JVI Accepts, published online ahead of print on 7 May 2014
J. Virol. doi:10.1128/JVI.00543-14

Deep transcriptional sequencing of mucosal challenge compartment from rhesus macaques acutely infected with simian immunodeficiency virus implicates loss of cell adhesion preceding immune activation

Fredrik Barrenas[1,2], Robert E. Palermo[1], Brian Agricola[3], Michael B. Agy[3], Lauri Aicher[1], Victoria Carter[1], Leon Flanary[3], Richard R. Green[1], Randy McLain[3], Qingsheng Li[4], Wuxun Lu[4], Robert Murnane[3], Xinxia Peng[1], Matthew J. Thomas[1], Jeffrey M. Weiss[1], David M. Anderson[3], Michael G. Katze[1,3]

[1] *Department of Microbiology, University of Washington, Seattle, WA 98195*

[2] *Department of Cell and Molecular Biology, Uppsala University, Box 596, 751 24 Uppsala*

[3] *Washington National Primate Research Center, University of Washington, Seattle, WA 98195*

[4] *Nebraska Center for Virology, School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68583*

Running title: Deep mRNA-seq of rectal mucosa in acute SIV infection

†Correspondence should be addressed to Michael G. Katze, Department of Microbiology, Washington National Primate Research Center, University of Washington, Box 358070, Seattle, WA 98195-8070, USA

Phone: (206) 732-6135

Fax: (206) 732-6056

E-Mail: honey@u.washington.edu

Abstract

Pathology resulting from HIV infection is driven by protracted inflammation; the primary loss of CD4+ T cells is caused by activation-driven apoptosis. Recent studies of nonhuman primates (NHPs) have suggested that during the acute phase of infection, antiviral mucosal immunity is restricting viral replication in the primary infection compartment. These studies imply that HIV achieves systemic infection as a consequence of a failure in host antiviral immunity. Here, we used high-dose intrarectal inoculation of rhesus macaques with SIV_{mac251} to examine how the mucosal immune system is overcome by SIV during acute infection. The host response in rectal mucosa was characterized by mRNA deep sequencing (mRNA-seq) at 3 and 12 days post inoculation (DPI) in 4 animals for each time point. While we observed a strong host transcriptional response at 3 DPI, functions relating to antiviral immunity were absent. Instead, we observed a significant number of differentially expressed genes relating to cell adhesion and reorganization of the cytoskeleton. We also observed down regulation of genes encoding members of the claudin family of cell adhesion molecules, co-expressed with genes associated with pathology in the colorectal mucosa, and a large number of noncoding transcripts. By contrast, the differentially expressed genes at 12 DPI were enriched in immune system functions, in particular functions relating to T cells, B cells, and NK cells. Our findings indicate that host responses that negatively affect mucosal integrity occur before inflammation. Consequently, when inflammation is activated at peak viremia, mucosal integrity is already compromised, potentially enabling rapid tissue damage, driving further inflammation.

Importance

The HIV pandemic is one of the major threats to human health, causing over a million deaths per year. Recent studies have suggested that mucosal antiviral immune responses play an important role in preventing systemic infection after exposure to the virus. Yet, despite their potential role in decreasing transmission rates between individuals, these antiviral mechanisms are poorly understood. Here, we carried out the first deep mRNA sequencing analysis of mucosal host responses in the primary infection compartment during acute SIV infection. We found that during acute infection, a significant host response was mounted in the mucosa before inflammation was triggered. Our analysis indicated that the response has a detrimental effect on tissue integrity, causing increased permeability, tissue damage and recruitment of SIV target cells. These results emphasize the importance of mucosal host responses preceding immune activation in preventing systemic SIV infection.

Introduction

HIV-induced immune cell depletion is primarily caused by protracted inflammation. However, recent studies indicate that the earliest time period following HIV infection, the acute phase, is critical in HIV pathogenesis (1). Nonhuman primate (NHP) models have revealed that during this phase, mucosal immunity can inhibit viral replication and prevent systemic infection. The low rate of sexual transmission among humans (≤ 0.5% occurrences per sexual contact) also implies that the virus runs a high risk of dying in the primary infection compartment, or adjacent tissues, before it can infect a sufficient number of CD4+ cells to spread systemically. Even after exposure to the virus, early antiretroviral medication can clear the virus from the host (2, 3). This makes the acute phase a critical time point in the transmission process.

Currently, simian immunodeficiency virus (SIV) infection of NHPs is the most accurate model to study the early events following HIV infection in humans. In this study, we used deep RNA sequencing to study the host transcriptional response at the site of inoculation during early SIV infection of rhesus macaques (RMs). The aim was to identify aspects of the host response that contribute to early viral control, or loss thereof, leading to subsequent systemic HIV/SIV infection. Understanding these processes could greatly contribute to the development of therapies to decrease transmission rates between individuals. To our knowledge, this is the first in-depth look at the mucosal host response to SIV infection during the acute phase.

The acute phase of HIV/SIV infection follows a characteristic time course, which provides an opportunity to identify key events that could be modulated by vaccines or other therapeutics to limit subsequent pathogenesis. Sexual transmission of HIV/SIV is followed by a time period when the virus is undetectable in circulation, termed the *eclipse phase* (4). NHP studies have led to the discovery of several mechanisms by which the host controls SIV, including viral entry

77 blockage (e.g. SDF-1, MIP1a/b), IFNa/b expression by mucosal dendritic cells and host
78 restriction factors (e.g. BST2, CD317) (1). If these mechanisms fail, a small number of virions
79 eventually escape host restriction and infect mucosal $CD4^+$ T cells, macrophages and dendritic
80 cells. Carried by DCs, the virus eventually reaches the draining lymph nodes, which brings it into
81 contact with a large number of target $CD4^+$ $CCR5^+$ T cells (5). Viral replication then increases
82 rapidly as the infection spreads first to other lymph nodes, then to the bloodstream – this process
83 takes about 1 week in NHPs. The rising viremia is accompanied by an adaptive immune
84 response, involving specific B cells and CD8+ and CD4+ T cells (6). Around day 12, viremia
85 reaches its peak, and then decreases to the *viral set point*. At the viral set point, although viral
86 load is reduced, the immune system is gradually depleted through inflammation-driven
87 apoptosis. The protracted inflammation at the viral set point is a key driver in the development of
88 AIDS, as shown by three main findings: (a) most of the immune cells that are lost in HIV
89 infection are themselves not infected by the virus, but rather are bystander cells (7), (b) NHP
90 species that are natural carriers of SIV (e.g. African green monkeys) resolve the inflammatory
91 response after peak viremia and do not develop immunodeficiency despite high viral load (8, 9)
92 and (c) levels of inflammatory markers predict disease progression more accurately than viral
93 load (10, 11).

94 In the interest of learning how viral spread from the primary infection compartment can be
95 prevented, we focused this study on the eclipse phase. Building upon *in vivo* studies of NHPs
96 that found mucosal antiviral mechanisms capable of restricting SIV replication, (12) we
97 investigate the critical question of how this control is maintained and lost. The fact that the
98 sexual transmission rate of HIV is low implies that the physical barriers and immune systems in
99 the mucosal tissues are normally successful at preventing infection (13, 14). In apparent

contradiction, activation of the innate immune system during the acute phase can also be harmful to the host; early responses by NK cells, macrophages and dendritic cells attract CD4+ T lymphocytes and other target cells to the site of infection (15). Mucosal immune activation has additional harmful effects, as it can result in compromise to the integrity of the mucosal barrier, thereby leading to bacterial translocation into the lamina propria, enhancing protracted inflammation (16). It has not been determined whether the loss of epithelial integrity is merely a consequence of ongoing inflammation, or if the virus itself can cause epithelial damage. While *in vitro* studies have found that HIV/SIV can affect the integrity of epithelial explants without the influx of immune cells (17), our use of NHPs provides a unique opportunity to explore this issue *in vivo*.

