www.sciencemag.org/cgi/content/full/1126531/DC1

Supporting Online Material for

Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1

Brandon F. Keele, Fran Van Heuverswyn, Yingying Li, Elizabeth Bailes, Jun Takehisa, Mario L. Santiago, Frederic Bibollet-Ruche, Yalu Chen, Louise V. Wain, Florian Liegeois, Severin Loul, Eitel Mpoudi Ngole, Yanga Bienvenue, Eric Delaporte, John F. Y. Brookfield, Paul M. Sharp, George M. Shaw, Martine Peeters, Beatrice H. Hahn*

*To whom correspondence should be addressed. E-mail: bhahn@uab.edu

Published 25 May 2006 on *Science* Express
DOI: 10.1126/science.1126531

This PDF file includes:

SUPPORTING ONLINE MATERIAL

Materials and Methods

Study sites, sample collection and shipment. Fecal samples (n=599) were collected from wild-living (non-habituated) ape communities at 10 remote forest sites in southern Cameroon (Fig. 1). Forests were surveyed in the vicinity of base camps by collection teams (usually comprised of one team leader and two local guides) who followed elephant paths or forest transects in search for signs of wild apes (e.g., vocalization, night nests, feeding areas). The EK, CP, BB, and LB sites were located in National Parks or Forest Reserves, while the remaining field sites were in un-protected areas with considerable hunting pressure. Fecal samples were identified to be of likely chimpanzee origin by experienced trackers. For each sample, approximately 20g of feces were placed into a 50 ml tube containing 20 ml of RNA*later* (Ambion, Austin, TX). Time, date and collection site were recorded. In addition, samples were inspected to estimate their likely time of deposition. In some instances, global positioning system (GPS) coordinates were obtained. Specimens were kept at base camps at ambient temperature for up to 20 days, and subsequently transported to a central laboratory facility in Yaounde for storage at –20^{o}C. Samples were shipped to Europe and the US at ambient temperatures, and stored at –80^{o}C upon receipt.

Detection of SIVcpz antibodies in RNAlater preserved fecal samples. Fecal samples were examined for the presence of SIVcpz antibodies using an enhanced chemiluminescent Western immunoblot assay specifically developed for RNA*later* preserved specimens. RNA*later* is a high salt solution (25 mM Sodium Citrate, 10 mM EDTA, 70 g ammonium sulfate/100 ml solution, pH 5.2) that preserves nucleic acids, but precipitates immunoglobulin and thus generally precludes antibody detection by immunoblot analysis (*S1*). However, we found that

RNA*later* precipitated immunoglobulin can be resolubilized by diluting fecal/RNAlater mixtures (1.5 ml) with PBS-Tween-20 (8.5 ml), inactivating the mixture for 1 hr at 60ºC, clarifying it by centrifugation (3500 x *g* for 10 min), and then dialyzing it against PBS overnight at 4ºC. The reconstituted extracts were then subjected to immunoblot analysis as reported previously (*S2*).

Sensitivity and specificity of fecal antibody detection. The sensitivity and specificity of the fecal immunoblot assay were determined using test results from wild-living but habituated chimpanzees in Gombe National Park (Tanzania) where SIVcpz*Pts* prevalence rates range from 5 to 30% in different communities (*S2, S3*). Fecal samples (n=148) were collected from chimpanzees of known infection status and subjected to Western blot analysis. Antibodies were detected in 47 of 51 samples from infected chimpanzees, but in none of 97 samples from uninfected chimpanzees. Using the fraction of positive samples (rather than infected individuals), this yielded a sensitivity of 0.92 with confidence limits of 0.83 - 0.97 (as determined assuming binomial sampling). The specificity of fecal antibody detection was calculated using test results from uninfected chimpanzees and determined to be 1.00 with a lower confidence limit of 0.97. For these calculations, it was assumed that successive test results from the same individual were not correlated.

Nucleic acid extraction from fecal samples. Fecal RNA was extracted from RNA*later*-preserved fecal specimens using the RNA*queous* Midi-Kit (Ambion, Austin, TX) as described (*S2-S4*). Fecal DNA was extracted using the QIAamp Stool DNA Mini kit (Qiagen, Valencia, CA). Briefly, 1.5 ml of fecal RNA*later* mixture was resuspended in stool lysis buffer, clarified by centrifugation, reacted with an InhibitEx tablet (Qiagen, Valencia, CA), treated with

proteinase K, and passed through a DNA binding column. Bound DNA was eluted in 50-150µl elution buffer.

Amplification of SIVcpz sequences from fecal RNA. Diagnostic reverse transcriptase polymerase chain reaction (RT-PCR) was performed as described, using SIVcpz/HIV-1 consensus primers in *pol* (integrase) and *env* (gp41 ectodomain) regions (*S2, S3, S5*). cDNA was synthesized using the R1 primer, followed by nested PCR using primers F1/R1 and F2/R2. The *pol* primers included CPZ-pol-F1a (5'-ACCTGGATNCCWGANTGGGA-3'), CPZ-pol-R1a (5'-ACBACYGCNCCTTCHCCTTTC-3'), CPZ-pol-F2a (5'- TWYTATGTWGATGGRGCAGC -3') and CPZ-pol-R2 (5'-CCCAATCCCCCCTTTTCTTTTAAAATT-3'). In some instances, CPZ-pol-F1b (5'- GTTACCTGGGTACCTGAGTGGGA -3'), CPZ-pol-F1c (5'-TGGTGGDCWGANTAYTGGCA-3'), CPZ-pol-R1b (5'-ACTGCHCCYTCWCCTTTCCACAG -3'), and CPZ-pol-F2b (5'-ACCTGGATHCCHGANTGGGA-3') were also used. The *env* (gp41 region) primers included CPZ-gp41-F1 (5'-TCTTAGGAGCAGCAGGAAGCACTATGGG-3'), CPZ-gp41-R1 (5'-AACGACAAAGGTGAGTATCCCTGCCTAA-3'), CPZ-gp41-F2 (5'-ACAATTATTGTCTGGTATAGTGCAACAGCA-3') and CPZ-gp41-R2a (5'-TTAAACCTATCAAGCCTCCTACTATCATTA-3'). In some instances, CPZ-gp41-R2b (5'-TCCTACTATCATTATGAATATTTTTATATA-3') was also used. RT-PCR products from *pol* (~890 bp; HXB2 coordinates 3887-4775) and gp41 regions (~390 bp; HXB2 sequence 7880-8264) were gel purified (Qiagen, Valencia, CA) and sequenced directly (GenBank accession numbers DQ370366 to DQ370419).

