Identification of Subtype B, Multiple Circulating Recombinant Forms and Unique Recombinants of HIV Type 1 in an MSM Cohort in China

Wanhai Wang,^{1,2} Shulin Jiang,³ Shenwei Li,¹ Kai Yang,¹ Liying Ma,¹ Fengmin Zhang,²
Xiaoyan Zhang,¹ and Yiming Shao¹

Abstract

To study HIV-1 genetic diversity among HIV-positive men who have sex with men (MSM) in China, a cohort consisting of HIV-positive MSM was established in 2005 and was monitored every 2 years. In 2007, 44 HIV-positive MSM subjects were genotyped, and the results showed HIV-1 subtype B decreased from 77.5% to 41.9%, but non-B subtypes increased rapidly represented by CRF01_AE from 3.7% to 30.2% compared to 2005. In addition, one case of CRF07_BC was first identified in this population, which mainly circulated among HIV-1-infected injection drug users (IDUs) in China. There were 11 unique recombinant forms (URFs) consisting of a recombination of CRF01_AE with subtype B or CRF07_BC. The subtype-specific phylogenetic tree analysis showed that the genetic distance within subtype B group viruses was larger than the genetic distance within the CRF01_AE group ($p < 0.001$). Of the identified viruses in the Chinese MSM population, over 80% of subtype B viruses might originate from the United States and Brazil, and over 85% of the CRF01_AE viruses might originate from Thailand. In addition, epidemic study data showed that some of the HIV-1-infected MSM had foreign sexual partners (13.6%) and heterosexual activities (43.2%). The patients infected with HIV-1 URF viruses had more heterosexual encounters (54.5%) and more sexual partners (72.7%) compared to those infected with subtype B (44.4%; 33.3%) and CRF01_AE (23.1%; 69.2%) viruses. Taken together, we suspected that the genetic complexity of HIV-1 viruses identified in Chinese MSM populations was more likely a result of multiple introductions of viruses from the general population infected with HIV-1 through IDUs or heterosexual transmission.

Introduction

HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) has a high level of genetic variation, which originates from the high mutation and recombination rates of the reverse transcriptase enzyme combined with a high turnover rate.^{1,2} According to the current nomenclature, 9 “pure” genetic subtypes (A–D, F–H, J, and K) and 43 circulating recombinant forms (CRFs), namely CRF01–CRF43, have been identified so far within group M, which accounts for more than 90% of the reported HIV/AIDS cases worldwide.³

HIV-1 prevalent strains in China were also discovered with increasing complexity. Data from two national molecular epidemic surveys conducted from 1996 to 1998 and 2002 to 2004 indicated that HIV-1 subtypes and CRFs, including

A, B', B, C, D, F, G, J, CRF07_BC, CRF08_BC, CRF01_AE, and CRF02_AG, circulated in China.^{4–6} CRF_BC (07BC, 08BC), B', and CRF01_AE are the major forms. At the end of 2007, the total estimated cases of HIV/AIDS were more than 700,000, and the HIV-1-infected people who acquired HIV-1 from the sexual route increased to 57%. The percentage of men who have sex with men (MSM) increased rapidly from 7.3% in 2005 to 12.2% in 2007.⁷ Unpublished data predicted that the MSM proportion might reach as high as 15% of HIV infections by 2010.⁸ Engagement in high-risk behaviors, such as unprotected sex with multiple male or female partners, has rendered this group one of the most vulnerable to HIV infection in China.

Genetic differences among HIV-1 variants can influence the biological properties of the virus, the susceptibility to an-

¹State Key Laboratory for Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 100050, China.

²Department of Microbiology, Harbin Medical University, Harbin 150086, China.

