Identifying host factors that inhibit HIV-1 infection in primary $CD4^+$ T cells

Hannah L. Itell

A dissertation

submitted in the partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Michael Emerman, Chair

Julie Overbaugh

Jennifer Lund

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Identifying host factors that inhibit HIV-1 infection in primary $CD4^+$ T cells

Hannah L. Itell

Chair of the Supervisory Committee:

Michael Emerman

Department of Microbiology and Department of Global Health

Cellular antiviral factors comprise one of the first lines of defense against viral infection. In the case of HIV-1, it is well established that the early stages of natural infection are inefficient, characterized by infrequent transmission, a severe bottleneck in viral genetic diversity during transmission, and variable viral loads between individuals in the first weeks of infection. Because HIV-1 transmission and acute infection occur prior to the onset of adaptive immune responses, cell-intrinsic factors that inhibit HIV-1 replication may contribute to the inefficiency of early HIV-1 infection.

Many studies have identified cellular factors with potent antiviral activity against HIV-1, which work has shed light on the cell types permissive to HIV-1 infection, the host range of HIV-1, and the role of HIV accessory proteins. Though these prior reports emphasize the significance of restriction factors, they do not provide direct insight into the genes that impact natural HIV-1 transmission and infection, as nearly all were performed in cell lines with lab-adapted viruses.

This caveat is largely due to the fact that the main target cells of HIV-1 infection, primary $CD4^+$ T cells, were not amenable to reliable experimental manipulation until recently.

To understand which host factors are relevant during the early stages of HIV-1 infection *in vivo*, we adapted the high throughput HIV-CRISPR screening approach for use in primary $CD4^+$ T cells. This development allowed us to comprehensively interrogate hundreds of genes upregulated by the innate interferon response (Chapter 2) as well as endogenously expressed genes (Chapter 3) for inhibitory activity against a primary HIV-1 virus in a physiologically relevant cell type. In Chapter 2, we identified novel, previously undescribed HIV-restricting interferon-stimulated genes (HM13, IGFBP2, LAP3). We also found that two factors characterized in other HIV-1 infection models (IFI16 and UBE2L6) mediated interferon restriction in primary target cells, although several others (MX2, IFITM1, TRIM5α) did not impact the primary HIV-1 strain in this cell type. Moreover, our results from inactivating single and combinations of antiviral interferon-stimulated genes suggest that interferon restriction of HIV-1 is multifaceted, resulting from several effectors functioning collectively as opposed to one potent gene. In Chapter 3, I specifically explored host factors that influence the HIV-1 transmission bottleneck, which results in the preferential transmission of CCR5-tropic viruses as opposed to CXCR4-tropic variants. We identified SLC35A2, a constitutively expressed gene involved in glycosylation, as an CXCR4-tropic-specific restriction factor. Inactivation of SLC35A2 impacted HIV-1 in a tropism-dependent manner by truncating host cell glycans, underscoring a previously unappreciated role for host cell glycosylation on HIV-1 that may contribute to the bottleneck during HIV-1 transmission.

Together, these projects shed light on the cellular factors that likely influence HIV-1 transmission and acute infection and demonstrate the value of a high throughput screening approach for host genes that impact HIV-1 in primary target cells. Future studies should investigate the mechanisms of the restriction factors described here and leverage HIV-CRISPR screens in primary cells to study host genes that influence HIV-1 in other contexts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This work would not have been possible without the mentorship, support, insight, and effort of many individuals. First and foremost, I would like to thank Julie Overbaugh for being an incredibly dedicated mentor and an aspirational role model to me. I'm so grateful that I joined your lab and cannot thank you enough for the time, advice, and encouragement you've given me during my PhD. I have learned so much from you during this time – not only technical skills, such as grant writing, prioritizing experiments, or tailoring effective talks, but also personal lessons from of the example you set, such as maintaining optimism and motivation when a project isn't moving forward or in the face of critical reviewer comments and being generous with your time when it comes to training or offering advice to others. Overall, your passion, resilience, thoughtfulness, and grace are so inspiring to me, and I feel extremely lucky to have you as a mentor and role model – thank you!

I would also like to thank Dara Lehman and all current and former members of the Overbaugh and Lehman labs. Dara, you are the one of the most kind, reassuring, and insightful people I know. I am so grateful for our conversations about science, life, travel (really any topic!), and I cherish your perspective and feedback on all matters. In addition to Julie and Dara, I feel like I hit the lottery in terms of overall lab environment. From your incredibly helpful feedback during lab meetings and practice talks, to lunches together, lab bake off, and dog meetups, it's been a blast to get to spend the past few years with such a supportive, impressively smart, and fun group of people. To Daryl Humes – I would not have had this PhD project without all of the work you did before I even started in the lab, let alone without you training me and working together every day of the first two years of my PhD. Thank you for your time training and advising me, for your work on all of the projects described here, and for being a great role model. To Team IFN, which over the years has included Daryl, Ted, Caroline K.,

Joshua, Nell, Kevin, Alex, and Caitlin, thank you for your helpful discussions, feedback, and comradery. Nell, thank you for being such a generous and thoughtful lab mate and for your work on Chapter 3. To Nicole, Meghan, Ted, Zak, Alex, Ryan, Feli, and Caroline P., it's been so fun to be grad students together and learn from and with each of you. Meghan, thank you for being the lab cheerleader, the best rotation mentor, and such an optimistic person – you continue to be such a role model to me. To the postdocs in the lab, Laura, Joshua, Mackenzie, Delphine, and Caitlin, I have relied on each of you for advice, feedback, and mentorship at various times of my PhD and I'm so grateful for the time you gave me. To Jamie, I have loved working with you on recent primary cell projects and can't wait to see where you take the Chapter 3 story. To Vrasha, thank you for being such a kind lab mate and so generous with your time in helping with cohort log and inventory questions. To the techs in the lab, Dana, Noah, Sonja, Carolyn, Caroline K., Haidyn, Nell, Michelle, and Morgan, none of our lab's work would be possible without you all keeping the lab functioning and doing so much work behind the scenes. I really appreciate all of your help. So to everyone in the Overbaugh and Lehman labs – thank you!

Beyond our lab, I am so grateful to have pursued my PhD at Fred Hutch and UW. To the virology community at the Hutch, thank you for your challenging, supportive, and engaging feedback at Thursday Virus Meetings every week – these meetings have been an invaluable source of new ideas and insightful questions. To "Team Screen", including Michael Emerman, Molly Ohainle, Emily Hsieh, and Vanessa Montoya – first of all, both projects described here could not have been possible without Molly and Michael's work developing the HIV-CRISPR approach. Thank you for this major contribution to the HIV field that I was able to leverage in my graduate work. Moreover, I am so grateful for the feedback, advice, and tools I received during our Team Screen meetings during the pandemic. These discussions have shaped my PhD work and I've learned so much from each of you. To Jasmine Gonzalez, our lab's administrator, you are one of the most organized, prompt, and helpful people and we are so lucky to work with you. To MCB, particularly Andrea Brocato, Maia Low, and Denise Barnes, one of the main reasons I

chose MCB is because I was so impressed by how much the program supported its students and that has certainly been the case throughout my time in grad school, so thank you. And to my thesis committee – Julie Overbaugh, Michael Emerman, Jennifer Lund, Michael Gale, and Jesse Bloom – thank you for your advice, questions, and support throughout my PhD. An additional thanks to my reading committee (Julie Overbaugh, Michael Emerman, Jennifer Lund) for taking the time to thoroughly read my thesis and provide me with thoughtful feedback.

I would not be pursuing a PhD without the time, encouragement, and inspiration of prior mentors, particularly Malcolm Campbell, Genevieve Fouda, and Sallie Permar. You ignited my curiosity and passion for scientific research, gave me the opportunity to learn, grow, and pursue my aspirations, and supported and guided me at every stage (and still do now!). Thank you!

Finally, I could do not have made it through grad school without my friends and family. To my MCB cohort and grad school friends, it was so fun getting to know all of you and go through grad school together. To my friends from high school and college, you support every aspect of my life and grad school is no exception, so thank you as always. To the “Zoo Squad”, you all have become my family in Seattle. I can’t believe I went from knowing no one here to having such an incredible, tight-knit group of friends. You all are a constant source of joy, fun, and community in my life, and I am so grateful for you. To Molly and Laila, you two were the best roommates I could have asked for. I’m so thankful for our built-in friend time over the years and for our living room practice talks and hallway vent sessions about grad school. I’m not sure how I would have gotten through the first few years of grad school without you two. And a particular shout out to Molly – we have done literally everything together in Seattle, from hikes and taking trips for concerts and shows, to just walking to the Hutch every morning. You’re an incredible friend and have made every day of these past five years an absolute blast. To my parents, you are two of the most hard-working, considerate, detail-oriented, and driven people I know. You’ve not only taught me those qualities, but you’ve given me the freedom and opportunity to apply them towards whatever career path I’m passionate about. I am forever

grateful for your example, confidence in me, and continual encouragement. And to my brothers – being the youngest allowed me to watch and learn from both of you. Will, your creativity, curiosity, patience, and compassion. James, your resilience, charisma, dependability, and work ethic. Thanks for being the best role models from day one. And fortunately those I consider family has grown over the past several years, so to Ashleigh, Vincent, and the Scrivens, thank you all for your support. And finally, to my partner Aidan. You have supported me every day through every stage of my scientific journey – offering me advice and encouragement in times of frustration and stress, working around my chaotic experimental schedules, and celebrating my successes. I don't think I would have made the move out to the West Coast without you encouraging me to pursue my career aspirations and take that leap of faith. I am so grateful that I did, and I certainly could not have made it through grad school without you.

Chapter 1

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) is a retrovirus that establishes life-long infection upon transmission and causes acquired immunodeficiency syndrome (AIDS) if left untreated. Nearly 39 million people were reported to be living with HIV-1 globally in 2022, with an estimated 630,000 AIDS-related deaths and 1.3 million new infections occurring each year [1]. Therefore, despite the development of effective antiretroviral therapy (ART) to suppress viral replication in those living with HIV-1 and preventative drugs for those at high-risk of acquiring the virus, HIV-1 continues to cause a substantial burden of disease and mortality worldwide. Elucidating the interactions between HIV-1 and host cells during HIV-1 transmission, replication, and pathogenesis are critical to informing innovative prophylactic and therapeutic strategies to combat this global burden.

HIV-1 genetic diversity

Like all retroviruses, HIV-1 has a positive-sense, single-stranded RNA genome. In order to replicate upon infection of host cells, retroviruses must convert their single-stranded RNA genetic material back to double-stranded, proviral DNA using a viral enzyme called reverse transcriptase. However, reverse transcriptase is highly error-prone and frequently involves strand-transfer between two viral RNA templates, which can lead to recombination between strains if a cell is infected with more than one virus [2]. Reverse transcription therefore leads to high rates of mutations, recombination events, insertions, and deletions. Given this high mutation rate, coupled with the rapid rate of HIV-1 replication in infected, untreated individuals, it is estimated that mutations will occur in every position of the HIV-1 genome each day within a

single infected individual [3]. With 9.2 million untreated individuals living with untreated HIV globally, HIV-1 is one of the most genetically diverse pathogens, with high diversity both within a single host and globally.

HIV-1 group M, which is responsible for the global HIV/AIDS epidemic, is further classified into nine clades or subtypes (A-D, F-H, J, K), as well as an increasing number of circulating recombinant forms (CRFs), based on genetic identity. At the amino acid level, HIV-1 isolates from the same clade vary by 8-17% (maximum 30%) [4]. This variation is higher between clades, 17-35% (maximum 42%). A systematic review estimated that, between 1990-2015, clade C (46.6% of global cases), clade B (12.1%), clade A (10.3%), CRF02_AG (7.7%) and CRF01_AE (5.3%) were the top five most common HIV-1 subtypes globally [4]. Clade C is the main contributor of HIV-1 infections in southern Africa and India, whereas clade B dominates in western and central Europe and North America and Latin America. Clade A accounts for the majority of cases in east Africa, eastern Europe, and central Asia [4]. Despite this diversity and distribution, HIV-1 research has disproportionately focused on studies of clade B viruses. For instance, of the 19,988 complete HIV-1 genome sequences in the Los Alamos National Laboratory's database for HIV-1 sequences as of September 2023 [5], 54% (10,855) were clade B isolates, with only 5% (1046) and 12% (2,420) being clade A and C viruses, respectively. The high genetic diversity of HIV-1 within individuals and globally presents a major challenge to HIV-1 research efforts and must be considered during experimental design in order to differentiate findings that are strain- or clade-specific and those that are broadly relevant.

Early stages of HIV-1 infection

HIV-1 transmission

HIV-1 is primarily transmitted via viral exposure at mucosal sites such as the female genital tract, male genital tract, and intestinal tract (estimated 85% of cases) and can also be transmitted vertically (7%) and by intravenous exposure (8%) [6]. Upon exposure, the virus must

cross the mucosal epithelial barrier to reach potential target cells. Importantly, HIV-1 uses CD4 as a receptor for viral attachment and then further interacts with either of the chemokine receptors CXCR4 or CCR5 as a coreceptor to initiate viral fusion and entry. Potential target cells are therefore those that at least express CD4 and CXCR4 or CCR5, which implicates $CD4^+$ T cells, monocytes/macrophages, and dendritic cells. Because it is impossible to study transmission sites and this temporal window in humans in real time, researchers have relied on non-human primate and mouse infection models and human *ex vivo* tissue explants, primarily with cervicovaginal tissue, to understand how infection is initially seeded and propagated. These models have identified memory $CCR5^+$ $CD4^+$ T cells in the submucosal epithelium as the earliest target cells of productive HIV-1 infection [reviewed in [6, 7]]. Virus is rapidly propagated in these cells locally for a few days, until it disseminates to draining lymph nodes and secondary lymphoid tissue, which occurs within about a week after transmission in macaques. Mucosal and epithelial Langerhans cells, which are tissue-resident macrophages, and submucosal immature dendric cells may also contribute to early local viral propagation by binding and capturing HIV-1 and infecting $CD4^+$ T cells by cell-to-cell transmission [8]. $CD4^+$ T cells are thus the main cells that HIV-1 infects and productively replicates in during natural infection.

Importantly, HIV-1 transmission is an inefficient process. Despite heterosexual transmission accounting for nearly 70% of HIV-1 infections worldwide [6], the probability of transmission via this route in the absence of ART in the index case ranges from 1 in 200 to 1 in 3000 exposure events, with the highest estimated frequency being 1 in 10 exposure events depending on cofactors that influence risk, such as viral load, genital inflammation, circumcision, and clinical HIV stage [6]. In addition to the relatively low likelihood of HIV-1 transmission upon exposure, it is also well established that HIV-1 undergoes a severe population bottleneck during transmission. As previously described, HIV-1 has high genetic diversity even within a single infected individual. However, only one or a few of these viral variants present in the transmitting partner are detected in the recipient soon after transmission [6, 7, 9-14]. It is conceivable that

the inefficiency of HIV-1 transmission could be solely due to effective physical barriers, such as mucus and the epithelium of the genital tract, which would result in infrequent, stochastic transmission events. Nevertheless, transmitted/founder (T/F) HIV-1 variants share distinct characteristics, which suggests that this bottleneck is not entirely random. For instance, HIV-1 strains circulating later in infection can often use either CXCR4 or CCR5 as a coreceptor for viral entry [15]. However, transmitted variants nearly always use CCR5 [13, 16]. In fact, mutations in CCR5 that prevent surface expression of the coreceptor, such as the well-studied CCR5Δ32 mutation, confer near resistance to HIV-1 infection in individuals that carry these mutated alleles, underscoring the importance of CCR5 coreceptor tropism for transmitting viruses [15, 17-19]. Interestingly, preferential transmission of CCR5-tropic (R5) viruses occurs despite broad expression of CXCR4 on target $CD4^+$ T cells at transmission sites [20]. It therefore remains unclear what host or biological pressures are driving selection of R5 variants during HIV-1 transmission.

While CCR5 usage is the most characterized and consistent feature of T/F viruses, these variants also share additional attributes. For instance, the envelope (Env) proteins of transmitted viruses have shorter and less glycosylated variable loops, though this trend has been less consistent with clade B infections [6, 7]. Moreover, transmission seems to select for variants with increased viral and replicative fitness [21, 22] as well as those that are less sensitive to inhibition by interferon (IFN) [23-25], a component of the innate immune response against viral infection. Though a transmission bottleneck that selects for viral fitness and resistance to innate immune restriction is biologically plausible, the data are inconsistent across studies [26] and thus additional evidence with larger cohorts from various clades and geographical regions is required to understand whether these are obligatory features of T/F variants.

Ultimately, it is clear that HIV-1 variants that are successful in the transmission bottleneck have at least one shared characteristic, CCR5 coreceptor tropism, and may also

share the features of short variable loops, low levels of Env glycosylation, high replicative fitness, and partial resistance to IFN-mediated inhibition. These similarities suggest that there are consistent selective pressures at play during transmission and that this process is not entirely stochastic. Nevertheless, it remains unknown why certain T/F variant features are selected and whether there are additional viral characteristics that determine transmission likelihood. Understanding the naturally occurring restrictive pressures that bottleneck HIV-1 transmission – from the overall low frequency of mucosal HIV-1 transmission to the preferential transmission of a single T/F virus with certain features – remains a high priority as this might inform innovative interventions to block new HIV-1 infections. Experimental models of HIV-1 transmission and acute infection should focus on the subset of R5-tropic viruses isolated near the time of transmission to best recapitulate the biology of this stage of HIV-1 infection.

HIV-1 pathogenesis during acute infection

Once viral production at the initial foci of infection has sufficiently amplified, virus disseminates to the draining lymph node and secondary lymphoid tissues and establishes a self-propagating systemic infection [6]. Systemic infection, denoted by the detection of virus in the peripheral blood, occurs roughly 1-2 weeks after transmission and signals the end of the first phase of HIV-1 infection, termed the “eclipse phase”, as the infection is clinically silent during this period. Infection continues to rapidly expand and cause immune activation and massive depletion of $CD4^+$ T cells, particularly in the gut but also from the circulation and lymph nodes [27]. Plasma viral load exponentially increases during this period, until it reaches a peak about 2-4 weeks after infection. HIV-1-specific $CD8^+$ T cells responses first appear right around this time and continue to increase over the next few weeks, correlating with a decrease in plasma viral load [28]. Within the first 3-6 months of infection, plasma viral load decreases to a “set point” plateau, a balance between restrictive host immune responses and virus propagation [28].

While a peak in plasma viral load is a hallmark of early HIV-1 infection, the magnitude of this peak varies considerably between individuals. In a study that assessed two or more blood collection time points in the early window of acute infection prior to the detection of HIV-specific antibodies, peak viremia ranged from 4.5 to 8.5 $\log_{10}$ copies per milliliter across 50 individuals – a 10,000-fold difference [29]. Notably, peak viral load positively correlated with viral load set point in this dataset. Because a higher set point viral load is predictive of faster disease progression and mortality [30-33], this data suggests that peak viral load may also impact disease outcomes. Unlike set point viral load, which is established in part due to host adaptive immune pressures, peak viral load occurs prior to or concurrently with the onset of adaptive immune responses and is therefore not likely predominantly influenced by these factors. This suggests that acute infection replication dynamics, virus-host interactions, and innate immune pressures may determine the magnitude of peak viral load and impact clinical disease outcomes, though these relationships have not yet been elucidated.

Emergence of CXCR4 tropism in chronic infection

Interestingly, the viral population at the time of peak plasma viral load remains fairly homogenous, closely matching the R5-tropic T/F genetic variant [28]. However, mutations rapidly accumulate after the onset of HIV-1-specific $CD8^+$ T cell responses and directly correspond to amino acids in $CD8^+$ T cell epitopes [28]. Around 12 weeks post infection, mutations in Env that confer escape to neutralizing antibodies also appear, with stochastic mutations accumulating all the while [28]. Therefore, following peak plasma viral load, HIV-1 undergoes substantial genetic diversification, which continues into chronic infection (>100 days post infection) [28].

For many untreated HIV-1 infections, the high extent of genetic diversity results in functional diversity. Most notably, some viral variants begin to utilize CXCR4 as a coreceptor for viral entry as opposed to CCR5. The mutations that confer coreceptor switching have been

mapped to the third variable loop (V3) of Env [34-36], a target of both cellular and humoral immune responses [37, 38]. The emergence of CXCR4-tropic (X4) viruses is associated with the rapid decline of $CD4^+$ T cells and progression to AIDS [34, 39, 40]. Whether mutations conferring CXCR4 usage are specifically selected for via a virological or immunological pressure or are simply a byproduct of a high mutation rate and immune escape remains unclear [19].

A study from 2010 assessed the frequency of X4 tropism in 498 individuals living with chronic HIV-1, with strains representing six different clades, by aggregating data from 29 different studies. CXCR4 usage was found in 66% of non-clade C infections (clade A: 60%, clade B: 67%, clade D: 61%, CRF01_AE: 77%, CRF02_AG: 65%), with clade C infections having a lower frequency of 15% [35]. However, the occurrence of X4 tropism in clade C infections may be evolving. In the same analysis, only 8% of clade C samples collected prior to 2000 demonstrated CXCR4 usage, compared to 33% in clade C samples collected after 2000, suggesting that X4 tropism may be increasing in the clade responsible nearly half of all HIV-1 infections [35]. Therefore, although there are differences in the frequency of X4 tropism between clades, X4 HIV-1 viruses are characteristic of many chronic stage HIV-1 infections.

Host restriction factors and early HIV-1 infection

Because the course of early HIV-1 infection, from transmission through peak viral load, occurs prior to the onset of effective adaptive immune responses, it is plausible that interactions between HIV-1 and antiviral cellular factors intrinsically expressed in HIV-1 target cells, $CD4^+$ T cells, may contribute to the inefficient and variable nature of this period. In fact, several germline-encoded host proteins with potent, inherent antiviral activity against HIV-1 have already been described. Many of these studies have identified host- or cell type-specific antiviral factors and have found that some of these factors are rendered nonfunctional in certain contexts due to antagonism by HIV-1 accessory proteins or HIV-1 evolving to escape their restriction. These studies have thus helped uncover how viral proteins and cellular factors interact to either

permit or restrict HIV-1 replication. The host proteins that impact natural HIV-1 infection in humans, where $CD4^+$ T cells are the main host cell type targeted for viral replication, are thus likely to be those that are active in $CD4^+$ T cells, not antagonized by HIV-1, and for which HIV-1 has not evolved to escape. The aim of this thesis is therefore to identify the host factors that actively inhibit HIV-1 infection in primary $CD4^+$ T cells, as these genes may influence early viral infection and replication dynamics.

HIV-1 restriction factors can be generally categorized based on expression patterns – those whose expression is triggered by the type I IFN innate immune response to viruses, such as IFN-stimulated genes (ISGs), and those that are endogenously expressed. The following sections will explore what is known about HIV-1 restriction factors, how they were identified, and their potential role for impacting HIV-1 transmission and acute infection.

Type I IFN and HIV-1 infection

Type I IFNs, which include IFNα and IFNβ, are small signaling proteins secreted by host cells upon innate immune sensing of viral infection (type I IFNs reviewed in [41, 42]). Secreted IFNα and IFNβ can then bind to their cell surface receptor, termed the IFNα receptor (IFNAR), which initiates a signaling cascade that results in the formation and activation of the ISGF3 transcription factor complex comprised of STAT1, STAT2, and IRF9 (**Figure 1.1**). In the nucleus, this complex binds to its cognate DNA sequences, called IFN-stimulated response elements (ISREs), which are located upstream of hundreds of genes. ISGF3 binding triggers transcription of these IFN-stimulated genes (ISGs), many of which encode cellular proteins that block viral replication within target cells. Importantly, the proteins involved in type I IFN signaling (IFNAR, JAK1, TYK2, STAT1, STAT2, IRF9) are widely expressed, allowing nearly all cell types to mount this IFN-induced antiviral state.

Like other viruses, HIV-1 is sensitive to IFN-mediated restriction. Cell culture work has provided extensive evidence that type I IFN confers robust inhibition of HIV-1 replication [43-47], including in the main target cells of natural HIV-1 infection, primary $CD4^+$ T cells. Clinical studies administering type I IFN to chronically infected patients also report reduced viral loads [48-51]. Moreover, type I IFN is detectable in the plasma of individuals within days of primary HIV-1 infection, supporting the clinical relevance of IFN during this early window [52, 53]. Plasmacytoid dendritic cells are thought to be the main cellular sources of type I IFN in the blood during this period due to sensing of HIV-1 and HIV-infected cells by TLR7 [54], though IFI16 and cGAS have also been implicated as sensors of HIV-1 infection depending on cell type [55]. Overall, these findings suggest that the antiviral IFN response mediated by cellular restriction factors may pose a major barrier to HIV-1 during transmission and acute infection.

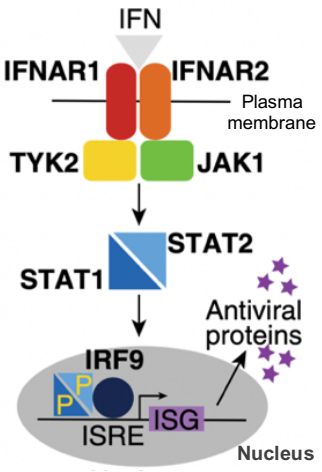


Figure 1.1. Type I IFN pathway schematic. Secreted IFNα/β binds to IFNAR, which is present on most nucleated cells, and triggers a signaling cascade that upregulates hundreds of ISGs, some of which encode antiviral proteins. Adapted from Richardson et al., 2018 [56].

IFN-stimulated HIV-1 restriction factors

Interferon-stimulated antiviral genes represent important drivers of the cell-intrinsic block to HIV-1 infection. Numerous studies have already discovered several HIV-1-restricting ISGs, including TRIM5α, tetherin/BST2, MX2, and more (HIV-1-restricting ISGs reviewed in [55, 57-60]). By elucidating the mechanisms and contexts of restriction for these cellular proteins, this vast body of research has informed several fundamental aspects of HIV-1 biology. For instance,

it was known that HIV-1 is species-specific, unable to infect Old World monkeys such as rhesus macaques. To identify the cellular target responsible for this barrier, investigators expressed a macaque complementary DNA (cDNA) library in human cells to select for a gene that conferred protection against HIV-1 but not the related retroviruses murine leukemia virus (MLV) or simian immunodeficiency virus (SIV) [61]. This led to the discovery of TRIM5α, an IFN-stimulated E3 ubiquitin ligase that binds to the capsid of retroviruses to trigger premature uncoating of capsid, which disrupts reverse transcription. TRIM5α thus represents a major determinant of the block against HIV-1 infection in Old World monkeys.

Just as the TRIM5α investigations shed light on a cellular factor that influences the host range of HIV-1, the discovery of tetherin, also known as BST2, informed the function of an HIV-1 accessory protein. Studies had observed that the HIV-1 accessory protein Vpu was required for HIV-1 to fully replicate in certain cell lines (considered non-permissive) but not others (permissive). An investigation into the cellular target of Vpu determined the genes uniquely expressed in non-permissive cell lines, ectopically expressed those genes in permissive cells, infected with Vpu-deficient HIV-1, and selected candidates that conferred restriction [62]. The target of HIV-1 Vpu was found to be tetherin/BST2, a transmembrane protein that prevents the release of budding HIV-1 virions from the surface of cells. Tetherin/BST2 is thus an ISG that presents such a potent barrier to HIV-1 that the virus has had to evolve countermeasures in the form of an accessory protein to overcome its restriction.

TRIM5α and Tetherin/BST2 were identified in part due to a previous understanding of the hosts or cell types that could or could not support HIV-1 replication. Similarly, MX2 was discovered by researchers investigating why the magnitude of IFN restriction of HIV-1 differed among cell lines. These studies determined genes whose expression was IFN-inducible and uniquely elevated in cell lines capable of potent IFN-mediated inhibition of HIV-1 [63, 64]. MX2 was thus identified and demonstrated to restrict HIV-1 via ectopic expression and silencing

experiments. MX2 is a GTPase that directly interacts with the HIV-1 capsid to inhibit HIV-1 infection after reverse transcription [65, 66]. Therefore, mutations in the HIV-1 capsid can confer resistance to MX2 restriction.

Other ISGs including IFITMs [67], UBE2L6 [68, 69], MARCH2/8 [70, 71], CH25H [72], GPB5 [73], and IFI16 [74], have also been shown to inhibit HIV-1 in *in vitro* cell culture models via diverse mechanisms that act on various stages of the HIV-1 replication cycle. While IFITMs and CH25H interfere with viral entry, IFI16 prevents viral transcription, MARCH2/8 and GPB5 act on HIV-1 after translation, and UBE2L6 impacts viral assembly and release. Despite the extensive nature of this previous work, it is unclear if these ISGs are clinically relevant due to the nature of the investigations in which their function was described. A later section will consider the translational potential of these studies.

Endogenously expressed HIV-1 restriction factors

In addition to the cell-intrinsic mediators of the antiviral type I IFN response, there are also endogenously expressed genes that inhibit HIV-1. For instance, APOBEC3G is a cytidine deaminase that is packaged into budding HIV-1 virions and, upon infection of a new target cell, hypermutates viral DNA intermediates during reverse transcription, specifically causing a high frequency of G-to-A mutations in proviral DNA. These mutations often introduce premature stop codons or otherwise render new virions nonfunctional, strongly inhibiting HIV-1 replication. Other APOBEC family members have subsequently been identified and found to suppress HIV-1 infection. APOBEC3G was identified by studies aiming to find the antiviral gene antagonized by the HIV-1 accessory protein Vif, as it was found that Vif was required for HIV-1 infection in certain cell types but not others. APOBEC3G was found to the target of HIV-1 Vif [75] and thus represents a potent antiviral gene that is rendered less effective against HIV-1 due to viral antagonism.

SAMHD1 is another endogenously expressed host protein that impacts HIV-1 infection. SAMHD1 reduces the availability of cytosolic deoxynucleotide triphosphates (dNTPs) in non-dividing cells, which prevents the generation of proviral DNA during reverse transcription ([76-78], reviewed in [57]). Though its expression is induced by innate immune responses in some cell types, SAMHD1 is constitutively highly expressed in $CD4^+$ T cells, monocytes, and dendritic cells. Notably, its restrictive activity can be overcome in activated cells, as these cells have an abundant supply of cellular dNTPs and activation causes the phosphorylation of SAMHD1, which inhibits its antiviral activity [79]. Therefore, SAMHD1 is in part responsible for the non-permissive nature of resting $CD4^+$ T cells, monocytes, and dendritic cells to HIV-1 infection, which substantially limits the prevalence of target cells during HIV-1 transmission and early viral propagation. SAMHD1's restrictive activity against HIV-1 was discovered by a study aiming to identify the restriction factor targeted by the HIV-2 accessory protein Vpx [77]. It was known that HIV-2, unlike HIV-1, was able to infect myeloid cells and that ectopic expression of HIV-2 Vpx rendered myeloid cells permissive to HIV-1. SAMHD1 was found to be the main binding partner of HIV-2 Vpx in myeloid cells and, ultimately, the antiviral target of HIV-2 Vpx antagonism.

Similar to SAMHD1 and HIV-2 Vpx, the host inhibitory factors SERINC3 and SERINC5 (SERINC3/5) were identified via investigations into an HIV accessory protein, in this case HIV-1 Nef. HIV-1 Nef was known to influence HIV-1 replication and pathogenesis *in vivo* and enhance virion infectivity when packaged into budding virions, but the mechanism of Nef's effect on infectivity was not understood (reviewed in [80]). Researchers used Nef-intact and -deficient HIV-1 infections to determine cellular factors either uniquely packaged into budding Nef-deficient virions [81] or highly expressed in cell lines requiring Nef for HIV-1 infection [82]. Both approaches identified SERINC3/5 as the inhibitory factors counteracted by this accessory protein. In the absence of Nef, SERINC3/5 proteins are incorporated into new HIV-1 virions and disrupt the conformation of HIV-1 Env, thereby inhibiting viral entry into new target cells [83].

Though SERINC3/5 are unlikely to influence HIV-1 infection *in vivo* due to Nef antagonism, they exemplify constitutive HIV-1 restriction factors with a unique mechanism and robust effect.

The discovery and characterization of APOBEC3G, SAMHD1, and SERINC3/5 were prompted and guided by the knowledge that there were unknown antiviral host proteins targeted by HIV-1 Vif, HIV-2 Vpx, and HIV-1 Nef, respectively. While there are likely many other endogenously expressed antiviral host genes yet to be uncovered, identifying such genes without a guiding phenotype or hypothesis represents a broad, challenging endeavor that will require a high throughput method to interrogate the antiviral function of hundreds to thousands of host genes simultaneously.

Caveats of prior studies of HIV-1 restriction factors

The described host genes highlight the remarkable potency and diversity of HIV-1 restriction factors. Despite being germline-encoded, these cellular factors are able to influence the host range of HIV-1, the cell types that are permissive to HIV-1 infection, and HIV-1 protein evolution. While this body of work supports the robust nature of cellular restriction factors and how studying these genes impact our understanding of HIV-1 biology, these studies also highlight how important context is for both viral restriction and viral escape from restriction factors. For instance, SAMHD1 restriction is overcome in activated cells, thus is unlikely to impact HIV-1 in the main target cells of infection, activated $CD4^+$ T cells. MX2 potency is diminished with certain HIV-1 capsid mutations. Human TRIM5α is less potent against HIV-1 than rhesus TRIM5α. APOBEC3G and tetherin/BST2 are only expressed in certain cell lines and are less effective against full-length virus due to viral antagonism. The cellular factors that are thus able to inhibit HIV-1 replication vary in potency depending on the host, cell type, and HIV-1 strain due to viral antagonism and escape and variability in restriction factor expression and activity between hosts and cell types. To understand the role of cell-intrinsic factors on the

early stages of HIV-1 infection, it is therefore necessary to investigate their expression and inhibitory function directly in the main target cells of HIV-1 infection *in vivo*, primary human $CD4^+$ T cells, against the R5-tropic primary viruses that are characteristic of this stage of infection.

Despite the extensive amount of work that has gone into studying HIV-1 restriction factors, very few studies have utilized both primary $CD4^+$ T cells and primary isolates of HIV-1 in their *in vitro* experiments to understand the biological significance of these genes on early HIV-1 infection. First, the vast majority of these studies performed infections with lab-adapted X4 HIV-1 strains. Lab-adapted HIV-1 tends to be X4-tropic and is not as sensitive to CD4 expression levels as primary HIV-1 viruses [84], as a result of being passaged on cells expressing low levels of CD4. Primary HIV-1 isolates refer to viruses that were produced by culturing cells from an individual living with HIV-1 and thus reflect strains that contribute to *in vivo* replication and disease kinetics. X4-tropic, lab-adapted viruses are commonly used in *in vitro* settings because CXCR4 expression is more common on cell lines than CCR5 expression and transfection of lab-adapted viruses tends to yield high-titer virus than that of primary isolates. However, as previously discussed, R5 tropism is selected for during HIV-1 transmission and the viral population remains fairly homogenous through the advent of adaptive immune responses, around the time of peak viral load, which temporal window corresponds with detectable levels of IFN. Therefore, R5-tropic primary variants should be utilized when studying the role of cellular restriction factors, particularly ISGs, during early HIV-1 infection. Restriction factors exclusively tested against X4-tropic, lab-adapted strains will need to be evaluated in the context of R5-tropic, primary HIV-1 infection to understand their *in vivo* significance. However, HIV-1 diversity is far more extensive than the dichotomy of primary versus lab-adapted strains and CXCR4 versus CCR5 coreceptor usage. To elucidate the breadth and functionality of a cellular gene, it is ideal to test its inhibitory activity against several HIV-1 viruses, accounting for differences in tropism, source (primary/lab-adapted), clade, and geographic region.

In addition to the selection of virus, many studies investigating HIV-1 restriction factors, particularly those using high throughput methods, have relied on immortalized or cancer cell line models, as these tend to be far more experimentally tractable, low maintenance, readily available, and cost effective. As previously highlighted, the experiments that lead to the discovery and characterization of a novel HIV-1 restriction factor often require manipulating the expression of that gene to interrogate its function. These types of experiments have not been possible to perform in primary cells until recently, as a result of advances will be discussed in the next section. Though they are easy to experimentally alter, cell lines do not always faithfully recapitulate their primary cell counterparts. This was even evidenced in the identification of APOBEC3G – primary T cells could not support Vif-deficient virus but the SupT1 (T cell lymphoma) cell line could, which was found to be due a lack of APOBEC3G expression in SupT1s [62]. Similarly, in one of the original MX2 studies, MX2 expression via immunoblot was barely detectable in Jurkat T cells (T cell leukemia) despite clear expression in primary $CD4^+$ T cells [64]. In fact, the protective effect of type I IFN treatment is substantially weaker in T cell lines compared to primary $CD4^+$ T cells, primary macrophages, and the THP-1 monocytic cell line (monocytic leukemia) [46]. Many studies have therefore leveraged THP-1 cells to interrogate ISGs with activity against HIV-1. Due to the many differences in basal expression and IFN induction of host genes across various cell types and cell lines, it is the premise of this thesis that restriction factors that have been described in cell line models must be assessed in primary $CD4^+$ T cells to understand their role during natural HIV-1 infection.

