JOURNAL OF VIROLOGY, Mar. 1997, p. 2320–2330 Vol. 71, No. 3
0022-538X/97/$04.00+0


# Study of the Dynamics of Neutralization Escape Mutants in a Chimpanzee Naturally Infected with the Simian Immunodeficiency Virus SIVcpz-ant

PHILLIPE N. NYAMBI,[1] PAUL LEWI,[2] MARTINE PEETERS,[3] WOUTER JANSSENS,[1] LEO HEYNDRICKX,[1] KATRIEN FRANSEN,[1] KOEN ANDRIES,[2] MARLEEN VANDEN HAESEVELDE,[4] JONATHAN HEENEY,[5] PETER PIOT,[6] AND GUIDO VAN DER GROEN[1]*

*Institute of Tropical Medicine, Antwerp,[1] Center for Molecular Design, Janssen Research Foundation, Beerse,[2] and Innogenetics n.v., Ghent,[4] Belgium; ORSTOM/SIDA, Laboratoire de Rétrovirus, Montpellier, France[3]; BPRC, Laboratory of Viral Pathogenesis, Department of Virology, Rijswijk, The Netherlands[5]; and UNAIDS, Geneva, Switzerland[6]*

Received 21 August 1996/Accepted 2 December 1996

**Here we report on the use of spectral map analysis of time-paired sequential neutralization data of 11 serum samples of a chimpanzee naturally infected with a simian immunodeficiency virus (SIVcpz-ant) and 8 primary consecutive SIVcpz-ant isolates, taken at about 4-month intervals. The analysis reveals the existence of three SIVcpz-ant isolate and serum neutralization clusters. Each cluster groups virus isolates and/or sera based on similarities of their neutralization spectra. On average, neutralization escape mutants emerged after 15 months and mounted a neutralization response approximately 8 months later. The entire gp160 regions of eight consecutive isolates were sequenced and analyzed by a new statistical method called polygram, which allowed the deduction of amino acid sequence motifs of gp160 which were specific for SIVcpz-ant isolates belonging to the same isolate neutralization clusters. Changes in specific amino acid quadruplets in V1, V2, C3, V4, V5, and CD4 domains of gp120 and gp40 were seen to correlate with the neutralization clusters with most of the specific changes occuring in the V4 region. This method of analysis may facilitate an understanding of the study of the dynamic interplay between human immunodeficiency virus (HIV) and host neutralization responses as well as providing possible insights into mechanisms of persistence of HIV-1-related lentiviruses in their natural hosts.**

The simian immunodeficiency virus (SIVcpz), isolated from wild-captured, naturally infected chimpanzees, is the most closely related isolate of nonhuman origin to the human immunodeficiency virus (HIV) type 1 (20, 44, 53) and thus provides an opportunity to study the interactions between primate lentiviruses and their host. Isolates of HIV, even those taken from the same individual at different times postinfection, are genetically and biologically variable (12, 33, 42, 57, 58). In particular, individual variants of HIV differ in phenotype, and the appearance of more virulent strains within a patient has been associated with the development of AIDS (4, 7, 13, 48, 52). In HIV-1-infected persons, neutralization depends on the association of antibodies with various linear epitopes, such as those of the V3 loop, but importantly with other discontinuous and conformational epitopes of the HIV-1 envelope glycoproteins (11, 17, 18, 34, 35, 51). The SIVcpz-ant isolate was isolated from a naturally infected chimpanzee which was impounded by customs officers in Brussels upon illegal arrival in Belgium from Zaire (44) and, based on sequence information, is closely related to HIV-1 (53). How sequence variation in SIVcpz-ant envelope glycoproteins might affect the occurrence of neutralization escape mutants is not yet well documented. Insights into this question may provide a better understanding of the mechanism of HIV persistence.

HIV persistence is a complex kinetic phenomenon of simultaneous interaction of virus, cells, and humoral and cellular immune responses. According to Nowak and May (37), HIV persistence may be explained in terms of escape mutants emerging faster than the immune system can induce high-quality neutralizing antibodies. Despite the development of virus variants with reduced sensitivity to neutralization by time-paired sequential sera as observed in our previous study (45), a remarkably consistent nucleic acid sequence of the V3 region was observed in the consecutive SIVcpz-ant isolates (45). This suggested that escape from the humoral immune response could be due to changes in the envelope gene other than in the V3 loop. In the context of developing a protective vaccine capable of inducing neutralizing antibodies, the causal relationship between escape from neutralization and the eventual outgrowth of virus variants needs to be demonstrated. Furthermore, little is known about the time required to generate sufficient genetic diversity in the envelope protein of SIV isolates for neutralization escape mutants to emerge, as well as the time required to mount an adequate neutralization response against newly appearing virus variants in a naturally infected chimpanzee.

Recently, we applied spectral map analysis, a multivariate statistical method of data analysis (25–29), to study the inter- and intraclade cross-neutralization of primary HIV-1 isolates by their homologous and heterologous polyclonal sera from individuals infected with different genetic subtypes of HIV-1 groups M and O (40). In the present study, we apply this analysis to determine the kinetics of the appearance of SIVcpz-ant neutralization escape mutants and correlate neutralization escape with amino acid mutation patterns in the gp160 of sequential SIVcpz-ant strains. Sera and isolates were tested in a time-paired sequential manner in neutralization assays with human peripheral blood mononuclear cells (PBMCs) as target cells. The gp160 *env* sequences of the isolates were determined. The term "time-paired sequential" means that sera and

* Corresponding author. Mailing address: Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerp, Belgium. Phone: 32 3 247 63 19. Fax: 32 3 247 63 33.

isolates were collected from the same sample taken during a consecutive bleeding schedule from a particular individual. At intervals of about 15 months, neutralization escape mutants evolved, and these eventually formed three clusters. Approximately 8 months later, neutralizing antibodies were detected against the escape mutants. Furthermore, polygram analysis (8) revealed that changes in the V1, V2, C3, V4, V5, CD4, and gp40 domains of the envelope in the sequential isolates correlated with the neutralization escape mutants in each cluster.

## MATERIALS AND METHODS

**Chimpanzee sera and isolates.** The naturally infected chimpanzee (chimp NOAH) has been described previously (44, 45, 53) and is currently housed at the Biomedical Primate Research Center in Rijswijk, The Netherlands. This animal is approximately 9 years of age at the time of writing, maintains a profound thrombocytopenia, but is otherwise in relatively good health. Whole-blood samples as a source for virus isolates and polyclonal sera were obtained consecutively at 4-month intervals over a 41-month period. Eleven serum samples and eight virus isolates were jointly obtained as previously described (44, 45). Virus was isolated by cocultivation of PBMCs from the infected chimpanzee with 3-day-phytohemagglutinin-stimulated lymphocytes from healthy (HIV-negative) human donors.

**Neutralization assay.** The neutralizing capacity of the sequential sera to the sequential isolates was tested in a time-paired sequential manner in a classical neutralization assay with human PBMCs as target cells as previously described (45). Briefly, tissue culture supernatants of PBMC infected with virus (50 50% tissue culture infective doses) and twofold serial dilutions of heat-inactivated (56°C for 30 min) serum were incubated for 1 h at 37°C in a 5% $CO_2$ atmosphere before addition of $5 \times 10^4$ 3-day-phytohemagglutinin-stimulated PBMC for 2 h. The cells were washed and incubated in RPMI 1640–interleukin-2 medium (GIBCO, Paisley, United Kingdom) supplemented with 10% fetal calf serum, 0.03% L-glutamine, (Janssen Chimica, Geel, Belgium), and antibiotics. The presence of residual anti-p24 antibodies after the last washing was monitored carefully to minimize interference with the p24 antigen measurement. Extreme precautions were taken to guarantee the reproducibility of the neutralization data as described previously (39, 40). Viral replication was assessed after 7 days by an in-house p24 antigen capture enzyme-linked immunosorbent assay. Neutralizing titers of <1:10 were not considered significant and were scored as negative, $ID_{90}$ values were defined as the highest serum dilutions that produced ≥90% reduction in p24 antigen production.

**Spectral map analysis.** The sequential neutralization data were subjected to spectral map analysis. This multivariate statistical method of data analysis has been used previously to study the interaction between drugs and receptors (24, 25) and between viruses and antiviral compounds (2) and to study data on inter- and intraclade neutralization of primary isolates of HIV-1 by polyclonal sera of individuals infected with different HIV-1 genetic subtypes (A to H) in group M and O viruses (40). For this analysis, logarithms were taken from the reciprocal neutralizing-antibody titers. Logarithms of the titers were then double-centered to determine the specificity of interaction between sera and isolates. We applied principal-component analysis to the logarithmic and double-centered data to determine the modes of interaction and then plotted the results graphically.

A detailed description of the method of spectral map analysis has been provided in a previous report (40), and only the main steps are briefly discussed here.

**(i) Step 1.** Logarithmic transformation of the original dilution titers ($x$) into neutralization activities ($y$) is given as

$$y_{ij} = \log x_{ij}, \text{ where } i = 1, \ldots, n \text{ and } j = 1, \ldots, p$$

($n$ is the number of sera, and $p$ is the number of isolates). The logarithmic transformation is a standard procedure in bioassays.

