JOURNAL OF VIROLOGY, Nov. 1995, p. 7349–7353 Vol. 69, No. 11
0022-538X/95/$04.00+0

Regulatory Genes of Simian Immunodeficiency Viruses from West African Green Monkeys (*Cercopithecus aethiops sabaeus*)

VERONIQUE JUBIER-MAURIN,[1]* PASCALE SARNI-MANCHADO,[1]† FRANCISCO VEAS,[1] NICOLE VIDAL,[1] FREDERIC BIBOLLET-RUCHE,[1] JEAN-PAUL DURAND,[2]‡ ANH GALAT-LUONG,[3] AND GERARD CUNY[1]

Laboratoire Rétrovirus, Institut Français de Recherche Scientifique pour le Développment en Coopération (ORSTOM), 34032 Montpellier Cedex 1, France,[1] and Institut Pasteur[2] and Laboratoire de Primatologie, ORSTOM,[3] Dakar, Senegal

Received 21 March 1995/Accepted 9 August 1995

The high seroprevalence of simian immunodeficiency viruses (SIVs) in African green monkeys (AGMs) without immunological defects in their natural hosts has prompted consideration of SIV-infected AGMs as a model of apathogenic SIV infection. Study of the molecular mechanisms of SIVagm asymptomatic infection could thus provide clues for understanding the pathogenesis of human immunodeficiency viruses. Regulatory genes could be candidates for genetic control of SIVagm apathogenicity. We have characterized Vpr, Tat, Rev, and Nef genes of two SIVagm strains isolated from naturally infected sabaeus monkeys captured in Senegal. The results provide further evidence that SIVagm from West African green monkeys is the most divergent class of AGM viruses, with structural features in long terminal repeat sequences and Vpr and Tat genes that distinguish them from viruses isolated from other AGM species (vervet, grivet, and tantalus monkeys).

Simian immunodeficiency viruses (SIVs) are primate retroviruses closely related to the etiologic agents of human AIDS, human immunodeficiency virus types 1 and 2 (HIV-1 and HIV-2, respectively).

Distinct viruses have been identified from several different African simian species since the initial isolation of SIV from immunodeficient captive Asian macaques (SIVmac) (5). Five subgroups of primate viruses have been characterized according to sequence data: HIV-1 and SIVcpz (25, 40); HIV-2, SIVmac, and SIVsm (24); SIVmnd (47); SIVsyk (22); and SIVagm (6, 28).

It should be noted that African monkey viruses do not cause immunologic defects in their natural hosts. Nevertheless, in a heterologous host, the pig-tailed macaque, two SIVagm strains isolated from vervet monkeys (*Cercopithecus aethiops pygerythrus*) induced an immunodeficiency syndrome: the SIVagm 155 uncloned virus (16) and the very recently characterized molecular clone SIVagm 9063, as well as its parental virus (23). The only well-documented simian AIDS-like illness model for human disease is that observed in macaques experimentally infected with SIVsm, but no SIV-infected macaques have been found in nature.

Understanding the features of simian viruses may thus shed some light on the pathogenesis of human AIDS. SIVagm infection of wild-caught African green monkeys (AGMs) could be considered as a model of persistent and asymptomatic simian retrovirus infection because of the unusually high seroprevalence of SIVagm infection. Up to 80% seroprevalence has been reported in naturally infected adults (14) from West African species (*Cercopithecus aethiops sabaeus*). SIVagm strains were previously identified in four AGM species: in vervets found from Southern Ethiopia to South Africa (6, 30), grivets (*C. aethiops aethiops*) found only in Ethiopia and the Sudan (28), sabaeus monkeys from West Africa (commonly called green monkeys [1, 32]), and tantalus monkeys (*C. aethiops tantalus*) (32) from the Central African Republic. Four East African SIVagm strains have been completely sequenced (2, 12, 13, 28). The characterization of a full-length provirus clone from a sabaeus SIVagm strain was recently published (27).

At least two nonexclusive reasons could explain the apparent apathogenicity of SIVagm: a natural host resistance or the genetic characteristics of the viruses. We carried out studies with the SIVagm D30 isolate (89030 in reference 8) from *C. aethiops sabaeus* caught in West Africa (Senegal). The lack of clinical signs in a heterologous host (*Erythrocebus patas*) 2 years after experimental infection (7) could support the hypothesis of genetic control. Another SIVagm strain, D42 (89042 in reference 8), from a sabaeus monkey captured in the same geographic region was also selected for the present study.

In order to understand which regulatory genes could be responsible for the pathogenicity patterns shown by these two isolates, we first characterized these genes.

(Results were presented in part at the VIIIth International Conference on AIDS in Africa, Marrakech, Morocco, 12 to 16 December 1993 [28a].)

