

Host cell glycosylation selects for infection with CCR5- versus CXCR4-tropic HIV-1

Hannah L. Itell^{1,2}, Daryl Humes^{2,3}, Nell E. Baumgarten², Julie Overbaugh^{2*}

¹ Molecular and Cellular Biology PhD Program, University of Washington, Seattle, WA, 98109, USA

² Division of Human Biology, Fred Hutchinson Cancer Center, Seattle, WA, 98109, USA

³ Present address: Tr1X Inc, La Jolla, CA, 92037, USA

* Corresponding Author: joverbau@fredhutch.org (J.O.)

ABSTRACT

HIV-1 infection involves a selection bottleneck that leads to transmission of one or a few HIV variants, which nearly always use CCR5 as the coreceptor (R5 viruses) for viral entry as opposed to CXCR4 (X4 viruses). The host properties that drive this selection are not well understood and may hold keys to factors that govern HIV susceptibility. In this report, we identified SLC35A2, a transporter of UDP-galactose, as a candidate X4-specific restriction factor in CRISPR-knockout screens in primary target CD4⁺ T cells. SLC35A2 inactivation in CD4⁺ T cells, which resulted in truncation of glycans due to the absence of galactose, not only increased X4 infection levels, but also consistently decreased infection levels of R5 HIV strains. Single cycle infections demonstrated that the effect is host cell dependent. SLC35A2 is expressed in CD4⁺ T cells at different tissue sites, with high levels in the genital tract – the site of most HIV infections. These data support a role for a host cell protein that regulates glycan structure on HIV infection, with enhanced R5 infection but reduced X4 infection associated with SLC35A2-mediated glycosylation. Host cell glycosylation may therefore contribute to R5 selection and host susceptibility during HIV transmission.

MAIN TEXT

HIV-1 undergoes a severe viral population bottleneck during transmission. HIV's high mutation rate coupled with host selective pressures leads to a large diversity of viral variants circulating during chronic infection, yet only one or a few viral variants are detected soon after infection¹⁻⁶. While transmitted viruses are thought to be successful in part due to their high replicative fitness^{7,8} and reduced susceptibility to interferon (IFN)-mediated inhibition^{9,10}, the most consistent, defining feature of transmitted variants is their strong coreceptor tropism preferences. HIV can use either CCR5 or CXCR4 as a coreceptor for viral entry¹¹, yet transmitted variants nearly always use CCR5^{6,12} despite broad expression of CXCR4 at transmission sites¹³, with CXCR4-tropic (X4) viruses arising in the months to years after primary infection¹⁴. It has remained a longstanding question in the HIV field why CCR5-tropic (R5) variants, but not X4 viruses, can successfully navigate this bottleneck, leading to new infections. In this study, we aimed to identify host factors that preferentially inhibit X4 HIV, as these may represent selective pressures that contribute to R5 selection during HIV transmission and thus to HIV susceptibility.

To identify X4-specific restriction factors, we performed CRISPR-knockout (KO) screens in the main target cells of HIV infection *in vivo*, primary CD4⁺ T cells, and infected with an X4 strain (LAI¹⁵) and an R5 virus (Q23.BG505¹⁶) for comparison. We leveraged the HIV-CRISPR screening strategy¹⁷, which we recently adapted for use in primary T cells¹⁸, and interrogated ~2,000 host genes, including many IFN-stimulated genes (ISGs) as they are a major driver of antiviral immunity (**Fig. 1A; Supplementary Table 1**). Screens were therefore performed in the context of type I IFN treatment. Genes required for IFN signaling were enriched for all screens and we observed strong agreement between screens using primary T cells from different donors (**Fig. 1B; Supplementary Table 2**). By comparing Q23.BG505 and LAI screens and selecting LAI-specific hits, we identified 81 candidate X4-specific restriction factors, of which the top-scoring hit was SLC35A2 (**Fig. 1C**). This gene was of particular interest because it was highly enriched for both LAI screens and conversely scored below background for both Q23.BG505 screens (**Fig. 1D**). SLC35A2 was not upregulated by IFN in primary CD4⁺ T cells (**Fig. 1E**). These data support SLC35A2's candidacy as a non-ISG host factor that may preferentially inhibit X4 HIV.

To validate our CRISPR-KO screen findings, we performed single gene knockouts in primary CD4⁺ T cells, inactivating SLC35A2 and the B cell marker CD19 as a negative control. Across several unique donors

and independent experiments, we found that SLC35A2 KO not only significantly increased infection levels of X4-tropic LAI (5.9-fold decrease, $p=0.004$), but, surprisingly, it also significantly decreased those of R5-tropic Q23.BG505 (4.8-fold increase, $p<0.0001$) (**Fig. 2A**). These findings were consistent for two separate assays of HIV infection: quantifying reverse transcriptase (RT) activity in infection supernatants (**Fig. 2A, Extended Data Fig. 1A**) and measuring the percentage of cells staining positive for HIV-Gag (**Fig. 2B, Extended Data Fig. 1B, C**). Notably, based on Gag staining, an average of 73% of SLC35A2 KO cells across four unique donors (range: 64.8-83.2%) were infected with LAI compared to 13% in CD19 KO control cells (range: 9.6-16%) (5.9-fold decrease, $p=0.001$) (**Fig. 2B, Extended Data Fig. 1C**). The opposite trend was true for Q23.BG505, with 1.8% of SLC35A2 KO cells (range: 1-2.8%) staining positive for HIV-Gag compared to 11% of CD19 KO cells (range: 7.8-16.4%) (7.6-fold increase, $p=0.03$). SLC35A2 KO therefore causes substantial, opposing phenotypes for these two HIV strains that differ in their coreceptor usage.

To determine whether the observed phenotypes were specific to SLC35A2 inactivation and not due to off-target effects, we exogenously expressed SLC35A2 cDNA in SLC35A2 KO CD4⁺ T cells. Complementation led to full or partial rescue of wildtype HIV infection levels (**Fig. 2C**). Cases of partial rescue are likely due to the polyclonal nature of both SLC35A2 KO and complementation, as these primary cell culture experiments are not amenable to antibiotic or clonal selection. These data support the specificity of SLC35A2 KO.

Though the opposing SLC35A2 KO phenotypes for Q23.BG505 and LAI were consistent across experimental platforms, donors, and infection metrics, it is unclear what biological viral characteristics are driving these effects. One major difference between these viruses is their coreceptor usage; however, they are also from different HIV clades and LAI is lab-adapted whereas Q23.BG505 is not. To understand which viral features represent determinants of the phenotype of SLC35A2 KO on HIV infection, we infected SLC35A2 KO and CD19 KO CD4⁺ T cells with five X4 and five R5 full-length HIV strains from various clades and consisting of both primary and lab-adapted viruses. We observed a strikingly consistent pattern – SLC35A2 KO increased infection levels for all X4 strains whereas all R5 viruses exhibited lower levels of infection (**Fig. 2D, Extended Data Fig. 2A**). Coreceptor tropism for each strain was confirmed by individually inactivating CXCR4 and CCR5 in CD4⁺ T cells (**Extended Data Fig. 2B**). In addition to X4 and R5 HIV, we investigated the effect of SLC35A2 KO on two viruses considered dual-tropic (93BR020¹⁹ and 89.6²⁰). These

results closely recapitulated those of X4 HIV (**Fig. 2D, Extended Data Fig. 2A**), which agrees with our findings from coreceptor KO cells that indicated that both 93BR020 and 89.6 rely on CXCR4 substantially more than CCR5 (**Extended Data Fig. 2B**). Inactivation of SLC35A2 therefore differentially impacts HIV based on whether viruses use CXCR4 or CCR5 as a coreceptor for viral entry.

