Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 6693–6697, June 1996
Medical Sciences

Pattern of gp120 sequence divergence linked to a lack of clinical progression in human immunodeficiency virus type 1 infection

WEI-KUNG WANG*, MAX ESSEX*, MARY FRAN MCLANE*, KENNETH H. MAYER†‡§, CHUNG-CHENG HSIEH¶, HUNTER G. BRUMBLAY*, GEORGE SEAGE‖**, AND TUN-HOU LEE*††

*Department of Cancer Biology, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115; †Research Department, Fenway Community Health Center, Boston MA 02115; ‡Division of Infectious Diseases, Department of Medicine, Memorial Hospital of Rhode Island, Pawtucket, RI 02860; §Brown University AIDS Program, Providence, RI 02916; ¶Department of Epidemiology, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115; ‖Institute for Urban Health Policy and Research, Boston Department of Health and Hospitals, Boston, MA; and **Department of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA

Communicated by Elkan Blout, Harvard Medical School, Boston, MA February 20, 1996 (received for review June 6, 1995)

ABSTRACT **Differential rates of AIDS development and/or T4 lymphocyte depletion in HIV-1-infected individuals remain unexplained. The hypothesis that qualitative differences in selection pressure *in vivo* may account for different rates of disease progression was addressed in nine eligible study participants from a cohort of 315 homosexual men who have been followed since 1985. Disproportionately fewer changes in variable regions and more in C3 of gp120 were found to be significantly associated with slower disease progression. Our finding provides the first example to demonstrate that differential selection pressure related to the emergence of HIV-1 variants is associated with long term nonprogression. Candidate vaccines that elicit strong selection pressure on C3 of gp120 are likely to provide better protection than those targeting variable regions.**

The natural history of HIV-1 infection is characterized by an asymptomatic phase of variable duration (1–3). This phase is likely to represent a finely balanced equilibrium between the virus and the immune system (4). A critical issue in AIDS pathogenesis is why the eventual breakdown of this finely balanced steady state occurs at different rates in different HIV-1-infected individuals (5–8). The possibility that variations in virus–host interaction account for different rates of disease progression has not been studied in detail.

In the case of two other lentiviruses, equine infectious anemia virus and visna virus, antigenically distinct variants containing alterations in the envelope glycoprotein have been known to arise progressively during infection (9, 10). It has been suggested that this is a mechanism by which the virus evades the host immune response (11).

Intrasubject sequence divergence of the HIV-1 envelope gene is known to be influenced by positive selection pressure, a notion supported by the high ratio of nonsynonymous nucleotide substitutions per site (d_N) to synonymous nucleotide substitutions per site (d_S) (12–15). Sequence divergence of the envelope gene may therefore be studied to elucidate the virus–host interaction within HIV-1-infected individuals. This study addresses the hypothesis that qualitative differences in selection pressure, as reflected by different patterns of sequence divergence, may account for different rates of disease progression.

MATERIALS AND METHODS

Study Participants. Participants in this study were from a cohort of 315 homosexual men who enrolled at the Fenway Community Health Center in Boston between January 1985 and June 1986 (16). Of the 315 men, 85 were HIV-1 seropositive and 230 were HIV-1 seronegative at entry. Each study participant has been followed subsequently every 6 months for blood sampling and clinical evaluation. As of July 1994, 39 of the 85 men seropositive at entry were deceased, 13 had declined to continue in the study, and six were lost to follow-up. Of the 27 men who remained in the study, only six were asymptomatic and clinically classified at CDC-II (17). Among the 230 men seronegative at entry, 14 seroconverted during the follow-up period. Of these 14, three progressed to AIDS as of July 1994. The pattern of gp120 sequence divergence was investigated for all six of the asymptomatic participants from the original 85 seropositive men, as well as the three seroconverters who progressed to AIDS.

