JOURNAL OF VIROLOGY, Nov. 1999, p. 9256–9265 Vol. 73, No. 11
0022-538X/99/$04.00+0

Genetic Stability of Foamy Viruses: Long-Term Study in an African Green Monkey Population

MATTHIAS SCHWEIZER,* HUBERTUS SCHLEER, MICHAEL PIETREK, JÜRGEN LIEGIBEL, VALERIA FALCONE, AND DIETER NEUMANN-HAEFELIN

Abteilung Virologie, Institut für Medizinische Mikrobiologie und Hygiene, University of Freiburg, Freiburg, Germany

Received 13 May 1999/Accepted 2 August 1999

The genetic variability of the envelope surface domain (SU) of simian foamy virus (FV) of African green monkeys was studied. To assess the interindividual diversity of FV, isolates were obtained from 19 animals living together in a monkey house. The monkeys had been imported from Kenya prior to being placed in long-term housing in the research institute. In addition, a simian FV isolate and proviral DNA were obtained from an animal caretaker infected in this setting. DNA of the complete SU (1779 to 1793 bp) was analyzed by PCR and sequencing. The sequences revealed four clusters with high homologies (>95%). Between the clusters, divergencies ranged from 3 to 25%. Obviously, the clusters reflect four different strains or subtypes of simian FV type 3 that were prevalent in the colony. In contrast to lentiviruses, hypervariable regions could not be detected in the FV SU. Furthermore, to analyze the intraindividual diversity of FV, we investigated the virus population within an individual monkey at a given time point and its evolution over 13 years. For this purpose, 22 proviral SU clones generated by PCR from one oral swab and seven isolates obtained from the same animal between 1982 and 1995 were examined. These sequences revealed exceptionally high homology rates (99.5 to 100%), and only a minimal genetic drift was recognized within the series of isolates. In conclusion, the low in vivo divergency of FV SU suggests that genetic variability is not important for the maintenance of FV persistence.

Genetic variability of various viruses is thought to be instrumental in viral persistence in vivo. In particular, the extraordinarily high variability of lentiviruses, like human and simian immunodeficiency viruses (HIV and SIV, respectively), results in the rapid development of a complex viral quasispecies. It has been proposed that genetic variability might counteract host immune reactions, for example by generating viral escape mutants (3) or by exceeding a threshold of antigenic diversity beyond which the immune system cannot cope (31). However, these hypotheses have not been confirmed by recent reports (20, 32), and the role of variability in viral persistence and in the development of disease is discussed controversially (45). Retroviruses of the human T-lymphotropic virus/bovine leukemia virus (HTLV/BLV) group (*Deltaretrovirus* genus) reveal very stable genomes in vivo (13), questioning the assumption that genetic variability is essential for retrovirus persistence. However, substantial data on intra- and interanimal genetic variation of retroviruses other than the primate lentiviruses or members of the HTLV/BLV group are missing (45).

In this study, we analyzed the genetic variability of foamy viruses (FV; *Spumavirus* genus of the *Retroviridae* family [18]), which are complex retroviruses utilizing a replication cycle that shares features of retro- and hepadnaviruses (25, 34, 49), both prone to genetic variability. FV cause persisting infections in various mammalian species and are prevalent at high rates in nonhuman primates (29). Early reports on FV prevalence in the general human population could not be confirmed (1, 40). However, human infections resulting from accidental transmissions of nonhuman primate FV are well documented (16, 29, 41). FV pathogenicity does not occur in naturally and experimentally infected animals nor in accidentally infected humans. This points to a highly host-adapted mode of persistence. Developmental deficits in FV-transgenic mice have been described (4), but the relevance of these findings for the natural course of infection is questionable.

The degree of the genomic variability of FV in vivo is not known at present. Only four complete sequences of FV isolates from different primate species are available and can be compared. Small portions (e.g., from the *pol* region [38]) that were obtained by PCR from FV isolates or from peripheral blood lymphocyte (PBL) samples of infected primates have been used to compute phylogenetic relationships of various primate FV. However, up to now the genomic diversity of viruses within one species has been analyzed exclusively by Southern blot hybridization of isolates from one monkey colony (39), and data on variability within an infected individual are completely missing. Therefore, we analyzed the genomic variability of the FV present in a stable colony of African green monkeys (AGM) and in a single animal of this colony. We decided to focus our work on the complete surface domain (SU) of the *env* gene for several reasons. (i) SU is the major target for antiviral immunity, and highest degrees of variability can be expected in this region. (ii) Most variability studies in other retroviruses have been done on SU sequences, facilitating comparison of data. (iii) Sequencing of complete domains minimizes the risk of overlooking hypervariable regions.

According to our results, the genomic variability of FV within one animal appears to be too low to play an essential role for viral persistence but does allow distinct FV strains within one monkey species to be defined.

* Corresponding author. Mailing address: Abteilung Virologie, Institut für Medizinische Mikrobiologie und Hygiene, Hermann-Herder-Str. 11, D-79104 Freiburg, Germany. Phone: 49 761 203-6619. Fax: 49 761 203-6634. E-mail: mschweiz@ukl.uni-freiburg.de.

MATERIALS AND METHODS

Animals. All AGM (*Chlorocebus* [*Cercopithecus*] *aethiops*) investigated in this study had been caught in the wild in Kenya and kept in single cages at the Department of Virology in Freiburg, Germany, ever since (Table 1).

TABLE 1. Features of FV present in the AGM colony investigated

FV sequence[a]	Year of FV isolation	Date of shipment of host monkey	Cluster[b]	Group[b] (SU length)
agm5	1988	4 Oct. 1980	A	1 (563 aa[c])
agm8	1992	4 Oct. 1980	A	1 (563 aa)
agm9	1990	4 Oct. 1980	A	1 (563 aa)
agm11	1988	4 Oct. 1980	A	1 (563 aa)
agm18	1982	4 Oct. 1980	A	1 (563 aa)
agm26	1988	19 July 1979	A	1 (563 aa)
agm37	1992	4 Oct. 1980	A	1 (563 aa)
agm1[d]	1982–1995[d]	4 Oct. 1980	B	1 (563 aa)
agm3	1988	4 Oct. 1980	B	1 (563 aa)
agm4	1982	4 Oct. 1980	B	1 (563 aa)
agm6	1991	19 July 1979	B	1 (563 aa)
agm10	1986	4 Oct. 1980	B	1 (563 aa)
agm12	1992	4 Oct. 1980	B	1 (563 aa)
agm16	1982	4 Oct. 1980	B	1 (563 aa)
agm25	1990	19 July 1979	B	1 (563 aa)
SFV-3/LK-3	1978	1976		1 (563 aa)
agm17	1994	4 Oct. 1980	C	2 (568 aa)
agm22	1994	8 March 1978	C	2 (568 aa)
agm24	1992	22 Feb. 1979	C	2 (568 aa)
agm20	1982	23 Jan. 1975	D	2 (568 aa)
SFVhum	—[e]		D	2 (568 aa)
SFVka	1996		D	2 (568 aa)

[a] For accession numbers of the sequences obtained in this study, see Materials and Methods.
[b] Assignment to clusters and groups is according to Fig. 3.
[c] aa, amino acid.
[d] For the seven isolates from agm1, see Table 4.
[e] —, sequence in PBL.

Virus isolation. Isolation of FV from the oral mucosa of AGM has been described previously (41). Briefly, human diploid fibroblasts (strain alpha-1) were inoculated with throat swab material freshly obtained from FV-seropositive monkeys and cultivated until typical FV cytopathic effect occurred. Infection was proven by indirect immunofluorescence by using FV-positive AGM sera (30) and FV-specific *pol* PCR (38).

