

## Generation of Infectious Molecular Clones of Simian Immunodeficiency Virus from Fecal Consensus Sequences of Wild Chimpanzees<sup>∇</sup>

Jun Takehisa,<sup>1</sup> Matthias H. Kraus,<sup>1</sup> Julie M. Decker,<sup>1</sup> Yingying Li,<sup>1</sup> Brandon F. Keele,<sup>1</sup> Frédéric Bibollet-Ruche,<sup>1</sup> Kenneth P. Zammit,<sup>1</sup> Zhiping Weng,<sup>1</sup> Mario L. Santiago,<sup>2</sup> Shadrack Kamenya,<sup>3</sup> Michael L. Wilson,<sup>3</sup> Anne E. Pusey,<sup>4</sup> Elizabeth Bailes,<sup>5</sup> Paul M. Sharp,<sup>5</sup> George M. Shaw,<sup>1,6</sup> and Beatrice H. Hahn<sup>1,6\*</sup>

Departments of Medicine<sup>1</sup> and Microbiology,<sup>6</sup> University of Alabama at Birmingham, Birmingham, Alabama 35294; Gladstone Institute of Virology and Immunology, San Francisco, California 94158<sup>2</sup>; Gombe Stream Research Centre, The Jane Goodall Institute, Kigoma, Tanzania<sup>3</sup>; Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, Minnesota 55108<sup>4</sup>; and Institute of Genetics, University of Nottingham, Queens Medical Centre, Nottingham, United Kingdom<sup>5</sup>

Received 15 March 2007/Accepted 30 April 2007

**Studies of simian immunodeficiency viruses (SIVs) in their endangered primate hosts are of obvious medical and public health importance, but technically challenging. Although SIV-specific antibodies and nucleic acids have been detected in primate fecal samples, recovery of replication-competent virus from such samples has not been achieved. Here, we report the construction of infectious molecular clones of SIVcpz from fecal viral consensus sequences. Subgenomic fragments comprising a complete provirus were amplified from fecal RNA of three wild-living chimpanzees and sequenced directly. One set of amplicons was concatenated using overlap extension PCR. The resulting clone (TAN1.24) contained intact genes and regulatory regions but was replication defective. It also differed from the fecal consensus sequence by 76 nucleotides. Stepwise elimination of all missense mutations generated several constructs with restored replication potential. The clone that yielded the most infectious virus (TAN1.910) was identical to the consensus sequence in both protein and long terminal repeat sequences. Two additional SIVcpz clones were constructed by direct synthesis of fecal consensus sequences. One of these (TAN3.1) yielded fully infectious virus, while the second one (TAN2.69) required modification at one ambiguous site in the viral *pol* gene for biological activity. All three reconstructed proviruses produced infectious virions that replicated in human and chimpanzee CD4<sup>+</sup> T cells, were CCR5 tropic, and resembled primary human immunodeficiency virus type 1 isolates in their neutralization phenotype. These results provide the first direct evidence that naturally occurring SIVcpz strains already have many of the biological properties required for persistent infection of humans, including CD4 and CCR5 dependence and neutralization resistance. Moreover, they outline a new strategy for obtaining medically important “SIV isolates” that have thus far eluded investigation. Such isolates are needed to identify viral determinants that contribute to cross-species transmission and host adaptation.**

Of all human infectious diseases that have emerged in recent history, AIDS is one of the most devastating. With more than 65 million people infected and over 25 million deaths, human immunodeficiency virus type 1 (HIV-1) represents one of the most serious public health threats of the 21st century (11). There are three major variants of HIV-1, termed groups M, N, and O, which are the result of independent cross-species transmission events of simian immunodeficiency viruses (SIVcpz/SIVgor) that naturally infect chimpanzees (*Pan troglodytes troglodytes*) and gorillas (*Gorilla gorilla gorilla*) in west central Africa (22, 31, 51, 54). Although each cross-species transmission generated an AIDS-causing pathogen, the subsequent spread of these infections in the human population has been variable (22, 51, 58). Group M is the cause of the vast majority of HIV-1 infections worldwide and the source of the AIDS pandemic (9, 23, 35). In contrast, group N and O infections

account for a much smaller fraction of AIDS cases and are restricted to west central Africa (1, 7, 44, 61, 62). To elucidate the circumstances of their emergence, recent efforts have focused on characterizing their natural ape reservoirs. These studies have traced the closest SIV relatives of HIV-1 group M to geographically isolated chimpanzee (*P. t. troglodytes*) communities in southeastern Cameroon (31). The closest SIV relatives of HIV-1 group N have also been identified in wild-living chimpanzees, but in *P. t. troglodytes* communities in south-central Cameroon (31). Finally, HIV-1 group O-like viruses have been identified in western lowland gorillas (*G. g. gorilla*) (54). Since these gorilla viruses are also of *P. t. troglodytes* origin, gorillas may have served as an intermediary host for the human infection. Alternatively, *P. t. troglodytes* apes harboring group O-like viruses may have been the source of independent gorilla and human infections (54).

Previous studies have classified common chimpanzees into four subspecies on the basis of differences in mitochondrial DNA sequences, including *P. t. verus* in west Africa; *P. t. vellerosus* in Nigeria and northern Cameroon; *P. t. troglodytes* in southern Cameroon, Gabon, Equatorial Guinea, and the Republic of Congo; and *P. t. schweinfurthii* in the Democratic

\* Corresponding author. Mailing address: Department of Medicine, University of Alabama at Birmingham, 720 20th Street South, Kaul 816, Birmingham, AL 35294. Phone: (205) 934-0412. Fax: (205) 934-1580. E-mail: bhahn@uab.edu.

<sup>∇</sup> Published ahead of print on 9 May 2007.

Republic of Congo, Uganda, Rwanda, Burundi, and Tanzania (10). Although the validity of this taxonomy has recently been questioned by primatologists (17), the differentiation of chimpanzees into geographically isolated populations has been of use to virologists. For example, of the four previously defined subspecies only *P. t. troglodytes* and *P. t. schweinfurthii* apes are naturally infected with SIVcpz (51). Moreover, their viruses fall into two divergent lineages (SIVcpzPtt and SIVcpzPts) that differ in up to 40% of their nucleotide sequences and exhibit lineage-specific protein signatures (46, 48). This extent of genetic diversity is comparable to that observed for SIVs from related monkey species (2, 24). However, the divergence of SIVcpzPtt and SIVcpzPts strains does not reflect different origins. Evolutionary studies have shown that SIVcpz has a mosaic genome structure comprising SIV lineages infecting monkey species on which chimpanzees prey (3). Since all known SIVcpz strains exhibit the same mosaic genome structure, it is clear that they share a common ancestry. The observed diversity between the SIVcpzPtt and SIVcpzPts lineages must therefore reflect their evolution in geographically isolated hosts. Thus, until a new taxonomy is formally adopted, we will use the existing subspecies designations to describe SIVcpz infection in distinct chimpanzee populations.

Although the origins of the AIDS pandemic have now been determined, important questions remain concerning the circumstances of the cross-species transmission events and the factors that contributed to HIV-1's emergence. In particular, it is perplexing that only certain strains of SIVcpzPtt, but not of SIVcpzPts, have been found in humans. Recent surveys of wild-living *P. t. schweinfurthii* apes in the Democratic Republic of Congo have shown that SIVcpzPts is as widely distributed as SIVcpzPtt, with high levels of endemicity in some communities and rare or absent infection in others (48, 59; Y. Li and B. H. Hahn, unpublished). Thus, the lack of SIVcpzPts-like viruses in humans cannot be explained by a paucity of naturally occurring SIVcpzPts infections. It is possible that environmental factors and/or human behavior may have rendered SIVcpzPts transmissions less likely. It is also possible that SIVcpzPts transmissions have occurred but subsequently gone extinct. However, given the genetic diversity of SIVcpzPts and SIVcpzPtt strains, viral properties may also have played a role. The latter possibility is testable but requires a panel of diverse SIVcpzPtt and SIVcpzPts isolates for detailed biological characterization. To date, only five SIVcpz isolates and one infectious molecular clone have been obtained, all of which were derived by *in vitro* cultivation (13, 38, 40, 41). Four of these (GAB1, CAM3, CAM5, and CAM13) are from *P. t. troglodytes* apes, while the fifth represents the only SIVcpzPts (ANT) isolate (51). The single infectious molecular clone of SIVcpz (GAB1) was derived from a productively infected human lymphocyte culture (28). Given that all existing SIVcpz isolates have been adapted to growth in human T cells, they are not suitable for studies aimed at identifying viral determinants of host adaptation.

Given the highly endangered status of chimpanzees and gorillas throughout equatorial Africa, obtaining blood samples for virus isolation studies from wild-living apes is neither ethical nor practical. We recently developed noninvasive methods that permit the detection of virus-specific antibodies and nucleic acids in RNA*later*-preserved fecal samples collected from the forest floor (46–50, 59). While this approach has been

instrumental in determining the prevalence, geographic distribution, and evolutionary history of SIVs infecting wild chimpanzee, gorilla, and sooty mangabey populations (31, 49, 54), it is not suitable for isolation of infectious virus. In this paper, we have thus explored an alternative approach. Reasoning that virion RNA must represent the progeny of recent virus infection and expression, we constructed SIVcpzPts genomes based on viral consensus sequences present in the feces of three wild-living *P. t. schweinfurthii* apes. Our results show that this form of virus recovery indeed yielded infectious SIVcpz molecular clones, thus providing proof of concept for a new isolation strategy applicable especially to situations when direct human-animal contact is infrequent or discouraged.

## MATERIALS AND METHODS

**Chimpanzee samples.** The amplification and sequence analysis of the TAN1 genome from a wild chimpanzee in Gombe National Park, Tanzania (Ch-06) have previously been reported (46). To amplify two additional SIVcpz genomes, fecal samples were obtained from a female member of the nonhabituated Kalande community (Ch-64, sampled 11 November 2001) and a high-ranking male member of the Mitumba community (Ch-45, sampled 3 January 2002). These two chimpanzees had previously been shown to harbor divergent strains of SIVcpzPts, termed TAN2 and TAN3, respectively (48). Fecal RNA was extracted using the RNAqueous Midi-Kit (Ambion, Austin, TX) as described previously (31, 46, 48, 54).

**Generation of fecal SIVcpzPts population sequences.** TAN2 and TAN3 full-length fecal consensus sequences were obtained by amplifying and sequencing partially overlapping subgenomic fragments (367 bp to 2,197 bp) from fecal RNA as described previously (46). Briefly, cDNA was synthesized by adding fecal viral RNA (10  $\mu$ l) to a reverse transcriptase PCR (RT-PCR) master mixture containing 1 $\times$  buffer (Invitrogen; Carlsbad, CA), 0.5 mM deoxynucleoside triphosphate (dNTP), 5 mM dithiothreitol, 2 pmol of primer, 20 U of RNase inhibitor (Ambion), and 200 U of SuperScript RT III (Invitrogen) and by incubating the mixture for 1 h at 50°C. cDNA (10  $\mu$ l) was then added to a PCR mixture consisting of 1 $\times$  Expand Buffer II (Roche Diagnostics, Indianapolis, IN), 0.2 mM dNTP, 300 nM of first-round PCR primers, 25  $\mu$ g of bovine serum albumin, and 2.6 U of Expand high-fidelity polymerase mixture (Roche Diagnostics). Amplification conditions for both first- and second-round PCRs included 45 cycles of denaturation (94°C, 0.5 min), annealing (50°C, 0.5 min), and extension (68°C, 1 min). Amplicons were gel purified and sequenced directly using an ABI 3730 DNA analyzer. Population sequences were analyzed using Sequencher version 4.6 (Gene Codes Corporation, Ann Arbor, MI), and chromatograms were carefully examined for positions of base mixtures.

