Origin and Biology of Simian Immunodeficiency Virus in Wild-Living Western Gorillas

Citation for published version:

Digital Object Identifier (DOI):
10.1128/JVI.02311-08

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Journal of Virology

Publisher Rights Statement:
RoMEO blue

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Origin and Biology of Simian Immunodeficiency Virus in Wild-Living Western Gorillas

Jun Takehisa, Matthias H. Kraus, Ahidjo Ayouba, Elizabeth Bailes, Fran Van Heuverswyn, Julie M. Decker, Yingying Li, Rebecca S. Rudicell, Gerald H. Learn, Cecile Neel, and Beatrice H. Hahn

Departments of Medicine and Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294; Institut de Recherche pour le Développement (IRD) and Department of International Health, University of Montpellier 1, 34994 Montpellier Cedex 5, France; Institute of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, United Kingdom; Projet Prévention du Sida ou Cameroun (PRESICA), Yaoundé, Cameroon; and Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Received 4 November 2008/Accepted 2 December 2008

Western lowland gorillas (Gorilla gorilla gorilla) are infected with a simian immunodeficiency virus (SIVgor) that is closely related to chimpanzee and human immunodeficiency viruses (SIVcpz and HIV-1, respectively) in west central Africa. Although existing data suggest a chimpanzee origin for SIVgor, a paucity of available sequences has precluded definitive conclusions. Here, we report the molecular characterization of one partial (BQ664) and three full-length (CP684, CP2135, and CP2139) SIVgor genomes amplified from fecal RNAs of wild-living gorillas at two field sites in Cameroon. Phylogenetic analyses showed that all SIVgor strains clustered together, forming a monophyletic lineage throughout their genomes. Interestingly, the closest relatives of SIVgor were not SIVcpzPtt strains from west central African chimpanzees (Pan troglodytes troglodytes) but human viruses belonging to HIV-1 group O. In trees derived from most genomic regions, SIVgor and HIV-1 group O formed a sister clade to the SIVcpzPtt lineage. However, in a tree derived from 5 pol sequences (~900 bp), SIVgor and HIV-1 group O fell within the SIVcpzPtt radiation. The latter was due to two SIVcpzPtt strains that contained mosaic pol sequences, pointing to the existence of a divergent SIVcpzPtt lineage that gave rise to SIVgor and HIV-1 group O. Gorillas appear to have acquired this lineage at least 100 to 200 years ago. To examine the biological properties of SIVgor, we synthesized a full-length provirus from fecal consensus sequences. Transfection of the resulting clone (CP2139.287) into 293T cells yielded infectious virus that replicated efficiently in both human and chimpanzee CD4+ T cells and used CCR5 as the coreceptor for viral entry. Together, these results provide strong evidence that P. t. troglodytes apes were the source of SIVgor. These same apes may also have spawned the group O epidemic; however, the possibility that gorillas served as an intermediary host cannot be excluded.
Second, the route by which gorillas could have acquired SIV was no prior evidence of SIV infection in captive gorillas. Thus, while chimpanzees were believed to avoid physical interaction with other primates (41, 60, 67). In contrast, chimpanzees are avid hunters that prey on smaller monkeys (13, 29, 40, 61). Thus, while chimpanzees likely acquired SIV through predation (6, 57), this mode of transmission is improbable for SIVgor. Third, the range of western lowland gorillas (G. g. gorilla) overlaps that of central chimpanzees (Pan troglodytes troglodytes) (Fig. 1), and SIVgor is significantly more closely related to SIVcpz from the central subspecies than it is to SIVcpzPtt from eastern chimpanzees (Pan troglodytes schweinfurthii) (71). These data would suggest that central chimpanzees transmitted their virus to sympatric gorillas, albeit by an as-yet-unknown route. However, under such a scenario, gorilla viruses might be expected to fall within the radiation of SIVcpzPtt in evolutionary trees, which is not the case (71). Finally, given that gorillas acquired a chimpanzee SIV (71), one could ask whether this has occurred only once and only in west central Africa. As shown in Fig. 1, eastern chimpanzees and gorillas also live in sympatry, thus providing potential transmission opportunities. Given these uncertainties, it is clear that additional chimpanzee and gorilla viruses need to be analyzed to elucidate the ancestry of SIVgor.

In the present study, we sought to gain new insight into the evolutionary origin of SIVgor and to begin to probe its biological properties. To accomplish this, we amplified partial (BO664) and full-length (CP684, CP2135, and CP2139) genomes from fecal samples previously shown to contain SIVgor-specific viral RNA. We then performed analyses to (i)
characterize the phylogenetic relationships of SIVgor over the entire length of its genome, (ii) investigate whether SIVgor has undergone recombination, (iii) identify the primate reservoir that gave rise to SIVgor, (iv) examine the possibility of local chimpanzee-to-gorilla SIV transmission, and (v) generate the first replication-competent molecular clone of SIVgor for biological characterization. Our results strongly suggest that gorillas acquired SIVcpzPtt from P. t. troglodytes apes and that the current SIVgor lineage is the result of a single cross-species transmission event. Although it is still unclear whether chimpanzees or gorillas were the source of HIV-1 group O, in vitro studies demonstrate that SIVgor has many of the biological properties necessary for establishing a persistent infection in humans.