**(ii) Step 2.** Double-centering of neutralization activities ($y$) into neutralization specificities ($z$) is given as

$$z_{ij} = y_{ij} - mr_i - mc_j + mg$$

where

$$mr_i = \frac{1}{p}\sum_j^p y_{ij}, \text{ the } i\text{th row mean of } y$$

$$mc_j = \frac{1}{n}\sum_i^n y_{ij}, \text{ the } j\text{th column mean of } y$$

$$mg = \frac{1}{np}\sum_i^n \sum_j^p y_{ij}, \text{ the global mean of } y$$

Double-centering corrects for differences in average neutralization activity of the sera and of the isolates. As a result, the row, column, and global means of the doubly centered dots are all equal to zero.

**(iii) Step 3.** Extraction of eigenvectors ($u$ and $v$), also called principal components or factors, from the table of neutralization specificities ($z$) is given as

$$u_{ik} = \sum_j^p z_{ij}\, v_{jk}\, \lambda_k^{-1} \text{ with } k = 1, \ldots, r$$

$$v_{jk} = \sum_i^n z_{ij}\, v_{ik}\, \lambda_k^{-1}$$

where $r$ is the number of eigenvectors that have been extracted. The coefficients $\lambda$ are the square roots of the associated eigenvalues, which can be interpreted as the variances in the table of neutralization specificities ($z$) accounted for by the corresponding eigenvectors ($u$ and $v$). Furthermore, eigenvectors ($u$ and $v$) are required to be orthonormal, which can be expressed by means of the conditions

$$\sum_i^n u_{ik}\, u_{ik'} = \delta_{kk'}$$

$$\sum_j^p v_{jk}\, v_{jk'} = \delta_{kk'}$$

where $\delta_{kk} = 1$ if $k = k'$ and $\delta_{kk} = 0$ otherwise. The eigenvectors ($u$ and $v$) are the principal components or factors of the neutralization specificities ($z$) and are, by convention, arranged in decreasing order of their contributions ($\lambda^2$) to the global variance of specificities. The calculations can be performed by means of standard programs for eigenvector extraction, which are available in most scientific programming libraries.

**(iv) Step 4.** Scaling of eigenvectors ($u$ and $v$) into coordinates of the sera ($s$) and the isolates ($l$) in factor space is given as

$$s_{ik} = u_{ik}\, \lambda_k^{1/2}, \text{ factor coordinates or scores of sera}$$

$$l_{jk} = v_{jk}\, \lambda_k^{1/2}, \text{ factor coordinates or loadings of isolates}$$

Factor scores and factor loadings are compiled into a factor table, which provides the corresponding coordinates in factor space for each serum and each isolate. The horizontal and vertical axes of a factor plot of the sera and of the isolates usually represent the first two most important factors. A third factor can be represented on such a plot by varying the thickness of the contours of the symbols (circles for sera and squares for isolates, by convention).

**gp160 sequence analysis.** For 10 SIVcpz-ant strains, the provirus DNA gp160 sequence was determined. In addition, the V3 sequences of the sequential isolates in viral RNA of plasma were determined. RNA was extracted from 100 μl of plasma (5) and reverse transcribed (9, 54). A 277-bp fragment encompassing the V3 loop was amplified and sequenced as described previously (45).

To determine the gp160 sequence from the proviral DNA, virus was isolated by cocultivation of PBMCs from the chimpanzee with phytohemagglutinin-stimulated lymphocytes from healthy (HIV-negative) human donors, as described previously (45). Primers for PCR of the envelope gene were designed by using sequence information of the first sequenced isolate (Nh1) (53). A nested PCR was carried out in a 100-μl reaction mixture containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2.5 mM $MgCl_2$, 0.01% gelatin, 0.2 mM each deoxynucleoside triphosphate, 0.4 μM each primer, and 2.5 U of *Taq* DNA polymerase (Perkin-Elmer Cetus Corp., Zaventem, Belgium). An aliquot corresponding to 1 μg of DNA was subjected to a first PCR round of 35 cycles with the outer primers 93C285 (5′ CCTCCATGGTTACACCCTCCTGC 3′) and SIV3968 (5′ GGAA CTTGAGGTCTCACAGGAAAGCC 3′). Cycle conditions were 1 min each at 94 and 50°C and 5 min at 72°C, with an extension of 7 min at 72°C in a termal DNA cycler machine (Perkin-Elmer Cetus Corp.). Amplified DNA (5μl) was subjected to a second round of 25 cycles with conditions as noted above with inner primers SCE1601 (5′ ggtctagAATAGAAGAAGATGAGGAAGCCG 3′) and SIVE3776 (5′ ccctcgagATTTTAGACCATGCAGAACCCAT 3′). Restriction sites (indicated in lowercase type) were added to the 5′ end of the primers for cloning purposes.

For three isolates (Nh3, Nh8, and Nh9), the total envelope gene was amplified successfully. Because of failure to amplify the remaining isolates with the nested primer combination described above, PCR fragments were obtained for the gp120- and part of the gp41-encoding envelope region and for the gp41-encoding envelope region, seperately (isolates Nh2, Nh4, Nh5, Nh6, Nh7, and Nh10). For the envelope fragments encoding gp120 and part of gp41, the outer primers were 93C285 and 93C284 (5′ GGTGCTGAGATTGTTCTCCAGATGG 3′) and the inner primers were SCE1601 and SCE1602 (5′ ggaagcttGGTGTTGGAGGAC TGCCC 3′). For the envelope fragments encoding envelope gp41, the outer primers were SIV3384 (5′ TCATTTCAGATCCCTACCCAAAACCAG 3′) and

TABLE 1. Neutralizing antibody titers of sequential serum samples from a chimpanzee with SIVcpz-ant infection to sequential SIVcpz-ant isolates[a]

| Time of serum collection (mo) | Titer[b] for isolate obtained at time (mo): | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0 | 7 | 15 | 18 | 22 | 25 | 29 | 33 | Mean |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 20 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.5 |
| 7 | 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.6 |
| 15 | 160 | 80 | 40 | 40 | 40 | 1 | 10 | 1 | 13 |
| 18 | 80 | 160 | 80 | 40 | 80 | 40 | 1 | 1 | 25 |
| 22 | 40 | 80 | 160 | 80 | 40 | 1 | 1 | 1 | 14 |
| 25 | 40 | 80 | 160 | 40 | 20 | 20 | 1 | 1 | 17 |
| 29 | 40 | 160 | 320 | 80 | 40 | 20 | 1 | 1 | 25 |
| 33 | 80 | 160 | 160 | 40 | 320 | 20 | 40 | 20 | 67 |
| 38 | 40 | 320 | 160 | 80 | 80 | 40 | 20 | 20 | 62 |
| 41 | 80 | 320 | 160 | 40 | 160 | 40 | 40 | 40 | 80 |
| Mean | 37 | 38 | 35 | 18 | 21 | 6.2 | 2.6 | 2.4 | 13 |

[a] Data from reference 45.

[b] Data recorded are $ID_{90}$ neutralization values, i.e., the reciprocal serum dilutions causing 90% neutralization of the primary SIVcpz-ant isolates. Titers lower than 10 are depicted as 1. Sera and isolates were named according to the time of collection (in months) after initiation of the study. The sera collected at the start of the study (time 0) and at 1, 7, 15, 18, 22, 25, 29, 33, and 41 months were tested for neutralizing antibodies against the corresponding SIVcpz-ant isolates obtained at the same moment. The corresponding isolates at 3 months (1, 38, and 41) were not included in the analysis.

SIV3968 and the inner primers were BSIV3456 (biotin-5′ AAAGACAGGATC AGGTGGAGGG 3′) and SIV3776. Cycle conditions were 1 min each at 94, 50, and 72°C, with an extension of 5 min at 72°C. Amplification products were cloned and sequenced on a A.L.F. automatic sequencer (Pharmacia Biotech), except for the gp41 amplification products, for which direct sequencing was done.

**Polygram analysis.** The entire gp160 sequences of 10 serial SIVcpz-ant strains were scanned for the presence of specific quadruplets formed by four adjacent amino acids as previously described (8). A quadruplet is considered specific for a given sequence if it occurs in that sequence and only rarely or not at all in all other sequences.

The measure of specificity of a given quadruplet i for a sequence of the particular type $k$ is given by means of a chi-square statistic ($\chi^2$):

$$\chi^2_{ik} = \frac{N(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

where $N$ is the total number of sequences in the set that is studied, $a$ is the number of sequences of type $k$ in which quadruplet $i$ occurs, $b$ is the number of sequences of type $k$ in which quadruplet $i$ does not occur, $c$ is the number of sequences not of type $k$ in which quadruplet $i$ occurs, and $d$ is the number of sequences not of type $k$ in which quadruplet $i$ does not occur.

Quadruplets of amino acids are preferred because they were previously found to be the most sensitive in representing specific amino acid changes in the test sequence (8). To visualize the pattern of specific quadruplets, we represented each SIVcpz-ant gp160 sequence by a horizontal line segment. If the specificity of a particular amino acid quadruplet is significant, a vertical mark was placed along the line at the position where the quadruplet occurred in the test sequence. This process was repeated for each of the 10 serial gp160 SIVcpz-ant sequences. Finally, we obtained a set of 10 parallel tracings, which is called a polygram, because it resembles the tracings produced by a polygraph recorder. Similarities among the 10 serial gp160 SIVcpz-ant sequences, in terms of specific amino acid quadruplet distribution, can be identified visually or by counting the number of significant quadruplets for each gp160 SIVcpz-ant sequence studied. This method was extended to SIVcpz-ant gp160 sequences of isolates which belong to the same SIVcpz-ant isolate neutralization cluster. In this extended case, an amino acid quadruplet is specific for a given isolate neutralization cluster, if it occurs in most of the members of that neutralization cluster and rarely or not at all in the members of other isolate neutralization clusters. This extended polygram approach allowed us to identify amino acid quadruplet patterns that are specific to the previously determined isolate neutralization clusters.

