Effect of Major Splice Site deletions in Latency and Infectivity of 8 HIV Constructs
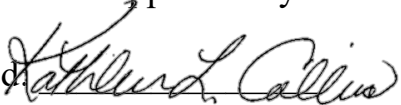
By: Rebecca Nitschelm

Honors Thesis, Biochemistry 4/21/21

This thesis has been read and approved by Kathleen Collins.

Signed:

Date: __04_ / _20__ /_2021_

Faculty advisor email: klcollin@med.umich.edu

Phone:

Table of Contents

I.   ABSTRACT

HIV infection is treatable but remains uncurable today, partially due to the virus's integration and latency in the host cell. Latent infection can later cause an active infection in the body if the virus resurfaces. Two donors on antiviral therapy for HIV infection are studied, donor 449 and donor 436. The HIV sequences studied, expressed in the donor's peripheral blood plasma or via an induction of cells via a viral outgrowth assay, contain a deletion or mutation in the major splice donor site (MSD) of the 5' leader sequence (Fig.1). As demonstrated in other retroviruses like murine leukemia virus, loss of a primer binding motif in the 5' leader sequence, which also acted as a repressor binding site, lead to an increase in viral expression.[1] So, we hypothesized that the genetic differences in these sequences may also explain why they were detectable even in donors on antiviral therapy. This thesis aims to elucidate possible explanations behind sequence differences that lead to differences in infectivity and latency.

Through construction of eight donor derived dual reporter vectors, *in-vitro* transfection and infection experiments were performed on human cell lines. We have found that our constructs are viable to express both mCherry and GFP, and that addition of the HIV trans-activator Tat increases transcription in our constructs. In addition, a correlation between low infection rate and a high degree of latent infection, possibly driven by the size of the deletion, has been found in the donor derived sequences. We attempted to activate the long terminal repeats (LTRs) during infection using a PMA and ionomycin drug treatment to show that the constructs were still functional, yet data was inconclusive due to a loss in cell viability. However, because the constructs showed activity during transfection experiments as quantified via mCherry reporter expression off of donor LTRs, we can conclude that the leader sequences are functional to produce viral particles. We suggest that the differences in infectivity were not due to differences in virion protein or HIV RNA production during transfection.

II. INTRODUCTION

A. Retroviruses

Retroviruses are characterized by reverse transcription, via the enzyme reverse transcriptase, and integration into the host cell DNA. Integration means that the viral RNA is reverse transcribed into DNA and inserted into the host genome, allowing for replication and transcription via the host cell machinery. This integration also allows persistent infection as well as vertical transmission into offspring, making retroviruses very difficult to cure.[4,5] The unique properties of retroviruses can also lead to other problems during integration, such as altering gene expression, leading to tumor formation.[5] Retrovirus integration is a very coordinated step in viral proliferation and latency in cells. It has even been shown that upon entry into the nucleus, unintegrated viral genomes undergo histone packaging and modifications prior to insertion into the host DNA that can affect its expression patterns later.[6] Integration of the viral genome is performed by the viral integrase enzyme. Integrase regulates the formation of an intasome complex that aligns the linear viral DNA (vDNA) with host protein machinery allowing insertion into DNA in the host chromosomes.[5,7] Use of non-catalytic site integrase inhibitors (NCINIs) or integrase strand inhibitors has also been shown to not affect viral processing or reverse transcription, yet specifically blocks the integration of the vDNA into the host DNA, thus providing more evidence of the specific role of integrase in retroviral integration.[5,8]

All retroviruses have three genes known as *gag*, *pol*, and *env*, that encode the necessary structural proteins of the virus. Specifically, *gag* encodes proteins for viral formation, *pol* encodes many important viral enzymes such as the viral protease (Pro), reverse transcriptase (RT) and the enzyme needed for integration into the host DNA integrase , and *env* encodes a transmembrane protein needed for entry into host cells.[5] Various retroviruses have different

splicing and different open reading frames to create other structural proteins; those for HIV will be discussed below.

There are 2 subfamilies of retroviruses orthoretrovirinae and spumaretrovirinae, with HIV belonging to the former. Of these two subfamilies, there are 6 genera and 1 genus respectively. The main differences between genera is defined by their morphology, assembly, open reading frames, transcription, and splicing methods.[5] HIV, a lentivirus, assembles at the cell membrane and contains conical core morphology; the core encases the RNA and is made of capsid protein. Lentiviruses have also been shown to infect cells throughout the cell cycle, during both mitosis and non-dividing stages.[5,9] Other examples of lentiviruses include SIV, caprine arthritis encephalitis virus, and visna virus.[5]

B. Human Immunodeficiency virus (HIV)

The history of HIV first began in 1981 when cases of severe immune deficiency began to rise, initially reported as clusters of Kaposi's sarcoma, a type of cancer, and Pneumocystis pneumonia in California and New York city among gay men as described by the Centers of Disease Control and Prevention.[10] Following its initial characterization, research on acquired immunodeficiency syndrome (AIDS), a result of HIV infection, as well as HIV, initially called Lymphadenopathy-Associated Virus (LAV) or HTLV-II retrovirus, increased as cases grew exponentially.[11]

HIV-1 is the most common type of HIV in humans and is studied in this thesis. The Human Immunodeficiency virus (HIV) genome is a 9.3 kb long RNA that complexes as a dimer. If left untreated, HIV can lead to AIDS. HIV is a positive sense retrovirus, which is characterized by reverse transcription of its single stranded linear RNA genome and integration of the provirus into the host DNA via the viral enzyme integrase. This integration allows persistent infection that

6

can be reactivated even after lying dormant in the host during treatment.[5,14] HIV specifically

forms a spherical core containing many gaps at the immature phases of assembly, eventually

forming an ordered core containing nucleocapsid protein that surrounds the RNA dimer. The

capsid proteins are hexamer or pentamer multimers, and are enveloped by a transmembrane Env

protein, as well as the host cell's bilayer during cell budding.[13]

      People on antiviral therapy still maintain their reservoir of latent infection, whose cell

proliferation makes HIV incurable today.[5,12] Latent infection of cells is characterized by

integration of provirus into the host DNA, proliferation through host cell replication machinery,

and then possible reactivation. A detectable viral load in the body can cause the host's CD4+ T

cells to die and could lead to AIDS; Antiretroviral drugs prevent establishment of this detectable

viral load. Although HIV+ individuals may not contain a detectable viral load on drug therapy,

targeting the latent reservoirs for eradication is required for a cure to HIV. Resting CD4+ T cells

are the largest contributors to this latent reservoir population, either through direct infection of

resting T cells, or infection of activated of CD4+ T cells, who return to a quiescent state after

infection.[15, 16, 17]

      In addition, other cell types such as peripheral blood mononuclear cells (PBMCs) and

hematopoietic stem cells (HSPCs) have been shown to be sources of the latent virus

reservoir.[2,3,15] Actively proliferating HSPC's that have been infected have been shown to

produce progeny containing the integrated provirus, and thus could be sources of residual virus if

the progeny's latent infection turns active. [14] With treatment, the viral load in the human body

can remain low, but because not all virus is active, if treatment becomes unavailable, latent virus

may reactivate and cause serious infection and medical complications in the host. This need for a

reliable source of treatment highlights the inaccessibility for healthcare treatment for many

people due to cost, environment of treatment, stigma, family, gender, or societal pressures and further exacerbates the disparity in health between groups.[18]

In the study of HIV latency, there are two main approaches to combatting latent virus in an effort to find a cure. One hypothetical method attempts to permanently turn off the latent virus, meaning the integrated viral genome will not reactivate and hurt the host; antiretroviral therapy will therefore not be required to diminish reactivated virus. The second option is to "turn on" all the provirus in the cells to become active, so that the virus can be eradicated with current or future treatment methods. Current antiretroviral therapy mostly targets either the viral protease or reverse transcriptase which prevent virus maturation or viral replication via integration, respectively, as well as a combination of treatments.[19] Overall, the study of what mechanisms allow HIV to be latent in cells, as well as the difference in active vs latent virus between different HIV sequences will help increase the understanding of how HIV functions in-vivo.

BI: 5' Leader Region of HIV

HIV contains many different important mechanisms and proteins that drive host cell binding, transcription, translation, and assembly. Specifically, this thesis focuses on the role of the 5' leader region in infectivity and latency of 8 HIV constructs. The 5' region of interest in HIV contains the primer binding site, stem loops 1, 2 and 3 (SL1, SL2, SL3), the dimerization initiation sequence in stem loop 1 (DIS), and the major splice donor site in stem loop 2 (MSD) (Fig 1). The primer binding site is complementary to an initiation tRNA, whose binding initiates DNA synthesis from genomic viral RNA (gRNA).[5] It has also been shown that binding to the primer binding site may also induce dimerization of the gRNA, allowing for RNA packaging.[20]

The three stem loops found in this region are important for assembly and packaging of the gRNA. Specifically, the Psi sequence of the 5' leader sequence mediates interaction with the

8

nucleocapsid protein that is required for viral assembly, and is also referred to as stem loop 3.[5, 21,22] Psi interacts with other stem loop regions of the 5' region in order to mediate other important roles in packaging; Stem loop 1 contains the dimer initiation site, made of a 6 nucleotide palindromic sequence that is important in dimerization of the gRNA prior to packaging. DIS has been shown to allow for Watson-Crick base pairing between two full length HIV gRNA, allowing for co-packaging of the genome, as well as mixing of HIV sequences if more than one exists in the cell.[22,23] Stem loop 2, which contains the major splice donor site, has been shown to form a platform conformation, whose long-range interaction ability may help explain its role in Psi-RNA stabilization, packaging, and splicing.[24]

The major splice donor site is responsible for making many RNA spliced products.

Figure 1: Donor 5' Leader region
The 8 donor constructs show a deletion or mutation in the major splice donor site. The constructs also show deletions in stem loop 1 (SL1), stem loop 2 (SL2), and stem loop 3 (SL3). The construct sequences are of interest because their deletion may give insight into why they were detectable in donors on antiviral therapy.
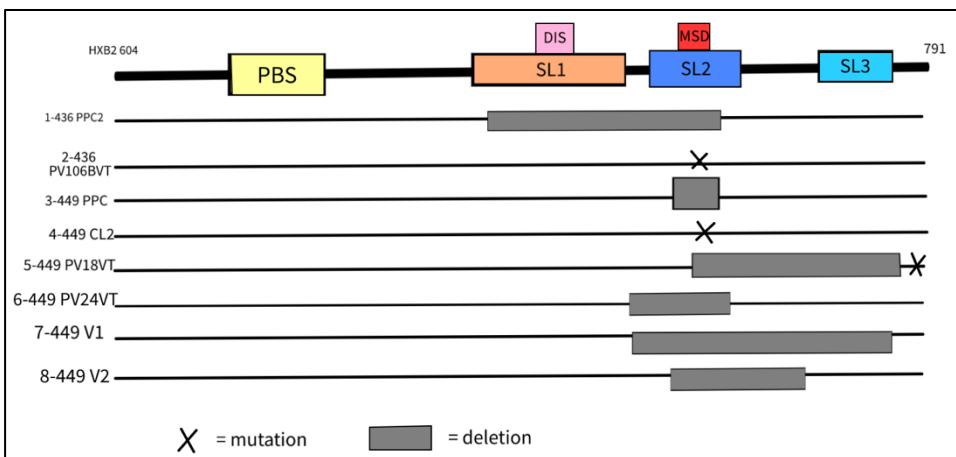
Transcription initially occurs after integration in the host cell DNA and is performed by the host RNA polymerase II. Cleavage of proteins is also required for HIV to make other proteins out of the three precursor proteins Gag, Pol, and Env.

Both Gag and Gag-Pol proteins are made from full length mRNA, while Env is made from splicing from the major splice donor site (D1) to various acceptor sites termed A4b, A4c, and A5.[5,25] The Gag-Pol polyprotein is made due to a -1 frameshift during transcription that occurs at about a 5-10% rate due to interaction of the RNA frameshift stimulatory signal and the ribosome,

9

making the stop codon that separates the two open reading frames no longer in frame.[5,26] Specifically, the MSD is required to create the spliced transcripts for many important proteins such as Env, Tat, Rev, and Nef.[27] Although the constructs studied contain a deletion in the MSD, they may still be infectious, as HIV has been shown to contain alternative MSD sites that may facilitate splicing even if the major splice donor site is mutated or missing, as well as 3 other donor sites termed D2, D3, and D4.[27,28] Specifically, HIV with alternative MSD sites such as AC|GG, TA|GT, ATGG|GT that also contain a deleted or defective MSD have been shown to still create spliced products needed for essential viral proteins.[28]

The Tat protein is of important interest in this thesis, as it is created via splicing from the MSD site, and drives an increase in transcription in HIV-1. At the start of transcription after infection, only a small amount of mRNA production occurs and produces Tat, which then drives the activation of transcription.[29] Tat has also been shown to play a role in cell to cell activation, being secreted by infected cells and taken up by uninfected cells leading to gene regulation of many immunoregulatory genes.[30]

Later protein cleavage of the Gag protein creates the matrix, capsid, and nucleocapsid proteins, from Pol the protease, reverse transcriptase, integrase, and dUTPase, and from Env the surface and transmembrane proteins.[5] Therefore, the major splice donor site, along with the other stem loops and primer binding site of the 5' leader region of HIV play a very important role in the function and processing of HIV, yet may have alternative sites if the main site is mutated or defective.

BII: Experimental Background

Eight viral donor constructs from 2 donors on antiviral therapy were taken from peripheral blood plasma or induced in donor cells via a viral outgrowth assay. These sequences

10

are of interest because their detection at all in donors who had clinically undetectable plasma virus level shows a possible explanation for how a MSD site deletion or mutation could lead to an increase in expression levels; the MSD site may also contain some sort of repressor binding site, and thus when deleted, leads to increased levels of expression. Thus, the study of these constructs may also elucidate explanations behind active vs latent mechanisms of HIV. Sequence differences and the size of the deletion or mutation could lead to conclusions on how varying sites in the 5' leader sequence can lead to expression or latency. It is also important to note that we are unable to quantify how many cells in the donors expressed these specific sequences, yet their detection at all is interesting.

It is unknown information from the donors whether the virus from which our donor constructs came from is infectious or not; this can only be determined in-vitro. Thus, we investigated the role of these various deletions in latency, infectivity, and ability of the viral constructs, in the presence of helper plasmid and VSVg, to express our reporter, mCherry; mCherry expression during transfection can be used to measure
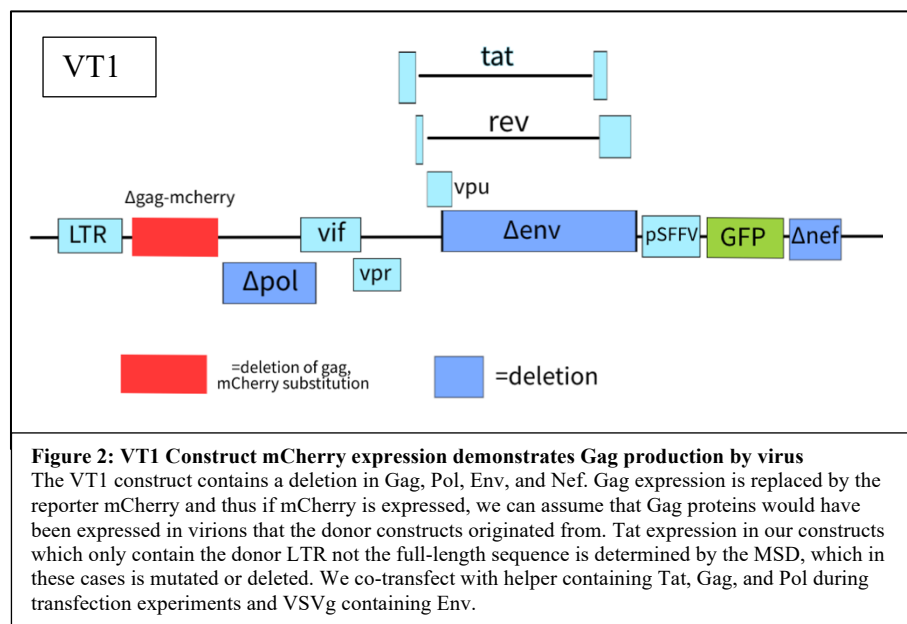
Figure 2: VT1 Construct mCherry expression demonstrates Gag production by virus
The VT1 construct contains a deletion in Gag, Pol, Env, and Nef. Gag expression is replaced by the reporter mCherry and thus if mCherry is expressed, we can assume that Gag proteins would have been expressed in virions that the donor constructs originated from. Tat expression in our constructs which only contain the donor LTR not the full-length sequence is determined by the MSD, which in these cases is mutated or deleted. We co-transfect with helper containing Tat, Gag, and Pol during transfection experiments and VSVg containing Env.

activity of the 5' leader region. Specifically, it is interesting to note that although all constructs studied contain a deletion or mutation in the major splice donor site of stem loop 2, some contain deletions in other parts of the 5' leader sequence such as in stem loop 1 and stem loop 3 that

11

could also affect their latency, activity, and infectivity (Fig 1). Expression of mCherry, the reporter downstream of each specific construct's leader sequence, demonstrates that the viruses would have made Gag, and were packaged in the case of mCherry expression during infection (Fig. 2). These sequences were found in virions, meaning that the provirus in the host DNA could be transcribed and synthesize the proteins needed for genome packaging. In addition, constructs V1 and V2, acquired through a viral outgrowth assay, because they were created by induced cells, are at least structurally functional to secrete virions. Therefore, we can assume that the constructs are functional, and detection of latent versus active virus in infection can be attributed to their sequence differences.

As our constructs are deleted for the MSD, we hypothesize that there will be an increase in activity in our constructs when Tat is provided during transfection, because they are not able to produce it on their own; whether they can produce Tat on their own in-vivo remains unknown. However, we would hypothesize that because our constructs were detectable in plasma virus, they were able to make Tat through alternative downstream splicing methods that our constructs do not contain, as we only studied the donor LTRs, not the full-length sequences. In our transfection experiments, we provide Tat, and therefore would expect mCherry expression to increase in comparison to
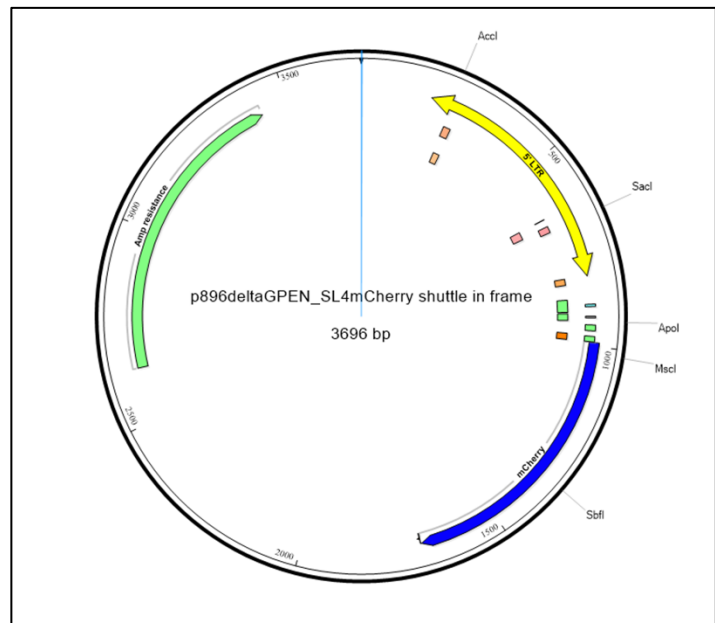
Figure 3: Shuttle mCherry Vector.
There is ampicillin resistance that allows for selection of bacteria containing the plasmid; we used this in our transformation reactions. mCherry expression is following the 5' leader sequence and thus allows visualization of the activity of inserted donor construct sequences. Restriction sites are labeled in the plasmid and used in diagnostic digests as well as for cloning into the full VT1 vector.

when no helper is provided because high levels of transcription require Tat expression.[29]

Five main steps were taken to study these donor constructs. First, stocks of all donor constructs were grown in bacteria (DH5a) and sequence confirmed. DH5a is an E. coli strain that via a bacterial transformation, allows researchers to grow up more plasmid. Then, the donor constructs were inserted into a shuttle containing the mCherry reporter inserted after the 5' donor leader sequence in the same LTR of the deleted Gag (Fig 2, 3). After this, a segment of the shuttle vector containing the donor 5' leader region was cloned into a full VT1 vector with a green fluorescent protein (GFP) reporter with its own constitutively expressed promoter (Fig. 4).

If GFP is expressed in cells after infection, this would show a successful infection of the cell, while if mCherry was expressed this would show that the 5' leader is active. This dual reporter model of study works by assuming that if the mCherry reporter is expressed, we would assume in vivo, the 5' region would be functional and express the viral proteins. All constructs were inserted into the full VT1 vector and checked using restriction digests and sequencing. Then, each construct was transfected into 293T cells, whose supernatants were then harvested

Figure 4: The full length VT1
The VT1 vector is deleted for gag, pol, env, and nef, making it non-infectious. For infection and transfection experiments, VSVg for env and pCMV-HIV to provide the structural proteins of HIV are used to allow for viral production during transfection. GFP expression is off of the Spleen focusing forming virus promoter (SFFVp) which is constitutively expressed.

and used to infect CEM-SS cells. The data on infectivity and active versus latent infection was

13

determined via flow cytometric analysis from the infection experiments. Flow cytometric

analysis of transfected cells demonstrates how well plasmids transfected into the 293T cells and

shows whether the constructs are able to express the fluorescent reporters in the presence of

helper, a substitute for expression of the viral genes if they were there in each construct (Fig. 5).

