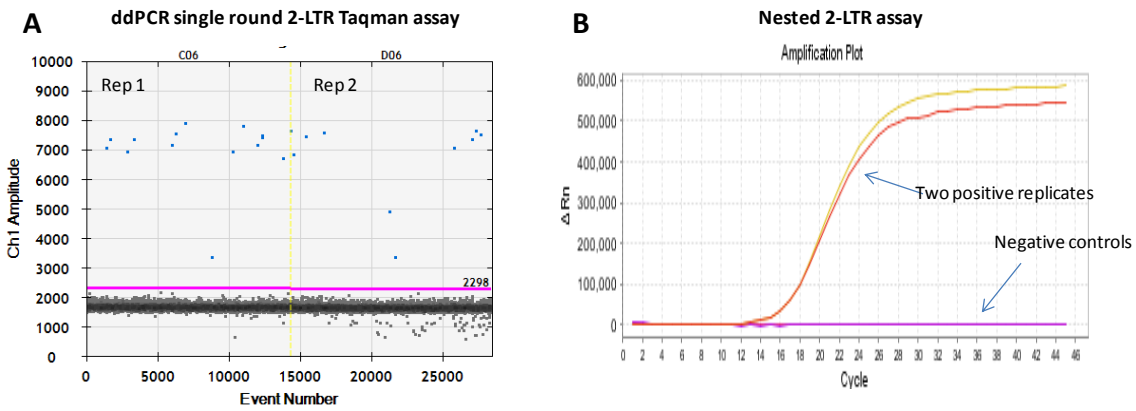
Supplementary Materials:

Part 1: Assay to detect HIV-1 subtype C 2-LTR episomal DNA

An assay to detect subtype C 2-LTR episomal HIV-1 DNA was developed. The 2-LTR junction (≈640bp amplicon spanning junction) was PCR amplified from a subtype C primary isolate (South Africa) cultured in donor PBMCs using a nested PCR approach: first round primers were designed to bind to within the *gag* and *nef*genes and second round primers within the 5' and 3' LTRs. This 640bpPCR fragment was gel purified and subsequently cloned into the TOPO TA cloning vector (Invitrogen, Thermo Fisher Scientific, California, USA). Plasmid DNA harbouring the junction sequence was subsequently sequenced using Sanger sequencing (3100 Genetic Analyzer; Applied Biosystems, Thermo Fisher Scientific) to confirm junction sequence. A TaqMan probe-based assay was designed over the junction sequence and tested on using a ten-fold dilution series of the purified 640bp 2-LTR amplicon (standard curve: r^2=1, Eff=0.87). In addition the assay was tested using a ddPCR platform on PBMC-derived gDNA from a subset of well described infected adult women. We subsequently designed a nested assay of the 2-LTR TaqMan assay by incorporating two outer primers in a first round PCR (standard PCR) and then using the complete 1st round PCR mixture to set up the second round qPCR (Taqman probe-based assay). Primers and probe sequences are as follows: **1st round standard PCR primers** – Forward Primer 5'-GAC TGC TGA CAC AGA AGG GAC T-3'; Reverse Primer 5'-CTT TGT AAT ACT CCG GAT GTA TCT C-3'; **2nd round TaqMan qPCR** – Forward Primer 5'-GCT TGC CTT GAG TGC TCT AAG TAG TG-3'; Reverse Primer 5'-AAT TAA CCC TTC CAG TAC TGC TAG AGA T-3'; Probe 5'-FAM-CTA GAG ATC CCT CAG AC-MGB-3'. The 1st round PCR was set up using the Roche Expand HiFi PCR system and the 2nd round qPCR was set up usingRoche Light Cycler 480 Probes Master Mix. Negative controls include both water and gDNA extracted from healthy HIV-1-uninfected donors. Positive controls include gDNA extracted from known HIV-1 infected individuals. Figure below shows an example of subtype C HIV-1-infected gDNA extracted from same individual run on a ddPCR platform (single round 2-LTR Taqman assay) and using the nested 2-LTR assay.



Calculated copies ≈260 2-LTR copies/1x10^6 PBMCs

Figure showing the same sample (subtype C HIV-1 infected adult extracted PBMC gDNA) run with (**A**) the 2-LTR single round TaqMan assay on a droplet digital PCR (quanitative) platform and (**B**) on a qPCR platform using the nested 2-LTR assay (qualitative).