**Construction of SIVcpzPts proviral clones.** The TAN1.24 provirus was generated by concatenating 13 RT-PCR products using extension overlap PCR (16, 65). Amplicons were linked in a stepwise manner using the GeneAmp high-fidelity PCR system (Applied Biosystems, Foster City, CA) as shown in Fig. 1. Each concatenation reaction included 30 cycles of denaturation (94°C, 15 s), annealing (55°C, 30 s), and extension (68°C, 2 to 5 min). PCR products representing the 5' and 3' halves of the TAN1 genome were subcloned into pCR-XL TOPO (Invitrogen) and ligated to generate a complete provirus (TAN1.24).

TAN2 and TAN3 proviruses were chemically synthesized (purchased from Blue Heron Biotechnology, Bothell, WA) using their DNA consensus sequences as a guide (see Results). Synthesized 3' and 5' proviral halves were subcloned in pCR-XL TOPO and ligated to generate full-length proviral clones, termed TAN2.1 and TAN3.1, respectively.

**Large-scale plasmid preparation.** Plasmid clones containing full-length SIVcpzPts genomes were propagated in STBL2 (Invitrogen) cells at 30°C in a Forma orbital benchtop shaker (model 420/21) using 2-liter disposable Erlenmeyer flasks (Corning; 16 cm diameter, vented caps) at low levels of agitation (225 rounds per minute). Bacterial cultures were harvested prior to reaching saturating growth density and purified endotoxin free (QIAGEN, Valencia, CA). Because proviral plasmid clones are inherently unstable, each large-scale plasmid preparation was completely sequenced prior to use in biological experiments. All replication-competent SIVcpzPts (TAN1.910, TAN2.69, and TAN3.1) clones have been submitted to the National Institutes of Health Research and Reference Program (Rockville, MD).

**Site-directed mutagenesis.** Site-directed mutagenesis of the TAN1.24 clone was performed using the QuikChange XL site-directed mutagenesis kit (Strat-

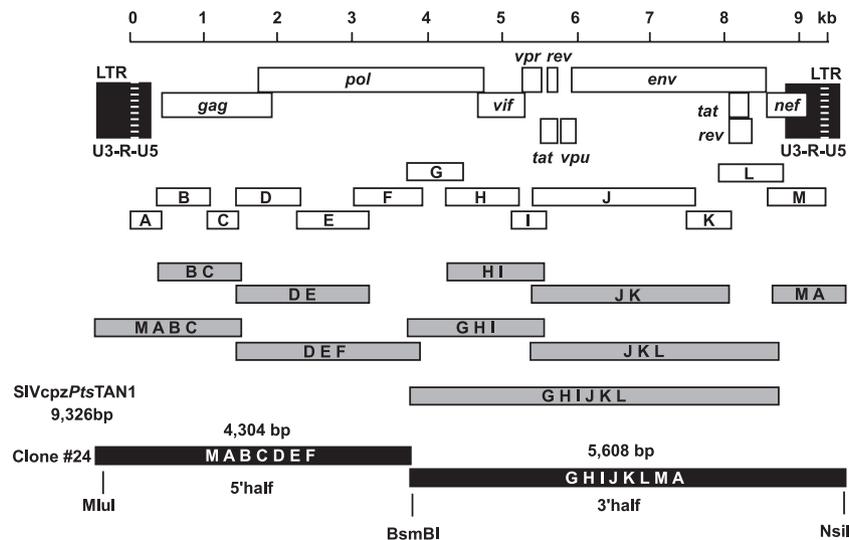


FIG. 1. Construction of a full-length SIVcpzPtt provirus by concatenation of subgenomic RT-PCR fragments. RT-PCR products (open boxes) previously reported to comprise the complete TAN1 genome (46) are shown in relation to sequential concatenation products (gray boxes) generated by extension overlap PCR. All fragments are drawn to scale, and the derivation of each concatenation product is indicated by capital letters. The final amplicons comprising 5' (4,304-bp) and 3' (5,608-bp) proviral halves were ligated at an internal BsmBI site, generating the TAN1.24 provirus. Nucleotide sequences are numbered starting at the beginning of the R region in the 5' LTR (see scale bar).

agene, La Jolla, CA). Up to 3 nucleotide substitutions were introduced simultaneously. Mutagenesis primers (15 pmol each) complementary to both DNA strands and containing the intended mutations were mixed with 5 to 10 fmol of supercoiled plasmid DNA as the template. PCR amplification was conducted for 50 s at 95°C, 50 s at 60°C, and 9 min at 68°C. Following 18 amplification cycles, template DNA was digested by methylation-specific cleavage using 10 U of DpnI. Unmethylated circular plasmid DNA containing the primer-mediated mutations was used for transformation of STBL2 cells (Invitrogen). Half-genome clones were directionally ligated into PCR-XL TOPO, and proviral constructs were confirmed by nucleotide sequence analysis.

Site-directed mutagenesis was also used to replace a guanine at position 3057 in the TAN2.1 genome with an adenine. Recombinant PCR was used to introduce this single nucleotide substitution, which was centered within the overlap of two primary amplification products using the following primers (mutations underlined and in boldface): F1 (5'-CAGAGAGTTAAATAAGAGAACTCAGG A-3'), R1 (5'-GGTCAGTCAATCCCTTACTCCTCTAATTAG-3'), F2 (5'-C TAATTAGAGGAGTAAAGGGATTGACTGACC-3'), and R2 (5'-CTGCTGC ACTAGTAAAGTTGGCCATTATC-3'). The two fragments were then linked by overlap extension PCR using 30 fmol of each amplicon and nested outside primers (F3, 5'-GGAACATGGTAACAGTGCTGGATGTAGG-3', and R3, 5'-GGTATGACCTCTGCTTCTAGGAAACCAC-3') using the following conditions: 3 cycles of 1 min at 95°C, a 2-min ramp from 37°C to 72°C, and 2 min at 72°C followed by 15 cycles of 1 min at 95°C, 1 min at 60°C, and 2 min at 72°C. The resulting fragment was cloned into TAN2.1 using unique BstZ171 and PshAI restriction enzyme sites to generate TAN2.69.

**Phylogenetic analyses.** SIVcpzTAN2 and -TAN3 were compared to representative members of the HIV-1/SIVcpz lineage using the following full-length proviral sequences (GenBank accession numbers are in parentheses): HIV-1 group M, U455A, subtype A (M62320), and BRU, subtype B (K02013); HIV-1 group N, YBF30 (AJ006022) and YBF106 (AJ271370); HIV-1 group O, MVP5180 (L20571) and ANT70 (L20587); SIVcpzPtt, LB7 (DQ373064), MB66 (DQ373063), EK505 (DQ373065), MT145 (DQ373066), CAM3 (AF115393), CAM5 (AJ271369), CAM13 (AY169968), GAB1 (X52154), GAB2 (AF382828), and US (AF103818); SIVcpzPtt, ANT (U42720) and TAN1 (AF447763). Amino acid sequences were aligned using CLUSTAL W (52). Sites that could not be aligned unambiguously or sites with a gap in any sequence were excluded. For full-length genome analysis, Gag/Pol and Pol/Vif overlaps were removed from the C termini of the deduced Gag and Pol protein sequences, respectively. Trees were inferred by the Bayesian method (64) implemented in MrBayes version 3.1 (27) using the Jones et al. matrix (29) with gamma-distributed rates at sites (63) and 1 million generations. Average standard deviation of split frequencies was 0.002.

**Transfection-derived viral stocks.** Full-length SIVcpz and HIV-1 proviral clones were transfected into 293T cells using the Lipofectin-based FuGene6 reagent (Roche Diagnostics). Supernatants were collected 48 to 72 h following transfection and clarified by low-speed centrifugation. Cell-free virus stocks were stored in aliquots at -80°C.

To determine particle-associated RT activity, supernatants (1 to 5 ml) were pelleted at 40,000 × g for 1 h at 4°C, solubilized in lysis buffer, titrated using serial 10-fold dilutions, and analyzed using a colorimetric enzyme-linked immunosorbent assay (Roche Diagnostics). RT activity was determined by quantitative immunodetection of digoxigenin-labeled nucleotides incorporated during cDNA synthesis using a homopolymeric poly(A)/oligo(dT) primer complex. Background RT activity was measured in supernatants of mock-transfected cells and subtracted. The amount of enzymatic activity (expressed as ng RT/ml) was determined using the linear range of a standard curve generated from decreasing amounts (2 ng to 3 pg) of recombinant HIV-1 RT analyzed in parallel (Roche Diagnostics).

**Single-round infectivity assay and coreceptor usage.** Viral infectivity was tested using JC53-BL cells, which express high levels of CD4, CXCR4, and CCR5 and also harbor Tat-responsive reporter genes for *Escherichia coli* β-galactosidase as well as firefly luciferase under the control of an HIV-1 long terminal repeat (LTR) sequence (14, 42, 56). JC53-BL cells were seeded in 24-well plates at a density of 50,000 cells/well in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum 24 h prior to infection. Serial dilutions of virus stocks were added in a final volume of 250 μl/well containing 40 μg/ml DEAE-dextran (Sigma-Aldrich, St. Louis, MO). Forty-eight hours postinfection, plates were stained, and β-galactosidase-positive cells were counted. Infectivity (infectious units [IU]/ml) scored at different dilutions was averaged and standardized based on particle-associated RT activity of viral input (IU/ng RT).

JC53-BL cells were also used to determine the coreceptor preference of the reconstructed viruses. Cells were seeded at 8,300 cells/well in 96-well plates overnight and then treated with AMD3100 (1.2 μM), TAK-779 (10 μM), or a combination of both for 1 h. Virus was added (1 ng of particle-associated RT = ~2,000 IU) in the presence of 40 μg/ml DEAE-dextran. After 48 h of incubation at 37°C, supernatant was removed, and cells were processed for detection of luciferase activity according to the manufacturer's recommendations (Promega, Madison, WI). Light intensity of cell lysates was quantitated using a Tropic luminometer and WinGlow version 1.24 software.

