and unique recombinant form (URF) for 2.2% (2/93) (Figure 1 and Table 1). The distribution of subtypes and sampling data in different regions are shown in Supplementary Figure S1, available in www.besjournal.com.

Bayesian analysis was performed on the major epidemic strains CRF07_BC, CRF01_AE, subtype C, and CRF08_BC in Tibet to investigate the region and time of their introduction into Tibet. We found that 62.9% of CRF07_BC strains in Tibet came from Sichuan province, 20.0% from Zhejiang province, 11.4% from Chongqing city, and 5.7% from Shenzhen city. Notably, 75.0% of CRF01_AE in Tibet came from Shenzhen city, 12.5% from Shanghai city, 8.3% from Thailand, and 4.2% from Sichuan province. We found that subtype C in Tibet was clustered with subtype C in India. All subtype C strains in Tibet came from India, and the time of most recent common ancestor (tMRCA) may be 2007 (95% highest posterior density HPD, 1998–2011) and 2011 (95% highest posterior density HPD, 2001–2013). All CRF08_BC strains in Tibet came from Yunnan province (Figure 2). The results showed that CRF07_BC strains came from other provinces in China, CRF01_AE strains came from other provinces in China and Thailand, subtype C strains came from India, and CRF08_BC strains came from Yunnan province.

This is the first study to determine the distribution of the genotypes of patients with HIV-1 infection in Tibet as well as the sources of the major subtypes by investigating the patients who had not yet received antiviral treatment in Tibet from 2018 to 2019. The subtypes of patients with HIV-1 infection in Tibet are mainly CRF07_BC, CRF01_AE, subtype C, and CRF08_BC. The distribution of HIV strains in Tibet is similar to that of inland China, but there is a high prevalence of subtype C[4].

The HIV-1 epidemic strains in Tibet have been affected by inland areas of China and the neighboring countries of India and Thailand (Sichuan, Zhejiang, Chongqing, Shenzhen, Shanghai, Yunnan, India, and Thailand). Combined with the results of Bayes analysis and sociodemographic information, the heterosexual transmission of CRF07_BC strains mainly originated in Sichuan and Chongqing, while the transmission *via* MSM mainly originated in Zhejiang and Shenzhen. There were also patients from other provinces,

namely, Sichuan and Chongqing, in Tibet. CRF07_BC subtype mainly originated in Liangshan and Sichuan, accounting for 98.3% of the strains. The strains identified in Liangshan, where called the Tibetan area belonging to Sichuan were mainly derived from heterosexual transmission. Tibet is bordered by Sichuan, the province with a high incidence of AIDS. CRF01_AE subtype originated in Shenzhen, Shanghai, Sichuan, and Thailand. As a pilgrimage site in China, Tibet is visited by many people annually. A considerable proportion of the population in both Thailand and Tibet follow Buddhism, and China's CRF01_AE strain was introduced by Thailand[5]. Coupled with the development of tourism, this may have caused the CRF01_AE strain in Thailand to spread to Tibet. Subtype C in Tibet originated in India. In the context of the 'Belt and Road' initiative, Tibet, as an important link in its strategy, has increased border trade with border countries, such as India and Nepal, which may be one of the factors driving the cross-border transmission of HIV-1. CRF08_BC strains were mainly found in Yunnan Province, where there are Tibetan Autonomous Prefectures. Therefore, it is possible that CRF08_BC in Tibet originated in Yunnan Province. At the same time, with the rapid development of China's economy and the

