Parallelized Genome-Wide Detection of Intact HIV Provirus with Bimodal Phenotyping in Single Cells

Yi Qi[1], Leena Babiker[2], Hanzhi Wang[1], Sam Barnett Dubensky[2,3], Hunter Courtney[4], Jonathan Tedesco[5], Parker C. Wilson[6], Michael R. Betts[7], Laura A. Vella[5]

[1]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA; [2]Immunology Graduate Group, Biomedical Graduate Studies, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [3]Medical Scientist Training Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [4]UCLA-Caltech Medical Scientist Training Program, David Geffen School of Medicine, University of California, Los Angeles, CA, USA; [5]Division of Infectious Diseases, Children's Hospital of Philadelphia, Philadelphia, PA, USA; [6]Division of Diagnostic Innovation, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [7]Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

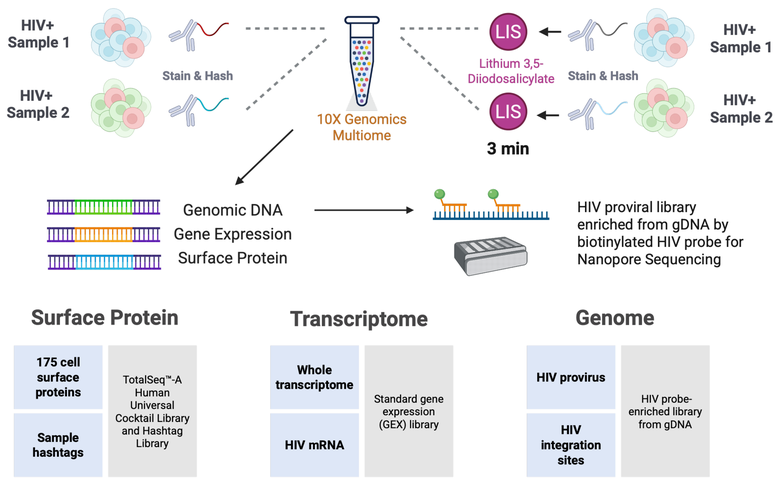
341

Background

Precise definition of viral and cellular features of HIV reservoir can reveal persistence mechanisms to advance targeted therapies. Prolonged cART has been shown to shift reservoir landscape and is associated with a greater proportion of provirus in less accessible chromatin. However, current single cell reservoir approaches are unable to detect a heterochromatic provirus while maintaining the ability to simultaneously define the features of its host CD4⁺ T cell. To enable rational targeting of the reservoir, we hypothesized that chemically relaxing the heterochromatin would enable detection of heterochromatic provirus as cell identity was maintained. Further, we hypothesized that long-read sequencing would allow detailed virologic readouts of these proviruses, including estimates of intactness and integration site (IS), in single cells.

Methods

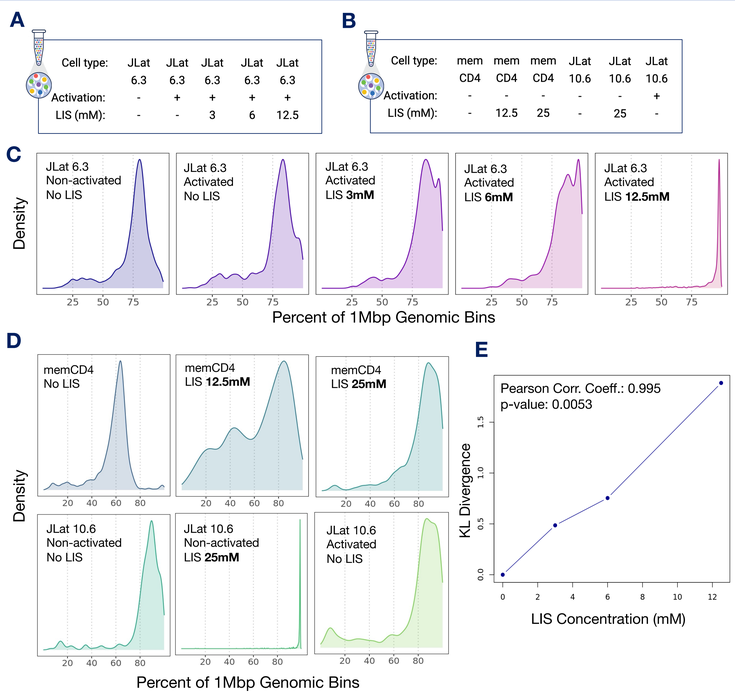
Here, we developed **SCRIPT-seq**: Single Cell detection of Retroviral Integration and cellular PhenoTyping. This method utilized a brief lithium 3,5-diiodosalicylate treatment to dissociate DNA-histone interaction and enable sequencing gDNA, mRNA, and surface protein in single cells.



