

Cocirculation of Two env Molecular Variants, of Possible Recombinant Origin, in Gorilla and Chimpanzee Simian Foamy Virus Strains from Central Africa

Léa Richard, Réjane Rua, Edouard Betsem, Augustin Mouinga-Ondémé, Mirdad Kazanji, Éric Leroy, Richard Njouom, Florence Buseyne, Philippe V.

Afonso, Antoine Gessain

► To cite this version:

Léa Richard, Réjane Rua, Edouard Betsem, Augustin Mouinga-Ondémé, Mirdad Kazanji, et al.. Cocirculation of Two env Molecular Variants, of Possible Recombinant Origin, in Gorilla and Chimpanzee Simian Foamy Virus Strains from Central Africa. Journal of Virology, 2015, 89 (24), pp.12480 - 12491. 10.1128/JVI.01798-15. pasteur-01380758

HAL Id: pasteur-01380758 https://pasteur.hal.science/pasteur-01380758

Submitted on 20 Mar 2017 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

- Co-circulation of two *env* molecular variants, of possible recombinant origin, in gorilla
 and chimpanzee simian foamy virus strains from Central Africa.
- 3

4 Léa Richard^{a,b,c}#, Réjane Rua^{a,b,c*}, Edouard Betsem^{a,b,d**}, Augustin Mouinga-Ondémé^e, Mirdad
5 Kazanji^{e***}, Eric Leroy^f, Richard Njouom^g, Florence Buseyne^{a,b}, Philippe V. Afonso^{a,b}, Antoine
6 Gessain^{a,b}

Unité d'Epidémiologie et Physiopathologie des Virus Oncogènes, Institut Pasteur, Paris,
France ^a; Centre National de la Recherche Scientifique (CNRS), UMR 3569, Paris, France^b;
Université Paris Diderot, Cellule Pasteur, Paris, France^c; Faculty of Medicine and Biomedical
Sciences, University of Yaounde I, Yaounde, Cameroon^d; Unité de rétrovirologie, Centre
International de Recherches Médicales de Franceville, Franceville, Gabon^e; Unité des
Maladies Virales Emergentes, Centre International de Recherches Médicales de Franceville,
Franceville, Gabon^f; Centre Pasteur du Cameroun, Yaounde, Cameroon^g

14 Running Head: Genetic variability of Apes-foamy virus *envelope* gene.

15

16 #Address correspondance to Léa Richard, <u>lea.richard@pasteur.fr</u>

17 *Present address: Rejane Rua, Viral immunology and intravital imaging section, National

18 Institute of Neurological Disorders and Stroke, National Institute of Health, Bethesda,

19 Maryland, United States.

20 **Present address: Edouard Betsem, Laboratoire Mixte International de Vaccinologie,

- 21 Agence de Médecine Préventive, Bobo-Dioulasso, Burkina Faso.
- 22 ***Present address: Mirdad Kazanji, Institut Pasteur de la Guyane, Cayenne, French Guiana.
- 23 Word count for the abstract: 246 + 104. Word count for the text: 4546.

25 Simian foamy virus (SFV) is a ubiquitous retrovirus in non-human primates (NHPs) that can 26 be transmitted to humans, mostly through severe bites. In the past few years, our 27 laboratory has identified more than 50 hunters from Central Africa infected with zoonotic 28 SFVs. Analysis of the complete sequences of five SFVs obtained from these individuals 29 revealed that *env* was the most variable gene. Furthermore, recombinant SFV strains, some 30 of which involve sequences in the *env* gene, have been recently identified. Here, we 31 investigated the variability of the env gene of zoonotic SFV strains, and searched for 32 possible recombinants. We sequenced the complete env gene or its surface glycoprotein 33 region (SU) from DNA amplified from the blood of: 1) a series of 40 individuals from 34 Cameroon or Gabon infected with a gorilla or chimpanzee FV strain; and 2) one gorilla and 35 three infected chimpanzees living in the same areas as these hunters. Phylogenetic analyses 36 revealed the existence of two env variants among both the gorilla and chimpanzee FV 37 strains that also were present in zoonotic and NHP strains. These variants differ greatly 38 (more than 30% variability) in a 753 bp long region located in the receptor-binding domain 39 of the SU whereas the rest of the gene is very conserved. Although the organization of Env 40 protein sequences is similar, the potential glycosylation patterns differ between variants. 41 The analysis of recombination suggests that the variants emerged through recombination 42 events between different strains, although all parental strains could not be identified.

43

44 Importance

45 SFV infection in humans is a great example of a zoonotic retroviral infection that has not 46 spread among human populations, in contrast to human immunodeficiency viruses (HIV)

47	and human T-lymphotropic viruses (HTLV). Recombination was a major mechanism leading
48	to the emergence of HIV. Here, we show that two SFV molecular <i>envelope</i> variants circulate
49	among ape populations in Central Africa and that both can be transmitted to humans. These
50	variants differ greatly in the SU region that corresponds to the part of the Env protein in
51	contact with the environment. These variants may have emerged through recombination
52	events between SFV strains infecting different NHP species.

54 The emergence of zoonotic viruses is a multi-step process involving the transmission 55 of viruses from domestic or wild animals to humans and their subsequent spread among 56 human populations. Such viral infections are frequent (1). In particular, several viruses 57 originating from non-human primates (NHPs) have had a major impact on human health. 58 Among them, the retroviruses, simian immunodeficiency virus (SIV) and simian T-59 lymphotropic virus (STLV), have crossed the species barrier from NHPs to humans, leading 60 to the emergence of human immunodeficiency virus (HIV) and human T-lymphotropic virus 61 (HTLV), respectively, in human populations (2).

62 Another retrovirus that can be transmitted from NHPs to humans is the simian foamy virus (SFV). SFVs are complex retroviruses from the Spumaretrovirinae subfamily that are 63 64 ubiquitous in both Old and New World NHPs, with a seroprevalence of up to 75-100% in 65 adult NHPs. Phylogenetic studies suggest that SFV have evolved by co-speciation with Old 66 World primates over more than 85 million years (3, 4). Consequently, host-specific groups 67 have emerged. Recombination can also be involved in the genetic diversity of SFV. Indeed, 68 NHPs can become co-infected with strains either from the same group (5-7), or from 69 different groups, as shown by some chimpanzees, which were found to be co-infected with 70 SFVcpz strains and with SFV from colobus monkeys, or Cercopithecus monkeys (5, 8). Co-71 infection may create recombinant strains, as suggested within the *env*, *gag* and *pol* genes (5, 72 7, 9). Of note, both co-speciation and host switching may have contributed to the 73 evolutionary history of SFVs infecting New World prosimians (10).

Humans are not considered to be natural SFV hosts; however, more than 100 cases of
human SFV infection have been reported mostly among individuals exposed to NHPs either

Z

76 in a professional context (e.g. veterinarians or zookeepers) (11-15) or natural settings (e.g. 77 hunters in Africa, and monkey temple workers or visitors, pet owners, or people living 78 around free-ranging macaques in Asia) (16-22). The clinical significance of zoonotic SFV 79 infection remains unknown. This may be due to the very small number of persons yet 80 studied with a limited follow-up (23) and also to a possible recruitment bias (among healthy 81 individuals). SFV is mainly transmitted through biting (22). Indeed, in African green 82 monkeys and macagues, the oral mucosa is a major site for viral replication (24-26) and SFV 83 viral RNA accumulates at high concentrations in saliva (25, 27). In humans, SFV infection is 84 persistent. SFV DNA is detectable in peripheral blood and saliva cells. However, cell-85 associated viral RNA has not been detected (28), (21) and secondary human-to-human 86 transmission (22, 29) has not been reported. Hence, human infection with zoonotic SFVs 87 represents a unique natural model to study the role of viral and immunological factors in the 88 restriction of viral transmission.

89 In the past few years, we have undertaken a large epidemiological and molecular 90 study involving hunters of monkeys and apes, living in Central Africa, to obtain insight into 91 the natural history of SFV emergence in humans. We identified a large series (more than 50) 92 of hunters that had been infected with SFVs, mostly following severe bites from NHPs (17, 93 22). We previously showed that the virus mostly localizes to CD8+, CD4+ T, and B 94 lymphocytes in the blood of infected humans (30). Our study also revealed natural 95 polymorphisms in gag and bet between SFV strains originating from different chimpanzee 96 subspecies and polymorphisms in U3 and tas at the inter-individuals level (31). Based on the 97 complete sequence of five replication-competent strains, we found no evidence of viral 98 adaptation among SFV strains isolated in humans (31).

99 Our previous study revealed some genetic diversity in the *env* gene of zoonotic strains 100 of SFV from Apes. The *env* gene is involved in different viral cycle steps such as receptor 101 binding (32, 33), fusion (34-36), budding (37-39) and site of particle formation (40, 41). The 102 broad tropism of SFV suggests an ubiquitous cellular receptor (42, 43). While all SFVs seem 103 to use the same cellular receptor (44-46), only heparin sulfate has been identified as an 104 attachment factor so far (47, 48).

Here, we aimed to investigate the genetic variability of *env* among SFVs infecting humans or NHPs living in Cameroon and Gabon. We found that viral strains from both gorilla and chimpanzee FVs segregate into two *env* variants. These variants co-circulate among humans and NHPs, and may have arisen by recombination.