VT1 was our wild type control as it was the parental plasmid that we cloned all donor leader

regions into. dGPE and mCherry were our single positive controls that showed GFP only and

mCherry only reporters, respectively. Data was compiled and trends were observed between size

of the genomic sequence deletion and which regions were deleted that affected their infectivity

and latency for each construct. Three main constructs showed trends of lower transfection rates

and infectivity coupled with an increased percentage of infected cells showing only latent

infection, with very little proportion of infected cells containing an active infection; they were

often only single positive for GFP but lacked high amount of double positive GFP and mCherry.
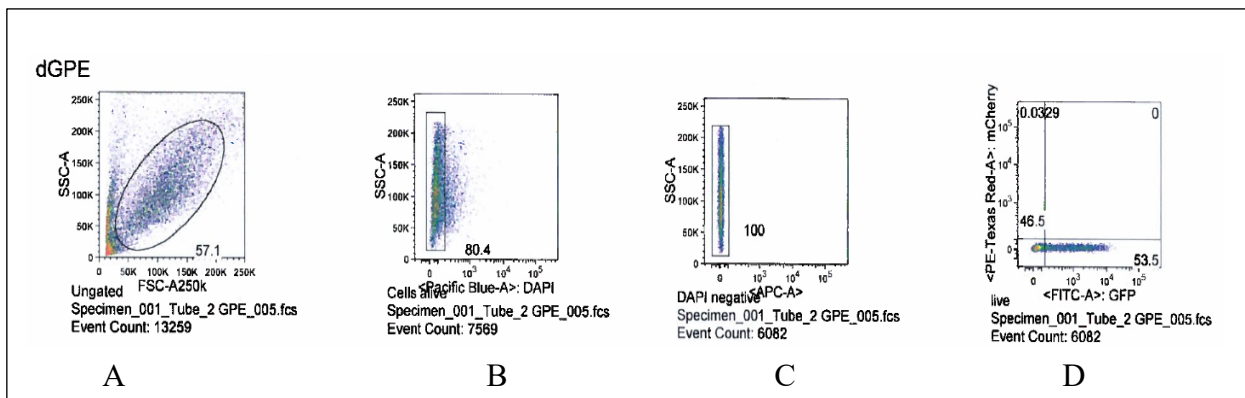
Figure 5: Example Flow Cytometric Analysis
This figure shows an example of flow cytometric analysis used in this thesis for one of our single-color controls, dGPE, which shows GFP expression. (A) We first gate on cells vs debris by analyzing the side scatter (SSC) vs forward scatter (FSC). (B) We then gate on alive cells using DAPI expression, who's channel for detection is often called "Pacific Blue." Alive cells are DAPI negative. (C) We then gate on an open channel, to make sure we are not detecting autofluorescence. (D) Our final gate for our experiments is looking for the reporter expression. Therefore we create our analysis graphs for mCherry expression versus GFP expression. The channel used for detection of these reporters in the flow cytomeric machines are often called "PE-Texas Red" and "FITC-A" respectively. We gate on Mock cells, which should have no expression of either reporter, and use those same gates in order to analyze our samples and controls. Above, we see that our single positive dGPE control had 53.5% GFP expression during transfection.

These constructs also contained the largest deletions, with deletions in more than one stem loop.

Constructs are numbered 1-8 and may be referred to as "Construct X" in this thesis (Fig. 1).

However, we saw that p24 concentration in viral supernatants, a protein in the HIV capsid, as

14

well as concentration of RNA produced in transfection, did not explain the trends that we saw in infectivity and latency, although some data was inconclusive and must be repeated. Finally, we observed that when Tat was provided in 293T transfection experiments there appeared to be no differences between constructs activity, as expected, showing that when Tat is provided, the deletion in the MSD that leads to decrease in Tat production may be recovered. However, we did see much lower activity in the infection experiments, which we hypothesize may be because Tat is no longer packaged with the virions, and thus construct's LTRs are not as activated. Further investigation of full-length donor sequences may be required to determine how these constructs are activated in-vivo and is described below.

III.   MATERIALS AND METHODS

A. Reagents

Buffers P1, P2, P3, N3, PB, PE, QBT, QC, and QF were acquired from a QIAGEN kit; either for the QIAprep spin miniprep kit, QIAGEN Plasmid Midi kit, QIAquick PCR purification kit, or QIAquick Gel extraction kit. Buffer TE is a common molecular biological solution (10mM Tris•Cl , 1 mM Ethylenediaminetetraacetic acid (EDTA)) and buffer EB is an elution buffer (10 mM Tris•Cl). Both of these reagents are also found in QIAGEN kits.

TAE buffer is made from 40 mM Tris Base, 20 mM acetic acid, and 1mM EDTA. LB broth consists of 5 g/L Yeast extract, 10g/L NaCl and 10g/L Tryptone. SOC growth media is similar to LB broth, but contains 20g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, and 20 mM glucose.

Carbonate coating buffer contains water as the solvent, with $NaHCO_3$ at a concentration of 100 mM and Na_2CO_3 at a concentration of 37.7 mM. Blocking buffer is made by addition of Casein at a final concentration of 1% in PBS. The casein is heated and stirred to dissolve. The wash buffer contains 2 M NaCl, Tween-20 to a final volume percentage of 0.5% into PBS as the solvent. This is then further diluted 1:10 with water. Lysis buffer is made via addition of Casein to a final concentration of 0.5%. Triton-X and Tween-20 make up 0.5%, and 0.05% respectively of the final volume. They are added to PBS, heated, and stirred to dissolve the Casein. Dilutant buffer is the same as Lysis buffer, in the same ratios, except no Triton-X is added.

B. Cell Maintenance of 293T and CEM-SS

293 T cells are cultured in D10 media, which is made of DMEM high glucose media with addition of Glutamine at a final concentration of 2 mM and fetal bovine serum (FBS) so it is

16

10% of the final volume, at 37°C and 5% CO_2 with a humidifying atmosphere. 293T cells are seeded at 10^5 cells for 4 days, or 2 x 10^5 cells for 3 days in a T25 tissue culture flask. For a T75 flask, 5 x 10^5 cells for 4 days can be seeded or 10^6 cells for 3 days. The seeding is done by aspirating the old media followed by a 2 or 5 mL phosphate buffered saline (PBS) rinse depending on the size of the flask. Then, the cells are trypsinized, which breaks them from their connection to the bottom of the container, with 1 mL of 0.025% Trypsin/EDTA media. These cells then sit for 3 min at 37°C, followed by trypsin neutralization by 4 mL of the D10 media. The cells are counted via microscope and reseeded as described above.

CEM-SS are grown in R10 media containing RPMI 1640 media, with addition of glutamine to a final concentration of 2mM, HEPES at a final concentration of 10mM, and 10% FBS. These cells are re-seeded every couple of days at a concentration of 1 x 10^5 cells per mL for 4 days and 2 x 10^5 cells per mL for 3 days in a T25. Cells are also grown in the same cell incubator conditions.

C. Sanger Sequencing

Genscript synthetically cloned each of the 8 donor constructs into a plasmid backbone containing restriction sites that we could use in our later cloning steps. We confirmed these synthetic plasmids via Sanger sequencing. Sanger sequencing is done through the Biomedical Research Core Facility DNA Sequencing Core at the University of Michigan. Following their guidelines, we created samples containing the DNA for sequencing at a concentration of 50 ng/µl. Buffer TE was added to make the total volume of the sample 10 µl. Primers were chosen that would confirm the validity of the sequence, or in later ligation experiments, the junction sites. Sequencing data was analyzed via the SeqMan program. Ends of the sequencing data often contained high background and unknown base pairs, so these were cut out from analysis.

17

Sequence data was analyzed in comparison with a reference sequence. We used this technique in order to confirm the junction sites in both cloning steps into the mCherry shuttle as well as the full length VT1 HIV vector. For confirmation of the Genscript synthesized insert sequences, we used the M13 reverse primer. For confirmation of the mCherry Shuttle ligation we used the primers HIV61-81d and mCh385r. For confirmation of VT1 correct ligation with shuttle vector we used the primers pUC2521 and mCh385r.

D. Bacterial Transformation of Genscript synthetic DNA

We received synthetically constructed DNA plasmids from the company Genscript Biotech. In order to create our own stock and grow up more plasmid of the constructed DNA plasmids, we transformed the plasmids into the E. coli competent cell DH5∝. Prior to performing this transformation, the Genscript plasmids were centrifuged at 6000g for 1 min at 4°C. Then, 20 µl of sterile nuclease free water was added to dissolve the plasmids. Finally, the samples were vortexed for 1 minute.

DH5∝ was thawed on ice as it is stored at -80°C. The competent cells were then aliquoted in 50 µl aliquots, followed by addition of the resuspended plasmid DNA. The samples were then placed on ice for 30 minutes. After 30 minutes, the vials were mixed and then heated at 42°C for 20 seconds, followed by putting the samples back on ice for 2 minutes. Finally, we added 950 µl of prewarmed super optimal broth (SOC) growth media and then put the samples in a 37° shaker for 1 hour. After 1 hour, we removed the samples and plated 100 µl onto 100 µg/ml Carbenicillin agar plates and stored them overnight at 37°C to allow colonies to grow.

The plasmids contain ampicillin resistance, and therefore the only colonies that survive on the agar plate must have taken up the plasmid. In addition, a pUC19 positive control plasmid was used. This cloning plasmid also contains ampicillin resistance and should transform into the

18

DH5∝ cells. Comparison with the number of colonies in the positive control shows that even if the sample transformations did not show colonies, the competent cells and reagents are still working if the positive control shows colonies. In addition, it allows calculation of the transformation efficiency of each test sample in comparison to the positive control.

The next day, a colony from each plate was picked and added to 3 mL of a Luria Broth (LB)/Carb mix made of 3 mL of LB and 3 µl of 100 mg/ml Carbenicillin. If the cultures were for a miniprep, they were put into the shaker overnight at 37°C. If these cultures were for a larger scale midiprep, they were put into the shaker for 6-8 hours at 37°C and then diluted, taking 25 µl of the original culture and adding it to 25 mL of the LB/Carb mix used above. These cultures were shaken overnight at 37°C. The cultures were stored as glycerol stocks at -80°C containing or were purified via minipreps or midipreps (see below). These techniques for transformation were utilized to make glycerol stocks of plasmids for future use in plasmid purification or to supply plasmids for purification by mini or midiprep.

In the shuttle vector transformation into DH5∝, the same steps were used, except 100 µl of the competent cells were used, as well as 900 µl of SOC media. After the SOC media was added, the shaking step was 30°C for 1.5 hours prior to plating on the Carbenicillin plates. The plates were stored at 30°C overnight, inverted. For the transformation of the VT1 ligated plasmid into competent cells, the same procedures were used as the shuttle vector transformation, except the competent cells Stbl2 were used. These competent cells are better for transformations of larger plasmids, like the full length VT1 HIV sequence, rather than DH5∝. This is because the full VT1 HIV construct is large, with repeating toxic elements. Thus, using a bacteria competent cell like Stbl2 that does not allow for as much recombination and does not expel the larger plasmid is required.

19

E. Glycerol Stocks

Glycerol stocks were made from transformed bacterial cultures that were grown overnight as described above. Samples were put into 1.5 ml snap cap tubes. Each snap cap was centrifuged at 6800g for 3 minutes at room temperature (RT), discarding the supernatant. Then the pellet was resuspended in LB broth+ 20% glycerol, transferred to a cryovial, and frozen in a dry ice/ethanol bath. These are stored at -80°C.

F. Plasmid Miniprep

Cultures that were not diluted after 6-8 hours in the shaker could undergo plasmid purification by QIAGEN miniprep. The samples were centrifuged at 6800g for 3 minutes at RT, discarding the supernatant.

Then, the samples were resuspended in 250 µl of chilled Buffer P1 containing RNAse A and Lyseblue in order to resuspend the cells to decrease viscosity as well as chew up any RNA contamination that may be present in the samples. The lyse blue allows for visualization of when the cells have been lysed. Then, 250 µl of buffer P2 was added, in order to lyse the cells, inverting 10-12 times until the solution turned blue, letting the solution sit for 5 minutes at RT after. After 5 minutes, 350 µl of buffer N3 was added, inverting 10-12 times until the solution is colorless; this buffer neutralizes the lysis buffer. Then, the samples were run at 17,000g for 10 minutes on a microcentrifuge. The supernatant was pipetted onto QIAprep spin column and spun again for 30-60 seconds. Then, 500 µl of buffer PB was added, spinning again for 30-60 seconds and discarding the flow through in order to bind DNA to the column. After this step, 750 µl of buffer PE was added and spun again for 30-60 seconds in order to wash the column, discarding the flow through. Finally, the column was spun for 1 min to dry, and the column was transferred

20

to a dry tube. The DNA was eluted into the new tube using 50 µl of buffer EB, letting the column sit for 1 minute, and then spinning at the same speed for 1 minute.

G. Plasmid Midiprep

Using QIAGEN plasmid midiprep kit instructions and buffers, we purified plasmid DNA in a larger quantity to the plasmid miniprep. First, we pelleted the competent cells at 6000g for 15 minutes at 4°C that had been transformed with plasmid DNA (Genscript DNA, shuttle plasmid DNA, or VT1 vector with insert DNA) as described above. We then used 4 mL of buffer P1 containing Lyseblue to resuspend the pellet. We then added 4 mL of buffer P2 and inverted the samples, making sure that the samples turned blue showing cell lysis. After letting the samples sit at room temperature for 5 min, we added 4 mL of chilled P3, again inverting to mix; samples should no longer be blue because P3 is acting to neutralize the lysis buffer. We then precipitated the DNA along with cells and proteins out by placing them on ice for 15 min. The samples were then centrifuged for 30 minutes at 20,000g at 4°C, transferring the supernatant to a fresh tube after spinning. The supernatant was then re-spun at 20,000g at 4°C for 15 minutes in order to make sure all cell particulates were removed from the supernatant in order to not clog up the column. The QIAGEN tip 100 was set up by running 4 mL of buffer QBT through it to equilibrate it. Following the spin and equilibration, the supernatant was poured onto the QIAGEN tip. The Qiagen spin column is made of a silica membrane that can bind DNA, as well as elute DNA in the presence of a low salt buffer.[31] Two wash cycles of 10 mL buffer QC were performed on the column. The purified DNA was then eluted off the column using 5 mL of buffer QF and precipitated using 3.5 ml of room temperature isopropanol. The precipitated DNA was then centrifuged at 6000g for 1 hour at 4°C, followed by a 2 mL 70% room temperature

21

ethanol wash. The pellet was spun again at 15,000g for 10 min at room temperature. Finally, the pellet was airdried and dissolved in 50 µl buffer TE or 50 µl buffer EB.

H. Restriction Digest for Diagnostic or Ligation Preparation

We used restriction digests as a diagnostic in order to confirm plasmid or vector ligation steps throughout our experiments. This is because if the correct ligation is performed, we are able to predict the sizes of the bands after running the digests on a gel; if there are 3 restriction cut sites in the ligated product, we should expect to see 3 bands at their expected sizes on the gel. This helps to support that our ligation occurred in the correct orientation. The first restriction diagnostic digest we performed was to confirm that our synthetic purified plasmids received from Genscript didn't contain any deletions after stocks were created by transformation of DH5∝ and midipreps performed (see above). The restriction enzymes AatII and HaeII were used in this experiment. 1 µg of each plasmid was treated with 10 units of restriction enzyme, corresponding to 1 µl of AatII and 1µl of HaeII, as well as 7 µl of water, and 2 µl of 10X buffer NEB CutSmart, for a total of 20 µl per reaction. Then, the digests were loaded into a PCR machine, set to run for 1 hour at 37°C, then 80°C for 20 minutes; samples were then put on ice. The restriction digests were then loaded onto a gel to visualize (see below). This digest should create single banding when run on a gel for the AatII samples as only 1 AatII cut site is present in our synthetic plasmids, and three bands created for the HaeII as there are 3 restriction sites of HaeII.

After growing stocks of our ligated mCherry shuttle containing the donor derived sequences, another diagnostic digest was performed on the purified plasmids in order to confirm that our ligation experiment created the restriction sites that we expect it to, supporting the hypothesis that the correct ends of each construct ligated into the shuttle. We treated each sample with AatII or BsrBI, following the same experimental procedures as above in the previous diagnostic digest.

22

AatII creates 1 band when run on a gel, while BsrBI creates 3, based on the number of cut sites each of these restriction enzymes contain in the insert-shuttle ligated product. We repeated this digest for both the mini and midprep purified plasmid DNAs.

The final diagnostic restriction enzyme digest we performed was on the full length VT1 HIV dual reporter vector containing our donor derived sequences. We utilized the restriction enzymes AatII and BsaAI combined with SaII. These restriction digests also contained the same procedures as seen above in the previous diagnostic digests. However, for the double digest of BsaAI and SaII, 1 µl of each enzyme, corresponding to 10 units per 1 µg of DNA was added to each sample, and only 6 µl of water; the 2 µl of 10X NEB CutSmart buffer was kept constant because the amount of DNA digested was the same. We used the same experimental settings when running our samples on the PCR machine as seen above. There should have been 1 band for the AatII digest of VT1 and 4 bands for the double digest when the restriction digests were run on a gel. We also checked that the SbfI restriction site was present in all final VT1 constructs in order to confirm correct ligation following the same experimental procedures described for the single diagnostic digests.

Other restriction digests on both the vector and insert were also performed prior to ligation experiments, to create complementary sticky ends to allow for ligation. When creating the mCherry shuttle vector and insert junction sites, we utilized the restriction enzymes SacI, MscI, and SacI combined with MscI in combination with either the vector or insert to a final concentration of 50 ng/µl. Single digests were utilized in order to see if the restriction enzyme could cut our samples; if each single digest shows the correct banding when run on a gel, we can assume that our double digest does in fact have the correct double cut from each enzyme. For construct 8, 449 V2, we substituted SacI restriction enzyme with AfIII due to different restriction

23

sites found in the donor HIV insert. Similar to the procedure above, we digested each sample of DNA (1µg of DNA) with 12 µl of water, 2 µl of 10X NEB CutSmart buffer, and 1 µl of the restriction enzyme (or 1 µl of each enzyme in the double digest mix and only 11 µl of water). We then combined these with 5 µl of the template DNA and ran the same PCR digest of 37°C for 1 hour followed by 80°C for 20 minutes. These specific digests are expected to create 1 band for the single digest, or 2 bands for the double digest in the gel analysis (explained below).

In order to clone the shuttle vector into the full length VT1 HIV vector we used the restriction enzymes AatII and SbfI single digest, as well as the double digest combination of the two. We followed the same ratios of water, buffer, and enzyme as described above in the shuttle vector cloning step. One of the two products of the double digest for both VT1 and each shuttle vector were ligated together in the cloning experiment (see below).

I. Agarose Gel Preparation and Gel purification

We created 0.8% agarose/TAE/GelRed agarose gels using 100 mL of 1X TAE buffer, 0.8 g of UltraPure agarose and 10 µl of 10,000X GelRed; for creation of larger gels, we doubled each compound in the mixture. For each gel analysis, we used λHindIII marker as our DNA ladder. For uncut samples, we loaded a sample containing DNA at a final concentration of 100ng/µl, 4 µl of buffer TE and 1 µl of 10X DNA loading buffer in order to visualize the banding. For each restriction digest treated sample, we took the 20 µl of the sample and added 2 µl of 10X DNA Loading buffer. We ran each gel at 120 Volts for approximately 1 hour or more until separation of bands can be seen. Gels were visualized using a UV gel visualization machine.

Gel imaging was used to analyze the diagnostic restriction digests performed to confirm correct synthetic sequences or ligation of DNA. In addition, gels were used to isolate specific fragments of DNA made by restriction digests described above based on size in a more

24

preparative manner. Many of the restriction digests mentioned above created multiple fragments of DNA, only one of which was of interest containing the insert, shuttle, or full-length vector with sticky ends, that will be cloned together. In order to isolate the correct piece of DNA, sequences of the shuttle, insert, or shuttle ligated to insert were used to determine the expected sizes of the DNA. These sizes could be compared to the bands seen on the gel in reference to the DNA ladder to determine the correct piece of DNA needed for cloning. This piece of DNA was then cut from the gel and purified following a QIAGEN QIAquick Gel extraction protocol.

The extracted gel was first weighed. The weight, in mg, was then used to determine the amount of buffer QG added to the gel. Buffer QG was added to each sample in the ratio of 3 µl of buffer for every 1 mg of gel. The samples were then heated at 50°C for 10 minutes to melt the gel, vortexing every 2-3 minutes to facilitate melting. The color should have appeared yellow at this point, and if they did not, then 10 µl of 3M sodium acetate was added. Then, 1 gel volume of isopropanol was added to the sample, 1 µl for every 1 mg of gel. The samples were then applied to QIAquick spin column and spun for 1 minute at max speed, discarding flow through after. This spin cycle was repeated until the total volume of the samples was run through the column and flow through discarded. The column was then washed with 500 µl of Buffer QG and spun for 1 minute at max speed. After the spin, the flow through was discarded and 650 µl of Buffer PE was added to the column and let stand for 4 minutes. After the 4 minutes, the column was once again spun at 1 minute at 17,000g, discarding the flow through after the spin. This spin was repeated to dry out the column. Finally, DNA was eluted into a clean 1.5 ml tube but 100 µl of buffer EB, incubating the samples for 4 minutes, finishing with another max spin for 1 minute.

25

J. PCR Purification

QIAGEN QIAQuick PCR purification kit and buffers were used to remove residual salts from the gel extraction of vectors and inserts as described above. For each sample, 5 volumes of buffer PB were added for every 1 volume of DNA; because we eluted with 100 µl of buffer EB in the gel extraction protocol, we added 500 µl of buffer PB. We then applied the samples onto the QIAQuick column. We then spun the samples at 17,000g for 1 min, discarding flowthrough. Following this, we performed a wash using 0.75 mL of buffer PE, following by a spin at the same speed and timing as before, discarding flowthrough. We repeated this spin again to dry out the column. We then eluted the purified DNA into a clean tube using 30 µl of buffer EB by incubating the buffer on the membrane for 1 min, followed by a spin at 17,000g for 1 minute. The concentration of resulting purified DNA was determined using a NanoDrop spectrophotometer.

K. DNA ligation of Vector and Inserts

Ligation of purified DNA was a vital step in the cloning process of making the full length VT1 plasmid vector. The two main ligation experiments performed involved ligating the mCherry Shuttle vector to the synthetically created donor derived 5' sequences. The purified product of this ligation was then also ligated to the full VT1 vector, creating a dual reporter vector containing the donor derived 5' sequences. For each of these experiments, two controls were utilized. The no ligase control acted as a negative control that should show that there is no ligation when the mediating enzyme is not present; if colonies appear on the no ligase control plate, the plates/samples may have been contaminated. The no insert control also acted as a negative control, and controls for the presence of uncut DNA (not digested correctly by the restriction digest) and for the presence of re-ligation of cut DNA not containing the insert.