Comparisons between regulatory proteins of D30 and D42 strains and other AGM viruses. To determine the nucleotide sequences, genomic DNAs were prepared from persistently SIVagm D30- and D42-infected Molt-4 (clone 8) cells obtained by coculture with peripheral blood mononuclear cells from the two animals. One microgram of genomic DNA was used as the template for PCR. PCR was performed with primers first chosen on the basis of published SIVagm sequences (33, 34) and then from the D30 and D42 isolate sequences we established. Most of the nucleotide sequences were obtained by direct sequencing of PCR products. Depending on the specificity of the reaction, the amplified DNA was purified directly or after the bands had been separated on low-melting-point agarose gels (GTG NuSieve; FMC, Rockland, Maine) with Gene Clean

* Corresponding author. Mailing address: Laboratoire Rétrovirus, ORSTOM, 911 Av. Agropolis, B. P. 5045, 34032 Montpellier Cedex 1, France. Phone: (33) 67 61 74 21. Fax: (33) 67 61 75 58 or (33) 67 54 78 00. Electronic mail address: maurin@citi2.fr.

† Present address: Laboratoire des polymères et des techniques physicochimiques, IPV, INRA, 34060 Montpellier, France.

‡ Present address: DIMT, PHARO IMTSSA, 13007 Marseille, France.

TABLE 1. Percent amino acid identities between regulatory proteins of SIVagm strains

SIVagm strain	% Amino acid identity[a]																							
	Vpr						First Tat exon						Rev						Nef					
	D42	sab-1	155	TYO	3	677	D42	sab-1	155	TYO	3	677	D42	sab-1	155	TYO	3	677	D42	sab-1	155	TYO	3	677
D30	88	88	42	44	44	43	76	69	37	44	39	45	81	82	45	44	43	42	84	84	50	51	51	51
D42		91	40	43	43	43		75	37	42	40	44		78	45	43	42	42		82	50	51	51	52
sab-1			41	42	42	45			35	38	36	47			46	45	43	40			55	54	55	54
155				84	87	68				71	67	50				57	77	50				82	86	68
TYO					90	70					70	54					59	43					84	71
3						69						53						46						71

[a] Values were obtained from the multialignment program CLUSTAL (20, 21). East African SIVagm sequences were obtained from the Los Alamos sequence database (33–35).

(BIO 101, La Jolla, Calif.). Some of these blunt-ended fragments were cloned in *Eco*RV-digested M13 BM20 (Boehringer Mannheim), and subsequent single-stranded templates were sequenced with the M13 universal primer. Cycle sequencing was performed by dye primer or dye deoxyterminator techniques with the kit protocols from Applied Biosystems (Foster City, Calif.). Sequences were analyzed with the automated Applied Biosystems model 373A DNA sequencing system.

Four of the five regulatory genes were completely sequenced (i.e., Vpr, Tat, Rev, and Nef [nucleotide accession numbers in GenBank, U20811 to U20814, U20965 to U20968, and U21093]). The predicted gene products were compared with those of other AGM viruses, and the deduced amino acid identities are given in Table 1. This comparison revealed several unexpected findings. There was very little homology between the D30 or D42 isolates and the other AGM virus regulatory proteins. Tat, Vpr, and Rev showed the lowest average homologies: 41% for Tat and 43% for Vpr and Rev. These values are much lower than those obtained by comparisons between other SIVagm proteins (50 to 71% in Tat, 68 to 90% in Vpr, and 43 to 77% in Rev). More precisely, considering species-specific divergence, the degrees of identity between the two viruses and viruses from vervet (strains 155, TYO, and 3) or grivet (strain 677) monkeys were lower than those between viruses from vervet and grivet monkeys (52, 69, and 46% on average for the same proteins, respectively). Moreover, this homology is quite similar to that between SIVagm (strain 155) and sooty mangabey (*Cercocebus atys*) virus (sm H4) (28). In contrast, comparisons with another virus (sab-1) isolated from a sabaeus monkey (27) revealed high levels of amino acid identity (72, 90, and 80% on average for the Tat, Vpr, and Rev proteins, respectively). Nef, on the other hand, was the best-conserved protein (51% amino acid identity on average). These two strains, D30 and D42, are therefore quite distantly related to the earliest-studied AGM viruses.

Despite the high level of divergence, amino acid stretches previously described to be conserved in any SIV or HIV isolates were identified in the D30 and D42 protein sequences. In Rev amino acid sequences, the best-conserved, although mutated, region was the putative arginine-rich nuclear localization domain. The cysteine-rich metal binding motif was highly conserved in the D30 and D42 first coding exon of Tat (see Fig. 3). Concerning Nef gene products of both viruses (Fig. 1), the myristoylation signal was found to be flanking the initiation codon. Two domains, KGGLEG and WK/RFD (underlined in Fig. 1), previously defined as being essential for interactions with GTP and therefore for potential GTPase activity (17, 18), were also found in all of the SIVagm Nef proteins. However, this HIV-1 Nef property was highly controversial (29, 37). It was noteworthy that in the best-conserved part of the Nef protein between HIV-1 and HIV-2, the vervet and grivet viruses possessed specific motifs (boxed in Fig. 1) which were absent from the SIVagm D30 and D42 amino acid sequences. Moreover, in these areas, the highly divergent AGM viruses D30 and D42 were more homologous to HIV-1 or HIV-2 than to other AGM viruses.