**Nucleotide sequence accession numbers.** The nucleotide sequence data were deposited in the EMBL, GenBank, and DDBJ nucleotide sequence databases.

## RESULTS

**Serial SIVcpz-ant isolate-serum neutralization spectra.** The serial SIVcpz-ant isolate-serum neutralization data depicted in Table 1 has been published previously (45) and is presented here for this analysis. The first 11 rows and first 8 columns represent the serum and isolate neutralization spectra, respectively, with progression of time. The (geometrical) mean neutralization titers observed from each serum sample are listed in the right-hand column of the table. These values provide a measure of the neutralizing potency of the corresponding sera. Note the increasing progression of the neutralizing potency of the sera between months 0 and 41. Initially, at month 0, the serum did not possess any neutralizing capacity. As time progressed, the serum became increasingly more capable of neutralizing isolates. The (geometrical) mean neutralization titers of each isolate are shown in the bottom row of the table. These constitute a measure of the sensitivity for neutralization of the corresponding isolates. We observe that the sensitivity to neutralization of the isolates varied over time.

Two spectra are considered to be similar if they agreed according to their shape, i.e., independently of the potency of the serum or the sensitivity of the isolate. For example, the spectra of isolates at months 7 and 15 are rather similar if one makes allowance for their difference in sensitivity. On the other hand, the spectra of isolates at 15 and 29 months are different. The same reasoning can be applied to the spectra of the sera at 33 and 41 months, which are similar, and to those at months 15 and 41 which are very different.

**Spectral map analysis of time-paired sequential sera and isolates.** The neutralization data ($ID_{90}$) in Table 1 were subjected to spectral map analysis. The logarithmic titers were double-centered, and a table of specificities (Table 2) was obtained. Double-centering consisted of subtracting simultaneously the corresponding row and column means from each element of the table. It is seen as a correction of the titers for differences in potency of the sera and for differences in sensitivity of the isolates. Note that specificities are relative values which add to zero both row- and columnwise. The table of specificities indicates the specificity (positive or negative) of the interactions between sera and isolates. Positive specificity

TABLE 2. Specificities of interaction between sequential SIVcpz-ant sera and isolates

| Time of serum collection (mo) | Specificity[a] for interaction with isolate obtained at time (mo): | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0 | 7 | 15 | 18 | 22 | 25 | 29 | 33 | Mean |
| 0 | −1.1 | −1.1 | −1 | −0.3 | −0.51 | 0.72 | 1.6 | 1.7 | 0.00 |
| 1 | 1.6 | −1.5 | −1.4 | −0.7 | −0.89 | 0.35 | 1.2 | 1.3 | 0.00 |
| 7 | 2.2 | −1.5 | −1.5 | −0.8 | −0.98 | 0.26 | 1.1 | 1.2 | 0.00 |
| 15 | 1.5 | 0.74 | 0.11 | 0.8 | 0.61 | −1.8 | −1 | −0.9 | 0.00 |
| 18 | 0.12 | 0.8 | 0.16 | 0.16 | 0.67 | 1.2 | −1.6 | −1.5 | 0.00 |
| 22 | 0 | 0.65 | 1.4 | 1.4 | 0.52 | −1.9 | −1 | −1 | 0.00 |
| 25 | −0.2 | 0.45 | 1.2 | 0.51 | −0.37 | 0.87 | −1.3 | −1.2 | 0.00 |
| 29 | −0.6 | 0.8 | 1.6 | 0.86 | 0 | 0.52 | −1.6 | −1.5 | 0.00 |
| 33 | −0.9 | −0.2 | −0.2 | −0.9 | 1 | −0.5 | 1.1 | 0.45 | 0.00 |
| 38 | −1.5 | 0.57 | 0 | 0 | −0.25 | 0.29 | 0.48 | 0.54 | 0.00 |
| 41 | −1.1 | 0.31 | −0.3 | −1 | 0.18 | 0 | 0.91 | 0.98 | 0.00 |
| Mean | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

[a] The specificities were obtained by taking logarithms and double-centering of the neutralizing antibody titers in Table 1. Note that the mean specificity of all sera and all isolates is zero. This is due to the operation of double-centering, as described in Materials and Methods. Double-centering corrects for differences of mean specificity between sera and between isolates.

TABLE 3. Factor coefficients for sequential SIVcpz-ant isolates and sera[a]

| Time (mo) | Factor | | |
|---|---|---|---|
| | I | II | III |
| Isolates | | | |
| 0.00 | −0.25 | −1.68 | −0.28 |
| 7 | 0.9 | 0.34 | 0.2 |
| 15 | 1.07 | 0.27 | 0.08 |
| 18 | 0.75 | −0.26 | 0.12 |
| 22 | 0.42 | 0.17 | 0.53 |
| 24 | −0.28 | 0.59 | −1.45 |
| 29 | −1.32 | 0.32 | 0.47 |
| 33 | −1.29 | 0.24 | 0.34 |
| Sera | | | |
| 0.00 | −1.08 | 0.68 | −0.01 |
| 1 | −1.2 | −0.71 | −0.31 |
| 7 | −1.23 | −1.03 | −0.38 |
| 15 | 0.68 | −1.07 | 0.73 |
| 18 | 0.8 | 0.01 | −0.88 |
| 22 | 1.05 | −0.37 | 0.93 |
| 25 | 0.79 | 0.13 | −0.75 |
| 29 | 1.14 | 0.23 | −0.52 |
| 33 | −0.35 | 0.54 | 0.71 |
| 38 | −0.12 | 0.86 | 0.13 |
| 41 | −0.47 | 0.73 | 0.36 |

[a] The factor table, obtained by spectral map analysis of Table 1, defines the coordinates (factor loadings) of eight isolates and (factor scores) of 11 sera in terms of the first three most important factors (I, II, and III) for the isolates and sera. The method by which the factor scores and factor loadings were derived from the original table of neutralization titers is described in Materials and Methods.

indicates that neutralization is greater than would be expected from the corresponding potencies and sensitivities. Likewise, negative specificity indicates that neutralization is smaller than expected.

To determine the principal modes of interaction between sera and isolates, we applied the principal-component analysis, referred to as factor analysis, to the data in Table 2. These modes are often referred to as factors, principal components, eigenvectors, etc. The number of factors may give an indication, for example, of the number of different epitopes (or antigenic determinants) involved in neutralization. The importance of a factor is defined by the amount of variance that can be attributed to it. Variance is defined here as the mean square specificity in Table 2. By convention, factors are arranged in descending order of their contribution to the variance. In the case of the serial neutralization data, we find that the first four factors account respectively for 61, 20, 13, and 4% of the variance. A residual 2% of the variance is distributed over the three higher-order factors. The result of the analysis, referred to as factor coefficients, is summarized in Table 3 for the first three dominant factors. The coefficients in Table 3 indicate how much each isolate "loads" and how much each serum "scores" on each factor. Each factor represents a mode of interaction between the sera and the isolates. It is therefore a common dominant factor which is shared by both the sera and the isolates. A mathematical property of the two-factor tables is that their algebraic product reproduces the data in the table of specificities (Table 2).

The factor coefficients are then used as coordinates that represent isolates and sera (Table 3) of the corresponding points in a rectangular (Cartesian) diagram referred to as a spectral map, as shown for isolates (Fig. 1) and for sera (Fig. 2). The horizontal and vertical axes of the spectral maps are defined by the first two factors. A third factor, which is perpendicular to the plane of the diagram, is encoded by the variable thickness of the symbols. According to this convention, the isolate at month 25 is located markedly below the plane of Fig. 1. The serum of month 22 is also positioned markedly above the plane of Fig. 2. (The exact locations of the isolates and sera are defined by their coordinates in Table 3.) For reasons that will be pointed out below, we attach no relevance here to the third- and higher-order factors. By convention of the spectral map, we represent isolates by means of squares and sera by means of circles. The areas of the squares and circles indicate the sensitivity and potency of the corresponding isolates and sera, respectively. In this case, the areas of the symbols were made proportional to the corresponding (geometrical) mean neutralization titers (Table 1).

If the spectra of two isolates are similar (in shape, irrespective of a difference in sensitivity), their positions will also be close on the spectral map (this applies to the isolates at months 7 and 15) (Fig. 1). If they are dissimilar (as is the case with isolates at months 15 and 29), they will appear to be separated. Likewise if the spectra of two sera are similar (in shape, irrespective of a difference in potency), the sera will be found close together on the spectral map (this is the case for sera at months 33 and 41) (Fig. 2). Spectra that are dissimilar (such as the sera at months 15 and 41) will be found at a distance from one another.