Potential for high throughput studies in primary $CD4^+$ T cells

Due to our incomplete understanding of the host genes that inhibit HIV-1 *in vivo*, there remains a need to not only determine the potency of previously identified genes in a more translatable model, but to also evaluate a broader pool of antiviral candidates, which may lead to the identification of new HIV-1-restricting genes. The experimental system that would most

efficiently address this would involve interrogating the antiviral function of hundreds or thousands of genes simultaneously in primary $CD4^+$ T cells against a R5 primary virus. Fortunately, recent advances, namely a high throughput approach for evaluating HIV-1 restriction factors and techniques to manipulate $CD4^+$ T cells, have made this system a possibility.

The techniques utilized by prior HIV-1 restriction factor studies, described previously, were low throughput, often relying on comparing gene expression between permissive and non-permissive cells to select a few candidates or using cDNA libraries for ectopic gene expression. Though some high throughput methods were implemented, overexpression screens [85, 86] do not capture the function of genes at physiological levels and siRNA screens [87] require transfection of siRNA pools, which limits the cell types that can be assessed. A 2018 paper overcame these limitations by developing a CRISPR-Cas9 knockout screening approach, called HIV-CRISPR, in which a library of guide RNAs are delivered by lentivirus that has a repaired long terminal repeat region, such that, upon infection with replication-competent HIV-1, guides are transcribed and packaged into budding HIV-1 virions [69]. By deep sequencing and comparing the guides present in cells and in the viral supernatant, genes that restrict (over-represented in the viral supernatant) or promote (depleted from the viral supernatant) HIV-1 infection can be determined. Therefore, this system not only efficiently inactivates thousands of genes, but it also uses HIV-1 infection as the direct read out, as opposed to screens that use indirect read outs like cell death, which often introduce experimental noise. Of note, this approach was developed to evaluate the antiviral function of ISGs and thus utilized IFN-treated THP-1 cells for robust IFN restriction of HIV-1.

We reasoned that adapting HIV-CRISPR screens to be compatible with primary $CD4^+$ T cells would allow for high throughput assessment of cellular factors in a physiological relevant context. However, primary $CD4^+$ T cells have historically been difficult to manipulate. Lentiviral transduction to overexpress or ectopically express a gene is often prohibitively inefficient in

$CD4^+$ T cells. This not only prevents primary cells from being used in HIV-CRISPR screens, but also makes them incompatible with typical, targeted CRISPR-Cas9 approaches, as both methods deliver Cas9 and the guide RNA(s) by lentivirus to silence the gene(s) of interest. However, a 2019 *Nature Protocol* publication offered an invaluable tool to the HIV-1 field – a step-by-step protocol for the isolation, activation, editing, and infection of primary $CD4^+$ T cells that does not rely on lentiviral transduction [88]. Specifically, $CD4^+$ T cells are isolated from whole blood and are edited by nucleofecting, or electroporating, cells with pre-made CRISPR-Cas9 complexes as opposed to lentiviral delivery. This editing strategy is not common because, in cell line models that can be indefinitely propagated, it is advantageous to have constitutive expression of Cas9 and the guide RNA for continuous editing. Nevertheless, because primary cells can only be cultured for a few weeks, one round of editing is sufficient. This protocol thus enabled researchers to perform gene characterization experiments in primary $CD4^+$ T cells.

In 2022, this approach was essentially scaled up for the semi-high throughput assessment of 426 host factors implicated in the HIV-1 replication cycle [89]. CRISPR-Cas9 complexes were arrayed into wells of 96-well plates containing $CD4^+$ T cells, for a total of eighteen 96-well plates. Though this strategy was convincing in its designation of positive and negative controls, it is limited in the number of genes that can be evaluated due to the amount of plates that would need to be managed, and it is likely too expensive for most labs to implement due to the high cost of both Cas9 protein and purified RNA.

Another group experimented with methods to edit primary $CD4^+$ T cells on a larger scale and found that they achieved moderate editing when they combined some of the above techniques – delivering the guide RNA by lentivirus but nucleofecting in Cas9 protein [90]. For the scale of high throughput screens, delivering guides by lentivirus as opposed to purchasing purified RNA dramatically reduces cost. The drawback of this approach, however, is in its editing efficiency, as it yielded only moderate knockout scores. This is likely due to the inefficiency of lentiviral delivery of guide RNA.

The final fundamental technical challenge to overcome in primary $CD4^+$ T cells thus became efficient lentiviral delivery. By significantly modifying the vector commonly used for lentiviral delivery of guides (lentiGuide), a 2020 paper developed the TOP vector (T cell-optimized for packaging), a lentiviral vector specifically optimized to deliver guide RNA sequences to primary T cells [91]. This vector not only improved transduction efficiency, which is necessary for guide delivery, but also increased the titer of the lentivirus, which is important for high throughput screens in order to maintain coverage. With this technical advancement, it thus became conceivable that HIV-CRISPR screens could be performed in primary $CD4^+$ T cells by transducing cells with a library of guides via the TOP vector and nucleofecting in Cas9 protein.

Thesis overview

Overall, the goal of this thesis is to identify cellular factors, those that are IFN-stimulated or constitutively expressed, that inhibit replication of primary HIV-1 viruses in primary $CD4^+$ T cells, as this may shed light on the host barriers that contribute to the inefficiency and variability of acute HIV-1 infection *in vivo*. Chapter 2 will investigate the function of antiviral ISGs by adapting HIV-CRISPR screens for use in primary $CD4^+$ T cells and interrogating the function of genes specifically IFN-stimulated in this cell type using a custom single guide RNA library. Also leveraging the primary cell HIV-CRISPR approach, Chapter 3 will describe the identification of an endogenously expressed host gene, SLC35A2, and will propose an important, previously underappreciated role for host cell glycosylation on restricting X4 HIV-1 infection while supporting that of R5 viruses. Finally, Chapter 4 will discuss the implications of these findings and priorities for future investigations.

Chapter 2

SEVERAL CELL-INTRINSIC EFFECTORS DRIVE TYPE I INTERFERON-MEDIATED RESTRICTION OF HIV-1 IN PRIMARY CD4$^+$ T CELLS

Sections of text in this chapter have been modified from the following manuscript:

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).

Abstract

Type I IFN upregulates proteins that inhibit HIV-1 within infected cells. Prior studies have identified ISGs that impede lab-adapted HIV-1 in cell lines, yet the ISG(s) that mediate IFN restriction in HIV-1 target cells, primary CD4$^+$ T cells, are unknown. Here, we interrogate ISG restriction of primary HIV-1 in CD4$^+$ T cells by performing CRISPR-knockout screens with a custom library that specifically targets ISGs expressed in CD4$^+$ T cells. Our investigation identifies previously undescribed HIV-restricting ISGs (HM13, IGFBP2, LAP3) and finds that two factors characterized in other HIV-1 infection models (IFI16, UBE2L6) mediate IFN restriction in T cells. Inactivation of these five ISGs in combination further diminishes IFN's protective effect against diverse HIV-1 strains. This work demonstrates that IFN restriction of HIV-1 is multifaceted, resulting from several effectors functioning collectively, and establishes a primary cell ISG screening model to identify both single and combinations of HIV-restricting ISGs.

Introduction

The type I IFN response is one of the first defenses of the host innate immune system that viruses must overcome during transmission and acute infection. IFN signaling, triggered by viral infection, upregulates hundreds of ISGs, many of which inhibit viral replication in a cell-intrinsic manner within infected cells. The ISGs that are effective at preventing viral replication

vary in their magnitude of restriction as well as by virus, host, and cell type [42]. Like other viruses, HIV-1 is sensitive to IFN restriction. Cell culture work has extensively characterized type I IFN-mediated inhibition of HIV-1 replication [43-47] and studies of IFN treatment in chronically infected patients report reduced viral loads [48-51], demonstrating the efficacy of IFN restriction *in vivo*. Because type I IFN is secreted early after natural HIV-1 infection [52, 53] and inhibits HIV-1 replication *in vivo*, the IFN response may significantly contribute to HIV-1 transmission barriers, transmission bottlenecks, and early viral load differences, all of which influence patient outcomes.

Due to this clinical potential, there is considerable interest in identifying the cellular ISGs that inhibit HIV-1 replication. Numerous studies have pursued this goal and described HIV-1 restriction mediated by ISGs such as TRIM5α, IFITMs, SAMHD1, SLFN11, MX2, Tetherin, APOBEC3s, and many others (reviewed in [57-59]). This extensive body of research has informed determinants of species-specificity for HIV-1 and related retroviruses [61, 92-94] and uncovered diverse mechanisms of cell-intrinsic inhibition of HIV-1. However, two key issues remain that limit our understanding of the proteins that are clinically relevant to IFN-mediated HIV-1 restriction. First, the majority of this work was performed in cell lines with lab-adapted HIV-1 strains or pseudoviruses [63-65, 69, 71, 73, 74, 95-99], and the restrictive capacity of some ISGs has varied based on these experimental conditions [69, 100-103]. Thus, to understand the role of these factors in HIV-1 pathogenesis, the antiviral effects of previously described HIV-restricting ISGs need to be examined in the main cell type that supports HIV-1 replication *in vivo*, $CD4^+$ T cells, against replicating viruses that have not been lab-adapted (primary viruses) and that best represent globally circulating viruses. Second, the reported inhibition of known ISGs does not fully account for the overall magnitude of IFN restriction of HIV. This suggests that the ISGs studied thus far exhibit more pronounced restriction in primary target cells, additional HIV-restricting ISGs have yet to be identified, and/or full IFN restriction

results from several ISGs working synergistically to mount an effect greater than that of each alone.

To determine the ISGs that have relevance restricting HIV-1 in primary target cells, we inactivated candidate ISGs via CRISPR-Cas9 editing in primary $CD4^+$ T cells and infected edited cells with two diverse replication-competent HIV-1 strains, including a primary virus and a lab-adapted strain. We first targeted eight ISGs previously studied in other infection systems but found that only one gene (IFI16) contributed to detectable IFN restriction in our model. To identify additional candidate ISGs, we performed comprehensive CRISPR-knockout (KO) screens in $CD4^+$ T cells using a custom library reported here (CD4-ISG library) that targets ISGs specifically upregulated in this important cell type. Screens with the CD4-ISG library identified several candidate ISGs. Half of the eight hits we pursued further (HM13, IGFBP2, LAP3, UBE2L6) demonstrated IFN-dependent restriction of the primary HIV-1 strain, including three factors not previously known to inhibit HIV. Because each IFN effector alone did not explain the entirety of IFN restriction, we inactivated all five ISGs in $CD4^+$ T cells and found that IFN inhibition was more fully ablated as additional genes were targeted. These findings demonstrate that IFN blocks HIV-1 by upregulating several effector proteins that work collectively. This study also describes a valuable primary $CD4^+$ T cell ISG screening model for the identification of both single and combinations of antiviral ISGs, including factors that have been overlooked in culture-adapted models of HIV-1 infection.

Results

Ex vivo $CD4^+$ T cell model to define ISG restriction of HIV

To identify effectors of IFN-mediated restriction of HIV-1 in primary target cells, we interrogated ISGs of interest by inactivating them via CRISPR-Cas9 editing in human $CD4^+$ T cells isolated from whole blood (**Figure 2.1A**). Activated $CD4^+$ T cells were nucleofected with commercial triple-guide CRISPR-Cas9 ribonucleoproteins (RNPs) to achieve high levels of ISG

KO. To recapitulate natural infection, we established spreading infections (MOI=0.02) with a replication-competent, R5-tropic primary HIV-1 strain (Q23.BG505, clade A) [104, 105]. To better align our results with previous studies, we also infected with a commonly used X4-tropic lab-adapted HIV-1 strain (LAI, clade B) [106] for comparison. IFN-β was chosen as the type I IFN for this model as it elicits a more robust, prolonged transcriptional response as compared to IFN-α [107]. To capture the effect of ISGs on any stage of viral replication, we evaluated infection by measuring reverse transcriptase (RT) activity in the supernatant.

In non-targeting control (NTC)-delivered $CD4^+$ T cells from several donors, both viruses exhibited consistent sensitivity to IFN treatment (Q23.BG505 median: 62-fold reduction; LAI median: 55-fold reduction; **Figure 2.1B and 2.1C**). Editing of IRF9, a component of the ISGF3 transcription factor complex that enables expression of ISGs, served as a positive control and led to high KO levels for all donors (median: 94.5% KO, range: 62-100% KO; **Supplementary Figure 2.1A**). As expected, IRF9 KO effectively blocked ISG protein upregulation as measured by Western blot analysis of the MX2 ISG (**Supplementary Figure 2.1B**), and completely ablated the inhibitory effect of IFN for both viruses (**Figure 2.1B and 2.1C**). Of note, we did not observe a substantial increase in HIV-1 levels with IRF9 KO in the absence of IFN treatment, which suggests that infected cells are not endogenously secreting inhibitory levels of type I IFN in this system (**Supplementary Figure 2.1C and 2.1C**). This *ex vivo* model of $CD4^+$ T cells infected with an IFN-sensitive primary HIV-1 strain therefore represents a robust, biologically tractable approach for measuring the contribution of ISGs to IFN restriction of HIV.

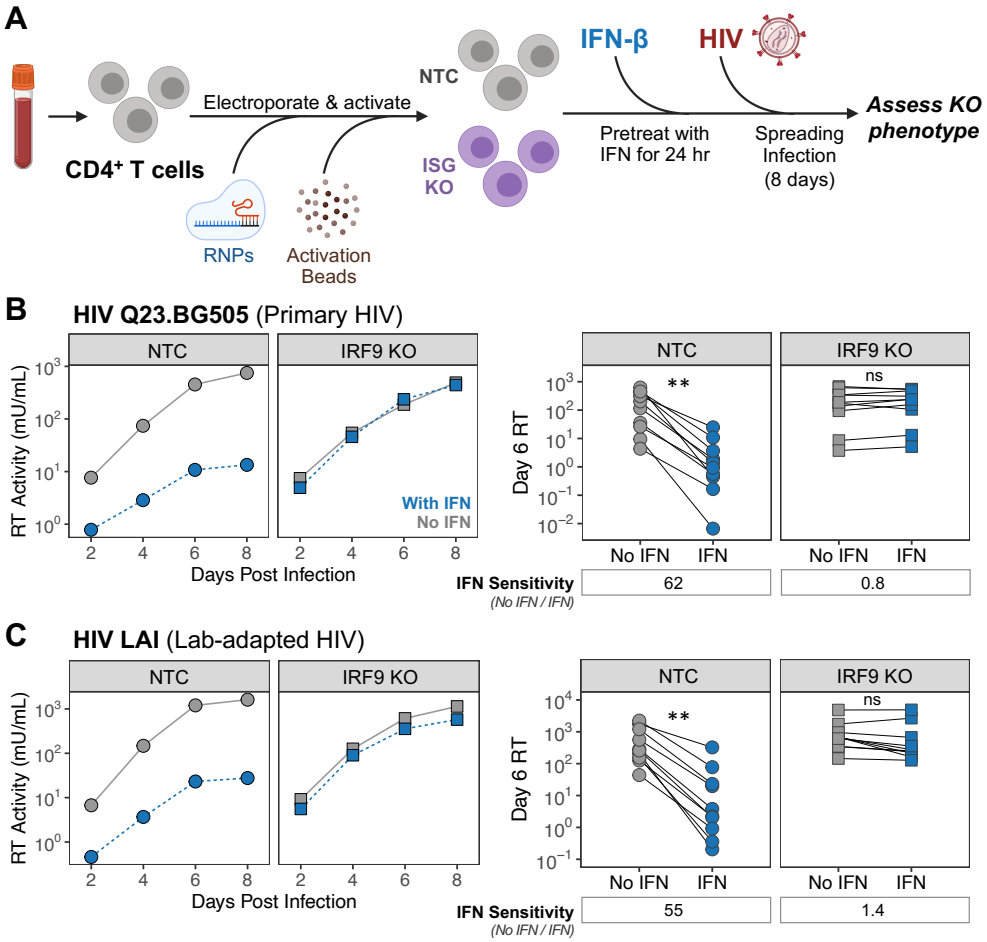


Figure 2.1. Ex vivo CD4+ T cell model to define ISG restriction of HIV.
(A) Schematic of approach for evaluating ISG function in primary CD4$^+$ T cells. (B, C) NTC-delivered and IRF9-edited CD4$^+$ T cells were infected HIV-1 strains Q23.BG505 and LAI, respectively, at an MOI of 0.02. Reverse transcriptase (RT) activity was measured over the course of infection. Left, representative infection curves from one donor from infections with and without IFN treatment (With IFN, No IFN). Right, 6 days post infection (dpi) RT activity from infections with and without IFN from 10 independent experiments that used 8 unique donors. Dot plots are paired by experiment and the median IFN sensitivity across experiments is depicted below (IFN sensitivity = untreated day 6 RT / with IFN day 6 RT). **p<0.002, paired Wilcoxon rank test. Non-significant (ns) indicates p>0.05. See also **Supplementary Figure 2.1**.

IFI16 contributes to IFN restriction of primary HIV-1 in CD4$^+$ T cells

We first interrogated the inhibitory function of eight proteins previously characterized as possible effectors of IFN restriction of HIV-1 in other systems. IFI16, IFITM1, MX2, and TRIM5α were selected based on their ability to confer cell-intrinsic restriction of HIV-1 in studies using lab-adapted viruses, pseudoviruses, or cell line models [61, 63-65, 67, 69, 74, 99, 108, 109]. Of

note, IFITM1 has been evaluated in primary $CD4^+$ T cells in one study that found it contributed to IFN restriction of one out of two IFN-sensitive primary strains [110]. APOBEC3A (A3A), BCL2L14, CMPK2, and IFI27, as well as the previously selected IFITM1 and MX2, were targeted because their increased expression in $CD4^+$ T cells correlated with decreased plasma viral loads in people living with HIV-1 treated with IFN [50].

Because of the inherent variability of primary cells from different donors, we performed KO experiments with 4-5 donors for each of the eight genes of interest. Nucleofection with Cas9 RNPs yielded median gene KO scores ranging from 79% (A3A) to 100% (IFI16, IFITM1, MX2), with all genes and donors exceeding 30% KO, a previously established functional cutoff for genomic editing [89] (**Figure 2.2A**). Negative control cells were IFN-treated and infected in parallel to ISG KOs to contextualize KO results. To determine whether each ISG meaningfully contributed to IFN inhibition, we applied two criteria based on IFN restriction and HIV-1 infection phenotypes at 6 days post infection (dpi), as illustrated in **Supplementary Figure 2.2A**. First, to assess whether ISG KO is contributing to IFN restriction, we required that ISG KO must increase HIV levels in the context of IFN treatment by at least two-fold as compared to the intra-assay negative control. We refer to this metric as “HIV Levels with IFN”. In cases where this condition was met, we applied a second criterion to eliminate genes whose inhibition is not specifically associated with IFN. For this, we calculated each KO’s relative IFN sensitivity, which is the ratio of gene KO without IFN to gene KO with IFN. We required that the IFN sensitivity of ISG KO must be at least two-fold lower than that of the negative control. We refer to this metric as “Relative IFN Sensitivity”. These criteria effectively select for genes that are IFN effector proteins as opposed to non-ISG restriction factors or genes whose KO impairs HIV infection (**Supplementary Figure 2.2B**). For instance, KO of a non-ISG restriction factor would increase HIV Levels with IFN, but would not decrease Relative IFN Sensitivity, because infection levels would be increased by the KO independent of IFN treatment. Likewise, if gene KO hinders HIV replication, such as in cases of diminished cell growth or viability, infection levels both with and

without IFN would be decreased, which may misleadingly decrease Relative IFN Sensitivity. By first requiring increased HIV Levels with IFN, this gene would not be categorized as an IFN effector. Importantly, this analysis was applied in order to capture the biological importance of ISGs in datasets with limited and variable donor replicates, as statistical analysis is less informative in this context. This approach was used to make comparisons about ISG restriction between viruses and ISGs, recognizing that additional donor replicates may be required to detect small effect sizes and to definitively conclude that an ISG is an IFN effector or not for a given virus. IRF9 KO appropriately meets the criteria of an IFN effector (**Supplementary Figure 2.1E**), causing a 160-fold increase in HIV Levels with IFN and a 78-fold decrease in Relative IFN Sensitivity for Q23.BG505.

As expected for primary $CD4^+$ T cells, we observed some variability in infection phenotypes between donors, such as A3A, BCL2L14, CMPK2, and IFI27 KOs demonstrating much stronger impacts on Q23.BG505 HIV Levels with IFN in Donor 2 as compared to the four other donors (**Figure 2.2B**). We therefore evaluated median fold effects of the 4-5 donors tested by our two analysis criteria to select ISGs with restrictive activity across donors. Surprisingly, only IFI16 met both criteria for the primary HIV-1 strain Q23.BG505, with knockout causing a median 2.3-fold increase in HIV Levels with IFN, which corresponded to a median 4.2-fold decrease in Relative IFN Sensitivity (**Figure 2.2B and 2.2C**). IFI16, as well as A3A, CMPK2, MX2, and TRIM5α, met both criteria for LAI. MX2 KO had the largest impact on the Relative IFN Sensitivity for this virus (median 4-fold decrease). These data suggest that IFI16 is a broad effector of IFN restriction of HIV-1 and that other tested ISGs, most notably MX2, IFITM1, and TRIM5α, have less relevance in primary target cells against a primary virus than in the models in which they were previously described.

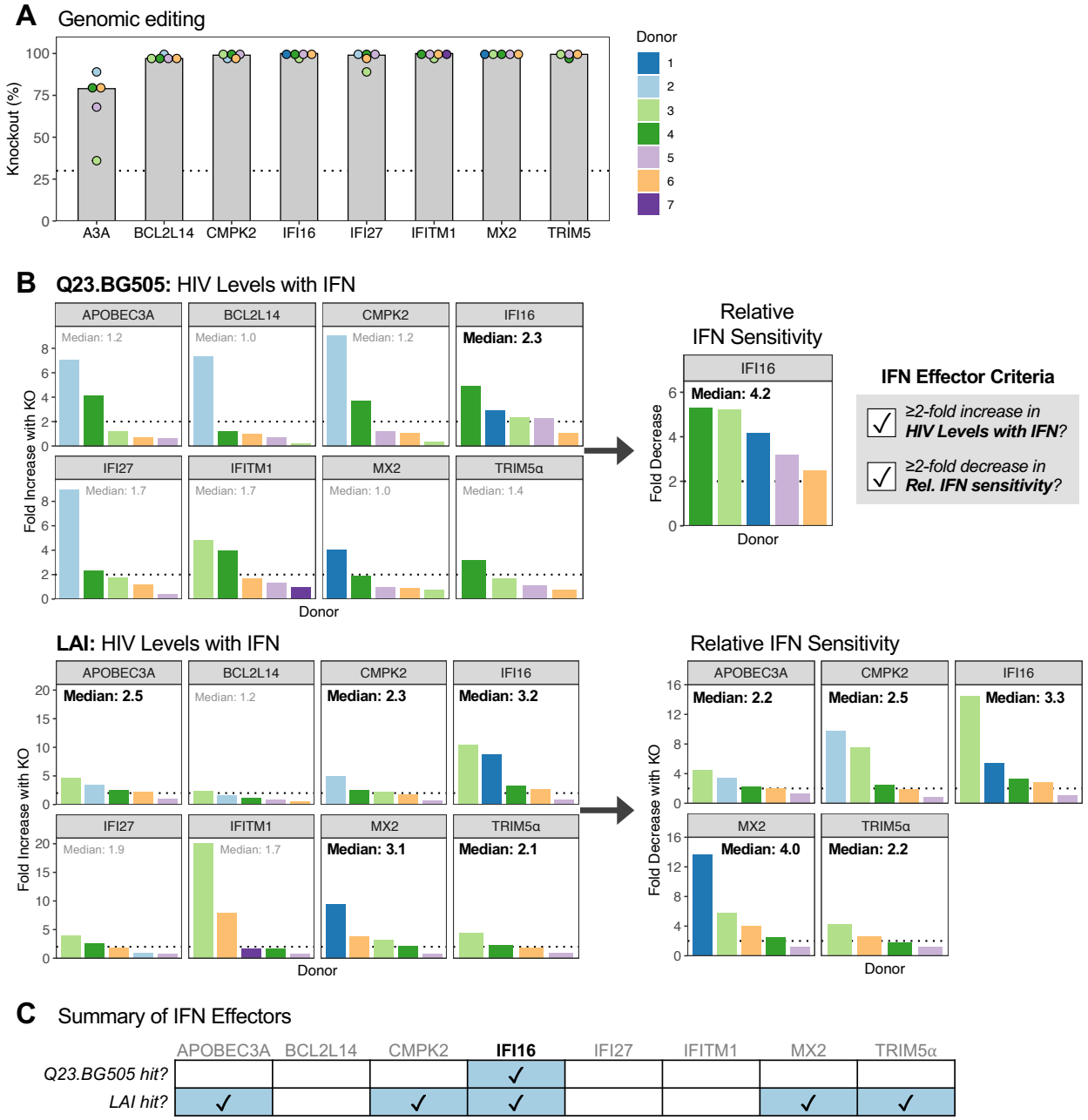


Figure 2.2. IFI16 contributes to IFN restriction of primary HIV-1 in CD4⁺ T cells.
(A) Median percent KO in nucleofected CD4⁺ T cells across 4-5 donors. Dotted line denotes 30% KO. Donor color is consistent within the figure but does not correspond with other figures unless otherwise noted. (B) ISG-edited and NTC-delivered cells were treated with or without IFN and infected with Q23.BG505 (top) and LAI (bottom). Left, increase in HIV Levels with IFN (6 dpi RT levels) with ISG KO as compared to NTC cells (ISG KO/NTC). The median increase across donors is depicted at the top of each graph. Dotted line indicates a fold difference of two. Right, for genes on the left with median increases in HIV Levels with IFN ≥2-fold, the decrease in Relative IFN Sensitivity with ISG KO compared to NTC is shown (NTC/ISG KO). (C) Summary of ISG KO infection phenotypes. See also **Supplementary Figures 2.2 and 2.3**.

Custom CRISPR sgRNA library to target ISGs in $CD4^+$ T cells

Though IFI16 KO reduced the Relative IFN Sensitivity of Q23.BG505 (median 4.2-fold, **Figure 2.2B**), it only captured a fraction of the effect seen with IRF9 KO (median 78-fold, **Supplementary Figure 2.1E**). This discrepancy indicates that other ISGs must contribute to IFN inhibition. To evaluate the restrictive capacity of all ISGs expressed in primary $CD4^+$ T cells, we sought to carry out comprehensive, high throughput CRISPR-KO screens with a customized CRISPR sgRNA library relevant to these cells. We therefore characterized ISG expression in our *ex vivo* model by performing bulk RNA-seq on primary $CD4^+$ T cells collected 0, 3-, 6-, 12-, and 24-hours post IFN treatment (**Figure 2.3A, Supplementary Table 2.1**). Principle component analysis revealed separation between IFN-treated and untreated cells and appreciable differences between collection time points (**Figure 2.3B**). We therefore assessed sampling times individually. For results from 6, 12, and 24 hours post-IFN, we selected genes that were IFN-stimulated by at least two-fold in the majority of donors for a given time point (**Figure 2.3C**). Because we only had one donor with cells collected at 3 hours post-IFN, we applied a more stringent cutoff of four-fold induction for this sample. This approach yielded 261, 449, 261, and 167 genes from the 3-, 6-, 12-, and 24-hour time points, respectively, with many ISGs overlapping across sampling times (**Figure 2.3D**). To create the sgRNA library, we designed eight guides per gene and included 200 NTC guides as controls. The final "CD4-ISG" library targets 555 ISGs represented across 4,579 guides (**Supplementary Table 2.2**). We compared this library to a larger, less focused library used to identify HIV antiviral ISGs in other studies (PIKA [69]) and found that the majority of genes in the CD4-ISG library (57%) were not present in the PIKA library (**Figure 2.3E**). Likewise, only 12.5% of the PIKA library represents genes that are IFN-stimulated in primary T cells. The CD4-ISG library therefore encompasses a set of ISGs tailored specifically to primary $CD4^+$ T cells.

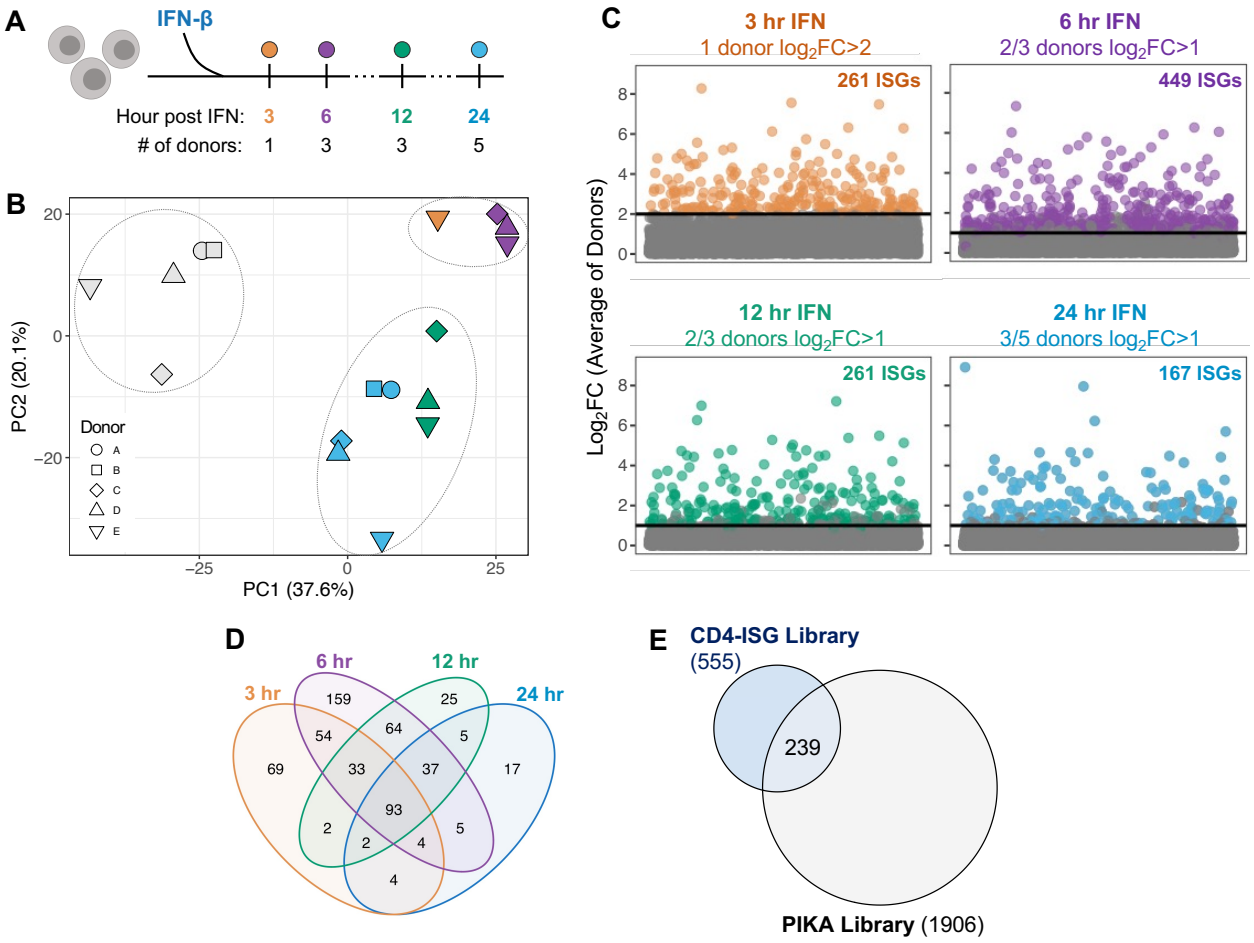


Figure 2.3. Custom CRISPR sgRNA library to target ISGs in CD4⁺ T cells.
(A) Schematic describing CD4⁺ T cells evaluated by bulk RNA-seq. (B) Principal component analysis of RNA-seq data. (C) To select ISGs to include in a new CRISPR library, thresholds were applied to each RNA-seq time point, as indicated by the solid line. Genes above thresholds are shown in color. Total ISG counts are included in each plot. (D) Venn diagram comparing ISGs selected in (C) by collection time. (E) Euler diagram comparing genes represented in the CD4-ISG libraries to those in the broad PIKA ISG library [69].

Strategy for HIV-CRISPR screens in primary CD4⁺ T cells

To perform CRISPR-KO screens with the CD4-ISG library in primary CD4⁺ T cells (**Figure 2.4A**), we leveraged the HIV-CRISPR approach, a robust strategy developed in the THP-1 monocytic cell line and reported in Ohainle *et al* [69]. HIV-CRISPR screens deliver sgRNA libraries to cells using a lentiviral vector with a repaired, complete HIV long-terminal repeat (LTR). This repair allows the integrated vector, including the sgRNA sequence, to be transcribed and packaged into budding virions during replication-competent HIV infection.

Therefore, guide sequences enriched in the supernatant viral RNA as compared to the cellular genomic DNA would be expected to target and inactivate genes that restrict HIV.

CRISPR screening in primary $CD4^+$ T cells presents three challenges that we had to overcome. First, lentiviral delivery of sgRNAs to $CD4^+$ T cells has been prohibitively inefficient in the past. To circumvent this issue, we utilized the TOP vector, which yields high-titer lentivirus with improved transduction of primary T cells and increased packageability as compared to the widely used lentiGuide construct [91]. To amend the TOP vector for HIV-CRISPR screening in primary T cells, the HIV LTR was repaired to restore packaging. The tracrRNA sequence was also optimized for more efficient sgRNA expression [111], an adjustment that was essential to achieve sufficient editing (data not shown). Second, lentiviral delivery of Cas9 is insufficient for high levels of editing in primary T cells [112]. We adopted a previously utilized strategy that delivers Cas9 to primary T cells by nucleofection of NTC Cas9 RNPs after lentiviral transduction with the guide library [90]. Lastly, most CRISPR screens transduce cells at low MOIs followed by antibiotic selection over several days to generate a pool of cells containing one sgRNA. Because primary T cells have short *ex vivo* culture windows, antibiotic selection is not practical. We therefore transduced cells at a high MOI, aiming to deliver roughly 3 vector copies per cell to maximize the fraction of cells expressing sgRNAs. These modifications enabled us to perform HIV-CRISPR screens in primary $CD4^+$ T cells.

HIV-CRISPR screens with the CD4-ISG library identify HIV-restricting ISGs

We conducted HIV-CRISPR screens with the custom CD4-ISG library in $CD4^+$ T cells from two donors (**Figure 2.4A**). Lentiviral transductions delivered 3.0 and 5.2 vector copies per cell for each donor, respectively, and library coverage was maintained (>1000-fold) at every stage. We infected transduced cells with HIV-1 at higher MOIs (0.1, 1) in screens than in validation experiments (MOI=0.02) to ensure sufficient read out. Genomic DNA and viral RNA deep sequencing was analyzed by MAGeCK-Flute [113] (**Supplementary Table 2.3**) to quantify

gene-level results and identify guides enriched in the supernatant, which would reflect the inactivation of restriction factors. Positive control IFN signaling genes (IRF9, STAT1, STAT2, IFNAR1) were highly enriched in all screens and we observed strong agreement between donors (Pearson R: 0.83-0.89; **Figure 2.4B**). A high proportion of genes (20.5-30%) scored above background for both donors, with LAI screens yielding slightly more hits (Q23.BG505 MOI 0.1: 114/555; Q23.BG505 MOI 1: 134/555; LAI MOI 0.1: 140/555; LAI MOI 1: 166/555). Overall, 56 ISGs scored above background in all eight screens.

We next aimed to validate genes with the highest likelihood of having broad restrictive activity in primary $CD4^+$ T cells (**Figure 2.4C**). To select this subset, we rationalized that genes with consistent robust expression across donors would be less likely to arise as donor-specific artifacts. Therefore, from the list of 56 top-scoring genes (**Supplementary Table 2.4**), we filtered out those with low average counts per million (CPMs). We also removed one long non-coding RNA, as they are often not amenable to CRISPR approaches, and one ISG that we had previously interrogated (IFITM1, **Figure 2.2**). From the remaining candidates, we identified and removed strain- or donor-specific hits: LAI-specific BST2 (Tetherin), Q23.BG505-specific BATF and OASL, and donor-specific IRF4 (**Supplementary Figure 2.4**). We selected eight genes to validate based on MAGeCK scores and putative function, as we aimed to identify IFN effectors that directly restrict HIV. Despite being the top-scoring hit, NLRC5 was not chosen for validation as it is known to regulate, as opposed to mediate, host antiviral responses [114].

