JOURNAL OF VIROLOGY, Mar. 2009, p. 2715–2727 Vol. 83, No. 6
0022-538X/09/$08.00+0 doi:10.1128/JVI.01960-08
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

Human Immunodeficiency Virus Type 1 Population Genetics and Adaptation in Newly Infected Individuals▽†

M. Kearney,[1,2]* F. Maldarelli,[1] W. Shao,[3] J. B. Margolick,[4] E. S. Daar,[5] J. W. Mellors,[6] V. Rao,[2] J. M. Coffin,[7] and S. Palmer[1]‡

HIV Drug Resistance Program, National Cancer Institute, National Institutes of Health, Frederick, Maryland[1]; Department of Biology, Catholic University of America, Washington, D.C.[2]; Advanced Biomedical Computing Center, SAIC, Frederick, Maryland[3]; Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland[4]; Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California[5]; Division of Infectious Diseases, University of Pittsburgh, Pittsburgh, Pennsylvania[6]; and Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts[7]

Received 17 September 2008/Accepted 22 December 2008

Studies on human immunodeficiency virus type 1 (HIV-1) diversity are critical for understanding viral pathogenesis and the emergence of immune escape variants and for design of vaccine strategies. To investigate HIV-1 population genetics, we used single-genome sequencing to obtain *pro-pol* and *env* sequences from longitudinal samples (n = 93) from 14 acutely or recently infected patients. The first available sample after infection for 12/14 patients revealed HIV-1 populations with low genetic diversity, consistent with transmission or outgrowth of a single variant. In contrast, two patients showed high diversity and coexistence of distinct virus populations in samples collected days after a nonreactive enzyme-linked immunosorbent assay or indeterminate Western blot, consistent with transmission or outgrowth of multiple variants. Comparison of PR and RT sequences from the first sample for all patients with the consensus subgroup B sequence revealed that nearly all nonsynonymous differences were confined to identified cytotoxic T-lymphocyte (CTL) epitopes. For HLA-typed patients, mutations compared to the consensus in transmitted variants were found in epitopes that would not be recognized by the patient's major histocompatibility complex type. Reversion of transmitted mutations was rarely seen over the study interval (up to 5 years). These data indicate that acute subtype B HIV-1 infection usually results from transmission or outgrowth of single viral variants carrying mutations in CTL epitopes that were selected prior to transmission either in the donor or in a previous donor and that reversion of these mutations can be very slow. These results have important implications for vaccine strategies because they imply that some HLA alleles could be compromised in newly acquired HIV infections.

Human immunodeficiency virus (HIV) diversity in vivo contributes to immune escape, viral persistence, antiviral drug resistance, and pathogenesis. Characterizing this diversity over the early months and years of infection is therefore fundamental to the development of effective therapeutic and vaccine strategies. HIV populations in the acute infection stage have been described as being either mostly homogeneous, resulting from the productive transmission or outgrowth of a single viral variant (8, 12, 18, 27, 31, 33, 34, 37, 38), or heterogeneous, resulting from the productive transmission or outgrowth of multiple variants (18, 31, 33). A recent study characterized both genotypic and phenotypic traits of transmitted *env* variants, finding that about 80% of patients with acute infection carry homogeneous populations, while 20% carry two or more *env* variants (18). It was also found that Env proteins in acute HIV infection are highly fusogenic, likely resulting in their efficient transmission (18). Factors that influence the transmission or outgrowth of single versus multiple HIV variants have not been elucidated. Studies are also needed to determine the impact that single versus multiple variant transmissions have on HIV diversification, immune escape, and disease pathogenesis. To date, it is unknown if additional similarities (genetic or otherwise) exist in newly HIV-infected patients or what impact the specific transmitted strains may have on diversification, divergence, and adaptation to a new host. Understanding the factors that influence HIV transmission and evolution will aid in determining mechanisms in viral pathogenesis, establishment of the latent reservoir, and emergence of drug resistance.

Viral diversity in vivo is strongly influenced by natural selection, especially by pressures imposed by our immune systems. Both antibody- and cytotoxic-T-cell (CTL)-mediated immune selection has been shown to induce escape mutations, influencing viral evolution and contributing to persistence (1, 4, 7, 9, 13–16, 20–23, 32, 36). However, there are conflicting data on the transmission and reversion of CTL-selected mutations in early infection. A number of studies have demonstrated rapid reversion of mutations in CTL epitopes after transmission to a new host (3, 11, 22), whereas other studies show evidence of CTL imprinting or persistence of CTL epitope mutations after transmission (13, 14, 20, 32). These studies have been limited by either the number of patients and/or samples observed, the methods used to detect mutations, or the gene fragment analyzed. Further studies are required to determine the role that

* Corresponding author. Mailing address: HIV Drug Resistance Program, National Cancer Institute at Frederick, 1050 Boyles Street, Building 535, Room 109, Frederick, MD 21702-1201. Phone: (301) 846-6796. Fax: (301) 846-6013. E-mail: kearneym@ncifcrf.gov.

† Supplemental material for this article may be found at http://jvi.asm.org/.

‡ Present address: Swedish Institute for Infectious Disease Control, Karolinska Institute, Stockholm, Sweden.

▽ Published ahead of print on 30 December 2008.

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

TABLE 1. Characteristics of patients in study

Patient ID	Sex[c]	Risk[d]	Age (yr)	Date of seroconversion (mo/day/yr) (days since last negative ELISA or WB)	Observation period	Duration of infection (days)[e]	No. of samples	HLA type
3021[a]	M	Heterosexual	36	11/26/98 (47)	8/4/99–10/28/03	249	9	A01 and 23, B35 and 44
3024[a]	F	Heterosexual	56	6/10/99 (10)	6/28/99–8/27/02	18	6	A68, B52 and 53
3036[a]	M	Heterosexual	37	7/5/99 (126)	1/4/00–3/13/01	179	4	A02 and 11, B18 and 35
3037[a]	M	MSM	36	9/12/99 (98)	5/17/00–11/13/02	235	5	Unknown
3041[a]	F	IDU	36	2/23/00 (14)	5/31/00–11/20/02	98	7	A03 and 29, B18 and 44
3062[a]	M	MSM	38	6/10/00 (132)	1/25/01–10/22/03	225	5	A11 and 29, B35 and 44
3077[a]	M	Heterosexual	34	3/20/01 (28)	8/21/01–5/15/03	151	9	A02 and 03, B41 and 42
3088[a]	M	MSM	25	8/13/01 (3)	2/26/02–2/26/03	163	8	Unknown
1001[b]	M	MSM	30	3/2/00 (7)	3/9/00–7/25/00	7	2	A03 and 24, B35 and 40
1002[b]	M	MSM	30	8/28/02 (8)	8/28/02–6/11/03	0	8	Unknown
1003[b]	M	MSM	38	10/23/01 (13)	10/23/01–2/20/02	0	6	Unknown
1004[b]	M	MSM	38	7/8/98 (0)	7/1/98–11/2/99	0	9	Unknown
1005[b]	M	MSM	28	2/1/01 (9)	2/1/01–7/8/04	0	7	Unknown
1006[b]	M	MSM	36	6/1/01 (14)	6/1/01–2/27/02	0	7	Unknown

[a] Patient samples from Johns Hopkins University.
[b] Patient samples from UCLA.
[c] M, male; F, female.
[d] MSM, men who have sex with men; IDU, intravenous drug user.
[e] Approximate duration of infection at first sampling, based on the day of seroconversion.

the immune system has in influencing the transmission of specific HIV variants, virus diversification after transmission to a new host, and global diversity of HIV.

In the present study, we used single-genome sequencing (SGS) to characterize viral diversity in longitudinal samples from treatment-naïve individuals with acute or early HIV type 1 (HIV-1) infection. Using SGS rather than cloning to study viral genetics eliminates the effects of PCR-based recombination and also eliminates template resampling, leading to more accurate phylogenetic analyses and measurements of diversity and divergence (29, 34). Our results show that population structure and diversification vary in early HIV infection among patients infected with single and multiple variants. However, in all patients studied, transmitted variants of HIV carried CTL epitope-associated mutations in protease (PR) or reverse transcriptase (RT) relative to the consensus subtype B virus. These mutations persisted for years, even at CTL epitopes that were not recognized by the new host's major histocompatibility complex (MHC). The persistence of these CTL mutations was concurrent with the selection of new escape mutations in epitopes recognized by the recipient's MHC. We demonstrate here that HIV variants carrying mutations at CTL epitopes are readily transmitted between hosts and that regardless of whether infection resulted from the transmission of a single variant or multiple variants, these mutations in PR or RT are not likely to revert quickly.

(This work was done in partial fulfillment of thesis requirements for M. Kearney.)

MATERIALS AND METHODS

Clinical specimens. Plasma samples (n = 93) from 14 patients with acute or recent HIV-1 subtype B infection (estimated date of seroconversion, 0 to 5 months prior to the first sample being obtained) were studied. All patients were naïve to antiretroviral therapy. Samples from six patients (patients 1001 to 1006) were obtained from the Acute HIV and Early Disease Research Program (AIEDRP) at the Los Angeles Biomedical Research Institute, Torrance, CA (Table 1). The last negative enzyme-linked immunosorbent assay (ELISA) or Western blot for these patients was less than 1 month before the first sample analyzed for this group (Table 1). Samples were obtained frequently postseroconversion (daily, biweekly, or weekly), followed by approximately monthly sampling for up to several years after infection. Samples from the other eight patients (subjects 3021, 3024, 3036, 3037, 3041, 3062, 3077, and 3088) were obtained from the AIEDRP at Johns Hopkins University, Baltimore, MD (Table 1). The last nonreactive ELISA screening for these patients was within 5 months of the first positive sample. Two to six samples were collected each year after early HIV infection (<1 year postseroconversion) through the first several years of infection. All studies were approved by the institutional review board of NIAID, NIH, Bethesda, MD, and all study subjects provided written informed consent. Consent for HLA typing was obtained from 7 of 14 patients (Table 1).

