VIROLOGICA SINICA, June 2007, 22 (3):212-217
CLC number: R512.91 Document code: A Article ID: 1003-5125(2007)03-0212-06

Differences in Variation of Human Immunodeficiency Virus Type 1 Sequences from Henan and Shanghai Regions of China*

Chun-yu YIN[1, 2], Hong-zhou LU[1**], Wei-ming JIANG[1], Maria Pia De Pasquale[3], Yue-kai HU[1], Xiao-zhang PAN[1], Xin-hua WENG[1], Richard T. D'Aquila[3,5] and Yi-Wei TANG[3, 4 **]

(1.Department of Infectious Diseases,Fudan University Huashan Hospital, Shanghai 200040,China; 2. Department of Infectious Diseases, the Sixth People's Hospital, Hangzhou 310014, China; 3.Division of Infectious Diseases, Department of Medicine; Tennessee 37232, USA; 4. Departments of Pathology; Tennessee 37232, USA; 5. Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA)

Abstract: Illegally paid blood donation was a risk factor for HIV acquisition exclusively in Henan and Hubei Provinces of China, and not in Shanghai. Nucleotide sequences in the *gag* and *env* genes of HIV-1 were compared between isolates from Henan and Shanghai regions of China to test whether an expected higher degree of a common source of infections from this unique blood donation transmission risk would be evident as decreased variation among Henan isolates in an exploratory cross-sectional analysis. Among 38 isolates studied, 23 of 23 (100%) from Henan and 8 of 15 (54%) from Shanghai were subtype B. In addition, fewer sequence differences were found in gp41 of subtype B isolates from Henan than from Shanghai isolates. Further studies with additional controls are therefore warranted to confirm the role of the degree of a common source of infections in differences in HIV variation across populations.

Key words: HIV; Transmission; Phylogenetic analysis

Human immunodeficiency virus 1 (HIV-1) evolves by the accumulation of mutations and by recombination. Different subtypes co-circulating in the population of a geographical region provide the basis for the generation and distribution of inter-clade mosaic viruses. Each HIV-1 epidemic in distinct geographical regions and population groups has its own specific characteristics and dynamics. Phylogenetic analyses of nucleotide sequences and HIV genetic variation have become well-accepted as tools to analyze the HIV epidemic, while methologic research continues (4, 9). HIV transmission occurs around the globe due to unprotected sexual exposure, sharing of illegal drug injection equipment, and transfusion of infected blood products. However, illegally paid blood donors were at risk for HIV exclusively in Henan and Hubei

Received: 2006-11-03, Accepted: 2006-12-27

* Fundation items: The Vanderbilt-Meharry Center for AIDS Research (P30 AI 54999); R.T.D (R01 AI 29193); Start Fund of Ministry of Education of China (for Hong-zhou LU, 2004BA719A10).

** Corresponding author. Yi-wei TANG. Tel: +1-615-322-2035. Fax: +1-615343-8420. E-mail: yiwei.tang@vanderbilt.edu
Hong-zhou LU. Tel: +86-21-57248758. E-mail: Luhong zhou66@hotmail.com

Provinces of China for a defined period of time. Until recently, blood drawing equipment was re-used for multiple donors and pooled erythrocytes were transfused back to plasma donors only in those Provinces; this was the only route of infection identified (13, 16). We hypothesized that a higher degree of common source infections expected via this unique blood donation transmission route would be evident as decreased variation in a cross-sectional sampling of HIV sequences, relative to another nearby region lacking transmission via blood donation.

Among other genes, HIV-1 includes *gag*, and *env* genes, which code for the precursor proteins of the viral capsid and the virion envelope. The P24 protein is one of the main core protein produced by the *gag* gene. Sequences in the P24 region are highly conserved, and mutations of epitope sequences may result in immune escape (12). On the other hand, HIV-1 envelope proteins can differ in more than 30% of their amino acids (2). In this study, we determined nucleotide sequences in *gag* and *env* genes of HIV-1 isolates from both the Shanghai and Henan regions of China, in order to compare inter- and intra-subtype genetic diversity in the two regions.

MATERIALS AND METHODS

Patients and specimens

A total of 38 HIV-1 infected plasma specimens were included in the study. Among them, 23 were collected from the Henan Province and 15 from Shanghai. Diagnosis was made by an enzyme immunoassay screening followed by a western blot confirmation. CD4 cell counts and HIV-1 viral loads were measured as previously described (5, 14). Demographic, clinical and epidemiologic information of subjects from the Henan and Shanghai regions are contrasted in Table 1. There were no significant differences between subjects from the two regions except that most from Henan had illegally donated blood, while the major infection risks were blood product transfusions and unprotected sexual intercourse among Shanghai subjects ($P < 0.001$).

