

# SIVsm Quasispecies Adaptation to a New Simian Host

Linda J. Demma<sup>1‡a</sup>, John M. Logsdon, Jr.<sup>1‡b</sup>, Thomas H. Vanderford<sup>1</sup>, Mark B. Feinberg<sup>2‡c</sup>, Silvija I. Staprans<sup>2\*</sup>

**1** Department of Biology, Emory University, Atlanta, Georgia, United States of America, **2** Departments of Medicine and Microbiology and Immunology, and Emory Vaccine Center, Emory University School of Medicine, Atlanta, Georgia, United States of America

**Despite the potential for infectious agents harbored by other species to become emerging human pathogens, little is known about why some agents establish successful cross-species transmission, while others do not. The simian immunodeficiency viruses (SIVs), certain variants of which gave rise to the human HIV-1 and HIV-2 epidemics, have demonstrated tremendous success in infecting new host species, both simian and human. SIVsm from sooty mangabeys appears to have infected humans on several occasions, and was readily transmitted to nonnatural Asian macaque species, providing animal models of AIDS. Here we describe the first in-depth analysis of the tremendous SIVsm quasispecies sequence variation harbored by individual sooty mangabeys, and how this diverse quasispecies adapts to two different host species—new nonnatural rhesus macaque hosts and natural sooty mangabey hosts. Viral adaptation to rhesus macaques was associated with the immediate amplification of a phylogenetically related subset of *envelope* (*env*) variants. These variants contained a shorter variable region 1 loop and lacked two specific glycosylation sites, which may be selected for during acute infection. In contrast, transfer of SIVsm to its natural host did not subject the quasispecies to any significant selective pressures or bottleneck. After 100 d postinfection, variants more closely representative of the source inoculum reemerged in the macaques. This study describes an approach for elucidating how pathogens adapt to new host species, and highlights the particular importance of SIVsm *env* diversity in enabling cross-species transmission. The replicative advantage of a subset of SIVsm variants in macaques may be related to features of target cells or receptors that are specific to the new host environment, and may involve CD4-independent engagement of a viral coreceptor conserved among primates.**

Citation: Demma LJ, Logsdon JM Jr, Vanderford TH, Feinberg MB, Staprans SI (2005) SIVsm quasispecies adaptation to a new simian host. PLoS Pathog 1(1): e3.

## Introduction

At least 40 primate species in Africa are infected by diverse simian immunodeficiency viruses (SIVs) assigned to six major phylogenetic lineages; however, the mosaic nature of the SIV genomes attests to the common simian-to-simian transmission of SIVs [1,2]. These African nonhuman primate reservoir hosts maintain normal CD4 T cell counts and avoid AIDS, despite lifelong SIV infection [3–6]. Our studies of naturally SIV-infected sooty mangabeys (SMs) indicate that these hosts are highly viremic, yet manifest far lower levels of aberrant immune activation and apoptosis than are seen in pathogenic SIV and HIV infections; these latter observations help to explain how SMs maintain numerically and functionally intact T lymphocyte populations [3]. Zoonotic transmission and sustained propagation of SIVcpz and SIVsm from SIV-infected chimpanzees and SMs, respectively, to humans [2,7], resulted in the human HIV-1 and HIV-2 AIDS epidemics.

SIV and HIV *env* sequence variation, including variation in length and glycosylation patterns, enables these viruses to utilize different coreceptors for infection, and to adapt to variation in the relative levels of the viral receptor (CD4) and coreceptors (e.g., CC-chemokine receptor 5 [CCR5]) to gain efficient entry into cells [8,9]. *Env* variation also enables the virus to readily escape antibody responses [10–12]. Our studies of SIVsm *env* diversity in naturally infected SMs demonstrate high levels of intrahost *env* variable region 1 and 2 (V1V2) amino acid diversity (median, 5.6%; range, 0%–38%) that are maintained by continual positive selection, presumably antibody mediated (unpublished data). Consid-

erable V1V2 amino acid length variation and high and variable numbers of glycosylation consensus sequences are also observed. This high diversity of SIV V1V2 in the natural host environment may promote the potential for cross-species transmission by generating the *env* variants necessary to ensure successful infection of new hosts.

For successful cross-species transmission to occur, including the continued propagation of an infectious agent in a new host species, the agent must be able to replicate at levels in the new host that ensure its sustained passage to new individuals of that species; otherwise the newly infected

Received March 10, 2005; Accepted June 20, 2005; Published September 30, 2005  
DOI: 10.1371/journal.ppat.0010003

Copyright: © 2005 Demma et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: CCR5, CC-chemokine receptor 5; GTR, general time reversible; IV, intravenous; NJ, neighbor-joining; ML, maximum likelihood; p.i., postinfection; RM, rhesus macaque; SI, source inoculum; SIV, simian immunodeficiency virus; SM, sooty mangabey; V1, SIV *envelope* variable region 1; V2, SIV *envelope* variable region 2

Editor: Richard A. Koup, National Institutes of Health, United States of America

\*To whom correspondence should be addressed. E-mail: sstapr2@sph.emory.edu

‡a Current address: Foodborne and Diarrheal Disease Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America

‡b Current address: Department of Biological Sciences, Roy J. Carver Center for Comparative Genomics, University of Iowa, Iowa City, Iowa, United States of America

‡c Current address: Merck Vaccine Division, Merck and Company, West Point, Pennsylvania, United States of America

## Synopsis

Why do some infectious agents establish successful cross-species transmission while others do not? Despite the clear potential for diseases harbored by animals to become emerging human pathogens, this question remains unanswered. Certain simian immunodeficiency viruses (SIVs) responsible for the human HIV-1 and HIV-2 epidemics have succeeded in infecting new host species, including humans. This study provides clues to how an SIV adapts to a new host in an experimental cross-species transmission. Indeed, many emerging diseases are caused by highly mutation-prone RNA viruses like SIV, which exist not as a single species, but rather as a population of genetic variants within a single infection. The presence of numerous viral variants in an infected animal increases the chance that variants with the ability to enter into or multiply in a new host species are present. This study describes how an SIV population from a natural reservoir host, the sooty mangabey, adapts to a new monkey species, the rhesus macaque. A limited subset of SIV variants containing unique viral surface proteins appears well suited to multiply in the new host. This study documents how viral variation facilitates cross-species transmission, and highlights the particular importance of immunodeficiency virus envelope variants in infecting new hosts.

host(s) will simply represent a “dead-end” infection that does not lead to secondary and sustained infections in the new species. Alternatively, the infectious agent that has been recently transmitted to a new host may require the accumulation of mutations that enable it to replicate at levels high enough to ensure continued transmission to new individuals. Thus, SIVs that are capable of quickly adapting to new hosts and replicating to high levels are most likely to successfully breach the species barrier and continue to spread in the new species. Adaptation of naturally occurring SIV quasiespecies to new hosts has not been studied. In studies analyzing the adaptation of diverse HIV-1 quasiespecies from identifiable human donors to newly infected “recipients,” the early expansion of viruses that are homogeneous in *env* sequences, macrophage-tropic, and CCR5-utilizing is described [13,14]. This sequence homogenization is not observed in *gag* [15], suggesting that multiple variants are transmitted, followed by selection for particular *env* variants during primary infection. Selection for *env* homogeneity has also been reported after parenteral inoculation [13], suggesting that, separate from any selective processes taking place at the mucosal barrier, there is strong selection for particular *env* genotypes during acute infection. Recently, a study of heterosexual HIV-1 transmission demonstrated that viruses encoding compact, glycan-restricted Envs with exposed neutralizing epitopes were significantly favored in newly infected hosts [16]. Another report confirmed these findings in transmission of subtype A but not B [17]. Whether these observations extend to other clades, cohorts, or routes of HIV infection remains to be determined [18,19].

Here we describe the adaptation of diverse SIVsm quasiespecies to the new rhesus macaque (RM) host, and compare quasiespecies evolution in natural SM and non-natural RM hosts. During the first days of infection, SIVsm replicated as well in the RM host as in the original host, if not better, apparently due to the robust replicative capacity of a subset of viral variants containing a shorter V1 loop and lacking two specific glycosylation sites. This study demon-

strates how viral quasiespecies diversity, by providing multiple variants, some of which can replicate to high levels in new hosts, may facilitate cross-species transmission.