SLC35A2 is a multichannel membrane protein that transports UDP-galactose from the cytosol into the Golgi^{21,22}. This process is critical for normal glycosylation as galactose is a common residue on both N- and O-glycans. Therefore, SLC35A2 inactivation results in truncated glycans that lack galactose²³. To determine whether SLC35A2 KO disrupts glycosylation in our primary cell system, we stained SLC35A2 KO CD4⁺ T cells with two lectins that bind terminal glycan residues that would be expected to be exposed with the loss of galactose (**Fig. 3A**). Specifically, *Griffonia Simplicifolia* lectin II (GSL-II) preferentially binds terminal GlcNAcs on N-glycans and *Vicia Villosa* lectin (VVL) recognizes terminal GalNAcs, which are residues are specific to O-glycans²⁴. We found that 94% and 91% of SLC35A2 KO cells were positive for GSL-II and VVL binding, respectively, compared to 3% and 1% of CD19 KO cells (**Fig. 3B**). Therefore, despite the polyclonal nature of our KOs, editing of SLC35A2 in this cell type substantially disrupts host cell surface glycosylation by causing the truncation of both N- and O-glycans.

Though lectin staining demonstrates altered glycosylation on target cells, HIV itself is heavily glycosylated and viral glycans are likely similarly altered upon replication in SLC35A2 KO cells. Because our infections have thus far utilized replication-competent HIV, the observed tropism-specific SLC35A2 KO HIV phenotypes may be driven by altered glycans either on host cells or on HIV virions after at least one round of replication in SLC35A2 KO cells. To distinguish these possibilities, we made X4 and R5 HIV pseudoviruses, which are capable of infection but not making infectious progeny virions. We found that SLC35A2 KO impacted these HIV pseudoviruses within a single infection cycle in the same manner as replication-competent virus, with a reduction in R5 infection and an increase in X4 virus infection (**Fig. 3C**). The magnitudes of these effects (BG505 average: 10.9-fold decrease, LAI average: 3-fold increase) were similar to those from replication-competent infection (**Fig. 2B**). The effect of SLC35A2 KO on HIV is therefore driven by changes to glycans on target cells, not on progeny virus.

These data support an opposing, early role for CD4⁺ T cell glycans on R5 and X4 HIV infection, whereby R5 HIV prefers wildtype target cell glycosylation and X4 viruses are more successful when host

glycans are truncated, such as in the setting of SLC35A2 KO (**Fig. 3D**). Interestingly, these results may explain the cellular driver for recent observations by Ma *et al.* showing that memory CD4⁺ T cells from the blood and tonsils that have higher levels of binding to WGA, a lectin that binds sialic acid and GlcNAc, are more susceptible to R5 HIV infection, whereas removal of sialic acid reduced R5 infection²⁵. Together, our findings suggest that sialic acid residues, which are absent in the setting of SLC35A2 KO, may specifically potentiate R5 HIV infection of CD4⁺ T cells and identify a specific host gene central to this process.

Our results not only provide insight into determinants of R5 infection; they also uniquely demonstrate that target cell glycan features differentially impact HIV infection based on coreceptor tropism. Based on the existing literature and our working model (**Fig. 3D**), if host glycans were to influence R5 selection during HIV transmission, we would expect there to be high levels of SLC35A2 and sialic acid on target cells that HIV first encounters in the mucosa. To address this, we analyzed a previously published bulk RNA-seq dataset of CD4⁺ T cells isolated from the blood or three female genital tract sites²⁶ and found that *SLC35A2* was expressed in all CD4⁺ T cell populations, with elevated levels in CD4⁺ T cells from the vaginal tract, a common site of HIV transmission, as compared those in the blood (**Fig. 3E**). Notably, CXCR4 was highly expressed in CD4⁺ T cells from all tissue sites in this dataset, particularly in mucosal tissues, consistent with previous findings¹³ that an absence of CXCR4 expression on relevant target cells does not explain R5 selection during transmission (**Extended Data Figure 3**). In accordance with the *SLC35A2* expression level trends, memory CD4⁺ T cells from the endometrium were reported to have higher levels of WGA lectin binding (sialic acid/GlcNAc) than those from the blood²⁵. Target cells in HIV transmission sites therefore have SLC35A2 levels and glycan features that support infection of R5 viruses and may restrict that of X4 viruses. Host glycosylation may therefore contribute to R5 selection during the HIV transmission bottleneck.

It remains to be determined whether the opposing tropism phenotypes with SLC35A2 KO are due to changes to specific glycans on target cells, such as glycans on the coreceptors themselves, or are due to broad changes in the overall glycan landscape of target cells. CXCR4 and CCR5 both harbor glycans on their extracellular N-termini²⁷⁻²⁹, and this region of CCR5 is known to interact with HIV³⁰. It is therefore possible that the impact of SLC35A2 is on the composition and structure of glycans on the coreceptors, with subsequent coreceptor-dependent effects on HIV binding and/or fusion. However, CD4⁺ T cells express a variety of glycoproteins and a recent study concluded that general glycan-glycan interactions between HIV and the host

cell may enhance attachment³¹. These glycan-glycan interactions could also be a mechanism through which SLC35A2 impacts HIV infection, though there is at present no evidence for a tropism-specific effect with these broad glycan-glycan interactions. Overall, our findings provide a window into the mechanism behind the tropism-dependent impact of target cell glycosylation on HIV. Further, they address a long-standing question in the field that may inform the development of new prophylactic approaches to specifically target viral variants that are successful in the HIV transmission bottleneck and thus a key to HIV susceptibility.

MAIN FIGURES

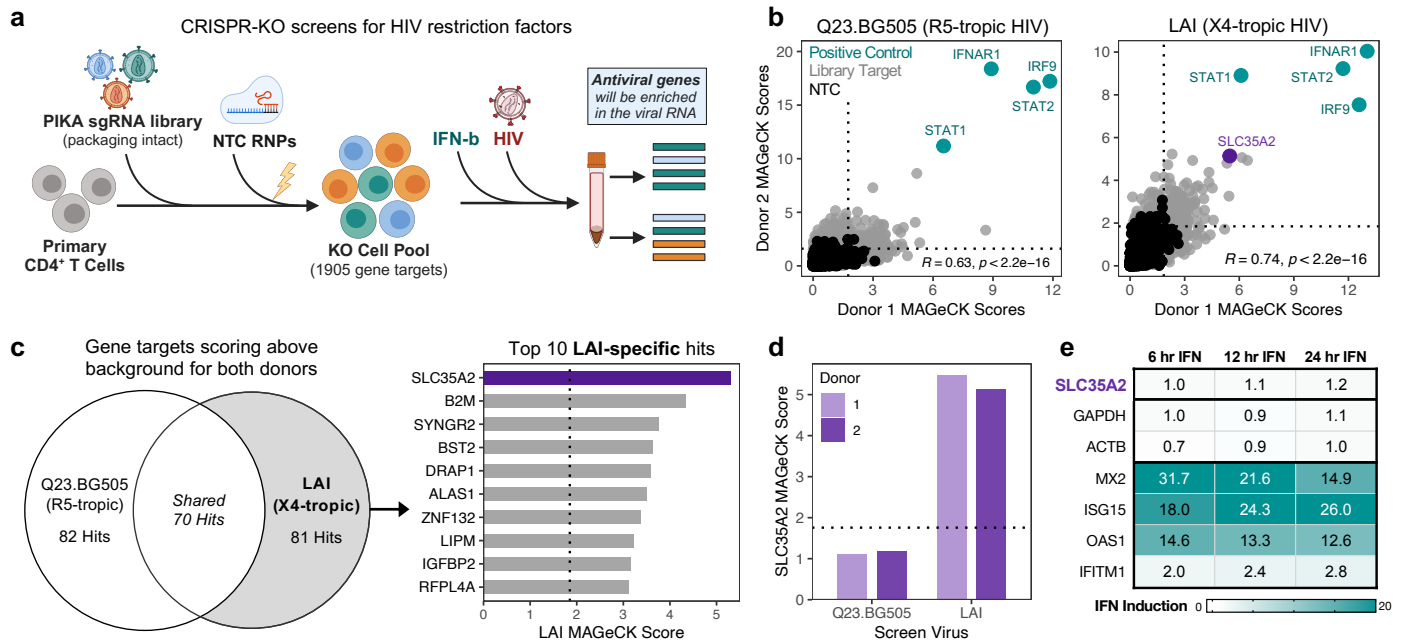


Figure 1. SLC35A2 is an X4-specific hit in a CRISPR-KO screen for HIV restriction factors in primary CD4⁺ T cells.