Structural differences between SIVagm D30 and D42 and other SIVagm Vpr and Tat proteins. As proposed by several authors (22, 36, 45, 46), the open reading frame juxtaposed between Vif and the first coding exon of Tat in the SIVagm genome has been more accurately designated as Vpr rather than Vpx. D30 and D42 putative Vpr proteins, like those of

```
D30   MGGKSSKQQQERSLWLWAKLRQVPEIQYDMLADPLLGQSSHIQEECAKSLKDGLIKQGD-   59
D42   ..A.....R.Q......S...EA.Q.R.....................E..........-   59
TYO   ..SQN..PAHKKYSK..QA.HKTHVTR.GL.....I.T..TV....D.A.RKS..RKQN-   59
155   ..LG...P.HKKQ.TI.RA.HATRHTR.GL.....I....TL....D.G.RKS..RKRN-   59
3     ..LGN..P.HKKQ.S..HA.HKTRATR.GL.....I....TL....D.A..ES..RKRN-   59
677   ..SSN..R..QGL.K..RG..GK.GADWVL.S...I....TV....G.A..KSWG.----   56
HIV1  ....W..SSVVGWPTVRERM.RAEPA------..GVGAA.RDLE----------------   38
HIV2  ..ASG..KHSRPPRG.QER.LRARAGACGGYWNESG.EY.RF..GSDREQ.SPSCEGRQY   60

D30   --SSRTEEGVKLKHQGRQPSWY---------DDDEETVGFPVRPQVPMRTMTFKLGVDFS  108
D42   --.....D.ITMR.P.......---------.....E.......RI.L.Q.....AI..G  108
TYO   --GNM...ERR.QEGDTWEE.----------S....E.......R..L.Q..Y..A....  107
155   --GNM.P..RR.QDGDQWDE.----------S.E.DE.......R..L.QI.Y..A....  107
3     --GKM.P..R..QEGDKWDE.----------S.E.DE.......R..L.Q..Y..A....  107
677   --GKM.PD.RR.QEGDTFDE.----------.....E.....Q.R..L.Q..Y..A....  104
HIV1  KHGAI.SSNTAATNAAC--A.--------LEAQE..E.....T....L.P..Y.AA..L.   88
HIV2  QQGDFMNTPW.DPAAE.EKNL.RQQNMDDVDS..DDQ.RVS.T.K..L.P..HR.AI.M.  120

D30   HFLKEKGGLEGIYFSERRKKILDLYALNEWGIVDGWQNYTDGPGTRFPKMFGWLFKLVPV  168
D42   ..........................................Y..Q...C......  168
TYO   .........D...Y.D..N...N.........I.D.NAWSK...I....C..FC......  167
155   .........D...Y.D..N...N.........I.D.NAWSK...I.Y.RC..FC......  167
3     .........D...Y.D..NQ..N.........I.D.NAWSE...I.Y.RC..FC......  167
677   ....S....D...Y....E...N.........I.D..A.SP...I.Y.RV..FC......  164
HIV1  ...........LIH.Q..QD....WIYHTQ.YFPD.....P...V.Y.LT...CY.....  148
HIV2  .LI.TR.....MFY....H...NI.LEK.E..IAD.....H...V.Y.MF....W.....  180

D30   DMS--DEAHDSDTHCLLHPAQIAYEDD--PWKEVLVWKFDPMLAVDYVAWRLFPEQVP--  222
D42   ..T--....NN.R.........S....--Q...T...R.............Y.....--  222
TYO   .LH--E..QTCER...V....MGEDP.GISHG.I..........IQ.DPN.EYFTDMHG-  224
155   ALH--E..ETCER...V....LHEDP.GINHG.I.A.........Q.DPS.EYFTDLYST  225
3     .LH--E..ETCER...V....VREDP.GINHG.............Q.DPN.KYLTDMHD-  224
677   .LH--E..RNCER...M....MGEDP.GIDHG.........K...E.RPDM--FKDMHEH  220
HIV1  EPDKVE..NKGENTS....VSLHGM..--.ER...E.R..SR..FHH..RE.H..YFKNC  206
HIV2  .VP--Q.GE.TE....V....TSKF..--.HG.T...E...L..YS.E.FIRY..EFGHK  236

D30   -GLS----------G  226
D42   -...----------A  226
TYO   LVKR----------K  229
155   V.TG----------N  230
3     L.KR----------K  229
677   --AK----------R  223
HIV1  ----------  206
HIV2  S..PEEEWKARLKARGIPFS  256
```