109

110 MATERIALS AND METHODS

111 **Population**

112 Human population: In three previous studies, we identified 48 individuals from Cameroon 113 who were infected either with a SFV from a Gorilla gorilla (Ggo) or from a Pan 114 troglodytes troglodytes (Ptr) as defined by the analysis of a 465 bp-long pol-integrase 115 fragment (17, 22, 49), In a few cases, SFV was detected by analyzing a smaller LTR fragment. 116 These individuals live in villages or settlements located in the rain forest of South and East 117 Cameroon. They belong to different Bantu tribes or the Baka and Bakola Pygmy tribes. 118 Fourteen Ggo- or Ptr-FV infected individuals from Gabon, all of whom were Bantus, were 119 also included in this study (20).

120 This study was approved by the research division of the Ministry of Public Health and the 121 National Committee of Ethics in Cameroon, the Ministry of Health and the Ethics Committee

Z

lournal of Virology

of the CIRMF in Gabon, and the Comité de Protection des Personnes and the Commission
Nationale de l'Informatique et des Libertés in France. All individuals provided written
informed consent.

<u>NHP population</u>: One *Ggo* from Cameroon and three *Ptr* (one from Cameroon and two from Gabon) were included in this study. NHPs from Cameroon were born and caught in the wild in the Southern rain forest area (50, 51). NHPs from Gabon were all born in the wild (20). Samples were collected in accordance with the rules of animal care committees.

129 Sampling and DNA preparation

Blood samples were collected from all individuals and NHPs in EDTA tubes. The buffy coat
(BC) was obtained after centrifugation and genomic DNA was extracted using the Qiamp
DNA Blood minikit (Qiagen, Courtaboeuf, France). DNA samples were quantified using a
Nanodrop instrument (ThermoScientific).

134 PCR amplification of env from genomic DNA

135 Regarding *Ggo*-FVs, three different overlapping *env* fragments (720, 1102 and 1490 bp-long) 136 were amplified from BC-DNA using specific primers (Table 1). Nested PCR was performed as 137 follows: 500 ng of DNA was mixed in the enzyme buffer with the external primers (0.25 μ M 138 of each), MgCl₂ (3.5 mM), dNTP (200 μ M of each) and 0.5 μ L of HotStarTaq polymerase 139 (Qiagen) in a final volume of 50 μ L. The external PCR consisted of a 15' long denaturation 140 step at 95°C, followed by 40 amplification cycles (45" at 95°C, 45" at 57°C and 1' per kb at 141 72°C) and a 7' long extension step at 72°C. The product (5 μ L) was then used as template for 142 a second internal PCR in the same conditions, but using the internal primers (table 1).

143 Regarding *Ptr*-FVs, two different overlapping *env* fragments (1645 and 1802 bp) were

144 amplified from BC-DNA using specific primers (Table 1). The PCR parameters were the same

as those used for the *Ggo*-FV strains, except for the annealing temperature of the PCR of the
second fragment (60°C instead of 57°C).

The *env* gene of the prototypic *Pan troglodytes verus* SFVpvr.SFV7 strain (52) was amplified
from SFVpvr.SFV7-infected BHK-21 cells (a gift from A. Rethwilm) using the CPZENVF1 and
CPZENVR2b primers (Table 1) with an annealing temperature of 57°C and an extension time
of 3'30".

PCR products were directly sequenced by MWG Operon (Ebersberg, Germany). Both sense and antisense sequences were obtained for each fragment and were found to be identical. To obtain complete *env* sequences, the different *env* fragments were concatenated (the overlapping regions were identical in all samples).

155 **Phylogenetic analyses**

156 Multiple sequence alignments on the previously known (Table 2) and the newly generated 157 sequences were performed using the DAMBE program (v4.2.13 158 [http://www.dambe.bio.uottawa.ca]). Modeltest v3.6 was used to select the most 159 appropriate nucleotide substitution model, based on the Akaike information criterion (AIC). 160 GTR was found to be the best-fitting model. Phylogenetic trees were constructed using the 161 Neighbor Joining method and bootstrap values were calculated on 1,000 replicates. 162 Phylogenetic tree topologies were confirmed using the maximum likelihood method with the 163 PAUP program (v4.0 [http://www.paup.csit.fsu.edu]). Percentage identity was calculated by 164 CLC software (CLC DNA Workbench 6 [http://www.clcbio.com]) on the alignment obtained 165 with DAMBE.

By definition, a clade is a monophyletic group supported by a strong bootstrap value. Herein, we used "group" as a host-specific clade, and "subgroup" as a clade defined on the central region of *env*.

169

170 **Recombinant analyses**

171 To search for potential recombination events, similarity plot and bootscanning analyses 172 were performed with Simplot software (v3.5.1 [http://www.simplot.com]). These analyses 173 were performed with default parameters, except for the bootscan repetitions (set to 1,000) 174 and the evolution model (Kimura 2-parameter). The similarity plot depicts a similarity score 175 between a group of interest (query) and other groups (defined based on the phylogenetic 176 tree) over a 200 bp-long region and the overall *env* gene was analysed by 20 bp-long steps. 177 The bootscan analysis reflects the phylogenetic relationship (bootstrap value) between a 178 group of interest (query) and the other groups (window 200 bp, step 20 bp).

179 The Recombination Detection Program (RDP4 (53)) was also used to investigate putative 180 recombination. Unlike the Simplot derived studies, the RDP method looks for recombination 181 by analyzing every possible sequence triplet (instead of considering groups). A window is 182 moved along the genome and a percentage of identity between each of the three possible 183 pairs is calculated at each position. A *p*-value is calculated to determine the likelihood that 184 the potential recombination event is due to chance.

185 **Protein analysis**

Percentage identity was calculated by CLC software (CLC DNA Workbench 6) on the
alignment obtained with DAMBE after the translation of the sequences into amino acids.
Putative glycosylation sites were identified by the NetGlyc 1.0 program and consisted of

189 NXS/T sites, where X denotes any amino acid except proline [<u>http://www.cbs.dtu.dk/</u>
190 <u>services/NetNGlvc</u>].

191 Nucleotide sequence accession number

All complete and partials *env* sequences generated in this work are published in GenBank.
Complete sequences are published through accession numbers KT211246 to KT211269, the
latter corresponding to SFVpvr.SFV7 sequence; partial sequences through accession
numbers KT211270 to 211290.

196

197 **RESULTS**

198 We previously identified 48 SFV-infected individuals in Cameroon (17, 22, 49) and 14 199 SFV-infected individuals in Gabon (20). The strains belonged to the Gorilla gorilla gorilla 200 (Ggo)- or Pan troglodytes troglodytes (Ptr)-FV groups according to their pol and/or LTR 201 sequences. In addition, we previously obtained complete SFV sequences (Bad468, Bak74, 202 AG15 and Bad327) from four Cameroonian individuals (31). Buffy coat or DNA samples were 203 available for 51 of the remaining 58 individuals. From these samples, 40 new env sequences 204 were generated by PCR and direct sequencing (19 env sequences were complete and 21 205 partial). Low viral load may explain the absence of amplification in the 11 remaining 206 samples. Thus, overall, our study population yielded 44 env sequences from zoonotic SFVs 207 (33 Ggo-FV and 4 Ptr-FV from Cameroon, five Ggo-FV and two Ptr-FV from Gabon) (Table 3, 208 Figure 1).

In addition, we also obtained complete *env* sequences from four wild-born NHPs
living in the same region as the infected individuals (one gorilla, and three chimpanzees) and

211 we sequenced the env gene of a prototypic SFVpvr.SFV7 Pan troglodytes verus strain (Table

212 3).

All newly obtained sequences contained an Open Reading Frame of 2958 to 2970 bpand were unique.

215

216 Complete env sequences define host-specific SFV groups

217 We analyzed phylogenetically the complete *env* sequences using the neighbor joining 218 (Figure 2) and the maximum likelihood methods (data not shown). The newly generated 219 complete env sequences were included together with the available SFV env sequences in 220 Genbank (Table 2). The topologies of both phylogenetic trees were comparable. Host-221 specific groups (Gorilla, Chimpanzee, Cercopithecus, Macaque) were identifiable and 222 supported by strong bootstraps (>90). Moreover, among the strains from the chimpanzee-FV 223 group, we identified three different clades corresponding to different subspecies: Pan 224 troglodytes troglodytes, P. t. schweinfurthii and P. t. verus.

Interestingly, the gorilla-specific FV group was composed of two monophyletic subgroups supported by strong bootstraps. These subgroups did not correspond to geographic segregation. Indeed, each subgroup contained sequences from both Cameroon and Gabon (underlined strains on the tree; Figure 2), as well as both zoonotic and NHP strains (SFVggo or SFVggo.Cam7). We named the subgroup containing the previously described Bad468 *env* FV sequence, SFV-GorI, and the one containing the Bak74 *env* FV and the prototypic SFVggo sequences, SFV-GorII.

Similarly, two phylogenetic subgroups were identifiable within the *Ptr*-FV group. We
named the one containing the previously described Bad327 and AG15 *env* sequences, SFVCpzII, and the other one, SFV-CpzI.

235

236 Definition of a central variant region within the surface env sequence

We performed sequence alignments to understand better the origins of the two
subgroups present in the *Ggo*-FV and *Ptr*-FV groups.

239 Concerning the Ggo-FV strains, SFV-GorI env sequences were 2970 bp long, whereas 240 SFV-GorII sequences were 2964 bp long. Complete sequences within a subgroup were 96.4 241 to 99.8% identical, whereas those belonging to different subgroups were only 88.2 to 90.4% 242 identical (Figure 3.A). SFV-GorI and SFV-GorII sequences differed the most in a central 243 region beginning at position 769 bp (red line) and ending at position 1473 bp (green line; 244 the positions were defined on the env PFV strain) (Figure 3.B). In the Ptr-FV strains, SFV-245 CpzI sequences were 2967 bp long whereas SFV-CpzII sequences were 2958 bp long. These 246 two Ptr-FV subgroups differed the most between positions 721 (red line) and 1440 (green 247 line) (Figure 3.C). Interestingly, this corresponds closely to the same variable region defined 248 for Ggo-FV strains.