In this study, we carried out high-dose intrarectal inoculation of Indian-origin rhesus macaques (RMs) with SIV_{mac251} and sacrificed animals at 3 or 12 days post-inoculation (DPI), corresponding to the eclipse phase and peak viremia respectively. Rectal specimens were obtained at the site of inoculum deposition and analyzed by deep mRNA sequencing (mRNA-seq) to capture the host response at the infection site. During the eclipse phase (3 DPI), we observed the activation of genes encoding cytoskeletal remodeling and cell adhesion proteins. This response was followed by a strong inflammatory response during peak viremia (12 DPI).

A notable of advantage of RNA-seq is that it can quantify both coding gene and noncoding RNA (ncRNA) expression. The role of ncRNA in SIV or HIV pathogenesis has not been extensively explored, but it has been shown that the suppression of enzymes required for microRNA biogenesis leads to enhanced HIV replication in PBMCs (18). Likewise, knockdown of particular long ncRNA can affect HIV replication (19), similar to the knockdown of coding genes known as HIV dependency factors. In addition, RNA sequencing of HIV-infected SUP-T1 cells has

123 shown the differential regulation of several classes of ncRNA (20). Whereas these previous
124 studies were carried out *in vitro*, the design of the present study allowed us to examine ncRNA
125 expression *in vivo* on a genomic scale.

126 This study suggests that even during the eclipse phase, the rectal mucosa is mounting a response
127 that compromises epithelial integrity. The fact that these changes preceded both local
128 inflammation and viremia raises the possibility that these changes are independent of
129 inflammation, and play an important role in the loss of viral containment within the challenge
130 compartment. This is the first study to evaluate mucosal host responses associated with epithelial
131 compromise as early as 3 days after SIV inoculation.

Materials and Methods

Animals

Ethics statement: All animal procedures were performed using standard protocols and according to guidelines approved by the University of Washington Environmental Health and Safety Committee, the Occupational Health Administration, the Primate Center Research Review Committee, and the Institutional Animal Care and Use Committee. The eight male RMs that underwent intrarectal SIV challenge were housed at the Washington National Primate Research Center. Control tissues from three uninfected animals were obtained from the tissue distribution program run by the National Primate Research Centers. All challenged animals were specific pathogen free (SPF) and were negative for the protective MHC Class I alleles Mamu A01, B01 and B17.

Intrarectal SIV challenge

The eight RMs were intrarectally challenged with SIVmac251 using 1 mL of a high-dose inoculum (6000 $TCID_{50}$/mL). SIV inoculates were deposited at a rectal depth of 25 mm from the anus. Baseline rectal samples were obtained 14 days prior to viral challenge by pinch biopsy.

Tissue preparation and RNA extraction

At necropsy, all rectal samples were taken at a depth of 25 mm from the anus, corresponding to the depth at which inoculates were deposited. Rectal tissues were immediately perfused in RNAlater and stored at -80°C until further processing. Tissues were homogenized in 20 volumes

of RLT reagent (Qiagen) using an Omni TH tissue homogenizer (Omni International, Kennesaw, GA). RNA was extracted from rectal tissue homogenate using the AllPrep DNA/RNA/Protein kit (Qiagen). RNA concentrations were quantified using an ND-2000c UV-Vis spectophotometer (NanoDrop, Wilmington, DE) and controlled for integrity and purity on a capillary electrophoresis system (Agilent 2100 Bioanalyzer; Agilent Technologies, Santa Clara, CA).

Blood samples for viral load measurement were taken (a) at the time of baseline sampling 14 days prior to inoculation, (b) at the 3 DPI necropsy, (c) at 6 DPI in the animals to be sacrificed at 12 DPI and (d) at the 12 DPI necropsy. Whole blood was collected into EDTA tubes (Becton Dickinson, Franklin Lakes, NJ) for use in plasma isolation. Tubes were mixed by inversion and subsequently centrifuged at 1,300 × *g* for 10 minutes. The upper (plasma) layer was carefully removed and stored at -80°C for later analysis. Viral RNA was prepared from EDTA-anticoagulated, cell-free plasma using the Gentra Puregene RNA isolation kit according to the manufacturer's instructions (Gentra Systems, Minneapolis, MN, USA). RNA was precipitated in the presence of glycogen, resuspended in 50 µl of nuclease-free water and analyzed immediately.

Viral load measurement

Plasma viral load was determined by real-time reverse transcription polymerase chain reaction (RT-qPCR) based on published methods (21). Intracellular viral RNA load in rectal mucosa was quantified as previously described (22).

mRNA library preparation

mRNA libraries were constructed using the Illumina TruSeq RNA Preparation Kit (Illumina, San Diego, CA) according to the manufacturer's guide. Libraries were quality controlled and quantitated using the BioAnalzyer 2100 system and qPCR (Kapa Biosystems, Woburn, MA). The libraries were clonally amplified on a cluster generation station using Illumina version 4 cluster generation reagents to achieve a target density of approximately 700,000 (700K)/mm^2 in a single channel of a flow cell.

Next generation sequencing and read mapping

The resulting libraries were sequenced on a Genome Analyzer IIx (Illumina, San Diego, CA) using Illumina version 5.0 sequencing reagents which generated paired-end reads of 75 nucleotides (nt). Image analysis, base calling, and error estimation were performed using Illumina Analysis Pipeline (version 2.8). Raw reads were trimmed to 50 bp and adapter sequences were removed. The 50-bp reads were mapped to known ribosomal sequences (human, mouse, rat) using the short-read aligner software Bowtie to remove potential rRNA sequences to maximize the coverage of mapped reads to our annotation (23). Viral reads were then determined by mapping to the SIVmac251 genome (GenBank accession no. M19499.1) using the gapped aligner software TopHat, which predicts splicing junctions and maps intron-spanning reads to known splicing junctions (24). We then mapped all the remaining reads to the rhesus macaque reference genome (Source: Ensembl, Build: Mmul_1) from Illuminia's igenomes (http://support.illumina.com/sequencing/sequencing_software/igenome.ilmn) using TopHat. After mapping, we assigned aligned read counts from BAM files to exons and genes using the python package HT-Seq (25). HT-Seq provided the most accurate way of aligning read counts to

overlapping exons. Reads that mapped to multiple positions were removed. Annotation for human large intergenic noncoding RNA (lincRNA) were obtained from a previously published catalogue (26). Annotation for novel macaque ncRNA were obtained from The Nonhuman Primate Reference Transcriptome Resource (27, 28).

For visualization, BAM files were generated using TopHat and SAMtools (29) and displayed using the IGV Genome Browser. Read count refers to the number of sequenced cDNA fragments that map to a particular genomic feature. Normalization and differential expression analysis were carried out using R (version 2.14.1) and software package edgeR. Normalization comprised of calculating a size factor for each sample (as the median ratio of read counts for each feature and sample to the geometric mean of read counts for each feature across samples) and dividing all of the read counts in a particular sample by the sample size factor (30). The data is available at Gene Expression Omnibus (accession GSE56845) and SRA.

