Near Full-Length Genome Identification of a Novel HIV-1 Recombinant Form (CRF01_AE/B'/C) Among Heterosexuals in Jilin, China

Xingguang Li,^{1,2} Chuanyi Ning,^{1,3} Yanli Chen,¹ Yi Feng,¹ Min Wei,¹ Hui Xing,¹ and Yiming Shao¹

Abstract

Recombinant forms contribute significantly to the genetic diversity of HIV-1. Here we report a novel HIV-1 recombinant form (CRF01_AE/B'/C) detected from a comprehensive HIV-1 molecular epidemiologic study among heterosexuals in Jilin province of northeastern China. Recombinant analyses of the near full-length genome (NFLG) of the novel HIV-1 recombinant isolate (JL.RF01) showed that the backbone of the genome was CRF01_AE, and three insertions of subtype B' (242, 370, and 233 bp) and C (1142, 230, and 271 bp), respectively, were inserted along the genome. Phylogenetic analyses revealed that the novel HIV-1 recombinant form (CRF01_AE/B'/C) more likely originated from Thailand subtype B' and CRF01_AE and India subtype C. We report a unique mosaic structure that is distinct to HIV-1 CRF01_AE/B'/C recombinant viruses reported to date. The emergence of this novel recombinant form (CRF01_AE/B'/C) suggests the increasing significance of heterosexual transmission contributing to the complexity of the HIV-1 epidemic in northeastern China.

ONE OF THE SIGNIFICANT FEATURES of human immunodeficiency virus type 1 (HIV-1) is its extremely high level of genetic variation. Although mutation is the ultimate force of genetic variation, in addition to mutation, genetic recombination is considered the mechanism most important to generating novel HIV-1 variants, and thereby undoubtedly leading to its genetic diversity. To date, through phylogenetic analysis of the genomic sequences of HIV-1 from globally circulating viral strains, HIV-1 strains have been currently divided into four distinct groups: M (major), O (outlier), N (non-M, non-O or new), and the newly described P. Group M is further subclassified into nine subtypes (A, B, C, D, F, G, H, J, K) and two sets of intersubtypes (A1, A2, and F1, F2). In addition, at the sequence search interface tool of the Los Alamos National Laboratory HIV Sequence Database (www.hiv.lanl.gov/components/sequence/HIV/search/search.html), currently 61 circulating recombinant forms (CRFs) have been reported. Apart from those CRFs, innumerable unique recombinant forms (URFs) have also been reported, especially in regions where multiple subtypes/CRFs of HIV-1 among the same high-risk group are severely epidemic.

China, located in East Asia, is the country in the world whose population is increasing most. In China the majority of prevalent genotypes of HIV-1 are subtype B' (Thailand variant of subtype B, also referred as Thai B), CRF01_AE, CRF07_BC, and CRF08_BC, which constituted 92.8% of HIV-1 infections in China in 2006; these were detected in various high-risk populations including former plasma donors (FPDs), injection drug users (IDUs), and blood transfusion recipients (BT), as well as in mother-to-child transmission (MCT), those infected through men having sex with men (MSM), and heterosexual transmission.¹ The opportunity for cocirculation and dual infection with different subtypes/CRFs of HIV-1, especially in the same high-risk group, can result in the emergence of various novel intersubtypes/CRFs recombinants. To the best of our knowledge, to date, seven NFLG sequences of URFs originating from subtype B' (Thai B), C (India C), and CRF01_AE (Thai CRF01_AE) lineages have been reported in Myanmar and China among IDUs, MSM, and heterosexual populations.²⁻⁴

Jilin province is located in the central part of northeast China, and Changchun city is the capital of Jilin province.

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.

³School of Public Health and Guangxi Key Laboratory Cultivation Base of AIDS Prevention and Treatment, Guangxi Medical University, Guangxi, China.

FIG. 1. Phylogenetic analyses of the near full-length genome (NFLG) nucleotide sequence of the JL.RF01 isolate. The neighbor-joining phylogenetic tree was constructed based on the NFLG nucleotide sequences of HIV-1 isolates using the MEGA 5.05 software package. All the reference strains were retrieved from the Los Alamos National Laboratory HIV Sequence Database. The JL.RF01 isolates identified among heterosexuals from Jilin province of northeastern China are labeled with a black solid circle (●) throughout the article. Bootstrap analysis was performed with 1,000 replications and only bootstrap values $\geq 70\%$ are shown. The scale bar represents 2% genetic distance.