Thus, we defined two regions in the *env* sequence (Figure 3.D): a conserved region, which comprises nucleotides 1 to 720 and 1474 to 2967 (positions defined in the PFV strain), and a variant region, from 721 to 1473. We confirmed a high sequence similarity (96.4 to 99.3% identity) in the conserved region between the SFV-GorI and SFV-GorII subgroups, and a very important divergence (62.2-64.7% identity) in the variant region (Figure 3.D). Similarity of the variant region was very high between strains from the same subgroup (94.8-100% identity). This was also true for the two *Ptr*-FV subgroups (SFV-CpzI
and SFV-CpzII). Strikingly, in the variant region, SFV-GorI and SFV-CpzI showed an identity
score of 71.7-73% which is higher than that between SFV-GorI and SFV-GorII (62.2-64.7%)
(Figure 3.D).

259

260 Two variants are present within SFV-groups

261 Given that a conserved and a variant region were defined, we performed phylogenetic262 analyses on each segment separately.

In the conserved region (Figure 4.A), sequences segregated following host-species and subspecies. The four major monophyletic and highly supported groups corresponded to SFV strains of gorilla, chimpanzee, macaque or Cercopithecus origin. Moreover, *Ggo*-FV strains were now subdivided into groups reflecting geographic origin (two Cameroonian clades and one Gabonese clade, underlined in Figure 4.A). This distribution resembles that of the *pol*-based phylogenetic tree established in previous studies on the same individuals (20, 22).

270 The phylogenetic tree based on the variant region (Figure 4.B) was quite different. 271 The Ggo- and Ptr-FV segregated according to the subgroups identified previously (SFV-Gorl, 272 SFV-GorII, SFV-CpzI and SFV-CpzII,). In this phylogenetic tree, we could define two clades for 273 African Ape SFVs (clade 1 and clade 2), each comprising both Ggo-FV and Ptr-FV strains. Of 274 note, P.t.verus-FV sequences (SFV7 and SFVcpz) were also split: the SFV7 env sequence was 275 closer to the SFV-CpzII sequences, whereas the SFVcpz env was closer to the 276 P.t.schweinfurthii PFV and the SFV-CpzI strains. Finally, macaque- and cercopithecus-FV 277 isolates were also divided into two distinct subgroups.

278

279 partial *env* sequences obtained from other individuals infected by a *Ggo*-FV strain (Table 3) 280 and 21 previously published African Green Monkey (AGM, cercopithecus group) env SU 281 sequences (Table 2)(54). Phylogenetic analysis of the variant region (Figure 4.C) confirmed 282 the robustness of the two major African Ape clades (clade 1 and clade 2) and of the different 283 subgroups SFV-GorI, SFV-GorII, SFV-CpzI and SFV-CpzII. Interestingly, cercopithecus strains 284 separated into two major subgroups: SFV-CercI and SFV-CercII. Macague-FV strains were 285 also split into two distinct subgroups (SFV-MacI and SFV-MacII). Of note, the unrooted 286 phylogenetic tree was star-shaped (Figure 4.C), suggesting that, although SFV-GorI and SFV-287 CpzI are the closest groups, the sequences are still genetically distant and do not share a 288 recent common ancestor.

To confirm these findings, in the analysis of the variant region, we included the 21

289 Thus, the strains from four simian species (gorilla, chimpanzee, macaque, 290 Cercopithecus) systematically segregate into two subgroups that differ regarding the env 291 central region.

292

293 Genomic and functional organization of the envelope gene

294 All protein sequences generated from the new env sequences showed typical Env 295 organization with the leader peptide (LP), the surface glycoprotein (SU) and transmembrane 296 glycoprotein (TM) regions separated by RXXR cleavage signals (data not shown). Other 297 important domains such as the fusion peptide (FP) or the membrane-spanning domain 298 (MSD) were also present (figure 5.A, data not shown). Every cysteine residue (essential for 299 proper folding) described for PFV Env (32) was also conserved.

lournal of Virology

300 The variant region was located within the SU domain, more specifically within the 301 Receptor Binding Domain (RBD) identified on PFV (32). Given that glycosylation is 302 important for RBD function, we examined the conservation of the putative glycosylation 303 sites. Although amino acid sequences diverge in the central region between variants (figure 304 5.B), 13 of the potential glycosylation sites (previously defined on PFV) were conserved 305 among Ggo- and Ptr-FV strains (55)(figure 5.C). Ggo-FV subgroups differed at position of 306 N11 and chimpanzee-FV subgroups differed at positions of N11, N5 and N8. Thus, the 307 different subgroups might have different glycosylation patterns.

308

309 Could variants be generated by recombination?

We performed bootscan analysis on the *env* sequences grouped by host specific subgroups to search for potential recombinants among *Ggo-* and *Ptr-*FV strains (Figure 6). Consistent with the phylogenetic trees, the conserved region of SFV-GorII *env* sequences was closely related to SFV-GorI (more than 99% of permuted trees grouped them together), whereas the variable region was closer to sequences from SFV-CpzII (Figure 6.A). Similarly, the conserved region of SFV-CpzI strains was closer to that of SFV-CpzII strains whereas the variable region of SFV-CpzI strains was closer to that of SFV-GorI strains (Figure 6.C).

Such a pattern suggests that strains from the SFV-GorII subgroup originated from a recombination event between SFV-GorI and SFV-CpzII strains. However, similarity plot analysis argues against this conclusion. Indeed, the similarity score between SFV-GorII and SFV-CpzII strains is low throughout the whole gene, and is lowest in the variant region (between 52 and 62% similarity)(Figure 6.B). Thus, if the strains from the SFV-GorII

lournal of Virology

322 subgroup did indeed appear by recombination, one of the parental strains remains 323 unknown.

324 To investigate further this hypothesis, we performed a Recombinant Detection 325 Program (RDP) analysis to detect recombinant sequences. This analysis strongly suggests 326 that the strains from the SFV-GorI and SFV-GorII subgroups diverged upon recombination 327 (p-value=4.745.10⁻¹²). Similarly, the SFV-CpzI and SFV-CpzII subgroups probably 328 recombined (*p*-value=2.074.10⁻¹²). However, in both cases, the parental strain from which 329 the central region was obtained could not be identified.

330 We conclude that the two variants observed among Ggo- and Ptr-FV strains probably 331 arose by a recombination event, and that one of the parental strain is still unknown.

332

333 DISCUSSION

334 Here, we studied the variability of the envelope gene of SFVs infecting Apes (gorillas 335 and chimpanzees) and humans living in Cameroon or Gabon. We demonstrate the co-336 circulation of two SFV env molecular variants in both the gorilla- and chimpanzee-FV groups. 337 In a given group, the complete nucleotide sequence of *env* differs by more than 10% between 338 the variants. These differences are mostly located in a central region of the *envelope* that we 339 have called the "variant" region, in contrast with the "conserved" region that corresponds to 340 the rest of the *env* gene. The genetic diversity of *env* may have arisen from recombination.

341 These findings raise several important issues and questions:

342

343 1) Genetic variability in the conserved region reflects host-species origin and 344 geographical clustering.

lournal of Virology

345 In the conserved region, the nucleotide variability among *Ggo*-FV strains was low, 346 ranging from 0.2% to 3.6%. This genetic stability is consistent with the low *in vivo* variation rate estimated to be around 1.7.10⁻⁸ substitutions per site per year (3). The genetic 347 348 variability of the conserved regions of the different *Ggo*-FV reflects geographic location. 349 Indeed, we defined two clusters of Cameroonian sequences and a cluster from strains 350 isolated in Gabon by phylogenetic analysis. Interestingly, these clusters were observed 351 previously when the same strains were divided according to their *pol* sequence (20, 22). 352 Similarly, geographic segregation was also observed in SFV strains in macaque populations 353 in Bangladesh (7), and in mandrill populations in Gabon (56). Indeed, geographic 354 segregation and consequent isolation lead to the accumulation of distinct mutations and 355 speciation.

By contrast, Cameroonian and Gabonese *Ptr*-FV sequences did not segregate together, consistent with the findings of studies on the *pol-integrase* gene (20, 22). This may suggest that, unlike Gorillas, *Ptr* populations are not genetically or virologically isolated, despite being geographically distant. However such results are based on few sequences.

In the central variant region, we found two subgroups for each FV host-specific group with more than 35% nucleotide variability. This variability did not reflect geographic location. Strains from every *Ggo-* and *Ptr-*FV subgroup were present in both humans and NHPs.

364

365 2) What is the mechanism leading to the emergence of *env* diversity among the SFV
 366 variants?

We first examined whether recombination, which is a frequent evolutionary event in retroviruses, could explain the diversity of *env* sequences. (57-60). Indeed, recombination is also frequent among SFVs. *In vitro* studies show that it can occur between PFV-based vectors and reveal that the probability of a template-switching event within a 1 kb-long region is 27% (61).

Our analyses suggested that a *Ggo*-FV variant originated from recombination between *Ggo*-FV strains and unknown FV strains (the closest known strain being the chimpanzee strains). Similarly one of the *Ptr*-FV variants may have arisen from recombination between chimpanzee- and gorilla-like FV strains. Until the parental strains are found, the recombination origin of these sequences might be debated. However this hypothesis is the most realistic.