## Results

### High Diversity of the SIVsm Quasiespecies Inoculum

The uncloned SIVsm inoculum consisting of plasma from a naturally infected SMs contained  $4 \times 10^6$  SIV RNA copies/ml. To characterize the diversity of this source inoculum (SI), and the molecular behavior of the quasiespecies upon transmission to new hosts, we analyzed a 456-nucleotide region spanning the variable V1V2 region of *env* and a 421-nucleotide region of the p27 capsid region of the more functionally conserved *gag* gene (GenBank accession numbers AY852284–AY853166). We chose to sequence only portions of the coding sequences of these two genes, as efforts to amplify full-length coding sequences resulted in poor RT-PCR amplification efficiencies that were not compatible with the reliable sampling of multiple quasiespecies variants. Sequences representing actively replicating SIV were amplified directly from virion RNA in the plasma by RT-PCR. 29 V1V2 and 7 *gag* clone sequences analyzed using maximum parsimony, neighbor-joining (NJ), and maximum likelihood (ML) phylogenetic tree constructions demonstrated that the SI was phylogenetically distinct from commonly used laboratory SIV isolates (Figure S1).

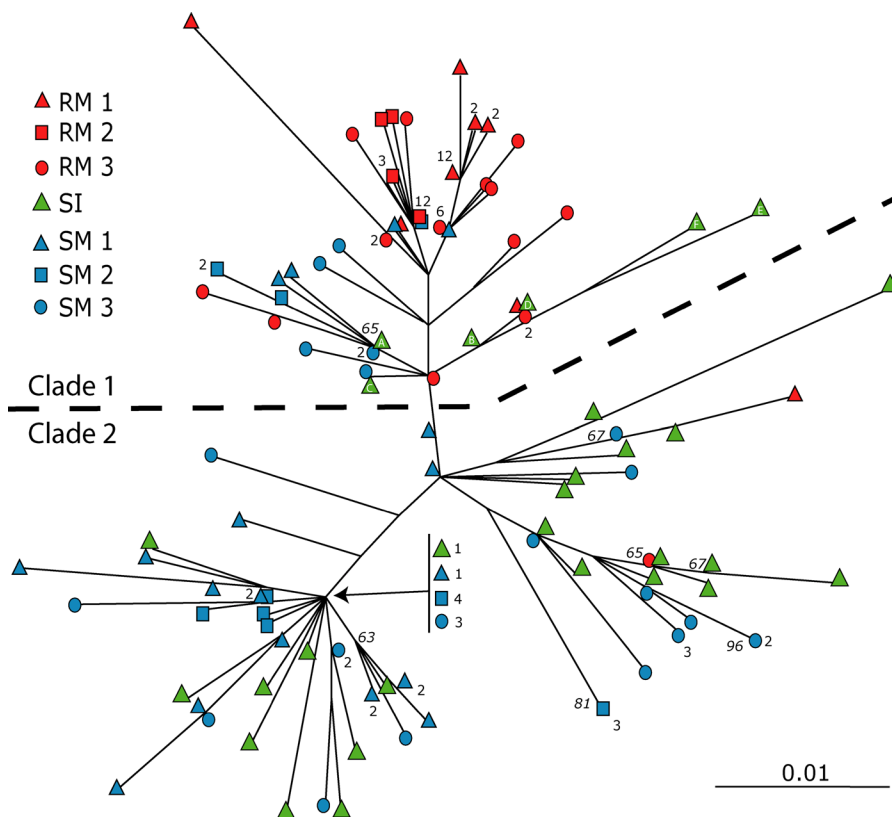
SI V1V2 sequence length varied between 139 and 143 aa (Table 1). The range of pairwise nucleotide diversity calculated for the SI population was 0.3%–5.1% for V1V2 (mean, 2.7%; median, 2.7%) and 0.7%–4.6% for *gag* (mean, 2.4%; median, 2.3%). The amino acid diversity ranged from 0% to 12.8% in V1V2 (mean, 5.9%; median, 6.3%) with only six of 406 identical pairwise comparisons. The amino acid diversity in *gag* p27 was lower, ranging from 0% to 0.7% (mean, 0.2%; median, 0%). The minimal diversity detected in *gag*, which was PCR-amplified using identical conditions, suggests that the observed V1V2 diversity is not the result of PCR-introduced mutation. The average viral diversity ob-

**Table 1.** V1V2 Amino Acid Variation in RMs and SMs at Days 10 and 14 and in SI

Variant Type	SI	Day 10					Day 14						
		RM1	RM2	SM1	SM2	RM1	RM2	RM3	SM1	SM2	SM3		
Length (aa)	143	0.03	0	0	0.05	0.00	0	0	0	0	0	0	
	142	0.38	0.05	0.05	0.60	0.29	0	0	0	0.71	0.53	0.30	
	141	0.03	0	0.05	0.10	0.15	0	0	0	0.05	0.20	0.07	
	140	0.34	0.86	0.75	0.20	0.12	0.90	1.00	0.86	0.24	0.27	0.30	
	139	0.21	0.05	0.05	0.05	0	0.10	0	0.10	0	0	0.26	
	138	0	0.05	0.05	0	0	0	0	0.05	0	0	0.07	
	137	0	0	0.05	0	0	0	0	0	0	0	0	
Glyco-sylation	Glyc.	0.52	0.05	0.05	0.80	0.44	0	0	0.05	0.71	0.73	0.74	
	Pos. 7												
	Glyc.	0.52	0.05	0	0.70	0.44	0	0	0.05	0.71	0.73	0.37	
	Pos. 19												

Shown are the frequencies within each set of sequences for each V1V2 amino acid length variant (137–143 aa) and the frequencies of glycosylation at amino acid positions 7 and 19. The amino acid length present at the highest frequency is underlined. For SI, the two amino acid lengths present at similarly high frequency are underlined and highlighted, corresponding to those seen in RMs (blue) and SMs (red). Note the homogeneity and shorter length of RM V1V2.

DOI: 10.1371/journal.ppat.0010003.t001



**Figure 1.** Phylogenetic Tree of V1V2 Variants

An NJ tree of all day 14 and SI V1V2 variants, constructed using a GTR model of evolution. Bootstrap support values greater than 50% are shown in italics at nodes and the number of multiple clones from the same animal at the ends of branches is indicated beside the symbol.  
DOI: 10.1371/journal.ppat.0010003.g001

served in this study is similar to that reported in other studies of SIV infections of natural hosts [4,20,21]. The within-host extremes of V1V2 diversity observed in this and another study (unpublished data) is a novel observation resulting from the large number of sequences analyzed.

#### Robust Virus Replication Demonstrates Immediate Quasispecies Adaptation to the New Nonnatural RM Host

SMs and RMs were intravenously (IV) inoculated with 1 ml of the SI described above (SIVsm). IV injection may partially recapitulate the circumstances of cross-species SIV transmission, which are thought to involve exposure to bloody flesh during hunting or butchering [22,23]. It ensures reproducible infection and enables the study of host-specific differences in response to SIV infection that lead to AIDS in RMs but not SMs [24]. At day 7 postinfection (p.i.), SIVsm replication was detected in all animals except RM2 (Figure S2). Peak viremia occurred between days 10 and 14 for all animals except RM2, whose peak likely occurred between days 14 and 28, an interval when no sampling was performed. RM1 and RM3 manifested peak viremia levels of  $1.6 \times 10^9$  copies/ml of plasma and  $6.1 \times 10^8$  copies/ml respectively, higher than the peak viremia for the three SMs ( $5.0 \times 10^7$  to  $1.5 \times 10^8$  copies/ml). Viral loads declined to similar set point levels of  $\sim 1 \times 10^6$  copies/ml, except for RM2, which maintained fewer than 1,000 copies/ml after day 60. RM1 and RM3 developed AIDS 2.5 and 3.5 y p.i. and were

euthanized. Divergent host responses and disease outcomes during primary SIVsm infection of SMs and RMs are described elsewhere [24].