26

The ligation samples were created by adding the vector to insert in a molar ratio of 1:3. This ratio favors the 1:1 vector to insert ligation, and skews away from incorrect ligations such as re-ligation of the vector. For the ligation experiment, the vector and insert concentration added was 1-10 ug/ml following this ratio. Four hundred units of the enzyme T4 DNA ligase were added to each sample except for the no ligase control. In addition, 2 μl of 10X NEB T4 DNA ligase buffer with ATP was added to each sample. The volume of each specific insert or vector added to each sample was determined using the concentration of each insert, determined by Nanodrop, as well as the ratio mentioned above that; based on average length of the inserts in comparison to the length of the vector, in the case of the mCherry shuttle vector ligation to insert 0.7 picomoles (pmol) of vector were added to each sample (corresponding to the 3.3 μl) as well as 0.21 pmol of insert. The μl of insert added thus depended on the concentration of each purified insert. Water was then added until each sample contained 20 μl total volume. The ligase was the last thing to be added to each sample. The samples were then incubated at 4°C for greater than 12 hours, followed by 16°C for 2 hours. The ligase enzymes in each sample were then heat inactivated for 10 minutes at 65°C. These ligated plasmids were then transformed into competent cells, DH5a for the shuttle and Stbl2 for full length VT1, as described above. In the case of the VT1 ligation with the shuttle vector, the average shuttle length in comparison to the VT1 length corresponded to 0.6 pmol of insert added to each sample and 0.2 pmol of vector. The 0.2 pmol of vector corresponded to 4.1 μl in our experiments based off the concentration of our purified plasmid, and the volume of insert varied. Again, water was added so the total volume of the samples was 20 μl.

27

L. Shuttle and VT1 vector Construction and usage

The shuttle vector is 3696 base pairs, and is from Kathleen Collins laboratory, construct 89.6 ΔGPEN_SL4 mCherry shuttle. The VT1 vector used for cloning is from Kathleen Collins laboratory, construct VT1 89.6 ΔGPEN_SL4mCherry_SFG. Spleen focus forming viral promoter (SFFVg) is the alternate promoter site for the GFP reporter. These constructs contain deletions in gag, pol, env, and nef that encode important proteins of HIV. This shuttle requires helper (pCMV-HIV-1) and VSV-G addition during transfection and infection tests. Helper provides structural proteins such as gag and pol and VSV-G provides Env, all of which our construct is missing.

M. Transfection in 293T cell line

293T cells were used as the cell line for transfection of the purified VT1 dual reporter plasmid containing the donor derived sequences in order to produce virus particles from each plasmid construct.

The first step of transfection is to seed the 293T cells the day before transfection. In a 6-well plate, this means seeding the cells at 5×10^5 cells per well. If a larger transfection is being used, 2×10^6 cells can be seeded per p100 plate. On the day of the transfection, the cells are first visualized to see if the cells are 50-70% confluent. D10 placed in the warm water bath. Each construct, as well as helper (pCMV-HIV-1), pVSVG, single color control GFP (NL-dGPE-GFP) and r mCherry (EF1a-mCherry), and positive control VT1 parental plasmid are nano-dropped to determine their concentration. In the small-scale transfection using 6-well plate, per well besides Mock, all wells have 400 µl of 150 mM NaCl, 2 µg of pCMV-HIV helper, 1.2 µg of pVSVg, 20.8 µl of 1 mg/ml PEI, as well as 2 µg of their plasmid DNA added. This plasmid DNA could be VT1, dGPE, mCherry, or one of the 8 constructs in question. For the large-scale transfections,

28

4 μg of pCMV-HIV helper, 4 μg of pVSVG, 48 μl of PEI and 4 μg of the DNA plasmid of interest are used, as well as addition of NaCl to 1 mL; only 1 mL is added to each p100 plate. The ratio of PEI used per plate is based on a ratio to the DNA amount;1 μg of DNA to 12 μL of PEI in the large scale transfections, where 1 μg DNA to 10.4 μl PEI is used in the small transfection.

First, the DNA and helper/VSVG are added to the NaCl in a sterile polystyrene tube. These samples are vortexed for 2-3 seconds to mix for small scale transfections or 10 seconds for larger scale. Then the PEI is added and vortexed immediately for 10 seconds. The samples let sit at room temperature for 15 min, and then were added drop wise to the 293T cells. The cells are put into the cell incubator containing 5% CO_2 at 37°C for 6-8 hours, until their media is replaced with 2 mL of fresh D10 media. In the case of the large-scale transfection, 8 mL of fresh D10 media is replaced 6-24 hours after addition. Media replacement is so that no viral supernatant collected at the end of the 2 days post transfection is merely what we added to the cells; we want to collect the viral particles that were created due to the transfection of the 293T cells.

At 2 days post transfection, the virus supernatant is collected. First the cells are spun out at 300g for 6 minutes and filtered through a 0.45 μm sterile syringe filter. Aliquots of the filtered supernatant are saved in aliquots used for later infection, enzyme linked immunosorbent assay (ELISA), or real time quantitative reverse transcription polymerase chain reaction (qRT-PCR) experiments and stored at -80°C. The 293T cells are then harvested to analyze using flow cytometric analysis. First, the cells are washed with 2 mL of PBS for small scale transfections, or 5 mL for larger scale transfections. The cells are then trypsinized with 1 mL 0.25% trypsin/EDTA. The cells are counted for viability and cell number under the microscope.

29

Each construct has 0.2 mL of cells transferred to a 96 well plate, including cells for compensation controls, which include Mock, Mock heat killed, dGPE and mCherry. Mock heat killed cells contain 50% mock cells that are head shocked for 1 minute at 65°C and then placed on ice, and 50% Mock cells. Then the cells are pelleted at 700g for 2 minutes and supernatant tapped off. The cells are then washed with 200 µl of fluorescent activated cell sorting buffer (FACS) buffer and spun again, with supernatant tapped off. All samples plus the compensation control Mock heat kill cells will be stained with 4′,6-diamidino-2-phenylindole (DAPI) which is a cell viability dye, while the solely compensation controls, Mock, mCherry, and dGPE will not be stained with DAPI. Thus, cells are then resuspended in 100 µl of DAPI/FACS buffer. This DAPI/FACS buffer consists of DAPI at a final concentration of 40 ng/mL. The compensation controls are resuspended with just 100 µl of FACS. Then the cells are kept on ice in the dark for 15 minutes.

After 15 minutes, the cells are spun at 700g for 2 minutes and supernatant discarded. The cells are then fixed with 100 µl of 2% paraformaldehyde/PBS and kept at RT for 30 minutes for fixation to occur. Finally, the cells are pelleted at 800g for 2 minutes and supernatant discarded. The cells are resuspended in 150 µl of FACS buffer once; the wells are rinsed again with 150 µl of FACs and pooled for flow cytometric analysis.

N. Infection in CEM-SS cell line

Similar to transfection experiments, prior to infection, cells are seeded the day before at 1:1 with R10 media in order have cells in the log phase at the time of transduction. On the day of the infection, 5×10^5 cells are aliquoted into sterile snap cap tubes. These are spun in a tube holder at 2500g for 5 minutes, and then the supernatant is aspirated. Following this, each sample is resuspended in 0.5 mL of the 293T viral supernatant of interest. Then, 5 µl of 400 µg/ml

30

polybrene is added in order to increase the efficiency of virus infection of the cell line. Finally, the samples are mixed gently and transferred to a 24 well plate. The spin infection is at 4160g at room temperature for 2 hours, after which the supernatant is aspirated and 0.5 mL of warm R10 media is added back to each well. The cells are harvested 3 days later after being grown at 37°C in the cell incubator as described above.

On day 3, the cells are counted for their viability. Then, 150 µl of each culture (including extra for compensation controls) are added to a 96 well plate. The cells are pelleted at 700g for 2 minutes at 4°C, discarding supernatant after the spin. Then, the plate is washed with 200 µl of FACS buffer. The same procedure for DAPI staining described above in the transfection experiment is used here; once again, the compensation controls do not receive DAPI staining.

O. PMA/ionomycin test

Addition of PMA/ionomycin 2 days post infection of CEM-SS cells with viral supernatant is a technique used to show the functionality of the construct leader sequences. In our experiments, we stimulated CEM-SS spin infected cells (0.5 mL of volume, see protocol above) on day 2 of infection. We added a volume of pma/ionomycin of 1:1 with the cell volume, thus adding also 0.5 mL. We added a 2X concentration of pma and ionomycin, which was PMA at a concentration of 100 ng/mL and ionomycin at a concentration of 6 µg/mL. The rest of the volume up to the 0.5 mL was R10 media; because we were adding this 0.5 mL of drug treatment to an equal volume of the cells, the final concentrations of pma/ionomycin were 1X, or 50 ng/mL PMA and 3 µg/mL ionomycin. The infected cells were still harvested on day 3, one day later.

We noticed that many of the sequences did not show an increase in mCherry expression compared to a control lacking pma/ionomycin, yet rather they showed significant decrease in cell viability. Therefore, we hypothesized that the drug treatment of PMA/ionomycin at 1X

31

concentration may have been too high and caused a large enough decrease that inhibited our ability to see an increase in mCherry activity in the drug treated cells. Therefore, we performed a pma/ionomycin titration in an attempt to determine the correct concentration to treat cells that did not decrease cell viability.

We followed very similar steps to the infection in CEM-SS cell line (above). However, instead of resuspending the cells in viral supernatant, we merely resuspended the cells in D10 media, which is what the viral supernatant is usually found in. On day 2, we performed the pma/ionomycin test. We had the conditions untreated, 1X DMSO which contained 2.1% DMSO, and then performed serial 1:2 dilutions, taking 1 mL of each stock out and adding it to 1 mL of R10 media to create 1/2X DMSO, 1/4X DMSO as our control drug treatment. DMSO is usually toxic to cells at a concentration greater than 1%, so we could use the dilutions to compare cell toxicity of this drug to our pma/ionomycin test. 1X PMA/ionomycin contained 50 ng/mL PMA and 3 µg/mL ionomycin. We also performed 1:2 serial dilutions down to 1:4, with addition of R10.

We then followed the same procedure for all samples; removal of 0.5 mL from the well was followed by addition of 0.5 mL of the respective solution. On day 3, these cells were harvested by the same protocol as noted in Infection in CEM-SS cell line (above).

P. Enzyme-Linked Immunosorbent Assay (ELISA)

ELISA was used to measure the p24 protein concentration in viral supernatant. P24 is a HIV viral capsid protein of gag, and thus its detection is possible though the use of 2 antibodies that detect the two different epitopes of the protein via ELISA and can show the concentration in ng/ml of HIV, whether it be infectious or not.

32

We used anti-p24 antibody, diluted in carbonate coating buffer at a concentration of 1 μg/ml by adding 5 μl of 2 mg/ml stock of anti-p24 in 10 ml of Carbonate coating buffer. The p24 antibody we used was purified 183-H12-5c, non-biotinylated. We then put 100 μl per well in a NUNC Maxisorp plate and let is sit overnight in order to coat the plates with the p24 capture antibody.

The next day we started the ELISA protocol. First, we washed the plate 3 times with 200 μl wash buffer, followed by addition of 200 μl of blocking buffer per well in order to stop nonspecific binding. The plates then sat at room temperature for 1 hour, no more, and were washed once again 3 times with wash buffer using the plate washer. Next, we prepared the 3, 10-fold serial dilutions on the samples, as well as dGPE, mCherry and VT1. The dilutions of each sample are important in order to see which concentration yields an ELISA concentration that is accurate; if the concentration of p24 is very low, the least dilute sample may be the most accurate reading, and vice versa. Each sample was diluted in lysis buffer, because it contains triton that breaks down the HIV envelope and allows p24 to be detected. For Mock, we performed a 1:4 lysis buffer dilution, while the other samples were a 1:10 serial dilution. In order to create a p24 standard, we also performed 6 2X serial dilutions of a Virogen p24 standard, starting with a concentration of 50 ng/ml and ending with 0.78125 ng/ml. We then added 100 μl of the sample or standard to each well plate in triplicate and sealed for a 2-hour RT incubation or overnight incubation at 4°C.

After incubation, we washed the plate 4 times in wash buffer with 1-minute incubations. Then, we added 100 μl of diluted anti-Gag-biotin antibody at a final concentration of 0.5 ug/ml. We diluted in diluent buffer. We let this detection antibody then sit for 1 hour at room temperature. We followed this incubation by washing the plate 4 times with wash buffer, with 1-

33

minute incubations. We then added 100 µl of diluted poly-streptavidin-HRP to each plate and let the plate sit for 30 minutes at room temperate, which binds to the biotin antibody, and also contains a the horseradish peroxidase conjugation (HRP), which will allow for color visualization and detection when bound by its substrate tetramethylbenzidine (TMB). The final concentration of poly-streptavidin-HRP added to each plate was determined by a 1:10,000 dilution, 2 µl of polySA-HRP into 10 mL of diluent buffer. We then once again washed the plate 4 times with wash buffer, with 1-minute incubations. Then, we added 100 µl of TMB substrate and let incubate at room temperature until the highest concentration appear to be at saturation, seen via a blue color change. Then 25 µl of 0.5 M H_2SO_4 was added and mixed by tapping; the wells should turn yellow as the acid stops the HRP-TMB reaction. Finally, the color of each well was read at 450 nm with a reference of 650 nm on a spectrophotometer. A standard curve was created based on the p24 standards in order to determine the concentration of p24 in each sample.

Q. Real Time Quantitative Reverse Transcription PCR (qRT-PCR)

Real time quantitative reverse transcription PCR (qRT-PCR) was used to determine the amount of RNA present in viral supernatants, in order to equalize amount of viral RNA used to infect in CEM-SS infections. First, 750 µl of Trizol LS was added to 250 µl of viral supernatant. Benches were wiped down with bleach, followed by RNAse- away in order to break down any RNA on surfaces that could contaminate our samples. After the workspace was set up, 1 µg of control Raji RNA, extracted from a human cell line, was added to each sample and inverted to mix. The control RNA is used to calculate the efficiency of the RNA precipitation step because we know the amount of RNA we added. In addition, in our unknown samples, if there is not a lot of RNA present, RNA precipitation may not be as effective; addition of the control RNA aids in

34

ensuring there is enough RNA present to make the RNA precipitation effective. The samples then sat at room temperature for 1-2 minutes.

Organic extraction of the RNA first began with addition of 200 µl of chloroform to each sample, shaking vigorously for 15 seconds after addition. Then, we let the samples sit for 2-15 minutes at room temperature. We then spun the samples at 12,000g for 15 min at 4°C. We removed the upper aqueous layer phase to a new pre-chilled tube, being careful to not take any part of the interface.

We then precipitated the RNA. First, we added 10 µg of glycoblue in order to increase visibility of the RNA pellet later. Then, we added 0.5 mL of isopropanol per sample and let the samples sit for 10 minutes at room temperature. Following this, we centrifuged the samples at 12,000g for 15 minutes at 4°C. We then washed the RNA pellet with 1 mL of 75% ethanol and vortexed to mix. We centrifuged the samples once again at 7500g for 5 min at 4°C and discarded the supernatant. We let the pellet air dry for 5-10 minutes and then resuspended it in 10 µl of RNAse free water. We finished the procedure by incubating the samples for 10-15 minutes at 55-60°C. We could stop here and store the samples at -80°C.

Then, we moved onto the cDNA synthesis phase of qRT-PCR because we need to synthesize DNA from RNA in order to quantify the samples using DNA polymerase in our PCR reaction. First, we ran the UV treatment to destroy any DNA contaminants that may be present in the PCR hood. We had a no template control, a no reverse transcriptase control, and a Raji RNA control. We added 11 µl of water, 4 µl of 5X qScript cDNA SuperMix to a final concentration of 1X, and 5 µl of the RNA to all samples except for the Raji direct and no reverse transcriptase controls. For the control RNA, we added 500 ng of the RNA, 4 µl of the SuperMix, and only 6 µl of water. For the no reverse transcriptase control, we added 5 µl of RNA (one sample chosen), and

35

15 μl of water. We flicked the samples to mix. We ran the cDNA synthesis reaction at 25°C for 5 minutes, followed by 42°C for 30 min, 85°C for 5 min, and then 4°C until collection.

Then, we performed the Gene expression Assay for beta-actin. We used the cDNA created from our control RNA described above to create our standards, creating 1:10 serial dilutions by taking 2 μl from our control Raji sample at a concentration of 50 ng/μl and adding it to 18 μl of nuclease free water. After thoroughly mixing, we would repeat, taking 2 μl from the new dilution and adding it to 18 μl of water. We used neat, 10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4} concentrations in order to make our standard curve. For our beta-actin reaction, we used 1 μl of 20X Taqman Beta-actin (ACTB) Gene Expression Assay which allows for quantification of Beta-actin in cDNA samples. ACTB contains the probe labelled with Fluorescein dye (FAM) that allows for quantification. We also added 7 μl of water, 10 μl of 2X Taqman Gene Expression Master Mix, and 2 μl of our template.

For our HIV expression assay, we used the primers 9501-mRNA-F as our forward primer and 9629-polyA-R for our reverse primer. A final concentration of 1μM of primer, 10 μl of 1X Taqman Gene expression Master Mix, 7.4 μl of water, 0.25 μM concentration of our probe for the reaction, referred to as L9531FM, were added, and 2 μl of our template sample. FAM is also the fluorescent marker and the minor groove binder (MGB) is our quencher. For our standard curve creation for HIV we used a pVQA Standard plasmid developed by Bullen et al. which is at a concentration of 10 ng/μl in 10mM tris buffer.[32] We got this plasmid from the National Institute of Health's AIDs Reagent Program. We know that the length of the HIV standard is 4338 base pairs, which means we can use the concentration, molar mass (660 g/mol bp), and Avogadro's number of copies per mol to determine the number of copies per volume. This standard then allows us to determine the copies of HIV cDNA in our samples. We created serial

36

dilutions just as we did in the beta-actin assay. Specifically, we took 5 µl of the stock solution and added it to 5.5 µl of buffer EB to make a solution with 1×10^9 copies per 2 µl. We did 1:10 dilutions by taking 4 µl of the starting solution and adding it to 36 µl of buffer EB.

We ran the qPCR reaction using the optical plates with an optical film layer at 50°C for 2 minutes, followed by 95°C for 10 minutes. Then we ran 45 cycles of 95°C for 15 seconds allowing for denaturation of strands, followed by 60 seconds at 60°C to allow for primer binding and DNA elongation. The probe fluorescence is measured during the DNA elongation stage when the polymerase breaks down the quencher-probe connection and allows the probe to fluoresce; as more cDNA is present, more probes can bind and subsequently be broken off from the quencher to fluoresce and be detected and quantified. Based on the known standards, the fluorescence of each sample can be compared to the fluorescence of the standards, specifically the cycle threshold value (CT), which corresponds to a specific quantity (ng) in the Beta-actin assay or a specific copy number in the HIV assay per volume, measured during the elongation stage.

R. Flow Cytometric Analysis

We utilized the Biomedical Research Flow Cytometry Core's equipment at the University of Michigan or the Microbiology Department at the University of Michigan in order to run samples from transfection and infection experiments. FlowJo allows data analysis via grouping of samples, compensation control adjustment, and layout for display and reporting. In addition, we utilized Flowjo to gate first away from debris by plotting Side-scatter (SSC) vs Forward-scatter (FSC). Then cells were gated for live cells (those that are DAPI negative) by SSC vs DAPI fluorescence. Once live cells were selected, we utilized an SSC vs an open channel such as APC-A to make sure that our final population does not contain any background due to

37

autofluorescence. Finally, we plotted our population of cells mCherry vs GFP in order to

visualize which cells were GFP positive, mCherry positive, or both. Therefore, Flowjo analysis

allowed us to see the percent active versus latent virus, because the active virus would be

expressing both GFP and mCherry, while latent virus would only be expressing GFP.

IV.   RESULTS AND DISCUSSION

A. Synthetic Constructs Are Cloned Successfully into VT1

Figure 6: Diagnostic Digest of Genscript Synthetic donor constructs
Diagnostic digest of miniprep purified plasmids of constructs 5-8. Expected band sizes for each diagnostic digest of Genscript synthetic plasmids (A). Diagnostic digest of construct 1-4 was performed prior to joining the lab. Constructs show streaky bands in digests 1-3; the third digest attempt is shown in (B), which may be evidence of residual ethanol from miniprep purification of the plasmids. The final digest and gel show the clearest bands (C). Extra DNA fragments may show differences in coiled or relaxed states of the DNA. (B) Upper Comb: Lanes 1-4 AatII digest on constructs 5-8. Lanes 5-8, HaeII digest on constructs 5-8. Lane 9, Molecular Weight marker λHindIII digest. Lower comb: Lanes 1-4 uncut constructs 5-8. Lane 5, Molecular Weight marker λHindIII digest. (5C) Upper comb Lanes 1-4 AatII digest on constructs 5-8. Lane 5, 6 empty. Lane 7, VT1 shuttle HaeII digest. Lanes 8-11 HaeII digest on constructs 5-8. Lane 12, Molecular Weight marker λHindIII digest. Lower comb: Lanes 1-4 uncut constructs 5-8. Lane 5, Molecular Weight marker λHindIII digest.

We were unable to completely confirm by restriction digest and gel analysis that the synthetic sequences obtained from Genscript contained the correct synthetic construct sequences as bands were very streaky and contained more bands than expected (Fig 6a,b).We repeated the digests 4 times, and only on the 4th digest did streaks decrease, even though extra bands were still present (Fig 6c). We concluded that the extra bands may have been from supercoiled DNA which migrates faster than the linear DNA or from nicked DNA which migrates slower. Therefore, we moved forward with sequencing the plasmids in order to confirm that the synthetic sequences

39

contained our donor leader regions as well as restriction sites needed for later steps in cloning. As determined by sequencing, the donor constructs and restriction sites were present in our plasmids after DH5a transformation and miniprep spin purification.

