Genotypic and Phenotypic Characterization of HIV-1 in Patients with Primary Infection

Tuofu Zhu, Hongmei Mo, Ning Wang, Daniel S. Nam, Yunzhen Cao, Richard A. Koup, David D. Ho*

Better characterization of human immunodeficiency virus–type 1 (HIV-1) in patients with primary infection has important implications for the development of an acquired immunodeficiency syndrome (AIDS) vaccine because vaccine strategies should target viral isolates with the properties of transmitted viruses. In five HIV-1 seroconverters, the viral phenotype was found to be uniformly macrophage-tropic and non–syncytium-inducing. Furthermore, the viruses were genotypically homogeneous within each patient, but a common signature sequence was not discernible among transmitted viruses. In the two cases where the sexual partners were also studied, the sequences of the transmitted viruses matched best with minor variants in the blood of the transmitters. There was also a stronger pressure to conserve sequences in gp120 than in gp41, *nef*, and p17, suggesting that a selective mechanism is involved in transmission.

Primary (acute) HIV-1 infection (*1*) offers the opportunity to examine the initial virus replication in an immunologically naïve host. Indeed, a burst of HIV-1 replication has been documented in a number of patients undergoing the acute seroconversion syndrome (*2, 3*). However, the high level of viremia is often brought under control in the ensuing weeks, presumably by an effective immune response. Koup *et al.* (*4*) have shown that the rapid decline in HIV-1 replication is temporally correlated with the occurrence of cytotoxic T lymphocyte responses. Patients with primary infection also provide another opportunity to study the type of virus that is successfully transmitted from person to person. As the transmitted virus should be the target of AIDS vaccine strategies, information on its biological properties could be directly relevant to the development of a vaccine.

We therefore undertook this study to characterize the genotype and phenotype of HIV-1 in five patients with primary infection syndrome (seroconvertors) as well as in two of the corresponding sexual partners (transmitters). Three patients (R, V, and A) acquired the infection from homosexual contacts; their partners were unavailable for comparative studies. Two female patients, F and L, were infected sexually by their chronically infected male partners, M and C, respectively. All patients had an acute, self-limited symptomatic illness with measurable viremia that was followed by seroconversion. Blood was obtained from R, V, and A before evident seroconversion, whereas blood was obtained from F and L within 4 weeks of seroconversion, at which time blood was also obtained from their partners (M and C).

Peripheral blood mononuclear cells (PBMCs) were isolated from each subject by Ficoll-Hypaque density centrifugation. We then extracted DNA from each sample and used it in nested polymerase chain reaction (PCR) to amplify sequences in the C2 to V5 region of gp120 as well as sequences in gp41 of *env* and *nef* and p17 of *gag* (*5*). The amplified products were subsequently cloned into M13 phage and 9 to 20 independent clones were sequenced (*6*). All nucleotide sequences from this study were aligned by means of the Multiple Alignment Sequence Editor (MASE) program (*7*) and deposited in GenBank (accession numbers L21224–L21592) as well as the Los Alamos National Laboratory Database. As an example, the amino acid sequences of gp120 (C2 to V5 region) are shown in Fig. 1. Marked sequence homogeneity (>99% similarity) was seen in the seroconvertors. In contrast, substantially more sequence heterogeneity (90 to 94% similarity) was found in the chronically infected partners (M and C), consistent with previous reports (*8*). In the two cases (F and L) where the sexual partners were also studied, the gp120 sequences of the transmitted virus matched best with minor, instead of major, variants in the blood of the transmitters (Fig. 1).

The sequence homogeneity in the seroconvertors was not limited to gp120 (Fig. 2). A high degree of homogeneity was also found in gp41, *nef*, and p17. The relative homogeneity of the virus found in seroconvertors is consistent with findings in patients reported by several groups (*9–11*). Similarly, sequence homogeneity has also been documented in the blood of infants who were vertically infected (*12*). In addition, in this study (*12*), it was found that a minor virus variant present in the blood of the mother represented the transmitted virus, a result similar to what was found in F-M and L-C pairs discussed above.

