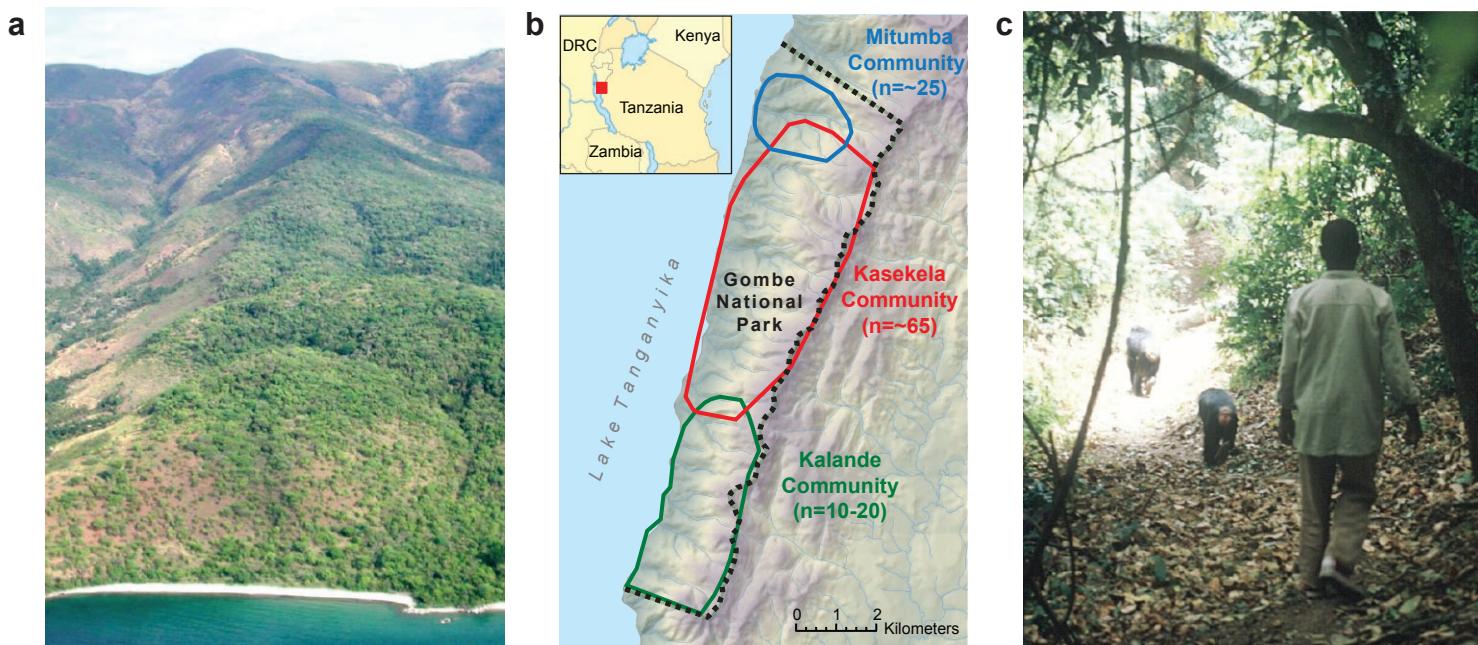
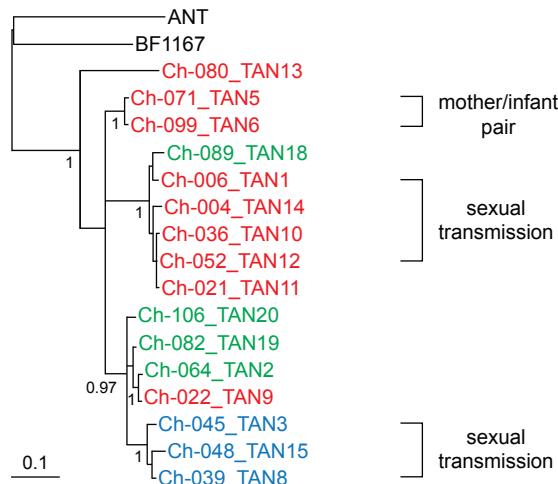
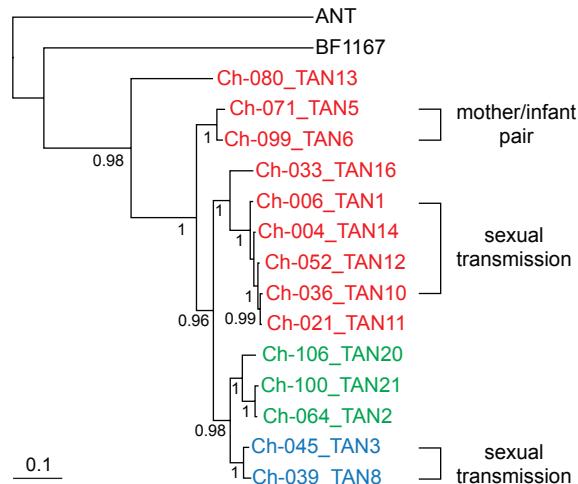
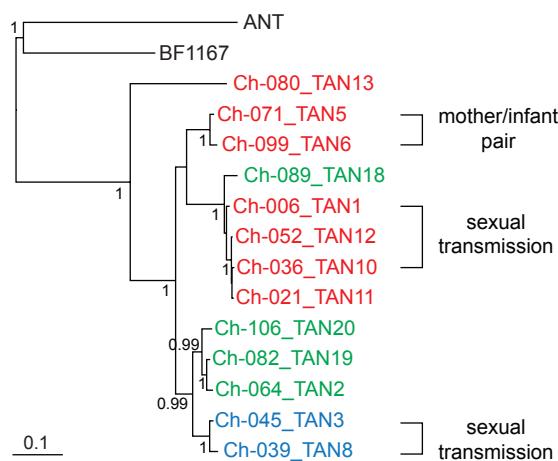
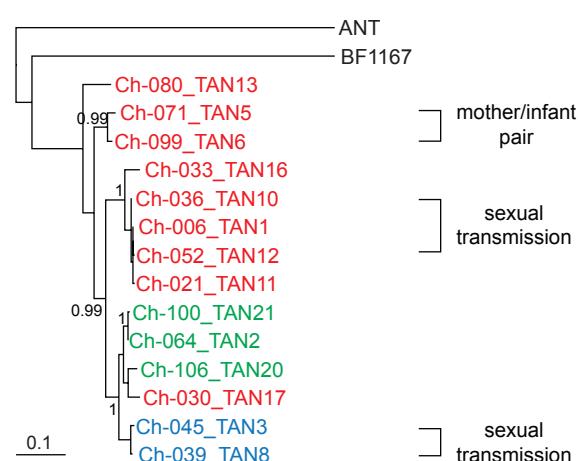