**Chimpanzee and human PBMC cultures.** Blood was obtained from healthy (HIV-1-uninfected) human volunteers (Research Blood Components, Boston, MA) as well as uninfected chimpanzees housed at the Yerkes Regional Primate Research Center. Chimpanzee blood was collected from anesthetized individuals

TABLE 1. Nucleotide and amino acid sequence differences between the TAN1 consensus sequence and individual TAN1 molecular clones

TAN1 clone	No. of differences																	
	Nucleotides (LTR)			Amino acids										Nef				
				Gag		Pol			Vif	Vpr	Tat		Rev		Vpu	Env		
	U3	R	U5	p17	p24	PR	RT	IN			Exon 1	Exon 2	Exon 1	Exon 2		gp120	gp41	
24	4	0	0	1	4	2	9	4	4	0	0	0	0	0	1	1	6	1
177	4	0	1 <sup>a</sup>	1	3	0	2	0	2	0	0	1 <sup>b</sup>	0	0	0	2	3 (1 <sup>b</sup> )	0
602	4	0	0	0	0	0	0	0	2	0	0	0	0	0	2	2	2	0
801	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	0
910	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

<sup>a</sup> Error introduced by PCR-mediated mutagenesis at position 138.

<sup>b</sup> Error introduced by PCR-mediated mutagenesis at position 8048.

as part of their annual health survey, a procedure approved by the Emory Institutional Animal Care and Use Committee. Peripheral blood mononuclear cells (PBMCs) were isolated by density separation using Ficoll-Hypaque Plus (GE Healthcare, Piscataway, NJ) and centrifuged at 1,800 rpm for 25 min at 22°C. Buffy coat cells were washed once at room temperature in Hanks balanced saline solution (HBSS) plus 4 mM EDTA and once at 4°C in HBSS plus 1% fetal calf serum (FCS). CD4<sup>+</sup> cells were enriched using CD4-containing microbeads and magnetic cell sorting (Miltenyi Biotec, Auburn, CA). For activation, CD4<sup>+</sup> cells were allowed to adhere in six-well plates for 30 min and then stimulated for 12 to 15 h with 3 µg/ml of staphylococcal enterotoxin B (Sigma-Aldridge, St. Louis, MO) in RPMI medium plus 15% FCS. Cells were washed once with HBSS, resuspended in DMEM with 10% giant cell tumor-conditioned media (BioVeris Corp., Gaithersburg, MD) and 10% human AB serum (Fisher Bioreagents, Fair Lawn, NJ), returned to the same well in which they were activated, and incubated for 5 to 6 days at 37°C. Cells were placed in DMEM containing 10% FCS and 30 U/ml interleukin-2 (IL-2) 1 day prior to infection. One million stimulated CD4<sup>+</sup> cells were incubated with transfection-derived viral stocks equilibrated by particle-associated RT activity (50 ng) in 300 µl DMEM containing 10% FCS and 30 U/ml IL-2 for 16 h at 37°C. Cells were washed three times with HBSS and plated on 24-well plates in 2 ml DMEM/well containing 10% FCS and 30 U/ml IL-2. To determine viral growth kinetics, 20-µl aliquots of supernatant were collected on the first day and every other day thereafter following cell seeding and analyzed for RT activity.

**Immunoblot analysis.** The protein content of SIVcpz and HIV-1 virions was examined by immunoblotting using anti-HIV-1 Gag-, Pol-, and Env-specific antibodies. Cell-free supernatants of transfected 293T cells were pelleted. Virions were solubilized in StaphA buffer (25 mM sodium phosphate buffer, pH 7.4, 100 mM NaCl, 1% Triton X-100, 0.1% sodium dodecyl sulfate, 0.5% deoxycholate, 5 mM EGTA) containing 1 mM phenylmethylsulfonyl fluoride and 10 µg/ml aprotinin protease inhibitors. Viral proteins were separated on 4 to 12% precast Bis-Tris polyacrylamide gradient gels (Invitrogen) and transferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA) by electroblotting. Protein loading was equilibrated by RT activity. Membrane blocking and antibody treatment were conducted in Tris-buffered saline, pH 7.4 (TBS), containing 0.05% Tween 20 and 5% nonfat dry milk. Membranes were washed three times following antibody incubations in TBS containing 0.05% Tween 20 and once in detergent-free TBS prior to addition of chemiluminescent substrate. Bound antibodies were visualized with horseradish peroxidase-conjugated secondary antibodies (Pierce, Rockford, IL, and Southern Biotech, Birmingham, AL), followed by West Femto chemiluminescence detection (Pierce). Polyclonal rabbit anti-HIV-1 p24 (4250) and integrase (756) (21) antibodies were obtained through the NIH AIDS Research and Reference Reagent Program and used at dilutions of 1:50,000 and 1:2,000, respectively. An anti-HIV-1 RT antiserum was provided by X. Wu and J. Kappes (18) and used at a dilution of 1:8,000. Finally, an anti-SIVcpz gp120 antiserum generated in immunized guinea pigs (see below) was used at a dilution of 1:2,000.

**Generation of SIVcpz lineage-specific anti-gp120 antibodies.** SIVcpz Env proteins cross-react only poorly with anti-HIV-1 Env antibodies. To permit detection of the TAN1.910, TAN2.69, and TAN3.1 Env proteins on immunoblots, we raised SIVcpz Env-specific antisera in guinea pigs. Guinea pigs were housed according to Accreditation of Laboratory Animal Care guidelines. Two female Hartley guinea pigs (Harlan Sprague, Indianapolis, IN) were immunized intramuscularly six times at 3-week intervals with 400 µg of a codon usage-optimized TAN1 env (gp160) plasmid (donated to the National Institutes of Health Research and Reference Program). Subsequently, animals were boosted

twice at 3-week intervals with 50 µg of recombinant TAN1 gp120 protein in the presence of trehalose dicorynomycolate adjuvant (1:1, vol/vol; Sigma; M6536). Recombinant TAN1 gp120 protein was generated by transfection of 293T cells with a codon usage-optimized TAN1 gp120 env expression vector purified from serum-free medium 72 h following transfection by lectin affinity chromatography, dialyzed against phosphate-buffered saline, and concentrated by membrane centrifugation (700-kDa cutoff).

**Nucleotide sequence accession numbers.** GenBank accession numbers for the sequences determined in this study are EF394357 (TAN2 consensus), EF394358 (TAN3 consensus), EF394356 (TAN1.910), DQ374657 (TAN2.69), DQ374658 (TAN3.1), and EF451055 (codon usage-optimized TAN1 gp160 env gene).

## RESULTS

**Construction of an infectious SIVcpz clone by concatenation of subgenomic fragments.** In a first attempt to generate an infectious molecular clone of SIVcpz, we concatenated a proviral genome (TAN1) that had previously been amplified as 13 partially overlapping cDNA fragments from fecal virion RNA of a wild-living chimpanzee (Ch-06) in Gombe National Park (46, 50). Using existing amplicons and the full-length TAN1 population sequence as a guide (46), we employed strain-specific primers and extension overlap PCR to construct a full-length provirus. As outlined in Fig. 1, this approach involved multiple (up to four) consecutive rounds of PCR amplification of the same genomic region, each consisting of 30 cycles. The final products comprised 5' and 3' proviral halves, which were ligated at an internal BsmBI site to generate a full-length provirus. The resulting clone (TAN1.24) contained complete LTR elements on both ends of its genome (Fig. 1).

To examine the integrity of the TAN1.24 genome, the clone was sequenced in its entirety. Like the TAN1 consensus sequence (46), TAN1.24 encoded a complete set of viral proteins and intact regulatory elements, including promoter and enhancer regions in the LTR, the transactivating response (TAR) element, the packaging signal, and the primer binding site. However, TAN1.24 differed from the TAN1 consensus sequence in 76 nucleotides. Thirty-three of these represented missense mutations which encoded amino acid substitutions in Gag, Pol, Vif, Vpr, Env, and Nef protein sequences (Table 1). In addition, there were four nucleotide sequence differences in the 3' unique (U3) region of the viral LTR (Table 1). Inspection of the original TAN1 chromatograms indicated that only 17 of the 76 nucleotide differences represented previously identified mismatches in regions of fragment overlap (46). The remainder were thus likely introduced during the PCR-mediated concatenation process.

TABLE 2. Biological activity of TAN1 molecular clones<sup>a</sup>

Virus	Infectious titer <sup>b</sup> (IU/ml)	RT <sup>c</sup> activity (ng/ml)	Virion infectivity (IU/ng RT)
TAN1.24	0	0	0
TAN1.177	1,100	66	17
TAN1.602	55,800	55	1,121
TAN1.801	69,050	57	1,235
TAN1.910	19,900	8	3,344
Mock	0	0	0
HIV-1/SG3	3,096,250	294	10,496

<sup>a</sup> Values are averaged from two independent experiments.

<sup>b</sup> Determined as the number of  $\beta$ -galactosidase-expressing JC53-BL cells per ml of transfection supernatant.

<sup>c</sup> Nanograms of particle-associated RT activity per ml of transfection supernatant.

To examine its biological activity, TAN1.24 was transfected into 293T cells and the resulting supernatants were tested in a single-round infectivity (JC53-BL) assay (14, 42, 56). JC53-BL cells express high levels of CD4, CCR5, and CXCR4 receptor molecules and are stably transfected with  $\beta$ -galactosidase and luciferase reporter genes under the control of the HIV-1 LTR. The infectious titer of any given virus stock can thus be determined by staining cultures for  $\beta$ -galactosidase expression and counting the number of blue cells. Analysis of transfection supernatants in this assay indicated that TAN1.24-derived virus was noninfectious (Table 2). Moreover, pelleted virion preparations lacked detectable RT activity (Table 2), and cocultivation with human PBMCs failed to yield replicating virus (Fig. 2). In all of these assays, the (CXCR4-tropic) HIV-1/SG3 strain was used as a positive control (19). Thus, despite encoding a complete set of uninterrupted reading frames, TAN1.24 virus was replication defective.

Plasma-derived viral RNA is known to represent recently produced (replicating) virus because of the short half-life of infected CD4<sup>+</sup> T lymphocytes (25, 57). Assuming that viral RNA in fecal samples is the result of ongoing replication in the gut-associated lymphatic tissue (6, 34, 55), we reasoned that fecal population sequences may represent a suitable template for functional genome design. To test this hypothesis, we employed a stepwise mutagenesis approach that changed the TAN1.24 sequence to that of its corresponding consensus sequence. Three proviral clones were generated; in these, the number of amino acid differences from the consensus sequence was reduced from 33 in TAN1.24 to 15 in TAN1.177, 6 in TAN1.602, and 4 in TAN1.801 (Table 1). A fourth clone (TAN1.910) was identical to the consensus sequence in all protein coding regions as well as its entire LTR. All of these viral constructs were found to produce virions that contained measurable RT activity and mediated infection of JC53-BL cells (Table 2). Moreover, there was a marked inverse correlation between the number of missense mutations and virion infectivity: the clone with the greatest number of amino acid sequence differences from the consensus sequence (TAN1.177) produced the least infectious virus (Table 2); conversely, the clone that matched the consensus sequence in both protein and LTR sequences (TAN1.910) yielded the most infectious virus (IU-per-particle ratio; Table 2). Thus, changing the TAN1.24 sequence to that of its consensus led to an improvement in viral infectivity of 3

orders of magnitude. Similar results were observed in primary human CD4<sup>+</sup> T-cell cultures, where TAN1.910 consistently replicated to the highest titers, while TAN1.177 was severely attenuated (Fig. 2). The titers of the remaining clones (TAN1.602 and TAN1.801) reached intermediate values. These results suggested that the predominant viral species in fecal RNA is replication competent.