A. Schematic of HIV-CRISPR screening approach.

B. Correlation of positive MAGeCK scores between CD4⁺ T cell donors in Q23.BG505 (left) and LAI (right) screens. Pearson correlation results are included within each plot. Dotted lines denote background for each screen based on non-targeting control (NTC) results (background = average NTC + 3 standard deviations).

C. Left, Euler diagram comparing hits between Q23.BG505 and LAI screens. Screen hits are defined as gene targets that scored above background for both donors (top right quadrants from Panel B). Right, top 10 scoring LAI-specific hits, ordered by their average MAGeCK score in LAI screens. Dotted line reflects average background from LAI screens.

D. MAGeCK scores for SLC35A2 in each HIV-CRISPR screen. Dotted line indicates average background for all screens.

E. Heatmap of average IFN-induced fold changes in RNA expression for SLC35A2, housekeeping genes (GAPDH, ACTB), and canonical ISGs (MX2, ISG15, OAS1, IFITM1) in primary CD4⁺ T cells at 6-, 12-, and 24-hours post-IFN treatment (n = 3, 3, and 5 donors, respectively by time point).

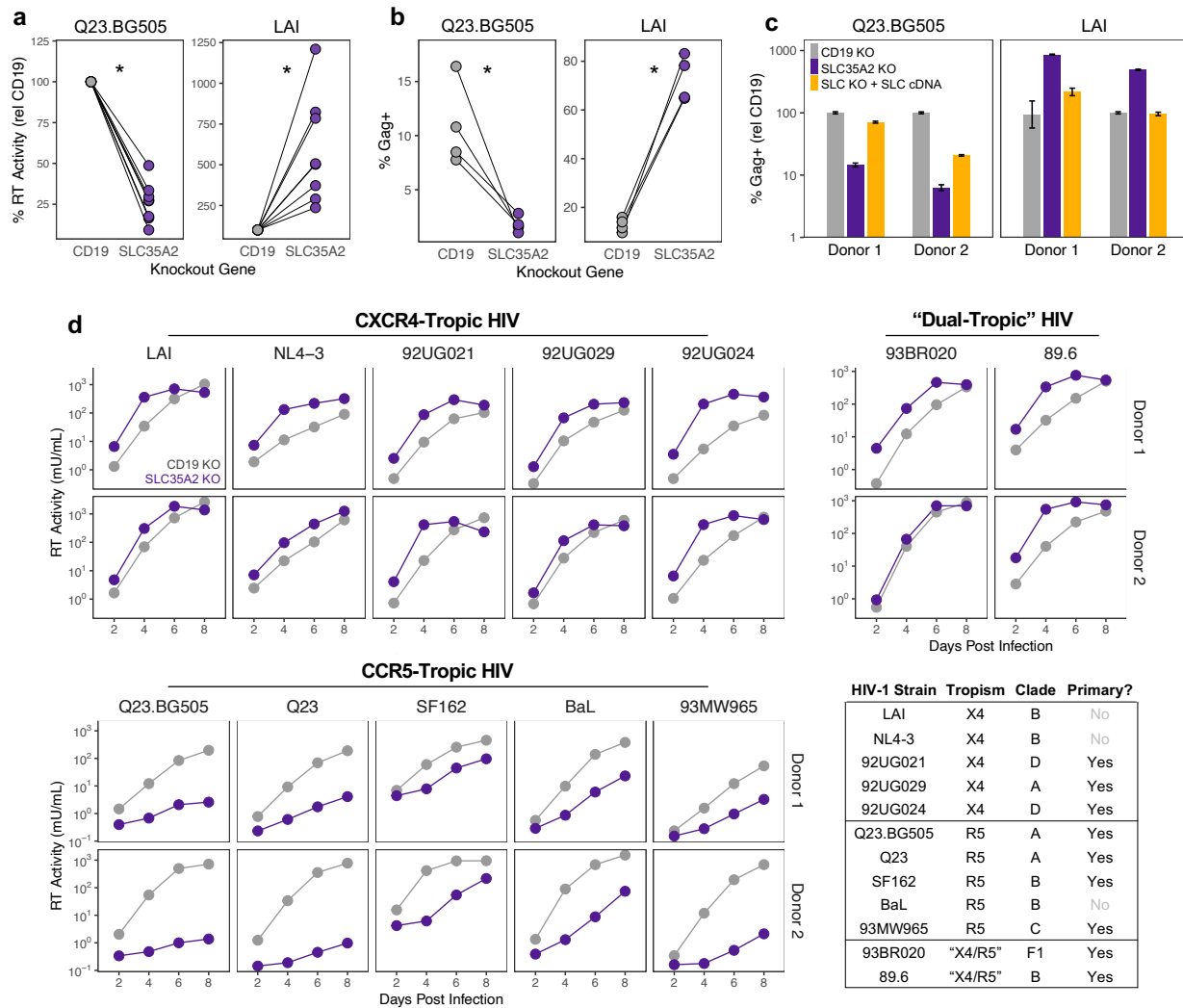


Figure 2. SLC35A2 KO differentially impacts CXCR4-tropic and CCR5-tropic HIV.

A. Reverse transcriptase (RT) activity relative to CD19 KO at 2 days post infection (dpi) (MOI=0.02). Knockouts were performed in CD4⁺ T cells from five donors across single or replicate knockout and infection experiments for a total of 8 independent experiments. Editing efficiency ranged from 52% to 96% KO (82% KO average, n=8). Two-tailed paired t-test, asterisk indicates p<0.05.

B. Percentage of CD4⁺ T cells staining positive for HIV-Gag at 3 dpi (MOI=1) (n=4 independent donors). Two-tailed paired t-test, asterisk indicates p<0.05.

C. Percentage of CD4⁺ T cells staining positive for HIV-Gag at 3 dpi (MOI=1), relative to CD19 KO. Data are shown as mean ± standard deviation of duplicate infections.

D. RT activity over time from spreading infections (MOI=0.02) in primary CD4⁺ T cells from two donors. Results are separated based on coreceptor tropism. Bottom right table, characteristics of the HIV strains used for infections.

