HHS Public Access

Author manuscript

J Acquir Immune Defic Syndr. Author manuscript; available in PMC 2016 August 01.

Published in final edited form as:

J Acquir Immune Defic Syndr. 2015 August 1; 69(4): 403–412. doi:10.1097/QAI.0000000000000649.

Comprehensive Characterization of the Transmitted/founder *env* Genes from a Single MSM Cohort in China

Yue Chen, PhD[1,2], **Ning Li, MD**[3], **Tong Zhang, MD, PhD**[3], **Xiaojie Huang, MD, PhD**[3], **Fangping Cai, MS**[1,2], **Nathan Vandergrift, PhD**[1,2], **Ruolei Xin, PhD**[4], **Zhefeng Meng, PhD**[5], **Xiaoyan Zhang, PhD**[5], **Chunlai Jiang, PhD**[6], **Xiaoning Xu, PhD**[7], **David C Montefiori, PhD**[8], **Feng Gao, MD**[1,2,6,*], and **Hao Wu, MD**[3,*]

[1]Duke Human Vaccine Institute, Duke University, Durham, NC 27710, USA

[2]Departments of Medicine, Duke University, Durham, NC 27710, USA

[3]Beijing You'an Hospital, Capital Medical University, Beijing 100069, China

[4]Beijing Center for Disease Control and Prevention, Beijing 100013, China

[5]Shanghai Public Health Clinical Center, Institutes of Biomedical Sciences, Fudan University, Shanghai 201508, China

[6]National Engineering Laboratory For AIDS Vaccine, College of Life Science, Jilin University, Changchun, 130012, Jilin, China

[7]Department of Immunology, Chelsea & Westminster Hospital, Imperial College London, SW10 9NH, UK

[8]Departments of Surgery, Duke University, Durham, NC 27710, USA

Abstract

Background—The men having sex with men (MSM) population has become one of major risk groups for HIV-1 infection in China. However, the epidemiological patterns, function of the *env* genes, and autologous and heterologous neutralization activity in the same MSM population have not been systematically characterized.

Methods—The *env* gene sequences were obtained by the single genome amplification (SGA). The time to the most recent common ancestor (tMRCA) was estimated for each genotype using the Bayesian MCMC approach. Coreceptor usage was determined in NP-2 cells. Neutralization was analyzed using Env pseudoviruses in TZM-bl cells.

Results—We have obtained 547 full-length *env* gene sequences by SGA from 30 acute/early HIV-1-infected individuals in the Beijing MSM cohort. Three genotypes (Subtype B, CRF01_AE, and CRF07_BC) were identified and 20% of the individuals were infected with multiple transmitted/founder (T/F) viruses. The tight clusters of the MSM sequences regardless of geographic origins indicated nearly exclusive transmission within the MSM population and

*Correspondence to: Feng Gao, MD, Phone: 1-919-668-6433; Fax: 1-919-681-8992; fgao@duke.edu, Hao Wu, MD, Phone: 86-10-6305-3963; Fax: 86-10-6329-4417; whdoc@sina.com.

Conflicts of Interest: The authors declare no conflicts of interest.

limited number of introductions. The tMRCA for each genotype was 10-15 years after each was first introduced in China. Disparate preferences for coreceptor usages among three genotypes might lead to the changes in percentage of different genotypes in the MSM population over time. The genotype-matched and -mismatched neutralization activity varied among the three genotypes.

Conclusions—Identification of unique characteristics for transmission, coreceptor usage, neutralization profile and epidemic patterns of HIV-1 is critical for the better understanding of transmission mechanisms, development of preventive strategies, and evaluation of vaccine efficacy in the MSM population in China.

Keywords

Genotype; genetic diversity; transmission; neutralization; coreceptor usage

Introduction

While heterosexual transmission is the primary mode for HIV-1 transmission in China, homosexual transmission accounts for a large portion of infection cases [1,2]. The proportion of reported cases from the men who have sex with men (MSM) population alarmingly increased over the years, from 0.4% before 2005 to 13.7% in 2011 [3]. Thus, it is essential to study the MSM population in order to understand epidemiological patterns of transmission, unique biological characteristics, pathogenesis, and potential vaccine strategies to better control its continuous spread.

Analysis of HIV-1 sequences from acute infection using the single genome amplification (SGA) method showed that approximately 20% of heterosexual infections are established by multiple transmitted/founder (T/F) viruses [4,5]. A recent study demonstrated that individuals infected with subtype B viruses in the MSM cohort in the United States had a significantly higher frequency (36%) of infections with multiple T/F viruses [6], suggesting that higher risk sexual activity could lead to infections by multiple T/F viruses, which might have a significant impact on disease progression [7-9]. Since viral genotypes, epidemiological factors and social behaviors differ greatly between the MSM populations in China and US, it is important to understand if MSM who were infected with different genotypes in China are infected with multiple T/F viruses at a higher percentage than those infected through heterosexual transmission.

Many viruses from MSM cohorts have been studied in China. However, most of the sequences were small partial sequences and only a few complete functional *env* sequences have been characterized. No T/F *env* sequences have been fully characterized from the MSM cohorts in China. Therefore, the biological characteristics of the *env* genes from MSM have not been well studied. Furthermore, whether cross-clade neutralization can be elicited among different genotypes in the same MSM cohort is not known. Thus, genetic and phenotypic analysis of multiple genotypes in the same MSM cohort has not been systematically performed. We recently have established an acute MSM infection cohort in Beijing, China [10]. Since the majority of infections were identified at early Fiebig stages in this cohort, we determined the number of T/F viruses in most of the subjects, the timing for introduction of each genotype into the MSM population, coreceptor preference of different

genotypes, and neutralization susceptibility to genotype-matched and -mismatched sera. The results obtained in this study may provide a foundation for a better understanding of HIV transmission in the MSM population and improve prevention programs in the future.

Materials and Methods

Plasma samples

Plasmas were obtained from 30 subjects with acute/early HIV-1 infection between 2007 and 2008 at You'an hospital in Beijing, China. The written consent was obtained from all patients who participated in the study, and the study was approved by the ethics committee of You'an Hospital and the Duke University Institutional Review Board. Fiebig stages were determined by measurement of viral loads by real-time PCR and HIV-1 specific antibodies by EIA and Western Blot as previously described[5,11]. Nine samples were at Fiebig stage I/II, six at Fiebig stage IV, 11 at Fiebig stage V, and four at Fiebig stage VI.

Analysis of the *env* gene sequences

Viral RNA was extracted from 200 μl of plasma using Pure Link viral RNA/DNA mini kit (Invitrogen, Carlsbad, CA). The reverse transcription and single genome amplification (SGA) of the 3′half genome were performed with the primers and conditions as previously described [12]. The sequence analysis, phylogenetic tree and Highlighter plot analyses were carried as reported in our previous study[13].