**Generation of replication-competent SIVcpzP<sub>ts</sub> proviruses by chemical synthesis of fecal consensus sequences.** To extend these observations and to develop a more efficient strategy of generating infectious molecular clones, we amplified SIVcpz genomes from two additional chimpanzees and determined their population sequences. Fecal samples were obtained from wild-living members of the Kalande (Ch-64) and Mitumba (Ch-45) communities in Gombe National Park, which had previously been shown to harbor diverse SIVcpzP<sub>ts</sub> strains, termed TAN2 and TAN3, respectively (48). An RT-PCR strategy similar to that originally employed for TAN1 (46) was used to amplify complete TAN2 and TAN3 genomes as sets of 15 and 14 partially overlapping subgenomic fragments (Fig. 3). Direct sequence analysis of these RT-PCR amplicons yielded very few positions of base mixtures (defined as "N" in sequence chromatograms of both DNA strands; Table 3). However, a considerable number of sequence differences between adjoining fragments in regions of sequence overlap were observed, suggesting primer selection and/or limited template input (Table 3). Approximately one-half of the observed ambiguities affected the encoded amino acid (Table 3), but none introduced stop codons or frameshift mutations. Both TAN2 and TAN3 consensus sequences contained uninterrupted reading frames for all structural and regulatory proteins as well as intact regulatory elements. In phylogenetic trees of Gag, Pol, Env, and Nef gene coding regions, TAN2 and TAN3 clustered with TAN1, forming a well-supported geographic subclade within the SIVcpzP<sub>ts</sub> radiation. A representative tree of envelope amino acid sequences is depicted in Fig. 4.

To test whether the TAN2 and TAN3 population sequences could be used to generate infectious molecular clones, we chemically synthesized their genomes as proviral halves (purchased from Blue Heron Biotechnology, Bothell, WA) (Fig. 3). For this purpose, it was necessary to select one nucleotide at each position of sequence ambiguity, including differences in

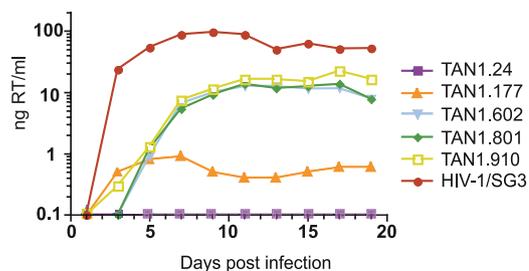


FIG. 2. Replication potential of TAN1-derived molecular clones in human CD4<sup>+</sup> T cells. The replication kinetics of TAN1.24, TAN1.177, TAN1.602, TAN1.801, and TAN1.910 are shown in relation to a positive HIV-1 (SG3) control (y axis, nanograms of RT activity per ml of culture supernatant). T cells enriched in CD4<sup>+</sup> from a human donor were infected with transfection-derived culture supernatants equilibrated by particle-associated RT activity (50 ng input).

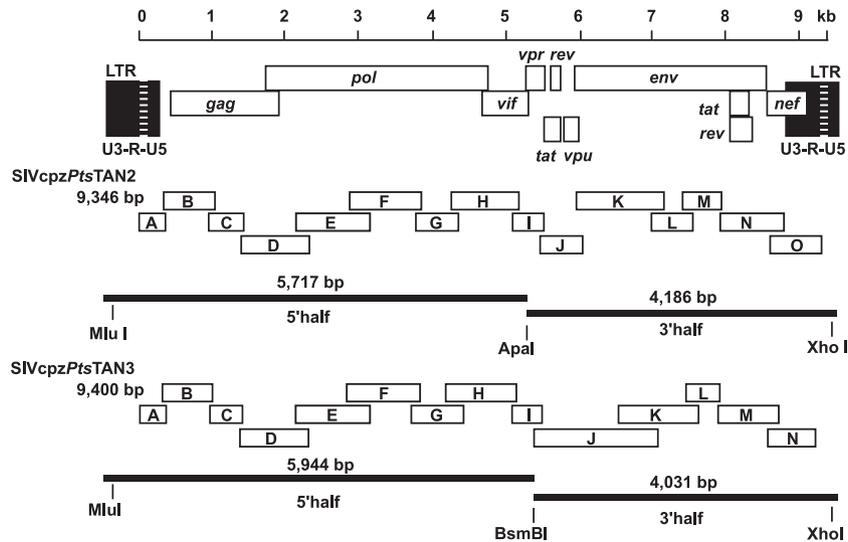


FIG. 3. Synthesis of SIVcpzPts proviruses from fecal RNA consensus sequences. The complete TAN2 and TAN3 genomes were amplified as partially overlapping RT-PCR fragments (open boxes) from fecal virion RNA and sequenced without interim cloning (fragments are drawn to scale). The resulting population sequences were used to synthesize the TAN2 and TAN3 genomes as proviral halves (black lines). Restriction enzymes introduced into the sequence allowed their directional cloning into an appropriate plasmid vector (pCR-XL TOPO).

regions of fragment overlap (Table 3). The following criteria were used to make these choices. For ambiguous positions that did not affect the corresponding amino acid, the most predominant nucleotide (highest amplitude in the sequence chromatogram or most frequent nucleotide in repeat sequencing reactions) was chosen. For mixtures that altered the corresponding amino acid, the base that encoded the most common amino acid residue in alignments of existing SIVcpz and HIV-1 protein sequences was selected. In the absence of an apparent conserved amino acid (e.g., in hypervariable protein regions), the nucleotide with the highest amplitude in sequence chromatograms was selected. Both TAN2 and TAN3 genomic halves were synthesized with a restriction site introduced to facilitate subsequent cloning into an appropriate plasmid vector. The integrity of the assembled proviruses (TAN2.1 and TAN3.1) was confirmed by sequence analysis.

To examine their biological activity, transfection-derived TAN2.1 and TAN3.1 viruses were tested for replication competence in primary CD4<sup>+</sup> T-cell cultures from different human donors (one representative experiment is shown in Fig. 5). TAN3.1 virus replicated to high titers, with kinetics very similar to those of HIV-1 (SG3) and SIVcpz (GAB1.1a) laboratory controls (19, 28). In contrast, TAN2.1 was replication defective in these same cultures (Fig. 5). To explore the reason(s) for this phenotype, transfection-derived TAN1.910, TAN2.1, and TAN3.1 stocks were purified and the resulting particles analyzed for protein composition by immunoblot analysis. These experiments revealed no obvious defect in TAN2.1 particle production and/or release: Western blots of pelleted virions revealed comparable quantities of properly processed Gag (p24), integrase (p32), and envelope (gp120) glycoproteins for all three SIVcpzPts strains (Fig. 6A). However, TAN2.1-derived virions lacked p51 and p66 bands (Fig. 6A, lane 2) and contained only very low levels of particle-associated RT enzymatic activity (reduced by 3 orders of magnitude compared to

TAN1.910 and TAN3.1; not shown). These data suggested a specific defect in the TAN2.1-encoded RT heterodimer.

To identify the underlying mutation, we reinspected all chromatograms containing TAN2 *pol* gene sequences. This analysis identified one ambiguous site that had initially gone unrecognized (Fig. 6B). At position 287 of the HIV-1 RT protein, the TAN2.1 *pol* sequence chromatogram contained an A/G double peak, encoding either a lysine (K) or a glutamic acid (E) in the corresponding protein. Alignment of published SIVcpz and HIV-1 RT sequences revealed that the lysine, but not the initially selected glutamic acid, was highly conserved (Fig. 6C). In fact, the lysine was invariant in all known SIVcpz strains as well as HIV-1 group M, N, and O consensus sequences, suggesting that it was essential for RT function (Fig. 6C). We thus used site-directed mutagenesis to change the glutamic acid to the conserved lysine residue. This substitution restored particle-associated RT activity as well as viral infectivity. The resulting clone (TAN2.69), which differed from TAN2.1 in a single nucleotide, produced virions that contained stable p66/p51 RT heterodimers (Fig. 6A, right lane) and wild-type levels of RT activity (not shown). Moreover, virus derived from the TAN2.69 clone replicated efficiently in human CD4<sup>+</sup> T cells, reaching maximal titers several days before virus derived from TAN1.910 and TAN3.1 (Fig. 5). Thus, modification at a single ambiguous site in the TAN2 population sequence restored its biological activity. The crystal structure of the p66/p51 heterodimer places the lysine at position 287 in the "thumb" of the HIV-1 RT complex (26). In the p51 subunit, this region is involved in forming the dimer interface. One possibility, thus, is that the glutamic acid replacement destabilized the p51/p66 interface, thereby rendering the heterodimer susceptible to premature degradation. The absence of a detectable p51/p66 RT complex in pelleted TAN2.1 virions is consistent with this hypothesis.

TABLE 3. Sequence ambiguities in TAN2 and TAN3 population sequences

Virus	Amplicon <sup>a</sup>	Sequence ambiguity <sup>b</sup>	Position in genome	Protein	Codon	Encoded amino acid <sup>c</sup>
SIVcpzTAN2	B-C	C-T	1068	p24 Gag	239	<b>P</b>
	E-F	T-C	2969	RT	257	<b>I</b>
	E-F	A-C	3017	RT	273	<b>G</b>
	E-F	C-T	3047	RT	283	<b>I</b>
	*F	G/A <sup>d</sup>	3057	RT	287	<b>E/K</b>
	F-G	G-A	3869	RT	557	<b>R</b>
	G-H	A-G	4397	IN	173	<b>R</b>
	G-H	A-G	4407	IN	177	<b>I-A</b>
	G-H	T-C	4408	IN	177	<b>I-A</b>
	H-I	A-G	5149	Vif	154	<b>I-V</b>
	H-I	T-C	5154	Vif	155	<b>V</b>
	J-K	T-C	6104	Env	74	<b>P</b>
	J-K	G-C	6123	Env	81	<b>D/A-H</b>
	*J-K	A/C-A	6124	Env	81	<b>D/A-H</b>
	J-K	T-C	6125	Env	81	<b>D/A-H</b>
	J-K	G-A	6153	Env	91	<b>E-K</b>
	*K	G/T	6417	Env	179	<b>V-F</b>
K-L	A-G	7165	Env	428	<b>K-R</b>	
L-M	G-A	7625	Env	581	<b>K</b>	
SIVcpzTAN3	*B	G/A	894	p24 Gag	27	<b>V</b>
	*B	C/T	909	p24 Gag	32	<b>F</b>
	E-F	G-T	3137	RT	313	<b>E-D</b>
	*F	A/T	3221	RT	341	<b>I</b>
	*F	G/A	3658	RT	487	<b>R/K</b>
	G-H	T-C	4241	IN	121	<b>F</b>
	G-H	G-A	4277	IN	133	<b>A</b>
	G-H	A-G	4280	IN	134	<b>Q</b>
	G-H	G-A	4292	IN	138	<b>E</b>
	G-H	A-G	4334	IN	152	<b>E</b>
	G-*H	T-C/T	4350	IN	158	<b>L</b>
	G-H	A-C	4406	IN	176	<b>V</b>
	H-I	A-G	5147	Vif	153	<b>K-R</b>
	J-K	C-T	6584	Env	236	<b>Y</b>
	J-K	A-C	6652	Env	259	<b>Y-S</b>
	J-K	G-C	6668	Env	264	<b>V</b>
	J-K	C-G	6728	Env	284	<b>L</b>
	J-K	A-C	6736	Env	287	<b>K-T</b>
	J-K	G-A	6737	Env	287	<b>K-T</b>
	J-K	A-G	6800	Env	308	<b>E</b>
	J-K	A-G	6831	Env	319	<b>N-D</b>
	J-K	G-A	6841	Env	322	<b>R-K</b>
	J-K	T-C	6872	Env	332	<b>T</b>
	J-K	C-T	6883	Env	336	<b>T-I</b>
	J-K	T-A	6890	Env	338	<b>N-K</b>
	J-K	T-G	6894	Env	340	<b>F-V</b>
	J-K	A-C	6957	Env	361	<b>N-H</b>
J-K	G-T	6989	Env	371	<b>K-N</b>	
J-K	A-G	7002	Env	376	<b>N-D</b>	
J-K	A-G	7011	Env	379	<b>R-E</b>	
J-K	G-A	7012	Env	379	<b>R-E</b>	
J-K	G-A	7015	Env	380	<b>R-K</b>	
J-K	C-A	7022	Env	382	<b>D-E</b>	

<sup>a</sup> The locations of TAN2 and TAN3 amplicons are shown in Fig. 3.