378 Such a scenario implies that NHPs can be co-infected with SFV strains of different host-379 species, and that these strains have recombined. Co-infection has been documented in 380 chimpanzees (5, 6) and rhesus macaques (7). However, in these cases, co-infection occurred 381 with strains from the same host-specific group. Interestingly, in these populations, 382 recombination in FVs has been observed in the *pol* or *gag* gene (5, 7). Furthermore, co-383 infection with strains belonging to different NHP-related groups has been documented in 384 wild chimpanzees. These animals were infected with both Colobus monkey- and 385 chimpanzee-related FVs (5, 8) but no recombinant strain was detected. Recently, 386 recombination between macaque- and Cercopithecus-FV strains has been reported (9, 62). 387 However, macaques and Cercopithecus live in different continents, suggesting that the 388 recombination event occurred in captivity. Indeed, animals from various species and/or 389 geographical areas are frequently mixed in primate centers. The variants we report would

390 be the first recombinants detected between SFVs of different primate hosts found in a391 natural setting.

Interestingly, the putative recombination sites were quite similar between the gorillaand the chimpanzee-FV variants. This site also corresponds closely to the same region described in macaque (9) and African Green Monkey FV sequences (54). A similar variant region, with similar recombination points, has been described in feline FVs (63). Together, this suggests that a recombination hot-spot may exist in the *env* region of foamy viruses.

397 Our data suggest the different *env*-based subgroups arose by recombination, but 398 many questions remain unsolved. First, the identity of one of the parental strains is still 399 undetermined. This second parental strain, which purportedly introduced the central 400 variable region into the env gene, may be either an existing but yet undescribed FV, or a 401 strain that has disappeared. Second, the date of the recombination events has not been 402 addressed. If the recombination event was ancient, genetic drift would have segregated SFV-403 GorI and SFV-GorII even in the conserved region. Given that the conserved region is very 404 similar between SFV-GorI and SFV-GorII strains and that they do not form separate 405 monophyletic groups in the conserved region, the recombination event probably occurred 406 fairly recently.

407

408 3) Why are only two *env* genetic variants present?

Strikingly, we observed two *env* molecular variants for each NHP species studied (i.e.
gorilla, chimpanzee, macaque and Cercopitecus). If the defined sites were hot spots of
recombination, we would expect many more variants to be generated. Strong functional

19

2

412 constraints may be an effective source of purifying selection that would only allow for the413 emergence of two subgroups per host.

414 The variant central region is located in the putative receptor-binding domain of the 415 SU domain, as identified on PFV (32). SFV uses heparin sulfate to attach to target cells (47, 416 48), but the receptor for SFV is still unknown (64). Interference studies showed that one 417 cellular receptor is used by all SFV (44, 45, 46). Binding studies of recombinant envelope 418 protein suggest that SFVcpz interacts with two receptors, one low affinity and one high 419 affinity (65). Although most glycosylation sites are conserved, we found that the pattern of 420 glycosylation may differ between the Ggo- and Ptr-FV subgroups. Consequently, it is possible 421 that the two FV variants correspond to viruses that use different receptor complexes, as this 422 is the case for Murine Leukemia Virus (66). Alternatively, the SFV receptor may be expressed 423 in different conformations at the cell surface, allowing two modes of engagement by the two 424 Env variants. Indeed, such conformational heterogeneity was described for the CCR5 425 molecule and its interaction with the envelope of HIV and chemokines (67).

426 Finally, env molecular variants will probably induce distinct immune responses. The Env SU is indeed the target of neutralizing antibodies (65). Two env genotypes 427 428 corresponding to two serotypes have been described among feline FVs (63, 68). The two 429 genotypes differ in a central region corresponding nearly to the one described in our study. 430 SFV strains from macaques and chimpanzees also segregate into two serotypes as defined by 431 neutralization assay with immune plasma or sera, serotype 1 and 2 for macaques (69), 432 serotype 6 and 7 for chimpanzees (52, 70, 71). We show here two cases in which strains 433 belonging to different serotypes belong to distinct genetic subgroups. Indeed SFV-mcy1 and 434 SFV-mcy2 are from serotype 1 and 2 respectively while belonging to SFV-MacI and SFV-

435 MacII respectively (fig. 4C); PFV/SFVcpz and SFVpvr.SFV7 are from serotype 6 and 7 436 respectively while segregating with SFV-CpzI or SFV-CpzII respectively (fig. 4C). Preliminary 437 results on neutralizing antibodies present in the plasma of infected persons support the 438 correspondence between serotype and genotype for SFV from the chimpanzee group (72). 439 Studies are currently ongoing to characterize immune responses induced by SFV strains

441

440

442 **Funding Information**

from the gorilla group.

443 L. Richard was personally supported by the Bourse de l'Ecole Normale Supérieure, Faculté 444 Paris Diderot. E. Betsem was supported by the association Virus Cancer Prevention and by 445 the Institut National pour le Cancer. This work was supported by the Institut Pasteur in 446 Paris, France, by Programme Transversal de Recherche 437 from the Institut Pasteur, and by 447 the French government program Investissement d'Avenir (grant ANR-10-LABX-62-IBEID). 448 The funders had no role in study design, data collection and interpretation, or the decision to 449 submit the work for publication.

450

451 Acknowledgments

452 We thank the Institut de Recherche pour le Développement (IRD) and the Centre Pasteur du 453 Cameroun for collaborating with us concerning the field work in Cameroon. We thank Axel 454 Rethwilm for providing SFV7-infected BHK cells and Arifa Khan for helpful discussions about 455 foamy virus taxonomy.

456

457

Bibliography 458

459	1.	Wolfe ND, Dunavan CP, Diamond J. 2007. Origins of major human infectious diseases. Nature
460		447: 279-283.
461	2.	Peeters M, D'Arc M, Delaporte E. 2014. Origin and diversity of human retroviruses. AIDS reviews
462		16: 23-34.
463	3.	Switzer WM, Salemi M, Shanmugam V, Gao F, Cong ME, Kuiken C, Bhullar V, Beer BE, Vallet D,
464		Gautier-Hion A, Tooze Z, Villinger F, Holmes EC, Heneine W. 2005. Ancient co-speciation of
465		simian foamy viruses and primates. Nature 434: 376-380.
466	4.	Rethwilm A, Bodem J. 2013. Evolution of foamy viruses: the most ancient of all retroviruses.
467		Viruses 5: 2349-2374.
468	5.	Liu W, Worobey M, Li Y, Keele BF, Bibollet-Ruche F, Guo Y, Goepfert PA, Santiago ML, Ndjango
469		JB, Neel C, Clifford SL, Sanz C, Kamenya S, Wilson ML, Pusey AE, Gross-Camp N, Boesch C,
470		Smith V, Zamma K, Huffman MA, Mitani JC, Watts DP, Peeters M, Shaw GM, Switzer WM,
471		Sharp PM, Hahn BH. 2008. Molecular ecology and natural history of simian foamy virus infection
472		in wild-living chimpanzees. PLoS pathogens 4 :e1000097.
473	6.	Blasse A, Calvignac-Spencer S, Merkel K, Goffe AS, Boesch C, Mundry R, Leendertz FH. 2013.
474		Mother-offspring transmission and age-dependent accumulation of simian foamy virus in wild
475		chimpanzees. Journal of virology 87: 5193-5204.
476	7.	Feeroz MM, Soliven K, Small CT, Engel GA, Pacheco MA, Yee JL, Wang X, Hasan MK, Oh G,
477		Levine KL, Alam SR, Craig KL, Jackson DL, Lee E, Barry PA, Lerche NW, Escalante AA, Matsen IV
478		FA, Linial ML, Jones-Engel L. 2013. Population dynamics of rhesus macaques and associated
479		foamy virus in Bangladesh. Emerging Microbes and Infections:e29.
480	8.	Leendertz FH, Zirkel F, Couacy-Hymann E, Ellerbrok H, Morozov VA, Pauli G, Hedemann C,
481		Formenty P, Jensen SA, Boesch C, Junglen S. 2008. Interspecies transmission of simian foamy
482		virus in a natural predator-prey system. Journal of virology 82:7741-7744.
483	9.	Galvin TA, Ahmed IA, Shahabuddin M, Bryan T, Khan AS. 2013. Identification of recombination
484		in the envelope gene of simian foamy virus serotype 2 isolated from Macaca cyclopis. Journal of
485		virology 87: 8792-8797.
486	10.	Katzourakis A, Aiewsakun P, Jia H, Wolfe ND, LeBreton M, Yoder AD, Switzer WM. 2014.
487		Discovery of prosimian and afrotherian foamy viruses and potential cross species transmissions
488		amidst stable and ancient mammalian co-evolution. Retrovirology 11 :61.
489	11.	Schweizer M, Turek R, Hahn H, Schliephake A, Netzer KO, Eder G, Reinhardt M, Rethwilm A,
490		Neumann-Haefelin D. 1995. Markers of foamy virus infections in monkeys, apes, and
491		accidentally infected humans: appropriate testing fails to confirm suspected foamy virus
492		prevalence in humans. AIDS research and human retroviruses 11: 161-170.
493	12.	Heneine W, Switzer WM, Sandstrom P, Brown J, Vedapuri S, Schable CA, Khan AS, Lerche NW,
494		Schweizer M, Neumann-Haefelin D, Chapman LE, Folks TM. 1998. Identification of a human
495		population infected with simian foamy viruses. Nature medicine 4: 403-407.
496	13.	Brooks JI, Rud EW, Pilon RG, Smith JM, Switzer WM, Sandstrom PA. 2002. Cross-species
497		retroviral transmission from macaques to human beings. Lancet 360: 387-388.
498	14.	Sandstrom PA, Phan KO, Switzer WM, Fredeking T, Chapman L, Heneine W, Folks TM. 2000.
499		Simian foamy virus infection among zoo keepers. Lancet 355: 551-552.
500	15.	Switzer WM, Bhullar V, Shanmugam V, Cong ME, Parekh B, Lerche NW, Yee JL, Ely JJ, Boneva R,
501		Chapman LE, Folks TM, Heneine W. 2004. Frequent simian foamy virus infection in persons
502		occupationally exposed to nonhuman primates. Journal of virology 78: 2780-2789.