Virus Isolation. Virus isolates were obtained from the peripheral blood mononuclear cells (PBMC) of the nine study participants by co-cultivation with PBMC of HIV-1 seronegative donors as described previously (18). Briefly, PBMCs from the study participants were separated from heparinized whole blood by Ficoll-Hypaque density centrifugation and were cocultivated with 5×10^6 PBMC from an HIV-1-seronegative donor. HIV-1-negative PBMC were previously stimulated with phytohemagglutinin A for 3 days. Cocultivation was carried out in RPMI 1640 medium containing 20% heat-inactivated fetal bovine serum, 20 units/ml recombinant human interleukin 2 (Collaborative Research), and antibiotics. The cultures were fed with fresh media twice a week. Virus production in the culture supernatant was assessed twice a week using a p24 antigen capture ELISA (Dupont/NEN). When the optical density of virus-infected cultures was greater than twice the mean of negative control wells for two consecutive determinations, an additional 3×10^6 phytohemagglutinin A-stimulated PBMC was added to each culture and expanded for 72 h. Culture supernatants were collected as stock virus, and cells were pelleted for DNA analysis. The cultures were kept for an average of 17.5 days before proviral DNA was extracted.

PCR Amplification, Cloning, and Sequencing. High molecular weight DNA was purified by a standard proteinase K digestion–phenol extraction technique. Two micrograms of DNA was subjected to 30 cycles of PCR amplification, using 2.5 units of ampliTaq polymerase (Perkin-Elmer/Cetus) under conditions recommended by the manufacturer [10 mM Tris·Cl (pH 8.3), 50 mM KCl, 0.01% gelatin, 0.2 mM each of the four deoxynucleotide triphosphates, 1.5 mM $MgCl_2$, 100 pmol of each primer in a volume of 100 μl]. Oligonucleotide primers 117A, 5′-AAAGCCATGTGTAAAATTAACC-3′

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.

Abbreviations: d_N, nonsynonymous nucleotide substitutions per site; d_S, synonymous nucleotide substitutions per site; PBMC, peripheral blood mononuclear cells.

Data deposition: The sequences reported in this paper have been deposited in the GenBank data base (accession numbers U19694–U19711).

††To whom reprint requests should be addressed.

[corresponding to positions 6571–6592 of the HXB2 genome (19)] and 439B, 5′-ATGGGAGGGGCATACATTGC-3′ (positions 7539–7520) were designed to amplify a 0.97-kb fragment of gp120, starting from the V1 region to the N-terminal half of C4. Each cycle consisted of three steps: denaturation (94°C, 1 min), annealing (58°C, 1 min), and extension (72°C, 3 min), with a final extension at 72°C for 10 min. The error rate of ampliTaq enzyme under our PCR conditions was estimated to be about 1.64 per 1,000 nucleotides (18).

Each PCR product was cloned to TA cloning vector PCR-II (Invitrogen), which was then transformed to *Escherichia coli* JM109 cells. Multiple independent clones containing amplified products from each isolate were sequenced completely. For each participant in Group I, 10 independent clones of each of the two isolates were sequenced. For each participant in Groups II and III, five independent clones of each of the two isolates were completely sequenced.

Sequence Analysis and Calculation of d_N/d_S Ratio. Sequences were aligned and analyzed using GENEWORKS, version 2.0 (IntelliGenetics) Refinement of sequence alignments was carried out manually when necessary. Consensus sequences at both nucleotide and protein levels were generated for each individual isolate and were compared to determine the pattern of gp120 divergence. The program MEGA was used to calculate the ratio of d_N to d_S for each study participant according to the method of Nei and Gojobori (20).

RESULTS

Correlation Between the Rate of CD4 Decline and Clinical Grouping. As shown in Fig. 1, the nine participants were divided into three groups. Group I consisted of two participants, ID1 and ID2, who remained asymptomatic and whose absolute CD4 counts generally remained above 1000/mm^3 for at least 9 years after infection. Group II consisted of four participants, ID3 to ID6, who remained asymptomatic but had most of their absolute CD4 counts below 1000/mm^3. Each of these participants experienced a reduction in CD4 counts during the follow-up period. Group III consisted of the three seroconverters who progressed to AIDS over the course of the study. All participants within the cohort of 315 men who met these clinical criteria were included in this study.