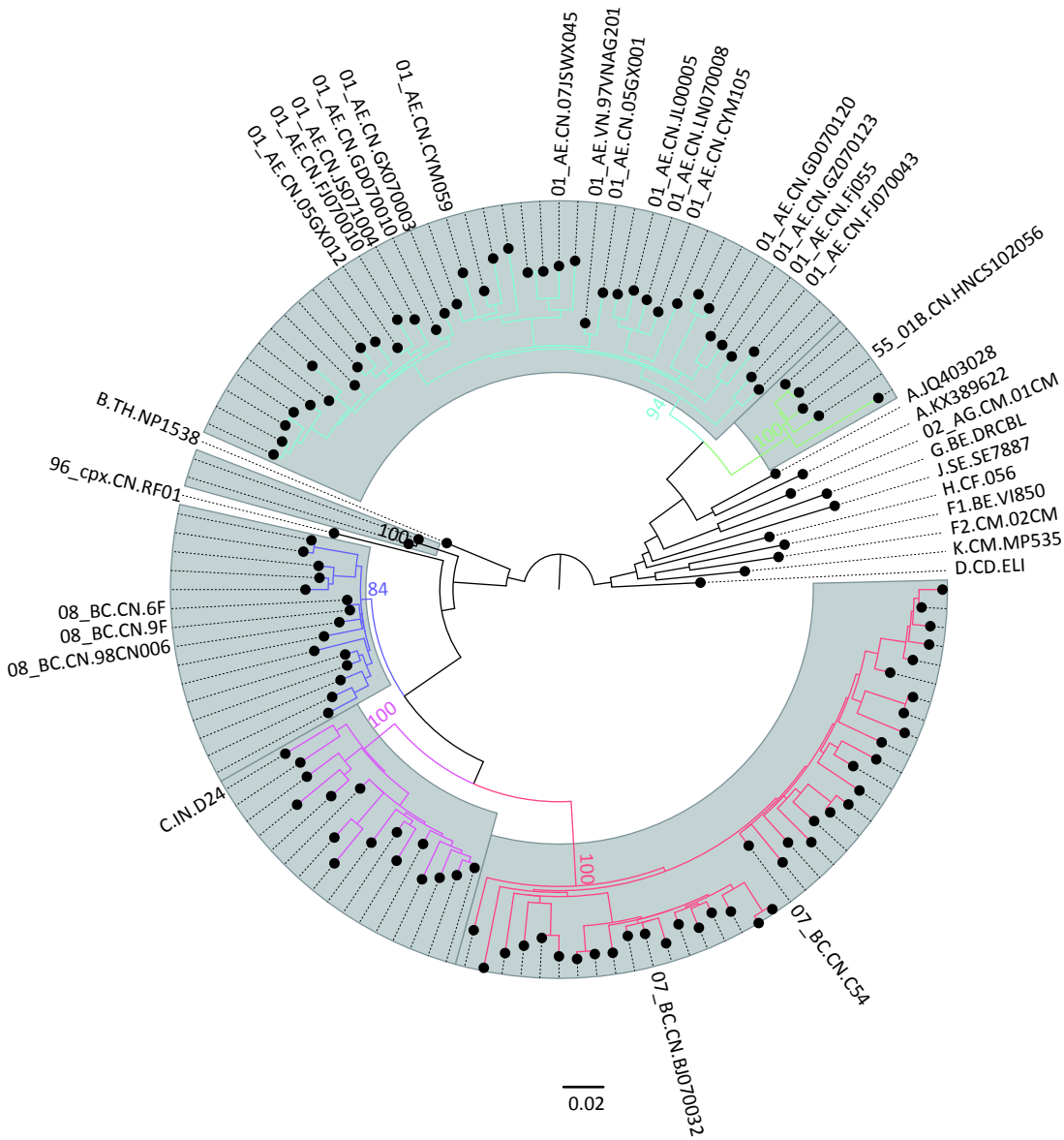


Figure 1. Maximum-likelihood phylogenetic trees of HIV-1 *pol* sequences from 93 HIV-infected people who had not received any antiviral treatment in Tibet from 2018 to 2019.

development of Tibet's tourism industry, epidemic strains in other provinces have spread to Tibet. This study identified another cross-border HIV transmission route in China besides Yunnan and Guangxi.

There are many HIV-1 subtypes, including single subtypes A-D, F-H, and J-K, and 101 circulating recombinant forms (CRFs), shown in the HIV databases. Among them, nearly half of all patients with HIV infection worldwide are subtype C, the most common subtype. Subtype C is mainly distributed in southern Africa, East Africa, Ethiopia, and India, showing a regional trend of transmission[6]. In the past 20 years, new recombinant strains have constantly emerged and spread throughout the population[7]. The 2015 national molecular epidemiological survey showed that China is dominated by two subtypes, CRF01_AE and CRF07_BC. Non-recombinant strain subtypes, such as subtype B, has low distribution in China and is gradually decreasing. There is almost no subtype C of non-recombinant strain. Since 2007, subtype C

Table 1. Subtypes and sociodemographic characteristics of HIV-infected people who had not received any antiviral treatment in Tibet from 2018 to 2019

