

1 **HIV Diversity Considerations in the Application of the Intact Proviral DNA Assay (IPDA)**

2  
3 Natalie N. Kinloch<sup>1,2\*</sup>, Yanqin Ren<sup>3\*</sup>, Winiffer D. Conce Alberto<sup>3</sup>, Winnie Dong<sup>2</sup>, Pragya  
4 Khadka<sup>3</sup>, Szu Han Huang<sup>3</sup>, Talia M. Mota<sup>3</sup>, Andrew Wilson<sup>4</sup>, Aniqa Shahid<sup>1,2</sup>, Don Kirkby<sup>2</sup>,  
5 Marianne Harris<sup>2,5</sup> Colin Kovacs<sup>6</sup>, Erika Benko<sup>6</sup>, Mario A. Ostrowski<sup>7</sup>, Perla M. Del Rio  
6 Estrada<sup>8</sup>, Avery Wimpelberg<sup>9</sup>, Christopher Cannon<sup>9</sup>, W. David Hardy<sup>9</sup>, Lynsay MacLaren<sup>9</sup>,  
7 Harris Goldstein<sup>10</sup>, Chanson J. Brumme<sup>2,5</sup>, Guinevere Q. Lee<sup>3</sup>, Rebecca M. Lynch<sup>4</sup>, Zabrina L.  
8 Brumme<sup>1,2#</sup>, R. Brad Jones<sup>3,4#</sup>

9  
10 <sup>1</sup> Faculty of Health Sciences, Simon Fraser University, Burnaby, Canada

11 <sup>2</sup> British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada

12 <sup>3</sup> Infectious Diseases Division, Department of Medicine, Weill Cornell Medical College, New  
13 York, USA

14 <sup>4</sup> Department of Microbiology, Immunology and Tropical Medicine, George Washington  
15 University, Washington DC, USA

16 <sup>5</sup> Faculty of Medicine, University of British Columbia, Vancouver, Canada

17 <sup>6</sup> Maple Leaf Medical Clinic, Toronto, Canada

18 <sup>7</sup> Department of Medicine, University of Toronto, Toronto, Canada.

19 <sup>8</sup> Center for Research in Infectious Diseases, National Institute of Respiratory Diseases, Mexico  
20 City, Mexico

21 <sup>9</sup> Whitman Walker Health, Washington DC, USA.

22 <sup>10</sup> Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx,  
23 NY, USA

24  
25 \* These authors contributed equally

26 # These are co-senior/co-corresponding authors.

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

1    **Opening Paragraph (serves as abstract for submission) and Body**  
2

3            The Intact Proviral DNA Assay (IPDA) was developed to address the critical need for a  
4            precise and scalable method for intact HIV reservoir quantification<sup>1</sup>. This duplexed droplet  
5            digital PCR (ddPCR) assay simultaneously targets the HIV Packaging Signal ( $\Psi$ ) and the Rev  
6            Responsive Element (RRE) within Envelope (*env*) to distinguish genomically intact proviruses  
7            against a large background of defective ones<sup>2</sup>. The IPDA requires less time, resources, and  
8            biological material than the gold standard for replication-competent HIV reservoir measurement,  
9            the Quantitative Viral Outgrowth Assay (QVOA)<sup>3</sup>, and is being adopted in research and clinical  
10          studies<sup>4-7</sup>. In our cohort of HIV-1 subtype B-infected individuals from North America however,  
11          the IPDA yielded a 28% failure rate due to HIV polymorphism. We further demonstrate that  
12          within-host HIV diversity can lead the IPDA to underestimate intact HIV reservoir size, which  
13          could negatively impact clinical trial results interpretation. While the IPDA represents an  
14          important methodological advance, HIV diversity should be addressed before its widespread  
15          adoption.

16            We applied the IPDA to 46 HIV subtype B-infected, virally-suppressed individuals from  
17          North America, yielding a median of 29 (interquartile range [IQR] 0-93) intact  
18          proviruses/million CD4+ T-cells (Extended Data Figure 1). Of note, the IPDA did not detect any  
19          intact (*i.e.*  $\Psi$  and *env* double-positive) proviruses in 17 participants (37%). In four of these  
20          individuals, both  $\Psi$ - and *env*- single-positive proviruses were detected, suggesting a true negative  
21          result (Extended Data Figure 3) In the remaining 13 individuals however, the IPDA did not  
22          detect  $\Psi$ - and/or *env*- single-positive proviruses above background levels despite recovery of  
23          replication competent HIV in 11/11 cases where QVOA was performed, suggesting that the  
24          assay failed to detect autologous proviruses. Specifically, in eight of these participants only  $\Psi$ -

1 positive proviruses were detected, in four only *env*-positive proviruses were detected, and in one  
2 participant no proviruses were detected (Figure 1B; Extended Data Figure 3). This non-detection  
3 rate is consistent with Gaebler *et al* who reported a lack of detectable  $\Psi$  and *env* double-positive  
4 proviruses in 2/6 individuals (33%) using a quadruplexed qPCR assay that employs the IPDA  
5 probes<sup>8</sup>. Near-full-length single-genome proviral sequencing performed on IPDA *env*-negative  
6 participant BC-004 revealed mismatches to the *env* probe, where *in silico*-predicted reservoir  
7 distributions that took these polymorphisms into account (Figure 1C, left) differed markedly  
8 from the experimentally-obtained result (center). Substituting an autologous *env* probe rescued  
9 detection to *in silico*-predicted levels (Figure 1C, right), confirming that HIV polymorphism can  
10 cause the IPDA to fail. HIV sequencing revealed mismatches in the probe and/or at the 3' end of  
11 a primer in all cases of presumed assay detection failure (Extended Data Figure 4), yielding an  
12 overall estimated failure rate of 13/46 (28%). Indeed, exclusion of these 13 datapoints markedly  
13 improved the correlation between IPDA and QVOA results among the subset of participants for  
14 whom sufficient biological material was available to perform the latter, from  $p= 0.03$ ,  $p=0.83$  to  
15  $p= 0.35$ ,  $p=0.08$  (Figure 1C), bringing this correlation more in line with that reported by the  
16 original authors<sup>1</sup>. The correlation between total HIV *gag* DNA and QVOA by contrast was  
17 significant without exclusion of these datapoints (Extended Data Figure 2). Our findings thus  
18 reinforce the value of the IPDA but underscore the need for strategies to overcome HIV diversity.