MATERIALS AND METHODS

Ape fecal samples. To amplify full-length SIVgor genomes, fecal samples representing four naturally infected (female) western lowland gorillas were selected for analysis. These included samples CP684 (collected on 18 April 2004) and BQ664 (9 August 2004), which were previously reported to contain SIVgor viral RNA (71), as well as samples CP2135 (11 February 2007) and CP2139 (11 February 2007), which were previously reported to contain SIVgor sequences. Reverse transcriptase PCR (RT-PCR) analysis was performed as described previously (33, 51, 52, 65, 71), with some modifications. For CP684, 10 µl of RNA was first incubated with 40 pmol of the outer reverse primer (1 µl) for 10 min at 65°C, rapidly cooled on ice, and then added to the remaining components in a 20-µl reaction volume containing 1 mM deoxynucleoside triphosphate (dNTP), 5 mM dithiothreitol, 1 unit (U)/µl RNase inhibitor (Ambion, Austin, TX), and 2.5 U/µl Expand reverse transcriptase (Roche Diagnostics, Indianapolis, IN). This reaction mixture was incubated for 60 min at 42°C and heat inactivated at 95°C for 2 min, and 10 µl was used for PCR. For CP2135 and CP2139, cDNA was synthesized in a 20-µl reaction containing 10 µl fecal RNA, 1× reaction buffer, 1 mM dithiothreitol, 1 unit (U)/µl RNase inhibitor (Ambion, Austin, TX), and 2.5 U/µl Expand reverse transcriptase (Roche Diagnostics, Indianapolis, IN). The reaction mixture was incubated at 55°C for 90 min at 55°C. Following heat inactivation of the enzyme at 70°C for 15 min, 7 µl cDNA was then added to a PCR mixture containing 1× Expand buffer 2, 0.5 mM dNTP, 5 mM dithiothreitol, 2 pmol outer reverse primer, and 200 U of SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA). The reaction mixture was incubated at 50°C for 90 min and, after addition of another 200 U of enzyme, for an additional 90 min at 55°C. Following heat inactivation of the enzyme at 70°C for 15 min, 7 µl cDNA was then added to a PCR mixture containing 1× Expand buffer 2, 0.5 mM dNTP, 300 nM of first-round PCR primers, 25 µg bovine serum albumin, and 3.75 U of Expand Long Template enzyme mixture (Roche Diagnostics, Indianapolis, IN). Different combinations of consensus sequence- and strain-specific primers were used to amplify between 8 and 15 subgenomic fragments per proviral genome (Fig. 2);
primer sequences and corresponding fragment lengths are listed in Table S1 in the supplemental material. The amplification conditions for first- and second-round PCR included 35 cycles of denaturation (94°C, 0.5 min), annealing (55°C, 0.5 min), and extension (68°C, 0.5 min). Amplicons were gel purified and sequenced directly using an ABI 3730 DNA analyzer. Chromatograms were examined for positions of base mixtures by using Sequencher version 4.7 (Gene Codes Corporation, Ann Arbor, MI) or Lasergene Sequenator Pro version 7.1.0 (DNASTAR, Inc., Madison, WI). In addition, nucleotide differences between adjoining fragments in regions of sequence overlap were recorded (see Table S2 in the supplemental material). The genome of CP2139 was amplified twice (Fig. 2), the first time using a combination of consensus sequence- and strain-specific primers (CP2139.1) and the second time using only strain-specific primers designed to generate larger fragments (CP2139.2). Consensus sequences from these two amplifications were then used to generate a master consensus sequence (see Table S3 in the supplemental material), which served as the template to generate a full-length SIVgor molecular clone.

Amplification of SIVcpz sequences. For samples CP1973 and CP2680, a highly conserved SIVcpz pol fragment (220 bp) was amplified, using degenerate HIV-1 SIVcpz/SIVgor consensus primers (F1, 5'-CAGTRGHCAGTGGNTWCMT AGARCGACARGT-3'; R1, 5'-ACBACYGCNCTTCHCCTTTT-3'; F2, 5'-A YAYCCCCTAGCCTAGAAGTGTG-3'; and R2, 5'-GTCCTCTCCAAAC DDGRCTTCTCTGTC-3'). The amplification conditions for both first- and second-round PCR included 35 cycles of denaturation (94°C, 0.25 min), annealing (55°C, 0.5 min), and extension (68°C, 0.5 min).

Construction of a full-length SIVgor clone. To obtain a full-length infectious molecular clone of SIVgor, the CP2139 master consensus sequence was used as a template to chemically synthesize three subgenomic fragments (Blue Heron Biotechnology). Sequencing of the corresponding fragments confirmed the absence of internal rearrangement. The three adjoining fragments were cloned into the MluI and NotI sites of pBR-MCS. Following digestion of this plasmid with MluI and NotI, both 3.0 kb MluI- and 2.6 kb NotI-adjacent fragments were simultaneously ligated to generate a full-length proviral clone. Transformation of XL2-MRF bacteria (Stratagene, La Jolla, CA) initially yielded only defective clones containing either single-nucleotide substitutions or larger insertions/deletions. One functional clone contained an approximately 150-bp insert in its env gene. Since the sequence of this clone was otherwise identical to the CP2139 master consensus sequence, it was digested with Apal and recircularized. Recircularized XL2-MRF transforms were selected for appropriate antibiotic resistant clones, transferred into 293T cells, and tested for infectivity in the JCS3-BL assay (21, 77). One functional clone (pBR-CP2139.287) was identified and grown large scale.

Large-scale plasmid preparation. The pBR-CP2139.287 clone containing the full-length SIVgor genome was propagated in XL2-MRF cells (Stratagene) at 30°C in a Forma orbital benchtop shaker (model 4262), using 500 mL LB containing 100 µg/ml ampicillin in 2 liter disposable Erlenmeyer flasks (16-cm diameter, vented cap; Corning) at low levels of agitation (225 rounds per minute). Bacterial cultures were harvested prior to reaching saturating growth density and purified to be endotoxin free (Qiagen, Valencia, CA). Due to the extreme instability of this clone in bacteria, all large-scale plasmid preparations were sequenced. The replication-competent pBR-CP2139.287 clone has been submitted to the National Institutes of Health Research and Reference Program (Rockville, MD).