However, we were still able to isolate the target insert and vector sequences containing the correct junction sites by gel purification from our restriction digests of the mCherry shuttle vector and insert sequences. From sequences, we knew the expected band sizes of each insert or
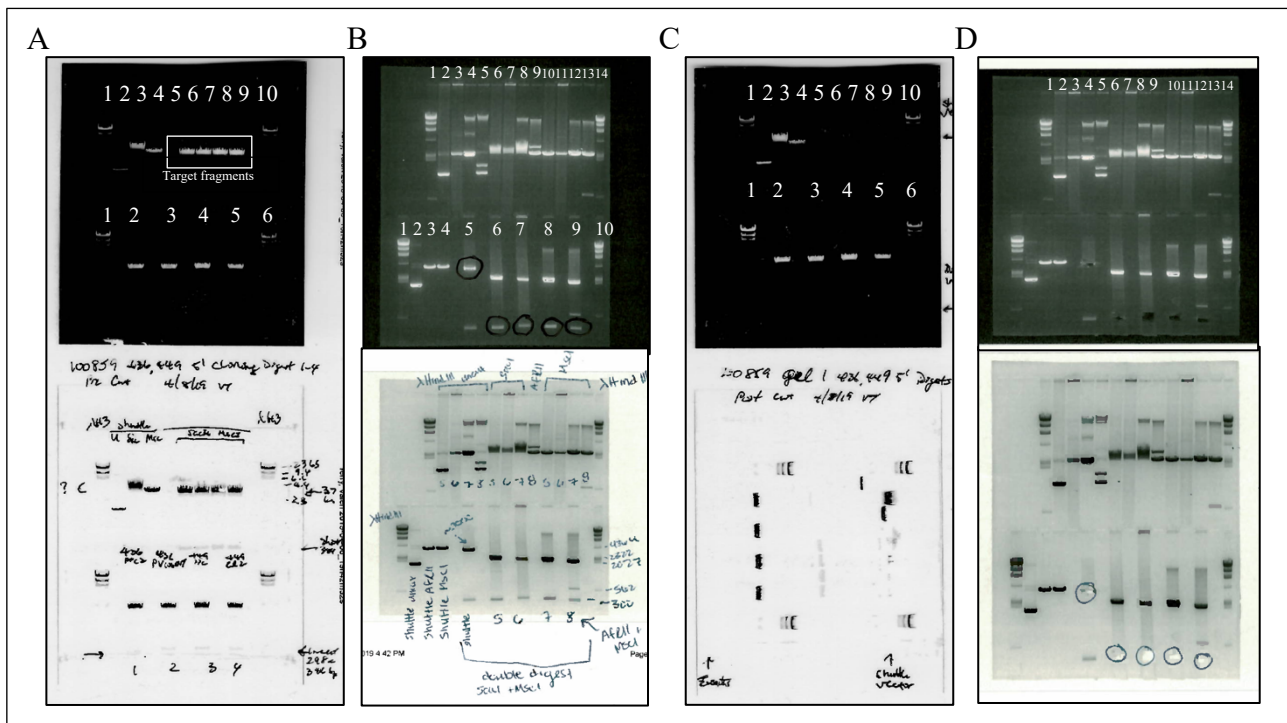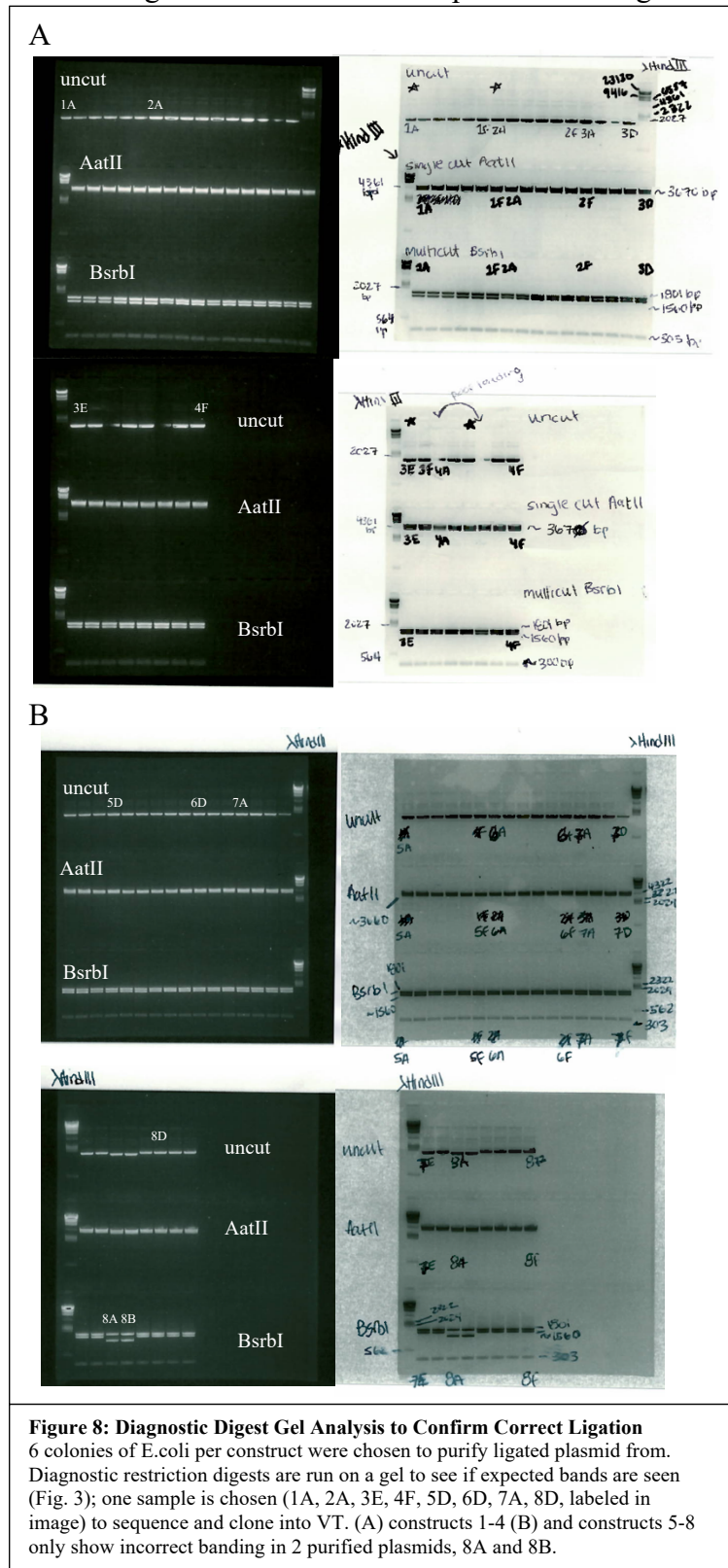

Figure 7: Successful isolation of DNA fragments for Cloning
(A-B) Pre-cut gels for construct 1-4 (A) and 5-8 (B). The target insert sequences are from the double digest, with band length of around 300 bp. (C-D) Post-gel cut. The removed DNA fragments are no longer seen on the gel. These will be purified and used for cloning. Constructs 1-4 on left (C) 5-8 on right (D). (6A/C) Upper comb: Lanes 1, 10 Molecular Weight marker λHindIII digest. Lanes 2, Shuttle uncut. Lane 3, SacI digest Shuttle. Lane 4, MscI digest. Lanes 6-9 Shuttle double digest SacI, MscI. Lower comb: Lanes 1, 6 Molecular Weight marker λHindIII digest. Lanes 1-4, double digest constructs 1-4. (6B/D) Upper comb: Lanes 1, 14 Molecular Weight marker λHindIII digest. Lanes 2-5, uncut constructs 5-8. Lanes 6-8, SacI digest on constructs 5-7. Lane 9, AflII digest on construct 8. Lanes 10-13, MscI digest on constructs 5-8. Lower comb: Lanes 1, 10 Molecular Weight marker λHindIII digest. Lane 2, Shuttle uncut. Lane 3, Shuttle AflII digest. Lane 4, Shuttle MscI digest. Lane 5, MscI/AflII double digest of Shuttle. Lanes 6-9, MscI/AflII double digest of constructs 5-8.

shuttle that was our target. We used UV and DNA loading dye to visualize the gel, before and after cutting out the target band containing our correctly digested vector and inserts. Although once again there appeared to be extra DNA fragments in the uncut DNA of constructs 5-8, we still isolated the DNA fragments that corresponded to our expected insert size (Fig 7a-d). We

40

then confirmed that these correctly sized DNA fragments did in fact correspond to our target vectors and inserts via sequencing, after ligation of the two. Expected insert size is around 300 bp, while the mCherry shuttle target DNA fragment should be about 3300 bp; we cut the gel for the bands corresponding to this size in order to purify the fragments for ligation of the insert and shuttle. Our restriction digest, gel purification, and ligation were successful in cloning the donor HIV leader constructs into the mCherry shuttle for 46/48 of our clones (Fig 8a-b); 6 transformed E.coli colonies were picked per construct in order to ensure that we got the correct ligation of shuttle and insert. We visualized the ligated products using a diagnostic digest run on an agarose gel (Fig 8a-b). Our ligation of insert into mCherry shuttle was shown because the

Figure 8: Diagnostic Digest Gel Analysis to Confirm Correct Ligation
6 colonies of E.coli per construct were chosen to purify ligated plasmid from. Diagnostic restriction digests are run on a gel to see if expected bands are seen (Fig. 3); one sample is chosen (1A, 2A, 3E, 4F, 5D, 6D, 7A, 8D, labeled in image) to sequence and clone into VT. (A) constructs 1-4 (B) and constructs 5-8 only show incorrect banding in 2 purified plasmids, 8A and 8B.

expected DNA fragments as determined by the restriction digests performed were present on the gel. There are 3 BsrbI cut sites and 1 for AatII; we saw the correct number of DNA fragments corresponding to these cut sites on our gel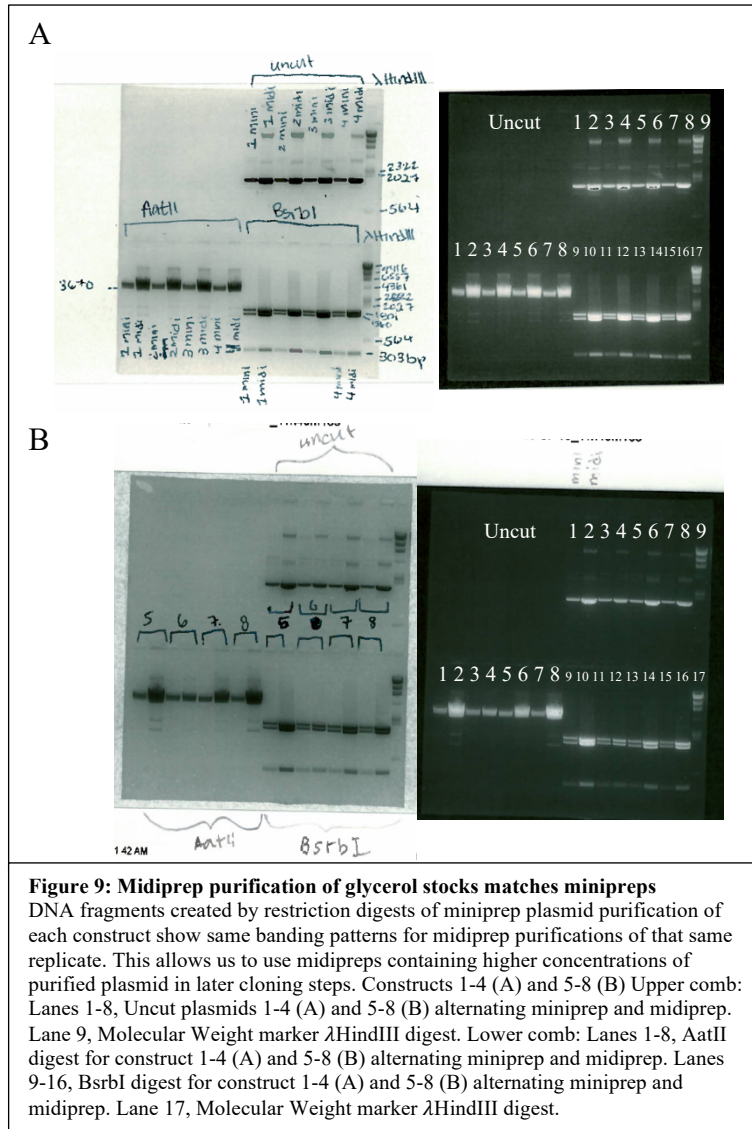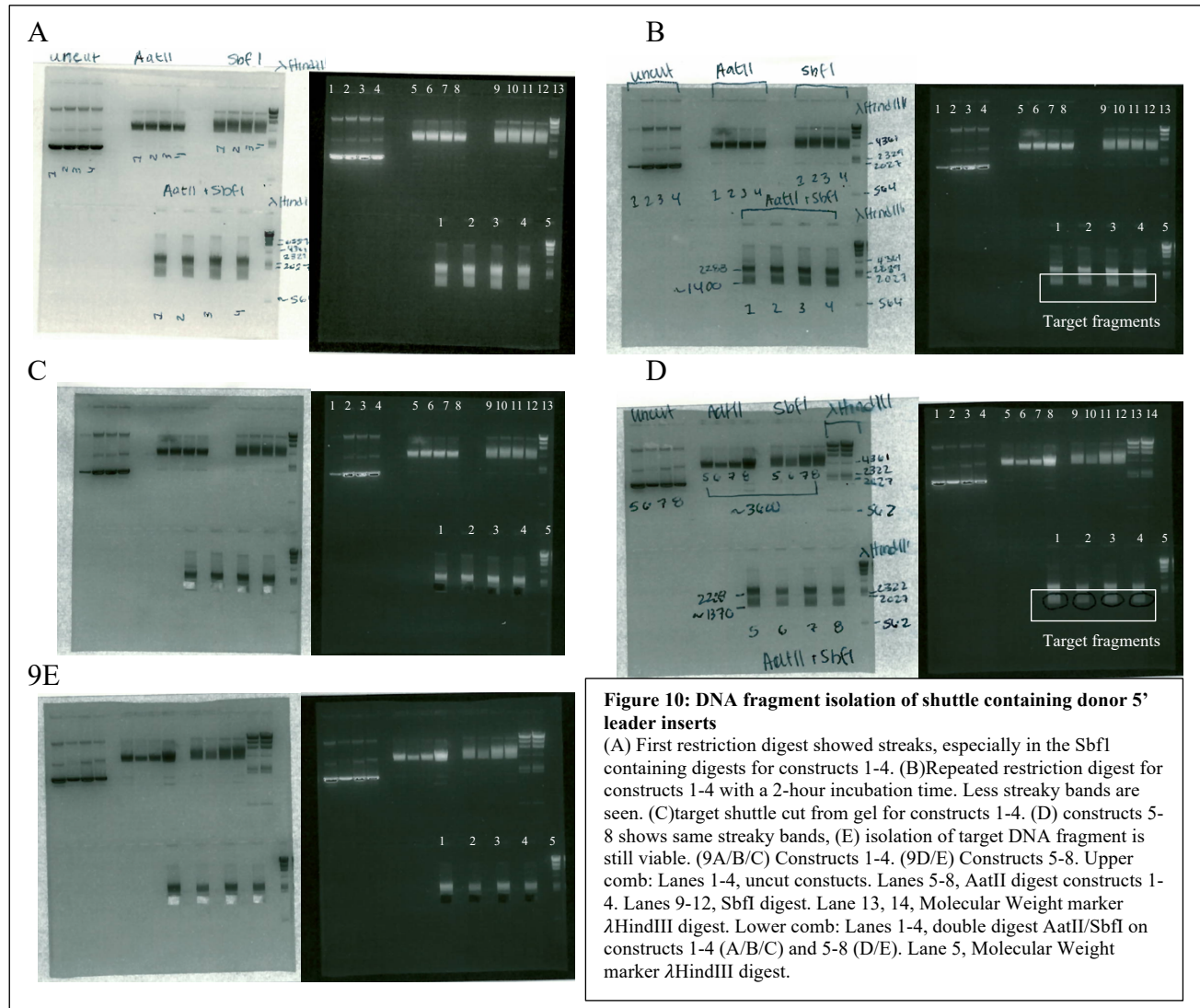 (Fig. 8a-b). We picked the ligated product with correct bands on the gel whose concentration was the highest to sequence and move forward in the cloning process (Fig 8a-b). We also performed midipreps from glycerol stocks of the constructs that we are moving forward with and confirmed that they showed the same DNA fragments as the minipreps for those constructs when restriction digest was applied; the midipreps showed the same banding pattern as the minipreps, so we can conclude that the midipreps can be used for VT1 cloning (Fig 9a-b).

Figure 9: Midiprep purification of glycerol stocks matches minipreps
DNA fragments created by restriction digests of miniprep plasmid purification of each construct show same banding patterns for midiprep purifications of that same replicate. This allows us to use midipreps containing higher concentrations of purified plasmid in later cloning steps. Constructs 1-4 (A) and 5-8 (B) Upper comb: Lanes 1-8, Uncut plasmids 1-4 (A) and 5-8 (B) alternating miniprep and midiprep. Lane 9, Molecular Weight marker λHindIII digest. Lower comb: Lanes 1-8, AatII digest for construct 1-4 (A) and 5-8 (B) alternating miniprep and midiprep. Lanes 9-16, BsrbI digest for construct 1-4 (A) and 5-8 (B) alternating miniprep and midiprep. Lane 17, Molecular Weight marker λHindIII digest.

Our restriction digests for VT1 cloning also showed streaky banding patterns in all constructs, especially in the SbfI and double digest DNA fragments; this may be due to the SbfI restriction enzyme needing a longer digest than we performed because if SbfI maintains its attachment to the DNA it's digesting, it can create smears on the gel (Fig 10a).[33] We repeated the digest for the constructs with a longer incubation time (2 hours) and ran the products on a gel.

The DNA fragments still showed some streaks, but isolation of the correct DNA fragments was still possible (Fig 10b-e).

Figure 10: DNA fragment isolation of shuttle containing donor 5' leader inserts
(A) First restriction digest showed streaks, especially in the Sbf1 containing digests for constructs 1-4. (B)Repeated restriction digest for constructs 1-4 with a 2-hour incubation time. Less streaky bands are seen. (C)target shuttle cut from gel for constructs 1-4. (D) constructs 5-8 shows same streaky bands, (E) isolation of target DNA fragment is still viable. (9A/B/C) Constructs 1-4. (9D/E) Constructs 5-8. Upper comb: Lanes 1-4, uncut constucts. Lanes 5-8, AatII digest constructs 1-4. Lanes 9-12, SbfI digest. Lane 13, 14, Molecular Weight marker λHindIII digest. Lower comb: Lanes 1-4, double digest AatII/SbfI on constructs 1-4 (A/B/C) and 5-8 (D/E). Lane 5, Molecular Weight marker λHindIII digest.

Ligation of the insert containing mCherry shuttle and VT1 construct was successful in 46/48 of our selected E.coli transformed colonies, as visualized through diagnostic digest on a gel (11a-b). Clear separation of DNA fragments demonstrates the expected band sizes in reference to the expected ligated constructs restriction sites. We confirmed constructs by sequencing constructs 1A, 2C, 3B, 4E, 5D, 6E, 7D, 8F. In addition, we confirmed that the SbfI site was still present in the VT1 plasmids for sequences 1-4 (Fig. 11c). We also confirmed that