³Chaoyang District Center for AIDS/STD Control and Prevention, Beijing 100021, China.

tiretroviral drugs, as well as vaccine development. Monitoring HIV genetic characteristics and subtyping prevalent isolates in China will provide essential clues for the development of an effective HIV vaccine, treatment regimens, as well as HIV/AIDS prevention and control strategies.²

We started the molecular epidemiological study on HIV-1 infection among MSM in Beijing in 2005, and our previous study evidenced that the prevalent type of HIV-1 among MSM in Beijing emerged with diversity. Except for HIV-1 subtype B initially reported in 2002,⁹ several non-B subtypes and CRFs were also identified.¹⁰ In order to systemically monitor the range of HIV-1 genetic evolution among HIV-positive MSM in Beijing, a cohort's study was conducted starting in 2005. After enrollment, all the subjects were followed up and sampled in 2005 and 2007. The collected blood samples were subjected to sequencing and phylogenetic analysis.

Materials and Methods

Study subjects

The cohort consisted of HIV-positive MSM subjects. They were screened for eligibility (age 18 years old or older and Beijing residence). Men who were eligible and willing to participate in the study provided written informed consent and completed a 30-min questionnaire administered by trained interviewers. The questionnaire elicited information about demographic characteristics (age, marital status, education, self-reported sexual orientation, having a Beijing residence card), lifetime history of HIV infection and sexually transmitted diseases (STDs), as well as number of male and female sexual partners lifetime and in the past 3 months. Their sexual experience with foreigners, frequency of anal and vaginal intercourse, condom use during anal and vaginal intercourse, and reasons for inconsistent use in the past 3 months were also investigated. The study was approved by the Committee for Human Research in the National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention.

Sample collection

Whole blood samples, anticoagulated with EDTA, were collected at Chaoyang District Center for AIDS/STD Control and Prevention, Beijing in China from July to September 2007. The plasma was separated no more than 4 h after collection, and stored at -80°C until further analysis.

HIV-1 RNA extraction, amplification, and sequencing

RNA extraction from plasma was performed using a High Pure Viral Nucleic Acid Kit (Roche, Germany), and HIV-1 cDNA was obtained by reverse transcriptase polymerase chain reaction (RT-PCR) using the TaKaRa One Step RNA PCR kit (TaKaRa Biotechnology Co., Ltd., Dalian, China) at 50°C for 30 min. HIV-1 cDNA was then subjected to nested PCR for amplification of the *gag*, *pol*, and *env* genes. For the *gag* and *pol* regions, the first and second rounds of PCR reagents, primer pairs, and thermocycling profiles have been previously described.¹⁰ New primers were used to amplify *env* fragments: first round: 44F (gp120) forward, 5'-ACAGTRCARTGYACACATGG-3' (R = A or G, Y = C or T) and 35R (gp120) reverse, 5'-CACTTCTCCAATTGTCCITCA-

3'; and second round: 33F (gp120) forward, 5'-CTGTTT-AATGGCAGICTAGC-3' and 48R (gp120) reverse, 5'-RATGGGAGGRGYATACAT-3' (540 bp for the C2 ~ V3 region of the *env* region). For the *env* thermocycling profile, cycling conditions were as follows: the first-round PCR was an initial melting step at 94°C for 2 min, followed by 35 cycles at 94°C for 15 s, 52°C for 45 s, and 72°C for 2 min; the second-round PCR was an initial melting step at 94°C for 2 min, followed by 30 cycles at 94°C for 30 s, 52°C for 30 s, and 72°C for 1.5 min. All reactions were followed by a final extension for 10 min at 72°C . The second round of PCR products was separated by agarose gel electrophoresis and purified using a Qiagen Gel Extraction Kit (QIAGEN Inc., Germany). Purified PCR products were sequenced bidirectionally with fluorescent dye terminators (Prism BigDye terminator cycle sequencing ready reaction kit; Applied Biosystems) and an automated DNA sequencer (Applied Biosystems model 377).