Items	CRF07_BC, *N* (%)	CRF01_AE, *N* (%)	C, *N* (%)	CRF08_BC, *N* (%)	Others[a], *N* (%)
Total	36 (38.7)	24 (25.8)	15 (16.1)	12 (12.9)	6 (6.5)
Age (years)					
18–25	8 (22.2)	5 (20.8)	2 (13.3)	4 (33.3)	1 (16.7)
26–49	27 (75.0)	17 (70.9)	11 (73.4)	7 (58.4)	5 (83.3)
≥ 50	1 (2.8)	2 (8.3)	2 (13.3)	1 (8.3)	0 (0.0)
Sex					
Male	21 (58.3)	17 (70.8)	10 (66.7)	6 (50.0)	5 (83.3)
Female	15 (41.7)	7 (29.2)	5 (33.3)	6 (50.0)	1 (16.7)
Ethnicity					
Zang	22 (61.1)	18 (75.0)	14 (93.3)	12 (100.0)	4 (66.7)
Han	14 (38.9)	4 (16.7)	1 (6.7)	0 (0.0)	2 (33.3)
Hui	0 (0.0)	2 (8.3)	0 (0.0)	0 (0.0)	0 (0.0)
Marital status					
Married	18 (50.0)	12 (50.0)	9 (60.0)	6 (50.0)	4 (66.6)
Single	14 (38.9)	11 (45.8)	2 (13.3)	5 (41.7)	1 (16.7)
Divorced/widowed	4 (11.1)	1 (4.2)	3 (20.0)	1 (8.3)	1 (16.7)
Unknown	0 (0.0)	0 (0.0)	1 (6.7)	0 (0.0)	0 (0.0)
Risk group[b]					
Heterosexual	20 (55.6)	12 (50.0)	9 (60.0)	10 (83.3)	6 (100.0)
MSM	2 (5.5)	3 (12.5)	0 (0.0)	0 (0.0)	0 (0.0)
Unknown	14 (38.9)	9 (37.5)	6 (40.0)	2 (16.7)	0 (0.0)
City[c]					
Lhasa	21 (58.3)	13 (54.2)	11 (73.3)	7 (58.4)	4 (66.7)
Changdu	7 (19.4)	7 (29.2)	1 (6.7)	0 (0.0)	0 (0.0)
Linzhi	1 (2.8)	1 (4.2)	0 (0.0)	4 (33.3)	2 (33.3)
Naqu	5 (13.9)	2 (8.3)	1 (6.7)	0 (0.0)	0 (0.0)
Shigatse	0 (0.0)	1 (4.1)	2 (13.3)	0 (0.0)	0 (0.0)
Shannan	2 (5.6)	0 (0.0)	0 (0.0)	1 (8.3)	0 (0.0)

Note. [a]Others include CRF55_01B and URF; [b]MSM men who have sex with men; [c]Ali did not collect samples.