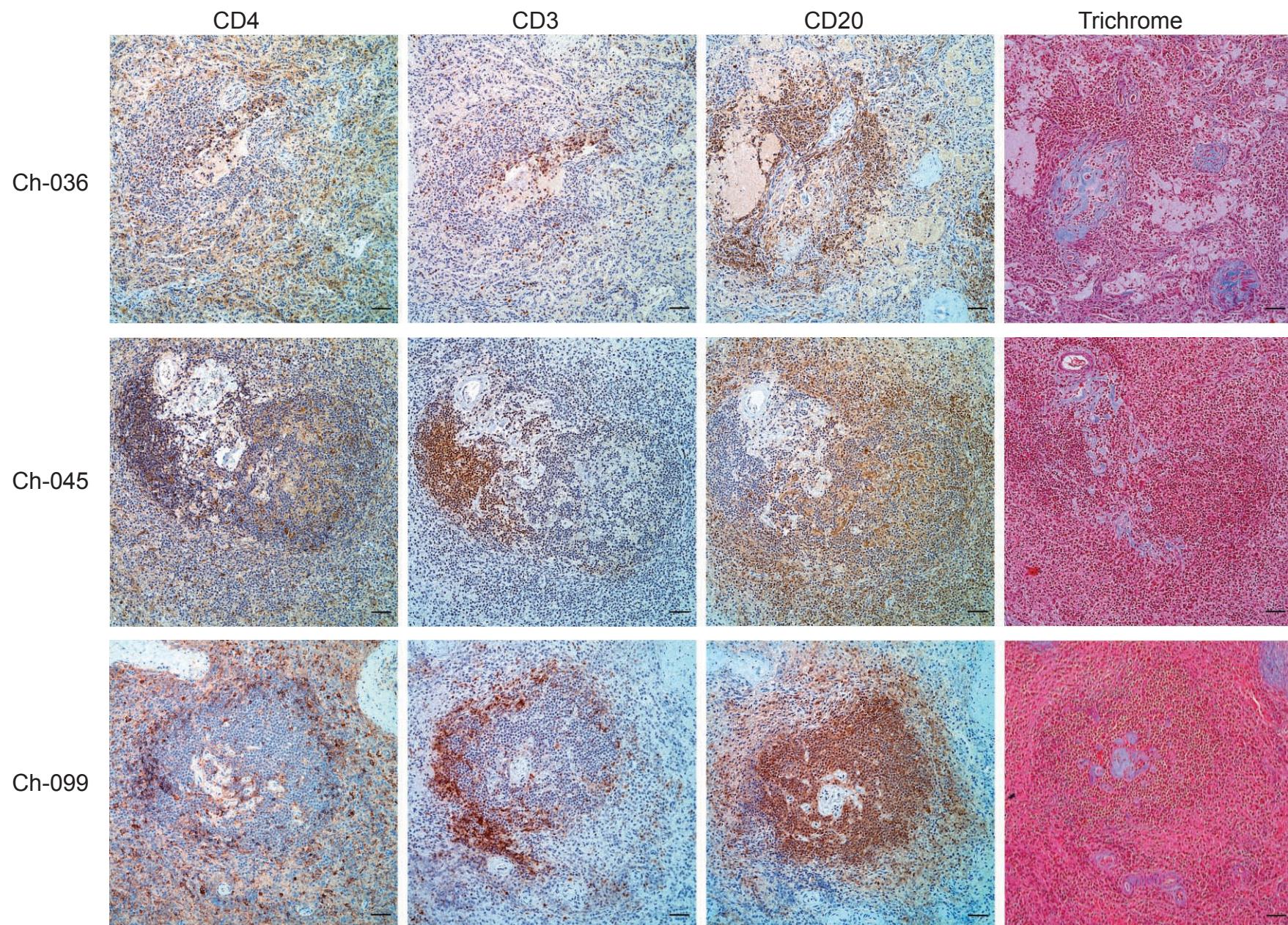
SUPPLEMENTARY INFORMATION

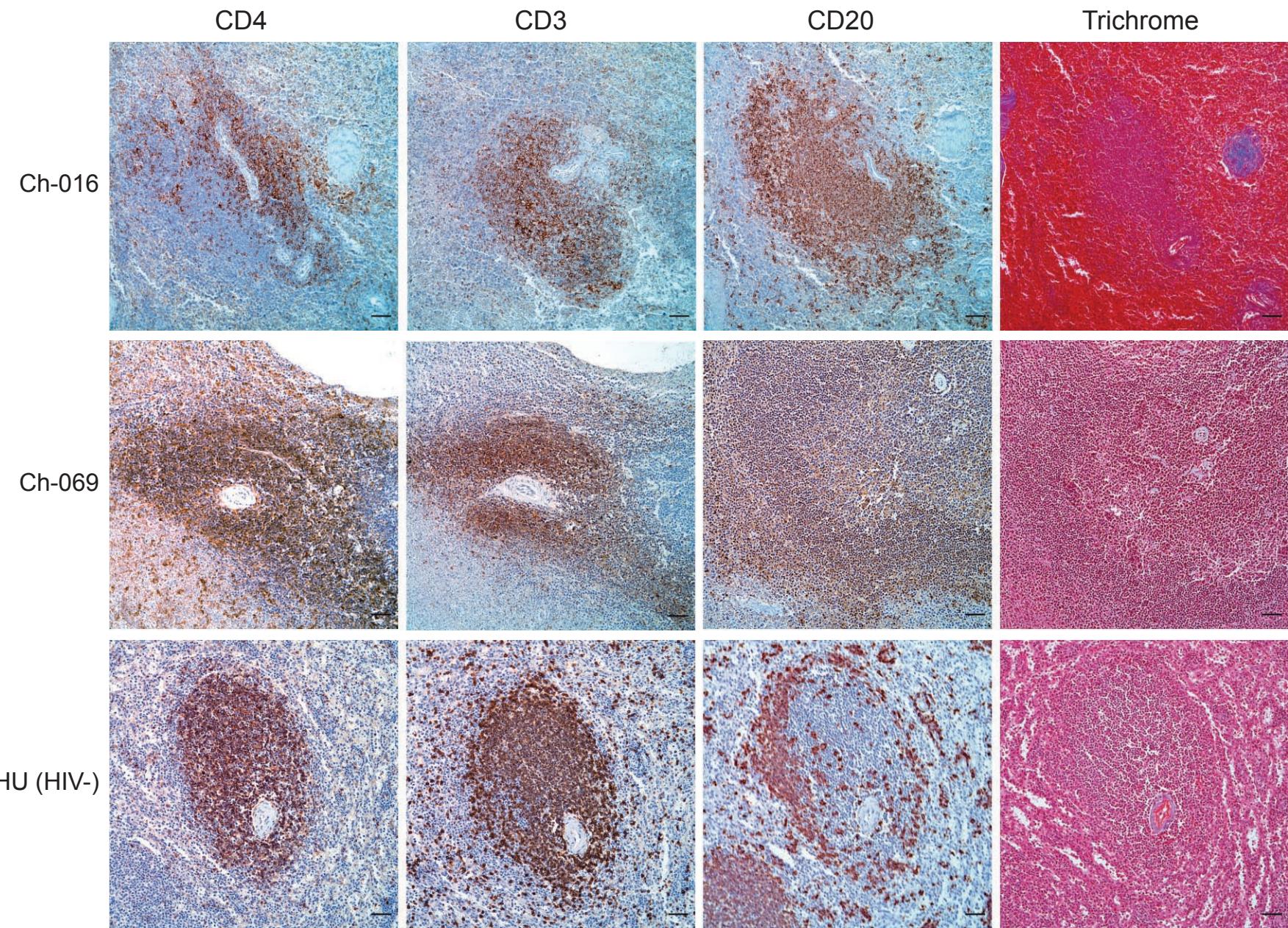


Supplementary Figure 1. Studies of wild chimpanzees in Gombe National Park. **a**, Aerial view of the northern border of the park from Lake Tanganyika depicting extensive deforestation outside the protected area (photograph courtesy of John MacLachlan, The Jane Goodall Institute). **b**, Approximate ranges of the northern Mitumba (blue), the central Kasekela (red), and the southern Kalande (green) communities in relation to park boundaries (black); the inset depicts the location of Gombe National Park (red square) within Tanzania (map courtesy of Lilian Pintea, The Jane Goodall Institute). **c**, Daily observation of habituated chimpanzees in Gombe (photograph courtesy of Magdalena Lukasik, The Jane Goodall Institute).

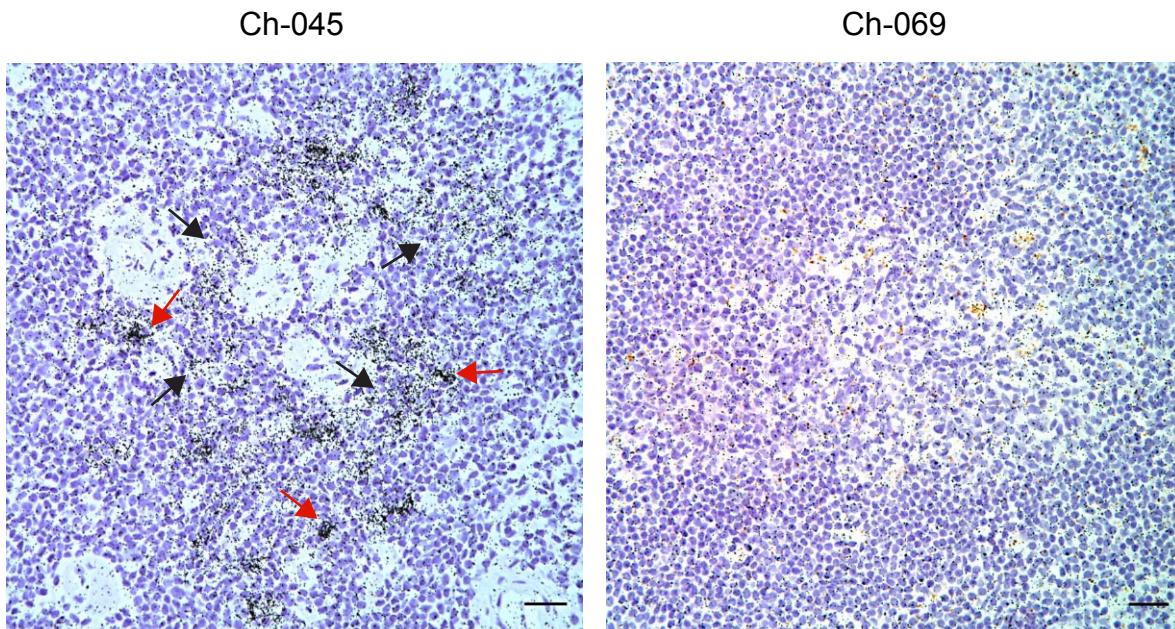
a pol (481bp)**b env-nef (609bp)****c pol (854bp)****d env (279bp)**

Supplementary Figure 2. Phylogeny of SIVcpz in Gombe. **a-d**, Evolutionary trees were constructed from (a) 481-bp pol, (b) 609-bp env-nef, (c) 854-bp pol and (d) 279-bp env (gp41 region) sequences, using SIVcpzP^ts strains from the Democratic Republic of Congo (ANT and BF1167) as outgroups. Viruses infecting members of the Mitumba, Kasekela and Kalande communities are shown in blue, red and green, respectively (GenBank accession numbers are provided in Supplementary Table 1). Trees were inferred by Bayesian²⁸ (a, b) and maximum likelihood³⁴ (c,d) methods; numbers on nodes are posterior probabilities (only values above 0.95 are shown). The scale bar represents 0.1 substitutions per site. Brackets indicate suspected sexual and vertical transmission clusters (note that Ch-021 was most likely infected through aggression; Supplementary Table 2).

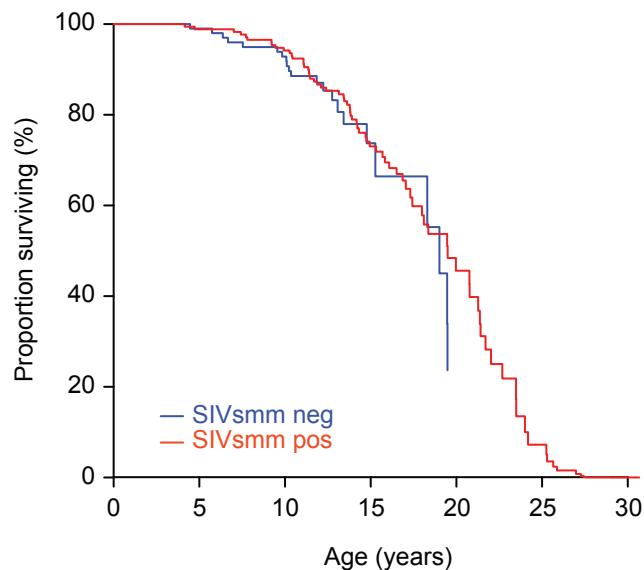
Supplementary Figure 3a.

Supplementary Figure 3b.

Supplementary Figure 3. CD4+ T cell depletion and collagen deposition in the spleen of SIVcpz infected chimpanzees. **a, b,** Representative images of CD4 (T helper cells), CD3 (pan T cells), CD20 (B cells) and Trichrome (collagen) stained spleen sections are shown for **(a)** three SIVcpz infected chimpanzees (Ch-036, Ch-045 and Ch-099), as well as **(b)** two uninfected chimpanzees (Ch-016, Ch-069) and one normal human (Hu HIV-) control. Each row represents consecutive sections of the same periarteriolar lymphoid sheath (PALS) region (the scale bar indicates 50 μ m). In sections stained with CD4, CD3 and CD20 antibodies, positive cells are dark brown (cell nuclei are blue). In Masson's Trichrome stained sections, collagen is dark blue (cytoplasm and nuclei are red-magenta). Note the varying degrees of CD4+ T cell depletion and associated collagen deposition in Ch-036 (severe), Ch-045 (intermediate) and Ch-099 (minor), and the fact that the CD4+ T cell loss in Ch-045 and Ch-099 occurred in the absence of a concomitant depletion of other CD3+ T cell and B cell populations.