Phylogenetic methods. Newly derived full-length SIVgor sequences were translated and compared to previously published full-length HIV-1 and SIVcpz sequences from the database as follows: for HIV-1 group M subtype A, U455A (GenBank accession numbers are listed in Table S4 in the supplemental material); for subtype B, HXB2; for group N, YBF30 and YBF106; for group O, MVP5180 and AN70; for SIVcpzPrb, MB987, D943, LB7, MB66, EK505, MT145, CAM3, CAM5, CAM13, GAB1, GAB2, and US; and for SIVcpzPrs, ANT, TAN1, TAN2, and TAN3. Amino acid sequences were aligned using CLUSTAL W (66). Sites that could not be aligned unambiguously and sites with a gap in any sequence were discarded. In regions of gene overlap (e.g., Gag/Pol and Pol/Vif genes), the carboxy termini of the overlapping protein sequences were discarded. Trees were inferred by the Bayesian Markov chain/Monte Carlo method of phylogenetic estimation (79), implemented in MrBayes version 3.1.2 (50) by using the mixed model of amino acid evolution with gamma-distributed rates at sites and 1 million generations with 25% burn in. Average standard deviations of split frequencies were 0.01 or lower. Four major regions of the proteome were analyzed: Gag, Pol, Pol2, and Env. The Pol sequence was divided into two segments (Fig. 2A) and a combination breakpoint previously identified in HIV-1 group N (26). Trees of partial pol sequences for SIVcpzPrb strains CP1973 and CP2680 were inferred as described above from an alignment of 67 amino acids, using 10 million generations in MrBayes.

To calculate the time to the most recent common ancestor (MRCA) of the SIVgor clade, regions of an HIV-1/SIVcpz/SIVgor nucleotide sequence alignment corresponding to available BQ664 sequences were concatenated (total length at the nucleotide alignment was 2,952 nucleotides). This alignment included new and previously published BQ664 sequences (see Table S4 in the supplemental material), sequences listed above, and additional sequences from HIV-1 group M: for ET02220 (subtype C), ELI (D) and 02CM.006BB (F2); for HIV-1 group N, 04CM.1015-04, 04CM.1131-03, and DJO131; and for HIV-1 group O, SEMP129, SEMP130, and VAU. A maximum-likelihood phylogenetic tree was constructed with PAUP* version 4.0b10 (64), using a general time-reversible model with gamma-distributed rates of nucleotide substitution and invariant sites. Trees were inferred from summed branch lengths and taking account of dates of virus isolation.

For the analysis of SIVgor V3-loop sequences, a phenetic dendrogram was estimated using the unweighted-pair group method with average linkages from a distance matrix of uncorrected sequence differences, using NEIGHBOR from PHYLIP version 3.64 (24). The following additional V3 sequences were included: for HIV-1 group M, A1, 92UG037 (see Table S4 in the supplemental material); for HIV-1 group N, B6-HIV-1, 77DKDT84 (sequences in this clade were aligned with the clade previously described by Quan et al. in reference 38); for SIVcpz, CM240 (01-AE), V1850 (F1), 02CM.0016BBY (F2), DRCBL (G), V9791 (H), and EQUITB11C (K); for SIVcpzPrb, BM1034, LB714, LB715, MB801, MB802, MB776, MB803, DP395, and SL995; and for HIV-1 group O, SEMP130 (O1), 97US08692A (O2), 96CMABB637 (O3), AN70 (O4), and MVP1810 (O5). To test for recombination among the SIVcpz ancestors of HIV-1 groups M, N, and O and SIVgor, phylogenetic trees were examined across a concatenated Gag and Pol alignment in windows of 300 residues moved in steps of 5 to 100 amino acids (larger step sizes were used to screen for recombination; smaller step sizes were used to map recombination breakpoints). To test for recombination among the three newly derived SIVgor isolates, full-length nucleotide sequences of the three SIVgor genomes were aligned and tested using GENECONV (http://www.math.wlu.ca/~sawyer/).

Infectivity, coreceptor usage, and neutralization phenotype analyses. Full-length molecular clones of SIVgor, SIVcpz, and HIV-1 were transfected into 293T cells and supernatants equilibrated by particle-associated reverse transcriptase activity as described previously (65). These included the SIVcpzPrs clones TAN1, TAN2, and TAN3 (65); the SIVcpzPrb clones GAB2 (11), MT145, MB987, and EK505 (I. Decker and B. H. Hahn, unpublished); and the HIV-1 clones SG3 (28), NL4-3 (1), YU2 (37), and WEAU1.6 (18). In addition, human peripheral blood mononuclear cells (PBMC)-derived viral isolates YBF30 (58) and 97US08692A (78) were also used. Viral infectivity was assessed for JCS3-BL-13 cells (TZM-bl; NIH AIDS Research and Reference Reagent Program catalogue no. 8129), a HeLa-derived line which has been genetically modified to constitutively express CD4, CCR5, and CXCR4 and to contain integrated luciferase and β-galactosidase reporter genes under the control of an HIV-1 LTR (21, 77). For coreceptor analysis, JCS3-BL cells were seeded in 96-well plates at 8,300 cells/well overnight and then treated with the CCR5 antagonist TAK-779 (10 mM) (4), the CXCR4 antagonist AMD3100 (1.2 mM) (20), or a combination of both for 1 hour (NIH AIDS Research and Reference Reagent Program catalogue no. 4983 and 8128, respectively). Virus was added in the presence of 40 µg/ml DEAE-dextran and removed 48 h later. Cells were then lysed and analyzed for luciferase activity (Promega, Madison, WI), using a Tropix luminometer with Windows version 1.24 software.