Z

503	16.	Wolfe ND, Switzer WM, Carr JK, Bhullar VB, Shanmugam V, Tamoufe U, Prosser AT, Torimiro
504		JN, Wright A, Mpoudi-Ngole E, McCutchan FE, Birx DL, Folks TM, Burke DS, Heneine W. 2004.
505		Naturally acquired simian retrovirus infections in central African hunters. Lancet 363 :932-937.
506	17.	Calattini S, Betsem EB, Froment A, Mauclere P, Tortevoye P, Schmitt C, Njouom R, Saib A,
507		Gessain A. 2007. Simian foamy virus transmission from apes to humans, rural Cameroon.
508		Emerging infectious diseases 13: 1314-1320.
509	18.	Jones-Engel L. Engel GA. Schillaci MA. Rompis A. Putra A. Suarvana KG. Fuentes A. Beer B. Hicks
510		S. White R. Wilson B. Allan JS. 2005. Primate-to-human retroviral transmission in Asia. Emerging
511		infectious diseases 11:1028-1035.
512	19	Iones-Engel I. May CC. Engel GA. Steinkraus KA. Schillaci MA. Fuentes A. Romnis A. Chalise MK.
513	10.	Aggimarangsee N. Feeroz MM. Grant R. Allan IS. Putra A. Wandia IN. Watanabe R. Kuller I.
514		Thongsawat S Chaiwarith R Kyes RC Linial MI 2008 Diverse contexts of zoonotic transmission
515		of simian foamy viruses in Asia Emerging infectious diseases 14:1200-1208
516	20	Mouinga-Ondeme A. Caron M. Nkoghe D. Telfer P. Mary P. Saih A. Leroy F. Gonzalez IP.
517	20.	Gossain A. Kazanii M. 2012. Cross species transmission of simian feamy virus to humans in rural
510 510		Gebon Control Africa, Journal of virology 96:1255, 1260
510	21	Engel CA. Small CT. Solivon K. Ecorez MMA. Wang Y. Hacan MK. Oh G. Alam SP. Craig Kl
520	21.	Liger GA, Shidii CT, Soliven K, Feeloz Will, Walig X, Hasan Wik, On G, Alam SK, Claig KL,
520		Jackson DL, Malsen IV FA, Liniai WL, Jones-Engel L. 2013. 2001010 Similari Toamy Virus III
521		bangladesh reflects diverse patterns of transmission and co-infection. Emerging Microbes and
522	22	Infections 658.
525	22.	Betsem E, Rua R, Tortevoye P, Froment A, Gessain A. 2011. Frequent and recent human
524		acquisition of simian foamy viruses through apes bites in central Africa. PLoS pathogens
525	22	
520	23.	Boneva RS, Switzer WM, Spira IJ, Bhullar VB, Shanmugam V, Cong ME, Lam L, Heneine W,
527		Folks TM, Chapman LE. 2007. Clinical and virological characterization of persistent human
528		infection with similar foamy viruses. AIDS research and human retroviruses 23:1330-1337.
529	24.	Falcone V, Leupold J, Clotten J, Urbanyi E, Herchenroder O, Spatz W, Volk B, Bohm N, Toniolo
530		A, Neumann-Haetelin D, Schweizer M. 1999. Sites of similar foamy virus persistence in naturally
531		infected African green monkeys: latent provirus is ubiquitous, whereas viral replication is
532		restricted to the oral mucosa. Virology 257: 7-14.
533	25.	Murray SM, Picker LJ, Axthelm MK, Linial ML. 2006. Expanded tissue targets for foamy virus
534		replication with simian immunodeficiency virus-induced immunosuppression. Journal of virology
535		80: 663-670.
536	26.	Murray SM, Picker LJ, Axthelm MK, Hudkins K, Alpers CE, Linial ML. 2008. Replication in a
537		superficial epithelial cell niche explains the lack of pathogenicity of primate foamy virus
538		infections. Journal of virology 82:5981-5985.
539	27.	Soliven K, Wang X, Small CT, Feeroz MM, Lee EG, Craig KL, Hasan K, Engel GA, Jones-Engel L,
540		Matsen FAt, Linial ML. 2013. Simian foamy virus infection of rhesus macaques in Bangladesh:
541		relationship of latent proviruses and transcriptionally active viruses. Journal of virology
542		87: 13628-13639.
543	28.	Rua R, Betsem E, Gessain A. 2013. Viral latency in blood and saliva of simian foamy virus-
544		infected humans. PloS one 8:e77072.
545	29.	Gessain A, Rua R, Betsem E, Turpin J, Mahieux R. 2013. HTLV-3/4 and simian foamy retroviruses
546		in humans: discovery, epidemiology, cross-species transmission and molecular virology. Virology
547		435: 187-199.
548	30.	Rua R, Betsem E, Montange T, Buseyne F, Gessain A. 2014. In vivo cellular tropism of gorilla
549		simian foamy virus in blood of infected humans. Journal of virology 88:13429-13435.

550	31.	Rua R, Betsem E, Calattini S, Saib A, Gessain A. 2012. Genetic characterization of simian foamy
551		viruses infecting humans. Journal of virology 86: 13350-13359.
552	32.	Duda A, Luftenegger D, Pietschmann T, Lindemann D. 2006. Characterization of the prototype
553		foamy virus envelope glycoprotein receptor-binding domain. Journal of virology 80: 8158-8167.
554	33.	Stirnnagel K, Luftenegger D, Stange A, Swiersy A, Mullers E, Reh J, Stanke N, Grosse A, Chiantia
555		S, Keller H, Schwille P, Hanenberg H, Zentgraf H, Lindemann D. 2010. Analysis of prototype
556		foamy virus particle-host cell interaction with autofluorescent retroviral particles. Retrovirology
557		7: 45.
558	34.	Pietschmann T, Zentgraf H, Rethwilm A, Lindemann D. 2000. An evolutionarily conserved
559		positively charged amino acid in the putative membrane-spanning domain of the foamy virus
560		envelope protein controls fusion activity. Journal of virology 74: 4474-4482.
561	35.	Picard-Maureau M, Jarmy G, Berg A, Rethwilm A, Lindemann D. 2003. Foamy virus envelope
562		glycoprotein-mediated entry involves a pH-dependent fusion process. Journal of virology
563		77: 4722-4730.
564	36.	Stirnnagel K, Schupp D, Dupont A, Kudryavtsev V, Reh J, Mullers E, Lamb DC, Lindemann D.
565		2012. Differential pH-dependent cellular uptake pathways among foamy viruses elucidated using
566		dual-colored fluorescent particles. Retrovirology 9:71.
567	37.	Lindemann D, Pietschmann T, Picard-Maureau M, Berg A, Heinkelein M, Thurow J, Knaus P,
568		Zentgraf H, Rethwilm A. 2001. A particle-associated glycoprotein signal peptide essential for
569		virus maturation and infectivity. Journal of virology 75: 5762-5771.
570	38.	Pietschmann T, Heinkelein M, Heldmann M, Zentgraf H, Rethwilm A, Lindemann D. 1999.
571		Foamy virus capsids require the cognate envelope protein for particle export. Journal of virology
572		73 :2613-2621.
573	39.	Swiersy A, Wiek C, Zentgraf H, Lindemann D. 2013. Characterization and manipulation of foamy
574		virus membrane interactions. Cellular microbiology 15 :227-236.
575	40.	Goepfert PA, Shaw K, Wang G, Bansal A, Edwards BH, Mulligan MJ. 1999. An endoplasmic
576		reticulum retrieval signal partitions human foamy virus maturation to intracytoplasmic
577		membranes. Journal of virology 73: 7210-7217.
5/8	41.	Goepfert PA, Shaw KL, Ritter GD, Jr., Mulligan MJ. 1997. A sorting motif localizes the foamy
5/9	10	virus glycoprotein to the endoplasmic reticulum. Journal of virology 71: //8-/84.
580	42.	Mergia A, Leung NJ, Blackwell J. 1996. Cell tropism of the simian foamy virus type 1 (SFV-1).
201	40	Journal of medical primatology 25:2-7.
582 E02	43.	WIKOVITS JA, HOTTMAN PWI, RETAVIIM A, RUSCETTI FW. 1996. In vitro infection of primary and
202		retrovirus-infected numan leukocytes by numan loamy virus. Journal of virology 70: 2774-2780.
504 505	44.	Moedes A, Enssie J, Bieniasz PD, Heinkelein M, Lindemann D, Bock M, Micclure MO, Rethwilm
202 202		A. 1997. Human loamy virus reverse transcription that occurs late in the viral replication cycle.
500 E07	45	Journal of Virology 71:7305-7311.
507	45.	development as a vector for gone therapy. The lownal of general virology 80 (Dt 8) -2002 2000
200	16	development as a vector for gene therapy. The Journal of general virology 80 (Pt 8) :2003-2009.
509	40.	glycoprotoin modiated resistance to superinfection. Virology 214 :242, 252
501	47	Biycopi oteni mediated resistance to superimection. Virology 314. 243-252.
502	47.	for prototype feamy virus in human menkey and redent cells. Melecular therapy : the journal of
592		the American Society of Gone Therapy 20 :1158-1166
593	18	Chemann K Horn Δ Gerhmack F Armhruster N Krieg I Wiktorowicz T Weber C Stimpage
595	40.	K Lindomann D. Rethwilm A. Scheller C. 2012. Henaran sulfate is an attachment factor for
596		foamy virus entry Journal of virology 86:10028-10035
		A CONTRACT A THE MERICAL AND A CONTRACT A CONT