We next examined the biological phenotype of the HIV-1's isolated from the patients with primary infection and from their partners. Each viral isolate was obtained as described (*2, 13*) and tested by published assays (*14*) to determine its capacity for replication in normal PBMCs, primary macrophages, and H9 and MT-2 cell lines as well as its syncytium-inducing (SI) property (Table 1). Every isolate replicated efficiently in macrophages in vitro but not detectably in T cell lines, even in cases where the partners did harbor viruses that were capable of infecting T cell lines and inducing syncytium formation. In general, these findings are in agreement with a recent report by Roos *et al.* (*15*).

How does one explain the observation that seroconvertors generally have a relatively homogeneous HIV-1 population, which is typically macrophage-tropic and non–syncytium-inducing (NSI), in spite of the fact that transmitters have a mixture of different viruses with a spectrum of phenotypes? Three models can be pro-

Table 1. Phenotype analysis of transmitted HIV-1 isolates.

HIV-1 replication	Seroconvertors					Partners	
	A	R	V	L	F	C	M
PBMC	+	+	+	+	+	+	+
Macrophage	+	+	+	+	+	+	+
H9	−	−	−	−	−	+	+
MT-2*	−	−	−	−	−	−	+

*Replication in MT-2 cells correlated with the ability to form syncytium in these cells.

Table 2. Comparison of the mean sequence diversity in gp120, gp41, *nef*, and *gag* among seroconvertors and partners. The difference in mean diversity between gp120 and gp41 or p17 of seroconvertors is statistically significant with *P* values of <0.005 or <0.001, respectively.

Region	Number of clones sequenced	Variations/ nucleotides	Mean diversity (%)
	Seroconvertors		
C2-V5	67*	106/43,293	0.24
gp41	44†	67/17,293	0.39
nef	42†	28/9,362	0.30
p17	24‡	37/7,344	0.50
	Partners		
C2-V5	34	1151/24,375	4.72
gp41	22	226/9,048	2.50
nef	22	77/4,914	1.57
p17	24	157/7,005	2.24

*From R, V, A, F, and L. †From R, V, F, and L. ‡From F and L.

Aaron Diamond AIDS Research Center, New York University School of Medicine, 455 First Avenue, New York, NY 10016.

*To whom correspondence should be addressed.