Molecular evolution clock analysis

The rates of evolutionary substitution and the divergence times for CRF01_AE, CRF07_BC, and subtype B were estimated using the Bayesian Markov Chain Monte Carlo (MCMC) approach available in the BEAST v1.5.4 package [14,15]. Each MCMC analysis was run for 20 or 30 million steps and sampled every 10,000 states. Posterior probabilities were calculated with a 10% burn-in and checked for convergence using Tracer v1.5. The maximum clade credibility tree was generated using Tree Annotator v1.5.4 available in BEAST and Fig Tree 1.3.1 was used for visualization of the annotated trees [16].

Determination of coreceptor usage

Coreceptor usage was determined in NP-2 cells expressing CD4 along with CCR5, CXCR4, or other coreceptors using Env pseudoviruses as previously described [13,17]. Virus replication was determined by measuring p24 at day 0 and day 3 using the Alliance HIV-1 p24 Antigen ELISA kit (PerkinElmer, Boston, MA).

Neutralization assays

Neutralization activity was measured in a luciferase reporter system in TZM-bl cells as described previously [18,19]. Plasma samples were heat-inactivated at 56°C for1 h and were then diluted at a 1:3 serial dilution starting at 1:20. The diluted plasma samples in duplicates were incubated with pseudoviruses for 1 h at 37°C and then used to infect TZM-bl cells. The 50% inhibitory dose (ID_{50}) was defined as the plasma dilution at which relative luminescence units (RLU) were reduced by 50% compared to RLU in virus control wells

after subtraction of background RLU in cell control wells. A response was considered positive for neutralization if the ID_{50} titer was >1:20 dilution.

Statistical analysis

Statistical analyses were performed using SAS v9.3 (SAS Institute Inc., Cary, NC). The coreceptor usage data were analyzed using a Wilcoxon (Mann-Whitney) rank sum test. The analysis of the autologous neutralization data in each genotype over time was performed using a linear mixed effects model. The heterologous neutralization data was analyzed in two by two blocks using Fisher's exact test.

Nucleotide Sequence Accession numbers

The Gen Bank accession numbers for all sequences generated in this study are KM217583-KM218333.

Results

The majority of individuals in the MSM cohort were infected with single T/F viruses

A total of 547 *env* SGA sequences were derived from 30 individuals (an average of 18 per person) in this MSM cohort (Table 1). Over 20 SGAs were obtained from 17 individuals. Fewer SGAs (2-18) were generated from other 13 individuals, due to the low viral loads and/or limited sample volumes for the screening samples. Phylogenetic tree analysis showed that all sequences from each individual clustered tightly together (Figure S1). Two T/F viruses were identified in three individuals (BJOX25, BJOX011 and BJOX07), while three T/F viruses were detected in BJOX28 (Figure 1A). The T/F viral sequences could be unambiguously inferred in 20 individuals from whom at least sixteen SGA sequences. Thus 20% (4/20) MSM individuals in this cohort were infected with more than one T/F virus. The other 10 individuals were not included for analysis because the numbers of T/F viruses in these individuals could not be unequivocally inferred due to higher levels of genetic variability (Fiebig stages V or later) or small numbers of SGAs (<15).

MSM sequences in China formed tight clusters within each genotype

To understand the genetic relationship between the newly characterized MSM sequences with those previously reported in China, a phylogenetic tree was constructed with one T/F or representative *env* gene sequence from each individual and all full-length *env* sequences from China available in the HIV-1 sequence database (http://www.hiv.lanl.gov). The phylogenetic tree analysis showed that 13 (43.3%) were CRF01_AE, 11 (36.7%) were subtype B, and 6 (20.0%) were CRF07_BC (Figure 1B and Table 1). Interestingly, all new MSM sequences were found in four closely related clusters: two in CRF01_AE, one in CRF07_BC and one in subtype B. Examination of the transmission routes for the other sequences clustered with the newly obtained sequences showed that nearly all of them were MSM derived (Figure 1B), strongly suggesting that the MSM sequences in China were highly related to each other. Analysis of additional shorter *env* sequences from China confirmed that the vast majority of the MSM sequences in China were tightly clustered together (Figure S2). Analysis of the geographic origins of those viruses showed that they were from provinces geographically far from each other; south provinces (Guangdong,

Yunnan and Hainan), north province (Liaoning), west province (Xinjiang), east province (Jiangsu) and central provinces (Hebei, Sichuan, Tianjin and Beijing) (Figure 1B and S2), similar to those recently reported by others [20-22]. These results suggested that the majority of MSM viruses, irrespective of large geographic distances, shared the same most recent common ancestors (MRCAs).

CRF01_AE MSM sequences formed three independent clusters, of which clusters I and II were equally predominant while cluster III was only represented by one MSM sequence (09LNA379) (Figure 1B). CRF07_BC MSM sequences from this and other studies formed one tight cluster. Although subtype B MSM sequences were more divergent than CRF01 and CRF07 sequences, they still tended to cluster together. Thai B′ sequence were much more predominant than US/Europe B sequences in China, but only one MSM *env* sequence was found to be B′. These results showed that only some subtypes or CRFs were introduced into the MSM population from the general HIV-1-infected population, irrespective of predominance of the genotypes in the general population. However, after the viruses were transmitted into the MSM population, they mainly spread within the MSM population nation wide.

We next sought to investigate how the newly characterized and previously reported MSM viral sequences related to viruses in the general population in Beijing. Since there were only very few non-MSM sequences from Beijing available for analysis, we obtained partial *env* sequences from 31 non-MSM subjects from Beijing and compared them with the MSM sequences from the same city. Phylogenetic tree analysis showed that three genotypes from the MSM population had different relationships with viruses from the general HIV-1-infected populations in Beijing. The subtype B MSM sequences intermingled with eight non-MSM viral sequences from Beijing in the phylogenetic tree (Figure S3A). However, all non-MSM subtype B sequences were within the large cluster formed by the MSM sequences, suggesting that they might be the result of transmission from the MSM population in Beijing. This is in good agreement with results from the full-length *env* sequence analysis (Figures 1B and 2). CRF01_AE sequences from non-MSM subjects formed three clusters. In two clusters, MSM and non-MSM sequences intermingled in the phylogenetic tree (Figure S3B), suggesting that they became indistinguishable from each other after introduction from outside. It was also possible that there were on-going transmissions between the two populations. Both MSM and non-MSM CRF07_BC viruses generally intermingled with those from other cities (Figure S3C). However, more than two-thirds of MSM CRF07_BC sequences formed a tight cluster with one non-MSM sequence and one sequence with unknown transmission route, suggesting that the viruses in this cluster resulted from a single introduction from the non-MSM population and had spread in the MSM population in Beijing. These results suggested that multiple transmissions occurred between MSM and non-MSM populations in Beijing but one introduction of CRF07_BC resulted in fast expansion in the Beijing MSM cohort. Since Beijing MSM CRF07_BC sequences clustered tightly with those from other cities (Figure 1B), it was likely that they would become one predominant genotype in the MSM population in the country. Although these results showed more frequent HIV-1 transmissions between MSM and non-MSM populations in Beijing, the majority of epidemic MSM sequences tended to

tightly cluster together, indicating that only a limited founder viruses was able to quickly spread and predominate the MSM population even in the city, possibly due to the exclusive relationship among partners within the MSM population. Some MSM CRF07_BC viruses were represented by single sequences. This indicated that they might represent independent introductions from the general population but might not become major epidemic strains. More sequences from MSM and non-MSM populations in other cities are required to confirm this epidemic pattern observed in Beijing. More importantly, the analysis of those sequences will reveal whether frequent transmissions between two populations can result in less distinction between viral sequences in MSM and non-MSM populations.