PCR. Infected cell cultures of freshly obtained throat swab material were lysed by sarcosyl, and DNA was isolated by phenol-chloroform extraction as described previously (37). The whole SU was amplified by nested PCR with primer sequences from segments next to SU which have been found to be highly conserved in all published sequences of primate FV (human FV [HFV] [11], simian FV from chimpanzees [SFVcpz] [17], SFV type 1 [SFV-1] [22], and SFV-3/LK-3 [33]). The primers for the outer PCR were pM3, 5′-GGCCAATTAGTCCAGG AGAGGGT-3′ (nucleotides 6911 to 6933 in SFV-3/LK-3 [33]), and pMA4, 5′-CTTCCATCAAAGTGACAACATGATC-3′ (nucleotides 9017 to 8994), and the primers for the inner PCR were pM1, 5′-ATTTTGGACCATCTTGGCAA CA-3′ (nucleotides 7013 to 7034), and pM2 (nucleotides 8836 to 8814). Amplification conditions have been previously described (38).

Cloning and sequencing of the amplification products. Amplification products were cloned by using a TA cloning kit (Invitrogen, San Diego, Calif.) and sequenced by the dideoxy chain termination method by using USB DNA sequencing kit 2.0 (U.S. Biochemicals, Cleveland, Ohio), as recommended by the suppliers. The following oligonucleotides were used as primers for sequencing: SP6 and T7, located on the TA vector and flanking the cloned amplification product; pM1 and pM2, which had been used for PCR; and p1 to p9, located in the amplification product (Fig. 1) generated according to the ongoing sequencing results. Primer sequences and positions on the resulting consensus sequence were as follows: for p1, CGCGTGTTATGTGTTGGT (nucleotides 245 to 262); for p2, CCTGTTTCGTTACTATTGCTA (nucleotides 302 to 322); for p3, GTT ACTCATCAGGCCACATACC (nucleotides 376 to 397); for p4, GTCTAGCA TACCACAAGGTG (nucleotides 474 to 493); for p5, CCTCTAGGAGATCCT AGAGATC (nucleotides 685 to 706); for p6, CATGTGCTATTCTGTTCTGA TC (nucleotides 988 to 1009); for p7, TCCACAGTCTCCTTCCCA (nucleotides 1266 to 1249); for p8, CAAAGTGGTGATGGAAATGC (nucleotides 1511 to 1492); and for p9, AAGCCCTAGCTGTAGGGAT (nucleotides 1678 to 1660).

Sequence analysis. Only the sequences between the PCR primers, spanning the region of the *pol*/*env* overlap, the complete SU, and the beginning of the transmembrane region (TM) were analyzed by using different software packages of HUSAR (Heidelberg Unix Sequence Analysis Resources). CLUSTALW (43) was used for multiple sequence alignments. Similarity plots were generated by PlotSimilarity, which calculates the average similarity among all members of a group of aligned sequences at each position in the alignment. The symbol comparison table used was plotsimpep.cmp, which has a similarity score of 1.5 for perfect symbol matches. Pairwise homology rates were calculated by Homologies; the symbol comparison table used for the estimation of amino acid sequence similarity was comparpep.cmp; for amino acid identity, uniquepep.cmp; and for nucleic acid homology, compardna.cmp. Phylogenetic analysis was performed with PHYLIP version 3.5 (9). Evolutionary distances were calculated by DNAdist by using Kimura's two-parameter method (21); phylogenetic relationships and trees were computed from the distance matrix by KITSCH by using the Fitch-Margoliash method, version 3.572c (10), or by Neighbor by using the neighbor-joining method (36).

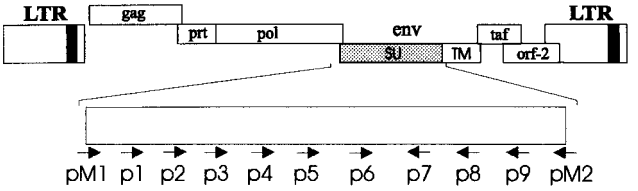

FIG. 1. Genomic organization of AGM FV. The region investigated is enlarged. Oligonucleotides used as PCR primers (pM1 and pM2) or sequencing primers (p1 to p9) are indicated by arrows. The exact positions are given in Materials and Methods. LTR, long terminal repeat.

Nucleotide sequence accession numbers. The nucleotide sequences of the 50 AGM FV SU sequences obtained in this study have been submitted to the EMBL Nucleotide Sequence Database under accession no. AJ244067 to AJ244116.

RESULTS

Characterization of the *env* SU of FV from different individuals. From 1982 to 1994, simian FV isolates were obtained from oral swabs of 19 animals living in a closed AGM colony. A recent simian FV isolate and proviral DNA from PBL of an animal caretaker were included in the study, since the infection had originated from an animal of this colony, probably transmitted by a bite in 1976 (38, 41). Furthermore, the AGM FV prototype SFV-3, which had been isolated before 1964 from an AGM kept in New York (42), was analyzed for comparison. The features of all FV investigated in this study are listed in Table 1. Proviral DNA of infected cells was amplified by a nested PCR with highly conserved primers flanking the SU of the FV *env* gene (Fig. 1). Amplification products were cloned and sequenced with the sequencing primers shown in Fig. 1. The nucleotide sequences were aligned by comparing them with the published sequence for SFV-3/LK-3 (an AGM FV isolated before 1978 from a lymph node of another animal of this colony [30, 33]) and other published sequences of primate FV prototypes (Table 2 and data not shown). The respective amino acid alignment is shown in Fig. 2.

All sequences contained an intact SU open reading frame. The lengths of the amplification products varied from 1779 to

TABLE 2. Features of FV prototypes

FV prototype	Host species	Reference	Accession no.	SU length[b]
SFV-3[a]	AGM	42		567 aa
SFV-3/LK-3	AGM	33	M74895	563 aa
SFV-1	Macaca cyclopsis	22	X54482	569 aa
SFVcpz	Chimpanzee	17	U04327	568 aa
HFV	Human	11	M38712	569 aa

[a] For accession no. of the sequences obtained in this study, see Materials and Methods.
[b] aa, amino acid.