SGS and analysis of data. Viral RNAs were extracted from plasma samples, and SGS of the approximately 1.2-kb p6, protease, and RT (*p6-rt*) fragment was performed as previously described (17). In addition, single genome sequences from an approximately 1-kb fragment of *gp120*, including the V1-V2 and V3 regions, were also obtained from a subset of samples (n = 73). The SGS protocol was conducted as previously described, except using envelope (*env*) primers E20 and E115 for PCR and E30 and E125 for nested PCR (see the supplemental material) (35). Five to 94 single genome sequences were obtained from each plasma sample (mean = 36). In total, we obtained 3,340 sequences from 14 patients. Data resulting from low-quality electropherograms and multiple sequences were eliminated by an automated quality control step prior to the generation of contigs. Contigs were required to have sequence reads in two directions to be included in the automated alignment process. Sequence alignments were obtained using Clustal W and additionally edited by hand when needed. Envelope sequences were first translated and amino acid sequences were aligned to ensure that no frameshift differences were included in the V1-V2 region. Phylogenetic analyses to determine population structure were performed using a neighbor-joining P distance method in MEGA3 (http://www.megasoftware.net). Analyses of synonymous versus nonsynonymous mutations relative to the HIV-1 consensus B reference sequence (http://hivdb.stanford.edu) were performed using an in-house software program and Highlighter (http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter.html). Sequences from the most recent common ancestors (MRCA) were determined phylogenetically using a maximum likelihood method in PAUP. Measurements of diversity (average pairwise distance) and divergence were calculated using MEGA4. CTL epitopes and mutations in PR, RT, and Env were determined based on the sequence and immunology databases from Los Alamos (http://www.hiv.lanl.gov/content/index). Insertions in patient sequences compared to the consensus subgroup B sequence were deleted for mutational analyses. $CD4^+$ T-cell slopes were determined using linear regression of 4 to 29 measurements of $CD4^+$ T cells/μl plasma (see Fig. 4) over the observation period for each patient (4 to 59 months). The statistical correlation between the frequency of transmitted CTL-associated mutations and $CD4^+$ T-cell slopes was determined using a nonparametric correlation (Spearman r) test with no Gaussian assumptions.

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

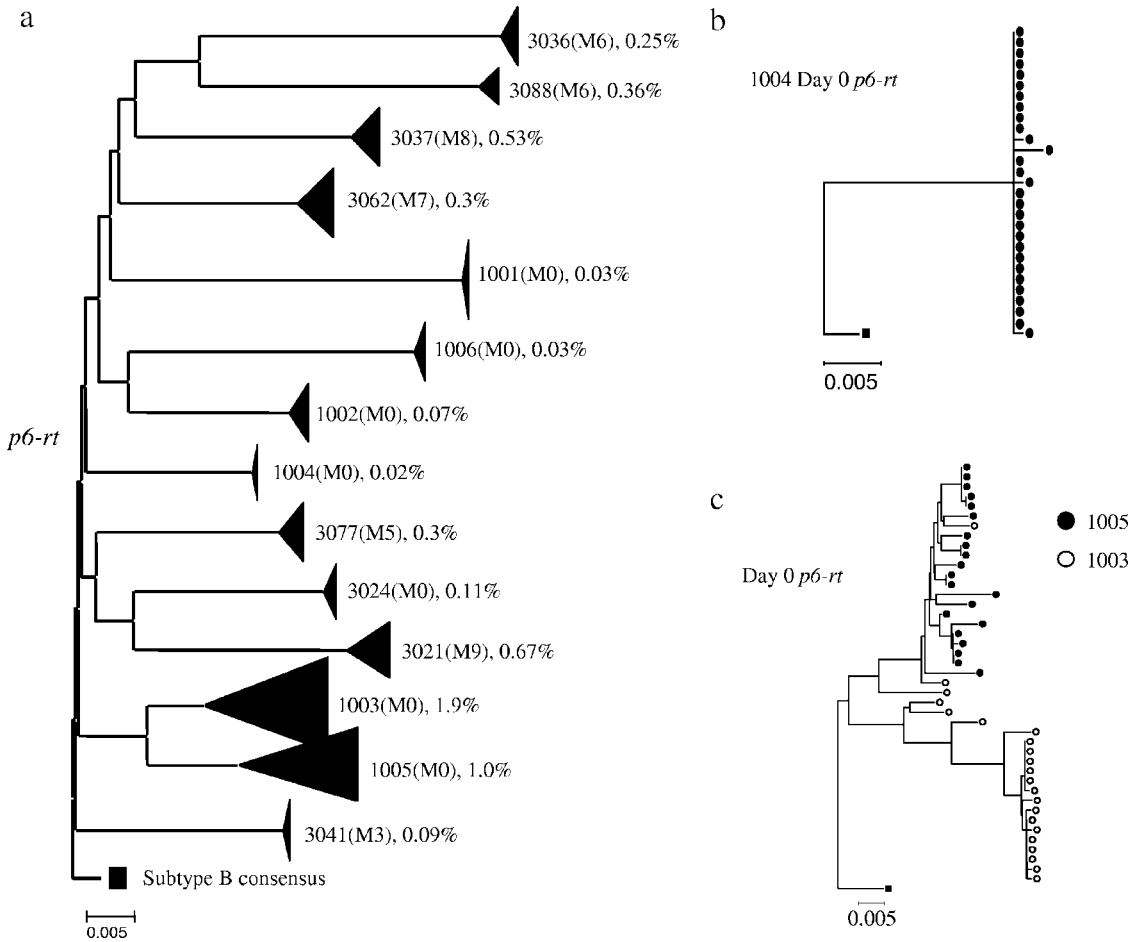

FIG. 1. HIV-1 *p6-rt* diversity in acute and early infections. (a) Population structure and diversity (measured as percent APD per site) of the *p6-rt* region in single genome sequences. Sequences obtained from the first available plasma sample postseroconversion from all 14 patients are shown on a schematic neighbor-joining phylogenetic tree rooted on the HIV-1 consensus B sequence. (b) Population structure in acute infection after single-variant transmission. A neighbor-joining phylogenetic tree was constructed for the *p6-rt* fragment from the first available plasma sample from one single-variant-infected patient (1004). (c) Population structure in acute infection after multiple-variant transmission. A neighbor-joining phylogenetic tree was constructed for the *p6-rt* fragment from the first available plasma samples from the two multiple-variant-infected patients, 1003 and 1005.

RESULTS

Population structure and diversity in acute and early HIV infections. To assess viral characteristics in acute and early HIV-1 infection (<1 year after transmission), we analyzed and compared single genome sequences obtained for *gag-pro-pol* (*p6-rt*) and *env* from the first available plasma sample for 14 recently infected patients (Table 1). Figure 1a shows the phylogenetic structure and diversity, measured as percent average pairwise distance (APD), of the *p6-rt* fragment in the earliest available sample for each patient. Sequences from 12 of 14 patients formed independent populations on the phylogenetic tree, with no intermingling of sequences (Fig. 1a), demonstrating that the viruses found in these patients were genetically distinct. The viruses in these 12 patients had simple population structures (an example is shown in Fig. 1b) and low diversity (0.02 to 0.67% in *p6-rt* and 0.05% to 0.9% in *env* [not shown]), with many sequences being identical. These data are consistent with infection having been seeded by or growing out of a single HIV variant. Two of the 14 patients (1003 and 1005) had viruses in acute infection (9 and 13 days after nonreactive ELISA or indeterminate Western blot) with intermingling sequences in the phylogenetic tree (Fig. 1a and c), indicating that these patients either were a donor-recipient pair or had a common donor. The viruses in these patients also exhibited high diversity and complex population structures (Fig. 1c) in the *p6-rt* fragment (1.9% in patient 1003 and 1.0% in patient 1005) (Fig. 1a) and in *env* (5.4% and 2.0%, respectively) (not shown), consistent with transmission of multiple HIV variants from the same diverse population.

