

# Subtype B Was the Dominant Strain Among HIV Type 1 Infections Except for the Population of Men Who Have Sex with Men in Harbin City, China

Bing Shao,<sup>1,2</sup> Wen-Jing Li,<sup>2</sup> Ting Liu,<sup>3</sup> Qing-Hai Li,<sup>4</sup> Hang Li,<sup>1</sup> Man-Li Chang,<sup>5</sup>  
Chao-Qun Huang,<sup>2</sup> Fu-Xiang Wang,<sup>2</sup> and Bin-You Wang<sup>1</sup>

## Abstract

We sought to identify the prevalent subtypes and study the genetic variation of HIV-1 circulating in HIV infections in Harbin City, China. Forty-seven samples from the *env* V3-V4 region were successfully sequenced and analyzed, which involved thirty-one men who have sex with men (MSM), eight heterosexuals, seven former plasma donors (FPD)/blood transfusion recipients (BT), and one injection drug user (IDU). In all, 46.8% of CRF01\_AE, 40.4% of subtype B, and 12.8% of CRF07\_BC were identified. CRF01\_AE (64.5%) was the dominant strain in MSM, and subtype B (81.2%) was the chief strain in other infected subjects except for the MSM population. Among all the genotypes, the B subtype possesses greater diversity of the tetramer on the tip of V3 loop than CRF07\_BC and CRF01\_AE, in which the peculiar GWGR was commonly found. Because nationwide there is a trend toward the increasing presence of CRF01\_AE, a consecutive surveillance campaign was necessary among all HIV vulnerable populations in this locality.

UNTIL 2011, THERE WERE 780,000 people in China infected with HIV: 46.5% of individuals received the HIV virus through heterosexual transmission, 17.4% through homosexual transmission, 28.4% through injection drug use, and 6.6% through blood transfusion or donation.<sup>1</sup> Among the 48,000 new infections estimated for 2011, heterosexuals accounted for 52.2%, men who have sex with men (MSM) accounted for 29.4%, and injection drug users (IDUs) accounted for 18.0%.<sup>1</sup> Compared to 2006, the proportion of reported cases infected through homosexual and heterosexual transmission in 2011 has increased from 2.5% to 13.7% and 30.6% to 62.2%, respectively.<sup>1</sup> From the viewpoint of molecular epidemiology, HIV-1 with subtype B, CRF01\_AE, and CRF07\_BC were coepidemic in the population of MSM in China.<sup>2-6</sup> While among IDUs, CRF07\_BC and CRF08\_BC are circulating widely, and subtype B' is dominant among former plasma donors,<sup>5</sup> CRF01\_AE is the principal strain circulating among heterosexuals in China.<sup>7</sup>

In recent years, MSM have been responsible for the tremendous increase in the outbreak of the HIV epidemic in China, and have attracted the attention of more and more

researchers currently working in China. It was reported based on a recent prospective cohort study that the incidence of HIV among MSM in Harbin increased from 1.0% in 2006 to 7.5% in 2010.<sup>8</sup> And another study conducted in Harbin in 2011 showed that the prevalence of HIV among MSM has risen to 9.5%.<sup>9</sup> The incidence of HIV in MSM has entered a stage of rapid growth in Harbin. Although epidemiological surveys of the incidence of HIV and related risk factors for HIV infection in MSM have been carried out repeatedly in this region, molecular investigations of HIV in Harbin MSM and the other specific high-risk groups are scarce.

To identify the circulating HIV-1 subtypes and study the variations of genes, we conducted this study in Harbin in 2011. All study subjects signed informed consent forms for the collection of blood samples and subsequent analyses. This study was approved by the Medical Research Ethics Committee of the No. 4 Hospital of Harbin Medical University.

Anticoagulant peripheral blood samples were collected and peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Hypaque density gradient centrifugation (Amersham Biosciences AB, Uppsala, Sweden) from 47

<sup>1</sup>Department of Epidemiology, Public Health College of Harbin Medical University, Harbin, China.