Table 1. Characteristics of HIV-1 infected subjects from the Henan and Shanghai regions of China

Characteristic	No. (%) of subjects	
	Henan N=23)	Shanghai (N=15)
Gender: male	12 (52)	12 (80)
Age (years) [a]	37±15	33±13
Presumed duration of infection (years) [a]	9±1	6±4
CD4 cells (counts/mm^3) [a]	204±111	233±102
HIV-1 viral load (log_{10}copies/mL) [a]	5.44±1.5	4.91±1.3
Antiretroviral therapy	3 (13)	3 (20)
Self-reported HIV-1 exposure risk [b]		
Unprotected sexual intercourse	3 (13)	10 (67)
Illegal blood donation	18 (78)	0 (0)
Blood product transfusion receipt	1 (4)	5 (33)
Mother to child	1 (4)	0(0)

[a] Mean±S.D. [b]A statistical difference between the two regions was indicated ($P < 0.001$).

Reverse transcription, PCR and sequencing

Total viral RNA was extracted from plasma specimens by using a QIAamp Viral RNA Mini Kit (Qiagen Inc., Valentia, Cali.) as previously published (14). Viral cDNA was synthesized by the use of Moloney-murine leukemia virus reverse-transcriptase (Invitrogen, Gaithersburg, MD) at 37℃ for 30 min according to the manufacturer's instructions. The primers used for cDNA synthesis were G01 for *gag* and E01 for *env* respectively. A nested-PCR was used to amplify the *gag* P24 region by using G01 and G10 for the first PCR round. Two fragments were amplified in the second PCR round by using primers G20, G55, G70 and G25, respectively. The 50 μL of reaction

mixture contained 25mmol/L $MgCl_2$, 100nmol- /L dNTPs, 1.25U *Pfu* polymerase, 25pmol/L each primer and 10μL cDNA templet (first round) or 2 μL amplicon (second round). Both first and second rounds of the PCR reaction were run using a three-step PCR procedure which included: i) an initial denaturation at 95℃ for 3 min; ii) 35 cycles of 95℃ for 30 s, 52℃ for 1 min and 72℃ for 3 min; and iii) a final extension at 72℃ for 10 min. The same PCR protocol was used for *env* gp41 IDR amplification by using E01 and E10 as the first round primer set and E170 and E35 as the second round primer set. Sequences of the above primers were published previously (11). Final PCR products included two fragments with 680 and 480 base pairs each for the *gag* gene and a 540 base-pair fragment for the *env* gene.

Amplicon sequencing and phylogenetic analysis

DNA fragments amplified by PCR were purified by using a QIAquick PCR Purification Kit (Qiagen Inc., Valentia, Cali.) according to the manufacturer's instructions. A reaction mixture including 200 ng of purified DNA fragment and 5 pmol/L of primer was used for cycle sequencing on an ABI PRISM 377 DNA sequencer (Applied Biosystems, Foster City, Cali). Two orientated sequences were assembled and analyzed through blasting HIV-1 databases (http://hiv-web.lanl.gov). Nucleotide sequences were aligned by the Clustal X program (version 1.81). Phylogenetic analysis was inferred using the Phylogeny Inference Package (PHYLIP, version 3.6a3) (http: //evolution. genetics. washington. edu/phylip/general.html) (1). Phylogenetic trees were constructed by the neighbor-joining method using DNA distance matrices, calculated with the Kimura two-parameter model. Robustness of the tree was evaluated by bootstrap analysis on 500 replicates. Phylogenetic trees were edited using a TreeView (version 1.6.5). The sequence of the reference HxB strain was used as a standard for determining the number and sites of polymorphisms (10). Genetic distances values were obtained by a Kimura two-parameter model. Representative sequence of different subtypes and of B subtype viruses from different countries was used as reference for inference of subtype differences and within-B subtype variation, respectively.

