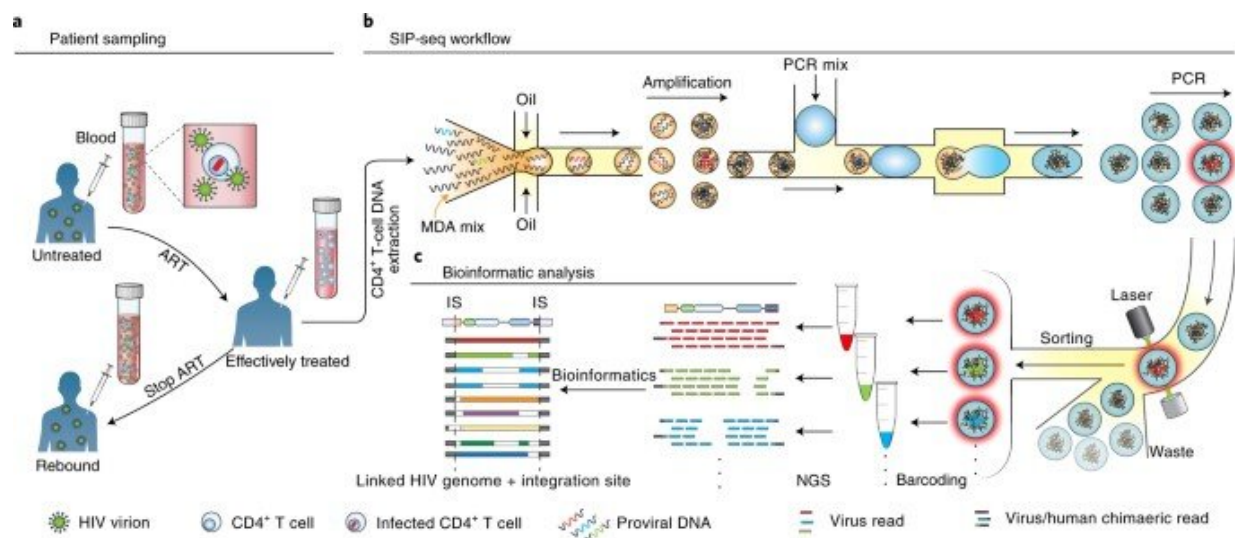


# Sequencing HIV proviruses from people on antiretroviral therapy using droplet-microfluidics

April 12 2022, by Thamarasee Jeewandara



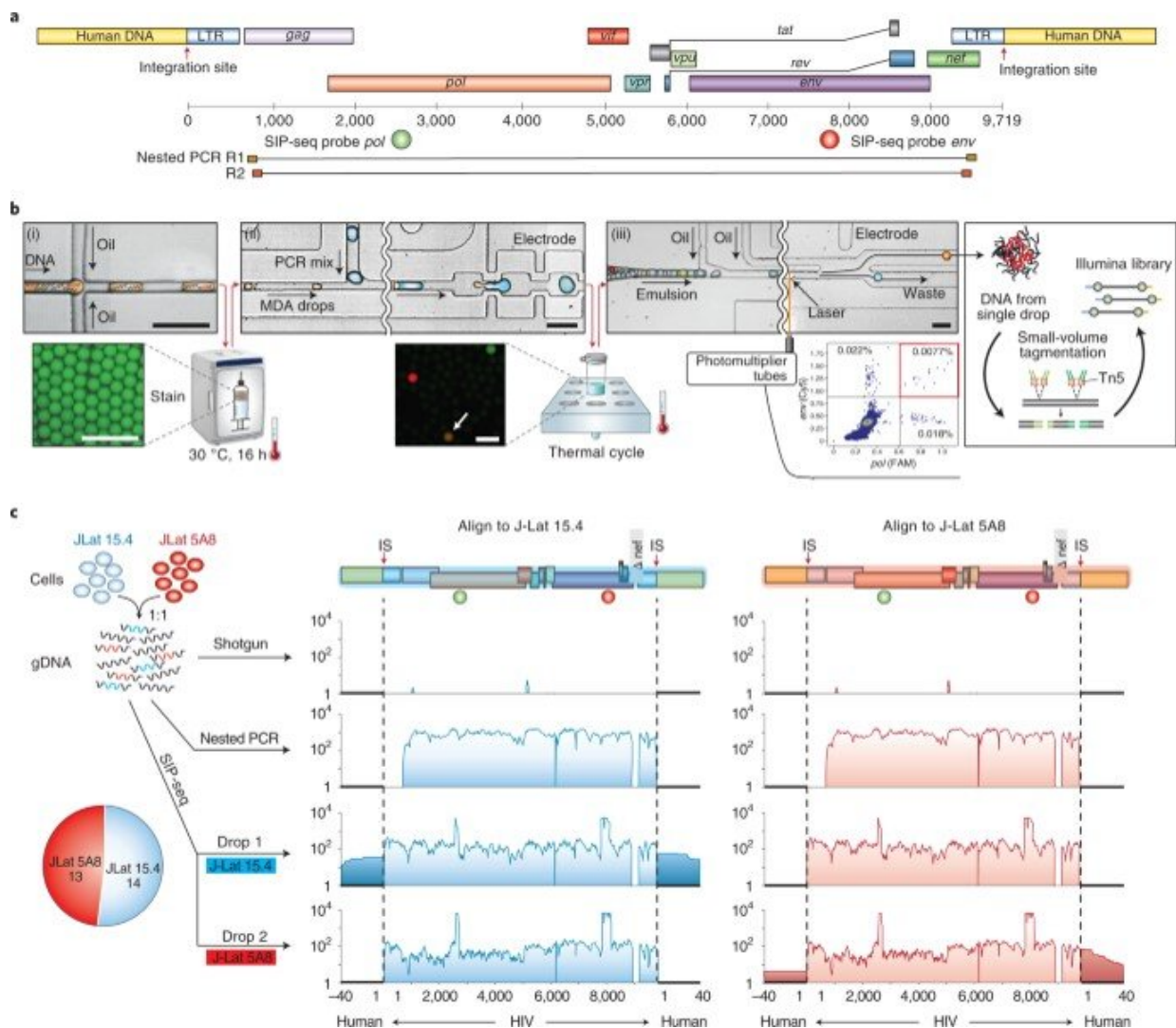
Application of SIP-seq to ART-treated individuals. (a) To study HIV persistence during ART, CD4<sup>+</sup> T cells are sampled from participants with undetectable viral load in the plasma. (b) CD4<sup>+</sup> T-cell DNA samples are processed with SIP-seq to generate single provirus sequencing libraries. Genomic DNA (gDNA) is extracted and compartmentalized in droplets. DNA is MDA-amplified in each droplet, followed by PCR detection and sorting of HIV-positive droplets. DNA from each sorted droplet is barcoded and sequenced. (c) Sequence data are mapped to a reference HIV genome for simultaneous determination of individual virus genomes and integration sites (IS). Credit: *Nature Biomedical Engineering* (2022). DOI: 10.1038/s41551-022-00864-8

The infection of human immunology viruses (HIV) occurs by integrating its genome into infected cells to enter an inactive state of reversible latency that evades [anti-retroviral therapy](#). The capacity to sequence such a provirus and adjacent host junctions in individual cells can highlight their mechanism of action and persistence in infected cells. However, this experiment is difficult to perform due to the 150-million-fold higher amount of background human DNA. In a new report now published in *Nature Biomedical Engineering*, Chen Sun and a research team in bioengineering and viral persistence and dynamics studies at the University of California, San Francisco, NIH and the Chan Zuckerberg Biohub, in the U.S., showed how full-length proviruses are connected to the HIV-host DNA junctions via high-throughput microfluidic assays, for droplet-based whole genome amplification of HIV DNA in its native context. The team conducted a polymerase chain reaction (PCR) to tag droplets containing proviruses to sequence and assay infected cells from people with HIV, receiving suppressing antiretroviral therapy at the time to detect and sequence paired proviral genomes and integration sites. The work sought to improve genetic analysis of persistent HIV-infected cell reservoirs.

