



Published in final edited form as:

Virology. 2008 May 10; 374(2): 229–233. doi:10.1016/j.virol.2008.01.029.

Env length and N-linked glycosylation following transmission of Human Immunodeficiency Virus Type 1 subtype B viruses

Yi Liu¹, Marcel E. Curlin¹, Kurt Diem², Hong Zhao¹, Ananta K. Ghosh^{1,¶}, Haiying Zhu², Amanda S. Woodward², Janine Maenza³, Claire Stevens³, Joanne Stekler³, Ann C. Collier³, Indira Genowati¹, Wenjie Deng¹, Rafael Zioni^{2,§}, Lawrence Corey^{2,3,4}, Tuofu Zhu^{1,2}, and James I. Mullins^{1,2,3,*}

¹Department of Microbiology, University of Washington School of Medicine, Seattle, Washington, 98195

²Department of Laboratory Medicine, University of Washington School of Medicine, Seattle, Washington, 98195

³Department of Medicine, University of Washington School of Medicine, Seattle, Washington, 98195

⁴Program in Infectious Diseases, Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109

Abstract

Whether there is selection for specific viral Env variants upon HIV-1 transmission is controversial. We examined the V1V2 and V1V4 regions of Env in 10 new and 8 previously described transmission pairs infected with HIV-1 subtype B, including a total of 9 pairs in which the infecting partner had developed substantial viral diversity prior to transmission. We found that during transmission of HIV-1 subtype B, as well as for other subtypes reported in the past, viral populations in recipients undergo substantial genetic bottlenecks, as well as weak evidence for a propensity to replicate viruses with shorter variable loops and fewer potential N-linked glycosylation sites.

Keywords

HIV-1; Env sequence length; N-linked glycosylation sites; transmission; primary infection

Introduction

Understanding the properties of viruses capable of establishing infection during transmission of HIV-1 is important for vaccine design and the prevention of AIDS. In a study of 8 heterosexual transmission pairs infected with HIV-1 subtype C, Derdeyn *et al.* reported selection for viruses upon transmission with shorter length and fewer N-linked glycosylation sites predicted within the gene segment encoding the V1V4 variable regions of the Env gp120 protein (Derdeyn *et al.*, 2004). These authors suggested that longer V1V4 and more extensive glycosylation in this region might serve to protect viruses from neutralizing antibodies. They reasoned that upon transmission into an immunologically naïve host, the shield provided by

*Corresponding author at the Department of Microbiology, University of Washington School of Medicine, Seattle, WA, 98195-8070. Telephone: (206) 732-6163; Fax: (206) 732-6167; E-mail: jmullins@u.washington.edu.

¶Present addresses: Biotechnology Centre, Kharagpur-721302, Midnapore, WEST BENGAL, INDIA

§Present addresses: Div. Infectious Diseases, Stanford University Medical Center, 300 Pasteur Drive Stanford, CA 94305

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

glycosylation was not initially required, and that such variants would be outgrown by strains with a more compact V1V4 region. A second study examined gp120 V1V2 sequences from early infection in 35 individuals heterosexually infected with HIV-1 subtype A (Chohan et al., 2005). The corresponding viral segments in these individuals were found to be significantly shorter and less glycosylated than other subtype A sequences found in the Los Alamos HIV Sequence Database (HIVDB), if database sequences derived within one year of infection were excluded. This led the authors to suggest that there was selection for viral variants with shorter V1V2 loops following transmission of HIV-1 subtype A. However, they did not find shorter or less glycosylated subtype B sequences from early infection (even using early sequences from heterosexual transmissions from the HIVDB) when compared to those from the HIVDB, again excluding subtype B sequences derived within one year of infection. In a third study involving eight subtype B infected male to male transmission pairs, no selection for shorter or less glycosylated V1V4 loops was observed in recipients (Frost et al., 2005). Of note, however, at least four of the eight infecting partners in this study were themselves newly infected at the time of transmission, and they had no evident variation in V1V4 sequence length. Reliance on rapid transmission pairs could have accounted for the failure to detect selection based on V1V4 length in the latter study. To more extensively test whether selection for virus with shorter and less glycosylated variable loops occurs following transmission of subtype B HIV-1, we studied an additional ten subtype B transmission pairs (four were identified previously (Truong et al., 2002)), including one female-to-male and nine male-to-male transmissions. The infecting partners in seven pairs had developed substantial viral diversity prior to transmission.

