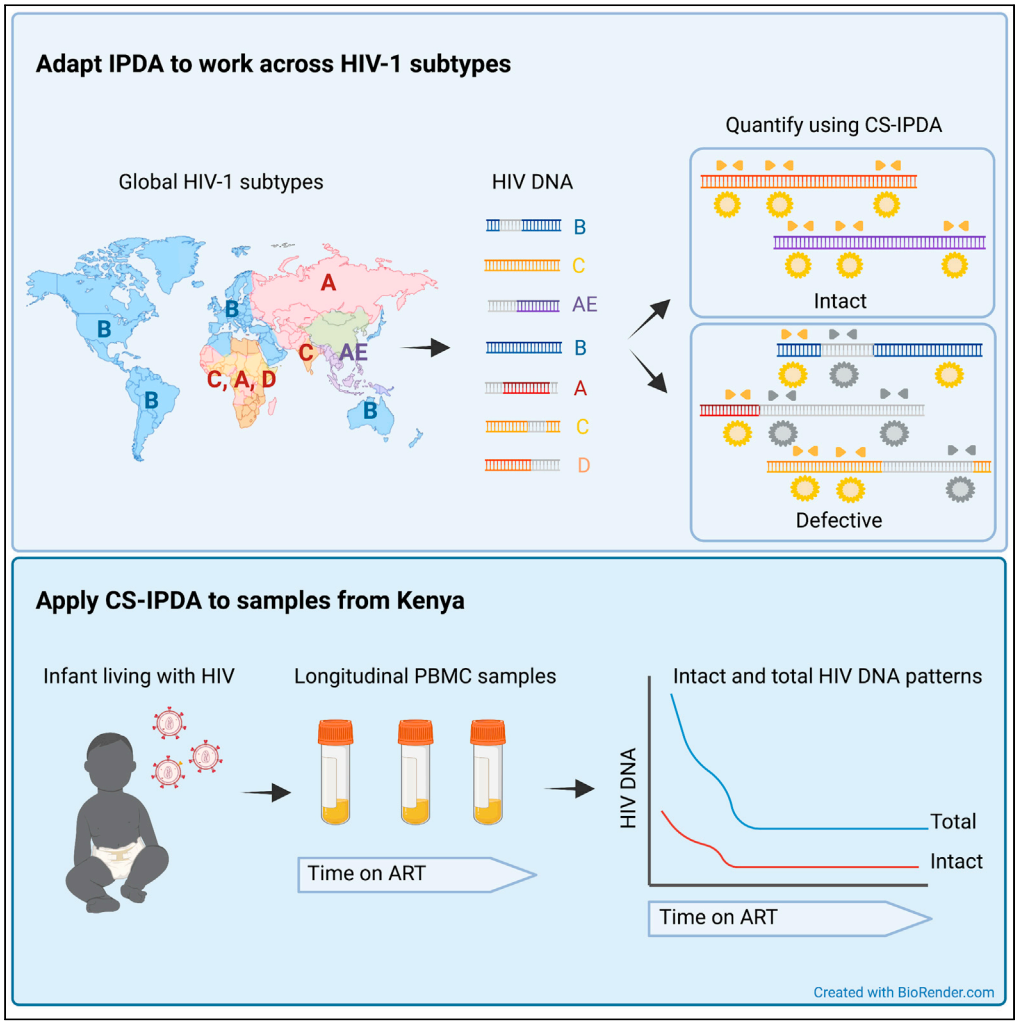
iScience

CellPress
OPEN ACCESS

Article

HIV reservoir quantification using cross-subtype multiplex ddPCR



Noah A.J. Cassidy, Carolyn S. Fish, Claire N. Levy, ..., Julie Overbaugh, Florian Hladik, Dara A. Lehman

dlehman@fredhutch.org

Highlights

High-throughput HIV reservoir assays for diverse HIV subtypes are needed

Cross-subtype intact proviral DNA assay (CS-IPDA) works across global HIV-1 subtypes

CS-IPDA can be used in cure studies in Africa and Asia where HIV prevalence is high

CS-IPDA on samples from Kenyan infants on ART show expected decay of intact HIV DNA

Cassidy et al., iScience 25, 103615
January 21, 2022
https://doi.org/10.1016/j.isci.2021.103615

iScience

CellPress
OPEN ACCESS

Article

HIV reservoir quantification using cross-subtype multiplex ddPCR

Noah A.J. Cassidy,[1] Carolyn S. Fish,[1,10] Claire N. Levy,[2,10] Pavitra Roychoudhury,[3,4] Daniel B. Reeves,[3] Sean M. Hughes,[2] Joshua T. Schiffer,[3,6] Sarah Benki-Nugent,[5] Grace John-Stewart,[5,6,7,8] Dalton Wamalwa,[9] Keith R. Jerome,[3,4] Julie Overbaugh,[1] Florian Hladik,[2,3,6] and Dara A. Lehman[1,5,11,*]

SUMMARY

A major barrier to conducting HIV cure research in populations with the highest HIV burden is the lack of an accurate assay to quantify the replication-competent reservoir across the dominant global HIV-1 subtypes. Here, we modify a subtype B HIV-1 assay that quantifies both intact and defective proviral DNA, adapting it to accommodate cross-subtype HIV-1 sequence diversity. We show that the cross-subtype assay works on subtypes A, B, C, D, and CRF01_AE and can detect a single copy of intact provirus. In longitudinal blood samples from Kenyan infants infected with subtypes A and D, patterns of intact and total HIV DNA follow the decay of plasma viral load over time during antiretroviral therapy, with intact HIV DNA comprising 7% (range 1%–33%) of the total HIV DNA during HIV RNA suppression. This high-throughput cross-subtype reservoir assay will be useful in HIV cure research in Africa and Asia, where HIV prevalence is highest.

INTRODUCTION

Globally effective HIV cure strategies will require cure research focused on the most affected populations. HIV prevalence is highest in sub-Saharan Africa where subtypes A, C, and D predominate; however, most HIV cure research is conducted in resource-rich countries on populations infected with HIV-1 subtype B (Ndung'u et al., 2019). A significant obstacle to including populations infected with non-B subtypes in HIV cure studies is the lack of an accurate assay that can quantify the HIV reservoir across the global diversity of HIV subtypes (Abdel-Mohsen et al., 2020; Ndung'u et al., 2019), given up to 30% sequence diversity across subtypes (Korber et al., 2001).

The HIV reservoir contains latently infected cells that persist even during suppressive combination antiretroviral therapy (ART). Cells productively infected with HIV carry a proviral DNA copy of the HIV genome that is stably integrated into the host cell chromosome. The process of generating the proviral DNA through reverse transcription is error prone and most integrated proviruses contain mutations, insertions, or deletions, rendering them unable to produce new infectious virus. The remaining cells that contain full-length intact proviral DNA and produce replication-competent virus are the source of viral rebound upon treatment cessation and are the major barrier to a cure (Bruner et al., 2016; Hiener et al., 2017; Ho et al., 2013).

For many years, quantitative viral outgrowth assays (QVOAs) that rely on stimulation and culture of replication-competent virus were the gold standard for HIV reservoir quantification (Laird et al., 2013, 2016). However, QVOAs are labor intensive, low-throughput, and underestimate the replication-competent reservoir size (Ho et al., 2013). Assays that quantify intact proviral DNA can measure the replication-competent HIV reservoir more easily and accurately than QVOA, reviewed in (Abdel-Mohsen et al., 2020). To date, assays that characterize proviruses as intact or defective have predominately been designed for subtype B and include full-length sequencing assays (Hiener et al., 2017) as well as intact proviral DNA assays (IPDAs) (Bruner et al., 2019; Levy et al., 2021). IPDAs detect multiple targets across the HIV genome in a multiplex droplet digital PCR (ddPCR) reaction and are higher throughput than full-length sequencing. Here, we optimized an IPDA (Levy et al., 2021) to enable accurate cross-subtype reservoir quantification, which will allow this high-throughput assay to be used in areas of the world with high prevalence of HIV, such as sub-Saharan Africa, where non-B subtypes predominate—an essential step in the progress toward HIV cure.