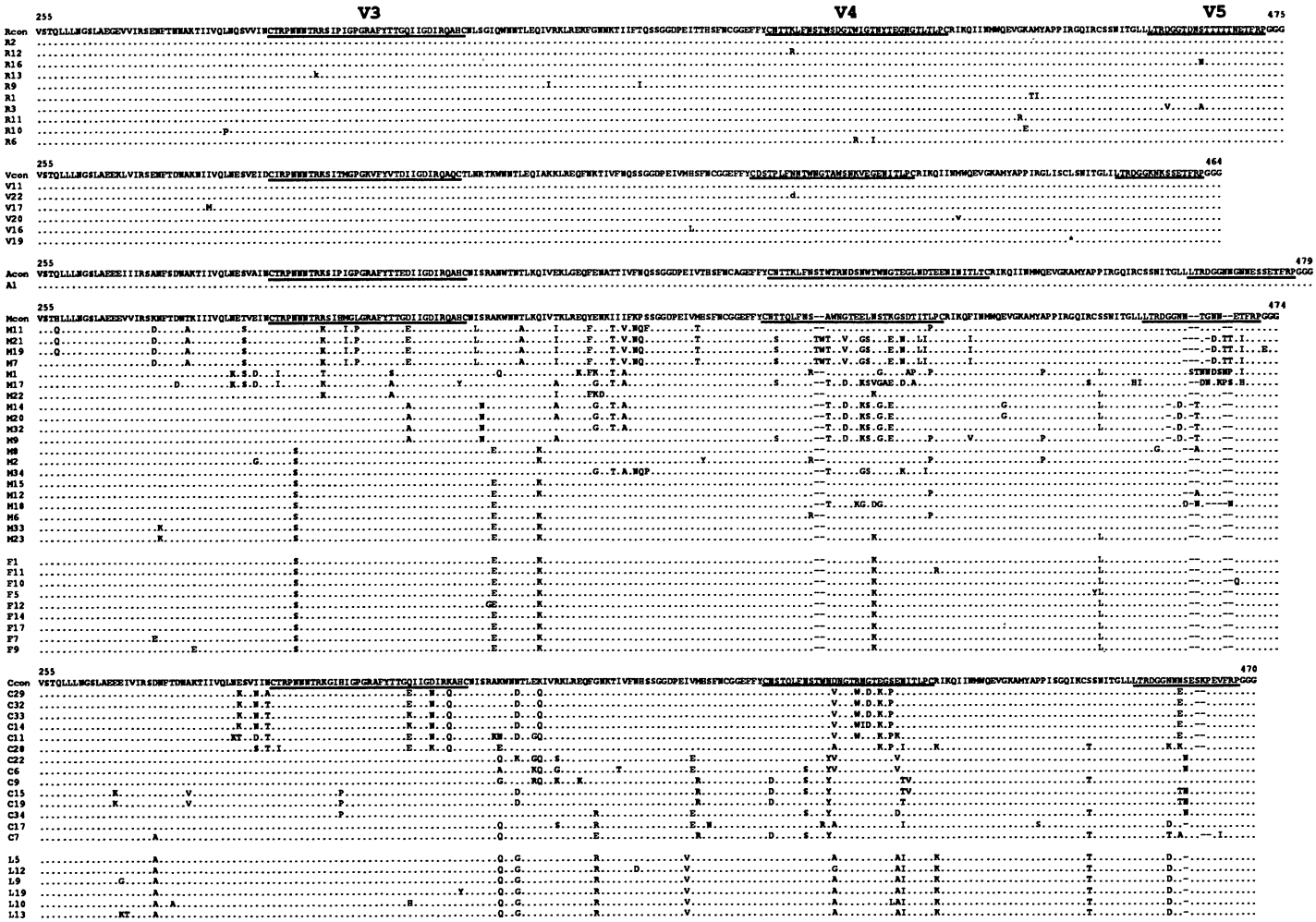

Fig. 1. Deduced amino acid sequence alignment of the C2 to V5 region of gp120 from seroconvertors (R, V, A, F, and L) and their sexual partners (M and C). The numbering starts with amino acid 255 of the HXB2 envelope sequence. Each sequence represents that obtained from one M13 clone except for R2, V11, A1, F1, and L5 where 5, 11, 10, 7, and 8 clones, respectively, were found to be identical. The sequences from R, V, and A are aligned with their own consensus sequence, whereas sequences from F and L are aligned with the consensus sequence of the corresponding partner, M or C. Dots indicate sequence identity; dashes indicate deletions; and asterisks indicate stop codons.

posed to account for this discrepancy. First, the homogeneity of the viral population found in the seroconvertor could reflect a low inoculum from the transmitter. That the transmitted virus represents only a minor variant in the blood of the transmitter argues against this explanation, because probability alone would favor the transmission of major variants. The possibility remains, however, that the minor variant in the blood could be the major variant in the genital secretion of the transmitter. Nevertheless, when one compares the sequences in gp120 with those of other genomic regions, a stronger argument against the low-inoculum effect becomes evident. As illustrated in three examples shown in Fig. 2, the degree of sequence similarity in the seroconvertors is greatest in the C2 to V5 region of gp120. Although relatively homogeneous, sequences in gp41, *nef*, and p17 show greater diversity in comparison. In fact, when all available sequences from seroconvertors were analyzed together as shown in Table 2, the mean sequence diversity (0.24%) in gp120 was significantly less than the mean values found for gp41 and p17 (0.39% and 0.50%, respectively). This finding is in contrast with the situation typically seen in chronically infected persons where the sequence diversity is most pronounced in gp120 (C and M, Fig. 2 and Table 2). Zhang *et al.* have also found greater sequence homogeneity in gp120 compared to p17 in their patients with primary infection (*11*). Taken together, these data suggest that there is a stronger selective pressure on conserving gp120 sequences at the time of HIV-1 transmission or in the interval between exposure and seroconversion. The existence of selection, in turn, argues against the random dilutional effect of a low inoculum as a sole explanation for the observed homogeneity.