Cross-sectional study of diversity in acute and early HIV infections. A cross-sectional analysis of diversity versus time since seroconversion with the first available samples from the 12 patients likely to have been infected by the transmission and expansion of single viral variants (Fig. 2) revealed a linear relationship (Fig. 2) between the time of the first sample tested (0 to 9 months postseroconversion) and the APD, with a slope of 0.002%/day, extrapolating back to the clonal virus (APD = 0) at about 18 days prior to seroconversion (not shown). The

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

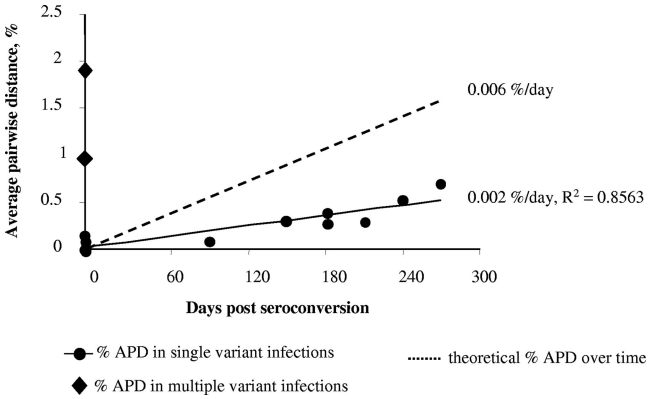

FIG. 2. Diversity of single genome sequences obtained from the first available plasma samples for all 14 patients. APD values for the first available samples from single-variant-infected patients are shown with circles, and those for multiple-variant-infected patients are shown with diamonds. Neutral accumulation of diversity predicted from the estimated mutation rate of HIV-1 (assumed to be 3×10^{-5} mutation per base per replication cycle and 1 replication cycle per day) is also shown.

accumulation of diversity determined in this analysis was somewhat slower than that expected from the theoretical mutation rate (0.006%/day), perhaps reflecting purifying selection pressure in early infection in vivo. Alternatively, the replication cycle time may be greater than the rate of 1 cycle/day used in this estimate. The APDs of the virus populations in the first sample tested from each of the two patients infected with multiple viral variants (1003 and 1005) were significant outliers in this cross-sectional analysis (Fig. 2). The viruses in these patients exhibited high diversity, far outside the range of the expected mutation accumulation rate for HIV (Fig. 2), again consistent with transmission of multiple variants.

CTL epitope-associated mutations in acute and early HIV infections. We sought to identify genetic similarities among virus populations in acute and early HIV infections. For this purpose, we analyzed synonymous and nonsynonymous mutations compared to the wild-type consensus B sequence (http://hivdb.stanford.edu) as well as to the phylogenetically determined MRCA of the combined patient sequences (nearly identical to the consensus B sequence). A summary of the *pro* sequence comparisons for the earliest time points for all patients (Fig. 3) reveals a striking pattern: over 95% (72/75 differences) of nonsynonymous differences formed the clusters shown in Fig. 3 (shaded regions), encompassing about half of the protein, whereas only about one-half (22/45 differences) of the synonymous differences were in the same regions. Further examination (Fig. 4 and 5) revealed that these variable regions in PR as well as similar regions in RT (see Fig. S2 in the supplemental material) corresponded to predicted CTL epitopes (19) in these genes. The same pattern was seen for patients infected with single (Fig. 4a and 5; see Fig. S1 in the supplemental material) or multiple (Fig. 4b; see Fig. S1 and S2 in the supplemental material) variants. Most mutations found within CTL epitopes were at anchor residues or within the proximal three amino acids, primary locations for diminishing antigen presentation (5, 26). In contrast, synonymous differences were distributed randomly across the sequence, with no obvious relationship to CTL epitopes (Fig. 3, green bars).

Among individuals infected with a single variant and where a sample was available within 5 months of seroconversion, most CTL-associated mutations in the first available samples were found in 100% of sequences obtained by SGS (Fig. 4a and c, maroon bars; see Fig. S1 and S2 in the supplemental material), consistent with the initial homogeneity of the virus population. In contrast, the virus populations in acute infections after transmission of multiple HIV variants consisted of dominant populations with mutations in CTL epitopes and minority populations that were wild type (matching the consensus B sequence) at several of the same residues. For example, the dominant population in patient 1003 carried CTL-associated mutations in 80% of the genomes obtained (Fig. 4b, maroon bars), while a minority population (20% of genomes) was wild type in the same epitopes in PR (Fig. 4b). Thus, sequence heterogeneity observed in acute infection after transmission of multiple HIV variants was due in part to amino acid differences in CTL epitopes among the transmitted variants. Diversity in *env* in patients infected with multiple variants was not as apparently related to mutations in CTL epitopes as that in PR and RT but was due primarily to insertions and deletions in the V1 and V2 regions and to amino acid substitutions in the V3 loop (data not shown).

Persistence of CTL epitope-associated mutations in longitudinal samples. It has been proposed that mutations in HIV-1 that lead to escape from CTLs are somewhat deleterious to the virus in the absence of an immune response and are therefore likely to revert to the wild type following transmission to an individual of a different MHC type (3, 11, 22). To investigate this issue, we analyzed nonsynonymous mutations in sequences from longitudinal samples (up to 5 years after infection) obtained from the same set of patients. Among the large number of mutations relative to consensus subtype B present within CTL epitopes, we found little evidence for reversion of transmitted mutations in PR and RT following infection in either single- or multiple-variant-infected patients, even in epitopes lacking selection according to the recipient MHC type (Fig. 4 and 5; see Fig. S1 and 2 in the supplemental material). In fact, in patient 1003, who was infected with multiple variants that were either mutant or wild type at specific residues within CTL epitopes, there was a loss of the wild-type populations in the *p6-rt* fragment over time, while the mutant population persisted. This effect can be seen by the increased frequency of these mutations in Fig. 4b (yellow bars versus maroon bars). In contrast, all transmitted *env* variants persisted in patient 1003, indicating that selection pressure upon transmitted CTL mutations may not have a strong impact on Env compared to that on PR and RT in early HIV infection. New mutations in epitopes recognized by the patients' MHC did emerge with time in several individuals (Fig. 5; see Fig. S1 in the supplemental material), without reversion of transmitted mutations.

Superinfection. The only evidence we found for early loss of transmitted CTL epitope mutations was in patient 1001 (Fig. 4c), who was previously reported to be superinfected 4 months after primary infection (30). Figure 4c shows the different CTL mutation profiles for this patient in primary infection (day 0) and in a sample collected after superinfection (day 139) (30). Although variants carrying CTL mutations were found in both

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

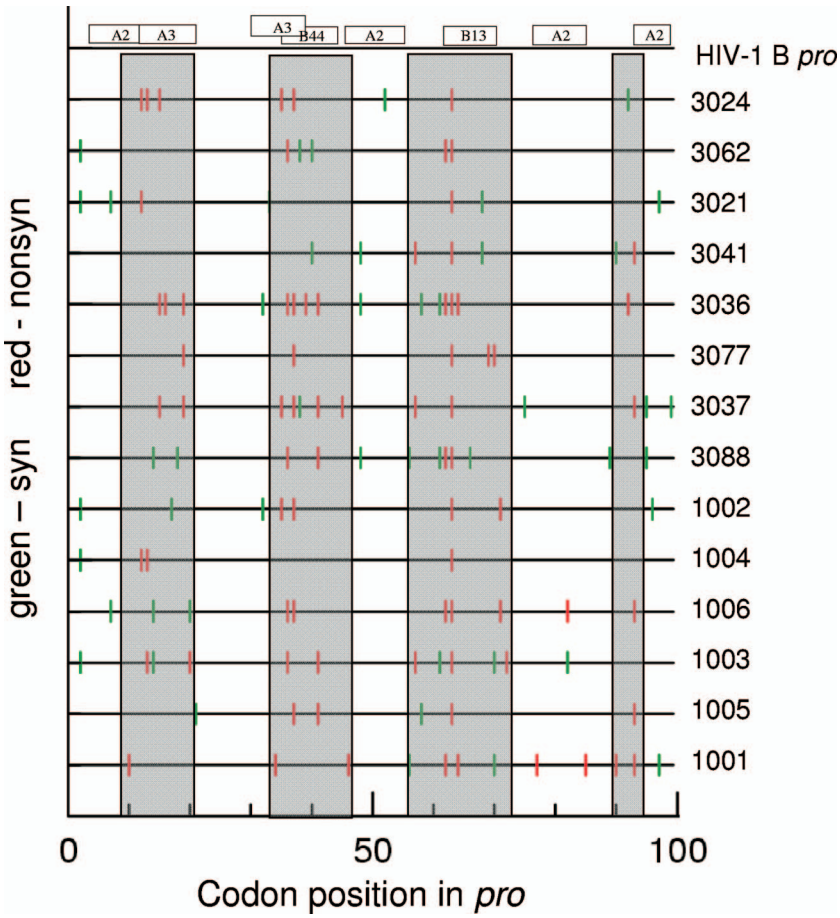

FIG. 3. Synonymous and nonsynonymous differences in *pro* compared to consensus subtype B sequence. The Highlighter plot shows the locations of amino acid changes (red) and synonymous mutations (green) in the *pro* gene relative to the consensus of the single genome sequences obtained from the first available sample after infection for each patient in the study and compared to the consensus B sequence. The shaded areas show the clustering of nonsynonymous changes. Locations of known CTL epitopes are shown above the plot.

transmitted strains, mutations associated with virus from the primary infection were scarcely detectable in samples collected after superinfection, consistent with nearly complete replacement of one virus by the other. SGS showed that this patient was initially infected with a clonal HIV-1 subtype B infection but was subsequently superinfected with a second single variant of HIV-1 subtype B (see Fig. S3 in the supplemental material) (39). The newly transmitted strain rapidly became the dominant viral population, and 98 of 100 sequences obtained from the first available sample after superinfection were descendants of the secondary exposure. No virus from the primary infection was detected in subsequent samples, even using a very sensitive allele-specific PCR assay (30; data not shown). Primary infection in patient 1001 was with a multidrug-resistant virus, while superinfection was with a virus that has wild-type sequences at all known sites of drug resistance (39). No recombination between the two populations was detected, indicating that the drug-resistant virus was less fit than the wild-type virus in vivo.