19         Although the original study reported no detection failures due to HIV polymorphism, the  
20 authors acknowledged that triage primer/probe sets would be required in such cases<sup>1</sup>. A follow-  
21 up study similarly acknowledged HIV diversity as a potential limitation, but reported no  
22 participants (of  $n = 81$ ) for whom detection consistently failed for either  $\Psi$  or *env* amplicons<sup>5</sup>. By  
23 contrast, our results indicate that, as the IPDA is applied to diverse cohorts<sup>4,5,7</sup>, detection failures

1 will occur, and not infrequently. The first step towards mitigation is awareness: samples that  
2 yield no  $\Psi$  or *env* single-positive proviruses should be flagged as ‘unreportable’ until HIV  
3 polymorphism has been addressed (e.g. using autologous primers/probes<sup>1,5</sup>). Analysis of HIV  
4 subtype B sequences from unique individuals in the Los Alamos HIV database revealed that 23%  
5 of 9,360 *env* sequences harbored at least one *env* probe mismatch (which is similar to our *env*  
6 detection failure rate of 9/46 or 20%), while 50% of 1,489 sequences harbored at least one  $\Psi$   
7 probe mismatch (which is substantially greater than our  $\Psi$  detection failure rate of 5/46 or 11%).  
8 This suggests that the *env* reaction may be particularly sensitive to polymorphism.

9 Though laborious to correct, complete detection failures due to HIV polymorphism are  
10 nevertheless easy to flag. In contrast, within-host HIV diversity (see Figure 1C for an example)  
11 could lead the IPDA to underestimate intact reservoir size if the within-host variants were  
12 differentially detectable by the assay. Such partial detection failures would *not* be easy to  
13 identify. Moreover, if IPDA-detectable and non-detectable reservoir subpopulations were  
14 differentially susceptible to HIV cure interventions, this could lead to erroneous conclusions  
15 regarding intervention efficacy.

16 We illustrate this using HIV-specific broadly neutralizing antibodies (bNAbs). bNAbs  
17 can facilitate elimination of HIV-infected cells<sup>9</sup> in part by targeting them for antibody-dependent  
18 cellular cytotoxicity (ADCC)<sup>9,10</sup>, and are being evaluated in clinical trials, some of which use the  
19 IPDA as a readout (e.g: ACTG A5386<sup>6</sup>). Participant 91C33 from a published trial<sup>11</sup> provides a  
20 hypothetical example. This individual did not respond to (off-ART) infusions of the bNAbs  
21 3BNC117 and 10-1074 because they harbored a plasma HIV subpopulation that was resistant to  
22 both bNAbs<sup>11</sup> (Figure 2A). Of note, this plasma HIV subpopulation also harbored a mismatch to  
23 the IPDA *env* probe. We confirmed that the published IPDA could not detect templates

1 harboring this mismatch, while those representing the bNAb-sensitive strains were detected  
2 readily (Figure 2A; Extended Data Figure 5). If a person harboring such diversity in their  
3 reservoir were to be successfully treated with one or both bNAbs, the IPDA could over-estimate  
4 the intervention's effect (Figure 2C).

5 ARV-suppressed participant OM5346 provides another example. Pre-ART HIV drug  
6 resistance genotyping identified subtype B infection, however single-genome sequencing of pre-  
7 ART plasma and proviruses sampled during long-term ART revealed co-infection with a non-B  
8 strain (Extended Data Figure 6) that harbored a G-to-A mismatch at position 13 (G13A) of the  
9 *env* probe but no other critical mismatches to the IPDA primers. While the IPDA is only  
10 designed for subtype B, G13A is the most frequent *env* probe polymorphism [ $\sim 5\%$ ] in subtype B  
11 (and was observed in 2/10 study participants for whom *env* was sequenced). Moreover, the  
12 original report indicated that the IPDA could detect G13A, at least when present on a plasmid  
13 template<sup>1</sup>. Using QVOA, we isolated replication-competent subtype B ("virus 3") and co-  
14 infecting ("virus 4") viruses from OM5346's reservoir, confirming both of these as true intended  
15 IPDA targets (Figure 2B). We further observed that these viruses were differentially susceptible  
16 to 3BNC117- and 10-1074-mediated ADCC: while virus 3-infected cells could be eliminated by  
17 10-1074-, but not 3BNC117-, mediated ADCC (25% and 0% relative reduction, respectively),  
18 the opposite was true for virus 4 (0% and 32% relative reduction, respectively) (Figure 2B,  
19 Extended Data Figure 7). At concentrations approximating participant samples, the IPDA could  
20 readily detect provirus 3, but could not distinguish provirus 4 (Figure 2B, Extended Data Figure  
21 8A). Similarly, at participant-like concentrations, the IPDA could not detect a synthetic DNA  
22 template encoding the virus 4 sequence, though noticeable signal elevation occurred at  $\sim 100$ -fold  
23 higher template concentrations (Figure 2B, Extended Data Figure 8B). Importantly, a secondary

1 *env* primer/probe set (see below and methods) readily detected both viruses at all concentrations,  
2 regardless of template type. If this individual were to be successfully treated with 10-1074 the  
3 IPDA would overestimate the intervention's effect on the reservoir, whereas if the individual  
4 were to be successfully treated with 3BNC117, the IPDA would erroneously conclude that the  
5 intervention had no effect (Figure 2C). Our findings, taken together with Gaebler *et al*'s  
6 observation that 4 of 9 (44%) studied reservoirs were heterogeneous in an IPDA probe region<sup>8</sup>,  
7 suggests that the impact of within-host HIV diversity on IPDA accuracy may be non-negligible.

8 During our investigation we identified another source of potential error, resulting from  
9 variable spillover of fluorescence from the  $\Psi$  channel into the *env* channel, which has  
10 implications for the appropriate placement of thresholds defining negative and positive  
11 populations. Importantly, the extent of spillover is HIV sequence-specific. We demonstrate this  
12 by applying the IPDA to synthetic templates encoding the  $\Psi$  regions of OM5346 viruses 3 and 4,  
13 without a corresponding *env* template present (Extended Data Figure 9). Virus 3, which harbors  
14 mismatches to the  $\Psi$  probe, produced modest yet clearly discernible  $\Psi$  signal that did not spill  
15 over into the *env* channel. By contrast, virus 4 yielded high-amplitude  $\Psi$  signal that spilled over  
16 noticeably. Thus, drawing a tight threshold based on a template-negative (or virus 3-containing)  
17 sample and applying this threshold to virus 4 creates *env* signal when no *env* template was  
18 present. This  $\Psi$  channel spillover for virus 4 thus likely contributed to the small amount of 'env  
19 signal' in Fig. 2B (events adjoining the negative population). This observation underscores the  
20 instrument manufacturer's instruction that appropriate thresholds can only be drawn if negative  
21 and positive populations clearly separate from one another.

