

Article

The mucosal barrier and anti-viral immune responses can eliminate portions of the viral population during transmission and early viral growth

Athena E. Golfinos ¹, Dane D. Gellerup ², Hannah Schweigert ¹, Jaffna Mathiapparanam ¹, Alexis J. Balgeman ¹, Andrea M. Weiler ^{2,3}, Thomas C. Friedrich ^{2,3}, Brandon F. Keele ⁴, Miles P. Davenport ⁵, Vanessa Venturi ⁵ and Shelby L. O'Connor ^{1,2}

¹ Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, Wisconsin, USA; slfeinberg@wisc.edu

² Wisconsin National Primate Research Center, Madison, Wisconsin, USA; thomasf@primate.wisc.edu

³ Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA; thomasf@primate.wisc.edu

⁴ AIDS and Cancer Virus Program, Frederick National Laboratory, Frederick, MD 21769; keelebf@mail.nih.gov

⁵ Infection Analytics Program, Kirby Institute for Infection and Immunity, UNSW Sydney, Sydney, NSW 2052, Australia; mdavenport@kirby.unsw.edu.au

* Correspondence: slfeinberg@wisc.edu; Tel.: +1 (608) 890-0843

Abstract: Little is known about how individual virus lineages replicating during acute Human Immunodeficiency Virus or Simian Immunodeficiency Virus (HIV/SIV) infection persist into chronic infection. In this study, we use molecularly barcoded SIV (SIVmac239M) to track distinct viral lineages for 12 weeks after intravenous and intrarectal challenge in macaques. Two *Mafa-A1*063+* cynomolgus macaques (*Macaca fascicularis*) were challenged intravenously (IV), and two *Mamu-A1*001+* rhesus macaques (*Macaca mulatta*) were challenged intrarectally (IR) with 200,000 Infectious Units (IU) of SIVmac239M. We deep sequenced the molecular barcode from all animals over 12 weeks to characterize the diversity and persistence of virus lineages, as well as the sequences of T cell epitopes during acute SIV infection. During the first three weeks post-infection, we found ~175-950 times more unique virus lineages circulating in the animals challenged intravenously than those challenged intrarectally, suggesting that challenge route is the primary driver restricting the transmission of individual viral lineages. Additionally, the emergence of escape variants can occur on multiple virus templates simultaneously, but elimination of some templates is likely a consequence of additional host factors. These data imply that virus lineages present during acute infection can be eliminated from the virus population even after initial T cell selection.

Keywords: Simian Immunodeficiency Virus; SIVmac239; barcoded viruses; intravenous challenge; intrarectal challenge

1. Introduction

While many studies have thoroughly examined the number and diversity of unique viral lineages during acute HIV/SIV infection, few studies focus on the persistence of specific individual virus lineages within a host over time. Defining the number and diversity of these lineages may shed light on the natural progression of how an original virus lineage persists and evolves in the absence of interventions and treatments. In HIV, intravenous (IV) infection is associated with the transmission of numerous unique systemic virus lineages. Researchers have modeled this in SIV studies, where the number of uniquely replicating viruses in plasma correlated with the infectious inoculum dose [1–3].

The molecular tools to characterize virus transmission and evolution have improved over the last ten years. One widely used tool to study the transmission and evolution of

HIV and SIV is Single Genome Amplification (SGA) (Reviewed in [4]). However, SGA is expensive and time-intensive, making large-scale studies using SGA quite challenging [5, 6]. Instead, viruses containing small molecular tags within their genomes have been developed for animal studies of infectious pathogens. These viruses can be enumerated more efficiently in the blood and tissues with deep sequencing [1, 3, 7–12]. In contrast to the small number of unique virus templates that can be sequenced by SGA, thousands of barcodes can be deep sequenced from a large population of virus templates to improve the detection of low-proportion virus variants. These “pseudo-swarm” viruses are isogenic outside the molecular barcode but highly diverse within the barcode, providing each viral lineage with similar fitness *in vitro* and *in vivo* [1, 3].

SIVmac239M is a barcoded virus stock derived from clonal SIVmac239 (Nef-open) containing nearly 10,000 viral variants differing only by the 34bp insert [3]. Previous studies have used SIVmac239M to understand viral reservoirs in SIV infection [2, 3], measure reactivation rates following antiretroviral treatment (ART) interruption [3], enumerate the number of viral lineages during acute infection after intravenous challenge [1, 3], and assess acute immune escape in multiple SIV viral lineages [8]. However, these studies challenged animals intravenously with SIVmac239M rather than mucosally, even though mucosal transmission is the primary route of HIV transmission. SIVmac239X, a predecessor to SIVmac239M, was previously used to determine the number of transmitted/founder lineages following intrarectal (IR) infection [2] and assess viral dynamics following intravaginal infection [13], but this virus stock contains only ten distinct lineages, precluding its usefulness in evaluating lineage-specific viral dynamics.

In this study, we compared the number of unique viral lineages circulating systemically during acute infection of macaques challenged either mucosally via IR challenge or IV with 200,000IU of SIVmac239M. We also examined the relationship between early immune escape and the number of persistent barcoded viral lineages. We followed animals for 12 weeks to determine whether the same virus lineages persisted in the blood throughout this time, as well as the associated variants in acutely targeted T cell epitopes.

2. Materials and Methods

2.1. Research animals

Two Mauritian cynomolgus macaques (MCMs; *Macaca fascicularis*) and two rhesus macaques (RMs; *Macaca mulatta*) were housed and cared for at the Wisconsin National Primate Research Center (WNPRC) in Madison, WI. All animals were cared for under IACUC protocol #G00680, approved by the University of Wisconsin Graduate School Institutional Animal Care and Use Committee. All procedures, including blood draws and SIVmac239M inoculation, were performed under anesthesia, and we made every effort to minimize distress and suffering. According to previous methods, the RMs were positively genotyped for MHC class I allele *Mamu-A1*001* [14], and the MCMs were positively genotyped for *Mafa-A1*063* [15]. Both RMs (r10001 and r04103) were challenged intrarectally under anesthesia with 200,000 Infection Units (IU) (or 4.89 billion viral copies (vc)) of SIVmac239M, and the MCMs (cy0575 and cy0428) were challenged IV with 200,000 IU of SIVmac239M.

2.2. Viral load quantification

Viral loads (VLs) were measured by the Virology Services unit of the WNPRC longitudinally throughout SIVmac239M infection. VLs were reported from the WNPRC as log viral copies (ceq) per milliliter of plasma. Viral RNA was isolated from plasma samples using the Maxwell 16 Viral Total Nucleic Acid Purification Kit on the Maxwell 16 MDx instrument (Promega, Madison WI). Viral RNA was then quantified using a highly sensitive QRT-PCR assay based on the one developed by Cline et al. [16]. RNA was reverse transcribed and amplified using the Superscript III Platinum One-Step qRT-PCR kit (Invitrogen) on the LightCycler 480 or LC96 instrument (Roche, Indianapolis, IN) and quantified by interpolation onto a standard curve made up of serial tenfold dilutions of *in vitro*

transcribed RNA. RNA for this standard curve was transcribed from the p239gag_Lifson plasmid kindly provided by Dr. Jeffrey Lifson, NCI/Leidos. The final reaction mixtures contained 5 mM MgSO₄, 150 ng random primers (Promega, Madison, WI), 600 nM each primer, 100 nM probe, and 0.8 µl enzyme mix. Primer and probe sequences are as follows: forward primer: 5'- GTCTGCGTCATCTGGTGATTC, reverse primer: 5'-CAC-TAGCTGTCTCTGCACTATGTGTTTTG-3' and probe: 5'-6-carboxyfluorescein-CTTCCTCAGTGTGTTTCACTTTCTCTTCTGCG-BHQ1-3'. The reactions cycled with the following conditions: 37°C for 15 min, 50°C for 30 min, 95°C for 2 min followed by 50 cycles of 95°C for 15 seconds and 62°C for 1 min. The limit of detection of this assay is 100 copies/ml.

2.3. Isolation of nucleic acids for sequencing

vRNA was isolated from either 500µL or 1ml of plasma using the QIAamp MinElute Virus Spin Kit (Qiagen) according to the manufacturer's protocol. We eluted the RNA from the column in Buffer AVE.

2.4. Deep sequencing of the barcode from plasma viral RNA

5µL of vRNA was reverse-transcribed using SuperScript III One-Step RT-PCR System with forward primer 5'-CAAGAAAGGCTGTAGATGTC-3' and reverse primer 5'-GACAAATAATCAACACTTGGCA-3'. Viral loads ranged from 2.1e3-1.51e8 ceq/ml. This step produced a 2,250 base pair cDNA amplicon that included the 34-nucleotide barcode. Then, 1-2ng of the purified 2,250 bp amplicon was used as a template to generate a smaller amplicon (range of 416-499 base pairs) with PCR primers containing Common Sequence (CS) tags derived from the Access Array Barcode Library for Illumina Sequences (Fluidigm Corporation) (Table S1). We generated this amplicon with Phusion High-Fidelity PCR Master Mix with High-Fidelity (HF) Buffer. PCR conditions were as follows: 98°C for 3 minutes (x1); 98°C for 5 seconds, 60°C for 10 seconds, 72°C for 30 seconds (x29); 72°C for 5 minutes (x1). PCR products were quantified using a Qubit Fluorometer (Invitrogen) and normalized to 1 ng/µL. Each PCR product was purified using the MinElute Gel Extraction Kit (Qiagen) and quantified using the Qubit double-stranded DNA (dsDNA) High Sensitivity Assay Kit (Invitrogen). Samples were pooled and sequenced with a 2 x 300 V3 sequencing kit on an Illumina MiSeq.