## **The current state of HIV treatment and evolving therapeutic strategies**

While approximately 37 million people worldwide are infected with [human immunodeficiency virus](#) (HIV), the cure remains unknown. The mechanism of infection of the virus imposes a central obstacle and also holds the key to curing the disease, as HIV integrates its genome into the genome of infected cells. While antiretroviral therapy (ART) can suppress virus replication to undetectable levels and prevent the progression of [acquired immunodeficiency syndrome](#) (AIDS), it lacks activity against [cellular reservoirs](#). As a result, the therapy must be continued indefinitely to prevent virus rebound and disease progression.

It is difficult to address HIV-associated stigma, due to the cost of lifelong ART administration, with toxic potential. It is therefore important to eradicate or suppress HIV-infected cellular reservoirs to obtain a functional cure. To examine the persistence of cellular reservoirs, scientists must conduct paired analysis of integration sites and full-length provirus sequences. By establishing a method that reliably and cost-effectively isolates and sequences rare pro-viruses, the researchers can comprehensively characterize HIV reservoirs. Sun et al developed such an approach to characterize proviruses and their cellular genomic context within HIV reservoirs using simultaneous integration site and provirus sequencing (SIP-Seq).



Demonstration and validation of SIP-seq with HIV-infected cell line samples. (a) TaqMan assays targeting the pol (green) and env (red) genes were used by SIP-seq to detect droplets containing HIV genomes. By comparison, nested PCR amplified near-full-length HIV genomes, but could not simultaneously recover the genome and flanking host sequences. Primer binding sites are shown relative to the schematic of an HIV genome integrated into a human genome. (b) SIP-seq microfluidics was used to identify HIV genome sequences and comprised the following steps. (i) Encapsulation and amplification of DNA via MDA. An aliquot was stained with EvaGreen for visualization. (ii) Droplet merging was used to add TaqMan PCR reagents and carry out in-droplet PCR. The fluorescence image shows representative drops after thermocycling, indicating a double positive that would be isolated by sorting. (iii) Droplet sorting was used to select double PCR positives. A representative scatter plot of pol (FAM dye) and env (Cy5 dye) intensities is shown. DNA from each sorted drop was processed for sequencing. Scale bars, 100  $\mu\text{m}$ . (c) SIP-seq recapitulated HIV virus genomes and integration sites from a mixture of DNA from two J-Lat cell lines. gDNA from a mixture of the cell lines was processed with shotgun sequencing, sequencing of nested PCR products, and two SIP-seq drops containing either of two J-Lat proviral sequences. Both J-Lat proviruses had a deletion in the nef gene and sequences mapping to the full-length HXB2 genome, so no reads mapped to this region. The pie chart indicates the recovery of genomes from a 50:50 mixture of DNA from the two cell lines. Credit: Nature Biomedical Engineering (2022). DOI: 10.1038/s41551-022-00864-8

### **Beating the odds: A method to recover single provirus genomes**

By using whole-genome sequencing amplification, Sun et al amplified the HIV genome in its native context within microfluidic droplets, then tagged the droplets containing proviruses for sequencing. The resulting method provided a full-length virus genome connected to its host-cell junctions via a single continuous assembly. The speed and efficiency of

the method allowed the recovery of single provirus genomes in a [150-million-fold higher background of DNA](#). The team used SIP-Seq (simultaneous integration site and provirus sequencing) to profile the provirus population in a variety of ART-related individuals to understand the latent HIV reservoir. While the team focused on HIV, this approach is applicable to diverse viruses that invade the host's genome, to provide a general platform to characterize genetics of a variety of infections.