The second model to explain the homogeneity of viruses during seroconversion is selective amplification. It is possible that multiple HIV-1 variants from the transmitter penetrate into the new host, but only one is selectively amplified to become the dominant population because of its biological characteristics. This explanation is supported by the observed sequence homogeneity in seroconverting hemophiliacs who were presumably inoculated with multiple HIV-1 variants parenterally (*11*). The preponderance of macrophage-tropic, NSI viruses in seroconvertors is unexpected given this model, however, because SI viruses generally replicate faster in vitro (*14, 16*) and would be expected to be the dominant virus more frequently.

The third model to account for the observed homogeneity is selective transmission. It is possible that one viral variant has a selective advantage in penetrating the

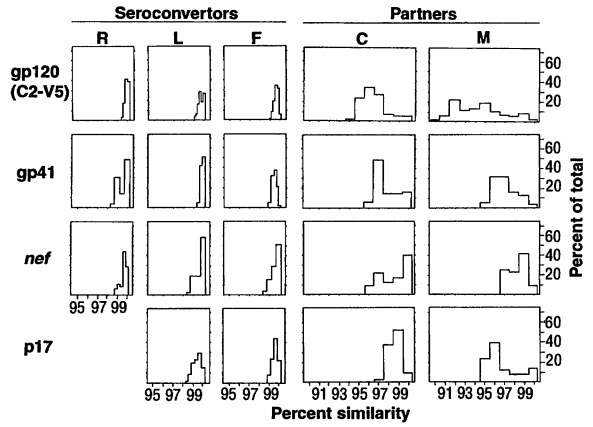

Fig. 2. Intrapatient nucleotide sequence similarity in gp120, gp41, *nef*, and p17 regions for three seroconvertors and two of the sexual partners. Each sequence within a set was compared to every other sequence in that set. For each comparison, the percent similarity for the partners' sequences is rounded off, and the fraction of the total number of sequences with a given similarity is represented as a histogram. Although the sequences in the seroconvertors showed a high degree of homogeneity, two or more subgroups were found by examining the gp41 region of patient R and *nef* region of patient L.

mucosal barrier of the new host. This model would explain not only the sequence homogeneity but perhaps also the phenotype of the transmitted viruses. That the transmitted HIV-1 isolates are all macrophage-tropic and NSI suggests that macrophages or related antigen-processing cells in the submucosal space may be responsible for the selection, allowing those viruses with the most efficient replication in these cells to penetrate. This possibility is consistent with the observation that the selective pressure is greatest on gp120, as it is known that macrophage-tropism is dictated by determinants in the HIV-1 envelope protein (*17*).

Some investigators have suggested that HIV-1 strains found in seroconvertors may have a common gp120 sequence, particularly in the V3 region (*10, 11, 18*). Closer examination of the gp120 sequences in our patients shown in Fig. 1 did not reveal a common pattern or a distinguishing feature, and no common feature was observed in the amino acid sequences. Specifically, a unique V3 signature sequence is not apparent. Comparison of the gp120 sequences from our five seroconvertors shows that the transmitted viruses differed from each other by about 10% and that they do not cluster together phylogenetically (*19*). Therefore, we conclude that, despite a common phenotype, the transmitted HIV-1 strains do not have a common genotype.