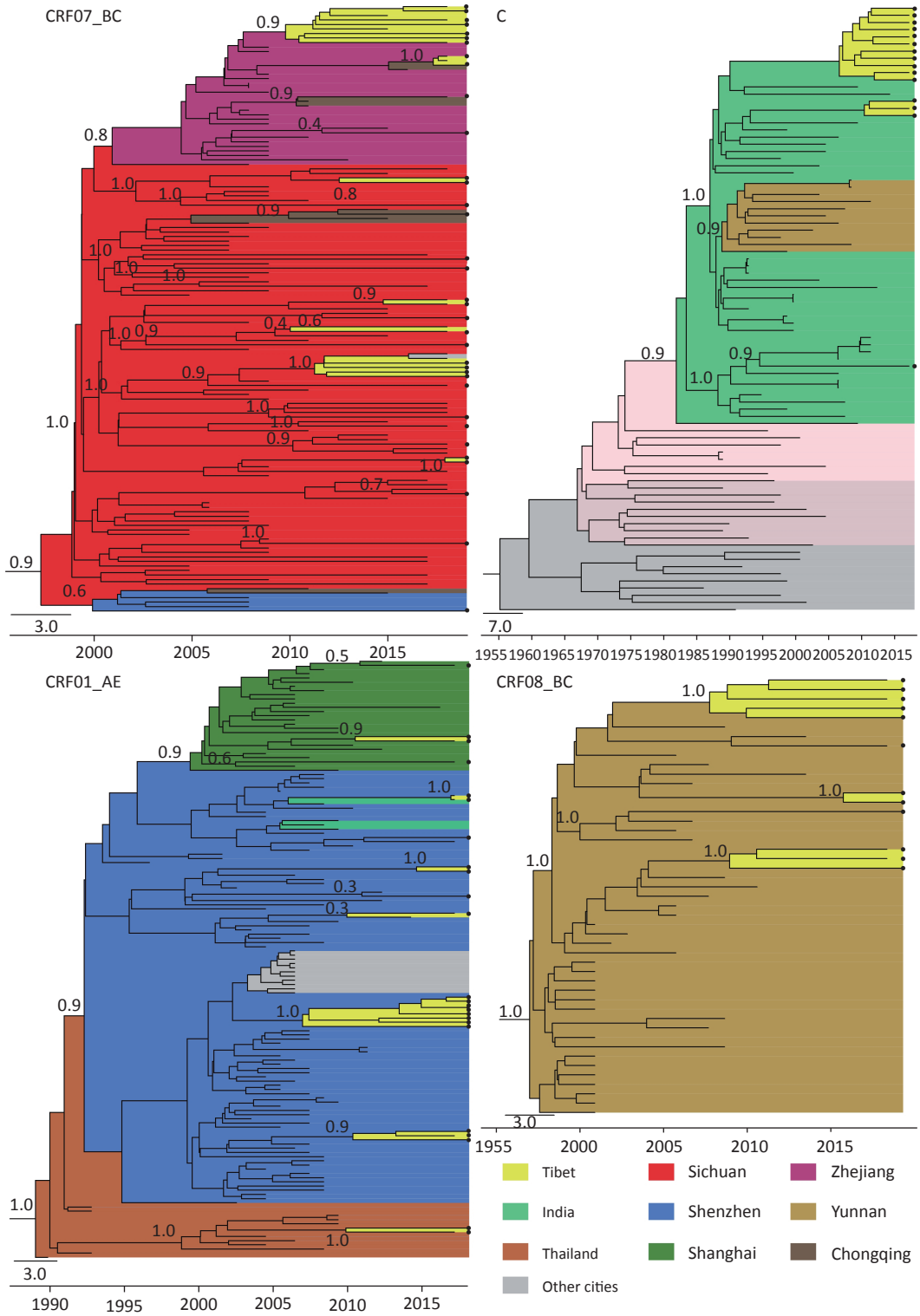


Figure 2. Bayesian time-scaled phylogeographic history of major subtypes of HIV-infected people in Tibet from 2018 to 2019. Background colors around branches represent the most probable source of the parental node of each branch. The values of probability inferred by Bayesian skygrid demographic model estimation for the most probable ancestral region were shown at the corresponding nodes. The small dots in the picture represent the sequences of Tibet.

has not accounted for more than 3% in China, and most of these strains have been transmitted in China in recombinant form, such as CRF07_BC and CRF08_BC[4]. The investigation in Tibet found that there were many subtypes C; thus, we focused on the spread of subtype C in Tibet.

We found that subtype C in Tibet was not transmitted from other provinces in China, instead directly from India to Tibet. To more accurately describe the spread of subtype C in Tibet, we considered the following two aspects: First, we added many sequences from Africa, India, and other locations to more accurately restore the transmission route of subtype C. Previous studies have shown that the origins of subtype C in South African, India, and China are 1960 (95% HPD, 1956–1964)[8], 1975 (95% HPD 1968–1981)[9], and 1990s[10]. This study found that the origin of subtype C in Yunnan, China was 1989 (95% HPD 1987–1993). Second, an epidemiological survey (mainly questionnaires and face-to-face interviews) of 15 patients with subtype C HIV-1 in Tibet was conducted, and important information was obtained. No direct links were identified between the 15 cases. Two patients reported that they had high-risk sexual behaviors with Indian women. Although these two patients did not know whether the women they had sex with had HIV-1 infection, and this information implies that patients with subtype C have sexual behaviors with Indian women. As a result, we conclude that the source of subtype C in Tibet is very reliable. Moreover, subtype C was dominated by Tibetans, suggesting that Tibetans are the main spreaders of subtype C. Additionally, there remain undiagnosed infections. Therefore, we should strengthen the management of the Tibet border and simultaneously strengthen the surveillance of HIV in the 'Belt and Road'.

The subjects in this study were sourced from primary treatment sites. Therefore, the study results may have a certain deviation. Most of the 15 subjects with subtype C infection were married. The questionnaire showed that these people were unwilling to mobilize those around them for active detection. Therefore, the study results on subtype C may be incomplete; thus, there is an urgent need to strengthen HIV-1 education and expand testing among Tibetans.

The proportion of Tibetans in this study was more than three-quarters; thus, this study is the first to describe, to our knowledge, not only the distribution of HIV subtypes in Tibet but also the HIV-1 strain subtypes in Tibetans. Due to the genotype of patients with HIV-1 infection in Tibet and the fact that subtype C was mainly transmitted from India to Tibetans in Tibet, the strengthening of HIV-1 surveillance as well as AIDS publicity and education among Tibetans is recommended to reduce high-risk sexual behavior between Tibetans as well as with individuals in other countries at the border. We also encourage active detection, early detection, and early treatment to control the rapid spread of subtype C and other HIV-1 strains among Tibetans.