Differential expression analysis

Differentially expressed mRNA and noncoding RNA were determined using a generalized linear model implemented in the Bioconductor package edgeR (R version 2.15.3; edgeR version 1.8.3). To avoid bias between samples obtained by necropsy and baseline samples obtained by pinch biopsy, differential expression was determined by two tests. First, each gene underwent a paired test between necropsy samples and baseline samples from the same animal. Second, each gene underwent a group-wise unpaired test for the time point of interest (3 DPI or 12 DPI) vs. all eight pinch biopsy baseline samples and uninfected rectal necropsy samples. P-values were adjusted for multiple testing by FDR (Figure S1). Differentially expressed coding genes and noncoding

RNA were defined as having an adjusted p-value of <0.05 and absolute fold change >1.5 in both tests.

Functional enrichment analysis

Functional enrichment of differentially expressed genes was carried out using Ingenuity Pathway Analysis (Ingenuity Systems, Inc). Predefined, manually curated functional categories, containing given genes, were tested for statistical enrichment with differentially expressed genes using the Fisher's Exact test. The functional categories have a hierarchical organization, with more specific sub-categories (e.g. *Activation of lymphocytes*) being aggregated into more generic categories (e.g. *Cellular signaling and interaction*). Since the specific sub-categories can be highly overlapping, we only present the most enriched sub-category in each broad category. In addition to determining the enrichment of DE genes in functional categories, the analysis includes information on genes that inhibit or activate each function and uses a regularization z-score to predict whether an enriched functional category is inhibited or activated (31). The categories shown in Figure 2B were chosen by manual survey of each generalized category and selection of specific sub-categories that included a large number of DE genes and exhibited a strong enrichment p-value.

Functional enrichment analysis of co-expressed gene sets was carried out using the GOSim library.

Co-expression analysis

All differentially expressed coding RNA and ncRNA from both time points were binned together, amounting to 4,015 transcripts. Co-expression between all pairs of transcripts was determined using biweighted midcorrelation, a measure which has shown good performance compared to alternative methods (32). Co-expressed transcripts were organized into modules by hierarchical clustering using the Ward method (33) and adaptive branch pruning (34).

To avoid the risk that a small set of outlier samples play a dominating role in generating co-expressed gene sets, we evaluated the hierarchical clustering by a bootstrap test, wherein the hierarchical clustering was repeated on randomized subsets of the data (35).

The co-expression network was constructed by connecting each transcript to the two other transcripts with which it shared the highest biweighted midcorrelation. This method does not require a correlation cutoff, and it also avoids the construction of networks consisting of large completely connected groups of genes, or groups of genes that are completely unconnected (36).

Quantitative PCR

RNA from rectal samples was reverse transcribed using the QuantiTect reverse transcription kit (Qiagen, Valencia, CA). The resulting cDNA samples were diluted 50x. SYBR green qPCR assays were run for each sample in triplicate. Relative expression was calculated using the $\Delta\Delta CT$ method with averaged ΔCT values (where *CT* stands for threshold cycle) for the Rhesus macaque 18s (FJ436026.1) and ACTB1 (NM_001033084.1) genes as a calibrator, as the expression of either did not significantly change over time in the mRNA sequencing data.

Analysis of public datasets related to ulcerative colitis

Four datasets contrasting colon biopsies from patients with ulcerative colitis (UC) to healthy controls were obtained from Gene Expression Omnibus (GSE9686, GSE10191, GSE22619, GSE38713). All datasets were individually normalized using quantile normalization; DE genes were determined using the Bioconductor package *limma*. Genes were classified as up- or down-regulated if their p-value was below 0.05 and their fold change showed the same direction (up or down) in all four analyses.

Results

Viral reads detectable in rectal mucosa 3 days post inoculation

Eight Indian-origin rhesus macaques (RMs) were infected intrarectally with SIV_{mac251}, using an inoculum containing 6,000 $TCID_{50}$. The inoculation protocol was designed to avoid abrasions in the rectal mucosa. Four animals were sacrificed at 3 days post inoculation (DPI) and four additional animals at 12 DPI (Figure 1A). Rectal tissues were obtained at necropsy from all animals at the site of inoculum deposition. Rectal tissues were examined to ensure that no visible damage had been caused by the inoculation procedure. Uninfected baseline samples were obtained by pinch biopsy 14 days prior to inoculation (referred to as -14 DPI). To avoid bias arising from the different sampling techniques, we also included rectal mucosal tissue from necropsy of three uninfected RMs. Host responses in rectal mucosa were examined on the transcriptomic level using deep mRNA sequencing.

To examine viral levels at these time points, virus was quantified in (a) peripheral blood and rectal mucosa by quantitative RT-PCR and (b) rectal mucosa by mapping sequenced mRNA reads to the SIV genome (Figure 1). Viral DNA was quantified by quantitative RT-PCR (see supporting information). Viral RNA was undetectable in blood samples taken at baseline (data not shown) and at 3 DPI. At 6 DPI, low levels of viremia were detected, ranging from 355 to 8,730 RNA copies/ml of plasma (geometric mean = 3,020 RNA copies/ml). At 12 DPI, viremia had risen to between 1.41×10^7 and 2.59×10^7 RNA copies/ml of plasma (geometric mean = 2.10×10^7 RNA copies/ml; Figure 1B, left panel). Similarly, at 3 DPI in rectal mucosa, no virus was detected in three out of four animals while high viral load was found at 12 DPI (ranging

from 3.00×10^5 to 6.99×10^5; Figure 1B, center panel). By contrast, using deep mRNA-seq, low numbers of SIV reads we detected at 3 DPI in mucosal samples from all four animals, ranging from 1 to 18 reads/sample (geometric mean = 7.54). Although these numbers were very low, no reads mapping to the SIV genome were found in any of the baseline mucosa samples or in mucosa from uninfected control animals. At 12 DPI, the number of reads in rectal mucosa had almost risen above 7,000 in all animals (geometric mean N_{reads} = 15,882; Figure 1B, right panel).

These analyses imply that at 3 DPI, the virus was present in rectal mucosa, but had not reached the bloodstream, corresponding to the eclipse phase. Virus in blood was detectable at 6 DPI; by 12 DPI, the viral load had increased by several orders of magnitude in both the challenge compartment and in blood.

Strong mucosal transcriptional response detected during both the eclipse phase and at peak viremia

To characterize the mucosal host response at 3 and 12 DPI, mRNA-seq data was mapped to the rhesus macaque (RM) genome, enabling the quantification of coding gene and noncoding RNA (ncRNA) expression. To limit technical noise, we first summed up the number of mapped reads for each annotated gene and ncRNA across all 19 samples and removed those with less than 20 detected reads; 18,926 of 30,246 annotated coding genes were expressed in the rectal samples (i.e., genes with >20 total reads). This cutoff was chosen to enable detection of transcripts only expressed in a small subset of cells in the mucosal samples (e.g. immune cells).

Coding genes and ncRNA that were differentially expressed at 3 and 12 DPI were identified by contrasting post-inoculation necropsies to (a) matched baseline pinch biopsies from each animal

using a paired test and (b) a pool of all baseline pinch biopsies and necropsy samples from uninfected animals by an unpaired test (Table S1). Differential expression was defined as having an adjusted p-value (FDR) of <0.05 and an absolute fold change ≥1.5 ($|\log_2 FC| \geq 0.58$) in both of these tests. Despite these stringent criteria, we identified a large number of differentially expressed coding genes as well as ncRNA at both time points. The number of differentially expressed coding genes at 3 and 12 DPI was 1,507 and 2,905, respectively (Figure 2A, supporting information). Notably, over 85% of all DE genes at each time point showed up-regulation.