Phylogenetic analysis, subtyping, and genetic distances

The sequences were aligned with reference sequences of HIV-1 strains of all subtypes (<http://www.hiv.lanl.gov>). Multiple alignments were done automatically by CLUSTAL W¹¹ with minor manual adjustments. The alignment was edited using the BioEdit program version 5.0.9.¹² The phylogenetic analysis of the aligned sequences was performed by the neighbor-joining method of MEGA version 3.1.¹³ To analyze trends in subtype genetic diversity over time, genetic distances were calculated using the Kimura two-parameter model of MEGA version 3.1 and we reconstructed separate phylogenies for subtype B and CRF01_AE sequences, after removing all putative intersubtype recombinants. These phylogenies were used to calculate the divergence from a single root node for each sequence. For the subtype B phylogeny, the root was the node connecting the one subtype B reference to the B tree. For the subtype CRF01_AE phylogeny, the root was that node connecting the one subtype CRF01_AE reference to the CRF01_AE tree. The reliability of the branching patterns was evaluated by bootstrapping (1000 replicates).

Statistical analysis

All statistical calculations were done using the Prism program version 5.00 (GraphPad, San Diego, CA). Statistical significance was defined as a *p* value of <0.05 .

Results

Characteristics of the cohort of HIV-positive MSM

The molecular epidemic evolution of HIV-1 was monitored every 2 years. The study population of HIV-positive MSM was firstly recruited in 2005 ($n = 54$), and 10 subjects consecutively enrolled in 2007. Together with 34 newly identified HIV-positive MSM, the final cohort consisted of 44 subjects. Their major demographic characteristics are shown in Table 1. The average age of the study population was 30.0 ± 8.0 years old (ranging from 19 to 56 years), and the median of HIV-infected time was 8 months (ranging from 0.5 to 144 months). According to their self-report, 34.1% were bisexual/heterosexual, 43.2% had heterosexual intercourse in

TABLE 1. DEMOGRAPHIC CHARACTERISTICS OF THE STUDY COHORT POPULATION ACCORDING TO B/NON-B HIV-1 SUBTYPE

Variable	Total (n = 44 ^a)	Subtype B (n = 18)	Non-B subtype	
			CRF01_AE (n = 13)	URFs ^b (n = 11)
Age [years mean (SD)]	30.0 (8.0)	28.2 (6.5)	39.7 (9.8)	33.0 (8.4)
HIV infected time (months) [median (range)]	8 (0.5–144)	6.5 (0.5–53)	5 (1–65)	23 (5–144)
Marital status				
Never married	79.5%	77.8%	92.3%	72.7%
Ever married	20.5%	22.2%	7.7%	27.3%
Sexual orientation				
Homosexual and gay	65.9%	66.7%	84.6%	54.5%
Bi/heterosexual	34.1%	33.3%	15.4%	45.5%
Anal sex roles (ever)				
Insertive	25.0%	22.2%	23.1%	36.4%
Receptive	31.8%	16.7%	46.2%	45.4%
Both	43.2%	61.1%	30.8%	18.2%
Ever had sex with women				
No	56.8%	55.6%	76.9%	45.5%
Yes	43.2%	44.4%	23.1%	54.5%
Ever had sex with foreigners				
No	86.4%	77.8%	84.6%	90.9%
Yes	13.6%	22.2%	15.4%	9.1%
Lifetime number of male sex partners				
1–10	45.5%	66.7%	30.8%	27.3%
11 or more	54.5%	33.3%	69.2%	72.7%

^aThere were two cases not displayed in the non-B subtype: CYM027 (classified as CRF07_BC, the only one genetic form identified in this population) and CYM013 (only the *gag* gene was amplified).

^bURFs, unique recombinant forms.

their lifetime, and 20.5% had married women. More than 50% of the subjects (54.5%) had 11 or more male sexual partners and 11.4% had participated in group sex in their lifetime. Out of the 44 respondents, 6 (13.6%) had sex with foreigners from the United States, Australia, India, or Europe (England, Germany, the Netherlands, Italy, Romania, and Belgium). Study results of sexual behavior showed that 43.2% of MSM engaged in both insertive and receptive anal intercourse, and the others preferred insertive (25.0%) or receptive (31.8%) anal intercourse.