597	49.	Calattini S, Betsem E, Bassot S, Chevalier SA, Tortevoye P, Njouom R, Mahieux R, Froment A,
598		Gessain A. 2011. Multiple retroviral infection by HTLV type 1, 2, 3 and simian foamy virus in a
599		family of Pygmies from Cameroon. Virology 410: 48-55.
600	50.	Calattini S, Nerrienet E, Mauclere P, Georges-Courbot MC, Saib A, Gessain A. 2006. Detection
601		and molecular characterization of foamy viruses in Central African chimpanzees of the Pan
602		troglodytes troglodytes and Pan troglodytes vellerosus subspecies. Journal of medical
603		primatology 35: 59-66.
604	51.	Calattini S, Nerrienet E, Mauclere P, Georges-Courbot MC, Saib A, Gessain A. 2004. Natural
605		simian foamy virus infection in wild-caught gorillas, mandrills and drills from Cameroon and
606		Gabon. The Journal of general virology 85: 3313-3317.
607	52.	Rogers NG, Basnight M, Gibbs CJ, Gajdusek DC. 1967. Latent viruses in chimpanzees with
608		experimental kuru. Nature 216: 446-449.
609	53.	Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P. 2010. RDP3: a flexible and fast
610		computer program for analyzing recombination. Bioinformatics 26: 2462-2463.
611	54.	Schweizer M, Schleer H, Pietrek M, Liegibel J, Falcone V, Neumann-Haefelin D. 1999. Genetic
612		stability of foamy viruses: long-term study in an African green monkey population. Journal of
613		virology 73: 9256-9265.
614	55.	Luftenegger D, Picard-Maureau M, Stanke N, Rethwilm A, Lindemann D. 2005. Analysis and
615		function of prototype foamy virus envelope N glycosylation. Journal of virology 79: 7664-7672.
616	56.	Mouinga-Ondeme A, Betsem E, Caron M, Makuwa M, Salle B, Renault N, Saib A, Telfer P, Marx
617		P, Gessain A, Kazanji M. 2010. Two distinct variants of simian foamy virus in naturally infected
618		mandrills (Mandrillus sphinx) and cross-species transmission to humans. Retrovirology 7:105.
619	57.	Elder JH, Gautsch JW, Jensen FC, Lerner RA, Hartley JW, Rowe WP. 1977. Biochemical evidence
620		that MCF murine leukemia viruses are envelope (env) gene recombinants. Proceedings of the
621		National Academy of Sciences of the United States of America 74 :4676-4680.
622	58.	Cloyd MW, Hartley JW, Rowe WP. 1980. Lymphomagenicity of recombinant mink cell focus-
623		inducing murine leukemia viruses. The Journal of experimental medicine 151: 542-552.
624	59.	Robertson DL, Sharp PM, McCutchan FE, Hahn BH. 1995. Recombination in HIV-1. Nature
625		374: 124-126.
626	60.	Desrames A, Cassar O, Gout O, Hermine O, Taylor GP, Afonso PV, Gessain A. 2014. Northern
627		African strains of human T-lymphotropic virus type 1 arose from a recombination event. Journal
628		of virology 88: 9782-9788.
629	61.	Gartner K. Wiktorowicz T. Park J. Mergia A. Rethwilm A. Scheller C. 2009. Accuracy estimation
630		of foamy virus genome copying. Retrovirology 6:32.
631	62.	Blochmann R. Curths C. Coulibaly C. Cichutek K. Kurth R. Norley S. Bannert N. Fiebig U. 2014. A
632	•	novel small animal model to study the replication of simian foamy virus in vivo. Virology 448: 65-
633		73.
634	63.	Winkler IG. Flugel RM. Lochelt M. Flower RL. 1998. Detection and molecular characterisation of
635	05.	feline foamy virus serotynes in naturally infected cats. Virology 247 :144-151
636	64	Berka II. Hamann MV. Lindemann D. 2013. Early events in foamy virus-host interaction and
637	01.	intracellular trafficking Viruses 5:1055-1074
638	65	Herchenroder O. Moosmaver D. Bock M. Pietschmann T. Rethwilm A. Bieniasz PD. McClure
639	05.	MO Weis R Schneider I 1999 Specific hinding of recombinant foamy virus envelope protein to
640		host cells correlates with suscentibility to infection. Virology 255 :228-236
641	66	Rattini II Heard IM Danos O 1992 Recentor choice determinants in the envelope
642	00.	alyconrotains of amphotronic venotronic and polytronic murine leukemia virusos. Journal of
64.2		giveoproteinis or amprotropic, xenotropic, and polytropic mutilie leukenild viruses. Journal or
043		VILUIORY UU. 1400-1473.

646 647		inhibition by chemokines. Proceedings of the National Academy of Sciences of the United States of America 110 :9475-9480.				
648	68.	Zemba M, Alke A, Bodem J, Winkler IG, Flower RL, Pfrepper K, Delius H, Flugel RM, Lochelt M.				
649		2000. Construction of infectious feline foamy virus genomes: cat antisera do not cross-neutralize				
650	60	feline foamy virus chimera with serotype-specific Env sequences. Virology 266 :150-156.				
651	69.	Johnston PB. 1961. A second immunologic type of similar foamy virus: monkey throat infections				
652	70	and unmasking by both types. The Journal of Infectious diseases 109: 1-9.				
654	70.	nooks JJ, Gibbs CJ, Jr., Cutchins EC, Rogers NG, Lampert P, Gajdusek DC. 1972. Characterization				
655		Virusforschung 29 -38-55				
656	71	Rieniasz PD. Rethwilm A. Pitman R. Daniel MD. Chrystie I. McClure MO. 1995. A comparative				
657	, 1.	study of higher primate foamy viruses including a new virus from a gorilla. Virology 207 :217-				
658		228.				
659	72.	Materniak M. Kubis P. Rola-Luszczak M. Khan AS. Busevne F. Lindemann D. Lochelt M. Kuzmak				
660		J. 2015. Tenth international foamy virus conference 2014-achievements and perspectives.				
661		Viruses 7: 1651-1666.				
662						
663	Figur	a_{1} Coographic location of the 44 individuals infected with Cao_{1} or Ptr_{1} formula				
005	I igui	e 1. deographic location of the 44 marvialars infected with 090° of 1 tratoany				
664	virus	s strains. Individuals infected with <i>Ggo</i> -FV strains are indicated by a circle: purple for				
665	SFV-0	Gorl strains and blue for SFV-GorII strains. Individuals infected with a <i>Ptr</i> -FV strain are				
666	indic	ated by a square: yellow for SFV-CpzI strains and orange for SFV-CpzII strains. These				
667	colors will be used in all figures.					
668						
669	Figu	re 2. Phylogenetic analysis of complete simian foamy virus <i>envelope</i> sequences.				
670	Phylo	ogenetic analysis of complete <i>env</i> sequences (2967 bp long for the PFV strain) from 43				
671	SFV i	solates including the 24 complete sequences generated in this study (in bold) and 19				
672	previ	ously published sequences. SFVmarm, a marmoset strain from South America, was the				
673	outgr	roup. Gabon isolates are underlined and asterisks indicate FVs isolated from non-				
674	huma	in primates. The phylogenic tree was derived by the neighbor-joining method using				

Colin P, Benureau Y, Staropoli I, Wang Y, Gonzalez N, Alcami J, Hartley O, Brelot A, Arenzana-

Seisdedos F, Lagane B. 2013. HIV-1 exploits CCR5 conformational heterogeneity to escape

675 the GTR model (gamma=0.6749). Horizontal branch lengths are drawn to scale, with the

644

645

67.

 \sum

676 bar indicating 0.1 nucleotide replacements per site. Numbers on each node indicate the 677 bootstrap value (calculated on 1,000 replicates) supporting the group.

678

679 Figure 3. Definition of a conserved and a variant region in the *envelope* gene of *Ggo*-680 and *Ptr*-foamy virus. A, D: (%) nucleotide identity between the different subgroups of *Ggo*-681 and *Ptr*-FV in the complete (A), the conserved (D) or the variant region (D) of the *envelope* 682 gene. Percentage identity was determined after alignment by the DAMBE program and 683 using CLC software. B, C: Gorilla and chimpanzee foamy virus sequences were aligned using 684 DAMBE software. The region from 698-798 bp and 1398-1498 bp (positions defined 685 according to the PFV env sequence) are shown for Ggo-FV sequences (B) and Ptr-FV 686 sequences (C), respectively. Dots indicate identity. The red vertical lines indicate the 687 position between two codons where the env sequences diverge into separate variants 688 (defined on the right of the alignment) and the green vertical lines show the position where 689 the variants become very similar again.

690

691 Figure 4. Phylogenetic analysis of the envelope conserved and variant regions. 692 Phylogenetic trees corresponding to (A) the conserved region (2214 bp for PFV) or (B) the 693 variant region (753 bp for PFV) were derived from the neighbor joining method (GTR; 694 gamma = 0.6749). Horizontal branch lengths are drawn to scale, with the bar indicating 0.1 695 nucleotide replacements per site. Bootstrap values were calculated on 1,000 replicates. The 696 SFVmarm strain was used as an outgroup; Gabon isolates are underlined and asterisks 697 indicate FVs isolated from non-human primates.