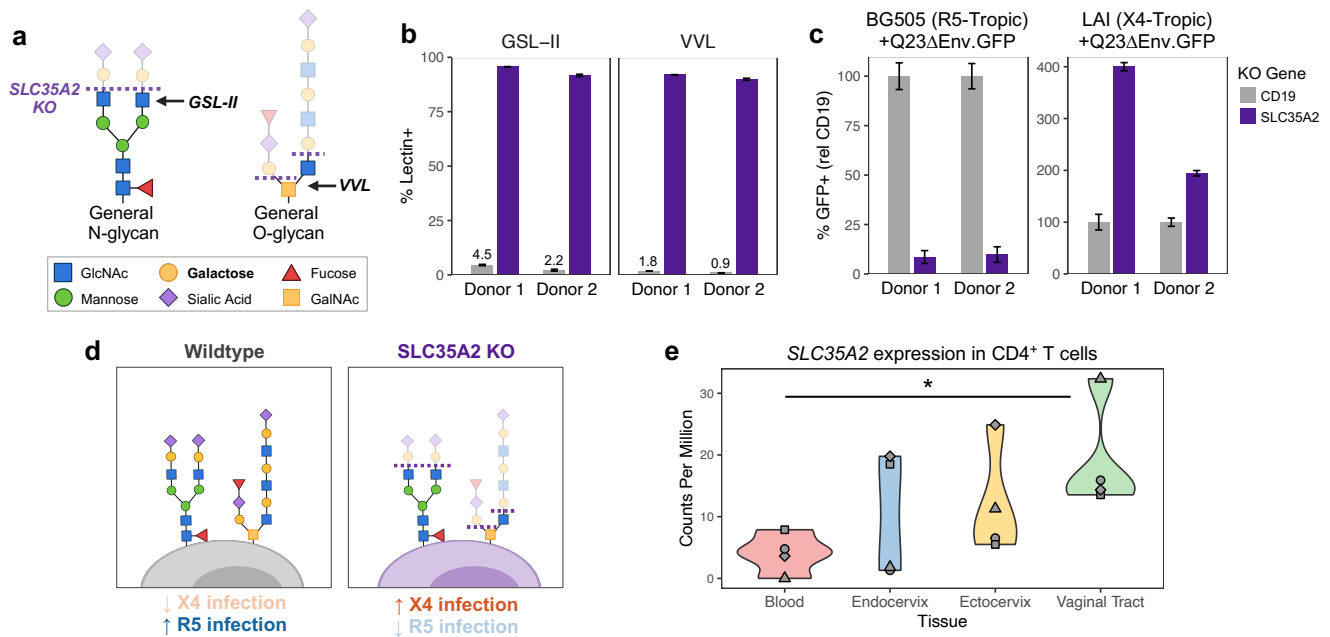


Figure 3. SLC35A2 inactivation causes truncated glycans on host cells and impacts HIV infection within a single round of infection.

A. General schematic of glycans, including the expected truncation when SLC35A2 is inactivated. The terminal residue binding sites of GSL-II and VVL lectins are indicated with arrows.

B. Percentage of cells staining positive for GSL-II and VVL lectins in CD4⁺ T cells from two donors. Data are shown as mean ± standard deviation of 2-3 replicate stains.

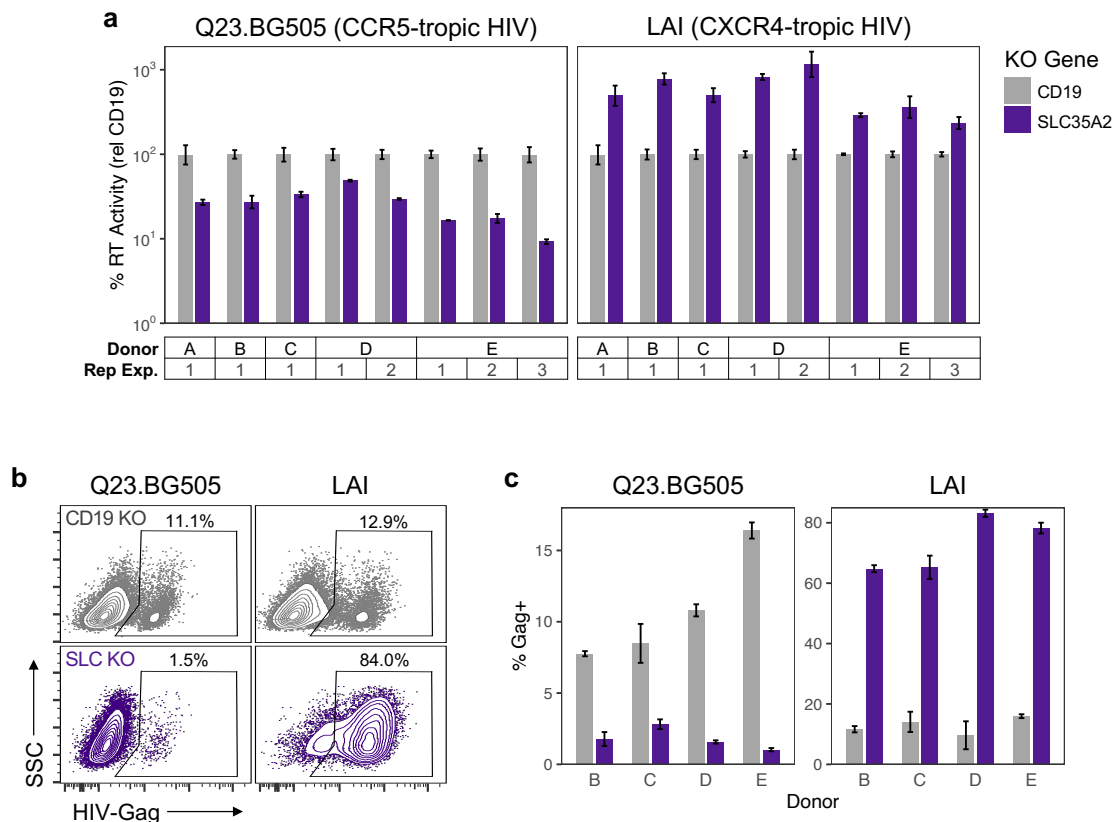
C. Percentage of GFP⁺ CD4⁺ T cells, relative to CD19 KO, after 2 days of infection with GFP-expressing HIV pseudoviruses (MOI=1). Data are shown as mean ± standard deviation of 3 replicate infections.

D. Working model of the impact of CD4⁺ T cell glycosylation on HIV infection.

E. SLC35A2 expression in CD4⁺ T cells isolated from different anatomical compartments from four donors.

Shapes denote donors. One-tailed paired t-test, asterisk indicates p < 0.05.

