Near Full-Length Genome Sequence of a Novel HIV-1 Recombinant Form (CRF01_AE/B) Detected Among Men Who Have Sex with Men in Jilin Province, China

Xingguang Li,^{1,2} Yi Feng,¹ Yao Yang,¹ Yanli Chen,¹ Qi Guo,³ Liuyan Sun,³
Xihui Zang,³ Hui Xing,¹ and Yiming Shao¹

Abstract

We report here a novel HIV-1 recombinant form (CRF01_AE/B) detected from a comprehensive HIV-1 molecular epidemiologic study among men who have sex with men (MSM) in Jilin province of northeastern China. The near full-length genome (NFLG) analyses showed that the novel HIV-1 recombinant isolate (JL.RF07) was composed of CRF01_AE cluster 5 (northeastern China origin) and subtype B (U.S. and European origin), with six recombinant breakpoints observed in the *pol*, *vif*, *tat*, *rev*, and *env* gene regions. To the best of our knowledge, this is the first detection of a novel HIV-1 recombinant form (CRF01_AE/B) in Jilin, which may indicate an active transmission network of HIV-1 infection among MSM in the region. Further studies of the molecular epidemiology of the HIV-1 epidemic among MSM in northeastern China are necessary to gain a fuller understanding of the transmission network and potential public health impact of HIV-1 among MSM in this region.

HUMAN IMMUNODEFICIENCY VIRUS type 1 (HIV-1) exhibits an extensive genetic diversity that is primarily driven by high rates of mutation, recombination, and replication. Through phylogenetic analyses of viral genomes, all globally circulating HIV-1 genotypes are currently categorized into four groups: M, O, N, and P. HIV-1 group M strains play a major role in the global HIV-1 epidemic and are further classified into 11 subtypes and subsubtypes (A1, A2, B, C, D, F1, F2, G, H, J, and K), a number of circulating recombinant forms (CRFs), and innumerable unique recombinant forms (URFs). To date, 61 CRFs have been reported at the Los Alamos National Laboratory HIV Sequence Database website (www.hiv.lanl.gov/conent/sequence/HIV/mainpage.html). Cocirculation and dual infection with CRF01_AE and subtype B in high-risk populations in Asia provide opportunities to generate various types of new CRFs. Up to now, 10 CRFs composed of subtype B and CRF01_AE have been reported in Asia: CRF15_01B and CRF34_01B identified in Thailand^{1,2}; CRF33_01B, CRF48_01B, CRF53_01B, and CRF54_01B in Malaysia³⁻⁶; CRF51_01B in Singapore⁷; CRF52_01B in Thailand and Malaysia⁸; and CRF55_01B and CRF 59_01B in China.⁹

Among them, CRF51_01B, CRF55_01B, and CRF59_01B were the three CRFs, known to date, reported among men having sex with men (MSM) in Asia.

The emergence of novel HIV-1 CRFs and URFs is an indication of multiple genotypes (subtypes/subsubtypes, CRFs and URFs) infecting one population of people. This unique recombinant virus was detected in a male who self-reported his risk as homosexual contact. HIV infections in China are rising faster in MSM than they are in the general population.¹⁰ In this study, we detected a novel HIV-1 recombinant form (CRF01_AE/B), designated as JL.RF07, by near full-length genome (NFLG) analyses, isolated from an HIV-positive male subject infected through homosexual transmission in Jilin province of northeastern China.

In this study, plasma was collected from an HIV-positive consenting MSM in China (Patient JL.RF07). JL.RF07, residing in Jilin city of Jilin province, was recruited during our nationwide cross-sectional HIV-1 molecular epidemiology survey among an MSM cohort and HIV screening in 2010, from the Jilin Center for Disease Control and Prevention (CDC). The subject was a 48-year-old male who was

¹State Key Laboratory for 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, China.

²HIV Molecular Epidemiology and Virology Research Group, The State Key Laboratory of Virology, Wuhan Institute of Virology, University of Chinese Academy of Sciences, Wuhan, China.

³Jilin Provincial Center for Disease Control and Prevention, Jilin, China.