2.5. Deep sequencing of the complete SIV genome from plasma

vRNA was isolated from plasma using the MinElute virus spin kit (Qiagen) as described above. Five microliters of vRNA were reverse transcribed and amplified using Superscript III one-step RT-PCR system with High-Fidelity Platinum Taq (Invitrogen). This step produced four overlapping amplicons spanning the entire SIV coding sequence, detailed previously [17, 18]. PCR products were purified using the MinElute gel extraction kit (Qiagen), then quantified using the Quant-IT double-stranded DNA high-sensitivity (HS) assay kit on a Qubit Fluorometer (Invitrogen). Libraries were generated from 1 ng of pooled amplicons and then tagged using the Nextera XT kit (Illumina). We quantified tagged libraries with the Quant-IT dsDNA HS assay kit, and we assessed the average fragment size with an Agilent Bioanalyzer. Libraries were pooled and sequenced on an Illumina MiSeq using either 2 x 250 or 2 x 300 sequencing kits.

2.6. ELISPOT assays

Fresh peripheral blood mononuclear cells (PBMCs) were subjected to an IFN-γ enzyme-linked immunosorbent spot (ELISPOT) assay, as previously described [19]. We briefly blocked a precoated monkey IFN-γ ELISPOTplus plate (Mabtech, Mariemont, OH), and added peptides listed in the figures to each well in duplicate at a final concentration of 1µM. We added 10⁵ PBMCs to each well, and plates were incubated overnight at 37°C. The positive control for this assay was 10µM Concanavalin A, and the plates were developed according to the manufacturer's protocol (Sigma Aldrich). The wells were

imaged with an AID ELISPOT reader (Autoimmun Diagnostika GmbH). We calculated the number of spot forming cells (SFCs) per 10^6 PBMCs by subtracting the average number of background spots (average of spots from four wells not stimulated with any peptide, which served as a negative control) and then multiplying this value by 10. We considered a response positive if it exceeded the no stimulation average threshold plus two times the standard deviation of the no stimulation wells or if it had 50 SFCs per 10^6 PBMCs, whichever was greater. The values for duplicate wells were then averaged and graphed.

2.7. Quality control and analysis of molecular barcode sequences

Paired-end reads from each time point were mapped to SIVmac239M, and we used the read containing the barcode during subsequent steps. We did not use the other read for barcode analysis. The reads containing the barcode were trimmed to Q>15 [20] and analyzed using the Barcode Virus Analysis tool [21], which analyzes barcoded viral lineage populations originating from SIVmac239M data. These methods are further described in Fennessey et al. [3]. For each sample, we calculated the number and frequency of individual viral lineages. We determined the authenticity of viral lineages based on extensive testing of the SIVmac239M barcode population in the stock and in macaques [1,3]. Samples were excluded if the FASTQ file contained <20,000 reads due to inadequate depth of coverage. Replicates were compared to one another to screen for adequate similarity using Morisita-Horn Similarity Index (MHSI), which accounts for overlap in barcodes and their corresponding proportions. Here, >75% similarity (or MHSI >0.75) was required to determine that replicates were similar enough to be analyzed further. An additional table details Morisita-Horn values for all replicate pairs (Table S2). We performed all computation and statistical analyses using in-house Python code [22]. Individual barcodes were only included in the analysis if they were found in both replicates from a given time point to prevent erroneous data from being enumerated and analyzed. Unique barcodes were determined based on the number of unique lineages present in both samples, with barcodes present in only a single sample eliminated. The number of barcodes eliminated by this method consisted of 0-9% per sample, and therefore comprised a limited proportion of the barcode population.

3. Results

3.1. Fewer circulating viral lineages are detectable following mucosal challenge with SIVmac239M compared to intravenous challenge

Two Mafa-A1*063+ MCMs were challenged IV, and two Mamu-A1*001+ RMs were challenged IR with 200,000 IUs of SIVmac239M. We saved plasma and PBMC from each animal according to the study timeline (Figure 1A). Peak plasma viral loads (VL) were similar between groups (Figure 1B) and are comparable to peak VLs observed in MCMs and RMs with similar MHC genetics in previous studies [23–27]. This observation suggests that regardless of inoculation route, SIVmac239M exhibits similar peak viremia in both groups.

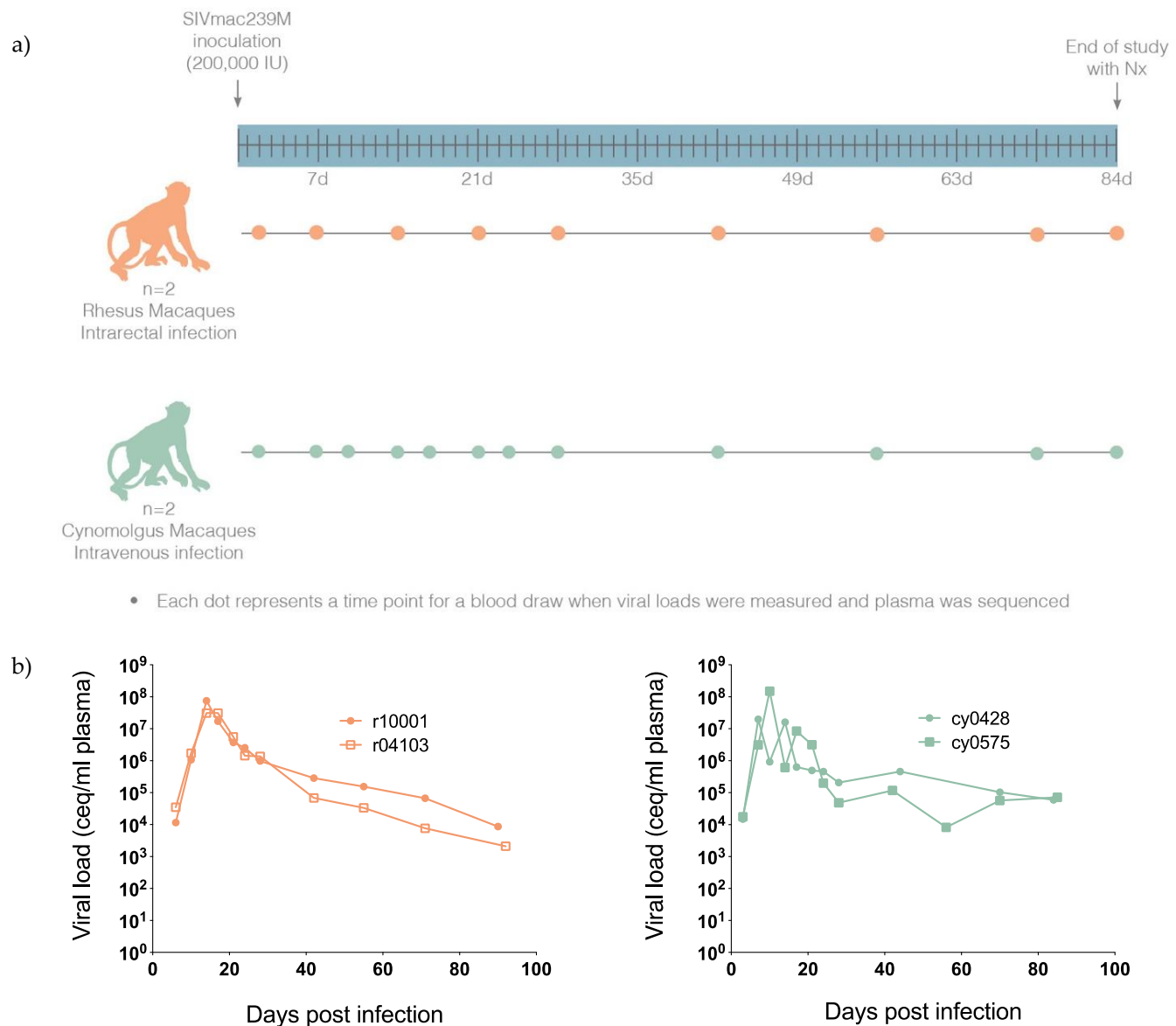


Figure 1. Study timeline and viral loads. (A) Two Mamu-A1*001+ rhesus macaques were inoculated intrarectally (IR), and two Mafa-A1*063+ cynomolgus macaques were inoculated intravenously (IV), with 200,000 Infectious Units (IU) of SIVmac239M. Whole blood was drawn at the time points indicated. (B) Viral loads in both groups of animals were calculated by qRT-PCR from plasma.

