

Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline

RICHARD B. MARKHAM*†‡, WEI-CINN WANG*, ANTON E. WEISSTEIN§, ZHE WANG*, ALVARO MUNOZ¶, ALAN TEMPLETON§, JOSEPH MARGOLICK*, DAVID VLAHOV¶, THOMAS QUINN||, HOMAYOON FARZADEGAN¶, AND XIAO-FANG YU*†

*Department of Molecular Microbiology and Immunology and †Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health, Baltimore, MD 21205; §Department of Biology, Washington University, St. Louis, MO 63130; and ¶National Institute for Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD 20892, and Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD 21205

Communicated by John W. Littlefield, Johns Hopkins University School of Medicine, Baltimore, MD, August 10, 1998 (received for review March 23, 1998)

ABSTRACT Evolution of HIV-1 *env* sequences was studied in 15 seroconverting injection drug users selected for differences in the extent of CD4 T cell decline. The rates of increase of either sequence diversity at a given visit or divergence from the first seropositive visit were both higher in progressors than in non-progressors. Viral evolution in individuals with rapid or moderate disease progression showed selection favoring nonsynonymous mutations, while nonprogressors with low viral loads selected against the nonsynonymous mutations that might have resulted in viruses with higher levels of replication. For 10 of the 15 subjects no single variant predominated over time. Evolution away from a dominant variant was followed frequently at a later time point by return to dominance of strains closely related to that variant. The observed evolutionary pattern is consistent with either selection against only the predominant virus or independent evolution occurring in different environments within the host. Differences in the level to which CD4 T cells fall in a given time period reflect not only quantitative differences in accumulation of mutations, but differences in the types of mutations that provide the best adaptation to the host environment.

The high mutation and replication rates of HIV-1 (1–3) permit rapid adaptation to changes in the host environment. In a stable environment the “best fit” virus would predominate rapidly and all subsequent mutations would be minimally represented in the gene pool (4). Samples of different viruses in this environment would show a generally uniform genetic makeup with minimal representation of variants arising because of the appearance of nearly neutral mutations. This pattern of evolution has been observed in other systems in which environmental changes are accommodated rapidly by organisms with high mutational capability, as seen, for example, with *Escherichia coli* grown in chemostats (5, 6).

An unstable host environment could have variable effects on the genetic composition of the virus pool. Such instability could be generated, for example, by a dynamic host immune response or, in the case of HIV-1, by differential display of coreceptors (7–15). If the destabilizing force selected powerfully, but indiscriminately, against the broad array of existing variants, diversity would be reduced to that present in the few surviving variants, which likely would represent those variants that were originally most numerous (16, 17). However, if selective forces such as the immune response target only the most abundant viral variant in a genetically heterogeneous population (frequency-dependent selection), there could be significant reduction in the viral load

without a corresponding reduction in the genetic diversity present in the viral population, due to the diversity contributed by the less frequent viral strains. As these diverse minority populations continue to mutate, diversity would continue to expand. Such continuing expansion of diversity eventually could give rise to viral variants that have exceeded the capacity of a selective force such as the immune response to control virus infection, as has been proposed by Nowak *et al.* (18–21). Thus, examination of the patterns of diversity during HIV-1 evolution can provide insights into the type and efficiency of selection forces influencing viral evolution and into how the virus is adapting to those forces.

Previous studies of HIV-1 genetic evolution examined very small cohorts of infected subjects (22), characterized HIV-1 genetic evolution using techniques that did not involve direct examination of sequence patterns (23, 24), or analyzed a very limited number of time points in each subject (25, 26). In the current work, HIV-1 evolution was analyzed in 15 subjects followed from seroconversion at frequent intervals over a period of up to 4 years. This analysis demonstrates that different patterns of selection are observed between nonprogressor and moderately or rapidly progressing subjects and that, contrary to previous reports, the attainment of higher levels of genetic diversity is most frequently associated with more rapid CD4 T cell decline.

METHODS

The Study Population. The 15 participants were selected from a cohort of injection drug users participating in the AIDS Linked to Intravenous Experiences (ALIVE) study in Baltimore, MD (27). This ongoing cohort study follows infected or at-risk injection drug users at 6-month intervals (visits), at which time blood is obtained for virologic and immunologic studies. The subset of individuals selected for this study were followed from the point of HIV-1 seroconversion and had attained different levels of CD4 T cells. Rapid progressors were defined as having attained a level of fewer than 200 CD4 T cells within 2 years of seroconversion; moderate progressors' CD4 T cell levels declined to 200–650 during the 4-year period of observation, and nonprogressors maintained CD4 T cell levels above 650 throughout the observation period.

Sequencing of HIV-1 *env* Genes. Nested PCR was used to amplify a 285-bp region of the *env* gene from peripheral blood mononuclear cells (PBMC). Several studies (28–30) have indicated that the preponderance of viral DNA is obtained from recently infected, unactivated PBMC, which constitute the vast majority of PBMC in human subjects. Such DNA has not yet

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9512568-7\$2.00/0
PNAS is available online at www.pnas.org.

Abbreviation: PBMC, peripheral blood mononuclear cells.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF016760–AF016825 and AF089109–AF089708).

†R.B.M. and X.-F.Y. contributed equally to this work.

‡To whom reprint requests should be addressed. e-mail: rmarkham@jhsph.edu.

integrated into the host genome, is unstable, and is capable of persisting only for several days in unactivated T cells (31). Because of the recent infection of the cells from which the DNA is derived, it should be closely related to the RNA in viruses concurrently circulating in the plasma.