FIG. 1. Alignment of Nef protein sequences from SIVagm and HIV isolates. Nef nucleotide sequences from sabaeus viruses (D30 and D42) were translated and aligned with Nef amino acid sequences of vervet (TYO, 155, and 3), grivet (strain 677), and HIV-1bru and HIV-2rod viruses (33, 34). Dots denote sequence identity with the D30 protein, while dashes represent gaps introduced to maximize the alignment. The bracket above the D30 sequence indicates the myristoylation signal GXXXS found in each Nef sequence shown in the figure. The two domains proposed as putative interacting sites with GTP (17, 18) are underlined. The amino acid stretches specific to the sabaeus Nef sequences, which are absent in East African SIVagm viruses, are boxed.

```
D30    MASGGWLPPTGGDPPRDPPQNPREEVPGWLETWDLDREPFDEWLRDMLQDLNQEAQYHFP   60
D42    .........M.....K...K.....I.........A.....................C...   60
TYO    ....---------------RD...AR..EV.I...S...W.K.............RL..G   45
155    ....---------------RD....R..G..I...S...W.......VEEI.N..KL..G   45
3      ....---------------RD...AR..E..I...S...W........E.I....KM..G   45
677    ....---------------RD...PL.....I.......W....Q...R...E..RR..G   45

D30    RNLLFRLWWNIVEEPAIDRGQSRLEGWYKYYRVLQKALFVHMKGRCCKPKTHPAYGPGGG  120
D42    ......................P.I........I.....Y...................A.  120
TYO    .E...QV.-.YCQ.EGERH.TPMM.RA.....LV.......FRCG.RRRQPFEP.E----  100
155    .E..YQV.-.YCQ.EGERQ.RPIA.RA.....LV.......FRCG.RRRQPFEP.E----  100
3      .E...QV.-.YCQ.EGERNRTPM..RA....KLV.......FRCG.RRRQPFEP.E----  100
677    M.M.I.V.-.YCV.EGRRHNTPWN.IG.....IV..SM...FRCG.RRRGPFSP.E----  100

D30    --PPGLGGAPGGAAAAPPGL-  138
D42    GP......T....S..A...-  140
TYO    ERRD.Q..--.R.NRV....E  119
155    ERRN.Q..--.RPGRV....D  119
3      ERRD.Q..--.R.GRV....D  119
677    ERRN.Q..--.APPP-....A  118
```

FIG. 2. Alignment of Vpr amino acid sequences of AGM viruses. Vpr nucleotide sequences from sabaeus viruses (D30 and D42) were translated and aligned with Vpr protein sequences of previously described SIVagm strains (TYO, 155, 3, and 677) (33, 34). Dots denote sequence identity with the D30 protein, while dashes represent gaps introduced to optimize the alignment. The 15-amino-acid insertion of D30 and D42 Vpr proteins is shown on the top line by a solid bar.

other SIVagm strains, lack the proline-rich region specific to the carboxy terminus of the HIV-2, SIVmac, and SIVsm Vpx proteins.

Surprisingly, the Vpr gene product displayed a very low level of identity (40 to 44%), whereas it was relatively well conserved within the SIVagm group (68 to 90%) (Table 1). The major insertion (15 amino acids) noted in the amino terminus of the D30 and D42 Vpr proteins (Fig. 2) was partly responsible for this low percentage of identity. This insertion took place with no shift in the open reading frame. The two isolates showed the same amino acid insertions with only two different residues, one being an equivalent amino acid (K [lysine] versus R [arginine]). In other parts of the Vpr product, the D30 and D42 amino acid sequences could be aligned with those of previously reported SIVagm strains (Fig. 2).

Because two almost identical decamers were found to be bordering nucleotides inserted in the D30 Vpr DNA sequence, we suggest that this insertion occurred in a common progenitor of D30 and D42 viruses. The fact that the 3′ repeated decamer accumulated mutations in the D42 Vpr gene could indicate that this isolate diverged before D30. Another explanation could be that it might have mutated faster under a lower level of selective pressure.

This large insertion probably modified the structure of these proteins compared with those of the other SIVagm strains. The recently described full-length sabaeus provirus also contained such a longer gene, which the authors called Vpx (27). Concerning the secondary structure prediction of the Vpr protein, in both D30 and D42 strains, these extra amino acids lead to a β turn, which could influence the function of this protein. While the Vpr protein is quite well conserved within the SIVagm group and also among HIV-1, HIV-2, and SIVsm, little is known about its function. Until now, Vpr has not seemed to be essential for replication in cell culture (15, 39). Studies of SIVmac Vpr mutants have not yet demonstrated that Vpr is absolutely essential for disease progression in vivo (31). Nevertheless, the virion association of Vpr (4, 48) suggests that this protein may play a role in early steps of the virus cycle, possibly in facilitating migration of the preintegration complex to the nucleus of nondividing cells such as macrophages (19).

Another interesting feature of the two isolates described here concerned generation of a stop codon in the Tat gene by splicing (not shown). The putative Tat protein produced from the spliced mRNAs thus only consisted of the first coding exon product (Fig. 3) in contrast to that of the previously described SIVagm. In the recently sequenced sabaeus SIVagm strain (27), the splice junction was mutated in exactly the same manner. Even if splicing did not occur, the D30 and D42 Tat open reading frames would have been precisely interrupted at the 5′ splice site by two in-frame stop codons. In HIV-1, HIV-2, SIVmm, and SIVagm isolates (33), stop codons following the splice donor were also found. Because it is well documented that the HIV-1 first Tat exon is fully functional in Tar-dependent activation of the virus promoter (26, 43), transactivation of D30 and D42 viruses could be expected to occur. However, the basic domain involved in RNA binding (boxed in Fig. 3) is less conserved, lacking the arginine amino acid required for correct binding (3). This is also the case for Tat amino acid sequences of other SIVagm strains (Fig. 3). Because previous studies (44) demonstrated that basic amino acids flanking the arginine residue are determinant in the sequence-specific RNA interaction and in the transactivation activity, we suggest that these Tat products could not optimally transactivate the D30 and D42 promoters relative to the high level of HIV-1 transactivation. In fact, long terminal repeats (LTRs) of SIVagm strains from vervet and grivet monkeys were found to be inefficiently activated by their Tat proteins (42).