[1]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[2]Department of Obstetrics and Gynecology, University of Washington, Seattle, WA, USA

[3]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[4]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

[5]Department of Global Health, University of Washington, Seattle, WA, USA

[6]Department of Medicine, University of Washington, Seattle, WA, USA

[7]Department of Epidemiology, University of Washington, Seattle, WA, USA

[8]Department of Pediatrics, University of Washington, Seattle, WA, USA

[9]Department of Pediatrics and Child Health, University of Nairobi, Kenyatta National Hospital, Nairobi, Kenya

[10]These authors contributed equally

[11]Lead contact

*Correspondence: dlehman@fredhutch.org

https://doi.org/10.1016/j.isci.2021.103615

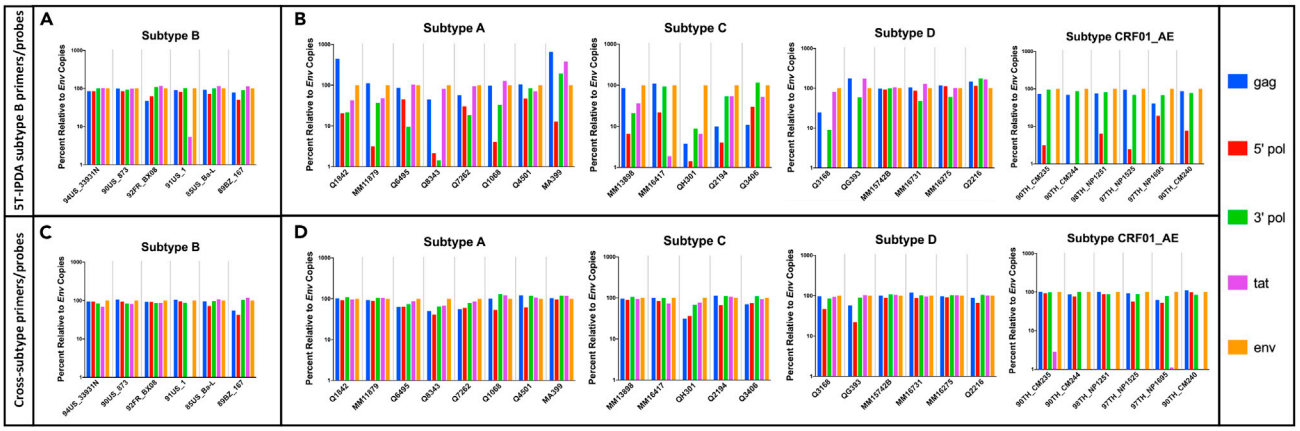


Figure 1. Quantification of a panel of HIV DNA subtype samples

(A–D) Quantification across the 5T-IPDA target regions (colors as in key) relative to the number of copies quantified by the *env* primer/probe set. The 5T-IPDA subtype-B primers and probes were used to quantify HIV DNA samples shown on the x axis derived from (A) patients infected with HIV subtype B viruses and (B) patients infected with HIV subtypes (A, C, and D) and CRF01_AE. The cross-subtype primers and probes were used to quantify HIV DNA samples derived from (C) patients infected with HIV subtype B viruses and (D) patients infected with HIV subtypes (A, C, and D) and CRF01_AE.

RESULTS

Multiplex ddPCR assay for reservoir quantification in subtype B infections

We previously developed a multiplex ddPCR assay designed to estimate the copy number of intact proviral DNA in patients infected with HIV subtype B (Levy et al., 2021). The 5 target IPDA (5T-IPDA) includes two complementary multiplex ddPCRs, each targeting 3 HIV genomic regions. Two of these regions use the same dye for probe detection at different concentrations to enable distinct detection and quantification of each region using a 2-color instrument, and one region is identical between the two multiplex ddPCRs (5 HIV target regions total) (Figure S1). The 5T-IPDA targets an 89bp region in the 3′ end of *pol*, a 72bp region in *tat*, and a 95bp region in *env* in one triplex assay (Assay1), and a 65bp region at the 5′ LTR/*gag* junction, a 107bp region in the 5′ end of *pol*, and the same 95bp region in *env* in a second triplex assay (Assay2). The frequency of intact provirus in the reservoir is estimated by normalizing the number of HIV triple-positive droplets to the number of T cells. The T cell number is determined by a third multiplex ddPCR with 2 targets in a cellular housekeeping gene (*RPP30*) to quantify the total number of cells, and 1 target in a T cell receptor gene that is deleted during T cell development. Together, these values provide an estimate of the number of T cells interrogated (see STAR Methods). Details about the development and use of this subtype B 5T-IPDA have been described (Levy et al., 2021).

Performance of the subtype B 5T-IPDA on non-B subtypes

We tested the performance of the 5T-IPDA subtype B primers and probes on HIV-1 DNA samples derived from short-term cultures of viruses from 31 patients, which included subtypes A (n = 8), B (n = 6), C (n = 5), D (n = 6), and CRF01_AE (n = 6). Because viruses that replicate and propagate in culture are derived from intact proviruses, we expect deletions in these sequences to be rare and limited to those that occur during virus spread in culture plus the few residual patient cells that harbored deleted proviruses. Thus, we expected quantification of all 5 target regions to be roughly equivalent. Indeed, for 5 of the 6 DNA samples from subtype B infections, the subtype B 5T-IPDA primers/probes were able to detect and quantify all 5 target regions within <2.5-fold of each other (Figure 1A). For a single subtype B variant, 91US_1, quantification of *tat* was >15-fold below the other 4 targets, suggesting that in some individuals sequence diversity in *tat* may cause an underestimation of intact proviral DNA when using the subtype B 5T-IPDA. When the subtype B 5T-IPDA primers/probes were used on the 25 non-B subtype DNA samples, quantification was less consistent overall, with an average 841-fold (range: 1–22,367) difference across the 5 target regions (Figure 1B), suggesting the 5T-IPDA subtype B assay required optimization for accurate reservoir quantification of non-B subtypes.

Adaptation of the multiplex IPDA to quantify intact provirus across diverse HIV subtypes

We analyzed the cross-subtype diversity within the binding site sequences of the subtype B 5T-IPDA primers and probes using the Los Alamos National Library (LANL) HIV Database (www.hiv.lanl.gov). Our

analysis included over 2,400 full-length HIV-1 sequences, including subtypes A (n = 171), B (n = 1179), C (n = 719), D (n = 71), and CRF01_AE (n = 292). We identified specific nucleotide positions in which the 5T-IPDA primer/probe sequences differed in >10% of sequences for any given subtype. To generate "cross-subtype" primers spanning the same sequence region, we used an iterative process (see STAR Methods) in which we changed up to six positions total per primer, prioritizing the three bases at the 3′ end, and at most one position per probe (Table S1 and Figure S2).

