Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 1264–1269, February 1997
Evolution

Evolutionary mechanisms and population dynamics of the third variable envelope region of HIV within single hosts

(molecular evolution/positive selection)

YUMI YAMAGUCHI*† AND TAKASHI GOJOBORI*‡

*Center for Information Biology, National Institute of Genetics, Mishima 411, Japan; and †New Energy and Industrial Technology Development Organization, Tokyo 170, Japan

Communicated by Susumu Ohno, Beckman Research Institute of the City of Hope, Duarte, CA, December 6, 1996 (received for review August 20, 1996)

ABSTRACT **Clonal diversifications of HIV virus were monitored by periodic samplings on each of the six patients with regard to 183- to 335-bp segments of the *env* gene, which invariably included the functionally critical V3 region. Subsequently, six individual phylogenetic trees of viral variants were constructed. It was found that at one time or another during the course of disease progression, viral variants were inexplicably released from a strong negative selection against nonsynonymous base substitutions, possibly indicating positive selection. This resulted in concentrated amino acid substitutions at five specific sites within the V3 region. It was noted that these sites were often involved as antigenic determinants that provoked the host immune response and that these sites were also involved in the determination of viral phenotypes as to their cell tropism, syncytium formation capability, and replication rates.**

Because of the high error rate of reverse transcriptase (1–3), the genome of HIV, like that of other retroviruses, mutates at a very high rate. Accordingly, the rate of nucleotide substitution for retroviruses is about a million times higher than that of the host genes (4–9). Due to this high mutation rate, genomic sequences of HIVs within a single host are not homogeneous; rather, they comprise a heterogeneous population in which all members are closely related to each other (10).

Of those sequence variants constantly generated by successive viral replication, some are destined to be eliminated because of their defects in the replication machinery (11), whereas others may enjoy selective advantage because of newly acquired endowments that facilitate escape from immune persecution by the host. To elucidate the mechanisms of viral evolution, we need to characterize the viral variants that can survive in a host and can successfully infect another host. In this study, we intended to elucidate the evolutionary mechanisms of viruses within a host at the molecular level and understand the dynamics of the viral population within a host by analyzing sequence variations of HIV, periodically sampled from a single host.

The genome of HIV is composed of three major protein-coding regions (*gag*, *pol*, and *env*), long terminal regions at both the 5′ and 3′ ends, and regulatory genes. In particular, it is known that the *env* region has a great many variations.

The third variable envelope (V3) region (12) of HIV contains a target for neutralizing antibodies (13), a recognition site for the T cells of HIV (14), and the determinants of infectivity (15, 16), cell tropism (17), syncytium formation (SI) capability (18, 19), and replication rates (19). We examined the patterns of nucleotide substitutions in this region to see when and at which amino acid sites positive selection is operating on the V3 region within a single host. We collected nucleotide sequence data of HIV clones isolated from a single host at several points in time after infection (20–22). By estimating the rates of synonymous and nonsynonymous substitutions for the V3 region of HIV within a single host at each period of time, we found that the rate of nonsynonymous substitution for the V3 region was significantly higher than that of synonymous substitution at one time or another after infection. We also estimated the number of amino acid substitutions that occurred at each amino acid site in the V3 region within single hosts. As a result, we found that amino acid substitutions dominantly occurred at five specific amino acid sites where the substitutions are known to be possibly responsible for production of antigenic variation and determination of the viral phenotypes. These observations indicate a strong possibility that positive selection is taking place by changing particular amino acid sites in the V3 region of HIV within a single host at one time or another after the infection.

MATERIALS AND METHODS

Nucleotide Sequence Data and Clinical Status of the Patients. We collected the nucleotide sequence data of the V3 region for clones that were periodically isolated from each of six patients (20–22). In practice, we used the nucleotide sequence data of those clones stored in the international DNA databases (DDBJ/EMBL/GenBank). The clinical status of the patients and methodological information about the isolation of viral clones and nucleotide sequencing are summarized in Table 1.

Construction of Phylogenetic Trees. Phylogenetic trees were constructed by the neighbor-joining method (23) using the number of nucleotide substitutions. The numbers of nucleotide substitutions were estimated by the six-parameter method (24). Bootstrap probabilities for major clusters were examined (25). The resampling procedure was repeated 2000 times.

Estimation of Rate of Nucleotide Substitutions. For each clone, the numbers of synonymous and nonsynonymous substitutions from the ancestral sequence were estimated using the method of Nei and Gojobori (26). When nucleotide sequences in the first sample were completely homogeneous, the sequence was regarded as the ancestral sequence. When nucleotide sequences in the first sample were not completely homogeneous, we deduced the ancestral sequence using a maximum parsimony principle. The deduced ancestral sequence was used as the reference sequence. The average rates of synonymous and nonsynonymous substitutions between two points in time were estimated (27). A statistical test was done by use of the bootstrap resampling test (28). The resampling procedure was repeated 2000 times.