For neutralization assays, 3,000 infectious units of virus were combined in a total volume of 60 µl with or without a 1× concentration of sCD4 in Dulbecco's modified Eagle's medium (DMEM) with 6% fetal calf serum (FCS) and 80 µg/ml DEAE-dextran. After 1 h at 37°C, an equal volume of human plasma (10% [vol/vol] in DMEM plus 6% FCS or fivefold dilutions thereof), monoclonal
antibodies, or fusion inhibitor was added. Monoclonal antibodies described in reference 12 were kindly provided by the following individuals: Dennis Burton (b12 and 2G12), Michael Zwick and Dennis Burton (Z13e1), Herman Katunger (2FS and 4E10), Susan Zolla-Pazner (447-52D), Lisa Cavacini (F425-B4e8), James Robinson (17b, 19c, and 21c), and David Montefiori (HIVIG). The following reagents were obtained commercially: soluble CD4 (514-CD, R&D Systems), T1249 (Triangle Pharmaceuticals), and anti-CD4 monoclonal antibody (catalogue no. 555344, clone RPA-T4; BD Pharmingen). The addition of ligand or antibody brought the final concentration of DEAE dextran to 40 μg/ml. Virus was incubated with test antibodies (with or without sCD4) for 1 h at 37°C and then added to JCS3-BL cells. Cells were incubated at 37°C for 2 days and then analyzed for lucerase expression. Controls included cells exposed to no virus or to virus pretreated with normal human plasma. Relative infectivity was calculated by dividing the number of lucerase units at each dilution of test plasma on monoclonal antibodies by values in wells containing normal human plasma. Neutralization was assessed by 50% inhibitory concentration (IC50) determined by linear regression using a least-squares method. All samples were tested in duplicate.

Chimpanzee and human PBMC cultures. Blood was obtained from normal human volunteers (Research Blood Components, Boston, MA) as well as healthy (HIV-1-uninfected) chimpanzees housed at the Yerkes Regional Primate Center as described previously (65). Briefly, PBMCs were isolated using Ficoll Hypaque Plus (GE Healthcare, Piscatway, NJ). CD4+ T cells were enriched using CD4 microbeads and magnetic cell sorting (Miltenyi Biotec, Auburn, CA), stimulated with staphylococcal enterotoxin B (Sigma-Aldridge, St. Louis, MO) for 12 to 15 h, and then washed three times and plated in 24-well plates in DMEM containing 10% FCS and 30 U/ml IL-2. After 24 h, 5 x 105 CD4+ T cells were incubated with transfection-derived viral stocks at a multiplicity of infection of 0.1 (as determined for JC53-BL cells) in 300 μl DMEM containing 10% FCS and 30 U/ml IL-2, and reverse transcriptase activity in culture supernatants was measured every 3 days to monitor viral replication.

Nucleotide sequence accession numbers. The GenBank accession numbers for the sequences determined in this study are FJ424871 (CP684 consensus sequence), FJ424863 (CP2135 consensus sequence), FJ424864 (CP2139.1 consensus sequence), FJ424865 (CP2139.2 consensus sequence), FJ424866 (pBR-CP2139.287), FJ424867 (BQ664 vif-env), FJ424868 (BQ664 gag), FJ424869 (CP1973 pol), and FJ424870 (CP2680 pol).

RESULTS

SIVgor genome structure. To amplify complete SIVgor genome sequences, we selected fecal samples from four infected western lowland gorillas, all of which were sampled in the wild. Two of these (CP684 and BQ664) were previously shown to harbor divergent SIVgor strains based on subgenomic pol and gp41 sequences (71). The other two (CP2135 and CP2139) were identified more recently, in a follow-up study of gorilla communities at the CP field site (C. Neel and M. Peeters, unpublished). All samples were positive for SIVgor antibodies, and their species origins were determined by mitochondrial DNA analysis. All samples were also shown to represent different individuals, based on microsatellite analysis (not shown). Fecal RNA was extracted and subjected to RT-PCR analysis using consensus sequence- as well as strain-specific primers (see Table S1 in the supplemental material). For CP684, CP2135, and CP2139, this approach yielded 12 to 15 partially overlapping fragments, which in each case spanned an entire provirus (Fig. 2A). Amplification of SIVgor sequences from sample BQ664 was more difficult. Despite repeated RT-PCR attempts and multiple primer combinations, only two new fragments in gag and the accessory gene region could be amplified (Fig. 2A). All amplicons were sequenced directly to generate fecal consensus sequences. As expected, these contained a limited number of ambiguous sites, including base mixtures and sequence differences in regions of fragment overlap (see Table S2 in the supplemental material). Although some of the ambiguous sites affected the encoded amino acid, none introduced stop codons or frameshift mutations.

The concatenated genomes of CP2139, CP2135, and CP684 were 9,252, 9,246, and 9,143 nucleotides in length, respectively. All three carried full-length open reading frames for all nine genes found in SIVcpz/HIV-1 (Fig. 2A) as well as all major regulatory sequences. Although the deduced SIVgor, SIVcpz, and HIV-1 protein sequences varied considerably, we found no obvious SIVgor-specific signatures. For example, all three gorilla viruses encoded additional cysteine residues in their fourth variable (V4) envelope domain (two in CP2139 and four in CP2135 and CP684), indicating a more diversified V4 loop structure; however, this was not unique to SIVgor, since additional cysteine residues (usually two) were also observed in the V4 domain of HIV-1 group O and some SIVcpz strains (not shown). Similarly, SIVgor Vpu sequences differed from those of SIVcpz and HIV-1 (including group O) in more than 60% of amino acid residues; however, the predicted secondary structure and hydrophobicity profiles of the deduced SIVgor Vpu proteins suggested a very similar function (not shown). Interestingly, there were some protein domains where SIVgor strains resembled chimpanzee and not human viruses. At position 30 of the Gag protein matrix, the three gorilla viruses (like all strains of SIVcpz and HIV-1) encoded a methionine rather than the arginine found in the ancestors of HIV-1 groups M, N, and O (75). Similarly, in a phenogram of V3 loop sequences, all gorilla viruses clustered with SIVcpz strains (Fig. 3), rather than grouping with HIV-1 group O viruses to which they are much more closely related (see below). Thus, at two sites known to be under host-specific selection pressures in humans and chimpanzees (70, 75), SIVgor resembled SIVcpz rather than HIV-1.