Full-length viral sequences were obtained for four SIVcpz*Ptt* strains (LB7, MB66, EK505 and MT145) using a combination of consensus and virus specific primers as described

previously (*S3*, *S5*). A series of overlapping subgenomic fragments (ranging in length from 400 bp – 1,500 bp) was generated and directly sequenced. In regions of fragment overlap, the 5' sequence was arbitrarily chosen. The concatenated full-length SIVcpz*Ptt* sequences were 9,132 bp to 9,281 bp in length (GenBank accession numbers DQ373063 to DQ373066).

Phylogenetic analyses. Deduced amino acid sequences of SIVcpz genes (or gene fragments) were aligned with corresponding sequences of representative SIVcpz and HIV-1 strains using CLUSTAL W (*S6*) (GenBank accession numbers were as follows: HIV-1 M U455, M62320; HIV-1 M HXB2, K03455; HIV-1 N YBF30, AJ006022; HIV-1 N YBF106, AJ271370; HIV-1 O ANT70, L20587; HIV-1 O MVP5180, L20571; SIVcpzGAB1, X52154; SIVcpzGAB2, AF382828; SIVcpzCAM3, AF115393; SIVcpzCAM5, AJ271369; SIVcpzCAM13, AY169968; SIVcpzUS, AF103818; SIVcpzTAN1, AF447763; SIVcpzTAN2, DQ374657; SIVcpzTAN3, DQ374658; and SIVcpzANT, U42720). 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-terminus of the deduced Gag and Pol protein sequences, respectively. In addition, the concatenated Pol and Vif alignment was divided into two regions around a previously reported recombination breakpoint between the SIVcpzCAM3/CAM5/US and HIV-1 group N lineages (*S7, S8*). Trees were inferred by the Bayesian method implemented in MrBayes (*S9*), using the Jones, Taylor and Thornton model of evolution and gamma distributed rates at sites, with one million generations and burn-in of 10%. Bayesian parameters were examined with the Tracer program (http://evolve.zoo.ox.ac.uk/software.html/id=tracer) and all estimated sample sizes were greater than 563.

Species and subspecies determinations. For species and subspecies analysis, a 498-bp region of the mtDNA genome (D loop) was amplified from fecal DNA as described (*S2*). PCR was performed using primers L15997 (5'-CACCATTAGCACCCAAAGCT-3') and H16498 (5'-CCTGAAGTAGGAACCAGATG-3') and amplification products were sequenced directly. Identical sequences were identified by phylogenetic analysis and grouped into unique mtDNA haplotypes. These were classified based on their relatedness to species and subspecies specific reference sequences from GenBank (accession numbers; AF102684, AF102685, AF102687, AF249683, AF249684, AF249685, AF249686, AF249687, AF249688, AF249689, AF290602, AF290604, AF290605, AF290606, AF290607, AF290608, AF290609, AY126684, AY126685, AY126686, AY126687, AY126691, AY126692, AY126693, AY126695, X93335) as well as other collection sites in Ivory Coast (TA), Tanzania (GM), Rwanda (NY), Democratic Republic of Congo (KS), and Republic of Congo (GT) as well as sanctuaries (MG) and primate centers (YK) (GenBank accession numbers: DQ370307 to DQ370365). Sequences were aligned using CLUSTAL W (*S6*). Sites that could not be aligned unambiguously or sites with a gap in any sequence were excluded. Phylogenetic trees were inferred by the Bayesian method implemented in MrBayes (*S9*), with the GTR model for gamma distributed rates at sites and 10 million generations with a burn in of 10%. Estimated sample sizes from Tracer (http://evolve.zoo.ox.ac.uk/software.html/id=tracer) were greater than 220. This analysis identified 82 new mtDNA haplotypes among the fecal samples from wild chimpanzees in Cameroon (fig. S1). These are listed in table S2 and available at GenBank under accession numbers DQ367532 to DQ367613.

Microsatellite and gender analyses. Microsatellite analyses were performed as described (*S4*). All samples were genotyped at four highly polymorphic loci (D18s536, D4s243, D10s676,

D9s922), with additional loci amplified for select samples (D2s1326, D2s1333, D4s1627, D9s905, D1s548). All PCR reactions were performed in duplicate. Individuals whose genotype appeared homozygous were amplified a minimum of seven times to exclude allelic drop out.

For gender determination, a region of the amelogenin gene was amplified that contains a deletion in the X, but not the Y chromosome (*S10*). Two sets of primers were used (AMEL-F106 5'-CCCTGGGCTCTGTAAAGAATAGTG-3' and AMEL-R106 5'-ATCAGAGCTTAAACTGGGAAGCTG-3' as well as AMEL-F212 5'-ACCTCATCCTGGGCACCCTGG-3' and AMEL-R212 (5'-AGGCTTGAGGCCAACCATCAG-3'). These generated a 106bp or 212 bp fragment from the X chromosome, and a 112bp or 218bp fragment from the Y chromosome, respectively. Amplification products were visualized and sized as the microsatellite amplicons.

SIVcpz prevalence determination. For DP, EK, BB, MB and LB sites, SIVcpz prevalence rates were estimated based on the proportion of infected individuals as determined by microsatellite analyses (tables S3 and S4). For each chimpanzee, the probability that it would be detected as being infected, if it was truly infected, was calculated, taking into consideration the number of specimens analyzed as well as the test sensitivity (these values varied between populations from 0.9298 and 0.9825, depending on the kind of tests performed for each individual). SIVcpz prevalence was then estimated as the proportion of chimpanzees identified as infected, with 95% confidence limits calculated based on binomial sampling (table S1).