Since the first AIDS case was reported in Jilin province in northeastern China in 1993, the number of reported HIV infections has been increasing each year; a total of 1,477 HIV infections had been identified by the end of 2010, detected in various high-risk populations. In the present study, we identified a novel HIV-1 intersubtype recombinant form (CRF01_AE/B'/C) composed of subtype B', C, and CRF01_AE among the heterosexual population, which is distinct from the HIV-1 CRF01_AE/B'/C recombinants previously reported.

In the present study, a plasma sample was collected from an HIV-positive consenting heterosexual in China (Patient JL.RF01). JL.RF01 was a Chinese citizen residing in Liuhe county of Jilin province, and was recruited during our HIV-1 molecular epidemiology survey among a heterosexual cohort and HIV screening in 2010, from the Jilin Center for Disease Control and Prevention (CDC). The patient was a 38-year-old male and was diagnosed as HIV-1 positive in June 30, 2008. The CD4⁺ T cell numbers of the patient were 361 and 413 cells/ μ l in March 23, 2010 and November 10, 2010, respectively, before and after the plasma sample was collected (June 29, 2010). Written informed consent was obtained from patient JL.RF01 at the time of sample collection. The study was approved by the institutional review boards of the

National Center for AIDS/STD Control and Prevention. The methods used in the present study for amplifying, analyzing, and characterizing the near full-length genome (NFLG) sequence of JL.RF01 were the same as previously described.⁵

In the present study, we finally obtained a length of 9,137 bp (relative to the HXB2 nucleotide numbering system: positions 554 to 9627) of the NFLG sequence of JL.RF01. Using phylogenetic analysis, the NFLG sequence of JL.RF01 clustered with CRF01_AE, but formed a distinct monophyletic branch distantly related to CRF01_AE (Fig. 1). Similarity plot analysis of the NFLG sequence of JL.RF01 revealed that JL.RF01 was a unique mosaic structure composed of subtype B', C, and CRF01_AE, with each of three regions of subtype B' and C inserted into the CRF01_AE backbone (Fig. 2A). Bootscanning analysis of the NFLG sequence of JL.RF01 revealed the recombinant breakpoints of JL.RF01 at nucleotide positions 1221, 1463, 2605, 2975, 4130, 4360, 4593, and 4864 (Fig. 2B, numbered according to the nucleotide sequence of HXB2, GenBank accession number K03455).⁶

Subgenomic phylogenetic analyses confirmed the genotype of each region of JL.RF01: region I [HXB2 nucleotides (nt) 790 to 1220]=CRF01_AE; region II (HXB2 nt 1221 to 1462)=B'; region III (HXB2 nt 1463 to 2604)=C; region IV

FIG. 2. Recombinant analyses of the NFLG nucleotide sequence of the JL.RF01 isolate. Reference sequences for subtype A1 (92UG037), subsubtype A2 (97CDKTB48), subtype B' (RL42), subtype C (95IN21068), subtype D (ELI), subsubtype F1 (VI850), subsubtype F2 (CM53657), subtype G (HH8793_12_1), subtype H (VI991), subtype J (SE9280_7887), subtype K (96CM_MP535), and CRF01_AE (CM240) were used in the similarity plot analysis (A) and subtype B' (RL42), subtype C (95IN21068), subtype J (SE9280_7887), and CRF01_AE (CM240) were used in the bootscanning analysis (B). The recombinant map results of JL.RF01 with breakpoints based on HXB2 nucleotide sequence numbering are shown in (C).

(HXB2 nt 2605 to 2974)=B'; region V (HXB2 nt 2975 to 4129)=CRF01_AE; region VI (HXB2 nt 4130 to 4359)=C; region VII (HXB2 nt 4360 to 4592)=B'; region VIII (HXB2 nt 4593 to 4863)=C; and region IX (HXB2 nt 4864 to 9411)=CRF01_AE (Fig. 3). To the best of our knowledge, the recombinant structure is different from any known HIV-1 subtypes/CRFs and URFs reported to date. Subgenomic phylogenetic analyses also indicated that the subtype B' regions (II, IV, and VII) originated from Thailand, but formed a monophyletic branch distinct from them. A potential reason for this would be that the subtype B' progenitor of JL.RF01 had diverged significantly from the Thailand subtype B' lineage associated with the bloodborne epidemic in Asia.⁷ Another potential reason could be the presence of very short regions of nonsubtype B' (for instance, the subtype C and CRF01_AE regions) within the subtype B' regions as undetected recombinant regions. In addition, the ancestral origin of the subtype C regions (III, VI, and VIII) was more likely from an Indian subtype C lineage, whereas the CRF01_AE regions (I, V, and IX) originated from Thailand CRF01_AE but were not related to the seven distinct clusters of CRF01_AE variants circulating among various high-risk populations and geographic regions in China (Fig. 3).⁸