Subtyping and phylogenetic analysis

The partial genes of HIV-1 *gag* (671 bp, encoding a portion of p24 and p17), *env* (540 bp of the C2 ~ V3 region), and *pol* (1315 bp, encoding the protease gene and part of the reverse transcriptase gene region) were amplified with different successful percentage (*gag*, 100%; *env*, 93.2%; and *pol*, 81.8%). The corresponding subtypes of *gag*, *env*, and *pol* were determined by phylogenetic analysis using the neighbor-joining method. According to the results of phylogenetic analysis (Fig. 1), 47.7% (21/44) of *gag* sequences were classified as subtype B and 52.3% (23/44) as non-B subtypes including CRF01_AE (22/44) and CRF07_BC (1/44). For 41 *env* and 36 *pol* sequences, there was a higher percentage of subtype B (56.1%, 69.4%) and a lower percentage of non-B subtypes (43.9%, 30.6%) including CRF01_AE (16/41, 9/36) and CRF07_BC (2/41, 2/36), respectively.

Taken together, 18 out of the 43 samples (excluding CYM013 that was the only *gag* gene region amplified) were sorted into subtype B (41.9%) and 25 (58.1%) into non-B subtypes. CRF01_AE (13/25) dominated in the non-B subtype

group, and one CRF07_BC, usually seen in HIV-1-infected injecting drug users, was also identified (Table 2). Compared with the sequences identified in 2005, the percentage of CRF01_AE (typical heterosexually transmitted in China) increased significantly from 3.7% (2/54) to 30.2% in HIV-positive MSM in China, and the percentage of subtype B sequences decreased more than 20% to 86.2%, 71.1%, and 92.5% in the *env*, *gag*, and *pol* genes, respectively, in 2005. However, those were 56.1% (*env*), 47.7% (*gag*), and 69.4% (*pol*) in samples collected in 2007 (Fig. 2).

There were 11 sequences that did not match known CRFs, and these were classified as URFs (44.0%). The most common putative recombinants between URFs were CRF01_AE_{gag}/B_{pol/env}, which were found in four cases. CRF01_AE was more frequently observed among viruses with discordant genotypes than was subtype B. This was also the case for the single CRF07_BC_{pol/env} sequence, which was CRF01_AE in *gag* (Table 2).

To better understand the genetic relationship of the identified genetic forms in our study with HIV-1 prevalent in the general population, we further performed phylogenetic tree analysis using subtype B viruses identified in our study as well as reference strains in the HIV-1 database. The selected subtype B reference strains (*n* = 324) were identified from 32 countries worldwide (tree not shown). The results showed that HIV-1 subtype B in HIV-positive MSM could be sorted into three to five clusters, although sequences of *gag*, *env*, and *pol* from one subject were not always sorted into one cluster. The different clusters related to reference strains from different geographic areas. Over 80% of subtype B viruses identified in our study population had a closer relationship to reference strains from Brazil and the United States (*gag*,

FIG. 1. Phylogenetic tree analysis of HIV-1 *gag*, *env*, and *pol* gene sequences. The phylogenetic trees were constructed using neighbor-joining methods for the *gag* (A), *env* (B), and *pol* (C) sequences. The sample sequences and HIV-1 subtype reference sequences available in the Los Alamos database were aligned using CLUSTAL W with minor manual adjustments. The statistical robustness of the neighbor-joining tree and reliability of the branching patterns were confirmed by bootstrapping (1000 replicates). Values at the nodes indicate the percent bootstraps in which the cluster to the right was supported. Bootstraps of 70% and higher only are shown. All samples analyzed in the trees are shown by black symbols (triangles for the *gag* region, tetragons for the *env* region, and circles for the *pol* region). Square brackets at the right indicate the sample sequence subtypes. Horizontal branch lengths are drawn to scale. The reference sequences also used included the previously reported representative group O (O.BE.87.ANT70) and N (N.CM.95.YBF30) sequences (not shown).