Results and Discussion

We obtained full or partial-length HIV-1 subtype B gp120 sequences in ten self-reported transmission pairs (Table 1). Recipients were sampled at a single time point between 0 and 199 days after onset of symptoms of primary infection, while infecting partners were sampled between 41 days before to 206 days after onset of symptoms of primary infection in the recipients. To be comparable to the results of Derdeyn *et al.* (Derdeyn et al., 2004) and Frost *et al.* (Frost et al., 2005), we performed our analysis on the V1V4 region of *env*. Consistent with reported contact histories, phylogenetic analysis of V1V4 nucleotide sequences (Figure 1) showed that for each transmission pair, putative infecting partner and recipient sequences were closely related, and formed a monophyletic cluster with respect to random sequences taken from the HIVDB. As with prior reports (Delwart et al., 1994; van't Wout et al., 1994; Wolfs et al., 1992; Zhu et al., 1993) of viral populations early in infection, we generally found a large reduction of sequence diversity in the recipients (Figure 1, Table 1). However, in three transmission pairs (Pairs A, B and C), sequences from both individuals were intermingled (Figure 1) and sequence diversities in the putative infecting partners and recipients were similar and quite low (Table 1). For pairs A and B, the dates of onset of symptoms of primary infection between the putative infecting partner and recipient are only 13 and 11 days apart; for pair C, the dates of onset of symptoms of primary infection or seroconversion of the putative infecting partner are unknown and there was no clinical evidence indicating that he was infected earlier than the recipient. Therefore, we could not assign the directionality of transmission and the three pairs were omitted from subsequent analyses.

We combined the seven pairs (Pairs D to J) above with the eight pairs from the study by Frost *et al.* (Table 2) (Frost et al., 2005). As described in Derdeyn *et al.* (Derdeyn et al., 2004), median values of sequence length and number of potential N-linked glycosylation sites (PNLGS) were calculated in each infecting partner, and the number of sequences below, equal to, or above the medians was determined. Compared to the median V1V4 sequence length within infecting partners, we found that lengths of the dominant viruses (the most common genotype) in the recipients were smaller in seven pairs, identical in five pairs, and greater in three pairs. Compared to the median number of PNLGS over V1V4 in infecting partners, the dominant

viruses in the recipients had fewer in eight pairs, identical numbers in four pairs, and more in three pairs (Table 2). In nine pairs, the V1V4 length and glycosylation levels were both concordant. To assess the likelihood of this distribution, we performed the non-parametric statistical test described in Derdeyn *et al.* (perl script kindly provided by B. Korber) (Derdeyn *et al.*, 2004). Assuming that all infecting partner sequences were equally likely to be transmitted and grow out in the recipient, and given the distribution of the sequence length and number of PNLGS in V1V4 in the infecting partners, the probabilities of observing virus in seven or more recipients with shorter V1V4 lengths, and eight or more with a lower number of PNLGS in fifteen pairs were $p = 0.021$ and $p = 0.003$, respectively. This result indicates a selection for shorter and less glycosylated V1V4 loops. However, as shown below, a total of 8 tests were performed, and the statistical significance we reported only remained for the PNLGS result (for V1V2 and V1V4 regions) after Bonferroni correction for multiple tests.