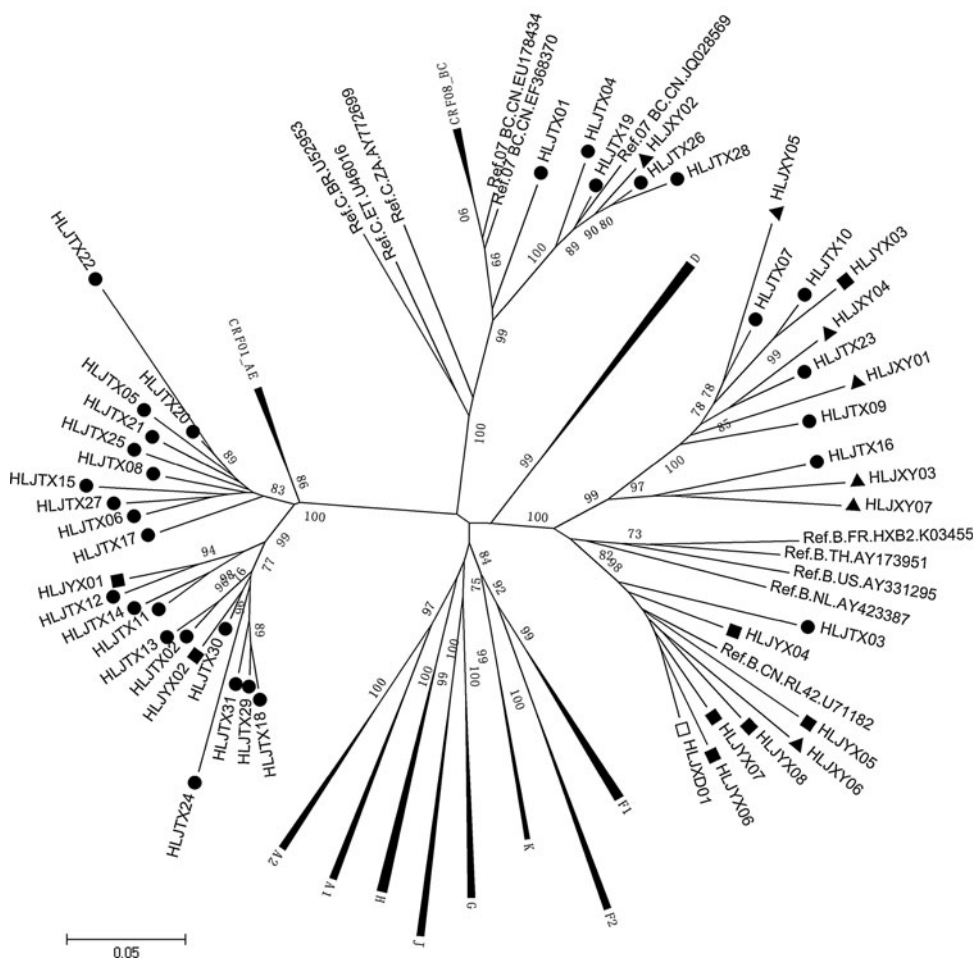
<sup>2</sup>Department of Infectious Diseases, Fourth Affiliated Hospital of Harbin Medical University, Harbin, China.

<sup>3</sup>Harbin Municipal Center for Disease Control and Prevention, Harbin, China.

<sup>4</sup>Center for Genome Research of Harbin Medical University, Harbin, China.

<sup>5</sup>Department of Clinical Laboratory, Second Affiliated Hospital of Harbin Medical University, Harbin, China.

**FIG. 1.** Phylogenetic analysis of the V3-V4 region of the *env* gene in HIV-1 strains from infections in Harbin City. The bootstrap probability (1,000 replicates) is indicated at the corresponding nodes of the tree. The scale bar indicates the evolutionary distance of 5% (0.05 substitutions per site). The reference sequences of subtypes A-D, F-H, J, K, CRF01\_AE, CRF07\_BC, and CRF08\_BC were obtained from the HIV database ([www.hiv.lanl.gov/](http://www.hiv.lanl.gov/)). The black solid circles, black diamonds, and black triangles of the sequences represent men who have sex with men (MSM), heterosexuals, and former plasma donors/blood transfusion (FPD/BT) recipients, respectively, and the open box of sequence represents intravenous drug user (IDU) analyzed in this study.



HIV-1-positive infections confirmed by Western blot assay, including 31 cases involving MSM, eight cases involving heterosexuals, seven cases involving former plasma donors/blood transfusion (FPD/BT) recipients, and one case involving IDU. They were all newly diagnosed cases and none involved any antiviral drug treatment. Viral DNA was extracted from PBMCs using the QIAamp Viral DNA Mini kit (Qiagen, Valencia, Roche, Germany). The V3-V4 region of the *env* gene fragment was amplified by two rounds of nested polymerase chain reaction (PCR) to obtain sufficient amounts of DNA followed by sequencing by the Beijing Invitrogen Life Technologies Corporation.