43

the midiprep plasmid purifications for constructs 5-8 matched the minipreps for those same selected full length VT1 clones (Fig. 11d). We se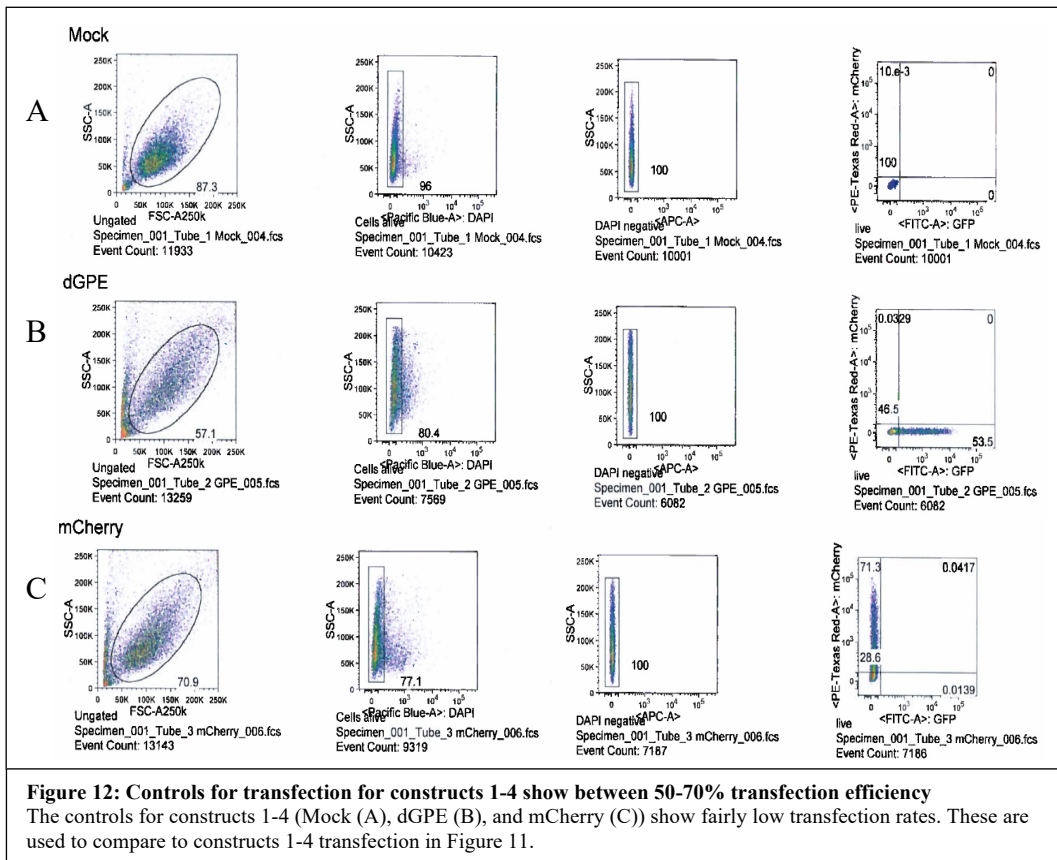quenced the clones to confirm correct ligation. Therefore, through diagnostic digests, gel analysis, and sequencing, we confirm that we have successfully cloned the 8 synthetic 5' donor leader regions into our dual report VT1. These sequences are stored in glycerol stocks and can be used
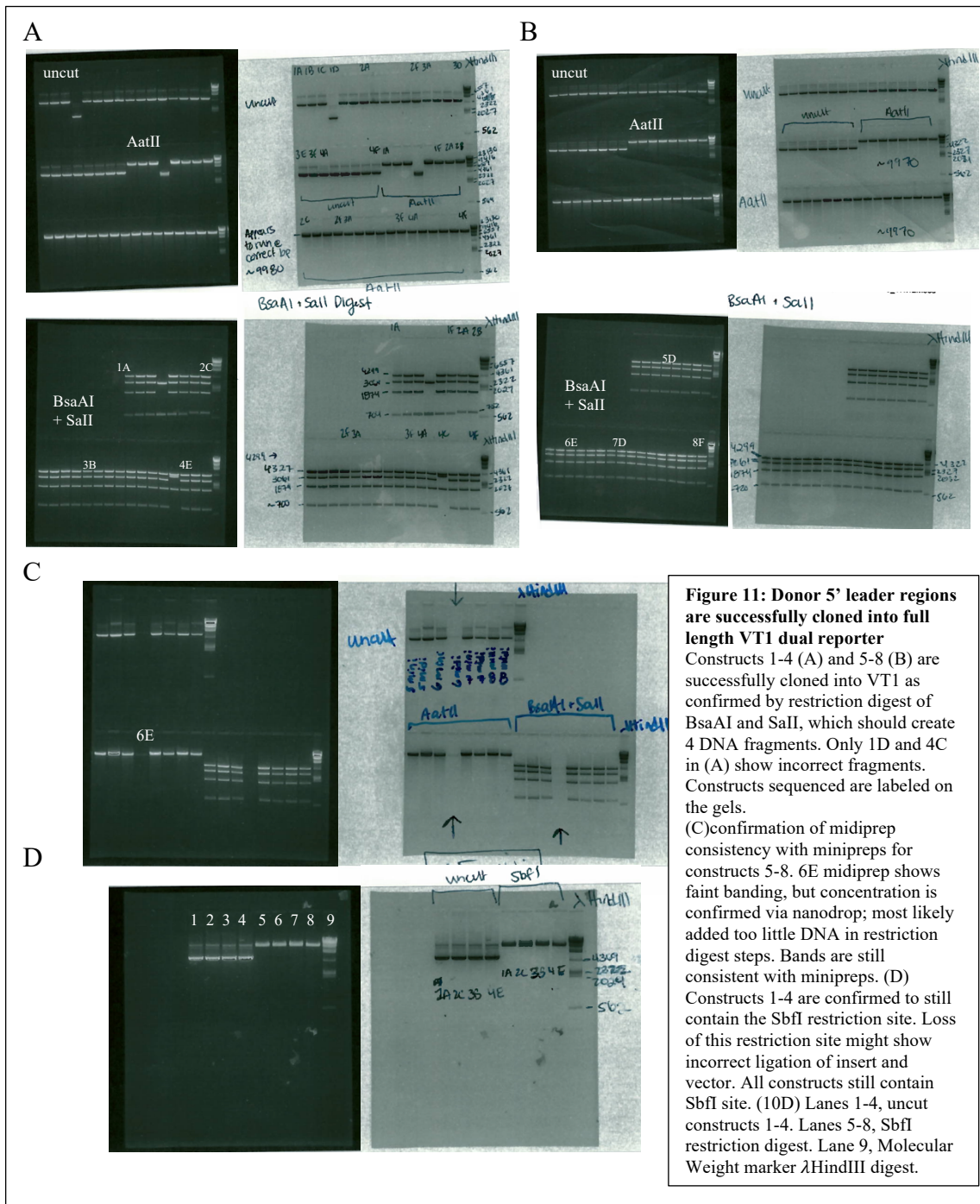
Figure 11: Donor 5' leader regions are successfully cloned into full length VT1 dual reporter Constructs 1-4 (A) and 5-8 (B) are successfully cloned into VT1 as confirmed by restriction digest of BsaAI and SaII, which should create 4 DNA fragments. Only 1D and 4C in (A) show incorrect fragments. Constructs sequenced are labeled on the gels. (C)confirmation of midiprep consistency with minipreps for constructs 5-8. 6E midiprep shows faint banding, but concentration is confirmed via nanodrop; most likely added too little DNA in restriction digest steps. Bands are still consistent with minipreps. (D) Constructs 1-4 are confirmed to still contain the SbfI restriction site. Loss of this restriction site might show incorrect ligation of insert and vector. All constructs still contain SbfI site. (10D) Lanes 1-4, uncut constructs 1-4. Lanes 5-8, SbfI restriction digest. Lane 9, Molecular Weight marker λHindIII digest.

in future investigations of the 5' leader region's role in latency and infectivity using the dual reporter vector.

44

B. Constructs Show that Recovery of Activity occurs when Helper containing Tat is provided in 293T Cells
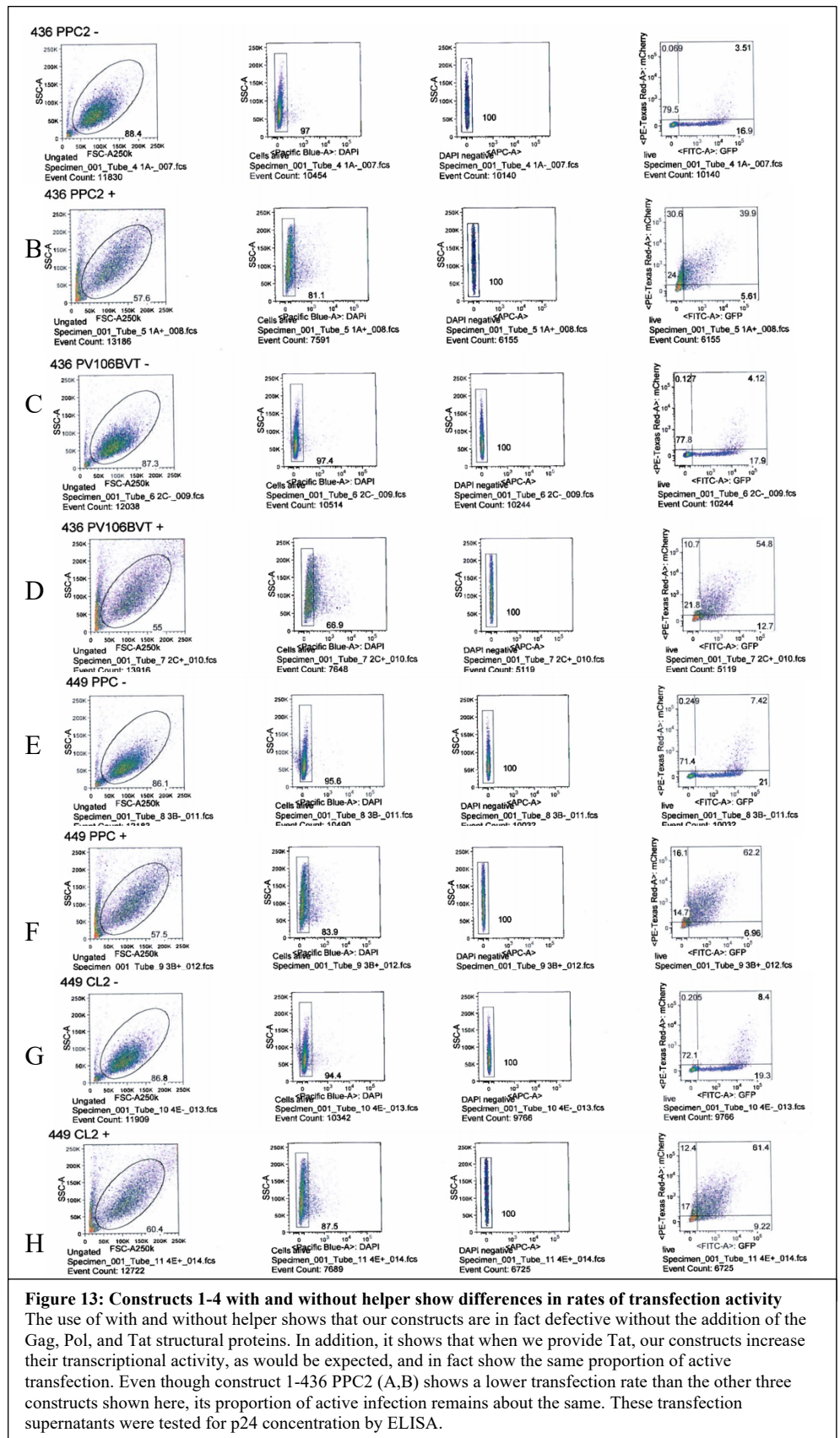
Transfection of 293T cells allows us to create the viral particles used to infect CEM-SS cells in order to study latency in packaged viral constructs. In addition, it allows us to quantify the expression of the dual reporters, and thus the expression of HIV genes in normal full-length donor sequences, in human cell lines. We utilized Mock and VT1 as controls, as well as the single color reporter plasmids dGPE, which s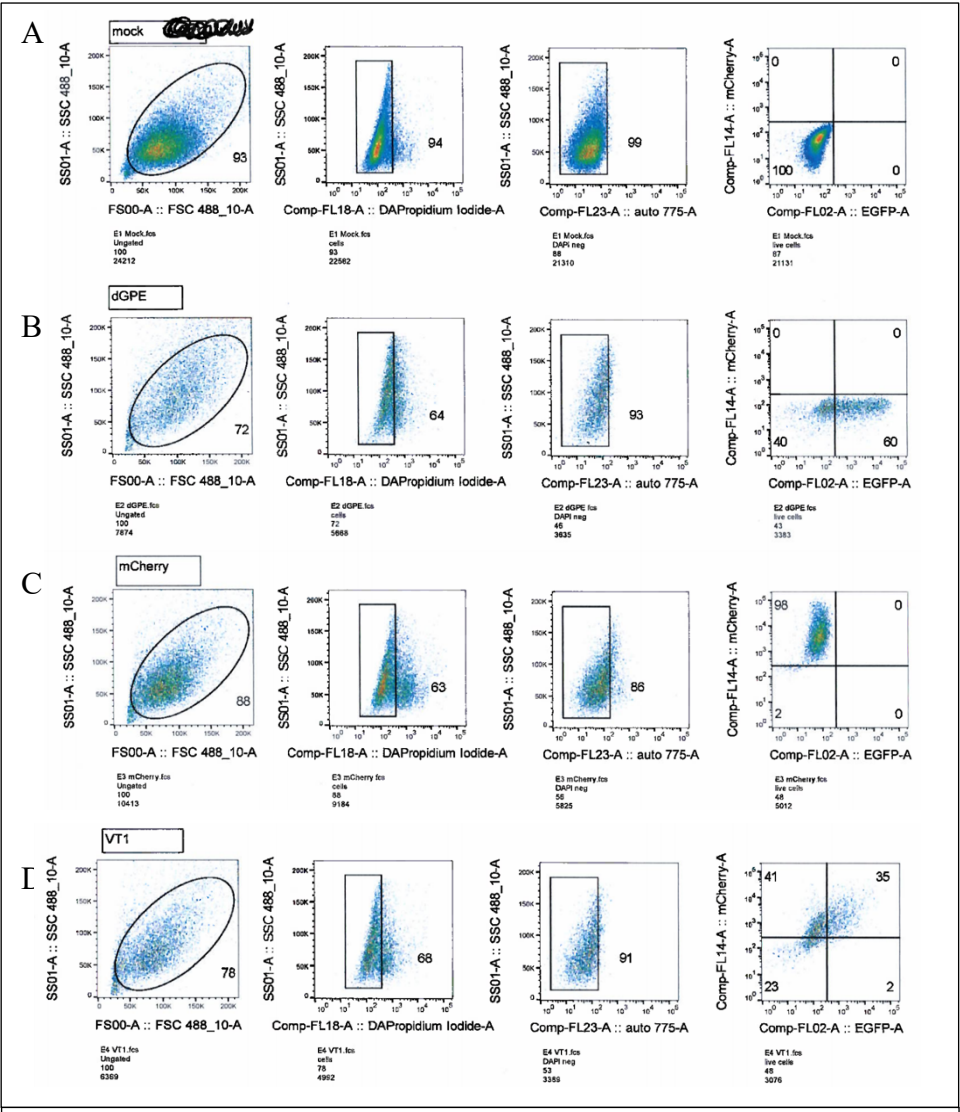hows only GFP expression, and mCherry, showing only mCherry reporter expression, as compensation controls (Fig. 12a-c, 14a-d); transfections and infections were performed in 2 rounds due to the number of constructs so we had controls for each experiment (12a-c, 14a-c).

Figure 12: Controls for transfection for constructs 1-4 show between 50-70% transfection efficiency
The controls for constructs 1-4 (Mock (A), dGPE (B), and mCherry (C)) show fairly low transfection rates. These are used to compare to constructs 1-4 transfection in Figure 11.

Transfections can help establish that our leader sequences are functional at expressing the viral particles; expression of mCherry in transfected cells verifies the functionality of the donor leader sequences. Different trends in rates of transfection are seen throughout the construct sequences

which can show that changes in sequences may affect transfection efficiency, viral gene expression, DNA plasmid quality, and even accuracy of concentrations used in transfection (Fig. 13a-h, 15a-h). However, the proportion of active transfection remains about the same between constructs containing helper (Table 1). These transfections were performed with and without helper in order to ensure that our constructs were able to express the reporters and viable to produce virus in the presence of helper containing their missing structural proteins Gag, Pol, and Tat (Fig. 13a-h, 15a-h).

Figure 13: Constructs 1-4 with and without helper show differences in rates of transfection activity
The use of with and without helper shows that our constructs are in fact defective without the addition of the Gag, Pol, and Tat structural proteins. In addition, it shows that when we provide Tat, our constructs increase their transcriptional activity, as would be expected, and in fact show the same proportion of active transfection. Even though construct 1-436 PPC2 (A,B) shows a lower transfection rate than the other three constructs shown here, its proportion of active infection remains about the same. These transfection supernatants were tested for p24 concentration by ELISA.

Interestingly, many of the constructs, including the parental plasmid VT1 showed an increase in mCherry expression when co-transfected with helper. As the helper plasmid helps to encode Tat, we hypothesize, as expected, that use of the helper lead to an increase in transcriptional activity, which lead to an increase in expression of mCherry. Some of our cells showed only an mCherry singly positive population, which may show that overexpression of Tat lead to mCherry expression to a greater extent than the constitutively expressed SFFV promoter of GFP. This could show that expression of the leader region of VT1 may affect the SFFV promoter or GFP expression or, more likely, GFP is just dim in these cells, and hard to detect. It is unclear what this means for our constructs, and future investigation may be needed in order to elucidate how there is a single mCherry positive

Figure 14: Controls for constructs 5-8 show higher rates of transfection
We added the VT1 wild type (WT) parent plasmid control to our experiment in order to see the transfection rate without the donor 5' leader inserts and how well expression of reporters in the parental strain is compared to donor constructs. We saw in this transfection that controls had a higher transfection efficiency than seen in the controls for constructs 1-4.

population without GFP expression as GFP should be constitutively expressed in transfected or infected cells.

We repeated the transfection of 293T cells by all constructs. Once again, we saw high mCherry expression in each construct containing helper Tat. Just the mCherry vs GPE populations are shown in Figure 16 and 17. As described in Table 1, the proportion of active transfection in the constructs in comparison to the wild type (WT) VT1 parental plasmid when Tat is provided is around the same. This is expected because the loss of the MSD means transcription of the HIV genome, quantified by mCherry expression, would be low; but when we provide Tat, we should see an increase in the activity of our reporter to the same extent in each construct.
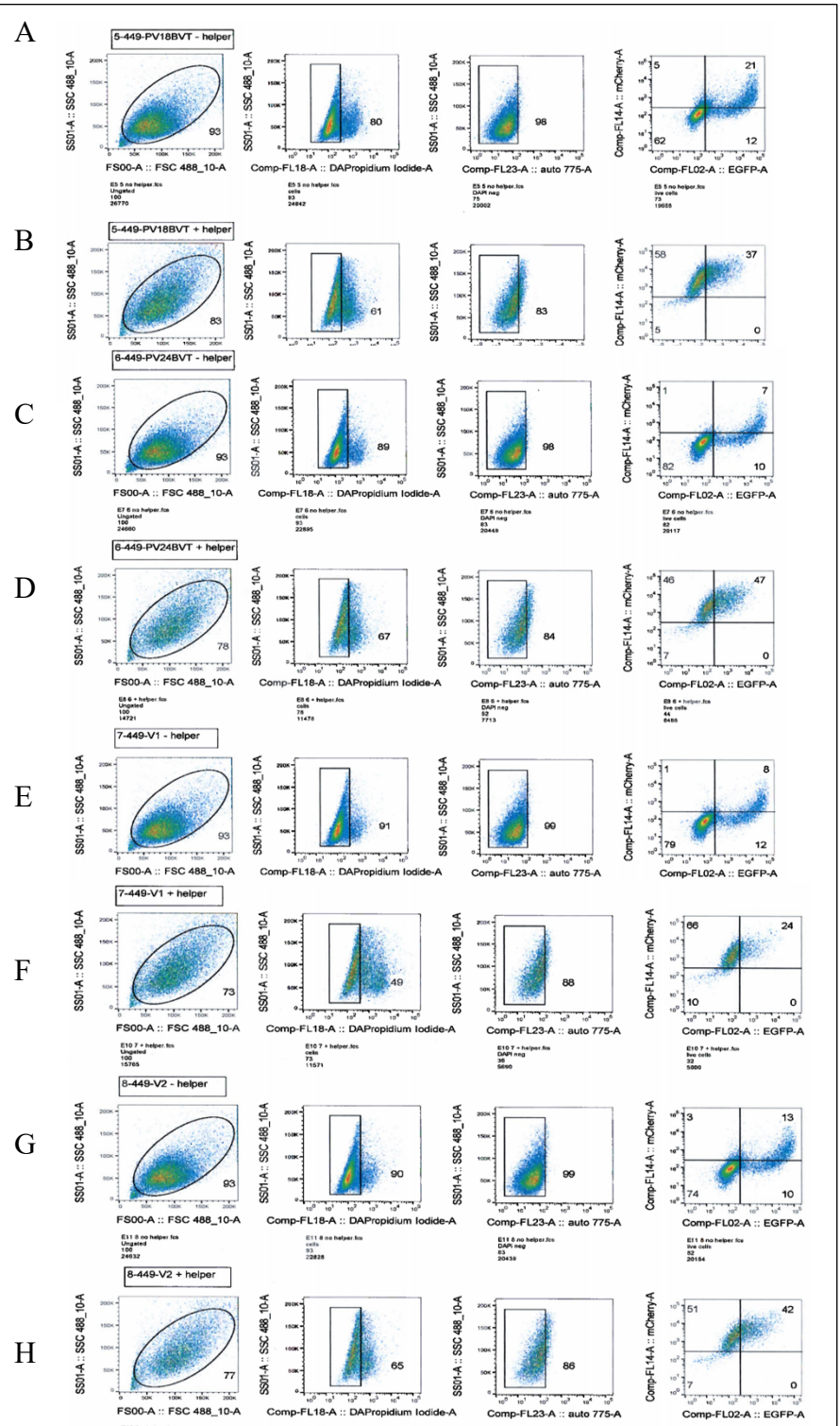
Figure 15: Providing helper during transfection leads to an increase in activity
The proportion of active transfection in constructs 5-8 is the same; all constructs that were positive for GFP were also positive for mCherry in the presence of helper. This is expected because the MSD site deletion does not allow Tat to be spliced, unless there is an alternative splice site. When we provide Tat, transcription should be activated to the same extend in all constructs as differences in sequences are seen in stem loops important for packaging and may not affect transcription.