22 The IPDA offers major scalability advantages over existing molecular<sup>8</sup> or culture-based  
23 approaches<sup>3</sup>, and Gaebler *et al*. confirmed that, of any two probes, those of the IPDA offered the

1 greatest selectivity for intact proviruses<sup>8</sup>. Any molecular assay targeting a genetically variable  
2 pathogen however must address polymorphism. While the requirement to discriminate proviral  
3 defects in the  $\Psi$  region limits the placement and sequence of this probe<sup>2</sup>, we developed a  
4 secondary primer/probe set in the intact-discriminating RRE region<sup>1</sup>, approximately 50 bases  
5 downstream of the original location, as a first step towards addressing diversity. This  
6 primer/probe set rescued detection of *env*-positive proviruses in 9/9 participants with IPDA *env*  
7 detection failure (Extended Data Figure 10A). When we applied it to 36 participants for whom  
8 the IPDA detected *env*-positive proviruses (excluding OM5346 who harbored within-host  
9 diversity), it failed to detect *env*-positive proviruses in 3 (8%) individuals, indicating that it is not  
10 a universal solution. For the remaining 33 (92%) however it yielded measurements that were not  
11 significantly different from the IPDA (Extended Data Figure 10B). The secondary primer/probe  
12 set could therefore be used to identify instances of detection failure, though users must note that  
13 it cannot discriminate hypermutated sequences.

14 We identify HIV diversity as a source of error in the IPDA. Within-host HIV diversity  
15 represents a challenge given the HIV reservoir's dynamic nature<sup>12</sup> and particularly when this  
16 diversity is linked to intervention susceptibility. Given the clear value of the IPDA, iterative  
17 efforts to refine this assay are a priority. In the meantime, proviral sequencing should be  
18 performed for clinical trial participants, and 'backup' and/or autologous primers/probes used to  
19 mitigate HIV diversity, an all-too-familiar challenge in viral quantification.

20

1 **Supplementary Information (Methods)**

2

3 **Participants and Ethics Statement**

4 We studied 46 individuals with HIV-1 subtype B infection who were recruited to cohorts  
5 in Toronto (N=19), Vancouver (N=15), New York City (N=4), Washington, D.C. (N=7), and  
6 Mexico City (N=1). For Vancouver participants, peripheral blood mononuclear cells (PBMCs)  
7 were isolated by density gradient separation and cryopreserved (-150°C, 90% Fetal Bovine  
8 Serum+ 10% DMSO). Participants recruited at other sites provided a leukapheresis sample from  
9 which PBMC were isolated and cryopreserved as above. All participants were on long-term,  
10 virally suppressive combination antiretroviral therapy (cART) at time of sampling, with the  
11 exception of one elite controller. Participants' duration of untreated HIV infection ranged from 1  
12 month to >10 years, though the exact duration was unknown for most participants. Ethical  
13 approval to conduct this study was obtained from the Institutional Review Boards of Simon  
14 Fraser University, Providence Health Care/University of British Columbia, Weill Cornell  
15 Medicine, and the George Washington University. All participants provided written informed  
16 consent.

17 **Quantitative Viral Outgrowth Assay (QVOA)**

18 For participants for whom sufficient biological material was available, the Quantitative  
19 Viral Outgrowth Assay (QVOA) was performed as previously described<sup>3</sup>. Briefly, CD4+ T-cells  
20 were isolated from PBMCs by negative selection and plated in serial dilution at either 4 or 6  
21 concentrations (12 wells/concentration, 24-well plates). CD4+ T-cells were stimulated with  
22 phytohemagglutinin (PHA, 2µg/mL) and irradiated allogeneic HIV-negative PBMCs were added  
23 to further induce viral reactivation. MOLT-4/CCR5 cells were added at 24hrs post-stimulation as  
24 targets for viral infection. Culture media (RPMI 1640 + 10% FBS + 1% Pen/Strep +50U/mL IL-

1 2 + 10ng/mL IL-15) was changed every 3 days and p24 enzyme-linked immunosorbent assay  
2 (ELISA) was run on day 14 to identify virus-positive wells. Infectious Units per Million CD4+  
3 T-cells (IUPM) was determined using the Extreme Limiting Dilution Analysis (ELDA) software  
4 (<http://bioinf.wehi.edu.au/software/elda/>)<sup>13</sup>. Culture supernatants from virus-positive wells were  
5 frozen (-80°C) for future use.

6 **Intact Proviral DNA Assay (IPDA)**

7 Genomic DNA was isolated from a median 4.5 (interquartile range [IQR] 4-5) million  
8 CD4+ T-cells using the QIAamp DNA Mini Kit (Qiagen) with precautions to minimize DNA  
9 shearing. Intact HIV copies/million CD4+ T-cells were determined by droplet digital PCR  
10 (ddPCR) using the Intact Proviral DNA Assay (IPDA)<sup>1</sup>, where HIV and human RPP30 reactions  
11 were conducted independently in parallel and copies were normalized to the quantity of input  
12 DNA. In each ddPCR reaction, a median 7.5ng (IQR 7- 7.5ng) (RPP30) or a median 750ng (IQR  
13 700- 750ng) (HIV) of genomic DNA was combined with ddPCR Supermix for Probes (no  
14 dUTPs, BioRad), primers (final concentration 900nM, Integrated DNA Technologies), probe(s)  
15 (final concentration 250nM, ThermoFisher Scientific) and nuclease free water. Primer and probe  
16 sequences (5' -> 3') were: RPP30 Forward Primer- GATTGGACCTGCGAGCG, RPP30  
17 Probe- VIC-CTGACCTGAAGGCTCT- MGBNFQ, RPP30 Reverse Primer-  
18 GCGGCTGTCTCCACAAGT; RPP30 Shear Forward Primer- CCAATTGCTGCTCCTGGG,  
19 RPP30 Shear Probe- FAM- AAGGAGCAAGGTTCTATTGTAG- MGBNFQ, RPP30 Shear  
20 Reverse Primer- CATGCAAAGGAGGAAGCCG; HIV  $\Psi$  Forward Primer-  
21 CAGGACTCGGTTGCTGAAG, HIV  $\Psi$  Probe- FAM- TTTGGCGTACTCACCAGT-  
22 MGBNFQ, HIV  $\Psi$  Reverse Primer- GCACCCATCTCTCCTCTAGC; HIV *env* Forward  
23 Primer- AGTGGTGCAGAGAGAAAAAGAGC, HIV *env* Probe- VIC-

