Single Cell resolution of early HIV latency using Novel HIV dual reporter system

Donald V. Doanman[1], Namita Satija[2], Tia Robinson[1], Gerrit Schmidt[1], Foramben Patel[1], Aditi Valada[3], Tova Y Lambert[3], Ying-Chih Wang, Kristin G. Beaumont, Robert Sebra, Schahram Akbarian[3], and Benjamin K. Chen[1].

1 Division of Infectious Diseases, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029
2 Present affiliation, Cadent Medical Communications, New York, NY
3 Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Icahn School of Medicine at Mount Sinai

BACKGROUND

Current Human Immunodeficiency Virus (HIV) treatments can effectively suppress viral replication but fail to cure patients of infection due to the persistence of latently infected cells. Latently infected cells are nearly indistinguishable from uninfected cells and evade clearance by the host immune system. There is no effective treatment to eliminate these cells. Current primary cell models of latency utilize in vitro infection often employ reporter genes that are carried and expressed by HIV, to monitor active or downregulated HIV infection. A limitation of these models is that because reporter genes carried by the HIV clone, they are susceptible to integration site-specific silencing mechanisms and since they are not replication competent, do not allow tracking of selective pressures of the reservoir driven by pathogenic or immunologic selective pressures.

METHODS

We designed an infectious model system that independently reports on a cell's HIV infection status and viral gene expression and used this model to study the kinetics of infection and early latency establishment. The system employed a cre-lox activated genetic switch, along with replication-competent HIV clones that co-express cre-recombinase and reporter gene for viral gene expression. This system, described as Enhanced HIV-Induced Lineage Tracing (EHILT), identifies cells with a "history" of HIV infection through an irreversible genetic marking system of HIV-infected cells. HIV infections were performed T cell lines stably transduced with a cre-activated red-to-green genetic switch or CD4+ T cells isolated from PBMCS of healthy donors transduced with Lenti-URG and then infected with a replication competent HIV designed to express cre recombinase and miRFP670nano fluorophore as a reporter for viral gene transcription. The system efficiently distinguishes latently infected cells from infected cells with active viral transcription and infection is easily measured using flow cytometry. Infected cells are sorted based on expression of EHILT reporters and RNA and DNA extraction is performed. qRT-PCR and DNA qPCR using viral genes can be used to measure the viral gene expression and presence of HIV DNA. Latently infected and actively infected cells sorted by flow cytometry were analyzed using 10x single cell RNA sequencing. A reference genome was developed by adding HIV-Nano Cre and Lenti-URG sequence based on HIV splice donor and acceptors. Analysis was conducted using Seurat 4.2.0. Standard QC using mitochondrial gene cut offs and unique feature counts ranges was applied. Multimodal Reference Mapping was used to subset the datasets for CD4 T Cells based on predictive cell type scores. Non-linear dimensionality reduction analysis was used to produce unsupervised clustering. Differential gene expression analysis was conducted comparing Actively infected and latently infected cells. Differential gene expression data was analyzed using Ingenuity Pathway Analysis.

RESULTS

Figure 1. Construct design of Enhanced HIV-1 induced lineage tracing (EHILT) marking system

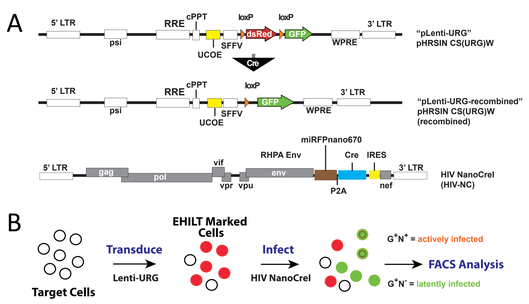

Figure 1. Lenti-URG is a lentiviral vector that constitutively expresses a floxed dsRed gene upstream of an untranslated GFP gene. In the presence of cre, recombination induces a switch from red to green. HIV-Nano Cre is a R5-Tropic HIV-1 clone that carries reporter gene RFP670nano and cre in place of nef, and nef expression restored with an IRES. A schematic of the workflow for transduction with Lenti-URG and infection with HIV-Nano Cre is shown with phenotypes of infected cells.

Figure 2. HIV Infection Kinetics of EHILT marked cells reveals immediate latent phenotype

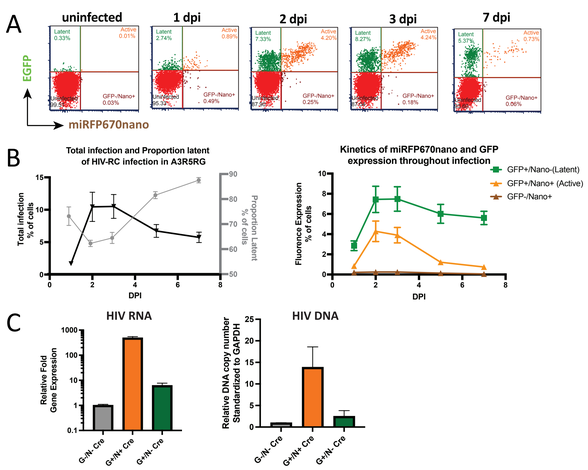

Figure 2. Infection of A3R5RG cells over a 7 day time course followed by Flow Cytometric analysis. A3R5RG cells are a human CD4+ lymphoblastoid cell line that expresses CCR5 and the red-to-green switch gene cassette. Infection by spinoculation was carried out using 10 ng of p24 per 50,000 cells. Expression of GFP and miRFP670nano was measured by flow cytometry. Infection was measured over 7 days and breakdown of infection by percentage of cells expressing GFP and HSA or miRFP670nano is plotted over time. GFP is a marker for HIV infection history and HSA or miRFP670nano is a marker for viral gene transcription. Cells were sorted at 7dpi and qRT PCR and DNA qPCR was conducted with primers specific for HIV cre and GAPDH. Relative fold gene expression and DNA copy number was calculated by normalizing to GAPDH.