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.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/941264-6$2.00/0
PNAS is available online at **http://www.pnas.org**.

Abbreviation: AZT, azidothymidine.

‡To whom reprint requests should be addressed. e-mail: tgojobor@genes.nig.ac.jp.

Table 1. Data used in our analyses

Designation in our analyses	Patient A	Patient B	Patient C	Patient D	Patient E	Patient F
Designation in the original paper (ref. no.)	p82 (20)	495 (21)	1 (21)	s2 (22)	s1 (22)	s4 (22)
Presumed transmission mode	A single batch of factor VIII	Homosexual contact	Homosexual contact	NA	NA	NA
Clinical status	Asymptomatic	AIDS in 1989	Asymptomatic	NA	NA	NA
CD4 counts during the study	Decreasing	Decreasing	Decreasing	Decreasing	Fluctuating	Decreasing
Antiviral therapy	None	AZT from 1989	None	None	None	None
Term for the study	1984–1991	1985 until after 56 months	1985 until after 59 months	May 1985–Oct. 1987	Nov. 1985–May 1989	Jan. 1985–June 1989
Nucleotide sequences determined	234 nt	183–276 nt	183–276 nt	332–335 nt	332–335 nt	332–335 nt
Tissue	Plasma	Serum	Serum	Peripheral blood leukocytes	Peripheral blood leukocytes	Peripheral blood leukocytes
Molecular type	Viral RNA	Viral RNA	Viral RNA	DNA	DNA	DNA

NA, not available.

RESULTS

Phylogenetic Relationships Among HIV Clones Isolated from a Single Host. We constructed phylogenetic trees of HIVs within a host using nucleotide sequence data covering a part of the *env* gene, including the V3 region (Fig. 1). Though the nucleotide sequences in the initial samples were completely or nearly homogeneous in and around the V3 region, an accumulation of nucleotide changes was observed at all successive time points in all six patients.

In patient A, the nucleotide sequences of the first samples were completely homogeneous. In 1987, 3 years after the first samples were taken, 13 variants were observed. However, the minor cluster in 1987 seemed to have survived in the successive period. After 1988, two major clusters corresponding to preferentially macrophage-tropic or T cell line-adapted phenotypes (29) were observed (20).

In the case of patient B, though the nucleotide sequences of the first samples were nearly homogeneous, the accumulations of nucleotide substitutions were observed every year. Patient B was diagnosed as having AIDS in 1989, and 55 months after the primary infection, azidothymidine (AZT) treatment was started (21). Clones sampled right after AZT treatment were closely related to each other. The hypothesis that only a certain variant was resistant to AZT can explain this reduction of sequence variation. Several studies reported that HIV variants resistant to AZT often have amino acid changes in reverse transcriptase (30, 31). In this case, though the nucleotide sequences of reverse transcriptase were not available, these variants might have mutations in the *pol* gene on the same genome, which enabled the viruses to be resistant to AZT.

In patient C, sequential shifts in the viral population (32) were observed at a later stage after infection. In patient D, some clones sampled after 29 months from the first sample, differed greatly from the ancestral clone at the nucleotide sequence level, while the other two clones sampled at that time were closely related to the ancestral clone. Such differences in the rate of nucleotide substitution among clones sampled at the same time were observed in patients E and F.

Rates of Nucleotide Substitution for the V3 Region Within a Single Host. To examine whether positive selection is operating on the V3 region within the human body, we estimated the average rates of synonymous and nonsynonymous substitutions (26, 27) for the V3 region within a single host between two points in time. The results are summarized in Fig. 2. During some periods of time after infection, the rate of nonsynonymous substitution for the V3 region was significantly higher than that of synonymous substitution (Fig. 2). These data cannot be attributed to random genetic drift (33), indicating that positive selection is taking place in the V3 region of HIV within a single host at one time or another after infection. However, we also noticed that for other periods of time, the rate of synonymous substitution for the V3 region was significantly higher than that of nonsynonymous substitution. This suggests that natural selection may take place at one time or another after the infection. The average rates of synonymous substitutions for the V3 region also greatly fluctuated with time. It may be that not all viral particles replicate at a uniform rate during the course of time after infection. The existence of substantial variation in replication rates among the viral variants may affect fluctuations in the rate of synonymous substitution in HIV.