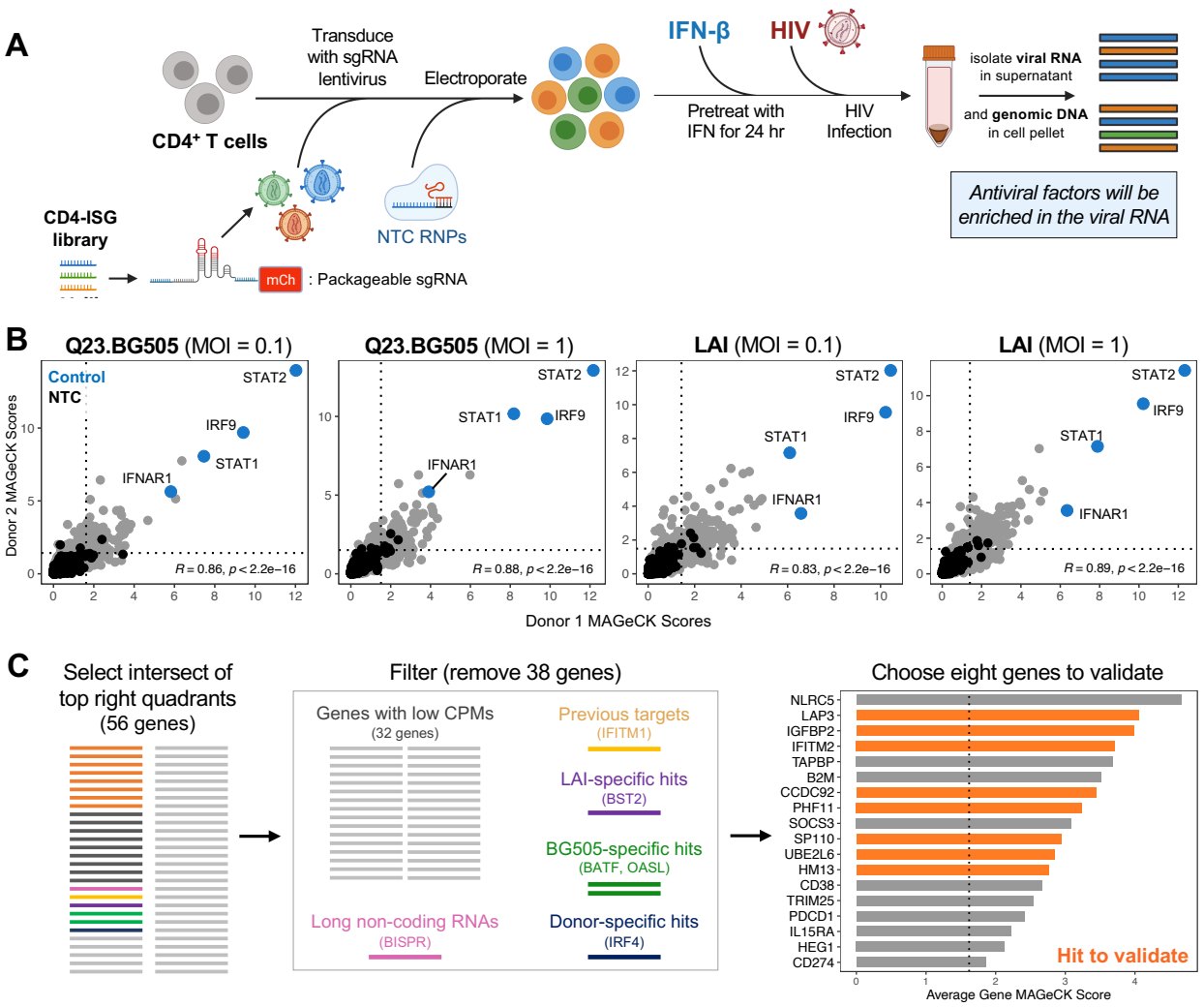


Figure 2.4. HIV-CRISPR screens with the CD4-ISG library identify HIV-restricting ISGs. (A) Schematic of HIV-CRISPR screens. (B) HIV-CRISPR screens were conducted in CD4$^+$ T cells from two donors. Gene-level MAGeCK enrichment scores were correlated between donors for each virus and MOI. Dotted lines indicate the NTC average + 3 standard deviations for individual screens. Pearson correlation coefficients are included within plots. (C) Strategy for selecting eight CD4-ISG screen hits to validate. Right, the average MAGeCK score across all eight screens for the top 18 screen hits. Dotted line denotes the highest background threshold from panel B plots. See also **Supplementary Figure 2.4**.

ISG editing exceeded 30% KO for all genes (**Figure 2.5A**) and did not impact cell growth (**Figure 2.5B**), indicating that these ISGs did not artificially arise as screen hits due to their inactivation causing cell death. Compared to the negative control, inactivation of HM13, IGFBP2, LAP3, and UBE2L6 increased Q23.BG505 HIV Levels with IFN ≥2-fold, which corresponded with ≥2-fold decreased Relative IFN Sensitivity, which results categorized them

as IFN effectors by our analysis approach (**Figure 2.5C**). Editing of HM13 and UBE2L6 also met both IFN effector criteria for LAI, suggesting that these genes may broadly restrict diverse HIV-1 strains. Though IFITM2 KO did not increase Q23.BG505 HIV Levels with IFN, it did demonstrate IFN-mediated restrictive activity against LAI, though results were variable between donors. Ultimately, half of the CD4-ISG screen hits we interrogated (4/8) validated as mediators of IFN restriction of a primary virus (**Figure 2.5D**).

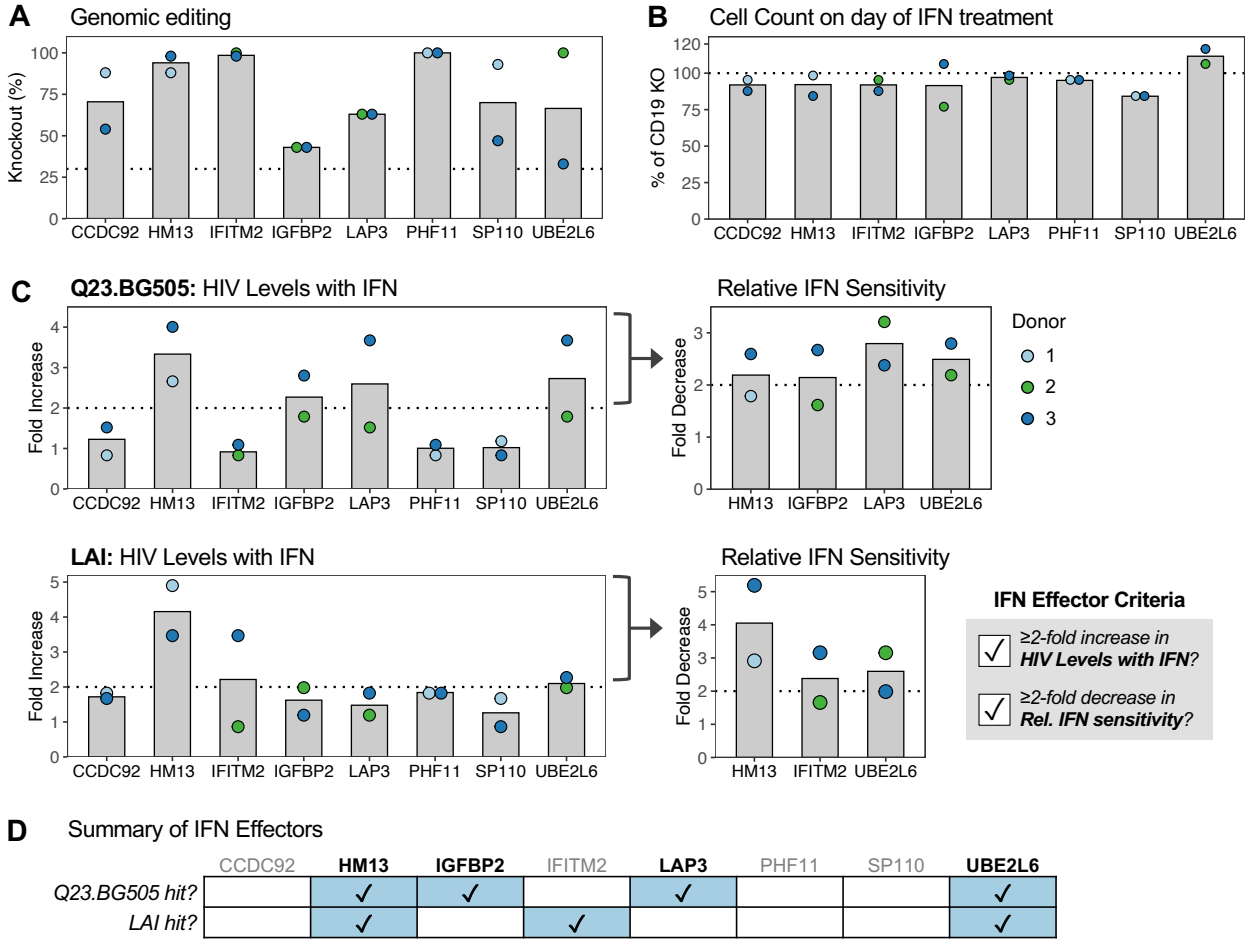


Figure 2.5. CD4-ISG HIV-CRISPR screen hits validate as IFN effectors.
ISGs from Figure 4C (in orange) were validated in CD4$^+$ T cells from two donors. (A) ICE KO percentages. Dotted line indicates 30% KO. (B) The number of cells for each ISG KO on the day of IFN treatment are shown relative to CD19-edited, negative control cells. Dotted line at 100%. (C) Left, increase in HIV Levels with IFN with ISG KO compared to CD19 KO cells (ISG KO/CD19 KO). Right, for ISG KOs that exceeded a two-fold increase in HIV Levels with IFN, the decrease in Relative IFN Sensitivity is depicted (CD19 KO/ISG KO). (D) Summary of ISG KO infection phenotypes.

IFN restriction of HIV-1 is driven by a combination of IFN effectors

We have thus far identified HM13, IFI16, and UBE2L6 as mediators of IFN restriction of both tested HIV-1 viruses. Additionally, IGFBP2 and LAP3 demonstrate IFN-dependent inhibition of the primary HIV-1 strain Q23.BG505. However, the effect of each single knockout on either virus (**Figures 2.2B and 2.5C**) is less substantial than IFN's entire restrictive effect (**Figure 2.1**). We therefore hypothesized that IFN inhibition of HIV-1 results from several IFN effectors functioning collectively. To test this hypothesis, we performed multi-gene knockouts, targeting two (HM13, IFI16), three (HM13, IFI16, UBE2L6), or all five ISGs (HM13, IFI16, UBE2L6, IGFPB2, LAP3) within the same cell pool. HM13 and IFI16 were selected for the double KO condition because they had the largest impacts on Q23.BG505 HIV Levels with IFN and Relative IFN Sensitivity, respectively. The triple KO includes all three genes that were classified as IFN effectors for both viruses in single KO experiments. To extend this study and address whether these factors inhibit diverse HIV-1 strains, we infected cells with a broader panel of six IFN-sensitive viruses, including both primary-isolated and lab-adapted viruses and viruses from different HIV-1 clades (**Figure 2.6A**).

We selected HM13 as the intra-experiment single ISG KO control, as HM13 editing led to the greatest increase in HIV Levels with IFN for both viruses (Q23.BG505: 3.3-fold; LAI: 4.2-fold). Like we previously observed with Q23.BG505 and LAI, HM13 met both IFN effector criteria for all additional HIV-1 strains we tested, with its KO leading to an increase in HIV Levels with IFN and a decrease in Relative IFN Sensitivity that was generally similar across viruses (HIV Levels with IFN: 2.5- to 4-fold, Relative IFN Sensitivity: 2.3- to 4.8-fold; **Figure 2.6B**). These data confirm that HM13 broadly contributes to IFN restriction of a diverse set of HIV-1 strains and is therefore a stringent single ISG KO control.

Multi-gene knockouts were achieved by nucleofecting primary $CD4^+$ T cells with RNPs targeting multiple ISGs. This strategy was successful, yielding editing efficiencies similar to prior respective single-gene KOs (**Figure 2.6C, Supplementary Figure 2.5C**). Though UBE2L6

editing fell below the 30% threshold for two multi-ISG KOs (Donor 1 triple KO in **Figure 2.6C**: 25% KO; Donor 2 KO with all five ISGs in **Supplementary Figure 2.5C**: 24%), this level of inactivation is similar to one of the previous UBE2L6 single-gene KOs (Donor 3 in **Figure 2.5A**: 33%), which impacted HIV Levels with IFN and Relative IFN Sensitivity to similar magnitudes as Donor 2 despite having lower levels of editing. Notably, editing of up to three ISGs in one cell pool did not impact cell viability (**Supplementary Figure 2.5A**). We did however observe a 34% reduction in the number of cells in the five-ISG KO condition, similar to a prior report that tested a six-gene KO in $CD4^+$ T cells [115]. To assess whether multi-gene editing influences HIV-1 infection or IFN restriction non-specifically, we compared the increase in HIV Levels with IFN for HM13 single KO cells to that of cells edited for both HM13 and the negative control B cell marker CD19 (**Supplementary Figure 2.5C**). Double KO cells closely recapitulated the infection phenotype observed in HM13 single KO cells, and there was no trend in terms of which KO caused greater effects on HIV Levels with IFN. Therefore, inactivation of two or three ISGs in the same cell pool does not impact cell viability, HIV-1 infection, nor IFN restriction non-specifically. Moreover, it is unlikely that the lower cell counts with five-ISG KO would influence whether this condition is classified as an IFN effector or not, as our analysis criteria account for such situations (**Supplementary Figure 2.2B**).

We next evaluated whether inactivation of several ISGs impacts IFN restriction of HIV-1 more potently than HM13 KO alone. We observed that with each additional ISG inactivated, HIV Levels with IFN increased (**Figure 2.6D**) and Relative IFN Sensitivities decreased (**Figure 2.6E**). This trend was consistent not only across six diverse viruses (**Figure 2.6D and 2.6E**), but also for Q23.BG505 and LAI in another donor (**Supplementary Figure 2.5D**), indicating that the ISGs identified in this report represent broad IFN effectors, likely with independent functions. Though the multi-ISG KOs do not fully account for IFN's entire restrictive effect, these data provide evidence that IFN restriction of HIV-1 is mediated by several individual antiviral ISGs functioning collectively.

A Infection panel

Virus	Tropism	Clade	Primary Isolate
Q23.BG505	R5	A	Yes
BaL	R5	B	No
93MW965	R5	C	Yes
LAI	X4	B	No
92UG024	X4	D	Yes
92UG029	X4	A	Yes

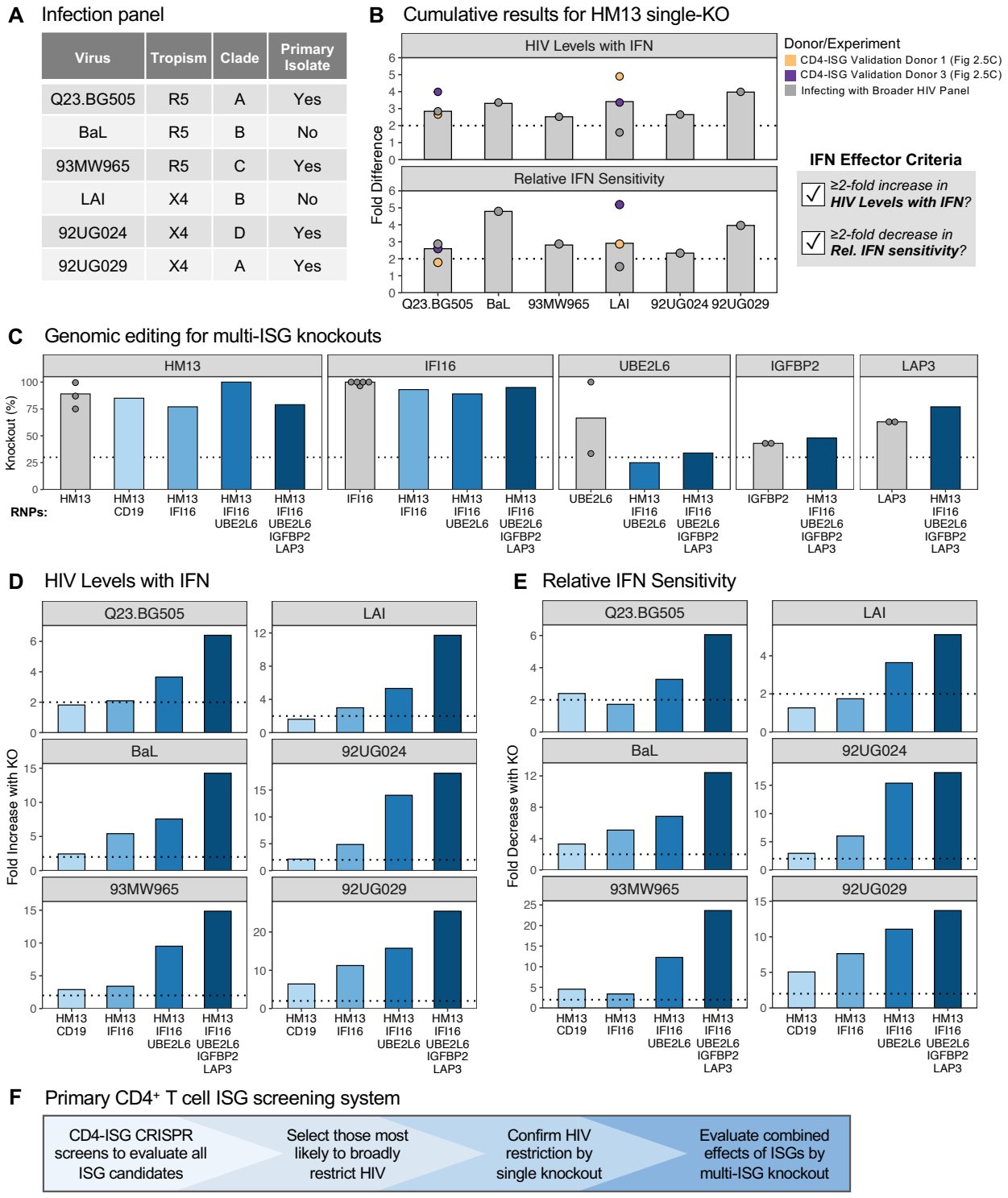


Figure 2.6. IFN restriction of HIV-1 is driven by a combination of IFN effectors.
CD4$^+$ T cells from one donor were nucleofected with 1-5 RNPs to inactivate several ISGs in combination. (A) HIV-1 viruses used for infections. (B) Fold difference in HIV Levels with IFN (HM13 KO/CD19 KO) and Relative IFN Sensitivities (CD19 KO/HM13 KO) between HM13 single KO and CD19 single KO cells for each virus tested.

(figure legend continued on next page)

Dotted lines indicate a fold difference of two. Colored points indicate results from CD4-ISG screen hit validation experiments (Figure 6C), which are presented again here to contextualize results from infecting HM13 single KO cells with the broader viral panel. (C) Percent KO in cells nucleofected with single (gray) or multiple (blue) RNPs, as denoted along the x-axis. Results are separated by editing locus, as indicated above each panel. Single KO bars include results from previous experiments with donors represented as points (n=2-5). Dotted line denotes 30% KO. (D) The increase in HIV Levels with IFN (ISG KO/CD19 KO) and (E) the decrease in Relative IFN Sensitivities (CD19 KO/ISG KO) for each multi-ISG KO compared to CD19 KO, for all viruses tested. Dotted lines indicate a fold difference of two. (F) Schematic of the screening model established in this report. See also **Supplementary Figure 2.5**.

Discussion

In this study, we shed light on the multifaceted nature of cell-intrinsic IFN restriction of HIV-1 and identify effector ISGs with likely relevance *in vivo* by establishing and employing a primary $CD4^+$ T cell ISG screening system. This strategy involves performing comprehensive CD4-ISG CRISPR screens to assess all ISG candidates, selecting and validating the factors most likely to have broad antiviral activity, and evaluating ISG synergy by inactivating several genes in combination (**Figure 2.6F**). We found that HM13, IFI16, IGFBP2, LAP3, and UBE2L6 each partially mediate IFN-dependent restriction of HIV-1 in primary $CD4^+$ T cells. Most of these factors (HM13, IGFBP2, LAP3) were not previously known to inhibit HIV, underscoring the advantage of leveraging unbiased CRISPR screens in primary target cells for ISG candidate selection. Inactivation of the five ISGs we identified in combination demonstrated that these factors collectively contribute to the antiviral effect of IFN for a diverse panel of HIV-1 strains, including several primary isolates. Altogether, this report comprehensively examines ISGs that inhibit primary HIV-1 in primary target cells and demonstrates that the inhibitory effect of IFN is the result of several antiviral ISGs working in concert.

Of the five ISGs we found to mediate IFN restriction of HIV-1, most (HM13, IGFBP2, and LAP3) have not been previously recognized as having anti-HIV activity. HM13 is of particular interest as it is known to directly interact with HIV-1 gp160 [116, 117] and, in this study, demonstrated broad antiviral activity against all tested HIV-1 strains both individually and in

combination with other ISGs. HM13, also referred to as signal peptide peptidase (SPP), cleaves signal peptides in the endoplasmic reticulum after they have been released from polyproteins [118], preproteins, and misfolded membrane proteins [119]. It has been implicated as a host dependency factor for HCV [120, 121] and HSV-1 [122, 123] by mediating HCV core protein processing and binding to HSV glycoprotein K, respectively. Moreover, it may have dual roles both promoting and inhibiting DENV replication at different stages [124, 125]. HM13 therefore interacts with several other viruses, and, in the case of HCV, this interaction relies on HM13's peptidase activity. Collectively, these findings suggest that future studies should investigate whether HM13 cleaves HIV-1 gp160 to inhibit viral replication.

Unlike HM13, the potential function of LAP3 and IGFBP2 during viral replication is less clear. LAP3 is a aminopeptidase primarily located in the cytosol that catalyzes the removal of N-terminal hydrophobic amino acids from various peptides and proteins [126]. IGFBP2 is known for binding insulin-like growth factors when secreted but it has many additional intracellular ligands and intrinsic roles, typically associated with cell survival [127]. Interestingly, IGFBP2 was the second highest-scoring hit in a study examining HIV-1 Vpu-mediated suppression of cytokines, suggesting that its expression may impede viral replication [128]. Because LAP3 and IGFBP2 have not been studied in the context of infection previously, additional work is required to elucidate their mechanisms as HIV-1 restriction factors.

In addition to identifying effector proteins not previously known to inhibit HIV-1, our findings indicate that IFI16 and UBE2L6, which have been studied in other HIV-1 infection models, mediate IFN restriction of several primary HIV-1 isolates in primary $CD4^+$ T cells. IFI16 has two known mechanisms of HIV-1 restriction – it can function as a cytosolic immune sensor of HIV-1 DNA intermediates, which results in IFN secretion and pyroptosis [129-131] to indirectly block HIV-1 replication, and it sequesters the transcription factor Sp1 to broadly inhibit retroviral transcription [74]. If IFI16 was impacting HIV-1 infection in our model via sensing and IFN production, we would expect to see increased levels of HIV-1 with IFI16 KO in the absence

of exogenous IFN treatment. Instead, we observed minimal effects, with IFI16 KO causing a median 1.4- and 1.0-fold decrease in HIV without IFN for Q23.BG505 and LAI, respectively. Therefore, our findings are likely the result of IFI16 directly restricting HIV-1 by reducing the availability of Sp1. This mechanism was originally characterized in macrophages, which also support HIV-1 replication both *in vitro* and *in vivo*; thus, our findings support that IFI16 also directly inhibits HIV-1 infection in primary $CD4^+$ T cells. Like IFI16, the mechanism of UBE2L6 restriction has been previous described. UBE2L6 conjugates ubiquitin or the highly IFN-inducible ISG15 (ISGylation) to other proteins, leading to broad biological impacts [132] including preventing HIV-1 budding and egress [133]. HIV-1 Vpu targets UBE2L6 for degradation to impair ISGylation, but this antagonism is incomplete as evidenced by residual UBE2L6 protein expression and restrictive activity in THP-1 cells [68, 69]. Our data support that Vpu-mediated antagonism of UBE2L6 is incomplete and UBE2L6 contributes to IFN restriction of replication-competent virus in primary target cells.

While HM13, IFI16, IGFBP2, LAP3, and UBE2L6 have consistent antiviral activity across several HIV-1 strains, our results also highlight cell type- and virus-specific trends for other ISGs. For instance, IFITM1 KO did not meet our criteria as an antiviral ISG for either virus in primary $CD4^+$ T cells, despite its potent restrictive effect in SupT1 [67] and U87 [110] cell line models. Our findings suggest that IFITM1 does not contribute to IFN restriction of HIV-1 in primary $CD4^+$ T cells as strongly or consistently as other ISGs we tested, unlike in the models for which it was previously studied. While insufficient protein knockout could be a confounding factor, this is unlikely to explain the results here given the consistently high editing scores for IFITM1. In the case of HM13, IGFBP2, and LAP3, which have not been described as HIV-1 inhibitors in cell lines, it remains to be seen whether they are functional in other cell types. The comprehensive nature of this screen, including the use of the tailored CD4-ISG library, may well reveal antiviral factors that are active in different cell types.

Our findings also shed light on trends in the antiviral effects of ISGs between HIV-1 strains, which we were able to capture by infecting with two viruses for all experiments: Q23.BG505 (clade A, R5-tropic, primary) and LAI (clade B, X4-tropic, lab-adapted). Most notably, we observed higher levels of restriction for IFITM2, MX2, and TRIM5α against LAI than Q23.BG505. These results for IFITM2 and TRIM5α are consistent with the existing literature, as IFITM2 has been shown to preferentially restrict X4 viruses [110, 134] and a previous report demonstrated that TRIM5α mediates restriction of LAI in a THP-1 monocytic cell line, overcoming CypA antagonism [69]. We were particularly surprised that MX2 KO did not impact Q23.BG505 replication to the same extent as LAI, as MX2 has been shown to mediate IFN restriction of several HIV-1 strains in cell lines [63-66, 99, 103, 135-137]. This led us to identify 18 amino acids in HIV-1 CA, which is the main viral determinant of MX2 sensitivity, that differ between Q23-17 and eight MX2-sensitive viruses. Four of the mutations reside in CA sites or regions known to influence MX2 restriction and thus are most likely to be responsible for Q23.BG505's resistance to MX2: V11I, E71D, A92P, E187D. Interestingly, these Q23-17 CA residues are present in many circulating viruses. D71 and P92 are utilized in the group M consensus sequence, as is I11 for clade A and D187 for clade C [138]. Until the use of clade A Q23.BG505 in this report, all MX2 studies have solely tested clade B viruses, suggesting the possibility that MX2 may primarily target clade B viruses *in vivo*. Our findings thus indicate that to best understand the impact of MX2, future studies should include additional strains from the globally dominant clades to investigate whether MX2 restriction is clade-specific in primary $CD4^+$ T cells. Given that TRIM5α also interacts with the HIV-1 CA, it is possible that CA differences between Q23-17 and LAI may also contribute to Q23.BG505's apparent resistance to TRIM5α restriction.

While our approach identified five IFN effectors with broad antiviral activity, we may have only scratched the surface on the ISGs that contribute to IFN-mediated inhibition of HIV-1 due to

the stringency of our analysis and the scope of our investigation. For instance, each CD4-ISG HIV-CRISPR screen enriched for over 100 ISGs (range: 114-166 hits), which we narrowed down to 17 candidates based on their consistency of enrichment in screens and RNA expression across donors. We then selected 8 of the highest-scoring hits to validate. In fact, IFI16 did not meet our filtering criteria despite it scoring above background in half of the screens and being a validated IFN effector. Therefore, this screen identified many more ISGs that have yet to be tested, including virus-specific hits that were not pursued in this study. These data also suggest that our criteria for defining an ISG as an IFN effector was highly stringent because two genes (IGFBP2, LAP3) that met our criteria as Q23.BG505 IFN effectors did not do so for LAI (average increase in HIV Levels with IFN: 1.6-fold for IGFBP2 KO and 1.5-fold for LAP3 KO), yet multi-gene KO experiments demonstrated that editing of these factors in addition to HM13, IFI16, and UBE2L6 further decreased IFN restriction of LAI. The stringency of our IFN effector classification criteria may have therefore excluded “real” HIV-restricting ISGs. Moving forward, inactivating factors alone as well as in combination may provide the most robust data to determine whether ISGs meaningfully contribute to IFN restriction.

Type I IFN upregulates hundreds of ISGs; thus, it is not surprising that we found IFN-mediated inhibition of HIV-1 to result from many ISG effectors as opposed to one extremely potent ISG. While nearly all prior reports have inferred this conclusion due to their ISG of interest not accounting for full IFN restriction, particularly a study based on overexpressed pairs of ISGs prior to HIV-1 pseudovirus challenge in a cell line [86], none to our knowledge have inactivated multiple ISGs within the same cell pool to directly support this hypothesis. Our report therefore establishes that many cell-intrinsic ISGs contribute to IFN restriction of replication-competent HIV-1 in primary target cells, $CD4^+$ T cells. Some of these effectors are genes targeted by HIV-1 accessory proteins (UBE2L6, IGFBP2), suggesting that their antagonism is incomplete. Moreover, viral sensitivity to ISGs can vary by clade and tropism/co-receptor use (MX2, IFITM2, TRIM5α), which is emphasized by the findings presented here. This underscores

the need to test diverse HIV-1 strains when assessing ISG inhibitory potencies and to consider the viruses tested when extrapolating *in vitro* studies to *in vivo* mechanisms.

There is now considerable interest in identifying the minimal network of ISGs needed to ablate the majority of IFN restriction and it remains to be determined whether a potent pan-HIV ISG network exists or if ISG sets would need to be clade-specific. Leveraging the primary $CD4^+$ T cell ISG screening system reported here provides the opportunity to identify additional HIV-1 restriction factors with relevance in the cells that are responsible for the majority of HIV-1 produced *in vivo*. In addition to the hits from our screens that have yet to be validated, future iterations of this system could perform HIV-CRISPR screens with the CD4-ISG library with different HIV-1 strains or in the background of an ISG KO to select for ISGs with synergistic effects. Resolving the group of ISGs that account for IFN-mediated protection against HIV-1 in primary $CD4^+$ T cells would directly inform the cell-intrinsic barriers that HIV-1 must overcome during transmission and acute infection.

Limitations

Though we aimed to maximize biological relevance and experimental robustness, our approach has important caveats to consider. First, primary cell models do not have unlimited cell availability. This made it impossible to use the same blood donors for all experiments, which could contribute to experimental variation and make it difficult to detect antiviral genes with small effects. However, we did not observe notable patterns in editing or KO phenotypes for specific donors and we controlled for differences in HIV-1 infection across donors by determining within-assay fold effects of ISG KO relative to negative controls. This allowed for direct, quantitative comparisons across independent experiments and ISG targets. Second, despite this analysis strategy, we did observe substantial variability across donors for some genes. Biological differences between donors are impossible to avoid in primary cell systems and, to account for this, we evaluated median values to identify IFN effectors. In a primary cell

system, there are trade-offs between having more donor replicates and testing more gene/virus combinations. Our approach evaluated 16 ISGs with 2-5 donors, as opposed to interrogating less ISGs with additional replicates. Lastly, our single KO approach with triple-guide RNPs was relatively low throughput compared to nucleofections with arrayed single-guide RNPs [88, 89, 139]. Though our approach limited the number of donors and viruses we could test as well as the number of screen hits we could validate, it also had valuable tradeoffs. Smaller scale nucleofections allowed us to closely monitor cell counts to ensure all KOs were accurately infected at a low MOI of 0.02. Moreover, we achieved reliable sequencing and high KO scores for all sixteen single ISG targets, whereas a recent arrayed CRISPR screen was unable to assess 18% (77/436) of targets due to incomplete sequencing information or insufficient editing [89].

Methods

Experimental model and subject details

Primary $CD4^+$ T cells

Primary $CD4^+$ T cell isolation, activation, single gene editing, and culturing procedures largely follow a previously published detailed protocol [88]. For primary cell isolation, whole blood from healthy donors collected in EDTA tubes was obtained from BloodWorks Northwest. Peripheral blood mononuclear cells (PBMCs) were isolated by centrifugation over Ficoll-Paque Plus (Cytiva #17144002). $CD4^+$ T cells were isolated using negative selection (STEMCELL Technologies #17952). $CD4^+$ T cells were immediately cultured in complete RPMI media (RPMI 1640 supplemented with 10% FBS, 2 mM L-glutamine, 1X antibiotic-antimycotic that contains penicillin, streptomycin, and Amphotericin B) with 100 U/mL recombinant interleukin (IL)-2 (Roche #11147528001) and activated as described below or frozen in freezing media (90% FBS, 10% DMSO). For all procedures, cell counts were obtained using a hemocytometer.

Cell lines

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

Plasmids

pMD2.G (Addgene #12259) and psPAX2 (Addgene #12260) were gifts from Didier Trono. The TOPmChΔW vector [91] was modified to utilize it in HIV-CRISPR screens. First, the intact 3'UTR was reconstituted as described for the HIV-CRISPR vector [69]. To do this, a gene fragment was synthesized (Integrated DNA Technologies (IDT)) and sub-cloned into the XcmI and PmeI sites using Gibson Assembly. Next, to allow for cloning in an sgRNA library and to incorporate an optimal tracrRNA sequence [111], a cassette was synthesized and assembled into the EcoRI and KpnI sites. In doing this, the spacer region between BsmBI sites used for cloning in of the sgRNA library was reduced from the 1885 bp of other commonly used CRISPR vectors to 355 bp. This modified version of the TOP vector is referred to as HIV-TOP opt mChΔW.

HIV-1 viruses

HIV Q23.BG505 [105] is a clade A, CCR5-tropic HIV-1 chimeric virus derived from a full-length provirus isolated early in infection (Q23) [104] and bears the BG505.C2 envelope [140] of a well-studied clade A transmitted/founder virus. HIV LAI is a clade B, CXCR4-tropic HIV-1 strain, as previously described [106]. Q23.BG505 and LAI viral stocks were prepared by transfection as described below. The following viral stocks were obtained through the NIH HIV

Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 Ba-L (ARP-510) contributed by Dr. Suzanne Gartner, Dr. Mikulas Popovic and Dr. Robert Gallo [141]; HIV-1 93/MW/965 (ARP-2913) contributed by Dr. Paolo Miotti and the UNAIDS Network for HIV Isolation and Characterization; HIV-1 92/UG/029 (ARP-1650) contributed by UNAIDS Network for HIV Isolation and Characterization and the DAIDS, NIAID; HIV-1 92/UG/024 (ARP-1649) contributed by UNAIDS Network for HIV Isolation and Characterization.

Method Details

Primary $CD4^+$ T cell activation and culture

$CD4^+$ T cells were used directly after isolation for the first set of single KO experiments (**Figure 2.2**) and for bulk RNA-seq experiments (**Figure 2.3**) and were thawed from frozen aliquots for all other experiments. After isolation or thawing, cells were resuspended at $2.5e10^6$ 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 assembly for single- and multi-gene KOs

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). NTC RNPs were prepared in advance by preincubating 1 µL of 160 nM CRISPR-RNA (crRNA, IDT) with 1 µL of 160 nM tracrRNA (IDT) at 37C for 30 min. The crRNA/tracrRNA complex was then incubated for 15 min at 37C with 2 µL of 40 nM Cas9-NLS protein (UC Berkeley QB3 MacroLab) for a final molar ratio of 4:1 RNA:Cas9. Complexes were frozen at -80C until ready for use. For all other single gene KOs, RNPs were assembled on the day of nucleofection by combining 6 µL of 30 pmol/µL sgRNAs (Synthego), 18 µL of P3 Primary Cell

Nucleofector solution (Lonza #V4SP-3096), and 1 µL of 20 µM Cas9-NLS protein (UC Berkeley QB3 MacroLab). Complexes were gently mixed, incubated for 10 min at room temperature, and then stored on ice until ready for same-day use.

For multi-gene KOs, RNPs for each gene were prepared separately and respective to the number of genes being targeted. For each gene, 1.5 µL of 60 pmol/µL sgRNAs (Synthego) was combined with 1 µL of 20 µM Cas9-NLS protein (UC Berkeley QB3 MacroLab) and the following volume of P3 Primary Cell Nucleofector solution: 8.75 µL for two-gene KOs, 5.67 µL for three-gene KOs, and 3.2 µL for five-gene KOs. Complexes were gently mixed, incubated for 10 min at room temperature, and then stored on ice until ready for same-day use. Immediately prior to nucleofection, RNPs for different genes were combined.