Two sets of primers were designed. The outer primers were 5'-TGGGATCAAAGCCTAAAGCCATGTG-3' and 5'-AGTGCTTCCTGTGCTCCCA-3'; the inner primers were 5'-CTGTTAAATGCCAGTCTAGC-3' and 5'-ACTTCTCCAATTGTCCTCA-3'. The first amplification round was performed at 94°C for 2 min, 55°C for 45 s, and 72°C for 1 min with one cycle, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min. The total volume of the PCR system was 30 µl; products were then kept at 4°C. The total volume of the second round of the PCR system was amplified to 50 µl, using inner primers with conditions similar to the first round: 94°C for 2 min, 55°C for 45 s, and 72°C for 1.5 min with one cycle, followed by 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min for 30 cycles; products were then kept at 4°C. The successfully

obtained nucleotide sequences were aligned by using CLUSTAL W (included in Mega 5.05 software), and the alignments were edited manually and amino acid sequences were deduced by BioEdit (version 7.00). Genetic distances were computed using the Kimura two-parameter substitution model with both transitions and transversions.

Phylogenetic analyses were performed with the neighbor-joining statistical method based on the maximum composite likelihood model as implemented by Mega 5.05 software with subtyping reference sequences obtained from the Los Alamos HIV sequence database. The reliability of the topologies was evaluated by performing bootstrap analysis with 1,000 replicates.

**TABLE 1.** DISTRIBUTION OF HIV-1 SUBTYPES IDENTIFIED BASED ON THE V3-V4 REGION OF THE *env* GENE IN EACH RISK GROUP

	CRF01_AE	CRF07_BC	B	Total
Homosexuals	20	5	6	31
Heterosexuals	2	—	6	8
Former plasma donors/blood transfusion recipients	—	1	6	7
Injection drug users	—	—	1	1
Total	22	6	19	47

CRF01\_AE

HXB2	CTRPNNNTRKRIRIQF	GPGRA	FVTIGK-IGNMRQAHC
HLJTX02	.I..S....TS.K.--	...CT..YRTXDIX.DI.K.F.	
HLJTX05	....S....TS.P.--	...Q...YRT.DI..DI...F.	
HLJTX06	....S....S.T.--	...QTL.YRT.DI..DI...Y.	
HLJTX08	....S....TS.N.--	...C...YRT.DI..DI...Y.	
HLJTX11	....S....TS.S.--	...XT..YRT.DI..DI.K.F.	
HLJTX12	....S....TS...--	...QM..YRT.DI..DI.K.Y.	
HLJTX13	....S..I.VST...--	...QV..YKT.DI..DI.K.F.	
HLJTX14	....F....TS...--	...QM..XRT.DI..DI.K.Y.	
HLJTX15	....S....ITS.S.--	...YRT.DIK.DI...Y.	
HLJTX17	....Y....TS.N.--	...YRT.DI..DI...Y.	
HLJTX18	....S..X.TST...--	...QV..YRT.DIT.DI.X.Y.	
HLJTX20	....S....TS.X.--	...XX..YRT.DI..DI...Y.	
HLJTX21	....S....TS.Q...--	...YRT.DI..DI....	
HLJTX22	....S....TS.PM--	...Q...YRT.DI..I...Y.	
HLJTX24	....FT.R.IGT...--	...HL..YRTEGIT.DI.KTY.	
HLJTX25	....S....TS.H.--	...QT..YRT.DI..DI...Y.	
HLJTX27	....S....TS.SL--	...YST.DI..DI...Y.	
HLJTX29	....S....TST...--	...QV..YRT.DIT.DI.K.F.	
HLJTX30	....S....TSX...--	...QV..YRT.XIT.DI.K.Y.	
HLJTX31	....S..V.TSS...--	...QV..YRT.AIT.DI.K.R.	
HLJYX01	...XF....TS...--	...QM..YRT.DI..DI.K.Y.	
HLJYX02	....S....SX...--	...QV..YRT.DIV.DI.K.Y.	