The LTR sequences were determined (not shown): 767 and 778 bp in D30 and D42 viruses, respectively, 332 of which overlapped their Nef genes. Regulatory sequences highly conserved in all primate lentiviruses were found in the D30 and D42 3′-terminal repeat. As is the case in SIVagm strains from other sabaeus and tantalus monkeys, only one NF-κB site was noted in the U3 region (27). In contrast, vervet and grivet viruses were shown to contain two NF-κB sites (27, 28). D30 and D42 were found to be the genetically closest AGM viruses (Table 1), with 87% identity in the LTR. One striking difference between these two strains was the presence of four potential Sp-1 binding sites in the D42 LTR, as in the HIV-2, SIVmac, and SIVsm group of viruses and SIVsyk. Only three were found in the D30 and most other SIVagm strains, including sabaeus viruses, as reported for HIV-1 (27, 28). Another structural feature common to the LTR of HIV-2, SIVmac, and SIVsm viruses was the presence of duplicated TAR sequences in both of our viruses. LTRs from all sabaeus isolates studied to date, in contrast to those from other SIVagm strains, contain two copies of the TAR motif (27). These findings underline the variability already described among LTR sequences in the SIVagm group.

```
D30    MAQ--EEARQVWEELQEELQRPLQACDNPCYCKKCCYHCSLCFTRKALGICYYVPRKRRRAK   60
D42    .D.VR.GRP.........R.....N.T....A..F..V..NQ...........T..HS.    62
155    .DKGE.DQDVSHQD.IKQYRK..ET.T.K.F........QF..L..G...T.HAF.T..KKI  62
TYO    .DKGEA.QIVSHQD.S.DY.K...T.K.K.F........Q..LQ.G..VT.HA..T..KKI  62
3      .DKGED.QGAYHQD.I.Q.KA..KR.T.K....C.....Q..LQ.G..VT.HA..I..KKI  62
677    .DKEE.-----PHP.LQD.H....P.T.K..........E..LQ.G..VR.H.S....KTS  57
HIV1   .EPVD----PRL.PWKHPGSQ.KT..T-T.......F..QV...T.....S.GRKKR.Q.RR  57

D30    KIQN---NKVPVHN   71
D42    ....---.......   73
155    ASAD---RIPVPQQ   73
TYO    RSL.---LAPLQ.Q   73
3      APLD---RFPEQKQ   73
677    TQD.---QDPIRQQ   68
HIV1   PP.GSQTHQ.SLSK   71
```

FIG. 3. Alignment of the first Tat coding exons from SIVagm and HIV-1 isolates. Tat nucleotide sequences from sabaeus virus isolates (D30 and D42) were translated and aligned with Tat amino acid sequences of vervet (155, TYO, and 3), grivet (677), and HIV-1bru isolates (33, 34). Dashes represent gaps introduced to optimize the alignment. Dots indicate sequence identity with the D30 protein. The basic domain involved in TAR RNA interaction is boxed in the HIV-1bru amino acid sequence. The arrowhead shows the arginine residue required for correct binding of Tat. The conserved cysteine-rich domain is also boxed.

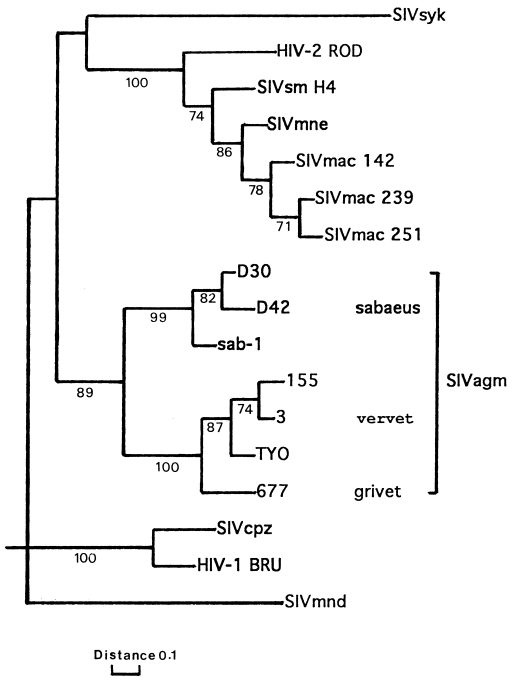

FIG. 4. Phylogenetic tree deduced from analysis of the Nef protein of SIVagm strains from sabaeus monkeys and various other primate lentiviruses. Amino acid sequences were obtained from the Los Alamos sequence database (33, 34) and were aligned with the D30 and D42 ones by the CLUSTAL program (20, 21). The alignment consensus length was 257 amino acids. The tree was built by the neighbor-joining method (41), with the SIV from mandrills (SIVmnd) (47) as the outgroup. Lengths of horizontal branches are proportional to the number of amino acid substitutions. The scale bar indicates 0.1 amino acid replacements per site. Clusters found in more than 70% of 1,000 bootstraps are indicated. The same tree was generated by two other methods, the bootstrap approach of PAUP (10) and parsimony analysis (PROTPARS) (11) with the maximal length alignment (325 residues).