We assessed the number of unique viral lineages following infection in both groups. We detected 2 unique barcoded lineages in r10001 from all 8 time points sampled and between 2 and 11 unique viral lineages in r04103 depending on the sampling time point (Figure 2A). All barcodes were present at a frequency of 0.1% or greater in the two IR-challenged RMs. The number of unique viral lineages remained consistent in r10001 throughout the infection, while r04103 had a higher number of unique circulating viral lineages (range 2-11), with a subtle decrease in the number of unique lineages present as infection progressed.

In contrast, the number of unique viral lineages detected in the IV-challenged MCMs was approximately 150-fold higher during the first three weeks of infection than in the IR-challenged RMs. Until day 17, the median number of unique viral lineages across all three time points was 2049 and 1845 for cy0428 and cy0575, respectively. These results were consistent with the number of barcodes previously detected in rhesus macaques challenged intravenously with the same dose of SIVmac239M [1]. The median number of unique viral lineages dropped to 38 and 11 for cy0428 and cy0575, respectively (Figure 2B)

from day 24 to necropsy on day 84. These data confirm the mucosal bottleneck theory of viral transmission since animals infected intrarectally with 200,000 IUs of SIVmac239M does not achieve the same diversity of circulating viral lineages during acute infection as animals challenged IV with the same dose of SIVmac239M. The average number of barcodes following IV infection is ~1900 barcodes, while the number of barcodes following IR infection is 2-11. This suggests that there are between 175-950-fold more detectable barcodes in IV infected animals than in the IR infected animals.

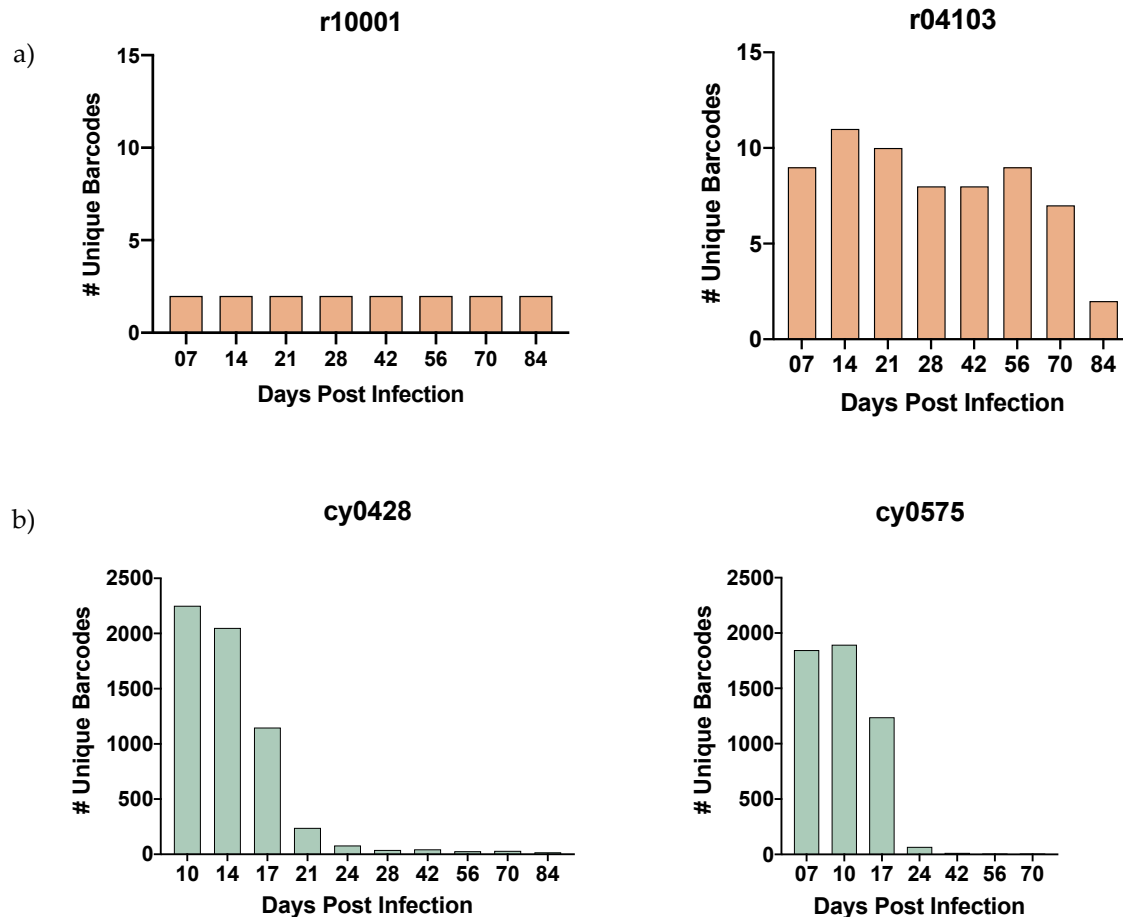


Figure 2. The number of unique viral lineages in IR and IV-challenged macaques. (A) The number of unique viral lineages was plotted in both rhesus macaques that were intrarectally challenged with SIVmac239M. (B) The number of unique viral lineages was plotted in both cynomolgus macaques that were intravenously challenged with SIVmac239M.

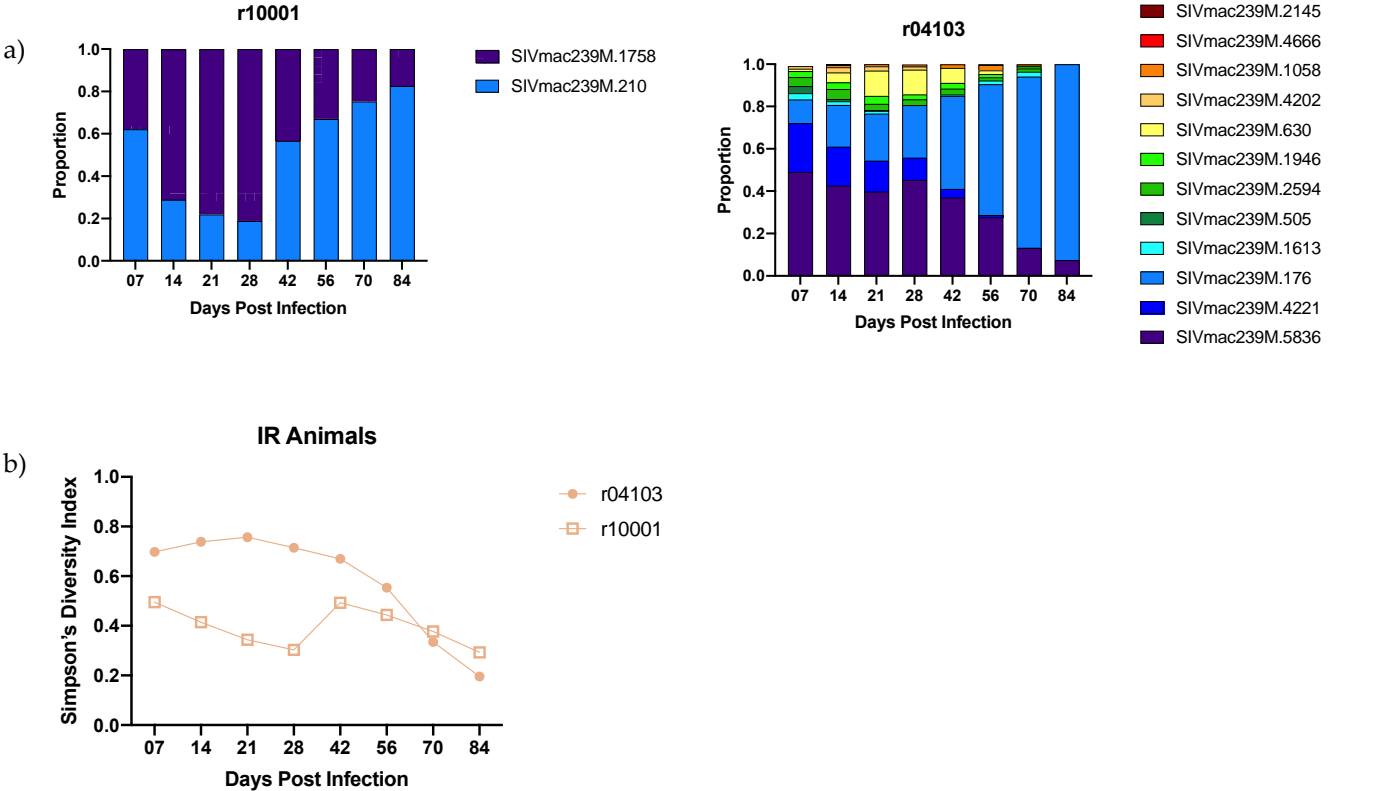
3.2. Simpson's Diversity Index of the molecular barcode declines rapidly at three weeks after infection in animals challenged intravenously but not intrarectally

We next plotted the frequency of individual barcoded virus lineages comprising the virus population in both IR and IV-challenged animals over time. In both IR-challenged RMs, the relative proportions of each barcode changed slowly over time (Figure 3A), such that one lineage was dominant (>75%) at euthanasia. We calculated Simpson's Diversity Index (SDI) over time as a proxy for the number and abundance of viral lineages within the total population. To calculate SDI, we used the number and frequency of unique barcodes as individual species comprising the virus population (26). We chose to calculate the SDI because it quantitatively assesses the evenness of individual virus lineages within the total virus population. A population with a high SDI (~1) will have an even distribution of many individual virus lineages in the population. In contrast, a population with a low SDI (~0) likely has a small number of high-frequency viral lineages dominating the population. In the two IR-challenged RMs, the small number of virus lineages comprising the total population was reflected by moderate diversity (SDI 0.5-0.7) during acute

infection that slowly declined throughout infection, until an SDI of 0.25 was reached at necropsy (Figure 3B).