FIG. 1. (Continued).

TABLE 2. HIV SUBTYPE DISTRIBUTION IN *gag*, *pol*, AND *env* FRAGMENTS SEQUENCED FOR THE 44 STRAINS^a

	<i>gag</i>	<i>pol</i>	<i>env</i>	n
Concordant genotypes	B	B	B	18
	CRF01_AE	CRF01_AE	CRF01_AE	7
	CRF07_BC	CRF07_BC	CRF07_BC	1
Discordant genotypes	CRF01_AE	B	B	4
	B	CRF01_AE	CRF01_AE	1
	CRF01_AE	B	CRF01_AE	1
	CRF01_AE	CRF07_BC	CRF07_BC	1
	B	CRF01_AE	B	1
Incomplete gene profiles	CRF01_AE	NS	CRF01_AE	6
	B	NS	CRF01_AE	1
	CRF01_AE	B	NS	2
	CRF01_AE	NS	NS	1

^aB, subtype B; NS, not sequenced.

B.BR.03.BREPM1040; *env*, B.US.95.5073.95; and *pol*, B.BR.03.BREPM1040).

A similar analysis was also performed for CRF01_AE viruses identified in our MSM population with 71 of CRF01_AE reference sequences from seven countries (tree not shown). The identified viruses also formed three to five clusters and the largest cluster (over 85% of the identified viruses) showed a closer connection to reference strains from Thailand including 01AE.TH.90.CM240 (*gag*), 01AE.TH.04.BKM (*env*), and,

01AE.TH.02.OUR769I (*pol*), as well as three reference sequences from China: 01AE.CN.04.04CNLN76 (*gag*) and 01AE.CN.06.FJ051 and FJ064 (*env*) from Liaoning and Fujian provinces.

Genetic distances of the distribution of HIV-1 subtypes among MSM

Pair genetic distances of intrasubtypes were calculated for *gag*, *env*, and *pol* sequences by phylogenetic tree analysis us-

FIG. 2. HIV-1 subtype diversification. Proportion between subtype B and CRF01_AE identified in samples collected from HIV-1-infected MSM in 2005 and 2007, respectively. The numbers of 2005 and 2007 indicated the year of sampling.

ing the Kimura two-parameter model of MEGA version 3.1. The mean of the genetic distance of the *gag* region was $5.2 \pm 2.6\%$ (ranging from 0.2% to 10.7%) within subtype B group, which was larger than that ($2.9 \pm 2.1\%$, ranging from 0% to 7.9%) within the CRF01_AE group ($p < 0.0001$). A similar trend was observed in the *pol* region, with the mean of the genetic distance being $3.9 \pm 1.3\%$ (ranging from 0.1% to 6.5%) within the subtype B group and $2.5 \pm 1.0\%$ (ranging from 0.2% to 4.6%) within the CRF01_AE group, respectively ($p < 0.0001$). For the *env* region, the diversity was enlarged with greater genetic distance: $13.5 \pm 5.4\%$ (ranging from 3.0% to 27.7%) within the subtype B group and $8.9 \pm 3.5\%$ (ranging from 0% to 17.0%) within the CRF01_AE group, respectively ($p < 0.0001$).