Performance of our cross-subtype IPDA assay on non-B subtypes

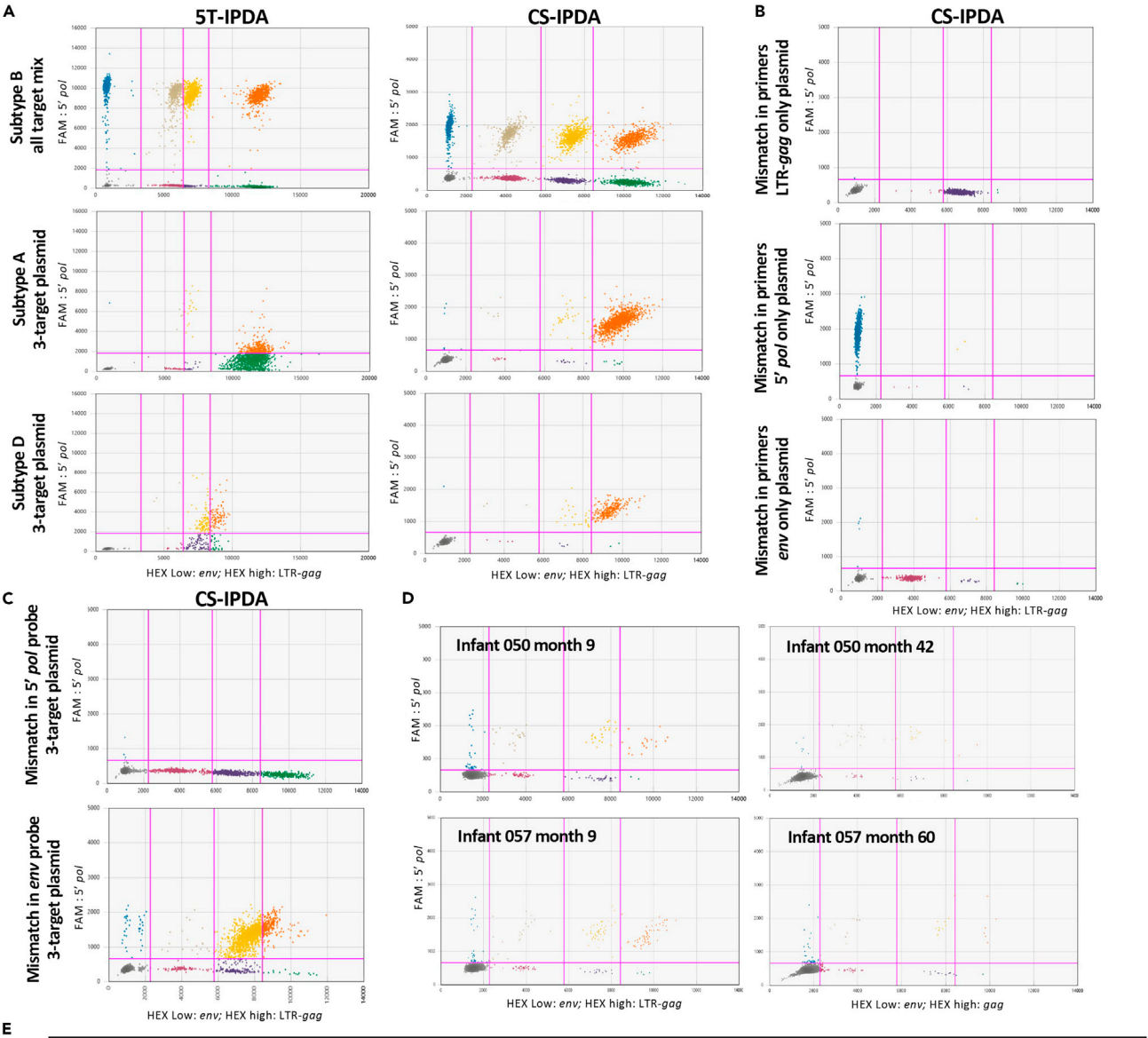
To assess the ability of our cross-subtype primer/probe sets to work across diverse HIV subtypes, we used the same panel of 31 HIV DNA samples described above. For subtype B samples, the cross-subtype primers worked comparably to the subtype B primers (compare Figures 1A and 1C), indicating they did not significantly compromise amplification of subtype B sequences. Importantly, the primer/probe adaptations improved performance on subtypes A, C, and D samples, with an average of 1.4-fold (range 1–4.7) difference between each of the 5 target regions (Figure 1D). Detection and quantification also greatly improved for subtype CRF01_AE viruses, with an average 1.2-fold difference between LTR/*gag*, 5′*pol*, and *env* (Figure 1D). However, quantification of the *tat* region in subtype B and CRF01_AE viruses was not improved by our changes and when we tried to accommodate the known sequence diversity in the *tat* primer/probe binding sites, the necessary changes compromised our ability to quantify the other non-B subtypes (data not shown). Thus, the cross-subtype primers/probes targeting 3′*pol*, *tat*, *and env* (Assay1 of the 5T-IPDA) quantify their targets well for A, C, and D subtypes (average 1.14-fold difference, range 1.0–1.55), but not as consistently for subtypes B and CRF01_AE. In contrast, our cross-subtype primers/probe targeting LTR/*gag*, 5′*pol*, and *env* (Assay2 of the 5T-IPDA) reliably quantify all 3 targets for all subtypes tested within an average of 1.2-fold (range 1–3.0). Thus, we chose to move forward with a 3-target assay using the LTR/*gag*, 5′*pol*, and *env* primers/probes, which we call cross-subtype IPDA (CS-IPDA).

Accuracy, sensitivity, and specificity of the CS-IPDA

In order to assess the accuracy and limit of detection of our CS-IPDA reservoir assay, we used HIV DNA controls from J-Lat 5A8 cells (Chan et al., 2013) (one subtype B HIV proviral genome per cell) estimated to be ~3000, 200, 150, 20, 10, and 1 copy per reaction (see STAR Methods). The intact HIV DNA copies in this series of J-Lat controls quantified by CS-IPDA were significantly correlated with the HIV copies measured by an independent *pol* qPCR assay (Benki et al., 2006) (Pearson's $r^2 > 0.99$, $p < 0.0001$, Figure S3). In addition, the CS-IPDA detected HIV DNA in 42% of 12 replicate J-Lat 5A8 controls estimated to contain an average of 1 copy per reaction, which is within range of the 63% positive reactions expected in single-copy controls based on Poisson distribution (Benki et al., 2006; Hughes and Totten, 2003). We also tested the specificity of the CS-IPDA using HIV-negative PBMC controls (~10^5 cells tested per replicate), and 0 of 34 replicates were positive for intact HIV DNA. Based on the J-Lat control data described, the absolute limit of detection is 1 copy per reaction. Notably, the LOD for HIV reservoir assays depends on the number of cells tested from a given specimen, and thus, if 1 million cells are tested, the LOD is 1 intact copy per million cells, while if only 100,000 total cells are tested, the LOD increases to 10 intact copies per million cells.

In-silico analysis of the CS-IPDA on full-length HIV proviral sequences

To test the ability of the CS-IPDA assay to detect and distinguish intact and defective HIV sequences of multiple subtypes, we performed an *in-silico* analysis of published full-length HIV proviral sequences. We first classified full-length sequences as intact if all 3 primer/probe pairs used in the CS-IPDA met previously published detection criteria, or defective if any of the primer/probe pairs failed to meet the criteria: 5 mismatches from the primer sequences and zero mismatches from the probe (Levy et al., 2021). We then compared that to classification of the same sequences by computational inference of intactness using the ProSeq-IT tool available through the HIV Proviral Sequence Database (https://psd.cancer.gov/tools/pvs_annot.php) (Shao et al., 2020). We did this first on the same set of subtype B proviral sequences used in (Levy et al., 2021). Of 1,071 subtype B proviral sequences analyzed, intactness classification of 967 (90.3%) sequences matched between the CS-IPDA protocol and the ProSeq-IT algorithm (Table S2A). Only 1 (0.1%) subtype B sequence classified intact by ProSeq-IT was classified defective by our assay, and 103 (9.6%) were defective by ProSeq-IT sequence analysis but classified as intact by CS-IPDA. We next performed a similar analysis of currently published non-B proviral sequences, which due to the lack of cross-subtype methods for full-length proviral sequencing, is currently limited to subtype C proviral sequences (n = 508) (Garcia-Broncano et al., 2019). Of the 508 subtype C sequences available, 428 (84.3%) matched by CS-IPDA and sequence analysis (Table S2B). Only 1 (0.2%) subtype C sequence was classified defective by



E

		Forward primer target region	Probe target	Reverse primer target region
gag	**CS-IPDA sequence**	GACTAGCGGAGGCTAGAAGGAGAGA	ATGGGTGCGAGA	CGTCARTATTAAGHGGSGGAAAATTAG
	Subtype B (5T-IPDA match)	GACTAGCGGAGGCTAGAAGGAGAGA	ATGGGTGCGAGA	CGTCAGTATTAAGCGGGGGAGAATTAG
	Subtype A	GACTAGCGGAGGCTAGAAGGAGAGA	ATGGGTGCGAGA	CGTCAGTATTAAGCGGGGGAAAATTAG
	Subtype D	GACTAGCAGAGGCTAGAAGGAGAGA	ATGGGTGCGAGA	CGTCGGTATTAAGCGGGGGAGAATTAG
	mismatch gag primers	GACTAGCGGAGGCTAGAAGGAGAGA	ATGGGTGCGAGA	CGTCAGTATTAAGCGGGGGAAAATTCG
5′ pol	**CS-IPDA sequence**	WCCYTTARYTTCCCTCARATCACTCT	TTGGCARCGACC	GARGCTCTMTTAGAYACAGGAGCAGATGATACAGTA
	Subtype B (5T-IPDA match)	TCCTTTAGCTTCCCTCAGATCACTCT	TTGGCAACGACC	GAAGCTCTATTAGATACAGGAGCAGATGATACAGTA
	Subtype A	ACCCTTAGTTTCCCTCAAATCACTCT	TTGGCAACGACC	GAAGCTCTATTAGATACAGGAGCAGATGATACAGTA
	Subtype D	TCCCTTAGCCTCCCTCAAATCACTCT	TTGGCAACGACC	GAAGCTCTATTAGATACAGGAGCAGATGATACAGTA
	mismatch 5′ pol primers	TCCTTCAATTTCCCTCAAATCACTCT	TTGGCAACGACC	GAAGCCCTATTAGACACAGGAGCAGATGATACAGTA
	mismatch 5′ pol probe	TCCTTTAGCTTCCCTCAGATCACTCT	TTGGGACAGACC	GAAGCTCTATTAGATACAGGAGCAGATGATACAGTA
env	**CS-IPDA sequence**	TVTTCMTTGGGTTCTTRGGAGCAGCAGG	AGCACKATGGG	GGTACAGGCCAGACARBTATTGTCTGGYATAGTGC
	Subtype B (5T-IPDA match)	TGTTCCTTGGGTTCTTGGGAGCAGCAGG	AGCACTATGGG	GGTACAGGCCAGACAATTATTGTCTGGTATAGTGC
	Subtype A	TCTTCCTTGGGTTCTTAGGAGCAGCAGG	AGCACTATGGG	GGTACAGGCCAGACAATTATTGTCTGGCATAGTGC
	Subtype D	TGTTCCTTGGGTTCTTGGGAACAGCAGG	AGCACGATGGG	GGTACAGGCCAGAAATTTATTGTCTGGTATAGTGC
	mismatch env primers	TGATCTTTGGGTTCTTAGGAGCAGCAGG	AGCACTATGGG	GGTACAAGCCAGACAATTGTTGTCTGGTATAGTGC
	mismatch env probe	TGTTCCTTGGGTTCTTGGGAGCAGCAGG	AGTACTATGGG	GGTACAGGCCAGACAATTATTGTCTGGTATAGTGC