We characterized the persistence of the individual virus lineages in the IV-challenged animals from three weeks post-infection to necropsy on day 84. Barcode 6128 (Figure 3C, left, purple) was present at 68% frequency in cy0428 at three weeks post-infection, but then nearly disappeared by week six and was replaced with barcode 2667 (Figure 3C, left, navy blue). Barcode 2667 remained prevalent until necropsy. Animal cy0575 exhibited more detectable lineages, with five barcodes present at >10% at three weeks post-infection, four of which persisted until necropsy (Figure 3C, right). This difference in population composition is mirrored in the SDI values of the virus populations replicating in both MCMs, with the population present in cy0575 consistently having a slightly higher SDI than cy0428 (Figure 3D). Beyond week 6, the SDI value remained between 0.4-0.7 and 0.1-0.25 in cy0575 and cy0428, respectively, until necropsy.

We next examined how the SDI of the virus population in the IV-challenged animals changed over time. The SDI of the virus population in these two MCMs from one to three weeks post-infection was approximately 1, as expected from the large number of unique virus lineages detected at these time points. The SDI rapidly declined to <0.5 between three and six weeks after infection, consistent with the decrease in the number of unique virus lineages detected during this time (Figure 3D). This decrease is consistent with other studies reporting viral escape from T cells during HIV or SIV infection [8, 28–30].



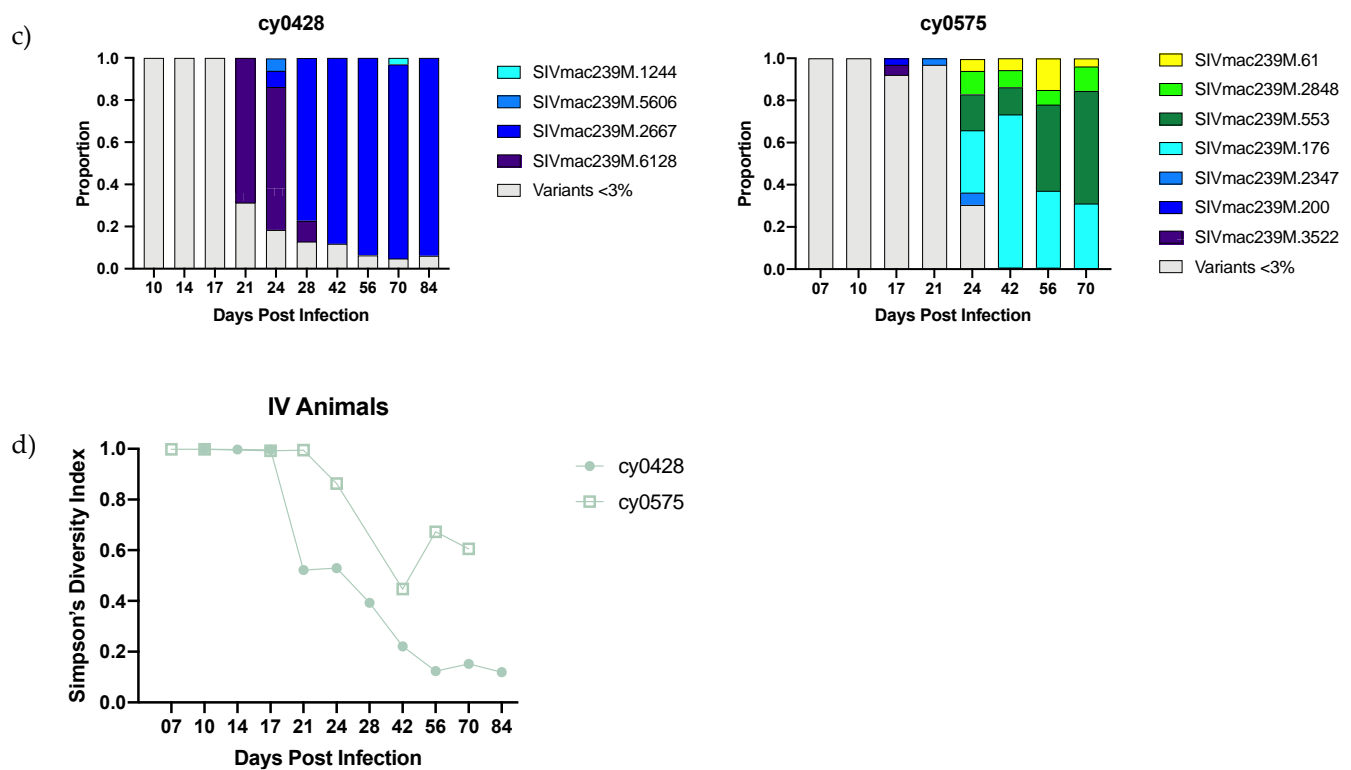


Figure 3. Diversity of the molecular barcodes present in the circulating virus populations. (A and C) The proportion of individual barcoded lineages circulating in the rhesus (A) and cynomolgus macaques (C) are shown over time. Variants that do not exceed 3% proportion at any point throughout infection were binned into the <3% category. (B and D) Longitudinal Simpson's Diversity Index of the population of barcoded lineages in the rhesus (B) and cynomolgus macaques (D) was calculated and plotted over time.

3.3. T cell responses in MHC-restricted epitopes are observed around weeks 3-4 post-infection in all animals

SIV-infected Mafa-A1*063+ MCMs rapidly develop CD8⁺ T cells targeting two peptides: Nef₁₀₃₋₁₁₁RM9 and Gag₃₈₆₋₃₉₄GW9 (30, 31). We performed IFN γ -ELISPOT assays using fresh PBMC from cy0428 and cy0575 using these two peptides as stimuli at days 21 or 28 post-infection. We identified detectable responses in the PBMC to the wild type (WT) sequence of both epitopes in both MCMs between three and four weeks post-infection (Figure 4A), similar to previous studies [31, 32].

Additionally, CD8⁺ T cells during acute SIV infection of Mamu-A1*001+ RMs target the SIV peptide Tat₂₈₋₃₅SL8 [8, 27, 33, 34]. We performed IFN γ -ELISPOT assays using fresh PBMC from r10001 and r04103. We detected responses in the PBMC to the WT and S1P variant of Tat₂₈₋₃₅SL8 on day 21 after infection (Figure 4B), indicating that there were T cell responses specific for both the WT and variant epitope sequences.

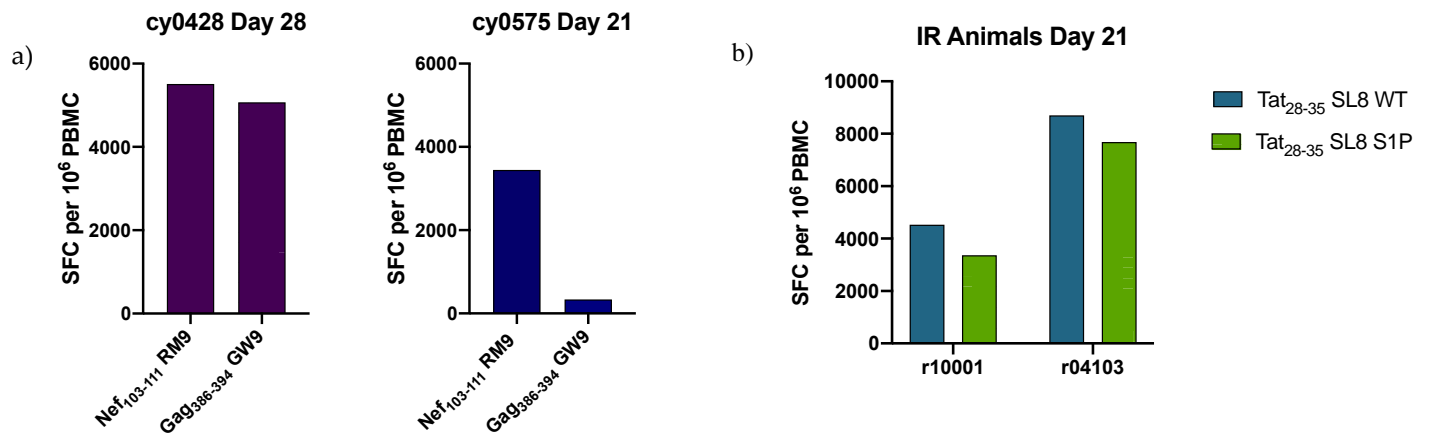


Figure 4. IFN γ -ELISPOT assays detect antigen-specific CD8⁺ T cells during acute SIV infection. (A) In IV-infected cynomolgus macaques, CD8⁺ T cell responses specific for Nef₁₀₃₋₁₁₁RM9 and Gag₃₈₆₋₃₉₄GW9 at day 28 in cy0428 and day 21 in cy0575 were measured by IFN γ -ELISPOT. (B) CD8⁺ T-cell responses to Tat₂₈₋₃₅SL8 wild-type (WT) and the S1P mutations assessed at day 21 were measured by IFN γ -ELISPOT in IR-infected rhesus macaques.