RESULTS AND DISCUSSION

Target gene fragments were amplified and sequenced from all specimens. Sequences of *gag* and gp41 IDR were used for subtype determination in comparison to reference isolates representative of the A, B, C, D, F, G, H clades, and the circulating recombinant forms CRF01 and CRF02 , respectively (Fig.1 A). Among the 38 isolates studied, 31 (81.6%) were subtype B, 1 (2.6%) was subtype A, 5 (13.2%) were CRF01_AE, and 1 (2.6%) was CRF02_AG. All 23 (100.0%) sequences from Henan and 8 of 15 (53.3%) from Shanghai were subtype B. Comparison of sequences in the *gag* P24 region indicated that all Henan and most Shanghai subtype B isolates clustered with a B'-Thai strain representing the most prevalent B-clade isolates in China (3). Only two subtype B isolates from Shanghai (SH06 and SH14) were close to subtype B sequences from the United States and Australia (7) (Fig.1.B). Two recombinant viruses (CRF01_AE and CRF02_AG) were recognized in Shanghai, indicating multiple HIV-1 viruses have been circulating in the region, which is consistent with previous findings (6, 8, 15).

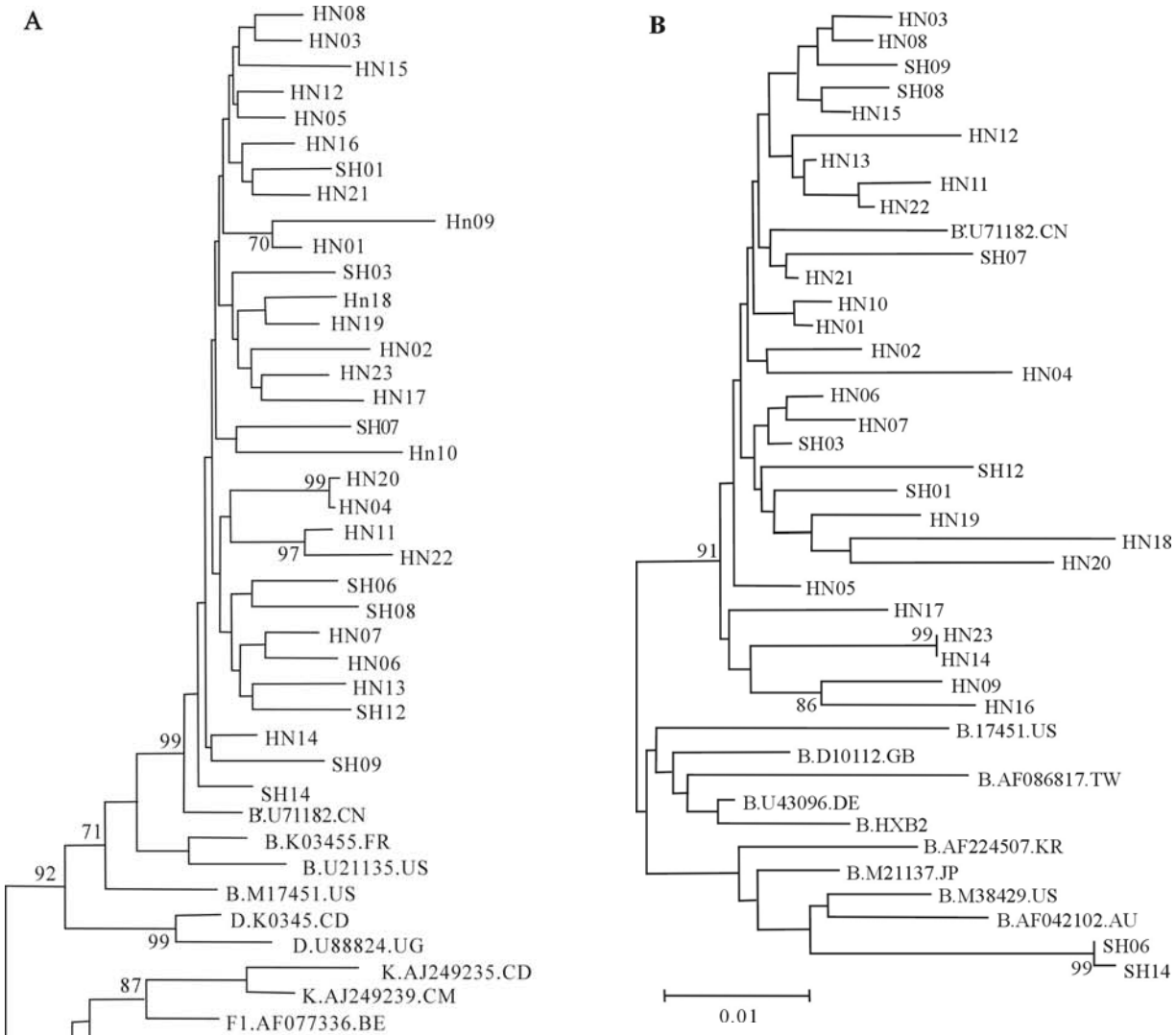

Fig. 1. Phylogenetic analysis of the HIV-1 *gag* P24 coding region of 38 isolates from Henan and Shanghai by the neighbor-joining method. Genetic distances were calculated by the Kimura two-parameter model. The scale indicates the relative phylogenetic distance. Bootstrap values were generated from 500 bootstrap replicates, and values greater than 70% are labeled. **A:** HIV-1 subtyping analysis. Reference isolates from clades A, B, C, D, F, G, H, J, K, and recombinant forms CRF01_AE, CRF02_AG, CRF07_BC and CRF08_BC were downloaded from the HIV-1 databases (http://hiv-web.lanl.gov). The bar represents 1% genetic distance. **B:** Nucleotide sequences in the *gag* P24 coding region of subtype B isolates. The two letters, US, CN, AU, JP, KR, TW, DE, and GB, following accession numbers indicated that these isolates were originally recovered from the U.S., China, Australia, Japan, Korea, Taiwan, Germany and England, respectively. The bar represents 1% genetic distance.

Among all 31 subtype B isolates, the mean intra-subtype sequence genetic distance in the *gag* p24 region was 3.6%, with a range from 0.4% to 6.7%. Inter-sequence genetic distance in the *gag* p24 gene in the 23 Henan sequences averaged 3.2%, in comparison to an average of 3.4% for the 8 isolates from Shanghai (P>0.05) (Table 2). In contrast, in the *env* gp41 IDR region, all subtype B isolates had an average intra-subtype distance of 4.4%, ranging from 2.0 to 7.6%. For the 23 isolates from Henan, the mean genetic distance in the *env* gp41 IDR was 3.7%, which was significantly less than the mean genetic distance in subtype B isolates from Shanghai (4.6%, P=0.038) (Table 2). When non-subtype B isolates were included in the analysis, the mean sequence variation in Shanghai reached 9.3%, which is also significantly more diversified than in the Henan sequences (Table 2). These data demonstrate that HIV-1 isolates from Henan were solely the B subtype and relatively more conserved in *env* gp41 IDR region sequences than the Shanghai isolates.