Not all computed factors are relevant to the analysis of the data. Some may explain the systematic variation in the data, and these are called structural factors. Others are due to random and nonsystematic variation in the data, and these are called noise factors. The number of structural factors in the data was determined by internal validation (1) by the leave-

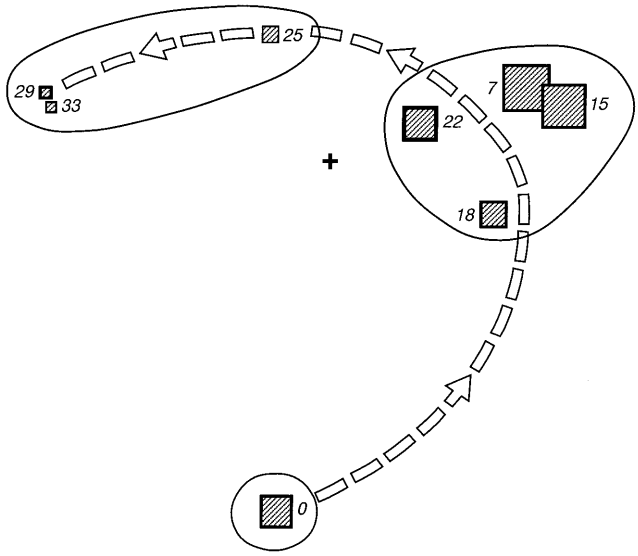


FIG. 1. Spectral map of eight consecutive SIVcpz-ant isolates. Squares represent individual isolates. The areas of the squares are indicative of the sensitivity to neutralization of the isolates, as defined by their geometrical mean titers in the bottom row of Table 1. The horizontal and vertical axes of the diagram represent the two most important factors of the specificities between sera and isolates in Table 2. As explained in the text, these two factors account for the structural information in the data, i.e., the information that is not caused by measurement errors and artifacts. They account for 61 and 20% of the variance, respectively. Other conventions are as described in the legend to Fig. 1. Three clusters of isolates (at 0, 7 to 22, and 25 to 33 months) have been determined by the UPGMA method of agglomerative cluster analysis. These clusters are specifically neutralized by three groups of sera (of 0 to 15, 15 to 29, and 33 to 41 months, respectively), as explained in the text. t, origin of the coordinate axes.

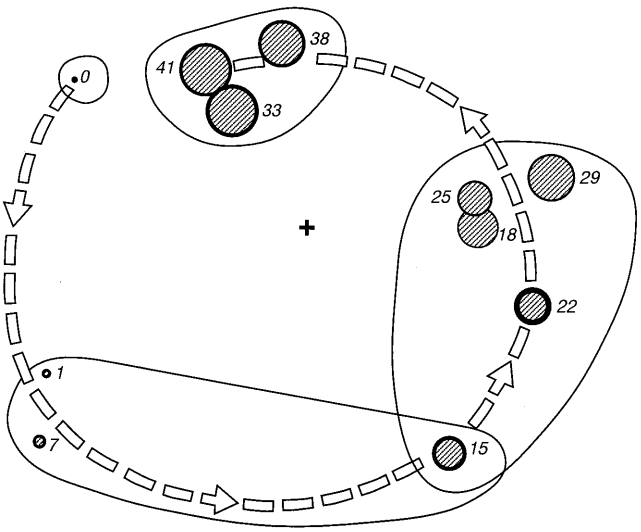


FIG. 2. Spectral map of 11 SIVcpz-ant consecutive sera. Circles represent individual sera. The areas of the circles are indicative of the neutralization potency of the sera, as defined by their geometrical mean titers in the rightmost column of Table 1. The horizontal and vertical axes of the diagram represent the two structural factors of the specificities between sera and isolates in Table 2. They account for 61 and 20% of the variance, respectively. A third nonsystematic factor, accounting for 13% of the variance, is encoded in the variable thickness of the contours of the circles. The coordinate axes are not shown explicitly. The small cross at the center defines the origin of the coordinate axes. The labels indicate the month at which the serum was taken. An arc has been added to indicate the progression of time. The grouping of the sera is according to their specificity for the clusters of isolates determined by the UPGMA method (Fig. 1). Note that the serum at 15 months is shared by two groups, as it shows mixed specificity for the cluster defined by the isolate at 0 months and by the cluster which comprises isolates at 7 to 22 months. The grouping of the isolates on this figure has been determined from inspection of Table 2, and the result is shown in Fig. 3. t, origin of the coordinate axes.

one-out method and by predicting their values by means of spectral map analysis, as described above, using only the data that remain. The process was repeated until predictions for all elements in the data table were obtained. When this was completed, the error between the observed and predicted values was plotted against the number of factors that were extracted (which ranges from 1 to 7 in this case). The error statistic is called PRESS (prediction error sum of squares) (1, 40). Initially, the error decreases because the first factors usually account for a large part of the structural variance in the data. The higher-order factors, however, add to the error, since they contain mostly noise. Hence, a plot of PRESS against the number of factors usually shows a minimum (plot not shown). The factor which produces minimal PRESS contains about equal parts of structural information and noise. Hence, structural factors are those that precede (in terms of contribution to the variance) the factor with minimal PRESS. In this case, we found that only the first two factors (accounting together for 81% of the variance) carry structural information. Hence, the third- and higher-order factors (which account together for 19% of the variance) are due to noise and nonsystematic variation in the data. The latter is caused by experimental error and some anomalies in the neutralization data mentioned above.

**Isolate and serum neutralization cluster analysis.** To understand the temporal relationships between isolates and sera, we first performed a hierarchiral (agglomerative) cluster analysis on the isolates to determine which of these can be grouped together. Cluster analysis by the unweighted pair-group mean arithmetic method (UPGMA) was applied, using only the two structural factors that were identified above (50). To this effect, only the first two columns of the factor table for isolates (Table 3) were submitted to cluster analysis. This effectively eliminated the noise and the nonsystematic variation in the data. The UPGMA analysis yielded three clusters (at 61% of the maximal agglomeration distance). These three clusters are composed of the isolate at 0 months, the isolates between 7 and 22 months, and the isolates between 25 and 33 months, respectively, as distinguished by contours on Fig. 1. Next, we defined the sera that are positively specific for these three clusters of isolates. For this purpose, we computed the mean specificities for the three clusters of isolates from the data in Table 2. The result is displayed in Fig. 3, which shows the mean specificity obtained at different times of serum collection (ranging from 0 to 41 months) for each of the three clusters of isolates. We observed three rather disjoined periods. The first period extended between 1 and 15 months, during which the sera showed positive specificity for isolate 0 and negative specificity for all other isolates. The second period ranged from 15 to 29 months, during which the sera possessed positive specificity for isolates taken between 7 and 22 months. Finally, the third period started at 33 months until the end of the observation period at 41 months, during which the serum showed positive specificity for isolates at months 25 to 33. Note that there is an overlap between the first two groups of sera at month 15. As can be seen from Table 2, the serum at month 15 is positively specific for all isolates between months 0 and 22.

The spectral maps for isolates and sera (Fig. 1 and 2) can also be superimposed, such that the origin (small cross at the center) and coordinate axes coincide. This produces a joint spectral map which represents both sera and isolates on the same diagram and provides a graphic interpretation of the interactions between sera and isolates. The origin of the map represents the point of zero specificity. Sera and isolates that are at a distance from the center and that point in the same direction possess positive specificity; similarly, sera and isolates in opposite directions indicate negative specificity. The group of sera at 1 to 15 months and the isolate at 0 months are

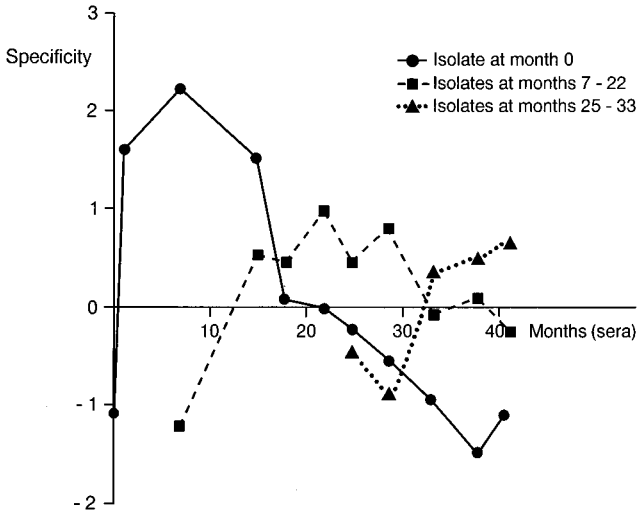


FIG. 3. Average specificity of the consecutive SIVcpz-ant sera for the three clusters of consecutive SIVcpz-ant isolates. The average specificity, as determined from Table 2, of the sera collected between 0 and 41 months for three clusters of consecutive isolates (at 0, 7 to 22, and 25 to 33 months, respectively) is shown. Specific sera for each of the three clusters of isolates have developed at times (months) when the corresponding specificity value is positive.

TABLE 4. Correlation between neutralization clusters of SIVcpz-ant isolates and groups of SIVcpz-ant sera[a]

| Cluster (mo) of isolates | Group (mo) of sera |
|---|---|
| 0 (0) | 1–15 (8) |
| 7–22 (14.5) | 15–29 (22) |
| 25–33 (29) | 33–41 (37) |

[a] The correlation was derived from the results of spectral map analysis (Fig. 3). Intervals refer to the months after the start of the experiment during which the isolates and sera where obtained. Values in parentheses represent the midpoints of the intervals.

positioned at a distance from the center and in the same direction. Hence, they possess a specific neutralizing interaction. Likewise, the group of sera at 18 to 29 months is positively specific for the cluster of isolates at 7 to 22 months. Finally, positive specificity is displayed between the group of sera at 33 to 41 months and the cluster of isolates at 25 to 33 months. These observations are in agreement with the results of cluster analysis, which were discussed above.