Journal of Virology

698 C- Unrooted phylogenetic tree of the variant region. Closely related strains that formed a 699 robust cluster (as seen on B) are represented with black triangles. AGM A, B, C and D are 700 clusters of African Green Monkey (Cercopithecus) sequences previously defined by 701 Schweizer (54).

702

703 Figure 5. Comparison of Env protein of Ggo-FV and Ptr-FV strains. A: Structure and 704 important domains of Env FV protein present in all sequences of this study. LP: leader 705 peptide, SU: Surface glycoprotein, TM: Transmembrane glycoprotein; WXXW motif: Trp-X-706 X-Trp motif, interaction site with Gag protein; RXXR: Arg-X-Arg cleavage sites; RBD: 707 Receptor Binding Site; FP: fusion peptide (i, i+3/4,i+7 pattern); MSD: Membrane Spanning 708 Domain; ERRS: Endoplasmic Reticulum Retrieval Signal (Lys at -3, Lys or Arg at -4 or/and -709 5 relative to C-terminus). B: Amino acid identity in the conserved and variant regions was 710 investigated using CLC software based on the alignment performed with DAMBE. C: 711 Location of N-glycosylation sites of the different *Ggo*- and *Ptr*-FV subgroups.

712

713 Figure 6. Recombination analysis of the SFV-GorII and SFV-CpzI envelope sequences.

714 A, B, C, D: Bootscan (A,C) analyses and (B,D) Simplot analyses using the SFV-GorII (A,B) or 715 SFV-CpzI (C,D) env sequences as a query against the other groups of SFV isolates were 716 performed.

- 717
- 718
- 719
- 720

Gorilla FV	Length		Name	Sequence 5'->3'	Site of hybridization /SFVggo (bp)
Fragment 1	720 bp Ex	xternal PCR I	AENVF1b	GAACTGTGGTAATTGTGGACC	6566-6586
		F	R AENVR1b	CCACGAGACCAAGAACAATA	7292-7311
	In	nternal PCR	AENVF1	GGTAATTGTGGACCATCTTGG	6573-6593
		F	R AENVR1	GGATCCACGAGACCAAGAAC	7288-7307
Fragment 2	1102 bp Ex	xternal PCR I	= BF3	CATCCACCCCTCCTGCCT	6431-6448
		F	R RR3	CCTGTAAATGAAATGCCTAAT	8228-8248
	In	nternal PCR	- DF3	CTATAATACACACGGAGAGG	7113-7132
		F	R RR3	CCTGTAAATGAAATGCCTAAT	8228-8248
Fragment 3	1490 bp Ex	xternal PCR I	AENVF2	TACGACAACAAGATTATGAAG	8109-8129
		F	R AENVR3	CTGAGTGAGCTTGTTGGTCC	9640-9659
	In	nternal PCR	AENVF2b	ATATCAAGAATGTAAGTTGG	8140-8159
		F	R AENVR3b	TCTGCAAACTCTGAGTGAGC	9630-9649
Chimpanzee FV	Length		Name	Sequence 5'->3'	Site of hybridization /PFV (bp)
Fragment 1	1645 bp Ex	xternal PCR I	CPZENVF1b	ACTGTTGTTATTTTGGACCA	7066-7085
		F	R CPZENVR1	CCTTTGTAGGCCTAGTAGAT	8763-8782
	In	nternal PCR	CPZENVF1	GGCAACAACAGAACTGTAAG	7090-7109
		F	R CPZENVR1b	CCTGTAAATGAAATGCCTAA	8735-8754
Fragment 2	1802 bp Ex	xternal PCR I	CPZENVF2	TTCTCTTTGTGGGAAGGAG	8330-8349
		F	R CPZENVR2	CTTAGTGAGCTTGTTGGTCC	10141-10160
	In	nternal PCR	CPZENVF2b	TCTTTGTGGGAAGGAGATTG	8333-8352
		F	R CPZENVR2b	CAGACTCTTAGTGAGCTTGT	10135-10154

Table 1. Polymerase chain reaction primers used to amplify the envelope gene of

F: Forward; R: Reverse; bp: base pair. Position / SFVggo or PFV complete sequences.

gorilla and chimpanzee foamy virus strains.

Accepted Manuscript Posted Online

 \leq

Isolate	Strain	Origin	Type of sequence	Genbank Accession no.
PFV	Chimpanzee (P. t. schweinfurthii)	human	complete genome	Y07725
SFVcpz	Chimpanzee (P. t. verus)	NHP	complete genome	U04327
AG15	Chimpanzee (P. t. troglodytes)	human	complete genome	JQ867462
Bad327	Chimpanzee (P. t. troglodytes)	human	complete genome	JQ867463
SFVggo	Gorilla (G. g. gorilla)	NHP	complete genome	HM245790
Bak74	Gorilla (G. g. gorilla)	human	complete genome	JQ867464
Bad468	Gorilla (G. g. gorilla)	human	complete genome	JQ867465
SFVora	Orangutan	NHP	complete genome	AJ544579
SFVAGMhu	African Green Monkey (Chlorocebus aethiops)	human	complete genome	AX575326
SFV-agm3	African Green Monkey (Chlorocebus aethiops)	NHP	complete genome	NC_010820
SFVka	African Green Monkey (Chlorocebus aethiops)	human	partial env	AJ244092
SFV3	African Green Monkey (Chlorocebus aethiops)	NHP	partial env	AJ244094
agm1 to agm37 (19 isolates)	African Green Monkey (Chlorocebus aethiops)	NHP	partial <i>env</i>	AJ244067 to AJ244091
AG16	Cercopithecus	NHP	complete genome	JQ867466
SFV-mcy1	Macaque (Macaca cyclopis)	NHP	complete genome	NC_010819
SFV-mcy2	Macaque (Macaca cyclopis)	NHP	complete genome	KF026286
SFVR289HybAGM	Macaque (Macaca mulatta)	NHP	complete genome	JN801175
SFVmmu-K3T	Macaque (Macaca mulatta)	NHP	env	KF026287
SFVmmu-A4W	Macaque (Macaca mulatta)	NHP	env	KF026288
SFVsp	Spider monkey	NHP	complete genome	EU010385
SFVmar	Marmoset	NHP	complete genome	GU356395
SFVsq	Squirrel monkey	human	complete genome	GU356394

733 Table 2. Genbank accession numbers and origin of other SFV isolates used in the

734 study.

735 NHP: Non Human Primate.

 \leq

Human individuals	Country	Ethnicity	Individual	Sex	Age at cont/samp	FV strain	Type of env FV sequence	FV variant	FV isolate (accession number)
	Cameroon	Pygmy	Pyl106	М	15/60	Chimpanzee	Complete	SFV-Cpzl	SFVptr-hu.Pyl106 (KT211263)
			BAK46	Μ	26/50	Gorilla	Complete	SFV-Gorl	SFVggo-hu.Bak46 (KT211255)
			BAK74	М	26/47	Gorilla	Complete	SFV-Gorll	Bak74 (JQ86746)
			BAK82	Μ	46/50	Gorilla	Complete	SFV-Gorl	SFVggo-hu.Bak82 (KT211247)
			BAK228	М	29/70	Gorilla	Complete	SFV-Gorll	SFVggo-hu.Bak228 (KT211246)
			BAK242	М	30/49	Gorilla	Complete	SFV-Gorl	SFVggo-hu.Bak242 (KT211254)
			Sabak36	М	40/68	Gorilla	Complete	SFV-Gorll	SFVggo-hu.Sabak36 (KT211249)
			BAK33	М	25/45	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak33 (KT211277)
			Bak55	Μ	30/65	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak55 (KT211273)
			BAK56	Μ	40/65	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak56 (KT211274)
			BAK132	Μ	30/64	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak132 (KT211276)
			BAK177	М	26/36	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak177 (KT211281)
			BAK224	Μ	19/38	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak224 (KT211280)
			BAK232	Μ	40/60	Gorilla	Partial	SFV-Gorll	SFVggo-hu.Bak232 (KT211278)
			BAK270	Μ	25/60	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bak270 (KT211272)
			Bobak153	М	53/59	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bobak153 (KT211279)
			Bobak237	М	?/68	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Bobak237 (KT211275)
			Lobak2	М	37/57	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Lobak2 (KT211282)
			Lobak89	М	20/50	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Lobak89 (KT211287)
			Mebak65	М	20/40	Gorilla	Partial	SFV-Gorl	SFVggo-hu.Mebak65 (KT211283)
			801001	M	35/60	Gorilla	Partial	SFV-Gorl	SFVago-hu.801001 (KT211284)
			CH29	M	49/50	Gorilla	Partial	SEV-Gorl	SEVago-hu CH29 (KT211289)
		Bantu	AG15	M	28/71	Chimpanzee	Complete	SFV-Cozll	AG15 (JQ867462)
			Bad316	M	35/51	Chimpanzee	Complete	SEV-CozII	SEVptr-hu Bad316 (KT211262)
			Bad327	M	30/33	Chimpanzee	Complete	SFV-Cozll	Bad327 (JQ867463)
			BAD348	M	19/27	Gorilla	Complete	SEV-Gorl	SEVggo-hu Bad348 (KT211252)
			BAD456	M	24/30	Gorilla	Complete	SFV-Gorl	SFVggo-hu.Bad456 (KT211251)
			BAD463	M	37/43	Gorilla	Complete	SEV-Gorl	SEVggo-hu Bad463 (KT211253)
			BAD468	M	25/35	Gorilla	Complete	SEV-Gorl	Bad468 (JQ867465)
			BAD551	M	37/38	Gorilla	Complete	SEV-Gorll	SEVggo-hu Bad551 (KT211248)
			CH101	M	65/76	Gorilla	Complete	SEV-Gorll	SEVggo-bu CH101 (KT211256)
			BAD332	M	25/37	Gorilla	Partial	SEV-Gorl	SEVago-hu Bad332 (KT211288)
			BAD349	M	32/40	Gorilla	Partial	SEV-Gorl	SEV/ggo-bu Bad349 (KT211270)
			BAD350	M	40/68	Gorilla	Partial	SEV_Gorll	SEVago-bu Bad350 (KT211290)
			BAD447	M	40/56	Gorilla	Partial	SEV_Corl	SEV/ggo-bu Bad447 (KT211271)
				M	53/53	Gorilla	Partial	SEV-Gorll	SEVago-bu AKO394 (KT211286)
			CH61	M	52/65	Gorilla	Partial	SEV-Gorl	SEVago-bu CH61 (KT211285)
	Gabon	Bantu	H3Gab56	M	47/48	Chimpanzee	Complete	SEV_CozII	SEVntr-hu H3Gab56 (KT211267)
	Gubon	Duniu	H4Gab50	M	50/51	Chimpanzee	Complete	SEV-CozII	SEV/ntr-bu H4Gab59 (KT211268)
			H2Gab54	M	52/53	Gorilla	Complete	SEV_Gorl	SEV/000-bu H2Gab54 (KT211208)
			H5Cab27	M	52/90	Gorilla	Complete	SEV Corl	SEVago by H5Cab27 (KT211257)
			H6Cab51	M	29/56	Gorilla	Complete	SEV Corl	SEVago by H6Cab51 (KT211250)
			H7Cab42	M	20/00	Gorilla	Complete	SEV Corll	SEVago by H7Cab42 (KT211261)
			H12Cab60	N/	20/00	Gorilla	Complete	SEV Corl	SFVgg0-III.II/Gab42 (KT211201)
Non Human Primates	Country	Species	Name	Sex	Age at sampling	Situation	Type of env FV sequence	FV variant	ST Vgg0-h0.1112Gab09 (K1211200)
	Cameroon	Gorilla	GooCam7SEV	F	7	wild-born 700	Complete	SEV-Gorl	SEV000 Cam7 (KT211250)
	20	Chimpanzee	CpzCam15SEV	M	6	wild-born zoo	Complete	SEV-Cnzl	SEVptr Cam15 (KT211264)
	Gabon	Chimpanzee	Cpz.JudWd	F	adult	wild-born zoo	Complete	SFV-Cnzll	SEVptr. JudWd (KT211266)
	Gubon	Chimpanzee	Cnz133W/d	2	adult	wild-born net	Complete	SEV-Cozil	SEVptr 133Wd (KT211265)
<u>لــــــــــــــــــــــــــــــــــــ</u>	1	1	Op2100440	1 :	auun	wiid-boiri, pet	Complete		Gr vpti.100vvd (K1211200)