All authors declare no competing interests.

The authors would like to thank SONG Yu Fei (Harvard TH Chan School of Public Health) for modifying the article language.

[&]These authors contributed equally to this work.

[#]Correspondence should be addressed to XING Hui, Professor, PhD, Tel: 86-10-58900647, E-mail: xingh@chinaaids.cn; FENG Yi, Tel: 86-10-58900642, E-mail: fengyi@chinaaids.cn

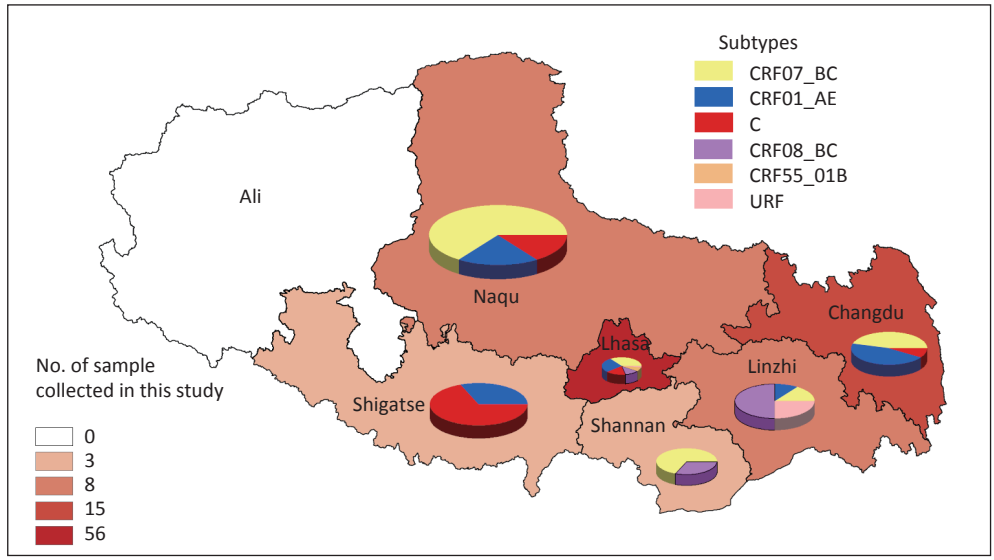
Biographical notes of the first authors: Dorjiwangmo, female, born in 1979, Master, majoring in STD/AIDS prevention; GAN Meng Ze, male, born in 1994, Master, majoring in HIV molecular epidemiology; DONG Ao Bo, female, born in 1989, Master, majoring in HIV drug resistance.

Received: February 9, 2020;

Accepted: November 9, 2020

REFERENCES

1. Liao LJ, Xing H, Li XP, et al. Genotypic analysis of the protease and reverse transcriptase of HIV type 1 isolates from recently infected injecting drug users in western China. AIDS Res Hum Retroviruses, 2007; 23, 1062−5.
2. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol, 2007; 7, 214.
3. Ronquist F, Teslenko M, van der Mark P, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol, 2012; 61, 539−42.
4. Li XS, Li W, Zhong P, et al. Nationwide trends in molecular epidemiology of HIV-1 in China. AIDS Res Hum Retroviruses, 2016; 32, 851−9.
5. Zeng HY, Li TT, Wang Y, et al. The epidemic dynamics of four major lineages of HIV-1 CRF01_AE strains after their introduction into China. AIDS Res Hum Retroviruses, 2016; 32, 420−6.
6. Hemelaar J, Gouws E, Ghys PD, et al. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS, 2011; 25, 679−89.
7. He X, Xing H, Ruan YH, et al. A comprehensive mapping of HIV-1 genotypes in various risk groups and regions across China based on a nationwide molecular epidemiologic survey. PLoS One, 2012; 7, e47289.
8. Wilkinson E, Engelbrecht S, de Oliveira T. History and origin of the HIV-1 subtype C epidemic in South Africa and the greater southern African region. Sci Rep, 2015; 5, 16897.
9. Shen CL, Craigo J, Ding M, et al. Origin and dynamics of HIV-1 subtype C infection in India. PLoS One, 2011; 6, e25956.
10. Luo CC, Tian CQ, Hu DJ, et al. HIV-1 subtype C in China. Lancet, 1995; 345, 1051−2.



Supplementary Figure S1. Distribution of HIV-1 genotypes and sample data of HIV-infected people who had not received any antiviral treatment in each prefecture across Tibet.