The temporal relationships between sera and isolates are summarized in Table 4. If we consider the midpoints of the clusters of isolates (0, 14.5, and 29 months), we find that they are separated by about 15-month intervals. If we look at the midpoints of the groups of sera (8, 22, and 37 months), we also find that they are separated by intervals of about 15 months. Furthermore, the differences in time between the midpoints of the clusters of isolates and the groups of sera that are positively specific for them amount to about 8 months. Our analysis of the data from this sequential neutralization experiment on a single individual suggests that major changes in the antigenic determinants of the isolates occurred at about 15-month intervals. It also suggests that effective neutralization may have taken up to 8 months to develop after appearance of a new antigenic determinant. To obtain a rough check on the last statement, we looked at the difference between the time at which each isolate was taken and the time that neutralization titers of at least 80 first appeared in the serum. The average difference amounts to 8.2 months (Table 5), which confirms the finding from the combined spectral map and cluster analysis.

**Sequence analysis.** In our previous study, we have shown that the nucleic acid sequence of the V3 region was remarkably stable over time for sequential isolates (45). These results suggest that regions other than the V3 loop may play a role in the emergence of neutralization-resistant viruses. Therefore, the entire gp160 regions of the sequential SIVcpz-ant isolates were sequenced. It is clear that the gp160 regions of all sequential isolates differ from each other. In addition, we applied cluster analysis by the UPGMA method to the 10 aligned sequences (data not shown). The clusters we identified from the sequence data did not agree, however, with the clusters that resulted from the neutralization data. We found three sequence clusters (at 74% of the maximal agglomeration distance) composed of isolates at months 0 to 15, months 18 to 29, and months 33 to 38. This indicates that neutralization, if it is directed to antigenic determinants of gp160, is directed to specific sites of this protein. The problem, then, was to localize regions within the protein sequence where specific changes occurred according to the pattern determined by the three neutralization clusters. In other words, we wanted to identify regions where an amino acid motif is present in sequences at months 0 and 1 and absent in all the others, and where another motif is present in sequences at months 7 to 22 and absent in all others, and where, finally, yet another motif is present in sequences at months 25 to 38 and absent in all others. For this purpose, a statistical method of sequence identification and elucidation, the polygram method, was applied to attempt to explain why certain sequential isolates escape neutralization. The three neutralization clusters of gp160 sequences were group 1, comprising two sequences (at 0 and 1 month); group 2, composed of four sequences (between 7 and 22 months); and group 3, containing four sequences (between 25 and 38 months). Each of the 10 individual sequences was tested for quadruplets that are specific for each of the three groups of sequences. The result of the analysis is represented in Fig. 4. In Fig. 4a, each of the 10 sequences is tested for specific quadruplets in group 1 (0 and 1 month). In Fig. 4b, the same 10 sequences are tested for specific quadruplets in group 2 (7 to 22 months), and Fig. 4c shows the same again in group 3 (25 to 38 months). In Fig. 4a, we observe that sequences at 0 and 1 month possess highly specific quadruplets for the group 1 isolates in the V1, V2, V4, V5, CD4, and gp40 regions. Figure 4b shows that sequences at months 7 to 22 possess highly specific quadruplets for group 2 isolates in the V4 region. Finally, Fig. 4c reveals that sequences at months 25 to 38 have highly specific quadruplets for group 3 isolates in the V1, V4, and gp40 regions. An extensive overlap of the specific changes in the C3 region is observed. These results suggest that the V1, V2, C3, V4, V5, CD4 and gp40 regions of gp160 may be involved as specific antigenic determinants in the neutralization process. Figure 5 also shows the amino acid sequences of the variable regions in the sequential SIVcpz-ant isolates. However, the entire gp160 of the sequential isolates will be described (unpublished data). The evidence in favor of this hypothesis is that the patterns of amino acid mutations in these regions coincide with the clusters of isolates determined from the neutralization assays. It is perhaps noteworthy that V4 is the only region where specific changes have occurred at the same location of the sequences in each of the three clusters of isolates.

The V3 loop sequence of SIVcpz-ant in PBMC of the sequential isolates was remarkably stable (45). To verify whether the viruses present in the PBMC and plasma were similar, we sequenced the V3 loop region of the virus present in plasma as well. We observed that the V3 loop of the virus present in plasma remained unchanged and similar to that in the PBMC (data not shown).

## DISCUSSION

The close relationship of SIVcpz-ant to HIV-1 may help to improve our understanding of the interaction between the host and the virus in HIV-1-infected individuals. The clinical relevance of neutralizing antibodies in HIV persistence remains uncertain, and their effect in controlling viral spread is not clear.

TABLE 5. Difference between the time when an isolate was taken and the subsequent time when a neutralization titer of at least 80 first appeared in the serum

| Time of isolate collection (mo) | Time to measurement of sera with titers ≥80 (mo) | Difference (mo)[a] |
|---|---|---|
| 0 | 15 | 15 |
| 7 | 15 | 8 |
| 15 | 18 | 3 |
| 18 | 22 | 4 |
| 22 | 33 | 11 |
| 25–33 | | |

[a] The mean value was 8.2 months.

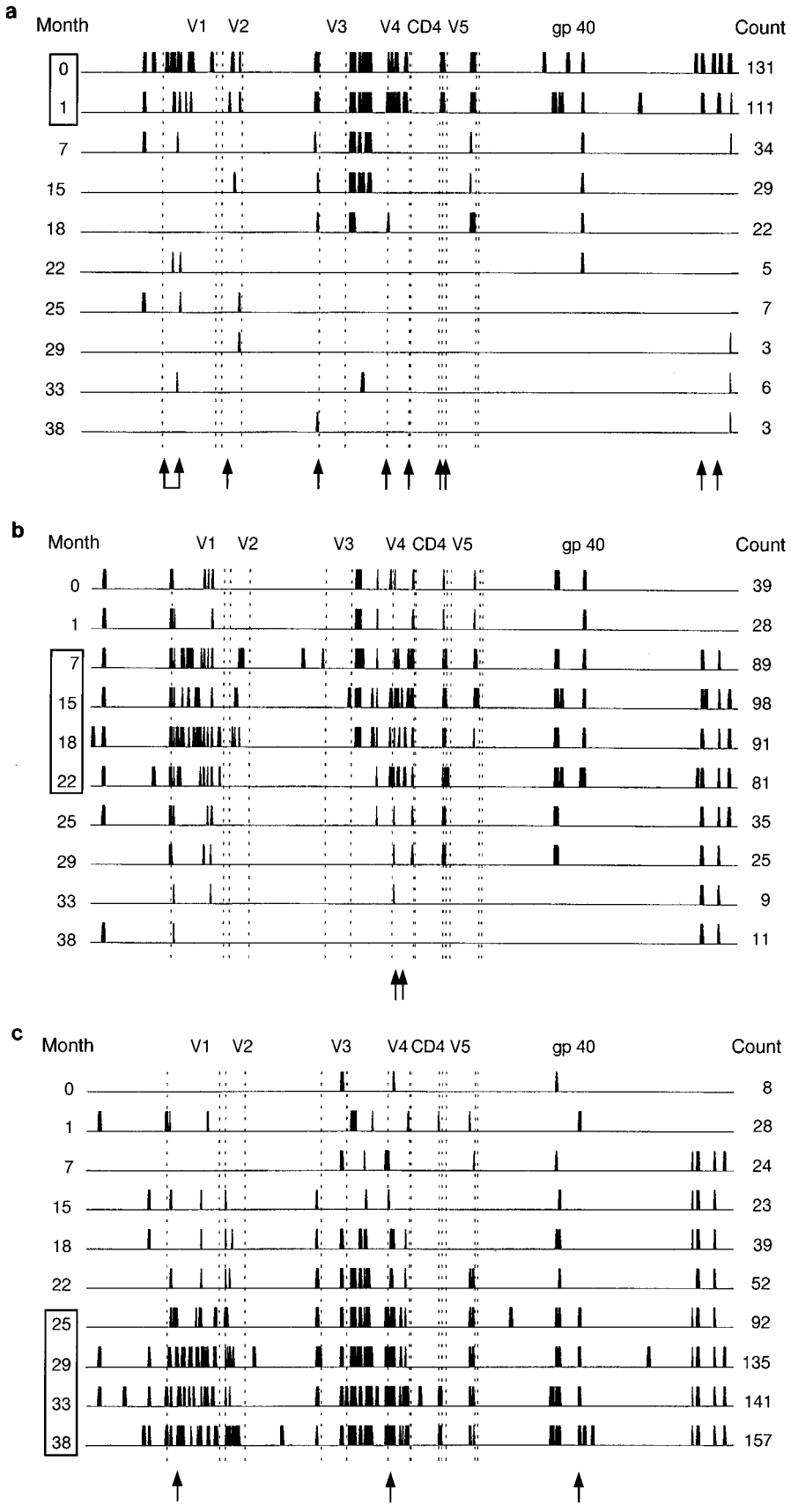
a
Month
V1 V2 V3 V4 CD4 V5 gp 40
Count
0 131
1 111
7 34
15 29
18 22
22 5
25 7
29 3
33 6
38 3
b
Month
V1 V2 V3 V4 CD4 V5 gp 40
Count
0 39
1 28
7 89
15 98
18 91
22 81
25 35
29 25
33 9
38 11
c
Month
V1 V2 V3 V4 CD4 V5 gp 40
Count
0 8
1 28
7 24
15 23
18 39
22 52
25 92
29 135
33 141
38 157

FIG. 4. Polygrams for the occurrence of specific amino acid quadruplets of 10 aligned gp160 sequences for each of three groups (a to c) of SIVcpz-ant isolates. The groups of isolates were determined from the spectral map and cluster analysis of the sequential neutralization data (Fig. 1). Each horizontal tracing represents 1 of 10 consecutive isolates which were collected between 0 and 38 months. Vertical marks indicate the occurrence of a quadruplet, at the corresponding position in the sequence, which is specific for a particular group of isolates. The numbers in the right margin provide the count of specific quadruplets for the particular group of isolates that have occurred in each of the 10 sequences. The points at which highly specific quadruplets occur for each of the three clusters of isolates are indicated on the panels by means of arrows at the bottom of the polygrams. The threshold of significance for quadruplets was fixed at 1.5.