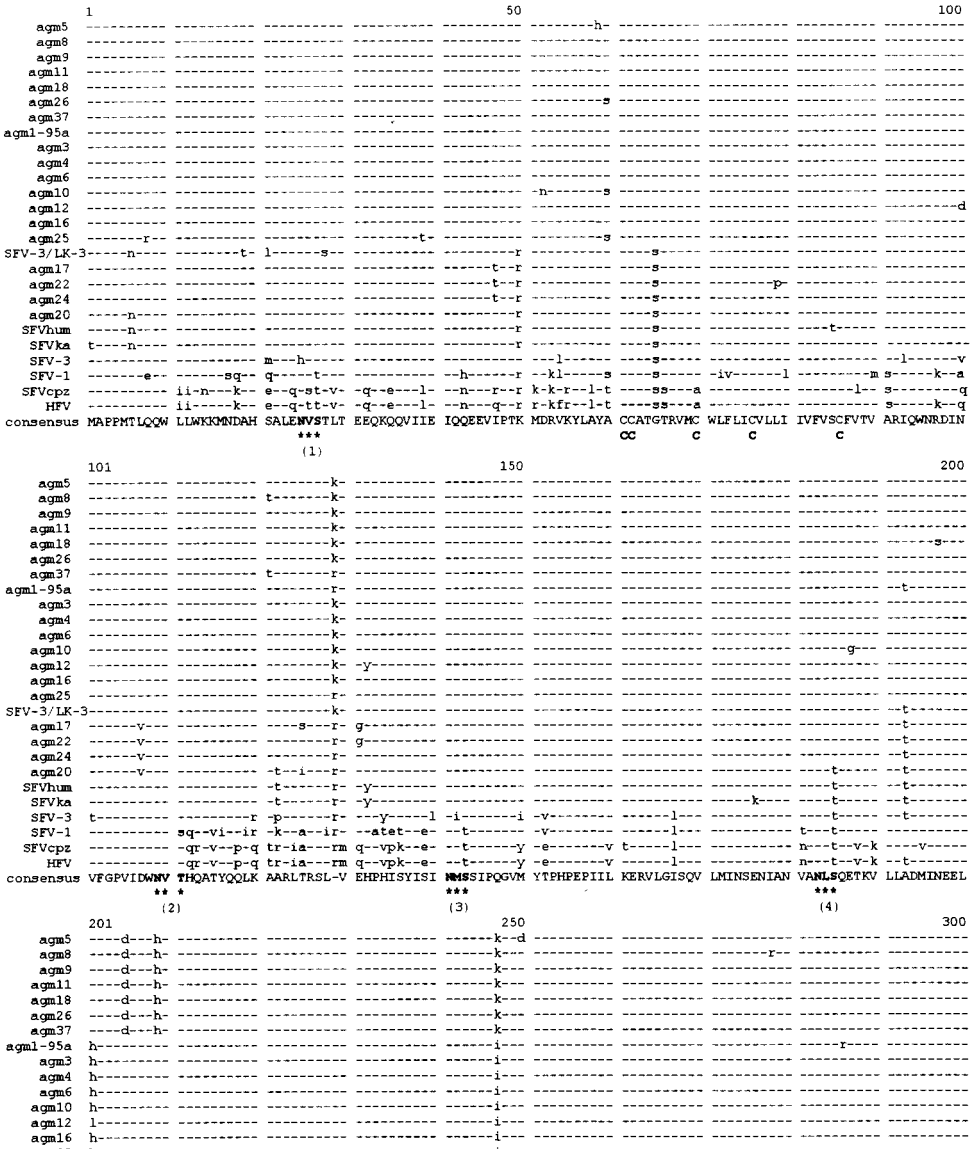

FIG. 2. Multiple amino acid sequence alignment of the SU regions of the FV listed in Table 1, generated by CLUSTAL. Highly conserved sequence motifs are indicated. All variant glycosylation sites are emphasized by bold letter. C, cysteines; ***, N-glycosylation sites (N-Xaa-S/T); cleavage, cleavage site between SU and TM.

1794 bp; the differences correspond to insertions or deletions of amino acids in the 3′ region of SU. The first ATG in the *env* open reading frame is located at positions 49 to 51 of the amplification products. Sequences coding for a motif similar to a subtilisin-like protease cleavage site typically placed between the SU and TM of retroviruses (R-X-K/R-R [46]) were found downstream of either nucleotide 1737 or 1752. Thus, the sizes of the SUs of the different FV range from 563 to 568 amino acids. Within SU, the predicted cleavage site C/F behind the putative signal peptide (46) at amino acid positions 86 and 87 as well as the positions of 11 glycosylation sites (N-X-S/T [44]) and 18 cysteines are highly conserved, pointing to an essential role of these motifs (Fig. 2). Remarkably, the positions of glycosylation sites 1 to 4, 7, 10, and 11 are completely conserved, whereas the positions of sites 5, 6, 8, and 9 differ by several amino acids in some of the sequences. In detail, site 5 of SFVcpz and HFV is shifted by two amino acid positions, versus all other sequences, and site 6 of SFV-1 is shifted by five amino acid positions. Sites 8 of the AGM isolates of group 2 and of SFV-1 are shifted by five amino acids. Sites 9 of AGM

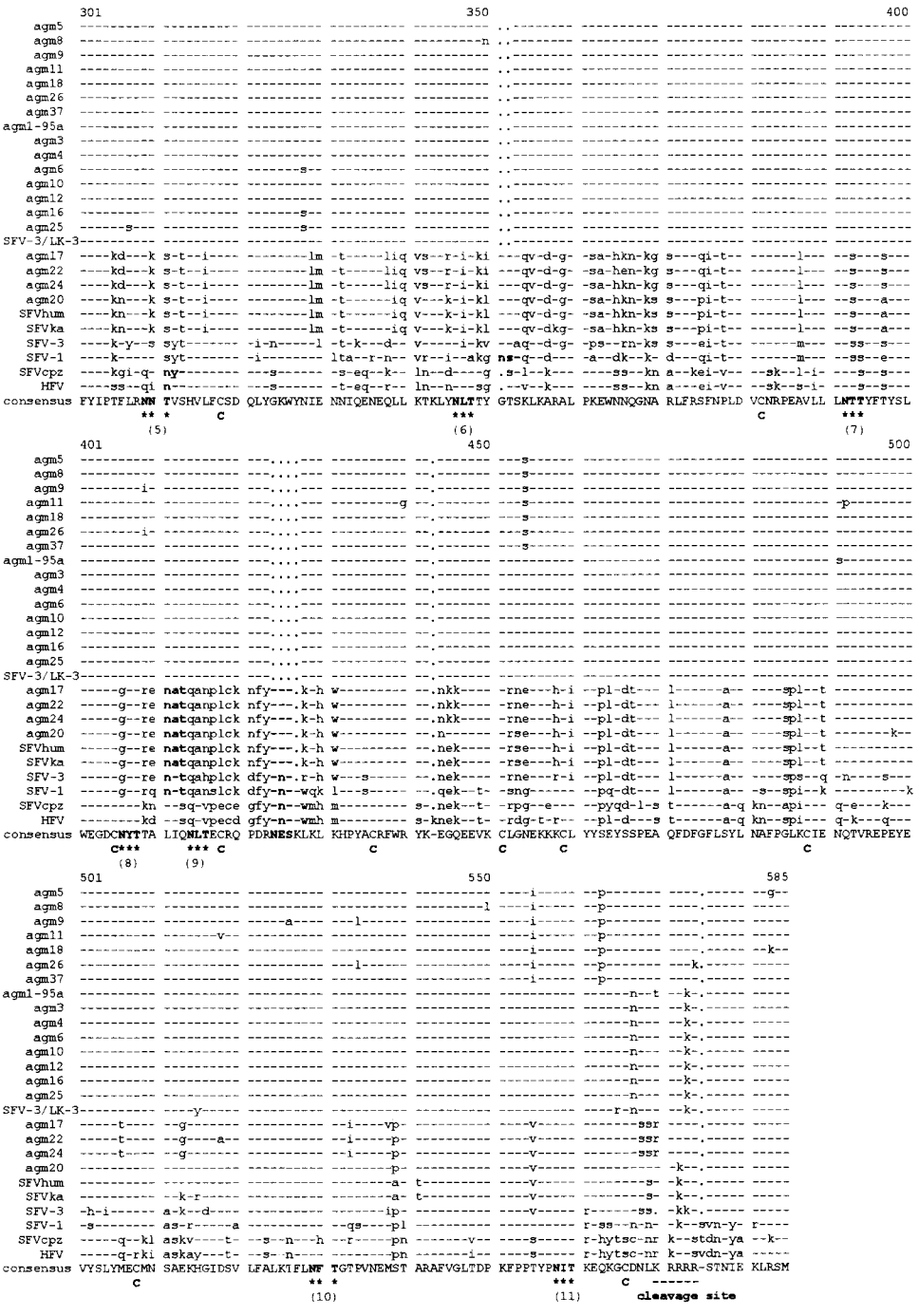

FIG. 2—*Continued.*

isolates of group 2 and of SFV-1, SFVcpz, and HFV are positioned 10 amino acids downstream of sites 9 of the AGM isolates of group 1.