Characteristics of MSM-infected different genetic forms of HIV-1

In order to explore the epidemiological and behavioral factors associated with the diversity of HIV-1 strains prevalent among the MSM population, we designed a comprehensive questionnaire and conducted the survey in a private manner. Among MSM infected with subtype B, the median of HIV-infected time was 6.5 months (range 0.5–53 months), and 33.3% had more than 11 sexual partners in their lifetime. Some of them had sex with foreigners (22.2%) and women (44.4%), and the majority (61.1%) engaged in both insertive and receptive anal intercourse. However, MSM infected with non-B subtypes were homosexually oriented (84.6%) with a shorter average time (median, 5 months; ranging from 1 to 65 months) to acquire the HIV-1 non-B subtype virus. They preferred receptive anal intercourse (46.2%) with multiple male sexual partners (69.2%) compared to those infected with HIV-1 subtype B virus (16.7%, 33.3%, respectively) (Table 1). The subject (CYM027) infected with HIV-1 CRF07_BC denied the high-risk use of intravenous drugs in the survey. However, he had lived in Yunnan province, a well-known place for HIV-1 CRF07_BC-infected IDUs, and had heterosexual activities according to his self-report.

Discussion

In China, there may be as many as 18 million (10.18–25.45 million, accounted for 10–15% of the entire male population of mainland China) gay and male bisexuals.¹⁴ MSM often engage in high-risk behaviors for HIV-1 infection such as unprotected sexual activities, commercial sexual activities, and bisexual activities with multiple sexual partners.^{15–17} Restricted by political and cultural elements, by custom, and by traditional values concerning family and marriage, many MSM in China eventually get married and have children and conceal their homosexuality from society and their families. MSM has become one of the highest risk groups for HIV infection and transmission in China.^{7,17}

The Chinese Molecular Epidemiology Study was initiated in 1996. The last nationwide survey (2002–2004) covered the HIV-1-infected population including IDUs, former plasma donors (FPDs), and patients who acquired HIV-1 from heterosexual transmission. The data indicated that the dominant genetic forms of HIV-1 changed from subtype B (primary in FPDs) to CRF01_AE (heterosexual transmission) and CRF07/08_BC (primary in IDUs).^{4–6} In contrast, knowledge

about the molecular epidemiology of HIV-1 among MSM in China was limited.

Our study results evidenced significant changes of HIV-1 genetic forms among the MSM population in China. HIV-1 subtype B accounted for 41.9% in our study, which previously was dominant in HIV-positive MSM in China^{9–10} and is still dominant in developed areas such as North America, Western Europe, and Australia.^{18,19} In addition, non-B subtypes became predominant (58.1%), and the majority of HIV-1 non-B subtype viruses were CRF01_AE (52%), which was usually transmitted through heterosexual activities in Southeast Asia and Chinese southeast coast provinces,²⁰ and was identified among Chinese HIV-infected MSM in 2005. Identification of CRF07_BC among the MSM studied further increased the genetic complexity of HIV-1 strains, which mainly circulated among IDUs in Northwest and Southwest China.^{21,22} The primary recombinant of URFs was CRF01_AE^{gag}/CRF07_BC^{pol/env} (Table 2). All this suggested that the MSM cohort studied must have had interactions with other high-risk populations for HIV infection. A similar trend was observed among the MSM population worldwide, with subtype D in South Africa,²³ subtype F in Colombia,²⁴ and CRF12-BF and recombinants of B, C, and F in Argentina.²⁵

It is curious to think about the power driving such genetic complexity of HIV-1 strains among MSM in China. Is it caused by genetic evolution from a single founder virus or by multiple viruses introduced at different time points? First, we calculated the genetic distances of intra-/inter-HIV subtype B and CRF01_AE, respectively, and found that the ge-

FIG. 3. Genetic distances of HIV-1 subtype B and CRF01_AE. The distribution of genetic distances between subtype B and the corresponding CRF01_AE is depicted with mean $\pm$ SD, respectively. The genetic distance in the *gag*, *env*, and *pol* gene regions between subtype B and the corresponding CRF01_AE was significantly different ($p < 0.0001$). The differences in genetic distances among the *gag*, *env*, and *pol* gene regions in subtype B were also significantly different ($p < 0.0001$). However, within CRF01_AE, the significant differences for genetic distance were observed only between *env* and *gag* or *pol* ($p < 0.0001$).