Estimation of the timing of introduction of different genotypes into the MSM cohort

The different clusters and various divergent levels of viruses indicated that HIV-1 was introduced into in the MSM cohort at different times. Since all complete *env* gene sequences from acute/early infection were obtained near the transmission time, we could accurately estimate the evolutionary rates and the most recent common ancestors (MRCA) of the different genotypes in the Beijing MSM cohort using the molecular clock approach implemented in BEAST v1.5.4 [16]. Both relaxed and strict molecular clocks showed similar evolutionary rates for subtype B, CRF01_AE and CRF07_BC *env* genes (Table S1). We then estimated the time to the most recent common ancestor (tMRCA) for each clusters (Table S1 and Figure 2), as previously described with the partial HIV-1 gene sequences [23-26]. Among all available Europe/US subtype B sequences, the MSM sequences were highly divergent from each other (11.78% [5.70%-15.67%]), while the threenon-MSM sequences were less divergent (9.79% [6.70%-11.65%]). The tMRCA of the subtype B viruses in China was estimated at 1971 (95HPD: 1956-1983). All sequences that closest to the root were MSM sequences and they are more divergent from each other (14.48% [13.00%-16.05%]) that all those sequences within the cluster I (9.64% [4.37%-12.66%]), suggesting that subtype B was first introduced into the non-MSM population. Soon after that, all but one Beijing MSM sequences formed a tight cluster with these three non-MSM sequences. The tMRCA for this subtype B cluster was estimated at 1985 (95% HPD: 1975-1993) (Figure 2A). These results indicated that the subtype B viruses were first introduced into the MSM population in 1971 in China and then they disseminated into the general population and into the MSM population in Beijing in 1985.

The CRF01_AE viruses were introduced into China at a later time point, with an estimated tMRCA at 1984 (95% HPD: 1977-1990). Both the CRF01_AE clusters were introduced into the MSM cohort roughly at the same time. The tMRCA for cluster I and II viruses were 1994 (95% HPD; 1989-1998) and 1996 (95%HPD: 1992-1999), respectively (Figure 2B). The CRF07_BC viruses were only recently identified in China [27]. The tMRCA of the CRF07_BC viruses in China was estimated at 1987 (95% HPD: 1980-1992). They were introduced into Beijing MSM cohort at 2002 (95% HPD: 2000-2004) (Figure 2C). These results demonstrated that different genotypes were introduced into the Beijing MSM cohort at different times, and all three genotypes were introduced about 10-15 years after the they were first introduced into China.

Disparate coreceptor usages among different genotypes

In addition to the primary CCR5 and CXCR4, other coreceptors(CCR3, APJ, GPR15 and FPRL-1) are also used by HIV-1[13,28,29]. CRF01_AE were mainly found in central Africa and Southeast Asia while CRF07_BE was primarily present in China [30,31]. However, little was known for their coreceptor usage. To investigate the coreceptor tropism of the three genotypes in the Beijing MSM cohort, we determined their coreceptor usages in NP-2 cell lines. All viruses infected CCR5+ cells but only BJOX35 (B) and BJOX37 (CRF07_BC) infected CXCR4+ cells (Figure 3). In addition, BJOX35 also used APJ and FPRL-1, while BJOX37 used FPRL-1. Both showed typical branching as other subtype B and CRF07_BC sequences in the phylogenetic tree (Figure 1B). Examination of amino acid sequences in V3 and other regions in both sequences did not reveal any unique signatures that differed from other subtype B and CRF07_BC sequences (data not shown). Subtype B viruses replicated significantly better in CCR5+ cells than CRF01_AE viruses ($p=0.011$). The subtype B viruses were most promiscuous for coreceptor usage as previously reported [29,32,33]; most of them (7/11) used CCR3 while some used APJ (3/11) and FPRL-1 (3/11). All six CRF07_BC viruses infected FPRL-1+ cells and replicated significantly better than the subtype B and CRF01_AE viruses ($p=0.001$ and $p<0.001$, respectively). None of CRF07_BC viruses used CCR3, APJ or GPR15. Interestingly, CRF01_AE viruses did not use any of those coreceptors other than CCR5, except one CRF01_AE virus (BJOX09) infected FPRL-1+ cells. Although GPR15 was often used by someHIV-1 strains [13,33,34], our results showed that no viruses in this cohort utilized it for infection, suggesting that different genotypes might affect the GPR15 tropism. Two and three T/F Env pseudoviruses generated from BJOX25 and BJOX28, respectively, did not show differences in coreceptor tropisms (data not shown). These results demonstrated that viruses in the Beijing MSM cohort could infect cells through CCR5, but they had different preferences for other coreceptors.

Autologous and heterologous neutralization among different genotypes

Since three different genotypes were identified in this MSM cohort, we first sought to investigate whether the development of autologous neutralization was different among them. The Env pseudoviruses were generated with the T/F or one representative *env* genes from 25 subjects for whom longitudinal plasma samples were collected (more three time points) and their neutralization susceptibility to autologous plasma (up to 90 weeks of infection) was determined. The day post infection was estimated for each individual according to the Fiebig stage [11,35]. Nineteen subjects developed autologous nAbs against autologous viruses (Figure 4A). No autologous nAbs were detected in five subjects (BJOX28, BJOX35, BJOX47, BJOX19 and BJOX27). The autologous Abs were detected as early as day 103 in BJOX25 and as late as day 533 in BJOX03. Multiple T/F viruses were identified in four subjects (Figure 1A). We determined the neutralization susceptibility of all identified T/F *env* genes in BJOX25 (2) and in BJOX28 (3) to longitudinal autologous plasma samples. The susceptibility of two BJOX25 T/F Env pseudoviruses to autologous neutralization was different, while no autologous neutralization to any of three T/F Envs was developed in BJOX28 (Figure S4A). Since there was only one sequence representing the second T/F virus in BJOX07 and BJOX11 (Figure 1A), T/F sequences could not be inferred and studied. Subjects infected with subtype B and CRF01_AE viruses developed relatively high titers of

autologous nAbs between days 200 and 400, while only 3 out of 5 subjects infected with CRF07_BC viruses developed autologous nAbs (Figure 4A). However, the differences in the autologous nAbs titers were not statistically significant among genotypes.