Distribution of Amino Acid Substitutions. To elucidate the kind of selection operating on each amino acid site, we estimated the number of amino acid substitutions that occurred at each amino acid site in the V3 region (Fig. 3) and examined the correspondence between biological function and relative frequency of amino acid substitution. It has been reported that in the V3 region, amino acid changes at the 13th and 25th amino acid sites from the cysteine on the 5′ side of the V3 region can alter antigen binding specificities (21, 34). It is also known that changes at the 11th, 24th, and 25th amino acid sites from the cysteine on the 5′ side of the V3 region, affect syncytium formation or cell tropism (35, 36). Determination of crystal structure revealed the amino acid stretch (Fig. 4) with which monoclonal antibodies for HIV interact (37, 38). We found that there were five amino acid sites (the 11th, 13th, 18th, 20th, and 25th amino acid sites from the cysteine on the 5′ side of the V3 region) where amino acid substitutions predominantly occurred. We also found that the frequencies of amino acid changes at these positions were significantly larger according to a statistical test based on the Poisson distribution. All five amino acid sites detected are sites where substitutions are known to be possibly responsible for antigenic or phenotypic variations in HIV. These are also sites with which monoclonal antibodies for HIV interact. These observations suggest that amino acid substitutions predominantly occurred at the particular amino acid sites where substitutions are possibly responsible for the production of antigenic variation and determination of viral phenotypes.

Population Dynamics of HIV Variants Within a Single Host. We wanted to observe the population dynamics of HIV variants within a single host (20). We classified the HIV variants into an appropriate number of types according to amino acid variation at