Microsatellite data were also used to estimate the extent of sample degradation and oversampling. Of 219 fecal samples collected at the DP, EK, BB, MB and LB sites, only 182 produced a complete (four loci) genotype. The remainder (n=37) yielded only a subset of loci, indicating partial degradation. Thus, on average, a proportion of 37/219ths (=0.169) of all

collected samples were degraded. Moreover, the 182 completely genotyped samples were derived from 106 individual chimpanzees. Thus, on average, each chimpanzee was sampled 182/106 =1.716 times. Finally, of the 106 fully genotyped chimpanzees, 66 were sampled once and 40 more than once. Given that the probability of a positive test for an infected chimpanzee sampled once is 0.92 and that of a chimpanzee sampled twice is (1-0.08^2 = 0.9936), the probability that a positive chimpanzee was detected as infected was (66 x 0.92 + 40 x 0.9936)/106 = 0.9478. For fecal samples not tested by microsatellite analyses, we thus estimated the number of individuals that were sampled as 106n/219 (where n is the number of samples collected), and for each chimpanzee included among the individuals yielding non-degraded samples, we estimated a 0.9478 chance of its being detected as positive if it was positive. From this, we calculated the prevalence rates and their confidence limits for the populations at the WE, MT, BQ, DG and CP sites (table S1).

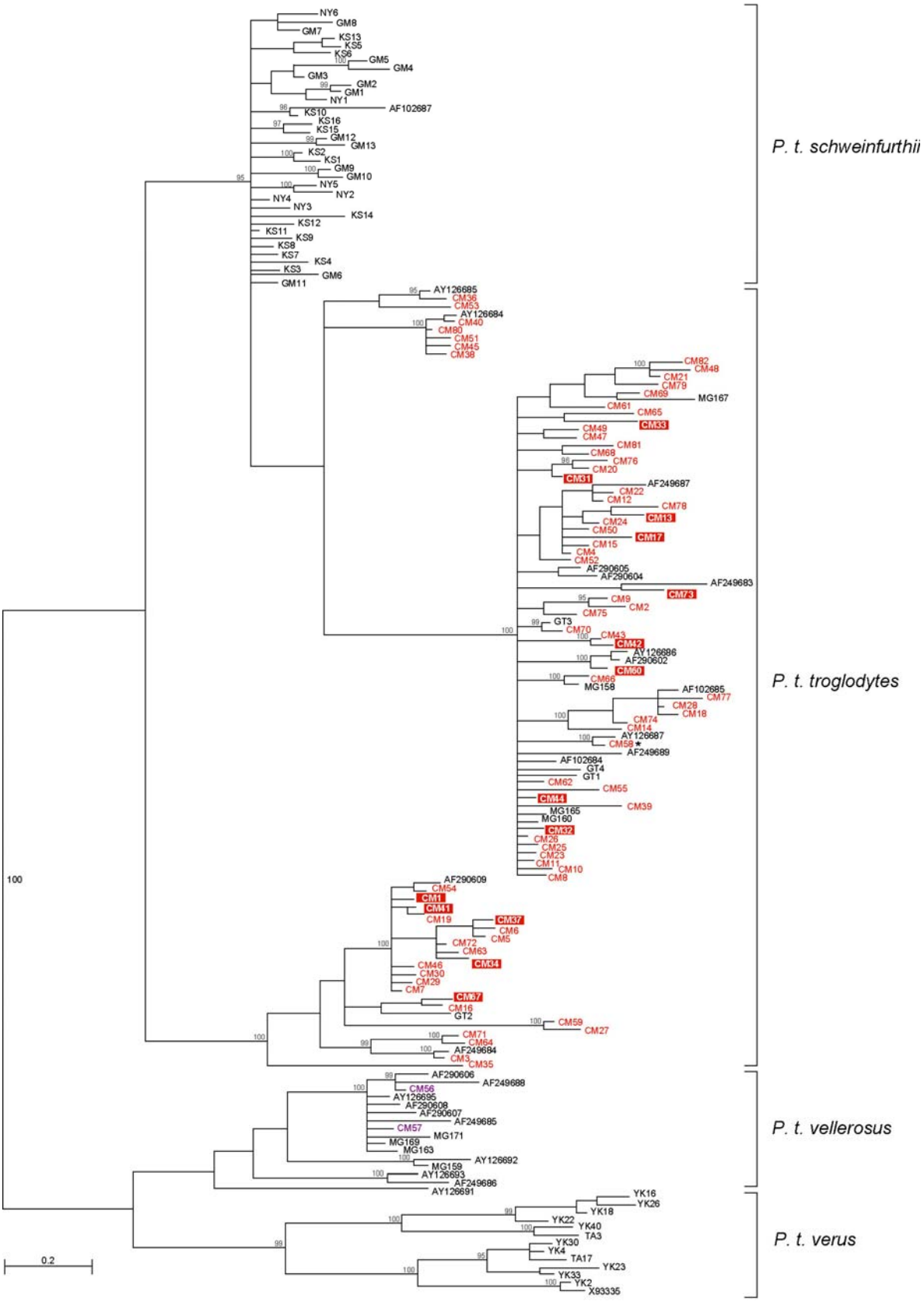
P. t. schweinfurthii
P. t. troglodytes
P. t. vellerosus
P. t. verus
0.2

fig. S1

Fig. S1. Subspecies origin of chimpanzee fecal samples. Mitochondrial DNA sequences (a 498 bp D loop fragment) from 446 chimpanzee fecal specimens were grouped into unique mtDNA haplotypes (listed in table S2) and then compared to subspecies specific reference sequences by phylogenetic tree analysis (indicated by brackets). The tree was constructed by the Bayesian method (*S9*), using the GTR model with gamma distributed rates at sites and 10 million generations with a burn in of 10%. Haplotypes from *P. t. troglodytes* are shown in red and those from *P. t. vellerosus* in magenta. An asterisk denotes the haplotype of a *P. t. troglodytes* sample (CM58) that was identified in the range of *P. t. vellerosus* apes north of the Sanaga River. Haplotypes from SIVcpz infected *P. t. troglodytes* apès are boxed. The numbers on nodes are percentage posterior probabilities (only values above 95% are shown). The scale bars represents 0.2 substitutions per site.