Origin of SIVgor. To compare the evolutionary relationships of the newly derived gorilla viruses to each other and to previously characterized SIVcpz and HIV-1 strains, phylogenetic trees were constructed from different regions of the proteome. For the fully sequenced CP2139, CP2135, and CP684 strains, these included Gag, the N-terminal half of Pol (Pol1), the C-terminal half of Pol plus Vif (Pol2), and Env (Fig. 4). For the partially characterized BQ664 strain, individual Gag, Pol, Vif-Env, and gp41 fragments were concatenated prior to analysis (Fig. 5). These studies confirmed and extended previous results from subgenomic SIVgor sequences (71). In all regions of their genome, the gorilla viruses clustered in a monophyletic clade that was most closely related to HIV-1 group O viruses. Within the SIVgor lineage, strains from the CP field site were much more closely related to one another than they were to BQ664. The branching orders of the three CP strains varied among trees from different parts of the genome, but these relationships were not significantly discordant, and closer examination of the nucleotide alignments of the CP strain sequences (using GENECONV) detected no evidence for recombination. Finally, the SIVgor/HIV-1 group O clade was significantly more closely related to SIVcpzPrt from central chimpanzees than to SIVcpzPts from eastern chimpanzees, strongly suggesting a west central African origin for SIVgor.

The phylogenetic relationships in Fig. 4 were derived from
large genomic regions, and so, to check for any evidence of recombination involving smaller fragments, we made numerous trees based on shorter (300 amino acids) windows from the alignments. This analysis identified a 900-bp region at the 5' end of the pol gene that appeared to have a different evolutionary history. While a tree derived from Pol sites 301 to 699 (from a gap-stripped Pol alignment) exhibited the conventional topology with SIVgor and HIV-1 group O viruses clustering exclusive of SIVcpzPtt strains and other HIV-1 groups (Fig. 6, right), in a tree of N-terminal Pol sequences (sites 1 to 300), this was no longer true (Fig. 6, left). Support for this discordant branching was high: the grouping of the SIVcpzPtt strains MT145 and GAB2 together with SIVgor/group O outside the remaining SIVcpzPtt strains had a posterior probability value of 98%. These results indicated that recombination had occurred during the divergence of these sequences.

Two different evolutionary scenarios could explain the discordant branching orders in Fig. 6. One possibility is that an ancestor of SIVgor acquired sequences from an ancestral SIVcpzPtt strain; in this case, SIVgor is recombinant. Alternatively, an ancestor of MT145 and GAB2 could have acquired sequences from a divergent SIVcpzPtt lineage that has not yet been sampled or is now extinct; in this case, MT145 and GAB2 are recombinant. To distinguish between these scenarios, we compared the divergences of the various strains in recombinant and nonrecombinant regions. Under the first scenario, the SIVgor/group O lineage would be expected to change its position in the two trees, i.e., this lineage would move inside the SIVcpzPtt clade. As a result, the common ancestor of the entire SIVcpzPtt/SIVgor/HIV-1 clade in the recombinant region (the red node in Fig. 6) would be equivalent, in terms of distance from the root or the tips of the tree, to the ancestor of the SIVcpzPtt/HIV-1 clade in the nonrecombinant region (the blue node in Fig. 6). Alternatively, if MT145 and GAB2 were recombinant, the red node would be equivalent to the ancestor of the entire SIVcpzPtt/SIVgor/HIV-1 clade in the nonrecombinant region (the green node in Fig. 6). To evaluate these alternative hypotheses, we calculated the evolutionary distances from each of these highlighted nodes to the tips of the four strains in the SIVcpzPtt clade (Table 1). All distances were measured as average branch lengths, and this was done independently for Gag (Fig. 4) as well as the recombinant (Fig. 6, left) and nonrecombinant (Fig. 6, right) portions of Pol. In order to correct for different rates of evolution in different proteins, each distance was divided by the total length of the tree for that genomic region. Finally, all HIV-1 strains and one SIVcpzPtt recombinant (LB7) were excluded. The results showed that the distance for Pol sites 1 to 300 (for the red node) was similar to those for the green node (and lower than those for the blue node) for Gag and Pol sites 301 to 699 (Table 1), indicating that SIVcpzPtt strains MT145 and GAB2 were recombinant. Although these two SIVcpzPtt strains did not form a statistically supported clade in the recombinant region (Fig. 6, left), the shared mosaicism implied that they were monophyletic. This finding was also supported by results from an analysis of internal branch lengths (not shown) and is consistent with our previous conclusion that GAB2 is recombinant and contains pol gene sequences from a divergent
SIVcpzPtt strain (11). Since it is clearly most parsimonious to assume that the recombination event in the ancestry of MT145 and GAB2 occurred during infection of a central chimpanzee, this implies that SIVgor and HIV-1 group O viruses are derived from a divergent SIVcpzPtt lineage that existed (and possibly still exists) in west central Africa.

Chimpanzee and gorilla viruses from the same field site. The finding of divergent SIVgor strains at two field sites located 400 km apart (CP and BQ) (Fig. 1) suggested that wild-living gorillas are endemically infected with SIVgor (71). However, the possibility that local chimpanzee-to-gorilla transmissions had generated the observed SIVgor diversity could not be formally excluded. To examine whether neighboring ape communities harbored epidemiologically linked viruses, we specifically targeted wild-living chimpanzees in one area of high gorilla density (39). Screening 77 chimpanzee fecal samples from the CP field site, we found 9 to be SIVcpz antibody positive. These were identified to represent two naturally infected P. t. troglodytes apes (CP1973 and CP2680) that were sampled 8 months apart. Although the results for initial RT-PCR analyses were negative for diagnostic pol and gp41 regions (33, 70), most likely because of sample degradation, a small, 220-bp pol fragment was eventually amplified. Phylogenetic analysis of the resulting sequences confirmed that the two chimpanzees were infected with distinct SIVcpzPtt strains which were only distantly related to SIVgor
strains identified in the same geographic area (Fig. 7). While we have analyzed only two chimpanzees from the CP site, previous surveys have shown that SIVcpz\textsubscript{Ptt} strains exhibit strong phylogeographic clustering (33, 70). The distant relationship between the chimpanzee and gorilla viruses from CP thus indicates that these SIVgor strains were not due to a local transmission from chimpanzees. Furthermore, all SIVgor strains form a single cluster, strongly suggesting that they all descended from the same chimpanzee-to-gorilla transmission. Thus, the diversity within the SIVgor clade most likely reflects viral evolution in gorillas.