CRF07\_BC

HXB2	CTRPNNNTRKRIRIQF	GPGRAF	FVTIGK-IGNMRQAHC
HLJTX01	.....S....--	...QT.YAT.DI..DI....	
HLJTX04	....G....S...--	...QT.YAT.DI..DI....	
HLJTX19	....G....S...--	...QT.YAT.DV..DI....	
HLJTX26	....G....S...--	...QT.YAT.EI..DI....	
HLJTX28	....G....S...--	...QT.YAT.DI..DI....	
HLJXY02	....G....S...--	...QT.YAT.DI..DI....	

B subtype

HXB2	CTRPNNNTRKRIRIQR	GPGR	AFVTIGK-IGNMRQAHC
HLJTX03	....TD..RGLHVGHV	...K..VY.T.-I..DL...Y.	
HLJTX07	....G....G.HT--	...W...XLYVT.QI..DI....	
HLJTX09	....G....XS.HM--	...W...SLYVT.QI..DI....	
HLJTX10	....G.....H.--	...L...SVYVTRQI...I.K...	
HLJTX16	.....S.HF--	...YSLYAT..I..DIXX...	
HLJTX23	.....G.HM--	...W...SLYVT.DI..DI....	
HLJYX03	....G....S.HM--	...W...SLLATGQVRDI.K...	
HLJYX04	....X....SVHX--	...Q...FAT.QI...I...X.	
HLJYX05	.....X.S.PL--	...WY.T.QIX.DI....	
HLJYX06	.....S.PL--	...Q...WYAT.QI..DI....	
HLJYX07	.....S.PL--	...Q...WY.S.QI..DI.R...	
HLJYX08	.....S.PL--	...K..WY.T.QI...I....	
HLJXY01	....A....S.HM--	...W...LYVT.QIV.KI....	
HLJXY03	....S....RS.HV--	...SLF.TXXI...I.L...	
HLJXY04	....D....G.HV--	...W...SLYVT.QIV.DI.K.Y.	
HLJXY05	....S....RG.HM--	...W...SIYVT.QI..DI....	
HLJXY06	.....S.PL--	...X...WXAT.EI..DI....	
HLJXY07	.....S.PT--	...V...YAT.DI..DI.K...	
HLJXD01	.....S.PL--	...X.X..WYXT.QI..XI....	

FIG. 2. Deduced amino acid residues of the V3 region derived from CRF01\_AE, CRF07\_BC, and the B subtype in HIV-1 strains in Harbin City versus the reference sequences (HXB2). The deduced amino acids are shown by the single-letter code. Tetramers on the tip of the V3 loop are indicated on the consensus sequence by open boxes. The dots indicate consensus of the amino acid identified and the dashes correspond to gaps introduced to maintain the correct alignment.

Phylogenetic analysis demonstrated that the HIV-1 subtypes identified among 31 MSM were as follows: 64.5% were CRF01\_AE, 19.4% were B subtypes, and 16.1% were CRF07\_BC. Among the eight heterosexual infections 25.0% were CRF01\_AE and 75.0% were subtype B. Among the seven HIV infections through blood transfusion/donation 14.3%

CRF07\_BC and 85.7% subtype B were identified. The B subtype was identified only in the IDU (Fig. 1). Our study showed that CRF01\_AE was the dominant strain circulating in MSM. However, other than the MSM population, subtypes of B were the principal strains (81.2%) in this region (Table 1). In addition, the largest mean genetic distance was found among

subtype B ( $12.6 \pm 1.1\%$ ), followed by CRF01\_AE ( $8.9 \pm 0.8\%$ ); the smallest mean genetic distance was found in CRF07\_BC ( $7.1 \pm 0.8\%$ ). It can be concluded that subtype B has the longest epidemic time among HIV infections in this region.

The nucleotide sequences of the *env* V3 sequences described in this study were translated and the corresponding putative products were aligned and compared with the reference sequence HXB2 (7110–7217) available in the Los Alamos HIV sequence database. V3 of the *env* gene is emphasized in this study because of the importance of the role of this domain in viral entry, eliciting antibodies.<sup>10</sup> V3 of *env* is a cysteine-bound loop structurally composed of 35 amino acids, with the principal neutralizing antibody determinant (PND) located in the crown of the V3 loop,<sup>10,11</sup> which is constituted of the specific tetramer (which locates the position of 15–18 of the V3 amino acid sequence). V3 in different HIV-1 subtypes has distinct antigenic properties.<sup>10,11</sup> In this study, the main tetramers on the tip of the V3 loop in CRF01\_AE and CRF07\_BC were GPGQ (15/22) and (6/6), respectively (Fig. 2), whereas the tetramer in CRF07\_BC was the most conservative in our study. However, in subtype B, tetramer structures on the crown of the V3 loop represented the maximum variation compared to CRF01\_AE and CRF07\_BC, in which GWGR accounted for the largest proportion (7/19) (Fig. 2).