3.4. Variation in T cell epitopes is coincident with reduced barcode diversity in IV-challenged MCMs

We next deep sequenced the Nef₁₀₃₋₁₁₁RM9 and Gag₃₈₆₋₃₉₄GW9 T cell epitopes in viruses circulating in the MCMs to determine the frequency of wild type and variant epitope sequences over time. In cy0428, the most dominant Nef₁₀₃₋₁₁₁RM9 variant was a K3R mutant present at three weeks post-infection, but this was almost completely replaced with a P2T mutant by week four (Figure 5A, left). In cy0575, we detected four Nef₁₀₃₋₁₁₁RM9 variants >10% beginning at week four (Figure 5A, right). However, by week six post-infection, the K3N Nef₁₀₃₋₁₁₁RM9 variant was present at greater than 75% frequency and remained that high until the animal was euthanized (Figure 5A, right). Interestingly, the K3N variant was comprised of two different nucleotide sequences coding for the same amino acid sequence.

When we assessed Gag₃₈₆₋₃₉₄GW9 variants, we found that both the P2L and P5S variants were present in both animals after week four (Figure 5B). We also found that dominant Gag₃₈₆₋₃₉₄GW9 variants emerged at weeks three and six post-infection in MCMs cy0428 and cy0575, respectively, similar to the detection of variants in Nef₁₀₃₋₁₁₁RM9 (Figure 5B). Our findings are consistent with our previous studies suggesting that immune escape is conditional [18, 31], because the detection of variants in Gag₃₈₆₋₃₉₄GW9 does not occur without variants detectable in Nef₁₀₃₋₁₁₁RM9. Conditional escape requires the presence of a particular mutation for a second mutation to lead to escape in that viral lineage.

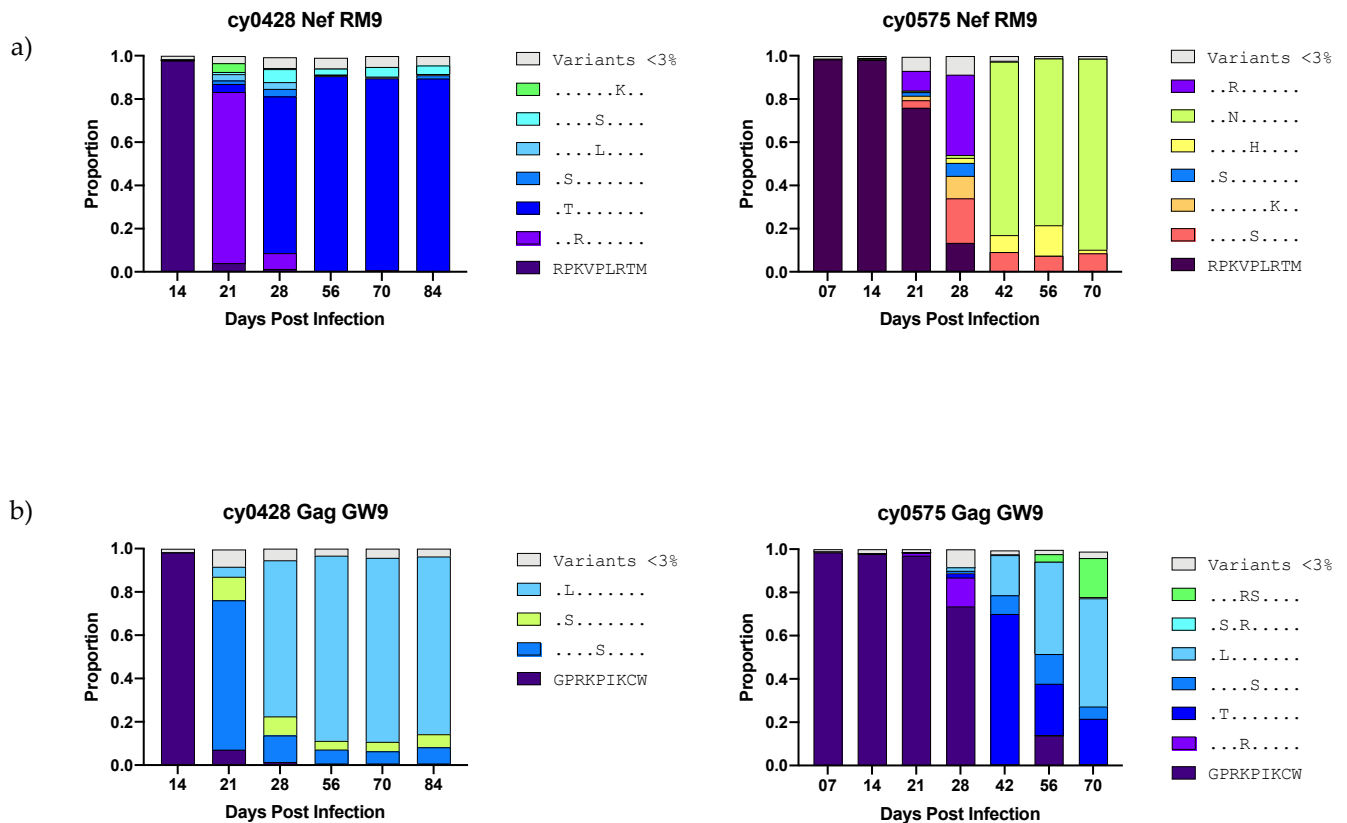


Figure 5. Detection of amino acid variants in Nef₁₀₃₋₁₁₁RM9 and Gag₃₈₆₋₃₉₄GW9 were coincident with reduced barcode diversity. (A) Nef₁₀₃₋₁₁₁RM9 variants identified in the intravenously challenged animals cy0428 and cy0575. The wild-type variant is indicated by RPKVPLRTM. (B) Gag₃₈₆₋₃₉₄GW9 variants identified in the intravenously challenged animals cy0428 and cy0575. The wild-type variant is indicated by GPRKPIKCW. Amino acids identical to the wild type are represented with dots, while mutant amino acids are indicated.

3.5. The same mutations in Tat₂₈₋₃₅SL8 occur on multiple viral lineages in IR-challenged RMs

We longitudinally assessed variants in the T cell epitope, Tat₂₈₋₃₅SL8, targeted in Mamu-A1*001+ RMs. The molecular barcode and Tat₂₈₋₃₅SL8 were captured on the same amplicon, allowing us to link the sequence of the Tat₂₈₋₃₅SL8 epitope with the individual virus lineages. In animal r10001, we identified two common viral lineages present throughout infection: 1758 and 210. Both lineages had a similar distribution of Tat₂₈₋₃₅SL8 variants (Figure 6A) throughout infection, even though lineage 210 became dominant in the plasma (>75%) at necropsy. Although the same Tat₂₈₋₃₅SL8 variant emerged in both lineages, the expansion of lineage 210 during chronic infection suggests that there are likely other adaptations outside of the Tat₂₈₋₃₅SL8 sequence contributing to the expansion of this lineage during chronic infection.

The composition of Tat₂₈₋₃₅SL8 variants replicating in animal r04103 was more complex than for r10001 (Figure 6B). Even though there were up to 11 viral lineages present during acute infection, lineage 176 was most common at necropsy. We found that the Tat₂₈₋₃₅SL8 variants linked to the barcode of lineage 176 were similar to those linked to barcodes of other lineages present during acute infection (e.g., 5836, 4221, and 630) (Figure 6B). However, these other barcoded virus lineages disappeared from the population, despite being linked to the same Tat₂₈₋₃₅SL8 variants as 176 at three weeks post-infection (Figure 6B). These data cumulatively suggest that additional sequences outside of Tat₂₈₋₃₅SL8 in lineage 176 contributed to its expansion during chronic infection, while the other lineages disappeared.