745

746 Table 3. Epidemiological data of the 44 individuals and four non-human primates

747 infected with simian foamy virus included in this study.

 \leq

Σ

748	The env of FVs from individuals in bold (BAK74, AG15, BAD327, BAD468) was sequenced
749	by our laboratory in a previous study (31). Cont : contact; samp : sampling; FV : Foamy
750	Virus.



Z



Complete env

		SF	V-Gorl	96.7-99.8	88.2-90.4	76.5-77	7.2 74.2-	75.4		
		SF	V-Gorll		96.4-98.9	74.1-74	4.7 75.3	-77		
		SF	V-Cpzl			95.9	86.5-	89.4		
		SF	V-Cpzll				90.7-	100		
В	700		720	740)	760)	780	Pos	sition / PFV e <i>nv</i> (bp)
SFVggo-hu.Bak242	CTCATTGTT	ΑΤΤΤΑGΤΑΑΑΑΤ	аталалстсст	CAACCATGGCCT	АСТСААСААСТ	TATTG CAGA	T CAA TG T <mark>C CA</mark> T	ΓΤΑ C CAGGA ΤΑ C	CA TG CCGGAG TA	GA.
SFVggo-hu.Bak82 SFVggo-hu.Bad456 SFVggo.Cam7 SFVggo-hu.Bad463 SFVggo-hu.Bad348		. c			G		· · · · · · · · · · · · · · · · · · ·			:: :: :: SFV-
Bad468 SFVggo-hu.Bak46 SFVggo-hu.H6Gab SFVggo-hu.H12Gab	51		G G G G						. c.	Gorl
SFVggo-hu.H5Gab2 SFVggo-hu.Bad551 Bak74 SFVggo-hu.Bak228	27				. A G				GCC.AT.TTAG GC.AT.TTAG GC.AT.TTAG	SFV-
SFVggo-hu.Sabak3 SFVggo-hu.CH101 SEVggo-hu.H7Gab	6		G	т	. A	c	т. т		GC AT. TTAG	Gorll
	1400		1420	14	10	146	50	1480		-
SFVggo-hu.Bak242 SFVggo-hu.Bak82	ТGAAGCTAC	ATATGACTTCGC	ATTTTTAGCTT	атсалалталт	TCCCAGCGCCT	ATTTGTATT	TCGAGTAAGC	АА ТА CGA CAA C	AAGATTATGAAG	
SFVggo-hu.Bad456 SFVggo.Cam7										··· SEV
SFVggo-nu.Bad463 SFVggo-hu.Bad348 Bad468		c	c.		T T		A			Gorl
SFVggo-hu.Bak46 SFVggo-hu.H6Gab	51	c	c				A			
H12Gab69 SFVggo-hu.H2Gab	54 • • • • • • • • • •	тт.					A			
SFVggo-hu.H5Gab2 SFVggo-hu.Bad551	27GTT		GC.TT		. T T. GA C	CAA	. TA			SEV-
Bak/4 SFVggo-hu.Bak228 SFVggo-hu.Sabak3 SFVggo-hu.CH101 SFVggo-hu.H7Gab4	GTT 6GTT 12GTT	T T T	GCC.TT GCC.TT GC.TT GC.TT	A TG . GA A TG . GA A TG . GA A TG . GA	. T T. GA CO . T T. GA CO . T T. GA CO . T T. GA CO	CAAA CAAA CAAA CAAA	GAATCA.CAT GAATCA.CAT GAATCA.CAT GAATCA.CAT	CA TA		Gorll
-									Pos	sition / PFV
С	700		720	74)	760)	780	(e <i>nv</i> (bp)
SFVptr-hu.Bad316 Bad327	CACATTGCT	ATTTGG TAAAA 1			AAAGATGGCCTT	ATAGCTGAT	ТСА БТ БТ ССАТ		A TGG CA CAAG T	SFV-
AG15 SFVptr-hu.H3Gab56	6			A	. c	c		c	¢	Cpzll
SFVptr-hu.H4Gab59 SFVptr.JudWd	9т.	C. A	· · · · · · · · · · · · · · · · · · ·	т		c	A	т	. с т. с	
SFVptr.133Wd SFVptr-hu.Pyl106 SFVptr.Cam15	. T T. . T T.		GA.T.AG		. GT AA. TT. G . GT GA. TT. G			AGGGCAT.	CAGGT. T./	SFV- Cpzl
	1400 I		1420 I	14 I	40	146 I	60	1480 I		
SFVptr-hu.Bad316 Bad327	AGAAGCTC	TTTTTGATTTTG		CA CAAAGGAAT	TTTCCTGCCCCA	ATTTGTATA		AAATTAGGGACG		SFV-
SFVptr-hu.H3Gab56	6 9		т							Cpzll
SFVptr.JudWd SFVptr.133Wd	G	c.	c.	G.AC				. G		∷ <mark>ਟ SF</mark> V
SFVptr-hu.Pyl106 SFVptr.Cam15	TTAC TTAC	C.AC.	.ACA	АТААС.С АТААС.С			GA.	т.		Cpzl
DΓ	Conserved	region	Variant regi	on	Conse	rved regi	ion	Positio	on / PFV	
L 1	1 720/721 1473/1474 2967									
nt identity (%)	SEVGor	SEV.Goril	SEV Cord	SEV Con-	I nt ident	ity (%)	SEV.Corl	SEV	SEV.Co-	SEV_Cort
SEV_Cord	96 8-00 P	96 1_00 2	78_78 8	77 6_79 C		Gorl	0/ 8-100	62 2_64 7	71 7-72	62 2_64 2
SEV-Goril	30.0-99.0	06 5 00	77 8 79 6	77 / 70 0		Gorl	34.0-100	02.2-04.1	62 2 62 4	68 2 71 0
SEV Cord		90.0-99	05.0	01/07/				90-99.0	02.2-03.4	623620
SEV_Cort			30.8	01 7 00					30.0	876.00.0
SFV-CPZII	0			31.7-90		<u>אר אר אר אר אין אין אין אין אין אין א</u>	ariant an	ragion		07.0-99.9
	(Conco	rven env r	enion			· · · · · · · · · · · · · · · · · · ·				-

nt identity (%) SFV-Gorl SFV-Gorll SFV-Cpzl SFV-Cpzll

Figure 3

Σ

Figure 4





SFVggo-hu.Bak82 SFVggo.Cam7 *

Came

roon

Gabon

Gorilla

no

Macaque

1/PFV

env (bp)

SFVsp

SFVpvr.SFV7

0.1

Clade 2

Chimpanzee

Clade 1





M



 \sum





Figure 6

 \leq