RNP nucleofection of $CD4^+$ T cells

Three days after activation, 1e6 $CD4^+$ T cells per nucleofection were pelleted at 100xg for 10 min and washed once with PBS. 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 µL of the complexes described above. For RNPs with Synthego guides (both single and multi-gene KO), cells were gently resuspended with 25 µL of the complexes described above. Cells and complexes were incubated for 5 min at room temperature before nucleofection in a 16-well Nucleocuvette (Lonza #V4SP-3096) in an Amaxa Nucleofector (Lonza) using pulse code EH-115. Cells were supplemented with 80 µL of RPMI complete media with 125 U/mL IL-2 and allowed to recover in the cuvette for 30 min to 2 h at 37C. Cells were then brought to 2.5e6 cells/mL in RPMI complete media with 100 U/mL IL-2 and plated in 96-well flat-bottom plates at a 1:1 ratio with activation beads (Miltenyi Biotec #130-091-441). Two days later, 200 µL of RPMI complete media with 100 U/mL IL-2 was added to each nucleofection. At four days post nucleofection, fresh media was added to each nucleofection by pelleting cells at 300xg for 10 min, removing 250 µL of supernatant (roughly half of the total

culture volume), and adding RPMI complete media with 100 U/mL IL-2 to bring cells 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. As needed, cells were transferred from 96-well plates (working volume: 100-300 µL/well) to 48-well plates (400-600 µL/well) or 24-well plates (1-2 mL/well).

HIV-1 and lentivirus production and concentration

HIV-1 and lentiviral stocks were prepared by transfecting HEK293T/17 cells (ATCC #CRL-11268) in a six-well format. Twenty hours prior to transfection, 5e5 cells/well were plated in 2 mL of DMEM complete. On the day of transfection, serum-free DMEM media, transfection reagents, and plasmids were equilibrated to room temperature. For each HIV-1 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). For each lentiviral transfection well, 667 ng of transfer plasmid, 500 ng of psPAX2 packing vector (Addgene #12260), 333 ng of the pMD2.G VSV-G expression vector (Addgene #12259), and 200 µL serum-free DMEM were combined, prior to the addition of 4.5 µL TransIT-LT1 transfection reagent (Mirus Bio #MIR-2305). 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, removed of debris using a 0.2 µM filter (Millipore Sigma #SCGP00525), and concentrated as described below.

For HIV-1 transfections, supernatants were passed through Amicon 100 KDa concentrators (Millipore Sigma #UFC910024) until concentrated ~100-fold. Concentrated virus was aliquoted and stored at -80C. HIV-1 stocks were titered on TZM-bl cells (NIH HIV Reagent Program #ARP-8129). Infections were carried out in the presence of 10 µg/mL DEAE-dextran

and viral titer was determined by staining fixed cells for β-galactosidase activity at 48 h post-infection and counting blue, infected cells.

For lentiviral transfections, ice-cold PBS was added to supernatants to bring the volume to 35 mL. Diluted supernatants were concentrated over 2 mL of cold 20% sucrose by ultra-centrifugation at 23,000 rpm for 1 h at 4C. After pouring the supernatant off, the sucrose pellet was resuspended in RPMI complete media, such that the virus was concentrated ~100-fold. Pellets were resuspended by periodically vortexing at a low intensity and were left at 4C overnight before aliquoting and freezing at -80C. Lentiviral stocks were titered by copy number as described below.

IFN treatment and HIV-1 infection of $CD4^+$ T cells

Nucleofected $CD4^+$ T cells were treated with 1000 U/mL IFN-β1a (PBL Assay Science #11410) 24 h prior to HIV infection. For the first set of single KO experiments (**Figure 2.2**) and bulk RNA-seq experiments (**Figure 2.3**), cells were IFN-treated six days post nucleofection. Otherwise, cells were IFN-treated nine days post nucleofection to allow for additional recovery. For targeted KO experiments, half of each nucleofection was IFN-treated, whereas all cells were IFN-treated for CRISPR screens. For ISG KO spreading infections, cells were resuspended to 1e6 cells/mL in RPMI complete media with 100 U/mL IL-2 and 100 µL was added in duplicate for each virus to 96-well flat-bottom plates. Cells were infected with 10 µL of virus at an MOI of 0.02 in the presence of 8 µg/mL polybrene (Millipore Sigma #TR-1003-G) by spinoculation at 1100xg for 90 min at 30C. After infection, cells were transferred to 96-well V-bottom plates, pelleted at 300xg for 10 min, and resuspended in 100 µL of fresh RPMI complete with 100 U/mL IL-2 that contained 1000 U/mL IFN as necessary. Cells were then transferred back to the 96-well flat-bottom infection plates and were incubated at 37C. Viral supernatants were harvested 2, 4, 6, and 8 days post-infection (dpi), stored at -80C, and replaced with RPMI complete media with 100 U/mL IL-2 containing 1000 U/mL IFN as needed.

Reverse transcriptase (RT) activity qPCR assay

HIV infection levels were assessed by measuring RT activity in viral supernatants using the previously described RT activity qPCR assay [142]. Upon thawing, 5 µL of viral supernatant was incubated with 5 µL lysis buffer (0.25% Triton X-100, 50 mM KCl, 100 mM Tris HCl, 40% glycerol, 0.8 U/µL RNAse Inhibitor (ThermoFisher #EO0381)) for 10 min at room temperature. This mixture was then diluted with 90 µL of ultrapure water and 9.6 µL was transferred to a 384-well MicroAmp Optical plate (Applied Biosystems #4309849) as input for the assay. Per well, 10.4 µL of the following reaction mix was used: 10 µL ROX SYBR 2X MasterMix (Eurogentec #UF-RSMT-B0701), 0.1 µL 10-fold diluted RNase Inhibitor (ThermoFisher #EO0381), 0.1 µL of MS2 RNA (Roche #10165948001), and 0.1 µL of each 100 µM primer (MS2-F: TCCTGCTCAACTTCCTGTCGAG, MS2-R: CACAGGTCAAACCTCCTAGGAATG). Plates were read on an Applied Biosystems QuantStudio 7 Real Time PCR machine.

To quantify the RT activity (mU/mL) of viral supernatants, the RT units of a concentrated stock of HIV-1 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 then used as the quantitative standard curve in all assays.

Genomic editing analysis

On the day of HIV-1 infection, 2.5e5 nucleofected cells were harvested by pelleting at 300xg for 10 min in PBS. Cell pellets were either frozen at -80C or immediately used for genomic DNA extraction (QIAGEN #51104). Edited loci were amplified from genomic DNA from edited and control cells using 10 µM primers specific to each targeted locus (**Supplementary Table 2.5**), Herculase II Fusion DNA Polymerase (Agilent #600679), and the thermocycler program detailed in **Supplementary Table 2.5**. PCR amplicons were purified (QIAGEN #28104) and Sanger sequenced (Fred Hutch Shared Resources Genomics Core). Chromatograms were

analyzed by ICE analysis tool v3 (Synthego) to determine the rate of editing predicted to cause KO at each genomic locus in the cell population.

Bulk RNA-seq of IFN-treated $CD4^+$ T cells

$CD4^+$ T cells were isolated and immediately activated on the same day, as described above. Three days later, cells were brought to 2.5e6 cells/mL in RPMI complete media with 100 U/mL IL-2 and plated in 96-well flat-bottom plates at a 1:1 ratio with activation beads (Miltenyi Biotec #130-091-441). Cells were maintained thereafter as usual. Six days after bead activation, $CD4^+$ T cells were treated with 1000 U/mL IFN-β1a (PBL Assay Science #11410) as described above. Cells were collected prior to IFN treatment and 3, 6, 12, and 24 h thereafter, depending on the donor (**Figure 2.3A**). During collection, 0.5-1e6 cells were washed with PBS, pelleted, lysed in 350 µL Buffer RLT Plus (QIAGEN #74134) supplemented with 1% β-mercaptoethanol, and stored at -80C. This procedure was repeated for five donors.

For RNA extractions, thawed cell lysates were homogenized using a QIAshredder column (QIAGEN #79654) and genomic DNA was removed on a gDNA Eliminator column. Total RNA was then extracted using the RNeasy Plus Mini Kit (QIAGEN #74134) and eluted in 30 µL RNase-free water. The quality and concentration of RNA preps was assessed by Agilent High Sensitivity RNA ScreenTape (Fred Hutch Shared Resources Genomics Core). After samples were found to have high RNA Integrity scores (RIN ≥ 9.4), they were sequenced on an Illumina HiSeq 2500 (Fred Hutch Shared Resources Genomics Core) with a sequencing depth of roughly 20 million reads per sample.

CD4-ISG sgRNA library construction

An oligo pool containing the CD4-ISG sgRNA library was synthesized (Twist Biosciences) and cloned into HIV-TOP opt mChΔW, as previously described in detail [69]. The library pool was amplified (Array-F primer: TAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG, optTracr-Array-R primer: GTTGATAACGGACTAGCCTTATTTCAACTTGCTATGCTGTTTCCAGCATAGCTCTGAAAC), gel-purified (QIAGEN #28704), and cloned into the HIV-TOP opt mChΔW following digestion with BsmBI (NEB #R0739L) using Gibson Assembly (NEB #E2611S) with a 5:1 molar ratio of insert to vector. Gibson reactions were transformed into Endura electrocompetent cells (Lucigen #60242-2) at a large scale, such that there was 230-fold coverage of the CD4-ISG library. Transformed bacteria were scraped from plates and plasmid DNA was isolated using the Endotoxin-Free Nucleobond Plasmid Midiprep Kit (Takara Bio #740422.10). The CD4-ISG library was sequenced and contains 4,574 of the 4,759 guides (99.9%) initially included in library design.

Titration of lentivirus stocks by copies per cell

$CD4^+$ T cells were isolated, thawed, and activated as described above but were plated in a 96-well plate. The next day (20-24 h), cells were transduced with a dilution series of the concentrated CD4-ISG library lentivirus. Each well was transduced with 10 µL of viral inoculum supplemented with protamine sulfate (Millipore Sigma #P3369) for a final concentration of 8 µg/mL. Cells were spinoculated at 1100xg for 90 min at 30C and maintained at 37C afterwards. Four days after transduction, roughly 1e5 cells were collected, pelleted in PBS, and subjected to genomic DNA isolation (QIAGEN #51104). Genomic DNA was digested with BglI (NEB #R0143L) and diluted 1:3 for droplet digital PCR (ddPCR) amplification to measure vector copies per cell. Per well of a 96-well ddPCR plate (Bio-Rad, #12001925), 3 µL of diluted

template DNA was used as input, along with 1.8 µL of each 10 µM vector-specific primer (ddPCR-cPPT-F: GTACAGTGCAGGGGAAAG, ddPCR-U6-R: ATGGGAAATAGGCCCTCG), 0.5 µL FAM probe (6-FAM/ZEN - AGACATAATAGCAACAGACATACAAAC -IBFQ), 1 µL RPP30 ddPCR copy number variation HEX assay mix (Bio-Rad, #100-31243), and 10 µL SuperMix with no dUTPs (Bio-Rad #1863024). The volume per well was brought to 20 µL with water, followed by droplet generation, PCR amplification, and reading of the plate on an Applied Biosystems QuantStudio 7 Real Time PCR machine. Vector copies per cell were calculated by dividing vector copies/µL by RPP30 copies/µL, and then dividing this value by two to correct for two RRP30 copies per cell. Titrations were performed in $CD4^+$ T cells from two unique donors during independent experiments. Cumulative results determined that 0.06 µL of CD4-ISG library lentivirus per 1e6 cells yielded an average of 2.7 vector copies per cell.

CD4-ISG HIV-CRISPR screening

$CD4^+$ T cells were isolated, thawed, and activated as described above. The next day (20-24 h), roughly 20e6 cells were transduced with the CD4-ISG library lentivirus, aiming to deliver 3 vector copies per cell as described in titration experiments above. Each well was transduced with 10 µL of diluted viral inoculum supplemented with protamine sulfate (Millipore Sigma #P3369) for a final concentration of 8 µg/mL. Cells were spinoculated at 1100xg for 90 min at 30C and maintained at 37C afterwards. Two days later, Cas9 was delivered to cells by nucleofection with NTC RNPs, as described above. This approach utilizes a guide-swap strategy [90] that replaces the nucleofected NTC guide with the library guides delivered by lentiviral transduction. Cells were maintained thereafter as usual. Nine days after nucleofection, all library cells were treated with 1000 U/mL IFN-β1a (PBL Assay Science #11410). On the following day, HIV-1 infections were established.

For infections, 2.5e5 $CD4^+$ T cells in 100 µL of RPMI complete media with 100 U/mL IL-2 and 1000 U/mL IFN were added to wells of a 96-well flat-bottom plate for each virus and MOI. In

total, 10e6 IFN-treated $CD4^+$ T cells were used for each CRISPR screen to maintain high library coverage. Cells were infected with 10 µL of virus at an MOI of 0.1 and 1 in the presence of 8 µg/mL polybrene (Millipore Sigma #TR-1003-G) by spinoculation at 1100xg for 90 min at 30C. After infection, 150 µL of RPMI complete media with 100 U/mL IL-2 and 1000 U/mL IFN was added to each well to bring cells to 1e6 cells/mL and plates were incubated at 37C. Three days post infection, all cells were split 1:2 with fresh RPMI complete media with 100 U/mL IL-2 and 1000 U/mL IFN. For high MOI screens (HIV MOI=1), viral supernatants were harvested 3 and 5 dpi and cells were collected 5 dpi. For low MOI screens (HIV MOI=0.1), viral supernatants were harvested 4 and 7 dpi and cells were collected 7 dpi. All viral supernatants were filtered to remove cell debris (Millipore Sigma #SCGP00525) and stored at 4C. Within four days of collection, viral supernatants were concentrated over a sucrose gradient as described above, resuspended in 100 µL PBS, and stored at 4C. Viruses harvested on different days were kept separate throughout this process. Cells were washed in PBS, pelleted, and stored at -80C.

CRISPR screen samples were prepared for Illumina sequencing as previously described [69]. Genomic DNA was extracted from thawed cell pellets using the QIAamp DNA Blood Midi Kit (QIAGEN #51185) and digested with BglI (NEB #R0143L). Viral RNA was isolated from viral pellets one day after concentration (QIAGEN #52904). RNA extracts were then treated with DNase I (Roche #04716728001) and reverse transcribed to generate cDNA with Superscript III (Invitrogen #18080044) and a primer specific to the HIV-TOP opt mCh∆W vector (SeqMSCV-R1-R: CTTGCTAAACCTACAGGTGG). Guide sequences were amplified from 20 µg of prepared genomic DNA (~400-fold coverage) and viral cDNA using Herculase II Fusion DNA Polymerase (Agilent #600679) and 10 µM primers specific to the HIV-TOP opt mCh∆W vector (PLC-seq-R1-F: GAGGGCCTATTTCCCATGATTCCTTCA, SeqMSCV-R1-R: CTTGCTAAACCTACAGGTGG). Amplicons were purified and concentrated (QIAGEN #28104). A second-round PCR was performed to add identifier sequences to each sample using indexing

primers (**Table S5**) and the PLC-Seq-R2-RS primer (CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTG CCACTTTTTCAAGTTGATAACGGACT). PCR products were gel isolated (QIAGEN #28704) and quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen #P7589). Samples were pooled at an equimolar ratio to generate a 2 nM pool. Pooled libraries were sequenced on an Illumina NextSeq 2000 with ≥500-fold coverage (Fred Hutch Shared Resources Genomics Shared Resource).

Western blotting

Cells collected on the day of HIV-1 infection were lysed in RIPA buffer (50 mM Tris pH 8.0, 0.1% SDS, 1% Triton-X, 150 mM NaCl, 1% deoxycholic acid, 2 mM PMSF). Cell extracts were prepared in 4X LDS Buffer (Invitrogen #NP007) and 2X MES SDS Running Buffer (Invitrogen #NP002). After boiling at 100C for 5 min, 20 µg of the prepared samples were loaded into a NuPAGE 4-12% Bis-Tris Gel (Invitrogen #NP0321) and resolved. Following transfer onto a nitrocellulose membrane (Invitrogen #LC2001), blots were blocked in blocking buffer (PBS, 0.1% Tween, 5% milk) for 1 h, incubated with the MX2 primary antibody (Santa Cruz Biotech #sc-271527; 1:100) overnight, probed with secondary antibody (Cytiva #NA931, 1:1000) for 1 h, and visualized by chemiluminescence (Cytiva #RPN2232) on a Bio-Rad ChemiDoc Touch. Membranes were washed in PBST (PBS, 0.1% Tween) in between each step. After imaging the MX2 stain, membranes were stripped (ThemoFisher #46430), washed, probed with GAPDH:HRP (Bio-Rad #MCA4739P, 1:1000) for 1 h, and imaged with chemiluminescence.

HIV-1 CA sequence alignments

CA sequences were obtained via http://www.hiv.lanl.gov/ in September 2022. Global alignments were performed using Geneious Prime version 2022.2 (RRID:SCR_010519).

Quantification and Statistical Analysis

RT activity qPCR assay analysis

The RT activity (mU/mL) of experimental wells was interpolated from a standard curve (concentrated Q23.BG505 stock, as described above) using the lm function in the R stats package. RT values from duplicate experimental wells were then averaged prior to analysis. To compare RT activity across donors for KO experiments, the following were determined for each KO, virus, and donor using day 6 RT activity: Relative IFN Sensitivity (no IFN RT / IFN RT), the fold increase in HIV Levels with IFN due to ISG KO as compared to the negative control (ISG KO / negative control), and the fold decrease in Relative IFN Sensitivity due to ISG KO as compared to the negative control (negative control / ISG KO). NTC-delivered cells were used as the negative control for the first set of single KOs (**Figure 2.2**). For all other KO experiments (**Figures 2.5 and 2.6**), we inactivated the B cell marker CD19 as the negative control because nucleofection with CD19 RNPs generates double-stranded DNA breaks, making it an even more robust negative control, and does not impact HIV-1 replication, which we verified (**Supplementary Figure 2.6**). Standard curve interpolation and statistical comparisons were performed in RStudio.

Bulk RNA-seq analysis and CD4-ISG library design

Illumina sequencing reads (SRA accession number: PRJNA921704) were aligned to the GRCh38 reference genome to generate raw counts (Fred Hutch Shared Resources Bioinformatics Core). Counts per million (CPM) were then calculated using the cpm function in the Bioconductor edgeR package [143] (**Supplementary Table 2.1**). For principal component analysis (PCA) of all 17 samples, data from genes with sufficient expression (CPM>1 in at least 8 of the samples) were corrected for donor variation with the combat_seq function [144] in the Bioconductor sva package prior to PCA generation.

Prior to calculating IFN induction, raw counts equal to 0 were adjusted to 1 and CPMs were re-generated. With roughly 20 million reads per sample, cases originally with zero counts ultimately had ~0.05 CPM, which allowed for fold change calculations but should not skew downstream analyses. After CPM generation, long intergenic non-coding RNAs genes (GeneName contains “LINC”) were removed as were mitochondrial genes (Chr = MT). ISG designation was then performed for each time point independently. For data from 6, 12, and 24 hours after IFN treatment, CPMs from IFN-treated samples from a given time point and their donor-matched untreated controls were aggregated and genes with low expression were filtered out (required CPM>1 in half of the samples). IFN fold induction was calculated from CPMs for each treated/untreated donor pair (**Supplementary Table 2.1**). Genes with a fold change (FC) ≥2 in the majority of donors for that time point were considered ISGs. This yielded 449, 261, and 167 ISGs for the 6, 12, and 24 h post IFN time points, respectively. For data from 3 hours after IFN treatment, we implemented a more stringent strategy for ISG designation because we only had data from one donor for this time point. IFN fold induction was calculated for the 3 h treated/untreated donor pair for genes that passed the CPM cutoff for any of the prior time point analyses. Genes with a FC≥4 were considered ISGs, yielding 261 ISGs. These analyses identified 573 genes in total (**Figure 2.3D**). Finally, to capture genes with high IFN stimulation in any donor or time point, we also included genes with FC≥5 in any sample, which added 27 genes for a total of 600 ISGs.

To design the CD4-ISG library, we generated four guides with two different tools, CHOPCHOP [145] and GUIDES [146], for a total of eight guides per gene. Of the 600 ISGs identified via bulk RNA-seq, we were unable to design guides for 107 due to transcripts encoding pseudogenes, RNA genes, or read-through regions. To complete the CD4-ISG sgRNA library, we added 21 control genes, 2 genes from the first set of single KOs, 38 top-scoring hits from preliminary CRISPR screens in $CD4^+$ T cells with the PIKA ISG library [69] (data not

shown), and 200 NTC guides for a total of 555 genes represented across 4,579 guides (**Supplementary Table 2.2**).

CD4-ISG HIV-CRISPR screen analysis

Illumina sequencing reads (SRA accession number: PRJNA919035) were analyzed by MAGeCK-Flute [113] to generate guide counts and gene-level enrichment data (**Supplementary Table 2.3**; Fred Hutch Center for Data Visualization). As opposed to using all 200 NTC guides to inform data for one NTC "gene", 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 CD4-ISG library (554) (**Supplementary Table 2.2**). 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. Results from viral supernatants collected 3 dpi for high MOI screens and 4 dpi for low MOI screens were used for the analyses presented in this report, as these collection days had higher donor correlation coefficients compared to the later collection time points (5 and 7 dpi, respectively by MOI; data not shown). Pearson R correlations between MAGeCK scores from different donors were determined using the stat_cor function in the ggpubr package [147]. Screen background was calculated as the average of the synNTC MAGeCK scores + 3 standard deviations.

Additional quantification and statistical analysis

The graphical abstract and Figures 2.1A, 2.3A, and 2.4A were generated with BioRender.

Addendum

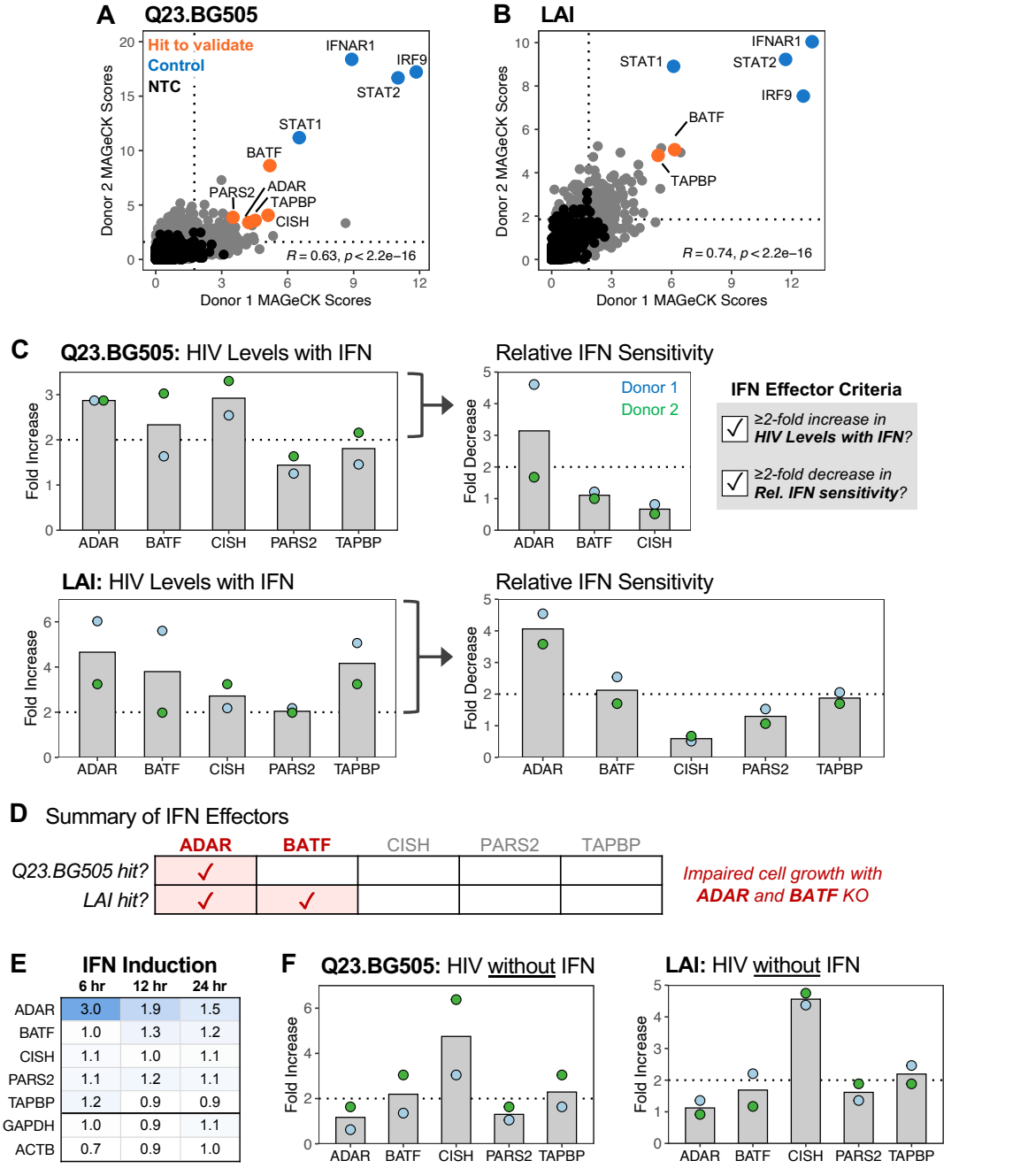
The initial HIV-CRISPR screens for this study were performed with the established PIKA ISG sgRNA library, which targets 1,906 genes included in previous ISG libraries and those found

to be IFN-stimulated in a variety of cell types and lines [69]. These screens were conducted before we understood that 87.5% of the PIKA library targets (1667 out of 1906) are not IFN-stimulated in our $CD4^+$ T cell model (**Figure 2.3E**), which prompted the design and construction of the tailored CD4-ISG library. This addendum will describe results from the HIV-CRISPR screens with the PIKA library, which data were not included in the manuscript for this project. Of note, Chapter 3 reports these same HIV-CRISPR screens with the PIKA library but focuses on just one gene (SLC35A2), which will not be discussed in this section.

To identify ISGs that restrict HIV-1 in the main target cells of infection, we performed HIV-CRISPR screens with the PIKA library in primary $CD4^+$ T cells from two independent donors in the context of IFN treatment (**Figure 2.4A**). Positive control IFN signaling genes (IRF9, STAT1, STAT2, IFNAR1) and negative control dependency factors (CCR5, CXCR4) were enriched or depleted as expected for all screens (**Figure 2.7A and 2.7B, Supplementary Figure 2.7A**). We observed strong agreement between donors and identified 152 and 151 genes scoring above background for both donors for Q23.BG505 and LAI, respectively (top right quadrant in **Figure 2.7A and 2.7B**). Therefore, our screening approach successfully categorized control genes and identified numerous possible antiviral hits for each virus.

We next selected top-scoring hits with consistent enrichment between donors to validate by single KO. Specifically, we chose three Q23.BG505-specific hits (ADAR, CISH, PARS2) and two hits shared by both viruses (BATF, TAPBP) (**Figure 2.7A and 2.7B, Supplementary Figure 2.7B**). We employed the same editing approach and analysis criteria described earlier (**Figure 2.1A, Supplementary Figure 2.2**), using Synthego Gene KO v2 Kits, which guide sequences are detailed in **Supplementary Table 2.6**. ADAR was the only knockout that met both IFN effector criteria for Q23.BG505, whereas editing of ADAR and BATF impacted IFN restriction of LAI (**Figure 2.7C and 2.7D**). However, on the day of IFN treatment, after the editing and culture period, we noticed that the populations of ADAR- and BATF-edited cells were much smaller than the rest, with cell counts that were 61% and 42%, respectively, lower than the average of the

other KOs (**Supplementary Figure 2.7C**). Notably, ADAR KO is lethal in mice and causes toxicity in other *in vitro* systems [148, 149], consistent with our findings. We were thus unable to determine whether ADAR and BATF directly restrict HIV-1 or if the infection results were a byproduct of a cell growth phenotype. Excluding ADAR and BATF as hits, our validation experiments ultimately did not identify any ISGs that contribute to IFN restriction of HIV-1 (**Figure 2.7D**).



E IFN Induction

	6 hr	12 hr	24 hr
ADAR	3.0	1.9	1.5
BATF	1.0	1.3	1.2
CISH	1.1	1.0	1.1
PARS2	1.1	1.2	1.1
TAPBP	1.2	0.9	0.9
GAPDH	1.0	0.9	1.1
ACTB	0.7	0.9	1.0

(figure legend on next page)

Figure 2.7. HIV-CRISPR screens with PIKA library enrich for many genes that are not IFN-stimulated in primary CD4$^+$ T cells.
(A, B) HIV-CRISPR screens with the PIKA library were conducted using CD4$^+$ T cells from two donors, infecting cells at an MOI of 1. Dotted lines indicate NTC average + 3 standard deviations. Pearson correlation coefficients are included within plots. (C) Five CRISPR screen hits (orange in Panels A and B) were validated via targeted gene KO in CD4$^+$ T cells from two donors. Left, increase in HIV Levels with IFN with ISG KO compared to CD19 KO cells (ISG KO/CD19 KO). Right, for ISG KOs that exceeded a two-fold increase in HIV Levels with IFN, the decrease in Relative IFN Sensitivity is depicted (CD19 KO/ISG KO). (D) Summary of ISG KO infection phenotypes. (E) Heatmap of average IFN-induced fold changes in RNA expression for screen hits and housekeeping genes (GAPDH, ACTB) in primary CD4$^+$ T cells at 6-, 12-, and 24-hours post-IFN treatment (n = 3, 3, and 5 donors, respectively by time point). (F) Increase in HIV Levels without IFN with ISG KO compared to CD19 KO cells (ISG KO/CD19 KO). See also **Supplementary Figure 2.7**.

We are now able to contextualize these validation results with our bulk RNA-seq dataset on IFN-treated CD4$^+$ T cells (**Supplementary Table 2.1**). Of the genes selected as hits in the PIKA HIV-CRISPR screens, ADAR is the only gene considered an ISG in CD4$^+$ T cells (**Figure 2.7E**). Given ADAR's IFN inducibility (3-fold induction at 6 hours post IFN treatment) and consistent, robust effect on IFN-mediated restriction of both viruses (**Figure 2.7C**), it is possible that ADAR represents an IFN effector. ADAR, which stands for adenosine deaminase RNA-specific, is an enzyme that mutates and consequently destabilizes double-stranded RNA by deaminating adenosine to inosine [150]. ADAR has also been shown to suppress the activity of PKR, which is a kinase that blocks the initiation of viral translation. Counter to our findings, some previous studies have found a proviral role for ADAR on HIV-1 infection at the level of protein translation, likely due to ADAR's effect on PKR [151-155]. However, other publications have reported the opposite trend, which we also observed, that ADAR inhibits HIV-1 replication [156, 157]. There is also evidence that ADAR is packaged into budding HIV-1 virions [158], thus it is possible that ADAR has different effects in producer versus target cells. Elucidating the nuances of the effect(s) of ADAR on HIV-1 infection will likely require controlling for cell growth phenotypes, performing ADAR and PKR double-knockouts, and assessing the impact of ADAR KO in producer and target cells with and without IFN treatment.

Despite suggesting a potential role for ADAR, the bulk RNA-seq results highlight that the rest of the PIKA screen hits we followed up on (BATF, CISH, PARS2, TAPBP) are not ISGs in $CD4^+$ T cells (IFN inductions of 0.9-1.3; **Figure 2.7E**) and thus explain why these genes did not validate as IFN effectors. We therefore assessed whether any of these genes impact HIV-1 in the absence of IFN treatment, as these genes could represent endogenously expressed restriction factors. Notably, CISH KO substantially increased infection levels of both viruses in the absence of IFN (**Figure 2.7F**), causing a 4.8-fold and 4.6-fold increase in Q23.BG505 and LAI infection, respectively. CISH, which stands for cytokine-inducible SH-2 domain protein, is a member of the suppressor of cytokine signaling (SOCS) family of protein and inhibits JAK-STAT and T cell receptor (TCR) signaling, with its inactivation in T cells leading to enhanced sensitivity to TCR and cytokine stimulation, longer survival, and higher cytokine secretion [159]. Consistent with our findings of CISH as an antiviral factor, polymorphisms in *CISH* are associated with susceptibility to bacteremia, malaria, and tuberculosis [160]. The mechanism of CISH protection against HIV-1 or other infectious diseases has not been explored. Such investigations would likely be informed by performing bulk RNA-seq on CISH-inactivated cells before and after infection, as it is likely that CISH indirectly impacts infection by regulating the expression of other genes.

Ultimately, the initial HIV-CRISPR screens with the PIKA library demonstrated the reliability of this screening approach in primary cells and shed light on two host factors that may broadly inhibit HIV-1 infection and warrant further investigation – the IFN-inducible ADAR and constitutively expressed CISH. Nevertheless, in a screen intended to identify antiviral ISGs, only one of the chosen hits was actually IFN-stimulated in $CD4^+$ T cells. This discrepancy underscores the importance of using a tailored guide library customized for ISGs expressed in the cell type of interest and demonstrates the utility of the bulk RNA-seq dataset reported here.

Chapter 3

HOST CELL GLYCOSYLATION SELECTS FOR INFECTION WITH CCR5- VERSUS CXCR4-TROPIC HIV-1

Sections of text in this chapter have been modified from the following manuscript:

Itell, H.L., Humes, D., Baumgarten, N.E., & Overbaugh, J. Host cell glycosylation selects for infection with CCR5- versus CXCR4-tropic HIV-1. *bioRxiv*, doi:10.1101/2023.09.05.556399 (2023).

Abstract

HIV-1 infection involves a selection bottleneck that leads to transmission of one or a few variants, which nearly always use CCR5 as the coreceptor for viral entry rather than CXCR4. The host properties that drive this selection are not well understood and may hold keys to factors that govern HIV-1 susceptibility. In this report, we identified SLC35A2, a transporter of UDP-galactose, as a potent CXCR4-tropic-specific restriction factor in primary target $CD4^+$ T cells. SLC35A2 inactivation, which resulted in truncated glycans, not only increased CXCR4-tropic infection levels, but also consistently decreased those of CCR5-tropic strains. Single cycle infections demonstrated that the effect is cell-intrinsic. These data support a role for a host protein that influences glycan structure on HIV-1 infection, with reduced CXCR4-tropic infection associated with SLC35A2-mediated glycosylation. Because SLC35A2 is expressed in target cells in both the blood and mucosa, it may also contribute to CCR5-tropic selection during HIV-1 transmission.

Main text

HIV-1 undergoes a severe viral population bottleneck during transmission. HIV-1'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

infection [6, 7, 9-13]. While transmitted viruses are thought to be successful in part due to their high replicative fitness [21, 22] and reduced susceptibility to interferon (IFN)-mediated inhibition [23, 25], the most consistent, defining feature of transmitted variants is their strong coreceptor tropism preferences. HIV-1 can use either CCR5 or CXCR4 as a coreceptor for viral entry [15], yet transmitted variants nearly always use CCR5 [13, 16] despite broad expression of CXCR4 at transmission sites [20], with CXCR4-tropic (X4) viruses typically arising in the months to years after primary infection in a subset of individuals [39]. While R5 viruses are, overall, more common during natural infection, the consistency of CCR5-tropic (R5) virus selection across transmission settings suggests that the HIV-1 transmission bottleneck is not solely stochastic. It has thus remained a longstanding question in the HIV-1 field why R5-tropic variants, but not X4 viruses, can more successfully navigate this bottleneck, leading to new infections.

Given that $CD4^+$ T cells are thought to be the first cells that HIV-1 productively infects after transmission [6, 7], host factors that inhibit HIV-1 in this important cell type may impose selective pressures that favor transmission of R5 over X4 HIV-1. Because we have previously found that genes that restrict HIV-1 in cell lines do not always have appreciable activity against HIV-1 isolates in $CD4^+$ T cells [161] (HIV-1 restriction factors reviewed in [55, 57-60]), here we sought to identify host genes that preferentially inhibit X4 HIV-1 specifically in primary $CD4^+$ T cells. By performing comprehensive CRISPR screens and inactivation experiments in primary target cells, we discovered SLC35A2, a host protein involved in glycosylation, as an X4-specific restriction factor that has an opposite, proviral role during R5 infection. This study sheds light on the robust, tropism-dependent effects of host cell glycosylation on HIV-1 and suggests that target cell glycosylation may contribute to R5 selection during HIV-1 transmission.

Results

CRISPR screening in primary $CD4^+$ T cells identifies a candidate X4-specific restriction factor.

To identify X4-specific restriction factors, we performed CRISPR-knockout (KO) screens in primary $CD4^+$ T cells and infected with an X4 strain (LAI [106]) and an R5 virus (Q23.BG505 [105]) for comparison. We leveraged the HIV-CRISPR screening strategy [69], which we recently adapted for use in primary T cells [161]. Unlike the smaller tailored library utilized in our previous study (555 genes), here we interrogated ~2,000 host genes, including many IFN-stimulated genes (ISGs) as they are a major driver of antiviral immunity (**Figure 3.1A; Supplementary Table 3.1**). 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 (**Figure 3.1B; Supplementary Table 3.2**). By comparing R5 Q23.BG505 and X4 LAI screens and selecting LAI-specific hits, we identified 81 candidate X4-specific restriction factors, of which the top-scoring hit was SLC35A2 (**Figure 3.1C**). 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 (**Figure 3.1D**). SLC35A2 was not upregulated by IFN in primary $CD4^+$ T cells (**Figure 3.1E**). These data support SLC35A2's candidacy as a non-ISG host factor that may preferentially inhibit X4 HIV-1.