EXTENDED DATA FIGURES

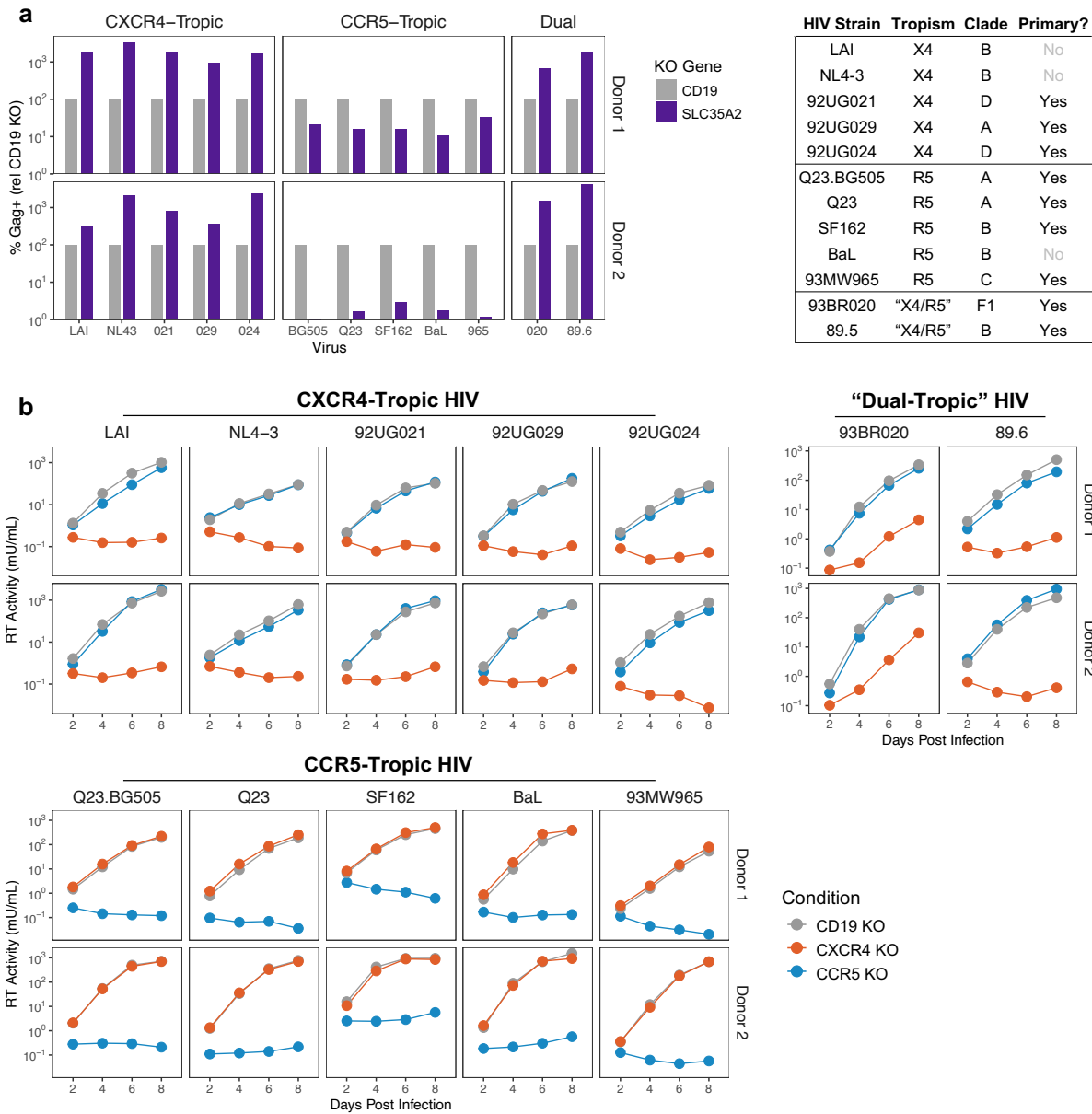


Extended Data Figure 1. SLC35A2 KO has opposite effects on two HIV strains that utilize different coreceptors, related to Figure 2A and 2B.

A. Technical and donor replicates for data depicted in Figure 2A. Reverse transcriptase (RT) activity, relative to CD19 KO for each primary CD4⁺ T cell donor/experiment, at 2 days post infection (dpi) (MOI=0.02). Knockouts were performed in five donors across independent replicate KO and infection experiments where indicated. Data are shown as mean \pm standard deviation of 2-3 replicate infections.

B. Representative flow plots of HIV-Gag-FITC staining CD4⁺ T cells from one donor after 3 dpi (MOI=1).

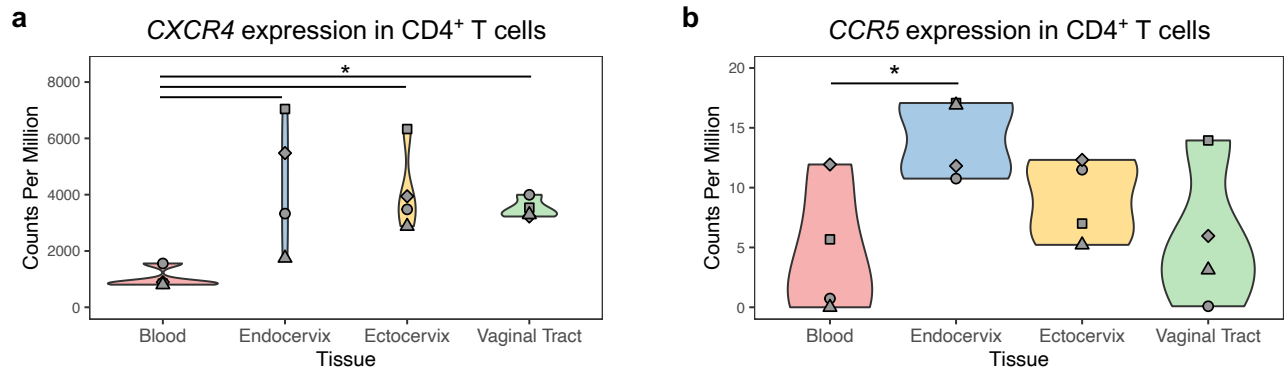
C. Technical and donor replicates for data depicted in Figure 2B. Percentage of CD4⁺ T cells staining positive for HIV-Gag-FITC at 3 dpi (MOI=1). Donor letters correspond with those in Panel A. Data are shown as mean \pm standard deviation of 2-3 replicate infections.



Extended Data Figure 2. SLC35A2 KO differentially impacts CXCR4-tropic and CCR5-tropic HIV, related to Figure 2D.

A. Percentage of CD4⁺ T cells staining positive for HIV-Gag, relative to CD19 KO, at 6 dpi (MOI=0.02, same infections as depicted in Figure 2A, different measurement of HIV infection). Donor numbering is consistent with Figure 2D.

B. RT activity over time from spreading infections (MOI=0.02) in primary CD4⁺ T cells from two donors to determine coreceptor tropism. Results are separated by expected coreceptor tropism based on prior literature. Donor numbering is consistent with Panel A and Figure 2D.



Extended Data Figure 3. Expression of HIV coreceptors in CD4⁺ T cells from the blood and common HIV transmission sites.

(A) *CXCR4* and (B) *CCR5* expression in CD4⁺ T cells isolated from different anatomical compartments from four donors. Shapes denote donors. One-tailed paired t-test, asterisk indicates $p < 0.05$.

SUPPLEMENTARY INFORMATION

Supplementary Table 1. PIKA guide library with synthetic NTC gene assignments.

Supplementary Table 2. HIV-CRISPR screen results.

Supplementary Table 3. Guide sequences for single-gene editing experiments.

METHODS

Cell lines

HEK293T/17 (ATCC #CRL-11268) and TZM-bl cells (NIH HIV Reagent Program #ARP-8129) were cultured in DMEM complete media (DMEM (Gibco #11965092) supplemented with 10% FBS (Gibco #26140079), 2 mM L-glutamine (Gibco #25030081), 1X antibiotic-antimycotic (Gibco #15240062) that contains penicillin, streptomycin, and Amphotericin B).

HIV-1 viruses

HIV Q23.BG505¹⁶ is a clade A, CCR5-tropic HIV chimeric virus derived from a full-length provirus isolated early in infection (Q23³²) and bears the BG505.C2 envelope³³ of a well-studied clade A transmitted/founder virus. HIV LAI is a clade B, CXCR4-tropic HIV strain, as previously described¹⁵. Infectious molecular clones (IMCs) of Q23.BG505¹⁶ and LAI¹⁵, as well as Q23³², SF162³⁴, and NL4-3³⁵ (NIH HIV Reagent Program ARP-2852, contributed by Dr. M. Martin) were generated by transfection using proviral plasmids. The remaining viruses were obtained as viral stocks through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: HIV 92/UG/021 (ARP-1648), HIV 92/UG/029 (ARP-1650), HIV 92/UG/024 (ARP-1649), and HIV 93/BR/020 (ARP-2329) contributed by UNAIDS Network for HIV Isolation and Characterization; HIV Ba-L (ARP-510) contributed by Dr. Suzanne Gartner, Dr. Mikulas Popovic, and Dr. Robert Gallo³⁶; HIV 93/MW/965 (ARP-2913) contributed by Dr. Paolo Miotti and the UNAIDS Network for HIV Isolation and Characterization; HIV 89.6 (ARP-1966) contributed by Dr. Ronald Collman³⁷.