In the continuum of clinical progression associated with HIV-1 infection, the participants in Group I and Group III are at opposing ends. The former are at the slower progression end of the spectrum, while the latter are at the faster progression end. The participants in Group II represent an intermediate group. In agreement with previous reports that CD4 count is generally correlated with disease status in HIV-1 infection, the correlation between the rate of CD4 loss per year for these nine participants and their clinical grouping was highly significant (Table 1; Spearman's rank correlation coefficient, $r_s = -0.811$; $P < 0.00006$).

Pattern of gp120 Sequence Divergence Over Time. To analyze gp120 sequence divergence over time, two sequential virus isolates were obtained from the PBMC of each study participant by short term co-cultivation with PBMC from HIV-1-seronegative donors (18). Since a substantial proportion of HIV-1 proviruses in PBMC or of plasma virions is likely to be defective (21–24), the approach of short term co-cultivation enhances the likelihood that the sequences analyzed are derived from infectious viruses. The intrasubject isolates analyzed in this study were separated by at least 1 year. This time interval should allow for sufficient rounds of virus replication and significant accumulations of divergent sequences (25, 26). For each virus isolate, a 0.97-kb segment of

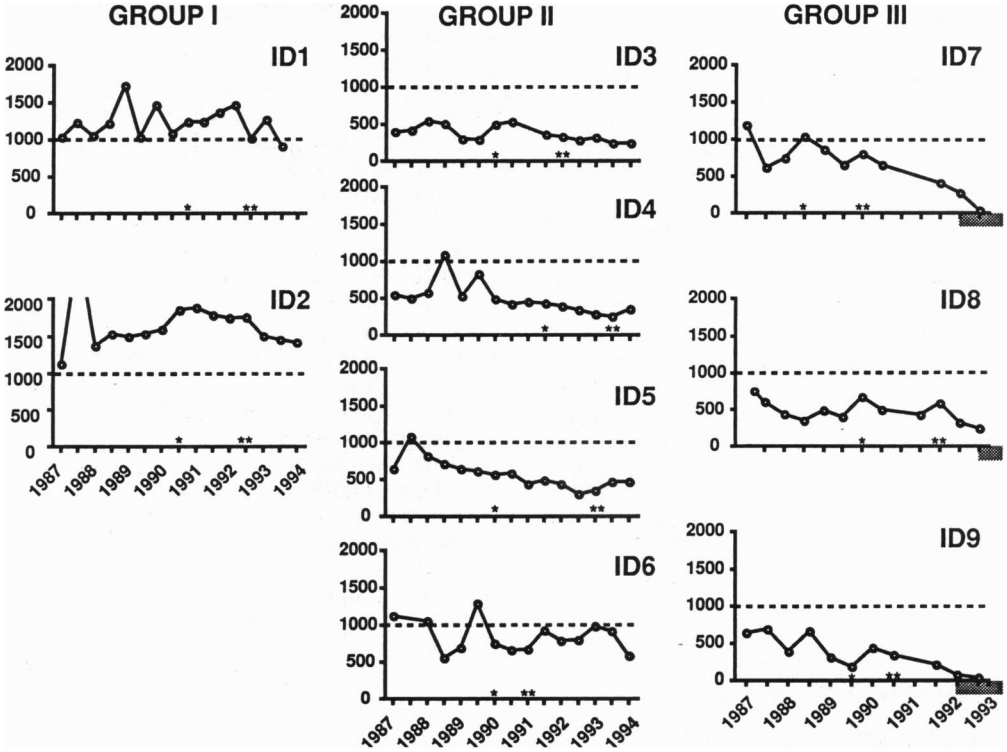

FIG. 1. CD4 counts (mm^{-3}) of HIV-1-infected study participants. Group I consists of asymptomatic HIV-1 seropositives with most of their CD4 counts generally above 1000/mm^3; Group II consists of asymptomatic HIV-1 seropositives with most of the CD4 counts below 1000/mm^3; and Group III is composed of HIV-1 seroconverters who have progressed to AIDS. Shaded areas indicate when patients were diagnosed with AIDS. Asterisks indicate the times at which isolates were derived. First isolates were obtained at 71, 66, 54, 65, 60, and 53 months after enrollment for ID1 through ID6, respectively (single asterisk). First isolates were obtained at 38, 39, and 20 months after seroconversion for ID7 through ID9, respectively (single asterisk). Second isolates were obtained at 97, 90, 77, 90, 96, and 66 months after enrollment for ID1 through ID6, respectively (double asterisks). Second isolates were obtained at 56, 63, and 33 months after seroconversion for ID7 through ID9, respectively (double asterisks).