Phylogenetic relationships with other lentiviruses. The percent amino acid identities of the D30 and D42 predicted regulatory proteins indicated that these two viruses are closely related (Table 1). Concerning the first coding exon of Tat (69 to 76%) and in Rev amino acid sequences (78 to 82%), D30, D42, and sab-1 are the most homologous AGM viruses described so far. Other values deduced from these comparisons (88 to 91% in Vpr and 82 to 84% in Nef) are quite similar to those obtained for viruses from vervets. No significant difference between the two sabaeus viruses described here could be detected, except in the LTR, where D42 possesses one more Sp-1-binding site than does D30. Furthermore, common structural differences and the high level of divergence of D30 and D42 isolates from other SIVagm revealed that they belong to a new class of viruses from AGM. This finding was clearly illustrated by the phylogenetic tree analysis performed with Nef amino acid sequences (Fig. 4). SIVagm D30 and D42 clustered together with the other AGM viruses. However, with the recently cloned sabaeus isolate SIVagm sab-1 (27), they are clearly genetically distinct from other AGM viruses and constitute an outgroup. Moreover, this group likely diverged before grivet and vervet viruses separated. These isolates appear to be the most divergent AGM viruses known, as shown by the percentages of homology given in Table 1. This phylogenetic tree is consistent with another based on Tat protein data (first coding exon). In a previous study (1), the tree constructed with Pol sequences also separated the four sabaeus viruses analyzed from vervet and grivet viruses within the SIVagm group. Our results could be explained by the origin of the animals from which D30 and D42 viruses were isolated. They were two sabaeus monkeys caught in the same region of West Africa, whereas SIVagm described previously originated from vervet and grivet monkeys from East Africa. As previously discussed (1, 27, 32), each AGM species might harbor its own distinct SIV, which could have coevolved with its natural host.

For each regulatory protein studied, the present results demonstrated that our two sabaeus viruses were more distantly related to vervet and grivet viruses than the latter viruses were to each other. This is in line with the phylogeny of these AGM species (1). In the opposite direction, Müller et al. (32) showed that viruses from sabaeus monkeys, including the D30 virus described here, are more closely related to vervet viruses. This could be due to data differences concerning part of the Env nucleotide sequences, including SIV isolates from tantalus monkeys. In fact, in the trees established for entire Env and Nef proteins (27), the relative positions of each SIVagm species were identical to the branching orders we obtained by analysis of Nef and Tat amino acid sequences (Fig. 4).

In conclusion, sequence analysis of four regulatory genes from two sabaeus viruses demonstrated that they are closely related and belong to a distinct class of viruses clearly divergent from other AGM virus groups. Moreover, some structural features of the regulatory proteins were specific to these sabaeus viruses.

Regulatory open reading frames are maintained in all viruses of the lentivirus family, which may be a sign of an important function of these proteins during the virus life cycle and disease development. Study of the corresponding genes of putative nonpathogenic virus genomes could provide clues for understanding features involved in the pathogenesis of the human viruses.

Several reasons could contribute to the asymptomatic infection of the sabaeus monkey, attributed both to the virus and to its host. While the functional immune response of AGM to infection with SIV was similar to that of humans to infection with HIV-1 (38), a striking difference was recently shown (9) in the proportion of $CD4^+$ (10%) and $CD8^+$ (80%) lymphocytes. Moreover, AGM $CD8^+$ lymphocytes, as do their human counterparts, secrete a factor with an inhibitory activity on viral replication in CD4 cells (9). This could be of importance in understanding why SIVagm does not induce immunological defects in its natural host. It is thus evident that investigation of the mechanisms of SIVagm asymptomatic infection could provide clues for the comprehension of AIDS induction by human viruses.

More experimental data concerning the mode of action of SIVagm genes are needed to better explain which genetic virus or host properties are involved in the apparent apathogenicity of AGM viruses. The absence of AIDS-like disease development in a heterologous species (*E. patas*) could indicate involvement of genetic control in the asymptomatic infection with SIVagm D30. This situation contrasts with the experimental infection of macaques with SIVsm, the only currently available animal model for HIV infection in humans.

We thank Eric Delaporte for carefully reviewing the manuscript.

Pascale Sarni-Manchado was the recipient of a fellowship from the Fondation pour la Recherche Médicale (FRM). This work was supported by grants from the Institut Français de Recherche Scientifique pour le Développement en Coopération (ORSTOM).

REFERENCES

1. **Allan, J. S., M. Short, M. E. Taylor, S. Su, V. M. Hirsch, P. R. Johnson, G. M. Shaw, and B. H. Hahn.** 1991. Species-specific diversity among simian immunodeficiency viruses from African green monkeys. J. Virol. **65:**2816–2828.