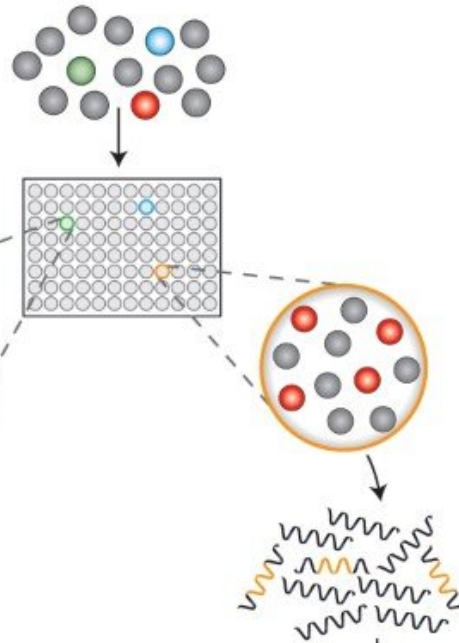
**a**

Patient resting  
CD4<sup>+</sup> T cells

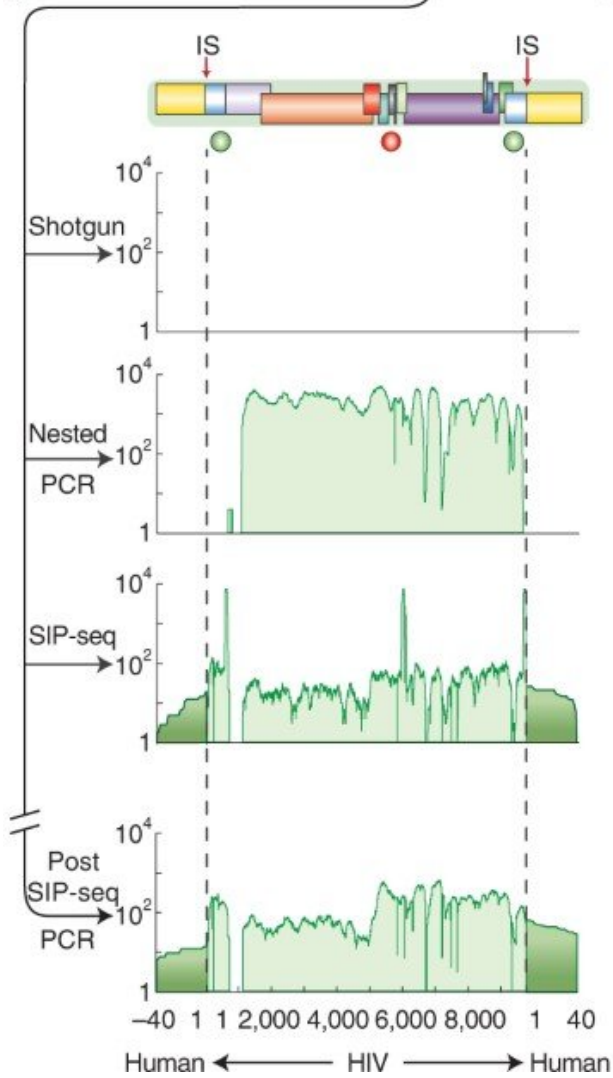
<1 infected  
cell per well

Clonal  
expansion

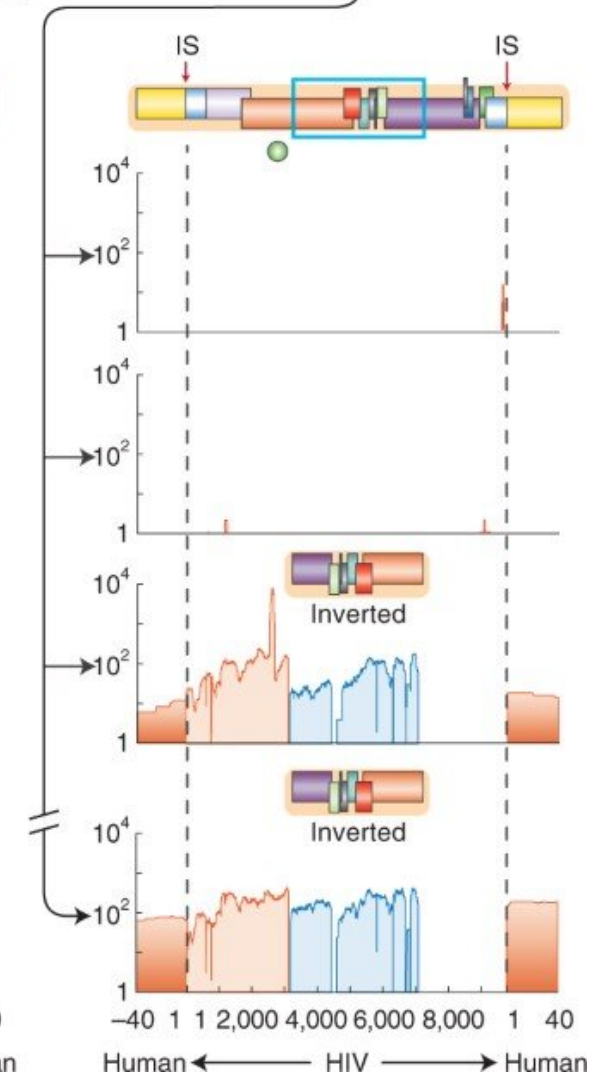
DNA  
isolation



**b**



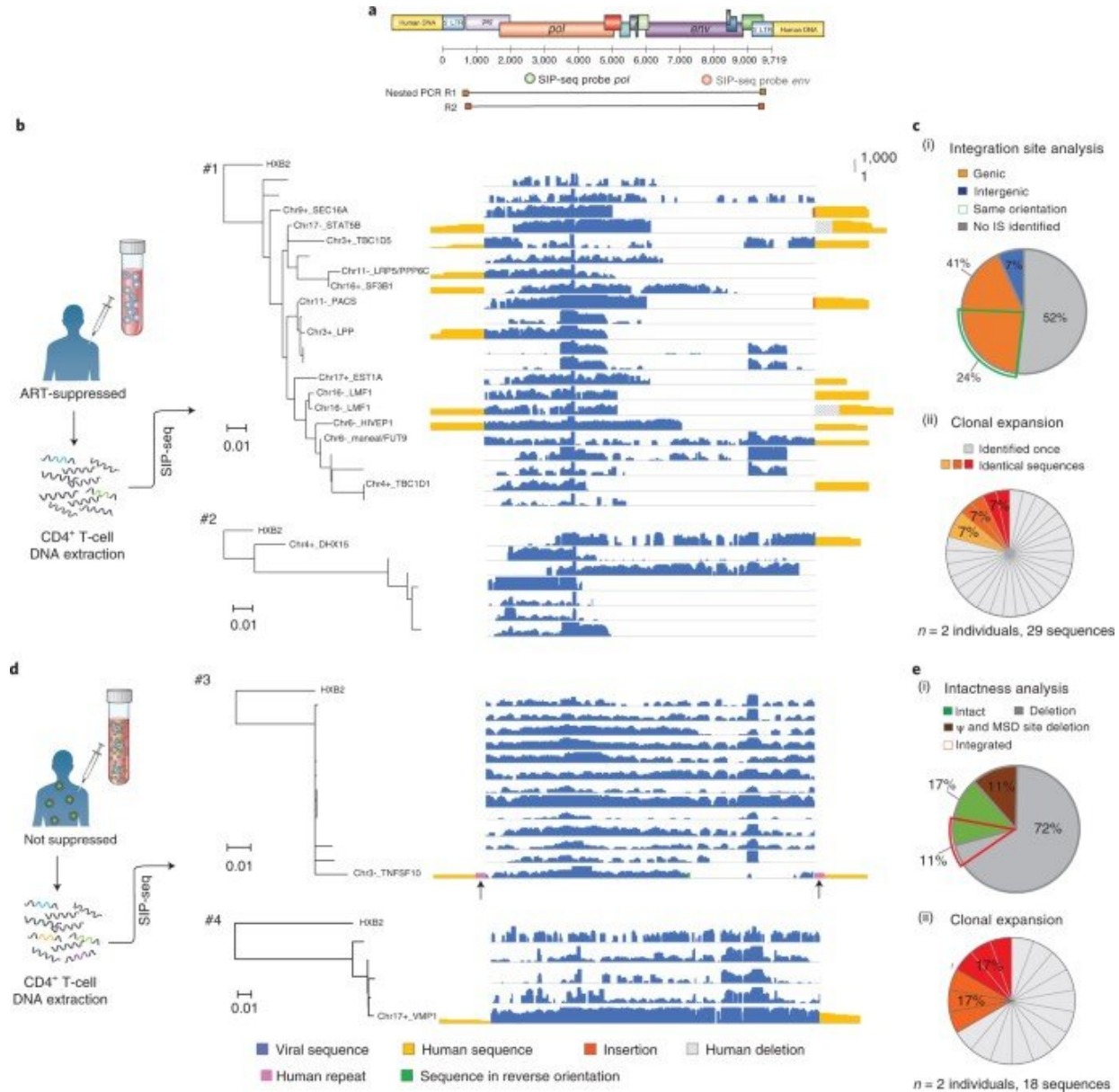
**c**



(a) ICE was prepared by seeding participant CD4<sup>+</sup> T cells to less than one infected cell per five wells and culturing for clonal expansion. Two ICE clones were separately processed for sequencing with different technologies. (b) SIP-seq detected an intact HIV genome integrated into a human genome. (c) SIP-seq identified an integrated HIV genome containing an inverted sequence and a large deletion including 3' LTR. Nested PCRs from unprocessed genomic DNA using primers designed on the basis of the SIP-seq results confirmed the HIV genomes (bottom panels of b and c). Credit: Nature Biomedical Engineering (2022). DOI: 10.1038/s41551-022-00864-8