**Timing of the MRCA of SIVgor.** The results described above suggest that all currently known SIVgor strains are derived from a single introduction of SIVcpz\textsubscript{Ptt} into western lowland gorillas. To gauge when this introduction might have occurred, we estimated the time to the MRCA of the SIVgor clade. It is not yet possible to obtain a reliable estimate of the rate of SIVgor evolution, because only a few strains have thus far been characterized, and all were isolated within a narrow time period (2004 to 2007). We therefore used an evolutionary rate estimated from closely related HIV-1 group O strains. A phylogeny was obtained from a nucleotide sequence alignment produced by concatenating the four genomic regions available for BQ664, and distances between sequences were calculated by summing branch lengths from that tree. The average estimated distance between BQ664 and the CP strains was 0.295 substitutions per site. The year of the MRCA of HIV-1 group O has been estimated as 1920, with a confidence (95% highest posterior density) interval of 1896 to 1942 (35). With these dates, the rate of nucleotide substitution for HIV-1 group O

![FIG. 5. Evolutionary relationships of newly derived SIVgor strains in partial genome regions. Partial Gag, Pol, Vpr-Vpu, and Env sequences corresponding to BQ664 were concatenated. SIVgor sequences are highlighted in red; HIV-1, SIVcpz\textsubscript{Ptt}, and SIVcpz\textsubscript{Pts} strains are as in Fig. 4. Numbers on internal branches indicate estimated posterior probabilities (only values above 90% are shown). The scale bars represent 0.1 replacements per site.](image)

![FIG. 6. Evidence for an SIVcpz\textsubscript{Ptt} lineage ancestral to SIVgor and HIV-1 group O. Trees were inferred for two parts of the Pol1 region in Fig. 4, separated at a putative recombination breakpoint at position 300. Strains are color coded as in Fig. 4. Numbers on internal branches indicate estimated posterior probabilities (only values above 95% are shown). The scale bars represent 0.05 replacements per site. Green, blue, and red circles indicate nodes that were used to calculate the distances shown in Table 1.](image)

<table>
<thead>
<tr>
<th>Node</th>
<th>Gag</th>
<th>Pol sites 1–300</th>
<th>Pol sites 301–699</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.27</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>0.22</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Red</td>
<td>0.22</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

*Distances were calculated as average branch lengths from the tips of each of the four SIVcpz\textsubscript{Ptt} strains to the relevant nodes depicted in Fig. 6.*
for the genomic regions used here was estimated to be 1.05 × 10^{-3} (range, 0.79 × 10^{-3} to 1.50 × 10^{-3}) per site per year. Then, assuming that SIVgor strains have evolved at the same rate as group O viruses, their MRCA was estimated to have existed around 1864 (range, 1818 to 1906). Thus, the chimpanzee-to-gorilla transmission that spawned the current SIVgor infections seems to have occurred at least 100 to 200 years ago.

Generation and biological characterization of a replication-competent SIVgor clone. We have previously shown that replication-competent molecular clones of SIVcpz can be derived from fecal viral consensus sequences (65). However, for this strategy to be successful, we found that it was critical to amplify the most predominant viral species in the fecal sample. To ensure that this was the case for SIVgor, we selected one of the three fully sequenced SIVgor strains (CP2139) and reamplified its entire genome by using reextracted fecal RNA as well as strain-specific primers (Fig. 2B). As expected, the second consensus sequence (CP2139.2) was not identical to the first (CP2139.1); there were 54 nucleotide sequence differences, the majority of which were located in the env gene. Table S3 in the supplemental material lists all sequence differences between CP2139.1 and CP2139.2 and indicates which nucleotides were selected for inclusion in the molecular clone.

The CP2139 master sequence was used as the template to chemically synthesize three subgenomic fragments. These were then ligated into a low-copy-number (pBR322-derived) plasmid to generate the proviral clone CP2139.287 (Fig. 2B). To examine its biological activity, CP2139.287 was transfected into 293T cells and the resulting supernatant tested for infectivity in the JC53-BL cell assay. This analysis showed that CP2139.287-derived virus was infectious (600 to 800 IU/μg reverse transcriptase) and entered JC53 cells via the CCR5 coreceptor. As shown in Fig. 8, infectivity of CP2139.287 was completely blocked by the CCR5 antagonist TAK-779 but not by the CXCR4 antagonist AMD3100. This was also true for the primary HIV-1 strains YU-2 (group M), YBF30 (group N), and 97US08692A (group O) but not for NL4-3- and WEAU 1.6-derived viruses, which served as X4 and R5/X4 dual tropic...
controls, respectively (Fig. 8). Culture experiments also showed that CP2139.287-derived virus replicated efficiently and to high titers in both human and chimpanzee CD4+ T cells, with kinetics very similar to those of previously characterized SIVcpz and HIV-1 strains (Fig. 9). Taken together, these data demonstrated that the reconstructed CP2139.287 clone carried a fully functional, replication-competent SIVgor genome.

To examine the antigenic properties of SIVgor, we tested CP2139.287-derived virus as well as panel of HIV-1 group M strains. 97US08692A and YBF30 are primary HIV-1 isolates. a Neutralization titers are expressed as the reciprocal 50% inhibitory plasma dilutions or the IC50 for the respective Env ligands in one representative experiment (performed in duplicate). All titers are given in μg/ml unless otherwise stated.