netic distance within subtype B group viruses was larger than that within the CRF01_AE group (Fig. 3, $p < 0.001$). Second, the genetic distance in the *env* region was 8.9% in the MSM population, which was comparable to that among the heterosexually transmitted population in Fujian (9.3%) but higher than that in Thailand (5.6%). However, the mean genetic distance of *gag* and *pol* within the MSM CRF01_AE group (2.9%, 2.5%) was significantly lower than that within the CRF01_AE virus prevalent in Thailand (4.2%, 3.2%) and the Chinese heterosexual population (6.3%, 5.3%).²⁰ Third, epidemiological data also suggested that the HIV-1 subtype B viruses (in 1992)⁹ entered into the Chinese MSM population earlier than CRF01_AE (around 2002 to 2005).¹⁰ In addition, 22.2% of subtype B-infected and 15.4% of CRF01_AE virus-infected MSM had foreign sexual partners, and 43.2% of the study subjects had heterosexual activity. The subtype-specific phylogenetic tree analysis with more reference sequences (324 of subtype B reference viruses and 71 of CRF01_AE reference sequences) provided more clues about the introduction of such complex genetic forms into our studied MSM population. Of the identified viruses in the Chinese MSM population, over 80% of subtype B virus might originate from the United States and Brazil, and over 85% of CRF01_AE viruses might originate from Thailand. In addition, the patients infected with HIV-1 URF viruses seem to have been engaged in higher risk behaviors for HIV infection such as heterosexual intercourse (54.5%) and having more than 11 sexual partners (72.7%) compared to those infected with the subtype B (44.4%; 33.3%) and CRF01_AE (23.13%; 69.2%) viruses. Therefore, they might have acquired HIV-1 from multiple infection routes.

Taken together, we suspected that the genetic complexity of HIV-1 viruses identified in Chinese MSM populations was more likely a result of the multiple introduction of viruses from the general population infected with HIV-1 through IDU use or heterosexual transmission, but not from a founder virus. Genetic diversity presents a great challenge for the current surveillance system; current monitoring technology needs to be improved to be able to predict future HIV-1 strains within high-risk populations including MSM.

Sequence Data

The nucleotide sequences shown in the present study have been submitted to GenBank with the following accession numbers: the *gag* region sequences EU921908–EU921951; the *pol* region sequences EU921952–EU921987; and the *env* region sequences EU921988–EU922028.

Acknowledgments

We gratefully acknowledge the contributions of the study participants. This study was supported by grants from the Ministry of Science and Technology of China (2004BA719A01), the Ministry of Health of China (WA200601), the Global foundation (IV) AIDS Program (2007OR04), and the National Natural Science Foundation of China (30671941).

Disclosure Statement

No competing financial interests exist.