Part 2: Validation of nested real-time PCR assay for integrase

We developed a nested real-time PCR assay for the highly-conserved integrase (IN) gene in pol designed using all subtype C sequences available on the Los Alamos HIV Database (https://www.hiv.lanl.gov/) [1]. As part of the validation of this assay, we tested samples from HIV-exposed infants recruited at Rahima Moosa Mother and Child Hospital in Johannesburg, South Africa [2]. Samples were selected from nine infants who had tested indeterminate on the nucleic acid amplification test (NAAT) (COBAS AmpliPrep/TaqMan HIV-1, v2.0, Roche Molecular Systems, Branchburg, NJ) as part of the routine early diagnosis program at the site. The indeterminate results were obtained on samples collected soon after birth as part of the routine program. The infants were followed with repeat blood sampling to determine their HIV status. Diagnostic NAAT described above and viral load tests (COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Test, Roche Molecular Systems, Branchburg, NJ) were repeated. Three of the nine infants with birth indeterminate results were later determined to be HIV-infected. The other six of the nine infants with birth indeterminate results were later determined to be uninfected. Six samples from the three infants determined to be infected tested positive on our in-house nested PCR for integrase in between 3 and 8 of 8 replicates for each sample. In contrast, none of the 8 replicates for each of the six samples from the six infants determined to be uninfected tested positive on the nested assay (Supplementary Table 1).

Supplementary Table 1. Testing of nested HIV-1 integrase assay on select indeterminate infant samples

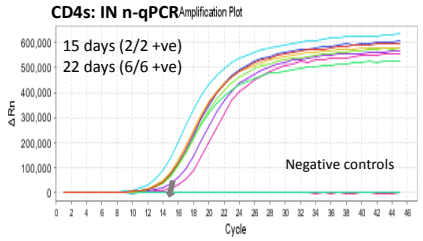
Samples Tested	Age of sample tested (days)	Sample Type	Replicates Positive/tested
Birth NAAT Indeterminate – Later tests indicate HIV infection			
Infant1	15	PBMCs	8/8
	30	PBMCs	8/8
Infant2	14	PBMCs	4/8
	32	PBMCs	3/8
Infant3	1	PBMCs	3/8
	33	PBMCs	4/8
Birth NAAT Indeterminate – Later tests indicate HIV uninfected			
Infant4	45	PBMCs	0/8
Infant5	28	PBMCs	0/8
Infant6	40	PBMCs	0/8
Infant7	271	PBMCs	0/8
Infant8	541	PBMCs	0/8
Infant9	220	PBMCs	0/8

Part 3: Select amplification plots showing the qualitative nested qPCR results

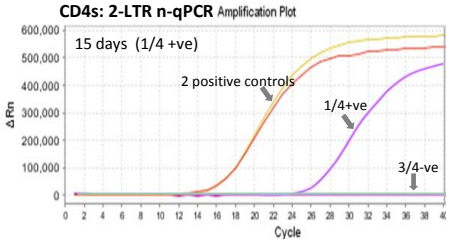
Figure S1. Select amplification plots showing the qualitative nested qPCR (n-qPCR) results (i and ii) for case 1 (A) and case 2 (B). Negative amplifications of either negative controls or samples are shown by flat lines and positive amplifications by distinct sigmoidal curves (where positive controls are shown they are indicated as such). Case 1 integrase-targeted n-qPCR results are shown for gDNA extracted from CD4+ T cells at days 15 and 22 (Ai), in addition n-qPCR targeting 2-LTR in gDNA extracted from CD4+ T cells at day 15 is shown (Aii). For case 2, integrase-targeted n-qPCR results are shown for gDNA extracted from PBMCs at days 282, 335 and 471 of age (Bi) and at days 117 and 233 days of age (Bii).