In this study, spectral map analysis of the neutralization data depicted in Table 1 allowed us to group both isolates and sera into a number of neutralization spectrum clusters. Each cluster groups those isolates or sera with sufficient similarity in the shape of their neutralization spectra (Fig. 1 and 2). Cluster analysis by UPGMA indicated that neutralization escape mutants emerged on average at about 15-month intervals. The differences between the midpoints of the isolate neutralization spectrum clusters and the serum neutralization spectrum clusters which are positively specific for each other amounts to about 8 months, suggesting that on average, 8 months is required to mount an adequate neutralization response against a neutralization escape mutant.

Our observations of the naturally infected chimpanzee seem to correlate with the predictions as formulated by Bremermann (6) about the mechanism of HIV persistence in humans, in that escape mutants can emerge and will gradually take over, as long as the avidity of neutralizing antibodies directed against a specific mutant is lower than that of antibodies against competing strains. Bremermann predicted that the takeover by an escape mutant would require months, which fits with our observations in the naturally infected chimpanzee.

HIV persistence by escape of mutants which emerge faster than the immune system can respond may eventually generate high-quality antibody responses (37). The chimpanzee we have studied is estimated to be approximately 9 years old at the time of writing. During the last 5 years of follow-up, the animal has remained clinically asymptomatic. Virus could frequently be isolated from PBMC and from plasma, while the $CD4^+$ cell counts remained relatively stable (45).

From this study, we hypothesize that the role of neutralizing antibodies in the clinical evolution of the infection in this naturally infected chimpanzee is determined by at least three factors. These include the speed at which SIVcpz-ant diversity takes place, by the speed at which neutralization escape mutants are generated, and the speed at which an appropriate neutralization response is induced against the escape mutants. However, in HIV-1-infected humans, at least four serial different virus neutralization patterns can be hypothesized, including (i) patients in whom the generation of neutralization

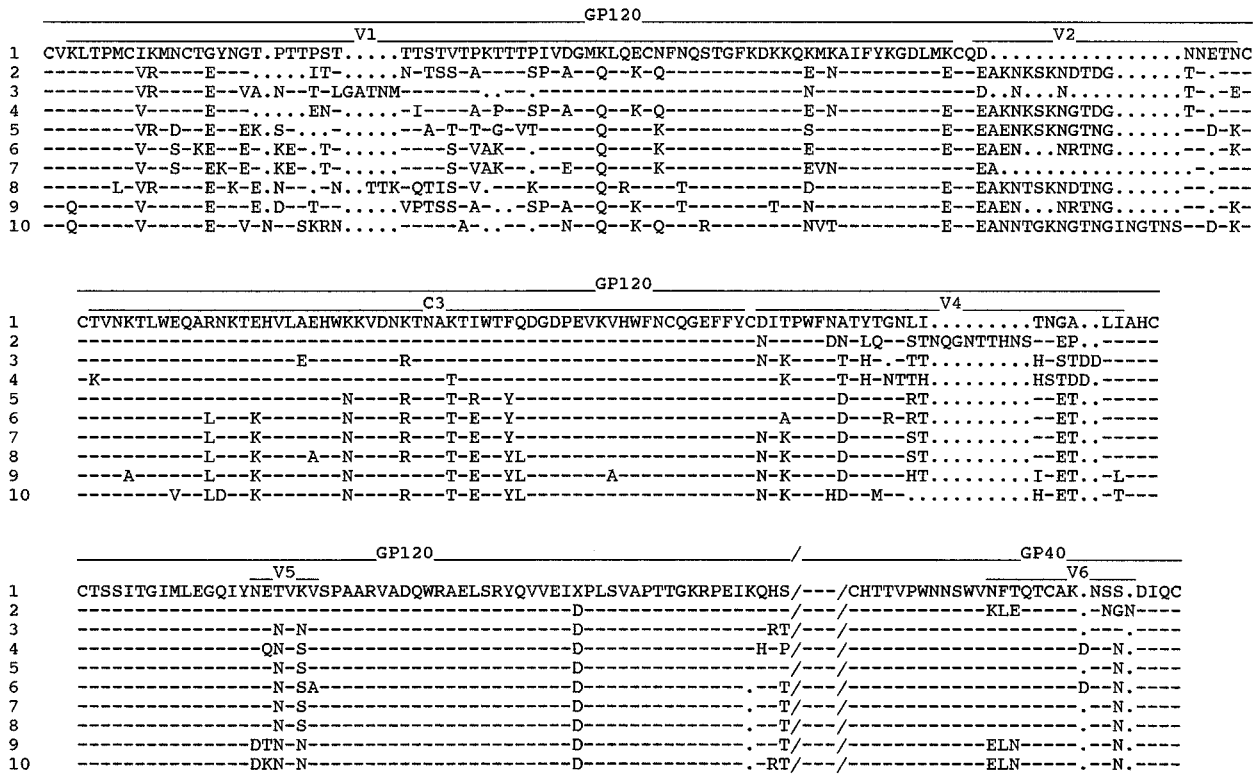


FIG. 5. Amino acid sequences of the envelope variable regions of SIVcpz-ant consecutive isolates. Sequence numbers 1 to 10 denotes isolates taken at months 0, 1, 7, 15, 18, 22, 25, 29, 33, and 38, respectively. Variation in V1, V2, and V4 is characterized by mostly nonsynonymous nucleotide substitutions, together with insertions and deletions, both involving the creation and deletion of potential N-glycosylation sites. As reported previously, the V3 loop is highly conserved (data not shown). Single mutations are observed in the C3 region downstream of the V3 loop. Variation in V5 involves only single mutations. An additional variable region is observed in gp120 at the gp120/gp40 cleavage site. In the cytoplasmic tail of gp41, one variable region involving single mutations and single insertions or deletions is observed. Amino acid identity between sequences is represented by dashes; points are introduced to align the sequences.

escape mutants is slow and the time required to mount an adequate neutralization response is slow; (ii) patients in whom the generation of neutralization escape mutants is fast and the time required to mount an adequate neutralization response is long; (iii) patients in whom the generation of neutralization escape mutants is slow and the time required to mount an adequate neutralization response is short; and (iv) patients in whom the generation of neutralization escape mutants is fast and the time required to mount an adequate neutralization response is short. The results of our analysis of the neutralization escape pattern in the naturally infected chimpanzee seem to correlate with the last two hypotheses. These suggest that this lentivirus in this host may be behaving as an HIV-1-infected human long-term nonprogressor. The 15-month period taken for evolution of a neutralization escape mutant and the 8 month taken to mount an appropriate neutralizing response could be compatible with the clinical status of a long-term nonprogressor.

Evidence exists that the immune response to HIV-1 results in the continuous selection of new viral variants (3, 10, 19, 32, 49, 55). A dynamic immune response should result in positive selective pressure on viral epitopes, resulting in a succession of epitopic variants. We believe that the data presented in this study support this hypothesis. Thus, if this same approach were applied to HIV-infected individuals, differences between rapid progressors and long-term nonprogressors based on the time required for neutralization escape mutant formation and for mounting an appropriate neutralizing antibody response might be quantified. In this context, studies are under way to compare the sequential neutralization pattern of the naturally lentivirus-infected chimpanzee with the patterns of HIV-1-infected individuals belonging to the rapid, intermittent, and long-term nonprogressor categories. The same study approach may also allow us to verify if a switch in sequential virus neutralization can help to predict disease progression in long-term asymptomatic patients. However, it must be emphasized that the persistence of HIV is a complex multifactorial interaction of the virus and the host humoral and cellular immune response (6). Nowak et al. (38) have developed a model based on the observed antigenic variation within CTL which may be responsible for viral persistence. However, the contribution of antigenic variation in neutralization and CTL epitopes to HIV viral persistence is still unclear.

The second objective of this report was to better understand the SIVcpz-ant neutralization escape with changes in the gp160 of SIVcpz-ant strains isolated at different moments of the infection. The lack of correlation between the phylogenetic clusters calculated by UPGMA and the isolate neutralization spectra clusters indicated that neutralization is directed to specific antigenic determinants in gp160 and not to gp160 as a whole. Interestingly, polygram analysis demonstrated that particular amino acid quadruplets in the hypervariable regions V1, V2, C3, V4, V5, and CD4, as well as gp40 regions, are involved as specific determinants in the neutralization process. This is supported by the pattern of amino acid mutations in these regions coinciding with the isolate neutralization clusters defined by the spectral map analysis.

Several reports have shown the V3 loop of HIV-1 to be the principal neutralization domain (15, 30, 33, 41, 47), and this region has been shown to be involved in neutralization escape in sequential isolates obtained from persons naturally infected with HIV-1 (33). It was striking that the V3 loop sequences of the sequential SIVcpz-ant in PBMC and in plasma were stable. However, this observation does not exclude the role of the V3 loop in neutralization, because it is possible that amino acid changes in other regions will change the microconformation of the V3 loop for the virus to escape neutralization.