Chohan *et al.* reported that sequences encompassing the V1V2 region of Env from early HIV-1 subtype A infections were significantly shorter and had fewer PNLGS than the overall population of circulating viruses (Chohan *et al.*, 2005). We therefore performed an analysis of our transmission pairs confined to the V1V2 region. Dominant V1V2 lengths in recipients were shorter than, identical to and longer than the infecting partner medians in five, nine and one pair(s), respectively. The numbers of PNLGS in dominant forms of virus in recipients were smaller than, identical to and greater than the infecting partner medians in six, seven, and two pairs, respectively. The probabilities of observing five or more recipients with shorter V1V2 sequences, and observing six or more recipients with fewer PNLGS in fifteen pairs, were $p = 0.086$ and $p = 0.0003$, respectively. For the three pairs (Pairs A to C) with uncertain transmission directionality, the sequence lengths and numbers of PNLGS in V1V2 and V1V4 were identical between the putative infecting partners and recipients in two cases, and greater in the recipient in one.

We reasoned that detection of selection was likely to be facilitated by high viral genetic diversity in infecting partners. The highest diversity among all putative recipients was 0.66% (Table 1), whereas viral diversities in the infecting partners of nine pairs in our combined data set were over 0.66%, ranging from 1.47–6.11%. We therefore performed a sub-analysis on these nine pairs (Table 2, the groups of 7 in upper panel, and 2 in lower panel, below the dotted lines). All these infecting partners had variance in sequence length and number of PNLGS. Surprisingly, we did not find selection for viral population with shorter or less glycosylated V1V2 or V1V4 loops in the recipients (all p values >0.1 ; data not shown).

In summary, our study reaffirmed the finding that substantial genotypic bottlenecks occur during and/or closely following transmission (Delwart *et al.*, 1994; van't Wout *et al.*, 1994; Wolfs *et al.*, 1992; Zhu *et al.*, 1993). We have also shown for the first time in HIV-1 subtype B, selection for viral variants with shorter variable regions and a reduced degree of glycosylation following transmission to a new host. However, this demonstration was dependent upon how pairs were chosen for inclusion. It should be stressed that we expected that this difference would be greater when choosing only infecting partners with substantial viral diversity prior to transmission, and this was not the case. Furthermore, no study so far has found these reductions to be consistent within all or a large majority of recent HIV-1 infections (e.g., shorter loop lengths and fewer PNLGS were found in 6/8 and 5/8 cases, respectively, reported by Derdeyn (Derdeyn *et al.*, 2004), and 7/15 and 8/15 cases, respectively, for V1V4 in this study). In addition, the distributions of both lengths and PNLGS numbers overlapped substantially between the two groups representing early and chronic infections, in the Chohan study (Chohan *et al.*, 2005).

A high proportion (nearly one-half) of transmissions are thought to occur when the infecting partner is in the early stages of infection (Pilcher *et al.*, 2004; Wawer *et al.*, 2005). This is

certainly a component of why the issue of selection in these regions of Env remain unsettled. The statistically significant selection we found for HIV-1 Subtype B infections occurred as sample sizes of pairs with substantial diversity in infecting partners have grown - all eight of the infecting partners evaluated by Derdeyn had high levels of diversity. The basis for this selection is thought to represent a trade-off between the advantages of antibody neutralization escape, afforded by addition of PNLGS, and replication fitness in a naïve immune environment (Derdeyn et al., 2004). The same phenomenon has been detected for SIV grown *ex vivo* in cell culture versus passaged in macaques (Edmonson et al., 1998). The failure to routinely detect this selection suggests that the fitness loss associated with neutralization escape is not always substantial.