Figure 2. Impact of subtype diversity on detection of intact sequences by CS-IPDA

(A) Droplet clusters on mixtures of all possible single-, dual-, and triple-combination plasmids of CS-IPDA targets with subtype B sequences (top), and on 3-target subtype A and D plasmids (middle and bottom) using 5T-IPDA (left) and CS-IPDA (right) primers/probes.

Figure 2. *Continued*

(B) CS-IPDA on single-target plasmids with a sequence from LANL that had the most mismatches relative to CS-IPDA and prevalent at >1% of sequences in the LANL HIV database.

(C) CS-IPDA on 3-target plasmids with mismatches in the 5′*pol* probe (top) and *env* probe (bottom).

(D) CS-IPDA results from 2 Kenyan infants at 2 timepoints each.

(E) Sequences used in the plasmid controls shown above with degenerate bases in the CS-IPDA primers and probes shown in green, bases that mismatch only the 5T-IPDA in blue, bases that mismatch only the CS-IPDA in orange, and bases that mismatch both 5T-IPDA and CS-IPDA in red.

our assay that was intact by sequence analysis, and 79 (15.5%) were classified as intact by CS-IPDA but were defective by sequencing. Of the 1,579 subtype B and C sequences combined, our CS-IPDA misclassified as intact 182 (11.5%) sequences that ProSeq-IT called defective based on premature stop codons, large deletions or insertions, a frameshift in a gene, or missing or mutated splice donor site. Notably, our CS-IPDA missed only 2 intact sequences of all the proviral subtype B and C sequences analyzed.

Impact of polymorphisms in the CS-IPDA primer and probe binding sites

As described above, we introduced degenerate bases at sites of diversity across the CS-IPDA primer/probe binding sites (Figure S2). To assess the ability of these degenerate primer/probe sets to clearly define clusters of single-, dual-, and triple-target positive droplets when known polymorphisms are present in the binding sites, we compared the use of the 5T-IPDA LTR/*gag*, 5′*pol*, and *env* primer/probes to the CS-IPDA primer/probes on engineered plasmids (Figure 2). First, we compared the amplitude and separation of droplet clusters for a mix of all target combinations of subtype B plasmids, which serve as a control to set gates in both our 5T-IPDA and CS-IPDA assays (Figure 2A, top). While the individual clusters were well defined for both assays, the amplitude of both the FAM and HEX signals for the CS-IPDA was significantly lower than that of the 5T-IPDA (compare 5T-IPDA vs CS-IPDA plots in top of Figure 2A). We assessed the independent effects of the degenerate bases in the primers versus the probes by using the 5T-IPDA primers in combination with the CS-IPDA probes (with degenerate bases) and observed lower amplitude signals similar to the CS-IPDA, while the combination of CS-IPDA degenerate primers with 5T-IPDA probes (no degenerate bases) resulted in higher amplitudes comparable to those observed with 5T-IPDA (data not shown), suggesting that it is predominately the degenerate bases in the probes and not the primers that cause the shift in amplitude.

To further understand how diversity across subtypes impacts the ability to detect intact proviruses, we used both the 5T-IPDA and CS-IPDA primer/probe sets on 3-target plasmids that were engineered to contain known sequences of subtypes A and D (Figure 2A). For both the subtype A and D 3-target plasmids, the CS-IPDA results in a distinct triple-positive cluster that falls within the gates defined by the subtype B plasmids. However, when using the 5T-IPDA primers/probes, the triple-positive cluster significantly shifts (Figure 2A), resulting in misclassification of triple-positive clusters as dual- and single-positive due to the polymorphisms relative to the 5T-IPDA primers (both red plus blue bases highlighted in Figure 2E).

We next wanted to determine whether the CS-IPDA could tolerate the known diversity across subtypes as documented in the LANL HIV database. For each of the 6 primers and 3 probes, we identified a sequence in LANL that had the most mismatches relative to our CS-IPDA and were at a prevalence in the HIV-1 database of $\geq$1%. We engineered single-target plasmids with these mismatches in the primers only, as well as triple-target plasmids with mismatches in the 5′*pol* or *env* probe as shown (Figure 2E). When we used the CS-IPDA on the single-target plasmids that contained the primer-only mismatches, we see distinct single-target clusters that did not shift in amplitude relative to the clusters in the subtype B gating controls (compare Figures 2B to 2A top right). We next tested 2 plasmids with no mismatches in the primers and either 3 mismatches in the 5′*pol* probe or 1 mismatch in the *env* probe (Figure 2E) and observed complete abrogation of the *pol* signal (Figure 2C top) or a shift in the *env* signal (Figure 2C bottom), respectively, thus causing misclassification of intact sequences as defective. Importantly, mismatches to our CS-IPDA probe sequences are rare (Figure S2). Indeed, when we queried all HIV-1 sequences covering our probe regions currently in the LANL HIV-1 database, we found only 73 (1.96%) of 3720 LTR/*gag* probe, 122 (2.8%) of 4310 5′ *pol* probe, and 158 (3.7%) of 4310 *env* probe sequences in LANL have 1 or more mismatches to our CS-IPDA probe sequences, and thus would be missed by our assay.

Application of the cross-subtype IPDA to samples from Kenyan infants on ART

We tested the CS-IPDA on longitudinal blood samples from Kenyan infants infected with HIV subtypes A (n = 5) and D (n = 1) that had HIV RNA levels suppressed below detection (<150 copies/mL) during up to

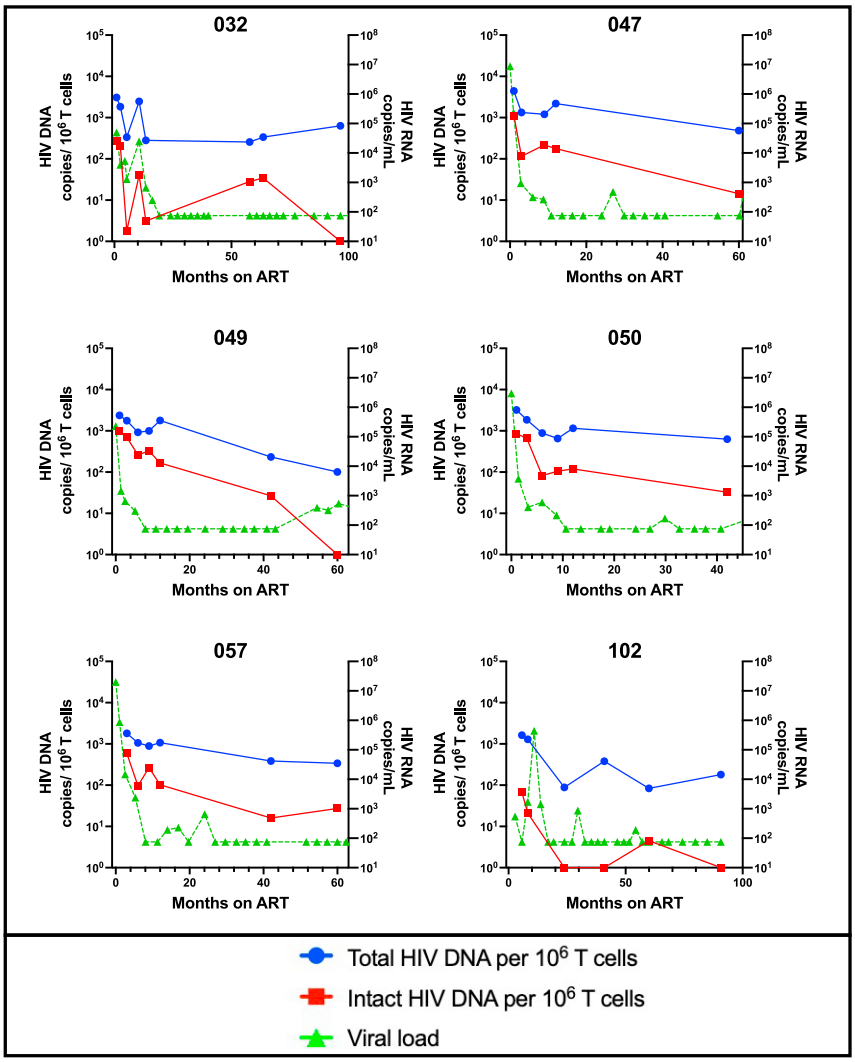


Figure 3. Longitudinal patterns of total and intact HIV DNA and HIV RNA during ART in Kenyan infants