DE genes at 3 DPI showed a strong tendency to remain DE at 12 DPI, particularly in two of the animals. Specifically there were 888 DE genes common to both time points, which amounted to a 3.84-fold enrichment compared to random ($P < 10^{-15}$; Fisher's Exact Test based on the 18,926 expressed genes). All of these genes showed the same direction of change at both time points, meaning that a quarter of the genes that were differentially expressed at peak viremia were already differentially expressed during the eclipse phase.

Mucosal immune response is preceded by the differential expression of genes associated with rearrangement of cytoskeleton and cell adhesion

For a comprehensive biological overview of the DE genes, we utilized Ingenuity Pathway Analysis (IPA) to determine functional gene categories that were enriched among the DE genes, as well as what categories showed activation or repression (Figure 2B, Table S2-S3). At 3 DPI, the functional enrichment primarily implicated genes associated with increased rearrangement of the cytoskeleton and the formation of cellular protrusions. These are associated with several aspects of HIV infection in humans, including endocytosis, exocytosis and recruitment of co-

receptors to a HIV-bound CD4 receptor. However, a significant number of DE genes at 3 DPI were also involved in differentiation of several types of connective tissue cells, including stromal cells, smooth muscle cells and adipocytes (Figure 2B, Table S2). We also found a significant number of DE genes involved in maintenance of epithelial tissue integrity (e.g. *Cell-to-cell adhesion*, *Formation of focal adhesions*, *Gap junction signaling*). This suggests that the reorganization of the cytoskeleton is not limited to intracellular structures, but also indicates early effects on epithelial cell adhesion. Notably, while our analysis showed significant transcriptional perturbation at the inoculation site early after SIV infection, we also observed a marked absence of inflammatory functions. We did observe a small number of DE genes at 3 DPI with a documented role in host response to HIV infection, including IFNA21 (37), TRIM22 (38) and ISG20 (39). These genes could represent the first signs of an antiviral immune response.

In contrast, the 12 DPI analyses showed a strong enrichment of DE genes in immune functions, predominantly T cell activation (Figure 2B, Table S3). Whereas many of these functions related to T cells in general, two T cell subsets were specifically implicated: Th1 cells and cytotoxic T lymphocytes – these specialized functions are subsets of generic T cell functions. Activation of B cells was also evident, including earlier phases of development (e.g., *development of pro-B lymphocytes*) and IgG production. The enriched functions also implicated activation and recruitment of several innate immune cell types, including NK cells, dendritic cells, eosinophils and macrophages. This suggests a well-developed antiviral immune response at 12 DPI involving the recruitment of innate and adaptive cell types. Many of the functions that were enriched at 3 DPI included the same number of DE genes at 12 DPI, but with the total number of DE genes at 12 DPI having increased by a factor of ~2, these functions were not statistically enriched at the later time point. If this study had focused exclusively on peak viremia, the

344 statistical enrichment of DE genes associated with the activation of cytoskeletal functions would
345 not have been detected.

346 We also more closely analyzed the underlying pathways, corresponding intracellular signaling
347 cascades that were enriched at the two time points. (Figure 2C). This analysis confirmed the
348 clear difference between the cytoskeletal and cell adhesion pathways that were activated at 3
349 DPI, and the immune system pathways that were activated at 12 DPI. Among the many immune
350 cell related pathways, interferon signaling ranked among the highest at 12 DPI, with DE genes
351 including six type I interferons (IFNA2, IFN6, IFN8, IFN10, IFN14 and IFNB1) and the type II
352 interferon IFNG. Consistent with the presence of interferon signaling, we identified 133 DE
353 interferon stimulated genes (ISGs) at 12 DPI, which constituted a 2.95-fold enrichment ($p < 10^{-15}$).
354 In fact, the 100 most significantly DE genes at 12 DPI included 54 ISGs, amounting to a
355 26.1-fold enrichment. Several of the most activated pathways also included the strong up-
356 regulation of genes encoding pattern recognition receptors, including TLR2 (9.64-fold), TLR3
357 (5.03-fold), and CLEC7A (7.41-fold). Whereas TLR3 activates IRF3/7 upon recognition of viral
358 antigens (double-stranded RNA), TLR2 and CLEC7A both activate NF-κB after recognizing
359 bacterial and fungal antigens, respectively.

360 **Co-expression analysis implicates coding genes and ncRNA in loss of cell**
361 **adhesion**

362 While the analyses presented above provided an overview of the functions associated with
363 differentially expressed coding genes, RNA-seq can also be used to quantify ncRNA. To predict
364 the functions of ncRNA in acute SIV infection, we carried out a co-expression analysis between

differentially regulated ncRNA and coding genes. This also assisted with the functional characterization of transcripts with specific expression patterns.

For a comprehensive overview of rhesus macaque ncRNA, sequenced reads were mapped to three annotations that explored different categories of ncRNA (Figure 3A). First, we used previously characterized ncRNA from the Ensembl rhesus macaque reference genome (build MMUL_1), which provided an overview of classes of ncRNA (e,g. rRNA or miRNA). We term these *known ncRNA*. The second annotation consisted of sequences orthologous to human large noncoding intergenic RNA (26), which we term *lincRNA*. The third annotation consisted of transcripts from the nonhuman primate reference transcriptome resource (27) that did not correspond to any characterized transcript in Ensembl and which showed low protein-coding potential. This was the most comprehensive annotation, containing 6,027 transcripts in total. We term these *unannotated ncRNA*. Given that these transcripts were not previously annotated, we carried out RT-qPCR validation of five strongly up-regulated unannotated ncRNA, showing high reproducibility of RNA sequencing results (Figure S2).

Due to the lack of functional annotation for ncRNA, we carried out a co-expression analysis between coding genes and ncRNA to associate ncRNA with the functions of their co-expressed coding genes. We pooled the 4015 differentially expressed transcripts (both coding and non-coding) and organized them into co-expressed groups by hierarchical clustering, which produced 8 co-expressed gene sets (referred to as clusters 1-8, Figure 3B, Table S4). The reproducibility of this analysis was tested by a bootstrapping test, which showed highly reproducible hierarchical structure ($p < 0.0001$ for all eight clusters). Most of these clusters showed predominant up-regulation at either 3 or 12 DPI. The most notable exception was cluster 4, which showed down-regulation at both 3 and 12 DPI. For each cluster, we performed a functional enrichment analysis

388 using the Gene Ontology Biological Processes of their coding genes (Table S5). Clusters 1, 2, 3
389 and 6, which were predominantly activated at 12 DPI, were associated with antiviral, innate
390 immunity. Cluster 8 was primarily activated at 3 DPI and is associated with microtubule
391 organization, cell spreading and cell adhesion.

392 The down-regulated cluster 4 is associated with wound healing, cell-cell adhesion, and tissue
393 formation, similar to the functions enriched in the full DE list at 3 DPI. The co-expression
394 analyses showed that a significant number of genes associated with these functions were down-
395 regulated at both time points. Notably, this co-expressed transcript set contained 17% ncRNA,
396 constituting the highest proportion of all clusters. In cluster 4, the largest family of genes
397 associated with cell adhesion proteins encoded claudin tight junction proteins of the epithelium
398 (CLDN3, CLDN4, CLDN5, CLDN23). The rest encoded non-voltage-sensitive sodium channels
399 (SCNN1B, SCNN1G), extracellular matrix proteins (COL5A1, LAMC2) and a gap junction
400 protein (GJB2). The expression pattern and functional associations of these genes suggests their
401 involvement in the loss of epithelial integrity.