Our findings from comparisons of viral sequences in the blood of transmitters raise the possibility that transmission is selective. However, it will be important in the future to carry out comparative sequence analysis on the genital secretions of the transmitters. Despite having a common biological phenotype, no common genotype was discernible among the transmitted viruses. These findings suggest that the determinants for selective transmission are not limited to one particular envelope domain. Instead, the determinants are likely to be complex and widely distributed throughout gp120.

It seems logical to think that AIDS vaccines should induce immune responses that would target the transmitted viruses. However, the current vaccine effort is largely based on viruses that are non–macrophage-tropic and SI, which are not properties of transmitted viruses. Given that the phenotypic differences may be associated with structural or antigenic differences (*20*), it is important that some of the vaccine effort should begin to focus on viruses with the biological properties of HIV-1 found in seroconvertors.

It has been found that SI viruses emerge in the course of HIV-1 infection, which then results in a more rapid clinical course (*14, 16*). It might, therefore, be expected, a priori, that SI viruses would be transmitted occasionally, leading to a faster clinical course to AIDS in the recipient (*15*). This scenario would then predict that the clinical latency period before the development of AIDS would shorten with time, which in turn would result in an acceleration of the epidemic. However, although there are anecdotal cases of rapid clinical progression following transmission of SI viruses (*3, 15*), the overall epidemiologic observations to date do not support this prediction. It is interesting to speculate that a transmission "bottleneck" that counterselects against SI viruses is perhaps an explanation. Selection in favor of NSI, macrophage-tropic viruses would "reset the clock" and force the evolution of HIV-1 quasi-species to begin anew in each newly infected individual.