The external *env* primers used were 5'-GTCAGCACAGTCAATGTACACATG (nucleotides 6947–6971, based on sequence numbering in the molecular clone HXB2) and 5'-AATTACAGTAGAAAAATTCCCCTC (nucleotides 7359–7382 of HXB2) and the nested primers were 5'-CGGGATCC-TGTTAAATGGCAGTCTAGCAGAAG (nucleotides 7002–7025 of HXB2) and 5'-CGGAATTCCTGTTACAATTC-TGGGTCCCCTCC (nucleotides 7319–7342 of HXB2). The underlined portion of the primers are, respectively, *Bam*HI or *Eco*RI restriction sites included in the nested primers for cloning purposes. Both first- and second-round PCR were run for 2 min at 95°, followed by 35 cycles of 94° for 30 sec, 60° for 30 sec, and 72° for 45 sec. At the end of the 35 cycles the samples were held at 72° for 10 min before returning to the 4° temperature at which they were held until subjected to further analysis. The amplified sequences from the nested PCR then were cloned into pUC19 using standard methods (32) and sequenced by using the Sanger chain termination method (33) with the appropriate nested primers described above. Limiting dilution single-round PCR was used initially to screen for input viral DNA copy number, and five samples that had PCR product detected only at the lowest cellular dilution were subjected to second-round PCR, as the samples used for sequencing had been. Even these samples, which initially could be detected at the lowest dilutions had, by this analysis, greater than 125 input copies of viral DNA for the initial round of PCR amplification. Based on this analysis it is likely that most or all of the clones generated as a result of PCR amplification were derived from a unique viral genome template (34). The sequences analyzed for this study may be obtained through GenBank (accession numbers AF016760–AF016825 and AF089109–AF089708).

Plasma Viral Load. Plasma viral load was determined by reverse transcription-PCR (Roche Molecular Systems, Branchburg, NJ).

Generation of Phylogenetic Trees. Trees were constructed by using the MEGA computer package (35) employing the neighbor-joining algorithm and the Tamura-Nei distance measure, which corrects for base composition and transition/transversion bias. Taxon labels indicate the time at which each strain was isolated and the number of identical replicates sampled; thus, for sequence S1V2–8 (2), two identical repli-

cates were obtained of clone 8 from subject S1's second visit. Note that data are lacking from individual visits of some subjects. Taxa are colored according to the time point at which they were observed: red (V1), orange (V2), green (V3), light blue (V4), dark blue (V5), purple (V6), brown (V7), gray (V8), and black (V9). Intermediate sequences were assigned the same color as their earliest-observed descendent sequence. A phylogenetic tree of clones from the different subjects (not shown) demonstrated the independent segregation of those clones, with the exception of clones from two subjects who were known epidemiologically to be related (S1 and S2).

Correlation Analysis. To determine at a given time point the correlation between genetic diversity or mutational divergence and CD4 T cell count 1 year later, the units of analysis were defined by pairs of visits within individuals. Seventy-six time points were available from the 15 subjects at which sequence data and CD4 T cell count 1 year later were known. Depending on which analysis was being performed, the term X_0 can be used to indicate the value of either diversity (π) or percent mutational divergence. Y_0 was assigned for the value of the CD4 T cell count at the visit at which X_0 was determined, and Y_1 represented the CD4 T cell count in that subject 1 year later. The analysis was then stratified by whether Y_0 was in one of each of the following categories of CD4 T cell count: <350, 350–499, 500–699, and >700. Stratification by the CD4 T cell count (Y_0) permits the determination of how different values of X_0 (either diversity or divergence) are related to the different CD4 T cell counts 1 year later (Y_1) in individuals that start at similar CD4 T cell counts.

Determination of dS/dN Ratios. Once the initial consensus sequence for each subject had been computed, it was compared with each subsequently observed strain. Each difference between the two strains was classified as synonymous or nonsynonymous. This was adjusted by the number of sites at which each class of mutation could occur, thereby correcting for the inherent bias toward nonsynonymous changes exhibited by totally random mutations. This procedure used the Jukes-Cantor correction for multiple hits, which would otherwise bias these estimates toward smaller values. The resulting values of dS and dN (and their ratio, dS/dN) were then averaged over all the strains observed at a particular visit, to remove possible bias because of unequal sample sizes at different visits. Because these values exhibited a skewed distribution, a median value was used for this average. These summary data formed the basis of all subsequent analysis.

Examination of Source of Greater Initial Visit Diversity in Subjects 9 and 15. Because of the high genetic variation observed at the first visit in subjects 9 and 15, we studied whether these