1 CCTTGGGTTCTGGGA- MGBNFQ, anti-Hypermutant *env* Probe-  
2 CCTTAGGTTCTTAGGAGC- MGBNFQ, HIV *env* Reverse Primer-  
3 GTCTGGCCTGTACCGTCAGC. For participant BC-004, an autologous *env* probe was  
4 designed by modifying the published IPDA *env* probe to match the participant's sequence (VIC-  
5 CCTTGGGTTCTGGGA- MGBNFQ). Droplets were prepared using either the Automated or  
6 QX200 Droplet Generator (BioRad) and cycled at 95°C for 10 minutes; 45 cycles of (94°C for  
7 30 seconds, 59°C for 1 minute) and 98°C for 10 minutes, as previously described<sup>1</sup>. Droplets were  
8 analyzed on a QX200 Droplet Reader (BioRad) using QuantaSoft software (BioRad, version  
9 1.7.4), where replicate wells were merged prior to analysis. Four technical replicates were  
10 performed for each participant sample, where a median (IQR) 243,100 (106,200- 265,950) cells  
11 were assayed in total. Intact HIV copies ( $\Psi$  and *env* double-positive droplets) were corrected for  
12 DNA shearing based on the frequency of RPP30 and RPP30-Shear double positive droplets. The  
13 median (IQR) DNA shearing index (DSI), measuring the proportion of sheared DNA in a sample,  
14 was 0.31 (0.28- 0.35), highly comparable to that reported by the original authors<sup>1</sup>. For the  
15 experiments that evaluated the IPDA's ability to detect the specific *env* sequences harbored by  
16 participants 91C33 (Figure 2A) and OM5346 (Extended Data Figure 8B), synthetic templates  
17 (purified PCR amplicons for 91C33 and commercially-synthesized gBlocks [IDT] for OM5346)  
18 were used as targets. Total HIV *gag* copies/million CD4+ T-cells (Extended Data Figure 2) were  
19 measured as previously described<sup>14</sup>.

20 **Single-template, near-full length HIV genome proviral amplification and sequencing**

21 Single-template, near-full-length proviral amplification was performed on DNA extracted  
22 from purified CD4+ T-cells by nested PCR using Platinum Taq DNA Polymerase High Fidelity  
23 (Invitrogen) as described<sup>15</sup> such that ~25% of the resulting PCR reactions yielded an amplicon.

1 First round primers were: Forward - AAATCTCTAGCAGTGGCGCCGAACAG, Reverse -  
2 TGAGGGATCTCTAGTTACCAGAGTC. Second round primers were: Forward -  
3 GCGCCCGAACAGGGACYTGAAARCGAAAG, Reverse-  
4 GCACTCAAGGCAAGCTTATTGAGGCTTA. Reactions were cycled as follows: 92°C for 2  
5 minutes; 10 cycles of (92°C for 10 seconds, 60°C for 30 seconds and 68°C for 10 minutes); 20  
6 cycles of (92°C for 10 seconds, 55°C for 30 seconds and 68°C for 10 minutes); 68°C for 10  
7 minutes<sup>15,16</sup>. Amplicons were sequenced using Illumina MiSeq technology and *de novo*  
8 assembled using an in-house modification of the Iterative Virus Assembler (IVA)<sup>17</sup>, using the  
9 custom software MiCall (<https://github.com/cfe-lab/MiCall>) or through collaboration with the  
10 Massachusetts General Hospital CCIB Core.

11 ***In silico* predicted IPDA results**

12 Near full-length proviral sequences from participant BC-004 were used to predict IPDA  
13 results *in silico* under the assumption that proviruses with intact probe regions would be detected  
14 by the assay. The IPDA  $\Psi$  and *env* probe regions were excised from each proviral sequence and  
15 were considered defective if the probe region was absent or if it contained mutations associated  
16 with defective proviruses (e.g. those consistent with hypermutation<sup>1</sup> or common splice donor site  
17 mutations<sup>2</sup>). All other sequences, regardless of whether they matched the published IPDA probe  
18 sequence, were considered intact in the probe region for the *in silico* reservoir composition  
19 assessment. A total of 381 proviruses were sequenced for BC-004, 338 of which contained at  
20 least one of the IPDA  $\Psi$  or *env* probe regions.

21 **Targeted sequencing of QVOA outgrowth viruses and participant-derived HIV sequences**

22  
23 For participant OM5346, HIV RNA was extracted from QVOA outgrowth viruses and  
24 pre-ART plasma, after which gp160 (QVOA) and gp120/Pol (plasma) were amplified from

1 endpoint-diluted templates by nested RT-PCR using HIV-specific primers and high fidelity  
2 enzymes. For participants with suspected IPDA detection failure, and for whom HIV sequences  
3 were not already available, the IPDA  $\Psi$  and/or *env* amplicon regions were bulk-amplified from  
4 extracted proviral DNA using HIV-specific primers. Amplicons were sequenced using either  
5 Sanger (3730xl, Applied BioSystems) or Next Generation (Illumina MiSeq) technologies. Sanger  
6 chromatograms were analyzed using Sequencher (version 5.0.1, Gene Codes), while Illumina  
7 MiSeq reads were *de novo* assembled as described above.