Phylogenetic analysis of FV from different individuals. Pairwise alignments of the SU sequences showed various degrees of genetic homologies. FV from AGM could be divided into two main groups, which were again composed of minor clusters (A to D) with sequence homologies of more than 95%. Remarkably, cluster B contained two pairs of 100% identical amino acid sequences (pair agm6 and agm16 and pair agm3 and agm4). In the phylogenetic tree established by the PHYLIP program, these groups and clusters can easily be distinguished (Fig. 3). The range of the pairwise homology rates within each cluster or between the different clusters is shown in Table 3. Remarkably, the first group (comprising clusters A and B and SFV-3/LK-3) contains all sequences with the smallest SU, i.e., 563 amino acids. All isolates of this group (except SFV-3/LK-3) originate from monkeys which had been obtained by two shipments in July 1979 and October 1980. All sequences of the second group (clusters C and D) revealed 568

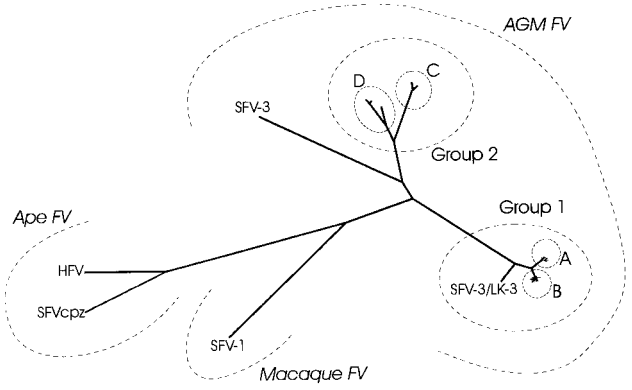

FIG. 3. Phylogenetic tree of the FV investigated, compiled by CLUSTAL and PHYLIP programs (DNAdist and KITSCH). Resulting clusters and groups are indicated. Sequences compiling the clusters are not labeled.

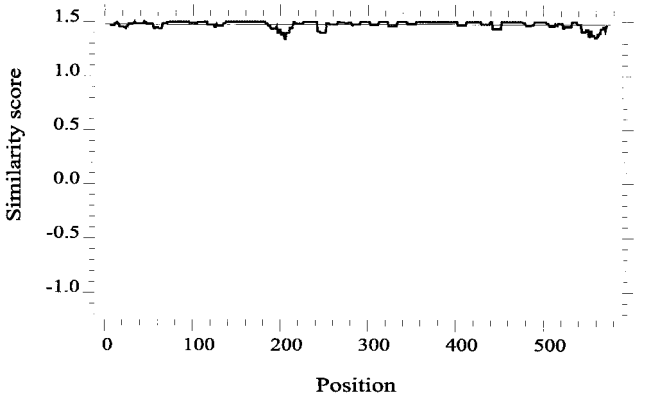

FIG. 4. Similarity plot of the amino acid sequences of the FV isolates of group 1. The symbol comparison table used was plotsimpep.cmp, which has a similarity score of 1.5 for perfect symbol matches.

amino acids. Host monkeys of these isolates had been delivered earlier than the previous ones, with the exception of agm17 which had also been shipped in October 1980. The sequence SFVhum derived from PBL of the infected animal caretaker is 99.3% (98.8% at the amino acid level) homologous to the respective isolate SFVka; these two sequences, together with the sequence agm20, form cluster D. The AGM sequence SFV-3 (which does not originate from this colony) has 567 amino acids and is about 82% homologous to clusters C and D but only about 71% homologous to clusters A and C. Nevertheless, the homology of SFV-3 to any FV of this AGM colony is higher than that to FV from other species (Fig. 3). The range of pairwise homologies between all AGM sequences and SFV-1 or SFVcpz/HFV is also shown in Table 3.

The sequence variability found in this colony was analyzed in detail. The two major groups differ in length by 5 amino acids which are present only in the sequences of the second group (clusters C and D). G and T are found at amino acid positions 351 and 352, and N, E, and S are found at amino acid positions 424 to 426. Furthermore, AGM prototype SFV-3 differs from the sequences of the second group in that it is missing one lysine at the beginning of the cleavage site. Moreover, analysis of single base exchanges (in addition to the amino acid insertions described) revealed that, altogether, 3,004 substitutions against the consensus sequence are present in the 22 FV sequences derived from this colony; 974 of them (32.4%) resulted in amino acid exchanges.

Figure 4 shows a similarity plot of the 16 amino acid sequences of the first group of AGM FV (clusters A and B and SFV-3/LK-3). It emphasizes the homogeneous distribution of variant amino acids over the entire SU; deviations of up to 5% were not concentrated in a hot spot. Thus, the SU of FV from AGM does not contain a hypervariable region. A similarity plot of all AGM FV, including the six sequences of the second group, and SFV-3 simulated a peak of variability in the 3′ region of SU which corresponds to the insertions described (not shown).

Genomic variability of the FV population within a single AGM. To monitor FV sequence variation over the lifetime of one animal, seven FV isolates were obtained from animal no. 1 (AGM1) at six time points between 1982 and 1995 (Table 4). To investigate the virus population without selection of cell culture-adapted strains, DNA was extracted directly from one defined throat swab that also gave rise to isolates agm1-95a and agm1-95b (isolates obtained in 1995) by parallel cultures. After PCR and cloning of the amplification products, 22 clones (ra28 to ra57) originating from one ligation reaction were analyzed as described previously.

Multiple alignment of all SU nucleotide sequences from this animal (data not shown) revealed highly conserved sequences: all of them were more than 99.5% homologous, and two of them were identical. Interestingly, these two identical sequences were ra46 (one proviral clone from the 1995 throat swab) and agm1-94a (an isolate obtained in 1994), whereas the two isolates obtained simultaneously in 1995 from the same throat swab were different. Each sequence revealed between one and five substitutions, compared to the consensus sequence. The multiple amino acid sequence alignment is shown

TABLE 3. Range of pairwise sequence homologies within and between the different clusters of FV SU sequences

Cluster(s) analyzed or compared	Nucleic acids (% identity)	Amino acids (% similarity)	Amino acids (% identity)
A	98.99–99.55	98.96–99.65	98.27–99.31
B	98.88–99.78	98.44–100	98.27–100
C	99.11–99.44	99.31–99.66	98.63–99.31
D	94.82–99.33	97.42–99.14	96.74–98.80
A and B	96.35–97.36	97.75–98.96	96.36–97.92
A and C	76.14–77.03	81.79–82.47	75.26–76.98
A and D	76.98–77.81	81.62–82.99	75.60–77.84
B and C	75.98–77.31	81.10–82.82	75.26–76.98
B and D	76.87–78.48	81.79–83.33	75.77–77.66
C and D	88.85–91.30	95.19–96.56	92.44–94.50
AGM FV and rhesus FV	68.94–73.11	76.54–84.42	65.24–72.77
AGM FV and ape FV	64.27–66.33	72.73–75.73	60.55–62.56
Rhesus FV and ape FV	67.67–68.05	76.07–76.37	63.18–63.42
Chimp FV and HFV	83.97	92.81	88.01