Impact of CTL-associated mutations on CD4 T-cell slope. Table 2 shows the frequencies of predicted mutant and wild-type CTL epitopes in PR and RT as well as the $CD4^+$ T-cell slopes for all 14 patients. Although patients with the highest frequencies of mutated CTL epitopes in PR and RT (3024, 3037, 1001, and 1003) (Table 2, sum of columns 5 and 6) tended to have $CD4^+$ T-cell counts that dropped to below 200 during the observation period, there was no significant trend ($P = 0.12$; Spearman rank correlation) relating the $CD4^+$ T-cell slope to the frequency of transmitted CTL mutations. However, the two patients carrying viruses resulting from the transmission of multiple HIV variants and the superinfected patient all had at least one $CD4^+$ T-cell count that was <200 cells/μl over the observation period, whereas only 3 of 11 single-variant-infected patients had $CD4^+$ cell counts as low ($P = 0.0096$; Spearman rank correlation), suggesting a significant correlation between infection with multiple variants and lower $CD4^+$ T-cell counts.

Diversification after HIV infection. In all single-variant-infected patients with sufficient follow-up, the initially homogeneous virus populations accumulated diversity in the early years of infection until essentially every *p6-rt* sequence obtained from the sampled populations was unique. Accumulation of genetic diversity with time is shown in Fig. 6 for the four patients infected with single HIV variants for whom the first available sample was collected within 3 months of the last nonreactive ELISA (1002, 1004, 1006, and 3041) and for the

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

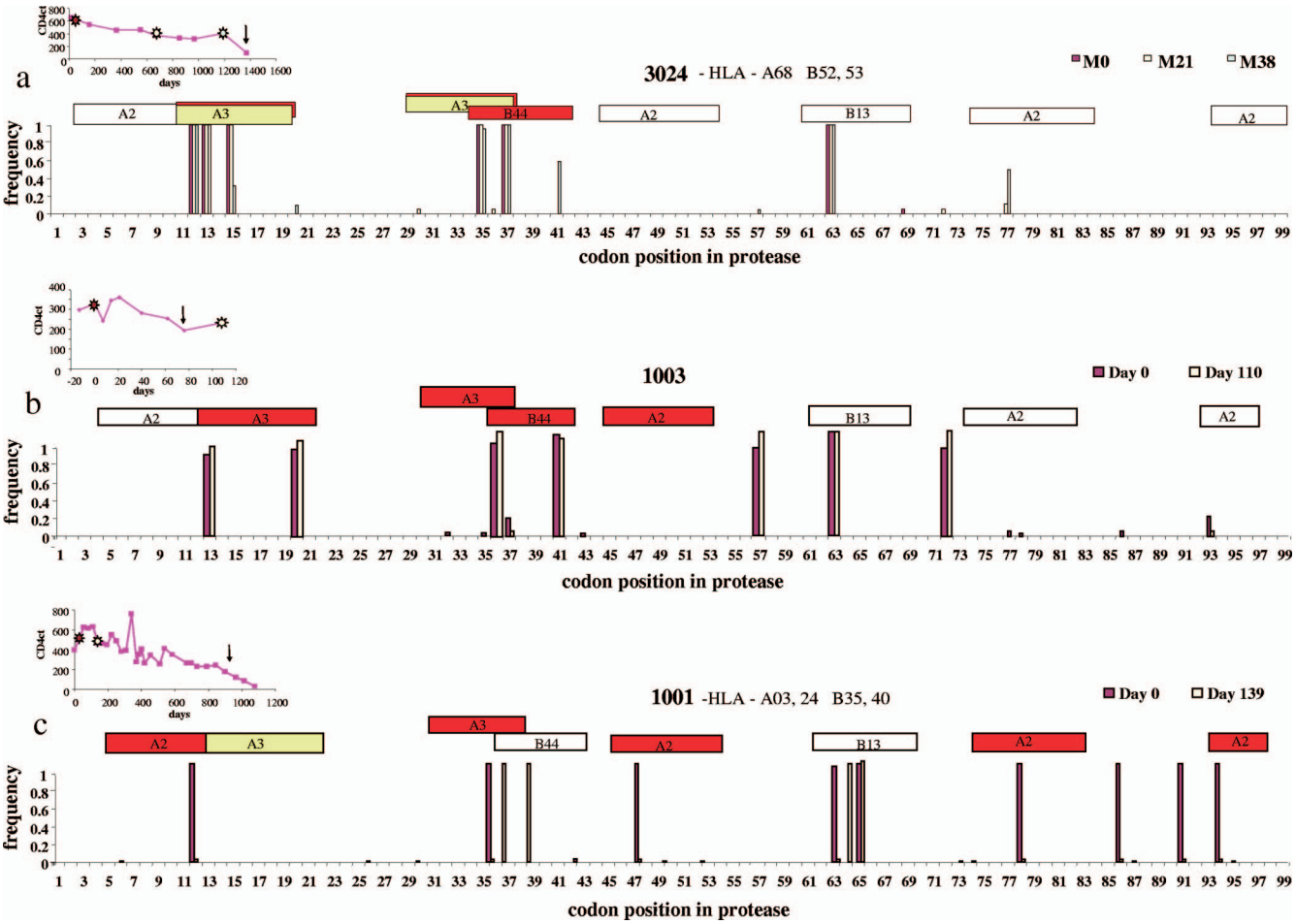

FIG. 4. Amino acid changes in protease compared to the consensus subtype B sequence. The bar graphs show the proportions of single genome sequences with amino acid changes at the indicated amino acid positions in *pro* compared to the consensus B sequence at early and late time points after infection, with different colors representing different times after infection, as shown. Horizontal boxes above the graphs represent known CTL epitopes in the subtype B consensus sequence. Red boxes designate CTL epitopes with probable escape mutations compared to the consensus. Yellow boxes designate epitopes expected to be recognized by each patient's MHC. Plots of $CD4^+$ T-cell counts as a function of time are shown in the inset graphs. Downward arrows indicate CD4 counts of <200. (a) Typical patient (3024) infected with a single variant; (b) patient (1003) infected with multiple variants; (c) patient (1001) who was superinfected about 4 months after primary infection. Plots for all patients are shown in Fig. S1 (PR) and S2 (RT) in the supplemental material.

two patients after infection with multiple variants (1003 and 1005) (note the different scales). While the *p6-rt* and *env* regions analyzed tended to show similar (if sometimes erratic) increases in diversity in the first year after transmission of a single variant (Fig. 6, patients 1002, 1004, and 1006), leveling off or increasing steadily after the first year, diversity in *env* continued to increase erratically over the course of infection (Fig. 6, patient 3041, and Fig. 7, patient 3024), consistent with greater selection and tolerance of diversity in this portion of the genome. Diversity was typically lower in *p6-rt* than in *env* after the first year of infection (Fig. 6, patient 3041, and Fig. 7, patient 3024). Figure 7 shows the slow accumulation of diversity over the first 3 years of infection in single-variant-infected patient 3024. In this patient, diversity increased in *p6-rt* at an initial rate of 0.0015%/day (similar to that observed in the cross-sectional analysis) (Fig. 1b), slowing to 0.0007%/day between 9 and 38 months. In *env*, diversity increased initially at a rate of 0.0026%/day and erratically thereafter, at an average rate of 0.002% between months 9 and 38 (Fig. 7, note the different scales on the trees). Divergence from the initial infecting virus was also plotted for patient 3024 (Fig. 7) and followed essentially an identical pattern to that of diversification.

Little is known about diversification in patients infected with multiple HIV variants. Diversification after infection in the two multiply infected patients (1003 and 1005) is shown in Fig. 6 (note the different scales). *env* diversity was approximately two- to sixfold higher than *p6-rt* diversity both after transmission and in longitudinal samples. In patient 1003, there was a decrease in diversity in *p6-rt* from 1.9% at seroconversion to 0.2% only 14 days later (Fig. 6, 1003, corresponding to the loss of the wild-type transmitted CTL epitopes in the *p6-rt* variants (Fig. 4b). In contrast, all transmitted viral variants in *env* persisted over the same observation period, resulting in little overall change in diversity beyond the typical variation seen in *env* diversity in all patients (Fig. 6, patient 1003). In patient 1005,