#### Early Amplification of a Phylogenetically Related Subset of *env* Variants in RMs Contrasts with Unrestricted *env* Diversity in SMs

At day 14 p.i., the replicating *env* V1V2 sequences for all six animals were compared to each other and to the SI. Despite the more robust replication of SIVsm in the RMs, few of the SI V1V2 variants appeared among the clade containing most of the variants replicating in RMs (clade 1; Figure 1), demonstrating that a subset of genetically related *env* variants was amplified during acute SIVsm infection of RMs. Specifically, the proportion of variants replicating in RMs in clade 1 (97%) was significantly higher than that of either the SI variants (21%; Marascuillo Procedure,  $p < 0.0001$ ) or the variants replicating in SMs (22%;  $p < 0.0001$ ). In contrast, there was little selection of specific SIVsm *env* variants upon transfer to new naïve SMs, with no significant difference (Marascuillo Procedure,  $p = 99\%$ ) in the distributions of SI variants and variants replicating in SMs between clades 1 and 2 (Figure 1). Mean intrahost pairwise nucleotide diversity in the RMs at day 14 was 0.9%, significantly lower than that of the SMs (2.1%; Tukey's HSD,  $p < 0.05$ ) and the SI (Newman-Keuls,  $p < 0.05$ ). Mean intrahost amino acid diversity in RMs at day 14 (1.6%) was lower than SM amino acid diversity

	10	20	30	40	50	60
	CIAMRCNKTE	TDKWGLTRNAT	ATTATTTTSTTTT	TKPDVTAKVIN	ESNPCIONNS	CAGLE
<b>Source</b>	.....R..GT..AT..--..P--	.....K..D..E..				
	.....T..T---..--A..A..	.....S..				
	.....E.....ATV..---	.....				
	.....R..GT..PT..---	.....T..K..E..				
<b>Inoculum</b>	.....T..T---..--A..A..	.....D..M..				
	.....T-.....T..K..	.....D..M..				
	.....R.....T.....T..K..	.....D..YM..				
	.....T..T---..T-S..T.G..	.....K..E..				
<b>RM1</b>	.....E.....ATV---	.....				
	.....A.....ATV---	.....D..				
	.....D.A.....ATV---	.....				
	.....D.A.....ATV---	.....A..				
<b>day 10 p.i.</b>	.....D.A.....ATV---	.....				
	.....D.A.....ATV---	.....V..				
	.....D.A.....ATV---	.....T..K..D..M..				
	.....G..A.....ATV---	.....				
<b>SM1</b>	.....D.A.....ATV---	.....E..				
	.....R..GT..PT..---	.....T..N..K..E..				
	.....S.....TA.T..IA.T-S..T.S..	.....Y..D..R..				
	.....D.A.....ATV---	.....T..K..D..M..				
<b>day 10 p.i.</b>	.....A..E.....ATV---	.....K..D..S..				
	.....A..E.....ATV---P-	.....K..D..S..				
	.....V.....T..K..D..M..	.....				
	.....V.P---AT-A..AT..TE..R..D..	.....				
<b>RM1</b>	.....R..GT---APT.....	.....K..E..				
	.....R.....A.P---AT..NE..	.....GS..				
	.....S.....SP---T-A..AT..TEAV..D..D..	.....ED--Q				
	.....T.....G---TA.P.....A.A..E..	.....S..R..S..				
<b>day 578 p.i.</b>	.....T.....G---TA.T.....A..A..K..	.....YL..K..E..				
	.....T.....AG---TA.T.....A..A..K..	.....YL..K..E..				
	.....T.....G---DA.T.A.T-S..A..A..K..	.....YL..K..E..				
	.....T.....G---TA.T.....A..A..K..	.....YL..K..E..				
<b>SM1</b>	.....T.....S-ATA.TA.....T..R..	.....Y..K..D..E..				
	.....T.....I---I..T..A..V..A..M..	.....				
	.....R.....P-I..T..N..E..	.....				
	.....R.....T..N..E..	.....				
<b>day 578 p.i.</b>	.....T.....Q..GT..T.....IT.....ND..S..	.....				
	.....R.....DT.....A.....A..M..	.....				
	.....T.....T.....I---I..T..A..A..M..	.....				
	.....T.....Q.....GT..AP---AT..E..K..D..M..	.....				
<b>SM1</b>	.....R.....DT.....A.....A..M..	.....				
	.....T.....I---I..T..A..RD..L..	.....				

**Figure 2.** Pattern and Temporal Dynamics of Protein Sequence Evolution in Envelope

V1V2 amino acid sequence for SI, SM1, and RM1 at days 10 and 578. The consensus of all sequences is indicated at the top with amino acid positions labeled above. Glycosylation consensus motifs (NXT/S) are highlighted in yellow.

DOI: 10.1371/journal.ppat.0010003.g002

(3.4%), although not significantly. This pattern of restriction was observed as early as 10 d p.i., with 37/39 (95%) of RM variants clustering with the same six SI variants seen at day 14 (unpublished data). Thus, only a subset of *env* genotypes appear well suited to replicate in the new RM host environment, but this subset replicates to surprisingly high levels.

#### SIVsm *env* Variants That Are Preferentially Amplified in Newly Infected RMs Contain Shorter V1 Regions and Lack Two Glycosylation Consensus Sequences

SIV *env* glycosylation is important in receptor and coreceptor utilization [25], and in evading neutralizing antibodies [26,27]. Eight predicted N-linked glycosylation sites (N-gly) containing the amino acid motif NXT/S were identified along the 125 aa of V1V2 analyzed, although direct

evidence of glycosylation at these sites is not explicitly demonstrated. Among the 29 SI clones analyzed, most of these positions encoded the consensus N-gly sequence (Figures 2 and S3), consistent with the observation that SIVsm V1V2 is highly glycosylated in its natural host (unpublished data).

Two of these predicted N-gly sites, at positions 7 and 19 in V1 (Figure 2), were immediately selected against in the newly infected RMs, before the anticipated development of an antibody response. (RMs and SMs demonstrated anti-SIV antibodies by ELISA at day 40 p.i., the first time-point assessed, with increasing titers by day 130 p.i. [unpublished data]). In the SI, 76% and 52% of V1V2 sequences exhibited the N-gly motif at positions 7 and 19. At 10 d p.i., the motif at position 7 was present in 80% of SM1 sequences, but in only



5% of RM1 sequences. At position 19, the motif was present in 70% of SM1 sequences compared to 5% of RM1 sequences (Figure 2 for SM1 and RM1 at day 10; Table 1 for frequencies in all animals). The near absence of the motif at positions 7 and 19 was observed in all RMs analyzed at days 10 and 14 (Figure S3). Furthermore, the predominant V1V2 variants in RMs at days 10 and 14 were shorter in length by two amino acids compared to the variants in SM (Table 1). Thus, a disadvantage of variants with longer V1 loops and two specific N-gly sites in V1 may explain the restricted outgrowth of specific *env* variants during early infection of RMs.

At days 10–100, SIVsm sequences from SMs had a greater mean number of N-gly sites per sequence than variants in RMs, but by day 578 the overall frequency of glycosylation consensus motifs increased in both species, and there was no difference between species (Figure S4;  $p < 0.001$ ). This increase in mean glycosylation over time in the RMs is in part due to the reemergence of variants containing the two specific N-gly sites that were absent in the majority of early RM variants. The range of V1 region amino acid length variations also increased over time, and no species-specific differences were seen at day 100 and thereafter (unpublished data). These data demonstrate continual evolution of V1 in both SMs and RMs.

### Increasing Positive Selection in SMs and RMs at Later Times Postinfection

To compare selection pressures between hosts and time points, nonsynonymous and synonymous nucleotide substitutions (dN and dS, respectively) at each codon of V1V2 were calculated for sequences at day 14 (prior to seroconversion) and day 578 (chronic infection) for each animal and the SI (Figure 3). The pattern of selection in SMs at both time points (Figure 3A and 3B) was similar to the SI (Figure 3E), suggesting few changes in selection pressure upon SIVsm transfer to naïve SMs. This is consistent with the phylogenetic analyses, which indicate that IV transfer of SIVsm does not subject the quasispecies to any significant selective pressures or bottleneck in the natural SM host, but results in considerable restriction of the SIVsm quasispecies diversity in RMs. In contrast to the SMs, the relative lack of sites under strong selection in RMs at day 14 (Figure 3C) corroborates the strong, early selection of a subset of variants from the SI. The subsequent substantial increase (Figure 3D) in the number of sites under selection and the magnitude of selection at those sites not only reflect the outgrowth of variants more similar to the SI, but also suggest the presence of immune-selective pressures in the RMs during the postacute phase of infection.