<sup>b</sup> Hyphens indicate nucleotide differences between adjoining fragments in regions of sequence overlap; slashes indicate positions of base mixtures; the latter are also highlighted by asterisks on the corresponding amplicons.

<sup>c</sup> Amino acids encoded by the TAN2.69 and TAN3.1 infectious molecular clones are boldfaced.

<sup>d</sup> Initially unrecognized double peak in TAN2.1 sequence chromatogram.

**Biological properties of the infectious SIVcpzPts clones.** Availability of the first infectious molecular clones of SIVcpzPts allowed us to compare their baseline biological properties, including coreceptor usage, replication capacity, cell tropism, and neutralization phenotype, with those of representative HIV-1 and SIVcpzPtt clones. Using the JC53-BL assay, we

found that the infectivity of TAN1.910, TAN2.69, and TAN3.1 (as well as that of GAB1.1a) was blocked by TAK779, an antagonist for the CCR5 coreceptor, but not by AMD3100, an antagonist specific for CXCR4 (Fig. 7). The three SIVcpzPts strains thus resembled primary strains of HIV-1 (as well as the single SIVcpzPtt clone) in their coreceptor preference. Virions

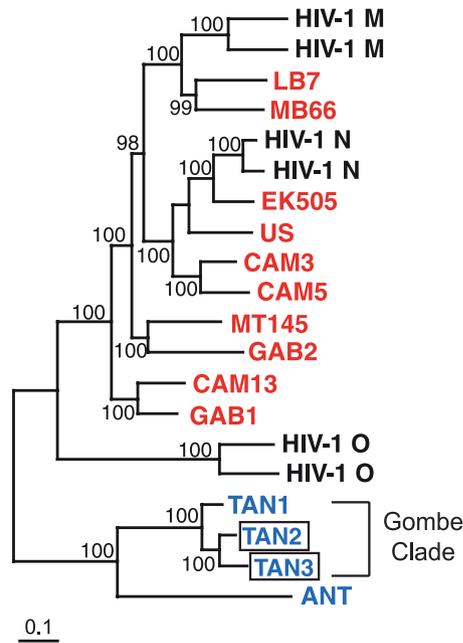


FIG. 4. Phylogenetic relationships of TAN2 and TAN3 inferred from Env amino acid sequences. The phylogenetic positions of the two newly derived SIVcpzPtt strains (boxed) are shown in relation to representative strains of HIV-1 (groups M, N, and O; black), SIVcpzPtt (red), and SIVcpzPtt (blue). Viruses identified in wild chimpanzees from Gombe National Park are indicated. Trees were inferred by the Bayesian method; numbers on nodes are percent posterior probabilities (only values of 95% and above are shown). The scale bar represents 0.1 substitutions per site.

containing the CXCR4-tropic NL4.3 Env, the CCR5-tropic YU2 Env, and the dual-tropic WEAU1.6 Env were analyzed in parallel for control (Fig. 7).

We next compared TAN1.910-, TAN2.69-, and TAN3.1-derived viruses in CD4<sup>+</sup> T-cell cultures from different human donors. All three SIVcpzPtt strains induced large syncytia and were capable of establishing a productive infection in human lymphocytes (Fig. 5 and not shown). However, in most CD4<sup>+</sup> T-cell cultures TAN1.910 replicated 1 to 2 orders of magnitude less efficiently than the other two SIVcpzPtt strains (Fig. 5 and not shown). Since TAN1.910 is not identical in its sequence to the TAN1 consensus, it is possible that some (or all) of the 40 remaining (albeit synonymous) nucleotide differences are responsible for this more attenuated phenotype. Alternatively, the observed differences may reflect the normal range of replication fitness of naturally occurring SIVcpz strains in human PBMCs. Biological characterization of additional SIVcpzPtt and SIVcpzPtt strains will be necessary to address this possibility.

The replication capacity of TAN1.910-, TAN2.69-, and TAN3.1-derived viruses was also tested in chimpanzee PBMCs. CD4<sup>+</sup> T-cell cultures were established from uninfected chimpanzees housed at the Yerkes Regional Primate Center and cocultivated with (RT-equilibrated) transfection-derived viral supernatants. As shown in Fig. 8, all three SIVcpzPtt strains replicated in chimpanzee CD4<sup>+</sup> T cells, although TAN1.910 again grew to only modest titers compared to the other two strains. Surprisingly, however, cultures from different chimpanzees

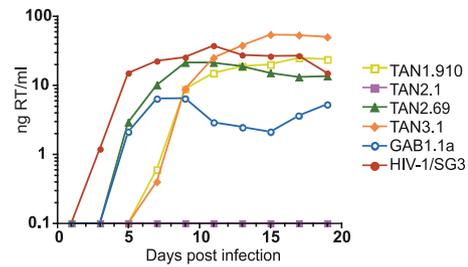


FIG. 5. Replication potential of TAN1, TAN2, and TAN3 molecular clones in human CD4<sup>+</sup> T cells. The replication kinetics of TAN1.910, TAN2.1, and TAN3.1 as well as TAN2.69, which encodes a Lys at position 287 of its RT protein, are shown in relation to those of HIV-1 (SG3) and SIVcpzPtt (GAB1.1a) reference strains (y axis, nanograms of RT activity per ml of culture supernatant). Enriched CD4<sup>+</sup> T cells from four human donors were infected with transfection-derived culture supernatants equilibrated by particle-associated RT activity (50 ng), except for TAN2.1, which was used at maximal RT activity. Results from one representative experiment are shown.

varied in their abilities to support SIVcpzPtt replication (Table 4): TAN1.910 and TAN3.1 replicated in only a small fraction of cultures tested (2/7 and 3/8, respectively), while TAN2.69 replicated in each of four cultures analyzed, including two that did not support the replication of TAN1.910 and TAN3.1. HIV-1/SG3 and SIVcpzGAB1 replicated in all, or all but one, donor cultures (8/8 and 5/6, respectively), thus excluding suboptimal culture conditions as the reason for these differences (Table 4). Based on these data, it seems clear that the host range of TAN1.910 and TAN3.1 is more restricted than that of TAN2.69 and that chimpanzee CD4<sup>+</sup> T cells appear to be overall more refractory to SIVcpz infection than their human counterparts.

Finally, we probed the three new SIVcpzPtt strains for sensitivity to neutralization by plasma from HIV-1-infected individuals, human neutralizing monoclonal antibodies, and soluble CD4. These experiments showed that TAN1.910, TAN2.69, and TAN3.1 were resistant to neutralization by a high-titer human plasma pool (not shown), the CD4 binding site monoclonal antibody b12 (39) (not shown), and the coreceptor binding site antibody 17b (32, 60) (Fig. 9A). However, the *P. t. schweinfurthii* viruses were sensitive to neutralization by two human monoclonal antibodies (2F5 and 4E10) directed against the membrane-proximal external region, indicating the conservation of these epitopes in SIVcpzPtt (8, 66): 2F5 neutralized TAN1.910, TAN2.69, and TAN3.1 with 50% inhibitory concentrations (IC<sub>50</sub>) of 5.5, 9.2, and 2.9 μg/ml, respectively (Fig. 9B), while 4E10 yielded comparable IC<sub>50</sub> values of 6.1, 6.7, and 1.7 μg/ml (Fig. 9C). Finally, TAN1.910, TAN2.69, and TAN3.1 were also neutralized by soluble CD4, albeit at high IC<sub>50</sub> values of 395, 65, and 235 nM, respectively (Fig. 9D). Overall, the neutralization phenotype of TAN1.910, TAN2.69, and TAN3.1 was very similar to that of primary HIV-1 isolates (5). In contrast, SIVcpzGAB1 was sensitive to neutralization by human plasma (1:180), soluble CD4 (3.5 nM), and the CD4i antibody 17b (1.1 μg/ml) in the absence of pretreatment with soluble CD4 (Fig. 9). The last indicates an increased formation and/or exposure of the coreceptor binding site, which is typical of tissue culture-adapted HIV-1 strains (Fig. 9, HIV-1/SG3).