Table 2. Neutralization phenotype of SIVgor

<table>
<thead>
<tr>
<th>Virus</th>
<th>sCD4 (nM)</th>
<th>HIVIG</th>
<th>Clade B plasma pool</th>
<th>Clade C plasma pool</th>
<th>Group O plasma</th>
<th>Group O plasma</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIVgor (CP2139.287)</td>
<td>154</td>
<td>&gt;1000</td>
<td>20</td>
<td>&lt;20</td>
<td>&lt;20</td>
<td>&lt;20</td>
</tr>
<tr>
<td>HIV-1_O (97US08692A)</td>
<td>21</td>
<td>&gt;1000</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>&lt;50</td>
</tr>
<tr>
<td>HIV-1_N (YBF30)</td>
<td>495</td>
<td>&gt;1000</td>
<td>&lt;20</td>
<td>&lt;20</td>
<td>&lt;20</td>
<td>&lt;20</td>
</tr>
<tr>
<td>HIV-1_M (NL4.3)</td>
<td>2</td>
<td>71</td>
<td>630</td>
<td>358</td>
<td>83</td>
<td>&lt;20</td>
</tr>
<tr>
<td>HIV-1_M (YU2)</td>
<td>14</td>
<td>891</td>
<td>178</td>
<td>267</td>
<td>&lt;20</td>
<td>&lt;20</td>
</tr>
</tbody>
</table>

*Neutralization titers are expressed as the reciprocal 50% inhibitory plasma dilutions or the IC50 for the respective Env ligands in one representative experiment (performed in duplicate). All titers are given in μg/ml unless otherwise stated. HIVIG, pooled human immunoglobulin purified from plasma samples of healthy HIV-1-infected individuals.*

DISCUSSION

A primary objective of this study was to determine the origin of SIVgor and to estimate where and when this virus might have been introduced into wild-living gorillas. By sequencing three new full-length SIVgor genomes and conducting detailed recombination analyses, we found evidence for an ancestral SIVcpzPtt lineage from which SIVgor and HIV-1 group O viruses evolved (Fig. 6). Although full-length representatives of this lineage were not identified, we documented its existence in the form of mosaic pol fragments in present-day SIVcpzPtt recombinants. This finding strongly suggests that P. t. troglodytes apes were the original source of SIVgor and that the cross-species transmission took place in the Cameroon/Equatorial Guinea/Gabon area of west central Africa (Fig. 1). To estimate when this event might have occurred, we calculated the time to the MRCA of the current SIVgor clade. This analysis yielded an estimate of 100 to 200 years for the divergence of BQ664 from the three other SIVgor strains (Fig. 5). The transmission that gave rise to SIVgor likely occurred much earlier than that since (i) currently available SIVgor sequences may not represent the entire diversity of extant gorilla viruses, (ii) SIVgor strains that diverged earlier may have gone extinct, and (iii) phylogenetic methods are notorious for underestimating the deeper divergence times of rapidly evolving RNA viruses (31, 56). Once introduced, SIVgor spread within its new host, as evidenced by the presence of viruses at field sites located 400 km apart (Fig. 1). Whether this occurred primarily by sexual or other (e.g., exposure to infectious saliva) routes is not known; however, it is likely that gorilla behavior and social structure facilitated virus dispersal. Western lowland gorillas live in family units of 2 to 30 individuals which typically comprise one dominant male (silverback), three or more sexually active adult females, their offspring, and possibly a few non-dominant males (27, 42, 62, 67). The silverback mates with all active adult females, their offspring, and possibly a few non-dominant males (27, 42, 62, 67). Silverbacks transfer between groups (63). Individual groups have home ranges of about 15 to 20 km² (8, 22, 47), which often overlap, resulting in frequent encounters of neighboring groups (8, 17, 22, 67). Thus, gorilla mating and social networks provide ample opportunity for a newly introduced SIV to spread, both within and between neighboring communities.

As illustrated in Fig. 4, chimpanzees have transmitted SIVcpz to gorillas and humans on multiple occasions. For humans, exposure to infected chimpanzee blood or mucosal
TABLE 2—Continued

<table>
<thead>
<tr>
<th>Neutralization titer</th>
<th>Coreceptor used</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>R5</td>
</tr>
<tr>
<td>0.02</td>
<td>R5</td>
</tr>
<tr>
<td>0.01</td>
<td>R5</td>
</tr>
<tr>
<td>0.17</td>
<td>R5</td>
</tr>
<tr>
<td>0.26</td>
<td>R5</td>
</tr>
<tr>
<td>0.1</td>
<td>0.034</td>
</tr>
<tr>
<td>0.17</td>
<td>0.007</td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

secrections in the context of bushmeat hunting is considered the most plausible scenario for cross-species transmission (30). For gorillas, the route and circumstances of transmission are much less clear. Gorillas are herbivores and do not hunt other mammals (41, 49, 68). However, recent studies have shown that gorillas and chimpanzees commonly feed in the same forest areas and sometimes even in the same fruiting trees (76). Thus, there are focal points for gorilla/chimpanzee encounters that may facilitate virus transmission. Moreover, SIV is not the only virus requiring physical contact for infection that has crossed the species barrier from chimpanzees to gorillas: screening ape fecal samples for hepatitis B virus DNA, we found that this pathogen has been transmitted from wild-living chimpanzees to gorillas on at least two occasions (W. Liu and B. H. Hahn, unpublished). Since hepatitis B virus, like HIV/SIV, is transmitted by direct contact with infectious blood or mucosal secretions (25), encounters between chimpanzees and gorillas that are conducive to cross-species infection must occur. Whether these are physical in nature (i.e., fighting or biting), involve exposure to infectious feces or urine in cofeeding areas, or occur through saliva in partially eaten, discarded fruit (76) will need to be determined. Whatever the circumstances, it appears that successful transmission events are rare.