CS-IPDA quantification of HIV in PBMC samples from 5 to 8 timepoints over 42–96 months of ART treatment in 6 HIV-infected Kenyan infants infected with subtypes A or D. Patterns of total HIV DNA (blue circles) and intact HIV DNA (red squares) in copies per million T cells (left Y axis). HIV RNA (green triangles) copies per mL (right Y axis, limit of detection 150 copies/mL).

96 months on ART (Figure 3). In all infants tested, the CS-IPDA detected both intact and defective HIV sequences at multiple timepoints (see examples of droplet clusters in Figure 2D). In 3 of the 6 cases, there were 1–3 timepoints during suppressive ART with no triple-positive droplets detected—that is, intact HIV DNA was not detected in these samples (ID [month]: 032[m99], 049[m60], and 102 [m24, 42, 99]), despite the fact that a median of 1.3×10^5 (range 1.1×10^5– 5.7×10^5) cells were interrogated and DNA shearing rates (which is corrected for in our assay, see STAR Methods) were low: median 5.2% (range 1.7%–8.5%) of DNA was sheared. Importantly, in all individuals with undetectable intact HIV DNA levels, including one case (049[m60]) with undetectable intact HIV DNA in the presence of detectable HIV RNA, the CS-IPDA did detect triple-positive droplets at earlier timepoints, and single- and/or dual-positive droplets were detected for all 3 targets at all timepoints, showing that the primers and probes were able to recognize the specific HIV sequences in these samples.

For all 6 infants, the levels of total HIV DNA were consistently higher than the levels of intact proviral DNA, with a median intact to total HIV DNA ratio of 7% (range 1%–33%) during HIV RNA suppression. The dynamics revealed by the CS-IPDA are in general agreement with typical trends of HIV DNA

(Blankson et al., 2000; Siliciano et al., 2003; Uprety et al., 2015, 2017) during ART (Figure 3). During the first year of ART, HIV RNA dropped in several phases, and DNA correspondingly decreased. Once RNA levels were below detectable limits, DNA levels tended to flatten. Taken together, this data strongly suggest the CS-IPDA is able to tolerate the proviral diversity present in these children known to be infected with subtypes A and D.

DISCUSSION

Recent recommendations for measuring the latent reservoir in clinical HIV cure studies place IPDAs in the highest priority category because they require a relatively small number of cells and are both high-throughput and cost-effective (Abdel-Mohsen et al., 2020). However, to date, IPDAs have been designed specifically for HIV subtype B, even though non-B subtypes predominate in areas of the world with the highest HIV prevalence such as sub-Saharan Africa and Asia, where the need for a cure is critical (Ndung'u et al., 2019). Thus, we adapted an IPDA to accommodate the known diversity across the major HIV-1 subtypes and demonstrate consistent amplification across the 3 CS-IPDA target regions (*LTR/gag*, 3′ end of *pol*, and *env*) when tested on subtypes A, B, C, D, and CRF01_AE (Figure 1). While our CS-IPDA uses only 3 targets instead of the 5 targets used in the 5T-IPDA, the use of 3 targets adds specificity beyond the 2-target IPDAs commonly used (Bruner et al., 2019). We utilize the same dye for 2 of the probes at different concentrations to distinguish these 3 genomic regions using a 2-color ddPCR instrument, which are more likely than ≥3-color instruments to be available in regions of the world with non-B subtypes—however, we acknowledge that adapting to a ≥3-color instrument as they become more available would be ideal. Based on detection of single-copy positive controls and the lack of false-positives in our negative controls, we set the absolute limit of detection to 1 intact copy per reaction. Of note, even if we use probit analysis (a more conservative approach), the lower limit is 3 intact copies per reaction (data not shown). Because the LOD for HIV reservoir assays depends on the number of cells tested, we are confident that our CS-IPDA is sensitive enough to detect as low as 1–3 or 10–30 intact copies per million cells when 1 million or 100,000 cells, respectively, are interrogated.

In-silico analysis of published subtype B and C proviral sequences showed that while our CS-IPDA overestimates the quantity of intact sequences due to mutations or deletions outside of the CS-IPDA target regions, only 2 of 1579 (0.13%) provirus sequences were incorrectly classified as defective by our assay (Table S2). To further confirm that the CS-IPDA can tolerate the expected diversity across subtypes, we created a series of in-vitro synthesized DNA templates with known sequences representing multiple subtypes, as well as templates with the maximum proviral polymorphisms in HIV-1 sequences prevalent at >1% in the LANL HIV database. We show that despite known polymorphisms in our primer binding sites, the ddPCR droplets cluster as expected based on the known template input and are well defined (Figure 2). When we introduced mismatches in probe sequences, the droplet clusters shifted, causing the misclassification of truly intact sequences as defective. However, all PCR-based assays are susceptible to false-negatives based on polymorphisms in the binding sites. The fact that our CS-IPDA primers can tolerate the expected number of polymorphisms in sequences present at >1% of those in the LANL database, and that polymorphisms that impact detection by our 3 probes are infrequent (<4% of sequences have ≥1 mismatch in our probe targets), we expect underestimation of the intact reservoir by our assay due to sequence diversity across subtypes to be minimal.

In addition, when the CS-IPDA was applied to longitudinal samples from Kenyan infants during suppressive ART, detectable intact HIV DNA levels were a median of 7% of the total HIV DNA levels, and the temporal patterns of both total and intact HIV DNA follow closely with the patterns of plasma HIV RNA in these non-B subtype infections. This is consistent with other recent studies that used other IPDA assays in ART-treated HIV-infected adults in the USA, which showed intact provirus levels are on average 8% of total HIV DNA and correlate with both total HIV DNA and plasma HIV RNA levels over time (Gandhi et al., 2021; Simonetti et al., 2020). As expected, the patterns of both total and intact HIV DNA decay generally follow the dynamics reported in both pediatric (Uprety et al., 2017) and adult studies (Blankson et al., 2000; Siliciano et al., 2003) (Figure 3), with faster decay during the initial year following ART initiation, slower decay during long-term HIV RNA suppression, and evidence of differential decay of total and intact proviruses (Antar et al., 2020; Gandhi et al., 2021; Peluso et al., 2020). Unfortunately, our data presented here is limited to 6 children with only 1–2 samples available after >6 months of HIV RNA suppression, and thus we were not able to accurately estimate the half-life of HIV reservoir decay. Larger studies on dynamics of intact HIV DNA decay in the HIV reservoir of children are warranted.

Taken together, our data suggest that the CS-IPDA will accurately estimate intact HIV DNA in blood samples without the need to determine the subtype first. The CS-IPDA enables the quantification of intact HIV DNA across the predominant global HIV subtypes, which is key to expanding HIV cure research to regions of the world where HIV prevalence is high and non-B subtypes predominate (Gartner et al., 2020).

Limitations of the study

We were not able to compare our CS-IPDA to the gold standard of full-length proviral sequencing (FLIPS) (Hiener et al., 2017), as cross-subtype primers for FLIPS have not yet been developed. It is known that IPDA assays can overestimate intact HIV DNA due to defective regions of the genome outside of the primer/probe binding sites. In addition, as with all PCR-based assays, sequence diversity in the primer and probe binding sites can cause amplification failure resulting in an underestimation of intact HIV DNA. While our in-silico analysis suggests that our CS-IPDA is concordant with ~90% of full-length sequencing data (Table S2), studies that directly compare CS-IPDA and FLIPS will better inform the accuracy of the CS-IPDA.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - HIVDNA from primary cells
  - HIVDNA from cell lines
  - Synthesized DNA plasmids
  - Human subject samples
  - Protection of human subjects
- METHOD DETAILS
  - Adaptation of 5T-IPDA for cross-subtype diversity
  - DNA extraction
  - HIV pol qPCR and HIVRNA viral loads
  - Multiplex ddPCR assays
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of ddPCR assay results
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103615.