Table S1. Prevalence of SIVcpz infection in ten wild-living chimpanzee communities.

Collection Sites[1]	Fecal samples collected	Samples positive for mtDNA	Gorilla	Chimpanzee	*P.t. vellerosus*	*P.t. troglodytes*	Samples containing SIVcpz antibodies	Samples with complete microsatellite genotypes[2]	Number of chimpanzees tested	Number of chimpanzees infected	SIVcpz prevalence (95% CI)[4]
WE	27	26	1	25	23	2	0	nd[3]	nd	0	0.0% (0.0% - 23.1%)
MT	94	81	0	81	0	81	10	nd	nd	2	5.4% (1.0% - 16.0%)
DG	71	54	25	29	0	29	0	nd	nd	0	0.0% (0.0% - 20.3%)
DP	139	125	0	125	0	125	7	100	46	2	4.4% (0.8% - 13.3%)
BQ	101	88	6	82	0	82	0	nd	nd	0	0.0% (0.0% - 7.4%)
EK	26	25	6	19	0	19	6	18	15	4	28.7% (10.4% - 54.9%)
BB	39	38	4	34	0	34	0	25	18	0	0.0% (0.0% - 16.4%)
MB	37	25	0	25	0	25	9	25	18	6	35.3% (16.5% - 58.7%)
LB	54	40	24	16	0	16	2	14	9	2	23.3% (4.3% - 57.7%)
CP	11	11	1	10	0	10	0	nd	nd	0	0.0% (0.0% – 48.7%)
Total	599	513	67	446	23	423	34	182	106	16	

[1]Location of field sites is shown in Fig. 1.
[2]Individual genotypes are shown in tables S3 and S4.
[3]nd, not done.
[4]CI, Confidence Interval.

Table S2. Subspecies origin of chimpanzee fecal samples.

Haplotype[1]	Fecal samples with identical mtDNA haplotype	GenBank accession number
CM1	BQ28, BQ33, BQ50, MT150, MT330, BQ497, BQ499, DP03, DP17, DP18, DP19, DP22, DP23, DP26, DP27, DP67, DP69, DP80, DP82, DP130, DP134, DP163, DP206, DP218, DP220	DQ367534
CM2	BQ29	DQ367535
CM3	BQ30	DQ367536
CM4	BQ32, BQ44, MT122, MT123, MT125, MT153, MT156, MT157, BQ193, BB241, MT335, MT371, MT373, MT436, MT437, EK504, EK508, EK509, DP01, DP05, DP103, DP104, DP137, DP151, DP159, DP162,	DQ367537
CM5	BQ38, BQ389, BQ391, BQ392, MB135, LB188, EK503, EK517, DG533, DG541, DP98	DQ367538
CM6	BQ39, BQ42, BQ52, BQ55, BQ474, BQ481, BQ484, BQ487	DQ367539
CM7	BQ40, BQ41, MT326, MT337, MT339, DG407, CP384, EK513, EK522	DQ367540
CM8	BQ43, BQ51, BQ83, BQ84, BQ387, BQ388, BQ393, MT334, DP14, DP83, DP89, DP90, DP91, DP92, DP94, DP99, DP100, DP101, DP110, DP140, DP142, DP219	DQ367541
CM9	BQ45, BQ47, MT121, MB140, MT141, LB310, MT348, MT357, MT365, MT385, BQ390, BQ396, BQ397, BQ398, BQ400, BQ401, BQ402, BQ403, BQ473, BQ475, BQ482, BQ489, BQ494, BQ495, BQ496, BQ498, BQ500, DG524	DQ367542
CM10	BQ46, BQ194, BQ195, BQ490	DQ367543
CM11	BQ48, BQ57, BQ394, BQ476, BQ483, BQ488, BQ488b, BQ491	DQ367544
CM12	BQ49, MB97, DP70	DQ367545
CM13	BQ54, MB23, MB24, LB307, LB308, BQ399, DG546, DP71, DP72, DP74, DP75	DQ367546
CM14	BQ56, BQ59, BQ471, BQ477, BQ478, BQ479, BQ480, BQ485	DQ367547
CM15	MT51, MT52, MT126	DQ367548
CM16	MT53, MT54, MT120, MT124, MT154, MT155, MT343, MT344, MT345, MT346, MT347, MT349, MT350, MT351, MT353, MT354, MT355, MT356, MT358, MT360, MT361, MT362, MT363, MT366, MT386	DQ367549
CM17	MB66	DQ367550
CM18	BB72, BB75, BB79, BB92, BB104	DQ367551
CM19	BB73, BB74, BB87, BB94, BB102, BB230, DG526, DG542, DG543	DQ367552
CM20	BB76, BB239, EK501	DQ367553
CM21	BB77	DQ367554
CM22	BB78, BB90, BB101, BB103, MT342	DQ367555
CM23	BQ81, MT127, LB204, LB205, LB208, MT329, MT331, MT332, MT333, MT336, MT340, DG531, DG547, DG548, DG549, DG562, DP62, DP65	DQ367556
CM24	BQ82, EK521, DP63, DP64, DP68, DP86, DP117, DP118, DP127, DP128, DP131, DP132,	DQ367557
CM25	BQ85, BQ86	DQ367558
CM26	BB88, BB95	DQ367559
CM27	BB89, BB237, BB240	DQ367560
CM28	BB91, BB99, BB100, LB174, LB175	DQ367561
CM29	BB93, BB229, BB234, BB238	DQ367562
CM30	BB106	DQ367563
CM31	MT114, MT115, MT116, MT117, MT148	DQ367564
CM32	MT119, MT151, EK516	DQ367565
CM33	MB138, MB139, MB189, MB191, MB192, MB250, LB311, LB312, LB313, MB318, MB319, MB320	DQ367566
CM34	MT142, MT143, MT145, MT147, MT149, LB176, MB324	DQ367567
CM35	MT144, DP125, DP129, DP133, DP135	DQ367568
CM36	MT146	DQ367569
CM37	LB186	DQ367570
CM38	MB190, LB309	DQ367571
CM39	BB235	DQ367572
CM40	BB236	DQ367573
CM41	MB245	DQ367574
CM42	MB246, MB247, MB248, EK511	DQ367575
CM43	MB315, MB316	DQ367576
CM44	MB317	DQ367577
CM45	MB323	DQ367578
CM46	MT325, MT338, DG525, DG527, DG530, DG532, DG534, DG535	DQ367579
CM47	MT327	DQ367580
CM48	MT341	DQ367581
CM49	MT352, MT359, MT364	DQ367582
CM50	MT379	DQ367583

Table S2. Subspecies origin of chimpanzee fecal samples (continued).