To quantify the magnitude of selection in the SIVsm *env* V1V2, cumulative dN and cumulative dS were calculated for each animal at days 14 and 578 (Figure 3F). SMs and RMs showed relative increases in cumulative dN and dS at day 578 (Wilcoxon rank sum test,  $p < 0.005$ ), especially in a region of V1 (amino acid positions 22–57) described as important in antibody escape [25,28–30]. At later times, despite increases of both dN and dS in RMs, cumulative dN-dS was greater in RMs than SMs, although the difference was not statistically significant, suggesting greater positive selection pressures in the non-natural host. However, continual evolution of V1V2 occurred in both species, consistent with observations of persistent within-host positive diversifying selection in SMs (unpublished data).

### Variants Related to the Original Inoculum Reemerge in RMs at Later Times Postinfection

Phylogenetic analyses of clones from day 100 p.i. showed V1V2 sequences beginning to diversify in RMs, although the variants remained clustered by host (Figure 4; see Figure S5 for parallel phylogenetic analysis of SM1). At this time, viral sequences in RM1 were more closely related to variants from the SI than to variants from 10 d p.i., suggesting a reemergence of the SI-related quasispecies during chronic infection. These results indicate that day 10 V1V2 variants are an evolutionary “dead-end,” as it is unlikely that directional evolution would result in viral quasispecies in all RMs that are highly related to the original SI quasispecies.

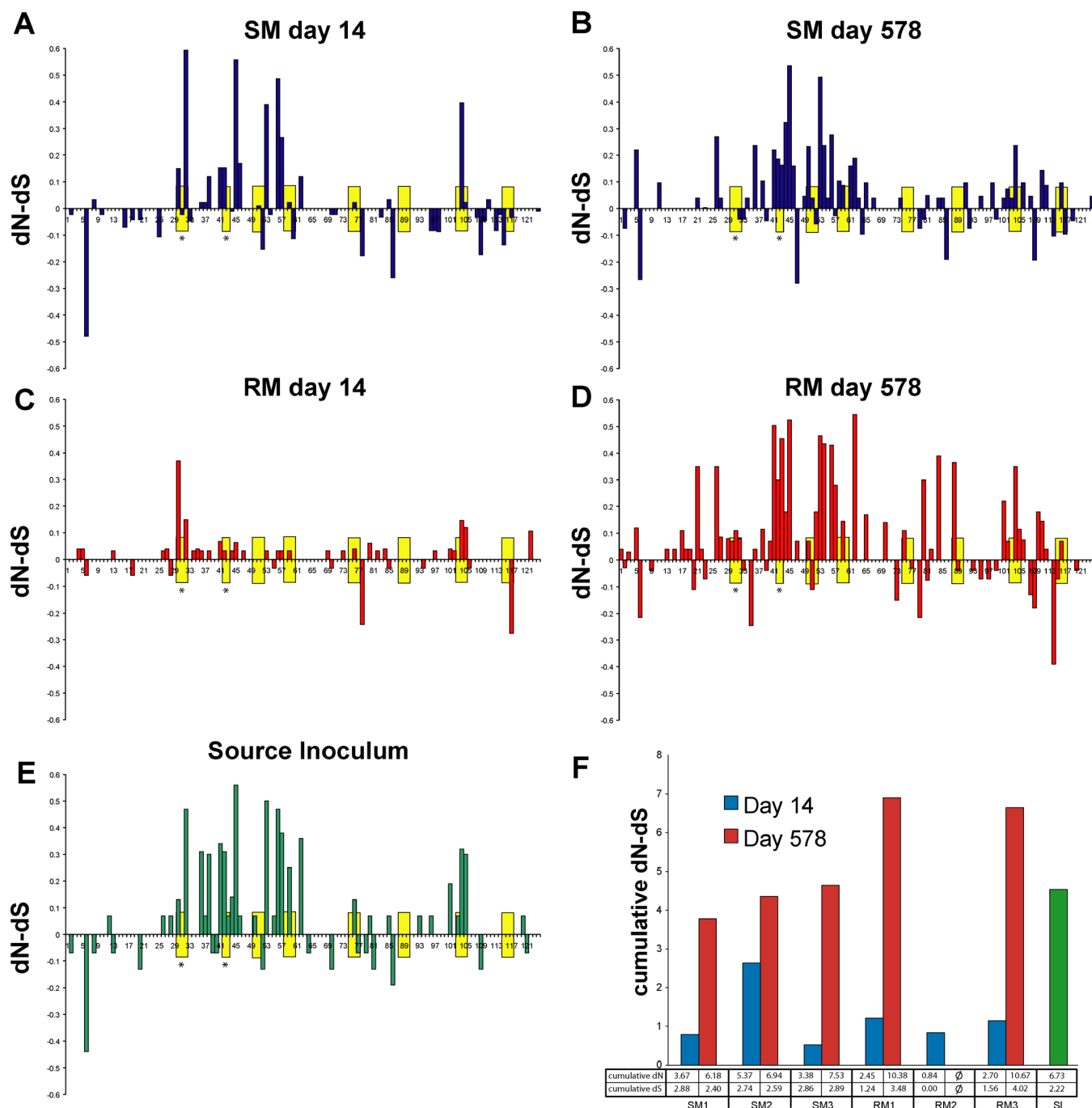
At all time points after infection of the three SMs, average nucleotide diversity of SIVsm V1V2 sequences remained similar to that of the original SI (~3.0%; Figure 5). In contrast, in RMs, nucleotide diversity increased after day 40 despite manifesting an initial restriction in viral diversity. Viral variants at day 578 became more animal-specific (unpublished data), as would be expected under host-specific selection pressure. The viral diversity in RMs at day 578 (averaging  $4.5\% \pm 0.8\%$ ) was greater than both that of the SI and that of the SIVsm variants observed in SMs at late times (t-test with Bonferroni adjustment,  $p = 0.03$ ). These data suggest that selection pressures change during the course of SIVsm infection of RMs; V1V2 variants that replicated to high levels in primary infection lost their replicative advantage, and previously undetected variants that were closely related to the SI became detectable. The increasing nucleotide diversity over time in RMs is consistent with the observed increase in positive diversifying selection pressures in RMs (see Figure 3).

### No Early Selection for Specific *gag* Variants Following Intra- or Cross-Species Transmission

Despite the high levels of selection in *env*, no species-specific phylogenetic relationships were observed for SIVsm *gag* variants at day 10 p.i. (Figure 6), indicating that there was no preferential amplification of specific *gag* variants in association with the establishment of infection in either SMs or RMs. The average nucleotide diversity of *gag* variants following transmission to both species was similar to that of the SI (unpublished data). These data suggest that most *gag* variants were equivalent in their ability to establish successful infection of either host. At later times, some amino acid changes in *gag* became apparent in individual animals (Figure S6). In both RMs analyzed at days 100 and 578, there was almost complete amino acid fixation at two sites (positions 39 and 68). Only one SM manifested any evidence of amino acid fixation in *gag*, and this was only partial (position 126 in SM2). Fixation of amino acid changes, particularly at position 68, which occurred in two RMs, could be due to cell-mediated immune response pressures, which are thought to be stronger in RMs compared to SMs [24]. However, that these changes are due to random amino acid fixation through genetic drift cannot be ruled out, because of the large population sizes involved and the limited number of *gag* clones analyzed.

### Discussion

Identification of specific characteristics that enable pathogens to infect new species may reveal why some emerging



**Figure 3.** Evidence for Greater Positive Selection in SIV *env* at Later Times Postinfection

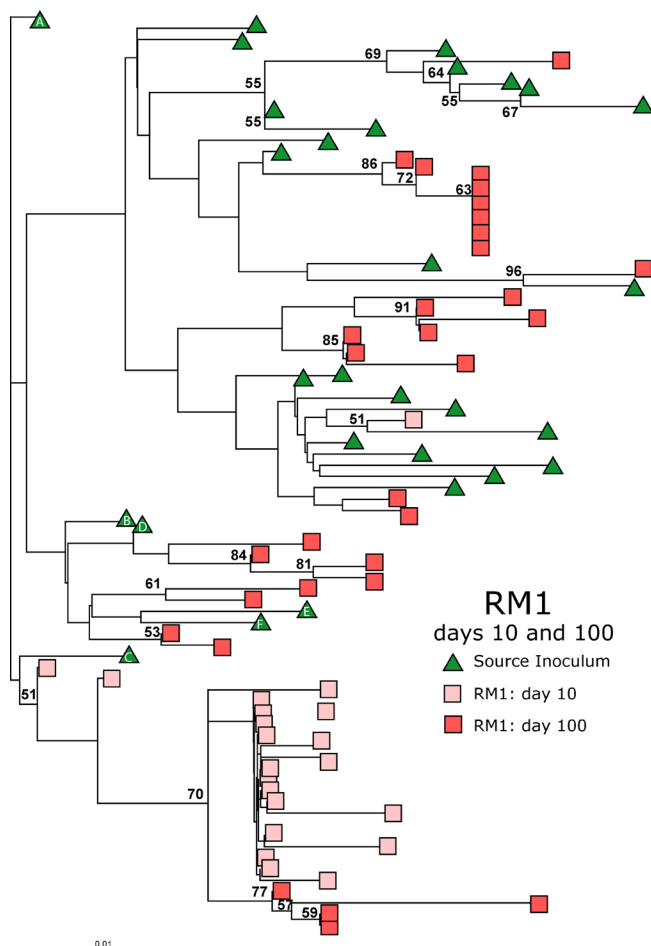
(A–E) Calculations for dN and dS were performed along the 124 amino acids of the V1V2 region using SNAP (<http://hiv-web.lanl.gov/>). The average dN and dS at each codon is shown for SMs at day 14 (A) and day 578 (B), as well as for RMs at day 14 (C) and day 578 (D), and for the SI (E). Yellow boxes indicate predicted N-gly sites, and asterisks indicate N-gly sites not present at early time points in RMs.