ACKNOWLEDGMENTS

We thank the study participants as well as the administrative, clinical, and data teams in Kenya involved in the clinical cohort studies for their dedication and support. This work was supported by the National Institutes of Health (NIH) grants R01 HD094718 (to DAL, GJS), R01 DA040386 and AI116292 (to FH), UM1 A126623 (to KRJ), P30 AI027757, K25 AI155224 (to DBR), R01 HD023412 (to GJS), and R37 A138518 (to JO).

AUTHOR CONTRIBUTIONS

DL and JO conceived the study. DL, JO, NC, KJ, FH, CL, and SH contributed to the design of the study. NC and CF designed and performed experiments. NC, CF, CL, PR, DR, SH, and DL contributed analysis tools and data analysis. DW, GJS, and SBN developed and managed the OPH cohort and contributed patient samples. DL, JO, NC, CF, CL, PR, DR, SH, JS, SBN, GJS, DW, JK, and FH contributed to analysis and interpretation of data. NC and DL wrote the manuscript with input from all authors.

DECLARATIONS OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as a member of the LGBTQ + community. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

Received: May 14, 2021
Revised: October 15, 2021
Accepted: December 9, 2021
Published: January 21, 2022

REFERENCES

Abdel-Mohsen, M., Richman, D., Siliciano, R.F., Nussenzweig, M.C., Howell, B.J., Martinez-Picado, J., Chomont, N., Bar, K.J., Yu, X.G., Lichterfeld, M., et al. (2020). Recommendations for measuring HIV reservoir size in cure-directed clinical trials. Nat. Med. *26*, 1339–1350.

Antar, A.A., Jenike, K.M., Jang, S., Rigau, D.N., Reeves, D.B., Hoh, R., Krone, M.R., Keruly, J.C., Moore, R.D., Schiffer, J.T., et al. (2020). Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. J. Clin.Invest. *130*, 3543–3559.

Benki, S., McClelland, R.S., Emery, S., Baeten, J.M., Richardson, B.A., Lavreys, L., Mandaliya, K., and Overbaugh, J. (2006). Quantification of genital human immunodeficiency virus type 1 (HIV-1) DNA in specimens from women with low plasma HIV-1 RNA levels typical of HIV-1 nontransmitters. J. Clin. Microbiol. *44*, 4357–4362.

Blankson, J.N., Finzi, D., Pierson, T.C., Sabundayo, B.P., Chadwick, K., Margolick, J.B., Quinn, T.C., and Siliciano, R.F. (2000). Biphasic decay of latently infected CD4+ T cells in acute human immunodeficiency virus type 1 infection. J. Infect. Dis. *182*, 1636–1642.

Brown, B.K., Darden, J.M., Tovanabutra, S., Oblander, T., Frost, J., Sanders-Buell, E., de Souza, M.S., Birx, D.L., McCutchan, F.E., and Polonis, V.R. (2005). Biologic and genetic characterization of a panel of 60 human immunodeficiency virus type 1 isolates, representing clades A, B, C, D, CRF01_AE, and CRF02_AG, for the development and assessment of candidate vaccines. J. Virol. *79*, 6089–6101.

Bruner, K.M., Murray, A.J., Pollack, R.A., Soliman, M.G., Laskey, S.B., Capoferri, A.A., Lai, J., Strain, M.C., Lada, S.M., Hoh, R., et al. (2016). Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat. Med. *22*, 1043–1049.

Bruner, K.M., Wang, Z., Simonetti, F.R., Bender, A.M., Kwon, K.J., Sengupta, S., Fray, E.J., Beg, S.A., Antar, A.A.R., Jenike, K.M., et al. (2019). A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. Nature *566*, 120–125.

Chan, J.K., Bhattacharyya, D., Lassen, K.G., Ruelas, D., and Greene, W.C. (2013). Calcium/calcineurin synergizes with prostratin to promote NF-kappaB dependent activation of latent HIV. PLoS One *8*, e77749.

Gandhi, R.T., Cyktor, J.C., Bosch, R.J., Mar, H., Laird, G.M., Martin, A., Collier, A.C., Riddler, S.A., Macatangay, B.J., Rinaldo, C.R., et al. (2021). Selective decay of intact HIV-1 proviral DNA on antiretroviral therapy. J.Infect. Dis. *223*, 225–233.

Garcia-Broncano, P., Maddali, S., Einkauf, K.B., Jiang, C., Gao, C., Chevalier, J., Chowdhury, F.Z., Maswabi, K., Ajibola, G., Moyo, S., et al. (2019). Early antiretroviral therapy in neonates with HIV-1 infection restricts viral reservoir size and induces a distinct innate immune profile. Sci. Transl. Med. *11*, eaax7350.

Gartner, M.J., Roche, M., Churchill, M.J., Gorry, P.R., and Flynn, J.K. (2020). Understanding the mechanisms driving the spread of subtype C HIV-1. EBioMedicine *53*, 102682.

Green, M.R., and Sambrook, J. (2016). Precipitation of DNA with ethanol. Cold Spring Harb.Protoc. *2016*. https://doi.org/10.1101/pdb.prot093377.

Hiener, B., Horsburgh, B.A., Eden, J.S., Barton, K., Schlub, T.E., Lee, E., von Stockenstrom, S., Odevall, L., Milush, J.M., Liegler, T., et al. (2017). Identification of genetically intact HIV-1 proviruses in specific CD4(+) T cells from effectively treated participants. Cell Rep. *21*, 813–822.

Ho, Y.C., Shan, L., Hosmane, N.N., Wang, J., Laskey, S.B., Rosenbloom, D.I., Lai, J., Blankson, J.N., Siliciano, J.D., and Siliciano, R.F. (2013). Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell *155*, 540–551.

Hughes, J.P., and Totten, P. (2003). Estimating the accuracy of polymerase chain reaction-based tests using endpoint dilution. Biometrics *59*, 505–511.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. Br. Med. Bull. *58*, 19–42.

Laird, G.M., Eisele, E.E., Rabi, S.A., Lai, J., Chioma, S., Blankson, J.N., Siliciano, J.D., and Siliciano, R.F. (2013). Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay. PLoS Pathog. *9*, e1003398.

Laird, G.M., Rosenbloom, D.I., Lai, J., Siliciano, R.F., and Siliciano, J.D. (2016). Measuring the frequency of latent HIV-1 in resting CD4$^+$ T cells using a limiting dilution coculture assay. Methods Mol. Biol. *1354*, 239–253.

Levy, C., Hughes, S., Roychoudry, P., Reeves, D., Amstuz, C., Zhu, H., Huang, M., Wei, Y., Bull, M., Cassidy, N., et al. (2021). A highlty multiplexed droplet digital PCR assay to measure the intact HIV-1 proviral reservoir. Cell Rep. Med. *2*, 100243.

Martin, H.L., Nyange, P.M., Richardson, B.A., Lavreys, L., Mandaliya, K., Jackson, D.J., Ndinya-Achola, J.O., and Kreiss, J. (1998). Hormonal contraception, sexually transmitted diseases, and risk of heterosexual transmission of human immunodeficiency virus type 1. J. Infect. Dis. *178*, 1053–1059.

Nduati, R., John, G., Mbori-Ngacha, D., Richardson, B., Overbaugh, J., Mwatha, A., Ndinya-Achola, J., Bwayo, J., Onyango, F.E., Hughes, J., et al. (2000). Effect of breastfeeding and formula feeding on transmission of HIV-1: a randomized clinical trial. JAMA *283*, 1167–1174.

Ndung'u, T., McCune, J.M., and Deeks, S.G. (2019). Why and where an HIV cure is needed and how it might be achieved. Nature *576*, 397–405.

Neilson, J.R., John, G.C., Carr, J.K., Lewis, P., Kreiss, J.K., Jackson, S., Nduati, R.W., Mbori-Ngacha, D., Panteleeff, D.D., Bodrug, S., et al. (1999). Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. J. Virol. *73*, 4393–4403.