We next determined whether nAbs in subjects infected with one genotype had different breadths of neutralization activities against other genotype viruses during the early infection stage. The one plasma sample with highest nAb titers from each individual and sufficient amounts was used to determine its cross-genotype neutralization potency for genotype-matched and -mismatched pseudoviruses (Figure 4B). The CRF01_AE plasmas could neutralize 47% of heterologous CRF01_AE viruses, but neutralized significantly fewer subtype B viruses (11%; $p<0.001$) and CRF07_BC viruses (19%; $p<0.001$). CRF07_BC plasmas neutralized 33% of heterologous CRF07_BC viruses, only 15% of CRF01_AE viruses ($p=0.093$), and none of the subtype B viruses ($p<0.001$). Subtype B plasmas neutralized all three genotype viruses (subtype B, CRF01AE and CRF07_BC) at similar levels (15%-21%; $p>0.38$). Analysis of multiple T/F Env pseudoviruses form same individual (BJOX28 and BJOX25) showed similar cross neutralization susceptibility to heterologous plasma (Figure S4B). These results showed that the development of nAbs that could neutralize genotype-matched and -mismatched viruses varied among individuals infected with different genotype viruses.

Discussion

By genetically and phenotypically characterizing the *env* genes from 30 acute/early HIV-1-infectedindividuals in the Beijing MSM cohort, we have found that: single T/F viruses were the majority in the HIV-1-infected MSM individuals in this cohort; there was a 10-15 year delay for introduction of HIV-1 into the MSM population after their first introduction in China; limited genetic variation was observed in the MSM cohort across the county; the preference for usage of various secondary coreceptors varied among different genotypes; and disparate genotype-matched and mismatched neutralization activity developed among different genotypes. These results can have important implications for understanding HIV-1 transmission mechanisms, developing means to control further introductions of new genetic variants into the MSM cohorts, and evaluating vaccine efficacy in this unique population.

A recent nationwide survey showed that CRF07_BC, CRF01_AE, CRF08_BE and “Thai” B′ were most predominant (35.5%, 27.5%, 20.2% and 9.6%, respectively), “US/Europe” B was only present at 1%, and other genotypes were found in less than 1% of the population in China [30]. However, only three genotypes (subtype B, CRF01_AE and CRF07_BC) were found in the current MSM cohort or in other studies[20,36-38]. These results indicate only limited independent introductions of HIV-1 into the MSM population, irrespective of the predominance of the genotypes in the general HIV-1-infected population in China. Future studies are warranted to investigate why predominant subtype B′ and CRF08_BC were rarely detected while rare “US/Europe” B in the general population was one of the predominant genotypes in the MSM population. Results from such studies may offer clues about the transmission mechanisms and lead to development means to prevent further introductions of new genetic variants into the MSM population.

The molecular evolutionary clock analysis of the complete *env* gene showed that subtype B, CRF01_AE and CRF07_BC viruses were first introduced into China in 1971, 1984 and 1987, respectively, while they were introduced into the Beijing MSM population in 1985, 1994-1996, and 2002, respectively. These estimated tMRCAs were similar to those reported by others using different gene sequences [22-26,39]. Interestingly, all three genotypes were introduced into the MSM populations in Beijing 10-15 years after it was first introduced in China. It will be important to understand the factors that determined this delay in future studies.

Analysis of all currently available HIV-1 sequences in China from the Los Alamos HIV Sequence Database showed the vast majority of MSM sequences, irrespective of geographic areas, formed closely related clusters. Similar results were reported with small sequence datasets when other parts of HIV-1 genome sequences were analyzed [20-22]. This epidemic pattern was different from those in MSM populations in other countries, in which the MSM sequences intermingled with sequences from general HIV-1 infected individuals [40,41]. This unique epidemic pattern indicated that only limited introductions of HIV-1 strains from the general population became predominant strains in the MSM populations in China and the close network resulted in transmission chains mainly among MSM, allowing spreading of the same viruses in all geographic regions across the country.

The number of the T/F *env* sequences could be unambiguously inferred from 20 individuals. The sequence analysis showed that four of those individuals were infected with 2 or 3 T/F viruses. Thus, 20% of individuals were infected with multiple T/F viruses, which was similar to the percentages for individuals infected with HIV-1 through heterosexual transmission[4,5]. However, this rate was only about half of the rate(36%) that was previously reported for the MSM individuals infected with subtype B viruses in the United States [6]. This result suggested that the number of T/F viruses in MSM individuals could be affected by different genotypes, social behaviors, and races.

All newly characterized viruses in this MSM cohort could infect $CCR5^+$ cells, while only two viruses (BJOX35 and BJOX37) infected CXCR4 cells. However, the three genotypes had different preferences for other coreceptors. Most subtype B viruses could infect CCR3+ cells while CRF01_AE and CRF07_BC did not. CRF01_AE viruses only used CCR5, not any of the other tested coreceptors. CRF07_BC viruses infected FPRL-1+ cells significantly better than Subtype B and CRF01_AE viruses. The disparate preferences for coreceptors usage between subtype B and CRF01_AE might play a role in the continuous decrease of CRF01_AE [21,30,38,39,42] and the gradual decrease of subtype B in the MSM population in China[37,38].

The CRF01_AE plasmas neutralized a significantly higher percentage of heterologous CRF01_AE viruses than the subtype B and CRF07_BC viruses, while the subtype B plasmas similarly neutralized subtype B, CRF01_AE, and CRF07_BC viruses. Interestingly, none of the subtype B viruses could be neutralized by the CRF07_BC plasmas. These results demonstrated that the three genotypes elicited disparate cross-genotype neutralization activities, indicating that nAbs induced in individuals infected with one genotype viruses might not have optimal protection of super-infection of other subtypes or CRFs, especially

for those infected with CRF07_BC viruses. More recombinants among subtype B and CRF01_AE, CRF07_BC were recently identified [30,37,43,44], confirming super- and dual-infections indeed occurred in the MSM population.

The limited genetic variation of each genotype in the MSM population across China and high incidence rate (2.6-10.2 per 100 person-years) in the MSM cohorts [45-47] make the MSM population ideal for evaluating vaccine efficacy. Since each genotype elicited genotype specific nAb responses, it is also possible to investigate how protective the immune responses elicited by one genotype vaccine can be against viruses of different subtypes or CRFs in future vaccine trials in this unique population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of Funding: We thank the Duke Immunology Virology Quality Assessment Center (IVQAC) for Fiebig staging analysis of plasma samples. This work was supported by Bill and Melinda Gates Foundation and the Duke Centers for AIDS Research (AI064518), the Major Project of Beijing Municipal Science and Technology Committee (D141100000314005, D141100000314002), the Chinese Government 12th Five-Year Plan (2012ZX10001-003, 2012ZX10001-006 and 2012ZX10004904-002-002), and Beijing Key Laboratory (NO: BZ0089, 2014AZYJ01).