Figure 3. EHILT Marking in Primary CD4 T cells allows for scRNAseq

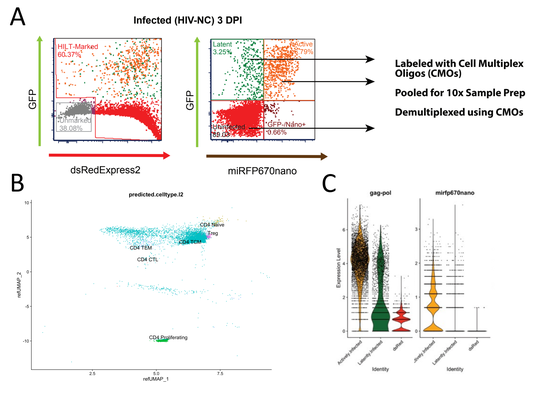

Figure 3. Primary CD4 T cells from healthy donors can be marked with EHILT by transduction with Lenti-URG and infection with HIV Nano Cre. Latently infected, Actively infected, dsRed+ uninfected, and unmarked cells were labeled with Cell Multiplexing Oligos and processed for 10x Genomics single cell RNA sequencing. Multimodal reference mapping was used to predicted cell type based on gene expression data and subset for CD4 T cells only. Higher levels of viral gene expression are detected in the Actively Infected cells in comparison to Latently Infected cells and dsRed Uninfected cells.

Figure 4. Unsupervised clustering based on Gene Expression is influenced by EHILT Infection Status, Cell Type, and HIV Expression

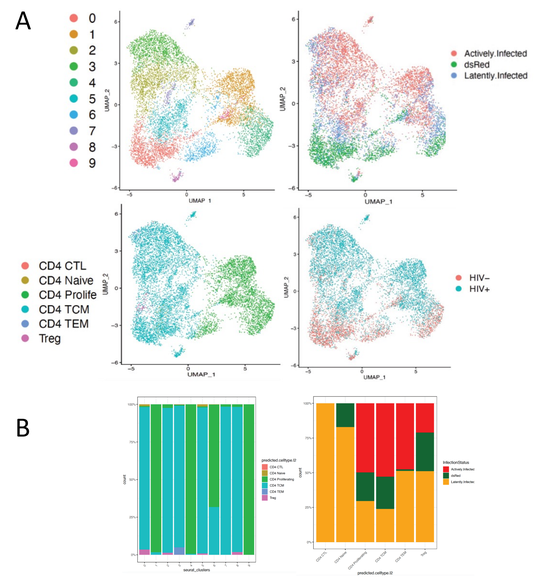

Figure 4. UMAP plot produced from unsupervised clustering reveals ten transcriptional clusters. EHILT infection status is determined by using the Cell Multiplexing Oligo labels from FACS sorting where actively and latently infected cells were sorted based on fluorescence of GFP and Nano. Predicted cell type is labeled by comparing gene expression data to a surface phenotype using Multimodal reference map provided by the Satija Lab (https://satijalab.org/seurat/articles/multimodal_reference_mapping.html) . HIV Expression status is determined by detection of HIV transcripts where any cell with 1 or more HIV transcript is considered HIV+. These data were added as metadata and overlayed the UMAP plot from unsupervised clustering. The fraction of EHILT phenotypes in each Seurat Cluster and CD4 T cell subset is represented as bar graphs.

Figure 5. Differential gene expression analysis reveals pathways differentially regulated in Active vs Latent infection

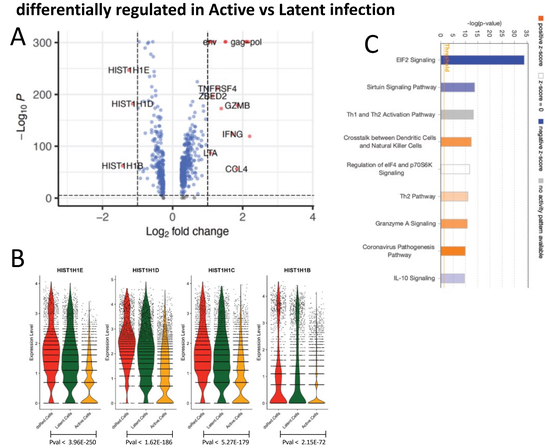

Figure 5. Active and latent cells were analyzed for differential gene expression using Wilcoxon Rank Sum test. These differentially expressed genes are represented in a volcano plot. The top differentially regulated genes include HIV genes and Histone 1 Linker genes. The Expression of Histone 1 Linker genes are represented in violin plots and show decreased expression in Actively infected cells. Genes with a P value of .05 or less were used in Ingenuity Pathway Analysis to determine transcriptional pathways differentially regulated by HIV latency.

SUMMARY & CONCLUSION

1. EHILT identifies infected cells with low transcription, putative early latent cells, in *in vitro primary* CD4 T cell infection without requiring prolonged quiescence or ART treatment
2. Efficient activation, transduction, infection and early analysis of CD4 T cells retains T cell phenotype and distinct transcriptional profiles
3. Multiple T cell types and transcriptional states can support early HIV latency during acute infection
4. Pathway analysis can identify common and distinct pathways associated with latency
5. Sirtuin and EIF2 signaling pathways are examples of pathways that may favor latency reversal in many T cell states
6. H1 linker protein is strongly downregulated in Active HIV transcription and may be relevant in maintaining transcriptional quiescence in proliferating cells

ACKNOWLEDGEMENTS

We would like to acknowledge and thank the following grants that have funded this project

Funding: NIH F-30 5F30AI162222-02