Haplotype[1]	Fecal samples with identical mtDNA haplotype	GenBank accession number
CM51	CP380, CP381, CP382	DQ367584
CM52	CP383	DQ367585
CM53	BQ395, BQ404, BQ492, BQ493	DQ367586
CM54	DG405, DP78, DP79, DP88, DP111, DP112, DP113, DP114, DP115, DP116, DP119	DQ367587
CM55	DG406	DQ367588
CM56	WE438, WE439, WE440, WE441, WE443, WE444, WE445, WE446, WE447, WE448, WE449, WE450, WE451, WE454, WE456, WE457, WE458, WE459, WE460, WE461, WE462	DQ367532
CM57	WE452, WE453	DQ367533
CM58	WE455, WE464	DQ367589
CM59	CP466	DQ367590
CM60	CP467, EK502, EK505, EK506, EK507	DQ367591
CM61	CP468, EK510	DQ367592
CM62	CP469	DQ367593
CM63	CP470	DQ367594
CM64	BQ472, DG523, DG529, DG536, DG537, DG538	DQ367595
CM65	EK512	DQ367596
CM66	EK518	DQ367597
CM67	EK519	DQ367598
CM68	DG528, DG539	DQ367599
CM69	DG540	DQ367600
CM70	DP04	DQ367601
CM71	DP06, DP13, DP102, DP120	DQ367602
CM72	DP07, DP08, DP12, DP16	DQ367603
CM73	DP09, DP10, DP11, DP15, DP20, DP24, DP25, DP108, DP109, DP126	DQ367604
CM74	DP66, DP154, DP156	DQ367605
CM75	DP77, DP95, DP97	DQ367606
CM76	DP81, DP85, DP121, DP122, DP124	DQ367607
CM77	DP93, DP96	DQ367608
CM78	DP105, DP106, DP107, DP123, DP160	DQ367609
CM79	DP136, DP141, DP143, DP144, DP145, DP146, DP148, DP149, DP150	DQ367610
CM80	DP152, DP153, DP164, DP165	DQ367611
CM81	DP155, DP157, DP158	DQ367612
CM82	LB7	DQ367613

[1]Haplotypes correspond to those shown in fig. S1. CM56 and CM57 represent *P. t. vellerosus* apes; all other haplotypes are from members of the *P. t. troglodytes* subspecies. CM58 denotes the haplotype of a *P. t. troglodytes* sample that was collected in the range of the *P. t. vellerosus* subspecies north of the Sanaga River. Haplotypes from SIVcpz infected chimpanzees are highlighted in red.

Table S3. Genetic identification of SIVcpz infected chimpanzees at five field sites.

Individual	Fecal sample	Date of sample collection	Fecal Western Blot Analysis[1]	Fecal virion RNA Detection[2]		mtDNA haplotype[3]	Sex[4]	Locus[5] D18S536	Locus[5] D4S243	Locus[5] D10S676	Locus[5] D9S922
ID1	MT114	08/06/03	Pos	pol	nd	CM31	M	145/161	187/199	177/177	293/297
	MT115	08/06/03	Pos	pol	gp41	CM31	M	145/161	187/199	177/177	293/297
	MT116	08/06/03	Pos	nd[6]	nd	CM31	M	145/161	187/199	177/177	293/297
	MT117	08/06/03	Pos	pol	gp41	CM31	M	145/161	187/199	177/177	293/297
	MT148	08/14/03	Pos	pol	gp41	CM31	M	145/161	187/199	177/177	293/297
ID2	MT142	08/14/03	Pos	pol	nd	CM34	M	161/169	195/227	181/181	289/309
	MT143	08/14/03	Pos	pol	gp41	CM34	M	161/169	195/227	181/181	289/309
	MT145	08/14/03	Pos	pol	gp41	CM34	M	161/169	195/227	181/181	289/309
	MT147	08/14/03	Pos	pol	gp41	CM34	M	161/169	195/227	181/181	289/309
	MT149	08/14/03	Pos	pol	gp41	CM34	M	161/169	195/227	181/181	289/309
ID3	DP22	03/6/03	Pos	nd	gp41	CM1	F	141/169	187/231	169/185	301/305
	DP23	03/6/03	Pos	nd	gp41	CM1	F	141/169	187/231	169/185	301/305
	DP26	03/6/03	Pos	nd	nd	CM1	F	141/169	187/231	169/185	301/305
	DP206	03/6/03	Pos	pol	gp41	CM1	F	141/169	187/231	169/185	301/305
	DP27	03/6/03	Pos	nd	nd	CM1	F	141/169	187/231	169/185	301/305
	DP80	03/6/03	Pos	nd	gp41	CM1	F	141/169	187/231	169/185	301/305
ID4	DP24	03/6/03	Neg[6]	pol	gp41	CM73	F	141/157	203/223	181/181	273/273
	DP25	03/6/03	Neg[6]	pol	gp41	CM73	F	141/157	203/223	181/181	273/273
	DP109	08/17/03	Pos	pol	gp41	CM73	F	141/157	203/223	181/181	273/273
ID5	EK505	07/11/04	Pos	pol	gp41	CM60	M	153/173	223/227	177/181	297/301
ID6	EK502	07/11/04	Pos	pol	gp41	CM60	M	153/161	195/223	177/189	297/305
	EK506	07/11/04	Pos	pol	nd	CM60	M	153/161	195/223	177/189	297/305
	EK507	07/11/04	Pos	pol	nd	CM60	M	153/161	195/223	177/189	297/305
ID7	EK516	08/10/04	Pos	nd	gp41	CM32	M	153/157	223/227	169/177	301/305
ID8	EK519	08/12/04	Pos	pol	gp41	CM67	F	149/169	227/227	177/185	293/301
ID9	MB66	06/12/03	Pos	pol	gp41	CM17	F	141/177	191/203	177/185	289/301
ID10	MB23	10/22/03	Pos	pol	gp41	CM13	F	149/153	247/247	177/189	293/301
	MB24	10/22/03	Pos	pol	gp41	CM13	F	149/153	247/247	177/189	293/301
ID11	MB189	03/10/03	Pos	nd	gp41	CM33	M	157/169	227/227	177/177	297/301
	MB191	03/10/03	Pos	nd	gp41	CM33	M	157/169	227/227	177/177	297/301
	MB192	03/10/03	Pos	pol	gp41	CM33	M	157/169	227/227	177/177	297/301
ID12	MB317	12/11/03	Pos	pol	gp41	CM44	F	165/173	223/227	181/185	297/305
ID13	MB245	04/18/03	Pos	pol	gp41	CM41	F	157/161	199/227	165/177	305/305
ID14	MB248	04/18/03	Pos	nd	gp41	CM42	F	145/149	191/203	165/165	289/293
ID15	LB186	03/07/03	Pos	pol	nd	CM37	M	153/173	191/223	177/189	289/297
ID16	LB7	~09/02	Pos	pol	gp41	CM82	F	161/173	227/231	185/189	297/301