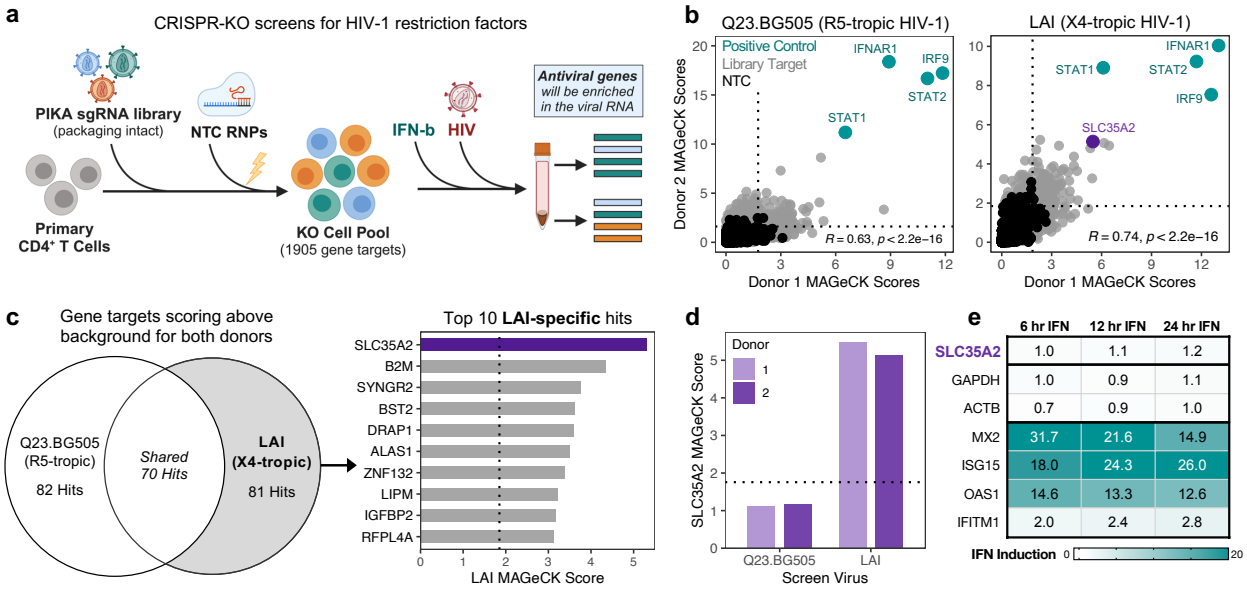


Figure 3.1. SLC35A2 is an X4-specific hit in a CRISPR-KO screen for HIV-1 restriction factors in primary CD4⁺ T cells.
(A) Schematic of HIV-CRISPR screening approach in primary CD4⁺ T cells. (B) Correlation of positive MAGeCK scores between CD4⁺ T cell donors in Q23.BG505 (left) and LAI (right) screens. 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 timepoint).

SLC35A2 KO differentially impacts CXCR4-tropic and CCR5-tropic HIV-1.

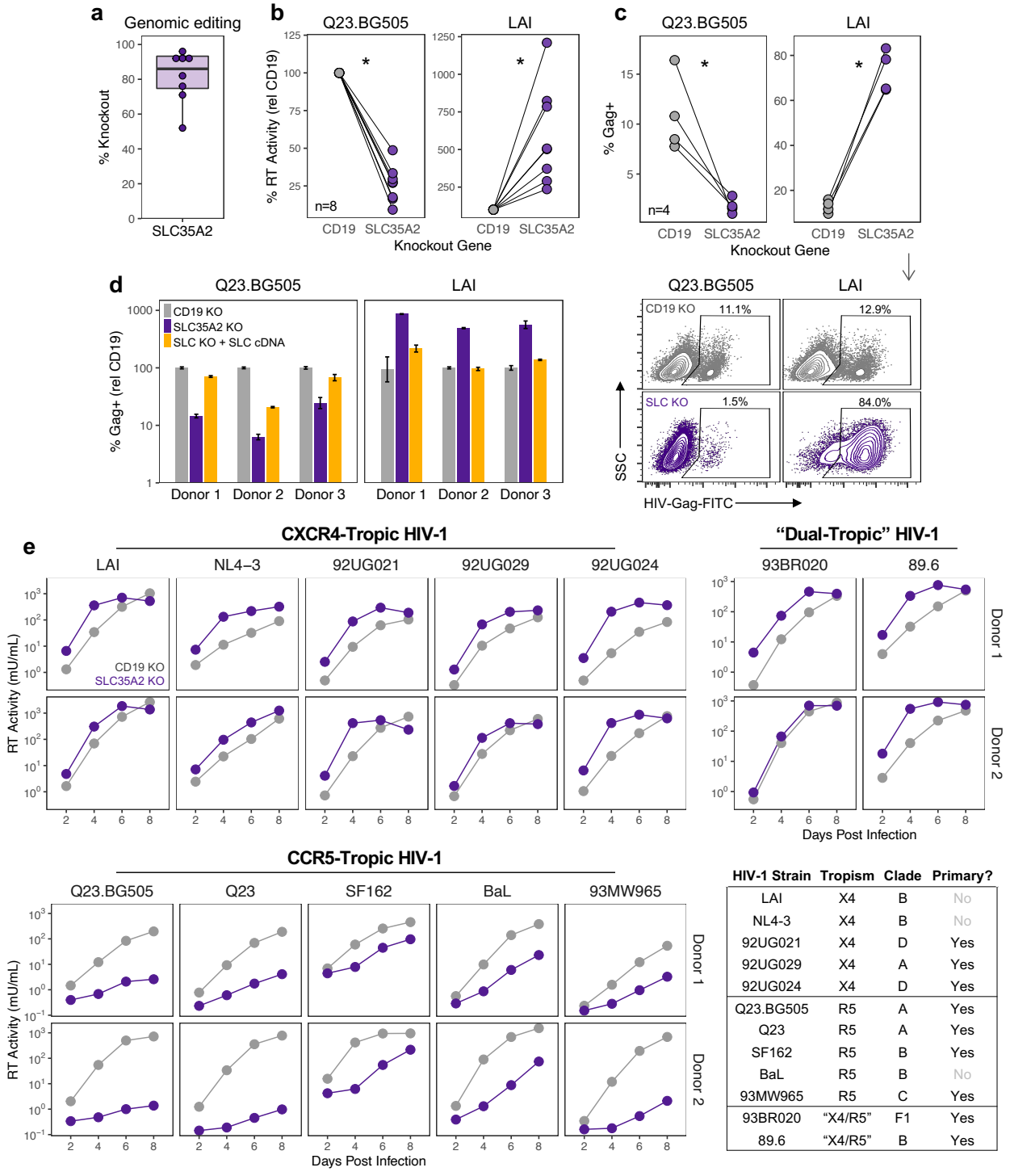
To validate our HIV-CRISPR screen findings, we performed single gene knockouts in primary CD4⁺ T cells, inactivating SLC35A2 and the B cell marker CD19 as a negative control. SLC35A2 editing led to consistently high levels of gene inactivation (average 82% KO) based on genomic DNA assessment (range: 52-96%; **Figure 3.2A**). 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) (**Figure 3.2B**). These

findings were consistent for two separate assays of HIV-1 infection: quantifying reverse transcriptase (RT) activity in infection supernatants (**Figure 3.2B, Supplementary Figure 3.1A**) and measuring the percentage of cells staining positive for HIV-Gag (**Figure 3.2C, Supplementary Figure 3.1B**). Notably, based on HIV-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) (**Figure 3.2C, Supplementary Figure 3.1C**). The opposite 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-1 strains that differ in their coreceptor usage.

To determine whether the observed phenotypes were specific to SLC35A2 inactivation, we exogenously expressed SLC35A2 cDNA in SLC35A2 KO $CD4^+$ T cells. Complementation led to full or partial rescue of wildtype HIV-1 infection levels (**Figure 3.2D**). 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-1 clades and LAI is lab-adapted whereas Q23.BG505 is not. To understand which viral features represent determinants of SLC35A2 KO's impact on HIV-1 infection, we infected SLC35A2 KO and CD19 KO $CD4^+$ T cells with five X4 and five R5 full-length HIV-1 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 (**Figure 3.2E, Supplementary Figure 3.2A**). Coreceptor tropism for each strain was

confirmed by individually inactivating CXCR4 and CCR5 in CD4$^+$ T cells (**Supplementary Figure 3.2B**). Notably, the effects of SLC35A2 KO were intermediate to those of coreceptor inactivation, suggesting that SLC35A2 partially inhibits, but does fully ablate, HIV-1 infection.



HIV-1 Strain	Tropism	Clade	Primary?
LAI	X4	B	No
NL4-3	X4	B	No
92UG021	X4	D	Yes
92UG029	X4	A	Yes
92UG024	X4	D	Yes
Q23.BG505	R5	A	Yes
Q23	R5	A	Yes
SF162	R5	B	Yes
BaL	R5	B	No
93MW965	R5	C	Yes
93BR020	"X4/R5"	F1	Yes
89.6	"X4/R5"	B	Yes

(figure legend on next page)

Figure 3.2. SLC35A2 KO differentially impacts CXCR4-tropic and CCR5-tropic HIV-1. (A) Percentage of SLC35A2 knockout based on evaluating genomic DNA editing by Synthego ICE. 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. (B) Reverse transcriptase (RT) activity relative to CD19 KO at 2 days post infection (dpi) (MOI=0.02) for the 8 independent experiments depicted in Panel A. Two-tailed paired t-test, asterisk indicates $p<0.05$. (C) (Top) Percentage of $CD4^+$ T cells staining positive for HIV-Gag at 3 dpi (MOI=1) (n=4 donors). Two-tailed paired t-test, asterisk indicates $p<0.05$. (Bottom) Representative flow plots from one donor. (D) 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 2-3 replicate infections. € 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-1 strains used for infections.

In addition to X4 and R5 HIV-1, we investigated the effect of SLC35A2 KO on two viruses considered dual-tropic (93BR020 [162] and 89.6 [163]). These results closely recapitulated those of X4 HIV-1 (**Figure 3.2E, Supplementary Figure 3.2A**), which agrees with our findings from coreceptor KO cells that indicated that both 93BR020 and 89.6 rely on CXCR4 substantially more than CCR5 (**Supplementary Figure 3.2B**). Inactivation of SLC35A2 therefore differentially impacts HIV-1 based on whether viruses use CXCR4 or CCR5 as a coreceptor for viral entry.

Host cell glycosylation impacts HIV-1 infection within a single round of infection.

SLC35A2 is a multichannel membrane protein that transports UDP-galactose from the cytosol into the Golgi [164-166]. 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 galactose. Naturally occurring polymorphisms in SLC35A2 are associated with a congenital disorder of glycosylation characterized by epileptic encephalopathy.

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 (**Figure 3.3A**). 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-glycans [167]. We

found that, on average, 92% and 90% of SLC35A2 KO cells were positive for GSL-II and VVL binding, respectively, compared to 3% and 2% of CD19 KO cells (**Figure 3.3B**). Complementation of SLC35A2 KO cells dramatically reduced lectin binding, demonstrating the specificity of this phenotype to SLC35A2 and further supporting the magnitude of SLC35A2 KO in our system (**Figure 3.3C**). Therefore, multiple lines of evidence indicate that editing of SLC35A2 in this cell type substantially disrupts host cell surface glycosylation by causing the truncation of both N- and O-glycans.

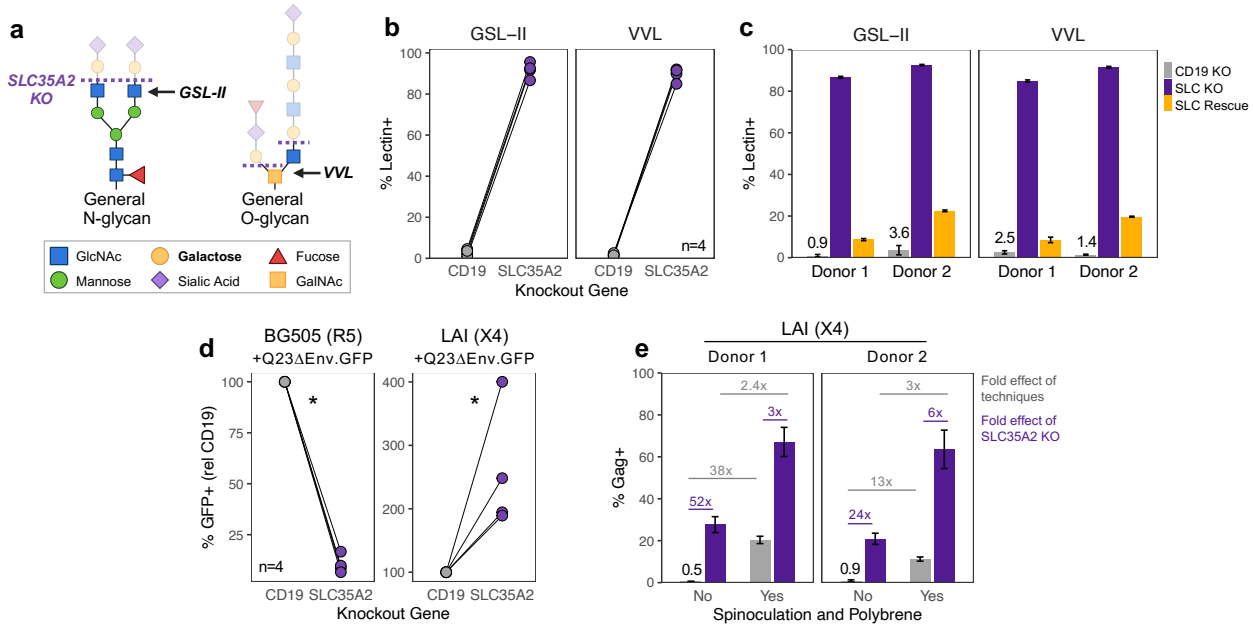


Figure 3.3. SLC35A2 inactivation causes truncated glycans on host cells and impacts HIV-1 infection within a single round of infection.
(A) General schematic of glycans, including the expected truncation when SLC35A2 is inactivated. The terminal residue binding sites of GSL-II and VVL lectins are indicated with arrows. (B) Percentage of cells staining positive for GSL-II and VVL lectins in $CD4^+$ T cells from three donors across single or replicate KO experiments for a total of four independent experiments. (C) Percentage of cells staining positive for GSL-II and VVL lectins in $CD4^+$ T cells in SLC35A2 KO cells delivered exogenous SLC35A2 cDNA (SLC Rescue). Data are shown as mean ± standard deviation of 3 replicate stains. (D) Percentage of GFP+ $CD4^+$ T cells, relative to CD19 KO, after 2 days of infection with GFP-expressing HIV-1 pseudoviruses (MOI=1) (n=4 unique donors). Two-tailed paired t-test, asterisk indicates $p<0.05$. (E) Percentage of $CD4^+$ T cells staining positive for HIV-Gag at 3 dpi (MOI=1) after infection with or without the use of spinoculation and polybrene. Data are shown as mean ± standard deviation of triplicate infections.

Though lectin staining demonstrates altered glycosylation on target cells, HIV-1 itself is heavily glycosylated. For instance, the envelope proteins (Envs) of the X4 and R5 viruses used in our study thus far (**Figure 3.2E**) contain a median of 28 putative N-glycan sites (X4 range: 28-34, R5 range: 25-30; p=0.14, two-tailed unpaired t-test). These viral glycans are likely similarly altered upon replication in SLC35A2 KO cells. Because our infections have utilized replication-competent HIV-1, the observed tropism-specific SLC35A2 KO HIV-1 phenotypes may be driven by altered glycans either on host cells or on HIV-1 virions after passaging in SLC35A2 KO cells. To distinguish these possibilities, we made X4 and R5 HIV-1 pseudoviruses, which are capable of infection but not making infectious progeny virions. We found that SLC35A2 KO impacted these HIV-1 pseudoviruses within a single infection cycle in the same manner as replication-competent virus, with reduced R5 infection (10.8-fold decrease, $p<0.0001$) and enhanced X4 virus infection (2.6-fold increase, p=0.049) (**Figure 3.3D, Supplementary Figure 3.1C**). The magnitudes of these effects were similar to those from replication-competent infection (**Figure 3.2C**) and consistent across four unique donors. Because single-cycle infections eliminate the possibility that SLC35A2 KO is acting on glycosylation of the virus, these results indicate that the effect of SLC35A2 KO on HIV-1 is driven by changes to glycans on target cells, not on progeny virus.

Our infections have utilized spinoculation and polybrene treatment, which are commonly used to increase infection efficiency by promoting viral attachment [168]. In their absence, SLC35A2 KO still conferred a robust X4 restriction phenotype – inactivation of SLC35A2 increased LAI infection by 52-fold (0.5% to 28% HIV-Gag+) and 24-fold (0.9% to 21% HIV-Gag+) in two donors (**Figure 3.3E**). In fact, this impact exceeded that of spinoculation and polybrene, which increased LAI infection by 38-fold and 13-fold for the two donors. Thus, SLC35A2 KO cells without the enhancement of infection by spinoculation and polybrene had higher infection levels than wildtype, CD19 KO cells with these interventions (Donor 1: 28% versus 20% HIV-Gag+, Donor 2: 21% versus 11% HIV-Gag+). These data not only demonstrate that SLC35A2 phenotypes are not dependent on *in vitro* techniques to promote viral infection, but they also

suggest that the impact of SLC35A2 KO on X4 HIV-1 infection exceeds that of the infection-enhancing methods of spinoculation and polybrene combined.

Changes to HIV-1 coreceptor levels do not reflect the pronounced SLC35A2 KO phenotypes.

Due to the tropism-dependent nature of SLC35A2's impact on HIV-1, we next investigated whether receptor (CD4) or coreceptor (CXCR4, CCR5) expression is impacted by SLC35A2. We assessed surface protein expression on wildtype and SLC35A2 KO cells on the day of HIV-1 infection by flow cytometry and found that essentially all cells expressed CD4 (**Figure 3.4A**). To capture relative protein levels, we examined the median fluorescence intensity (MFI) of cells expressing a given surface protein. Though SLC35A2 KO cells had a significantly elevated CD4 MFI in one donor ($p=0.04$), this effect was weak (increase of 12%) and not reproducible. SLC35A2 KO therefore does not alter CD4 expression in a magnitude or manner that would suggest an impact on HIV-1 infection.

Like CD4, there were also no significant differences in the percentage of cells staining positive for CXCR4 and CCR5, although modest effects were observed (**Figure 3.4B and 3.4C**). SLC35A2 KO cells displayed a slightly higher frequency of CXCR4+ cells than CD19 KO (average 98.3% and 96.5% CXCR4+, respectively; $p=0.08$) and less CCR5+ cells (average 17.9% and 24.5% CCR5+, respectively; $p=0.09$). The same trends were observed when MFI was examined, with an average 22.9% increase in CXCR4 MFI ($p=0.16$) and a 10.7% decrease in CCR5 MFI ($p=0.03$). These data suggest that SLC35A2 differentially impacts the expression of the HIV-1 coreceptors in the same direction as its effect on X4 and R5 HIV-1. However, the differences in expression based on SLC35A2 KO were modest and, notably, less pronounced than overall donor-specific differences in CCR5 expression. Because the effect sizes were low and not consistently significant, our results suggest that, while SLC35A2 KO's impact on coreceptor expression could be contributing to its effect on HIV-1 infection, it does not likely represent the dominant mechanism.

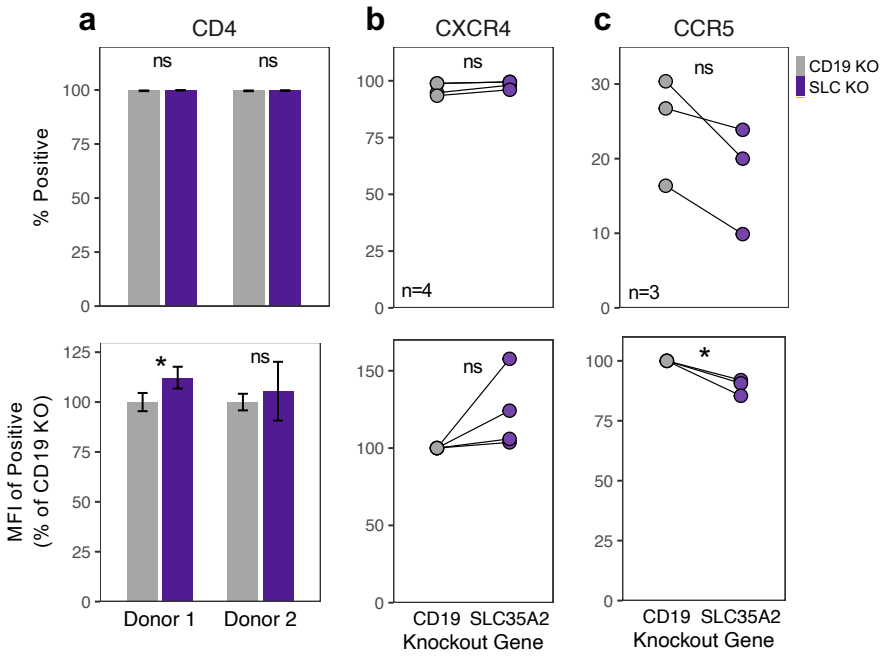


Figure 3.4. Changes to HIV-1 coreceptor levels do not reflect the pronounced SLC35A2 KO phenotypes.
(A, B, C) Surface protein expression of CD4, CXCR4, and CCR5, respectively, on $CD4^+$ T cells analyzed by flow cytometry. (Top) Percentage of cells staining positive for each surface protein. (Bottom) The median fluorescence intensity (MFI) of the positive population, relative to CD19 KO. The number of donors evaluated for CXCR4 and CCR5 are depicted within each paired dot plot. Data are shown as mean ± standard deviation of 3 replicate stains in bar plots. Two-tailed unpaired t-tests for Panel A, two-tailed paired t-tests for Panels B and C. Asterisk indicates $p<0.05$. Ns, not significant.

Discussion

In this study, we identify SLC35A2, a host protein involved in glycosylation, as a potent X4-specific cellular restriction factor of HIV-1 in the main target cells of HIV-1 infection, primary $CD4^+$ T cells. Interestingly, the presence of SLC35A2 has an opposite, enhancing effect on R5 infection. Our findings thus uncover an opposing role for $CD4^+$ T cell glycans on R5 and X4 HIV-1 infection, whereby R5 HIV-1 prefers wildtype target cell glycosylation and X4 viruses are more successful when host glycans are truncated, such as in the setting of SLC35A2 KO (**Figure 3.5**). Notably, the antiviral effect of SLC35A2 on X4 HIV-1 remained robust even in the absence of *in vitro* techniques commonly used to promote HIV-1 infection, further supporting the *in vivo* relevance of this restriction factor.

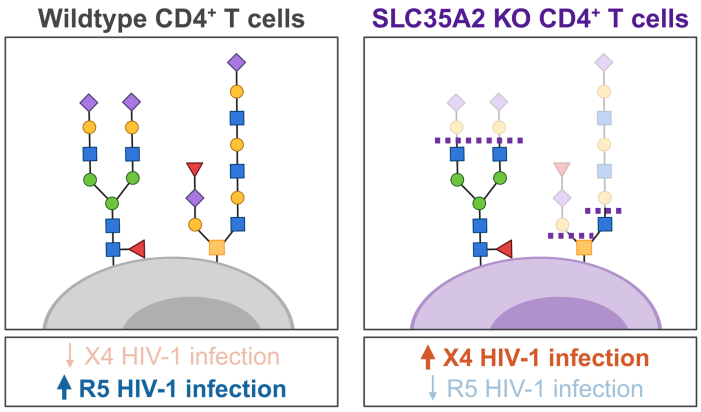


Figure 3.5. Working model of the impact of CD4⁺ T cell glycosylation on HIV-1 infection.

The results from our study uniquely demonstrate that the composition of glycans on host cells differentially impacts HIV-1 infection based on coreceptor tropism. Because this work was performed in a physiologically relevant, primary target cell model, it is plausible that SLC35A2-driven glycan features may contribute to R5 selection during natural HIV-1 transmission. In fact, mucosal CD4⁺ T cells in the vaginal tract, a common site of HIV-1 transmission, express *SLC35A2, CXCR4,* and *CCR5* RNA at levels comparable to or higher than those in the blood, which is the CD4⁺ T cell population used in our study (**Supplementary Figure 3.3**) [169]. These data indicate that SLC35A2 is present at functional levels in target cells at an important HIV-1 transmission site. This supports the possibility that this X4 restriction factor may contribute to R5 selection during transmission, which can be seeded by infection of blood and/or mucosal-derived CD4⁺ T cells depending on the transmission route and context of exposure. Future experiments using CD4⁺ T cells isolated from the mucosa or in *ex vivo* tissue explants will help address the role of SLC35A2 and host glycosylation on mucosal HIV-1 transmission suggested by this work. Notably, SLC35A2 KO reduces the efficiency of X4 infection but does not ablate it in the same way as coreceptor inactivation. These data are consistent with the fact that X4 viruses do evolve in the host, perhaps at times when the viral evolution pressures of adaptive immunity and host

cell availability outweigh that of SLC35A2-mediated restriction, resulting in the emergence of X4 viruses.

By uncovering a tropism-dependent role for target cell glycosylation on HIV-1, our study adds to recent evidence highlighting the importance of virus-glycan interactions during infection. Until now, the impact of host glycans on HIV-1 infection was singularly investigated in the context of R5 viruses. For instance, a recent study by Ma *et al* [170] showed 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-1 infection [170]. Further, removal of sialic acid directly reduced R5 infection. Because sialic acids are absent in the setting of SLC35A2 KO, our findings uncover a potential cellular driver for these findings. In addition to HIV-1, other viruses are also impacted by SLC35A2 and host glycans. SLC35A2 has been identified as a proviral factor for influenza A, which uses host sialic acids as receptors for entry [171, 172]. Additionally, in a CRISPR screen with VSV-G-pseudotyped HIV, SLC35A2 was a top antiviral hit [69], suggesting a possible role for this gene in VSV-G-mediated entry, which relies on the low-density lipoprotein receptor, a glycoprotein with extended O-glycans [173]. SLC35A2 may therefore broadly impact infection by viruses that rely on glycan interactions.

While our study highlights a fundamental role for host glycans on HIV-1, the precise mechanism(s) driving the tropism-dependent effect of SLC35A2 KO remain to be determined. Our single-cycle infection data point to a cell-intrinsic mechanism for the opposing effects of SLC35A2 on HIV-1 infection, as the magnitude of the effect of SLC35A2 inactivation on pseudovirus infection was similar to that observed using replicating viruses. Interestingly, SLC35A2 KO does differentially impact the prevalence and magnitude of CXCR4 and CCR5 protein expression in the same direction its effect on X4 and R5 infection. CXCR4 and CCR5 are both glycosylated [174-176], thus it is plausible that altering these glycans influences protein stability and/or trafficking to and expression on the cell surface. Nevertheless, the magnitude of the changes in coreceptor expression is weak, with SLC35A2 KO only increasing the frequency

of CXCR4+ cells from 96.5% to 98.3% (**Figure 3.4B**). It is therefore likely that SLC35A2's impact on HIV-1 is driven by another mechanism, although modulating coreceptor expression could contribute.

Investigating the interactions between HIV-1 and host surface glycans may provide insights into additional cell-intrinsic mechanisms contributing to SLC35A2's phenotypes. For instance, the glycans on both CXCR4 and CCR5 are on their extracellular N-termini [174-176] and this region of CCR5 is known to interact with HIV-1 [177]. It is therefore possible that the composition and structure of glycans on the coreceptors directly impact HIV-1 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-1 and the host cell may enhance attachment [178]. If broad glycan-glycan interactions were driving SLC35A2's tropism-specific effects, viral glycosylation levels would likely need to vary based on coreceptor usage. Because the X4 and R5 Envs in our study did not vary in their levels of glycosylation, general glycan-glycan interactions are unlikely to be driving SLC35A2's phenotypes. Finally, though our data support a cell-intrinsic mechanism for SLC35A2, glycans on viruses produced in SLC35A2 KO cells are also likely truncated. These changes may impact HIV-1 virion infectivity and compound the tropism-dependent effects on HIV-1 infection.

Elucidating the mechanisms driving the SLC35A2 KO phenotypes would be expedited by working in a cell line model. However, cancer cells, which are often used to establish cell lines, have a high prevalence of altered glycans [179]. In fact, Jurkat T cells, a T cell line commonly used in HIV-1 studies, have truncated O-glycans due to a mutation in a chaperone required for O-glycan synthesis, C1GALT1C1/Cosmc [180]. Efforts to study the impact of host cell glycans on HIV-1 in cell lines must therefore evaluate the suitability of the cell line and ensure that it recapitulates wildtype glycosylation of primary $CD4^+$ T cells.

We identified SLC35A2 by performing comprehensive HIV-CRISPR screens in primary $CD4^+$ T cells. This robust screening approach, coupled with the relevance of a primary cell model,

represents a powerful tool for identifying cellular factors that impact HIV-1 and likely influence natural infection. Though we focused on SLC35A2 in this study, there are additional candidate restriction factors that could be pursued. For instance, B2M, SYNGR2, and BST2 were the next top-scoring X4-specific hits in our screens. B2M is a component of major histocompatibility complex (MHC) class I molecules, which may suggest a previously unexplored, cell-intrinsic role for MHC I in inhibiting HIV-1 infection, perhaps X4 HIV-1 in particular. Of note, B2M has recently been found to facilitate Vaccinia virus entry, providing evidence for a role for MHC I during viral replication [181]. SYNGR2, a membrane protein implicated in regulating membrane trafficking, was identified as a proviral factor for bunyavirus infection and is hypothesized to promote the formation of bunyavirus replication complexes [182]. Unlike B2M and SYNGR2, BST2/Tetherin, is a well-studied HIV-1 restriction factor that prevents viral egress and is antagonized by the HIV-1 accessory protein Vpu (reviewed in [59]). Tetherin has now been classified as an X4 LAI-specific hit in three independent HIV-CRISPR screens – those reported here, our previous screens with a different library [161], and HIV-CRISPR screens in THP-1 cells [69]. Because Tetherin is more sensitive to antagonism by Vpu proteins from primary HIV-1 isolates [100], which are general R5 viruses, these screen results may be driven by Vpu differences or coreceptor tropism. Our findings thus identify additional candidate X4-specific restriction factors in $CD4^+$ T cells that can be validated against a broad panel of X4 and R5 viruses to understand whether they confer tropism-specific effects.

Overall, our findings identify SLC35A2 as a cell-intrinsic, X4-specfic restriction factor and reveal a tropism-dependent impact of target cell glycosylation on HIV-1. These results provide new insights and a promising avenue of investigation to address a long-standing question in the field regarding HIV-1 transmission. Understanding the interactions between host glycans and HIV-1 that confer differential impacts on R5 and X4 HIV-1 may inform the development of new prophylactic approaches to specifically target viral variants that are successful in the HIV-1 transmission bottleneck.

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.BG505 [105] is a clade A, CCR5-tropic HIV chimeric virus derived from a full-length provirus isolated early in infection (Q23 [104]) and bears the BG505.C2 envelope [140] of a well-studied clade A transmitted/founder virus. HIV LAI is a clade B, CXCR4-tropic HIV strain, as previously described [106]. Infectious molecular clones (IMCs) of Q23.BG505 [105] and LAI [106], as well as Q23 [104], SF162 [183], and NL4-3 [184] (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 Gallo [141]; 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 Collman [185].

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.C2 [140] or LAI (gift from Dr. Michael Emerman) and Q23ΔEnv.GFP [186] (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 µM filter (Millipore Sigma #SCGP00525). Transfection supernatants were concentrated using Amicon 100 KDa concentrators (Millipore Sigma #UFC910024) and titered on TZM-bl cells, as previously described [161].

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 plasmid [69] and cloned into the HIV-TOP opt mChΔW [16] plasmid for primary $CD4^+$ T cell screens, as previously described [161]. 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 vector [91]. Specifically, the codon-optimized SLC35A2 fragment, WPRE-LTR from the HIV-CRISPR PIKA plasmid [69], 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 mChΔW 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 µL serum-free DMEM, and 4.5 µL TransIT-LT1 transfection reagent (Mirus Bio #MIR-2305). Lentiviral transfections were concentrated over a sucrose gradient, as previously described [161].

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 3.3**. RNPs were prepared with Cas9-NLS protein (UC Berkeley QB3 MacroLab) as described previously [161].

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 group for different KOs and HIV-CRISPR screens than those reported here [161]. 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 described [161]. 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 µL of NTC complexes. For RNPs with Synthego guides, cells were gently resuspended with 25 µL 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 µL 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 µg/mL polybrene (Millipore Sigma #TR-1003-G) as previously described [161]. 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∆W lentivirus containing the PIKA library by spinoculation in the presence of 8 µg/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-β1a (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 µL 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 described[161]. Pooled libraries were sequenced on an Illumina HiSeq 2500 (Fred Hutch Shared Resources Genomics Shared Resource). Illumina sequencing reads were analyzed by MAGeCK-Flute [113] to generate guide counts and gene-level enrichment data (**Supplementary Table 3.2**; 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 3.1**). 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 µg/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 [142, 161]. 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

For high MOI infections (MOI=1), cells were collected for HIV-Gag staining 3 dpi (**Fig, 3.2C, 3.2D, 3.3E; Supplementary Figure 3.1B**). For low MOI infections (MOI=0.02), infections were allowed to proceed for 6 days before harvesting due to the slower nature of viral infection and spread (**Supplementary Figure 3.2A**). HIV-Gag staining is less optimal for evaluating low MOI infections due to reduced signal and therefore serves as a secondary read out for this experiment, supporting measurements of RT activity (**Figure 3.2E**).

After collection, cells were 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 detail [161]. In brief, regions containing SLC35A2 guide cut sites were amplified from genomic DNA (forward primer: 5' AAGCTGCCCACAAATAGCCT;

reverse primer: 5' AAGCTGCCCACAAATAGCCT) and sequenced (sequencing primer: 5' AAGCTGCCCACAAATAGCCT) 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.

Surface protein staining

On the day of HIV-1 infection, 2e5 edited $CD4^+$ T cells per stain replicate were collected and resuspended in 20 uL FACS buffer with either 2 uL anti-CD4-APC antibody (BD Biosciences #555349) or 1 uL anti-CXCR4-APC 12G5 antibody (BD Biosciences #555976) and 2.5 uL anti-CCR5-PE 2D7 antibody (BD Biosciences #555993). Cell-antibody mixtures were incubated for 1 h at room temperature in the dark. Cells were then resuspended in 1% paraformaldehyde in FACS buffer 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-β1a (SRA: PRJNA921704) [161]. 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 individuals [169]. 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 package[143]. 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.

Analysis of HIV-1 Env glycosylation

Env protein sequences for the X4 and R5 viruses used in **Figure 3.2E** were obtained via http://www.hiv.lanl.gov/ in October 2023. Putative N-glycan sites were considered sites with asparagine-X-serine/threonine, where X is any amino acid except proline.

Statistical analysis

The number of technical and biological replicates for each experiment are described in the figure legends. Paired dot plots and two-tailed paired t-tests were used to evaluate comparisons between CD19 KO and SLC35A2 KO cells in experiments with a least three donor replicates, pairing by donor (**Figure 3.2B, 3.2C, 3.3D, 3.4B, 3.4C**). For experiments with less than three donor replicates, bar plots and two-tailed unpaired t-tests were applied to evaluate within-donor differences (**Figure 3.3A**). All of these datasets were found to be normally distributed by Shapiro-Wilk tests. Prism 9 was utilized for statistical analyses.

Data availability

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

Addendum

This section will describe results from experiments that are not part of the published manuscript for this project. They provide some preliminary data to guide future studies.

Removing CCR5 glycans does not reproduce SLC35A2 KO effect on R5 HIV-1

Given the tropism-dependent effect of SLC35A2 KO and SLC35A2's role in glycosylation, we hypothesized that the changes to the glycans on the coreceptors are responsible for SLC35A2 KO's impact on HIV-1. CXCR4 has one N-glycosylation site (N11) [174] and CCR5 has two possible O-glycans (S6, S7) [175]. To test our hypothesis, we delivered wildtype and glycan-mutated versions of the coreceptors to primary $CD4^+$ T cells that had their endogenous coreceptors inactivated. The following mutations were chosen to remove glycans sites: CXCR4 N11Q, CXCR4 T13A, and CCR5 S6A S7A. Codon-optimized coreceptor cDNA was synthesized as gene fragments from Twist and cloned into the TOP-CAR vector, which was used for lentiviral transfections [91]. To test these constructs, primary $CD4^+$ T cells were transduced with lentivirus delivering the codon-optimized coreceptors and then were subsequently edited for their endogenous coreceptors by nucleofection, using the same lentiviral transduction and single-gene editing methods and experimental timeline described in the main manuscript.