We also developed **LENS-HIV**, an in-house pipeline for proviral reads: (Long-read, targeted Examination of geNomes in Single cells, for HIV provirus)

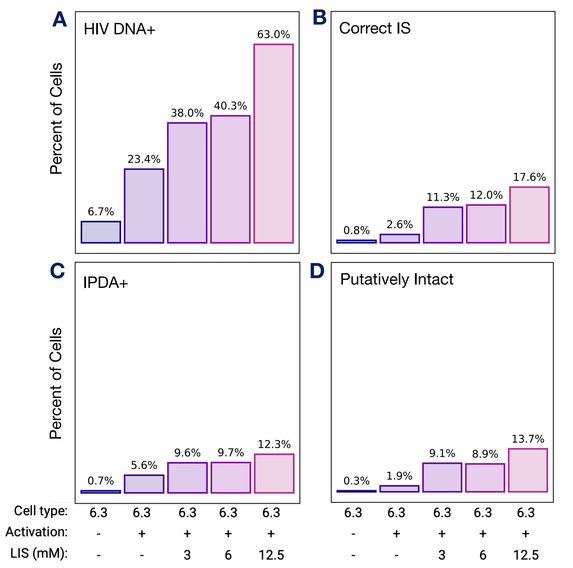
1. Cell barcode detection
2. Alignment to the human + HIV chimeric genome
3. Call HIV DNA+ cells: ≥ 10 chimeric or on-virus reads
4. Call HIV IS of cells: ≥ 10 chimeric reads, representing > 60% all chimeric reads, supporting one site
5. Call HIV IPDA+ cells: ≥ 10 reads containing exact matches to PSI and RRE motifs.
6. Estimate HIV putative intactness: ≥ 2 reads covering ≥ 75% of each HIV gene from a cell with a valid integration site (see #4)

Figure 1 – LIS treatments increased global genome opening and closed chromatin coverage in HIV-integrated cell lines and human peripheral CD4⁺ T cells.