Therefore, our sequence differences were not as important in the transfection experiments because we provided missing elements such as Tat. The sequence differences may cause differences in packaging of the gen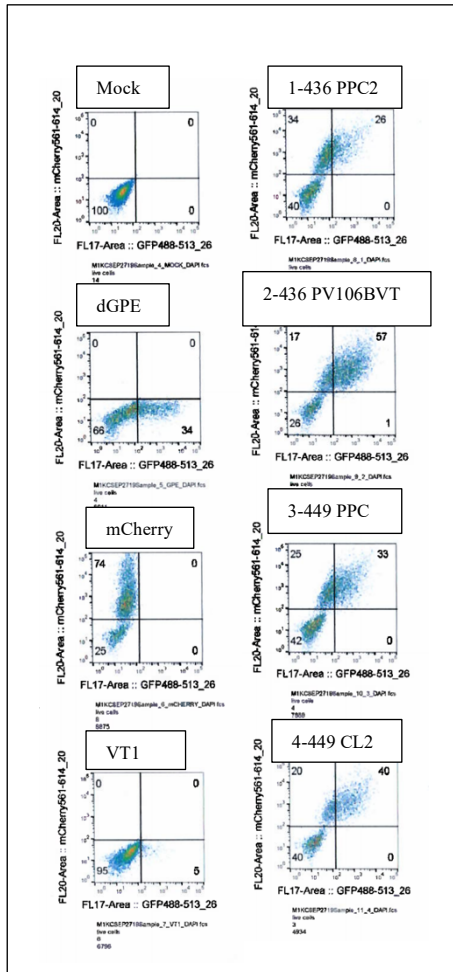ome, which we can ob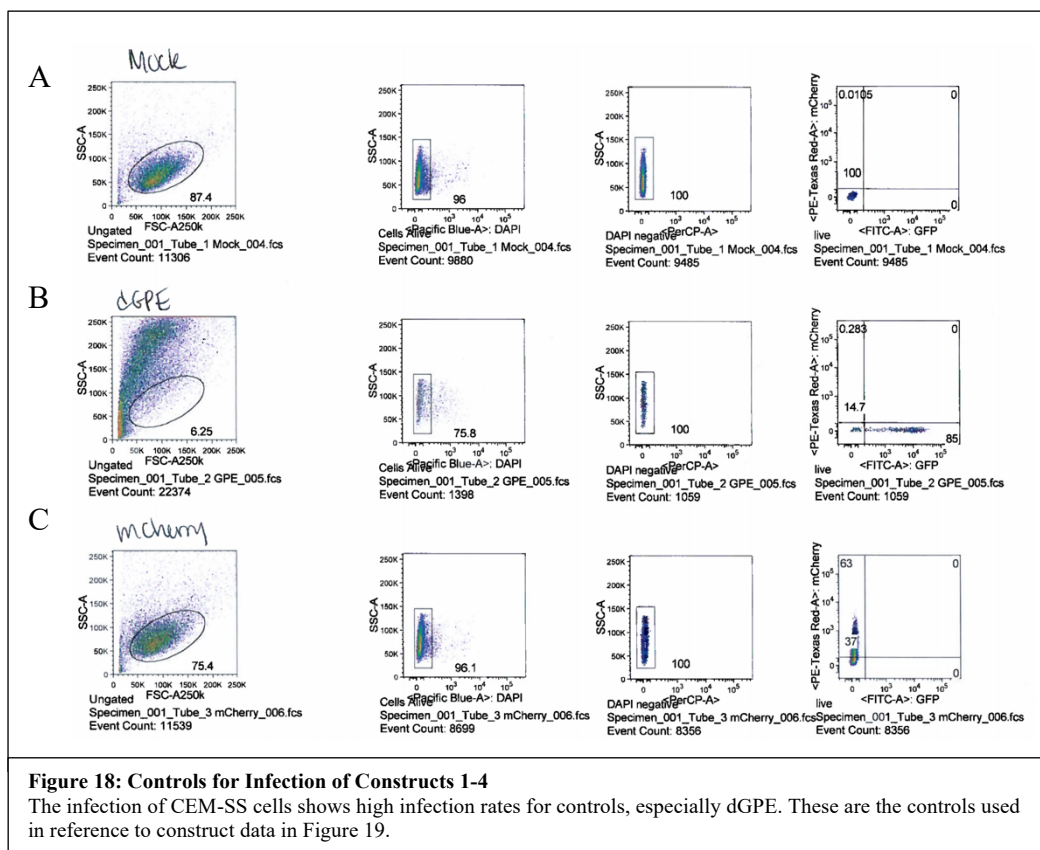serve in the infection experiments (see below). It is hard to determine why the expression is only singly positive and not both mCherry and GFP positive, although we hypothesize some of the single mCherry positive cells may actually be double negative if the gating was changed or double positive if GFP is too dim for detection but is actually expressed. Overall, the ability for each construct to express mCherry implies that they

Figure 16: Repeat Transfection of 293T Constructs 1-4
Our VT1 parental plasmid control did not work, but because other controls were successful, there may have been human error in transfection of VT1. We see once again that proportion of active transection is around the same across constructs when Tat helper is provided.
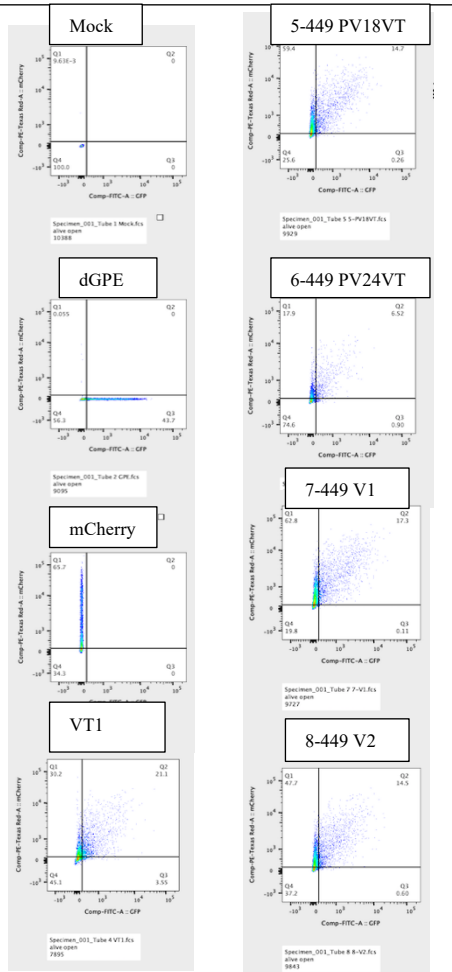
Figure 17: Repeat transfection of 293T Constructs 5-8
These transfection sups are utilized for qRT-PCR analysis of viral RNA concentration They show the same trends seen in the first transfection of constructs 5-8. We see that the percentrage of active transfection is more or less the same across constructs as well as WT VT1.

are functional at producing viral gene products, and thus the differences in active vs latent populations in CEM-SS infections can be attributed to sequence differences of SL1, 2, and 3, rather than a lack of functionality. The transfection experiment results show that with the use of
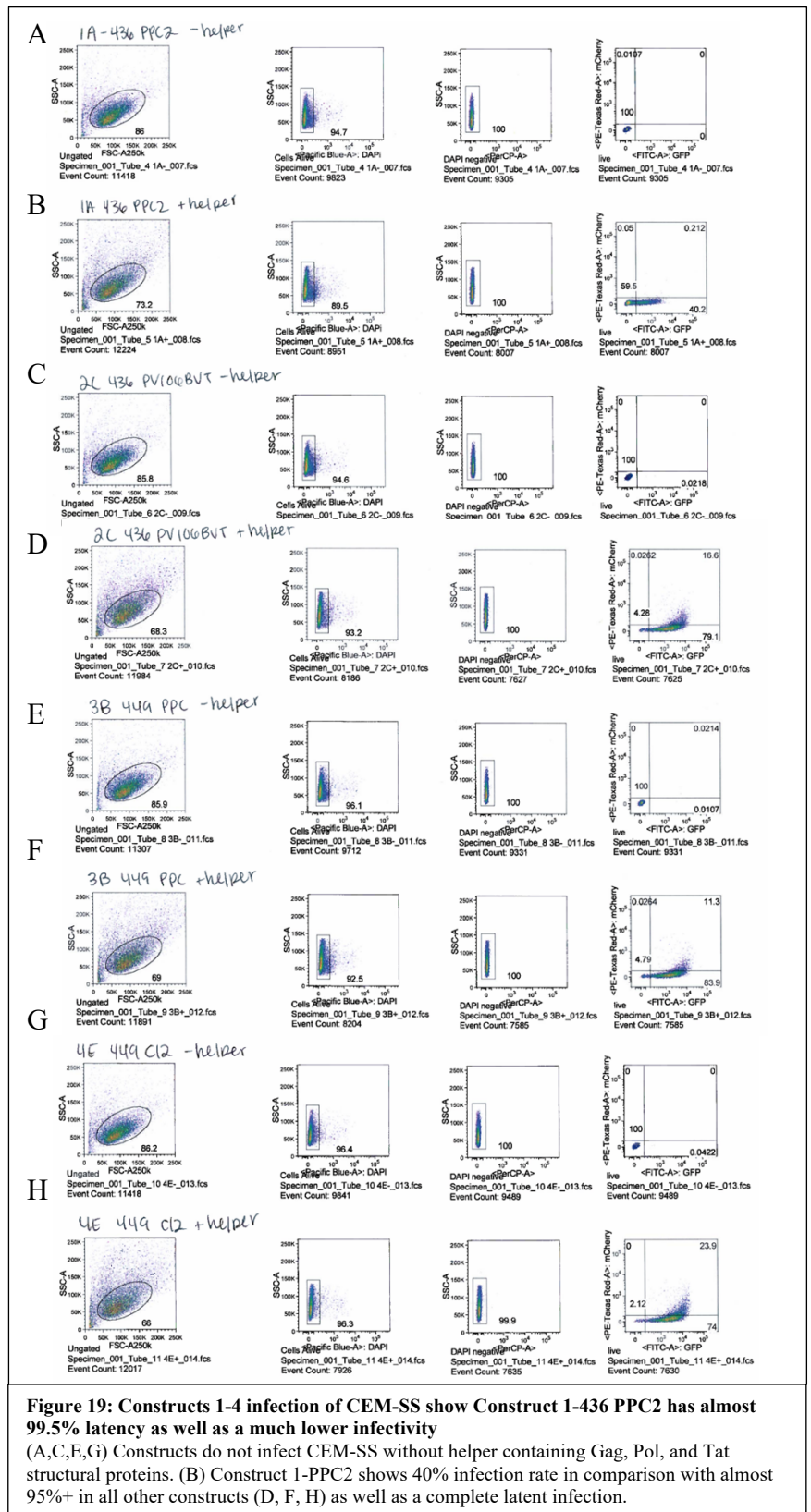
49

helper, the activity of the leader regions is higher, as would be expected. Future studies are needed to determine how in-vivo these constructs are able to produce Tat, as we clearly see that Tat is required for high transcription of these constructs. As we only studied the donor leader regions, later investigation of full-length sequences may be required to determine other splice sites needed Tat expression in these donors as the MSD is mutated or deleted.

C. Larger SL deletions lead to lower infectivity and higher latency

Figure 18: Controls for Infection of Constructs 1-4
The infection of CEM-SS cells shows high infection rates for controls, especially dGPE. These are the controls used in reference to construct data in Figure 19.

Infection of CEM-SS was performed in order to see the population of cells expressing just GFP, latently infected cells, versus those actively infected quantified by expression of mCherry and GFP. We took our viral supernatants from our transfection experiments for each

50

construct and infected CEM-SS cell lines in order to determine how different donor constructs viral particles infect human cell lines *in-vitro*. We performed infection experiments for all constructs, with and without helper (Fig. 19a-l, 21a-l), using the same controls as in the transfection experiments (Fig. 18a-c, 20a-d). Without helper, we saw no infection of CEM-SS as expected, as our constructs lack the structural proteins Gag and Pol. Constructs with the largest deletions that spanned into either SL1 or SL3, constructs 1-436 PPC2, 5-449 PV18VT and 7-449 V1, all showed the highest percentages of latency, or only GFP positive, in comparison with other constructs (Fig. 19b, 21b, 21f). In addition, these constructs showed the lowest rates of infection. When comparing these constructs

Figure 19: Constructs 1-4 infection of CEM-SS show Construct 1-436 PPC2 has almost 99.5% latency as well as a much lower infectivity
(A,C,E,G) Constructs do not infect CEM-SS without helper containing Gag, Pol, and Tat structural proteins. (B) Construct 1-PPC2 shows 40% infection rate in comparison with almost 95%+ in all other constructs (D, F, H) as well as a complete latent infection.

back to their transfection data, it appears that these three constructs did not show differences in

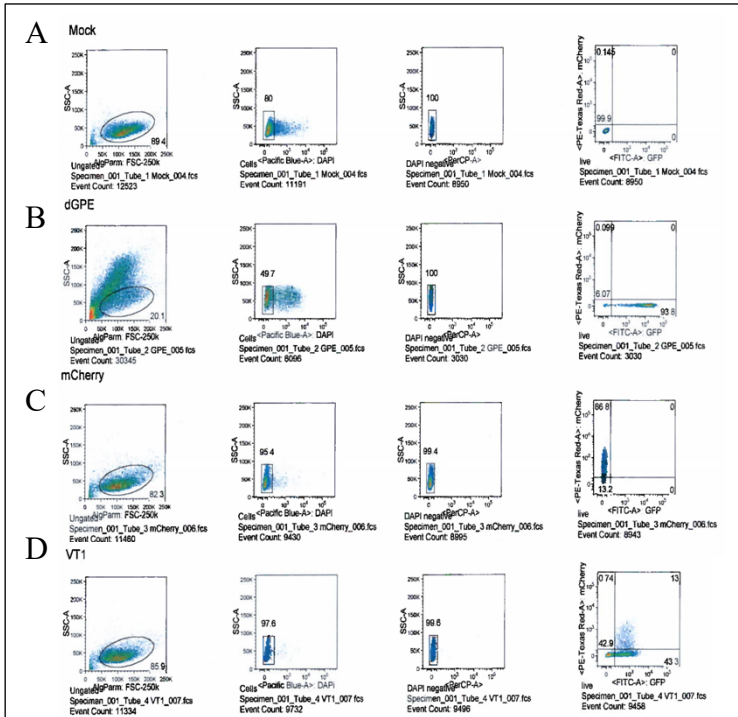proportion of active transfection when

Figure 20: Controls for Construct 5-8 Infection
Our single positive controls show very high rates of infection, while VT1 shows around 57% infection rate, and only 23% active infection.

provided Tat. (Table 1). As expected, because

these constructs contain deletions in more

SLs, those regions required for packaging of

the genome, we wouldn't expect a lower

transcription rate when provided Tat, but we

would expect a lower rate of infectivity as

seen.

Thus, a possible conclusion is that the

deletion of an extra stem loop yields the
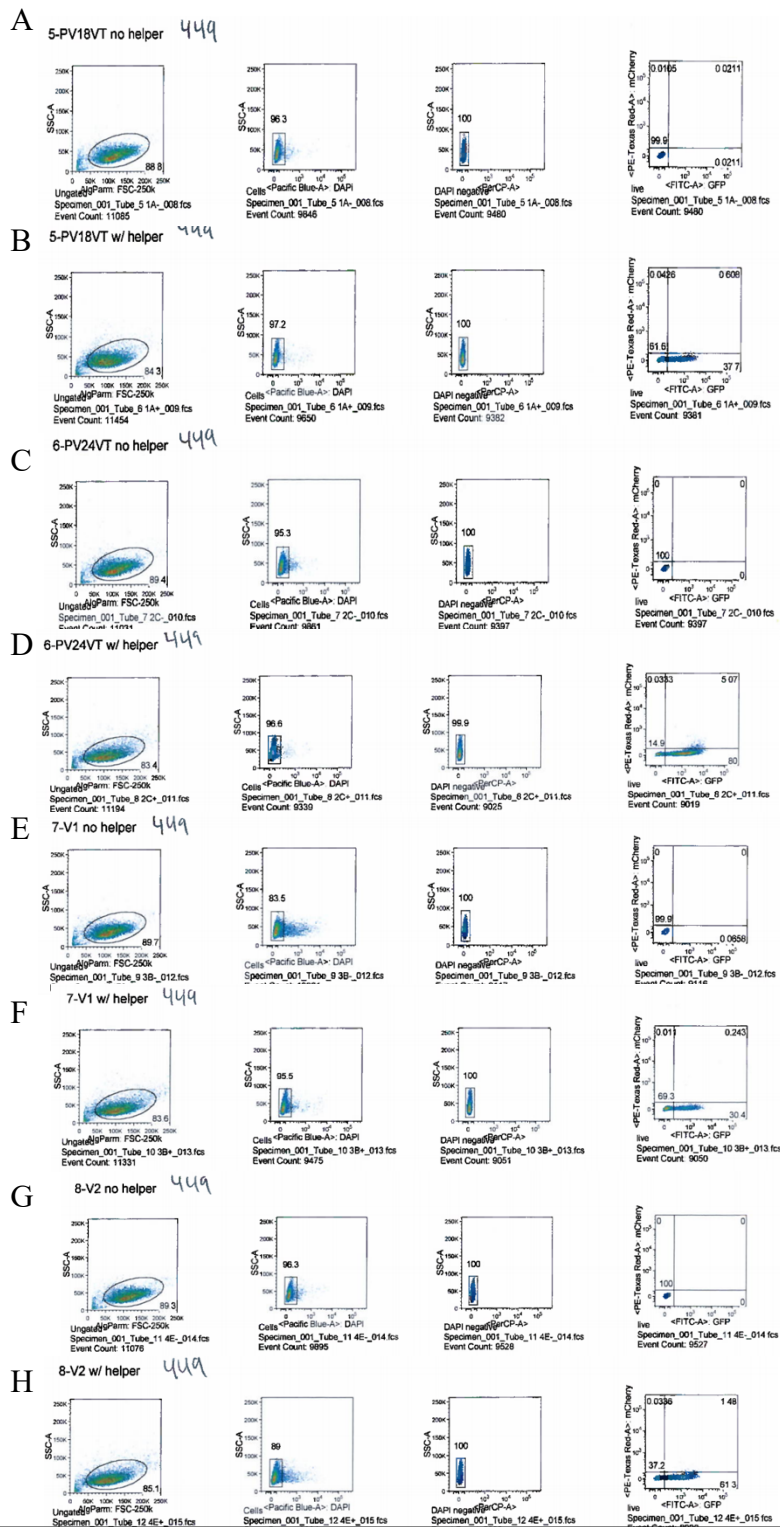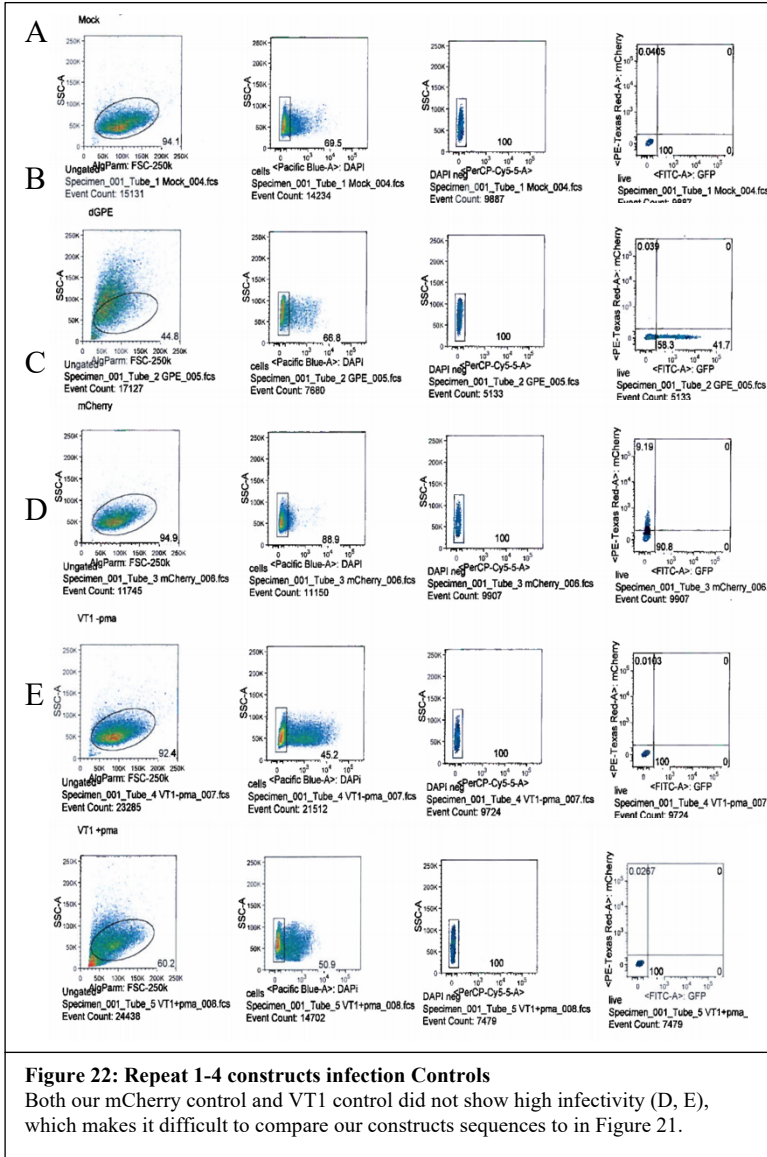
constructs less infectious, which also affects

Figure 21: Infection for Constructs 5-8
Constructs 5 and 7 show lower infection rate, and almost completely latent infection in comparison (B, F) with the other two constructs (D, H). As seen in constructs 1-4, infection does not occur without helper.

Figure 22: Repeat 1-4 constructs infection Controls
Both our mCherry control and VT1 control did not show high infectivity (D, E), which makes it difficult to compare our constructs sequences to in Figure 21.

their rate of active infection. We repeated the first four constructs but due to timing were unable to perform repeats for constructs 5-8. In this repeat, we included a pma/ionomycin stimulation at day 2, which was supposed to show that addition of drug treatment could stimulate the leader regions to be more active. However, our constructs did not show an increase in activity but rather showed a large loss of cell viability due to drug stimulation, which resulted in lower percentages of infectivity (Fig 23a-h). In the future, we would hope to repeat the infection experiments for constructs 5-8 as well as 1-4 again, using a lower concentra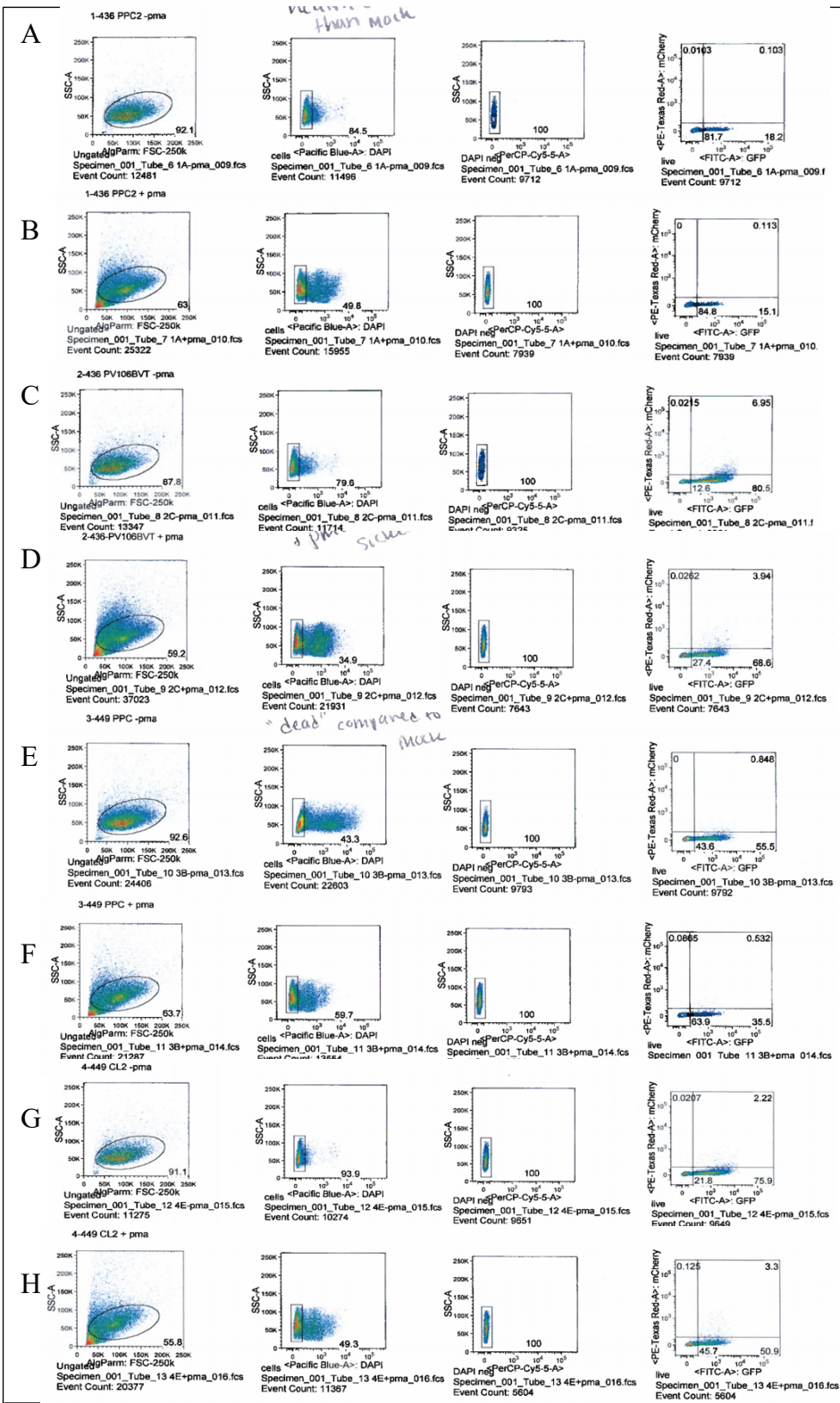tion of pma/ionomycin mentioned below, in order to support our initial transfection evidence that our leader sequences are functional and can be stimulated, and when they show latent infection they are in fact latent rather than just dysfunctional. This would help support our conclusions that when 2 stem loops are deleted in our 5' leader region sequences there is a lower rate of infection, and most infection is latent. A

Figure 23: Repeat of construct 1-4 Infection with PMA/ionomycin stimulation
PMA/ionomycin stimulation lead to a loss of viability in cells, leading to a decrease in infectivity. Therefore, we could not quantify whether our leader sequences could be stimulated by drug treatment at this concentration. Similar trends to the previous construct 1-4 infection show 1-PPC2 infects at the lowest rate with the highest percentage of latent infection in the pma absent infection.

summary of transfection and infection data of interest is in Table 1, showing that construct 1, 5, and 7, those that contain the largest deletions, also have the lowest rates of infection and highest rates of latency. Unfortunately, we did not use a WT VT1 control in our first transfection and infection experiments for constructs 1-4, and for our repeat of these constructs, our VT1 control was not viable. Therefore, it is difficult to compare to WT the

54

infection data for constructs 1-4 to see if their infection rate is in fact lower than WT.