We first examined whether transduction with the wildtype and glycan-mutated coreceptors led to similar protein levels, which is necessary to be able to accurately compare infection phenotypes. The percentage of cells expressing CCR5 in wildtype CCR5-transduced

and CCR5 S6A S7A-transduced cells were comparable (67% and 58%, **Figure 3.6A**). These results indicate that complementation was successful and provided cells with similar levels of CCR5 for comparative studies of wildtype and mutant receptors. However, CXCR4 levels were low in all CXCR4-transduced conditions, with levels that matched the staining pattern of CXCR4 KO cells. These data indicated that there was an unknown issue with the CXCR4 constructs. To further investigate, we performed droplet digital PCR on transduced cells to measure the number of integrated vector copies per cell (**Figure 3.6B**) and, in a separate transduction experiment, we used qPCR to detect whether the constructs were being transcribed (data not shown). Both demonstrated that the CXCR4 constructs were effectively delivered to cells, integrated, and transcribed. Of note, CXCR4-transduced cells were not permissive of X4 LAI infection, another indication that the coreceptor was not being expressed and that the lack of expression observed by flow cytometry was not due to technical issues with the stain or antibody (data not shown). Therefore, for an unknown reason, the CXCR4 constructs did not yield protein expression at the cell surface of primary $CD4^+$ T cells. It is worth noting that the CXCR4 cDNA used here began with the start codon and did not contain the 5' untranslated region. Theoretically, I would not expect this to be an issue, but the pBABE.CXCR4 vector (HIV Reagent Program #ARP-3332), which is commonly used for CXCR4 expression, does contain this region. Therefore, it could be worth cloning the exact sequence in the pBABE.CXCR4 vector into the TOP-CAR plasmid.

Because the CCR5-transduced cells displayed similar levels of CCR5 expression, we were able to directly compare their infection phenotypes. Interestingly, we observed no difference in R5 Q23.BG505 infection between the cells with wildtype CCR5 and the glycan-mutated CCR5, despite observing the expected phenotype with SLC35A2 KO within the same experiment (**Figure 3.6C**). These data suggest that completely removing the glycans on CCR5 does not impact HIV-1 infection nor recapitulate the effect of SLC35A2 KO.

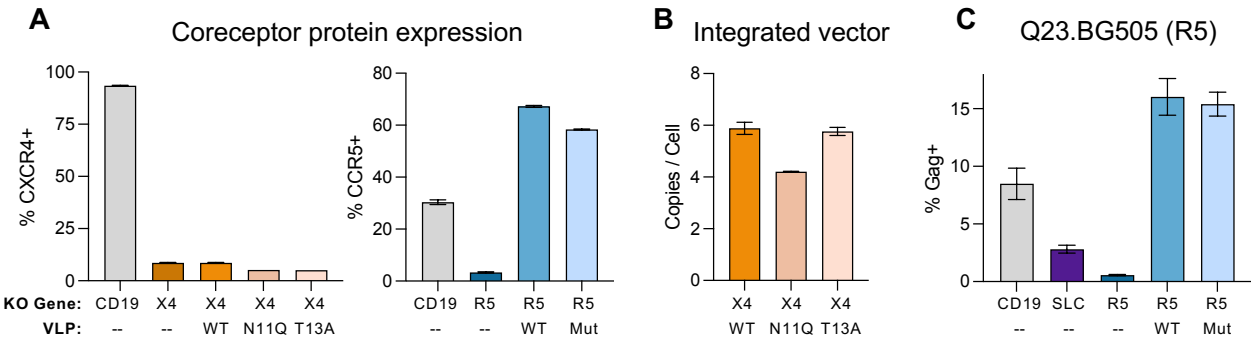


Figure 3.6. Removing CCR5 glycans does not reproduce SLC35A2 KO effect on R5 HIV-1. (A) CXCR4 and CCR5 surface protein expressed as measured by flow cytometry. (B) The number of integrated lentiviral vector copies per cell in CXCR4-transduced cells. (C) The percentage of cells staining positive for HIV-Gag at 3 dpi (MOI=1). All figures depict data from one donor, 2-3 technical replicates. WT, wildtype.

While informative, this experiment unfortunately does not directly test our original hypothesis that changes to the glycans on the coreceptors are driving the effects of SLC35A2 KO on HIV-1. Glycans on SLC35A2 KO cells are truncated, not absent, and this difference may be a critical determinant in the interactions between HIV-1 and the coreceptor. Because it is not possible to selectively truncate the glycans on just the coreceptors without globally impacting glycan structures on all host proteins, understanding the importance of the coreceptor glycans may require extracting these proteins from cells and performing biochemical binding assays. These potential approaches are discussed in Chapter 4.

Establishing a cell line model for SLC35A2 KO

The optimization and performance of mechanistic experiments would be much faster, convenient, and less resource intensive in a cell line as opposed to primary cells. Before adopting a cell line model, it is first necessary to demonstrate that SLC35A2 KO in the cell line recapitulates the phenotypes we observe in primary cells – truncation of glycans and an opposing effect of R5 and X4 HIV-1. We first evaluated whether Jurkat T cells would be a suitable cell line. We inactivated SLC35A2 in Jurkat cells by nucleofection using the same approach as in primary cells and assessed the status of glycosylation by lectin staining. Control

and SLC35A2-edited cells displayed the same striking trends in GSL-II binding as in primary cells, with 3% and 86% cells staining GSL-II+ for CD19 and SLC35A2 KO, respectively (**Figure 3.7A**). To our surprise, we found that wildtype Jurkat T cells were 100% VVL+, indicating that all of the O-glycans on these cells were already truncated prior to SLC35A2 KO. Indeed, we later found that Jurkat cells are known to have truncated O-glycans due to a mutation in a chaperone called C1GALT1C1 (also called Cosmc) that is required for O-glycan synthesis [180]. Due to this underlying abnormality in glycosylation, Jurkat cells are not a suitable cell line for SLC35A2 studies.

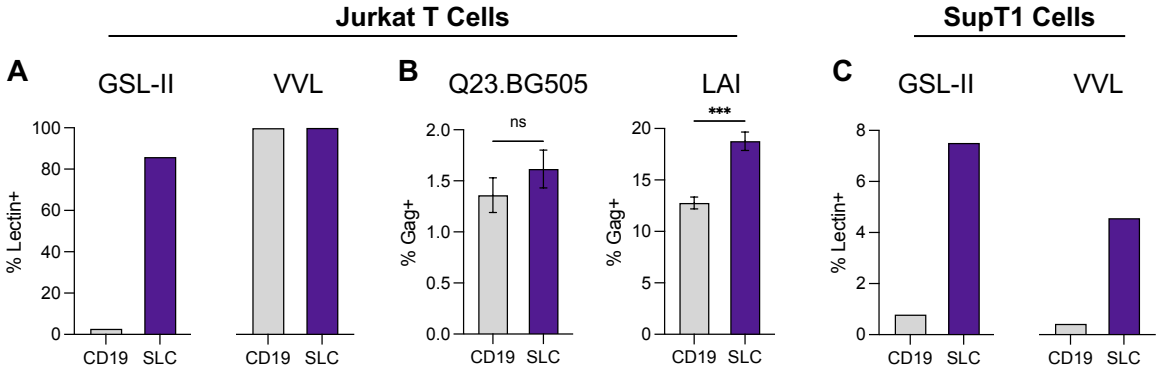


Figure 3.7. Establishing a cell line model for SLC35A2 KO.
(A) Percentage of cells staining positive for GSL-II and VVL lectins in Jurkat T cells.
(B) Percentage of cells staining positive for HIV-Gag at 3 dpi (MOI=1). Paired t-test, ***p<0.001.
(C) Percentage of cells staining positive for GSL-II and VVL lectins in SupT1 cells.

Despite Jurkat cells having atypical glycosylation, we were still interested in infecting these cells with HIV-1. We found that R5 Q23.BG505 infection was not significantly impacted by SLC35A2 KO in Jurkat cells and that both CD19 KO and SLC35A2 KO had very low levels of Q23.BG505 infection (average overall 1.5%) (**Figure 3.7B**), which is comparable that of SLC35A2 KO primary cells (1.8%, **Figure 3.2C**). Because CCR5 harbors O-glycans, these results are consistent with our hypothesis that the glycans on the coreceptors are driving SLC35A2 KO phenotypes. Unlike the lack of impact on Q23.BG505, SLC35A2 KO in Jurkat cells led to increased levels of LAI infection, in line with the trends we observed in primary cells (**Figure 3.7B**). Because CXCR4 has an N-glycan, this result is, again, consistent with our

hypothesis. It is worth noting, however, that the magnitude of SLC35A2 KO's impact on LAI (1.5-fold) is much less potent than that in primary cells (5.9-fold, **Figure 3.7B**). Whether this is due to aberrant O glycosylation in Jurkat cells or something else remains unclear.

Though Jurkat cells are not a suitable cell line for studies into the mechanism of SLC35A2's impact on HIV-1, the trends of these results are consistent with our hypothesis that the glycans on the coreceptors are driving the effects of SLC35A2 KO. Moreover, they offer a potential experimental approach – it may be informative to inactivate the C1GALT1C1/comsc chaperone in primary cells to specifically capture the impact of O-glycans on HIV-1.

We next evaluated whether the SupT1 cell line recapitulated SLC35A2 KO trends in primary cells. Though we found a clear increase in lectin binding in SLC35A2 KO SupT1 cells, the magnitude of this effect was very small (**Figure 3.7C**). SLC35A2 KO led to 7.5% and 4.6% of cells staining positive for GSL-II and VVL, respectively, compared to roughly 90% in primary cells (**Figure 3.3B**). Importantly, this difference cannot be explained by a lack of editing, as these SLC35A2 KO cells displayed 87% KO according to ICE analysis. These data demonstrate that SLC35A2 KO in SupT1 cells does not lead to the same level of truncated glycans as we see in primary cells and is therefore not a suitable model for SLC35A2 KO mechanistic studies.

Chapter 4

CONCLUSIONS AND FUTURE DIRECTIONS

For years, the HIV-1 field has relied on studying virus-host interactions using relatively low throughput approaches, immortalized or cancer cell lines, and lab-adapted viruses. These technical aspects have limited our understanding of the complex web of host factors that impact the course of natural HIV-1 infection. In the projects presented in this thesis, we describe the adaptation of the HIV-CRISPR screening approach [69] for use in the main target cells of HIV-1 infection, primary $CD4^+$ T cells, and the implementation of these screens, which uncover previously undescribed HIV-1 restriction factors. In Chapter 2, we identified three ISGs that had not been known to inhibit HIV-1 in other models of infection (HM13, LAP3, IGFBP2). Our targeted knockout experiments with multiple ISGs also demonstrated that IFN restriction of HIV-1 is the result of many antiviral genes working collectively, which is a model that have been proposed but not directly shown in prior studies as they tended to focus on singular ISGs. In Chapter 3, we identified an X4-specific restriction factor, SLC35A2, and uncovered an opposing role for host cell glycosylation on X4 and R5 HIV-1 infection by infecting cells with a broad range of HIV-1 strains. These studies highlight the value of utilizing primary cells, a diverse range of HIV-1 strains, and high throughput screens by providing new insights on the restriction factors that likely influence HIV-1 during transmission and early infection.

HIV-CRISPR screens in primary $CD4^+$ T cells

In addition to the biological findings reported in this thesis, the technical advancement of optimizing HIV-CRISPR screens for primary $CD4^+$ T cells also represents a major contribution to the HIV-1 field. We demonstrate that HIV-CRISPR screens in primary cells reliably enrich for positive control genes and demonstrate high reproducibility between independent screens with guide libraries targeting up to roughly 2,000 genes (15,414 guides), though we expect that

larger library sizes would be suitable as long as sufficient cell coverage is maintained. We performed HIV-CRISPR screens in primary cells with two libraries – a custom library that was designed based on in-house bulk RNA-seq data (CD4-ISG library, 555 genes) and a larger, pre-existing library informed by other ISG libraries and transcriptomic analyses of various cell lines (PIKA library, 1906 genes [69]). Interestingly, screens with the tailored CD4-ISG library yielded stronger agreement between donors (CD4-ISG screens Pearson R: 0.83-0.89; PIKA screens Pearson R: 0.63-0.74) and proportionally more hits (20.5-30% of genes were hits in CD4-ISG screens; 8% for PIKA screens) than those with the PIKA library (**Figure 2.4B, Figure 3.1B**). These trends may be the result of increased experimental noise in PIKA screens due to the PIKA library targeting a large proportion of genes that are not expressed in $CD4^+$ T cells. It may therefore be advantageous for future investigations to either use the CD4-ISG library or leverage the bulk RNA-seq dataset reported in Chapter 2 to inform the design of new custom libraries targeting genes expressed in $CD4^+$ T cells.

The unbiased nature of HIV-CRISPR screens allowed us to identify antiviral genes that had not been previously described (HM13, IGFBP2, LAP3, SLC35A2). To understand the genes that may work in concert with these new factors, it may be informative to perform HIV-CRISPR screens in, for instance, HM13 KO cells. Genes enriched in screens carried out in HM13 KO cells should represent factors that have non-redundant functions with HM13. This selection strategy would be particularly useful in identifying potent networks of antiviral ISGs. Future iterations of HIV-CRISPR screens in primary cells could also augment screen characteristics, such as IFN dose, viral MOI, length of infection, and viral strains, or analysis strategy, such as validating shared or virus-specific hits, to address particular research questions.

Though we performed HIV-CRISPR screens in activated bulk $CD4^+$ T cells to identify antiviral genes, future investigations could modify this system for other research purposes. For instance, the original HIV-CRISPR screening approach, which was developed in the THP-1 monocytic cell line for identifying HIV-restricting genes [69], has now been adapted for use in

the Jurkat T cell line to screen for proviral, dependency factors [187] and in J-Lat cell lines to select for latency maintenance factors [188]. Similarly, HIV-CRISPR screens in primary T cells could be adapted to select for proviral factors. Our screening approach should also be amenable to using specific subsets of $CD4^+$ T cells, as long the cells are activated and dividing for efficient CRISPR-Cas9 editing.

As with all screening platforms, HIV-CRISPR screens in primary cells have several important limitations. Screen results should always be validated by individually inactivating genes, ideally with different guides than those in the library, to verify that hits are not false positives. Such false positives may arise due to experimental noise, library guides having off-target effects, and gene KO impacting cell growth. Just as positive hits must be validated, genes that are not enriched cannot be assumed to not impact HIV-1 infection. Experimental design (virus, length of infection, viral MOI, donor, cell activation state, IFN/drug dosage) and the editing efficiency of library guides can vastly influence screen results, particularly in primary cells as they are highly sensitive to external stimuli. Therefore, one cannot definitively assert that a gene does not impact HIV-1 infection in primary T cells based on one or a few HIV-CRISPR screens. Rather, the HIV-CRISPR screening approach in primary target cells is a powerful tool for the unbiased screening of large gene libraries and for the selection of genes that impact HIV-1 replication. Screen hits can then be validated and mechanistically explored to understand new aspects of HIV-host interactions.

IFN-mediated restriction of HIV-1

In Chapter 2, we identified three previously undescribed cellular effectors of IFN restriction of HIV-1 (HM13, LAP3, IGFBP2) and found that two ISGs studied in other models of infection (IFI16, UBE2L6) inhibit primary HIV-1 in primary target cells. HM13, UBE2L6, and IFI16 stand out among this set of genes, as these factors impacted both Q23.BG505 and LAI in

single KO experiments. Furthermore, HM13 met our IFN effector criteria for four other diverse HIV-1 strains. HM13 thus represents a newly identified, broadly acting HIV-1 restriction factor.

Because the nature of HM13 restriction of HIV-1 remains unknown, HM13 is a high priority candidate for future mechanistic evaluation. HM13 is a peptidase that localizes to the endoplasmic reticulum and is known to physically interact with HIV-1 gp160 [116, 117]. HM13 may therefore inhibit HIV-1 by interfering with Env processing to reduce the amount and/or infectivity of virus produced during HIV-1 replication. While many experiments could inform this mechanism, it would first be helpful to determine the timing of HM13 restriction during the HIV-1 replication cycle by infecting HM13 KO and control cells with single-cycle HIV-1 pseudotyped-virus. With the hypothesis that HM13 impacts HIV-1 after viral translation, we would not expect HM13 to impact pseudotyped-virus infection. Next, HM13's impact on Env and virion infectivity should be examined. Virus passaged in HM13 KO and controls cells could be assessed by p24 ELISA to quantify the amount of virus produced, titered to measure relative infectivity, and utilized in a Western blot to stain for Env, using Gag staining as a loading control (similar to that in [73]). If HM13 is impacting gp160 processing, we would expect Env staining by Western blot to either be faint in virus collected from control versus HM13 KO cells due to less protein overall or skewed towards existing as unprocessed gp160. Results from these experiments should begin to uncover how HM13 inhibits HIV-1 in primary target cells.

Additional results reported in Chapter 2 warrant future investigation. For example, we were particularly surprised by the finding that MX2 KO did not substantially impact Q23.BG505, despite it increasing replication of LAI, and that all prior reports on MX2 and HIV-1 had exclusively used clade B viruses. It is therefore unclear whether Q23.BG505 is an outlier strain not impacted by MX2 or if MX2 is less effective against non-clade B HIV-1. Future work should infect MX2 KO and control primary cells with a broad panel of viruses, like that in Chapter 3, to understand the breadth of MX2 restriction and whether there are naturally occurring CA residues that impact HIV-1 sensitivity to MX2 restriction. In addition to following up on specific

genes like HM13 and MX2, there are also many more screen hits that we did not validate that could represent novel HIV-1 restriction factors. Finally, as more HIV-restricting ISGs are identified, it will be important to perform additional combinations of multi-ISG KOs to determine sets of ISGs that confer potent restriction of HIV-1, as these genes likely represent selection pressures on HIV-1 during acute infection.

Our multi-gene KO results support a model in which several ISGs contribute to IFN restriction of HIV-1. Accordingly, we find that the effect size of each individual ISG KO (2.1-4.2-fold) is much smaller than the overall impact of IFN treatment (55-62-fold). It is therefore important to note that single KO experiments could lead to false negatives when categorizing IFN effectors due to the low relative signal of individual ISGs. A possible approach to combat this issue could be to treat cells with a lower dose of IFN in KO experiments. To test this approach, one could perform bulk RNA-seq on primary cells treated with different doses of IFN, determine ISGs expressed at lower doses, identify those that are screen hits, and then perform validation experiments with various doses of IFN. I would expect that the relative contribution of these ISGs would be greater at lower IFN doses. Because of this challenge, studies evaluating many ISG candidates, like ours, must apply the same analysis criteria to all gene KOs to fairly compare antiviral potencies between genes and across viruses. Our IFN effector analysis criteria described in **Supplementary Figure 2.2** can thus serve as an example for future ISG studies.

Overall, Chapter 2 supports the feasibility of comprehensively interrogating ISG restriction of HIV-1 in primary target cells and demonstrates that such efforts can uncover previously unknown HIV-1 restriction factors. We hope future work will leverage these techniques, mechanistically explore the newly identified antiviral ISGs, and pursue the experiments we propose here to gain a better understanding of the cellular ISGs that inhibit HIV-1.

The role of glycosylation on HIV-1 infection

In Chapter 3, we identify an X4-specific restriction factor, SLC35A2, that influences host cell glycosylation and impacts HIV-1 infection in a tropism-dependent manner during early viral replication. Though host sialic acid residues have been previously shown to facilitate R5 infection [170], our study was able to uniquely capture the tropism-specific effect of host glycans by infecting with many X4 and R5 viruses. This differential effect is critical to the implications of our findings, as we are unaware of any cellular factors besides CD4 and the coreceptors themselves that influence HIV-1 infection according to tropism.

The findings reported in Chapter 3 represent just the beginning of our understanding of how host cell glycans impact HIV-1 infection. Many important mechanistic questions remain, including understanding what stage of the viral replication cycle is impacted by SLC35A2 KO, whether specific glycoproteins or the overall glycan landscape are causing these effects, how altered glycosylation impacts viral infectivity, and what precise virus-host interface is impacted with SLC35A2 KO. It is likely that SLC35A2 KO is impacting viral fusion due to the role of the coreceptors in this stage of HIV-1 infection and the early effect of SLC35A2 KO based on pseudotyped-virus infections. Future work could use the BlaM-Vpr fusion assay [189] with X4 and R5 viruses to address this. If fusion is impacted by SLC35A2 KO, it will be necessary to investigate the specific interactions between HIV-1 Env and the coreceptors. Such investigations could include binding assays between X4 and R5 Env with CXCR4 and CCR5, respectively, pulled down from wildtype and SLC35A2 KO cells. If SLC35A2 KO leads to increased binding between R5 Env and CCR5 and decreased binding between X4 Env and CXCR4, it would be informative to perform mass spectrometry on the coreceptors purified from wildtype and SLC35A2 KO cells. This approach would not only detail the composition of glycans on the coreceptors in wildtype cells, which has not previously been described, but it would also reveal the specific changes in glycosylation that confer the effects on HIV-1 observed with SLC35A2 KO.

Despite the relevance of a primary cell model, it would take a considerable amount of time to optimize and perform mechanistic studies in primary $CD4^+$ T cells. It would therefore be advantageous to develop a cell line model to perform initial mechanistic studies of the impact of SLC35A2 KO on HIV-1. While cell line systems are typically considered easier to implement than primary cell models, we have experienced obstacles in the context of this study (see Addendum for Chapter 4). Unfortunately, many cancer cells have altered glycosylation [179]. This includes the commonly used Jurkat T cell line, which has truncated O-glycans due to a mutation in a chaperone required for O-glycan synthesis [180]. We have also found that SLC35A2 KO in SupT1.CCR5 cells did not yield the same levels of truncated glycans as in primary cells. A transformed, immortalized cell line may therefore be more promising than a cancer cell line for studies concerning host cell glycosylation.

In addition to elucidating the mechanism of SLC35A2's impact on HIV-1, there is an immediate interest to understand the significance of these interactions *in vivo*. SLC35A2 inhibits infection of X4 viruses and is known to be expressed in $CD4^+$ T cells at common HIV-1 transmission sites. Therefore, it is plausible that SLC35A2 and consequently host cell glycosylation contribute to R5 selection during the HIV-1 transmission bottleneck. Though our data in primary target cells from the blood are consistent with this hypothesis, additional experiments are needed to directly support a role for host glycans on mucosal transmission of HIV-1. This work would require isolating $CD4^+$ T cells from human FGT tissue samples, such as from biopsies or elective hysterectomies, and inactivating SLC35A2. One could then perform competition experiments with reporter viruses, which would involve infecting wildtype and SLC35A2 KO tissue cells with an equal mixture of, for instance, X4.GFP and R5.RFP viruses. If SLC35A2 impacts mucosal transmission of HIV-1, we would expect the ratio of R5 to X4 infection in wildtype cells to be higher than that in SLC35A2 KO cells. These data would provide a direct role for SLC35A2 in the first target cells that HIV-1 encounters in the mucosa.

The critical next steps in this project are therefore to define the mechanism of SLC35A2's effect on HIV-1 and to understand its impact on mucosal HIV-1 transmission. Together, this work could inform the nature of one of the selection pressures influencing HIV-1 transmission and reveal features of T/F viruses that could be targeted for HIV-1 prevention strategies.

Conclusions

The projects described in this thesis provide new insights into the complex network of cellular factors that impact HIV-1 transmission and acute infection. The novelty and significance of these findings also emphasize the benefit of utilizing a primary cell system. There are likely more antiviral genes and unique mechanisms of HIV-1 restriction that remain to be discovered. Future studies that implement HIV-CRISPR screens in primary cells and continue the investigations described here will bring us closer to a more comprehensive understanding of the many ways host genes impact HIV-1 infection.

REFERENCES

1. UNAIDS. *Fact sheet - Latest global and regional statistics on the status of the AIDS epidemic.* 2023 July 13, 2023; Available from: https://www.unaids.org/en/resources/documents/2023/UNAIDS_FactSheet.

2. Hu, W.S. and S.H. Hughes, *HIV-1 reverse transcription.* Cold Spring Harb Perspect Med, 2012. **2**(10)DOI: 10.1101/cshperspect.a006882.

3. Perelson, A.S., *Modelling viral and immune system dynamics.* Nat Rev Immunol, 2002. **2**(1): p. 28-36. DOI: 10.1038/nri700.

4. Hemelaar, J., et al., *Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis.* Lancet Infect Dis, 2019. **19**(2): p. 143-155. DOI: 10.1016/S1473-3099(18)30647-9.

5. Laboratory, L.A.N. *HIV Sequence Database*. Available from: https://www.hiv.lanl.gov/.

6. Shaw, G.M. and E. Hunter, *HIV transmission.* Cold Spring Harb Perspect Med, 2012. **2**(11)DOI: 10.1101/cshperspect.a006965.

7. Ronen, K., A. Sharma, and J. Overbaugh, *HIV transmission biology: translation for HIV prevention.* AIDS, 2015. **29**(17): p. 2219-2227. DOI: 10.1097/QAD.0000000000000845.

8. Wu, L. and V.N. KewalRamani, *Dendritic-cell interactions with HIV: infection and viral dissemination.* Nat Rev Immunol, 2006. **6**(11): p. 859-68. DOI: 10.1038/nri1960.

9. Wolinsky, S.M., et al., *Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants.* Science, 1992. **255**(5048): p. 1134-7. DOI: 10.1126/science.1546316.

10. Wolfs, T.F., et al., *HIV-1 genomic RNA diversification following sexual and parenteral virus transmission.* Virology, 1992. **189**(1): p. 103-10. DOI: 10.1016/0042-6822(92)90685-i.

11. Zhu, T., et al., *Genotypic and phenotypic characterization of HIV-1 patients with primary infection.* Science, 1993. **261**(5125): p. 1179-81. DOI: 10.1126/science.8356453.

12. 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, 1993. **67**(6): p. 3345-56. DOI: 10.1128/JVI.67.6.3345-3356.1993.

13. 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, 2008. **105**(21): p. 7552-7. DOI: 10.1073/pnas.0802203105.

14. Haaland, R.E., et al., *Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1.* PLoS Pathog, 2009. **5**(1): p. e1000274. DOI: 10.1371/journal.ppat.1000274.

15. Wilen, C.B., J.C. Tilton, and R.W. Doms, *HIV: cell binding and entry.* Cold Spring Harb Perspect Med, 2012. **2**(8)DOI: 10.1101/cshperspect.a006866.

16. Long, E.M., et al., *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, 2002. **18**(8): p. 567-76. DOI: 10.1089/088922202753747914.

17. Balotta, C., et al., *Homozygous delta 32 deletion of the CCR-5 chemokine receptor gene in an HIV-1-infected patient.* AIDS, 1997. **11**(10): p. F67-71. DOI: 10.1097/00002030-199710000-00001.

18. Liu, R., et al., *Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection.* Cell, 1996. **86**(3): p. 367-77. DOI: 10.1016/s0092-8674(00)80110-5.

19. Moore, J.P., et al., *The CCR5 and CXCR4 coreceptors--central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection.* AIDS Res Hum Retroviruses, 2004. **20**(1): p. 111-26. DOI: 10.1089/088922204322749567.

20. 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, 2010. **3**(3): p. 280-90. DOI: 10.1038/mi.2010.2.

21. Carlson, J.M., et al., *HIV transmission. Selection bias at the heterosexual HIV-1 transmission bottleneck.* Science, 2014. **345**(6193): p. 1254031. DOI: 10.1126/science.1254031.

22. Sugrue, E., et al., *The apparent interferon resistance of transmitted HIV-1 is possibly a consequence of enhanced replicative fitness.* PLoS Pathog, 2022. **18**(11): p. e1010973. DOI: 10.1371/journal.ppat.1010973.

23. Fenton-May, A.E., et al., *Relative resistance of HIV-1 founder viruses to control by interferon-alpha.* Retrovirology, 2013. **10**: p. 146. DOI: 10.1186/1742-4690-10-146.

24. Iyer, S.S., et al., *Resistance to type 1 interferons is a major determinant of HIV-1 transmission fitness.* Proc Natl Acad Sci U S A, 2017. **114**(4): p. E590-E599. DOI: 10.1073/pnas.1620144114.

25. Parrish, N.F., et al., *Phenotypic properties of transmitted founder HIV-1.* Proc Natl Acad Sci U S A, 2013. **110**(17): p. 6626-33. DOI: 10.1073/pnas.1304288110.

26. Etemad, B., et al., *Characterization of HIV-1 envelopes in acutely and chronically infected injection drug users.* Retrovirology, 2014. **11**: p. 106. DOI: 10.1186/s12977-014-0106-8.

27. Lackner, A.A., M.M. Lederman, and B. Rodriguez, *HIV pathogenesis: the host.* Cold Spring Harb Perspect Med, 2012. **2**(9): p. a007005. DOI: 10.1101/cshperspect.a007005.

28. McMichael, A.J., et al., *The immune response during acute HIV-1 infection: clues for vaccine development.* Nat Rev Immunol, 2010. **10**(1): p. 11-23. DOI: 10.1038/nri2674.

29. Robb, M.L., et al., *Prospective Study of Acute HIV-1 Infection in Adults in East Africa and Thailand.* N Engl J Med, 2016. **374**(22): p. 2120-30. DOI: 10.1056/NEJMoa1508952.

30. Lavreys, L., et al., *Higher set point plasma viral load and more-severe acute HIV type 1 (HIV-1) illness predict mortality among high-risk HIV-1-infected African women.* Clin Infect Dis, 2006. **42**(9): p. 1333-9. DOI: 10.1086/503258.

31. Sterling, T.R., et al., *Initial plasma HIV-1 RNA levels and progression to AIDS in women and men.* N Engl J Med, 2001. **344**(10): p. 720-5. DOI: 10.1056/NEJM200103083441003.

32. Schacker, T.W., et al., *Biological and virologic characteristics of primary HIV infection.* Ann Intern Med, 1998. **128**(8): p. 613-20. DOI: 10.7326/0003-4819-128-8-199804150-00001.

33. Lefrere, J.J., et al., *The risk of disease progression is determined during the first year of human immunodeficiency virus type 1 infection.* J Infect Dis, 1998. **177**(6): p. 1541-8. DOI: 10.1086/515308.

34. Connor, R.I., et al., *Change in coreceptor use correlates with disease progression in HIV-1--infected individuals.* J Exp Med, 1997. **185**(4): p. 621-8. DOI: 10.1084/jem.185.4.621.

35. Esbjornsson, J., et al., *Frequent CXCR4 tropism of HIV-1 subtype A and CRF02_AG during late-stage disease--indication of an evolving epidemic in West Africa.* Retrovirology, 2010. **7**: p. 23. DOI: 10.1186/1742-4690-7-23.

36. Hwang, S.S., et al., *Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1.* Science, 1991. **253**(5015): p. 71-4. DOI: 10.1126/science.1905842.

37. Goonetilleke, N., et al., *The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection.* J Exp Med, 2009. **206**(6): p. 1253-72. DOI: 10.1084/jem.20090365.

38. Safrit, J.T., et al., *A region of the third variable loop of HIV-1 gp120 is recognized by HLA-B7-restricted CTLs from two acute seroconversion patients.* J Immunol, 1994. **153**(8): p. 3822-30.

39. 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, 2008. **198**(8): p. 1104-12. DOI: 10.1086/591623.

40. Penn, M.L., et al., *CXCR4 utilization is sufficient to trigger CD4+ T cell depletion in HIV-1-infected human lymphoid tissue.* Proc Natl Acad Sci U S A, 1999. **96**(2): p. 663-8. DOI: 10.1073/pnas.96.2.663.

41. Ivashkiv, L.B. and L.T. Donlin, *Regulation of type I interferon responses.* Nat Rev Immunol, 2014. **14**(1): p. 36-49. DOI: 10.1038/nri3581.

42. Schneider, W.M., M.D. Chevillotte, and C.M. Rice, *Interferon-stimulated genes: a complex web of host defenses.* Annu Rev Immunol, 2014. **32**: p. 513-45. DOI: 10.1146/annurev-immunol-032713-120231.

43. Michaelis, B. and J.A. Levy, *HIV replication can be blocked by recombinant human interferon beta.* AIDS, 1989. **3**(1): p. 27-31.

44. Yamamoto, J.K., et al., *Human alpha- and beta-interferon but not gamma- suppress the in vitro replication of LAV, HTLV-III, and ARV-2.* J Interferon Res, 1986. **6**(2): p. 143-52. DOI: 10.1089/jir.1986.6.143.

45. Pitha, P.M., *Multiple effects of interferon on HIV-1 replication.* J Interferon Res, 1991. **11**(6): p. 313-8. DOI: 10.1089/jir.1991.11.313.

46. Goujon, C. and M.H. Malim, *Characterization of the alpha interferon-induced postentry block to HIV-1 infection in primary human macrophages and T cells.* J Virol, 2010. **84**(18): p. 9254-66. DOI: 10.1128/JVI.00854-10.

47. Harper, M.S., et al., *Interferon-alpha Subtypes in an Ex Vivo Model of Acute HIV-1 Infection: Expression, Potency and Effector Mechanisms.* PLoS Pathog, 2015. **11**(11): p. e1005254. DOI: 10.1371/journal.ppat.1005254.

48. Asmuth, D.M., et al., *Safety, tolerability, and mechanisms of antiretroviral activity of pegylated interferon Alfa-2a in HIV-1-monoinfected participants: a phase II clinical trial.* J Infect Dis, 2010. **201**(11): p. 1686-96. DOI: 10.1086/652420.

49. Azzoni, L., et al., *Pegylated Interferon alfa-2a monotherapy results in suppression of HIV type 1 replication and decreased cell-associated HIV DNA integration.* J Infect Dis, 2013. **207**(2): p. 213-22. DOI: 10.1093/infdis/jis663.

50. El-Diwany, R., et al., *CMPK2 and BCL-G are associated with type 1 interferon-induced HIV restriction in humans.* Sci Adv, 2018. **4**(8): p. eaat0843. DOI: 10.1126/sciadv.aat0843.

51. Hubbard, J.J., et al., *Host gene expression changes correlating with anti-HIV-1 effects in human subjects after treatment with peginterferon Alfa-2a.* J Infect Dis, 2012. **205**(9): p. 1443-7. DOI: 10.1093/infdis/jis211.

52. Stacey, A.R., et al., *Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections.* J Virol, 2009. **83**(8): p. 3719-33. DOI: 10.1128/JVI.01844-08.

53. von Sydow, M., et al., *Interferon-alpha and tumor necrosis factor-alpha in serum of patients in various stages of HIV-1 infection.* AIDS Res Hum Retroviruses, 1991. **7**(4): p. 375-80. DOI: 10.1089/aid.1991.7.375.

54. Lepelley, A., et al., *Innate sensing of HIV-infected cells.* PLoS Pathog, 2011. **7**(2): p. e1001284. DOI: 10.1371/journal.ppat.1001284.

55. Altfeld, M. and M. Gale, Jr., *Innate immunity against HIV-1 infection.* Nat Immunol, 2015. **16**(6): p. 554-62. DOI: 10.1038/ni.3157.

56. Richardson, R.B., et al., *A CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus replication.* Nat Microbiol, 2018. **3**(11): p. 1214-1223. DOI: 10.1038/s41564-018-0244-1.

57. Colomer-Lluch, M., et al., *Restriction Factors: From Intrinsic Viral Restriction to Shaping Cellular Immunity Against HIV-1.* Front Immunol, 2018. **9**: p. 2876. DOI: 10.3389/fimmu.2018.02876.

58. Doyle, T., C. Goujon, and M.H. Malim, *HIV-1 and interferons: who's interfering with whom?* Nat Rev Microbiol, 2015. **13**(7): p. 403-13. DOI: 10.1038/nrmicro3449.

59. Malim, M.H. and P.D. Bieniasz, *HIV Restriction Factors and Mechanisms of Evasion.* Cold Spring Harb Perspect Med, 2012. **2**(5): p. a006940. DOI: 10.1101/cshperspect.a006940.

60. Gelinas, J.F., D.R. Gill, and S.C. Hyde, *Multiple Inhibitory Factors Act in the Late Phase of HIV-1 Replication: a Systematic Review of the Literature.* Microbiol Mol Biol Rev, 2018. **82**(1)DOI: 10.1128/MMBR.00051-17.

61. Stremlau, M., et al., *The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys.* Nature, 2004. **427**(6977): p. 848-53. DOI: 10.1038/nature02343.

62. Sheehy, A.M., et al., *Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein.* Nature, 2002. **418**(6898): p. 646-50. DOI: 10.1038/nature00939.

63. Kane, M., et al., *MX2 is an interferon-induced inhibitor of HIV-1 infection.* Nature, 2013. **502**(7472): p. 563-6. DOI: 10.1038/nature12653.

64. Goujon, C., et al., *Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection.* Nature, 2013. **502**(7472): p. 559-62. DOI: 10.1038/nature12542.

65. Liu, Z., et al., *The interferon-inducible MxB protein inhibits HIV-1 infection.* Cell Host Microbe, 2013. **14**(4): p. 398-410. DOI: 10.1016/j.chom.2013.08.015.