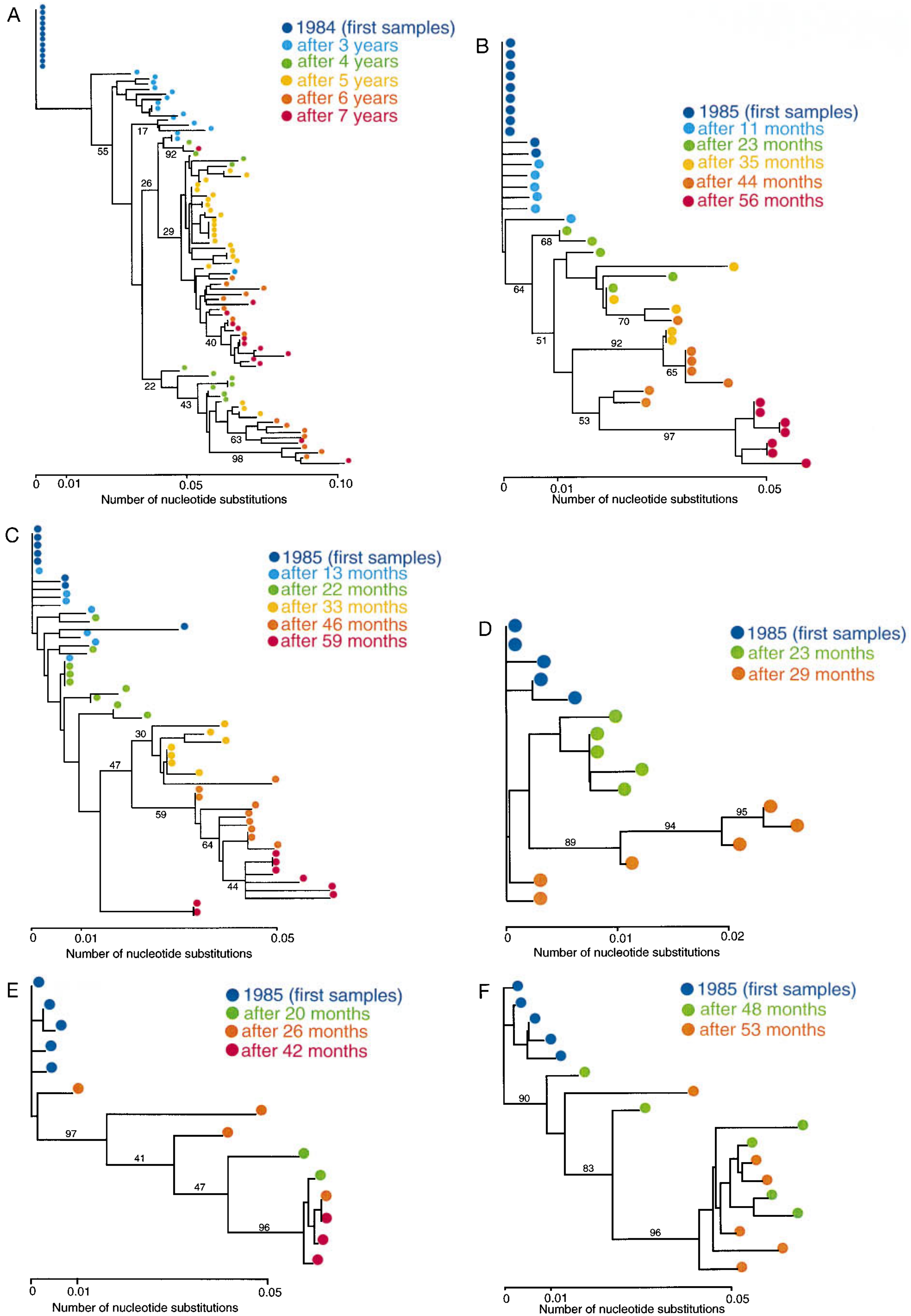

FIG. 1. Phylogenetic trees of HIVs within single hosts. Clones are shown by colored dots according to their sampling time. Bootstrap probabilities (25) for major clusters are also shown by percentages (%). (*A*) Phylogenetic tree of HIVs periodically sampled from patient A (20) using a 234-nt sequence. (*B*) Phylogenetic tree of HIVs periodically sampled from patient B (21) using a 183-nt sequence. (*C*) Phylogenetic tree of HIVs periodically sampled from patient C (21) using a 183-nt sequence. (*D*) Phylogenetic tree of HIVs periodically sampled from patient D (22) using a 332-nt sequence. (*E*) Phylogenetic tree of HIVs periodically sampled from patient E (22) using a 335-nt sequence. (*F*) Phylogenetic tree of HIVs periodically sampled from patient F (22) using a 332-nt sequence.

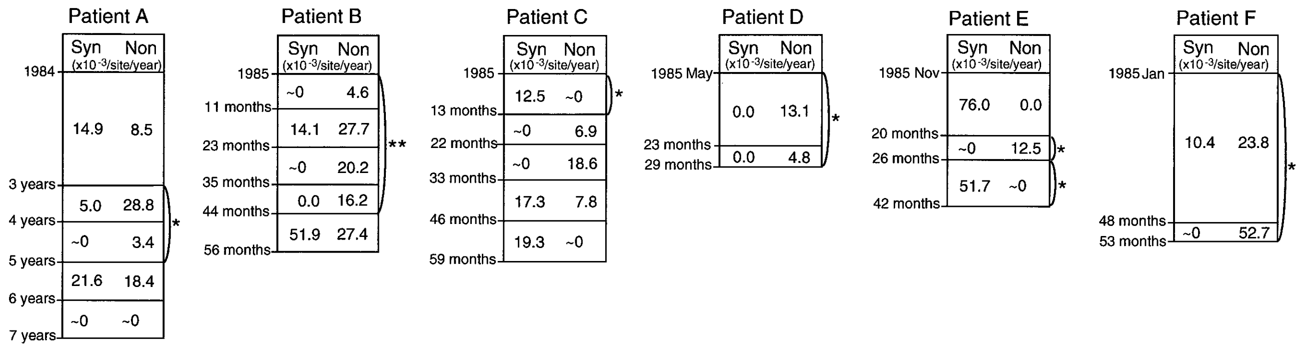

FIG. 2. Rates of synonymous and nonsynonymous substitutions of the V3 region within single hosts. One and two asterisks indicate that difference between the rates of synonymous and nonsynonymous substitutions has statistical significance at the 5% and 1% levels, respectively.

particular sites in the V3 region, where substitutions are responsible for antigenic variation or where monoclonal antibodies for HIV interact with (21, 34, 37, 38). We can trace the population dynamics of HIV variants by focusing attention on particular amino acid sites that are possibly responsible for natural selection. Let us examine the case of patient A (20). We illustrated the change of frequency for each type of HIV within a single host over time (Fig. 4). It is clear that the ancestral type was lost three years after the first sampling. We observed 19 types of HIV in 7 years in this patient. At the 25th amino acid site from the cysteine on the 5′ side of the V3 region, though the ancestral amino acid was glutamic acid, this type of variant was completely lost by 1987. In fact, viral types in which this position was occupied by aspartic acid, glutamine, glycine, and arginine were all found during the study. The substitutions at this site were known to be possibly responsible for antigenic and phenotypic variation (34), and the largest numbers of amino acid substitutions were observed in this patient. Positive selection on this site might affect the rapid change and population dynamics of HIV variants.

DISCUSSION

Elucidation of the mechanism for viral evolution within a human body is of particular importance in understanding the evolutionary mechanism of viruses and for developing effective strategies for antiviral therapy. We analyzed the nucleotide sequence data of HIV, periodically sampled from six patients (20–22). Our analyses, including phylogenetic analysis, periodic examination of rates of nucleotide substitution, and relative frequencies of substitution at each amino acid site, provided much information about the nature of evolution of HIV within the human body. In particular, we identified time periods in which positive selection may operate and the amino acid sites where positive selection takes place. Our attempts at tracing the frequency changes of HIV variants classified by particular amino acid sites enabled us to see the population dynamics of HIV variants very clearly.