It is remarkable that specific amino acid quadruplets in the V4 region correlate with the changes of the neutralization patterns of the isolates (Fig. 4). Our findings demonstrate the importance of highly specific amino acid quadruplets in the V4 region, which might have affected neutralization escape. These amino acid quadruplets in V4 perhaps contribute to a significant change in the overall conformation of the envelope protein of SIVcpz-ant, as previously reported in HIV-infected individuals (23, 46, 56). In these studies, it was demonstrated that V4 and V5 deletion mutations in gp120 affected the capacity to bind to CD4, as well as deletions and insertions in V4 which may have contributed to destabilization of the gp120 molecule. Perhaps the specific amino acid quadruplet changes in the V4 region contribute significantly to the generation of neutralization escape mutants. If so, they may be considered one of the two most important factors which rule the spectral map analysis of the neutralization data in Table 1.

The suggestion that broadening of the immune response to a particular region results in an increased rate of sequence evolution for that region implies that regions of HIV-1 envelope may evolve at differential rates. There is experimental evidence for differential rates of evolution between hypervariable regions from longitudinal studies of envelope sequence variation (21, 22, 43). In our study, based on the polygram analysis, the largest number of amino acid quadruplets specific for a particular isolate neutralization cluster was situated in the V4 region (seven quadruplets) followed by V1 (three), gp40 (three), V2 (one), C2 (one), and the region between CD4 (one) and V5 (one). These data support the notion that differential rates of evolution within the SIVcpz-ant envelope gene could reflect differential rates of selective forces on envelope epitopes. Furthermore, we believe that amino acid quadruplet changes in V4 contribute to perhaps the most important factor which rules the spectral map analysis data. In addition, the V1 and gp40 of SIVcpz-ant may contribute to the second factor which rules the spectral map analysis of the serial neutralization data, thus indicating the second most important factor in inducing neutralization escape SIVcpz-ant mutants. These observations corroborate the observations of others in HIV-infected individuals whereby the external portion of gp41 in HIV-1 is involved in antibody neutralization (56). Several studies have demonstrated that the hypervariable domains in V1 or V2 and in V4 or V5 of gp120 in HIV-1 could induce neutralizing antibodies (14, 16, 31, 36, 46, 56). These observations suggest that this lentivirus infection with its patterns of neutralization escape is analogous to the HIV-1 neutralization escape phenomenon observed in many infected humans.

In conclusion, the spectral map analysis of sequential neutralization data in a chimpanzee naturally infected with SIVcpz-ant demonstrated that an average of about 15 months is required to generate neutralization escape mutants and an average of about 8 months is required to mount an adequate neutralizing antibody response. Furthermore, the polygram analysis of specific amino acid quadruplets in the gp160 of serial sequences of SIVcpz-ant revealed that the changes which occurred in the V1, V2, C3, V4, V5, CD4, and gp40 regions correspond to the pattern of isolate neutralization clusters. Multivariate analysis of serial neutralization data of HIV-infected patients with different clinical progression are under way to monitor the dynamics of neutralization escape in infected humans and to compare these patterns with the observations we have made with this naturally lentivirus-infected chimpanzee.

## ACKNOWLEDGMENTS

We thank Betty Willems, Katleen Vereeken, and Patrick Nys for technical assistance; Ciska Maeckelbergh for typing; and David Davis for critical comments on the manuscript. We acknowledge Jul Van Hoof and Frits Daeyaert for assistance with the spectral map and polygram analysis, respectively.

This work was supported by Nationaal Fonds Wetenschappelijk Onderzoek (National Funds for Scientific Research) grants 3.3025.91, 3.3022.91, G.0329.96, and G.0330.96; Algemeen Bestuur voor Ontwikkelingssamenwerking (doctoral training fellowship to P.N.N.); and European Commision Directorate General XII grant ERB 3514 PL950995.

## REFERENCES

1. **Allen, D. M.** 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics **16:**125–127.
2. **Andries, K., B. Dewindt, J. Snoeks, L. Wouters, H. Moereels, P. J. Lewi, and P. A. J. Janssen.** 1990. Two groups of rhinoviruses revealed by a panel of antiviral compounds present sequence divergence and differential pathogenicity. J. Virol. **64:**1117–1123.
3. **Arendrup, M., A. Sonnerborg, B. Svennerholm, L. Akerblom, C. Nielsen, H. Clansen, S. Olofsson, J. O. Nielsen, and J. E. Hansen.** 1993. Neutralizing antibody response during human immunodeficiency virus type 1 infection: type and group specificity and viral escape. J. Gen. Virol. **74:**855–863.
4. **Asjö, B., L. Mörfeldt-Manson, J. Albert, G. Biberfeld, A. Karlsson, K. Lidman, and E. M. Fenyö.** 1986. Replicative capacity of human immunodeficiency virus from patients with varying severity of HIV infection. Lancet **ii:**660–663.
5. **Boom, R., C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. **28:**496–503.
6. **Bremermann, H. J.** 1995. Mechanism of HIV presistence: implications for vaccines and therapy. J. Acquired Immune Defic. Syndr. **9:**459–483.
7. **Cheng-Meyer, C., D. Seto, M. Tateno, and J. A. Levy.** 1988. Biologic features of HIV-1 that correlate with virulence in the host. Science **240:**80–82.
8. **Daeyaert, F., H. Moereels, and P. J. Lewi.** Classification and identification of proteins by means of common and specific amino acid n-tuples in unaligned sequences. Submitted for publication.
9. **De Beenhouwer, H., H. Verhaert, H. Claeys, and C. Vermylen.** 1992. Confirmation of hepatitis C virus positive blood donors by immunoblotting and polymerase chain reaction. Vox Sang. **63:**198–203.
10. **Delassus, S., R. Cheynier, and S. Wain-Hobson.** 1991. Evolution of human immunodeficiency virus type 1 nef and long terminal repeat sequences over 4 years in vivo and in vitro. J. Virol. **65:**225–231.
11. **Earl, P. L., C. C. Broder, D. Long, S. A. Lee, J. Peterson, S. Chakrabarti, R. W. Doms, and B. Moss.** 1994. Native oligomeric human immunodeficiency virus type 1 envelope glycoprotein elicits diverse monoclonal antibody reactivities. J. Virol. **68:**3015–3026.
12. **Fenyö, E. M., J. Albert, and B. Asjö.** 1989. Replicative capacity, cytopathic effect and cell tropism of HIV. AIDS **3**(Suppl. 1)**:**S5–S12.
13. **Fenyö, E. M., L. Mörfeldt-Manson, F. Chiodi, B. Lind, A. von Gegerfelt, J. Albert, E. Olausson, and B. Asjo.** 1988. Distinct replicative and cytopathic characteristics of human immunodeficiency virus isolates. J. Virol. **62:**4414–4419.
14. **Gorny, M. K., J. P. Moore, A. J. Conley, S. Karwowska, J. Sodroski, C. Williams, S. Burda, L. J. Boots, and S. Zolla-Pazner.** 1994. Human anti-V2 monoclonal antibody that neutralizes primary but not laboratory isolates of human immunodeficiency virus type 1. J. Virol. **68:**8312–8320.
15. **Goudsmit, J., K. Ljunggren, L. Smit, M. Jondal, and E. M. Fenyö.** 1988. Biological significance of the antibody response to HIV antigens expressed on the cell surface. Arch. Virol. **103:**189–206.
16. **Haigwood, N., J. Shuster, G. Moore, H. Lee, P. V. Skiles, K. W. Higgins, P. J. Barr, C. George-Nascimento, and K. S. Steimer.** 1990. Importance of hypervariable regions of HIV-1 gp120 in the generation of virus neutralizing antibodies. AIDS Res. Hum. Retroviruses **6:**855–869.
17. **Ho, D. D., G. M. Sarngadharan, M. S. Hirsch, R. T. Schooley, T. R. Rota, R. C. Kennedy, T. C. Chanh, and V. L. Sato.** 1987. Human immunodeficiency virus neutralizing antibodies recognize several conserved domains on the envelope glycoproteins. J. Virol. **61:**2024–2028.
18. **Ho, D. D., J. A. McKeating, X. L. Li, T. Moudgil, E. S. Daar, N. C. Sun, and J. Robinson.** 1991. Conformational epitope on gp120 important in CD4 binding and human immunodeficiency virus type 1 neutralization identified by a human monoclonal antibody. J. Virol. **65:**489–493.
19. **Holmes, E. C., L. Q. Zhang, P. Simmonds, C. A. Ludlam, and A. J. Leigh Brown.** 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. Proc. Natl. Acad. Sci. USA **89:**4835–4839.
20. **Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson.** 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature **345:**356–359.
21. **Lamers, S. L., J. W. Sleasman, J. X. She, K. A. Barrie, S. M. Pomeroy, D. J. Barrett, and M. M. Goodenow.** 1994. Persistence of multiple maternal genotypes of human immunodeficiency virus type 1 in infants infected by vertical transmission. J. Clin. Invest. **93:**380–390.
22. **Lamers, S. L., J. W. Sleasman, J. X. She, K. A. Barrie, S. M. Pomeroy, D. J. Barrett, and M. M. Goodenow.** 1993. Independent variation and positive selection in env V1 and V2 domains within maternal-infant strains of human immunodeficiency virus type 1 in vivo. J. Virol. **67:**3951–3960.
23. **Lee, C.-N., J. Robinson, G. Mazzara, Y. L. Cheng, M. Essex, and T. H. Lee.** 1995. Contribution of hypervariable domains to the conformation of a broadly neutralizing glycoprotein 120 epitope. AIDS Res. Human Retroviruses **11:**777–781.
24. **Lewi, P. J.** 1995. Spectral mapping of drug-test specificities, p. 219–253. *In* H. van de Waterbeemd (ed.), Chemometrics in molecular design. VCH, Weinheim, Germany.
25. **Lewi, P. J.** 1982. Multivariate data analysis in industrial practice. Research Studies Press and John Wiley & Sons, Ltd., Chichester, England.
26. **Lewi, P. J.** 1989. Spectral map analysis. Analysis of contrasts especially from log ratios. Chemometrics Intelligent Lab. Syst. **5:**105–116.
27. **Lewi, P. J.** 1996. Spectral mapping, a technique for classifying biological activity profiles of chemical compounds. Arzneim.-Forsch. **26:**1295–1300.
28. **Lewi, P. J.** 1995. Spectral mapping of drug-test specificities. Doctoral thesis. Vrije Universiteit Brussel, Brussels, Belgium.
29. **Lewi, P. J., H. Moereels, and D. Adriaensen.** 1992. The combination of dendrograms with plots of latent variables. An application to G-proteins coupled receptor sequences. Chemometrics Intelligent Lab. Syst. **16:**145–154.
30. **Matsushita, S., M. Robert-Guroff, J. Rusche, A. Koito, T. Hattori, A. Hoshino, K. Javaherian, K. Takatsuki, and S. Putneys.** 1988. Characterization of a human immunodeficiency virus neutralizing monoclonal antibody and mapping of the neutralizing epitope. J. Virol. **62:**2107–2114.
31. **McKeating, J. A., C. Shotton, J. Cordell, S. Graham, P. Balfe, N. Sullivan, M. Charles, M. Page, A. Bolmstedt, S. Olofsson, C. S. Kayman, Z. Wu, A. Pinter, C. Dean, J. Sodroski, and R. A. Weiss.** 1993. Characterization of neutralizing monoclonal antibodies to linear and conformation-dependent epitopes within the first and second variable domains of human immunodeficiency virus type 1 gp120. J. Virol. **67:**4932–4944.
32. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfelt-Manson, B. Asjo, and S. Wain-Hobson.** 1989. Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. Cell **58:**901–910.
33. **Michael, L. N., E. K. Davis, D. L. Loomis-Price, C. T. Van Cott, S. D. Burke, R. R. Redfield, and L. D. Brix.** 1996. V3 seroreactivity and sequence variation: tracking the emergence of V3 genotypic variation in HIV-1 infected patients. AIDS **10:**121–129.
34. **Moore, J. P., and D. D. Ho.** 1992. Conformationally sensitive epitopes on the gp120 glycoprotein of human immunodeficiency virus type 1 are highly prevalent in sera of infected humans. J. Virol. **67:**863–875.
35. **Moore, J. P., H. Yoshiyama, D. D. Ho, J. E. Robinson, and J. Sodroski.** 1993. Antigenic variation in gp120s from molecular clones of HIV-1 LAI. AIDS Res. Hum. Retroviruses **9:**1185–1193.
36. **Moore, J. P., Q. J. Sattentau, R. Wyatt, and J. Sodroski.** 1994. Probing the structure of the human immunodeficiency virus surface glycoprotein gp120 with a panel of monoclonal antibodies. J. Virol. **68:**469–484.
37. **Nowak, M. A., and R. M. May.** 1992. Coexistence and competition in HIV infections. J. Theor. Biol. **159:**329–342.
38. **Nowak, M. A., R. M. May, R. E. Phillips, S. Rowland-Jones, D. G. Lalloo, S. McAdam, P. Klenerman, B. Köppe, K. Sigmund, C. R. M. Bangham, and E. J. McMichael.** 1995. Antigenic oscillations and shifting immunodominance in HIV-1 infections. Nature **375:**606–611.
39. **Nyambi, P. N., J. Nkengasong, M. Peeters, F. Simon, J. Eberle, W. Janssens, K. Fransen, B. Willems, K. Vereecken, L. Heyndrickx, P. Piot, and G. van der Groen.** 1995. Reduced capacity of antibodies from patients infected with human immunodeficiency virus type 1 (HIV-1) group O to neutralize primary isolates of HIV-1 group M viruses. J. Infect. Dis. **172:**1228–1237.
40. **Nyambi, P. N., J. Nkengasong, P. Lewi, K. Andries, W. Janssens, K. Fransen, L. Heyndrickx, P. Piot, and G. van der Groen.** 1996. Multivariate analysis of human immunodeficiency virus type 1 neutralization data. J. Virol. **70:**6235–6243.
41. **Palker, T. J., M. E. Clark, A. Langlois, T. J. Matthews, K. J. Weinhold, R. R. Randall, D. P. Bologenesi, and B. F. Haynes.** 1988. Type-specific neutralization of the human immunodeficiency virus with antibodies to env-encoded synthetic peptides. Proc. Natl. Acad. Sci. USA **85:**1932–1936.
42. **Pang, S., Y. Shlesinger, E. S. Daar, T. Moudgil, D. D. Ho, and I. S. Y. Chen.** 1992. Rapid generation of sequence variation during primary HIV-1 infection. AIDS **6:**453–460.
43. **Pedroza Martins, L., N. Chenciner, and S. Wain-Hobson.** 1992. Complex intrapatient sequence variation in the V1 and V2 hypervariable regions of the HIV-1 gp120 envelope sequence. Virology **191:**837–845.
44. **Peeters, M., K. Fransen, E. Delaporte, M. Vanden Haesevelde, G. M. Gershy-Damet, L. Kestens, G. van der Groen, and P. Piot.** 1992. Isolation and characterization of a new chimpanzee lentivirus (simian immunodeficiency