HIV-1 production

HIV IMCs and pseudoviruses were prepared by transfecting HEK293T/17 cells in a six-well format. Twenty hours prior to transfection, 5e5 cells/well were plated in 2 mL of DMEM complete. For each HIV IMC transfection well, 1 µg of proviral DNA and 200 µL serum-free DMEM were combined, prior to the addition of 18 µL of FuGENE (Promega #E2692). Pseudoviruses were also made by transfections using plasmids expressing the envelope of BG505.C2³³ or LAI (gift from Dr. Michael Emerman) and Q23ΔEnv.GFP³⁸ (NIH HIV Reagent Program ARP-12647). For each HIV pseudovirus transfection well, 667 µg of envelope DNA, 1.334 µg of Q23ΔEnv.GFP, and 200 µL serum-free DMEM were combined, prior to the addition of 36 µL of

FuGENE. Transfection complexes were gently mixed, incubated for 15-30 minutes at room temperature, and added to each well in a drop-wise manner. At 20 h post-transfection, media was replaced with 1.5 mL of DMEM complete per well. Supernatants were harvested two days post-transfection and clarified using a 0.2 μ M filter (Millipore Sigma #SCGP00525). Transfection supernatants were concentrated using Amicon 100 KDa concentrators (Millipore Sigma #UFC910024) and titered on TZM-bl cells, as previously described¹⁸.

Lentiviral plasmids

Lentiviruses were used to deliver the CRISPR guide library to CD4⁺ T cells for HIV-CRISPR screens and SLC35A2 cDNA to SLC35A2 KO CD4⁺ T cells for rescue experiments. To generate the lentiviral vector for HIV-CRISPR screens, the PIKA guide library was amplified from the HIV-CRISPR PIKA library plasmid¹⁷ and cloned into the HIV-TOP opt mCh Δ W¹² plasmid for primary CD4⁺ T cell screens, as previously described¹⁸. To generate the lentiviral vector for SLC35A2 rescue experiments, SLC35A2 cDNA (393 aa, NCBI Reference Sequence: NP_001035963.1) was codon-optimized and synthesized by Twist Bioscience, amplified, and cloned into the TOP-CAR expression vector³⁹. Specifically, the codon-optimized SLC35A2 fragment, WPRE-LTR from the HIV-CRISPR PIKA plasmid¹⁷, and AgeI- and PmeI-digested TOP-CAR were assembled using HiFi DNA Assembly Master Mix (NEB #E2621L) to generate the TOP-SLC35A2 expression plasmid.

Lentivirus production

Lentiviruses were transfected as described above for HIV production, using the following transfection reagents per well: 667 ng of transfer vector (HIV-TOP opt mCh Δ W with the PIKA guide library or TOP-SLC35A2), 500 ng of psPAX2 packing vector (Addgene #12260), 333 ng of the pMD2.G VSV-G expression vector (Addgene #12259), 200 μ L serum-free DMEM, and 4.5 μ L TransIT-LT1 transfection reagent (Mirus Bio #MIR-2305). Lentiviral transfections were concentrated over a sucrose gradient, as previously described¹⁸.

Preparation of ribonucleoproteins (RNPs)

Non-targeting control (NTC) RNPs utilized a single sgRNA from IDT, whereas RNPs for all other single gene KOs were assembled using a mixture of three sgRNAs from Synthego (Synthego Gene KO Kit v2). Guide

sequences are available in **Supplementary Table 3**. RNPs were prepared with Cas9-NLS protein (UC Berkeley QB3 MacroLab) as described previously¹⁸.

Overview of primary CD4⁺ T cell experiments

In brief, this study reports three types of primary CD4⁺ T cell editing and infection experiments: (1) single gene editing, (2) HIV-CRISPR screens, and (3) SLC35A2 rescue experiments. All primary CD4⁺ T cell experiments utilize the following methods and timeline, with additional methods for specific experiments as indicated: CD4⁺ T cell isolation on Day 1, CD4⁺ T cell activation on Day 1, transduction of CD4⁺ T cells with lentivirus on Day 2 (*CRISPR screens and SLC35A2 rescue experiments only*), Cas9 delivery and genomic editing of CD4⁺ T cells by nucleofection on Day 4, and HIV infection on Day 14. This experimental timeline and all related methods were recently described in detail by our group¹⁸. General protocols, as well as methods specific to CRISPR screens and SLC35A2 rescue experiments, are described in the following five sections.

Primary CD4⁺ T cell isolation and activation

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood from healthy donors by centrifugation over Ficoll-Paque Plus (Cytiva #17144002). CD4⁺ T cells were isolated using negative selection (STEMCELL Technologies #17952) and were immediately cultured in complete RPMI media (RPMI 1640 (Gibco #22400089) supplemented with 10% FBS, 2 mM L-glutamine, 1X antibiotic-antimycotic) with 100 U/mL recombinant interleukin (IL)-2 (Roche #11147528001) and activated or frozen in freezing media (90% FBS, 10% DMSO). CD4⁺ T cells were used directly after isolation for HIV-CRISPR screens and were thawed from frozen aliquots for all other experiments. After isolation or thawing, cells were resuspended at 2.5e6 cells/mL in 48-well flat-bottom plates with complete RPMI media and 100 U/mL IL-2. Cells were activated with plate-bound anti-CD3 (Tonbo Biosciences #40-0038) at a concentration of 10 µg/mL and anti-CD28 (Tonbo Biosciences #40-0289) supplemented to the media at a concentration of 5 µg/mL.

RNP nucleofection of CD4⁺ T cells

Three days after activation (Experiment Day 4), nucleofections were performed using 1e6 CD4⁺ T cells per nucleofection as previously described¹⁸. For nucleofections with NTC RNPs, cells were resuspended in 20 µL

of P3 Primary Cell Nucleofector solution (Lonza #V4SP-3096) and mixed gently with 4 μ L of NTC complexes. For RNPs with Synthego guides, cells were gently resuspended with 25 μ L of RNPs. Nucleofections were carried out in a 16-well Nucleocuvette (Lonza #V4SP-3096) using an Amaxa Nucleofector (Lonza) and pulse code EH-115. Cells were allowed to recover for 30 min to 2 h at 37C and were then brought to 2.5e6 cells/mL in RPMI complete media with 100 U/mL IL-2 and at a 1:1 ratio with activation beads (Miltenyi Biotec #130-091-441). Two days later, 200 μ L of RPMI complete media with 100 U/mL IL-2 was added to each nucleofection. At four days post nucleofection, half of the culture media was replaced with fresh RPMI complete media with 100 U/mL IL-2 and cells were brought to 1e6 cells/mL. Starting the following day, cells were counted and resuspended to 1e6 cells/mL in RPMI complete media with 100 U/mL IL-2 every other day.

HIV infection of CD4⁺ T cells

Ten days post-nucleofection (Experiment Day 14), HIV infections were established via spinoculation at 1100xg for 90 min at 30C in the presence of 8 μ g/mL polybrene (Millipore Sigma #TR-1003-G) as previously described¹⁸. After spinoculation, infection inoculum was replaced with fresh RPMI complete with 100 U/mL IL-2. Viral supernatants were harvested at various days post infection as indicated in figure legends and stored at -80C.

HIV-CRISPR screening

On the day after CD4⁺ T cell isolation and activation (Experiment Day 2), cells were transduced with HIV-TOP opt mCh Δ W lentivirus containing the PIKA library by spinoculation in the presence of 8 μ g/mL protamine sulfate (Millipore Sigma #P3369). Two days later (Experiment Day 4), Cas9 was delivered to cells by nucleofection with NTC RNPs. Nine days after nucleofection (Experiment Day 4), all library cells were treated with 1000 U/mL IFN- β 1a (PBL Assay Science #11410). On the following day, infections with HIV IMCs (MOI=1) were established. Cells were split 1:2 three days post infection with fresh RPMI complete media with 100 U/mL IL-2 and 1000 U/mL IFN. Viral supernatants were harvested 3, 4, and 5 dpi and cells were collected 5 dpi. All viral supernatants were filtered to remove cellular debris (Millipore Sigma #SCGP00525) and stored at 4C up to four days post collection. Viral supernatants across collection days were combined and

concentrated over a sucrose gradient, resuspended in 100 μ L PBS, and stored at 4C. Cells were washed in PBS, pelleted, and stored at -80C.