Supplementary Figure 4. SIVcpz replication *in vivo*. Spleen sections from an SIVcpz infected (Ch-045) and uninfected (Ch-069) chimpanzee were subjected to *in situ* hybridization (ISH) using ^{35}S -labelled SIVcpz*Pts* specific RNA riboprobes. SIVcpz RNA is shown as black silver grains (nuclei are stained blue). Red arrows highlight individual cells that express copious amounts of viral RNA (areas of very dense silver grains); black arrows indicate extensive trapping of virus particles by follicular dendritic cells (areas of more diffuse silver grains). Note that ISH also identified a few productively infected cells in the spleen of Ch-036, but not Ch-099 (possibly due to partial tissue degradation). Original magnification 400x (the scale bar is 25 μm).



Supplementary Figure 5. SIVsmm associated mortality in captive sooty mangabeys. Kaplan-Meier survival curves are shown for SIVsmm infected ($n=167$; red) and uninfected ($n=62$; blue) sooty mangabeys housed at the Yerkes National Primate Research Center. The mortality rates of the two groups do not differ significantly ($p=0.55$). The proportion of surviving sooty mangabeys is shown in relation to their age in years.

Individual ¹	Sex ²	Community ³	Sample type	Sample Code	Collection Date ⁴	mtDNA haplotype ⁵	Microsatellite Loci ⁶				GenBank accession numbers ⁹					
							D18s536	D4s243	D10s676	D9s922	urine WB ⁷	fecal WB	fecal vRNA ⁸	<i>pol</i>	gp41/ <i>nef</i>	SIVcpz Strain
Ch-114	F	KK	fecal	1319	27-Jun-07	11	141/173	196/235	182/190	286/306	neg	neg	neg	neg	neg	DOB (9-Sep-06)
							141/173	196/235	182/190	286/306						
Ch-115	M	KK	fecal	1327	17-May-07	3	141/161	204/204	186/192	302/306	neg	neg	neg	neg	neg	DOB (16-Oct-05)
							141/161	204/204	186/192	302/306						
Ch-116	M	KK	fecal	1488	24-Oct-08	5	141/157	196/231	190/190	286/298	neg	neg	neg	neg	neg	DOB (9-Nov-06)
							141/157	196/231	190/190	286/298						
Ch-117	F	KK	fecal	1417	10-Sep-07	11	153/161	204/204	186/190	286/298	neg	neg	neg	neg	neg	DOB (4-Jun-05)

¹Black, uninjected; red, infected.²F, female; M, male.³Community in which sample was collected: KK, Kasekela; MT, Mitumba; KL, Kalande.⁴na, not available.⁵Numbers indicate mtDNA haplotypes as previously reported⁸.⁶Fecal samples that were genotyped at a minimum of four microsatellite loci are shown in bold.⁷pos, positive; neg, negative.⁸FL, full-length genome sequence.⁹Previously reported SIVcpz sequences are shown in italics^{7,32}¹⁰DOB, date of birth (estimated DOBs¹² are denoted by asterisk), DOD, date of death; DLS, date last seen.

Supplementary Table 2. Horizontal transmission of SIVcpz

Individual	Sex ¹	Community ²	Last SIVcpz negative sample	First SIVcpz positive sample	Age at infection (years)	Potential source of infection (sex) ³	Observed matings ⁴	Observed acts of aggression ⁵
Ch-004	M	KK	5-Nov-05	12-Aug-06	13.3	Ch-036 (F) Ch-021 (F) Ch-006 (M) Ch-052 (M)	none ⁶ none n/a n/a Ch-033 (22-Sep-05) Ch-022 (5-Sep-05) Ch-022 (7-Oct-05) Ch-022 (25-Jan-06) Ch-022 (24-May-06)	none
Ch-021	F	KK	5-Dec-04	11-May-05	15.9	Ch-052 (M) Ch-006 (M) Ch-004 (M) Ch-036 (F)	none ⁷ none none n/a	Ch-052 (6-Mar-05) ⁷ none none none UNF ⁸ (16-Feb-05) UNF ⁸ (2-Mar-05)
Ch-036	F	KK	10-Dec-04	15-Mar-06	22.7	Ch-006 (M) Ch-052 (M) Ch-004 (M) Ch-021 (F)	Ch-006 (7-Feb-05) Ch-006 (11-Feb-05) Ch-006 (19-Feb-05) Ch-006 (25-Mar-05) Ch-006 (29-Mar-05) Ch-006 (23-Feb-06) Ch-052 (7-Feb-05) none n/a	none
Ch-039	F	MT	15-Aug-02	27-Feb-04	11.7	Ch-045 (M) Ch-048 (M)	Ch-045 (19-Jun-03) Ch-045 (20-Jun-03) Ch-045 (29-Nov-03) Ch-048 (19-Jun-03) Ch-048 (13-Jul-03) Ch-048 (9-Sep-03) Ch-048 (2-Dec-03)	Ch-045 (3-Dec-03) none
Ch-048	M	MT	27-Jan-02	29-Oct-03	17.3	Ch-039 (F) Ch-045 (M)	Ch-039 (19-Jun-03) Ch-039 (13-Jul-03) Ch-039 (9-Sep-03) n/a	Ch-080 (26-Jul-02) none
Ch-052	M	KK	23-Dec-04	24-Oct-05	11.8	Ch-036 (F) Ch-021 (F) Ch-006 (M) Ch-004 (M)	Ch-036 (7-Feb-05) none n/a n/a Ch-022 (18-Feb-05) Ch-033 (9-Apr-05) Ch-071 (11-Jul-05) Ch-071 (20-Aug-05)	none Ch-021 (6-Mar-05) none none

¹F, female; M, male.²KK, Kasekela; MT, Mitumba.³SIVcpz infected chimpanzees who could have served as transmitters based on phylogenetic analysis of viral sequences (Supplementary Fig. 2).⁴Observed matings with SIVcpz infected chimpanzees prior to infection; n/a, not applicable.⁵Observed acts of aggression with SIVcpz infected or unknown chimpanzees prior to infection.⁶Ch-036 mated frequently with males other than Ch-004 during the infection window and is thus the most likely transmission source (note that not all matings are observed).⁷Ch-021 likely acquired SIVcpz by aggression because she gave birth to an infant on 5-Feb-04 and was thus either pregnant or lactating during her window of infection.⁸UNF, unknown female.