References

1. Lu L, Jia M, Ma Y, et al. The changing face of HIV in China. Nature. Oct 2; 2008 455(7213):609–611. [PubMed: 18833270]
2. Wu Z, Sullivan SG, Wang Y, Rotheram-Borus MJ, Detels R. Evolution of China's response to HIV/AIDS. Lancet. Feb 24; 2007 369(9562):679–690. [PubMed: 17321313]
3. Shang H, Xu J, Han X, Spero Li J, Arledge KC, Zhang L. HIV prevention: Bring safe sex to China. Nature. May 31; 2012 485(7400):576–577. [PubMed: 22660303]
4. Abrahams MR, Anderson JA, Giorgi EE, et al. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. J Virol. Apr; 2009 83(8):3556–3567. [PubMed: 19193811]
5. Keele BF, Giorgi EE, Salazar-Gonzalez JF, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A. May 27; 2008 105(21):7552–7557. [PubMed: 18490657]
6. Li H, Bar KJ, Wang S, et al. High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. PLoS Pathog. May.2010 6(5):e1000890. [PubMed: 20485520]
7. Sagar M, Lavreys L, Baeten JM, et al. Infection with multiple human immunodeficiency virus type 1 variants is associated with faster disease progression. J Virol. Dec; 2003 77(23):12921–12926. [PubMed: 14610215]
8. Gottlieb GS, Nickle DC, Jensen MA, et al. Dual HIV-1 infection associated with rapid disease progression. Lancet. Feb 21; 2004 363(9409):619–622. [PubMed: 14987889]
9. Tsai L, Tasovski I, Leda AR, Chin MP, Cheng-Mayer C. The number and genetic relatedness of transmitted/founder virus impact clinical outcome in vaginal R5 SHIVSF162P3N infection. Retrovirology. 2014; 11:22. [PubMed: 24612462]
10. Zhang HH, Sun Y, Liu YL, et al. Env gene sequencing and HIV-1 subtyping of infected MSM from Beijing. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi. Aug; 2009 23(4):275–277. [PubMed: 20108772]

11. Fiebig EW, Wright DJ, Rawal BD, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. Aids. Sep 5; 2003 17(13):1871–1879. [PubMed: 12960819]
12. Sanchez AM, DeMarco CT, Hora B, et al. Development of a contemporary globally diverse HIV viral panel by the EQAPOL program. J Immunol Methods. Jul.2014 409:117–130. [PubMed: 24447533]
13. Jiang C, Parrish NF, Wilen CB, et al. Primary Infection by a Human Immunodeficiency Virus with Atypical Coreceptor Tropism. J Virol. Aug 10; 2011 85(20):10669–10681. [PubMed: 21835785]
14. Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 1985; 22(2):160–174. [PubMed: 3934395]
15. Yang Z, Goldman N, Friday A. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Mol Biol Evol. Mar; 1994 11(2):316–324. [PubMed: 8170371]
16. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7:214. [PubMed: 17996036]
17. Shimizu N, Tanaka A, Oue A, et al. Broad usage spectrum of G protein-coupled receptors as coreceptors by primary isolates of HIV. Aids. Apr 27; 2009 23(7):761–769. [PubMed: 19307942]
18. Sarzotti-Kelsoe M, Bailer RT, Turk E, et al. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. J Immunol Methods. Jul.2014 409:131–146. [PubMed: 24291345]
19. Montefiori, DC. Measuring HIV neutralization in a luciferase reporter gene assay. In: Prasad, VR.; Kalpana, GV., editors. HIV Protocols: Second Edition, Methods in Molecular Virology. Humana Press; 2009. p. 395-405.
20. Li L, Sun G, Li T, et al. Multiple introductions of HIV into men who have sex with men were found in Zhengzhou City, China. AIDS Res Hum Retroviruses. Sep; 2012 28(9):1147–1151. [PubMed: 22236149]
21. Han X, An M, Zhang M, et al. Identification of 3 distinct HIV-1 founding strains responsible for expanding epidemic among men who have sex with men in 9 Chinese cities. J Acquir Immune Defic Syndr. Sep 1; 2013 64(1):16–24. [PubMed: 23542640]
22. 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. [PubMed: 23372706]
23. Chen JH, Wong KH, Chan KC, To SW, Chen Z, Yam WC. Phylodynamics of HIV-1 subtype B among the men-having-sex-with-men (MSM) population in Hong Kong. PLoS One. 2011; 6(9):e25286. [PubMed: 21966483]
24. Takebe Y, Liao H, Hase S, et al. Reconstructing the epidemic history of HIV-1 circulating recombinant forms CRF07_BC and CRF08_BC in East Asia: the relevance of genetic diversity and phylodynamics for vaccine strategies. Vaccine. May 26; 2010 28(Suppl 2):B39–44. [PubMed: 20510742]
25. Tee KK, Pybus OG, Li XJ, et al. Temporal and spatial dynamics of human immunodeficiency virus type 1 circulating recombinant forms 08_BC and 07_BC in Asia. J Virol. Sep; 2008 82(18): 9206–9215. [PubMed: 18596096]
26. 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. [PubMed: 23365653]
27. McClutchan FE, Carr JK, Murphy D, et al. Precise mapping of recombination breakpoints suggests a common parent of two BC recombinant HIV type 1 strains circulating in China. AIDS Res Hum Retroviruses. Oct 10; 2002 18(15):1135–1140. [PubMed: 12402948]
28. Cilliers T, Willey S, Sullivan WM, et al. Use of alternate coreceptors on primary cells by two HIV-1 isolates. Virology. Aug 15; 2005 339(1):136–144. [PubMed: 15992849]
29. Nedellec R, Coetzer M, Shimizu N, et al. Virus entry via the alternative coreceptors CCR3 and FPRL1 differs by human immunodeficiency virus type 1 subtype. J Virol. Sep; 2009 83(17):8353–8363. [PubMed: 19553323]