Pankau, M.D., Wamalwa, D., Benki-Nugent, S., Tapia, K., Ngugi, E., Langat, A., Otieno, V., Moraa, H., Maleche-Obimbo, E., Overbaugh, J., et al. (2018). Decay of HIV DNA in the reservoir and the impact of short treatment interruption in Kenyan infants. Open Forum Infect. Dis. *5*, ofx268.

Peluso, M.J., Bacchetti, P., Ritter, K.D., Beg, S., Lai, J., Martin, J.N., Hunt, P.W., Henrich, T.J., Siliciano, J.D., Siliciano, R.F., et al. (2020). Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. JCI Insight *5*, e132997.

Rainwater, S., DeVange, S., Sagar, M., Ndinya-Achola, J., Mandaliya, K., Kreiss, J.K., and Overbaugh, J. (2005). No evidence for

rapid subtype C spread within an epidemic in which multiple subtypes and intersubtype recombinants circulate. AIDS Res. Hum. Retroviruses *21*, 1060–1065.

Shao, W., Shan, J., Hu, W.S., Halvas, E.K., Mellors, J.W., Coffin, J.M., and Kearney, M.F. (2020). HIV proviral sequence database: a new public database for near full-length HIV proviral sequences and their meta-analyses. AIDS Res. Hum. Retroviruses *36*, 1–3.

Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J., and Siliciano, R.F. (2003). Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat. Med. *9*, 727–728.

Simonetti, F.R., White, J.A., Tumiotto, C., Ritter, K.D., Cai, M., Gandhi, R.T., Deeks, S.G., Howell, B.J., Montaner, L.J., Blankson, J.N., et al. (2020). Intact proviral DNA assay analysis of large cohorts of people with HIV provides a benchmark for the frequency and composition of persistent proviral DNA. Proc. Natl. Acad. Sci. U S A *117*, 18692–18700.

Uprety, P., Chadwick, E.G., Rainwater-Lovett, K., Ziemniak, C., Luzuriaga, K., Capparelli, E.V., Yenokyan, G., and Persaud, D. (2015). Cell-associated HIV-1 DNA and RNA decay dynamics during early combination antiretroviral therapy in HIV-1-Infected infants. Clin.Infect. Dis. *61*, 1862–1870.

Uprety, P., Patel, K., Karalius, B., Ziemniak, C., Chen, Y.H., Brummel, S.S., Siminski, S., Van Dyke, R.B., Seage, G.R., and Persaud, D. (2017). Human immunodeficiency virus type 1 DNA decay dynamics with early, long-term virologic control of perinatal infection. Clin.Infect. Dis. *64*, 1471–1478.

Wamalwa, D., Benki-Nugent, S., Langat, A., Tapia, K., Ngugi, E., Moraa, H., Maleche-Obimbo, E., Otieno, V., Inwani, I., Richardson, B.A., et al. (2016). Treatment interruption after 2-year antiretroviral treatment initiated during acute/early HIV in infancy. AIDS *30*, 2303–2313.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
90US_873	NIH HIV Reagent Program	Cat# ARP-11251
91US_1	NIH HIV Reagent Program	Cat# ARP- 7686
92FR_BX08	NIH HIV Reagent Program	Cat# ARP- 11420
94US_33931N	NIH HIV Reagent Program	Cat# ARP -11,250
85US_Ba-L	NIH HIV Reagent Program	Cat# ARP-510
89BZ_167	NIH HIV Reagent Program	Cat# ARP- 7692
90TH_CM235	NIH HIV Reagent Program	Cat# ARP-7701
90TH_CM244	NIH HIV Reagent Program	Cat# ARP-11268
98TH_NP1251	NIH HIV Reagent Program	Cat# ARP-11271
97TH_NP1525	NIH HIV Reagent Program	Cat# ARP-11274
97TH_NP1695	NIH HIV Reagent Program	Cat# ARP-11275
90TH_CM240	NIH HIV Reagent Program	Cat# ARP-7703
Biological samples		
Peripheral blood mononuclear cells (HIV negative)	Bloodworks Northwest	https://www.bloodworksnw.org
Chemicals, peptides, and recombinant proteins		
Buffer EB	QIAGEN	Cat# 19,086
3M guanidine HCl	Thermo Fisher	Cat# 24,115- 200
6M guanidinium isothiocyanate	Sigma	Cat# 50,983
Proteinase K solution, 20mg/mL, > 318mAu/mL at 30 C	QIAGEN	Cat# 19,131
3M sodium acetate	Thermo Fisher	Cat# S210
NEBuffer 3.1	New England BioLabs	Cat# B7203S
AatII	New England BioLabs	Cat# R0117S
ddPCR supermix for probes (no dUTP)	Bio-Rad	Cat# 1,863,024
BglI	New England BioLabs	Cat# R0143S
Experimental models: Cell lines		
J-lat 8.4 cell line	NIH HIV Reagent Program	Cat#9847-427; RRID: CVCL_8284
Oligonucleotides		
HIV-1 primers/probes	Integrated DNA Technologies	Table S1
RPP30 gene primers/probes	Levy et al.[10]	N/A
Recombinant DNA		
Seattle_IPDA_control_1_001	Levy et al. (2021)	Addgene ID: 167,347
Seattle_IPDA_control_6_002	Levy et al. (2021)	Addgene ID: 167,348
Seattle_IPDA_control_9_003	Levy et al. (2021)	Addgene ID: 167,349
Seattle_IPDA_control_14_004	Levy et al. (2021)	Addgene ID: 167,350
Seattle_IPDA_control_17_005	Levy et al. (2021)	Addgene ID: 167,351
Seattle_IPDA_control_29_006	Levy et al. (2021)	Addgene ID: 167,352
Seattle_IPDA_control_32_007	Levy et al. (2021)	Addgene ID: 167,353
Software and algorithms		
QuantaSoft analysis pro (version 1.0.596.0525)	Bio-Rad	https://www.bio-rad.com/en-us/product/qx200-droplet-digital-pcr-system
R (version 4.0.0)	http://www.r-project.org/	RRID: SCR_001905

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Microsoft office	https://www.microsoft.com/en-us/ N/A microsoft-365	N/A
Biorender	https://biorender.com/	N/A
Code for analyzing data	This paper	https://github.com/LehmanLab/CS-IPDA
Other		
QIAGEN QIAquick gel extraction kit	QIAGEN	Cat# 28,704
PCR plate heat seal, foil, pierceable	Bio-Rad	Cat#1814040
Bio-rad PX1 PCR plate sealer	Bio-Rad	Cat#1814000

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dara Lehman (dlehman@fredhutch.org)

Materials availability

All subtype-specific plasmid sequences generated for this study are available from the Lead Contact.

Data and code availability

- Data generated from this study are available from the Lead Contact upon request.
- All original code is available at https://github.com/LehmanLab/CS-IPDA
- Additional information required to reanalyze data reported in this paper is available from the Lead Contact upon request

EXPERIMENTAL MODEL AND SUBJECT DETAILS

HIVDNA from primary cells

The subtype A, C and D HIV DNA samples (n = 19) were obtained by short-term co-culture of archived peripheral blood mononuclear cell (PBMC) samples from HIV-infected Kenyan cohorts (Martin et al., 1998; Nduati et al., 2000) with HIV-negative donor PBMCs. Subtyping of these Kenyan samples was based on Sanger sequencing of *env* (Neilson et al., 1999; Rainwater et al., 2005). For subtype B and CRF01_AE DNA samples (n = 12), primary isolates obtained from short term PBMC co-cultures from the NIH International Panel of HIV-1 Isolates (Brown et al., 2005) were used to infect HIV-negative donor PBMCs at an MOI of 0.02. Infected PBMCs were collected on day 19 of virus culture and DNA was extracted as described below.

HIVDNA from cell lines

DNA was extracted as described below from 1 × 10^6 J-lat 5A8 cells (each containing a single subtype B HIV proviral genome) quantified by hemocytometer cell count (Chan et al., 2013). After extraction, DNA was diluted to an estimated 1, 10, 20, 150, 200 or 3,000 copies of HIV per 2μL. The concentration of HIV in each control was then refined based on quantification by a well-validated independent *pol* ddPCR (Benki et al., 2006; Pankau et al., 2018) that targets a region in *pol* that is completely non-overlapping with the CS-IPDA *pol* primers. Our single-copy of HIV per 2μL control was verified by the independent *pol* ddPCR detecting a positive copy in 5 of 12 (42%) replicates which is within the range of the 63% positive reactions expected based on both a true concentration of 1 copy of HIV per 2μL added to each reaction and the probability of pipetting at least one copy into a reaction based on the Poisson distribution (Benki et al., 2006; Hughes and Totten, 2003).