66. Matreyek, K.A., et al., *Host and viral determinants for MxB restriction of HIV-1 infection.* Retrovirology, 2014. **11**: p. 90. DOI: 10.1186/s12977-014-0090-z.

67. Lu, J., et al., *The IFITM proteins inhibit HIV-1 infection.* J Virol, 2011. **85**(5): p. 2126-37. DOI: 10.1128/JVI.01531-10.

68. Jain, P., et al., *Large-Scale Arrayed Analysis of Protein Degradation Reveals Cellular Targets for HIV-1 Vpu.* Cell Rep, 2018. **22**(9): p. 2493-2503. DOI: 10.1016/j.celrep.2018.01.091.

69. OhAinle, M., et al., *A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV.* Elife, 2018. **7**DOI: 10.7554/eLife.39823.

70. Zhang, Y., J. Lu, and X. Liu, *MARCH2 is upregulated in HIV-1 infection and inhibits HIV-1 production through envelope protein translocation or degradation.* Virology, 2018. **518**: p. 293-300. DOI: 10.1016/j.virol.2018.02.003.

71. Tada, T., et al., *MARCH8 inhibits HIV-1 infection by reducing virion incorporation of envelope glycoproteins.* Nat Med, 2015. **21**(12): p. 1502-7. DOI: 10.1038/nm.3956.

72. Liu, S.Y., et al., *Interferon-inducible cholesterol-25-hydroxylase broadly inhibits viral entry by production of 25-hydroxycholesterol.* Immunity, 2013. **38**(1): p. 92-105. DOI: 10.1016/j.immuni.2012.11.005.

73. Krapp, C., et al., *Guanylate Binding Protein (GBP) 5 Is an Interferon-Inducible Inhibitor of HIV-1 Infectivity.* Cell Host Microbe, 2016. **19**(4): p. 504-14. DOI: 10.1016/j.chom.2016.02.019.

74. Hotter, D., et al., *IFI16 Targets the Transcription Factor Sp1 to Suppress HIV-1 Transcription and Latency Reactivation.* Cell Host Microbe, 2019. **25**(6): p. 858-872 e13. DOI: 10.1016/j.chom.2019.05.002.

75. Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.* Nature, 2008. **451**(7177): p. 425-30. DOI: 10.1038/nature06553.

76. Baldauf, H.M., et al., *SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells.* Nat Med, 2012. **18**(11): p. 1682-7. DOI: 10.1038/nm.2964.

77. Laguette, N., et al., *SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx.* Nature, 2011. **474**(7353): p. 654-7. DOI: 10.1038/nature10117.

78. St Gelais, C., et al., *SAMHD1 restricts HIV-1 infection in dendritic cells (DCs) by dNTP depletion, but its expression in DCs and primary CD4+ T-lymphocytes cannot be upregulated by interferons.* Retrovirology, 2012. **9**: p. 105. DOI: 10.1186/1742-4690-9-105.

79. Cribier, A., et al., *Phosphorylation of SAMHD1 by cyclin A2/CDK1 regulates its restriction activity toward HIV-1.* Cell Rep, 2013. **3**(4): p. 1036-43. DOI: 10.1016/j.celrep.2013.03.017.

80. Basmaciogullari, S. and M. Pizzato, *The activity of Nef on HIV-1 infectivity.* Front Microbiol, 2014. **5**: p. 232. DOI: 10.3389/fmicb.2014.00232.

81. Usami, Y., Y. Wu, and H.G. Gottlinger, *SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef.* Nature, 2015. **526**(7572): p. 218-23. DOI: 10.1038/nature15400.

82. Rosa, A., et al., *HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation.* Nature, 2015. **526**(7572): p. 212-7. DOI: 10.1038/nature15399.

83. Leonhardt, S.A., et al., *Antiviral HIV-1 SERINC restriction factors disrupt virus membrane asymmetry.* Nat Commun, 2023. **14**(1): p. 4368. DOI: 10.1038/s41467-023-39262-2.

84. Kabat, D., et al., *Differences in CD4 dependence for infectivity of laboratory-adapted and primary patient isolates of human immunodeficiency virus type 1.* J Virol, 1994. **68**(4): p. 2570-7. DOI: 10.1128/JVI.68.4.2570-2577.1994.

85. Kane, M., et al., *Identification of Interferon-Stimulated Genes with Antiretroviral Activity.* Cell Host Microbe, 2016. **20**(3): p. 392-405. DOI: 10.1016/j.chom.2016.08.005.

86. Schoggins, J.W., et al., *A diverse range of gene products are effectors of the type I interferon antiviral response.* Nature, 2011. **472**(7344): p. 481-5. DOI: 10.1038/nature09907.

87. Liu, L., et al., *A whole genome screen for HIV restriction factors.* Retrovirology, 2011. **8**: p. 94. DOI: 10.1186/1742-4690-8-94.

88. Hultquist, J.F., et al., *CRISPR-Cas9 genome engineering of primary CD4(+) T cells for the interrogation of HIV-host factor interactions.* Nat Protoc, 2019. **14**(1): p. 1-27. DOI: 10.1038/s41596-018-0069-7.

89. Hiatt, J., et al., *A functional map of HIV-host interactions in primary human T cells.* Nat Commun, 2022. **13**(1): p. 1752. DOI: 10.1038/s41467-022-29346-w.

90. Ting, P.Y., et al., *Guide Swap enables genome-scale pooled CRISPR-Cas9 screening in human primary cells.* Nat Methods, 2018. **15**(11): p. 941-946. DOI: 10.1038/s41592-018-0149-1.

91. Humes, D., S. Rainwater, and J. Overbaugh, *The TOP vector: a new high-titer lentiviral construct for delivery of sgRNAs and transgenes to primary T cells.* Mol Ther Methods Clin Dev, 2021. **20**: p. 30-38. DOI: 10.1016/j.omtm.2020.10.020.

92. Duggal, N.K. and M. Emerman, *Evolutionary conflicts between viruses and restriction factors shape immunity.* Nat Rev Immunol, 2012. **12**(10): p. 687-95. DOI: 10.1038/nri3295.

93. Kirchhoff, F., *Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses.* Cell Host Microbe, 2010. **8**(1): p. 55-67. DOI: 10.1016/j.chom.2010.06.004.

94. Sharp, P.M. and B.H. Hahn, *Origins of HIV and the AIDS pandemic.* Cold Spring Harb Perspect Med, 2011. **1**(1): p. a006841. DOI: 10.1101/cshperspect.a006841.

95. Li, M., et al., *Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11.* Nature, 2012. **491**(7422): p. 125-8. DOI: 10.1038/nature11433.

96. Compton, A.A., et al., *IFITM proteins incorporated into HIV-1 virions impair viral fusion and spread.* Cell Host Microbe, 2014. **16**(6): p. 736-47. DOI: 10.1016/j.chom.2014.11.001.

97. Lodermeyer, V., et al., *90K, an interferon-stimulated gene product, reduces the infectivity of HIV-1.* Retrovirology, 2013. **10**: p. 111. DOI: 10.1186/1742-4690-10-111.

98. Barr, S.D., J.R. Smiley, and F.D. Bushman, *The interferon response inhibits HIV particle production by induction of TRIM22.* PLoS Pathog, 2008. **4**(2): p. e1000007. DOI: 10.1371/journal.ppat.1000007.

99. Betancor, G., et al., *Author Correction: MX2-mediated innate immunity against HIV-1 is regulated by serine phosphorylation.* Nat Microbiol, 2021. **6**(9): p. 1211. DOI: 10.1038/s41564-021-00960-6.

100. Pickering, S., et al., *Preservation of tetherin and CD4 counter-activities in circulating Vpu alleles despite extensive sequence variation within HIV-1 infected individuals.* PLoS Pathog, 2014. **10**(1): p. e1003895. DOI: 10.1371/journal.ppat.1003895.

101. Opp, S., et al., *MxB Is Not Responsible for the Blocking of HIV-1 Infection Observed in Alpha Interferon-Treated Cells.* J Virol, 2015. **90**(6): p. 3056-64. DOI: 10.1128/JVI.03146-15.

102. Lim, E.S., et al., *The function and evolution of the restriction factor Viperin in primates was not driven by lentiviruses.* Retrovirology, 2012. **9**: p. 55. DOI: 10.1186/1742-4690-9-55.

103. Liu, Z., et al., *The highly polymorphic cyclophilin A-binding loop in HIV-1 capsid modulates viral resistance to MxB.* Retrovirology, 2015. **12**: p. 1. DOI: 10.1186/s12977-014-0129-1.

104. Poss, M. and J. Overbaugh, *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, 1999. **73**(7): p. 5255-64. DOI: 10.1128/JVI.73.7.5255-5264.1999.

105. Haddox, H.K., et al., *Mapping mutational effects along the evolutionary landscape of HIV envelope.* Elife, 2018. **7**DOI: 10.7554/eLife.34420.

106. Peden, K., M. Emerman, and L. Montagnier, *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, 1991. **185**(2): p. 661-72. DOI: 10.1016/0042-6822(91)90537-l.

107. Bolen, C.R., et al., *Dynamic expression profiling of type I and type III interferon-stimulated hepatocytes reveals a stable hierarchy of gene expression.* Hepatology, 2014. **59**(4): p. 1262-72. DOI: 10.1002/hep.26657.

108. Kim, K., et al., *Cyclophilin A protects HIV-1 from restriction by human TRIM5alpha.* Nat Microbiol, 2019. **4**(12): p. 2044-2051. DOI: 10.1038/s41564-019-0592-5.

109. Selyutina, A., et al., *Cyclophilin A Prevents HIV-1 Restriction in Lymphocytes by Blocking Human TRIM5alpha Binding to the Viral Core.* Cell Rep, 2020. **30**(11): p. 3766-3777 e6. DOI: 10.1016/j.celrep.2020.02.100.

110. Foster, T.L., et al., *Resistance of Transmitted Founder HIV-1 to IFITM-Mediated Restriction.* Cell Host Microbe, 2016. **20**(4): p. 429-442. DOI: 10.1016/j.chom.2016.08.006.

111. Chen, B., et al., *Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system.* Cell, 2013. **155**(7): p. 1479-91. DOI: 10.1016/j.cell.2013.12.001.

112. Seki, A. and S. Rutz, *Optimized RNP transfection for highly efficient CRISPR/Cas9-mediated gene knockout in primary T cells.* J Exp Med, 2018. **215**(3): p. 985-997. DOI: 10.1084/jem.20171626.

113. Wang, B., et al., *Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute.* Nat Protoc, 2019. **14**(3): p. 756-780. DOI: 10.1038/s41596-018-0113-7.

114. Ranjan, P., et al., *NLRC5 interacts with RIG-I to induce a robust antiviral response against influenza virus infection.* Eur J Immunol, 2015. **45**(3): p. 758-72. DOI: 10.1002/eji.201344412.

115. Albanese, M., et al., *Rapid, efficient and activation-neutral gene editing of polyclonal primary human resting CD4(+) T cells allows complex functional analyses.* Nat Methods, 2022. **19**(1): p. 81-89. DOI: 10.1038/s41592-021-01328-8.

116. Luo, Y., et al., *HIV-host interactome revealed directly from infected cells.* Nat Microbiol, 2016. **1**(7): p. 16068. DOI: 10.1038/nmicrobiol.2016.68.

117. Jager, S., et al., *Global landscape of HIV-human protein complexes.* Nature, 2011. **481**(7381): p. 365-70. DOI: 10.1038/nature10719.

118. Weihofen, A., et al., *Identification of signal peptide peptidase, a presenilin-type aspartic protease.* Science, 2002. **296**(5576): p. 2215-8. DOI: 10.1126/science.1070925.

119. Schrul, B., et al., *Signal peptide peptidase (SPP) assembles with substrates and misfolded membrane proteins into distinct oligomeric complexes.* Biochem J, 2010. **427**(3): p. 523-34. DOI: 10.1042/BJ20091005.

120. McLauchlan, J., et al., *Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets.* EMBO J, 2002. **21**(15): p. 3980-8. DOI: 10.1093/emboj/cdf414.

121. Randall, G., et al., *Cellular cofactors affecting hepatitis C virus infection and replication.* Proc Natl Acad Sci U S A, 2007. **104**(31): p. 12884-9. DOI: 10.1073/pnas.0704894104.

122. Wang, S., U. Jaggi, and H. Ghiasi, *Knockout of signal peptide peptidase in the eye reduces HSV-1 replication and eye disease in ocularly infected mice.* PLoS Pathog, 2022. **18**(10): p. e1010898. DOI: 10.1371/journal.ppat.1010898.

123. Allen, S.J., et al., *Binding of HSV-1 glycoprotein K (gK) to signal peptide peptidase (SPP) is required for virus infectivity.* PLoS One, 2014. **9**(1): p. e85360. DOI: 10.1371/journal.pone.0085360.

124. Marceau, C.D., et al., *Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens.* Nature, 2016. **535**(7610): p. 159-63. DOI: 10.1038/nature18631.

125. Zanini, F., et al., *Single-cell transcriptional dynamics of flavivirus infection.* Elife, 2018. **7**DOI: 10.7554/eLife.32942.

126. Matsui, M., J.H. Fowler, and L.L. Walling, *Leucine aminopeptidases: diversity in structure and function.* Biol Chem, 2006. **387**(12): p. 1535-44. DOI: 10.1515/BC.2006.191.

127. Chen, X., et al., *IGF binding protein 2 is a cell-autonomous factor supporting survival and migration of acute leukemia cells.* J Hematol Oncol, 2013. **6**(1): p. 72. DOI: 10.1186/1756-8722-6-72.

128. Sauter, D., et al., *Differential regulation of NF-kappaB-mediated proviral and antiviral host gene expression by primate lentiviral Nef and Vpu proteins.* Cell Rep, 2015. **10**(4): p. 586-99. DOI: 10.1016/j.celrep.2014.12.047.

129. Jakobsen, M.R., et al., *IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication.* Proc Natl Acad Sci U S A, 2013. **110**(48): p. E4571-80. DOI: 10.1073/pnas.1311669110.

130. Jonsson, K.L., et al., *IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP.* Nat Commun, 2017. **8**: p. 14391. DOI: 10.1038/ncomms14391.

131. Monroe, K.M., et al., *IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV.* Science, 2014. **343**(6169): p. 428-32. DOI: 10.1126/science.1243640.

132. Zhang, D. and D.E. Zhang, *Interferon-stimulated gene 15 and the protein ISGylation system.* J Interferon Cytokine Res, 2011. **31**(1): p. 119-30. DOI: 10.1089/jir.2010.0110.

133. Pincetic, A., et al., *The interferon-induced gene ISG15 blocks retrovirus release from cells late in the budding process.* J Virol, 2010. **84**(9): p. 4725-36. DOI: 10.1128/JVI.02478-09.

134. Wu, W.L., et al., *Delta20 IFITM2 differentially restricts X4 and R5 HIV-1.* Proc Natl Acad Sci U S A, 2017. **114**(27): p. 7112-7117. DOI: 10.1073/pnas.1619640114.

135. Bulli, L., et al., *Complex Interplay between HIV-1 Capsid and MX2-Independent Alpha Interferon-Induced Antiviral Factors.* J Virol, 2016. **90**(16): p. 7469-7480. DOI: 10.1128/JVI.00458-16.

136. Busnadiego, I., et al., *Host and viral determinants of Mx2 antiretroviral activity.* J Virol, 2014. **88**(14): p. 7738-52. DOI: 10.1128/JVI.00214-14.

137. Kane, M., et al., *Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2.* Elife, 2018. **7**DOI: 10.7554/eLife.35738.

138. Troyano-Hernaez, P., R. Reinosa, and A. Holguin, *HIV Capsid Protein Genetic Diversity Across HIV-1 Variants and Impact on New Capsid-Inhibitor Lenacapavir.* Front Microbiol, 2022. **13**: p. 854974. DOI: 10.3389/fmicb.2022.854974.

139. Hultquist, J.F., et al., *A Cas9 Ribonucleoprotein Platform for Functional Genetic Studies of HIV-Host Interactions in Primary Human T Cells.* Cell Rep, 2016. **17**(5): p. 1438-1452. DOI: 10.1016/j.celrep.2016.09.080.

140. Wu, X., et al., *Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant.* J Virol, 2006. **80**(2): p. 835-44. DOI: 10.1128/JVI.80.2.835-844.2006.

141. Gartner, S., et al., *The role of mononuclear phagocytes in HTLV-III/LAV infection.* Science, 1986. **233**(4760): p. 215-9. DOI: 10.1126/science.3014648.

142. 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, 2012. **7**(12): p. e50859. DOI: 10.1371/journal.pone.0050859.

143. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40. DOI: 10.1093/bioinformatics/btp616.

144. Zhang, Y., G. Parmigiani, and W.E. Johnson, *ComBat-seq: batch effect adjustment for RNA-seq count data.* NAR Genom Bioinform, 2020. **2**(3): p. lqaa078. DOI: 10.1093/nargab/lqaa078.

145. Labun, K., et al., *CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing.* Nucleic Acids Res, 2019. **47**(W1): p. W171-W174. DOI: 10.1093/nar/gkz365.

146. Meier, J.A., F. Zhang, and N.E. Sanjana, *GUIDES: sgRNA design for loss-of-function screens.* Nat Methods, 2017. **14**(9): p. 831-832. DOI: 10.1038/nmeth.4423.

147. Kassambara, A. *ggpubr: 'ggplot2' Based Publication Ready Plots*. 2023; 0.6.0:[Available from: https://rpkgs.datanovia.com/ggpubr/.

148. Hartner, J.C., et al., *ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling.* Nat Immunol, 2009. **10**(1): p. 109-15. DOI: 10.1038/ni.1680.

149. Li, Y., et al., *Ribonuclease L mediates the cell-lethal phenotype of double-stranded RNA editing enzyme ADAR1 deficiency in a human cell line.* Elife, 2017. **6**DOI: 10.7554/eLife.25687.

150. Samuel, C.E., *ADARs: viruses and innate immunity.* Curr Top Microbiol Immunol, 2012. **353**: p. 163-95. DOI: 10.1007/82_2011_148.

151. Cuadrado, E., et al., *ADAR1 Facilitates HIV-1 Replication in Primary CD4+ T Cells.* PLoS One, 2015. **10**(12): p. e0143613. DOI: 10.1371/journal.pone.0143613.

152. Phuphuakrat, A., et al., *Double-stranded RNA adenosine deaminases enhance expression of human immunodeficiency virus type 1 proteins.* J Virol, 2008. **82**(21): p. 10864-72. DOI: 10.1128/JVI.00238-08.

153. Clerzius, G., et al., *ADAR1 interacts with PKR during human immunodeficiency virus infection of lymphocytes and contributes to viral replication.* J Virol, 2009. **83**(19): p. 10119-28. DOI: 10.1128/JVI.02457-08.

154. Doria, M., et al., *Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection.* Nucleic Acids Res, 2009. **37**(17): p. 5848-58. DOI: 10.1093/nar/gkp604.

155. Clerzius, G., et al., *The PKR activator, PACT, becomes a PKR inhibitor during HIV-1 replication.* Retrovirology, 2013. **10**: p. 96. DOI: 10.1186/1742-4690-10-96.

156. Biswas, N., et al., *ADAR1 is a novel multi targeted anti-HIV-1 cellular protein.* Virology, 2012. **422**(2): p. 265-77. DOI: 10.1016/j.virol.2011.10.024.

157. Weiden, M.D., et al., *Adenosine deaminase acting on RNA-1 (ADAR1) inhibits HIV-1 replication in human alveolar macrophages.* PLoS One, 2014. **9**(10): p. e108476. DOI: 10.1371/journal.pone.0108476.

158. Orecchini, E., et al., *The ADAR1 editing enzyme is encapsidated into HIV-1 virions.* Virology, 2015. **485**: p. 475-80. DOI: 10.1016/j.virol.2015.07.027.

159. Lv, J., et al., *Disruption of CISH promotes the antitumor activity of human T cells and decreases PD-1 expression levels.* Mol Ther Oncolytics, 2023. **28**: p. 46-58. DOI: 10.1016/j.omto.2022.12.003.

160. Khor, C.C., et al., *CISH and susceptibility to infectious diseases.* N Engl J Med, 2010. **362**(22): p. 2092-101. DOI: 10.1056/NEJMoa0905606.

161. Itell, H.L., D. Humes, and J. Overbaugh, *Several cell-intrinsic effectors drive type I interferon-mediated restriction of HIV-1 in primary CD4(+) T cells.* Cell Rep, 2023. **42**(6): p. 112556. DOI: 10.1016/j.celrep.2023.112556.

162. S, P.H., M. Krumbiegel, and F. Kirchhoff, *Coreceptor usage of BOB/GPR15 and Bonzo/STRL33 by primary isolates of human immunodeficiency virus type 1.* J Gen Virol, 1999. **80 (Pt 5)**: p. 1241-1251. DOI: 10.1099/0022-1317-80-5-1241.

163. 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, 1996. **85**(7): p. 1149-58. DOI: 10.1016/s0092-8674(00)81314-8.

164. Quelhas, D., et al., *SLC35A2-CDG: Novel variant and review.* Mol Genet Metab Rep, 2021. **26**: p. 100717. DOI: 10.1016/j.ymgmr.2021.100717.

165. in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2022: Cold Spring Harbor (NY).

166. Ng, B.G., et al., *SLC35A2-CDG: Functional characterization, expanded molecular, clinical, and biochemical phenotypes of 30 unreported Individuals.* Hum Mutat, 2019. **40**(7): p. 908-925. DOI: 10.1002/humu.23731.

167. Bojar, D., et al., *A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities.* ACS Chem Biol, 2022. **17**(11): p. 2993-3012. DOI: 10.1021/acschembio.1c00689.

168. O'Doherty, U., W.J. Swiggard, and M.H. Malim, *Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding.* J Virol, 2000. **74**(21): p. 10074-80. DOI: 10.1128/jvi.74.21.10074-10080.2000.

169. Woodward Davis, A.S., et al., *The human memory T cell compartment changes across tissues of the female reproductive tract.* Mucosal Immunol, 2021. **14**(4): p. 862-872. DOI: 10.1038/s41385-021-00406-6.

170. Ma, T., et al., *Single-cell glycomics analysis by CyTOF-Lec reveals glycan features defining cells differentially susceptible to HIV.* Elife, 2022. **11**DOI: 10.7554/eLife.78870.

171. Carette, J.E., et al., *Haploid genetic screens in human cells identify host factors used by pathogens.* Science, 2009. **326**(5957): p. 1231-5. DOI: 10.1126/science.1178955.

172. Moskovskich, A., et al., *The transporters SLC35A1 and SLC30A1 play opposite roles in cell survival upon VSV virus infection.* Sci Rep, 2019. **9**(1): p. 10471. DOI: 10.1038/s41598-019-46952-9.

173. Pedersen, H.L., et al., *Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research.* J Biol Chem, 2012. **287**(47): p. 39429-38. DOI: 10.1074/jbc.M112.396598.

174. Chabot, D.J., et al., *N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates.* J Virol, 2000. **74**(9): p. 4404-13. DOI: 10.1128/jvi.74.9.4404-4413.2000.

175. 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, 2001. **194**(11): p. 1661-73. DOI: 10.1084/jem.194.11.1661.

176. Farzan, M., et al., *Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry.* Cell, 1999. **96**(5): p. 667-76. DOI: 10.1016/s0092-8674(00)80577-2.

177. 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, 1999. **274**(14): p. 9617-26. DOI: 10.1074/jbc.274.14.9617.

178. Spillings, B.L., et al., *Host glycocalyx captures HIV proximal to the cell surface via oligomannose-GlcNAc glycan-glycan interactions to support viral entry.* Cell Rep, 2022. **38**(5): p. 110296. DOI: 10.1016/j.celrep.2022.110296.

179. Mereiter, S., et al., *Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading?* Cancer Cell, 2019. **36**(1): p. 6-16. DOI: 10.1016/j.ccell.2019.06.006.

180. Ju, T. and R.D. Cummings, *A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16613-8. DOI: 10.1073/pnas.262438199.

181. Matia, A., et al., *Identification of beta2 microglobulin, the product of B2M gene, as a Host Factor for Vaccinia Virus Infection by Genome-Wide CRISPR genetic screens.* PLoS Pathog, 2022. **18**(12): p. e1010800. DOI: 10.1371/journal.ppat.1010800.

182. Sun, Q., et al., *Synaptogyrin-2 Promotes Replication of a Novel Tick-borne Bunyavirus through Interacting with Viral Nonstructural Protein NSs.* J Biol Chem, 2016. **291**(31): p. 16138-49. DOI: 10.1074/jbc.M116.715599.

183. Cheng-Mayer, C., et al., *Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.* J Virol, 1990. **64**(9): p. 4390-8. DOI: 10.1128/JVI.64.9.4390-4398.1990.

184. Adachi, A., et al., *Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone.* J Virol, 1986. **59**(2): p. 284-91. DOI: 10.1128/JVI.59.2.284-291.1986.

185. 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, 1992. **66**(12): p. 7517-21. DOI: 10.1128/JVI.66.12.7517-7521.1992.

186. Provine, N.M., et al., *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, 2012. **427**(1): p. 25-33. DOI: 10.1016/j.virol.2012.02.001.

187. Montoya, V.R., et al., *A Virus-Packageable CRISPR System Identifies Host Dependency Factors Co-Opted by Multiple HIV-1 Strains.* mBio, 2023. **14**(1): p. e0000923. DOI: 10.1128/mbio.00009-23.

188. Hsieh, E., et al., *A modular CRISPR screen identifies individual and combination pathways contributing to HIV-1 latency.* PLoS Pathog, 2023. **19**(1): p. e1011101. DOI: 10.1371/journal.ppat.1011101.

189. Cavrois, M., C. De Noronha, and W.C. Greene, *A sensitive and specific enzyme-based assay detecting HIV-1 virion fusion in primary T lymphocytes.* Nat Biotechnol, 2002. **20**(11): p. 1151-4. DOI: 10.1038/nbt745.

Appendix A:

SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Supplementary Tables

Supplementary Table 2.1. Bulk RNA-seq of IFN-treated CD4$^+$ T cells.
URL accession: https://www.cell.com/cms/10.1016/j.celrep.2023.112556/attachment/f7b55ec7-80c0-410d-ba9d-9e5223a3f6b6/mmc2.xlsx

Supplementary Table 2.2. CD4-ISG sgRNA library.
URL accession: https://www.cell.com/cms/10.1016/j.celrep.2023.112556/attachment/3fa60848-845e-49fb-a711-cad487ecb5b5/mmc3.xlsx

Supplementary Table 2.3. CD4-ISG HIV-CRISPR screen results.
URL accession: https://www.cell.com/cms/10.1016/j.celrep.2023.112556/attachment/0c62204e-a3d5-496a-afd2-c9601b7f3b63/mmc4.xlsx

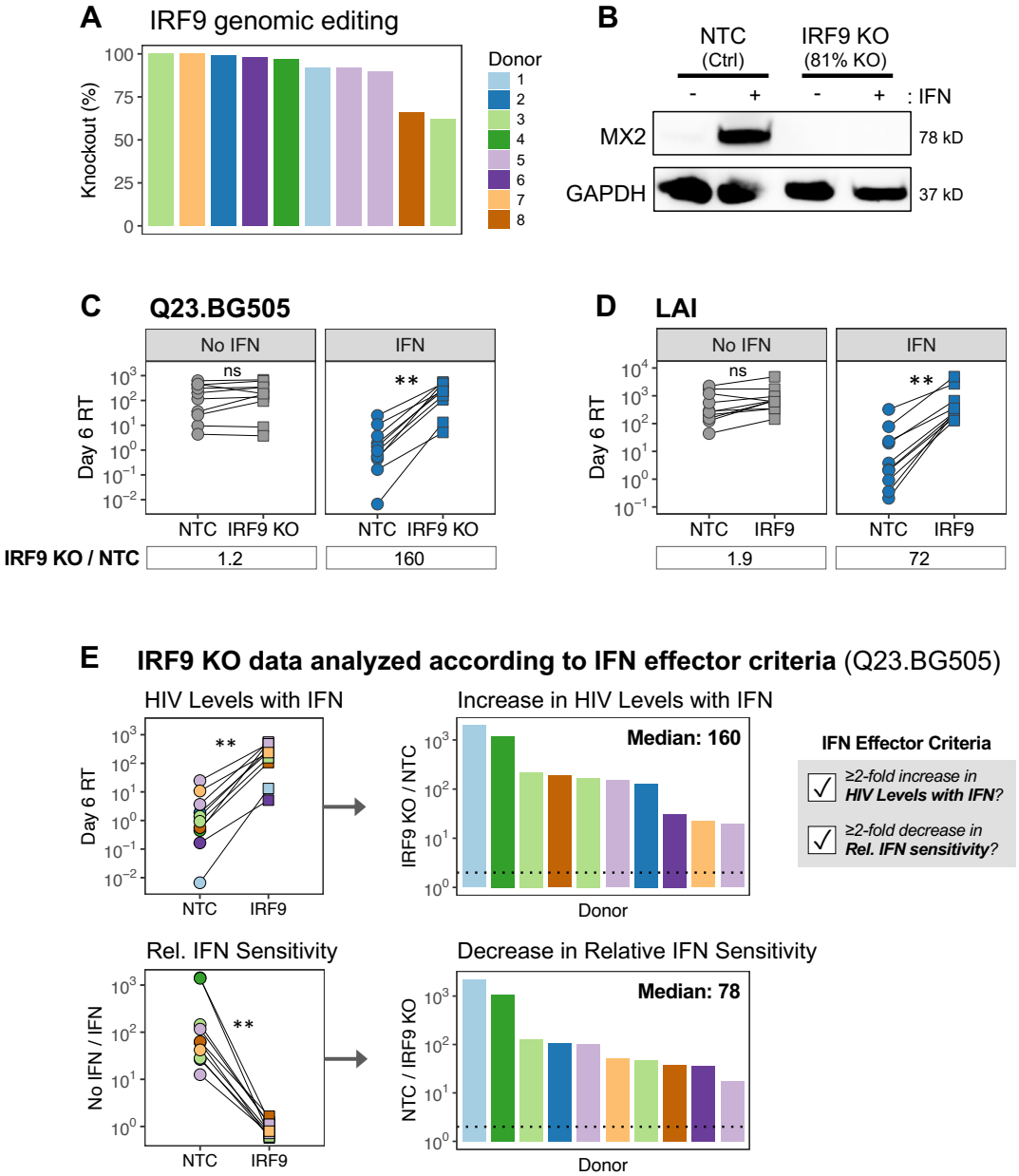
Supplementary Table 2.4. Intersect of non-control genes scoring above background in CD4-ISG screens.
URL accession: https://www.cell.com/cms/10.1016/j.celrep.2023.112556/attachment/9bf391b8-89dd-424d-abfc-4a77002ea5d8/mmc5.xlsx

Supplementary Table 2.5. Oligonucleotides and thermocycler programs.
URL accession: https://www.cell.com/cms/10.1016/j.celrep.2023.112556/attachment/ea3c426c-8a1e-4932-9832-abcd32ab62af/mmc6.xlsx

Supplementary Table 2.6. Guide RNAs used in PIKA screen validation experiments.

Gene Target	Guide ID	Guide Sequence
ADAR	ADAR_g1	UUAGCACAGCAUUUAUAUCU
ADAR	ADAR_g2	UGACUCCUCUGCCCUGAAUU
ADAR	ADAR_g3	UUUAGACAUGGCCGAGAUCA
BATF	BATF_g1	GCUGUCGGAGCUGUGAGGCA
BATF	BATF_g2	GGACUCUACCUGUUUGCCAG
BATF	BATF_g3	UUGUCCUGCCCAGGGAGCUG
CISH	CISH_g1	UCCUCUGGGUCCAGCACCUU
CISH	CISH_g2	CAAGGGCUGCAUGACUGGCU
CISH	CISH_g3	UUCUAGACCUCGUCCUUUGC
PARS2	PARS2_g1	GCCGCUGGCUCUUACAGGUC
PARS2	PARS2_g2	GGAACACACGAGACAGCAGC
PARS2	PARS2_g3	AAGGAACAUACCCAGAGAGC
TAPBP	TAPBP_g1	GCUCAGGGUCGAGGUCCGGC
TAPBP	TAPBP_g2	AAGGGCCUGGCCAAGAGACC
TAPBP	TAPBP_g3	CUGCUGAGACGGCGGUCGCC

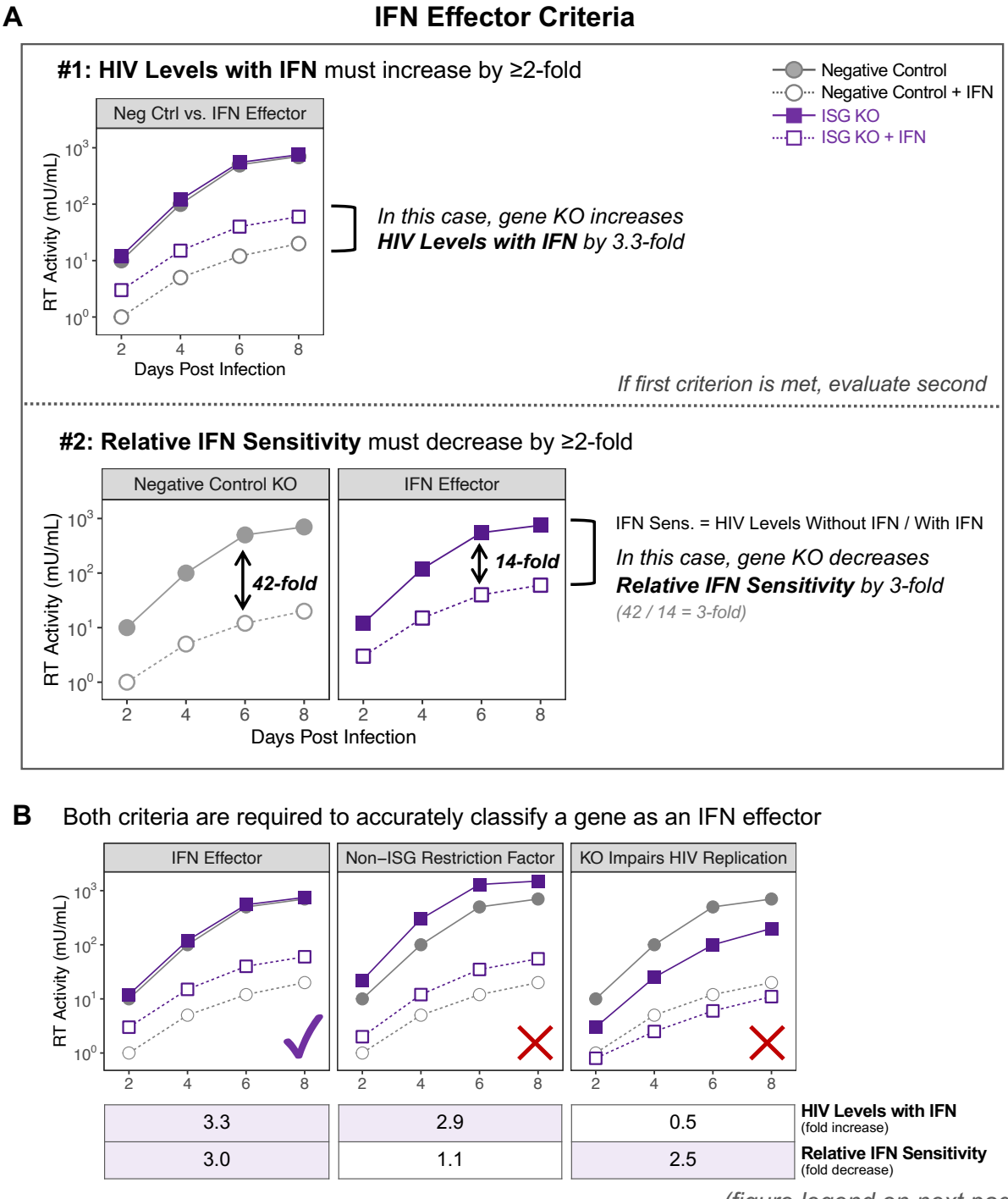
Supplementary Figures



Supplementary Figure 2.1. IRF9 editing ablates the antiviral ISG response in primary CD4⁺ T cells.
(A) IRF9 RNP-delivered CD4⁺ T cells were collected on the day of infection and assessed for genomic editing by Sanger sequencing. Synthego ICE analysis was applied to sequencing results to determine percent knockout. (B) Western blot with lysates from NTC and IRF9-edited CD4⁺ T cells from one donor to evaluate protein levels of the highly induced ISG MX2. (C, D) Paired data from Figure 2.1B and 2.1C, reformatted to compare infection results between NTC and IRF9-edited cells with and without IFN treatment. (E) HIV-1 infection levels according to the metrics used to evaluate whether ISGs are IFN effectors.