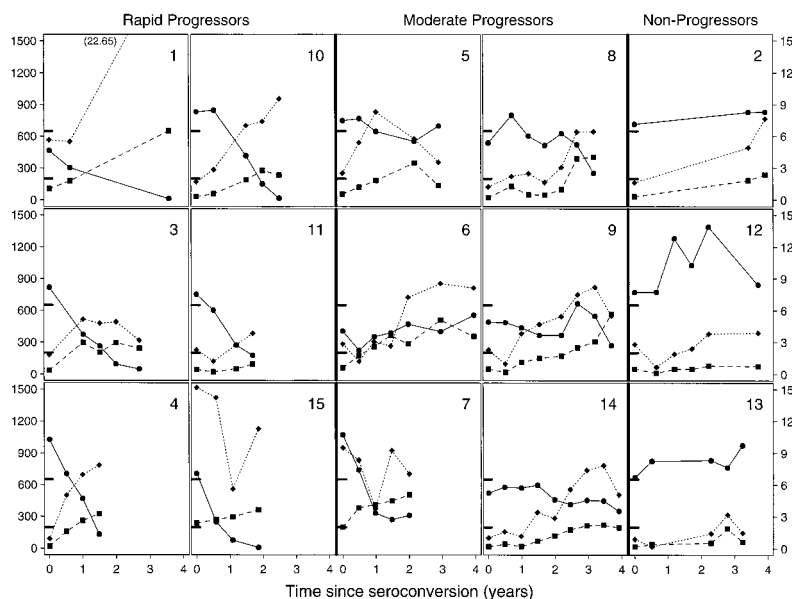


FIG. 1. CD4 T cell trajectory, diversity, and divergence over time since first seropositive visit ($t = 0$) in each of the 15 subjects. CD4 T cell counts (—●—) are shown on the left vertical axis. The solid dashes on the left vertical axis indicate the 200 and 650 CD4 T cell count levels that were used to distinguish the different progressor groups. Diversity (---◆---) and divergence (---■---) values are shown on the right vertical axis. For diversity, axis values indicate the mean of the nucleotide differences between intravisit clones. For divergence, the values indicate the median percent of nucleotides per clone that diverged from the original postseroconversion consensus sequence. For Subject 1, the diversity value at 3.57 years postseroconversion was 22.65 nt differences, which is offscale. Subject identification numbers are given in the upper right corner of each graph.

Table 1. Summary data on 15 seroconverters

| Subject | No. of observations | Baseline | | Annual rate of CD4 T cell decline | Slope of change in intraviral nucleotide differences per clone per year | Slope of divergence (% nucleotides mutated from baseline consensus sequence per year) | Median dS/dN | |
|----------------------------|---------------------|----------|---|-----------------------------------|---|---|--------------|-----|
| | | CD4 | Median intraviral nucleotide differences among clones | | | | | |
| Rapid Progressor | | | | | | | | |
| Subject 4 | 4 | 1,028 | 0.90 | 6.8 | -593 | 4.64 | 2.09 | 0.0 |
| Subject 10 | 5 | 833 | 1.71 | 99.3 | -363 | 3.16 | 1.00 | 0.2 |
| Subject 11 | 4 | 753 | 2.27 | 62.2 | -363 | 1.11 | 0.32 | 0.0 |
| Subject 15 | 4 | 707 | 15.16 | 171.0 | -362 | -2.94 | 0.68 | 0.7 |
| Subject 3 | 5 | 819 | 1.82 | 302.5 | -294 | 0.53 | 0.74 | 1.0 |
| Subject 1 | 3 | 464 | 5.64 | 307.6 | -117 | 5.10 | 1.55 | 0.3 |
| Moderate Progressor | | | | | | | | |
| Subject 7 | 5 | 1,072 | 2.27 | 317.6 | -392 | -0.79 | 1.35 | 1.3 |
| Subject 8 | 7 | 538 | 1.24 | 209.0 | -92 | 1.68 | 1.16 | 0.5 |
| Subject 14 | 9 | 523 | 1.00 | 50.9 | -51 | 1.69 | 0.60 | 0.0 |
| Subject 5 | 5 | 749 | 2.50 | 260.6 | -41 | 0.06 | 0.50 | 1.4 |
| Subject 9 | 8 | 489 | 9.49 | 265.0 | -11 | 1.58 | 1.21 | 0.0 |
| Subject 6 | 7 | 405 | 2.82 | 321.4 | 52 | 1.92 | 0.82 | 0.4 |
| Nonprogressor | | | | | | | | |
| Subject 2 | 5 | 715 | 1.64 | 21.6 | 30 | 1.32 | 0.49 | 1.8 |
| Subject 12 | 6 | 772 | 2.80 | 5.1 | 44 | 0.62 | 0.13 | 0.9 |
| Subject 13 | 5 | 671 | 0.87 | 1.7 | 53 | 0.53 | 0.28 | 3.5 |

Annual changes in CD4, intraviral nucleotide diversity, and percent nucleotide divergence from the first viruses sequenced after seroconversion reflect slopes of regression lines between individual visits. As slopes of CD4 T cell decline were quite variable between visits in the same subject, progressor categorization of subjects was based on the lowest level of CD4 T cell counts attained during the period of observation. Although subject 7 had a 392/year CD4 T cell decline, his CD4 T cell level never fell below 200 and therefore he was included in the moderate progressor group. His movement to the rapid progressor group would not have altered the statistical support for any of the conclusions reached.

Individuals might have been infected with two different viruses. Three different phylogenetic trees were constructed including the clones from the first visit of subject 9 and subject 15 and between 150–200 randomly selected clones from other individuals from the study. We then determined whether the viruses from the first visits of subjects 9 or 15 segregated as independent or as monophyletic viruses. They clearly grouped as monophyletic (data not shown).

In addition, reexamination of the preseroconversion history of these subjects confirmed that they were HIV-1 seronegative

up to 7 months before the first visit at which sequences were analyzed. It was also determined that the exclusion of all possible recombination events in subject 15 would still leave that subject with a π value of 11.4, considerably above the previously reported values for a 6-month postseroconversion visit. Exclusion of subject 15 did not change the conclusions of the dS/dN analysis.

Comparison of the Rate of Change of Divergence and Diversity. For each individual we fit a regression line of divergence/diversity over time and summarize it with the slope, β_i , for the i th