8 **Antibody-dependent cellular cytotoxicity (ADCC) assays**

9 Total CD4+ T-cells were isolated, as described above, from HIV-negative donors and  
10 activated with anti-CD3/anti-CD28 antibodies. Cells were then infected with the participant virus  
11 of interest collected from QVOA supernatant, and monitored by flow cytometry by intracellular  
12 staining for HIV-Gag (KC57-RD1, Beckman Coulter) until cells were >5% Gag+. CD4+ T-cells  
13 were washed and incubated with the broadly neutralizing antibody of interest (either 3BNC117  
14 or 10-1074, 10 $\mu$ g/mL) for 2hrs. Natural Killer (NK) cells were negatively selected (EasySep,  
15 StemCell Technologies) from PBMCs of allogeneic, HIV-negative donors and activated using  
16 interleukin-15 (IL-15) provided by the National Cancer Institute Biological Research Branch.  
17 Activated NK cells were then co-cultured with infected, antibody-treated CD4+ T-cells for 16  
18 hours at an effector-to-target (E:T) ratio of 1:1. Following co-culture, cells were stained with  
19 fluorophore-conjugated antibodies against human IgG, CD3, CD56, and CD4 (Biolegend), as  
20 well as intracellular Gag (KC57-RD1, Beckman Coulter) and LIVE/DEAD Fixable Aqua Stain  
21 amine-reactive dye (Invitrogen). The percentages (%) of infected cells remaining post-ADCC  
22 relative to no antibody control were then determined using the following formula: (1- ((%Gag+  
23 cells amongst viable CD3+ cells in no antibody condition)- (%Gag+ cells among viable CD3+

1    cells in test condition))/(%Gag+ cells amongst viable CD3+ cells under no antibody  
2    condition))\*100.

3    **Phylogenetic Analysis of participant OM5346's HIV RNA and reservoir diversity**

4            For participant OM5346, single-genome HIV amplification and sequencing was  
5    performed from pre-ART plasma (sampled in 2012), proviruses from purified CD4+ T-cells  
6    isolated during suppressive cART (sampled in 2017 and 2019), and replication-competent HIV  
7    strains isolated by QVOA (from the 2017 sample) as described above. The 2012 and 2017/2019  
8    samples were shipped directly to, and amplified in, two separate laboratories. A plasma HIV *pol*  
9    sequence derived from clinical drug resistance testing in 2012 was also incorporated into the  
10   analysis. *pol* and *gp120* sequences were specifically amplified or retrieved from near-full  
11   genome proviral sequences using GeneCutter  
12   ([https://www.hiv.lanl.gov/content/sequence/GENE\\_CUTTER/cutter.html](https://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html)). Sequences were  
13   multiply-aligned using MAFFT (version 7.427)<sup>18</sup> in a codon-aware manner and inspected in  
14   AliView<sup>19</sup>. Maximum-likelihood phylogenies were inferred using PhyML under a General Time  
15   Reversible nucleotide substitution model<sup>20</sup> and visualized using FigTree (version 1.3.1).

16   **Design of Secondary *env* ddPCR primer/probe set**

17            HIV-1 Subtype B *envelope* sequences, limited to one per participant, were retrieved from  
18   the HIV LANL database (N=4,670) and aligned. Sequence conservation was assessed across 16-  
19   25 base-pair intervals, and three intervals that maximized sequence conservation while meeting  
20   ddPCR assay specifications (*i.e.* amplicon length, T<sub>m</sub>) were chosen as secondary primer and  
21   probe locations. These were: Secondary *env* Forward Primer ACTATGGGCGCAGCGTC  
22   (representing nucleotides 7,809-7,825 in the HIV-1 genomic reference HXB2; predicted  
23   sequence conservation 79%), Secondary *env* Probe VIC-CTGGCCTGTACCGTCAG-

1 MGBNFQ (HXB2 nucleotides 7,833-7,849, 83% conservation), Secondary *env* Reverse Primer  
2 CCCCCAGACTGTGAGTTGCA (HXB2 nucleotides 7,939- 7,921, 88% conservation). Reaction  
3 composition and cycling conditions were same as those used for the IPDA as described above.  
4 Accurate quantification using the Secondary *env* primer/probe set was verified using DNA  
5 extracted from the J-Lat 9.2 cell line (obtained from the NIH AIDS Reagent Program, Division  
6 of AIDS, NIAID, NIH, contributed by Dr. Eric Verdin)<sup>21</sup> (Extended Data Figure 11).

## 7 **Statistical Analyses**

8 All statistical analyses were performed using GraphPad Prism (version 8).

## 9 **Data and Sequence Availability**

10  
11 All sequences collected for the present study have been deposited in GenBank (Accession  
12 Numbers pending). Sequences for participant 91C33 were previously deposited under Accession  
13 Numbers: MH632930- MH632955<sup>11</sup>. All other data are available from the corresponding authors  
14 upon request, in compliance with institutional and REB requirements.

## 15 **Author Contributions Statement**

16 NNK, YR, ZLB and RBJ designed the study. NNK, YR, WDCA, WD, PK, SHH, TMM,  
17 AS and AW performed experiments. NNK, YR, WDCA, WD, PK, SHH, TMM, AS, AW, DK,  
18 CJB, GQL and RML analyzed data. MH, CK, EB, MAO, PMDRE, AW, CC, WDH, LM, HG  
19 and ZLB provided participant samples. NNK, YR, ZLB and RBJ wrote the manuscript. All  
20 authors contributed to the critical revision of the manuscript.

## 21 **Acknowledgments**

22 This work was supported by the Martin Delaney ‘BELIEVE’ Collaboratory (NIH grant  
23 1UM1AI26617), which is supported by the following NIH Co-Funding and Participating  
24 Institutes and Centers: NIAID, NCI, NICHD, NHLBI, NIDA, NIMH, NIA, FIC, and OAR. It

1 was also supported in part by the NIH funded R01 grants AI31798 and AI147845 (to BRJ) and a  
2 Canadian Institutes of Health Research (CIHR) project grant PJT-159625 (to ZLB). NNK is  
3 supported by a CIHR Vanier Doctoral Award. ZLB is supported by a Scholar Award from the  
4 Michael Smith Foundation for Health Research.

5 We thank the National Cancer Institute (NCI) Biological Research Branch for provision  
6 of IL-15 and IL-2, the NCI's AIDS and Cancer Virus Program for provision of HIV p24 ELISA  
7 reagents, and the NIH AIDS Research and Reference Reagent Program for provision of the  
8 CCR5<sup>+</sup> MOLT-4 cells. We thank Marina Caskey for providing de-identified samples from ART-  
9 naïve donors used in this study. We thank the BC Centre for Excellence in HIV/AIDS for  
10 support. We thank the Massachusetts General Hospital Center for Computational & Integrative  
11 Biology DNA Core, specifically Nicole Stange-Thomann, Amy Avery, Kristina Belanger, and  
12 Huajun Wang for providing us with the Illumina MiSeq deep sequencing service used in this  
13 manuscript. We thank Bruce Ganase, Jeff Knaggs, Daniel MacMillan and Hanwei Suderuddin  
14 for research and bioinformatics support and Gursev Anmole, Mark Brockman and Robert Holt  
15 for helpful discussions. We thank Michel Nussenzweig, Christian Gaebler and Yotam Bar-On for  
16 provision of materials and helpful discussions. We gratefully acknowledge the contributions of  
17 the study participants, without whom this work would not be possible.