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

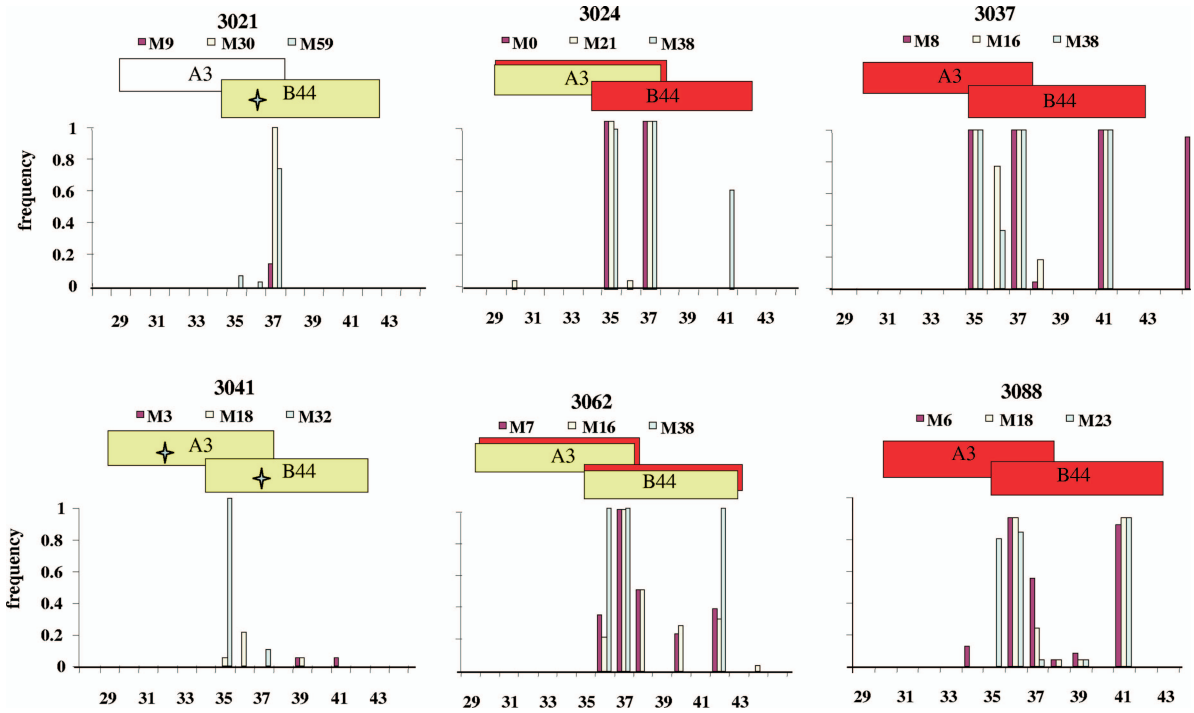

FIG. 5. Amino acid changes in protease compared to the consensus subtype B sequence. An enlarged view of a portion of a graph like those in Fig. 4 (shown in full in Fig. S1 in the supplemental material) is shown for the indicated singly infected patients. Stars indicate mutant epitopes matching patient MHC types that were wild type in the earliest available sample after infection.

little change in diversity was seen in either *p6-rt* or *env* for more than 1 year after seroconversion (Fig. 6, again note the different scale), despite divergence of the population into structures that lacked identical sequences (data not shown). However, 3 years after infection, genetic diversity increased in *env*, while in *p6-rt* it was only slightly higher (1.3%) than that seen in acute infection in this patient (1.0%). Although both multiple-variant-infected patients had persistently high diversity in *env* from acute HIV infection onward, early emergence of CXCR4 variants was not detected. Patient 1003, but not patient 1005, had higher viral loads at entry and at the set point (month 3) than did patients infected with single variants (Table 2).

TABLE 2. Virus loads, slopes of CD4 T-cell decline, and potential CTL escape mutations in patients in this study

Patient ID	Viral load at entry (copies/ml)	Viral load at set point (copies/ml)[a]	CD4 slope (cells/μl)	No. of CTL epitopes in first available sample[b]			
				Mutant CTL epitopes		Sensitive CTL epitopes	
				PR	RT	PR	RT
3021	3,782	63,265	−0.16	**1**	**8**[d]	1	1
3024	115,721	173,550	−0.31[c]	**3**[d]	**8**[d]	0	4
3036	40,912	61,459	−0.81	**4**[d]	**5**[d]	3	14
3037	47,272	82,106	−0.28[c]	5	7	NA	NA
3041	143,505	21,248	−0.07	**1**	**4**[d]	3	6
3062	34,299	19,476	−0.08[c]	**2**[d]	**3**[d]	1	7
3077	187,982	722,926	0.3	**3**[d]	**4**[d]	4	12
3088	1,392,011	470,189	−0.55	2	7	NA	NA
1001	300,815	136,647	−0.46[c]	**5**[d]	**10**[d]	1	2
1002	671,679	17,317	−1.21	2	4	NA	NA
1003	110,000,000	453,737[e]	−0.89[c]	4	8	NA	NA
1004	9,347,800	67,989	−0.26	1	8	NA	NA
1005	750,000	52,148	−0.11[c]	2	4	NA	NA
1006	500,000	22,876	0.46	3	6	NA	NA

[a] Mean viral load at 3 months postseronconversion.
[b] Values in bold indicate that the HLA type is known. NA, not available.
[c] CD4 count declined to <200 during the observation period.
[d] Mutant CTL epitopes matched patient HLA type.
[e] Only a single sample was available at 3 months postseroconversion.

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

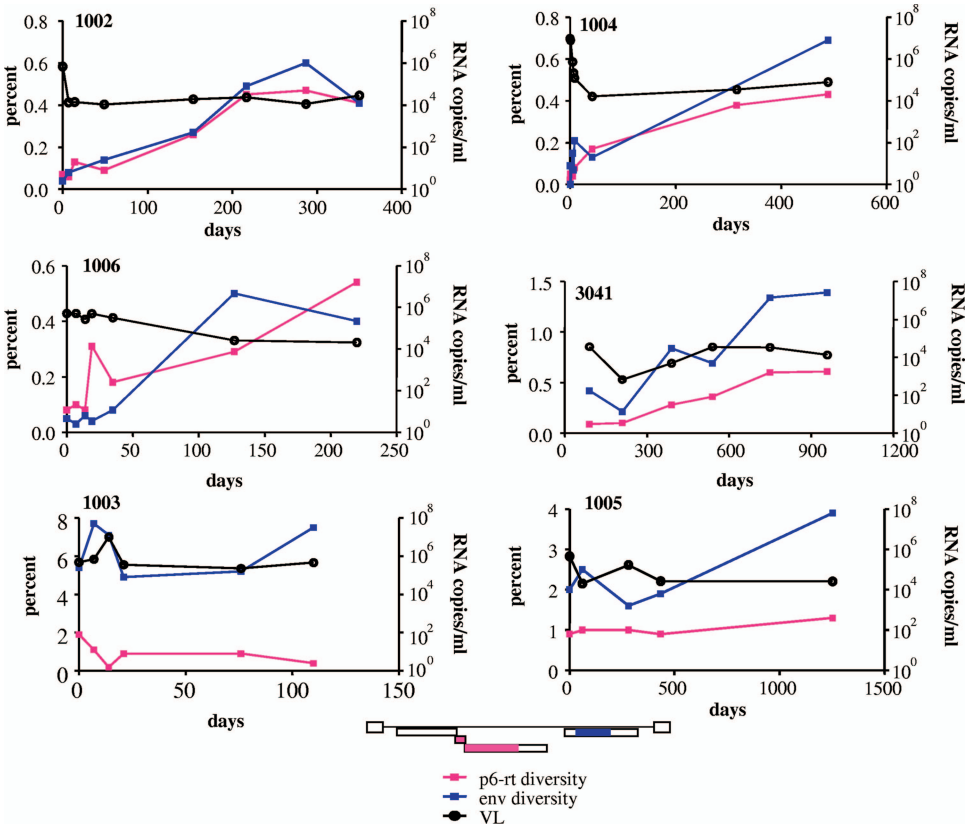

FIG. 6. Diversity in *p6-rt* and *env* in longitudinal samples from early HIV infections. The time after infection, the virus load, and the percent diversity are shown graphically for those patients for whom the first available sample was collected within 3 months of seroconversion.

DISCUSSION

We found that virus populations in acute and recent HIV infections in 12 of 14 patients had low initial genetic diversity, consistent with the transmission and/or outgrowth and dissemination of a single viral variant. Consistent with this conclusion, the slope of the best-fitting line resulting from a cross-sectional plot of the diversity versus time postseroconversion from these 12 patients was lower than the slope predicted from the neutral accumulation of mutations, as expected from the error rate of HIV replication. This result suggests that higher diversity in samples collected several months after infection is more likely to be due to early accumulation of mutations than to transmission of multiple variants. Included among the 12 was one patient from whom the first available sample was collected prior to seroconversion. The virus population in this patient was almost entirely monomorphic, supporting the conclusion that early homogeneity does not result from a genetic bottleneck at seroconversion. These results are in agreement with previous studies in which diversity was measured in *env* in patients prior to seroconversion (10, 18, 25). Our studies focused primarily on the *p6-rt* fragment, as it was plausible that early homogeneity seen in *env* in previous studies resulted from selection pressures imposed on the surface proteins during transmission. Such selection pressures are not imposed on PR and RT, and consequently, patients might have presented with multiple variants in *p6-rt* while carrying only single variants in *env*. Our results show that equal numbers of variants were present in both *p6-rt* and *env*, indicating that early homogeneity results from the outgrowth of a single HIV-1 virion and not only a specific *env* phenotype.