Supplementary Table 3. Suspected vertical transmission of SIVcpz

Individual	Date of Birth ¹	Sample Date	Age at Sampling (years)	SIVcpz infection ²		Relationship
				fecal antibodies	fecal vRNA	
Ch-099	2-Jul-77*	17-Aug-03	26	pos	neg	mother ⁴
		17-Aug-03	26	pos	neg	mother
		21-Aug-03	26	pos	neg	mother
		2-Oct-04	27	pos	pos	mother
		2-Oct-04	27	pos	pos	mother
		2-Oct-04	27	pos	pos	mother
		23-Dec-04	27	pos	nd	mother
		11-May-05	27	pos	nd	mother
		12-Jul-05	28	pos	pos	mother
		25-Oct-05	28	pos	nd	mother
Ch-071	2-Jul-92*	22-Jun-02	10	pos	pos	daughter ⁴
		17-Aug-03	11	pos	nd	daughter
		1-Jun-04	11	pos	neg	daughter
		28-Nov-04	12	pos	neg	daughter
		1-Mar-05	12	pos	pos	daughter
		22-Jan-06	13	pos	pos	daughter
		12-Feb-07	14	pos	nd	daughter
		7-Apr-07	14	pos	nd	daughter
		9-May-08	15	pos	nd	daughter
		15-Oct-08	16	pos	nd	daughter
Ch-021	2-Jul-89*	18-Aug-02	13	neg	nd	mother
		6-Aug-03	14	neg	nd	mother
		29-May-04 ³	14	neg	nd	mother
		2-Dec-04 ³	15	neg	nd	mother
		5-Dec-04 ³	15	neg	nd	mother
		11-May-05	15	pos	pos	mother
		20-Nov-05	16	pos	nd	mother
		16-Mar-06	16	pos	nd	mother
		14-Oct-06	17	pos	pos	mother
		11-Feb-07	17	pos	nd	mother
		23-Oct-08	19	pos	nd	mother
		19-Sep-06	2.6	pos ⁵	neg	daughter
Ch-103	5-Feb-04 ³	16-Oct-06	2.7	pos ⁵	neg	daughter
		23-Oct-01	16	pos	neg	mother
Ch-033	2-Jul-85*	4-Jan-02	16	pos	neg	mother
		6-Sep-04	19	pos	nd	mother
		13-Dec-04	19	pos	nd	mother
		12-Jul-05	20	pos	pos	mother
		31-Jan-06	20	pos	nd	mother
		23-Feb-06	20	pos	nd	mother
		9-Aug-07	22	pos	nd	mother
		18-Oct-00	5-Aug-02	1.8	pos ⁵	neg
Ch-062						son

¹Asterisks denote estimated birth dates¹².²pos, positive; neg, negative; nd, not done.³Ch-021 was SIVcpz negative for at least 10 months following the birth of Ch-103.⁴Presumed mother/daughter pair.⁵Western blot results may indicate the presence of maternal antibodies in the stool of nursing infants.

Supplementary Table 4. Infant mortality

Mother ¹	Mother's SIVcpz status at delivery ²	Infant ³	Infant sex ⁴	Infant birth date	Infant SIVcpz status ⁵	Infant age at death (years) ⁶
Ch-001	neg	Ch-077	M	15-Jan-04	neg	Alive ⁶
Ch-002	neg	BB-002a	M	4-Apr-03	n/a	1.6
Ch-002	neg	Ch-097	F	14-Dec-05	neg	Alive
Ch-009	neg	Ch-102	F	18-Apr-04	neg	Alive
Ch-009	neg	BB-009a	F	14-Nov-07	n/a	Alive
Ch-011	neg	BB-011a	F	11-Oct-02	n/a	1.9
Ch-015	neg	BB-015a ⁷	M	18-Jul-08	n/a	0.0
Ch-015	neg	BB-015b ⁷	M	18-Jul-08	n/a	0.0
Ch-015	neg	BB-015c	M	28-Apr-06	n/a	0.4
Ch-019	neg	BB-019a	M	25-Feb-04	n/a	0.2
Ch-019	neg	Ch-117	F	4-Jun-05	neg	Alive
Ch-021	neg	Ch-103	F	5-Feb-04	pos	Alive
Ch-025	neg	Ch-073	F	17-Jun-01	neg	Alive
Ch-025	neg	Ch-116	M	9-Nov-06	neg	Alive
Ch-026	neg	BB-026a	na	11-Sep-07	n/a	0.1
Ch-028	neg	Ch-072	M	25-Aug-04	neg	2.2
Ch-031	neg	Ch-056	M	9-Jun-01	neg	Alive
Ch-032	neg	Ch-114	F	9-Sep-06	neg	Alive
Ch-036	pos	BB-036a	F	21-Mar-07	n/a	0.6
Ch-039	pos	BB-039a	na	15-May-07	n/a	0.0
Ch-041	neg	Ch-104	F	10-Jul-04	neg	3.2
Ch-042	neg	BB-042a	F	18-Nov-04	n/a	0.7
Ch-042	neg	BB-042b	M	3-Aug-06	n/a	Alive
Ch-044	neg	Ch-094	F	31-Mar-06	neg	Alive
Ch-046	neg	Ch-112	F	7-May-05	neg	Alive
Ch-047	neg	BB-047a	F	8-Jan-06	n/a	Alive
Ch-049	neg	Ch-111	F	13-Jun-04	neg	Alive
Ch-050	neg	Ch-115	M	16-Oct-05	neg	Alive
Ch-066	neg	BB-066a	M	15-Jul-07	neg	Alive
Ch-071	pos	BB-071a	M	5-Aug-07	n/a	0.1
Ch-076	neg	BB-076a	M	28-Jan-07	n/a	Alive
Ch-078	neg	Ch-090	M	20-Nov-05	neg	Alive
Ch-099	pos	BB-099a	F	21-Oct-05	n/a	0.99
Ch-105	neg	BB-105a	M	6-Sep-07	n/a	Alive

¹Black, uninfected; red, SIVcpz infected.²pos, positive; neg, negative.³BB numbers denote unsampled infants who are not included in Supplementary Table 1.⁴F, female; M, male.⁵n/a, data not available.⁶Alive at the end of the study period (31-Dec-08).⁷Twins.

SUPPLEMENTARY MATERIALS AND METHODS

Chimpanzee demographic and behavioral data. Kasekela and Mitumba chimpanzees have been observed on a daily basis since the 1960s and 1980s, respectively^{12,13,26}. Each day, a team of field assistants follows one "focal" individual from each community from dawn to dusk (Supplementary Fig. 1c). Behavioral data are recorded for this particular chimpanzee as well as other chimpanzees interacting with this individual. In addition, conspicuous interactions (e.g., mating or aggression) involving any member of the group are also recorded. The types of data collected include the incidence and timing of mating, aggression, grooming and nursing, as well as geographic location, group composition, and foraging behavior. Since March 2004, Kasekela and Mitumba chimpanzees have also been routinely monitored for baseline health data, with general body condition, skin abnormalities, wounds and sores, lameness, respiratory signs (e.g., coughing and sneezing), and gastrointestinal symptoms (e.g., stool consistency) recorded for focal individuals on a daily basis³⁰. In addition, a team of veterinarians performs on-site necropsies and determines the cause of death for all chimpanzees who are found dead in the park.

Sample collection. Most Kasekela and Mitumba chimpanzees tolerate human observers at close distance which facilitates sample collection; however, a subset of apes, in particular infants, are more difficult to sample, thus resulting in less frequent analyses. Urine samples were collected by placing plastic bags under night nests and stored without preservatives⁷. Fecal samples (~20g) were collected under direct observation and placed into 50 ml tubes containing 20 ml of RNA/alter (Ambion, Austin, TX)⁷. Specimens were frozen on site, shipped at ambient temperatures, and stored at -80°C upon receipt^{7,8,31}.