18

## 1 References

2

3 1 Bruner, K. M. *et al.* A quantitative approach for measuring the reservoir of latent HIV-1  
4 proviruses. *Nature* **566**, 120-125, doi:10.1038/s41586-019-0898-8 (2019).

5 2 Ho, Y. C. *et al.* Replication-competent noninduced proviruses in the latent reservoir  
6 increase barrier to HIV-1 cure. *Cell* **155**, 540-551, doi:10.1016/j.cell.2013.09.020 (2013).

7 3 Laird, G. M. *et al.* Rapid quantification of the latent reservoir for HIV-1 using a viral  
8 outgrowth assay. *PLoS pathogens* **9**, e1003398, doi:10.1371/journal.ppat.1003398  
9 (2013).

10 4 Antar, A. A. R. *et al.* Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics  
11 during long-term antiretroviral therapy. *The Journal of clinical investigation*,  
12 doi:10.1172/jci135953 (2020).

13 5 Peluso, M. J. *et al.* Differential decay of intact and defective proviral DNA in HIV-1-  
14 infected individuals on suppressive antiretroviral therapy. *JCI insight* **5**,  
15 doi:10.1172/jci.insight.132997 (2020).

16 6 Wilkin, T. & Jones, R. B. (Personal Communication, 2020).

17 7 Cyktor, J., et, al. Intact proviral DNA levels decline in people with HIV on antitetroviral  
18 therapy (ART). *Journal of Virus Eradication* **5**, Supplement 3 (2019).

19 8 Gaebler, C. *et al.* Combination of quadruplex qPCR and next-generation sequencing for  
20 qualitative and quantitative analysis of the HIV-1 latent reservoir. *The Journal of  
21 experimental medicine* **216**, 2253-2264, doi:10.1084/jem.20190896 (2019).

22 9 Bruel, T. *et al.* Elimination of HIV-1-infected cells by broadly neutralizing antibodies.  
23 *Nature communications* **7**, 10844, doi:10.1038/ncomms10844 (2016).

24 10 Mendoza, P. *et al.* Combination therapy with anti-HIV-1 antibodies maintains viral  
25 suppression. *Nature* **561**, 479-484, doi:10.1038/s41586-018-0531-2 (2018).

26 11 Bar-On, Y. *et al.* Safety and antiviral activity of combination HIV-1 broadly neutralizing  
27 antibodies in viremic individuals. *Nature medicine* **24**, 1701-1707, doi:10.1038/s41591-  
28 018-0186-4 (2018).

29 12 Wang, Z. *et al.* Expanded cellular clones carrying replication-competent HIV-1 persist,  
30 wax, and wane. *Proceedings of the National Academy of Sciences of the United States of  
31 America* **115**, E2575-e2584, doi:10.1073/pnas.1720665115 (2018).

32 13 Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted  
33 and enriched populations in stem cell and other assays. *Journal of immunological  
34 methods* **347**, 70-78, doi:10.1016/j.jim.2009.06.008 (2009).

35 14 Huang, S. H. *et al.* Latent HIV reservoirs exhibit inherent resistance to elimination by  
36 CD8+ T cells. *The Journal of clinical investigation* **128**, 876-889, doi:10.1172/jci97555  
37 (2018).

38 15 Lee, G. Q. *et al.* Clonal expansion of genome-intact HIV-1 in functionally polarized Th1  
39 CD4+ T cells. *The Journal of clinical investigation* **127**, 2689-2696, doi:10.1172/jci93289  
40 (2017).

1 16 Li, B. *et al.* Rapid reversion of sequence polymorphisms dominates early human  
2 immunodeficiency virus type 1 evolution. *J Virol* **81**, 193-201, doi:10.1128/jvi.01231-06  
3 (2007).

4 17 Hunt, M. *et al.* IVA: accurate de novo assembly of RNA virus genomes. *Bioinformatics*  
5 (*Oxford, England*) **31**, 2374-2376, doi:10.1093/bioinformatics/btv120 (2015).

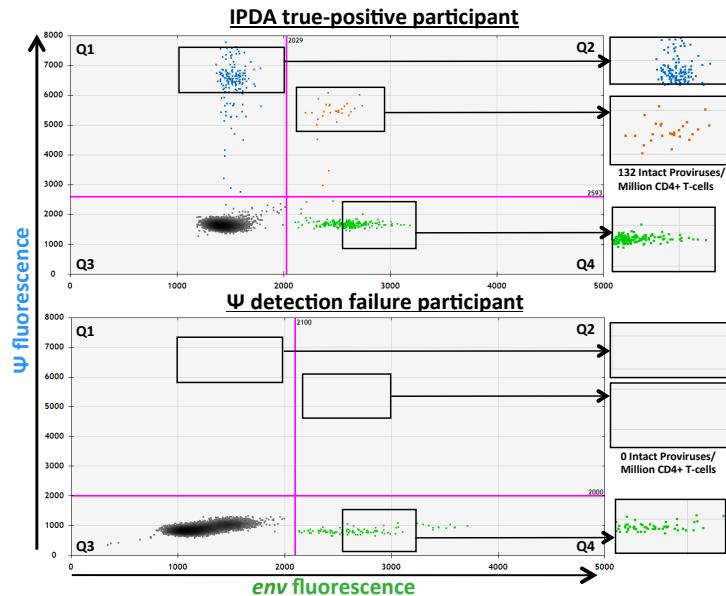
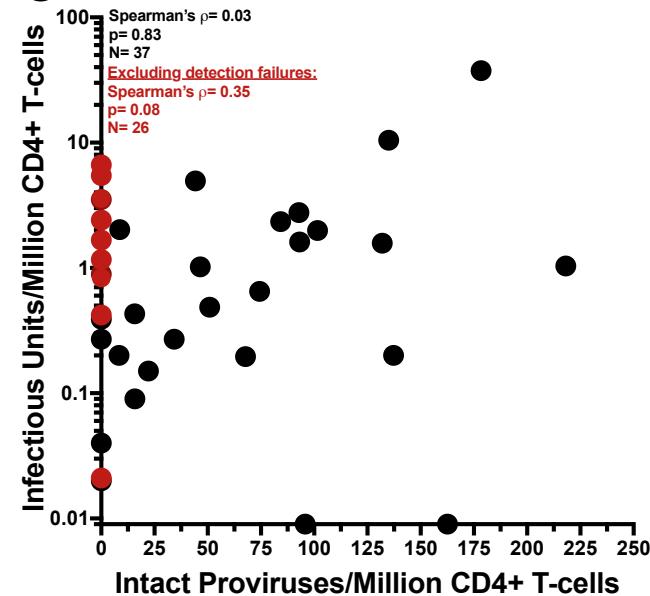
6 18 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:  
7 improvements in performance and usability. *Molecular biology and evolution* **30**, 772-  
8 780, doi:10.1093/molbev/mst010 (2013).