Table 1. Regression analysis of CD4 loss over time in study subjects.

Group	ID	CD4 loss*/Year
I	1	−5.21
	2	−15.57
II	3	−29.33
	4	−61.39
	5	−70.01
	6	−26.13
III	7	−136.62
	8	−40.09
	9	−106.83

*Slope is based on a simple regression analysis of semiannual CD4 counts for each subject. Spearman's rank correlation coefficient, $r_s = -0.811$; $P < 0.00006$.

gp120 spanning the first hypervariable region, V1, to the N-terminal 16 residues of the fourth conserved region, designated C4′ in this study, was analyzed. Multiple independent clones from each isolate were sequenced. Two consensus protein sequences were derived for each study participant, one for each sequential isolate. These were compared to determine the pattern of intrasubject gp120 sequence divergence over time (Table 2).

The number of amino acid substitutions increased in direct proportion to the length of the time between sequential samples for all study participants with the exception of ID4 (Fig. 2*A*). Amino acid changes involving insertions and deletions occurred exclusively in the hypervariable regions of gp120 for all participants (Table 2). This is compatible with the current understanding of gp120 sequence diversity summarized by Myers *et al.* (27). The number of insertions and deletions was not correlated with the length of time between sequential samples (Fig. 2*B*). However, a significant correlation was observed between the number of insertions and deletions and the clinical status of study participants. Study participants who progressed to AIDS had more insertions and deletions per year than those who remained asymptomatic (Spearman's rank correlation coefficient, $r_s = 0.768$; $P < 0.00006$).

Influence of Selection Pressure on Sequence Divergence. To analyze whether the gp120 sequence divergence observed in the study participants was influenced by selection pressure, the d_N/d_S ratio for each study participant was calculated (20). This ratio is commonly employed in the study of adaptive evolution at the nucleotide level and can be used to infer both negative (purifying) selection pressure and positive (Darwinian) selection pressure. As shown in Table 3, d_N/d_S ratios for the study participants ranged from 0.784 to 2.551. No correlation was observed between d_N/d_S ratio and clinical status. However,

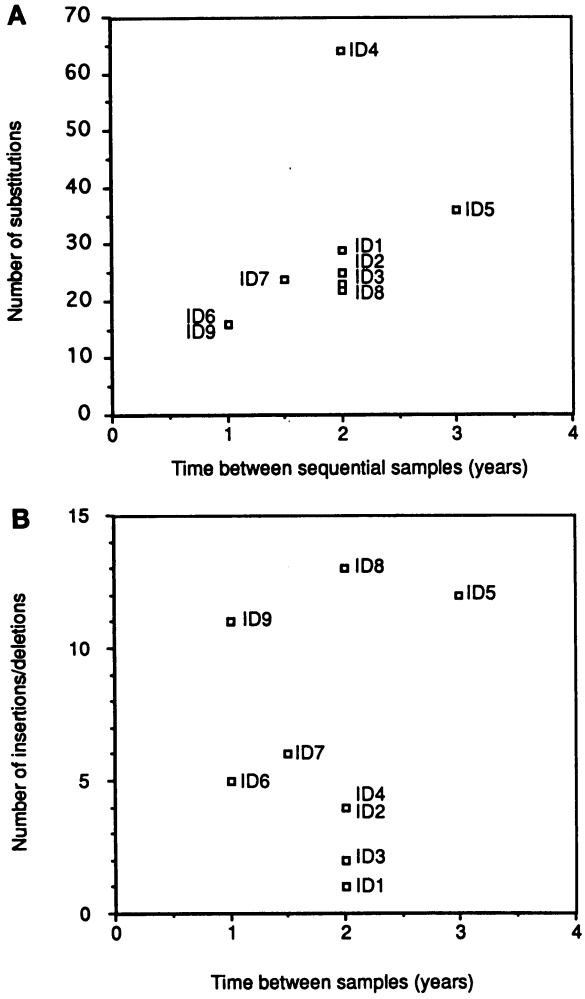

FIG. 2. The relationship between the number of amino acid changes and the length of time between sequential samples. (*A*) Substitutions; (*B*) insertions and deletions.

these ratios were significantly higher than the upper bound of d_N/d_S ratios calculated for 42 different mammalian genes (28). In addition, these intrasubject d_N/d_S ratios were higher than the average ratio, 0.44, reported for 119 intersubject HIV-1 gp120 sequences (27). Taken together, these comparisons suggest that the intrasubject gp120 sequence divergence was influenced by positive selection pressure.