FIG. 1. Phylogenetic analyses of the near full-length genome (NFLG) sequence of the JL.RF07 isolate from Jilin province of northeastern China. All HIV-1 group M reference sequences of all known subtypes/subsubtypes and the CRFs relevant to our study (CRF01_AE, CRF15_01B, CRF33_01B, CRF34_01B, CRF48_01B, CRF51 through CRF55, and CRF59_01B), as well as all NFLG or full-length genome sequences of URFs available from the Los Alamos National Laboratory HIV Sequence Database (www.hiv.lanl.gov/content/sequence.html), were initially used to construct the neighbor-joining phylogenetic tree; some references were later removed for clarity. A black solid circle (●) marks the JL.RF07 isolate throughout the article. The stability of the nodes was assessed by bootstrap analysis with 1,000 replications, and only bootstrap values $\geq 70\%$ are shown at the corresponding nodes of the tree. The scale bar represents 2% nucleotide sequence divergence.

diagnosed as HIV-1 positive on October 16, 2008 and the CD4⁺ T cell number was 327cells/ μ l in September 3, 2010 after plasma was collected. The study was approved by the institutional review boards of the National Center for AIDS/STD Control and Prevention. Viral RNA was extracted from plasma using the QIAamp Viral RNA Mini kit (Qiagen, Valencia, CA) and reverse transcribed by SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) following the manufacturer's instructions.

The NFLG sequence (related to the HXB2 nucleotide numbering system, position 552-9636) was amplified using primers B'-FL1.5 (5'-CCTTGAGTGCTCAAGTAGTGTG TGCCCGTCTGT-3', HXB2 nt 538 to 571), B'-FL1.3 (5'-ACT ACTTGAAGCACTCAAGGCAAGCTTTATTG-3', HXB2 nt 9642 to 9611), B'-FL2.5 (5'-AGTAGTGTGTGCCCGTCTGT TGTGTGACTC-3', HXB2 nt 552 to 581), and B'-FL2.3 (5'-TTAAGCACTCAAGGCAAGCTTTATTGAGGCTTA-3', HXB2 nt 9636 to 9604) for the first and second round nested polymerase chain reaction (PCR) amplification, respectively, as previously described.¹¹ The positive PCR products were purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and sequenced directly by an ABI 3730XL automated sequencer using BigDye terminators (Applied Biosystems, Foster City, CA). All Sequenced data were cleaned and assembled using Sequencher v.5.1 (Gene Codes Corporation, Ann Arbor, MI).

The NFLG sequences were aligned with the HIV-1 reference subtypes, subsubtypes, and CRFs obtained from the Los Alamos HIV Sequence Database (www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html) as well as all NFLG or full-length genome sequences of URFs available from the Los Alamos HIV Sequence Database (www.hiv.lanl.gov/content/sequence.html). The codon-aligned nucleotide sequence alignment was constructed using the Gene Cutter tool, from the Los Alamos HIV Sequence Database (www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html) and subsequently adjusted manually using BioEdit v.7.0.9.¹² Phylogenetic and subregion tree analyses were performed using the neighbor-joining method based on the Kimura two-parameter model implemented in MEGA v.5.05.¹³ The reliability of trees was evaluated by 1,000 bootstrap replicates. Bootstrap values above 70% were considered to be defined as a phylogenetic cluster. Recombination analyses were performed using RIP, jpHMM, and SimPlot v.3.5.1.¹⁴⁻¹⁶ The breakpoints

FIG. 2. Recombinant analyses of the JL.RF07 isolate. Genomic regions are indicated at the top of the plot with recombinant breakpoints (HXB2 numbering). At the bottom, the posterior probability of the JL.RF07 isolate is calculated by jpHMM software using all default settings.

FIG. 3. Subgenomic phylogenetic analyses of the novel HIV-1 recombinant isolate JL.RF07. Using methods described in Fig.1, the analyses were performed to confirm the genotype within each region, based on recombinant breakpoints shown in Fig. 2. The seven unique CRF01_AE lineages detected in China were labeled 01-1 through 01-7 and the Central African CRF01_AE sequences were labeled 01-CF. Genetic distance corresponding to the lengths of the branches of each phylogenetic tree is shown by the bottom line.

confirmation and the origin of each region were analyzed by subregion tree analyses.

The NFLG sequence from JL.RF07 in the present study was 8,996 bp (relative to the HXB2 nucleotide numbering system: positions 609 to 9600) in size, spanning the noncoding region (NCR), the *gag*, *pol*, *env*, *tat*, *rev*, *vif*, *vpr*, *vpu*, and *nef* gene regions, and a 3' part of the 5' long terminal repeat (LTR) and a 5' part of the 3' LTR. NFLG phylogenetic analyses showed that JL.RF07 formed a distinct monophyletic branch distantly related to all known HIV-1 subtypes/CRFs (Fig. 1).