30. 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. [PubMed: 23056619]
31. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. Aids. Mar 13; 2011 25(5):679–689. [PubMed: 21297424]
32. Edinger AL, Hoffman TL, Sharron M, et al. An orphan seven-transmembrane domain receptor expressed widely in the brain functions as a coreceptor for human immunodeficiency virus type 1 and simian immunodeficiency virus. J Virol. Oct; 1998 72(10):7934–7940. [PubMed: 9733831]
33. Pohlmann S, Krumbiegel M, Kirchhoff F. Coreceptor usage of BOB/GPR15 and Bonzo/STRL33 by primary isolates of human immunodeficiency virus type 1. J Gen Virol. May; 1999 80(Pt 5): 1241–1251. [PubMed: 10355771]
34. Isaacman-Beck J, Hermann EA, Yi Y, et al. Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. J Virol. Aug; 2009 83(16):8208–8220. [PubMed: 19515785]
35. Lee HY, Giorgi EE, Keele BF, et al. Modeling sequence evolution in acute HIV-1 infection. J Theor Biol. Nov 21; 2009 261(2):341–360. [PubMed: 19660475]
36. Zhang X, Li S, Li X, et al. Characterization of HIV-1 subtypes and viral antiretroviral drug resistance in men who have sex with men in Beijing, China. Aids. Dec; 2007 21(Suppl 8):S59–65. [PubMed: 18172393]
37. 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. Oct; 2008 24(10):1245–1254. [PubMed: 18834324]
38. 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. Mar; 2011 9(2):136–139. [PubMed: 21361866]
39. 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. Nov; 2012 86(22):12402–12406. [PubMed: 22933290]
40. Dennis AM, Hue S, Hurt CB, et al. Phylogenetic insights into regional HIV transmission. Aids. Sep 10; 2012 26(14):1813–1822. [PubMed: 22739398]
41. de Oliveira T, Pillay D, Gifford RJ. The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. PLoS One. 2010; 5(2):e9311. [PubMed: 20174561]
42. Li L, Chen L, Liang S, et al. Subtype CRF01_AE dominate the sexually transmitted human immunodeficiency virus type 1 epidemic in Guangxi, China. J Med Virol. Mar; 2013 85(3):388–395. [PubMed: 23341369]
43. Han X, An M, Zhang W, et al. Genome Sequences of a Novel HIV-1 Circulating Recombinant Form (CRF59_01B) Identified among Men Who Have Sex with Men in Northeastern China. Genome Announc. 2013; 1(3)
44. 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. Jan 2013.1(1)
45. Li HM, Peng RR, Li J, et al. HIV incidence among men who have sex with men in China: a meta-analysis of published studies. PLoS One. 2011; 6(8):e23431. [PubMed: 21887251]
46. Ruan Y, Jia Y, Zhang X, et al. Incidence of HIV-1, syphilis, hepatitis B, and hepatitis C virus infections and predictors associated with retention in a 12-month follow-up study among men who have sex with men in Beijing, China. J Acquir Immune Defic Syndr. Dec; 2009 52(5):604–610. [PubMed: 19710617]
47. Zhang M, Chu Z, Wang H, Xu J, Lu C, Shang H. A rapidly increasing incidence of HIV and syphilis among men who have sex with men in a major city of China. AIDS Res Hum Retroviruses. Nov; 2011 27(11):1139–1140. [PubMed: 21388334]

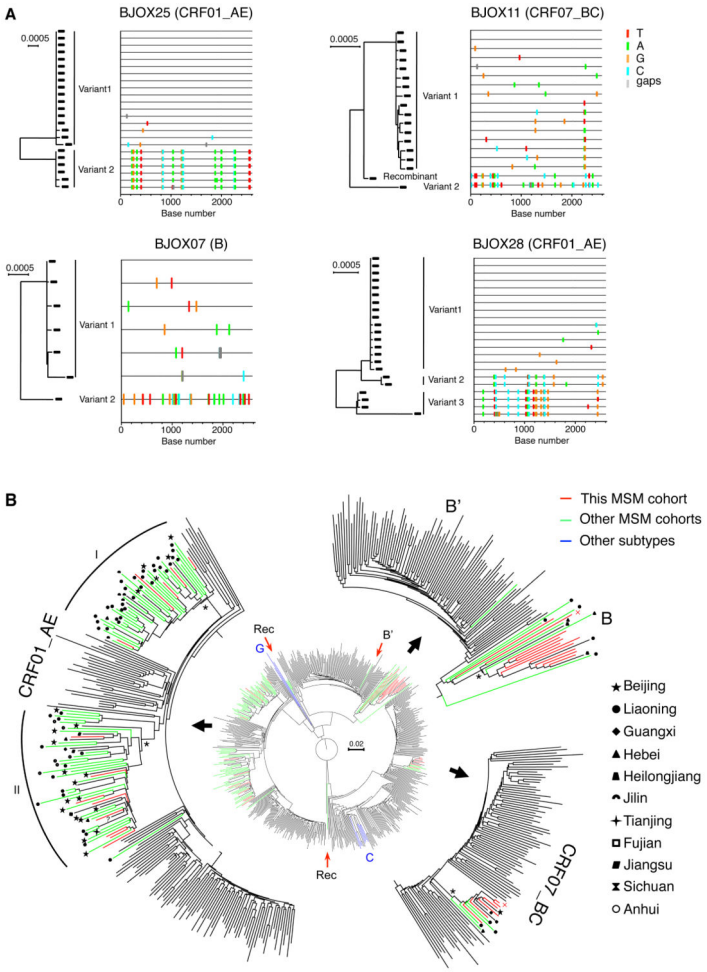


Figure 1. Genetic analysis of newly characterized MSM *env* sequences

(A) Phylogenetic tree and Highlighter plot analysis of the full-length *env* sequences. The *env* SGA sequences from each of the individuals who were infected with two or three T/F viruses were subjected to phylogenetic tree and Highlighter plot analysis. The neighbor-joining trees were constructed using the neighbor-joining method and the Kimura two-parameter model. The scale bar represents 0.0005 nucleotide substitutions per site. The Highlighter plot was done using the online tool (http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html). (B) Phylogenetic relationship of the full-length *env* sequences from the MSM cohort in Beijing and other cities. The phylogenetic tree was constructed with the newly characterized *env* sequences and all *env* sequences in China available from the Los Alamos HIV-1 Sequencer Database. The scale bar represents 0.02 nucleotide substitutions per site. The subtype B, CRF01_AE and CRF07_BC branches are enlarged to better show the phylogenetic relationship among the MSM and non-MSM sequences. Sequences from this Beijing MSM cohort, other MSM cohort are indicted by red and green, respectively. The MSM sequences from different provinces are indicated with different symbols. Non-MSM sequences were shown in black. Other subtypes are indicated by blue. MSM sequences from subtype B′ and recombinants (Rec) are indicated by red arrows. BJOX35 (B) and BJOX37 (CRF07_BC) that used CCR5, CXCR4 and other

coreceptors are indicated by red crosses. Asterisks indicate bootstrap values in which the node is supported in 80% or more replicates (out of 1000).