Divergence in Constant Regions Linked to Slower Disease Progression. To determine if particular patterns of gp120 divergence correlated with the rate of disease progression among the study participants, we analyzed the relative proportions of

Table 2. Number and distribution of amino acid changes in gp120.

		Substitutions							Insertions/deletions			
Group	ID	Total (per year)	V1V2	C2	V3	C3	V4	C4′	Total (per year)	V1V2	V3	V4
I	1	29 (14.5)	12	7	0	6	3	1	1 (0.5)	1	0	0
	2	25 (12.5)	8	4	2	5	6	0	4 (2.0)	2	0	2
II	3	23 (11.5)	10	7	0	1	4	1	2 (1.0)	2	0	0
	4	64 (32.0)	25	8	13	4	12	2	4 (2.0)	4	0	0
	5	36 (12.0)	12	6	6	4	5	3	12 (4.0)	12	0	0
	6	16 (16.0)	8	0	0	5	3	0	5 (5.0)	5	0	0
III	7	24 (16.0)	15	3	0	2	4	0	6 (4.0)	5	0	1
	8	22 (11.0)	8	2	4	0	6	2	13 (6.5)	13	0	0
	9	16 (16.0)	4	5	4	1	2	0	11 (11.0)	8	3	0

Table 3. d_N and d_S gp120 nucleotide substitutions per site.

Group	ID	d_N	d_S	d_N/d_S*
I	1	0.0449	0.0176	2.551
	2	0.0406	0.0518	0.784
II	3	0.0356	0.0303	1.175
	4	0.1202	0.0938	1.281
	5	0.0580	0.0513	1.131
	6	0.0272	0.0242	1.124
III	7	0.0379	0.0303	1.251
	8	0.0379	0.0407	0.931
	9	0.0255	0.0206	1.238

*Spearman's rank correlation coefficient, $r_s = -0.376$; $P = 0.2897$.

amino acid substitutions in variable regions, V1 to V4 and in constant regions, C2 to C4′, and determined their association with the clinical status of the study participants. Disproportionately more amino acid substitutions in variable regions were significantly correlated with disease progression (Spearman's rank correlation coefficient, $r_s = 0.695$; $P = 0.00028$).

The relative proportion of amino acid substitutions within individual regions of gp120 was also analyzed to determine if any particular region contributed disproportionately to the correlation observed. As shown in Table 4, no correlation was found between clinical status and the proportion of amino acid substitutions in V1/V2, C2, V3, V4, or C4′. In contrast, disproportionately more amino acid substitutions in the C3 region were associated with slower disease progression (Spearman's rank correlation coefficient, $r_s = -0.613$; $P < 0.0004$).

DISCUSSION

Recent studies of HIV-1-infected patients treated with potent inhibitors of virus replication have shed new light on AIDS pathogenesis (25, 26). HIV-1 undergoes an extraordinarily rapid rate of turnover *in vivo*, which is estimated to be about 1 to 2 days (25, 26). Assuming a burst size of two (29), the number of progeny virions produced over the course of a year can in theory approach $2^{365/2}$, or 8.67×10^{54}. Such a high rate of HIV-1 replication can be expected to generate enormous genetic diversity *in vivo* and very likely contributes to the appearance of "quasi-species" (30, 31) and the temporal and spatial divergence found in infected individuals (21, 32, 33). Even taking into consideration that a fraction of progeny virions may be defective (24), it remains likely that a genetically divergent pool of HIV-1 variants can be established rather quickly postexposure (4). Furthermore, variants with growth advantage, such as drug-resistant variants (4, 34) or immune-escape variants (25), are expected to emerge from such a large pool of variants to initiate each new round of infection. With this model of AIDS pathogenesis, the asymptomatic phase of HIV-1 infection represents a finely balanced equilibrium between the virus and the immune system (4, 25, 26).