## The experiments

Sun et al aimed to recover all DNA fragments containing HIV genomes from a population of millions of [CD4<sup>+</sup> T cells](#), to sequence the molecules individually, and their immediately adjacent junctions. The team extracted large DNA fragments from cells encapsulated in microdroplets to isolate the integrated HIV genomes. Using [multiplexed displacement amplification](#) (MDA), they non-specifically amplified each genomic fragment to obtain sufficient single proviruses for sequencing. They then followed this by isolating individual droplets with HIV proviruses, where each droplet was barcoded and sequenced. The virologists then mapped readouts from each droplet to an HIV reference genome to identify integration sites to facilitate clear-cut information on the complexity of the virus genome, and their integration sites—crucial to assess the cellular reservoir.



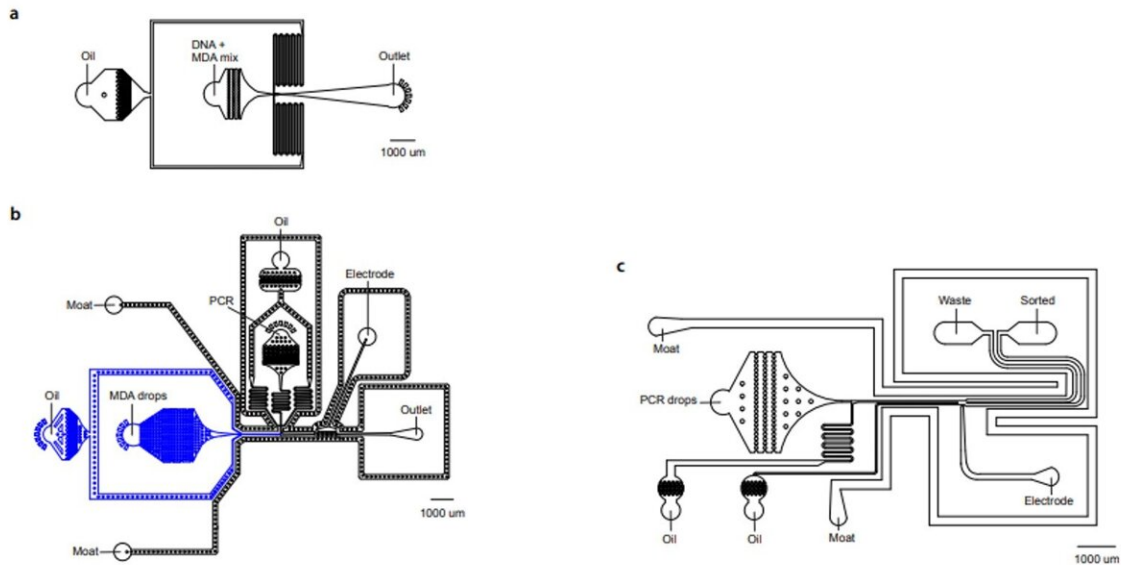
SIP-seq of HIV proviruses and flanking sequences from CD4+ cells directly from infected individuals. a, Schematic showing positions of TaqMan PCR probes in HIV pol and env, as well as the locations of the nested PCR primers used for near-full-length genome amplification. The sorting used here was based on pol solely. b, SIP-seq of DNA from T cells from individuals (1 and 2) with ART suppression. Phylogenetic trees were generated using an ~500-bp region of pol from all HIV sequences obtained from each participant. Host genes and provirus orientation relative to the host gene for each sequence are indicated unless no integration site was identified. Each bar represents a coverage map of



an individual HIV genome and its integration site. c, Integration sites were favoured in genic regions, with no preference in orientation relative to host genes (i) and expanded clones were found (ii) in the participants on ART suppression. d, SIP-seq of DNA from CD4+ T cells from donors without ART suppression (participant 3 had just started ART and participant 4 had ART suspended). e, Percentage of intact and defective HIV (i) and expanded clones (ii) in participants 3 and 4.  $\Psi$ , packaging signal; MSD, major splice donor. Credit: Nature Biomedical Engineering (2022). DOI: 10.1038/s41551-022-00864-8