REFERENCES AND NOTES

1. D. D. Ho *et al.*, *Ann. Intern. Med.* **103**, 880 (1985); D. A. Cooper *et al.*, *Lancet* **i**, 537 (1985).
2. E. S. Daar, T. Moudgil, R. D. Meyer, D. D. Ho, *N. Engl. J. Med.* **324**, 961 (1991).
3. S. Clark *et al.*, *ibid.*, p. 954; M. Piatek *et al.*, *Science* **259**, 1749 (1993).
4. R. A. Koup *et al.*, in preparation.
5. Nested PCR was performed using the following primers: gp120 outer P5-2 (6848–6868 according to the pNL4-3 sequence in the Los Alamos Database; 5′-CCAATTCCCATACATTATTGT-3′) and P2 (7815–7789; 5′-GACGCTGCGCCCATAGTGCTTCCTG-3′); gp120 inner P5 (6957–6978; 5′-ACACATGGAATTCGGCCAGTAGT-3′) and P4-3 (7661–7639; 5′-ATTCACTTCTAGAATTGTCCCTC-3′); p17 outer P51 (512–541; 5′-CACTGCTTAAGCCTCAATAAAGCTTGCCTT-3′) and P52′ (1658–1631; 5′-TTTGGTCCTTGTCTTATGTCCAGAATGC-3′); p17 inner P53 (618–643; 5′-GTGGGAATTCTCTAGCAGTGGCGCCC-3′) and P52 (1571–1545; 5′-ATTTCTTCTAGAGGGATAGGTGGATTA-3′); gp41/*nef* outer set #1 P31 (7695–7727; 5′-TAGGAGTAGCACCCACCAAGGCAAAGAGAAGAG-3′) and P18 (9163–9132; 5′-TTCTGCCAATCAGGGAAGTAGCCTTGTGTGTG-3′); gp41/*nef* outer set #2 P32′ (8016–8045; 5′-AACTCATTTGCACCACTGCTGTGCCTTGGA-3′) and P18-2 (9072–9042; 5′-CCCCCTTTTCTTTTAAAAAGTGGCTAAGATC-3′); and gp41/*nef* inner P32 (8322–8348; 5′-CTATAGTGAATTCAGTTAGGCAGGGAT-3′) and P16 (9028–8999; 5′-TAAGTCATTGGTCTAGAAGGTACCTGAGGT-3′). The sites for restriction enzymes within the inner primers are underlined. Mixtures for PCR consisted of 10 mM tris-HCl (pH 8.3), 50 mM KCl, 0.2 mM of each of four deoxynucleotide triphosphates, 2 mM $MgCl_2$, 10 pmol of each primer, 0.2 to 1.0 μg of DNA, and 2.5 U of AmpliTaq (Perkin-Elmer Cetus) in a volume of 100 μl. Amplification conditions for the first PCR were 96°C for 2 min, 95°C for 1 min, 55°C for 1 min, and 72°C for 1.5 min for three cycles followed by 94°C for 1 min, 60°C for 1 min, and 72°C for 1.5 min for 22 cycles. We used 5 μl of the first PCR product in a second PCR that consisted of 95°C for 2 min, 94°C for 1 min, 55°C for 1 min, and 72°C for 1.5 min for 25 cycles with a final extension at 72°C for 10 min. All PCRs were carried out in the Perkin-Elmer Model 9600 thermocycler. With this procedure, we estimated the rate of misincorporation in our PCR and cloning to be 1 out of 1690 bases.
6. The PCR products digested with Eco RI and Xba I were ligated to M13m19RF and transfected into JM103 cells. We then sequenced single-strand DNA using Sequenase 2.0 kits (USB Corporation) according to manufacturer's directions.
7. D. V. Faulkner and J. Jurka, *Trends Biochem. Sci.* **13**, 321 (1988).
8. M. S. Saag *et al.*, *Nature* **344**, 440 (1988); A. Meyerhans *et al.*, *Cell* **58**, 901 (1989).
9. S. Pang *et al.*, *AIDS* **6**, 453 (1992); T. F. W. Wolfs, G. Zwart, M. Bakkar, J. Goudsmit, *Virology* **189**, 103 (1992); K. Cichutek *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7365 (1992).
10. T. McNearney *et al.*, *ibid.*, p. 10247.
11. L. Q. Zhang *et al.*, *J. Virol.* **67**, 3345 (1993).
12. S. M. Wolinsky *et al.*, *Science* **255**, 1134 (1992).
13. D. D. Ho, T. Moudgil, M. Alam, *N. Engl. J. Med.* **321**, 1621 (1989).
14. R. I. Connor, H. Mohri, Y. Cao, D. D. Ho, *J. Virol.* **67**, 1772 (1993).
15. M. T. L. Roos *et al.*, *J. Infect. Dis.* **165**, 427 (1992).
16. M. Tersmette *et al.*, *J. Virol.* **62**, 2026 (1988); M. Tersmette *et al.*, *ibid.* **63**, 2118 (1989); H. Schuitemaker *et al.*, *ibid.* **66**, 1354 (1992).
17. W. A. O'Brien *et al.*, *Nature* **348**, 69 (1990); T. Shioda, J. A. Levy, C. Cheng-Mayer, *ibid.* **349**, 167 (1991); S. S. Hwang, T. J. Boyle, H. K. Lyerly, B. R. Cullen, *Science* **253**, 71 (1991).
18. B. Chesebro, K. Wehrly, J. Nishio, S. Perryman, *J. Virol.* **66**, 6547 (1992).
19. T. Zhu *et al.*, unpublished data.
20. E. S. Daar, X. L. Li, T. Moudgil, D. D. Ho, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6574 (1990); J. P. Moore *et al.*, *J. Virol.* **66**, 235 (1992).

21. We thank G. Myers, R. K. Lenroot, K. MacInnes, and M. M. McEvilly at the Los Alamos National Laboratory for assistance with sequence analysis, C. Farthing and W. Borkowsky for providing clinical materials, R. Connor, Y. Huang, and G. Gu for technical help, and W. Chen for preparation of figures. Supported by grants from the NIH (AI24030, AI25541, AI27742, AI28747, AI30358, AI32427, and AI27665), Ernst Jung Foundation, Pediatric AIDS Foundation, and Aaron Diamond Foundation.

6 May 1993; accepted 8 July 1993