Table 4. Proportion of amino acid substitutions in different regions of gp120*.

Group	ID	V1V2	C2	V3	C3	V4	C4′
I	1	41.38	24.14	0.00	20.69	10.34	3.45
	2	33.33	16.67	8.33	20.83	20.83	0.00
II	3	43.48	30.43	0.00	4.35	17.39	4.35
	4	39.06	12.50	20.31	6.25	18.75	3.13
	5	33.33	16.67	16.67	11.11	13.89	8.33
	6	50.00	0.00	0.00	31.25	18.75	0.00
III	7	62.50	12.50	0.00	8.33	16.67	0.00
	8	36.36	9.09	18.18	0.00	27.27	9.09
	9	25.00	31.25	25.00	6.25	12.50	0.00
	r_s†	0.018	−0.166	0.368	−0.613	0.054	0.019
	P	0.9625	0.6449	0.2139	0.00040	0.9005	0.9602

*Proportion expressed as the percentage of total amino acid substitutions for each subject.
†Spearman's rank correlation coefficient.

It is known that the rate of disease progression in HIV-1 infection can vary from a few months to more than a decade (5–8). The basis of this difference remains elusive and has not been adequately explained under the currently proposed model of AIDS pathogenesis. However, our finding that selection pressure focused disproportionately more on C3 than on variable regions of gp120 is linked to slower disease progression provides an example that qualitative differences in selection pressure should be considered an important determinant of AIDS pathogenesis.

Selection pressure focused on variable regions of gp120 is likely to perpetuate the emergence of variants with disproportionately more changes in variable regions. Because variable regions are known to be associated with phenotypic changes of HIV-1 (35–39), selection pressure on these regions may increase the possibility that variants with higher potential to disturb the finely balanced steady state will emerge from a predictably large pool of HIV-1 variants. In this regard, the emergence of more cytopathic variants such as syncytium-inducing viruses, which contain certain positively charged amino acid residues in V3 of gp120, could be the result of selection pressure on variable regions (12, 40, 41). Compatible with this interpretation, isolates from two of the three progressors in this study, ID8 and ID9, switched from the non-syncytium-inducing to the syncytium-inducing V3 genotype and phenotype (data not shown). The greater number of insertions and deletions in the variable regions of isolates from progressors also supports this interpretation.

In contrast, selection pressure focused disproportionately less on variable regions but more on C3 of gp120, may be less likely to select for variants with the capacity to disturb this balance. In the event that the selection pressure focused on C3 is mediated by some as yet to be defined immune response, it is conceivable that such selection pressure can also be more effective than that focused on the variable regions to reduce virus load by preventing successive rounds of infection.

One of the difficulties in studying HIV-1 sequence divergence *in vivo* arises because the majority of infection appears to be in the lymphoid tissues (42, 43). Studies with a thorough sampling of infectious HIV-1 from lymphoid tissues are more likely to reveal the entire spectrum of sequence divergence, although such studies are logistically difficult to carry out. Our sampling of the viruses derived from a single cycle of *in vitro* co-cultivation increases the likelihood that the viruses analyzed were infectious and would continue to be subjected to new rounds of selection pressure *in vivo*. This protocol creates the possibility that only selected variants were included in the analysis, although there is no evidence to suggest that virus populations with opposite patterns of genetic divergence were selectively isolated from the study participants in Groups I and III. Another difficulty that must be overcome in this type of study is the selection bias associated with studying representative progressors and nonprogressors from a large cohort. Mindful of this pitfall, all eligible subjects from a single cohort were included in order to investigate the concept that qualitative differences in selection pressure are linked to the rate of disease progression in HIV-1 infection.