Guide sequences were amplified from genomic DNA extracted from cell pellets and viral RNA extracted from concentrated viral pellets and were prepared for Illumina sequencing as previously described¹⁸. Pooled libraries were sequenced on an Illumina HiSeq 2500 (Fred Hutch Shared Resources Genomics Shared Resource). Illumina sequencing reads (SRA accession number: **TBD**) were analyzed by MAGeCK-Flute⁴⁰ to generate guide counts and gene-level enrichment data (**Supplementary Table 2**; Fred Hutch Center for Data Visualization). NTC guide sequences were iteratively binned to create a number of NTC “genes” (synthetic NTCs, synNTCs) with eight guides per gene to match the number of genes in the PIKA library (1906 synNTCs total; **Supplementary Table 1**). This approach captures variation across NTC guides when analyzing gene-level data more accurately than aggregating all 200 guides into one NTC gene-level data point. Pearson R correlations between MAGeCK scores from different donors were determined using the `stat_cor` function in the `ggpubr` package. Screen background was calculated as the average synNTC MAGeCK score + 3 standard deviations.

SLC35A2 rescue experiments

Primary CD4⁺ T cells were transduced with TOP-SLC35A2 lentivirus on the day after cells were initially thawed and activated. Transductions were performed via spinoculation at 1100xg for 90 min at 30C in the presence of 8 μ g/mL protamine sulfate (Millipore Sigma #P3369). Cells were nucleofected with SLC35A2 RNPs two days later, following the normal timeline of all other gene editing and HIV infection experiments. Importantly, SLC35A2 RNPs are unable to target and edit the exogenous codon-optimized SLC35A2 cDNA delivered via transduction with TOP-SLC35A2 lentivirus.

Reverse transcriptase (RT) activity qPCR assay

HIV infection levels were assessed by measuring RT activity (mU/mL) in viral supernatants using the previously described RT activity qPCR assay^{18,41}. The RT units of a concentrated stock of HIV Q23.BG505 virus were determined multiple times using a standard curve of purified HIV RT enzyme (Worthington

Biochemical Corp. #LS05003). Aliquots of this titered stock of Q23.BG505 were used as the quantitative standard curve in all assays. The RT activity (mU/mL) of experimental wells was interpolated from the quantitative standard curve using the `lm` function in the R stats package.

HIV-Gag staining

After HIV infection, as indicated in figure legends, cells were collected, fixed in 4% paraformaldehyde (Santa Cruz Biotechnology #sc-281692) for 10 min, and either stored at 4C for up to two days or immediately prepared for staining. Cell pellets were permeabilized for intracellular staining via resuspension in 0.5% Triton X-100 in PBS for 10 min and were subsequently resuspended in a 1:500 dilution of anti-p24-FITC KC57 antibody (Beckman Coulter #6604665) in 1% BSA in PBS for 1 h at room temperature in the dark. Washed cells were analyzed by flow cytometry on a BD LSRFortessa X-50 or BD FACSymphony A5.

Genomic editing analysis

SLC35A2 KO CD4⁺ T cells were collected on the day of HIV infection to assess for genomic editing as previously described in detail¹⁸. In brief, regions containing SLC35A2 guide cut sites were amplified from genomic DNA (forward primer: 5' AAGCTGCCACAAATAGCCT; reverse primer: 5' AAGCTGCCACAAATAGCCT) and sequenced (sequencing primer: 5' AAGCTGCCACAAATAGCCT) by Sanger sequencing (Fred Hutch Shared Resources Genomics Core). Chromatograms were analyzed by ICE analysis tool v3 (Synthego) to determine the rate of editing predicted to cause SLC35A2 KO.

Lectin staining

On the day of HIV infection, 2e5 edited CD4⁺ T cells per stain replicate were collected, fixed in 4% paraformaldehyde for 10 min, and resuspended in 20 µg/mL VVL-FITC (Vector Laboratories #FL-1231-2) and 10 µg/mL GSL-II-AF647 (ThermoFisher #L32451) in 2% FBS in PBS for 1 h at room temperature in the dark. Cells were washed, resuspended in 2% FBS in PBS, and were analyzed by flow cytometry on a BD LSRFortessa X-50 or BD FACSymphony A5.

Analysis of IFN-inducibility

We previously reported a bulk RNA-seq dataset of primary CD4⁺ T cells treated with IFN- β 1a (SRA: PRJNA921704)¹⁸. Here, we analyzed this dataset to determine the fold induction of a subset of genes at various times post IFN treatment. Fold inductions were averaged across unique donors (3 donors for 6 and 12 h timepoints, 5 donors for 24 h timepoint).

Analysis of gene expression at HIV transmission sites

To determine expression levels of SLC35A2 and the HIV coreceptors at different anatomical compartments, we analyzed a previously published bulk RNA-seq dataset that assessed CD4⁺ T cell populations isolated from the blood, endocervix, ectocervix, and vaginal tract of four individuals²⁶. Specifically, this study gated on Live, CD45+CD3+CD4+CCR7⁻ cells to isolate CD4⁺ T cells and further separated cells into CD103-CD69⁻, CD103-CD69⁺, and CD103+CD69⁺ populations before performing RNA-seq, as CD103 and CD69 are markers of resident memory T cells. We accessed raw counts from this dataset via GEO (GSE163260) and calculated counts per million (CPM) using the cpm function in the Bioconductor edgeR package⁴². CPMs were averaged across CD4⁺ T cell populations within a tissue site for each donor to determine gene expression levels in bulk CD4⁺ T cells at each site. SLC35A2 expression levels did not vary between T cell subsets within anatomical sites.

Data availability

Illumina sequencing reads from HIV-CRISPR screens can be accessed via the SRA accession number **TBD** (*SRA submission in progress*). Any additional information required to reanalyze the data reported in this paper is available upon request.

ACKNOWLEDGMENTS

We thank the Fred Hutch Shared Resources Genomics and Bioinformatics Cores, particularly Alyssa Dawson and Elizabeth Jensen for performing Illumina and Sanger sequencing, respectively; the Fred Hutch Center for Data Visualization, especially Michael Zager, Nathan Thorpe, and Sam Minot, for their support with MAGeCK analysis; Michael Emerman, Molly OhAinle, Caitlin Stoddard, and Alex Willcox for helpful discussions and technical assistance. Figures 1A, 3A, and 3D were generated with BioRender. This work was supported by the following NIH grants: NICHD R01 HD103571 to J.O. and NIGMS T32 GM007270 and NIAID F31 AI165168 to H.L.I.

AUTHOR CONTRIBUTIONS

J.O. conceived the project and J.O., H.L.I., and D.H. designed the methodology of the study. H.L.I., D.H., and N.E.B. performed experiments. H.L.I. and D.H. conducted the data analysis, with the supervision of J.O. H.L.I. generated data visualizations. J.O. and H.L.I. wrote the paper with input from D.H. and N.E.B.

COMPETING INTERESTS

The authors declare no competing interests.

MATERIALS & CORRESPONDENCE

Further information and requests for data, resources, and research materials should be directed to Julie Overbaugh (joverbau@fredhutch.org).