References

1. Cornelissen M, Kampinga G, Zorgdrager F, *et al.*: Human immunodeficiency virus type 1 subtypes defined by *env* show high frequency of recombinant *gag* genes. *J Virol* 1996;70:8209–8212.
2. Hemelaar J, Gouws E, Ghys PD, *et al.*: Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006;20:W13–23.
3. HIV Sequence Database Web Site, <http://hiv-web.lanl.gov>. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
4. Shao YM, Su L, Sun XH, *et al.*: Molecular epidemiology of HIV infection in China. Presented at the XII International Conference on AIDS, Geneva, Switzerland, June 28–July 3, 1998 (abstract No. 13132).
5. Shao YM: AIDS in south and southeast Asia. *HIV/AIDS: Perspective on China*. *AIDS Patient Care STDs* 2001;15:431–432.
6. Ministry of Science and Technology, China: Research progress: Nationwide HIV molecular epidemiology survey and its database setting (May 30, 2004). Available at <http://shgy.jhgl.org/shownews.asp?newsid=893>. Last accessed 05/25/2006 (in Chinese).
7. China State Council AIDS Working Committee Office and UN Theme Group on HIV/AIDS in China. A Joint Assessment of HIV/AIDS Prevention, Treatment and Care in China (2007). Available at <http://www.chinaids.org.cn>. Last accessed November 29, 2007.
8. Zhang BC and Chu QS: MSM and HIV/AIDS in China. *Cell Res* 2005;15:858–864.
9. Yao J, Zhang FJ, He ZP, *et al.*: Subtype and sequence analysis of the C2-V3 region of *env* gene among HIV-1 infected homosexual men in Beijing. *Chin J STD AIDS Prev Control* 2002;8:131–133.
10. Zhang XY, Li SW, Li XP, *et al.*: Characterization of HIV-1 subtypes and viral antiretroviral drug resistance in men who have sex with men in Beijing, China. *AIDS* 2007;21:S59–65.
11. Thompson JD, Higgins DG, and Gibson TJ: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
12. Hall T: Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 96/98/NT. *Nucleic Acid Symp Ser* 1999;41:95–98.
13. Kumar S, Tamura K, and Nei M: MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings Bioinform* 2004; 5:150–163.
14. Zhang BC, Li XF, Shi TX, *et al.*: A preliminary estimate to population size and HIV infection rate among male homosexuals/bisexuals in China. *Chin J STD AIDS Prev Control* 2002;8:197–199.
15. Qu SQ, Zhang DP, Wu YH, *et al.*: Prevalence of HIV infection and sexual risk behaviors among men who have sex with men in northeastern China. *Chin J STD AIDS Prev Control* 2002;8:145–147.
16. Liu H, Liu Y, and Xiao Y: HIV-related knowledge, attitudes, and behavior among men who have sex with men in Beijing, China. *Chin J STD AIDS Prev Control* 2001;7:289–291.
17. Choi KH, Gibson DR, Han L, *et al.*: High levels of unprotected sex with men and women among men who have sex with men: A potential bridge of HIV transmission in Beijing, China. *AIDS Educ Prevent* 2004;16:19–30.
18. Expert Group of the Joint United Nations Programme on HIV/AIDS.

19. Eshleman SH, Husnik M, Hudelson S, *et al.*: Antiretroviral drug resistance, HIV-1 tropism, and HIV-1 subtype among men who have sex with men with recent HIV-1 infection. *AIDS* 2007;21:1165–1174.
20. Huang HL, Zheng J, Yan PP, *et al.*: Genetic characterization of CRF01_AE full-length human immunodeficiency virus type 1 sequences from Fujian, China. *AIDS Res Hum Retroviruses* 2007;23:569–574.
21. Shao YM, Zhao F, Yang WZ, *et al.*: The identification of recombinant HIV-1 strains in IDUs in southwest and northwest China. *Chin J Exp Clin Virol* 1999;13:109–112.
22. Zhang Y, Lu L, Ba L, *et al.*: Dominance of HIV-1 subtype CRF01_AE in sexually acquired cases leads to a new epidemic in Yunnan province of China. *PLoS Med* 2006;3:e443.
23. Harmelen JV, Wood R, Lambrick M, *et al.*: An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. *AIDS* 1997;11:81–87.
24. Eyzaguirre L, Bautista CT, Ayala C, *et al.*: First case of HIV type 1 subtype F among men who have sex with men in Colombia. *AIDS Res Hum Retroviruses* 2006;22:808–811.
25. Pando MA, Eyzaguirre LM, Segura M, *et al.*: First report of an HIV-1 triple recombinant of subtypes B, C and F in Buenos Aires, Argentina. *Retrovirology* 2006;3:59.

Address reprint requests to

Yiming Shao

Xiaoyan Zhang

National Center for AIDS/STD Control and Prevention

Chinese Center for Disease Control and Prevention

No. 27 Nanwei Road

Xuanwu District

Beijing 100050, P.R. China

Fenming Zhang

Department of Microbiology

Harbin Medical University

Harbin 150006, China

E-mail: zhang_xycn2002@yahoo.com.cn