A: Case 1

i

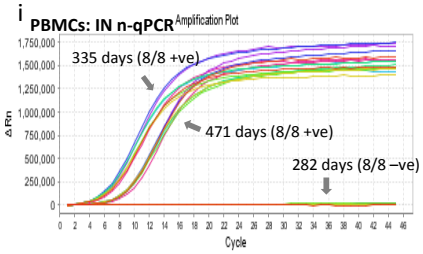


ii

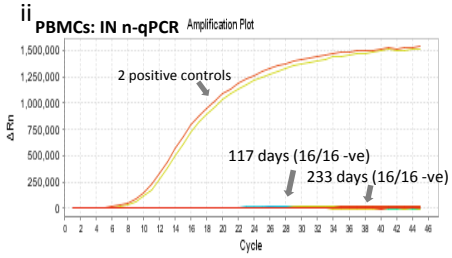


B: Case 2

i

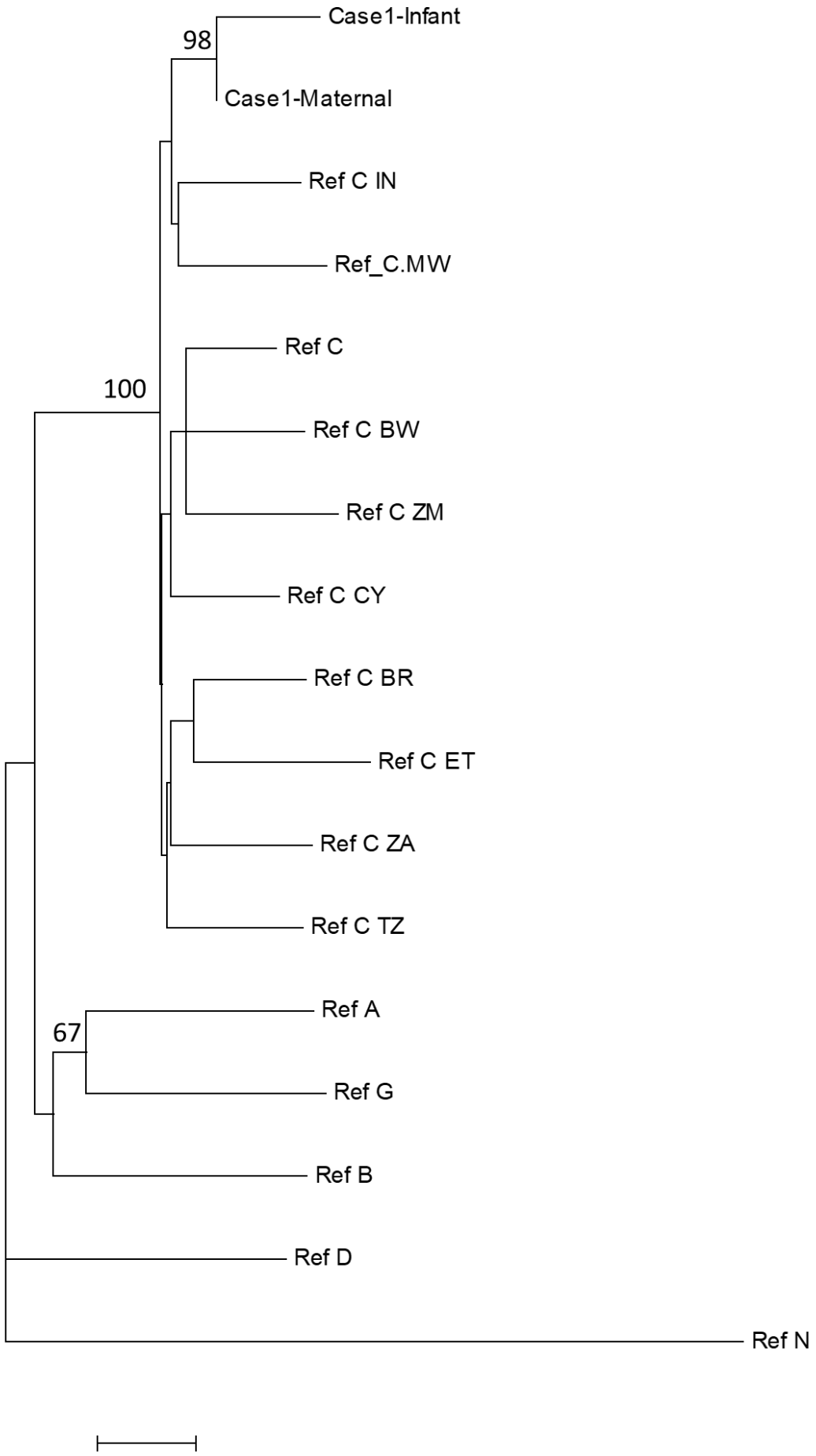


ii



Part 4: Phylogenetic tree

Figure S2: Neighbour-joining tree showing phylogenetically linked mother-child pair sequences from case 1 with a bootstrap support value of 98.0%. Phylogenetic analysis was performed by constructing neighbour-joining tree with 1000 bootstrap replicates using HKY85 model within PAUP v4.0b10. Phylogenetic tree reconstruction included HIV-1 subtype reference sequences from the Los Alamos database to illustrate subtype grouping. Bootstrap values >50% are shown.


Case1-Infant
98
Case1-Maternal
Ref C IN
Ref_C.MW
Ref C
100
Ref C BW
Ref C ZM
Ref C CY
Ref C BR
Ref C ET
Ref C ZA
Ref C TZ
Ref A
67
Ref G
Ref B
Ref D
Ref N
0.02

References:

1. Violari A, Cotton MF, Kuhn L, Schramm DB, Paximadis M, Loubser S, *et al.* A child with perinatal HIV infection and long-term sustained virological control following antiretroviral treatment cessation. *Nat Commun* 2019,**10**:412.
2. Technau KG, Mazanderani AH, Kuhn L, Hans L, Strehlau R, Abrams EJ, *et al.* Prevalence and outcomes of HIV-1 diagnostic challenges during universal birth testing - an urban South African observational cohort. *J Int AIDS Soc* 2017,**20**:21761.