The V3 loop of the *env* gene is a focus of HIV research because this region of protein is a principal target for neutralizing antibodies and a major determinant for cell tropism and syncytium formation.<sup>11</sup> Understanding how and why *env* V3 regions differ between subtypes is necessary to assess the genetic diversity of HIV-1 in vaccine design and to develop anti-HIV drugs. V3 has long been a target of interest for entry-based inhibitors because of its critical role in defining the specificity of the interaction of *env* with cellular coreceptor molecules, usually CCR5 or CXCR4, to facilitate entry into target cells.<sup>10,11</sup>

Our study has indicated that the tetramer on the crown of the V3 loop has greater diversity, especially for subtype B, compared with a previous study conducted in Harbin City published in 2007, which reported that 52.63% of the V3 tip amino motif was GPGQ and 21.53% was GPGR among 19 HIV-1-seropositive infectors that were identified with subtype B'.<sup>12</sup> Moreover, the impact of the peculiar GWGR on the tip of the V3 loop among the B subtypes is worth noting for the development of profiles of HIV pathogenesis and vaccine design. In addition, we also first determine the distribution of the tetramer on the tip of the V3 loop on CRF01\_AE and CRF07\_BC strains in this region, which were all highly conserved in both of the strains, with GPGQ identified as the primary epitope within them (Fig. 2).

The distribution of HIV subtypes and circulating recombinant forms (CRFs) has been broadly linked with geographic location and risk groups. Subtype distribution of the global HIV pandemic has diversified extensively through mutation and recombination, partly driven by a combination of population mobility, sexual mixing, and the impact of anti-retroviral therapies.<sup>13</sup> Monitoring the distribution of HIV-1 subtypes has important significance in determining the dynamics of HIV virus transmission as well as for diagnosis, treatment, and vaccine development. Our results reveal that the prevalence of HIV in this region has varied greatly, which is reflected not only by the diversity of genotypes but also by the extent of HIV gene variation.<sup>12</sup>

The CRF01\_AE in MSM in this study was grouped into two distinct clusters, similar to the results of other studies reported in China.<sup>3,14</sup> A previous study<sup>4</sup> conducted in Liaoning province indicated that one of the two CRF01\_AE clusters may have originated in the southern or southwestern provinces and then spread to North China along the southeast coast, and the other cluster of strains of geographic origin might have been in Yunnan, but that the strains might have disseminated directly into northeastern regions. However, these two clusters of CRF01\_AE strains were both transmitted into China from Thailand. Because of the close geographic distance, we inferred that the two distinct CRF01\_AE clusters in Harbin MSM may have been introduced from MSM in Liaoning province.

Our results, consistent with reports about MSM in other regions in China, showed that CRF01\_AE was the dominant strain circulating in the population of MSM.<sup>3–6</sup> Compared with other risk groups, individuals with HIV infection in this region were infected through injection drug use and blood and heterosexual contact, with the dominant strain being subtype B. The B subtype has historically been the primary subtype in Harbin.<sup>12,15</sup> However, the prevalence of CRF01\_AE seemingly showed an increasing trend in MSM.<sup>2,5</sup> The factors associated with the rapid increase of CRF01\_AE in MSM are not completely known. It has been reported that sexual transmission with the dominant CRF01\_AE strain is a major risk factor in the current outbreak of HIV-1 in the general population of Guangdong, and the drug-resistant variants are starting to emerge.<sup>16</sup> However, whether the CRF01\_AE strain will replace other subtypes and further expand to the general population and then become the dominant strain is uncertain. Ongoing surveillance of the molecular epidemiology of these HIV vulnerable populations is necessary, which will be an onerous task for disease control and prevention personnel and for all health-related staff.

#### Nucleotide Sequence Accession Numbers

The nucleotide sequences that we obtained for the V3-V4 region of *env* genes from each patient have been submitted to GenBank with accession numbers KC953044–KC953090.