Synthesized DNA plasmids

The subtype B control plasmid mixtures which include all combinations of the CS-IPDA single-, dual- and triple-targets were generated previously (Levy et al., 2021) and are described in the key resources table. Here we designed additional single-target and 3-target plasmids each with sequences as shown in

Figure 2E that were derived from known sequences as follows: Subtype A and D plasmids contain all 3 target regions with sequences from primary isolates from Kenyan cohorts (Q23 and 2059: GenBank accession numbers AF004885 and AF133821, respectively) (Neilson et al., 1999). The "mismatch" plasmids include mismatches in each primer and probe target region as found in the sequence with the greatest number of mismatches relative to the CS-IPDA primers/probes that is prevalent at >1% in the LANL database as determined by interrogating sequences in LANL using the pre-curated "Filtered Web Alignment" of HIV-1 sequences (https://www.hiv.lanl.gov/content/sequence/QUICK_ALIGNv2/QuickAlign.html). All plasmids were synthesized by Integrated DNA Technologies (IDT). To minimize intermediate droplets ("rain") associated with plasmid DNA, we linearized the plasmids with *Aat*II (New England BioLabs), gel isolated and then extracted the digested DNA using the QIAGEN QIAquick Gel Extraction Kit (Qiagen).

Human subject samples

PBMCs from blood samples were collected from Kenyan infants who initiated ART between 3 and 12 months of age during participation in the Optimizing Pediatric HAART (OPH) study (NCT00428116) as previously described (Wamalwa et al., 2016). Blood was collected at the time of ART initiation, 1 month after ART start, and 3-monthly intervals thereafter. PBMC samples for this laboratory sub-study were from 6 children (3 female, 3 male) with HIV RNA levels suppressed below detection (<150 copies/mL) that had samples available during up to 96 months on ART and were infected with subtypes A (n = 5) and D (n = 1). CS-IPDA was performed in duplicate on all samples, with additional replicates performed on samples with no intact HIV DNA detected until a minimum of 10^5 cells were tested.

Protection of human subjects

Written informed consent was obtained from all participants or caregivers of children. This study was approved by the ethical review committees at Kenyatta National Hospital, the University of Washington, and the Fred Hutchinson Cancer Research Center.

METHOD DETAILS

Adaptation of 5T-IPDA for cross-subtype diversity

Using full-length HIV sequences from the Los Alamos National Laboratories (www.hiv.lanl.gov), including subtypes A (n = 171), B (n = 1179), C (n = 719), D (n = 71) and CRF01_AE (n = 292), we identified specific nucleotide positions that differed from the 5T-IDPA subtype B primer/probe sequences in >10% of sequences for any given subtype. We started by changing up to 4 of the most diverse bases within each primer or probe to the appropriate single or degenerate base, prioritizing changes in the three bases at the 3′ end of each primer (Table S1). We then tested this first generation of our cross-subtype primers on subtype A, C and D HIV DNAs from the subtype DNA panel described above and assessed for amplification efficiency. For primer/probe sets in which the difference in quantification across the 5 target regions was still >3 fold, we used additional degenerate bases at up to 6 positions total per primer and at most 1 per probe (Table S1) to create a set of cross-type primers and probes. In addition, we switched the HEX concentrations of the *env* and LTR/*gag* probes from the subtype B 5T-IPDA, and used HEX low concentration for *env* and HEX high concentration for LTR/*gag* in the CS-IPDA to accommodate amplitude shifts caused by the degenerate base in the env probe and allow for distinct cluster separation along the HEX axis.

DNA extraction

With the exception of the samples noted below, DNA was extracted by a guanidine-based method optimized to minimize DNA shearing as described(Green and Sambrook, 2016). Briefly, cell pellets are resuspended in 3M guanidine HCL, 20mM proteinase K is added and samples are mixed and incubated for 1 hour at 56C, after which guanidine isothiocyanate is added to 4.6M and incubated at 42C for 10 minutes. DNA is then precipitated using isopropanol followed by a 70% ethanol wash step before resuspending in buffer EB (Qiagen). For the 19 subtype A, C and D samples in the subtype panel, DNA was extracted by a Qiagen column according to the manufacturer's instructions.

HIV pol qPCR and HIVRNA viral loads

HIV DNA was quantified in duplicate by an independent in-house cross-subtype *pol* quantitative PCR as previously described (Benki et al., 2006). Plasma HIV RNA was quantified using the Gen-Probe HIV-1 RNA assay (Gen Probe, San Diego, CA) with a limit of detection (LOD) of 150 copies/mL, and samples below the LOD were set to the midpoint between zero and the LOD (75 copies/mL).

Multiplex ddPCR assays

Extracted DNA was digested overnight with *BglI* (which cuts in <8% of subtype B HIV genomes or and not at all in the housekeeping gene targets), purified by ethanol precipitation and resuspended in Qiagen buffer EB. The ddPCR assays were set up following manufacturer instructions (Bio-Rad) modified for multiplex assays as follows for each reaction: 10uL Supermix for probes (no dUTP), 3uL of the primer/probe cocktail for either HIV Assay1 (3′*pol, tat, env*), HIV Assay 2 (LTR/*gag,* 5′*pol, env)* or *RPP30*/deltaD (5′*RPP30,* 3′*RPP30*, deltaD), template DNA for HIV or *RPP30* assays, and H_2O to 20uL total volume per reaction. Droplet generation was as recommended (Bio-Rad), and PCR was 65 cycles of 94°C for 30 seconds and 60°C for 60 seconds followed by 98°C for 10 minutes. HIV assays were run in either duplicate or triplicate and RPP30/delta D was run in singlet for each sample tested. In addition, 2 gating and 2 assay controls were included on each plate as described (Levy et al., 2021).

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of ddPCR assay results

Gates for single, dual- and triple-positive droplets were set in QuantaSoft Analysis Pro (version 1.0.596.0525) based on gating controls included on each plate. These gates were then applied to results from experimental samples included in that plate. Cluster and well data were exported from QuantaSoft Analysis Pro and all subsequent data manipulation was carried out using R version 4.0.0.

To assess the ability of our primers and probe sets to accurately quantify each of the 5 HIV target regions across the HIV subtype panel samples, we determined the total number of copies for each target region ddPCR in a sample and normalized the data to the number of copies detected by the *env* ddPCR (percent HIV target region copies relative to *env* in Figure 1). To determine the concentration of both total and intact HIV (Figure 3), DNA copies per uL for each "cluster" of either single, dual- or triple-combination of targets were calculated by modeling a Poisson distribution as described in https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf:

$$\textit{Target copies}/\mu L = -\ln(\textit{droplets not in target cluster}/\textit{total droplets in well})/\textit{droplet volume}$$

The total HIV DNA is the sum of the concentration of positive droplets in each cluster and intact HIV DNA is the concentration of triple-positive droplets. For the RPP30/deltaD assay, the *RPP30* targets are present in all cells and the deltaD region is present in all non-T cells. Thus, the total number of cells interrogated was the 5′*RPP30* positive droplets and the number of T cells in each sample was the total number of genomes (RPP30-positive droplets) minus the number of non-T cell genomes (deltaD-positive droplets). We then normalized the number of either total HIV DNA or intact HIV DNA copies to 10^6 T cells in the sample. To compensate for underestimation of the intact HIV DNA results due to DNA shearing during sample processing, results were corrected by applying a DNA shearing index (DSI) specific to each sample as described (Bruner et al., 2019; Levy et al., 2021). Briefly, because the two *RPP30* targets are 11.7kb apart, and thus similar in length to an intact HIV provirus, the DSI was determined by calculating the proportion of sheared DNA templates based on single positive droplet counts for the two *RPP30* targets (S1 and S2) relative to the total number of human genomes present. Total genomes were calculated as the average of S1 and S2 plus the dual- (D) positive *RPP30* droplets as in the equation below. The DSI is then used to correct the estimated intact HIV copies as noted:

$$DSI = \frac{S1 + S2}{2D + S1 + S2}$$

Such that the corrected triple positive copies (I^{3+}_{corr}) or the best estimate of the number of intact HIV DNA sequences in the sample are computed by rescaling the observed copies (I^{3+}_{obs}) by the shearing index as

$$I^{3+}_{corr} = \frac{I^{3+}_{obs}}{1 - DSI}$$

Statistical analysis

Analysis was performed using R version 4.0.0. Pearson's correlation coefficients were used to compare HIV DNA copies measured by the CS-IPDA and an independent qPCR assay.