(F) Cumulative dN and dS are shown across all sites for each animal at day 14 and day 578. Raw values of cumulative dN and cumulative dS are indicated below the graph.

DOI: 10.1371/journal.ppat.0010003.g003

infections become widespread while others do not. RNA virus quasispecies diversity has been posited to underlie their zoonotic success, yet no study had analyzed the behavior of diverse naturally occurring viral quasispecies upon inoculation into different host species. This study represents the first analysis, to our knowledge, of the evolution of a diverse naturally occurring SIV quasispecies, following its side-by-

side inoculation into a new nonnatural host species (rhesus macaques) and the natural host species from which it was derived (sooty mangabeys). Our studies, which focused on the intensive sequencing of large numbers of viral variants in the *env* V1V2 region, point to the importance of diversity in this region in initiating a successful cross-species infection event. However, this does not exclude the possibility that diversity in



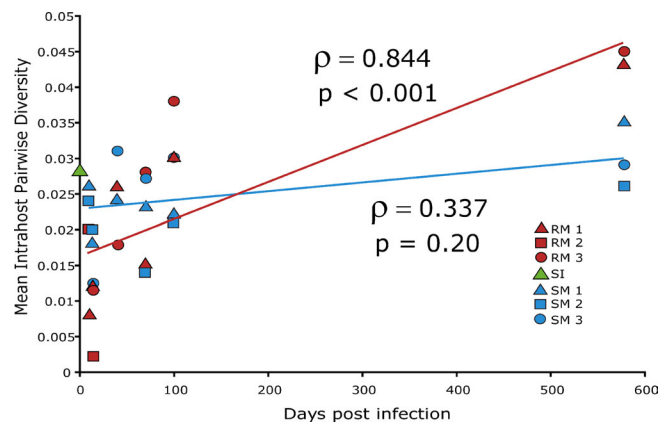
**Figure 4.** Reemergence of Donor-Related Variants in Late Infection of Nonnatural Hosts

ML tree of sequences from the SI and RM1 at days 10 and 100. Bootstrap values greater than 50% are shown at nodes. The SI variants are identified by the legend.

DOI: 10.1371/journal.ppat.0010003.g004

other genome regions, including diversity in other *env* regions, plays an important role in cross-species transmission events.

Upon inoculation of SMs with the diverse SIVsm quasispecies, little host restriction was observed during acute infection despite continued strong positive selection pressures consistent with host-specific viral evolution and similar to our findings in a study of natural infection in SMs (unpublished data). However, a restricted, genetically related subset of SIVsm *env* V1V2 variants that harbored a shorter V1 loop and lacked two specific glycosylation sites was preferentially amplified in all of the RMs during acute infection. This was observed despite IV inoculation, which would have bypassed mucosal barriers, and was observed as early as 10 days p.i., likely prior to the development of immune responses. While we cannot rule out that these variants hitchhiked to a high frequency in the RMs, the observed amplification of a subset of variants appeared to represent an advantage for these envelopes that was related to specific features of target cells in the RM but not the SM. Loss of one of these glycosylation sites (position 7) has been shown to



**Figure 5.** Temporal Changes in Nucleotide Sequence Evolution in SMs and RMs

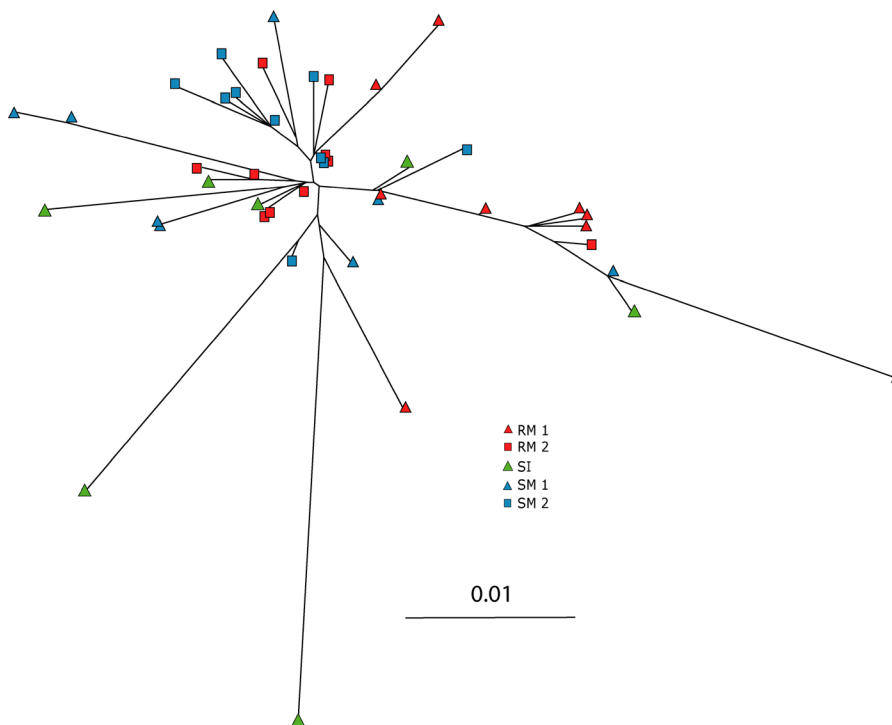
Mean pairwise nucleotide diversity of the V1V2 sequences for each animal at each time point, calculated using the Tamura-Nei model of nucleotide substitution in MEGA 2.1 [52]. The diversity of the SI is indicated on the y-axis. Trend lines are drawn for RMs (red) and SMs (blue).

DOI: 10.1371/journal.ppat.0010003.g005

result in CD4-independent SIVs in the SIVmac239 strain [25], suggesting the possibility that viral variants that preferentially use CCR5 independently of CD4 may be selected for during acute infection of the new RM host. It remains to be determined whether loss of this same glycosylation site in SIVsm also results in CD4 independence. Because CCR5 is more highly conserved between SMs and RMs than is CD4 [31,32], it might be anticipated that CD4-independent viral variants could overcome species differences in the primary viral receptor (CD4) and have a distinct advantage in the new host environment. The possibility that efficient coreceptor utilization independent of CD4 is an important factor in establishing cross-species transmission is a topic for further study.

Recently, the selection of more compact, glycan-restricted HIV *env*s after heterosexual transmission was described [16,17]. If the selection of slightly shorter, glycan-restricted SIVsm *env*s observed in this study of cross-species transmission is a related phenomenon, then, because our IV inoculations bypassed mucosal barriers, the advantage of such variants may be a posttransmission phenomenon. One possibility is that in an antibody-naïve host environment, more compact, less glycosylated Env conformations with more accessible receptor-binding domains have a replicative advantage. However, it is noteworthy that SIVsm encoding less glycosylated V1V2 regions do not appear to have any replicative advantage in newly infected, antibody-naïve SMs. The lack of selective pressure on the SIVsm quasispecies in acutely infected SMs suggests that highly glycosylated V1V2 variants are well adapted to initiate new infections of its natural host species.