Non-invasive SIVcpz testing. Chimpanzee fecal and urine samples were examined for the presence of SIVcpz antibodies using an enhanced chemiluminescent Western blot assay as described^{2,7,8}. A

subset of antibody positive fecal samples was also tested for the presence of SIVcpz sequences by reverse transcriptase polymerase chain reaction (RT-PCR) amplification of fecal RNA^{7,8,31}, using SIVcpz*Pts* specific *pol* and *env-nef* primers (Supplementary Table 1). cDNA was synthesized using the R1 primer, followed by nested PCR using primers F1/R1 and F2/R2. A ~890-bp *pol* fragment was amplified using PTS-*pol*-F1 (5'-GTTACCTGGTACCTGAGTGGGA-3'), PTS-*pol*-R1 (5'-ACTACTGCCCTTCACCTTC-3'), PTS-*pol*-F2 (5'-TTTTATGTAGATGGGCAGC-3'), and PTS-*pol*-R2 (5'-CAATCCCCCTTTCTTTAAAATTGTG-3'). A shorter ~500-bp *pol* fragment was amplified using PTS-*pol*-F1 (5'-TGGGTACCAGCMCATAAAGGYATAGG-3'), PTS-*pol*-R1, PTS-*pol*-F2 (5'-GARGACCATGAYAAATATCA-3'), and PTS-*pol*-R2. A ~780-bp *env/nef* fragment was amplified using PTS-*env/nef*-F1 (5'-ATTGGTTGAYATAACACAATG-3'), PTS-*env/nef*-R1 (5'-CCCATCCAGTCCCCCTTTC-3'), PTS-*env/nef*-F2 (5'-CACAATGGCTRTGGTATATAA-3') and PTS-*env/nef*-R2 (5'-CTTTCTTTAARAACCATGACAG-3'). RT-PCR products were gel purified and sequenced directly^{7,8}. For TAN17 (from Ch-030), only a 286 bp *env* (gp41 region) fragment was amplified due to a misaligned primer. All newly derived TAN sequences were submitted to GenBank and their accession numbers are listed in Supplementary Table 1.

Genotyping of fecal samples. Fecal DNA was extracted as described^{7,8} and used to amplify a 498-bp mitochondrial (mt) genome fragment (D loop) for haplotype determination⁸. Fecal DNA was also used to amplify highly polymorphic microsatellite loci as well as a region of the amelogenin gene for sex determination (Supplementary Table 1)⁸.

SIVcpz prevalence. The prevalence of SIVcpz infection in Kasekela and Mitumba was determined for each of the past nine calendar years (Fig. 1). For these analyses, individuals were considered SIVcpz positive if they had detectable antibodies in their urine or fecal samples (a positive Western blot is diagnostic for SIVcpz infection, except for nursing infants who may have breast-milk derived maternal antibodies in their stool). The proportion of SIVcpz infected apes was determined by dividing the

number of positive individuals by the total number of apes tested (Fig. 1b). Most chimpanzees were tested at least once every year (Fig. 1, Supplementary Table 1). If a chimpanzee was not sampled in a particular year, the individual was excluded from the analysis for that calendar year (e.g., Ch-036 in 2005); however, because SIVcpz causes a chronic life-long infection, it was possible to infer infection data for some of the missing time points. For example, Ch-006 was not sampled in 2006 and 2007, but known to be SIVcpz infected from prior analyses. Since this male was observed (and thus known to be alive) in January 2007, he was counted as SIVcpz infected in both 2006 and 2007 calendar years (indicated by asterisks in Fig. 1a). Similarly, Ch-012 tested SIVcpz negative in 2004, but was not sampled in 2005. Since this female tested negative again in 2006, she was counted as uninfected in 2005. Finally, Ch-098 immigrated to Mitumba in June 2004, but was first tested in that community in April 2005. Since several samples collected in Kalande in 2002 and 2003 were SIVcpz negative (see Supplementary Table 1), Ch-098 was counted as uninfected in 2004.

Phylogenetic analyses. To determine the evolutionary relationships of the newly derived Gombe viruses to previously characterized HIV-1 and SIVcpz strains, phylogenetic trees were constructed from available *pol* and *env-nef* sequences, including representatives of all major SIVcpz*Ptt*, SIVcpz*Pts*, SIVgor and HIV-1 (groups M, N and O) lineages. In all of these analyses, Gombe viruses formed a single monophyletic clade within the SIVcpz*Pts* radiation (not shown). For subsequent tree constructions, we thus selected two SIVcpz*Pts* strains (ANT and BF1167) from the Democratic Republic of Congo as outgroups (GenBank accession numbers: U42720 and FJ869116/FJ869117). Nucleotide sequences were aligned using CLUSTAL W³²; sites that could not be aligned unambiguously were excluded. Trees were inferred by Bayesian²⁸ (Supplementary Fig. 2a and b) or maximum likelihood³³ (Supplementary Fig. 2c and d) methods, using a general time reversible model with gamma-distributed site-to-site rate variation and allowing for invariable sites (GTR+I+G)³⁴. Posterior probabilities were calculated using MrBayes²⁸ with 10⁶ generations (average standard deviations < 0.01) and 25% burn in.

To calculate the time to the most recent common ancestor (tMRCA) of the Gombe SIVcpz clade, alignments of *pol* and *env-nef* sequences of TAN1, TAN2, TAN3, TAN6, TAN8, TAN10, TAN11, TAN12, TAN13 and BF1167 were concatenated (total length 1,470 nt) and analyzed using programs from the BEAST package³⁵. Due to the limited number and length of the Gombe sequences, we were unable to obtain a robust tMRCA estimate. We thus calculated the evolutionary rate of HIV-1 for the same concatenated region in a separate BEAST analysis. Using 37 group M sequences from the Los Alamos Subtype Reference Set and their sample times (in years) as priors in a relaxed clock model³⁶, we obtained a rate of 1.40×10^{-3} substitutions/site/year and a median HIV-1 group M tMRCA estimate of 94.7 years (95% Highest Probability Density interval [HPD]: 63.7-156.7) consistent with previous analyses^{37,38}. Assuming that SIVcpz has evolved at a similar rate as HIV-1, we then used this rate as well as the sample dates for the different TAN viruses to estimate tMRCA of the Gombe clade. Four independent chains of 10^7 iterations were analyzed with each chain showing results consistent with others from the same set. Times to MRCA were determined from the combined runs (with 10% burn in) using Tracer version 1.4.1. This yielded a median date of 1851 for the Gombe tMRCA (95% HPD: 1730-1923).