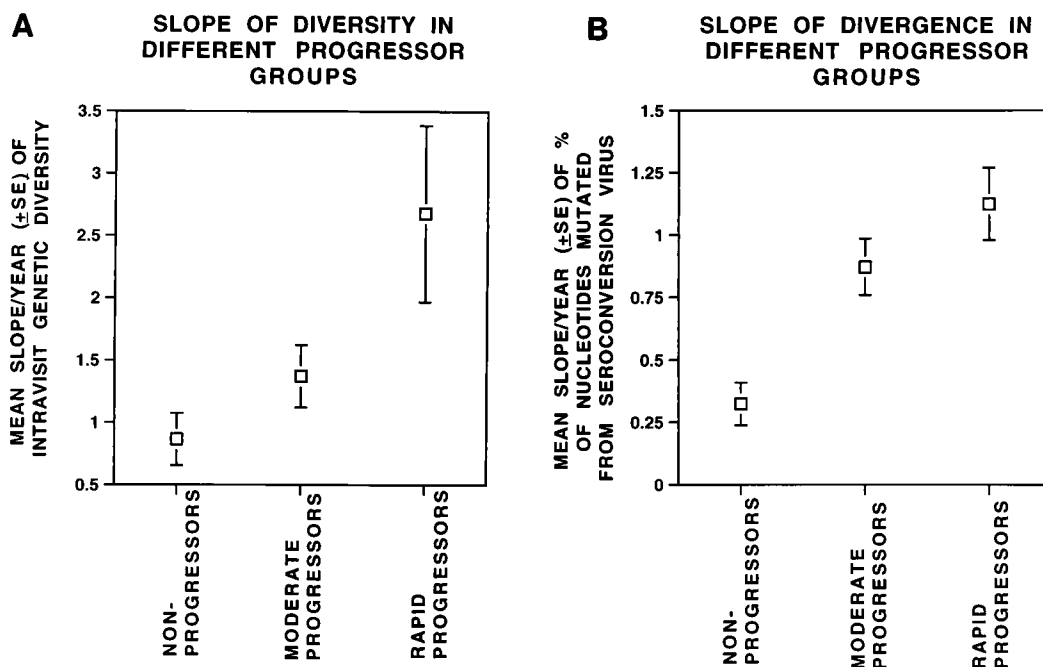


FIG. 2. Comparison among different progressor groups of the mean slope per year (+SE) of intraviral viral genetic diversity (A) and the percent of nucleotides that diverged from the original postseroconversion consensus sequence (B). Mean slope for each group is obtained from a weighted average incorporating the number of observations and the precision of the slopes of each subject shown in Table 1.

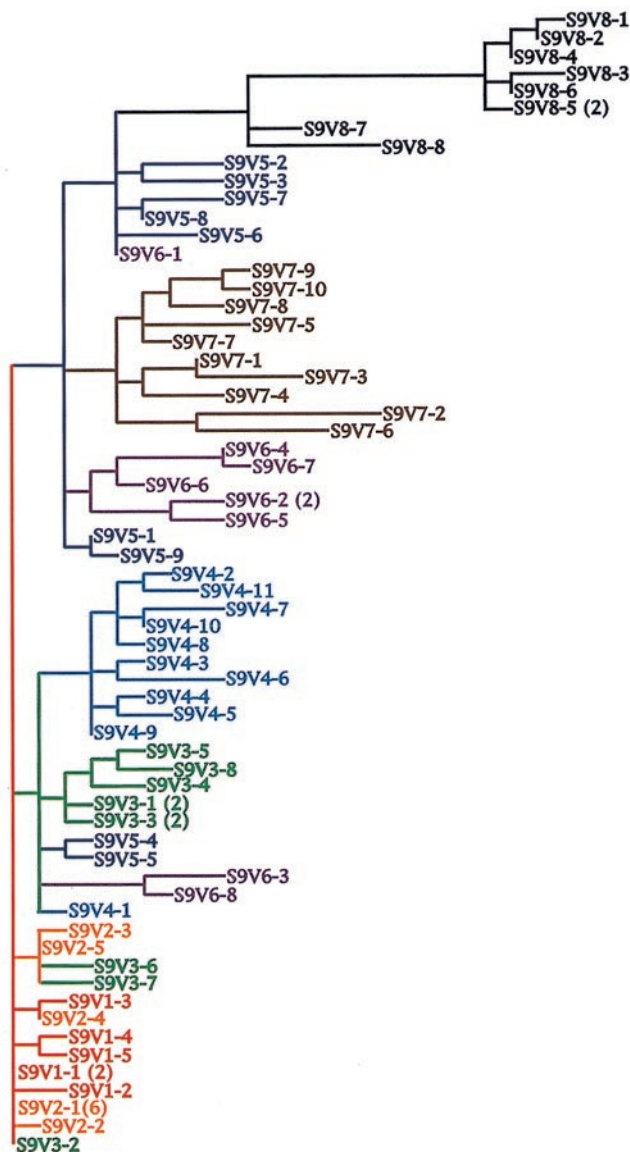


FIG. 3. Phylogenetic tree from subject 9. The horizontal distance at the bottom of the tree between S9V2-1 and S9V2-2 reflects a single mutation.

individual. The averages of the slopes for each of the three groups according to progression as measured by the decline of CD4 T cell counts were compared by the use of random effects models (36).

RESULTS

Patterns of CD4 decline were quite variable among the 15 subjects (Fig. 1), with median annual changes in the subjects' CD4 T cell number ranging from an increase of 53 cells per year to a decrease of 593 cells per year (Table 1). Serum viral load data were available for all subjects from one of the first three visits and ranged from 1,702 to 321,443 copies of viral genomic RNA/ml (Table 1). While the nonprogressor group generally had low viral loads at these early time points compared with the combined progressor groups ($P < 0.02$, ANOVA), the limited viral load data did not distinguish the moderate from the rapid progressors.

Genetic sequence analysis focused on the viral *env* region around the third hypervariable (V3) domain. This region was analyzed because it is an important site of host-virus interaction and is known to tolerate frequent mutations (37-45). Between 6 and 21 *env* sequences were amplified by PCR from each individual visit, yielding a total of 873 clones that were sequenced and analyzed.