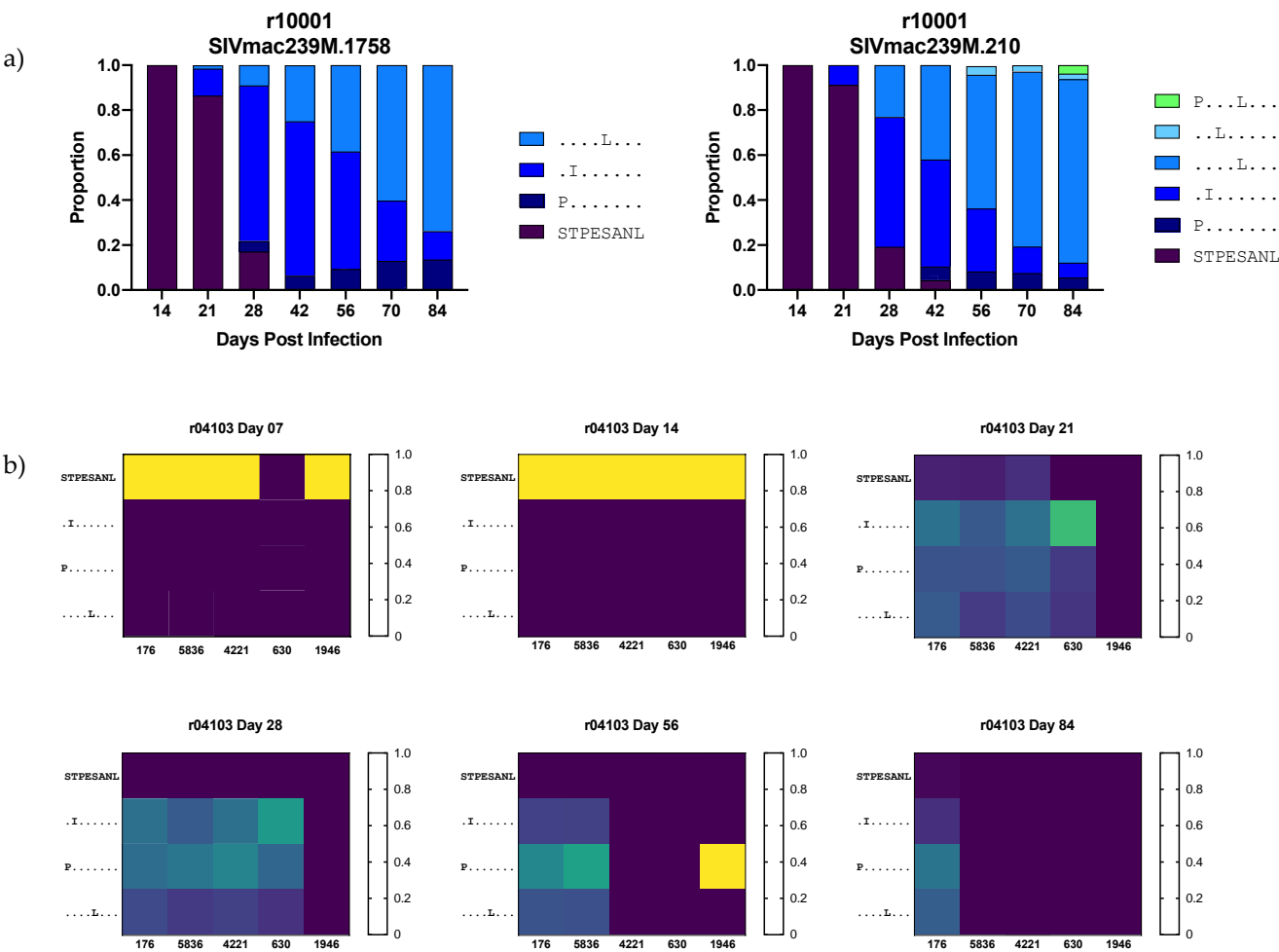


Figure 6. Tat₂₈₋₃₅SL8 escape variants are linked to multiple viral lineages in intrarectally challenged animals. Amplicon data linking the barcode region to the Mamu-A1*001+ restricted epitope Tat₂₈₋₃₅SL8 was examined. (A) The frequency of individual epitope variants linked to each barcode lineage is shown for r10001 and (B) r04103. The frequency of each variant in the heat maps is noted in the legend on the right. The wild-type variant is indicated by STPESANL. Amino acids identical to the wild type are represented with dots, while mutant amino acids are indicated.

4. Discussion

Here, two cynomolgus macaques were challenged IV, and two rhesus macaques were challenged IR with 200,000 IUs of SIVmac239M. Assessing lineage-specific viral dynamics within the viral population increases the power to make conclusions despite the small number of animals. Here we found from 175-950 times more virus lineages circulating during acute infection in the IV-challenged macaques than in the IR-challenged macaques. This difference in the number of unique viral lineages is consistent with historical studies showing that infection route influences the number of virus lineages circulating systemically soon after infection with swarm virus populations [36]. The number of unique virus lineages present after IR challenge in our study was similar to those detected in animals challenged IV with 100 IUs of SIVmac239M [1].

Previous studies comparing the number of transmitted viruses between different challenge routes were complicated by sequence variability in the virus population [36]. One previous study compared challenge routes with SIVmac239X, but that stock only had ten different lineages [2], thereby drastically limiting the power to detect differences in high dose settings. Here, we compared the number of lineages circulating after intrarectal vs. intravenous challenge with SIVmac239M that has thousands of different lineages but which is isogenic outside of the molecular barcode. Similar to challenging animals with a

swarm virus, we found reduced transmission of virus lineages upon intrarectal challenge with SIVmac239M, which is consistent with the hypothesis that the challenge route is the primary driver determining the number of individual virus lineages that initiate systemic SIV/HIV replication [36].

Other studies of antiretroviral-naïve macaques challenged with SIVmac239M tracked virus populations only up to 32 days post-infection [8]. In that study, they found that the virus population had drastically changed with at least a 3-fold change in detectable barcodes at day 32 [8]. Building on that study, we tracked the virus population up to 84 days post-infection in the two intravenously challenged cynomolgus macaques. We found an even greater reduction in the number of systemic lineages in both cynomolgus macaques at approximately week three. In cynomolgus macaques, the number of virus lineages in the IV challenged animals was <50 after detecting acute T cell responses. The species, depth of sequencing, time points selected, or prior experience of the animals are possible reasons for the increased reduction in the number of virus lineages circulating in cynomolgus macaques. Future studies comparing SIVmac239M dynamics in both of these species will be needed to identify which factors may be responsible for these differences.

There were too few virus lineages for the animals challenged intrarectally to accurately dissect out whether an early variant in a targeted epitope emerged on a single or multiple barcoded virus lineages. Instead, we determined the proportion of each Tat₂₈₋₃₅SL8 epitope variant linked to unique barcodes at a specific time. We captured the molecular barcode and Tat₂₈₋₃₅SL8 sequence on the same amplicon, which allowed us to track the persistence or elimination of particular combinations of barcodes and sequences in the Tat₂₈₋₃₅SL8 epitope. We found that multiple independent virus lineages contained the same nucleotide mutations in Tat₂₈₋₃₅SL8 from three to six weeks post-infection. However, one barcode lineage expanded to comprise the majority of the virus population at necropsy and was linked to the diverse array of mutations in Tat₂₈₋₃₅SL8. In contrast, the other barcode lineages present three weeks after infection and linked to the same Tat₂₈₋₃₅SL8 variants reduced in frequency or disappeared from the population. We hypothesize that the barcode lineage that persisted and expanded likely acquired additional variants outside of the sequenced region, but the lineages that disappeared did not acquire these additional variants.

Overall, this study revealed some exciting results that may expand future uses for SIVmac239M. Most importantly, we found that strong, early CTL responses eliminated the majority (97%) of viral lineages in MCMs. This data sets the stage for future studies that could explore whether augmentation of acute CTL responses (e.g. vaccination or activation) may be able to further eliminate virus lineages. This unique model allows for unprecedented resolution of viral population dynamics. With this large dynamic range, it is possible that interventions could be optimized so that even minor improvements are detectable and build upon current methods. In addition, the difference in virus dynamics between the IV-challenged cynomolgus macaques studied here versus previous studies of *Mamu-A1*001+* rhesus macaques infected intravenously with SIVmac239M suggests that perhaps the acute selection mechanisms of virus lineages may vary by species or host MHC genotype. It would be interesting to compare the selection of acute-phase barcode lineages between *Mamu-A1*001+* and *Mamu-B*08+* rhesus macaques challenged with SIVmac239M to determine how the quality of the earliest T cells affects the population of virus lineages present during acute infection. Further, additional studies comparing SIVmac239M dynamics between rhesus macaques and other macaque species may identify critical features of species-specific SIV pathogenicity that are not apparent when animals are challenged with clonal SIVmac239. Also, even though our study only challenged two animals intrarectally with SIVmac239M, our data is consistent with other studies finding that virus transmission is restricted by mucosal tissue [36]. Future studies that continue to explore how the number of transmitted SIVmac239M lineages by different routes may better quantify how mucosal tissue limits transmission under different types of stressors.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: RT-PCR and PCR primers, Table S2: Screening methods for barcode data replicates.

Author Contributions: BFK generated SIVmac239M and provided the R script for barcode identification and analysis. BFK and MPD identified and named the SIVmac239M lineages. DDG, HS, and JM generated sequence data. AJB collected samples from animals. AMW measured SIV viral loads as part of Virology services at the WNPRC that TCF manages. AJB performed ELISPOT assays. BFK and AEG generated the sequence analysis workflows. VV and AEG conceptualized mathematical analysis at the Infection Analytics Program at UNSW-Sydney, directed by MPD. AEG completed the statistical analyses. SLO and AEG wrote the manuscript.