Mortality analyses. To determine whether SIVcpz infection is associated with an increased mortality risk, we modeled chimpanzee death as a discrete event-history process^{29,39}. We adopted this particular approach (with the “chimpanzee-year-of-observation” as the unit of analysis) because data were collected from a free-ranging population, and vital events and SIVcpz infection data were frequently interval censored. Vital events (i.e., death, birth) and the change of time-varying covariates (i.e., SIVcpz infection) were recorded for yearly intervals. Individuals who had disappeared were classified as dead based on previously reported criteria^{15,16}, and none of the deceased chimpanzees was a known victim of human poaching. Only vital events from known members of the Kasekela and Mitumba communities were included. The mortality hazard, which is a strict conditional probability for discrete models, was then estimated as a function of covariates using a binomial generalized linear model with

a complementary log-log link. This represents the discrete-time equivalent of the proportional hazards assumption in continuous time Cox regression models^{29,40}. Exploration of age-effects using likelihood-based model selection criteria⁴¹ revealed that a cubic specification of age provided the best model fit. Because sex has previously been shown to affect the death hazard in chimpanzees¹⁵, we also included this parameter as a covariate. Two different event definitions were used. For a maximally conservative estimate, two fecal antibody positive but vRNA negative infants (Ch-062 and Ch-103) were classified as SIVcpz negative (i.e., assumed to have maternal antibodies in their stool), two incident infections (Ch-036 and Ch-039) were assumed to have occurred one year earlier (in 2005 and 2003, respectively; Fig. 1a and Supplementary Table 1), and two uninfected females who went missing for more than one year (Ch-038 and Ch-108) were counted as being dead. For a less stringent estimate, the two infants were assumed to be infected, the two incident infections were counted negative in the unsampled years, and the two missing females were censored, i.e. assumed to have migrated to Kalande or outside the park. Supplementary Table 5 lists the coefficients, hazard ratios, 95% confidence intervals (CI) for the hazard ratios (lower and upper), and *p*-values corresponding to the two event definitions.

Supplementary Table 5. SIVcpz associated mortality

Maximally conservative event definitions					
Variable	Coefficient	Hazard Ratio	Lower	Upper	<i>p</i> -value
Intercept	-2.34	0.10	0.02	0.59	0.011
age	-0.46	0.63	0.43	0.91	0.014
age ²	0.02	1.02	1.00	1.04	0.020
age ³	-0.0002	0.9998	0.9995	0.9999	0.057
male sex	0.40	1.50	0.56	4.05	0.424
SIVcpz pos	2.28	9.76	2.78	34.26	<0.0001
Less conservative event definitions					
Variable	Coefficient	Hazard Ratio	Lower	Upper	<i>p</i> -value
intercept	-2.86	0.06	0.01	0.39	0.003
age	-0.59	0.55	0.37	0.83	0.004
age ²	0.03	1.03	1.01	1.05	0.007
age ³	-0.0003	0.9997	0.9994	0.9999	0.022
male sex	0.69	1.99	0.68	5.88	0.211
SIVcpz pos	2.75	15.63	4.72	51.77	<0.0001

The results show that SIVcpz infection status has a strong and highly significant impact on the mortality hazard ($p<0.0001$). Under the maximally conservative event definitions, SIVcpz infection increased the hazard of dying by $\exp(2.28) = 9.76$ times; under the less conservative model, the hazard increased by a similar factor of $\exp(2.75) = 15.63$. 95% confidence intervals were wide due to the small sample size: For the maximally conservative approach, CI values were 2.8 and 34.3; for the less conservative approach, CI values were 4.7 and 51.8.

To put these results into perspective, we examined survival differences in a second SIV infected primate host. Vital data were obtained from 167 SIVsmm infected and 62 uninfected sooty mangabeys housed at the Yerkes Regional Primate Center. Like the Gombe chimpanzees, these sooty mangabeys were naturally infected; however, unlike the Gombe chimpanzees, the age at which they acquired SIVsmm was not determined. The absence of age-specific infection data precluded the need to analyze time-varying covariates. Moreover, there was no interval censoring since the sooty mangabeys were housed (and thus continuously observed) at a primate center. We thus employed a simple Cox proportional hazards regression model to examine potential survival differences, using only entry age and SIVsmm infection status as covariates^{29,42}. The results showed a significantly negative effect of entry age on sooty mangabey mortality (Supplementary Table 6), i.e., adult sooty mangabeys had lower mortality rates than infants ($p<0.0001$); however, no such effect was seen for SIVsmm infection: although the coefficient on SIVsmm status was negative, this was not statistically significant ($p=0.55$).

Supplementary Table 6. SIVsmm associated mortality

Variable	Coefficient	Hazard Ratio	se(Coef)	z value	Pr(> z)
entry age	-0.125	0.882	0.0287	-4.346	<0.0001
SIVsmm pos	-0.157	0.855	0.2651	-0.593	0.55000

Fertility analyses. To determine the impact of SIVcpz infection on female fertility, we compared the number of births to infected versus uninfected females, counting a pair of twins as one birth. The fertility of female chimpanzees is influenced by several factors. Adolescent females experience a period of subfecundity^{43,44}, as do females who have recently immigrated from another community⁴⁵, and

females nursing a young infant experience lactational subfecundity until the infant is weaned^{43,44}. To control for these factors, we compared births per female-years-at-risk for both infected and uninfected females, rather than simply comparing their total number of births. For each calendar year in which the infection status of the female was known, we determined whether she was at risk of giving birth based on long-term fertility data obtained for Gombe chimpanzees^{43,44,46}, using the minimum observed periods of adolescent sub-fecundity (min = 0.9 years; median = 2.5, n = 15 females), time between immigration and birth (min = 0.92 years; median = 2.66, n = 24 females), and inter-birth interval when the previous infant survived (min = 3.28 years; median = 5.51, n = 57 intervals). Thus, we considered nulliparous females who remained in their natal community to be at risk of birth if they had been showing full sexual cycles during which they mated with an adult male for at least one year before the midpoint of the year being scored. Likewise, we scored immigrant females without infants as being at risk if they had been in the community for at least one year before the mid-point of the year being scored. Finally, we scored females who had a previous surviving infant to be at risk if the previous infant was older than 3.28 years by the middle of the year being scored and the female had resumed sexual cycling by May 1 of that calendar year (given a gestation period of about 7.5 months, a female must have mated by May to produce an infant by the end of that calendar year). These measures are conservative, in that for each factor, the median interval was one to two years longer than the minimum interval. The results yielded 29 births per 90 years-at-risk for uninfected females, compared to 4 births per 30 years-at-risk for infected females, indicating that SIVcpz infected females were significantly less likely to give birth than uninfected females (Fisher's exact test: p=0.034; assuming that SIVcpz infection does not increase female fertility, a one-tailed test was used).

Infant mortality. All infants born between 2000 and 2008 to females of known SIVcpz infection status were scored as living or dying before one year of age. Four of four infants born to females who were SIVcpz positive at the time of delivery died before they reached 12 months of age, while 6 of 30 infants born to SIVcpz negative females died before their first birthday (Supplementary Table 4). Two of the

latter were twins, which generally have a higher mortality. The results showed that infants born to infected mothers had a significantly higher mortality rate than infants born to uninfected mothers (Fisher's exact test: $p < 0.005$; assuming that SIVcpz infection does not increase infant survival, a one-tailed test was used).