402 The early down-regulation of genes encoding claudins and other proteins involved in cell
403 adhesion (cluster 4) was accompanied by the up-regulation of genes associated with similar
404 functions (e.g. microtubule organization, cell spreading and cell adhesion) in cluster 8. However,
405 at 12 DPI, the expression of most of these genes had returned to near baseline, while the down-
406 regulation of cluster 4 persisted. Cluster 8 contains genes encoding the integrins ITFG2 and
407 ITGA1 and genes encoding ligands of integrin, collagen (COL11A1, COL12A1, COL24A1) and
408 laminin (LAMA2, LAMA4, LAMB1, LAMC1). Several other genes in cluster 8 can affect the
409 structure of the cytoskeleton, including caveolin (CAV1, CAV2) and cofilin (CFL2).

Co-expression network analysis implicates pathological mechanisms of other gastrointestinal diseases

To view the internal correlation structure of cluster 4, we organized the transcripts into a co-expression network. This allowed us to identify particularly strong correlations between coding genes and ncRNA to make functional inferences. This also allowed us to identify *hub* transcripts that shared high correlations with a large number of other transcripts, which would imply that they play a central role in the function of the cluster. The network was constructed by connecting each transcript to the two other transcripts with which it shared the strongest correlations. To identify genes with roles in the loss of cell adhesion, we examined the hubs in the resulting network. Several of the hub genes had known associations with pathology of colorectal mucosa. The most notable example was ABCB1, a transporter protein that interacts with several drugs and which has been genetically associated with the failure of first-line protease inhibitors in HIV-infected patients (40). Other important hubs included RPS5, a ribosomal protein associated with colorectal cancer, and RNF186, a ring finger protein that has been genetically associated with ulcerative colitis. The most connected gene was GPRC5A, a G-protein coupled receptor protein which has been associated with epithelial cell differentiation, followed by EMP1, a tight junction protein (41, 42). The most connected ncRNA was an unannotated transcript, XLOC_045516, connected to four other coding genes and two ncRNA. The coding genes included a transporter protein (SLC5A10), an actin cytoskeleton reorganizer (VAV1), a signaling protein (PRKAR1) and one uncharacterized protein.

Because a gene (RNF186) associated with ulcerative colitis (UC) was a hub gene in cluster 4, we compared the eight co-expressed gene clusters to genes affected by this disease. UC is an

inflammatory disorder of the gastrointestinal tract, associated with compromised epithelial integrity and bacterial translocation (43). The purpose of this analysis was to examine what co-expressed gene clusters were shared with other pathological transcriptional perturbations of the gastrointestinal tract. We used four datasets comparing patients with ulcerative colitis (UC) with healthy controls. A meta-analysis of these datasets produced one set of genes that were up-regulated (N = 483) in all four datasets, and one set of genes that were down-regulated (N = 521) in all datasets. We then tested these two gene sets for enrichment among the co-expressed gene clusters. We found that the up-regulated genes were associated with cluster 1 ($p = 1.12 \times 10^{-19}$), cluster 2 ($p = 0.0459$), cluster 3 ($p = 8.59 \times 10^{-16}$), cluster 6 ($p = 8.30 \times 10^{-10}$) and cluster 8 ($p = 3.48 \times 10^{-5}$). By contrast, the down-regulated genes were only enriched in cluster 4 ($p = 0.0384$, see Figure 3B).

In short, the co-expression analysis showed that genes encoding cell adhesion proteins were highly overrepresented among down-regulated genes at 3 DPI and many of these genes are also down regulated in other pathological states associated with intestinal epithelial damage.

Immunohistochemistry and quantitative RT-PCR confirm the down regulation of tight junction genes

As a validation of the down regulated co-expression network, we examined the differential expression of ten transcripts, protein coding as well as noncoding, from the network using quantitative RT-PCR. Candidates were selected among the genes that showed the strongest down regulation, and occupied important positions in the co-expression network (determined by their number of interactions). The candidates included tight junction genes CLDN3, CLDN4, and EMP1, the extracellular matrix gene LAMC2, and the network hub gene GPRC5A. We also

454 included four known noncoding RNA, one unannotated noncoding RNA. In most cases, the
455 selected RNA showed down regulation in all animals at both time points (Figure 4A).

456 In the co-expression network, CLDN3 shared interactions with other cell adhesion proteins,
457 including CLDN4 and PDLIM2. These genes were significantly down-regulated compared to
458 baseline at 12 DPI and, to a lesser degree, at 3 DPI as well. To examine whether the trends in the
459 mRNA levels were indicative of protein levels, we performed immunohistochemistry analysis of
460 the tight junction protein CLDN3. This is a major tight junction protein (44) that has been
461 implicated in the loss of mucosal epithelial integrity in several gastrointestinal diseases,
462 including colorectal cancer (45), celiac disease (46) and chronic SIV infection (21).

463 The quantification of CLDN3 protein showed a strong decrease in two animals at 3 DPI,
464 compared to tissue from uninfected animals. By 12 DPI, the level of CLDN3 protein had
465 significantly decreased in all animals ($p = 0.00327$, Fig 4B, first panel). The variation in CLDN3
466 protein expression between individual animals could indicate that the 3 DPI time point represents
467 a transition period, during which the tissue is undergoing the first changes that lead to subsequent
468 loss of tissue integrity. During this time, differences in response kinetics cause a high degree of
469 heterogeneity between individuals, while the strong antiviral immune response at peak viremia
470 caused a consistent down regulation.

471 In uninfected animals, CLDN3 is expressed primarily in the mucosal surface facing the intestinal
472 lumen, and surrounding the crypts. In the two animals that showed loss of CLDN3 protein
473 expression at 3 DPI, as well as at 12 DPI, CLDN3 expression was diminished in all these
474 locations. (Figure 4B, second to fourth panels). Thus, the contribution of CLDN3, and possibly

475 other down regulated tight junction proteins, to tissue integrity can be lost mere days after SIV
476 infection.

477 Taken together, these results confirm that the earliest mucosal response to SIV infection includes
478 down regulation of several genes with important roles in cell adhesion and tissue integrity. This
479 down regulation is observable in both RNA levels and protein levels as early as at 3 DPI.

Discussion

Here, we describe the first whole-genome transcriptional profiling of rectal mucosa from SIV-infected nonhuman primates. By focusing on early time points after mucosal challenge, we aimed to characterize mechanisms that contribute to viral spread from the inoculation site. In short, we found that at 3 DPI, corresponding to the eclipse phase, there was a strong transcriptional response at the site of inoculation. This response did not show a significant association with immune activation. Instead, it was predominantly associated with cytoskeleton reorganization, cell morphology and cell adhesion. In particular, we observed a down regulation of genes encoding claudins, a class of tight junction proteins. At peak viremia, we observed a strong inflammatory response, involving both innate and adaptive immune functions, and antibacterial in addition to antiviral functions.

The first challenge presented to the virus is crossing the mucosal epithelium into the lamina propria where it gains access to CD4+ target cells. Any circumstance that causes damage to the epithelium can elevate the risk of SIV/HIV transmission, including minor wounds caused by sexual intercourse and preexisting inflammation (47). The rate of sexual transmission is normally low, which implies that undamaged mucosal epithelium is an effective barrier to HIV infection. HIV-positive patients and SIV-infected NHPs show significant damage to the mucosal epithelium, which leads to translocation of bacterial antigens from the gastrointestinal tract to the lamina propria, further enhancing protracted inflammation and immune cell depletion, and driving increased epithelial damage (48). Yet, it has not been determined whether this degenerative cycle is first triggered by the antiviral inflammatory response or by an alternative pathway. Here, we found activation of cytoskeletal rearrangement, cell adhesion and, in particular, down-regulation of tight junction proteins before immune activation. This study

supports the hypothesis that the virus affects epithelial integrity directly, within days of mucosal challenge.