virus isolate cpz-ant) from a wild-captured chimpanzee. AIDS **6:**447–451.
45. **Peeters, M., W. Janssens, M. Vanden Haesevelde, K. Fransen, B. Willems, L. Heyndrickx, L. Kestens, P. Piot, G. van der Groen, and J. Heeney.** 1995. Virologic and serologic characteristics of a natural chimpanzee lentivirus infection. Virology **211:**312–315.
46. **Pollard, S. R., M. D. Rosa, and D. C. Wiley.** 1992. Truncated variants of gp120 bind CD4 with high affinity and suggest a minimum CD4 binding region. EMBO J. **11:**585–591.
47. **Rusche, J. A., K. Javaherian, C. McDanal, J. Petro, D. L. Lynn, R. Grimaila, A. Langlois, R. C. Gallo, L. O. Arthur, P. J. Fischinger, et al.** 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. Proc. Natl. Acad. Sci. USA **85:**3198–3202.
48. **Schwartz, S., B. K. Felber, E. M. Fenyö, and G. N. Pavlakis.** 1989. Rapidly and slowly replicating human immunodeficiency virus type 1 isolates can be distinguished according to target-cell tropism in T-cell and monocyte cell lines. Proc. Natl. Acad. Sci. USA **86:**7200–7203.
49. **Simmonds, P., L. Q. Zhang, F. Mcomish, P. Balfe, C. A. Ludlam, and A. J. Leigh Brown.** 1991. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 *env* sequences in plasma viral lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. **65:**6266–6276.
50. **Sneath, P. H., and R. R. Sokal.** 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco, Calif.
51. **Stamatatos, L., and C. Cheng-Meyer.** 1995. Structural modulations of the envelope gp120 glycoprotein of human immunodeficiency virus type 1 upon oligomerization and differential V3 loop epitope exposure of isolates displaying distinct tropism upon virion-soluble receptor binding. J. Virol. **69:** 6191–6198.
52. **Tersmette, M., R. A. Gruters, F. de Wolf, R. E. Goede, J. M. Lang, P. T. Schellekens, J. Goudsmit, H. G. Huisman, and F. Miedema.** 1989. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. J. Virol. **63:**2118–2125.
53. **Vanden Haesevelde, M., M. Peeters, G. Jannes, W. Janssens, G. van der Groen, P. M. Sharp, and E. Saman.** 1996. Sequence analysis of a HIV-1 related lentivirus isolated from a wild captured chimpanzee. Virology **221:** 346–350.
54. **Van Kerckhoven, I., K. Fransen, M. Peeters, H. De Beenhouwer, P. Piot, and G. van der Groen.** 1994. Quantification of human immunodeficiency virus in plasma by RNA PCR, viral culture, and p24 antigen detection. J. Clin. Microbiol. **32:**1669–1673.
55. **Wolfs, T. F., J. J. de Jong, H. Van den Berg, J. M. Tijnagel, W. J. Krone, and J. Goudsmit.** 1990. Evolution of sequences encoding the principal neutralization epitope of human immunodeficiency virus 1 is host dependent, rapid, and continuous. Proc. Natl. Acad. Sci. USA **87:**9938–9942.
56. **Wyatt, R., N. Sullivan, M. Thali, H. Repke, D. Ho, J. Robenson, M. Posner, and J. Sodroski.** 1993. Functional and immunologic characterization of human immunodeficiency virus type 1 envelope glycoproteins containing deletions of the major variable regions. J. Virol. **67:**4557–4565.
57. **Zhang, L. Q., P. MacKenzie, A. Cleland, E. C. Holmes, A. J. Leigh Brown, and P. Simmonds.** 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J. Virol. **67:**3345–3356.
58. **Zhu, T., H. Mo, N. Wang, D. S. Nam, Y. Cao, R. A. Koup, and D. D. Ho.** 1993. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. Science **261:**1179–1181.