Analyses of nonsynonymous changes in single genome sequences compared to the consensus B sequence revealed that all 14 patients carried viruses with CTL-associated mutations in PR and RT in the first available sample. These mutations probably accumulated in previous hosts in the chain of transmission and are not likely to be artifacts resulting from the phylogenetic history of the virus, as proposed to explain other cases of apparent transmission of CTL mutations (3), for the following reasons. First, both the MRCA and the consensus of all the patient sequences were identical to the published consensus B sequence, with the exception of position 63 in PR and 83 in RT, both of which are known to be variable in the U.S. HIV population (19). Second, the pattern of mutations in each patient was unique. Third, again with the exception of the common polymorphisms at PR position 63 and RT position 83, no single mutation was found in the initial virus for a majority of patients. Although it is possible that some of these mutations might have been selected by CTL pressure early in infection, many of them were in epitopes not recognized by the patient's MHC and were most likely selected somewhere along the chain of transmission leading from the MRCA to each individual patient studied. The majority of the mutations in CTL epitopes were stable over the early years of infection, even in the apparent absence of selection for escape. In fact,

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

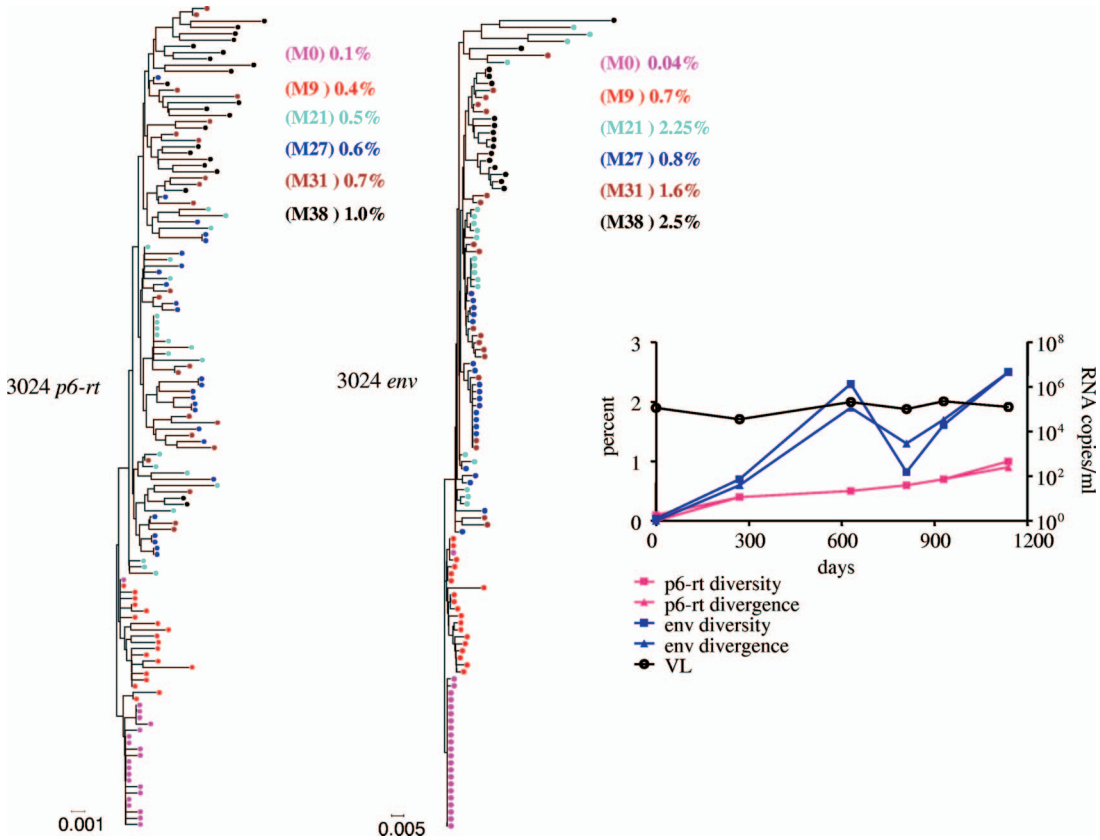

FIG. 7. Evolution in *p6-rt* and *env* in single-variant-infected patient 3024. The left panels show neighbor-joining phylogenetic trees of single genome sequences obtained from longitudinal samples for both *p6-rt* and *env* from this single-variant-infected patient. The trees are rooted on the sequences obtained from the first available sample. The time after infection, the virus load, and the percent diversity and divergence in each sample are also listed and shown graphically on the right.

only in the case of the superinfected patient (1001), in whom virus from the primary infection dropped to levels undetectable by allele-specific PCR (unpublished results), did we see the loss of transmitted CTL epitope mutations in early HIV infection. Furthermore, variants carrying mutations in CTL epitopes were preferentially selected over wild-type variants in patients infected with multiple HIV viruses from the same donor. These data indicate that viral variants carrying CTL-associated mutations are transmitted between hosts and that these CTL mutations do not readily revert to the consensus sequence in PR and RT. The lack of reversion of CTL mutations in PR and RT, two highly conserved regions of the HIV-1 genome (6), indicates that for the most part, such mutations are unlikely to strongly compromise the function of the viral enzymes or otherwise negatively affect the overall fitness of the virus. Consequently, although there is little net change in the consensus subtype B sequence, the majority of the CTL-associated mutations found in these recently infected patients are well represented as minority components of the HIV sequence database (19).

In this analysis, we compared each individual patient sequence with the consensus subgroup B sequence to investigate variation from circulating HIV-1 subtype B strains (Table 3). As a group, many of the sites varying from the consensus sequence in the recently infected patients were also sites that vary within the group itself (e.g., positions 36, 37, 63, 64, and 69). There is sufficient evidence based on the analysis of sequences comprising the consensus B group itself that CTL-associated mutations are accumulating in the population as a whole and contributing to the total diversity of HIV. However, identical CTL mutations are not typical between patients due to the diversity of HLA types in the population. As a result, specific mutations are not found in the majority and so are not represented in the overall subtype consensus B sequence. However, there were some variant sites in the recently infected patients that were significantly underrepresented in the consensus B group. For example, mutations at positions 39, 85, and 92 of protease were detected in the acute seroconverters, but these positions are invariant in the consensus B group. Similarly, the frequency of mutations at position 19 or 36 of protease was significantly higher in acute seroconverters than that found in the consensus B group ($P < 0.0004$ by chi-square analysis). It is not clear whether recent changes in transmitted viruses or selection following seroconversion is responsible for these strong differences from the consensus B sequence. However, their existence indicates that, if transmitted, reversion at these positions is likely to occur at a low rate.

Many of the CTL epitope mutations detected were found in anchor residues or in the epitopes' proximal three amino acids, primary sites for decreasing antigen presentation (5, 26). In a

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

TABLE 3. CTL mutations in protease

Codon in PR	Consensus B residue	Residue in virus from patient[a]:																	
		3024			3062			3021			3041			3036			3077		
		M0	M21	M38	M7	M16	M31	M9	M30	M59	M3	M18	M32	M6	M13	M22	M5	M22	M37
10	L																		
12	T	**E**	**E**	**E**				S	S	S/**P**									
13	I	V	V	V															
15	I	**V**	**V**	**V**/I										**V**	**V**	**V**			
16	G													E	E	E			
19	L													I	I/K	I/K	I	I	I
20	K																		
34	E																		
35	E	**D**	**D**	**D**/E	**D**/E	**D**/E	**D**/E												
36	M				I	I/T	I/L							L	L	L			
37	N	H	H/Q	H/P	S/N/D	S/N	S/N/D	S/N/A	S/N/T/H	S/N/A/Q				S	S/N	S	E	E	E
39	P				P/L	P/L	P							S	S	S			
41	R				R/K/G	R/K	K							K	K	K			
45	K																		
46	M																		
57	R										K	K	K						
62	I													V	V	V			
63	L	P	P	P	P	P	P	P/H	P/H/T	P/H	P	P	P	Q	Q	Q	P	P	P
64	I									I/L				V	V	V			
69	H																N	N	N
70	K																R	R	R
71	A																		
72	I																		
77	V																		
82	V																		
85	I																		
90	L																		
92	Q													K/R	K/R	K/R			
93	I										**L**	**L**	**L**						

[a] Residues in bold indicate previously defined CTL escape mutations (5, 25).
[b] Multiple-variant transmission or superinfection.

number of cases, there is functional evidence for CTL escape (Table 3), and it is likely that many nonsynonymous mutations in PR and RT that confer escape from the CTL response (26) were selected along the transmission chain prior to infection of the patients studied here. It is also possible that some of the CTL mutations were retained because they conferred a fitness advantage within the patient that was not previously known for that HLA allele. Although the HLA types were not known for all subjects and they were not tested for the ability to bind particular peptides, we found clear evidence that CTL-associated mutations were retained even in epitopes that would not likely be recognized by the patient's MHC type. Further studies are needed to determine which transmitted PR and RT mutations confer CTL resistance in patients with corresponding HLA types, which mutations are compensatory for fitness, and which, if any, are the result of random drift due to founder effects (3, 28).