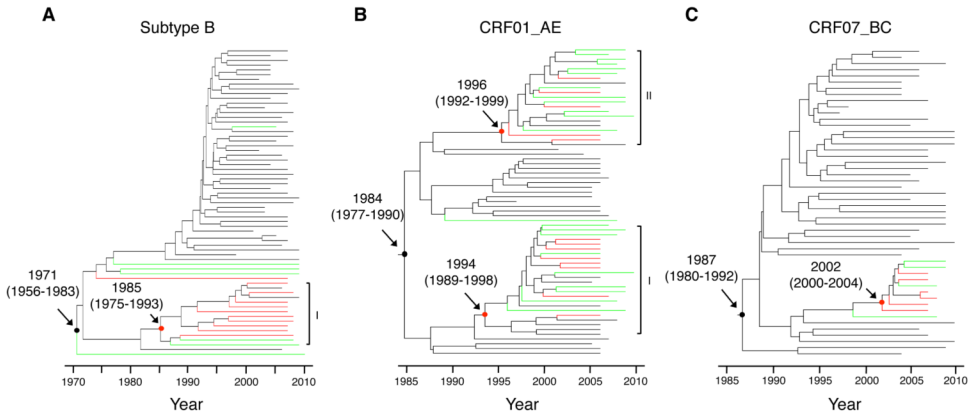


Figure 2. Estimated timescale of the spread of HIV-1 subtype B, CRF01_AE and CRF07_BC in the MSM population

Maximum-clade credibility trees were generated using Tree Annotator v1.5.4 for subtype B (A), CRF01_AE (B) and CRF07_BC (C) by Bayesian MCMC approach in BEAST1.5.4. The full-length *env* sequences from the Beijing MSM cohort (red), the MSM cohorts in other cities (green), the general population (black) were analyzed. Both strict and relaxed (uncorrelated lognormal) molecular clocks were enforced under the GTR and HKY nucleotide substitution models [14], respectively, with a gamma-distribution model of among site rate heterogeneity (with four rate categories) [15]. Each MCMC analysis was run for 20 or 30 million steps and sampled every 10,000 states. Posterior probabilities were calculated with a 10% burn-in and checked for convergence using Tracer v1.5. The maximum clade credibility tree was generated using Tree Annotator v1.5.4 available in BEAST and Fig Tree 1.3.1 was used for visualization of the annotated trees [16]. The mean time and 95% highest posterior density (HPD) of the most common ancestor (tMRCA: year) areshowed for the key nodes based on relaxed (uncorrelated lognormal) molecular clocks under HKY nucleotide substitution models in a gamma-distribution of among site rate heterogeneity with four rate categories (HKY+Γ_4). All posterior probability values for key nodes are 1.0.

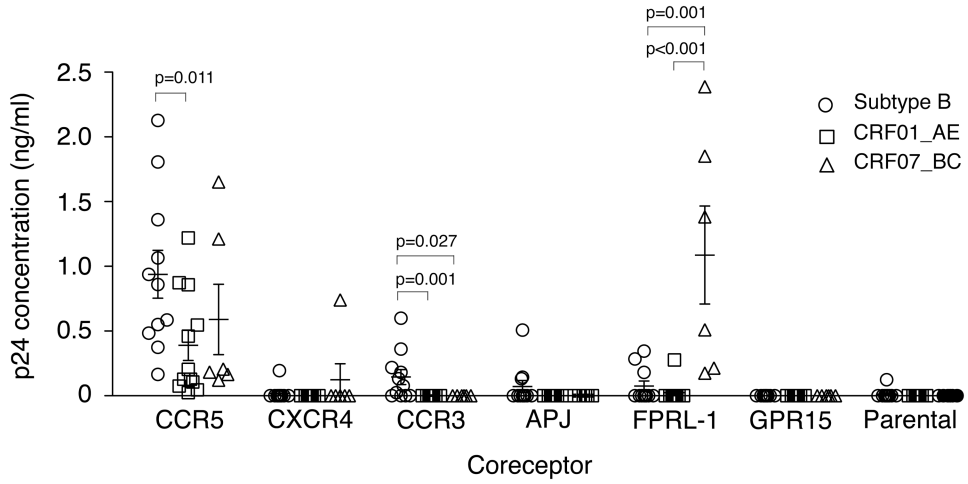
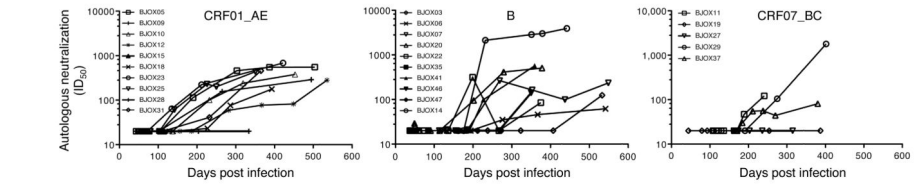


Figure 3. Disparate coreceptor usage of different genotypes

Same amount of the pseudoviruses (500 $TCID_{50}$) were used to infect NP-2 cell lines expressing CD4 and one of the coreceptors. The virus replication was determined by measuring p24 concentrations in the culture supernatant three days after infection. The coreceptor usage was determined in two independent experiments and data from one experiment is shown. The mean and standard error bars are shown for each genotype.