9 19 Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large datasets.  
10 *Bioinformatics (Oxford, England)* **30**, 3276-3278, doi:10.1093/bioinformatics/btu531  
11 (2014).

12 20 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood  
13 phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307-321,  
14 doi:10.1093/sysbio/syq010 (2010).

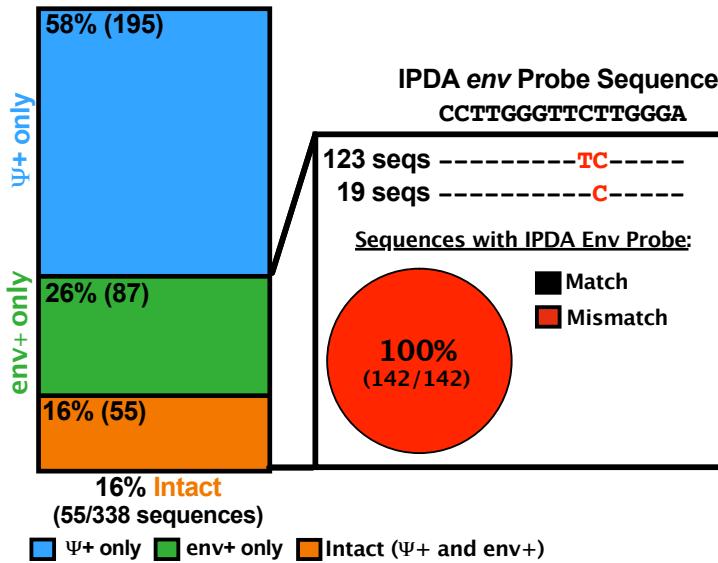
15 21 Jordan, A., Bisgrove, D. & Verdin, E. HIV reproducibly establishes a latent infection after  
16 acute infection of T cells in vitro. *The EMBO journal* **22**, 1868-1877,  
17 doi:10.1093/emboj/cdg188 (2003).

18

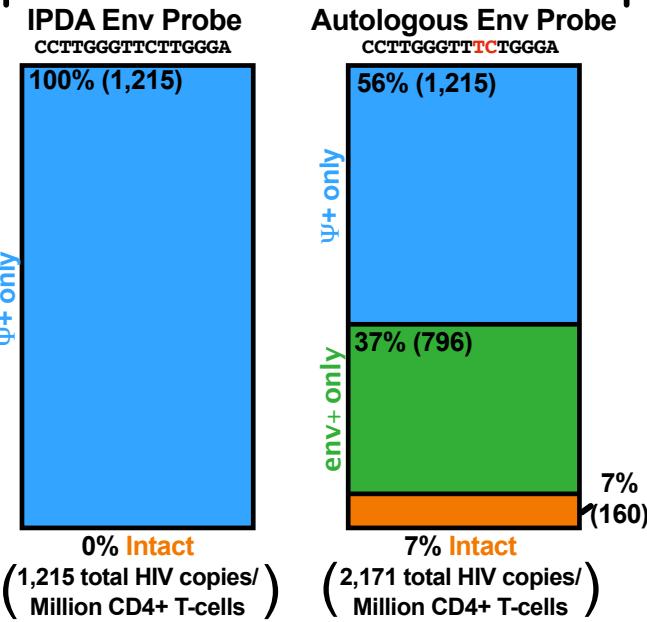
**Figure 1****A****C****B**

### IPDA-defined Reservoir Composition (BC-004)

#### *In silico* sequence-based result



#### Experimental Result



**Figure 1: Inter-individual HIV diversity can lead to detection failure by the IPDA**

**(A) Example IPDA 2D plots contrasting a true positive result from a case of detection failure.** 2D ddPCR plots showing  $\Psi$ -single positive events (Q1, blue),  $\Psi$ - and *env*- double positive events (Q2, orange), double-negative events (Q3, grey) and *env*-single positive events (Q4, green), for an IPDA true-positive individual (top) and a presumed case of  $\Psi$  detection failure (bottom).