Transmission of CTL-associated mutations has been reported in previous studies (2–4, 11, 13, 14, 20, 22–24, 32). However, there are conflicting data on the persistence and reversion of such mutations. Additionally, previous studies have been cross-sectional analyses, focused primarily on *gag*, *env*, or *nef*, or were based on findings for simian immunodeficiency virus rather than HIV. Only two studies have included longitudinal samples and *pro-pol* sequences in early infection. One was limited by the number of patients ($n = 1$) (24), and the other was limited by the method used (sequencing of genomic DNA) (23). To our knowledge, this is the most complete study of population structure, viral characteristics, including the transmission and persistence of CTL-associated mutations across several genes, and diversification in individuals infected with both single and multiple variants. Additionally, this study is the first to fully characterize the genetic profiles and diversity of PR and RT in frequent longitudinal samples spanning early HIV-1 infection. Our results show that although transmission of single HIV variants carrying CTL-associated mutations is most common, transmission of multiple variants from the same donor is not rare (2 of 14 patients) and results in high diversity, complex population structures, and different CTL profiles in acute infection. Similar frequencies of multiple transmission (4/12 sequences) were reported for a set of *env* sequences from Zambian patients recently infected by heterosexual contact with subgroup C HIV (34) and for a large cross-sectional set of subtype B *env* sequences studied by Keele et al. (18). Comparison of these studies supports the generality of the observation across subtypes, genes examined, geographical region, and mode of transmission. In the study by Keele et al., it was demonstrated that transmitted HIV variants contained functional and highly fusogenic Env proteins. Here we found that transmitted HIV variants also contain frequent mutations within CTL epitopes in PR and RT, that these transmitted mutations often persist for years after seroconver-

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

TABLE 3—*Continued*

Residue in virus from patient[a]:																	
3037			3088			1001[b]		1002		1003[b]		1004		1005[b]		1006	
M8	M16	M38	M6	M18	M23	D0	D138	D0	D287	D0	D110	D0	D489	D0	D1253	D0	D220
						I	L/I										
												A	A, T				
										I/V	I/V	V	V				
A/N/I	A/**V**/I	**V**/I															
I/T	I	I															
										K/R	K/R						
						D	E/D										
D	**D**	**D**						N	N/T								
S	S	S	I/T	I/T/V	I/M					M/I	M/I					I	I/V
			S/N/H	S/N/H/T	S/N			D	D					S	S	S	S
K	K	K	R/K	K	K					K	K			K	K		
R	R	R/K															
						I	M/I										
K	K	K								K/R	K/R						
			V	V	V	V	I/V									V	V
P	P	P	Q	Q	Q/P			P	P	P	P	P	P	P/L	P	P	P
						V	M/V										
																T	T,/A
								T	T								
										I/E	I/E						
						I	M/I							V/I	I		
																I	I
						V	I/V										
						M	L/M										
L	**L**	**L**				**L**	**L**							**L**	**L**	**L**	**L**

sion, and that new CTL epitope mutations accumulate against this background. Furthermore, acute infection after transmission of multiple variants can include variants with different CTL mutational profiles among the transmitted variants. Consequently, some transmitted variants may be selected preferentially by immune pressure from cytotoxic T cells. Indeed, in patient 1003, we saw a selective loss of variants that lacked CTL-associated mutations in CTL epitopes in PR, consistent with selection by CTL expressing the corresponding MHC (although the MHC type of this patient was not available).

In addition to identifying and characterizing variants in recently HIV-infected individuals, we also obtained frequent longitudinal samples to determine rates of diversification in *p6-rt* compared to those for *env*. We found that diversification occurs at similar rates in *p6-rt* and *env* in the first year of infection in single-variant-infected patients but subsequently increases at a higher rate in *env* at later time points. These differences most likely reflect both the structural constraints on PR and RT compared to Env and the stronger immune selection on Env due to the antibody response. More diversity was found in *env* than in *p6-rt* for multiple-variant-infected patients both after transmission and in longitudinal samples, reflecting the higher diversity found in this gene during chronic infection. Little change in diversity over the early months of HIV infection was found in multiple-variant-infected patients, with the exception of the rapid selection for variants with mutant CTL epitopes in PR and RT. This observation further supports the fitness of *pro-pol* variants carrying CTL-associated mutations after transmission to a new host.

Overall, our data show that HIV transmission in both single- and multiple-variant-infected patients frequently includes viral species carrying CTL-associated mutations, presumably selected for in prior hosts along the transmission chain, and that these mutations do not revert easily regardless of the MHC type of the recipient. Indeed, our longitudinal data show a slow accumulation of new amino acid changes in PR and RT CTL epitopes in response to the patients' MHC type even in the background of the transmitted mutations. These results suggest that evaluation of CTL-inducing vaccines should include challenge by HIV variants with clinically relevant sequence variations in CTL epitopes to ensure efficacy against newly acquired HIV strains. Further studies are needed to determine if the transmission of CTL-associated mutations affects the breadth of the CTL response in newly infected patients. Although we found a trend between the high frequencies of CTL-associated mutations in PR and RT and progression to $CD4^+$ T-cell counts below 200 and a significant correlation between the transmission of multiple HIV variants and lower $CD4^+$ T-cell counts, further studies are needed to determine the clinical impact of transmission of CTL-associated mutations in HIV infection as well as the impact of single- versus multiple-variant transmission events on disease progression.

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

ACKNOWLEDGMENTS

This work was supported in part by the intramural research program of the NIH, NCI. This work was also supported in part by a grant (AI 41532) to J.B.M. through the AIEDRP Program of the National Institute for Allergy and Infectious Diseases. E.D. was supported by grants AI043638, M01-RR004425, and CH05-SD-607-005. J.W.M. was supported by SAIC contract 25XS119. J.M.C. was a Research Professor of the American Cancer Society, with support from the George Kirby Foundation.

We thank Linda Apuzzo of Johns Hopkins Bloomberg School of Public Health for aid in obtaining clinical specimens and sample information, Christopher Kearney for aid in sample preparation, and Valerie Boltz and Ann Wiegand of NCI-Frederick for many helpful conversations. We are especially indebted to our patient volunteers, without whom this study would not have been possible.