[1]Western blot profiles are shown in Fig. 2.

[2]*pol* sequences recovered from samples MT148 and EK507 were 830 bp and 477 bp in length, respectively; all other *pol* fragments were 889 bp in length.

[3]As shown in fig. S1 and table S2.

[4]M, male; F, female.

[$]Microsatellite loci were amplified from fecal DNA; two alleles per locus are shown; homozygous loci were amplified a minimum of seven times to exclude allelic dropout.

[6]nd, not done.

[6]Microsatellite analysis identified two antibody negative (DP24 and DP25) and one antibody positive fecal sample (DP109) to share an identical genotype, indicating that they were derived from the same SIVcpz infected chimpanzee (ID4). This ape's fecal antibody levels were thus either below the limits of detection at the time of the first analysis, or she experienced acute SIVcpz infection and seroconverted in the interim between the first and last specimen collection.

Table S4. Genetic identification of uninfected chimpanzees at five field sites.

Individual	Fecal Sample[1]	Date	mtDNA haplotype[2]	Gender[3]	Locus[4] D18S536		Locus[4] D4S243		Locus[4] D10S676		Locus[4] D9S922	
ID17	DP69	04/12/03	CM1	F	141	173	199	203	181	193	297	301
ID18	DP3	02/15/03	CM1	M	157	169	227	247	185	185	301	305
	DP82	03/01/03	CM1	M	157	169	227	247	185	185	301	305
	DP130	11/07/03	CM1	M	157	169	227	247	185	185	301	305
	DP134	11/07/03	CM1	M	157	169	227	247	185	185	301	305
	DP163	11/17/03	CM1	M	157	169	227	247	185	185	301	305
ID19	DP67	04/12/03	CM1	M	173	177	227	235	177	181	297	297
	DP218	04/12/03	CM1	M	173	177	227	235	177	181	297	297
ID20	DP18	02/19/03	CM1	M	161	169	227	231	169	169	301	305
ID21	DP5	02/17/03	CM4	M	153	161	191	195	181	185	297	297
ID22	DP103	08/15/03	CM4	M	153	169	187	195	181	185	293	297
	DP104	08/15/03	CM4	M	153	169	187	195	181	185	293	297
	DP151	11/13/03	CM4	M	153	169	187	195	181	185	293	297
	DP159	11/17/03	CM4	M	153	169	187	195	181	185	293	297
	DP162	11/17/03	CM4	M	153	169	187	195	181	185	293	297
ID23	DP098	07/29/03	CM5	F	137	149	187	199	177	177	305	309
ID24	DP89	07/26/03	CM8	F	161	177	203	235	177	185	289	297
	DP090	07/26/03	CM8	F	161	177	203	235	177	185	289	297
ID25	DP091	07/26/03	CM8	M	137	177	203	223	169	177	289	297
	DP092	07/26/03	CM8	M	137	177	203	223	169	177	289	297
	DP110	08/19/03	CM8	M	137	177	203	223	169	177	289	297
ID26	DP142	11/11/03	CM8	M	141	145	227	231	181	181	297	301
ID27	DP140	11/11/03	CM8	F	145	169	231	235	181	185	301	305
ID28	DP14	02/18/03	CM8	M	149	177	227	235	185	189	289	297
	DP099	08/14/03	CM8	M	149	177	227	235	185	189	289	297
	DP100	08/14/03	CM8	M	149	177	227	235	185	189	289	297
	DP101	08/14/03	CM8	M	149	177	227	235	185	189	289	297
ID29	DP094	07/29/03	CM8	F	157	165	227	235	181	189	293	305
ID30	DP70	04/03/03	CM12	F	149	153	223	227	177	189	289	289
ID31	DP75	04/03/03	CM13	F	145	157	223	227	177	185	289	289
ID32	DP71	04/03/03	CM13	F	157	173	223	227	177	181	289	289
	DP72	04/03/03	CM13	F	157	173	223	227	177	181	289	289
	DP74	04/03/03	CM13	F	157	173	223	227	177	181	289	289