## **Detecting proviruses with SIP-Seq and regulating microfluidics**

To increase the specificity of full-length proviruses and prevent random DNA cleavage, Sun et al employed a dual specific [multiplexed Taqman PCR](#), to target regions on the HIV genome. During the experiments, they yielded a specific volume of DNA to sequence and confirm HIV provirus recovery. The microfluidics of the SIP-Seq setup contained three devices, including a droplet encapsulator, a merger and a sorter. The scientists loaded the human genomic DNA fragments with multiple displacement amplification (MDA) reagents to encapsulate up to 10 billion DNA fragments, approximating 75 kbp in length, in 20 million separate droplets, within 15 minutes. Each droplet contained 500 distinct fragments and Sun et al incubated them at 30 degrees Celsius, for non-specific amplification, prior to merging with Taqman PCR reagents. After repeating the steps for all detected HIV genomes, the team [prepared the tubes for sequencing](#).



Schematics of microfluidic devices used to: a) encapsulate DNA and MDA reagents into isolated droplets; b) merge MDA droplets with PCR mix droplets; and c) sort individual PCR positive droplets. Credit: Nature Biomedical Engineering (2022). DOI: 10.1038/s41551-022-00864-8

## Analyzing HIV proviruses and their genomic landscape in ART-treated persons

Sun et al next validated the SIP-Seq on HIV cell lines and analyzed HIV in persons treated with ART. During the experiments, they prepared cells by plating, resting CD4<sup>+</sup> T cells from an ART-treated person, followed by stimulation and in vitro culture for proliferation via three methods; [shotgun](#), [nested PCR](#), and SIP-Seq with Taqman probes. Both SIP-Seq and nested PCR yielded excellent results, the team specifically opted for SIP-Seq to characterize the full genetic diversity of infected cells in vivo. They followed this work by characterizing the genomic landscape of HIV proviruses, which they directly isolated from CD4<sup>+</sup> T cells of infected persons. The outcomes revealed three clonal lineages to

support clonal origin relative to [previous studies of the same individuals](#).

## Outlook

The research indicated how the methods enabled Chen Sun and colleagues to study HIV proviruses more efficiently than with existing PCR methods. The outcomes provided a comprehensive analysis of the HIV genetic landscape in vivo by characterizing the latent reservoir to highlight its role in HIV persistence. Existing methods were challenging due to the rarity and lack of distinct surface markers to identify latently infected cells. Using the described SIP-Seq (simultaneous integration site and provirus sequencing) method, Sun et al provided a fast, cost-effective and scalable process, which they applied to people receiving [antiretroviral therapy](#) (ART), to highlight the presence in latent reservoirs of clonally expanded cells. The strategy is applicable across HIV-infections, to other viral diseases with integrated states in their life cycle.

**More information:** Chen Sun et al, Droplet-microfluidics-assisted sequencing of HIV proviruses and their integration sites in cells from people on antiretroviral therapy, *Nature Biomedical Engineering* (2022). [DOI: 10.1038/s41551-022-00864-8](https://doi.org/10.1038/s41551-022-00864-8)

Katherine M. Bruner et al, A quantitative approach for measuring the reservoir of latent HIV-1 proviruses, *Nature* (2019). [DOI: 10.1038/s41586-019-0898-8](https://doi.org/10.1038/s41586-019-0898-8)

© 2022 Science X Network

Citation: Sequencing HIV proviruses from people on antiretroviral therapy using droplet-microfluidics (2022, April 12) retrieved 24 April 2024 from <https://medicalxpress.com/news/2022-04-sequencing-hiv-proviruses-people-antiretroviral.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.