Our phylogenetic analysis showed particular variants in patient B that might survive AZT treatment. Though we analyzed only a part of the *env* gene, the survivors may have mutations in the *pol* gene (30, 31) on the same genome that enable the virus to escape from AZT treatment. Phylogenetic analysis is very useful in tracing surviving lineages and revealing how each variant arises. Wolinsky *et al.* (32) constructed phylogenetic trees for viral clones periodically sampled from each of two patients. One patient showed clear sequential shifts in the population of viruses, while the other patient showed substantial intermingling of viral se-

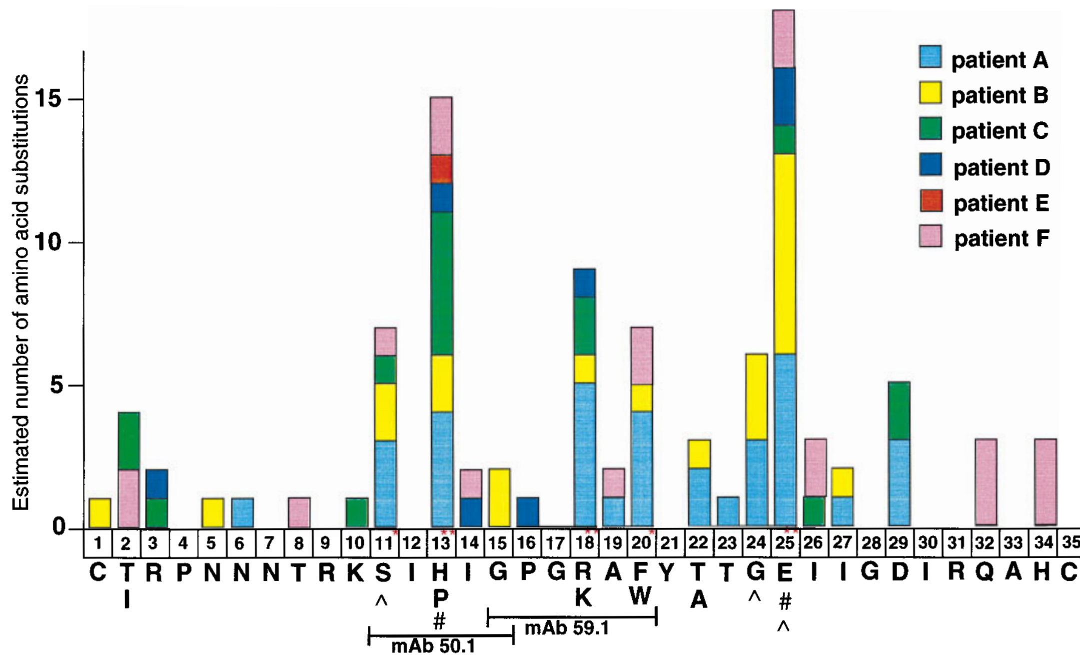

FIG. 3. Estimated number of amino acid substitutions at each amino acid site of the V3 region observed for the six patients. In estimating the number of amino acid substitutions for each patient, we inferred the ancestral amino acid sequences at all branching points from the phylogenetic tree. Amino acid positions with one or two red asterisks are the sites where amino acid substitutions dominantly occurred with statistical significance at the 5% or 1% level, respectively. Horizontal bars show the amino acid stretches with which monoclonal antibodies interact. #: Amino acid sites responsible for antigenic variation. ^, Amino acid sites responsible for cell tropism or syncytium-inducing phenotype.