Funding: This project was supported by the Office of The Director, National Institutes of Health under Award Number P51OD011106 to the Wisconsin National Primate Research Center, University of Wisconsin-Madison, including support through the WNPRC Pilot Grant program. This project was also supported through a supplemental task order to NIH R01 AI08415, the National Cancer Institute, National Institutes of Health, under Contract No. 75N91019D00024 (BFK), the Australian National Health and Medical Research Council (NHMRC) grants 1080001 and 1052979 (to MPD), and a NHMRC Career Development Fellowship 1067590 (to VV). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Institutional Review Board Statement: The study was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Wisconsin-Madison (protocol code #G00680, approved on 3/11/15).

Informed Consent Statement: Not applicable

Data Availability Statement: The dataset supporting the conclusions of this article is available in the Genbank repository, accession number PRJNA716664, <https://dataview.ncbi.nlm.nih.gov/object/PRJNA716664>. The in-house Python code used for this study is available on Github, https://github.com/agolfinos/SIV_Transmission_Study. The most updated version of this script is also available on Github, https://github.com/agolfinos/SIVmac239M_analysis.

Acknowledgments: The authors would like to thank the members of the WNPRC for their support with the animal care and other members of SLO's lab for critical discussions about the manuscript. AEG would also like to thank Shaun Owens, MS, for his input regarding computational workflow conceptualization and optimization.

Conflicts of Interest: The authors declare they have no conflicts of interest.

References

1. Khanal, S., C. M. Fennessey, S. P. O'Brien, A. Thorpe, C. Reid, T. T. Immonen, R. Smith, J. W. Bess, A. E. Swanstrom, G. Q. Del Prete, M. P. Davenport, A. A. Okoye, L. J. Picker, J. D. Lifson, and B. F. Keele. 2019. <i>In Vivo</i> Validation of the Viral Barcoding of Simian Immunodeficiency Virus SIVmac239 and the Development of New Barcoded SIV and Subtype B and C Simian-Human Immunodeficiency Viruses. *J Virol* 94, doi: 10.1128/JVI.01420-19
2. Del Prete, G. Q., H. Park, C. M. Fennessey, C. Reid, L. Lipkey, L. Newman, K. Oswald, C. Kahl, M. Piatak, O. A. Quiñones, W. G. Alvord, J. Smedley, J. D. Estes, J. D. Lifson, L. J. Picker, and B. F. Keele. 2014. Molecularly tagged simian immunodeficiency virus SIVmac239 synthetic swarm for tracking independent infection events. *J Virol* 88: 8077-8090. doi: 10.1128/JVI.01026-14
3. Fennessey, C. M. P., Mykola; Immonen, Taina T; Reynaldi, Arnold; Venturi, Vanessa; Nadella, Priyanka; Reid, Carolyn; Newman, Laura; Lipkey, Leslie; Oswald, Kelli; Bosche, William J; Trivett, Matthew T; Ohlen, Claes; Ott, David E; Estes, Jacob D; Del Prete, Gregory Q; Lifson, Jeffrey D; Davenport, Miles P; Keele, Brandon F. 2017. Genetically-barcoded SIV facilitates enumeration of rebound variants and estimation of reactivation rates in nonhuman primates following interruption of suppressive antiretroviral therapy. *PLoS Pathogens* 13. doi: 10.1371/journal.ppat.1006359
4. Shaw, G. M., and E. Hunter. 2012. HIV transmission. *Cold Spring Harb Perspect Med* 2: doi: 10.1101/cshperspect.a006965
5. Butler, D. M., M. E. Pacold, P. S. Jordan, D. D. Richman, and D. M. Smith. 2009. The efficiency of single genome amplification and sequencing is improved by quantitation and use of a bioinformatics tool. *J Virol Methods* 162: 280-283. doi: 10.1016/j.jviro.2009.08.002
6. Salazar-Gonzalez, J. F., E. Bailes, K. T. Pham, M. G. Salazar, M. B. Guffey, B. F. Keele, C. A. Derdeyn, P. Farmer, E. Hunter, S. Allen, O. Manigart, J. Mulenga, J. A. Anderson, R. Swanstrom, B. F. Haynes, G. S. Athreya, B. T. Korber, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 2008. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol* 82: 3952-3970. Doi: 10.1128/JVI.02660-07
7. Aliota, M. T., D. M. Dudley, C. M. Newman, J. Weger-Lucarelli, L. M. Stewart, M. R. Koenig, M. E. Breitbart, A. M. Weiler, M. R. Semler, G. L. Barry, K. R. Zarbock, A. K. Haj, R. V. Moriarty, M. S. Mohns, E. L. Mohr, V. Venturi, N. Schultz-Darken, E.

- Peterson, W. Newton, M. L. Schotzko, H. A. Simmons, A. Mejia, J. M. Hayes, S. Capuano, M. P. Davenport, T. C. Friedrich, G. D. Ebel, S. L. O'Connor, and D. H. O'Connor. 2018. Molecularly barcoded Zika virus libraries to probe in vivo evolutionary dynamics. *PLoS Pathog* 14: e1006964. Doi: 10.1371/journal.ppat.1006964
8. Immonen, T. T., C. Camus, C. Reid, C. M. Fennessey, G. Q. Del Prete, M. P. Davenport, J. D. Lifson, and B. F. Keele. 2020. Genetically barcoded SIV reveals the emergence of escape mutations in multiple viral lineages during immune escape. *Proc Natl Acad Sci U S A* 117: 494-502. Doi: 10.1073/pnas.1914967117
 9. Martin, C. J., A. M. Cadena, V. W. Leung, P. L. Lin, P. Maiello, N. Hicks, M. R. Chase, J. L. Flynn, and S. M. Fortune. 2017. Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *mBio* 8: Doi: 10.1128/mBio.00312-17
 10. Pinkevych, M., C. M. Fennessey, D. Cromer, M. Tolstrup, O. S. Sogaard, T. A. Rasmussen, B. F. Keele, and M. P. Davenport. 2018. Estimating Initial Viral Levels during Simian Immunodeficiency Virus/Human Immunodeficiency Virus Reactivation from Latency. *J Virol* 92: Doi: 10.1128/JVI.01667-17
 11. Pinkevych, M., C. M. Fennessey, D. Cromer, C. Reid, C. M. Trubey, J. D. Lifson, B. F. Keele, and M. P. Davenport. 2019. Predictors of SIV recrudescence following antiretroviral treatment interruption. *Elife* 8: Doi: 10.7554/eLife.49022
 12. Weger-Lucarelli, J., S. M. Garcia, C. Rückert, A. Byas, S. L. O'Connor, M. T. Aliota, T. C. Friedrich, D. H. O'Connor, and G. D. Ebel. 2018. Using barcoded Zika virus to assess virus population structure in vitro and in Aedes aegypti mosquitoes. *Virology* 521: 138-148. Doi: 10.1016/j.virol.2018.06.004
 13. Deleage, C., T. T. Immonen, C. M. Fennessey, A. Reynaldi, C. Reid, L. Newman, L. Lipkey, T. E. Schlub, C. Camus, S. O'Brien, J. Smedley, J. M. Conway, G. Q. Del Prete, M. P. Davenport, J. D. Lifson, J. D. Estes, and B. F. Keele. 2019. Defining early SIV replication and dissemination dynamics following vaginal transmission. *Sci Adv* 5: eaav7116. Doi: 10.1126/sciadv.aav7116
 14. Kaizu, M., G. J. Borchardt, C. E. Glidden, D. L. Fisk, J. T. Loffredo, D. I. Watkins, and W. M. Rehrauer. 2007. Molecular typing of major histocompatibility complex class I alleles in the Indian rhesus macaque which restrict SIV CD8+ T cell epitopes. *Immunogenetics* 59: 693-703. Doi: 10.1007/s00251-007-0233-7
 15. Wiseman, R. W., J. A. Wojcechowskyj, J. M. Greene, A. J. Blasky, T. Gopon, T. Soma, T. C. Friedrich, S. L. O'Connor, and D. H. O'Connor. 2007. Simian immunodeficiency virus SIVmac239 infection of major histocompatibility complex-identical cynomolgus macaques from Mauritius. *J Virol* 81: 349-361. 10.1128/JVI.01841-06
 16. Cline, A. N., J. W. Bess, M. Piatak, and J. D. Lifson. 2005. Highly sensitive SIV plasma viral load assay: practical considerations, realistic performance expectations, and application to reverse engineering of vaccines for AIDS. *J Med Primatol* 34: 303-312. Doi: 10.1111/j.1600-0684.2005.00128.x
 17. Adnan, S., A. D. Colantonio, Y. Yu, J. Gillis, F. E. Wong, E. A. Becker, M. Piatak, R. K. Reeves, J. D. Lifson, S. L. O'Connor, and R. P. Johnson. 2015. CD8 T cell response maturation defined by anentropic specificity and repertoire depth correlates with SIVΔnef-induced protection. *PLoS Pathog* 11: e1004633. Doi: 10.1371/journal.ppat.1004633
 18. Gellerup, D. D., A. J. Balgeman, C. W. Nelson, A. J. Ericson, M. Scarlotta, A. L. Hughes, and S. L. O'Connor. 2016. Conditional Immune Escape during Chronic Simian Immunodeficiency Virus Infection. *J Virol* 90: 545-552. Doi: 10.1128/JVI.02587-15
 19. Sutton, M. S., A. Ellis-Connell, R. V. Moriarty, A. J. Balgeman, D. Gellerup, G. Barry, A. M. Weiler, T. C. Friedrich, and S. L. O'Connor. 2018. Acute-Phase CD4+ T Cell Responses Targeting Invariant Viral Regions Are Associated with Control of Live Attenuated Simian Immunodeficiency Virus. *J Virol* 92: Doi: 10.1128/JVI.00830-18
 20. BBDuk. <https://sourceforge.net/projects/bbmap/>. Platform independent. Java. Compiled for Java 6 and higher. Accessed 24 January 2020.
 21. Dr. Brandon Keele laboratory, NCI-Frederick. Barcode Virus Analysis Tool. Platform independent. R Studio. <https://frederick.cancer.gov/science/compresources>.
 22. Golfinos, Athena. SIV Transmission Study Informatics Pipeline. Platform independent. iPython. https://github.com/agolfinos/SIV_Transmission_Study Accessed 19 January 2021.
 23. O'Connor, S. L., J. J. Lhost, E. A. Becker, A. M. Detmer, R. C. Johnson, C. E. Macnair, R. W. Wiseman, J. A. Karl, J. M. Greene, B. J. Burwitz, B. N. Bimber, S. M. Lank, J. J. Tuscher, E. T. Mee, N. J. Rose, R. C. Desrosiers, A. L. Hughes, T. C. Friedrich, M. Carrington, and D. H. O'Connor. 2010. MHC heterozygote advantage in simian immunodeficiency virus-infected Mauritian cynomolgus macaques. *Sci Transl Med* 2: 22ra18. Doi: 10.1126/scitranslmed.3000524
 24. Budde, M. L., J. M. Greene, E. N. Chin, A. J. Ericson, M. Scarlotta, B. T. Cain, N. H. Pham, E. A. Becker, M. Harris, J. T. Weinfurter, S. L. O'Connor, M. Piatak, J. D. Lifson, E. Gostick, D. A. Price, T. C. Friedrich, and D. H. O'Connor. 2012. Specific CD8+ T cell responses correlate with control of simian immunodeficiency virus replication in Mauritian cynomolgus macaques. *J Virol* 86: 7596-7604. Doi: 10.1128/JVI.00716-12
 25. Zhang, Z. Q., T. M. Fu, D. R. Casimiro, M. E. Davies, X. Liang, W. A. Schleif, L. Handt, L. Tussey, M. Chen, A. Tang, K. A. Wilson, W. L. Triona, D. C. Freed, C. Y. Tan, M. Horton, E. A. Emini, and J. W. Shiver. 2002. Mamu-A*01 allele-mediated attenuation of disease progression in simian-human immunodeficiency virus infection. *J Virol* 76: 12845-12854. Doi: 10.1128/jvi.76.24.12845-12854.2002
 26. Mothé, B. R., J. Weinfurter, C. Wang, W. Rehrauer, N. Wilson, T. M. Allen, D. B. Allison, and D. I. Watkins. 2003. Expression of the major histocompatibility complex class I molecule Mamu-A*01 is associated with control of simian immunodeficiency virus SIVmac239 replication. *J Virol* 77: 2736-2740. Doi: 10.1128/jvi.77.4.2736-2740.2003
 27. O'Connor, D. H., B. R. Mothe, J. T. Weinfurter, S. Fuenger, W. M. Rehrauer, P. Jing, R. R. Rudersdorf, M. E. Liebl, K. Krebs, J. Vasquez, E. Dodds, J. Loffredo, S. Martin, A. B. McDermott, T. M. Allen, C. Wang, G. G. Doxiadis, D. C. Montefiori, A. Hughes, D. R. Burton, D. B. Allison, S. M. Wolinsky, R. Bontrop, L. J. Picker, and D. I. Watkins. 2003. Major histocompatibility complex