Materials and Methods

Sequence generation and analysis

Plasma viral sequences were obtained using nested PCR as previously described (Liu et al., 2006). For *gp120*, the first round forward primer was ED3 (Delwart et al., 1995) (HIV-1_{HXB2} 5957–5986), and the reverse primer was nef3 (Shankarappa et al., 1999) (HIV-1_{HXB2} 9038–9015). The second round forward primer was gp120-5: GGCCGCGTCGACAAGAGCAGAAGACAGTGGCAATGA (HIV-1_{HXB2} 6194–6228), and the reverse primer was gp120-3: GGCCGCGGATCCGTGCTTCCTGCTGCTCCCAAGAAC (HIV-1_{HXB2} 7823 - 7787). For V1–V5, the first round forward primer was ED3 and the reverse primer was ED14 (Delwart et al., 1993) (HIV-1_{HXB2} 7961–7932). The second round forward primer was ED5 (Delwart et al., 1995) (HIV-1_{HXB2} 6557–6577), and the reverse primer was ED12 (Delwart et al., 1993) (HIV-1_{HXB2} 7811–7791). PCR products were cloned into TOPO vector (Invitrogen, San Diego, CA) and selected for sequencing as described previously (Shankarappa et al., 1999) to avoid template resampling (Liu et al., 1996). Sequences were determined with an automated DNA sequencer (Applied Biosystems, Foster City, CA) and edited using SEQUENCHER, version 3.0 (Gene Codes Corp., Ann Arbor, MI). Genbank access numbers will be provided. V1V4 nucleotide sequences from each subject were aligned using the program MacClade (Maddison and Maddison, 2001) with 114 randomly chosen HIV-1-B sequences from the Los Alamos Sequence Database (one sequence per individual). Neighbor-joining tree was constructed using HKY85 model with the program PAUP* (Swofford, 1999), and average pair-wise distance was calculated. Numbers of PNLGS were calculated using tool N-Glycosite (Zhang et al., 2004) from HIVDB.

ACKNOWLEDGEMENTS

This work was supported by grants from the US Public Health Services including for the Seattle Primary Infection Program (AI57005), AI47734, AI55336, AI49109 and the University of Washington Center for AIDS Research (AI27757).

REFERENCES

- Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, Overbaugh J. Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1–V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. *J Virol* 2005;79(10):6528–6531. [PubMed: 15858037]
- Delwart EL, Herring B, Rodrigo AG, Mullins JI. Genetic Subtyping of Human Immunodeficiency Virus Using a Heteroduplex Mobility Assay. *PCR Methods and Applications* 1995;4:S202–S216. [PubMed: 7580909]
- Delwart EL, Sheppard HW, Walker BD, Goudsmit J, Mullins JI. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J Virol* 1994;68(10):6672–6683. [PubMed: 8084001]

- Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, Mullins JI. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* 1993;262(5137):1257–1261. [PubMed: 8235655]
- Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, Heil ML, Kasolo F, Musonda R, Hahn BH, Shaw GM, Korber BT, Allen S, Hunter E. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* 2004;303(5666):2019–2022. [PubMed: 15044802]
- Edmonson P, Murphey-Corb M, Martin LN, Delahunty C, Heeney J, Kornfeld H, Donahue PR, Learn GH, Hood L, Mullins JI. Evolution of a Simian Immunodeficiency Virus pathogen. *J. Virol* 1998;72(1):405–414. [PubMed: 9420239]
- Frost SD, Liu Y, Pond SL, Chappey C, Wrin T, Petropoulos CJ, Little SJ, Richman DD. Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B. *J Virol* 2005;79(10):6523–6527. [PubMed: 15858036]
- Liu S-L, Rodrigo AG, Shankarappa R, Learn GH, Hsu L, Davidov O, Zhao LP, Mullins JI. HIV quasispecies and resampling. *Science* 1996;273(5274):415–416. [PubMed: 8677432]
- Liu Y, McNevin J, Cao J, Zhao H, Genowati I, Wong K, McLaughlin S, McSweyn M, Diem K, Stevens C, Maenza J, He H, Nickle DC, Shriner D, Collier AC, Corey L, McElrath MJ, Mullins JI. Selection on the human immunodeficiency virus type 1 proteome following primary infection. *J Virol* 2006;80(19):9519–9529. [PubMed: 16973556]
- Maddison, WP.; Maddison, DR. "MacClade - Analysis of Phylogeny and Character Evolution - Version 4.". Sunderland, MA: Sinauer Associates, Inc.; 2001.
- Pilcher CD, Tien HC, Eron JJ Jr. Vernazza PL, Leu SY, Stewart PW, Goh LE, Cohen MS. Brief but Efficient: Acute HIV Infection and the Sexual Transmission of HIV. *J Infect Dis* 2004;189(10):1785–1792. [PubMed: 15122514]
- Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, Gupta P, Rinaldo CR, Learn GH, He X, Huang XL, Mullins JI. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol* 1999;73(12):10489–10502. [PubMed: 10559367]
- PAUP* 4.0: Phylogenetic Analysis Using Parsimony (*and Other Methods) 4.0b2a. Sunderland, MA: Sinauer Associates, Inc.;
- Truong HM, Berrey MM, Shea T, Diem K, Corey L. Concordance between HIV source partner identification and molecular confirmation in acute retroviral syndrome. *J Acquir Immune Defic Syndr* 2002;29(3):232–243. [PubMed: 11873072]
- van't Wout AB, Kootstra NA, Mulder-Kampinga GA, Albrecht-van Lent N, Scherpbier HJ, Veenstra J, Boer K, Coutinho RA, Miedema F, Schuitemaker H. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. *J Clin Invest* 1994;94(5):2060–2067. [PubMed: 7962552]
- Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, Kiwanuka N, Kigozi G, Kiddugavu M, Lutalo T, Nalugoda F, Wabwire-Mangen F, Meehan MP, Quinn TC. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 2005;191(9):1403–1409. [PubMed: 15809897]
- Wolfs TF, Zwart G, Bakker M, Goudsmit J. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. *Virology* 1992;189:103–110. [PubMed: 1376536]
- Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C, Korber B. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* 2004;14(12):1229–1246. [PubMed: 15175256]
- Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, Ho DD. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. *Science* 1993;261:1179–1181. [PubMed: 8356453]