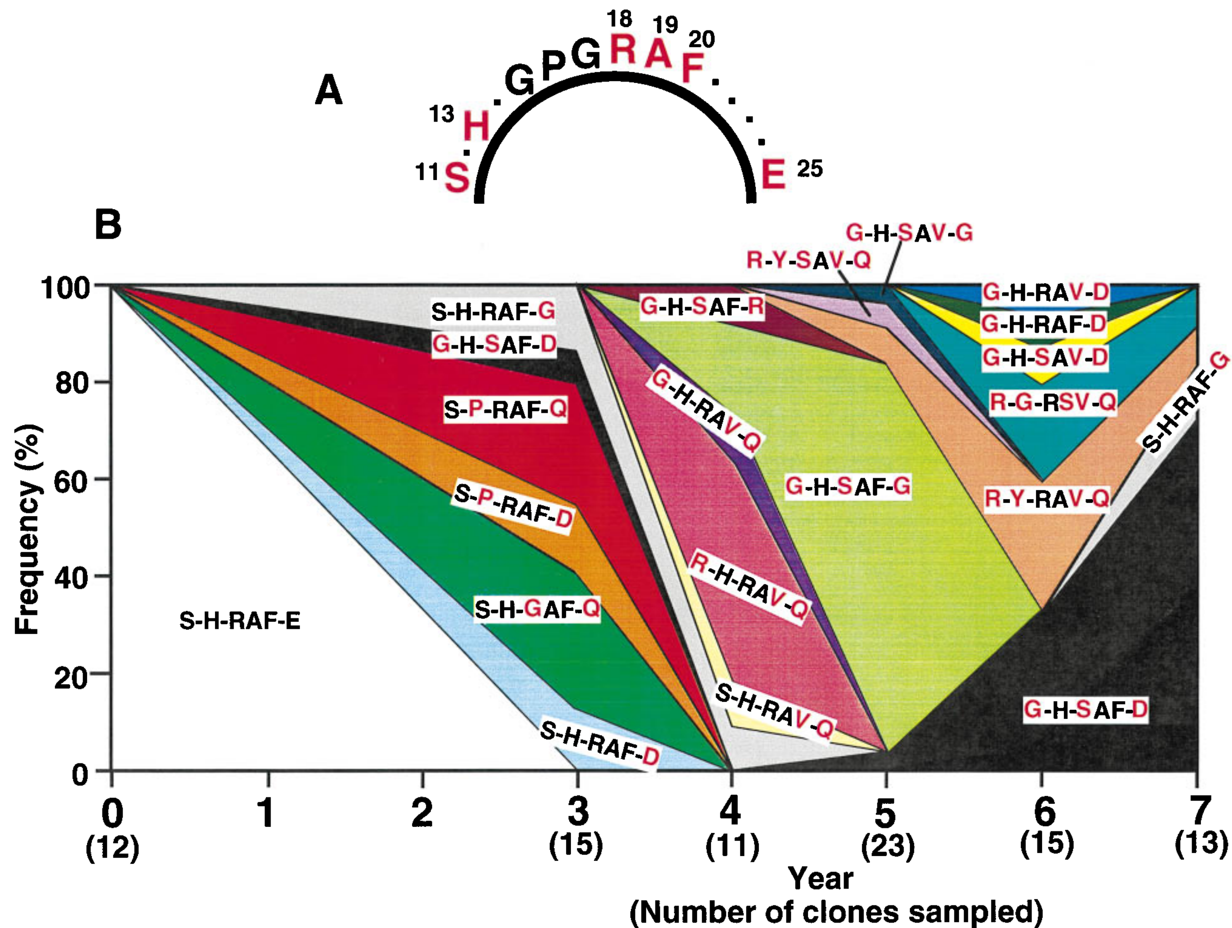

FIG. 4. Population dynamics of HIV variants within patient A. (*A*) The amino acid stretch including the crown of the V3 loop. The amino acid sites shown by single letters in red, where higher numbers of amino acid substitutions were observed, are thought to be responsible for antigenic variation in HIV. (*B*) Frequency change of HIV variants within patient A. We classified the HIV variants according to amino acid variation at the red sites in *A*.

quences from different time points (32). Our phylogenetic analyses showed more general aspects of viral evolution within a human body.

It is known that some types of HIV have the capability to induce syncytium formation. This syncytium-inducing capability of HIV depends on an amino acid sequence of the V3 region (35). We classified HIV clones into two categories, syncytium-inducing (SI) clones and non-syncytium-inducing (NSI) clones, by looking at amino acids occupying the 11th, 24th, and 25th sites of the V3 region. We found that the SI viruses evolved faster than the NSI viruses (data not shown). Our findings are consistent with several reports that SI viruses replicates themselves faster than NSI viruses and that NSI viruses are more resistant to neutralizing antibodies than SI viruses (39).

We estimated the rate of nucleotide substitution using the sequence data of HIV variants periodically sampled. We performed a statistical test using the bootstrapping resampling method. Statistical tests are necessary when we compare the number of synonymous substitutions with that of nonsynonymous substitutions to conduct a test of neutrality. We found periods in which the rate of nonsynonymous substitution was higher than that of synonymous substitution. Thus, it is possible that positive selection is operating on amino acid changes in the V3 region of HIV during those periods, to escape from the immune system of the host. This possibility is supported by the evidence of Phillips *et al.* (40) of immune escape from cytotoxic T lymphocyte recognition by HIV. It was also noted that positive selection was also found to take place in the *tax* gene of HTLV-1, which is the dominant antigen recognized by HTLV-1-specific cytotoxic T lymphocyte (41). However, we also noticed that for other periods, the rate of synonymous substitution was higher than that of nonsynonymous substitution. The type of selection and its intensity may not be constant all the time in the human body. Periodic analysis of the intrabody evolution of HIV may reveal the processes and forces of evolution.