Changes in HIV-1 sequences over time were quantified by two parameters: (i) the genetic diversity at each visit (termed π , Fig. 1), defined as the mean number of nucleotide differences between any two clones from that visit, and (ii) divergence, quantified as the median percentage of nucleotides per clone at a given visit that differed from those present in the consensus *env* sequence from the first visit of that subject (Fig. 1). The rate of change in the median diversity over time ranged from -2.94 to 5.10 nt per clone per year and the rate of change in divergence from 0.13% to 2.09% of the nucleotides per clone per year (Table 1).

Viruses from the initial visits of 13 of the 15 subjects were quite homogeneous (Table 1), consistent with previous reports (46-48). Virus from initial visits of two subjects (subjects 9 and 15) were more heterogeneous, raising the possibility of dual infection, the development of recombinant viruses, or mistiming of the point of seroconversion. Further analysis excluded these possibilities as the basis for the observed heterogeneity (see *Methods*).

Both diversity and divergence increased over time in all three progressor categories, and the increase in diversity per year or divergence per year was progressively greater in moving from the nonprogressor to the moderate progressor to the rapid progressor groups (Fig. 2). The rapid progressor group had a significantly higher rate of increase in divergence over time than the nonprogressor group ($P < 0.001$). While rapid progressors showed a greater rate of increase in divergence than moderate progressors, the difference between them did not achieve significance ($P = 0.17$). Differences in the slopes of the rate of increase in diversity were also significant between the rapid progressors and the nonprogressors ($P = 0.001$). The differences in slopes of diversity between nonprogressors and moderate progressors was significant ($P = 0.01$), while differences in the diversity slopes between moderate and rapid progressors showed a trend toward significance, but did not achieve it ($P = 0.08$).

The correlation between the viral genetic diversity or divergence at a given visit and the decline in CD4 T cell levels over the next year was also examined (see *Methods*). Controlling for the CD4 T cell count at a given visit, the relationship between the diversity or divergence at that visit and the decline in the CD4 T cell count over the next year was determined. Both the diversity and the divergence were significantly negatively correlated with the CD4 T cell count 12 months later (for diversity, $r = -0.43$, $P = .02$; for divergence $r = -0.30$, $P = 0.007$). Thus, subjects whose virus showed greater genetic diversity or divergence at a given visit were likely to have a greater CD4 T cell decline over the next year.

The ratio of the rate of synonymous mutations per potential site of synonymous mutation (dS) to the rate of nonsynonymous mutations per potential site of nonsynonymous (NS) mutation (dN) was determined for all the subjects from each of the progressor categories (Table 1). If mutation is occurring randomly without selection, this ratio should approximate 1. Rapid and moderate progressors both had a median dS/dN ratio of 0.4, indicating a selective advantage for NS changes in these progressor groups. These ratios differed significantly from 1 ($P = 0.01$ for rapid progressors, $P = 0.001$ for moderate progressors, Wilcoxon signed-rank test). The dS/dN value for nonprogressors was equal to 1.6, which, in this small sample size, did not differ significantly from 1, but showed a trend toward significant selection against NS changes ($P = 0.15$, Wilcoxon signed-rank test). This suggests that evolution in the nonprogressors does not demonstrate a selective advantage for those viruses carrying changes in the envelope protein structure. The median dS/dN ratios obtained at all visits were significantly different between the nonprogressors and either the moderate or the rapid progressors ($P = 0.02$ vs. either progressor group). These differences arose primarily from differences in dN ($P < 0.005$ vs. either progressor group), since dS was not significantly different among any of the groups ($P > 0.5$). Thus, nonprogressor was distinguished from progressor status by whether or not selection favored NS changes and demonstrated