2. **Baier, M., C. Garber, C. Müller, K. Cichutek, and R. Kurth.** 1990. Complete nucleotide sequence of a simian immunodeficiency virus from African green monkeys: a novel type of intragroup divergence. Virology **176:**216–221.
3. **Calnan, B. J., B. Tidor, S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Arginine-mediated RNA recognition: the arginine fork. Science **252:**1167–1171.
4. **Cohen, E. A., G. Dehni, J. G. Sodroski, and W. A. Haseltine.** 1990. Human immunodeficiency virus *vpr* product is a virion-associated regulatory protein. J. Virol. **64:**3097–3099.
5. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, A. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers.** 1985. Isolation of a T-cell tropic HTLV-III-like retrovirus from macaques. Science **228:**1201–1204.
6. **Daniel, M. D., Y. Li, Y. M. Naidu, P. J. Durda, D. K. Schmidt, C. D. Troup, D. P. Silva, J. J. MacKey, H. W. Kestler III, P. K. Sehgal, N. W. King, Y. Ohta, M. Hayami, and R. C. Desrosiers.** 1988. Simian immunodeficiency virus from African green monkeys. J. Virol. **62:**4123–4128.
7. **Diop, O. M.** 1992. Rapp. Inst. Pasteur Dakar **1992:**98–108.
8. **Durand, J.-P.** 1990. Rapp. Inst. Pasteur Dakar **1990:**88–99.
9. **Ennen, J., H. Findeklee, M. T. Dittmar, S. Norley, M. Ernst, and R. Kurth.** 1994. $CD8^+$ lymphocytes of African green monkeys secrete an immunodeficiency virus-suppressing lymphokine. Proc. Natl. Acad. Sci. USA **91:**7207–7211.
10. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
11. **Felsenstein, J.** 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.
12. **Fomsgaard, A., V. M. Hirsch, J. S. Allan, and P. R. Johnson.** 1991. A highly divergent proviral DNA clone of SIV from a distinct species of African green monkey. Virology **182:**397–402.
13. **Fukasawa, M., T. Miura, A. Hasegawa, S. Morikawa, H. Tsujimoto, K. Miki, T. Kitamura, and M. Hayami.** 1988. Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group. Nature (London) **333:**457–461.
14. **Galat-Luong, A., G. Galat, F. Bibollet-Ruche, J.-P. Durand, O. Diop, X. Pourrut, P. Sarni-Manchado, M. Senzani, and G. Pichon.** 1994. Social structure and SIVagm prevalence in two troops of green monkeys, *Cercopithecus aethiops sabaeus*, in Senegal, p. 259–262. *In* J. R. Anderson, J. J. Roeder, B. Thierry, and N. Herrmenschmidt (ed.), Current primatology, behavioural neuroscience, physiology and reproduction. ULP, Strasbourg, France.
15. **Gibbs, J. S., D. A. Regier, and R. C. Desrosiers.** 1994. Construction and *in vitro* properties of SIV_{mac} mutants with deletions in "nonessential" genes. AIDS Res. Hum. Retroviruses **10:**607–616.
16. **Gravell, M., W. T. London, R. S. Hamilton, G. Stone, and M. Monzon.** 1989. Infection of macaque monkeys with simian immunodeficiency virus from African green monkeys: virulence and activation of latent infection. Med. Primatol. **18:**247–254.
17. **Guy, B., M. P. Kieny, Y. Rivière, C. Le Peuch, K. Dott, M. Girard, L. Montagnier, and J.-P. Lecocq.** 1987. HIV F/3′ *orf* encodes a phosphorylated GTP-binding protein resembling an oncogene product. Nature (London) **330:**266–269.
18. **Guy, B., Y. Rivière, K. Dott, A. Regnault, and M. P. Kieny.** 1990. Mutational analysis of the HIV nef protein. Virology **176:**413–425.
19. **Heinzinger, N. K., M. I. Bukrinsky, S. A. Haggerty, A. M. Ragland, V. Kewalramani, M.-A. Lee, H. E. Gendelmann, L. Ratner, M. Stevenson, and M. Emerman.** 1994. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. Proc. Natl. Acad. Sci. USA **91:**7311–7315.
20. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene **73:**237–244.
21. **Higgins, D. G., and P. M. Sharp.** 1989. Fast and sensitive multiple sequence alignments on a microcomputer. Comp. Appl. Biosci. **5:**151–153.
22. **Hirsch, V. M., G. A. Dapolito, S. Goldstein, H. McClure, P. Emau, P. N. Fultz, M. Isahakia, R. Lenroot, G. Myers, and P. R. Johnson.** 1993. A distinct African lentivirus from Sykes' monkeys. J. Virol. **67:**1517–1528.
23. **Hirsch, V. M., G. Dapolito, P. R. Johnson, W. R. Elkins, W. T. London, R. J. Montali, S. Goldstein, and C. Brown.** 1995. Induction of AIDS by simian immunodeficiency virus from an African green monkey: species-specific variation in pathogenicity correlates with the extent of in vivo replication. J. Virol. **69:**955–967.
24. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIV_{sm}) closely related to HIV-2. Nature (London) **339:**389–391.
25. **Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson.** 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature (London) **345:**356–358.
26. **Jeyapaul, J., M. R. Reddy, and S. A. Khan.** 1990. Activity of synthetic tat peptides in human immunodeficiency virus type 1 long terminal repeat-promoted transcription in a cell-free system. Proc. Natl. Acad. Sci. USA **87:**7030–7034.
27. **Jin, M. J., H. Hui, D. L. Robertson, M. C. Müller, F. Barré-Sinoussi, V. M. Hirsch, J. S. Allan, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1994. Mosaic genome structure of simian immunodeficiency virus from West African green monkeys. EMBO J. **13:**2935–2947.
28. **Johnson, P. R., A. Fomsgaard, J. Allan, M. Gravell, W. T. London, R. A. Olmsted, and V. M. Hirsch.** 1990. Simian immunodeficiency viruses from African green monkeys display unusual genetic diversity. J. Virol. **64:**1086–1092.