We also found that amino acid substitutions predominantly occurred at five amino acid sites where substitutions may be responsible for production of antigenic variation and determination of viral phenotypes. This may suggest that nucleotide substitutions occurring in the V3 region may be important for survival of the virus. This analysis at each amino acid site enabled us to examine the relationships between the pattern of amino acid substitutions and biological function at the amino acid site. Lukashov *et al.* (42) suggested that rate of nonsynonymous substitution in the V3 region correlated with length of the immunocompetent period. We observed a preponderance of amino acid substitutions at sites responsible for antigenic variation. Our observations may be consistent with their suggestion (42). However, other types of positive selection may be operating in this region. HIV in early infection is mainly macrophage tropic and replicates slowly compared with T cell tropic HIV. Virulent HIV often has syncytium formation capacity. Our analyses showed that highly variable amino acid sites are not only sites responsible for antigenicity but also sites that are responsible for cell tropism. This indicates that positive selection may operate on mutations that alter viral cell tropism and replication rate.

Seibert *et al.* (43) deduced from observation of HIV clones isolated from different individuals that positive selection may be

operating in the V3 region of HIV-1, because nonsynonymous substitution had a higher rate than synonymous substitution. For antigenic sites of the hemagglutinin gene of the influenza A virus, a higher rate of nonsynonymous substitution was reported (44) from a comparison of 21 strains. Some nonsynonymous substitutions that occur in both regions of these viral genomes might enable the virus to escape from recognition by the host immune system, because both regions react with antibodies of the host immune system. Almost all examples of genes examined so far, on which positive selection operates, were molecules that take part in the interaction between antigen and antibody (44–46). Therefore, it may be common for positive selection to operate in epitope regions of many parasites and viruses. However, from these observations, we do not know whether viral variants having mutations on epitopes are more infectious or advantageous for survival in the host. We do not know much about how transmitting variants are selected within a heterogeneous viral population within a host.

Several studies (47, 48) suggest the possibility that functional constraints on the envelope protein exist upon primary infection with HIV. On the other hand, some studies (49, 50) have suggested that there is no signature amino acid sequence of the envelope protein upon primary infection of HIV. Therefore, it is difficult to understand the mechanism of evolution of a viral population from comparisons of rates of synonymous and nonsynonymous substitutions among strains isolated from different hosts. It is necessary to investigate whether viral variants having mutations on epitopes can be more infectious or advantageous for survival within the host. Our study showed that there are particular periods of time in which positive selection may operate on the V3 region within a single host.

Rapid changes in the viral genome are obstacles to developing effective vaccines. However, an understanding of the population dynamics of the viral population within a single host and the empirically obtained pattern of amino acid substitutions in the principal neutralizing domain would be helpful in developing effective strategies for antiviral therapy.

This study was supported by the Proposal-Based Advanced Industrial Technology Research and Development Program from the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

1. Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988) *Science* **242,** 1168–1171.
2. Roberts, J. D., Bebenek, K. & Kunkel, T. A. (1988) *Science* **242,** 1171–1173.
3. Ricchetti, M. & Buc, H. (1990) *EMBO J.* **9,** 1583–1593.
4. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982) *Science* **215,** 1577–1585.
5. Gojobori, T. & Yokoyama, S. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4198–4201.
6. Gojobori, T. & Yokoyama, S. (1987) *J. Mol. Evol.* **26,** 148–156.
7. Temin, H. (1989) *Genome* **31,** 17–21.
8. Gojobori, T., Moriyama, E. N. & Kimura, M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 10015–10018.
9. Gojobori, T., Yamaguchi, Y., Ikeo, K. & Mizokami, M. (1994) *Jpn. J. Genet.* **69,** 481–488.
10. Saag, M. S., Hahn B. H., Gibbons, J., Li, Y., Parks, E. S., Parks, W. P. & Shaw, G. M. (1988) *Nature (London)* **334,** 440–444.
11. Ohno, S. & Yomo, T. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 1218–1222.
12. Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-staal, F. & Wolf, H. (1987) *J. Virol.* **61,** 570–578.
13. Goudsmit, J., Debouck, C., Meloen, R. H., Smit, L., Bakker, M., Asher, D. M., Wolff, A. V., Gibbs, C. J. & Gajdusek, D. C. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4478–4482.
14. Takahashi, H., Cohen, J., Hosmalin, A., Cease, K. B., Houghten, R., Cornette, J. L., DeLisi, C., Moss, B., Germain, R. N. & Berzofsky, J. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3105–3109.
15. Page, K. A., Stearns, S. M. & Littman, D. R. (1992) *J. Virol.* **66,** 524–533.
16. Grimaila, R. J., Fuller, B. A., Rennert, P. D., Nelson, M. B., Hammarskjold, M.-L., Potts, B., Murray, M., Putney, S. D. & Gray, G. (1992) *J. Virol.* **66,** 1875–1883.
17. Shioda, T., Levy, J. A. & Chen-Mayer, C. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9434–9438.
18. Freed, E. O., Myers, D. J. & Risser, R. (1991) *J. Virol.* **65,** 190–194.
19. de Jong, J. J., Goudsmit, J., Keulen, W., Klaver, B., Krone, W., Tersmette, M. & de Ronde, A. (1992) *J. Virol.* **66,** 757–765.
20. Holmes, E. C., Zhang, L. Q., Simmonds, P., Ludlam, C. A. & Brown, A. J. L. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4835–4839.
21. Wolfs, T. F. W., Zwart, G., Bakker, M., Valk, M., Kuiken, C. L. & Goudsmit, J. (1991) *Virology* **185,** 195–205.
22. McNearney, T., Hornickova, Z., Markham, R., Birdwell, A., Arens, M., Saah, A. & Ratner, L. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10247–10251.
23. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
24. Gojobori, T., Ishii, K. & Nei, M. (1982) *J. Mol. Evol.* **18,** 414–423.
25. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
26. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
27. Li, W. -H., Tanimura, M. & Sharp, P. M. (1988) *Mol. Biol. Evol.* **5,** 313–330.
28. Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Plans* (Soc. Ind. Appl. Math., Philadelphia).
29. Bonhoeffer, S., Holmes, E. C. & Nowak, M. A. (1995) *Nature (London)* **376,** 125.
30. Larder, B. A. & Kemp, S. D. (1989) *Science* **246,** 1155–1158.
31. Kellam, P., Boucher, C. A. B. & Larder, B. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1934–1938.
32. Wolinsky, S. M., Korber, B. T. M., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D., Safrit, J. T. & Koup, R. A. (1996) *Science* **272,** 537–542.
33. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U.K.).
34. Shioda, T., Oka, S., Ida, S., Nokihara, K., Toriyoshi, H., Mori, S., Takebe, Y., Kimura, S., Shimada, K. & Nagai, Y. (1994) *J. Virol.* **68,** 7689–7696.
35. Fouchier, R. A. M., Groenink, M., Kootstra, N. A., Tersmette, M., Huisman, H. G., Miedema, F. & Schuitemaker, H. (1992) *J. Virol.* **66,** 3183–3187.
36. Chesebro, B., Wehrly, K., Nishio, J. & Perryman, S. (1992) *J. Virol.* **66,** 6547–6554.
37. Rini, J. M., Stanfield, R. L., Stura, E. A., Salinas, P. A., Profy, A. T. & Wilson, I. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6325–6329.
38. Ghiara, J. B., Stura, E. A., Stanfield, R. L., Profy, A. T. & Wilson, I. A. (1994) *Science* **264,** 82–85.
39. Cheng-Mayer, C., Quiroga, M., Tung, J. W., Dina, D. & Levy, J. A. (1990) *J. Virol.* **64,** 4390–4398.
40. Phillips, R. E., Rowland-Jones, S., Nixon, D. F., Gotch, F. M., Edwards, J. P., Ogunlesi, A. O., Elvin, J. G., Rothbard, J. A., Bangham, C. R. M., Rizza, C. R. & McMichael, A. J. (1991) *Nature (London)* **354,** 453–459.
41. Niewiesk, S. & Bangham, C. R. M. (1996) *J. Mol. Evol.* **42,** 452–458.
42. Lukashov, V. V., Kuiken, C. L. & Goudsmit, J. (1995) *J. Virol.* **69,** 6911–6916.
43. Seibert, S. A., Howell, C. Y., Hughes, M. K. & Hughes, A. L. (1995) *Mol. Biol. Evol.* **12,** 803–813.
44. Ina, Y. & Gojobori, T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8388–8392.
45. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335,** 167–170.
46. Endo, T., Ikeo, K. & Gojobori, T. (1996) *Mol. Biol. Evol.* **13,** 685–690.
47. Zhang, L. Q., MacKenzie, P., Cleland, A., Holmes, E. C., Brown, A. J. L. & Simmonds, P. (1993) *Virology* **67,** 3345–3356.
48. Zhu, T., Mo, H., Wang, N., Nam, D. S., Cao, Y., Koup, R. A. & Ho, D. D. (1993) *Science* **261,** 1179–1181.
49. Kuiken, C. L., Zwart, G., Baan, E., Coutinho, R. A., van der Hoek, J. A. R. & Goudsmit, J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 9061–9065.
50. Antonioli, I. M., Baumberger, C., Yerly, S. & Perrin, L. (1995) *AIDS* **9,** 11–17.