(figure legend continued on next page)

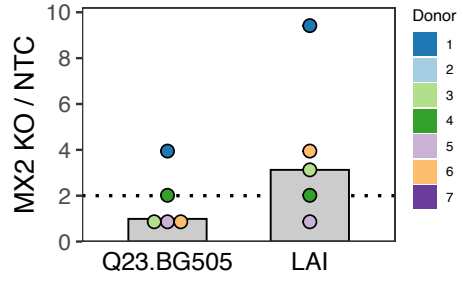
Left, paired data from panel B reformatted to compare "HIV Levels with IFN" and “Relative IFN Sensitivity” between NTC and IRF9-edited cells. Right, fold impact of IRF9 KO on these metrics, with bars depicting independent experiments. The median fold difference is denoted in the top right corner. The dotted line indicates a fold difference of two. **p<0.002, paired Wilcoxon rank test. Non-significant (ns) indicates p>0.05.



(figure legend on next page)

Supplementary Figure 2.2. Schematic of IFN effector criteria.
(A) Hypothetical HIV-1 infection curves illustrate how ISG KO data was compared to negative control results to categorize genes as IFN effectors. RT activity measured on 6 dpi are used to calculate fold changes compared to an intra-assay negative control. (B) Example of an IFN effector and two types of genes that might impact these criteria but would not be classified as IFN effectors.

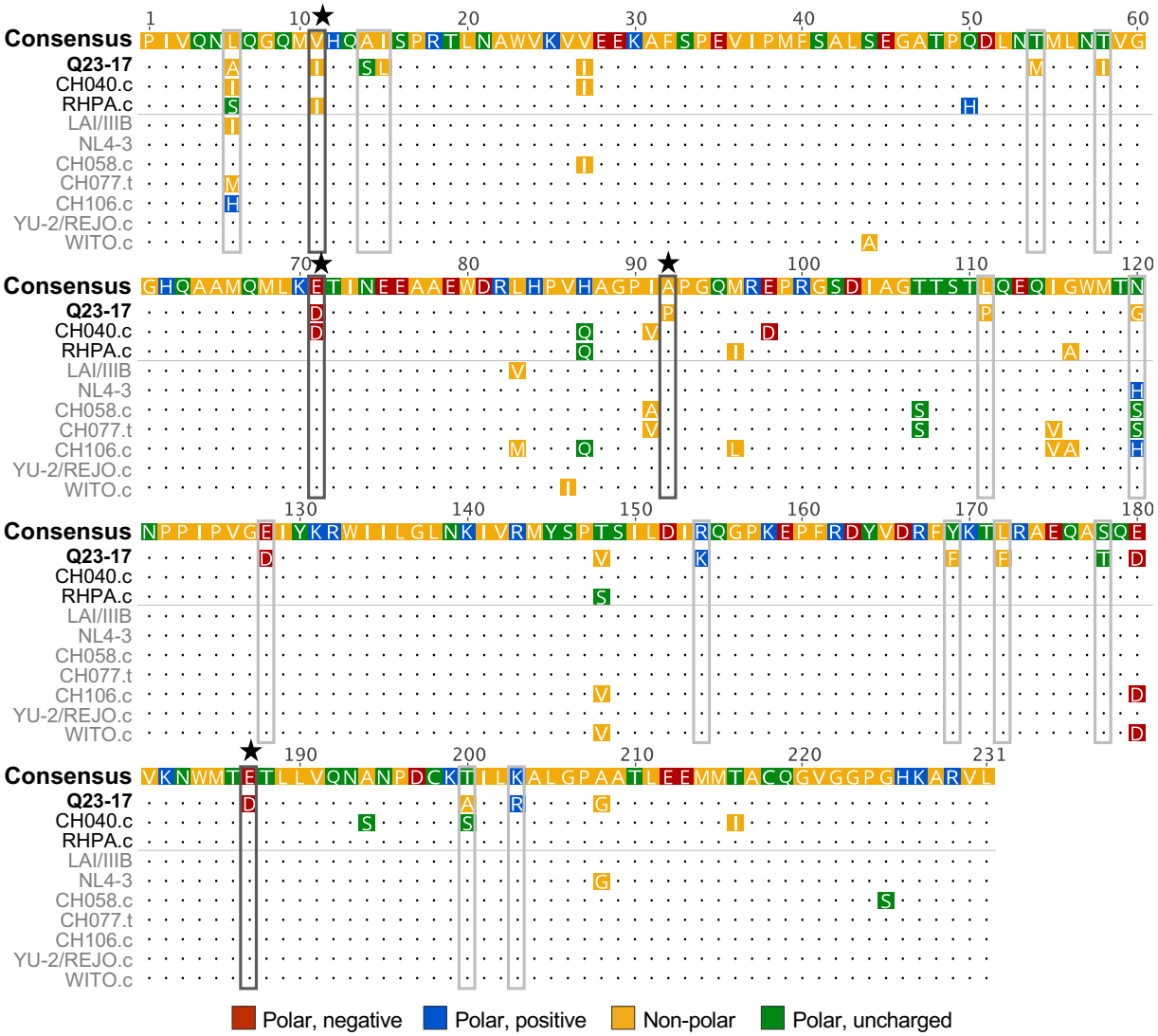
A HIV Levels with IFN



B Viruses previously tested for MX2 sensitivity

Virus	MX2 Sensitive	Clade	Tropism	Primary Isolate	Ref
Q23-17	**No**	**A**	R5	Yes	This study
CH040.c	**No**	B	R5	Yes	103
RHPA.c	**No**	B	R5	Yes	103
LAI/IIIB	Yes	B	X4	No	This study, 64, 69
NL4-3	Yes	B	X4	No	63-66, 99, 103, 134, 136, 137
CH058.c	Yes	B	R5	Yes	103
CH077.t	Yes	B	R5	Yes	64, 103
CH106.c	Yes	B	R5	Yes	64, 103
REJO.c	Yes	B	R5	Yes	64, 103
YU-2	Yes	B	R5	No	64
WITO.c	Yes	B	R5	Yes	103

C HIV CA sequence alignment

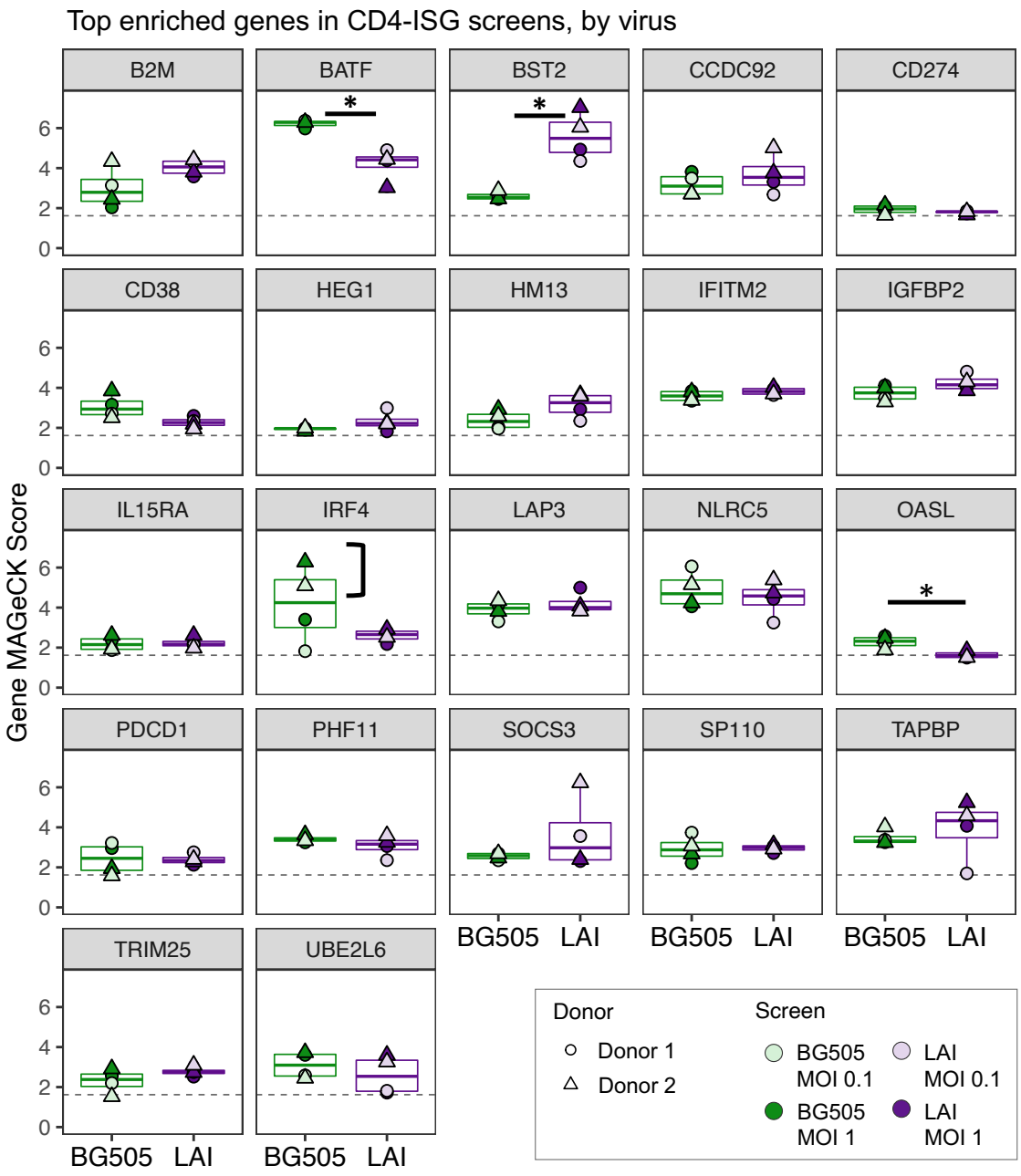


(figure continued on next page)

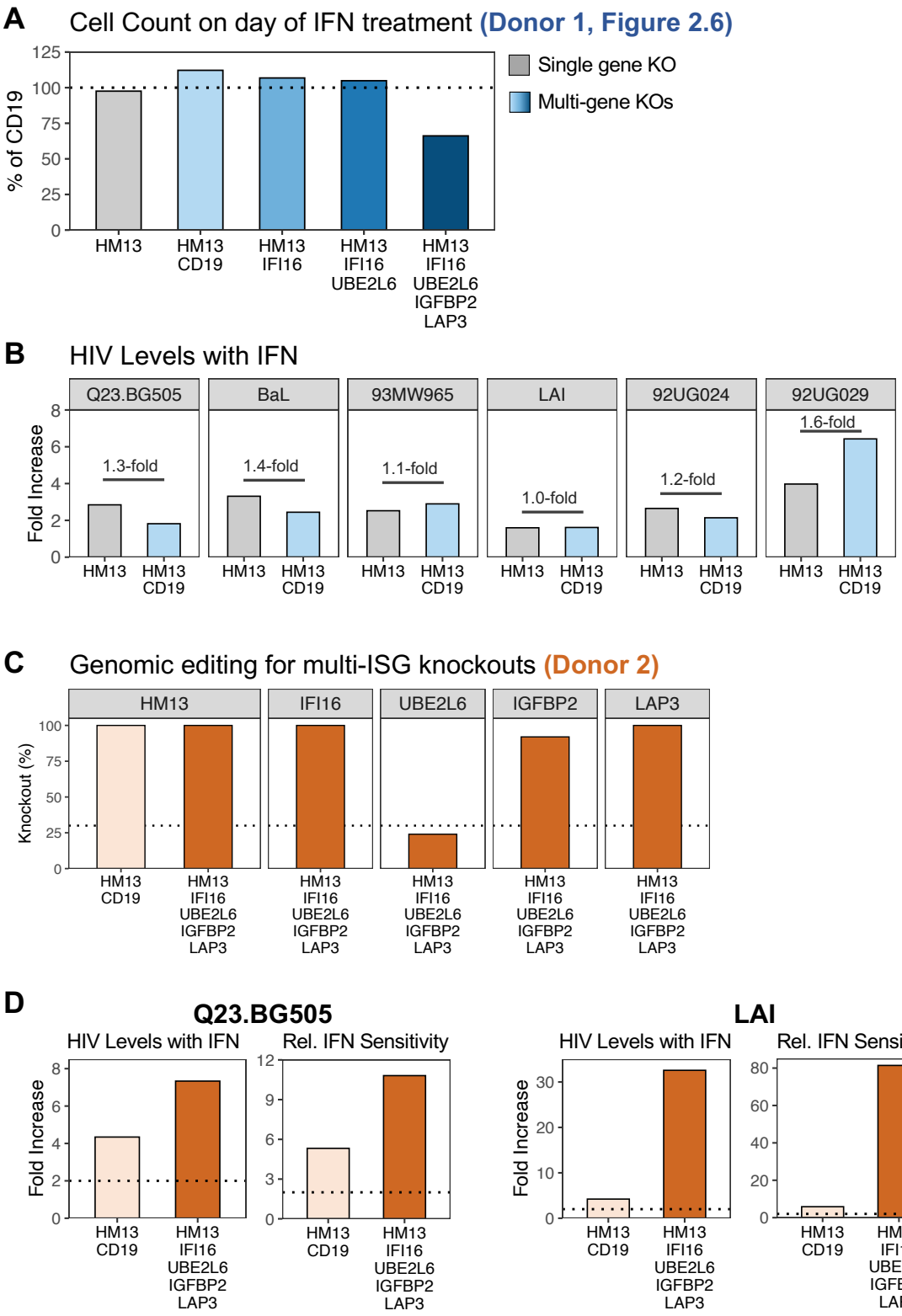
D HIV CA amino acid conservation

aa position	11	71	92	187
Q23-17	**I**	**D**	**P**	**D**
Group M	V	**D**	**P**	E
Clade A	**I**	**D**	**P**	E
Clade B	V	E	A	E
Clade C	V	**D**	A	**D**
Clade D	V	E	A	E
CRF01_AE	V	E	**P**	E
CRF02_AG	V	**D**	**P**	E
CRF07_BC	V	**D**	A	**D**

Supplementary Figure 2.3. HIV-1 Capsid amino acid residues differ between Q23-17 and MX2-sensitive viruses.
(A) Increase in HIV Levels with IFN as a result of MX2 KO compared to NTC (n=5 donors, data also in Figure 2.2). Donor colors correspond to those in Figure 2.2. (B) HIV-1 strains assessed for MX2 sensitivity in the current report and previous studies. (C) Alignment of HIV-1 CA protein sequences from viruses in panel B. HIV-1 strains are colored according to MX2 sensitivity (black: MX2-resistant, gray: MX2-sensitive). Amino acids are colored based on polarity. Locations in which the Q23-17 residue differs from all MX2-sensitive strains are boxed. Black, starred boxes indicate the four residues most likely to contribute to Q23-17's MX2 resistance. YU-2 and REJO.c are depicted together as their capsid sequences are identical. (D) Conservation of amino acids (aa) at sites likely to drive Q23.BG505's resistance to MX2 restriction. HIV-1 group and clade consensus sequences are based on Troyano-Hernáez *et al,* 2022 [138].



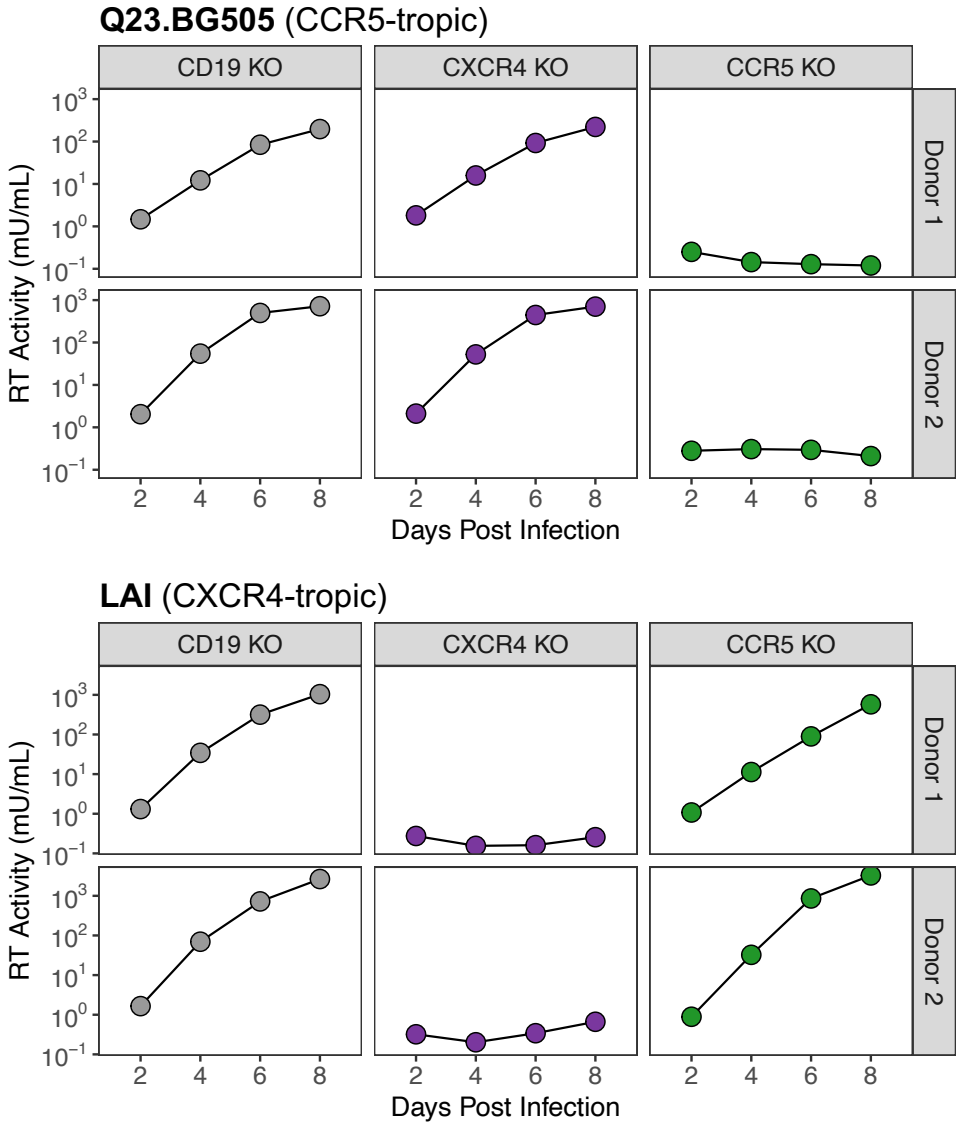
Supplementary Figure 2.4. Few hits from CD4-ISG screens are donor- or virus-specific. CD4-ISG screen hits were determined by selecting the genes that scored above background for every screen (56 genes, see Figure 2.4C). One long non-coding RNA, one previous ISG target (IFITM1), and genes with low expression in CD4$^+$ T cells (32 genes) were removed. MAGeCK scores for the remaining 22 genes are presented here. Graphs include data from all eight screens, separated by virus. Dotted line denotes the highest background threshold from Figure 2.4B plots. *p<0.05 by Wilcoxon rank test.



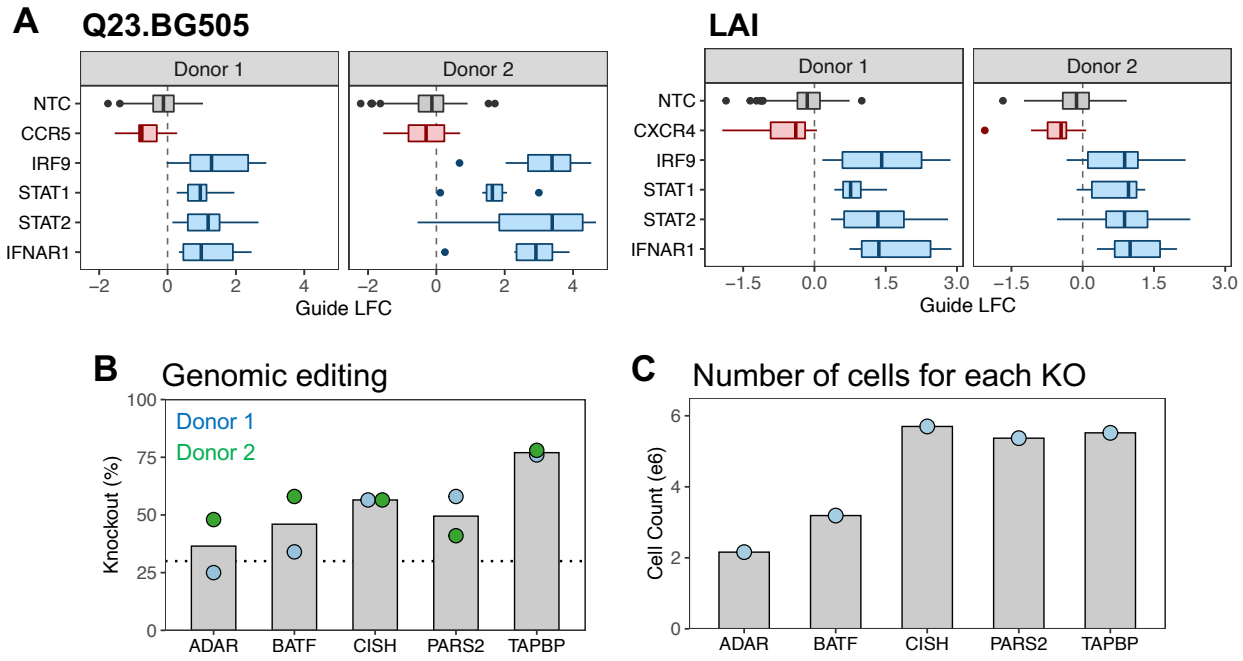
Supplementary Figure 2.5. Supporting data for multi-gene knockout experiment.
(A) The number of cells for each ISG KO on the day of IFN treatment are shown relative to CD19-edited cells for the Donor 1 multi-ISG KO experiment (related to Figure 2.6).

(figure legend continued on next page)

(B) The increase in HIV Levels with IFN (ISG KO/CD19 KO) for HM13 single or double KO as compared to CD19 KO for Donor 1. The fold difference between HM13 single and double KO is depicted within each plot. (C) Percent KO in cells nucleofected with multiple RNPs, as denoted along the x-axis, for the Donor 2 multi-ISG KO experiment. Results are separated by editing locus, as indicated above each panel. Dotted line denotes 30% KO. (D) The increase in HIV Levels with IFN (ISG KO/CD19 KO) and the decrease in Relative IFN Sensitivities (CD19 KO/ISG KO) for each multi-ISG KO compared to CD19 KO. Dotted lines indicate a fold difference of two.



Supplementary Figure 2.6. CD19 editing does not impact HIV-1 infection.
To determine whether CD19 editing impacts HIV-1 infection, we edited CD19 and HIV-1 co-receptors, CXCR4 and CCR5, in $CD4^+$ T cells from two donors. CXCR4 editing should not impact CCR5-tropic infection (Q23.BG505), and vice-versa for LAI infection. Spreading infection (MOI=0.02) data is shown for all treatments.



Supplementary Figure 2.7. Supporting data for PIKA HIV-CRISPR screens and validation experiments.
(A) Log_2 fold changes (LFC) for guides targeting control genes in PIKA HIV-CRISPR screens. Middle line of box-whisker plots indicates median of eight guides per gene (200 guides for NTC). (B) Knockout percentages for PIKA screen validation experiments. Donor colors in this figure correspond to those in Figure 2.7C. (C) The same number of cells were nucleofected for each ISG KO during PIKA screen validation (1e6). On the day of IFN treatment (9 days post nucleofection) each cell pool was counted. Results from Donor 1 are shown.

Appendix B:

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

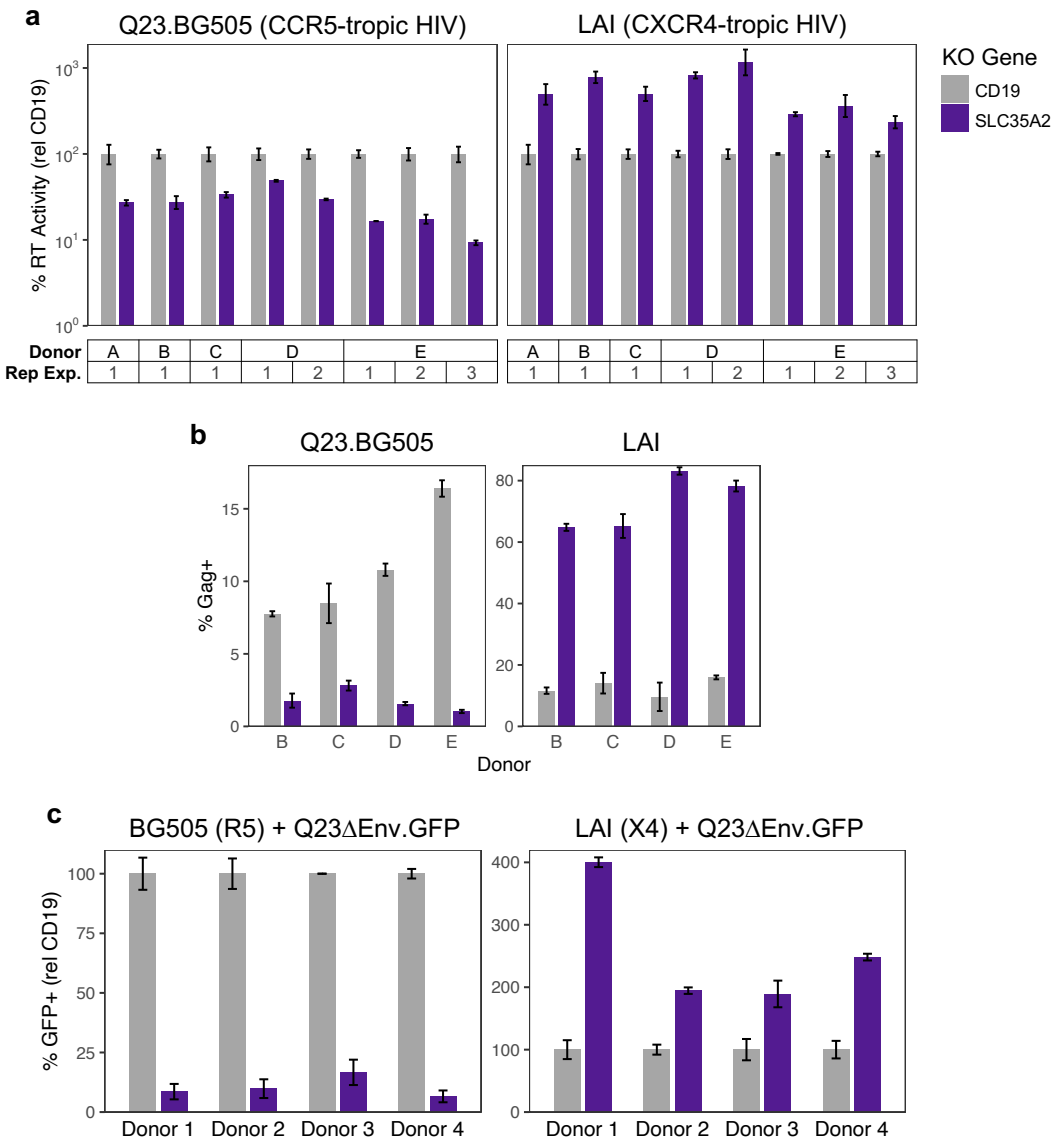
Supplementary Tables

Supplementary Table 3.1. PIKA guide library with synthetic NTC gene assignments.
URL accession:
https://www.biorxiv.org/content/biorxiv/early/2023/09/05/2023.09.05.556399/DC1/embed/media-1.xlsx?download=true

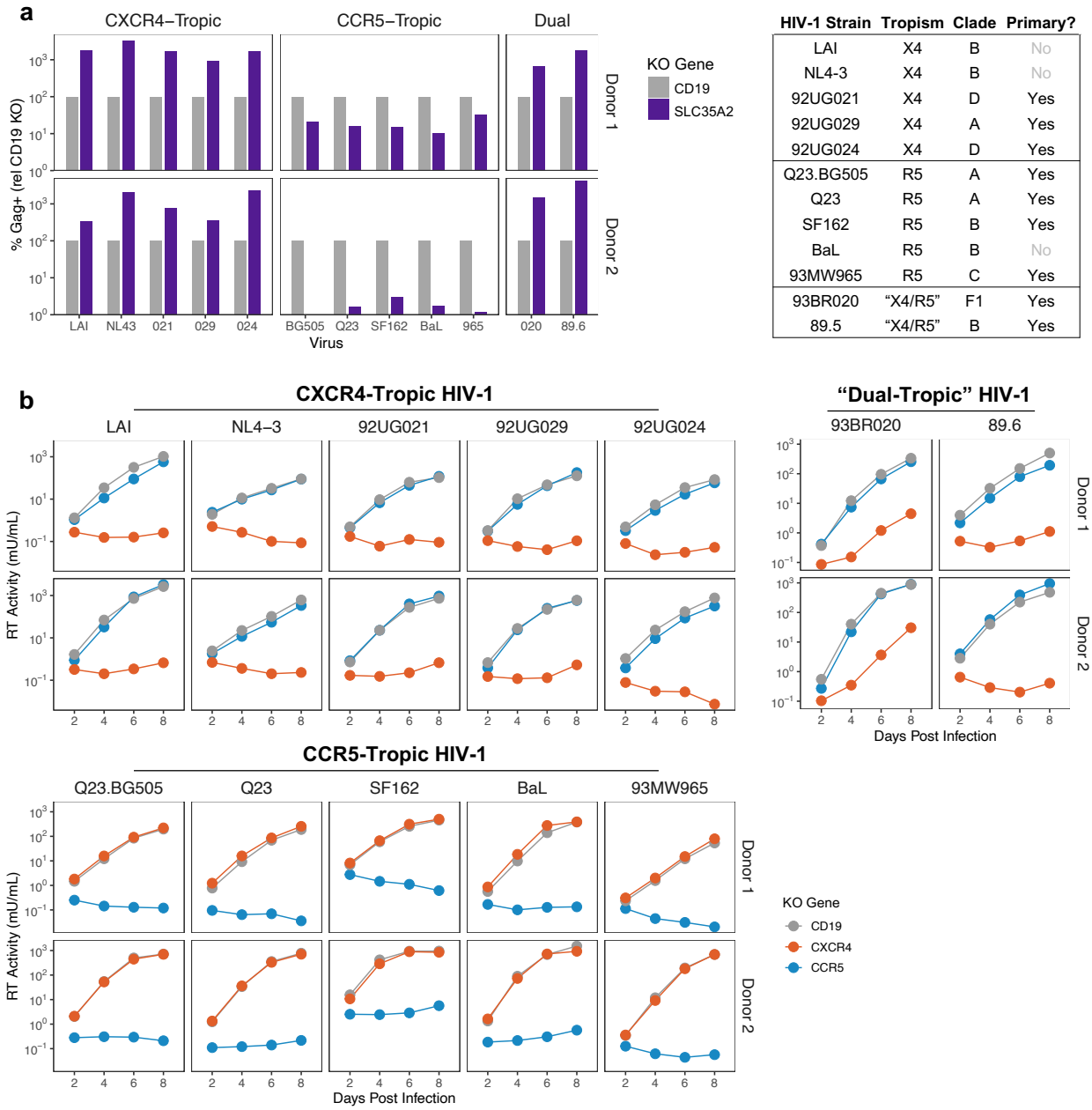
Supplementary Table 3.2. HIV-CRISPR screen results with the PIKA guide library.
URL accession:
https://www.biorxiv.org/content/biorxiv/early/2023/09/05/2023.09.05.556399/DC2/embed/media-2.xlsx?download=true

Supplementary Table 3.3. Guide sequences for single-gene editing experiments.
URL accession:
https://www.biorxiv.org/content/biorxiv/early/2023/09/05/2023.09.05.556399/DC3/embed/media-3.xlsx?download=true

Supplementary Figures

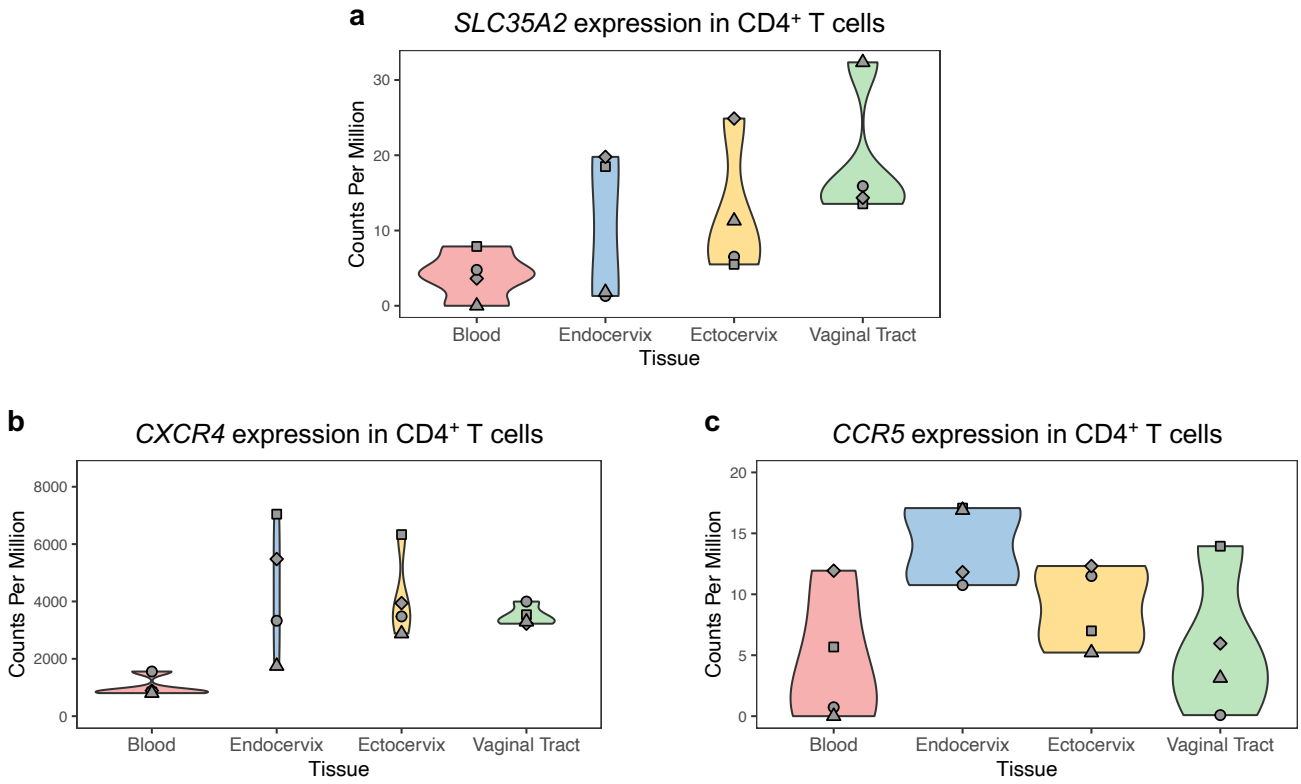


Supplementary Figure 3.1. SLC35A2 KO has opposite effects on two HIV-1 strains that utilize different coreceptors, related to Figures 2B, 2C, and 3D.
(A) Technical and donor replicates for data depicted in Figure 3.2B. Reverse transcriptase (RT) activity, relative to CD19 KO for each primary $CD4^+$ T cell donor/experiment, at 2 dpi (MOI=0.02). Knockouts were performed in five donors across independent replicate KO and infection experiments as indicated. (B) Technical and donor replicates for data depicted in Figure 3.2C. Percentage of $CD4^+$ T cells staining positive for HIV-Gag at 3 dpi (MOI=1). Donor letters correspond with those in Panel A. (C) Technical and donor replicates for data depicted in Figure 3.3D. Percentage of GFP+ $CD4^+$ T cells, relative to CD19 KO, after 2 days of infection with GFP-expressing HIV-1 pseudoviruses (MOI=1). All data are shown as mean ± standard deviation of 2-3 replicate infections.



HIV-1 Strain	Tropism	Clade	Primary?
LAI	X4	B	No
NL4-3	X4	B	No
92UG021	X4	D	Yes
92UG029	X4	A	Yes
92UG024	X4	D	Yes
Q23.BG505	R5	A	Yes
Q23	R5	A	Yes
SF162	R5	B	Yes
BaL	R5	B	No
93MW965	R5	C	Yes
93BR020	"X4/R5"	F1	Yes
89.5	"X4/R5"	B	Yes

Supplementary Figure 3.2. SLC35A2 KO differentially impacts CXCR4-tropic and CCR5-tropic HIV-1, related to Figure 2.2E.
(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 3.2E, different measurement of HIV-1 infection). Donor numbering is consistent with Figure 3.2E. (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 3.2E.



Supplementary Figure 3.3. Expression of SLC35A2 and the HIV-1 coreceptors in CD4⁺ T cells from the blood and common HIV-1 transmission sites.
(A) *SLC35A2,* (B) *CXCR4*, and (C) *CCR5* expression in CD4⁺ T cells isolated from different anatomical compartments from four donors. Shapes denote donors.