It is generally appreciated that prophylactic immunization against viral disease rarely protects the vaccinees from viral infection (44). Rather, immunization appears to offer sufficient memory response to prevent massive postexposure dissemination of viruses *in vivo* (44, 45). The ability to limit virus replication greatly ameliorates the pathological consequences of virus infection (46). Many HIV-1 envelope-based prophy-

lactic vaccines are currently under development. It is likely that the more achievable goal of these candidate vaccines is also to prevent massive dissemination of HIV-1 to distant lymphoid organs, rather than preventing infection per se. In this regard, our findings strongly suggest that candidate vaccines that elicit strong selection pressure on C3 of gp120 are likely to provide better protection than those targeting variable regions.

We thank B. Goldstein, B. Glaser, and the Fenway Community Health Center for invaluable support; Z. Matsuda and K. Chou for helpful discussions; and E. Conway for editorial assistance. This work was supported in part by Public Health Service Grants CA-39805 and HL-33774 from the National Institutes of Health, 17-90-C-0151 from the United States Army, and by a grant from the Commonwealth of Massachusetts Department of Public Health. W.-K.W. is supported by training grant 5 D43 TW00004 from the Fogarty International Center, National Institutes of Health.

1. Levy, J. A. (1993) *Microbiol. Rev.* **57,** 183–289.
2. Pantaleo, G., Graziosi, C. & Fauci, A. S. (1993) *N. Engl. J. Med.* **328,** 327–335.
3. Feinberg, M. B. & Greene, W. C. (1992) *Curr. Opin. Immunol.* **4,** 466–474.
4. Coffin, J. M. (1995) *Science* **267,** 483–489.
5. Isaksson, B., Albert, J., Chiodi, F., Furucrona, A., Krook, A. & Putkonen, P. (1988) *J. Infect. Dis.* **158,** 866–868.
6. Schrager, L. K., Young, J. M., Fowler, M. G., Mathieson, B. J. & Vermund, S. H. (1994) *AIDS* **8,** (Suppl. 1) S95–S108.
7. Cao, Y., Qin, L., Zhang, L., Safrit, J. & Ho, D. D. (1995) *N. Engl. J. Med.* **332,** 201–208.
8. Pantaleo, G., Menzo, S., Vaccarezza, M., Graziosi, C., Cohen, O. J., Demarest, J. F., Montefiori, D., Orenstein, J. M., Fox, C., Schrager, L. K., Margolick, J. B., Buchbinder, S., Giorgi, J. V. & Fauci, A. S. (1995) *N. Engl. J. Med.* **332,** 209–216.
9. Carpenter, S., Evans, L. H., Sevoian, M. & Chesebro, B. (1987) *J. Virol.* **61,** 3783–3789.
10. Narayan, O., Griffin, D. E. & Chase, J. (1977) *Science* **197,** 376–378.
11. Clements, J. E., Gdovin, S. L., Montelaro, R. C. & Narayan, O. (1988) *Annu. Rev. Immunol.* **6,** 139–159.
12. Wolfs, T. F. W., De Jong, J-J., Van Den Berg, H., Tijnagel, J. M. G. H., Krone, W. J. A. & Goudsmit, J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9938–9942.
13. Simmonds, P., Balfe, P., Ludlam, C. A., Bishop, J. O. & Leigh Brown, A. J. (1990) *J. Virol.* **64,** 5840–5850.
14. Lamers, S. L., Sleasman, J. W., She, J. X., Barrie, K. A., Pomeroy, S. M., Barrett, D. J. & Goodenow, M. M. (1993) *J. Virol.* **67,** 3951–3960.
15. Wolinsky, S. M., Wike, C. M., Korber, B. T. M., Hutto, C., Parks, W. P., Rosenblum, L. L., Kunstman, K. J., Furtado, M. R. & Munoz, J. L. (1992) *Science* **255,** 1134–1137.
16. McCusker, J., Stoddard, A. M., Mayer, K. H., Cowan, D. N. & Groopman, J. E. (1988) *Am. J. Pub. Health* **78,** 68–71.