Taken together, the findings of this study suggest a scenario whereby a low number of virions cross the thin rectal epithelium, giving them can access to CD4+ target cells in the mucosa. This triggers morphological and structural changes in the epithelium within days of infection, which includes the loss of cell adhesion by down-regulation of genes encoding tight junction proteins, in particular claudins. The co-expression network describing the down regulated genes associated with cell adhesion and tight junctions, showed that the central genes in this process are associated with other pathologies of intestinal mucosa. Specifically, the most interconnected *hub* genes in the network were associated with failure of protease inhibitors in HIV, colorectal cancer and ulcerative colitis.

Ulcerative colitis shares a number of disease mechanisms with SIV infection – inflammation in the intestinal mucosa leading to compromised epithelial integrity and bacterial translocation (43). This prompted us to compare genes that are affected by UC with genes affected by acute SIV infection. By this analysis, we found that expression changes in UC were primarily correlated with the down regulation of cell adhesion and tight junction proteins at 3 and 12 DPI, in addition to immune activation at 12 DPI.

While a separate set of genes associated with tissue integrity were activated at 3 DPI, even these cell adhesion molecules are down regulated at peak viremia, and when immune cells migrate to the site of infection, the tissue has already developed increased permeability. Thus, tissue inflammation facilitates viral spread by driving additional inflammatory activation and the recruitment of additional target cells. This scenario suggests that if the down-regulated

expression of claudins and other tight junction proteins could be prevented, the host would stand a better chance of restricting the virus to the primary infection compartment (and adjacent tissues) for a longer period of time. The potential benefit of delaying SIV infection is supported by previous studies that implicated slow CD8+ lymphocyte (6) and type I interferon (20, 49) responses in the loss of HIV/SIV control, by failing to manifest itself when the virus is the most vulnerable. Maintenance of mucosal integrity would provide additional opportunity for these antiviral mechanisms to come into play and increase the chance that the virus would be eliminated before it could infect enough host cells to spread systemically (Fig 5).

Aside from its role in epithelial tissue structure, the cytoskeleton plays an important role in intracellular events required for immunodeficiency virus infection; studies of HIV have shown that during engagement of HIV gp120 with the CD4 receptor, actin filaments participate in recruitment of co-receptors CCR5 and CXCR4 used by HIV for cellular entry (50-52). However, this local assimilation of polymerized actin can stabilize the plasma membrane and inhibit viral entry by endocytosis (53). To overcome this, HIV gp120 engagement to CXCR4 can induce cofilin to sever actin polymers, enabling invagination of the plasma membrane and endocytosis (53). The fact that alterations to the cytoskeleton can be both beneficial and detrimental to the virus could explain why disruption of the cytoskeleton during HIV infection *in vitro* has been shown to promote (54) or inhibit (50) viral infection. While these previous studies were carried out primarily in CD4+ cells, the strong enrichment in cytoskeletal functions in our study suggest that similar processes are activated by SIV in other cell types in the rectal mucosa, such as stromal cells, smooth muscle cells and adipocytes. Indeed, the transcriptional regulation of genes with cytoskeletal functions by HIV proteins has also been reported in DCs (55), and macrophages (56).

Together, our results emphasize the importance of mucosal integrity to control HIV/SIV infection. Through the down regulation of adhesion proteins, events leading to systemic infection take place within days of infection. The lack of immune activation at this time point suggests that future studies of vaccines and interventions should also focus on non-inflammatory mechanisms that could reduce viral replication rates in mucosal tissues. While this study highlights cell morphology and cell adhesion mechanisms in the loss of viral restriction, other mechanisms could be discovered by examining additional and still earlier time points. Furthermore, the insights gained from these studies of pathogenic SIV hosts should be compared to natural SIV hosts, to specifically explore the pathogenic aspects of this process. A previous study of sooty mengabeys chronically infected with SIV did not find signs of damaged epithelial mucosa (21), suggesting that the expression patterns of tight junction and cell adhesion genes would show different expression patterns during the acute phase.

These questions will be further investigated in future studies, where pathogenic hosts will be compared to non-pathogenic hosts. These studies also include other tissues, including lymph nodes, in addition to earlier time points, starting at day 1.

Acknowledgements

This work was supported by the National Institutes of Health, Office of the Director, P51OD010425, R24OD011157, and R24OD011172; NIAID Contract No. HHSN272201300010C; NIAID Simian Vaccine Evaluation Unit contract with the University of Washington, Contract No. N01-AI-60006; NIDDK, NIH R01 DK087625-01 (to Li, Q.) and by Preclinical Research & Development Branch, VRP, DAIDS, NIAID, NIH, Task Order under N01-AI-30018 (to Li, Q), DAIDS Reagent Resource Support Program for AIDS Vaccine Development, Quality Biological, Gaithersburg, Maryland, Division of AIDS Contract N01-A30018; and the Swedish Research Council, D0045701.

References

1. **Haase AT.** 2010. Targeting early infection to prevent HIV-1 mucosal transmission. Nature **464:**217-223.
2. **Smith DK, Grohskopf LA, Black RJ, Auerbach JD, Veronese F, Struble KA, Cheever L, Johnson M, Paxton LA, Onorato IM, Greenberg AE.** 2005. Antiretroviral postexposure prophylaxis after sexual, injection-drug use, or other nonoccupational exposure to HIV in the United States: recommendations from the U.S. Department of Health and Human Services. MMWR Recomm Rep **54:**1-20.
3. **Otten RA, Smith DK, Adams DR, Pullium JK, Jackson E, Kim CN, Jaffe H, Janssen R, Butera S, Folks TM.** 2000. Efficacy of postexposure prophylaxis after intravaginal exposure of pig-tailed macaques to a human-derived retrovirus (human immunodeficiency virus type 2). J Virol **74:**9771-9775.
4. **McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF.** 2010. The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol **10:**11-23.
5. **Zhang Z, Schuler T, Zupancic M, Wietgrefe S, Staskus KA, Reimann KA, Reinhart TA, Rogan M, Cavert W, Miller CJ, Veazey RS, Notermans D, Little S, Danner SA, Richman DD, Havlir D, Wong J, Jordan HL, Schacker TW, Racz P, Tenner-Racz K, Letvin NL, Wolinsky S, Haase AT.** 1999. Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. Science **286:**1353-1357.
6. **Reynolds MR, Rakasz E, Skinner PJ, White C, Abel K, Ma ZM, Compton L, Napoe G, Wilson N, Miller CJ, Haase A, Watkins DI.** 2005. CD8+ T-lymphocyte response to major immunodominant epitopes after vaginal exposure to simian immunodeficiency virus: too late and too little. J Virol **79:**9228-9235.
7. **Muro-Cacho CA, Pantaleo G, Fauci AS.** 1995. Analysis of apoptosis in lymph nodes of HIV-infected persons. Intensity of apoptosis correlates with the general state of activation of the lymphoid tissue and not with stage of disease or viral burden. J Immunol **154:**5555-5566.
8. **Gordon SN, Klatt NR, Bosinger SE, Brenchley JM, Milush JM, Engram JC, Dunham RM, Paiardini M, Klucking S, Danesh A, Strobert EA, Apetrei C, Pandrea IV, Kelvin D, Douek DC, Staprans SI, Sodora DL, Silvestri G.** 2007. Severe depletion of mucosal CD4+ T cells in AIDS-free simian immunodeficiency virus-infected sooty mangabeys. J Immunol **179:**3026-3034.

9. **Pandrea IV, Gautam R, Ribeiro RM, Brenchley JM, Butler IF, Pattison M, Rasmussen T, Marx PA, Silvestri G, Lackner AA, Perelson AS, Douek DC, Veazey RS, Apetrei C.** 2007. Acute loss of intestinal CD4+ T cells is not predictive of simian immunodeficiency virus virulence. J Immunol **179:**3035-3046.