While it seems clear that chimpanzees were the source of SIVgor, it is not known whether humans acquired HIV-1 group O from naturally infected chimpanzees or gorillas. Thus far, none of the chimpanzee and gorilla communities tested represent likely reservoirs. The known SIVgor strains are too divergent to have been the immediate source of HIV-1 group O (Fig. 4). In addition, the prevalence of SIVgor infection in Cameroon is very low (C. Neel and M. Peeters, unpublished). Finally, none of several hundred P. t. troglodytes apes tested in southern Cameroon, including those at the CP field site (Fig. 8), harbor SIVgor-like viruses (33, 70). Together, these data strongly suggest that the ape reservoir that gave rise to HIV-1 group O exists outside Cameroon. It will be important to determine whether this reservoir still exists and, if so, where it is located. Moreover, it will be important to differentiate between different transmission scenarios (Fig. 10). One possibility is that chimpanzees harboring SIVgor-like viruses infected gorillas and humans independently; identification of SIVcpz Ptt strains that join either the branch outside the HIV-1 group O or that outside the SIVgor clade would provide strong support for such a scenario (green lines in Fig. 10A). Alternatively, gorillas may have served as an intermediary host for the human infection; evidence for this would come from the finding of SIVgor strains that join either the branch outside group O or that outside the entire HIV-1 group O/SIVgor clade (Fig. 10B). The viruses invoked in either of these scenarios may exist in as-yet-unsampled apes in Equatorial Guinea, Gabon, or the Republic of the Congo. Indeed, the relative prevalence of group O infections has been reported to be much higher in Equatorial Guinea (9% of HIV-1 infections) than in Cameroon (2% of HIV-1 infections), pointing to this area as a possible starting point of the HIV-1 group O epidemic (3, 15, 43, 74, 80). Formally, there is a third possibility, namely, that chimpanzees harboring SIVgor-like viruses first infected humans, who then passed the virus to gorillas. However, this would imply that HIV-1 group O was far more divergent in the past than is currently the case (Fig. 10C). Moreover, it is hard

FIG. 10. Possible scenarios for the origin of HIV-1 group O. (A) Chimpanzees harboring SIVgor-like viruses infected gorillas and humans independently. (B) Chimpanzees harboring SIVgor-like viruses infected gorillas, which then passed the virus on to humans. (C) Chimpanzees harboring SIVgor-like viruses infected humans, who then passed the virus to gorillas. Support for each scenario would come from the discovery of one of the hypothetical viruses shown in green. In all three trees, nodes marked by circles indicate ancestral viruses that are assumed to have infected chimpanzees.
to imagine how humans could transmit a virus such as SIVgor to wild-living gorillas. Thus, the third scenario is implausible.

A second objective of this study was to generate a replication-competent molecular clone of SIVgor for biological analyses. We accomplished this by synthesizing the complete genome of one of the newly characterized gorilla viruses (CP2139) from fecal consensus sequences. The availability of this SIVgor clone allowed us to compare its baseline biological properties to those of SIVcpz and HIV-1 reference strains. As shown in Fig. 9, SIVgor replicated to high titers in CD4+ T cells from human (n = 4) and chimpanzee (n = 3) donors. We also assessed the sensitivities of SIVgor to a number of Env-specific ligands (Table 2). The results revealed that SIVgor was sensitive to neutralization by the membrane-proximal external region monoclonal antibody 2F5, consistent with resistance to the V3 monoclonal antibodies 447-52D and F425-B4e8, both of which recognize a highly conserved arginine residue at position 315 in the V3 loop of SIVcpz. However, the accessibility of V3 to antibodies has been shown to have a major impact on V3-mediated neutralization. The fact that an HIV-1 group O plasma (97US08692A) that encodes a similarly divergent V3 loop (GPMΔ) was resistant to entry inhibition by a monoclonal antibody against surface CD4 and was moderately sensitive to 447-52D (Table 2) suggests that the SIVgor V3 loop, like its CD4i bridging sheet, may be concealed in the functional Env trimmer. Finally, we found that SIVgor was resistant to neutralization by heterologous patient antibodies, including two group O plasmas. Altogether, these findings suggest that the sensitivity of SIVgor to Env-specific ligands resembles that of primary HIV-1 and SIVcpz strains in several ways, including those related to CD4 tropism, CCR5 coreceptor preference, and effective concealment of CD4i (and possibly also V3 epitopes) in the functional Env trimmer.

Given the biological properties of SIVgor and the extent of ape bushmeat hunting in west central Africa (14, 39, 69), it is clear that wild-living gorillas could represent a reservoir for human infection. Additional field studies are thus needed to determine the prevalence, geographic distribution, species association, and natural history of SIVgor throughout the entire gorilla habitat. Given the current rate of deforestation in central Africa and the extent of bushmeat hunting and consumption, it is critical to collect baseline data on existing sources of human zoonotic diseases now. This approach has been successful for Ebola hemorrhagic fever (36), monkeypox (46), and anthrax (34). Continued monitoring of wild ape populations is thus critical, not only to identify the primate origin of HIV-1 group O but also to ensure that additional reservoirs of HIVs are not overlooked.

ACKNOWLEDGMENTS

We thank Caroline Tutin, Peter Walsh, Kate Abernethy, and Bing Chen for helpful discussions; the Cameroon National Institute of Health, Environment and Forestry, and Research for permission to perform this study; the staff from the PRECIGA project for logistical support; the Yerkes Primate Center staff for shipping blood samples from wild-caught captive chimpanzees; Maria Salazar for technical assistance; and Jamie C. White for artwork and manuscript preparation. This work was supported in part by the National Institutes of Health (R37 AI50529, R01 AI87715, P30 AI27767, and P30 CA13148), Agence National de Recherches sur le SIDA (ANRS, France (ANRS 12125), the Institut de Recherche pour le Développement (IRD), the Yerkes Regional Primate Research Center (RR-00165), and the Bristol Myers Freedom to Discover Program.

REFERENCES