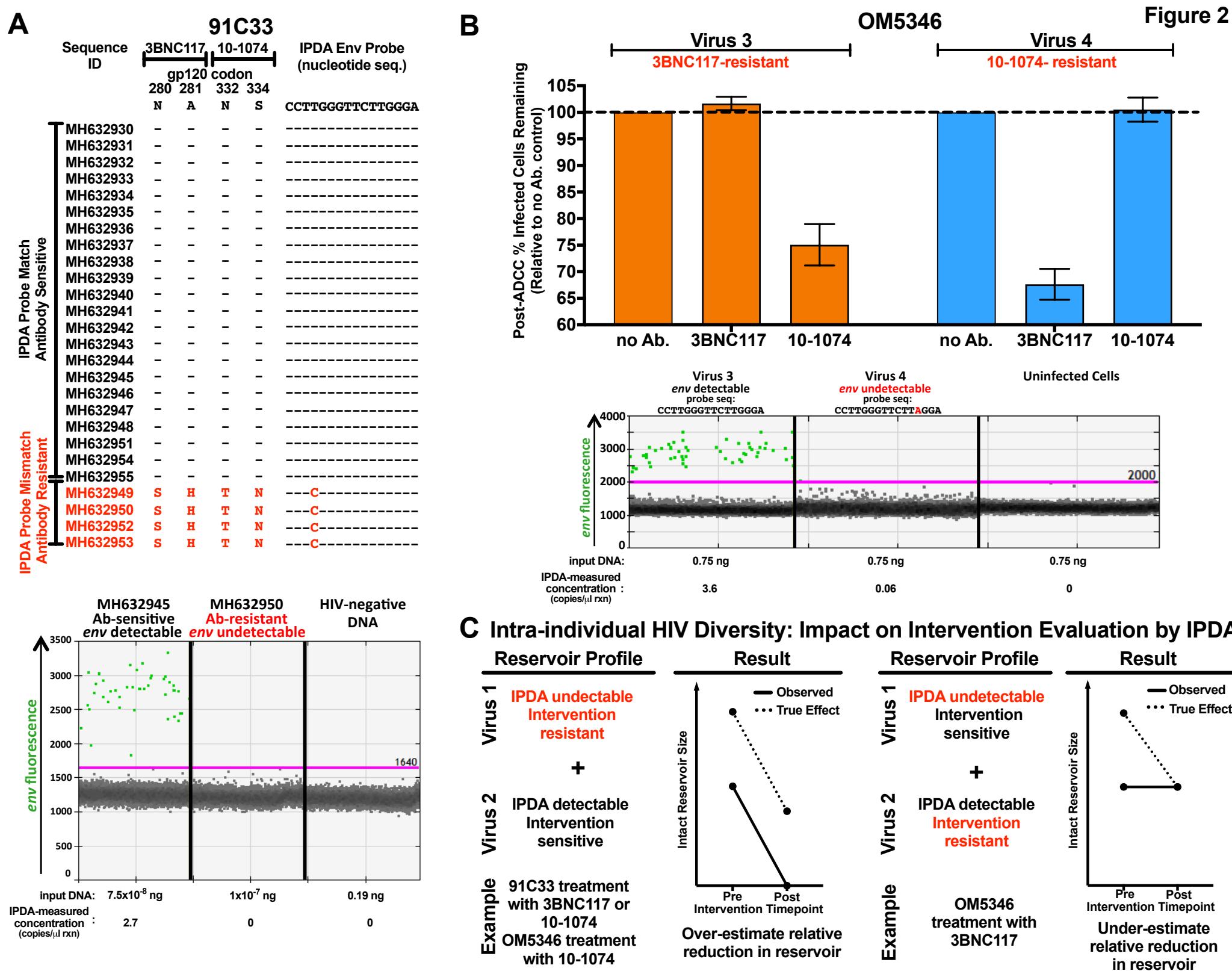
**(B) Inter-individual HIV diversity can lead to detection failure in the IPDA. (left)**

Analysis of 338 near-full length proviruses from participant BC-004, who originally failed detection in *env*, revealed single nucleotide mismatches (red, inset) to the IPDA *env* probe. These sequences yielded *in silico*  $\Psi$ + (blue), *env*+ (green) and intact ( $\Psi$ +*env*+; orange) provirus frequencies as shown. **(right)** Experimental results obtained with the published IPDA and autologous *env* probe, respectively.

**(C) IPDA detection failures obscure correlation between QVOA and IPDA in a**

**North American cohort.** Spearman's correlation between reservoir size as measured by IPDA (Intact Proviruses/Million CD4+ T-cells) and QVOA (Infectious Units per Million CD4+ T-cells) in 37 virally-suppressed participants for whom the latter measurement was available ( $\rho= 0.03$ ,  $p=0.83$ ) and excluding presumed instances of IPDA detection failure (red datapoints,  $n= 26$  remaining,  $\rho= 0.35$ ,  $p=0.08$ ). Two individuals for whom no replication competent viruses were detected (IUPM= 0) are plotted on the X-axis.

Figure 2



**Figure 2: Intra-individual HIV diversity can lead to reservoir underestimation by IPDA.**

**(A) Example of heterogeneous within-host HIV variants that are differentially detectable by IPDA and differentially sensitive to bNAbs. (top)** Participant 91C33 harbored two major plasma HIV subpopulations: one that matched the IPDA *env* probe and was sensitive to bNAbs 3BNC117 and 10-1074, and the other that harbored a mismatch at position 4 of the IPDA *env* probe (red) and was resistant to 3BNC117 and 10-1074. Sequences are identified by their GenBank accession numbers; hyphens (–) indicate matches to the nucleotide or amino acid reference; red letters denote mismatches (IPDA *env* probe) or amino acid substitutions (bNAb). **(bottom)** Representative ddPCR *env* plots for MH632945 and MH632950; positive droplets are green and negative droplets are grey. Templates were purified *env* PCR products of equal length and comparable quantities. HIV-negative donor DNA served as a negative control. See Extended data Figure 5 for additional experiments.

**(B) Example of heterogeneous within-host HIV variants that are differentially detectable by IPDA and differentially sensitive to bNAb-mediated ADCC. (top)** Two viruses ("virus 3" and "virus 4") isolated from participant OM5346's reservoir exhibited different sensitivities to bNAb-mediated ADCC: virus 3 was sensitive to 10-1074-mediated ADCC but resistant to 3BNC117, while virus 4 was sensitive to 3BNC117-mediated ADCC but resistant to 10-1074. Error bars indicate standard deviation from three technical replicates; see Extended Data Figure 7 for gating strategy and raw flow data. **(bottom)** Representative 1D IPDA *env* ddPCR plots for cells infected *in vitro* with virus 3, which matched the IPDA *env* probe sequence, and virus 4, which harbored the

G13A mismatch. Uninfected cells served as a negative control. Both viruses were readily detectable using an alternative primer/probe set (see Extended data Figure 8A).

**(C) Possible impacts of intra-individual HIV diversity on IPDA-measured changes in intact reservoir size following a hypothetical intervention.** **(left)** An intervention's impact on the reservoir could be *overestimated* when a subset of within-host HIV sequences are undetectable by IPDA and *resistant* to the treatment (*e.g.* OM5346 virus 4 during treatment with 10-1074, or 91C33 during treatment with 3BNC117 or 10-1074). The solid line indicates the observed effect by IPDA; dashed line indicates the true effect. **(right)** An intervention's impact on the reservoir could be *underestimated* when a subset of within-host HIV sequences are undetectable by IPDA and *sensitive* to treatment (*e.g.* OM5346 virus 4 during treatment with 3BNC117). In the worst-case scenario shown here, one could erroneously conclude that the treatment had no effect.