TABLE 4. FV isolation from AGM1

Isolate	Date of isolation
agm1-82	Jan. 1982
agm1-90	Oct. 1990
agm1-91	April 1991
agm1-94a	Jan. 1994
agm1-94b	Dec. 1994
agm1-95a	Jan. 1995
agm1-95b	Jan. 1995

in Fig. 5. Minimal variability was found homogeneously distributed over the SU segment with one exception: at nucleotide position 1031, four sequences (ra28, -30, -31, and -33) revealed the same substitution of A to G, resulting in an amino acid exchange from N to S (amino acid position 328). Altogether, 64 substitutions were located in 29 sequences represented by a total of 51,591 bases. Thirty-five substitutions (55%) were nonsynonymous; two of these resulted in stop codons (amino acid positions 147 in ra57 and 108 in agm1-91). At the amino acid level, one pair and one group of five identical sequences were detected (pair ra30 and ra31 and group ra43, ra46, ra51, ra55, and agm1-94a).

The sequence agm1-95a was used for the analysis of FV isolates from different monkeys and was found to fit into cluster B (Fig. 3). A dendrogram of the nucleotide sequences from cluster B, including all sequences from AGM1 (Fig. 6), shows that the sequences from AGM1 form a subcluster within cluster B. The isolates taken from AGM1 at different time points are distributed randomly over the whole cluster of proviral clones obtained directly from the throat swab; only the two oldest isolates (from 1982 and 1990) and two of the proviral clones (ra42 and ra57) were minimally separated from all other sequences from AGM1. Thus, only a minimal genetic drift, if any, could be detected from 1982 to 1991, whereas no drift could be recognized between 1991 and 1995.

To define the background of substitutions due to PCR, cloning, and sequencing artifacts, a molecular clone of SU was amplified, cloned, and sequenced by the same methods used for all other isolates and the throat swab. Twenty clones originating from the same ligation reaction were analyzed by sequencing the region between nucleotides 1000 and 1500 (data not shown). Within approximately 10,000 nucleic acids sequenced, only two substitutions were detected. This background is in the range of the error rate described for the employed *Taq* polymerase (19). Therefore, the minor variability found in the FV sequences from AGM1 (64 substitutions in 51,591 bases or 12 in 10,000 bp) was about sixfold higher than the inaccuracies of the amplification, cloning, and sequencing system used.

DISCUSSION

Primate FV reveal a particular mode of persistence characterized by apathogenicity in vivo in spite of severe cytopathogenicity in cell culture. To elucidate the role of FV genetic variability in the establishment and maintenance of persistence, we compared the FV SU sequences occurring in the population of an AGM colony and in an individual animal. It is known that FV of different primate species (chimpanzee, rhesus monkey, and AGM) reveal considerable (type-defining) sequence differences with the highest divergencies in the *gag* gene (46). Thus, comparing sequences obtained from different species was not the goal of this study.

Phylogenetic analysis of isolates from different AGM living in a closed colony revealed different clusters of FV addressable as strains and subtypes of AGM FV (Fig. 3). For example, clusters A to D might be designated as strains, and groups 1 and 2 might be designated as subtypes. The divergence between groups 1 and 2 ranged between 22 and 25%; this value is on the same order of magnitude as that reported for comparable genome segments of different subtypes of HIV or SIV (27) but significantly higher than that of subtypes of simian T-lymphotropic virus type 1 or 2 or HTLV type 1 or 2, which is 8% at most (13, 15). Obviously, genetic variability of FV has been sufficient to generate highly divergent strains within one species. On the other hand, the divergency within the strains or cluster A, B, or C is less than 1.2%, which is much lower than that reported for HIV or SIV (up to 15% [27]). Within cluster D, divergency is about 5%, pointing out how arbitrary an attempt to classify the clusters as strains or subtypes would be. Unfortunately, the period of time required to reach a certain level of diversity remained undefined since the time point and source of infection of each monkey are unknown. However, the phylogenetic tree proves that only a few strains have spread in this colony, as the homology rates within one cluster are remarkably high. It can be speculated that infection of the monkeys occurred before delivery to our institute, maybe in wild-living clans before capture or during transport; in fact, the animals had usually been kept in single cages in the final colony to avoid biting, which is considered to be the mode of primate FV transmission. Furthermore, the different clusters are associated with different capture and purchase dates of monkey herds: clusters A and B originate from monkeys obtained in July 1979 or October 1980, whereas the host monkeys of subtypes C and D (except isolate agm17) had been delivered between 1975 and February 1979. This interpretation of the phylogenetic tree is confirmed by the two sequences originating from the infected animal caretaker: they are closely related to AGM isolate agm20 obtained from a monkey that had been delivered in 1975, which is in accordance with the time of infection of the worker by a monkey bite in 1976 (41). Interestingly, agm17 had been obtained in 1980 but belongs to the earlier cluster C. This may be explained by infection in our institute, because housing in single cages was temporarily interrupted for experimental procedures or occasional breeding attempts.

To interpret the degree of sequence variability, only closely related viruses should be analyzed, as the comparison of nonrelated sequences may lead to overinterpretation of local or overall variability rates (45). Therefore, only clusters A and B were compared to identify hypervariable regions in the FV SU region. The high homology rates of more than 96% between isolates of this group point to common ancestors. Moreover, all sequences of clusters A and B reveal the same length of SU, whereas the main difference from group 2 is the insertion of 5 amino acids which would mimic hypervariable regions. However, the similarity plot of group 1 (Fig. 4) clearly revealed that hypervariable regions comparable to the five variable loops in the HIV type 1 gp120 do not exist in FV. Hypervariability in the immunodominant loops of HIV type 1 is considered as the prerequisite for the emergence of immune escape mutants (3). Such a mechanism is obviously not used by AGM FV. The difference between FV and HIV is illustrated by differences in the ratio of nonsynonymous versus synonymous substitutions: a ratio significantly greater than 1 is supposed to be indicative of positive selection for sequence changes (23, 24). In all AGM FV analyzed, the actual figure of the ratio was 0.48, which is in the range expected for random mutations or even for selection against mutations (23, 24). In contrast, HIV reveals ratios of nonsynonymous versus synonymous changes of up to 2.9 (23, 24), which implies positive selection of changes, at least within particular hypervariable domains.