A



B

		CRF01_AE										B											CRF07_BC						
		BJOX05 (d504)	BJOX09 (d495)	BJOX10 (d453)	BJOX18 (d393)	BJOX23 (d422)	BJOX31 (d365)	BJOX12 (d535)	BJOX21 (d530)	BJOX25 (d351)	BJOX28 (d334)	BJOX03 (d533)	BJOX04 (d649)	BJOX06 (d542)	BJOX07 (d549)	BJOX14 (d442)	BJOX20 (d378)	BJOX22 (d375)	BJOX35 (d85)	BJOX41 (d359)	BJOX46 (d350)	BJOX47 (d273)	BJOX11 (d241)	BJOX16 (d274)	BJOX19 (d387)	BJOX27 (d315)	BJOX29 (d402)	BJOX37 (d380)
CRF01_AE	BJOX05.11	1864	554	<20	<20	80	<20	24	27	<20	<20	96	<20	<20	<20	30	<20	<20	30	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
	BJOX09.02	<20	417	35	23	71	73	<20	59	49	<20	<20	<20	70	<20	<20	<20	<20	nt	<20	<20	<20	nt	<20	<20	<20	25	<20
	BJOX10.07	104	76	520	96	88	53	<20	184	86	<20	<20	<20	79	<20	<20	<20	<20	nt	<20	<20	<20	nt	<20	<20	<20	22	<20
	BJOX18.02	<20	77	36	167	<20	<20	<20	38	142	<20	<20	<20	95	<20	39	<20	<20	84	<20	<20	<20	nt	<20	<20	<20	44	<20
	BJOX23.05	<20	nt	<20	<20	72	<20	<20	33	84	<20	nt	nt	<20	<20	<20	nt	nt	nt	nt	<20	nt	nt	<20	nt	nt	<20	nt
	BJOX31.02	65	nt	78	<20	42	214	<20	23	160	<20	nt	nt	<20	<20	<20	nt	45	nt	nt	<20	nt	nt	<20	nt	nt	<20	nt
	BJOX12.04	37	77	<20	<20	26	<20	445	29	<20	<20	53	54	31	<20	<20	22	<20	nt	<20	<20	<20	27	<20	<20	<20	<20	<20
	BJOX21.03	<20	nt	104	26	<20	<20	<20	540	90	<20	nt	nt	<20	<20	<20	nt	nt	nt	nt	<20	nt	nt	<20	nt	nt	<20	nt
	BJOX25.01	<20	nt	70	<20	<20	<20	43	142	163	68	nt	35	<20	<20	<20	80	112	nt	nt	<20	nt	nt	<20	70	nt	<20	62
	BJOX28.10	<20	nt	<20	<20	<20	<20	<20	<20	30	<20	nt	nt	<20	<20	<20	nt	nt	nt	nt	<20	nt	nt	<20	nt	nt	<20	nt
B	BJOX03.23	<20	45	<20	<20	<20	<20	20	<20	<20	<20	353	<20	<20	<20	<20	<20	<20	58	<20	<20	<20	<20	<20	<20	<20	<20	<20
	BJOX04.22	81	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	1144	<20	<20	<20	59	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
	BJOX06.13	<20	<20	<20	<20	43	<20	24	<20	<20	<20	58	<20	173	<20	<20	27	<20	40	<20	<20	<20	<20	<20	<20	<20	<20	<20
	BJOX07.02	<20	<20	<20	<20	<20	<20	34	<20	<20	<20	74	80	<20	386	<20	<20	<20	30	<20	<20	<20	<20	<20	<20	<20	<20	<20
	BJOX14.09	77	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20	<20	<20	3471	41	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20
	BJOX20.03	<20	<20	<20	<20	<20	<20	<20	40	<20	<20	26	32	<20	<20	523	<20	21	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20
	BJOX22.02	<20	<20	<20	<20	<20	<20	<20	22	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20
	BJOX35.04	<20	<20	51	<20	37	<20	<20	74	<20	<20	<20	430	56	<20	<20	34	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20
	BJOX41.16	<20	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20	<20	<20	<20	21	<20	<20	424	<20	<20	nt	<20	<20	<20	<20	<20
	BJOX46.26	<20	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	32	<20	<20	<20	59	<20	<20	<20	99	<20	nt	<20	<20	<20	<20	<20
	BJOX47.20	<20	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20	<20	<20	<20	36	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20
CRF07_BC	BJOX1.011	<20	26	<20	<20	22	38	63	27	<20	<20	55	26	<20	<20	<20	<20	<20	nt	<20	<20	<20	254	30	<20	<20	42	<20
	BJOX16.15	<20	<20	<20	<20	<20	<20	nt	188	<20	<20	<20	<20	31	<20	<20	43	48	nt	<20	<20	<20	<20	134	<20	<20	155	<20
	BJOX19.02	<20	<20	<20	<20	<20	<20	24	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	nt	<20	<20	<20	<20	<20	<20	<20	30	<20
	BJOX27.02	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	46	nt	<20	<20	<20	<20	61	<20	<20	201	<20
	BJOX29.03	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	20	<20	<20	<20	<20	40	nt	<20	<20	<20	<20	53	<20	<20	832	<20
	BJOX37.02	<20	<20	44	<20	30	<20	nt	33	28	<20	<20	<20	24	<20	<20	<20	<20	nt	<20	<20	<20	151	<20	70	<20	84	44

Note: Days post infection are indicated in brackets. The first HIV-1 RNA positive samples of BJOX04, BJOX16 and BJOX21 were at Fiebig stage VI. Days post infection for those three subjects were estimated based on the date of the most recent negative blood samples.

Figure 4. Neutralization analysis of MSM Env-pseudoviruses against autologous and heterologous plasma

(A) Development of autologous neutralizing antibodies. Neutralizing antibody titers in the plasma were determined by measuring the luciferase activity in TZM-bl cells. One T/F or representative Env pseudovirus from each individual was assayed against autologous longitudinal plasmafor CRF01_AE, subtype B and CRF07_BC. (B)Neutralization of Env-pseudoviruses by genotype-matched and genotype-mismatched plasma. One T/F or representative Env pseudovirus from each individual was assayed against genotype-matched and genotype-mismatched plasma which had the highest autologous nAb titers. The values are the reciprocal plasma dilutions at which luciferase activity was reduced by 50% (ID_{50}) relative to no plasma control wells. Red: >1:500; Orange: 1:100-1:500; Green: 1:20-<1:100; White: <1:20 or nt (not tested).

J Acquir Immune Defic Syndr. Author manuscript; available in PMC 2016 August 01.

Table 1

Demographic characteristics of HIV-1 infected MSM individuals in Beijing

Subject	Subtype	Collection date	Viral load (copies/ml)	Fiebig stage	No. of SGAs	Average diversity (%)	No. of T/F sequences
BJOX03	B	05/22/07	2,150,000	I/II	25	0.10	1
BJOX04	B	06/27/07	17,560	VI	24	0.12	1
BJOX05	CRF01_AE	06/19/07	30,211	I/II	17	0.03	1
BJOX06	B	07/16/07	498,947	IV	22	1.07	1
BJOX07	B	07/26/07	1,558	V	6	0.74	2
BJOX09	CRF01_AE	08/06/07	77,600	IV	16	0.07	1
BJOX10	CRF01_AE	08/03/07	19,300	I/II	26	0.07	1
BJOX11	CRF07_BC	08/08/07	11,900	V	18	0.38	2
BJOX12	CRF01_AE	08/16/07	3,330	V	17	0.10	1
BJOX14	B	09/05/07	2,480,000	V	22	0.05	1
BJOX15	CRF01_AE	09/27/07	1,180,000	I/II	23	0.07	1
BJOX16	CRF07_BC	10/31/07	79,600	VI	27	0.78	na
BJOX17	CRF01_AE	10/10/07	138,000	VI	26	0.56	na
BJOX18	CRF01_AE	10/17/07	16,200	IV	9	0.16	na
BJOX19	CRF07_BC	10/16/07	330,000	I/II	18	0.06	1
BJOX20	B	10/17/07	64,700	I/II	26	0.08	1
BJOX21	CRF01_AE	11/01/07	14,100	VI	15	0.27	1
BJOX22	B	11/05/07	322,000	IV	20	0.07	1
BJOX23	CRF01_AE	11/08/07	4,380,000	IV	21	0.12	1
BJOX24	CRF01_AE	11/08/07	9,170	IV	2	0.14	na
BJOX25	CRF01_AE	11/12/07	1,740,000	I/II	23	0.34	2
BJOX27	CRF07_BC	11/30/07	3,270	V	4	0.11	na
BJOX28	CRF01_AE	12/07/07	70,200	I/II	22	0.56	3
BJOX29	CRF07_BC	12/09/07	7,440	V	8	0.13	na
BJOX31	CRF01_AE	01/18/08	14,700	V	24	0.26	na
BJOX35	B	02/26/08	11,600	I/II	17	0.03	1
BJOX37	CRF07_BC	03/04/08	1,430	V	2	na	na

Subject	Subtype	Collection date	Viral load (copies/ml)	Fiebig stage	No. of SGAs	Average diversity (%)	No. of T/F sequences
BJOX41	B	04/29/08	6,080	V	21	0.19	na
BJOX46	B	06/17/08	18,700	V	23	0.21	na
BJOX47	B	06/24/08	4,480	V	23	0.09	1
Average					18	0.24	

na: not analyzed