Therefore, taking the data together, it is tempting to hypothesize that the novel recombinant form (CRF01_AE/B'/C) of HIV-1 was more likely a cross-region (preferable Jilin–Yunnan), even cross-border (probably China–Myanmar) transmission event, and less likely a local HIV-1 recombinant form generated among heterosexuals in Jilin province of northeastern China. The recombinant map results revealed that the subtype B' (II, IV, and VII) and C (III, VI, and VIII) regions were located in the *gag* and *pol* gene regions, a unique mosaic structure distinct from CRF01_AE/B'/C recombinant isolates of HIV-1 previously reported (Fig. 2C).

In the present study, to the best of our knowledge, we first determined the NFLG sequence of a novel HIV-1 recombinant form (CRF01_AE/B'/C) composed of subtype B' (Thailand origin), subtype C (India origin), and CRF01_AE (Thailand origin) among heterosexuals in Jilin province of northeastern China. The novel recombinant form (CRF01_AE/B'/C) of HIV-1 reported in the present study may be a harbinger of more recombinant forms in this region. However, we cannot determine whether patient JL.RF01 studied here was dually or multiply exposed to each HIV-1 isolate through heterosexual behavior or was singly infected with the novel recombinant isolate itself by heterosexual behavior. Meanwhile, the public health importance of the novel recombinant form (CRF01_AE/B'/C) of HIV-1 identified among heterosexuals remains to be evaluated.

In summary, our previous and present studies have determined four novel recombinant forms (CRFs and URFs) of HIV-1 in Jilin province of northeastern China: CRF61_BC, B'/C and CRF01_AE/B'/C (identified in the present study) among heterosexuals and CRF01_AE/CRF07_BC among MSM.^{5,9,10} Of note, the emergence of these multiple novel recombinant forms of HIV-1 among various high-risk populations in Jilin province of northeastern China suggests that the concept of each genotype of HIV-1 associated almost exclusively with a specific high-risk population in a particular geographic location is being broken down, which can create great opportunities for the emergence of various novel recombinant isolates composed of subtype B', subtype C, and

CRF01_AE among heterosexuals in this region. This also suggests that the HIV-1 epidemic has spread into the general population from high-risk populations (for instance, IDUs, MSM, and FPDs) through heterosexual transmission networks in this region. Understanding the molecular epidemiologic properties of these newly emerging intersubtype or inter-CRF recombinants will provide pivotal information for designing effective prevention measures to control and limit the HIV/AIDS pandemic of these highly diverse recombinants of HIV-1 in China.

Sequence Data

The NFLG sequence of isolate JL.RF01 has been deposited in GenBank under accession number KF850149.

Acknowledgments

The authors would like to thank the staff from Jilin Provincial Center for Disease Control and Prevention (Jilin CDC) for their help with sample collection, as well as the patient who agreed to participate in the present study. 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), 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. 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.
2. Takebe Y, Motomura K, Tatsumi M, Lwin HH, Zaw M, and Kusagawa S: High prevalence of diverse forms of HIV-1 intersubtype recombinants in Central Myanmar: Geographical hot spot of extensive recombination. *AIDS* 2003;17(14):2077–2087.
3. Meng ZF, Zhang XY, Xin RL, *et al.*: A new approach for sequencing virion genome of Chinese HIV-1 strains subtype B and BC from plasma. *Chin Med J (Engl)* 2011; 124(2):304–308.
4. Han X, An M, Zhao B, *et al.*: High prevalence of HIV-1 intersubtype B'/C recombinants among injecting drug users in Dehong, China. *PLoS One* 2013;8(5):e65337.
5. 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.
6. 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.
7. 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.

8. Feng Y, He X, Hsi JH, *et al.*: The rapidly expanding CRF01_AE epidemic in China is driven by multiple lineages of HIV-1 viruses introduced in the 1990s. *AIDS* 2013;27(11):1793–1802.
9. 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).
10. 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