Although it might be expected that only a subset of SM-adapted SIVs could replicate well in a new species, it is intriguing that these variants replicated to levels exceeding those seen in the natural SM host. In this [24] and other studies of acute SIV infection [33,34], we have observed a relationship between the magnitude of early CD4 T cell activation and the magnitude of early virus replication. Given the higher levels of CD4 T cell activation observed in the



**Figure 6.** Phylogenetic Analysis of SIVsmm *gag* Variants in Natural and Nonnatural Hosts

An NJ unrooted phylogenetic tree of all day-10 *gag* variants was constructed in PAUP\* [43] using the GTR model with a gamma rate distribution of shape  $\alpha = 1.0$ . The SI variants are represented by triangles and identified by the legend. Bootstrap values greater than 50% are shown at nodes, and the number of multiple clones from the same animal at the ends of branches is indicated within the symbol.

DOI: 10.1371/journal.ppat.0010003.g006

acutely infected RMs as compared to the SMs in this study [24], it is conceivable that increased numbers of activated CD4 T cells provided additional cellular targets for infection. If this target cell-driven hypothesis of more robust SIVsm replication in RMs is correct, it raises the possibility that SIV infection-induced CD4 T cell activation in nonnatural hosts actually facilitates zoonotic transmission of these CD4 T cell-tropic lentiviruses. Additional studies are required to explore this hypothesis. It is also worth noting that activated CD4 T cells may up-regulate CCR5 (or other coreceptor) expression [35], and down-regulate CD4 expression [36]. This might provide another selective force for the observed glycan-restricted SIVsm variants that may be less CD4 dependent in newly infected RMs. Whatever the explanation for the selection of specific *env* V1V2 variants in RMs, selection pressures giving rise to these effects need not be strong, given the high level of diversity of the inoculum and the likely number of replication cycles involved. Nonetheless, the robust replication of these variants ensured the establishment of high viremia during infection of a new host, a characteristic that would be important for continued propagation of the virus in the new species.

At later times p.i. of RMs, SIVsm variants more closely related to the original SI quasispecies reemerged, suggesting that all variants were initially transmitted to the RMs, but that only a subset of variants replicated to high levels during the acute infection period. Variants related to the SI may have been physically sequestered in resting memory cells, as has been suggested for HIV-1 [37], or simply replicated at such low levels that they were not sampled. Studies have suggested that “archival” HIV variants are maintained in infected hosts

[38]. When host selection pressures change, such as with the termination of antiretroviral therapy, these archived variants may emerge, obviating the necessity for back mutation of the most predominant viral variants at the time of change in selection pressure. This capacity to archive the variants present in a diverse, infecting swarm, referred to as the “molecular memory” of the quasispecies [39], demonstrates the significant potential of lentiviral quasispecies to respond to changing selection pressures and presents significant hurdles when considering HIV prevention or treatment measures. SI V1V2 variant emergence at later times suggests that these viruses have replicative advantages in chronically infected RMs, perhaps due to their resistance to neutralizing antibodies. Compensatory changes in other regions of the genome (e.g., in the CD4-binding region of SIV *env*) could also have relieved initial selection pressures against these variants.

This study demonstrates how viral quasispecies diversity enables successful cross-species transmission by providing multiple variants, some of which are able to establish high-level viremia in new hosts, which, in turn, increases the probability of successful propagation within new species. Our studies point to SIVsm *env* diversity in its reservoir host as a likely required, although not necessarily sufficient prerequisite for successful cross-species transmission. These observations have implications for which infectious agents may be zoonotically transmitted and efficiently propagated in a new host species. Finally, the potential roles of CD4-independent SIVs and coreceptor sequence conservation in cross-species transmission are important topics for further study.



## Materials and Methods

**Experimental SIV infection.** SMs and RMs were housed at the Yerkes National Primate Research Center, Atlanta, Georgia, United States, and maintained in accordance with federal guidelines [40]. Prior to the study, the absence of SIV infection was confirmed by negative SIV PCR of plasma and negative HIV-2 serology for at least 1 y. Three RMs and three SMs were experimentally infected IV with a diverse inoculum of uncloned SIVsm from a naturally infected SM (individual FQj). SMs FLn, FCo, and FGU are referred to as SM1, SM2, and SM3, respectively. RMs RHt4, RQl4, and RZw4 are referred to as RM1, RM2, and RM3, respectively. The animals were followed at multiple time points following the infection, and quantitative PCR was carried out to determine the viral dynamics of their acute SIV infection [24].

**PCR.** Viral RNA was extracted from freshly thawed plasma samples from the three SMs and three RMs in this study using the Qiagen Viral RNA Kit. SIV sequences were amplified from 5  $\mu$ L of template in a PCR reaction using the Qiagen One-Step RT-PCR Kit (Qiagen, Valencia, California, United States).

To amplify the *env* V1V2 region, a mixture of two forward primers was used, FENV1 (5'-CTTGGGAGAATACAGTCACAG-3') corresponding to bp 6,780–6,800 of the SIVsmmH4 genome, and FENV2 (5'-CTTGGGAGAATACAGTAACAG-3') also corresponding to bp 6,780–6,800 but containing one different base at position 6,796. The reverse *env* V1V2 primer was also a mixture of RENV1 (5'-TAAATCTAATAGCATCCCAATAAT-3') and RENV2 (5'-TAAATCTAATAGCATCCCAATAGT-3') corresponding to bp 7,221–7,244 of the SIVsmmH4 genome, and differing at bp 7,222. The primer pair amplified a 456-bp fragments spanning the V1V2 hypervariable region of *env*.

The *gag* region was amplified using shortgagF1 (5'-TTAAGTC-CAAGAACATTAAATGC-3') and shortgagR (5'-GTAGAACCTGTC-TACATAGCT-3'), which correspond to bp 1,493–1,515 and 1,937–1,957 of SIVsmmH4, respectively, yielding a 421-bp product of the 5' end of the p27 capsid protein.

Conditions for each reaction were 30 min at 50 °C and 15 min at 95 °C, followed by 40 cycles of 94 °C for 1 min, 52 °C for 30 s, and 72 °C for 1 min. A final extension time was carried out for 5 min at 72 °C. Due to extremely low viral load, RNA from RM2 could not be amplified after day 14 for either V1V2 or *gag*. RT-PCR sensitivity was determined to be less than 500 copies per reaction. Viral loads from each of the animals did not significantly differ at each time point (with the exception of animal RM3, in which virus was undetectable using the RT-PCR protocol after day 14). Samples were not standardized for input copy number, potentially confounding the extent of change in viral diversity that was measured over time, although this would not confound comparisons between animals at each time point since viral loads were similar.

No-template controls and negative controls from the RNA extraction were used in each set of reactions to ensure that no cross contamination occurred at either step. In addition, samples from each pair of animals, SM1/RM1, SM2/RM2, SM3/RM3 were extracted at least 3 mo apart. This ensured that contamination within species was avoided. Contamination of negative-extraction controls was detected when extracting SM2 and RM2 samples from days 70 and 100. This extraction was repeated, and virus could not be amplified from RM2 due to very low copy numbers. On one occasion, the RT-PCR reaction was contaminated with a particular molecular clone, however these sequences were easy to identify with phylogenetic analysis due to their extensive divergence from the SI. These contaminants were excluded from the analysis. RNA extracted from day 10 plasma in SM1 and RM1 was RT-PCR amplified under the same conditions as above, except that 10  $\mu$ L of PCR product was removed at 25, 30, 35, and 40 cycles for cloning and sequencing to ensure that PCR bias during extended cycling was not a factor in sample diversity. Viral RNA from days 70 and 100 for RM1 and RM3 was extracted, PCR amplified, and cloned in duplicate to ensure experimental repeatability. A 1:10 dilution of SI was amplified under the same conditions, and 15 clones from this RT-PCR product were sequenced to ensure that input copy number did not bias diversity.

**DNA cloning and sequencing.** PCR products from each sample were run on a 1.5% low-melt agarose gel, and the 456-bp V1V2 or 421-bp *gag* product was extracted and cloned into the pCR4-TOPO vector (TOPO TA Cloning Kit, Invitrogen, Carlsbad, California, United States). Between 15 and 30 V1V2 clones and 5 and 10 *gag* p27 clones from each time point and each animal were randomly selected and sequenced using the M13F and M13R primers with the dye terminator cycle sequencing method.