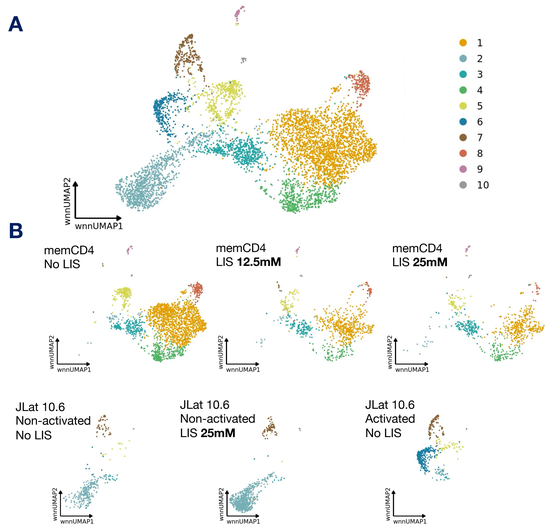
(A-B) Conditions of two SCRIPT-seq single cell experiments. (A) Jurkat latency (JLat) 6.3 cells were either non-activated or activated with PMA/Ionomycin. After short permeabilization with 0.01% digitonin, activated cells were treated with 0, 3, 6, or 12.5 mM LIS for 3 min. (B) Freshly thawed human memory CD4⁺ T cells were permeabilized and treated with 0, 12.5, or 25 mM LIS. Non-activated JLat 10.6 cells were permeabilized and treated with 0 or 25 mM LIS, with activated, non-treated cells as control. (C-D) Genomic reads of single cells in each condition were mapped to 2,772 GRCh38 1Mbp bins. Density plots show the single cell distribution of the percentage of bins overlapped by at least one read from a given cell. Right-shifting distributions of genomic bin coverage suggest LIS led to enhanced genomic accessibility in a dose-dependent manner in (C) JLat cells and in (D) human memory CD4⁺ T cells. (E) A positive correlation was shown between the LIS dose and the Kullback–Leibler (KL) divergence of each LIS-treated JLat 6.3 distribution from the distribution of the non-treated JLat 6.3 cells, as shown in panel C . (F-G) Genomic reads were mapped to the E043 human peripheral CD4⁺ T cell reference epigenome mnemonics (Roadmap Epigenomics Consortium) and assigned to Euchromatin, Heterochromatin, Polycomb, or the Quiescent group. Proportion of total reads assigned to each group was calculated in single cells. LIS-treated (F) JLat cells and (G) memory CD4⁺ T cells both had markedly higher proportion of reads mapped to Heterochromatin, Polycomb, and Quiescent groups, suggesting LIS effectively revealed less accessible regions. Euchromatin (TssA, TssFlnk, TssFlnkD, Tx, TxWk, EnhG1, EnhG2, EnhA1, EnhA2, EnhWk); Heterochromatin (ZNF/Rpts, Het); Polycomb (ReprPC, ReprPCWk); Quiescent (Quies).

Figure 2 – Targeted long-read sequencing with stringent thresholds detected HIV DNA, integration sites, IPDA motifs, and intactness.



Targeted long-read sequencing and analysis by the LENS-HIV pipeline with stringent thresholds enabled the detection of (A) HIV DNA, (B) integration site, and (C) IPDA motifs as well as (D) intactness estimation in single cells. LIS treatment increased the percent of cells detected in all metrics.

Figure 3 – Two-way WNN clustering by RNA and surface protein unchanged by LIS treatment.



(A) Weighted nearest neighbor (WNN) UMAP of RNA and surface protein clustering reveals (B) highly similar cluster participation of LIS-treated and non-treated memory CD4⁺ T cells and non-activated JLat 10.6 cells, regardless of LIS concentration, clear separation between JLat 10.6 vs. memory CD4⁺ T cells, and clear separation between activated vs. non-activated JLat 10.6 cells.

Conclusions

- To better understand mechanisms of HIV persistence during immunotherapy, it is critical to define the viral and cellular features of persisting reservoir. However, the most resistant reservoir may be in closed, inaccessible chromatin.
- LIS treatment enhanced genome accessibility and revealed regions associated with inactive chromatin states in HIV cell lines and primary CD4 T cells, without perturbing cell phenotypes.
- Targeted long-read sequencing enabled single-genome profiling of proviral DNA, IS, IPDA motifs, and intactness, which were more readily detectable in LIS-treated cells.
- SCRIPT-seq provided highly resolved virus-host information to decode the persistence mechanisms of the intact HIV reservoir.

Plain Language Summary

- We developed a sequencing technique to locate HIV hidden in a cell's DNA and study how the virus affects the cell, thus providing key information to develop better drugs that expose the virus to the immune system.

Acknowledgement

NIH Funding: R01AI176597, R21AI189303
Penn Center for AIDS Research
CHOP High-Throughput Sequencing Core
CHOP Flow Cytometry Core

Contact

Lab - www.vellalab.org
Email - qiy@chop.edu