REFERENCES

1. **Altfeld, M., E. S. Rosenberg, R. Shankarappa, J. S. Mukherjee, F. M. Hecht, R. L. Eldridge, M. M. Addo, S. H. Poon, M. N. Phillips, G. K. Robbins, P. E. Sax, S. Boswell, J. O. Kahn, C. Brander, P. J. Goulder, J. A. Levy, J. I. Mullins, and B. D. Walker.** 2001. Cellular immune responses and viral diversity in individuals treated during acute and early HIV-1 infection. J. Exp. Med. **193:**169–180.
2. **Bernardin, F., D. Kong, L. Peddada, L. A. Baxter-Lowe, and E. Delwart.** 2005. Human immunodeficiency virus mutations during the first month of infection are preferentially found in known cytotoxic T-lymphocyte epitopes. J. Virol. **79:**11523–11528.
3. **Bhattacharya, T., M. Daniels, D. Heckerman, B. Foley, N. Frahm, C. Kadie, J. Carlson, K. Yusim, B. McMahon, B. Gaschen, S. Mallal, J. I. Mullins, D. C. Nickle, J. Herbeck, C. Rousseau, G. H. Learn, T. Miura, C. Brander, B. Walker, and B. Korber.** 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. Science **315:**1583–1586.
4. **Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, and G. M. Shaw.** 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat. Med. **3:**205–211.
5. **Brumme, Z. L., C. J. Brumme, D. Heckerman, B. T. Korber, M. Daniels, J. Carlson, C. Kadie, T. Bhattacharya, C. Chui, J. Szinger, T. Mo, R. S. Hogg, J. S. Montaner, N. Frahm, C. Brander, B. D. Walker, and P. R. Harrigan.** 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. PLoS Pathog. **3:**e94.
6. **Cornelissen, M., R. van den Burg, F. Zorgdrager, V. Lukashov, and J. Goudsmit.** 1997. *pol* gene diversity of five human immunodeficiency virus type 1 subtypes: evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. J. Virol. **71:**6348–6358.
7. **da Silva, J., and A. L. Hughes.** 1999. Molecular phylogenetic evidence of cytotoxic T lymphocyte (CTL) selection on human immunodeficiency virus type 1 (HIV-1). Mol. Biol. Evol. **16:**1420–1422.
8. **Delwart, E. L., H. W. Sheppard, B. D. Walker, J. Goudsmit, and J. I. Mullins.** 1994. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. J. Virol. **68:**6672–6683.
9. **Fernandez, C. S., I. Stratov, R. De Rose, K. Walsh, C. J. Dale, M. Z. Smith, M. B. Agy, S. L. Hu, K. Krebs, D. I. Watkins, D. H. O'Connor, M. P. Davenport, and S. J. Kent.** 2005. Rapid viral escape at an immunodominant simian-human immunodeficiency virus cytotoxic T-lymphocyte epitope exacts a dramatic fitness cost. J. Virol. **79:**5721–5731.
10. **Freel, S. A., S. A. Fiscus, C. D. Pilcher, P. Menezes, J. Giner, E. Patrick, J. L. Lennox, C. B. Hicks, J. J. Eron, Jr., and D. C. Shugars.** 2003. Envelope diversity, coreceptor usage and syncytium-inducing phenotype of HIV-1 variants in saliva and blood during primary infection. AIDS **17:**2025–2033.
11. **Friedrich, T. C., E. J. Dodds, L. J. Yant, L. Vojnov, R. Rudersdorf, C. Cullen, D. T. Evans, R. C. Desrosiers, B. R. Mothe, J. Sidney, A. Sette, K. Kunstman, S. Wolinsky, M. Piatak, J. Lifson, A. L. Hughes, N. Wilson, D. H. O'Connor, and D. I. Watkins.** 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. Nat. Med. **10:**275–281.
12. **Furuta, Y., T. Bergstrom, G. Norkrans, and P. Horal.** 1994. HIV type 1 V3 sequence diversity in contact-traced Swedish couples at the time of sexual transmission. AIDS Res. Hum. Retrovir. **10:**1187–1189.
13. **Goulder, P. J., C. Brander, Y. Tang, C. Tremblay, R. A. Colbert, M. M. Addo, E. S. Rosenberg, T. Nguyen, R. Allen, A. Trocha, M. Altfeld, S. He, M. Bunce, R. Funkhouser, S. I. Pelton, S. K. Burchett, K. McIntosh, B. T. Korber, and B. D. Walker.** 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. Nature **412:**334–338.
14. **Goulder, P. J., C. Pasquier, E. C. Holmes, B. Liang, Y. Tang, J. Izopet, K. Saune, E. S. Rosenberg, S. K. Burchett, K. McIntosh, M. Barnardo, M. Bunce, B. D. Walker, C. Brander, and R. E. Phillips.** 2001. Mother-to-child transmission of HIV infection and CTL escape through HLA-A2-SLYNTVATL epitope sequence variation. Immunol. Lett. **79:**109–116.
15. **Guillon, C., K. Stankovic, Y. Ataman-Onal, F. Biron, and B. Verrier.** 2006. Evidence for CTL-mediated selection of Tat and Rev mutants after the onset of the asymptomatic period during HIV type 1 infection. AIDS Res. Hum. Retrovir. **22:**1283–1292.
16. **Karlsson, A. C., A. K. Iversen, J. M. Chapman, T. de Oliviera, G. Spotts, A. J. McMichael, M. P. Davenport, F. M. Hecht, and D. F. Nixon.** 2007. Sequential broadening of CTL responses in early HIV-1 infection is associated with viral escape. PLoS ONE **2:**e225.
17. **Kearney, M., S. Palmer, F. Maldarelli, W. Shao, M. A. Polis, J. Mican, D. Rock-Kress, J. B. Margolick, J. M. Coffin, and J. W. Mellors.** 2008. Frequent polymorphism at drug resistance sites in HIV-1 protease and reverse transcriptase. AIDS **22:**497–501.
18. **Keele, B. F., E. E. Giorgi, J. F. Salazar-Gonzalez, J. M. Decker, K. T. Pham, M. G. Salazar, C. Sun, T. Grayson, S. Wang, H. Li, X. Wei, C. Jiang, J. L. Kirchherr, F. Gao, J. A. Anderson, L. H. Ping, R. Swanstrom, G. D. Tomaras, W. A. Blattner, P. A. Goepfert, J. M. Kilby, M. S. Saag, E. L. Delwart, M. P. Busch, M. S. Cohen, D. C. Montefiori, B. F. Haynes, B. Gaschen, G. S. Athreya, H. Y. Lee, N. Wood, C. Seoighe, A. S. Perelson, T. Bhattacharya, B. T. Korber, B. H. Hahn, and G. M. Shaw.** 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. USA **105:**7552–7557.
19. **Korber, B. T., P. Marx, S. Wolinsky, B. H. Hahn, J. W. Mellors, F. McCutchan, and T. Leitner.** 2007. Los Alamos HIV database. Los Alamos National Laboratory, Los Alamos, NM.
20. **Leslie, A., D. Kavanagh, I. Honeyborne, K. Pfafferott, C. Edwards, T. Pillay, L. Hilton, C. Thobakgale, D. Ramduth, R. Draenert, S. Le Gall, G. Luzzi, A. Edwards, C. Brander, A. K. Sewell, S. Moore, J. Mullins, C. Moore, S. Mallal, N. Bhardwaj, K. Yusim, R. Phillips, P. Klenerman, B. Korber, P. Kiepiela, B. Walker, and P. Goulder.** 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. J. Exp. Med. **201:**891–902.
21. **Leslie, A., D. A. Price, P. Mkhize, K. Bishop, A. Rathod, C. Day, H. Crawford, I. Honeyborne, T. E. Asher, G. Luzzi, A. Edwards, C. M. Rousseau, J. I. Mullins, G. Tudor-Williams, V. Novelli, C. Brander, D. C. Douek, P. Kiepiela, B. D. Walker, and P. J. Goulder.** 2006. Differential selection pressure exerted on HIV by CTL targeting identical epitopes but restricted by distinct HLA alleles from the same HLA supertype. J. Immunol. **177:**4699–4708.
22. **Leslie, A. J., K. J. Pfafferott, P. Chetty, R. Draenert, M. M. Addo, M. Feeney, Y. Tang, E. C. Holmes, T. Allen, J. G. Prado, M. Altfeld, C. Brander, C. Dixon, D. Ramduth, P. Jeena, S. A. Thomas, A. St. John, T. A. Roach, B. Kupfer, G. Luzzi, A. Edwards, G. Taylor, H. Lyall, G. Tudor-Williams, V. Novelli, J. Martinez-Picado, P. Kiepiela, B. D. Walker, and P. J. Goulder.** 2004. HIV evolution: CTL escape mutation and reversion after transmission. Nat. Med. **10:**282–289.
23. **Li, B., A. D. Gladden, M. Altfeld, J. M. Kaldor, D. A. Cooper, A. D. Kelleher, and T. M. Allen.** 2007. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. J. Virol. **81:**193–201.
24. **Liu, Y., J. McNevin, J. Cao, H. Zhao, I. Genowati, K. Wong, S. McLaughlin, M. D. McSweyn, K. Diem, C. E. Stevens, J. Maenza, H. He, D. C. Nickle, D. Shriner, S. E. Holte, A. C. Collier, L. Corey, M. J. McElrath, and J. I. Mullins.** 2006. Selection on the human immunodeficiency virus type 1 proteome following primary infection. J. Virol. **80:**9519–9529.
25. **Long, E. M., H. L. Martin, Jr., J. K. Kreiss, S. M. Rainwater, L. Lavreys, D. J. Jackson, J. Rakwar, K. Mandaliya, and J. Overbaugh.** 2000. Gender differences in HIV-1 diversity at time of infection. Nat. Med. **6:**71–75.
26. **McMichael, A. J., and R. E. Phillips.** 1997. Escape of human immunodeficiency virus from immune control. Annu. Rev. Immunol. **15:**271–296.
27. **McNearney, T., Z. Hornickova, B. Kloster, A. Birdwell, G. A. Storch, S. H. Polmar, M. Arens, and L. Ratner.** 1993. Evolution of sequence divergence among human immunodeficiency virus type 1 isolates derived from a blood donor and a recipient. Pediatr. Res. **33:**36–42.
28. **Moore, C. B., M. John, I. R. James, F. T. Christiansen, C. S. Witt, and S. A. Mallal.** 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science **296:**1439–1443.
29. **Palmer, S., M. Kearney, F. Maldarelli, E. K. Halvas, C. J. Bixby, H. Bazmi, D. Rock, J. Falloon, R. T. Davey, Jr., R. L. Dewar, J. A. Metcalf, S. Hammer, J. W. Mellors, and J. M. Coffin.** 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J. Clin. Microbiol. **43:**406–413.
30. **Palmer, S., M. Kearney, V. Boltz, F. Maldarelli, G. Achaz, J. Mellors, E. Daar, and J. Coffin.** 2003. Population genetics in HIV-1 super-infection, abstr. 62. XIIth International HIV Drug Resistance Workshop: basic principles and clinical implications.
31. **Poss, M., A. G. Rodrigo, J. J. Gosink, G. H. Learn, D. de Vange Panteleeff, H. L. Martin, Jr., J. Bwayo, J. K. Kreiss, and J. Overbaugh.** 1998. Evolution of envelope sequences from the genital tract and peripheral blood of women

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010

infected with clade A human immunodeficiency virus type 1. J. Virol. **72:** 8240–8251.

32. **Price, D. A., C. A. O'Callaghan, J. A. Whelan, P. J. Easterbrook, and R. E. Phillips.** 1999. Cytotoxic T lymphocytes and viral evolution in primary HIV-1 infection. Clin. Sci. (London) **97:**707–718.
33. **Ritola, K., C. D. Pilcher, S. A. Fiscus, N. G. Hoffman, J. A. Nelson, K. M. Kitrinos, C. B. Hicks, J. J. Eron, Jr., and R. Swanstrom.** 2004. Multiple V1/V2 *env* variants are frequently present during primary infection with human immunodeficiency virus type 1. J. Virol. **78:**11208–11218.
34. **Salazar-Gonzalez, J. F., E. Bailes, K. T. Pham, M. G. Salazar, M. B. Guffey, B. F. Keele, C. A. Derdeyn, P. Farmer, E. Hunter, S. Allen, O. Manigart, J. Mulenga, J. A. Anderson, R. Swanstrom, B. F. Haynes, G. S. Athreya, B. T. Korber, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 2008. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. J. Virol. **82:** 3952–3970.
35. **Sanders-Buell, E. S. M., and F. E. McCutchan.** 1996. Sequencing primers for HIV-1, p. III-15–III-21. *In* G. Myers (ed.), Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, NM.
36. **Schutten, M., C. A. van Baalen, C. Guillon, R. C. Huisman, P. H. Boers, K. Sintnicolaas, R. A. Gruters, and A. D. Osterhaus.** 2001. Macrophage tropism of human immunodeficiency virus type 1 facilitates in vivo escape from cytotoxic T-lymphocyte pressure. J. Virol. **75:**2706–2709.
37. **Shankarappa, R., J. B. Margolick, S. J. Gange, A. G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C. R. Rinaldo, G. H. Learn, X. He, X. L. Huang, and J. I. Mullins.** 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. J. Virol. **73:**10489–10502.
38. **Wolfs, T. F., G. Zwart, M. Bakker, and J. Goudsmit.** 1992. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology **189:**103–110.
39. **Yang, O. O., E. S. Daar, B. D. Jamieson, A. Balamurugan, D. M. Smith, J. A. Pitt, C. J. Petropoulos, D. D. Richman, S. J. Little, and A. J. Brown.** 2005. Human immunodeficiency virus type 1 clade B superinfection: evidence for differential immune containment of distinct clade B strains. J. Virol. **79:**860–868.

Downloaded from jvi.asm.org at Scientific Library, NCI-FCRDC on May 14, 2010