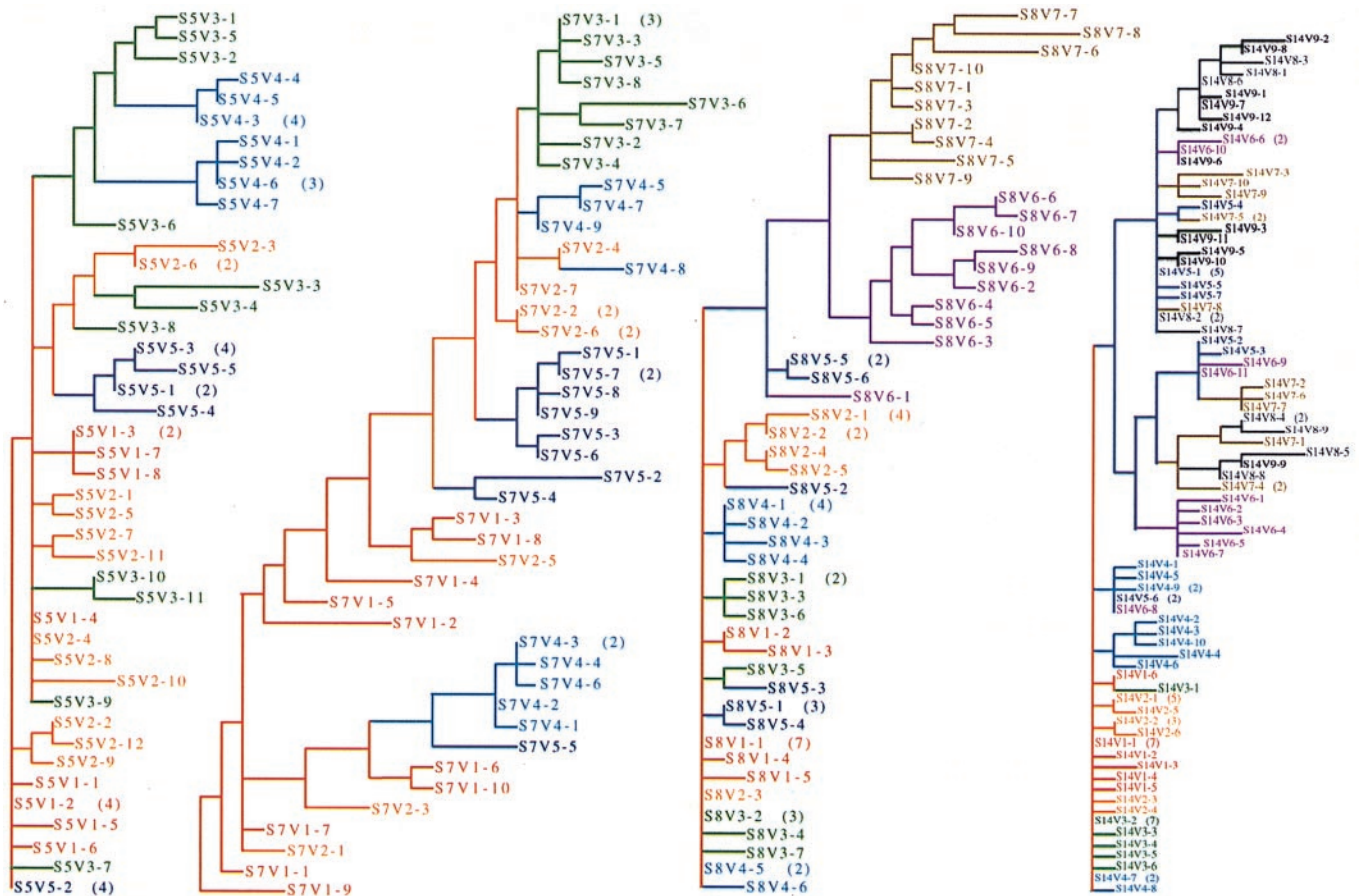


FIG. 4. Phylogenetic trees of viral evolution in four additional randomly selected individuals from the original cohort. In each subject evolution is not sustained along a single branch, being interrupted in subject 5 between visits 4 and 5, in subject 7 between visits 4 and 5, in subject 8 between visits 6 and 7, and in subject 14 between visits 5 and 6. The horizontal distance at the bottom of the tree between S14V4-7 and S14V4-8 reflects a single mutation.

a strong relationship between the selective pattern and the level to which CD4 T cells fell during this period of observation.

Phylogenetic trees from 10 of 15 subjects showed no evidence of predominance of a single strain throughout an extended period of time. For example, in subject 9 (Fig. 3) all viral isolates through visit 4 arise from a point close to the viruses in visit 1. The viruses at visit 5 are spread throughout the evolutionary tree. Viruses at visits 6 and 7 extend from the branch carrying clones 1 and 9 from visit 5. However, the viruses at visit 8 emerge from visit 6 clones, not from a visit 7 clone. This pattern of limited progression along a single branch followed by return to strains closely related to those present at an earlier visit is observed in nine other subjects (subjects 1, 3, 5–8, 10, 13, and 14). Phylogenetic trees randomly selected from the remaining 14 subjects show this pattern (Fig. 4). In these 10 subjects host factors that influence the evolution of early viruses appear potent enough to select against the few clones that predominate at any given visit. However, those same host factors are not broadly effective against the full range of viruses present at a given visit.

DISCUSSION

In this cohort of HIV-1-infected injection drug users, higher levels of both genetic diversity and divergence in the HIV-1 variants present in a given individual were associated with a greater decline in CD4 T cells. In subjects with similar CD4 T cell counts, those subjects with more genetically diverse viral clones experienced, as a group, greater CD4 T cell decline in the subsequent 12 months. Although synonymous substitution rates were comparable in rapid, moderate, and nonprogressors, non-synonymous substitution rates in progressors were roughly three times those in nonprogressors. The overall pattern is one in which

viral strains from nonprogressors showed possible selection against amino acid change, while those from progressors showed selection for such change (or against the absence of change). This finding is inconsistent with a model in which progression is caused by rapid proliferation of an initially "most fit" viral strain.

Two recent studies of the association in adult populations between genetic variation and disease progression have yielded somewhat conflicting results. McDonald *et al.* (25) compared *env* gene variation in rapid and slow progressors at two time points separated by 2.5 years in 10 subjects, all of whom had CD4 counts below 400 at the first time point. In agreement with the current study, these workers found greater genetic divergence in three of the five rapid progressors studied. The intravisit diversity in their rapid progressors was, however, less than that observed in their slow progressors at the second time point for their subjects. Conclusions on diversity varied depending on whether the analysis was restricted to the C2–C5 or the C2–V3 *env* region. The conclusions on diversity of that study probably differ from the current study because the subjects in the McDonald study were not followed from the time of seroconversion and because many fewer time points were evaluated in each subject. In a study of six adult subjects that applied analyses similar to those of the current study, Wolinsky *et al.* (22) observed less viral genetic diversity in two subjects with the most rapidly declining CD4 T cell counts compared with subjects with more slowly declining CD4 T cell counts. Virus from five of our six rapid progressors showed high levels of both diversity and divergence. One rapid progressor (subject 11, Table 1) showed the low divergence and diversity pattern seen in Wolinsky's two rapid progressors and in our nonprogressors. Unlike the nonprogressors, however, subject 11 had a higher viral load (62,179 copies/ml, Table 1). This subject

and the two subjects chosen by Wolinsky *et al.* (22), which also had extraordinarily rapid disease progression, may be exceptional. These exceptional subjects likely represent individuals who failed to develop any effective immune response to their viral infection and therefore a single "best fit," rapidly replicating virus could emerge in the immunologically stable environment. In fact, Wolinsky *et al.* (22) were not able to identify effective immune responses to the viruses in their very rapidly progressing subjects.