10. **Hunt PW, Cao HL, Muzoora C, Ssewanyana I, Bennett J, Emenyonu N, Kembabazi A, Neilands TB, Bangsberg DR, Deeks SG, Martin JN.** 2011. Impact of CD8+ T-cell activation on CD4+ T-cell recovery and mortality in HIV-infected Ugandans initiating antiretroviral therapy. AIDS **25:**2123-2131.

11. **Sandler NG, Wand H, Roque A, Law M, Nason MC, Nixon DE, Pedersen C, Ruxrungtham K, Lewin SR, Emery S, Neaton JD, Brenchley JM, Deeks SG, Sereti I, Douek DC.** 2011. Plasma levels of soluble CD14 independently predict mortality in HIV infection. J Infect Dis **203:**780-790.

12. **Borrow P.** 2011. Innate immunity in acute HIV-1 infection. Curr Opin HIV AIDS **6:**353-363.

13. **Varghese B, Maher JE, Peterman TA, Branson BM, Steketee RW.** 2002. Reducing the risk of sexual HIV transmission: quantifying the per-act risk for HIV on the basis of choice of partner, sex act, and condom use. Sex Transm Dis **29:**38-43.

14. 1992. Comparison of female to male and male to female transmission of HIV in 563 stable couples. European Study Group on Heterosexual Transmission of HIV. BMJ **304:**809-813.

15. **Li Q, Estes JD, Schlievert PM, Duan L, Brosnahan AJ, Southern PJ, Reilly CS, Peterson ML, Schultz-Darken N, Brunner KG, Nephew KR, Pambuccian S, Lifson JD, Carlis JV, Haase AT.** 2009. Glycerol monolaurate prevents mucosal SIV transmission. Nature **458:**1034-1038.

16. **Li Q, Estes JD, Duan L, Jessurun J, Pambuccian S, Forster C, Wietgrefe S, Zupancic M, Schacker T, Reilly C, Carlis JV, Haase AT.** 2008. Simian immunodeficiency virus-induced intestinal cell apoptosis is the underlying mechanism of the regenerative enteropathy of early infection. J Infect Dis **197:**420-429.

17. **Nazli A, Chan O, Dobson-Belaire WN, Ouellet M, Tremblay MJ, Gray-Owen SD, Arsenault AL, Kaushic C.** 2010. Exposure to HIV-1 directly impairs mucosal epithelial barrier integrity allowing microbial translocation. PLoS Pathog **6:**e1000852.

18. **Triboulet R, Mari B, Lin YL, Chable-Bessia C, Bennasser Y, Lebrigand K, Cardinaud B, Maurin T, Barbry P, Baillat V, Reynes J, Corbeau P, Jeang KT, Benkirane M.** 2007. Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science **315:**1579-1582.

19. **Zhang Q, Chen CY, Yedavalli VS, Jeang KT.** 2013. NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. MBio **4:**e00596-00512.
20. **Chang ST, Sova P, Peng X, Weiss J, Law GL, Palermo RE, Katze MG.** 2011. Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line. MBio **2**.
21. **Estes JD, Harris LD, Klatt NR, Tabb B, Pittaluga S, Paiardini M, Barclay GR, Smedley J, Pung R, Oliveira KM, Hirsch VM, Silvestri G, Douek DC, Miller CJ, Haase AT, Lifson J, Brenchley JM.** 2010. Damaged intestinal epithelial integrity linked to microbial translocation in pathogenic simian immunodeficiency virus infections. PLoS Pathog **6:**e1001052.
22. **Polacino P, Cleveland B, Zhu Y, Kimata JT, Overbaugh J, Anderson D, Hu SL.** 2007. Immunogenicity and protective efficacy of Gag/Pol/Env vaccines derived from temporal isolates of SIVmne against cognate virus challenge. J Med Primatol **36:**254-265.
23. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10:**R25.
24. **Trapnell C, Pachter L, Salzberg SL.** 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:**1105-1111.
25. **Anders S, Theodor Pyl P, Huber W.** 2014. HTSeq — A Python framework to work with high-throughput sequencing data. bioRxiv preprint.
26. **Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL.** 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev **25:**1915-1927.
27. **Pipes L, Li S, Bozinoski M, Palermo R, Peng X, Blood P, Kelly S, Weiss JM, Thierry-Mieg J, Thierry-Mieg D, Zumbo P, Chen R, Schroth GP, Mason CE, Katze MG.** 2013. The non-human primate reference transcriptome resource (NHPRTR) for comparative functional genomics. Nucleic Acids Res **41:**D906-914.
28. **Peng X, Pipes L, Xiong H, Green R, Jones D, Ruzzo W, Schroth G, Mason C, Palermo R, Katze M.** 2014. Assessment and improvement of Indian-origin rhesus macaque and Mauritian-origin cynomolgus macaque genome annotations using deep transcriptome sequencing data. Journal of Medical Primatology **In press**.
29. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078-2079.
30. **Anders S, Huber W.** 2010. Differential expression analysis for sequence count data. Genome Biol **11:**R106.

31. **Kramer A, Green J, Pollard J, Jr., Tugendreich S.** 2014. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics.

32. **Song L, Langfelder P, Horvath S.** 2012. Comparison of co-expression measures: mutual information, correlation, and model based indices. BMC Bioinformatics **13:**328.

33. **Ward JH.** 1963. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc **58:**236-&.

34. **Langfelder P, Zhang B, Horvath S.** 2008. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics **24:**719-720.

35. **Shimodaira H.** 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol **51:**492-508.

36. **Ruan J, Dean AK, Zhang W.** 2010. A general co-expression network-based approach to gene expression analysis: comparison and applications. Bmc Syst Biol **4:**8.

37. **Vazquez N, Schmeisser H, Dolan MA, Bekisz J, Zoon KC, Wahl SM.** 2011. Structural variants of IFNalpha preferentially promote antiviral functions. Blood **118:**2567-2577.

38. **Barr SD, Smiley JR, Bushman FD.** 2008. The interferon response inhibits HIV particle production by induction of TRIM22. PLoS Pathog **4:**e1000007.

39. **Espert L, Degols G, Lin YL, Vincent T, Benkirane M, Mechti N.** 2005. Interferon-induced exonuclease ISG20 exhibits an antiviral activity against human immunodeficiency virus type 1. J Gen Virol **86:**2221-2229.

40. **Coelho AV, Silva SP, de Alencar LC, Stocco G, Crovella S, Brandao LA, Guimaraes RL.** 2013. ABCB1 and ABCC1 variants associated with virological failure of first-line protease inhibitors antiretroviral regimens in Northeast Brazil patients. J Clin Pharmacol **53:**1286-1293.

41. **Bangsow T, Baumann E, Bangsow C, Jaeger MH, Pelzer B, Gruhn P, Wolf S, von Melchner H, Stanimirovic DB.** 2008. The epithelial membrane protein 1 is a novel tight junction protein of the blood-brain barrier. J Cereb Blood Flow Metab **28:**1249-1260.

42. **Lee HS, Sherley JL, Chen JJ, Chiu CC, Chiou LL, Liang JD, Yang PC, Huang GT, Sheu JC.** 2005. EMP-1 is a junctional protein in a liver stem cell line and in the liver. Biochem Biophys Res Commun **334:**996-1003.

43. **Johansson ME, Gustafsson JK, Holmen-Larsson J, Jabbar KS, Xia L, Xu H, Ghishan FK, Carvalho FA, Gewirtz AT, Sjovall H, Hansson GC.** 2014. Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. Gut **63:**281-291.