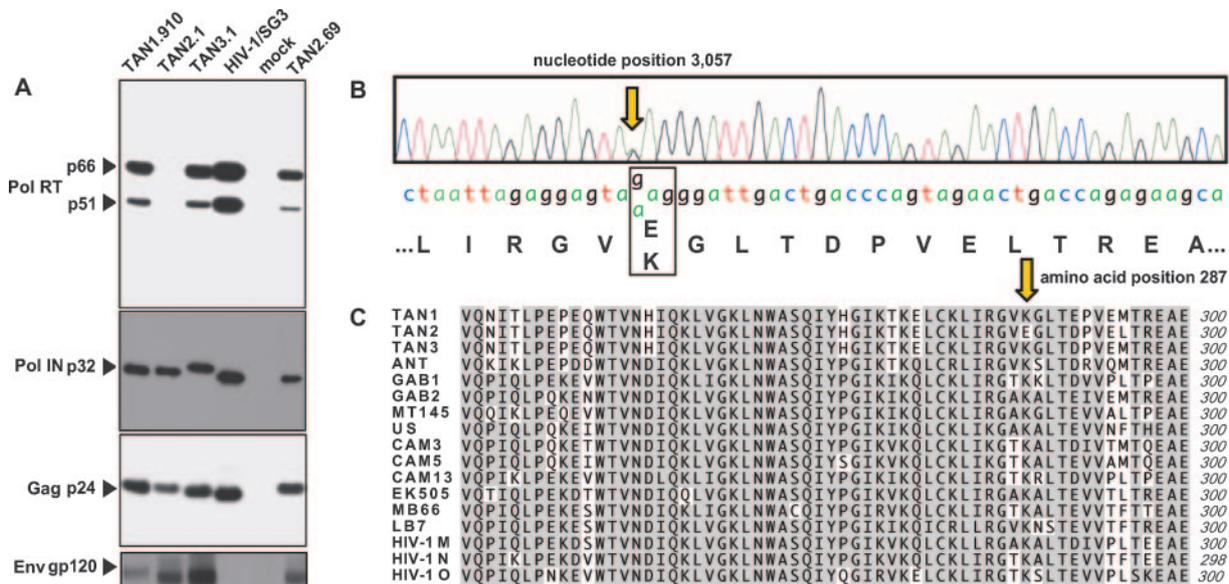


FIG. 6. Identification of an inactivating mutation in the TAN2.1 *pol* gene. (A) Western blot analysis of transfection-derived virions using Pol-, Gag-, and Env-specific antibodies. Protein loading was equilibrated by RT activity, with 100 ng loaded to detect Pol and Gag proteins and 400 ng loaded to detect Env (TAN2.1 was used at maximal RT activity). Bands corresponding to RT (p66 and p51), integrase (p32), capsid (p24), and the extracellular envelope (gp120) protein are indicated. Lane 4 lacks a gp120 band because the HIV-1/SG3 Env protein failed to cross-react with the SIVcpzPts-specific gp120 antiserum. Minor p32 migration differences correspond to predicted molecular weight differences of integrase proteins from the different viruses. (B) Sequence chromatogram of the TAN2 *pol* gene revealing a site of sequence ambiguity (G/A double peak) at position 3057 with differential coding potential (E or K at amino acid position 287 of the RT protein). (C) Alignment of available SIVcpzPts (TAN1, TAN2, TAN3, and ANT), SIVcpzPtt (GAB1, GAB2, MT145, US, CAM3, CAM5, CAM13, EK505, MB66, and LB7), and HIV-1 (group consensus) RT protein sequences (position 241 to 300) available from the Los Alamos HIV Sequence Database. Amino acids identical among all RT sequences are shaded, with position 287 highlighted by an arrow.

DISCUSSION

In this paper, we report the construction of infectious molecular clones of SIVcpzPts from fecal RNA population sequences of three wild-living chimpanzees. We show that viruses derived from these clones express all viral proteins required for productive infection of human and chimpanzee CD4<sup>+</sup> T lymphocytes. We also show that the resulting viruses are similar to primary HIV/SIV isolates in baseline biological properties, including receptor and coreceptor usage and neutralization phenotype. Although “live” viruses have previously

been generated by chemical synthesis of their genomic sequences (12, 53), this is the first example of the successful reconstitution of infectious primate lentiviruses from RNA sequences in stool samples collected from the jungle floor. This approach expands the existing arsenal of noninvasive virus detection methods and provides an important new tool to unravel the biology of viruses that are otherwise not accessible to investigation.

In our efforts to generate infectious SIVcpz clones from fecal RNA consensus sequences, we explored two alternative approaches. The first involved the concatenation of sub-

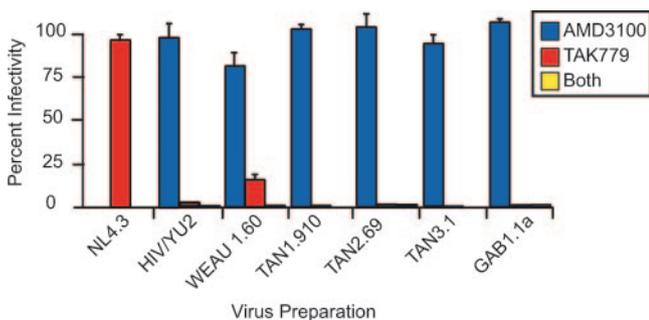


FIG. 7. Coreceptor usage of SIVcpzPts strains. JC53-BL cells were pretreated with AMD3100 (inhibitor of CXCR4), TAK779 (inhibitor of CCR5), or both prior to addition of the virus preparations indicated. Virus infectivity is plotted on the vertical axis as a percentage of the untreated control. Pseudovirions containing the CXCR4-tropic NL4.3, the CCR5-tropic YU2, and the dual-tropic WEAU envelope glycoproteins were included as controls.

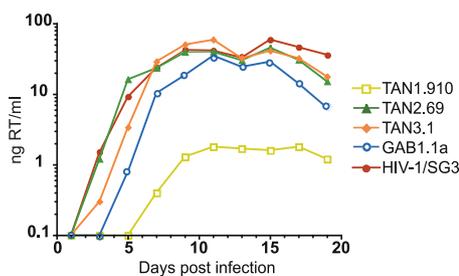


FIG. 8. Replication potential of TAN1, TAN2, and TAN3 molecular clones in chimpanzee CD4<sup>+</sup> T cells. The replication kinetics of TAN1.910, TAN2.69, and TAN3.1 are shown in relation to HIV-1 (SG3) and SIVcpzPtt (GAB1.1a) reference strains (y axis, nanograms of RT activity per ml of culture supernatant). Enriched CD4<sup>+</sup> T lymphocytes from chimpanzee C0462 were infected with transfection-derived culture supernatants equilibrated by particle-associated RT activity (50 ng).

TABLE 4. Replication potential of SIVcpzP<sub>ts</sub> strains in chimpanzee CD4<sup>+</sup> T cells<sup>a</sup>

Strain	Replication titers <sup>b</sup> (%) for chimpanzee:							
	C0415	C0526	C0522	C0726	C0495	C0724	C0535	C0462
HIV-1/SG3	100	100	100	100	100	100	100	100
GAB1.1a	8	100	Neg	25	ND	30	ND	59
TAN1.910	Neg	Neg	Neg	Neg	ND	Neg	50	3
TAN2.69	ND	ND	ND	74	47	50	ND	77
TAN3.1	Neg	Neg	Neg	Neg	12	Neg	120	100

<sup>a</sup> Replication titers are expressed relative to HIV-1/SG3 set to 100%; maximal titers are shown.

<sup>b</sup> ND; not done; Neg, negative.

genomic RT-PCR fragments by extension overlap PCR. Although this approach generated a full-length provirus with seemingly functional genes, the resulting particles were replication defective. Reasoning that the PCR-based concatenation process had introduced inactivating mutations, we explored whether changing the proviral sequence to that of the corresponding consensus sequence would restore its biological activity. This was done in a stepwise fashion to determine how many sequence changes were required. The results showed that the defective phenotype of TAN1.24 was caused by substitutions throughout its genome, all of which contributed to its replication defectiveness. Thus, the particular PCR-based concatenation scheme that we used does not seem suitable for generating infectious molecular clones of SIV. However, it may be possible to improve this approach by (i) concatenating multiple cDNA fragments simultaneously, (ii) reducing the number of cumulative amplification cycles, and (iii) using a

higher-fidelity polymerase. Such a revised protocol may produce a more manageable number of misincorporations (and other ambiguities) that can more reasonably be addressed by site-directed mutagenesis.

The second approach to generate infectious SIVcpz clones involved the direct synthesis of fecal consensus sequences. The rationale for this approach was severalfold. First, we inferred from the mutagenesis experiments of TAN1.24 that fecal population sequences were a suitable template for functional genome design. Second, we assumed that fecal viral RNA, like plasma viral RNA, represents replicating virus produced by CD4<sup>+</sup> T cells in the gut-associated lymphatic system (6, 34, 55). Finally, by generating a simple majority consensus sequence, i.e., by determining the most frequent nucleotide at each position within the viral population, we expected to eliminate most (if not all) inactivating mutations present among individual quasispecies members. These assumptions proved to be

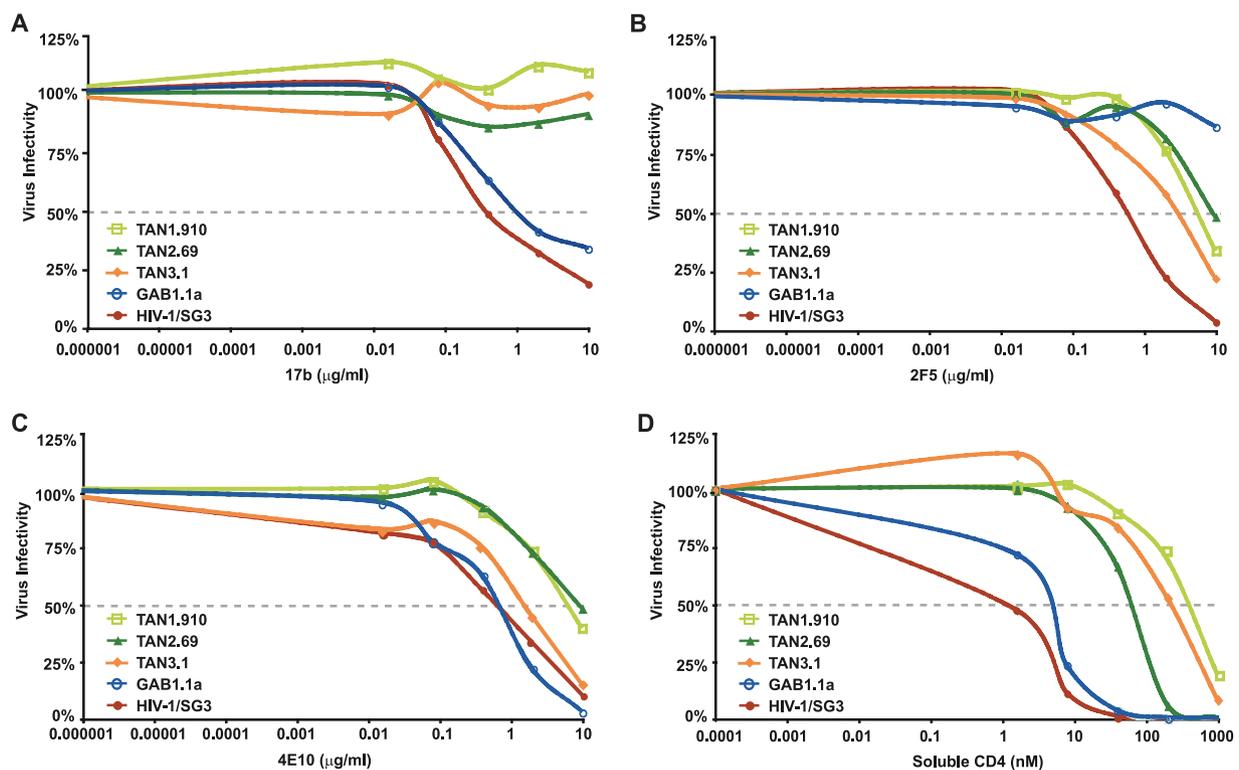


FIG. 9. Sensitivity of SIVcpzP<sub>ts</sub> strains to neutralization by human monoclonal antibodies and soluble CD4. Neutralization of virus infectivity in JC53-BL cells (y axis) was scored as the concentration of human monoclonal antibodies 17b (A), 2F5 (B), and 4E10 (C) as well as soluble CD4 (D) required to reduce virus infectivity by 50% (IC<sub>50</sub>) (gray line).

correct since the chemical synthesis of both TAN2 and TAN3 consensus sequences resulted in infectious viral genomes. The fact that TAN2 required modification at one ambiguous site highlights the need for careful inspection of the population sequences for positions of base mixtures and a suitable algorithm to resolve them. However, there are caveats to this virus reconstruction approach. Amplification of viral sequences under limiting dilution condition (e.g., because of partial sample degradation, low fecal viral loads, and/or primer mismatch) may generate sequences that are not fully representative of the population consensus and may thus not be functional. Thus, virus reconstruction from fecal consensus sequences should be performed under conditions of high template input with sufficiently cross-reactive primers.