**Sequence and phylogenetic analyses.** Sequences were aligned using the program CLUSTAL X [41], followed by manual adjustment using MacClade 4.0 [42]. Nonaligned regions of length variation in V1 and V2 were removed (corresponding to nucleotides 6,932–6,974), and sequences containing internal stop codons, single deletions, or double deletions were also excluded from analysis, as these are thought to be PCR artifacts [43]. Figures S1 and S2 show the resulting alignment of all sequences in V1V2 and *gag*, respectively.

For the SI, maximum parsimony and NJ were implemented using the PAUP 4.0b10\* package for V1V2 and *gag* [44]. For each of the resulting trees, bootstrap support was determined with 1,000 resamplings of the sequences. The most highly supported clade in both the NJ and the parsimony trees was used as the outgroup for all subsequent phylogenetic trees (Figure S3).

For tree construction, the Modeltest program [45] was used to construct and evaluate the DNA substitution models used. Based on the Modeltest results, phylogenetic analysis on sequences obtained from successive time points during the acute infection was performed by ML using the program Treefinder [46]. The general time reversible (GTR) model, which allows for rate variation between sites [47–49], was used, and the shape parameter ( $\alpha$ ) of the gamma distribution used in this model was estimated, as were base frequencies and substitution rate parameters. Bootstrap support was determined with 1,000 resamplings of the ML tree using distance methods in PAUP4.0b10\*, incorporating the estimated rate parameters.

The cumulative number of synonymous and nonsynonymous and nucleotide substitutions was estimated using Synonymous/Nonsynonymous Analysis (SNAP; <http://hiv-web.lanl.gov/>), which calculates rates of nucleotide substitution from a set of codon-aligned nucleotide sequences, based on the method of Nei and Gojobori [50], and incorporating a statistic developed in Ota and Nei [51]. Viral nucleotide diversity at each time point was determined by calculating the pairwise nucleic acid distances for each of the clones using the method of Tamura and Nei [52] in the program MEGA 2.1 [53]. This same method was also employed to quantify nucleotide divergence from the source, defined as the ratio of the difference in nucleotide diversity between SI and each sample of variants to the total diversity in the two groups. Amino acid diversity was calculated using the gamma distance method in the program MEGA 2.1. Phylogenetic trees constructed with synonymous or nonsynonymous sites only were constructed in MEGA 2.1 using distance methods incorporating the Tamura-Nei model of nucleotide substitution with gamma-distributed rates. All statistics were computed using SYSTAT 10 software (SPSS, Chicago, Illinois, United States).

## Supporting Information

**Figure S1.** Phylogenetic Analysis of Source Inoculum V1V2 Variants with Molecular Clones

(A) NJ tree showing the most highly supported clade of SI used as the outgroup for all subsequent phylogenetic trees. (B) An unrooted ML tree of 30 SI V1V2 variants and corresponding V1V2 sequences from clones SIVmac239, SIVsmmH4, and SIVmne (obtained from the HIV sequence database [<http://hiv-web.lanl.gov/content/index>]) was constructed with Treefinder [46] using a GTR model and estimated gamma rate distribution, base frequencies, and substitution rates. Bootstrap values greater than 50% are shown at nodes.

Found at DOI: 10.1371/journal.ppat.0010003.sg001 (53 KB PDF).

**Figure S2.** Viral Replication Dynamics following Infection with a Diverse SIVsmm

Three SMs and three RMs were inoculated with plasma obtained from a naturally infected SM. Viral replication was monitored in SMs and RMs by quantitative RT-PCR of plasma RNA (see Materials and Methods).

Found at DOI: 10.1371/journal.ppat.0010003.sg002 (210 KB PDF).

**Figure S3.** Amino Acid Sequences of All Animals and All Time Points in the V1V2 Region of Envelope

Region corresponds to nucleotides 6,801–7,220 of SIVsmmH4. Sequences were aligned using the program CLUSTAL X [41], followed by manual adjustment using MacClade 4.0 [42]. A non-aligned region of length variation in V1 was removed, corresponding to amino acids 129–137 of SIVsmmH4 *env*, and is indicated by “~”. The consensus of all sequences in this study is shown above all sample sets, with codon positions labeled above. A dot indicates amino acid identity with the consensus sequence, and any amino acid changes are indicated with the appropriate symbol. The V1V2 regions are

highlighted in blue on the consensus sequence, and glycosylation consensus motifs present in each sequence are highlighted in yellow.

Found at DOI: 10.1371/journal.ppat.0010003.sg003 (226 KB DOC).

**Figure S4.** Mean Number of Glycosylation Consensus Motifs in SMs and RMs for All Time Points

Frequency of glycosylation consensus motifs is lower in RMs (regression analysis,  $p < 0.001$ ) and increases over time in both SMs and RMs. The number of motifs in the SI is indicated with a star on the y-axis.

Found at DOI: 10.1371/journal.ppat.0010003.sg004 (114 KB PDF).

**Figure S5.** Phylogenetic Analysis of Natural Host VIV2 Variants at Days 10–100 Shows No Specific Pattern

ML phylogenetic tree of sequences obtained from SM1 at days 10 (pink) to 100 (red) is shown, constructed with Treefinder [46] using a GTR model and estimated gamma rate distribution, base frequencies, and substitution rates. Bootstrap values greater than 50% are shown at nodes, and the number of multiple clones from the same animal at the ends of branches is indicated within the symbol. The SI variants are represented by triangles and identified by the label within.

Found at DOI: 10.1371/journal.ppat.0010003.sg005 (77 KB PDF).

**Figure S6.** Amino Acid Sequences of All Animals and All Time Points in the p27 Region of *gag*

Region corresponds to nucleotides 1,516–1,936 of SIVsmmH4. Sequences were aligned using the program CLUSTAL X [41], followed by manual adjustment using MacClade 4.0 [42]. The top sequence in each set corresponds to the majority consensus sequence

from all sequences at all time points, with codon positions labeled above. A dot indicates amino acid identity with the consensus sequence, and any amino acid changes are indicated with the appropriate symbol.

Found at DOI: 10.1371/journal.ppat.0010003.sg006 (75 KB DOC).

**Accession Number**

The GenBank (<http://www.ncbi.nlm.nih.gov/>) accession number of the SIVsmmH4 genome is X14307.

**Acknowledgments**

We dedicate this paper to the memory of Dr. H. McClure, for his selfless devotion to advancing AIDS research in the nonhuman primate models, and for his genuine and warm collegiality. We thank Drs. F. Novembre and S. Garg for technical assistance, and Drs. E. Hunter and C. Derdeyn for valuable comments. This work was supported by National Institutes of Health grants AI4915502 and AI4476301 to MBF, RR00165 to the Yerkes National Primate Research Center, and National Institute on Allergy and Infectious Diseases Statistical Training on AIDS Grant T32-AI07442.

**Competing interests.** The authors have declared that no competing interests exist.

**Author contributions.** LJD, MBF, and SIS conceived and designed the experiments. LJD performed the experiments. LJD and THV analyzed the data. JML contributed reagents/materials/analysis tools. LJD, THV, and SIS wrote the paper. ■