This comparison of sequence variation in the *env* gp41 region between HIV-1 isolates recovered from Henan and Shanghai indicates that isolates from Shanghai were more heterogenous than those from Henan. This was attributable to more subtypes, as well as more sequence variation within *env* subtype B sequences from Shanghai. There are several potential reasons for differences in diversity in HIV sequences across groups in such cross-sectional analyses. One cause not controlled for in this report involves differences in the time since the individuals studied were infected, given the ongoing sequence change within infected persons. Other mechanisms for differences in genetic variation may include differrences in the relative replication rate/substitution rate of viruses, and the relative degree of superinfection and recombination. The difference noted here in populations with distinct transmission routes suggests that the difference in number of individuals infected independently with the same genetic variant needs to also be considered.

Subsequent studies are needed to confirm this hypothesis that sequence diversity differs in Henan versus Shanghai isolates because of the unique way in which there were common infection sources for many individuals in Henan who each received the same infecting inoculum when donating blood. Such confirmatory studies should compare viruses from subjects known to be infected for similar durations and also evaluate whether the rate of genetic change over time in individuals was similar in the different regions. If the hypothesis suggested by the data in this

Table 2. Nucleotide sequence variation between *gag* and *env* sequences from the Henan and Shanghai regions of China

Geographic region	Isolates (number detected)	Genetic distance (mean±S.D.%) [a]	
		gag p24 region	*env* IDR
Henan	All isolates (23, all subtype B)	3.2±0.2	3.7±0.2
Shanghai	All isolates (15, of which 8 were subtype B)	7.5±0.4b	9.3±0.5 c
	Subtype B isolates only (8)	3.4±0.2 b	4.6±0.3 c

[a]Genetic distances (%) were calculated by dividing the sum of the percent variation for each pair-wise comparison by the number analyzed. [b]Statistical analysis of *gag* gene sequences indicates a significant difference between all isolates from Henan and all isolates from Shanghai (P=0.007) and no significant difference in comparison between all subtype B viruses from each region (P>0.05). [c]Statistical analysis of *env* IDR gene sequences indicates a significant difference between all isolates from Henan and all isolates from Shanghai (P=0.004) and also a significant difference in comparison between all subtype B viruses from each region (P=0.038). Genetic distances between groups were statistically analyzed using the Wilcoxon Rank sum test.

report is confirmed, the degree to which there are common sources of infection in a population would be established as another explanation for differences in genetic diversity of HIV among different regions or population groups.