REFERENCES

- 1 Ronen, K., Sharma, A. & Overbaugh, J. HIV transmission biology: translation for HIV prevention. *AIDS* **29**, 2219-2227, doi:10.1097/QAD.0000000000000845 (2015).
- 2 Wolinsky, S. M. *et al.* Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. *Science* **255**, 1134-1137, doi:10.1126/science.1546316 (1992).
- 3 Wolfs, T. F., Zwart, G., Bakker, M. & Goudsmit, J. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. *Virology* **189**, 103-110, doi:10.1016/0042-6822(92)90685-i (1992).
- 4 Zhu, T. *et al.* Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science* **261**, 1179-1181, doi:10.1126/science.8356453 (1993).
- 5 Zhang, L. Q. *et al.* Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. *J Virol* **67**, 3345-3356, doi:10.1128/JVI.67.6.3345-3356.1993 (1993).
- 6 Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* **105**, 7552-7557, doi:10.1073/pnas.0802203105 (2008).
- 7 Carlson, J. M. *et al.* HIV transmission. Selection bias at the heterosexual HIV-1 transmission bottleneck. *Science* **345**, 1254031, doi:10.1126/science.1254031 (2014).
- 8 Sugrue, E. *et al.* The apparent interferon resistance of transmitted HIV-1 is possibly a consequence of enhanced replicative fitness. *PLoS Pathog* **18**, e1010973, doi:10.1371/journal.ppat.1010973 (2022).
- 9 Fenton-May, A. E. *et al.* Relative resistance of HIV-1 founder viruses to control by interferon-alpha. *Retrovirology* **10**, 146, doi:10.1186/1742-4690-10-146 (2013).
- 10 Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *Proc Natl Acad Sci U S A* **110**, 6626-6633, doi:10.1073/pnas.1304288110 (2013).
- 11 Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: cell binding and entry. *Cold Spring Harb Perspect Med* **2**, doi:10.1101/cshperspect.a006866 (2012).
- 12 Long, E. M., Rainwater, S. M., Lavreys, L., Mandaliya, K. & Overbaugh, J. HIV type 1 variants transmitted to women in Kenya require the CCR5 coreceptor for entry, regardless of the genetic complexity of the infecting virus. *AIDS Res Hum Retroviruses* **18**, 567-576, doi:10.1089/088922202753747914 (2002).
- 13 Saba, E. *et al.* HIV-1 sexual transmission: early events of HIV-1 infection of human cervico-vaginal tissue in an optimized ex vivo model. *Mucosal Immunol* **3**, 280-290, doi:10.1038/mi.2010.2 (2010).
- 14 Shepherd, J. C. *et al.* Emergence and persistence of CXCR4-tropic HIV-1 in a population of men from the multicenter AIDS cohort study. *J Infect Dis* **198**, 1104-1112, doi:10.1086/591623 (2008).
- 15 Peden, K., Emerman, M. & Montagnier, L. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. *Virology* **185**, 661-672, doi:10.1016/0042-6822(91)90537-I (1991).
- 16 Haddox, H. K., Dingens, A. S., Hilton, S. K., Overbaugh, J. & Bloom, J. D. Mapping mutational effects along the evolutionary landscape of HIV envelope. *Elife* **7**, doi:10.7554/eLife.34420 (2018).

- 17 OhAinle, M. *et al.* A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. *Elife* **7**, doi:10.7554/eLife.39823 (2018).
- 18 Itell, H. L., Humes, D. & Overbaugh, J. Several cell-intrinsic effectors drive type I interferon-mediated restriction of HIV-1 in primary CD4(+) T cells. *Cell Rep* **42**, 112556, doi:10.1016/j.celrep.2023.112556 (2023).
- 19 S, P. H., Krumbiegel, M. & Kirchhoff, F. Coreceptor usage of BOB/GPR15 and Bonzo/STRL33 by primary isolates of human immunodeficiency virus type 1. *J Gen Virol* **80 (Pt 5)**, 1241-1251, doi:10.1099/0022-1317-80-5-1241 (1999).
- 20 Doranz, B. J. *et al.* A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* **85**, 1149-1158, doi:10.1016/s0092-8674(00)81314-8 (1996).
- 21 Quelhas, D. *et al.* SLC35A2-CDG: Novel variant and review. *Mol Genet Metab Rep* **26**, 100717, doi:10.1016/j.ymgmr.2021.100717 (2021).
- 22 in *Essentials of Glycobiology* (eds A. Varki *et al.*) (2022).
- 23 Ng, B. G. *et al.* SLC35A2-CDG: Functional characterization, expanded molecular, clinical, and biochemical phenotypes of 30 unreported Individuals. *Hum Mutat* **40**, 908-925, doi:10.1002/humu.23731 (2019).
- 24 Bojar, D. *et al.* A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chem Biol* **17**, 2993-3012, doi:10.1021/acscchembio.1c00689 (2022).
- 25 Ma, T. *et al.* Single-cell glycomics analysis by CyTOF-Lec reveals glycan features defining cells differentially susceptible to HIV. *Elife* **11**, doi:10.7554/eLife.78870 (2022).
- 26 Woodward Davis, A. S. *et al.* The human memory T cell compartment changes across tissues of the female reproductive tract. *Mucosal Immunol* **14**, 862-872, doi:10.1038/s41385-021-00406-6 (2021).
- 27 Chabot, D. J., Chen, H., Dimitrov, D. S. & Broder, C. C. N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates. *J Virol* **74**, 4404-4413, doi:10.1128/jvi.74.9.4404-4413.2000 (2000).
- 28 Bannert, N. *et al.* Sialylated O-glycans and sulfated tyrosines in the NH₂-terminal domain of CC chemokine receptor 5 contribute to high affinity binding of chemokines. *J Exp Med* **194**, 1661-1673, doi:10.1084/jem.194.11.1661 (2001).
- 29 Farzan, M. *et al.* Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell* **96**, 667-676, doi:10.1016/s0092-8674(00)80577-2 (1999).
- 30 Lee, B. *et al.* Epitope mapping of CCR5 reveals multiple conformational states and distinct but overlapping structures involved in chemokine and coreceptor function. *J Biol Chem* **274**, 9617-9626, doi:10.1074/jbc.274.14.9617 (1999).
- 31 Spillings, B. L. *et al.* Host glycoalyx captures HIV proximal to the cell surface via oligomannose-GlcNAc glycan-glycan interactions to support viral entry. *Cell Rep* **38**, 110296, doi:10.1016/j.celrep.2022.110296 (2022).
- 32 Poss, M. & Overbaugh, J. Variants from the diverse virus population identified at seroconversion of a clade A human immunodeficiency virus type 1-infected woman have distinct biological properties. *J Virol* **73**, 5255-5264, doi:10.1128/JVI.73.7.5255-5264.1999 (1999).

- 33 Wu, X. *et al.* Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant. *J Virol* **80**, 835-844, doi:10.1128/JVI.80.2.835-844.2006 (2006).
- 34 Cheng-Mayer, C., Quiroga, M., Tung, J. W., Dina, D. & Levy, J. A. Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation. *J Virol* **64**, 4390-4398, doi:10.1128/JVI.64.9.4390-4398.1990 (1990).
- 35 Adachi, A. *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* **59**, 284-291, doi:10.1128/JVI.59.2.284-291.1986 (1986).
- 36 Gartner, S. *et al.* The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science* **233**, 215-219, doi:10.1126/science.3014648 (1986).
- 37 Collman, R. *et al.* An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1. *J Virol* **66**, 7517-7521, doi:10.1128/JVI.66.12.7517-7521.1992 (1992).
- 38 Provine, N. M., Cortez, V., Chohan, V. & Overbaugh, J. The neutralization sensitivity of viruses representing human immunodeficiency virus type 1 variants of diverse subtypes from early in infection is dependent on producer cell, as well as characteristics of the specific antibody and envelope variant. *Virology* **427**, 25-33, doi:10.1016/j.virol.2012.02.001 (2012).
- 39 Humes, D., Rainwater, S. & Overbaugh, J. The TOP vector: a new high-titer lentiviral construct for delivery of sgRNAs and transgenes to primary T cells. *Mol Ther Methods Clin Dev* **20**, 30-38, doi:10.1016/j.omtm.2020.10.020 (2021).
- 40 Wang, B. *et al.* Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nat Protoc* **14**, 756-780, doi:10.1038/s41596-018-0113-7 (2019).
- 41 Vermeire, J. *et al.* Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. *PLoS One* **7**, e50859, doi:10.1371/journal.pone.0050859 (2012).
- 42 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).