- class I alleles associated with slow simian immunodeficiency virus disease progression bind epitopes recognized by dominant acute-phase cytotoxic-T-lymphocyte responses. *J Virol* 77: 9029-9040. Doi: 10.1128/jvi.77.16.9029-9040.2003
28. Kijak, G. H., E. Sanders-Buell, A. L. Chenine, M. A. Eller, N. Goonetilleke, R. Thomas, S. Levityang, E. A. Harbolick, M. Bose, P. Pham, C. Oropeza, K. Poltavee, A. M. O'Sullivan, E. Billings, M. Merbah, M. C. Costanzo, J. A. Warren, B. Slike, H. Li, K. K. Peachman, W. Fischer, F. Gao, C. Cicala, J. Arthos, L. A. Eller, R. J. O'Connell, S. Sinei, L. Maganga, H. Kibuuka, S. Nitayaphan, M. Rao, M. A. Marovich, S. J. Krebs, M. Rolland, B. T. Korber, G. M. Shaw, N. L. Michael, M. L. Robb, S. Tovanabutra, and J. H. Kim. 2017. Rare HIV-1 transmitted/founder lineages identified by deep viral sequencing contribute to rapid shifts in dominant quasiespecies during acute and early infection. *PLoS Pathog* 13: e1006510. Doi: 10.1371/journal.ppat.1006510
 29. Loh, L., J. Petravic, C. J. Batten, M. P. Davenport, and S. J. Kent. 2008. Vaccination and timing influence SIV immune escape viral dynamics in vivo. *PLoS Pathog* 4: e12. Doi: 10.1371/journal.ppat.0040012
 30. Goonetilleke, N., M. K. Liu, J. F. Salazar-Gonzalez, G. Ferrari, E. Giorgi, V. V. Ganusov, B. F. Keele, G. H. Learn, E. L. Turnbull, M. G. Salazar, K. J. Weinhold, S. Moore, B. CHAVI Clinical Core, N. Letvin, B. F. Haynes, M. S. Cohen, P. Hrabner, T. Bhattacharya, P. Borrow, A. S. Perelson, B. H. Hahn, G. M. Shaw, B. T. Korber, and A. J. McMichael. 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med* 206: 1253-1272. Doi: 10.1084/jem.20090365
 31. O'Connor, S. L., E. A. Becker, J. T. Weinfurter, E. N. Chin, M. L. Budde, E. Gostick, M. Correll, M. Gleicher, A. L. Hughes, D. A. Price, T. C. Friedrich, and D. H. O'Connor. 2012. Conditional CD8+ T cell escape during acute simian immunodeficiency virus infection. *J Virol* 86: 605-609. Doi: 10.1128/JVI.05511-11
 32. Bimber, B. N., B. J. Burwitz, S. O'Connor, A. Detmer, E. Gostick, S. M. Lank, D. A. Price, A. Hughes, and D. O'Connor. 2009. Ultradeep pyrosequencing detects complex patterns of CD8+ T-lymphocyte escape in simian immunodeficiency virus-infected macaques. *J Virol* 83: 8247-8253. Doi: 10.1128/JVI.00897-09
 33. Allen, T. M., D. H. O'Connor, P. Jing, J. L. Dzuris, B. R. Mothé, T. U. Vogel, E. Dunphy, M. E. Liebl, C. Emerson, N. Wilson, K. J. Kunstman, X. Wang, D. B. Allison, A. L. Hughes, R. C. Desrosiers, J. D. Altman, S. M. Wolinsky, A. Sette, and D. I. Watkins. 2000. Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* 407: 386-390. Doi: 10.1038/35030124
 34. Vanderford, T. H., C. Bleckwehl, J. C. Engram, R. M. Dunham, N. R. Klatt, M. B. Feinberg, D. A. Garber, M. R. Betts, and G. Silvestri. 2011. Viral CTL escape mutants are generated in lymph nodes and subsequently become fixed in plasma and rectal mucosa during acute SIV infection of macaques. *PLoS Pathog* 7: e1002048. Doi: 10.1371/journal.ppat.1002048
 35. Greenier, J. L., C. J. Miller, D. Lu, P. J. Dailey, F. X. Lü, K. J. Kunstman, S. M. Wolinsky, and M. L. Marthas. 2001. Route of simian immunodeficiency virus inoculation determines the complexity but not the identity of viral variant populations that infect rhesus macaques. *J Virol* 75: 3753-3765. Doi: 10.1128/JVI.75.8.3753-3765.2001
 36. Keele, B. F., H. Li, G. H. Learn, P. Hrabner, E. E. Giorgi, T. Grayson, C. Sun, Y. Chen, W. W. Yeh, N. L. Letvin, J. R. Mascola, G. J. Nabel, B. F. Haynes, T. Bhattacharya, A. S. Perelson, B. T. Korber, B. H. Hahn, and G. M. Shaw. 2009. Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. *J Exp Med* 206: 1117-1134. Doi: 10.1084/jem.20082831