The association between CD4 T cell decline and increased genetic diversity and divergence observed in most of our subjects is consistent, in part, with the model of Nowak (18, 19, 21), who proposed that increasing viral genetic diversity is associated with CD4 T cell decline. Nowak hypothesized that ultimately this diversity results in viral clones developing critical epitopes that are outside the T cell repertoire of the host, resulting in failure of the host immune response to control HIV-1 infection. In such a model one might expect to see increasing diversity followed by a restriction of diversity limited to those clones carrying epitopes that have exceeded the T cell repertoire. In those six subjects who progressed to AIDS in this study, reduction in the level of diversity was only observed in two subjects, subjects 3 and 15, and the reduction was minimal in subject 3. In all other subjects who progressed to AIDS, diversity and divergence continued to increase throughout the observed course of infection. Progression in this cohort was associated with a shift away from the virus that predominated at the previous visit, a pattern most consistent with either frequency-dependent selection, in which only the most numerous genotypes are selected against, or independent evolution of viral variants from different geographical or environmental sites within the body, as described by Sala *et al.* (49).

By the latter model the virus observed in the PBMC pool would simply reflect at which site the cellular environment has become most conducive to the growth and spread of virus. The optimal site for virus growth could be continuously changing because of suppression of virus by a localized immune response at a specific site or by enhancement of viral growth due to T cell activation at a specific site. None of the observations made here are inconsistent with such a model. Ultimately, however, to overcome the infection the host must control it at the level of the whole organism. The persistent increase in genetic diversity and the appearance in many subjects of variants closely related to those that predominated earlier suggest that any host immune response that may have developed failed to eradicate HIV-1 infection, not because of the emergence of escape variants, but because, at any point in the infection, the immune response was not broadly targeting the full array of viruses present (50), targeting instead only the most frequent virus.

In the face of a more broadly effective immune response, as might exist in the nonprogressors, those viruses that are less replication competent therefore might be conserved precisely because they do not replicate to some critical threshold required for targeting by the immune response. Selection against nonsynonymous mutations, as was observed in our nonprogressors, therefore might be favored precisely because such nonsynonymous changes might result in the appearance of more replication-competent viruses that would be recognized by the host immune system.

We wish to thank Drs. Robert F. Siliciano, Margaret A. Liu, Stephen Gange, and James I. Mullins for critical review and helpful discussions, Mr. Glen McFarlane for technical assistance, and Ms. Edith Muth for secretarial assistance. This work was supported by the National Institute on Drug Abuse Grants DA 09973, 04334, and 09541.