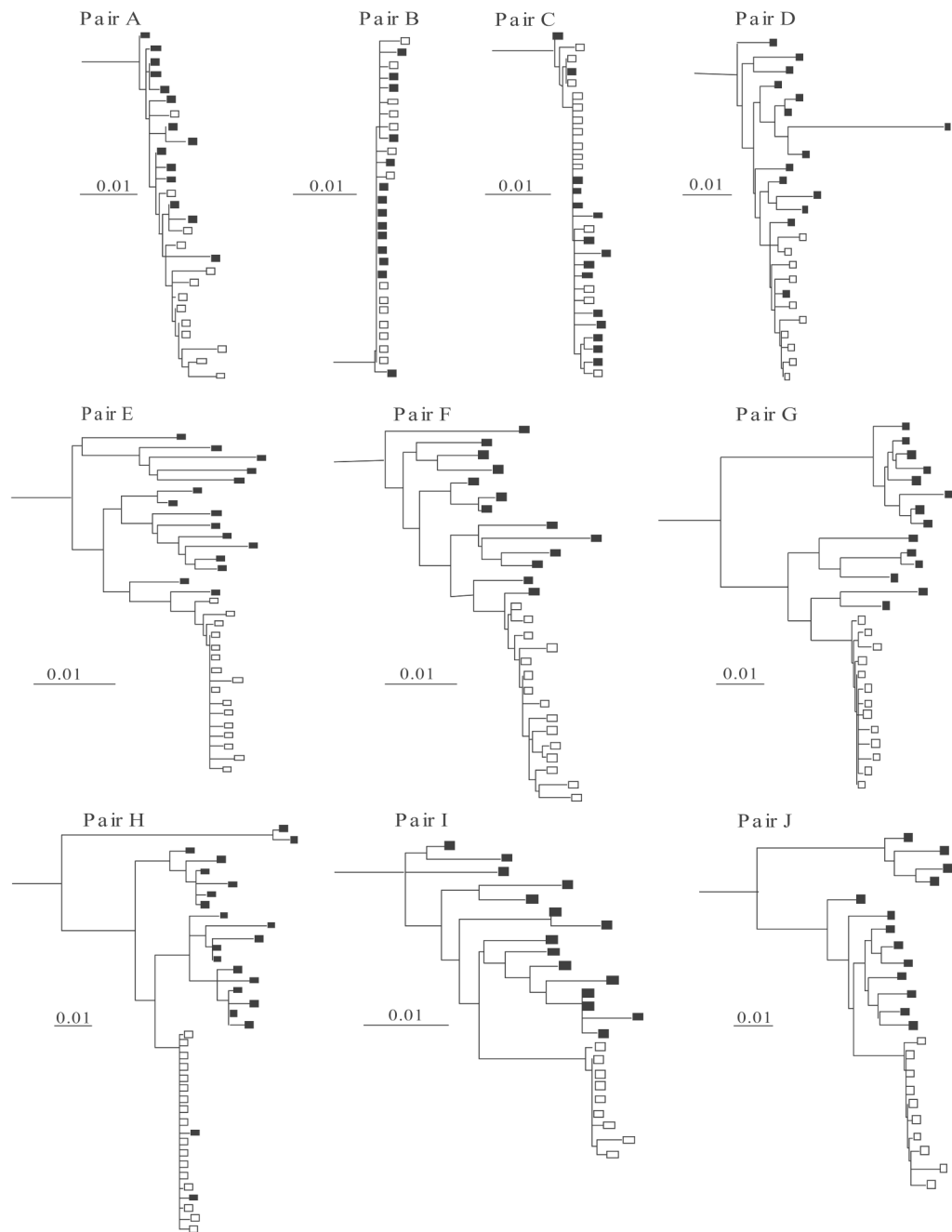


Figure 1. Rooted neighbor-join trees of the putative infecting partner-recipient pairs

To assess possible infecting partner-recipient relationships, V1V4 nucleotide sequences from each subject were aligned with 114 randomly chosen HIV-1-B sequences from the Los Alamos Sequence Database (one sequence per individual). Neighbor-joining trees were rooted with the 114 sequences from unrelated individuals (not shown in the trees). Filled rectangles represent sequences from reported infecting partners and empty rectangles represent those from reported recipients, which in each case formed a monophyletic cluster distinct from all reference sequences.

Table 1
Donors and recipients in HIV-1 subtype B infected transmission pairs

Pairs ^a	env region	Number of sequences		Estimated length of IP infection ^b	Sampling time (days) ^c		Pair-wise distance of V1V4 (%) ^e		
		IP ^d	R ^d		IP	R	IP	R	
A	gp120	14	13	94 d	81	92	0.61	0.66	
B	gp120	14	14	5 d	-6	0	0.25	0.11	
C	V1V5	15	14	<2yr	206	199	0.40	0.25	
D	gp120	15	10	~4 m	-41	56	2.11	0.62	
E	gp120	15	16	58 d	56	51	2.07	0.23	
F	gp120	15	15	2 y	49	29	2.60	0.49	
G	V1V5	14	13	~20 y	151	13	6.11	0.49	
H	V1V5	21	17	10 y	23	8	4.61	0.11	
I	V1V5	15	9	~6 y	65	42	2.33	0.23	
J	V1V5	13	10	~4 y	29	29	4.57	0.46	
Pairs from study by Frost et al. (Frost et al., 2005)									
IP	R	Days between sampling of IP and R		IP stage of infection	R	Days between sampling of IP and R	IP	R	R
		IP	R						
0007	0004	11	13	Recent	15	15	0.34	0.18	
0512	0558	11	14	Recent	-94	-94	0.15	0.26	
0512	0559	11	11	Recent	-94	-94	0.15	0.21	
0551	0550	13	12	Recent	0	0	0.31	0.32	
0206	0201	13	11	Chronic	16	16	0.62	0.17	
0206	0204	13	12	Chronic	3	3	0.62	0.32	
0465	0449	11	12	Chronic	32	8	2.78	0.21	
0564	0557	14	15	Chronic	14	14	1.47	0.34	