**References**

- Salemi M, De Oliveira T, Courgnaud V, Moulton V, Holland B, et al. (2003) Mosaic genomes of the six major primate lentivirus lineages revealed by phylogenetic analyses. *J Virol* 77: 7202–7213.
- Hahn BH, Shaw GM, Cock KMD, Sharp PM (2000) AIDS as a zoonosis: Scientific and public health implications. *Science* 287: 607–614.
- Silvestri G, Sodora DL, Koup RA, Paiardini M, O'Neil SP, et al. (2003) Non-pathogenic simian immunodeficiency virus infection of sooty mangabey monkeys is characterized by limited bystander immunopathology despite chronic high-level viremia. *Immunity* 18: 441–452.
- Rey-Cuille M-A, Berthier J-L, Bomsel-Demontoy M-C, Chaduc Y, Montagnier L, et al. (1998) Simian immunodeficiency virus replicates to high levels in sooty mangabey without inducing disease. *J Virol* 72: 3872–3886.
- Goldstein S, Ourmanov I, Brown CR, Beer BE, Elkins WR, et al. (2000) Wide range of viral load in healthy African green monkeys naturally infected with simian immunodeficiency virus. *J Virol* 74: 11744–11753.
- Broussard SR, Staprans SI, White R, Whitehead EM, Feinberg MB, et al. (2001) Simian immunodeficiency virus replicates to high levels in naturally infected African green monkeys without inducing immunologic or neurologic disease. *J Virol* 75: 2262–2275.
- Korber BTM, Muldoon M, Theiler J, Gao F, Gupta R, et al. (2000) Timing the ancestor of the HIV-1 pandemic strains. *Science* 288: 1789–1796.
- Puffer BA, Altamura LA, Pierson TC, Doms RW (2004) Determinants within gp120 and gp41 contribute to CD4 independence of SIV Envs. *Virology* 327: 16–25.
- Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, et al. (2004) Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* 14: 1229–1246.
- Reiter JN, Means RE, Desrosiers RC (1998) A role for carbohydrates in immune evasion in AIDS. *Nat Med* 4: 679–684.
- Richman DD, Wrin T, Little SJ, Petropoulos CJ (2003) Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* 100: 4144–4149.
- Wei X, Decker JM, Wang S, Hui H, Kappes JC, et al. (2003) Antibody neutralization and escape by HIV-1. *Nature* 422: 307–311.
- van't Wout AB, Kostrina NA, Mulder-Kampinga GA, Albrecht-van Lent N, Scherpbier HJ, et al. (1994) Macrophage tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. *J Clin Invest* 94: 2060–2067.
- Zhu T, Wang N, Carr A, Nam DS, Moor-Jankowski R, et al. (1996) Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: Evidence for viral compartmentalization and selection during sexual transmission. *J Virol* 70: 3098–3107.
- Zhang LQ, MacKenzie P, Cleland A, Holmes EC, Brown AJ, et al. (1993) Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. *J Virol* 67: 3345–3356.
- Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, et al. (2004) Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* 303: 2019–2022.
- Chohan B, Lang D, Sagar M, Korber B, Lavreys L, et al. (2005) Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. *J Virol* 79: 6528–6531.
- Cohen J (2004) Virology. HIV may shed some protection as it jumps to new hosts. *Science* 303: 1956.
- Sagar M, Kirkegaard E, Long EM, Celum C, Buchbinder S, et al. (2004) Human immunodeficiency virus type 1 (HIV-1) diversity at time of infection is not restricted to certain risk groups or specific HIV-1 subtypes. *J Virol* 78: 7279–7283.
- Beer BE, Bailes E, Gocken R, Dapolito G, Coulbaly C, et al. (1999) Simian immunodeficiency virus (SIV) from sun-tailed monkeys (*Ceropithecus solatus*): Evidence for host-dependent evolution of SIV within the *C. lhoesti* superspecies. *J Virol* 73: 7734–7744.
- Johnson PR, Fomsgaard A, Allan JS, Gravel M, London WT, et al. (1990) Simian immunodeficiency viruses from African green monkeys display unusual genetic diversity. *J Virol* 64: 1086–1092.
- Peeters M, Courgnaud V, Abela B, Auzel P, Pourrut X, et al. (2002) Risk to human health from a plethora of simian immunodeficiency viruses in primate bushmeat. *Emerg Infect Dis* 8: 451–457.
- Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, et al. (1999) Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* 397: 436–441.
- Silvestri G, Fedanov A, Germon S, Kozyn N, Kaiser W, et al. (2004) Divergent host responses during primary SIVsmm infection of natural mangabey and non-natural rhesus macaque hosts. *J Virol* 79: 4043–4054.
- Puffer BA, Pohlmann S, Edinger AL, Carlin D, Sanchez MD, et al. (2002) CD4 independence of simian immunodeficiency virus envs is associated with macrophage tropism, neutralization sensitivity, and attenuated pathogenicity. *J Virol* 76: 2595–2605.
- Rudensky LM, Kimata JT, Long EM, Chackerian B, Overbaugh J (1998) Changes in the extracellular envelope glycoprotein of variants that evolve during the course of simian immunodeficiency virus SIVmac infection affect neutralizing antibody recognition, syncytium formation, and macrophage tropism but not replication, cytopathicity, or CCR-5 coreceptor recognition. *J Virol* 72: 209–217.
- Burns DPW, Desrosiers RC (1994) Envelope sequence variation, neutralizing antibodies, and primate lentivirus persistence. *Curr Top Microbiol* 188: 185–219.
- Petry H, Pekrun K, Hunsmann G, Jurkiewicz E, Luke W (2000) Naturally occurring V1-env region variants mediate simian immunodeficiency virus SIVmac escape from high-titer neutralizing antibodies induced by a protective subunit vaccine. *J Virol* 74: 11145–11152.
- Jurkiewicz E, Hunsmann G, Schnaffner J, Nisslein T, Luke W, et al. (1997) Identification of the V1 region as a linear neutralizing epitope of the simian immunodeficiency virus SIVmac envelope glycoprotein. *J Virol* 71: 9475–9481.
- Chackerian B, Rudensky LM, Overbaugh J (1997) Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian

- immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies. *J Virol* 71: 7719–7727.
31. Fomsgaard A, Hirsch VM, Johnson PR (1992) Cloning and sequences of primate CD4 molecules: Diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus. *Eur J Immunol* 22: 2973–2981.
  32. Kunstman KJ, Puffer B, Korber BT, Kuiken C, Smith UR, et al. (2003) Structure and function of CC-chemokine receptor 5 homologues derived from representative primate species and subspecies of the taxonomic suborders Prosimii and Anthropoidea. *J Virol* 77: 12310–12318.
  33. Staprans SI, Barry AP, Silvestri G, Safrit JT, Kozyr N, et al. (2004) Enhanced SIV replication and accelerated progression to AIDS in macaques primed to mount a CD4 T cell response to the SIV envelope protein. *Proc Natl Acad Sci U S A* 101: 13026–13031.
  34. Garber DA, Silvestri G, Barry AP, Fedanov A, Kozyr N, et al. (2003) Blockade of T cell costimulation reveals interrelated actions of CD4+ and CD8+ T cells in control of SIV replication. *J Clin Invest* 113: 836–845.
  35. Ebert LM, McColl SR (2002) Up-regulation of CCR5 and CCR6 on distinct subpopulations of antigen-activated CD4+ T lymphocytes. *J Immunol* 168: 65–72.
  36. Paillard F, Sterckers G, Vaquero C (1990) Transcriptional and post-transcriptional regulation of TcR, CD4 and CD8 gene expression during activation of normal human T lymphocytes. *EMBO J* 9: 1867–1872.
  37. Siliciano JD, Siliciano RF (2004) A long-term latent reservoir for HIV-1: Discovery and clinical implications. *J Antimicrob Chemother* 54: 6–9.
  38. Albrecht D, Zollner B, Feucht HH, Lorenzen T, Laufs R, et al. (2002) Reappearance of HIV multidrug-resistance in plasma and circulating lymphocytes after reintroduction of antiretroviral therapy. *J Clin Virol* 24: 93–98.
  39. Ruiz-Jarabo CM, Arias A, Baranowski E, Escarmis C, Domingo E (2000) Memory in viral quasispecies. *J Virol* 74: 3543–3547.
  40. National Institutes of Health (1985) Guide for the care and use of laboratory animals, rev. ed. Department of Health and Human Services publication 85–23. Bethesda, Maryland: National Institutes of Health.
  41. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res* 22: 4673–4680.
  42. Maddison WP, Maddison DR (1989) Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatol* 53: 190–202.
  43. McAllister J, Casino C, Davidson F, Power J, Lawlor E, et al. (1998) Long-term evolution of the hypervariable region of hepatitis C in a common-source-infected cohort. *J Virol* 72: 4893–4905.
  44. Swofford DL (2002) PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). 4.0 ed. Sunderland, Massachusetts: Sinauer, Associates.
  45. Posada D, Crandall KA (1998) Modeltest: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
  46. Jobb G (2002) Treefinder [computer program]. Available at: <http://www.treefinder.de>. Accessed 01 January 2003.
  47. Gu X, Li W-H (1996) A general additive distance with time-reversibility and rate variation among nucleotide sites. *Proc Natl Acad Sci U S A* 93: 4671–4676.
  48. Gu X, Li W-H (1998) Estimation of evolutionary distances under stationary and nonstationary models of nucleotide substitution. *Proc Natl Acad Sci U S A* 95: 5899–5905.
  49. Yang Z (1994) Estimating the pattern of nucleotide substitution. *J Mol Evol* 39: 105–111.
  50. Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418–426.
  51. Ota T, Nei M (1994) Variance and covariances of the numbers of synonymous and nonsynonymous substitutions per site. *Mol Biol Evol* 11: 613–619.
  52. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512–526.
  53. Kumar S, Tamura K, Jakobsen I, Nei M (2001) MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244–1245.