Derivation of the first infectious molecular clones of SIVcpzP<sub>ts</sub> allowed us to compare their baseline biological properties to those of HIV-1 and SIVcpz reference strains. These analyses revealed a number of similarities with primary HIV-1 isolates, including receptor and coreceptor preference, cell tropism, and neutralization phenotype. However, differences in viral infectivity and replication potential were also observed (Fig. 5 and 8). Most notably, TAN1.910 exhibited reduced replication fitness compared to the other two SIVcpzP<sub>ts</sub> strains in both human and chimpanzee CD4<sup>+</sup> T cells, possibly because of the remaining 40 (silent) nucleotide sequence differences from the consensus sequence. The various SIVcpzP<sub>ts</sub> strains also differed in their abilities to infect and replicate in CD4<sup>+</sup> T cells from different chimpanzee donors (Table 4). Given that chimpanzees are the natural host of SIVcpz, these results were surprising. Moreover, the fact that PBMC cultures from every human donor tested supported the replication of these same SIVcpzP<sub>ts</sub> strains suggests that chimpanzees more than humans vary in their susceptibilities to HIV/SIV infection. This may explain the uneven distribution and low prevalence of SIVcpz in certain chimpanzee communities (31, 48, 51). The fact that GAB1 replicated in lymphocyte cultures of most chimpanzees tested (Table 4) does not contradict this possibility, since this in vitro-adapted virus is not representative of naturally occurring SIVcpz strains.

The noninvasive derivation of infectious molecular clones of SIVcpz has important implications. Most importantly, this new technology will permit the biological characterization of SIVs (and their gene products) that have thus far only been found in wild-living endangered primate species. Examples include the SIVcpzP<sub>tt</sub> precursors of pandemic (group M) and nonpandemic (group N) HIV-1 (31) and the newly discovered SIVgor lineage, infecting wild gorillas (54), as well as a multitude of other SIVs from different primate species that to date have been characterized only at the molecular level (4). Biological analysis of these viruses promises to yield new insight into their pathogenic potential as well as viral determinants that govern cross-species transmission and host adaptation. Specifically, the new technology should permit investigation of whether certain strains or lineages of SIV are more prone than others to infect, replicate, and spread in humans. To determine whether such virological differences exist, it will be important to test representative strains in more-complex biological systems, such as lymphoid (tonsil) explant cultures, which maintain their cell complexity and cytoarchitecture (20), are not dependent on exogenous stimulation (20), and require func-

tional accessory proteins for efficient viral replication and cytopathicity (37, 43, 45). Humanized mouse models may also represent an appropriate test system for dissecting biological differences. An example is the recently described BLT mouse, in which primary and secondary mouse lymphoid organs are reconstituted with functional human B and T cells, monocytes, macrophages, and dendritic cells (36). One prerequisite for these studies is the availability of a sufficiently representative set of infectious molecular clones that have not been subjected to in vitro selection pressures.

In summary, we describe here a new strategy of deriving replication-competent molecular clones of SIV from fecal consensus sequences. This approach is conceptually similar to the recent successful "resurrection" of biologically active ancestors of endogenous human and ape retroviruses (15, 30, 33) and emphasizes the advantage of consensus over individual viral quasispecies, the latter of which may contain deleterious mutations. In addition, the new strategy generates viruses that are truly representative of SIVcpz strains in wild ape populations since they are derived without in vitro propagation in human cells. Analysis of such unadulterated "isolates" will permit the identification of biological properties that are essential for cross-species transmission and human infection. The present study demonstrates the feasibility of such an undertaking and provides a framework for generating the necessary reagents.

#### ACKNOWLEDGMENTS

We thank field researchers from the Gombe Stream Research Centre for collecting fecal samples from wild chimpanzees; the Yerkes Primate Center staff for shipping blood samples from uninfected captive chimpanzees; Stephen Hughes, Bing Chen, Steven Harrison, Frank Kirchoff, and Victor Garcia for helpful discussions; 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 (R01 AI50529, R01 AI58715, P30 AI 27767), the Yerkes Regional Primate Research Center (RR-00165), and the Bristol Myers Freedom to Discover Program.

#### REFERENCES

1. Ayouba, A., P. Mauciere, P. M. Martin, P. Cunin, J. Mfoupoendoun, B. Njinku, S. Souquieres, and F. Simon. 2001. HIV-1 group O infection in Cameroon, 1986 to 1998. *Emerg. Infect. Dis.* 7:466-467.
2. Bailes, E., R. R. Chaudhuri, M. L. Santiago, F. Bibollet-Ruche, B. H. Hahn, and P. M. Sharp. 2002. The evolution of primate lentiviruses and the origins of AIDS, p. 65-96. *In* T. A. Leitner (ed.), *The Molecular Epidemiology of Human Viruses*. Kluwer Academic Publishers, Boston, MA.
3. Bailes, E., F. Gao, F. Bibollet-Ruche, V. Cournaud, M. Peeters, P. A. Marx, B. H. Hahn, and P. M. Sharp. 2003. Hybrid origin of SIV in chimpanzees. *Science* 300:1713.
4. Bibollet-Ruche, F., E. Bailes, F. Gao, X. Pourrut, K. L. Barlow, J. P. Clewley, J. Mwenda, D. K. Langat, G. K. Chege, H. M. McClure, E. Mpoudi-Ngole, E. Delaport, M. Peeters, G. M. Shaw, P. M. Sharp, and B. Hahn. 2004. New simian immunodeficiency virus infecting De Brazza's monkeys (*Cercopithecus neglectus*): evidence for a *Cercopithecus* monkey virus clade. *J. Virol.* 78:7748-7762.
5. Binley, J. M., T. Wrin, B. Korber, M. B. Zwick, M. Wang, C. Chappey, G. Stiegler, R. Kunert, S. Zolla-Pazner, H. Katinger, C. J. Petropoulos, and D. R. Burton. 2004. Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J. Virol.* 78:13232-13252.
6. Brenchley, J. M., T. W. Schacker, L. E. Ruff, D. A. Price, J. H. Taylor, G. J. Beilman, P. L. Nguyen, A. Khoruts, M. Larson, A. T. Haase, and D. C. Douek. 2004. CD4<sup>+</sup> T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J. Exp. Med.* 200:749-759.
7. Brennan, C. A., P. Bodelle, R. Coffey, B. Harris, V. Holzmayr, K. C. Luk, P. Swanson, J. Yamaguchi, A. Vallari, S. G. Devare, G. Schochetman, and J. Hackett, Jr. 2006. HIV global surveillance: foundation for retroviral discovery and assay development. *J. Med. Virol.* 78 Suppl. 1:S24-S29.
8. Brunel, F. M., M. B. Zwick, R. M. Cardoso, J. D. Nelson, I. A. Wilson, D. R.