However, we suggest that the trends we see for constructs 1-4 may still demonstrate that when there is a deletion that is only in one stem loop, the rate of infection is higher in comparison to when the deletion spans across more than one stem loop. As seen in Table 1 for constructs 1-4, constructs 2 and 4 have the highest active infection. These constructs only contain point mutations in the MSD, and thus, as we would expect, would have higher activity in comparison to construct 3 that has a deletion spanning across SL2, and even higher than construct 1 that has a deletion that spans from SL2 into SL1. We even hypothesize that the point mutation constructs may be able to produce some Tat expression on their own that could drive their higher activity and that the deletion of more than one stem loop affects the ability to package the HIV genome and thus decreases infectivity.

In constructs 5-8 we saw that in comparison to WT VT1, infection rate was lower for constructs 5 and 7 and higher for constructs 6 and 8. Constructs 5 and 7 contain a deletion that spans from SL2's MSD into SL3, whereas constructs 6 and 8 contain a deletion only in SL2. Therefore, we hypothesize that a deletion that spans into more than one stem loop, regions that are important for packaging, decreases the infectivity rate lower than WT, whereas if just the MSD is deleted, packaging of the virion may still be viable, as infectivity higher than WT is seen. In general, across all constructs, we see that when total infection rate is lower, the proportion of active infection also decreases. This has been shown in previous investigations by our lab, and thus helps support those conclusions that lower infectivity can lead to higher latency. More experiments are required to see if the rates of active infection, when infection rate is equalized, remains the same as the trends observed in these experiments.

55

Summary Transfection	Total transfection rate (GFP positive)		Percent Active Transfection (Double positive GFP, mCherry)	
Construct	Trial 1	2	1	2
1-436 PPC2	45.5	26.0	87.8	1
2-436 PV106BVT	67.5	57.0	81.2	1
3-449 PPC	69.2	33.0	89.9	1
4-449 CL2	70.6	40.0	86.9	1
WT VT1		0*		0*
5-449 PV18BVT	37.0	15.0	100	98.3
6-449 PV24BVT	47.0	7.4	100	87.9
7-449 V1	24.0	17.4	100	99.4
8-449 V2	42.0	15.1	100	96.0
WT VT1	37	24.7	94.6	85.6

Summary Infection	Total Infection rate (GFP positive)		Percent Active Infection (Double positive GFP, mCherry)	
Construct	Trial 1	2	1	2
1-436 PPC2	40.4	15.2	0.52	0.74
2-436 PV106BVT	95.7	72.5	17.35	5.43
3-449 PPC	95.2	36.0	11.87	1.48
4-449 CL2	97.9	54.2	24.41	6.09
WT VT1		0*		0*
5-449 PV18BVT	38.3		1.59	
6-449 PV24BVT	85.1		5.96	
7-449 V1	30.6		0.79	
8-449 V2	62.8		2.36	
WT VT1	56.3		23.1	

Table 1: Summary of Important Transfection and Infection Data is shown
Constructs 1, 5, and 7 show trends of lower transfection rate, lower infection rate and lower active infection in trials 1 and 2. For transfection data, we see that the proportion of active transfection when helper is provided is more or less the same across constructs, as expected, even if the total rate of transfection differs; the differences in LTR sequences for constructs are in sites important for packaging (SL1 and SL3) not transcription, so if we provide Tat needed for transcriptional activation, we should see activation in all constructs. For infection experiments, Tat may not be packaged within the virions because of the deletion of the MSD site, so we see differences between activity in constructs. In addition, the differences in deletion size between constructs can help explain the differences in infectivity that we see. The three constructs with the lowest active infection are constructs 1, 5 and 7. Construct 1 contains a large deletion spanning into Stem loop 1, and constructs 5 and 7 have deletions spanning into stem loop 3; these deletions are larger than the other constructs which only have deletions or mutations in stem loop 2. The most active constructs appear to be constructs 2 and 4, which only contain a point mutation in the MSD, and thus may even be able to express some Tat from the MSD on their own.
Unfortunately, VT1 controls were not viable in our construct 1-4 experiments, and thus we are unable to ensure that our rate of infectivity and active infection was the same as WT, and in fact did differ between constructs. In our infection for constructs 5-8, we see that deletions in more than one stem loop, SL2 to SL3 which are present in constructs 5 and 7, lead to a lower infection rate and lower activity rate. Although the other constructs 6 and 8 did not show as much activity as WT, it appears that they still shower a higher activity than 5 and 7, as would be expected because 6 and 8 only contain a deletion in SL2, not spanning into more than one SL.

D. qRT-PCR and ELISA Demonstrate Different RNA Packaging and Release into Medium and Protein Production

Construct	Average ng/mL p24
Mock 1	0
dGPE 1	433,877
mCherry 1	329,237
1-PPC2	364,768
2-PV106BVT	385,525
3-PPC	376,551
4-Cl2	352,558
Mock 2	0
dGPE 2	557,976
mCherry 2	467,063
VT1	457,657
5-PV18VT	495,171
6-PV24VT	461,452
7-V1	449,131
8-V2	313,047

Table 2: ELISA data shows no significant explanation for lower infectivity rates of constructs
ELISA data for all 8 constructs, including each control for infection experiments 1-4 and for 5-8 are shown to the left. Control "1" denotes for constructs 1-4 and "2" denotes constructs 5-8. 10-fold serial dilutions were performed in order to ensure that accurate measurements for protein concentration were determined. Some of the higher concentration dilutions were omitted from average calculations. These may have reached saturation prior to collection, meaning our concentration estimate would be undervalued. We noticed that the concentration of viral proteins in the viral supernatants was across the board similar between constructs, and thus cannot explain the differences in infectivity.

A possible explanation for the difference in infection rates could be that the less infectious constructs had produced less viral particles during transfection, thus when infecting with the same volume of supernatant across constructs, the less prolific would not infect as well. In order to test this, we took the viral supernatants from the transfection experiments shown in Figure 13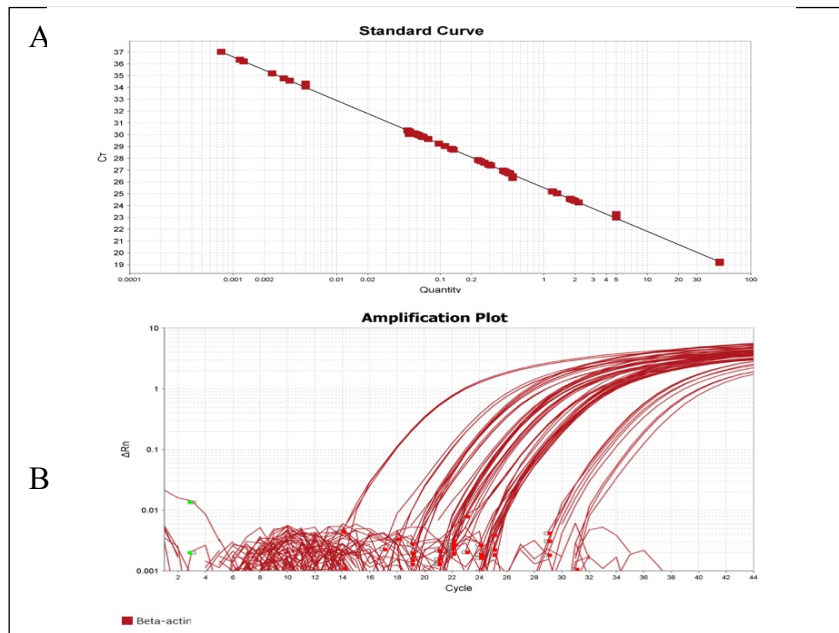 and 15 and performed ELISA in order to determine the concentration of p24 present in the supernatant (Table 2). We determined that there appeared to be no differences in the amount of p24, a protein that makes up the capsid protein of HIV, within each transfection supernatant and

Figure 24: Raw Data for Beta-Actin Assay of Viral Supernatants
The standard curve (A) made from Raji Control RNA shows where samples fall onto the standard curve line. The amplification plot data demonstrates varying levels of RNA concentration.

57

thus a lack of virion particles used in infection cannot explain the difference in rates of infection; the constructs contained more or less equal amounts of p24 in the viral supernatants. These same viral supernatants were used in the infection experiments shown in Figures 19 and 21, thus our differences in infectivity cannot be attributed to differences in number of viral particles in the supernatant.

We then wondered whether the differences in infection could be based on a difference in amount of HIV RNA packaged into virions. In order to test this, we took the transfection supernatants from transfections in Figure 16 and a transfection experiment repeat of constructs 5-8 (data not shown) and extracted RNA from each sample. We then synthesized cDNA from the RNA extracts and performed qRT-PCR in order to quantify the genomic concentrations of each transfection (Fig.

Figure 25: HIV qRT-PCR results of Viral Supernatants
The standard curve made of HIV standard (A) shows a high R^2 value as seen in the Beta-actin assay, as well as where samples fall onto the curve. The amplification plot shows varying data; the most right amplification curves are Mock cells, and would be expected to have no HIV detectable.

24a-b, Fig 25a-b, Table 3). Our goals were to use the genomic concentration data in order to infect CEM-SS cells with an equal concentration of HIV RNA rather than an equal volume of viral supernatant as we had in previous infection experiments. However, due to timing, we were unable to complete this final CEM-SS infection and hope to determine the role of RNA concentration in rates of infectivity and how it affects rates of latency in future research. These
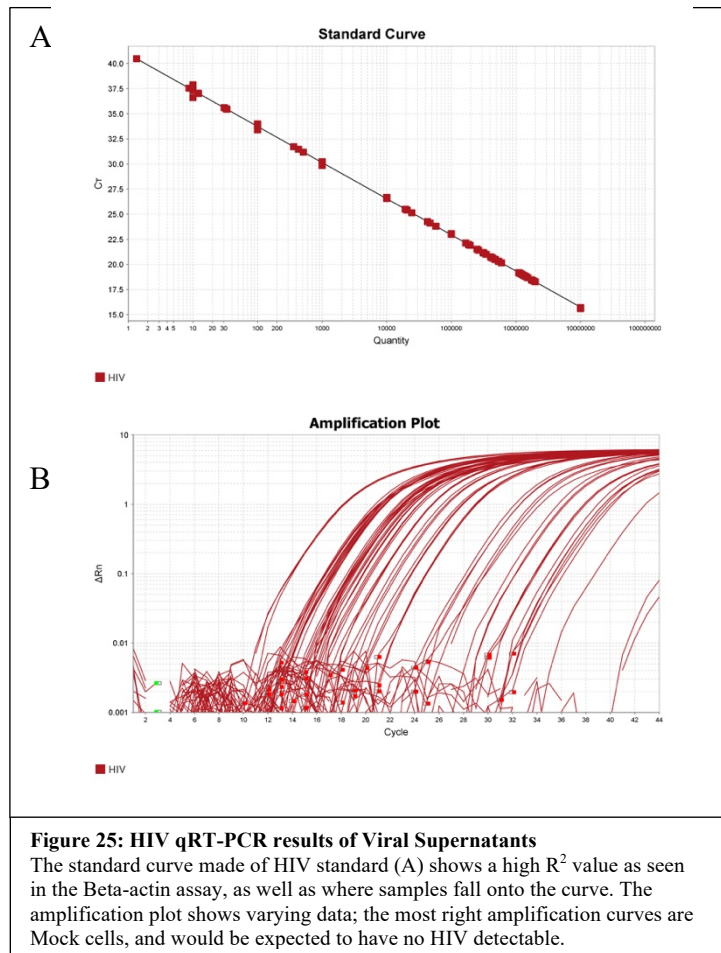
viral supernatants are not the same tested for ELISA, however, we hope that the trends observed in concentration of RNA would hold across all transfection supernatants.

We performed both a Beta-Actin Assay in our viral supernatants as a control in order to determine our RNA extraction efficiency's (Fig. 24a-b) as well as the experimental HIV targeted assay (Fig 25a-b). We calculated the quantity of RNA in each PCR reaction using the standard curves for Beta actin (Fig. 24a) and for HIV (Fig. 25a), and then used information on the dilutions we used, often a 1:10 dilution, and RNA extraction efficiency calculated using our control RNA in order to calculate the quantity of RNA in our viral supernatants; Results are shown in Table 3. The Beta actin assay demonstrated very poor RNA extraction, with the best RNA extraction of VT1 at 40% efficiency (Table 3). It is interesting that the RNA extraction for all constructs was so low and begs the question of whether data for HIV RNA quantity is accurate due to such a low extraction rate. Although our efficiency of extractions was so low that there may have been large variability, we still used the efficiency numbers to calculate HIV RNA copy number in the viral supernatants.

From our qRT-PCR experiment, it appears like there are no significant trends that show that our constructs that showed the lowest rate of infection and lowest proportion of active infection, constructs 1-436 PPC2, 5-449 PV18VT, and 7-449 V1, produced viral supernatant with a lower copy number of RNA genome, in either the raw data or data corrected for percent efficiency of RNA extraction. Although 1-436 PPC2 did contain a lower copy number than other constructs in its experimental cohort, constructs 5 and 7 did not show this trend. Repeating the RNA extraction and qRT-PCR may be necessary in order to determine trends in viral copy number in transfection supernatants, as well as using those number for infection of CEM-SS cells. Cycle threshold (CT) values in comparison with CT values of the standards are used to

59

calculate the concentration of cDNA in the PCR reactions. If we performed a 1:10 dilution on these samples, we then multiplied this concentration by 10 to get the concentration (or copy number) of the undiluted sample; these are shown in the left column of the Beta actin and HIV mRNA assay of Table 3.
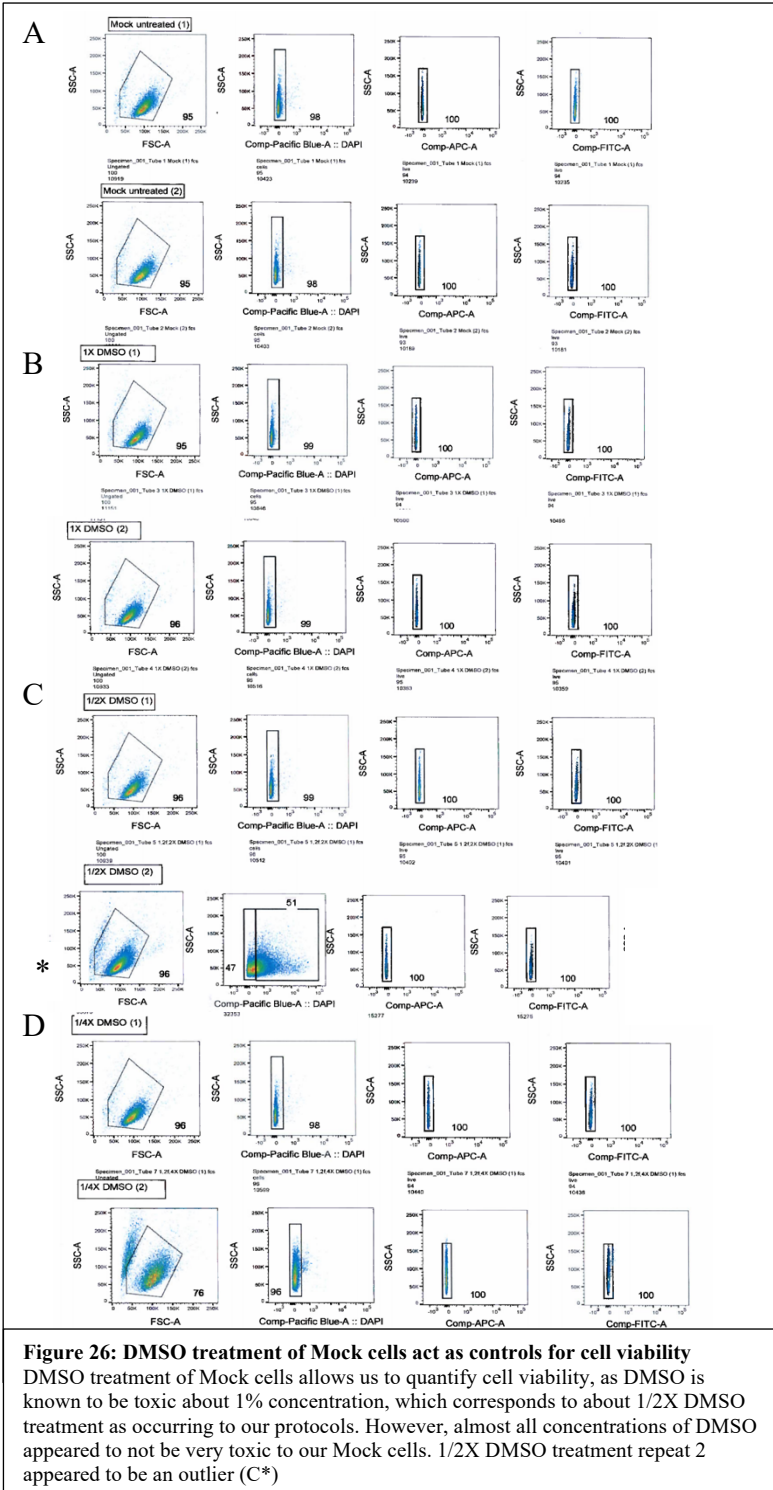
Although the RNA extraction was very inefficient, we also calculated the copy number for HIV mRNA using the percent efficiency of extraction determined by the Beta-Actin Assay. We took our copy number and divided it by the percent efficiency to get the expected total copy number per volume in our viral supernatants; this is shown in the right column of HIV mRNA Assay in Table 3. All values were in the dynamic range of the standard curve (1×10^7 cop/μl to 10 cop/μl), because we performed the 1:10 dilutions. The qRT-PCR data is suggestive that our conclusions that constructs with smaller deletions have a lower degree of latency is valid and is not because of a difference in HIV genome packaged. However, our data remains somewhat inconclusive due to low RNA extraction efficiency and must be repeated in future experiments.

60

Beta-Actin Assay	Average concentration of Beta-Actin (ng/μl)	Percent efficiency RNA extraction, (50 ng/μl expected)
Mock	0.61	1.22
dGPE	4.18	8.36
mCherry	2.99	5.98
VT1	2.49	4.98
1-436 PPC2	1.3	2.6
2-436 PV106BVT	0.52	1.04
3-449 PPC	12.57	25.14
4-449 CL2	0.01	0.02
Mock	4.65	9.3
dGPE	4.44	8.88
mCherry	18.16	36.32
VT1	20.29	40.58
5-449 PV18VT	0.03	0.06
6-449 PV24VT	0.85	1.7
7-449 V1	0.8	1.6
8-449 V2	0.67	1.34
HIV Assay	Average copy number HIV mRNA (cop/μl)	Copy number taking into account Percent efficiency (cop/μl)
Mock	0	0
dGPE	12,985,627	155,330,466.50
mCherry	110	1,839.50
VT1	3,545,050	71,185,743
1-436 PPC1	215,180	82,761,53.8
2-436 PV106BVT	2,784,869	267,775,865.40
3-449 PPC	4,616,749	18,364,156.70
4-449 CL2	4,684,601	23,423,005,000
Mock	0	0
dGPE	12,589,772	141,776,711.70
mCherry	4,348	11,971.40
VT1	15,581,428	38,396,816.20
5-449 PV18VT	17,352,727	28,921,211,667
6-449 PV24VT	491,804	28,929,647.10
7-449 V1	5,379,100	336,193,750
8-449 V2	1,829,612	136,538,209

Table 3: qRT-PCR results for transfection supernatants
There appear to be no significant trends in copy number per volume between constructs that can be releated back to their sequence differences or active infection rates. The RNA extractions across the board were very ineffective, so cop/ul was calculated from raw data and from using the RNA extraction efficiency. In both ways of calculating copy number, no trends can be seen that would explain the trends seen in lower infectivity of constructs 1, 5, and 7.