^aFour of the pairs have been reported previously (Truong et al., 2002). Pairs C, H, I and J in this study correspond to pairs G, E, K and B in the previous study, respectively. Pair J here is a heterosexual pair, with a female donor. In addition, additional chart reviews for pair C in this study revealed that the directionality of transmission reported previously (Truong et al., 2002) as well as the dates of infection reported in that paper were most likely in error. We were not able to determine the directionality of transmission for this pair using genetic analysis and clinical information.

^bDays (d), months (m) or years (y) between the onset of acute symptoms of primary infection or first positive antibody test and sampling in the donor.

^cDays between the onset of acute symptoms of primary infection in the recipient and sampling of donor or recipient.

^dIP, putative infecting partner; R, putative recipient.

Table 2
Comparison of V1V4 and V1V2 in donors and recipients in HIV-1 subtype B infected transmission pairs.

Pairs	Amino acid length of V1V4			PNLGS ^d in V1V4			Amino acid length of V1V2			PNLGS in V1V2		
	IP ^b Median (range) ^e	R ^b Dominant (range)	R ^b Dominant (range)	IP Median (range)	R Dominant (range)	R Dominant (range)	IP Median (range)	R Dominant (range)	R Dominant (range)	IP Median (range)	R Dominant (range)	R Dominant (range)
A	285(279-286)	285(278-285)	282(282-288)	20(18-21)	20(19-20)	67(60-67)	67(61-67)	66(66-72)	4(3-5)	4(3-5)	4(3-5)	4(3-5)
B	282(282-288)	282	290(290-293)	20(20-21)	20	66	66(66-72)	75(61-75)	5	5	5	5
C	290(290-293)	293 ^d (297-293)		21(20-22)	22(21-22)		72(72-75)					
D	278(275-282)	278(269-278)		21(18-22)	21(19-21)	58	58(55-62)					
E	297(283-313)	297		24(22-27)	23(22-23)	81	81(67-97)					4
F	287(243-296)	283(230-283)		24(18-26)	21(17-21)	68	69(66-78)					8(7-8)
G	307(295-317)	291(291-292)		24(20-27)	21(20-21)	75(75-76)	91(79-99)					6(5-6)
H	290(289-302)	292		23(22-24)	22(21-22)	70	70(69-81)					6
I	286(279-288)	288		21(19-22)	23(22-23)	69	69(67-69)					7(6-7)
J	287(283-289)	287		21(20-22)	22	68	68(67-68)					8
												6
Pairs from study by Frost et al.(Frost et al., 2005)												
IP												
0007	0004	294		23(22-23)	22(20-22)	77						6
0512	0558	295		25(24-25)	22(21-22)	66	73					7
0512	0559	295		25(24-25)	22(21-22)	66	73					4
0551	0550	288		22	22(20-22)	72	72					4(3-4)
0206	0201	283(281-283)		20(20-22)	20	63	63					6
0206	0204	283(281-283)		20(20-22)	20(19-20)	63	63					5
0465	0449	286(280-289)		21(19-21)	22(21-22)	79	70(64-73)					8
0564	0557	301(299-303)		25(23-26)	21(19-21)	72	82(81-82)					6(5-6)

^aPotential PNLGS.

^bIP, putative infecting partner; R, putative recipient.

^cRange shown if population was not monotypic.

^dLarger median or dominant numbers of the pairs are highlighted.