#### Acknowledgments

This work was supported by grants from the Fourth Affiliated Hospital of Harbin Medical University Funds for Distinguished Young Scientists and the Department of Education Foundation of Heilongjiang province, China (no. 12521227).

#### Author Disclosure Statement

No competing financial interests exist.

#### References

1. Ministry of Health of the People's Republic of China, Joint United Nations Programme on HIV/AIDS, World Health Organization (2011): The report of the HIV/AIDS epidemic estimation in China in 2011.
2. Wang W, Jiang S, Li S, *et al.*: Identification of subtype B, multiple circulating recombinant forms and unique recombinants of HIV type 1 in an MSM cohort in China. *AIDS Res Hum Retroviruses* 2008;24(10):1245–1254.

3. Li L, Lu X, Li H, *et al.*: High genetic diversity of HIV-1 was found in men who have sex with men in Shijiazhuang, China. *Infect Genet Evol* 2011;11(6):1487–1492.
4. An M, Han X, Xu J, *et al.*: Reconstituting the epidemic history of HIV strain CRF01\_AE among men who have sex with men (MSM) in Liaoning, northeastern China: Implications for the expanding epidemic among MSM in China. *J Virol* 2012;86(22):12402–12406.
5. He X, Xing H, Ruan Y, *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(10):e47289.
6. Wu J, Meng Z, Xu J, *et al.*: New emerging recombinant HIV-1 strains and close transmission linkage of HIV-1 strains in the Chinese MSM population indicate a new epidemic risk. *PLoS One* 2013;8(1):e54322.
7. 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(11):e443.
8. Wang K, Yan H, Liu Y, Leng Z, Wang B, and Zhao J: Increasing prevalence of HIV and syphilis but decreasing rate of self-reported unprotected anal intercourse among men who had sex with men in Harbin, China: Results of five consecutive surveys from 2006 to 2010. *Int J Epidemiol* 2012;41(2):423–432.
9. Zhang L, Zhang D, Yu B, *et al.*: Prevalence of HIV infection and associated risk factors among men who have sex with men (MSM) in Harbin, P. R. China. *PLoS One* 2013;8(3):e58440.
10. Galanakis PA, Spyroulias GA, Rizos A, Samolis P, and Krambovitis E: Conformational properties of HIV-1 gp120/V3 immunogenic domains. *Curr Med Chem* 2005;12(13):1551–1568.
11. Lynch RM, Shen T, Gnanakaran S, and Derdeyn CA: Appreciating HIV type 1 diversity: Subtype differences in Env. *AIDS Res Hum Retroviruses* 2009;25(3):237–248.
12. Wang FX, Zhou H, Ling H, *et al.*: Subtype and sequence analysis of HIV-1 strains in Heilongjiang Province. *Chin Med J (Engl)* 2007;120(22):2006–2010.
13. Hawke KG, Waddell RG, Gordon DL, Ratcliff RM, Ward PR, and Kaldor JM: HIV non-B subtype distribution: Emerging trends and risk factors for imported and local infections newly diagnosed in South Australia. *AIDS Res Hum Retroviruses* 2013;29(2):311–317.
14. Ye J, Xin R, Yu S, *et al.*: Phylogenetic and temporal dynamics of human immunodeficiency virus type 1 CRF01\_AE in China. *PLoS One* 2013;8(1):e54238.
15. Liu SY, Zeng FF, Huang GF, *et al.*: Analysis of human immunodeficiency virus type 1 nef gene sequences among inmates from prisons in China. *AIDS Res Hum Retroviruses* 2009;25(5):525–529.
16. Chen S, Cai W, He J, *et al.*: Molecular epidemiology of human immunodeficiency virus type 1 in Guangdong province of southern China. *PLoS One* 2012;7(11):e48747.

Address correspondence to:

*Fu-Xiang Wang*  
*Department of Infectious Diseases*  
*Fourth Affiliated Hospital of Harbin Medical University*  
*37 Yiyuan Street*  
*Harbin 150001*  
*China*

*E-mail: wangfuxiang999@hotmail.com*

*Bin-You Wang*  
*Department of Epidemiology*  
*Public Health College of Harbin Medical University*  
*157 Baojian Road*  
*Harbin 150081*  
*China*

*E-mail: diligent0689@hotmail.com*