- Mansky, L. M. & Temin, H. M. (1995) *J. Virol.* **69**, 5087–5094.
- Yokoyama, S. & Gojobori, T. (1987) *J. Mol. Evol.* **24**, 330–336.
- Hahn, B. H., Shaw, G. M., Taylor, M. E., Redfield, R. R., Markham, P. D., Salahuddin, S. Z., Wong-Staal, F., Gallo, R. C., Parks, E. S. & Parks, W. P. (1986) *Science* **232**, 1548–1553.
- Markham, R. B., Schwartz, D. H., Templeton, A., Margolick, J. B., Farzadegan, H., Vlahov, D. & Yu, X. F. (1996) *J. Virol.* **70**, 6947–6954.
- Dean, A. M. (1989) *Genetics* **123**, 441–454.
- Hartl, D. L., Dykhuizen, D. E. & Dean, A. M. (1985) *Genetics* **111**, 655–674.
- Zaitseva, M., Blauvelt, A., Lee, S., Lapham, C. K., Klaus-Kotvun, V., Mostowski, H., Manischewitz, J. & Golding, H. (1997) *Nat. Med.* **3**, 1369–1375.
- Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., Goedert, J. J., Buchbinder, S. P., Vittinghoff, E., Gomperts, E., *et al.* (1996) *Science* **273**, 1856–1862.
- Endres, M. J., Clapham, P. R., Marsh, M., Ahuja, M., Turner, J. D., McKnight, A., Thomas, J. F., Stoebenu-Haggarty, B., Choe, S., Vance, P. J., *et al.* (1996) *Cell* **87**, 745–756.
- Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P. D., Wu, L., Mackay, C. R., LaRosa, G., Newman, W., Gerard, N., Gerard, C. & Sodroski, J. (1996) *Cell* **85**, 1135–1148.
- Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M. & Berger, E. A. (1996) *Science* **272**, 1955–1958.
- Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., Goedert, J. J., O'Brien, T. R., Jacobson, L. P., Kaslow, R., *et al.* (1997) *Science* **277**, 959–965.
- Atchison, R. E., Gosling, J., Monteclaro, F. S., Franci, C., Digilio, L., Charo, I. F. & Goldsmith, M. A. (1996) *Science* **274**, 1924–1926.
- Huang, Y., Paxton, W. A., Wolinsky, S. M., Neumann, A. U., Zhang, L., He, T., Kang, S., Ceradini, D., Jin, Z., Yazdanbakhsh, K., *et al.* (1996) *Nat. Med.* **2**, 1240–1243.
- McKnight, A., Wilkinson, D., Simmons, G., Talbot, S., Picard, L., Ahuja, M., Marsh, M., Hoxie, J. A. & Clapham, P. R. (1997) *J. Virol.* **71**, 1692–1696.
- Bonnell, M. L. & Selander, R. K. (1974) *Science* **184**, 908–909.
- O'Brien, S. J., Roelke, M. E., Marker, L., Newman, A., Winkler, C. A., Meltzer, D., Colly, L., Evermann, J. F., Bush, M. & Wildt, D. E. (1985) *Science* **227**, 1428–1434.
- Nowak, M. A., Anderson, R. M., McLean, A. R., Wolfs, T. F., Goudsmit, J. & May, R. M. (1991) *Science* **254**, 963–969.
- Nowak, M., Anderson, R., Boerlijst, M., Bonhoeffer, S., May, R. & McMichael, A. (1996) *Science* **274**, 1008–1011.
- Wei, X., Ghosh, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifson, J. D., Bonhoeffer, S., Nowak, M. A., Hahn, B. H., *et al.* (1995) *Nature (London)* **373**, 117–122.
- Nowak, M. & Bangham, C. (1996) *Science* **272**, 74–79.
- Wolinsky, S. M., Korber, B. T., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D., Safritz, J. T., *et al.* (1996) *Science* **272**, 537–542.
- Delwart, E. L., Pan, H., Sheppard, H. W., Wolpert, D., Neumann, A. U., Korber, B. & Mullins, J. I. (1997) *J. Virol.* **71**, 7498–7508.
- Delwart, E. L., Sheppard, H. W., Walker, B. D., Goudsmit, J. & Mullins, J. I. (1994) *J. Virol.* **68**, 6672–6683.
- McDonald, R. A., Mayers, D. L., Chung, R. C., Wagner, K. F., Ratto-Kim, S., Birx, D. L. & Michael, N. L. (1997) *J. Virol.* **71**, 1871–1879.
- Lukashov, V. V., Kuiken, C. L. & Goudsmit, J. (1995) *J. Virol.* **69**, 6911–6916.
- Vlahov, D., Anthony, J. C., Munoz, A., Margolick, J., Nelson, K. E., Celentano, D. D., Solomon, L. & Polk, B. F. (1991) *J. Drug Issues* **21**, 759–776.
- Stevenson, M., Stanwick, T. L., Dempsey, M. P. & Lamonica, C. A. (1990) *EMBO J.* **9**, 1551–1560.
- Bukrinsky, M. I., Stanwick, T. L., Dempsey, M. P. & Stevenson, M. (1991) *Science* **254**, 423–427.
- Chun, T., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J., Taylor, H., Herman-kova, M., Chadwick, K., Margolick, J., Quinn, T., *et al.* (1997) *Nature (London)* **387**, 183–187.
- Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A. & Chen, I. S. Y. (1990) *Cell* **61**, 213–222.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY).
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Liu, S. L., Rodrigo, A. G., Shankarappa, R., Learn, G. H., Hsu, L., Davidov, O., Zhao, L. P. & Mullins, J. I. (1996) *Science* **273**, 415–416.
- Kumar, S., Tamura, K. & Nei, M. (1994) *Comput. Appl. Biosci.* **10**, 189–191.
- Laird, N. & Ware, J. (1982) *Biometrics* **38**, 963–974.
- Korber, B., Wolinsky, S., Haynes, B., Kunstman, K., Levy, R., Furtado, M., Otto, P. & Myers, G. (1992) *AIDS Res. Hum. Retroviruses* **8**, 1461–1465.
- Zwart, G., Langedijk, H., van der Hoek, L., de Jong, J. J., Wolfs, T. F., Ramautarsing, C., Bakker, M., de Ronde, A. & Goudsmit, J. (1991) *Virology* **181**, 481–489.
- Zwart, G., Wolfs, T. F., Valk, M., van der Hoek, L., Kuiken, C. L. & Goudsmit, J. (1992) *AIDS Res. Hum. Retroviruses* **8**, 1897–1908.
- Cann, A. J., Churcher, M. J., Boyd, M., O'Brien, W., Zhao, J. Q., Zack, J. & Chen, I. S. (1992) *J. Virol.* **66**, 305–309.
- Hwang, S. S., Boyle, T. J., Lyerly, H. K. & Cullen, B. R. (1991) *Science* **253**, 71–73.
- Schuitemaker, H., Kootstra, N. A., de Goede, R. E., de Wolf, F., Miedema, F. & Tersmette, M. (1991) *J. Virol.* **65**, 356–363.
- Schuitemaker, H., Kootstra, N. A., Groenink, M., de Goede, R. E., Miedema, F. & Tersmette, M. (1992) *AIDS Res. Hum. Retroviruses* **8**, 1679–1682.
- Sharpless, N. E., O'Brien, W. A., Verdin, E., Kufta, C. V., Chen, I. S. & Dubois-Dalcq, M. (1992) *J. Virol.* **66**, 2588–2593.
- Westervelt, P., Trowbridge, D. B., Epstein, L., Blumberg, B., Li, Y., Hahn, B. H., Shaw, G. M., Price, R. & Ratner, L. (1992) *J. Virol.* **66**, 2577–2582.
- Markham, R. B., Yu, X. F., Farzadegan, H., Ray, S. & Vlahov, D. (1995) *J. Infect. Dis.* **171**, 797–804.
- Zhang, L. Q., MacKenzie, P., Cleland, A., Holmes, E. C., Leigh Brown, A. J. & Simmonds, P. (1993) *J. Virol.* **67**, 3345–3356.
- Zhu, T., Mo, H., Wang, N., Nam, D., Cao, Y., Koup, R. A. & Ho, D. D. (1993) *Science* **261**, 1179–1181.
- Sala, M., Zambruno, G., Vartanian, J. P., Marconi, A., Bertazzoni, U. & Wain-Hobson, S. (1994) *J. Virol.* **68**, 5280–5283.
- Wilson, C. C., Kalams, S. A., Wilkes, B. M., Ruhl, D. J., Gao, F., Hahn, B. H., Hanson, I. C., Luzuriaga, K., Wolinsky, S., Koup, R., *et al.* (1997) *J. Virol.* **71**, 1256–1264.
- Margolick, J. B., Scott, E. R., Chadwick, K. R., Shapiro, H. M., Hetzel, D., Smith, S. J. & Vogt, R. F. (1992) *Cytometry* **13**, 198–203.