28a. **Jubier-Maurin, V., et al.** 1993. Regulatory genes of a new SIVagm, abstr. M.P.A. 006, p. 36. VIII International Conference on AIDS in Africa.

29. **Kaminchik, J., N. Bashan, D. Pinchasi, B. Amit, N. Sarver, M. I. Johnston, M. Fischer, Z. Yavin, M. Gorecki, and A. Panet.** 1990. Expression and biochemical characterization of human immunodeficiency virus type 1 *nef* gene product. J. Virol. **64:**3447–3454.
30. **Kraus, G., A. Werner, M. Baier, D. Binniger, F. J. Ferdinand, S. G. Norley, and R. Kurth.** 1989. Isolation of human immunodeficiency virus-related simian immunodeficiency viruses from African green monkeys. Proc. Natl. Acad. Sci. USA **86:**2892–2896.
31. **Lang, S. M., M. Weeger, C. Stahl-Hennig, C. Coulibaly, G. Hunsmann, J. Müller, H. Müller-Hermelink, D. Fuchs, H. Wachter, M. M. Daniel, R. C. Desrosiers, and B. Fleckenstein.** 1993. Importance of *vpr* for infection of rhesus monkeys with simian immunodeficiency virus. J. Virol. **67:**902–912.
32. **Müller, M. C., N. K. Saksena, E. Nerrienet, C. Chappey, V. M. A. Hervé, J.-P. Durand, P. Legal-Campodonico, M.-C. Lang, J.-P. Digoutte, A. J. Georges, M.-C. Georges-Courbot, P. Sonigo, and F. Barré-Sinoussi.** 1993. Simian immunodeficiency viruses from Central and Western Africa: evidence for a new species-specific lentivirus in tantalus monkeys. J. Virol. **67:**1227–1235.
33. **Myers, G., J. A. Berzofsky, A. B. Rabson, R. F. Smith, and F. Wong-Staal.** 1990. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
34. **Myers, G., B. Korber, J. A. Berzofsky, R. F. Smith, and G. N. Pavlakis.** 1991. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
35. **Myers, G., B. Korber, S. Wain-Hobson, K.-T. Jeang, L. E. Henderson, and G. N. Pavlakis.** 1994. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
36. **Myers, G., B. Korber, S. Wain-Hobson, R. F. Smith, and G. N. Pavlakis.** 1993. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
37. **Nebreda, A. R., T. Bryan, F. Segade, P. Wingfield, S. Venkatesan, and E. Santos.** 1991. Biochemical and biological comparison of HIV-1 NEF and *ras* gene products. Virology **183:**151–159.
38. **Norley, S. G., G. Kraus, J. Ennen, J. Bonilla, H. König, and R. Kurth.** 1990. Immunological studies of the basis for the apathogenicity of simian immunodeficiency virus from African green monkeys. Proc. Natl. Acad. Sci. USA **87:**9067–9071.
39. **Ogawa, K., R. Shibata, T. Kiyomasu, I. Higuchi, Y. Kishida, A. Ishimoto, and A. Adachi.** 1989. Mutational analysis of the human immunodeficiency virus *vpr* open reading frame. J. Virol. **63:**4110–4114.
40. **Peeters, M., C. Honoré, T. Huet, L. Bedjabaga, S. Ossari, P. Bussi, R. W. Cooper, and E. Delaporte.** 1989. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. AIDS **3:**625–630.
41. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
42. **Sakuragi, J., M. Fukasawa, R. Shibata, H. Sakai, M. Kawamura, H. Akari, T. Kiyomasu, A. Ishimoto, M. Hayami, and A. Adachi.** 1991. Functional analysis of long terminal repeats derived from four strains of simian immunodeficiency virus SIV_{AGM} in relation to other primate lentiviruses. Virology **185:**455–459.
43. **Seigel, L. J., L. Ratner, S. F. Josephs, D. Derse, M. B. Feinberg, G. R. Reyes, S. J. O'Brien, and F. Wong-Staal.** 1986. Transactivation induced by human T-lymphotropic virus type III (HTLV III) maps to a viral sequence encoding 58 amino acids and lacks tissue specificity. Virology **148:**226–231.
44. **Tao, J., and A. D. Frankel.** 1993. Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins. Proc. Natl. Acad. Sci. USA **90:**1571–1575.
45. **Tristem, M., C. Marshall, A. Karpas, and F. Hill.** 1992. Evolution of the primate lentiviruses: evidence from *vpx* and *vpr*. EMBO J. **11:**3405–3412.
46. **Tristem, M., C. Marshall, A. Karpas, J. Petrik, and F. Hill.** 1990. Origin of *vpx* in lentiviruses. Nature (London) **347:**341–342.
47. **Tsujimoto, H., A. Hasegawa, N. Maki, M. Fukasawa, T. Miura, S. Speidel, R. W. Cooper, E. N. Moriyama, T. Gojobori, and M. Hayami.** 1989. Sequence of a novel simian immunodeficiency virus from a wild-caught African mandrill. Nature (London) **341:**539–541.
48. **Yu, X.-F., M. Matsuda, M. Essex, and T.-H. Lee.** 1990. Open reading frame *vpr* of simian immunodeficiency virus encodes a virion-associated protein. J. Virol. **64:**5688–5693.