Necropsy and tissue collection. Complete necropsies were performed using strict safety precautions. Ch-016 was necropsied within 15 hours of death; The bodies of the other four chimpanzees (Ch-036, Ch-045, Ch-069 and Ch-099) were recovered within 6 to 18 hours of death and frozen (-20°C) for up to two months prior to necropsy. Sections of all tissues were fixed in 10% neutral buffered formalin, paraffin-embedded, cut at 3-5 mm, stained with hematoxylin and eosin, and evaluated by veterinary pathologists. Some tissue samples were also fixed in RNA/later. Both gross and histopathological findings were used to determine the cause of death: Ch-016 died of conditions related to old age; Ch-036 died of an AIDS-like syndrome; Ch-045 died of traumatic injuries following conspecific (intra-group) aggression involving two other males; Ch-069 died of trauma induced vertebral subluxation (C6-7) and associated spinal cord injury; and Ch-099 died of trauma induced non-union fracture of a vertebra (T5) and associated myelomalacia. Due to the warm climate and different intervals between death and carcass retrieval, tissues exhibited varying degrees of autolysis, most notably in the gastrointestinal mucosa, which precluded quantitative immunohistochemistry in the gut; however, lymph node and spleen tissues were generally well preserved. For the initial histological evaluation, veterinary pathologists were blinded to the infection status of the deceased ape, except for Ch-045 whose status was known prior to necropsy. This was also true for the first evaluation of all immunohistochemically stained slides which were analyzed under code.

In situ hybridization, immunohistochemistry and quantitative image analysis. SIVcpz specific *in situ* hybridization (ISH) was performed using SIVcpz anti-sense and control (sense) riboprobes. SIVcpz Pts specific probes were generated by amplifying *gag*, *pol*, *vif*, *env* and *nef* fragments from

TAN2 and TAN3 genomes³¹. Amplicons were cloned into the pCR-Blunt II-topo vector (Invitrogen), between SP6 and T7 promoters. These clones were linearized and transcribed with ³⁵S-labelled dUTP or dCTP to generate anti-sense and sense radio-labeled riboprobes depending on the orientation of the gene fragment and the promoter used. Transcripts were pooled (i.e. anti-sense fragments from all TAN2 and TAN3 clones) and ISH was performed as previously described^{20,47}. Immunohistochemistry was performed on paraffin sections (3-5 µm). Tissue sections were deparaffinized and heat induced epitope retrieval performed using a pressure cooker decloaking chamber (Biocare Medical). Endogenous peroxidase was blocked using hydrogen peroxide (1.5% - 3%) in methanol or in TBS (pH 7.4). Slides were either stained manually or using the intelliPATH FLX autostaining system using a non-biotin polymer detection system (Super Enhancer, BioGenex or Biocare Medical). Primary antibodies were incubated for 30 minutes to 1 hour at room temperature, followed by incubation with a polymer based detection system according to manufacturer's recommendations. After washing, sections were developed with the chromagen 3,3'-diaminobenzidine (DAB; BioGenex or Vector Laboratories) for 5 minutes, counterstained with hematoxylin (BioGenex or Biocare Medical), and mounted in Permount (Fisher Scientific). Primary antibodies used were mouse monoclonal anti-CD3 (clone F7.2.38; Dako), rabbit anti-CD3 (Cat no. A0452; Dako), mouse monoclonal anti-CD4 (clone 1F6; Vector Labs), mouse monoclonal anti-CD20 (clone L26; Labvision) and mouse monoclonal anti-CD79a (clone HM47/A9; Biocare Medical). Masson's Trichrome staining was performed as previously reported⁴⁸. Quantitative image analysis was performed on stained whole tissue sections scanned at high magnification (400x) using the ScanScope CS System (Aperio Technologies, Inc.) yielding high-resolution data from the entire stained tissue section. In order to determine CD4+ T cell levels within splenic white pulp, the periarteriolar lymphoid sheaths (PALS; splenic equivalent of the T cell zone of secondary lymph nodes) were manually extracted in Photoshop CS3 using the magnetic lasso tool from high magnification whole tissue images obtained from the Aperio ScanScope CS System. Individually isolated PALS were saved into separate files and the percent area that stained positive for CD4 was

determined using tools from Reindeer Graphics^{21,48}. The frequencies of CD4+ cells were compared among individual chimpanzees, and between infected and uninfected chimpanzees, using analyses of variance. An F-test indicated highly significant differences among individuals ($p<0.001$). For the comparison of the three infected individuals with the two uninfected individuals, it was assumed that SIVcpz infection would not lead to an increased CD4+ T cell count, and so a one-tailed test was used; the F-test yielded a p value of 0.024, indicating significantly lower counts in the SIVcpz infected individuals. Similarly, t-tests between individuals, with Tukey's adjustment for multiple comparisons, were significant in all 10 pairwise tests ($p<0.05$).

Supplementary References:

30. Lonsdorf, E.V., Travis, D., Pusey, A.E. & Goodall, J. Using retrospective health data from the Gombe chimpanzee study to inform future monitoring efforts. *Am J Primatol* **68**, 897-908 (2006).
31. Takehisa, J., et al. Generation of infectious molecular clones of simian immunodeficiency virus from fecal consensus sequences of wild chimpanzees. *J Virol* **81**, 7463-75 (2007).
32. Thompson, J.D., Higgins, D.G. & Gibson, T.J. 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 (1994).
33. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704 (2003).
34. Yang, Z. Estimating the pattern of nucleotide substitution. *J Mol Evol* **39**, 105-111 (1994).
35. Drummond, A.J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214 (2007).
36. Drummond, A.J., Ho, S.Y., Phillips, M.J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol* **4**, e88 (2006).
37. Korber, B., et al. Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789-1796 (2000).

38. Worobey, M., *et al.* Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**, 661-664 (2008).
39. Allison, P.D. *Event History Analysis: Regression Models for Event History Data*, (Sage Publications, Beverly Hills 1984).
40. Hosmer, D.W. & Lemeshow, S. *Applied Logistic Regression*, 2nd ed, (Wiley, New York 2000).
41. Burnham, K.P. & Anderson, D.R. *Model Selection and Inference: A Practical Information-Theoretic Approach*. 2nd ed, (Springer, New York 2002).
42. Klein, J.P. & Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*, (Springer-Verlag, New York 1997).
43. Goodall, J. Population dynamics during a 15-year period in one community of free-living chimpanzees in the Gombe National Park, *Z Tierpsychol* **61**, 1-60 (1983).
44. Wallis, J. A survey of reproductive parameters in the free-ranging chimpanzees of Gombe National Park. *J Reprod Fertil* **109**, 297-307 (1997).
45. Boesch, C. & Boesch, H. *The Chimpanzees of the Tai Forest*, (Oxford University Press, Oxford, England 2000).
46. Anne Pusey, unpublished data.
47. Li, Q., *et al.* Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* **434**, 1148-1152 (2005).
48. Schacker, T.W., *et al.* Collagen deposition in HIV-1 infected lymphatic tissues and T cell homeostasis. *J Clin Invest* **110**, 1133-1139 (2002).