ID33	DP64	04/12/03	CM24	F	141	149	199	227	177	177	293	297
	DP86	04/12/03	CM24	F	141	149	199	227	177	177	293	297
ID34	DP63	04/12/03	CM24	M	157	161	187	195	173	185	297	297
ID35	DP127	11/07/03	CM24	F	161	173	195	223	181	185	297	301
	DP132	11/07/03	CM24	F	161	173	195	223	181	185	297	301
ID36	DP68	04/12/03	CM24	F	141	157	195	199	177	177	297	301
ID37	DP65	04/12/03	CM23	F	141	177	227	235	177	181	297	301
ID38	DP129	11/07/03	CM35	F	141	173	231	251	177	177	297	313
	DP133	11/07/03	CM35	F	141	173	231	251	177	177	297	313
ID39	DP154	11/14/03	CM74	F	165	169	227	243	169	181	305	305
	DP156	11/14/03	CM74	F	165	169	227	243	169	181	305	305
ID40	DP79	03/06/03	CM54	F	149	153	195	227	169	177	281	281
ID41	DP88	07/17/03	CM54	F	161	169	187	227	177	181	301	313
ID42	DP112	11/04/03	CM54	M	161	169	223	227	169	181	313	317
	DP113	11/04/03	CM54	M	161	169	223	227	169	181	313	317
	DP114	11/04/03	CM54	M	161	169	223	227	169	181	313	317
	DP115	11/04/03	CM54	M	161	169	223	227	169	181	313	317
ID43	DP124	11/06/03	CM76	F	157	161	191	223	185	189	285	289
ID44	DP121	11/05/03	CM76	F	157	165	231	235	169	185	297	301
	DP122	11/05/03	CM76	F	157	165	231	235	169	185	297	301
ID45	DP81	03/06/03	CM76	M	157	169	231	235	165	169	293	297
ID46	DP105	08/15/03	CM78	F	157	161	227	231	177	185	289	301
	DP106	08/15/03	CM78	F	157	161	227	231	177	185	289	301
	DP107	08/15/03	CM78	F	157	161	227	231	177	185	289	301
	DP160	11/17/03	CM78	F	157	161	227	231	177	185	289	301
ID47	DP4	02/17/03	CM70	F	153	173	187	231	185	185	277	297
ID48	DP6	02/18/03	CM71	F	149	153	195	227	169	177	289	305
	DP13	02/18/03	CM71	F	149	153	195	227	169	177	289	305
	DP102	08/14/03	CM71	F	149	153	195	227	169	177	289	305
ID49	DP12	02/18/03	CM72	F	161	173	227	255	185	189	289	289
ID50	DP7	02/18/03	CM72	F	169	173	227	231	185	185	289	309
	DP8	02/18/03	CM72	F	169	173	227	231	185	185	289	309
	DP16	02/18/03	CM72	F	169	173	227	231	185	185	289	309
ID51	DP9	02/18/03	CM73	F	157	161	231	255	181	189	273	313
	DP10	02/18/03	CM73	F	157	161	231	255	181	189	273	313
	DP11	02/18/03	CM73	F	157	161	231	255	181	189	273	313
	DP15	02/18/03	CM73	F	157	161	231	255	181	189	273	313
	DP20	02/19/03	CM73	F	157	161	231	255	181	189	273	313
	DP108	08/17/03	CM73	F	157	161	231	255	181	189	273	313
ID52	DP77	03/31/03	CM75	F	153	153	235	239	177	185	301	305
	DP095	07/29/03	CM75	F	153	153	235	239	177	185	301	305
	DP097	07/29/03	CM75	F	153	153	235	239	177	185	301	305
ID53	DP096	07/29/03	CM77	F	153	173	191	227	177	185	301	309
ID54	DP093	07/29/03	CM77	F	169	173	223	227	169	185	301	309
ID55	DP136	11/08/03	CM79	F	137	153	227	231	181	189	289	305
ID56	DP143	11/11/03	CM79	F	153	173	215	223	177	185	301	305
	DP150	11/11/03	CM79	F	153	173	215	223	177	185	301	305
ID57	DP141	11/11/03	CM79	F	165	173	223	231	181	185	277	305
	DP145	11/11/03	CM79	F	165	173	223	231	181	185	277	305
	DP146	11/11/03	CM79	F	165	173	223	231	181	185	277	305
	DP148	11/11/03	CM79	F	165	173	223	231	181	185	277	305
	DP149	11/11/03	CM79	F	165	173	223	231	181	185	277	305

Table S4. Genetic identification of uninfected chimpanzees at five field sites (continued).