To further compare the degree of intrahost genetic diversity of FV to that of other retroviruses, a cross-sectional sample of virus produced in one animal at one time point was investigated. Nucleic acid analysis was applied directly to a throat swab without virus isolation avoiding selection of cell culture-adapted variants. Throat mucosa is the exclusive site where viral replication has been detected in naturally infected primates (8). PCR was done without previous reverse transcription of RNA, since besides intracellular proviral genomes sufficient DNA is also present in virions (25). Sequencing of 20

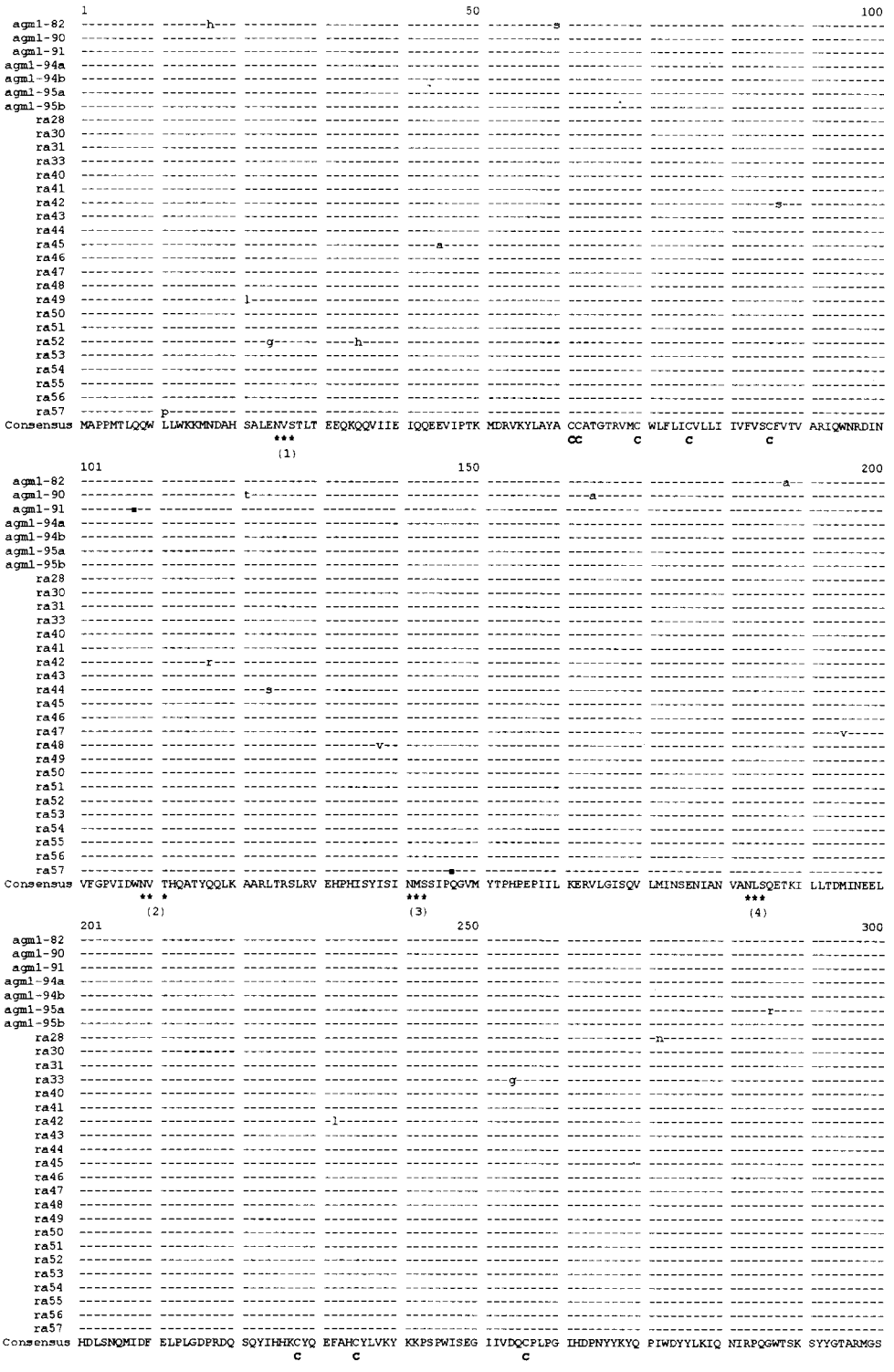

FIG. 5. Multiple amino acid sequence alignment of the SU regions of the proviral clones and isolates obtained from agm1, generated by CLUSTAL. ■, stop codons in agm1-91 and ra57; C, cysteines; ***, N-glycosylation sites (N-Xaa-S/T); cleavage, cleavage site between SU and TM.

molecular clones originating from one PCR has been considered to be sufficient to detect variants forming at least 10% of the virus population present in the respective sample (26). The divergency of 0.5% found between the 22 molecular clones obtained from one oral mucosa sample is very low but significant, indicating that the recovered variants indeed represent the range of genetic divergency of FV within this organ. This divergency is in the same range as that described for viruses of the HTLV/BLV group (<0.5% [13]) but considerably lower than that for lentivirus, which reaches 12% divergency at

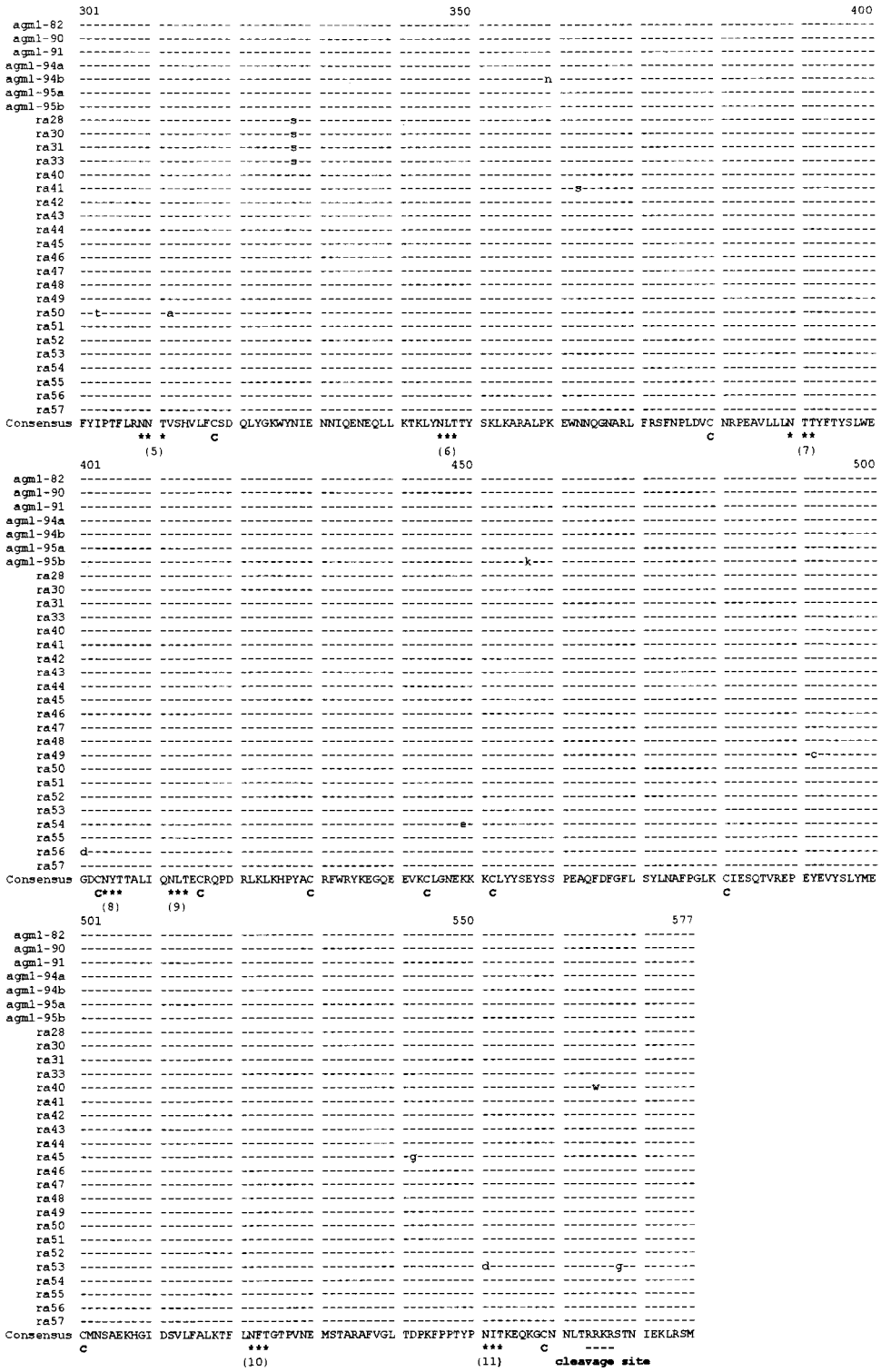

FIG. 5—*Continued.*

amino acid level within one host (5). Furthermore, absence of hypervariable regions and the low proportion of nonsynonymous substitutions again argue against a role of variability in the establishment of viral persistence, as already shown above for isolates from different monkeys. The two stop codons detected in isolate agm1-91 and in the proviral clone ra57 may be due to PCR artifacts or nonfunctional DNA present in cell culture or in the throat swab sample, respectively.