Nucleotide sequence accession numbers. The partial *gag* and *env* gene sequences included in the study have been deposited in the GenBank sequence database under the accession numbers AY934644-AY934680.

Acknowlegements

The authors thank Qingxia Zhao, Yun He, Guang-ming Li, and Haijing Li for their excellent assistance. We also acknowledge P30 AI 54999 (The Vanderbilt-Meharry Center for AIDS Research), R01 AI 29193 (R.T.D.), and 2004BA719A10 (A Ministry of Education of China Start Fund for Experts Who Returned from Abroad, L.H.Z.).

References

1. **Felsenstein J.** 1989. PHYLIP-Phylogeny Inference Package (Version 3.2). **Cladistics,** 5: 164-166.
2. **Gaschen B, Taylor J, Yusim K, *et al.*** 2002. Diversity considerations in HIV-1 vaccine selection. **Science,** 296 (5577): 2354-2360.
3. **Graf M, Shao Y, Zhao Q, *et al.*** 1998. Cloning and characterization of a virtually full-length HIV type 1 genome from a subtype B'-Thai strain representing the most prevalent B-clade isolate in China. **AIDS Res Hum Retroviruses,** 14 (3): 285-288.
4. **Hue S, Pillay D, Clewley J P, *et al.*** 2005. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. **Proc Natl Acad Sci U S A,** 102 (12): 4425-4429.
5. **Jiang W, Kang L, Lu H Z, *et al.*** 2004. Normal values for CD4 and CD8 lymphocyte subsets in healthy Chinese adults from Shanghai. **Clin Diagn Lab Immunol,** 11 (4): 811-813.
6. **Laeyendecker O, Zhang G W, Quinn T C, *et al.*** 2005. Molecular epidemiology of HIV-1 subtypes in southern China. **J Acquir Immune Defic Syndr,** 38 (3): 356-362.
7. **O'Brien W A, Koyanagi Y, Namazie A, *et al.*** 1990. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. **Nature,** 348 (6296): 69-73.
8. **Piyasirisilp S, McCutchan F E, Carr J K, *et al.*** 2000. A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant. **J Virol,** 74 (23): 11286-11295.
9. **Rambaut A, Posada D, Crandall K A, *et al.*** 2004. The causes and consequences of HIV evolution. **Nat Rev Genet**, 5 (1): 52-61.
10. **Ratner L**, **Haseltine W, Patarca R, *et al.*** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. **Nature,** 313 (6000): 277-284.
11. **Sanders-Buell E, Salminen M O, McCutchan F E.** 1995. Sequencing primers for HIV-1. In: **Human retroviruses and AIDS**. Myers G. ed. Los Alamos, New Mexico: Los Alamos National Laboratory, p15-21
12. **Seligmann M, Pinching A J, Rosen F S, *et al.*** 1987. Immunology of human immunodeficiency virus infection and the acquired immunodeficiency syndrome. An update. **Ann Intern Med,** 107 (2): 234-242.
13. **Su B, Liu L, Wang F, *et al.*** 2003. HIV-1 subtype B' dictates the AIDS epidemic among paid blood donors in the Henan and Hubei provinces of China. **AIDS,** 17 (17): 2515-2520.
14. **Tang Y W, Huong J T, Lloyd R M, *et al.*** 2000. Comparison of human immunodeficiency virus type 1 RNA sequence heterogeneity in cerebrospinal fluid and plasma. **J Clin Microbiol,** 38 (12): 4637-4639.
15. **Yu X F, Liu W, Chen J, *et al.*** 2002. Maintaining low HIV type 1 env genetic diversity among injection drug users infected with a B/C recombinant and CRF01_AE HIV type 1 in southern China. **AIDS Res Hum Retroviruses,** 18 (2): 167-170.
16. **Zhang L, Chen Z, Cao Y, *et al.*** 2004. Molecular characterization of human immunodeficiency virus type 1 and hepatitis C virus in paid blood donors and injection drug users in China. **J Virol,** 78 (24): 13591-13599.