E. PMA/ionomycin Tests Provide Future Directions of Study

Figure 26: DMSO treatment of Mock cells act as controls for cell viability
DMSO treatment of Mock cells allows us to quantify cell viability, as DMSO is known to be toxic about 1% concentration, which corresponds to about 1/2X DMSO treatment as occurring to our protocols. However, almost all concentrations of DMSO appeared to not be very toxic to our Mock cells. 1/2X DMSO treatment repeat 2 appeared to be an outlier (C*)

We have demonstrated that our current pma/ionomycin concentration used for stimulation of infected CEM-SS cells by constructs 1-4 lead to a large decrease in cell viability and no extra mCherry expression (Fig. 23). We sought to determine what concentration of drug treatment would maintain cell viability and thus allow us to show that our donor leader regions could be stimulated to produce mCherry by pma/ionomycin in order to show functionality. We performed CEM-SS infection by Mock cells, followed by pma/ionomycin stimulation two days post infection using different concentrations of pma/ionomycin (discussed in detail in materials and methods). We used DMSO, another solvent known to have cell toxicity at a concentration greater than 1%, as a control (Fig 26a-d), although we saw that DMSO did not appear to be very toxic even above the 1% concentration; 1/2X DMSO was at a
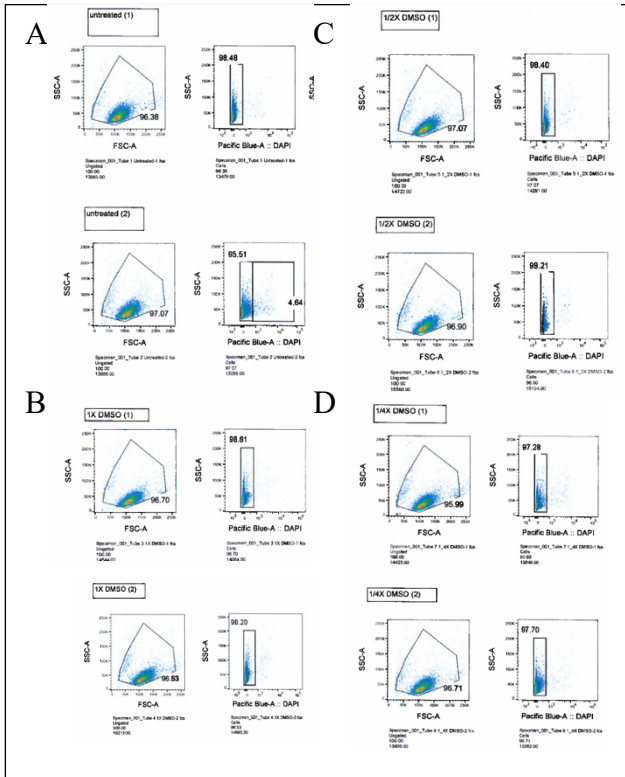
Figure 28: DMSO treatment for repeat pma/ionomycin titration
We see similar trends of high cell viability even with the control drug treatment.
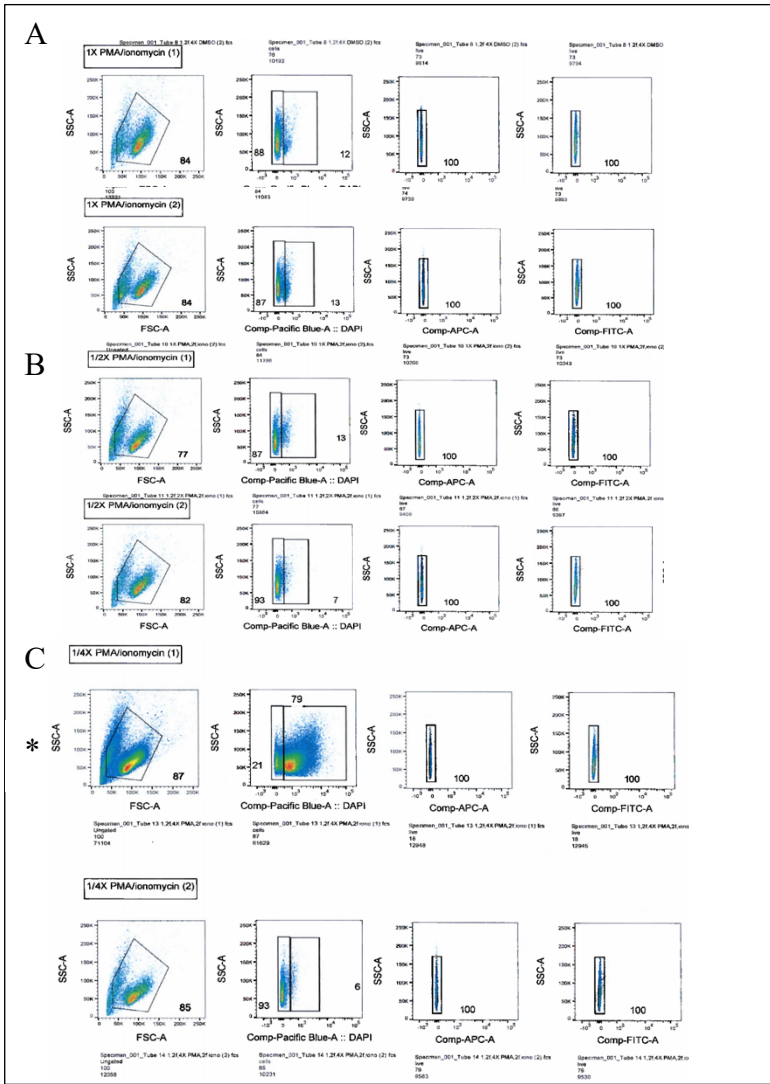

Figure 27: PMA/ionomycin titration shows cell viability increases with lower drug concentration
We determined cell viability using DAPI vs SSC gating shown in the second column. There also appears to be an outlier in the 1/4X PMA/ionomycin treatment (C*). These plots show the representative gating we used for the pma/ionomycin repeat as well.
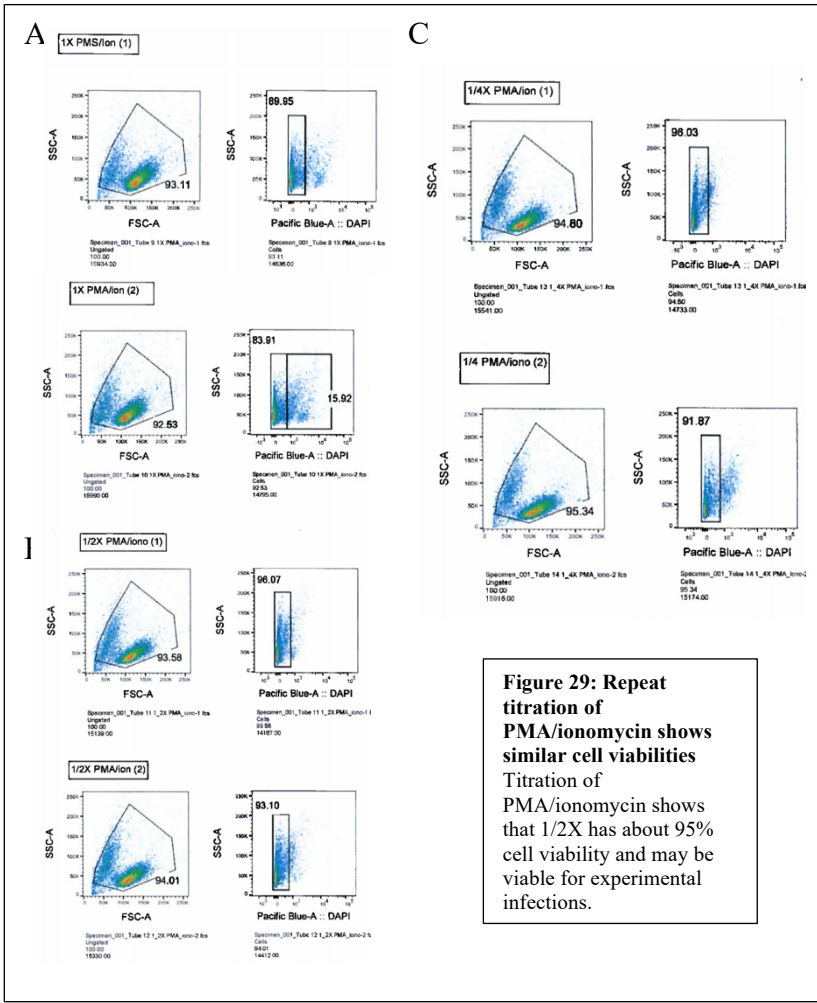
1.05% concentration. We saw that cell viability increased as concentration of pma/ionomycin decreased, as expected (Fig. 27a-c, Table 4). We also repeated the pma/ionomycin titration to confirm our results; Only the first two gates are shown that show cells and DAPI staining for dead vs alive cells for the repeat experiments in Figure 29a-c. From our data, it appears that using a concentration at 1/2 what we used before in our infection experiments increases cell viability to about 90% (see Materials and Methods for details on concentrations). We propose using a 1/2X PMA/ionomycin treatment, corresponding to 25 ng/mL PMA and 1.5 µg/mL ionomycin in future infection and drug stimulation experiments to determine functionality of our leader regions.

63

Figure 29: Repeat titration of PMA/ionomycin shows similar cell viabilities
Titration of PMA/ionomycin shows that 1/2X has about 95% cell viability and may be viable for experimental infections.

Table 4: Summary of PMA/ionomycin drug titrations
Use of 1/2X PMS/ionomycin concentration increased the cell viability, especially in the second trial. The numbers highlighted in yellow appear to be outliers, so they were not used in calculation of the averages, although they were used in standard deviation calculations. Future CEM-SS stimulation experiments should utilize 1% PMA and 0.5% ionomycin treatments. Refer to Materials and Methods for more information on concentration of drug treatments.

Take 1: Percentage Cell Viability							
	Mock untreated	1X DMSO	1/2X DMSO	1/4X DMSO	1X PMA/ionomycin	1/2X PMA/ionomycin	1/4X PMA/ionomycin
1	98	99	99	98	88	87	21
2	98	99	47	96	87	93	93
Avg	98	99	99	97	87.5	90	93
Standard Deviation	0	0	36.8	1.4	0.7	4.2	50.9
Take 2: Repeat Percentage Cell Viability							
	Mock untreated	1X DMSO	1/2X DMSO	1/4X DMSO	1X PMA/ionomycin	1/2X PMA/ionomycin	1/4X PMA/ionomycin
1	97	97	97	94	84	90	91
2	88	97	97	93	76	87	87
Avg	92.5	97	97	93.5	80	88.5	89
Standard Deviation	6.4	0	0	0.7	5.7	2.1	2.8

F. Final Conclusions

In conclusion, we have determined that substitution of donor LTR sequences with deletions in more than one stem loop were functional in expression of both mCherry and GFP reporters in transfection. We showed that our constructs, when Tat is provided, a protein that we would expect our constructs unable to make due to the MSD deletion or mutation, were able to increase activity; the constructs showed around equal proportion of active transfection. Therefore, the sequence differences in the packaging regions of the LTR appear to not affect rates of active transfection, only the fact that the MSD is deleted, and Tat is provided to activate transcription.

Therefore, our transfection data suggests that Tat is required for these constructs to have high activity, and when we do not provide Tat, our constructs cannot produce it off of the MSD. However, because we see these constructs expressed in-vivo, we propose that there may be a downstream splice site that allows for Tat expression. Future investigation of the role of Tat in MSD deleted or mutated constructs could involve adding a lentivirus that expressed Tat and a different color reporter during infection experiments. We would expect that if we provide Tat during infection to our constructs, the activity should increase as seen in transfection. Similarly, we could make full-length reporter constructs to see if the activity during infection increases in comparison to our current constructs, as would occur if there were a downstream splice site for Tat. Our constructs would not have included this possible downstream site, as they only contained the LTRs inserted into our dual reporter VT1.

For our infection experiments, we saw that when there was only a point mutation in the MSD, there was a higher degree of activity. In comparison, constructs with the lowest activity had deletions that spanned into more than one stem loop, and thus may have been unable to

65

package the genome correctly. We suggest that these differences in infectivity cannot be contributed to differences in virion particles or genomic concentration in the viral supernatants used to infect the human cell line, although these experiments need to be repeated to confirm this conclusion. We propose that there is a correlation between infectivity and latency in HIV-1, as suggested by other lab members work, where lower rates of infection are correlated with higher rates of latency. For future study, we are interested in determining what causes the correlation between infectivity and latency, as it is clear that an extra deletion in SL1 (construct 1) vs SL3 (constructs 5 and 7) does not affect the outcome, only the fact that there is a deletion in another SL besides the MSD in SL2.

This could be of interest in studying because the cure for HIV relies on diminishing the latent integrated population that is maintained in hosts on antiretroviral therapy; if future research could determine why the constructs with deletions in more than one stem loop are able to maintain their low infectivity rate, it could help elucidate mechanisms to how the latent population is maintained, and possibly how to eradicate it. A possible hypothesis for future study is that because the stem loops are required for assembly and packaging of the gRNA, when these are deleted or mutated, the virus is required to use alternative methods for packaging and assembly that yield the virus less infectious. We propose a future infection experiment that equalizes out the rate of infection, hopefully be infecting with equal numbers of HIV genome, in order to see how the rates of active infection compare across constructs and between constructs with deletions in SL1 vs SL3.

66

V.   ACKNOWLEDGEMENTS

VI.   REFERENCES

1. Wolf D, Goff SP. Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature. 2009 Apr 30;458(7242):1201-4. doi: 10.1038/nature07844. Epub 2009 Mar 8. PMID: 19270682; PMCID: PMC2676211.

2. Carter CC, Onafuwa-Nuga A, McNamara LA, Riddell J 4th, Bixby D, Savona MR, Collins KL. HIV-1 infects multipotent progenitor cells causing cell death and establishing latent cellular reservoirs. Nat Med. 2010 Apr;16(4):446-51. doi: 10.1038/nm.2109. Epub 2010 Mar 7. PMID: 20208541; PMCID: PMC2892382.

3. Sebastian NT, Zaikos TD, Terry V, Taschuk F, McNamara LA, Onafuwa-Nuga A, Yucha R, Signer RAJ, Riddell J IV, Bixby D, Markowitz N, Morrison SJ, Collins KL. CD4 is expressed on a heterogeneous subset of hematopoietic progenitors, which persistently harbor CXCR4 and CCR5-tropic HIV proviral genomes in vivo. PLoS Pathog. 2017 Jul 21;13(7):e1006509. doi: 10.1371/journal.ppat.1006509. Erratum in: PLoS Pathog. 2017 Sep 13;13(9):e1006617. PMID: 28732051; PMCID: PMC5540617.

4. Sharpe AH, Hunter JJ, Ruprecht RM, Jaenisch R. Maternal transmission of retroviral disease and strategies for preventing infection of the neonate. J Virol. 1989 Mar;63(3):1049-53. doi: 10.1128/JVI.63.3.1049-1053.1989. PMID: 2915376; PMCID: PMC247797.

5. Painter, Mark. "HIV and Retroviruses." *Encyclopedia of Microbiology*, edited by Kathleen Collins, Elsevier, 2019, pp. 613–628.

6. Wang GZ, Wang Y, Goff SP. Histones Are Rapidly Loaded onto Unintegrated Retroviral DNAs Soon after Nuclear Entry. Cell Host Microbe. 2016 Dec 14;20(6):798-809. doi: 10.1016/j.chom.2016.10.009. Epub 2016 Nov 17. PMID: 27866901; PMCID: PMC5159289.

7. Grandgenett DP, Aihara H. Oligomerization of Retrovirus Integrases. Subcell Biochem. 2018;88:211-243. doi: 10.1007/978-981-10-8456-0_10. PMID: 29900499.

8. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, Niedziela-Majka A, Stray KM, Sakowicz R, Cihlar T. Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. PLoS One. 2013 Sep 9;8(9):e74163. doi: 10.1371/journal.pone.0074163. PMID: 24040198; PMCID: PMC3767657.

9. Chang LJ, Gay EE. The molecular genetics of lentiviral vectors--current and future perspectives. Curr Gene Ther. 2001 Sep;1(3):237-51. doi: 10.2174/1566523013348634. PMID: 12109139.

10. Centers for Disease Control (CDC). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. MMWR Morb Mortal Wkly Rep. 1981 Jul 3;30(25):305-8. PMID: 6789108.

11. Centers for Disease Control (CDC). Update: acquired immunodeficiency syndrome-- United States. MMWR Morb Mortal Wkly Rep. 1986 Jan 17;35(2):17-21. PMID: 3079867.

12. Bale MJ, Kearney MF. Review: HIV-1 phylogeny during suppressive antiretroviral therapy. Curr Opin HIV AIDS. 2019 May;14(3):188-193. doi: 10.1097/COH.0000000000000535. PMID: 30882485; PMCID: PMC6482946.

13. Chen B. HIV Capsid Assembly, Mechanism, and Structure. Biochemistry. 2016 May 10;55(18):2539-52. doi: 10.1021/acs.biochem.6b00159. Epub 2016 Apr 26. PMID: 27074418.

14. Painter MM, Zaikos TD, Collins KL. Quiescence Promotes Latent HIV Infection and Resistance to Reactivation from Latency with Histone Deacetylase Inhibitors. J Virol. 2017 Nov 30;91(24):e01080-17. doi: 10.1128/JVI.01080-17. PMID: 29021396; PMCID: PMC5709582.

15. Pace MJ, Agosto L, Graf EH, O'Doherty U. HIV reservoirs and latency models. Virology. 2011 Mar 15;411(2):344-54. doi: 10.1016/j.virol.2010.12.041. Epub 2011 Feb 1. PMID: 21284992; PMCID: PMC3618966.

16. Han Y, Wind-Rotolo M, Yang HC, Siliciano JD, Siliciano RF. Experimental approaches to the study of HIV-1 latency. Nat Rev Microbiol. 2007 Feb;5(2):95-106. doi: 10.1038/nrmicro1580. PMID: 17224919.

17. Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A. 1998 Jul 21;95(15):8869-73. doi: 10.1073/pnas.95.15.8869. PMID: 9671771; PMCID: PMC21169.

18. Hudelson C, Cluver L. Factors associated with adherence to antiretroviral therapy among adolescents living with HIV/AIDS in low- and middle-income countries: a systematic review. AIDS Care. 2015;27(7):805-16. doi: 10.1080/09540121.2015.1011073. Epub 2015 Feb 23. PMID: 25702789.

19. Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. Lancet. 2006 Aug 5;368(9534):489-504. doi: 10.1016/S0140-6736(06)69157-5. PMID: 16890836; PMCID: PMC2913538.

20. Seif E, Niu M, Kleiman L. Annealing to sequences within the primer binding site loop promotes an HIV-1 RNA conformation favoring RNA dimerization and packaging.

RNA. 2013 Oct;19(10):1384-93. doi: 10.1261/rna.038497.113. Epub 2013 Aug 19. PMID: 23960173; PMCID: PMC3854529.

21. Yuan Y, Kerwood DJ, Paoletti AC, Shubsda MF, Borer PN. Stem of SL1 RNA in HIV-1: structure and nucleocapsid protein binding for a 1 x 3 internal loop. Biochemistry. 2003 May 13;42(18):5259-69. doi: 10.1021/bi034084a. PMID: 12731867.

22. Kuzembayeva M, Dilley K, Sardo L, Hu WS. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology. 2014 Apr;454-455:362-70. doi: 10.1016/j.virol.2014.01.019. Epub 2014 Feb 14. PMID: 24530126; PMCID: PMC6258065.

23. Clever JL, Wong ML, Parslow TG. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. J Virol. 1996 Sep;70(9):5902-8. doi: 10.1128/JVI.70.9.5902-5908.1996. PMID: 8709210; PMCID: PMC190608.

24. Amarasinghe GK, De Guzman RN, Turner RB, Summers MF. NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform. J Mol Biol. 2000 May 26;299(1):145-56. doi: 10.1006/jmbi.2000.3710. PMID: 10860728.

25. Ocwieja KE, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, Wang S, Link DR, Olson J, Travers K, Schadt E, Bushman FD. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. Nucleic Acids Res. 2012 Nov 1;40(20):10345-55. doi: 10.1093/nar/gks753. Epub 2012 Aug 25. PMID: 22923523; PMCID: PMC3488221.

26. Gendron K, Charbonneau J, Dulude D, Heveker N, Ferbeyre G, Brakier-Gingras L. The presence of the TAR RNA structure alters the programmed -1 ribosomal frameshift efficiency of the human immunodeficiency virus type 1 (HIV-1) by modifying the rate of translation initiation. Nucleic Acids Res. 2008 Jan;36(1):30-40. doi: 10.1093/nar/gkm906. Epub 2007 Nov 5. PMID: 17984074; PMCID: PMC2248755.

27. Tazi J, Bakkour N, Marchand V, Ayadi L, Aboufirassi A, Branlant C. Alternative splicing: regulation of HIV-1 multiplication as a target for therapeutic action. FEBS J. 2010 Feb;277(4):867-76. doi: 10.1111/j.1742-4658.2009.07522.x. Epub 2010 Jan 15. PMID: 20082634.

28. Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, Capoferri AA, Beg SA, Huang SH, Karandish S, Hao H, Halper-Stromberg E, Yong PC, Kovacs C, Benko E, Siliciano RF, Ho YC. Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. Cell Host Microbe. 2017 Apr 12;21(4):494-506.e4. doi: 10.1016/j.chom.2017.03.008. PMID: 28407485; PMCID: PMC5433942.

70

29. Debaisieux S, Rayne F, Yezid H, Beaumelle B. The ins and outs of HIV-1 Tat. Traffic. 2012 Mar;13(3):355-63. doi: 10.1111/j.1600-0854.2011.01286.x. Epub 2011 Oct 11. PMID: 21951552.

30. Huigen MC, Kamp W, Nottet HS. Multiple effects of HIV-1 trans-activator protein on the pathogenesis of HIV-1 infection. Eur J Clin Invest. 2004 Jan;34(1):57-66. doi: 10.1111/j.1365-2362.2004.01282.x. PMID: 14984439.

31. *Qiagen Plasmid Purification Handbook.* Qiagen , 2012.

32. Bullen CK, Laird GM, Durand CM, Siliciano JD, Siliciano RF. New ex vivo approaches distinguish effective and ineffective single agents for reversing HIV-1 latency in vivo. Nat Med. 2014 Apr;20(4):425-9. doi: 10.1038/nm.3489. Epub 2014 Mar 23. PMID: 24658076; PMCID: PMC3981911.

33. Biolabs, New England. "FAQ: Why Do I See a DNA Smear on an Agarose Gel after a Restriction Digest?" *NEB*, www.neb.com/faqs/2016/04/07/why-do-i-see-a-dna-smear-on-an-agarose-gel-after-a-restriction-digest.