Both RIP and jpHMM analyses showed that JL.RF07 was composed of subtype B and CRF01_AE, with three regions of CRF01_AE inserted into the subtype B backbone. jpHMM analyses identified six unique recombination breakpoints between CRF01_AE and subtype B at the nucleotide positions (relative to HXB2 genome) 2,873, 3,385, 4,745, 5,502, 5,917, and 88,486 nt, located in the *pol* (three breakpoints), *vif* (one breakpoint), 5' *tat* (one breakpoint), and 3' *rev* (or *env*) (one breakpoint) gene regions; similar results were obtained using SimPlot software (Fig. 2). Subregion phylogenetic analyses further confirmed the parental origin of each region of the recombinant genome as follows: region I (HXB2: 790–2872)=B; region II (HXB2: 2873–3384)=CRF01_AE; region III (HXB2: 3385–4744)=B; region IV (HXB2: 4745–5501)=CRF01_AE; region V (HXB2: 5502–5916)=B; region VI (HXB2: 5917–8847)=CRF01_AE; and region VII (HXB2: 8848–9407)=B. The recombinant structure is distinct from any known HIV-1 subtypes/CRFs reported to date. Subregion phylogenetic analyses also demonstrated that the parental origin of the CRF01_AE regions (II, IV, and VI) was the northern China CRF01_AE cluster 5, which is mainly circulating among MSM in northern China (especially in Beijing, Liaoning, and Jilin).¹⁷

Our previous study confirmed that the parental origin of the CRF01_AE regions (I, III, and V) of a novel HIV-1 second-generation recombinant form (CRF01_AE/CRF07_BC) was also from northern China CRF01_AE cluster 5, which indicated that recombination events may be more likely to occur between northern China CRF01_AE cluster 5 and other HIV-1 genotypes among MSM in northern China.¹⁸ On the other hand, the parental origin of the subtype B regions (I, III, V, and VII) was from northern China subtype B lineage, which is also mainly found among MSM in northeastern China. This is consistent with previous observations in MSM in China, where subtype B (U.S.-European origin) was the previously dominant strain in major cities of China (especially in Beijing, Shijiazhang, Liaoning, and Jilin) (Fig. 3).^{19–21} Furthermore, the ancestral origin of the subtype B regions (I, III, V, and VII) was of U.S.-European subtype B lineage, not the subtype B' (Thailand variant of subtype B is also referred as to Thai B) lineage associated with blood-borne epidemics in Asia.²²

The potential reason for this would be that as the CRF01_AE lineage was increasing in predominance and intermingled with the subtype B (U.S.-European origin) lineage within MSM in northern China, it led to an opportunity for generating this novel recombinant form (CRF01_AE/B) of HIV-1.

In the present study, to our best knowledge, we first identified the NFLG sequence of a novel HIV-1 second-generation recombinant form (CRF01_AE/B) composed of CRF01_AE (northern China origin) and subtype B (U.S.-European origin) among MSM in Jilin province of northeastern China. The other examples of novel recombinant

forms of HIV-1 identified from our previous study in this region were CRF61_BC and B'/C among heterosexuals and CRF01_AE/CRF07_BC among MSM.^{18,23,24} The emergence of multiple novel recombinant forms of HIV-1 in Jilin province of northeastern China suggests the presence of an active transmission networks of HIV infections among MSM in this region.

Sequence Data

The NFLG sequence of isolate JL.RF07 has been deposited in GenBank under accession number KF859773.

Acknowledgments

This work was supported by the National Science and Technology Major Project for Infectious Diseases Control and Prevention (2008ZX10001-004, 2012ZX10001-002, and 2012ZX10001-008), the National Natural Science Foundation of China (81261120379), SKLID Development Grant (2008SKLID203, 2012SKLID103), and International Cooperative Grant (2009DFB30420).

Author Disclosure Statement

No competing financial interests exist.