Individual	Fecal Sample[1]	Date	mtDNA haplotype[2]	Gender[3]	Locus[4] D18S536		Locus[4] D4S243		Locus[4] D10S676		Locus[4] D9S922	
ID58	DP152	11/13/03	CM80	M	153	165	227	231	185	185	309	313
	DP153	11/13/03	CM80	M	153	165	227	231	185	185	309	313
	DP165	11/17/03	CM80	M	153	165	227	231	185	185	309	313
ID59	DP155	11/14/03	CM81	F	165	169	199	227	181	181	293	297
ID60	DP157	11/14/03	CM81	F	169	173	223	227	181	185	285	285
	DP158	11/14/03	CM81	F	169	173	223	227	181	185	285	285
ID61	EK508	07/13/04	CM4	F	157	169	227	247	181	189	293	305
ID62	EK509	07/13/04	CM4	F	169	173	227	227	177	181	301	305
ID63	EK513	08/09/04	CM7	F	145	173	223	227	177	185	297	309
ID64	EK522	08/18/04	CM7	F	165	173	227	231	177	177	293	297
ID65	EK501	07/08/04	CM20	F	145	157	195	195	177	189	297	297
ID66	EK521	08/18/04	CM24	F	145	157	191	243	185	185	273	305
ID67	EK511	08/08/04	CM42	F	169	173	191	195	177	185	301	305
ID68	EK510	07/13/04	CM61	F	169	173	187	195	181	185	297	297
ID69	EK503	07/11/04	CM5	M	141	157	195	227	177	181	293	309
	EK517	08/10/04	CM5	M	141	157	195	227	177	181	293	309
ID70	EK512	08/08/04	CM65	F	169	173	191	223	177	185	305	309
ID71	EK518	08/10/04	CM66	F	169	173	195	227	177	177	289	293
ID72	BB241	04/13/03	CM4	F	157	169	191	227	169	197	305	313
ID73	BB79	06/06/03	CM18	F	157	173	227	231	177	185	297	309
ID74	BB104	06/07/03	CM18	F	145	157	231	255	177	185	297	297
ID75	BB106	06/08/03	CM30	F	145	173	195	215	177	185	293	297
ID76	BB235	04/13/03	CM39	F	145	173	195	207	181	185	289	289
ID77	BB236	04/13/03	CM40	F	157	173	227	227	177	189	305	305
ID78	BB74	06/06/03	CM19	F	157	173	211	227	173	181	285	297
ID79	BB230	04/10/03	CM19	M	157	169	187	223	169	169	285	289
ID80	BB94	06/07/03	CM19	F	157	165	227	235	181	185	297	301
ID81	BB102	06/07/03	CM19	F	165	173	223	227	185	189	293	301
ID82	BB239	04/13/03	CM20	M	145	173	195	235	177	185	293	305
ID83	BB76	06/06/03	CM20	F	141	173	227	231	181	189	281	297
ID84	BB77	06/06/03	CM21	F	157	173	187	227	181	189	305	309
ID85	BB101	06/07/03	CM22	F	141	145	187	195	177	185	289	301
ID86	BB88	06/07/03	CM26	F	141	173	211	227	181	185	293	309
	BB95	06/07/03	CM26	F	141	173	211	227	181	185	293	309
ID87	BB237	04/13/03	CM27	M	165	173	227	231	169	177	281	309
	BB240	04/13/03	CM27	M	165	173	227	231	169	177	281	309
	BB89	06/07/03	CM27	M	165	173	227	231	169	177	281	309
ID88	BB100	06/07/03	CM28	F	145	157	223	239	169	177	289	297
	BB99	06/07/03	CM28	F	145	157	223	239	169	177	289	297
ID89	BB229	04/10/03	CM29	F	145	173	195	195	177	185	293	297
	BB234	04/13/03	CM29	F	145	173	195	195	177	185	293	297
	BB238	04/13/03	CM29	F	145	173	195	195	177	185	293	297
	BB93	06/07/03	CM29	F	145	173	195	195	177	185	293	297
ID90	MB135	11/08/03	CM5	M	141	173	191	231	181	181	313	313
ID91	MB140	11/08/03	CM9	M	165	173	227	231	169	177	281	309
ID92	MB97	06/12/03	CM12	M	153	169	195	227	189	189	273	293
ID93	MB318	12/13/03	CM33	F	165	169	195	227	181	189	301	305
ID94	MB319	12/13/03	CM33	F	165	169	227	247	177	181	297	305
	MB320	12/13/03	CM33	F	165	169	227	247	177	181	297	305
ID95	MB138	11/08/03	CM33	F	145	173	195	235	177	185	293	305
	MB139	11/08/03	CM33	F	145	173	195	235	177	185	293	305
ID96	MB250	11/08/03	CM33	M	149	153	195	227	177	181	309	313
ID97	MB324	12/15/03	CM34	F	165	169	227	247	169	185	297	297
ID98	MB190	03/10/03	CM38	M	161	169	227	227	189	193	305	305
ID99	MB246	04/18/03	CM42	M	141	157	203	235	185	185	289	297
	MB247	04/18/03	CM42	M	141	157	203	235	185	185	289	297
ID100	MB315	12/11/03	CM43	F	153	165	187	227	177	185	293	309
	MB316	12/11/03	CM43	F	153	165	187	227	177	185	293	309
ID101	MB323	12/15/03	CM45	M	145	173	227	227	169	185	297	305
ID102	LB307	01/10/04	CM13	M	149	153	227	247	177	181	301	305
	LB308	01/10/04	CM13	M	149	153	227	247	177	181	301	305
ID103	LB311	01/10/04	CM33	F	149	153	227	239	177	181	293	293
	LB312	01/10/04	CM33	F	149	153	227	239	177	181	293	293
	LB313	01/10/04	CM33	F	149	153	227	239	177	181	293	293
ID104	LB176	03/04/03	CM34	M	157	161	219	231	177	177	285	297
ID105	LB309	01/10/04	CM38	M	153	157	195	227	185	189	297	305
ID106	LB204	06/07/03	CM23	F	145	149	187	187	181	205	273	273
ID107	LB205	06/07/03	CM23	M	157	161	191	227	169	177	293	305
	LB208	06/07/03	CM23	M	157	161	191	227	169	177	293	305
ID108	LB174	03/04/03	CM28	M	157	173	191	243	177	189	285	289
	LB175	03/04/03	CM28	M	157	173	191	243	177	189	285	289

[1]Fecal samples collected from five different field sites (DP, EK, BB, MB and LB) were subjected to mtDNA, gender and microsatellite analyses. Only data from uninfected chimpanzees are shown. The microsatellite genotype of SIVcpz*Ptt* infected chimpanzees from these same five sites is shown in table S3.

[2]mtDNA haplotype as listed in table S2.

[3]M, male; F, female.

[4]Microsatellite loci were amplified from fecal DNA; two alleles per locus are shown; homozygous loci were amplified a minimum of seven times to exclude allelic dropout.

References

S1. E. S. Lader, United States Patent Application 20010016312 (2001).

S2. M. L. Santiago *et al.*, *J. Virol.* **77**, 7545 (2003).

S3. M. L. Santiago *et al.*, *J. Virol.* **77**, 2233 (2003).

S4. M. L. Santiago *et al.*, *J. Virol.* **79**, 12515 (2005).

S5. E. Nerrienet *et al.*, *J. Virol.* **79**, 1312 (2005).

S6. J. D. Thompson *et al.*, *NAR* **22**, 4673 (1994).

S7. F. Gao *et al.*, *Nature* **397**, 436 (1999).

S8. F. Simon *et al.*, *Nature Medicine* **4**, 1032 (1998).

S9. J. P. Huelsenbeck, F. Ronquist, *Bioinformatics* **17**, 754 (2001).

S10. K. M. Sullivan *et al.*, *BioTechniques* **15**, 637 (1993).