Interestingly, viral isolates from various time points between 1991 and 1995, including two distinguishable isolates from the

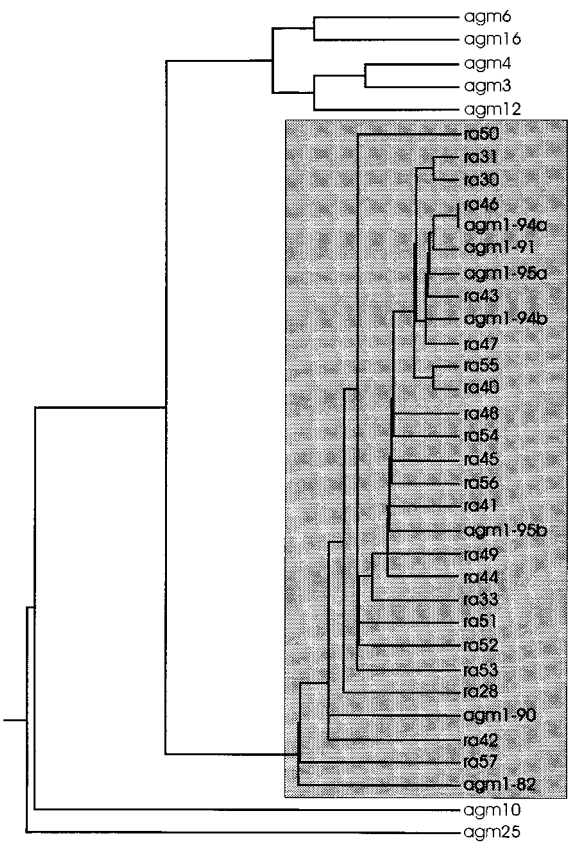

FIG. 6. Dendrogram of the SU nucleotide sequences of cluster B, including all sequences obtained from agm1 (shadowed box), compiled by CLUSTAL and PHYLIP programs (DNAdist and KITSCH).

swab used for generation of proviral clones, maintain the same range of sequence differences as the clones (Fig. 6). The fact that the virus population reveals this high degree of homogeneity indicates that sequence alterations due to cell culture adaptation can be neglected and any isolate represents proviral and viral population at the time point of isolation. Furthermore, no genetic drift could be recognized between 1991 and 1995. Only from 1982 to 1991 has a minimal drift of about 0.3% occurred, suggesting a variability rate of about 3×10^{-4} per site per year (six base substitutions per 1,779 bp per 9 years). This is slightly higher than that reported for the HTLV/BLV group (10^{-4} for BLV [48] and HTLV type 1 [12]) but much lower than that of lentiviruses, which ranges between 10^{-2} and 10^{-3} (14).

In conclusion, genetic variability of FV is sufficient to generate highly divergent strains in different species and also different strains within one primate species but probably too low in individual hosts to be instrumental in viral persistence. There may be different reasons for this degree of stability. In general, the amount of genetic variability of a gene is determined by the mutation rate and either positive or negative selection pressure. The mutation rate depends on the error rate of nucleic acid polymerases and on the replication rate. The fidelity of FV reverse transcriptase is not known. However, it has been argued that the mutation rate is the least important aspect of viral variation (6, 32, 45), since it is in the same range for all retroviruses tested (including the extraordinarily stable HTLV). Therefore, the most important factor may be the replication rate. For HIV and SIV, the high genomic variability has been reported to be due to the extraordinary high rates of viral replication in the host (7, 45), whereas the stability of HTLV/BLV has been explained by its low replication rate (47). The latter explanation is probably also true for FV, since the replication rate of FV is very low in infected monkeys: only low levels of viral RNA can be detected in the oral mucosa but not in other FV DNA-positive tissues, including blood cells (8). Finally, selection pressure may be considered: positive selection of variants could emerge from immune pressure, while negative selection might be favored by functional constraints. The low ratio of nonsynonymous against synonymous substitutions points to random mutation or even negative selection, preventing any changes (23, 24). This clearly contradicts the hypothesis that variability in the SU region might play an essential role for FV persistence.

Taken together, compared to lentiviruses, the genome of FV was found to be very stable, probably due to the low replication rate of FV in vivo and to missing selection pressure for sequence changes. Therefore, the probability of biological changes, e.g., expansion of pathogenic potential or host range, can be considered low for FV. From a practical point of view, the genetic stability proven by these investigations may encourage efforts to develop FV vectors for human gene therapy (2, 28, 35).

ACKNOWLEDGMENTS

We thank Otto Haller for continuous support and critical reading of the manuscript.

This work was supported by the Deutsche Forschungsgemeinschaft, grant Ne 213/5-1, and by the EC, grant BMH4-CT 97-2010 in the BIOMED 2 program.

REFERENCES

1. **Ali, M., G. P. Taylor, R. J. Pitman, D. Parker, A. Rethwilm, R. Cheingsong-Popov, J. N. Weber, P. D. Bieniasz, J. Bradley, and M. O. McClure.** 1996. No evidence of antibody to human foamy virus in widespread human populations. AIDS Res. Hum. Retroviruses **12:**1473–1483.
2. **Bieniasz, P. D., O. Erlwein, A. Aguzzi, A. Rethwilm, and M. O. McClure.** 1997. Gene transfer using replication-defective human foamy virus vectors. Virology **235:**65–72.
3. **Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, and G. M. Shaw.** 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CT escape virus. Nat. Med. **3:**205–311.
4. **Bothe, K., A. Aguzzi, H. Lassmann, A. Rethwilm, and I. Horak.** 1991. Progressive encephalopathy and myopathy in transgenic mice expressing human foamy virus. Science **253:**555–557.
5. **Bruce, C., C. Clegg, A. Featherstone, J. Smith, and J. Oram.** 1993. Sequence analysis of the gp120 region of the env gene of Ugandan human immunodeficiency proviruses from a single individual. AIDS Res. Hum. Retroviruses **9:**357–363.
6. **Coffin, J. M.** 1992. Genetic diversity and evolution of retroviruses. Curr. Top. Microbiol. Immunol. **176:**143–164.
7. **Coffin, J. M.** 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science **267:**483–489.
8. **Falcone, V., J. Leupold, J. Clotten, E. Urbanyi, O. Herchenröder, W. Spatz, B. Volk, N. Böhm, A. Toniolo, D. Neumann-Haefelin, and M. Schweizer.** 1999. Sites of simian foamy virus persistence in naturally infected African green monkeys: latent provirus is ubiquitous whereas viral replication is restricted to the oral mucosa. Virology **257:**7–14.
9. **Felsenstein, J.** 1989. PHYLIP—Phylogeny Inference Package (version 3.2). Cladistics **5:**164–166.
10. **Fitch, W. M., and E. Margoliash.** 1967. Construction of phylogenetic trees. Science **155:**279–284.
11. **Flügel, R. M., A. Rethwilm, B. Maurer, and G. Darai.** 1987. Nucleotide sequence analysis of the env gene and its flanking regions of the human spumaretrovirus reveals two novel genes. EMBO J. **6:**2077–2084.
12. **Gessain, A., R. C. Gallo, and G. Franchini.** 1992. Low degree of human T-cell leukemia/lymphoma virus type I genetic drift in vivo as a means of monitoring viral transmission and movement of ancient human populations. J. Virol. **66:**2288–2295.
13. **Gessain, A., R. Mahieux, and G. de Thé.** 1996. Genetic variability and molecular epidemiology of human and simian T cell leukemia/lymphoma