17. Centers for Disease Control, (1987) *MMWR* **36,** (Suppl. 1), 3S–15S.
18. Wang, W-K., Essex, M. & Lee, T-H. (1995) *AIDS Res. Hum. Retroviruses* **11,** 185–188.
19. Ratner, L., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313,** 636–637.
20. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
21. Meyerhans, A., Cheynier, R., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B. & Wain-Hobson, S. (1989) *Cell* **58,** 901–910.
22. Martins, L. P., Chenciner, N., Asjo, B., Meyerhans, A. & Wain-Hobson, S. (1991) *J. Virol.* **65,** 4502–4507.
23. Li, Y., Kappes, J. C., Conway, J. A., Price, R. W., Shaw, G. M. & Hahn, B. H. (1991) *J. Virol.* **65,** 3973–3985.
24. Piatak, M., Saag, M. S., Yang, L. C., Clark, S. J., Kappes, J. C., Luk, K-C., Hahn, B. H., Shaw, G. M. & Lifson, J. D. (1993) *Science* **259,** 1749–1754.
25. 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., Saag, M. S. & Shaw, G. M. (1995) *Nature (London)* **373,** 117–122.
26. Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. & Markowitz, M. (1995) *Nature (London)* **373,** 123–126.
27. Myers, G., Korber, B., Wain-Hobson, S., Smith, R. F. & Pavlakis, G. N. (1994) *Human Retroviruses and AIDS 1994: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences* (Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM).
28. Li, W-H., Wu, C-I. & Luo, C-C. (1985) *Mol. Biol. Evol.* **2,** 150–174.
29. Wain-Hobson, S. (1993) *Nature (London)* **366,** 22.
30. Holland, J. J., de la Torre, J. C. & Steinhauer, D. A. (1992) *Curr. Topics Microbiol. Immunol.* **176,** 1–20.
31. Wain-Hobson, S. (1992) *Curr. Topics Microbiol. Immunol.* **176,** 181–193.
32. Epstein, L. G., Kuiken, C., Blumberg, B. M., Hartman, S., Sharer, L. R., Clement, M. & Goudsmit, J. (1991) *Virology* **180,** 583–590.
33. Delassus, S., Cheynier, R. & Wain-Hobson, S. (1992) *J. Virol.* **66,** 5642–5645.
34. Wain-Hobson, S. (1995) *Nature (London)* **373,** 102.
35. Shoida, T., Levy, J. A. & Cheng-Mayer, C. (1991) *Nature (London)* **349,** 167–169.
36. Hwang, S. S., Boyle, T. J., Lyerly, H. K. & Cullen, B. R. (1991) *Science* **253,** 71–74.
37. Groenink, M., Fouchier, R. A. M., Broersen, S., Klaver, B., Koot, M., van't Wout, A. B., Huisman, H. G., Miedema, F., Tersmette, M. & Schuitemaker, H. (1993) *Science* **260,** 1513–1516.
38. Koito, A., Harrowe, G., Levy, J. A. & Cheng-Mayer, C. (1994) *J. Virol.* **68,** 2253–2259.
39. Westervelt, P., Trowbridge, D. B., Epstein, L. G., Blumberg, B. M., Li, Y., Hahn, B. H., Shaw, G. M., Price, R. W. & Ratner, L. (1992) *J. Virol.* **66,** 2577–2582.
40. De Jong, J-J., De Ronde, A., Keulen, W., Tersmette, M. & Goudsmit, J. (1992) *J. Virol.* **66,** 6777–6780.
41. Fouchier, R. A. M., Groenink, M., Kootstra, N. A., Tersmette, M., Huisman, H. G., Miedema, F. & Schuitemaker, H. (1992) *J. Virol.* **66,** 3183–3187.
42. Pantaleo, G., Graziosi, C., Demarest, J. F., Butini, L., Montroni, M., Fox, C. H., Orenstein, J. M., Kotler, D. P. & Fauci, A. S. (1993) *Nature (London)* **362,** 355–358.
43. Embretson, J., Zupancic, M., Ribas, J. L., Burke, A., Racz, P., Tenner-Racz, K. & Haase, A. T. (1993) *Nature (London)* **362,** 359–362.
44. Wigzell, H. (1991) *FASEB J.* **5,** 2406–2411.
45. Ada, G. L. (1990) *Lancet* **335,** 523–526.
46. White, D. O. & Fenner, F. (1986) *Medical Virology* (Academic Press, Orlando, FL), pp. 134–143.