- Burton, and P. E. Dawson. 2006. Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. *J. Virol.* **80**:1680–1687.
9. Butler, I. F., I. Pandrea, P. A. Marx, and C. Apetrei. 2007. HIV genetic diversity: biological and public health consequences. *Curr. HIV Res.* **5**:23–45.
10. Butynski, T. M. 2001. Africa's great apes. In B. Beck, T. Stoinski, M. Hutchins, T. L. Maple, B. G. Norton, A. Rowan, E. F. Stevens, and A. Arluke (ed.), *Apes and Humans at an Ethical Frontier*. Smithsonian Institution Press, Washington, DC.
11. CDC. 2006. The global HIV/AIDS pandemic, 2006. *Morb. Mortal. Wkly. Rep.* **55**:841–844.
12. Cello, J., A. V. Paul, and E. Wimmer. 2002. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297**:1016–1018.
13. Corbet, S., M. C. Muller-Trutwin, P. Versmisse, S. Delarue, A. Ayouba, J. Lewis, S. Brunak, P. Martin, F. Brun-Vezinet, F. Simon, F. Barre-Sinoussi, and P. Maucelere. 2000. *env* sequences of simian immunodeficiency viruses from chimpanzees in Cameroon are strongly related to those of human immunodeficiency virus group N from the same geographic area. *J. Virol.* **74**:529–534.
14. Derdeyn, C. A., J. M. Decker, J. N. Sfakianos, X. Wu, W. A. O'Brien, L. Ratner, J. C. Kappes, G. M. Shaw, and E. Hunter. 2000. Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *J. Virol.* **74**:8358–8367.
15. Dewannieux, M., F. Harper, A. Richaud, C. Letzelter, D. Ribet, G. Pierron, and T. Heidmann. 2006. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res.* **16**:1548–1556.
16. Fang, G., B. Weiser, A. Visosky, T. Moran, and H. Burger. 1999. PCR-mediated recombination: a general method applied to construct chimeric infectious molecular clones of plasma-derived HIV-1 RNA. *Nat. Med.* **5**:239–242.
17. Fischer, A., J. Pollack, O. Thalmann, B. Nickel, and S. Paabo. 2006. Demographic history and genetic differentiation in apes. *Curr. Biol.* **16**:1133–1138.
18. Gao, F., Y. Li, J. M. Decker, F. W. Peyerl, F. Bibollet-Ruche, C. M. Rodenburg, Y. Chen, D. R. Shaw, S. Allen, R. Musonda, G. M. Shaw, A. J. Zajac, N. Letvin, and B. H. Hahn. 2003. Codon usage optimization of HIV type 1 subtype C gag, pol, env, and nef genes: in vitro expression and immune responses in DNA-vaccinated mice. *AIDS Res. Hum. Retroviruses* **19**:817–823.
19. Ghosh, S. K., P. N. Fultz, E. Keddie, M. S. Saag, P. M. Sharp, B. H. Hahn, and G. M. Shaw. 1993. A molecular clone of HIV-1 tropic and cytopathic for human and chimpanzee lymphocytes. *Virology* **194**:858–864.
20. Glushakova, S., B. Baibakov, J. Zimmerberg, and L. B. Margolis. 1997. Experimental HIV infection of human lymphoid tissue: correlation of CD4<sup>+</sup> T cell depletion and virus syncytium-inducing/non-syncytium-inducing phenotype in histocultures inoculated with laboratory strains and patient isolates of HIV type 1. *AIDS Res. Hum. Retroviruses* **13**:461–471.
21. Grandgenett, D. P., and G. Goodarzi. 1994. Folding of the multidomain human immunodeficiency virus type I integrase. *Protein Sci.* **3**:888–897.
22. Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp. 2000. AIDS as a zoonosis: scientific and public health implications. *Science* **287**:607–617.
23. Hemelaar, J., E. Gouws, P. D. Ghys, and S. Osmanov. 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**:W13–W23.
24. Hirsch, V. M., B. J. Campbell, E. Bailes, R. Goeken, C. Brown, W. R. Elkins, M. Axthelm, M. Murphy-Corb, and P. M. Sharp. 1999. Characterization of a novel simian immunodeficiency virus (SIV) from L'Hoest monkeys (*Cercopithecus lhoesti*): implications for the origins of SIVmnd and other primate lentiviruses. *J. Virol.* **73**:1036–1045.
25. Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**:123–126.
26. Huang, H., R. Chopra, G. L. Verdine, and S. C. Harrison. 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**:1669–1675.
27. Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**:754–755.
28. Huet, T., R. Cheyner, A. Meyerhans, G. Roelants, and S. Wain-Hobson. 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* **345**:356–359.
29. Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**:275–282.
30. Kaiser, S. M., H. S. Malik, and M. Emerman. Restriction of an extinct retrovirus by the human Trim5 $\alpha$  antiviral protein. *Science*, in press.
31. Keele, B. F., F. Van Heuverswyn, Y. Li, E. Bailes, J. Takehisa, M. L. Santiago, F. Bibollet-Ruche, Y. Chen, L. V. Wain, F. Liegeois, S. Loul, E. M. Ngole, Y. Bienvenue, E. Delaporte, J. F. Brookfield, P. M. Sharp, G. M. Shaw, M. Peeters, and B. H. Hahn. 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* **313**:523–526.
32. Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**:648–659.
33. Lee, Y. N., and P. D. Bieniasz. 2007. Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog.* **3**:e10. doi:10.1371/journal.ppat.0030010.
34. Li, Q., L. Duan, J. D. Estes, Z. M. Ma, T. Rourke, Y. Wang, C. Reilly, J. Carlis, C. J. Miller, and A. T. Haase. 2005. Peak SIV replication in resting memory CD4<sup>+</sup> T cells depletes gut lamina propria CD4<sup>+</sup> T cells. *Nature* **434**:1148–1152.
35. McCutchan, F. E. 2006. Global epidemiology of HIV. *J. Med. Virol.* **78**:S7–S12.
36. Melkus, M. W., J. D. Estes, A. Padgett-Thomas, J. Gatlin, P. W. Denton, F. A. Othieno, A. K. Wege, A. T. Haase, and J. V. Garcia. 2006. Humanized mice mount specific adaptive and innate immune responses to EBV and TSST-1. *Nat. Med.* **12**:1316–1322.
37. Munch, J., D. Rajan, E. Rucker, S. Wildum, N. Adam, and F. Kirchhoff. 2005. The role of upstream U3 sequences in HIV-1 replication and CD4<sup>+</sup> T cell depletion in human lymphoid tissue ex vivo. *Virology* **341**:313–320.
38. Nerrienet, E., M. L. Santiago, Y. Foupouapouognigni, E. Bailes, N. I. Mundy, B. Njinku, A. Kfutwah, M. C. Muller-Trutwin, F. Barre-Sinoussi, G. M. Shaw, P. M. Sharp, B. H. Hahn, and A. Ayouba. 2005. Simian immunodeficiency virus infection in wild-caught chimpanzees from Cameroon. *J. Virol.* **79**:1312–1319.
39. Pantophlet, R., and D. R. Burton. 2006. GP120: target for neutralizing HIV-1 antibodies. *Annu. Rev. Immunol.* **24**:739–769.
40. Peeters, M., K. Fransen, E. Delaporte, M. Van den Haesevelde, G.-M. Gershy-Damet, L. Kestens, G. van der Groen, and P. Piot. 1992. Isolation and characterization of a new chimpanzee lentivirus (simian immunodeficiency virus isolate cpz-ant) from a wild-captured chimpanzee. *AIDS* **6**:447–451.
41. Peeters, M., C. Honore, T. Huet, L. Bedjabaga, S. Ossari, P. Bussi, R. W. Cooper, and E. Delaporte. 1989. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *AIDS* **3**:625–630.
42. Platt, E. J., K. Wehrly, S. E. Kuhmann, B. Chesebro, and D. Kabat. 1998. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J. Virol.* **72**:2855–2864.
43. Rajan, D., S. Wildum, E. Rucker, M. Schindler, and F. Kirchhoff. 2006. Effect of R77Q, R77A and R80A changes in Vpr on HIV-1 replication and CD4 T cell depletion in human lymphoid tissue ex vivo. *AIDS* **20**:831–836.
44. Roques, P., D. L. Robertson, S. Souquiere, C. Apetrei, E. Nerrienet, F. Barre-Sinoussi, M. Muller-Trutwin, and F. Simon. 2004. Phylogenetic characteristics of three new HIV-1 N strains and implications for the origin of group N. *AIDS* **18**:1371–1381.
45. Rucker, E., J. C. Grivel, J. Munch, F. Kirchhoff, and L. Margolis. 2004. Vpr and Vpu are important for efficient human immunodeficiency virus type 1 replication and CD4<sup>+</sup> T-cell depletion in human lymphoid tissue ex vivo. *J. Virol.* **78**:12689–12693.
46. Santiago, M. L., F. Bibollet-Ruche, E. Bailes, S. Kamenya, M. N. Muller, M. Lukasik, A. E. Pusey, D. A. Collins, R. W. Wrangham, J. Goodall, G. M. Shaw, P. M. Sharp, and B. H. Hahn. 2003. Amplification of a complete simian immunodeficiency virus genome from fecal RNA of a wild chimpanzee. *J. Virol.* **77**:2233–2242.
47. Santiago, M. L., F. Bibollet-Ruche, N. Gross-Camp, A. C. Majewski, M. Masozera, I. Munanura, B. A. Kaplin, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 2003. Noninvasive detection of simian immunodeficiency virus infection in a wild-living L'Hoest's monkey (*Cercopithecus lhoesti*). *AIDS Res. Hum. Retroviruses* **19**:1163–1166.
48. Santiago, M. L., M. Lukasik, S. Kamenya, Y. Li, F. Bibollet-Ruche, E. Bailes, M. N. Muller, M. Emery, D. A. Goldenberg, J. S. Lwanga, A. Ayouba, E. Nerrienet, H. M. McClure, J. L. Heeney, D. P. Watts, A. E. Pusey, D. A. Collins, R. W. Wrangham, J. Goodall, J. F. Brookfield, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 2003. Foci of endemic simian immunodeficiency virus infection in wild-living eastern chimpanzees (*Pan troglodytes schweinfurthii*). *J. Virol.* **77**:7545–7562.
49. Santiago, M. L., F. Range, B. F. Keele, Y. Li, E. Bailes, F. Bibollet-Ruche, C. Fruteau, R. Noe, M. Peeters, J. F. Brookfield, G. M. Shaw, P. M. Sharp, and B. H. Hahn. 2005. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (*Cercocebus atys atys*) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *J. Virol.* **79**:12515–12527.
50. Santiago, M. L., C. M. Rodenburg, S. Kamenya, F. Bibollet-Ruche, F. Gao, E. Bailes, S. Meleth, S. J. Soong, J. M. Kilby, Z. Moldoveanu, B. Fahey, M. N. Muller, A. Ayouba, E. Nerrienet, H. M. McClure, J. L. Heeney, A. E. Pusey, D. A. Collins, C. Boesch, R. W. Wrangham, J. Goodall, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 2002. SIVcpz in wild chimpanzees. *Science* **295**:465.
51. Sharp, P. M., G. M. Shaw, and B. H. Hahn. 2005. Simian immunodeficiency virus infection of chimpanzees. *J. Virol.* **79**:3891–3902.
52. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W—improving the sensitivity of progressive multiple sequence alignment through

- sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
53. **Tumpey, T. M., C. F. Basler, P. V. Aguilar, H. Zeng, A. Solorzano, D. E. Swayne, N. J. Cox, J. M. Katz, J. K. Taubenberger, P. Palese, and A. Garcia-Sastre.** 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**:77–80.
  54. **Van Heuverswyn, F., Y. Li, C. Neel, E. Bailes, B. F. Keele, W. Liu, S. Loul, C. Butel, F. Liegeois, Y. Bienvenue, E. M. Ngolle, P. M. Sharp, G. M. Shaw, E. Delaporte, B. H. Hahn, and M. Peeters.** 2006. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature* **444**:164.
  55. **Veazey, R. S., M. DeMaria, L. V. Chalifoux, D. E. Shvetz, D. R. Pauley, H. L. Knight, M. Rosenzweig, R. P. Johnson, R. C. Desrosiers, and A. A. Lackner.** 1998. Gastrointestinal tract as a major site of CD4<sup>+</sup> T cell depletion and viral replication in SIV infection. *Science* **280**:427–431.
  56. **Wei, X., J. M. Decker, H. Liu, Z. Zhang, R. B. Arani, J. M. Kilby, M. S. Saag, X. Wu, G. M. Shaw, and J. C. Kappes.** 2002. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob. Agents Chemother.* **46**:1896–1905.
  57. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**:117–122.
  58. **Worobey, M.** The origins and diversity of HIV. *In* M. Sande, J. Lange, P. Volberding, and W. C. Greene (ed.), *Global HIV/AIDS medicine*, in press. Elsevier, Philadelphia, PA.
  59. **Worobey, M., M. L. Santiago, B. F. Keele, J. B. Ndjango, J. B. Joy, B. L. Labama, B. D. Dheda, A. Rambaut, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 2004. Origin of AIDS: contaminated polio vaccine theory refuted. *Nature* **428**:820.
  60. **Xiang, S. H., L. Wang, M. Abreu, C. C. Huang, P. D. Kwong, E. Rosenberg, J. E. Robinson, and J. Sodroski.** 2003. Epitope mapping and characterization of a novel CD4-induced human monoclonal antibody capable of neutralizing primary HIV-1 strains. *Virology* **315**:124–134.
  61. **Yamaguchi, J., P. Bodelle, A. S. Vallari, R. Coffey, C. P. McArthur, G. Schochetman, S. G. Devare, and C. A. Brennan.** 2004. HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and group O infections. *AIDS Res. Hum. Retroviruses* **20**:944–957.
  62. **Yamaguchi, J., R. Coffey, A. Vallari, C. Ngansop, D. Mbanya, N. Ndembu, L. Kaptue, L. G. Gurtler, P. Bodelle, G. Schochetman, S. G. Devare, and C. A. Brennan.** 2006. Identification of HIV type 1 group N infections in a husband and wife in Cameroon: viral genome sequences provide evidence for horizontal transmission. *AIDS Res. Hum. Retroviruses* **22**:83–92.
  63. **Yang, Z.** 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**:306–314.
  64. **Yang, Z., and B. Rannala.** 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Mol. Biol. Evol.* **14**:717–724.
  65. **Young, L., and Q. Dong.** 2004. Two-step total gene synthesis method. *Nucleic Acids Res.* **32**:e59. doi:10.1093/nar/gnh058.
  66. **Zwick, M. B., R. Jensen, S. Church, M. Wang, G. Stiegler, R. Kunert, H. Katinger, and D. R. Burton.** 2005. Anti-human immunodeficiency virus type 1 (HIV-1) antibodies 2F5 and 4E10 require surprisingly few crucial residues in the membrane-proximal external region of glycoprotein *gp41* to neutralize HIV-1. *J. Virol.* **79**:1252–1261.