44. **Matsuda M, Kubo A, Furuse M, Tsukita S.** 2004. A peculiar internalization of claudins, tight junction-specific adhesion molecules, during the intercellular movement of epithelial cells. J Cell Sci **117:**1247-1257.

45. **de Oliveira SS, de Oliveira IM, De Souza W, Morgado-Diaz JA.** 2005. Claudins upregulation in human colorectal cancer. FEBS Lett **579:**6179-6185.

46. **Szakal DN, Gyorffy H, Arato A, Cseh A, Molnar K, Papp M, Dezsofi A, Veres G.** 2010. Mucosal expression of claudins 2, 3 and 4 in proximal and distal part of duodenum in children with coeliac disease. Virchows Arch **456:**245-250.

47. **Royce RA, Sena A, Cates W, Jr., Cohen MS.** 1997. Sexual transmission of HIV. N Engl J Med **336:**1072-1078.

48. **Marchetti G, Tincati C, Silvestri G.** 2013. Microbial translocation in the pathogenesis of HIV infection and AIDS. Clin Microbiol Rev **26:**2-18.

49. **Palermo RE, Tisoncik-Go J, Korth MJ, Katze MG.** 2013. Old world monkeys and new age science: the evolution of nonhuman primate systems virology. ILAR J **54:**166-180.

50. **Jolly C, Kashefi K, Hollinshead M, Sattentau QJ.** 2004. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. J Exp Med **199:**283-293.

51. **Jimenez-Baranda S, Gomez-Mouton C, Rojas A, Martinez-Prats L, Mira E, Ana Lacalle R, Valencia A, Dimitrov DS, Viola A, Delgado R, Martinez AC, Manes S.** 2007. Filamin-A regulates actin-dependent clustering of HIV receptors. Nat Cell Biol **9:**838-846.

52. **Iyengar S, Hildreth JE, Schwartz DH.** 1998. Actin-dependent receptor colocalization required for human immunodeficiency virus entry into host cells. J Virol **72:**5251-5255.

53. **Yoder A, Yu D, Dong L, Iyer SR, Xu X, Kelly J, Liu J, Wang W, Vorster PJ, Agulto L, Stephany DA, Cooper JN, Marsh JW, Wu Y.** 2008. HIV envelope-CXCR4 signaling activates cofilin to overcome cortical actin restriction in resting CD4 T cells. Cell **134:**782-792.

54. **Campbell EM, Nunez R, Hope TJ.** 2004. Disruption of the actin cytoskeleton can complement the ability of Nef to enhance human immunodeficiency virus type 1 infectivity. J Virol **78:**5745-5755.

55. **Wang JH, Wells C, Wu L.** 2008. Macropinocytosis and cytoskeleton contribute to dendritic cell-mediated HIV-1 transmission to CD4+ T cells. Virology **381:**143-154.

56. **Jouvenet N, Neil SJ, Bess C, Johnson MC, Virgen CA, Simon SM, Bieniasz PD.** 2006. Plasma membrane is the site of productive HIV-1 particle assembly. PLoS Biol **4:**e435.

Figure captions

Figure 1

FIG 1. (A) Experimental design. Eight animals were intrarectally inoculated with SIVmac251. Baseline samples were obtained by rectal pinch biopsy and blood draw 14 days prior to inoculation. Four animals were sacrificed at 3 DPI and 4 animals at 12 DPI. (B) Quantification of viral load by quantitative RT-PCR in plasma (log_{10} SIV RNA copies/mL, left panel), or rectal mucosa (log_{10} SIV RNA copies/µg total RNA, center panel) and of viral reads in rectal mucosa by deep RNA sequencing (log_{10} SIV reads, right panel). The X-axis represents the time point (days post inoculation); viral load is represented on the Y axis. Each animal is shown as an individual bar. Viremia was detectable after 6 days. By contrast, viral reads were detected in rectal mucosa by day 3 in all animals by deep RNA sequencing.

Figure 2

FIG 2. (A) Differential expression analysis at 3 DPI and 12 DPI. Differentially expressed genes were defined as having an adjusted p-value <0.05 and an absolute $\log_2$ fold change >1.5. Each column corresponds to an individual animal. Colors represent the fold changes of each gene. The upper section includes genes that were DE at 3 DPI only, the middle section includes genes that were DE at both time points, and the lower section includes genes that were DE at 12 DPI only. The middle section includes 888 genes. (B) Functional enrichment analysis, showing representative biological functions from the most enriched functional categories, shown on the right side. Circle sizes indicate the total number of differentially expressed genes in each function; color intensities indicate enrichment significance. The functional enrichment also distinguished between genes that activate and inhibit each function. Where the functional enrichment predicts activation or inhibition, this is indicated by an upward or downward pointer respectively. (C) Canonical pathway enrichment. Unlike the biological functions, canonical pathways describe signaling cascades that are activated by extracellular signals.

761 **Figure 3**

762 FIG 3. (A) Heatmaps giving an overview of differentially expressed ncRNA by three different
763 mappings: ncRNA annotated in Ensembl, macaque homologues of human long ncRNA and
764 previously unannotated ncRNA from the nonhuman primate reference transcriptome project. (B)
765 Co-expression analysis of differentially expressed coding RNA and ncRNA across all 19
766 biological samples. Differentially expressed transcripts formed eight co-expressed clusters. The
767 heatmap shows average $\log_2$ fold changes at the two time points compared to baseline. The
768 squares show enrichment of co-expressed gene sets among up or down regulated genes in
769 ulcerative colitis. Cluster 4, which was uniquely down-regulated at both time points, was
770 organized into a co-expression network for a more detailed view of its internal correlation
771 structure. Each gene was connected to the gene with which it shared the highest correlation (solid
772 line) and second highest (dashed line). Line thickness is dependent on correlation coefficient.
773 Nodes are color coded according to RNA class and sized according to their number of
774 interactions; color intensity is dependent on fold change at 3 DPI.

775

776

Figure 4

FIG 4: Quantitative RT-PCR and immunohistochemistry analysis of down regulated cell adhesion genes in rectal mucosa. (A) Validation of mRNAseq results using quantitative RT-PCR. 10 RNA, 5 coding 4 known noncoding and 1 unannotated noncoding RNA were analyzed. (B) First panel: Quantification of CLDN3 expression, measured in positive pixels/μm^2. CLDN3 showed strong down regulation at 3 DPI in two animals, compared to uninfected controls. At 12 DPI, CLDN3 was strongly suppressed in all four animals. Second-fourth panel: Representative slides showing CLDN3 (stained brown) in rectal necropsies from unchallenged animals, an animal that showed down regulation of CLDN3 at 3 DPI, and 12 DPI.

Figure 5

FIG 5: Hypothetical description of events leading up to systemic SIV infection. The events described in this study (red line) begin with down regulation of genes maintaining epithelial integrity in mucosal tissues, within days of exposure to SIV. This response is accompanied by a low rate of viral replication in rectal mucosa. When immune activation take place, the rectal mucosa is already compromised and tissue integrity is quickly lost, leading to microbial translocation, increased tissue damage, recruitment of SIV target cells and accelerated SIV replication. This process allows the virus to infect a large number of CD4+ cells, leading to systemic infection. This suggests that if expression of cell adhesion proteins (and mucosal integrity) could be maintained before inflammation is triggered (blue line), tissue damage, microbial translocation and recruitment of SIV target cells could be limited. Thereby, viral replication would be slower, delaying and perhaps allowing for the prevention of systemic infection.