References

1. Tovanabutra S, Watanaveeradej V, Viputtikul K, *et al.*: A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses* 2003; 19(7):561–567.
2. Tovanabutra S, Kijak GH, Beyrer C, *et al.*: Identification of CRF34_01B, a second circulating recombinant form unrelated to and more complex than CRF15_01B, among injecting drug users in northern Thailand. *AIDS Res Hum Retroviruses* 2007;23(6):829–833.
3. Tee KK, Li XJ, Nohtomi K, Ng KP, Kamarulzaman A, and Takebe Y: Identification of a novel circulating recombinant form (CRF33_01B) disseminating widely among various risk populations in Kuala Lumpur, Malaysia. *J Acquir Immune Defic Syndr* 2006;43(5):523–529.
4. Li Y, Tee KK, Liao H, *et al.*: Identification of a novel second-generation circulating recombinant form (CRF48_01B) in Malaysia: A descendant of the previously identified CRF33_01B. *J Acquir Immune Defic Syndr* 2010;54(2): 129–136.
5. Chow WZ, Al-Darraj H, Lee YM, Takebe Y, Kamarulzaman A, and Tee KK: Genome sequences of a novel HIV-1 CRF53_01B identified in Malaysia. *J Virol* 2012; 86(20):11398–11399.
6. Ng KT, Ong LY, Takebe Y, Kamarulzaman A, and Tee KK: Genome sequence of a novel HIV-1 circulating recombinant form 54_01B from Malaysia. *J Virol* 2012; 86(20):11405–11406.
7. Ng OT, Eyzaguirre LM, Carr JK, *et al.*: Identification of new CRF51_01B in Singapore using full genome analysis of three HIV type 1 isolates. *AIDS Res Hum Retroviruses* 2012;28(5):527–530.
8. Liu Y, Li L, Bao Z, *et al.*: Identification of a novel HIV type 1 circulating recombinant form (CRF52_01B) in Southeast Asia. *AIDS Res Hum Retroviruses* 2012;28(10): 1357–1361.

9. Han X, An M, Zhang W, *et al.*: Genome sequences of a novel HIV-1 circulating recombinant form, CRF55_01B, identified in China. *Genome Announc* 2013;1(1).
10. Hei F, Wang L, Qin Q, *et al.*: Epidemic characteristics of HIV/AIDS among men who have sex with men from 2006 to 2010 in China. *Zhonghua Liu Xing Bing Xue Za Zhi* 2012;33(1):67–70.
11. Li Z, He X, Wang Z, *et al.*: Tracing the origin and history of HIV-1 subtype B' epidemic by near full-length genome analyses. *AIDS* 2012;26(7):877–884.
12. Hall TA: BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999;41:95–98.
13. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S: MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;28(10):2731–2739.
14. Siepel AC, Halpern AL, Macken C, and Korber BT: A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res Hum Retroviruses* 1995;11(11):1413–1416.
15. Lole KS, Bollinger RC, Paranjape RS, *et al.*: Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999;73(1):152–160.
16. Schultz AK, Zhang M, Bulla I, *et al.*: jpHMM: Improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res* 2009;37(Web Server issue):W647–651.
17. Feng Y, He X, Hsi JH, *et al.*: The rapidly-expanding CRF01_AE epidemic in China is driven by multiple lineages HIV-1 viruses introduced in the 1990s. *AIDS* 2013;27(11):1793–1802.
18. Li X, Ning C, He X, *et al.*: Near full-length genome sequence of a novel HIV type 1 second-generation recombinant form (CRF01_AE/CRF07_BC) identified among men who have sex with men in Jilin, China. *AIDS Res Hum Retroviruses* 2013;29(12):1604–1608.
19. Wang W, Xu J, Jiang S, *et al.*: The dynamic face of HIV-1 subtypes among men who have sex with men in Beijing, China. *Curr HIV Res* 2011;9(2):136–139.
20. 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.
21. 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.
22. Li Y, Uenishi R, Hase S, *et al.*: Explosive HIV-1 subtype B' epidemics in Asia driven by geographic and risk group founder events. *Virology* 2010;402(2):223–227.
23. Li X, Ning C, He X, *et al.*: Genome sequences of a novel HIV-1 circulating recombinant form (CRF61_BC) identified among heterosexuals in China. *Genome Announc* 2013;1(3).
24. Ning C, Li X, He X, *et al.*: Near full-length genome identification of a novel hiv type 1 B'/C recombinant isolate JL100091 in Jilin, China. *AIDS Res Hum Retroviruses* 2013;29(12):1609–1612.

Address correspondence to:

Yiming Shao

Division of Research on Virology and Immunology

National Center for AIDS/STD

Control and Prevention

China CDC

Beijing 102206

China

E-mail: yshao@bnn.cn