virus type I. J. Acquired Immune Defic. Syndr. Hum. Retrovirology **13** (Suppl. 1)**:**132–145.
14. **Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks.** 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. Science **232:**1548–1553.
15. **Heneine, W.** 1996. The phylogeny and molecular epidemiology of human T-cell lymphotropic virus type II. J. Acquired Immune Defic. Syndr. Hum. Retrovirology **13**(Suppl. 1)**:**236–241.
16. **Heneine, W., W. M. Switzer, P. Sandstrom, J. Brown, V. Shanmugam, C. Schable, M. Schweizer, D. Neumann-Haefelin, L. E. Chapman, and T. M. Folks.** 1998. Identification of a human population endemically infected with simian foamy virus. Nat. Med. **4:**403–407.
17. **Herchenröder, O., R. Renne, D. Loncar, E. K. Cobb, K. K. Murthy, J. Schneider, A. Mergia, and P. A. Luciw.** 1994. Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFVcpz): high homology to human foamy virus (HFV). Virology **201:**187–199.
18. **Hooks, J. J., and C. J. Gibbs.** 1975. The foamy viruses. Bacteriol. Rev. **39:**169–185.
19. **Keohavong, P., and W. G. Thilly.** 1989. Fidelity of DNA polymerases in DNA amplification. Proc. Natl. Acad. Sci. USA **86:**9253–9257.
20. **Khanna, R., S. R. Burrows, and J. M. Burrows.** 1997. The role of cytotoxic T lymphocytes in the evolution of genetically stable viruses. Trends Microbiol. **5:**64–69.
21. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
22. **Kupiec, J. J., A. Kay, M. Hayat, R. Ravier, J. Peries, and F. Galibert.** 1991. Sequence analysis of the simian foamy virus type 1 genome. Gene **101:**185–194.
23. **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.
24. **Li, W. H., M. Tanimura, and P. M. Sharp.** 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. **5:**313–330.
25. **Linial, M. L.** 1999. Foamy viruses are unconventional retroviruses. J. Virol. **73:**1747–1755.
26. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfeldt-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.
27. **Myers, G., B. Korber, B. H. Hahn, K.-T. Jeang, J. W. Mellors, F. E. McCutchan, L. E. Henderson, and G. N. Pavlakis.** 1995. Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N. Mex.
28. **Nestler, U., M. Heinkelein, M. Lucke, J. Meixensberger, W. Scheurlen, A. Kretschmer, and A. Rethwilm.** 1997. Foamy virus vectors for suicide gene therapy. Gene Ther. **4:**1270–1277.
29. **Neumann-Haefelin, D., U. Fleps, R. Renne, and M. Schweizer.** 1993. Foamy viruses. Intervirology **35:**196–207.
30. **Neumann-Haefelin, D., A. Rethwilm, G. Bauer, F. Gudat, and H. zur Hausen.** 1983. Characterization of a foamy virus isolated from Cercopithecus aethiops lymphoblastoid cells. Med. Microbiol. Immunol. **172:**75–86.
31. **Nowak, M. A., R. M. Anderson, A. R. McLean, T. F. Wolfs, J. Goudsmit, and R. M. May.** 1991. Antigenic diversity thresholds and the development of AIDS. Science **254:**963–969.
32. **Pelletier, E., W. Saurin, R. Chenier, N. L. Letvin, and S. Wain-Hobson.** 1995. The tempo and mode of SIV quasispecies development in vivo calls for massive viral replication and clearance. Virology **208:**644–652.
33. **Renne, R., E. Friedl, M. Schweizer, U. Fleps, R. Turek, and D. Neumann-Haefelin.** 1992. Genomic organization and expression of simian foamy virus type 3 (SFV-3). Virology **186:**597–608.
34. **Rethwilm, A.** 1996. Unexpected replication ways of foamy viruses. J. Acquired Immune Defic. Syndr. Hum. Retrovirology **13**(Suppl. 1)**:**248–253.
35. **Russel, D. W., and A. D. Miller.** 1996. Foamy virus vectors. J. Virol. **70:**217–222.
36. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Schweizer, M., and D. Neumann-Haefelin.** 1995. Phylogenetic analysis of primate foamy viruses by comparison of *pol* sequences. Virology **207:**577–582.
39. **Schweizer, M., B. Corsten, and D. Neumann-Haefelin.** 1988. Heterogeneity of primate foamy virus genomes. Arch. Virol. **99:**125–134.
40. **Schweizer, M., R. Turek, H. Hahn, A. Schliephake, K. O. Netzer, G. Eder, M. Reinhardt, A. Rethwilm, and D. Neumann-Haefelin.** 1995. Markers of foamy virus (FV) infections in monkeys, apes, and accidentally infected humans: appropriate testing fails to confirm suspected FV prevalence in man. AIDS Res. Hum. Retroviruses **11:**161–170.
41. **Schweizer, M., V. Falcone, J. Gänge, R. Turek, and D. Neumann-Haefelin.** 1997. Simian foamy virus isolated from an accidentally infected human individual. J. Virol. **71:**4821–4824.
42. **Stiles, G. E., J. L. Bittle, and U. J. Cabasso.** 1964. Comparison of simian foamy virus strains including a new serological type. Nature **201:**1350–1351.
43. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. Improved sensitivity of profile searches through the use of sequence weights and gap excision. Comput. Appl. Biosci. **10:**19–29.
44. **Vogt, V. M.** 1997. Retroviral virions and genomes, p. 27–69. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
45. **Wain-Hobson, S.** 1996. Running the gamut of retroviral variation. Trends Microbiol. **4:**135–141.
46. **Wang, G., and M. J. Mulligan.** 1999. Comparative sequence analysis and predictions for the envelope glycoproteins of foamy viruses. J. Gen. Virol. **80:**245–254.
47. **Wattel, E., J. P. Vartanian, C. Pannetier, and S. Wain-Hobson.** 1995. Clonal expansion of human T-cell leukemia virus type 1-infected cells in asymptomatic and symptomatic carriers without malignancy. J. Virol. **69:**2863–2868.
48. **Willems, L., E. Thienpont, P. Kerkhofs, A. Burny, M. Mammerickx, and R. Kettmann.** 1993. Bovine leukemia virus, an animal model for the study of intrastrain variability. J. Virol. **67:**1086–1089.
49. **Yu, S., D. N. Baldwin, S. R. Gwynn, S. Yendapalli, and N. L. Linial.** 1996. Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. Science **271:**1579–1582.