Received: 19 June 2024 | Accepted: 23 September 2024

DOI: 10.1002/jmv.29951

RESEARCH ARTICLE

JOURNAL OF MEDICAL VIROLOGY WILEY

Comparison of short-read and long-read next-generation sequencing technologies for determining HIV-1 drug resistance

Camille Vellas[1,2] | Amira Doudou[3] | Sofiane Mohamed[3] | Stéphanie Raymond[1,2] | Nicolas Jeanne[1] | Justine Latour[1] | Sofia Demmou[1] | Noémie Ranger[1] | Dimitri Gonzalez[3] | Pierre Delobel[2,4] | Jacques Izopet[1,2]

[1]CHU de Toulouse, Laboratoire de Virologie, Toulouse, France

[2]INSERM UMR1291-CNRS UMR5051-Université Toulouse III, Toulouse Institute for Infectious and Inflammatory Diseases, Toulouse, France

[3]ABL Diagnostics (ABLD), Woippy, France

[4]CHU de Toulouse, Service de Maladies Infectieuses et Tropicales, Toulouse, France

Correspondence
Camille Vellas, Institut Fédératif de Biologie, Laboratoire de Virologie 330 av de Grande Bretagne, 31052 Toulouse, France.
Email: vellas.c@chu-toulouse.fr

Funding information
ABLD provided DeepChek® whole genome HIV-1 assay and NGS library preparation kits

Abstract

Accurate HIV-1 genome sequencing is necessary to identify drug resistance mutations (DRMs) in people with HIV-1 (PWH). Next-generation-sequencing (NGS) allows the detection of minor variants and is now available in many laboratories. Our study aimed to compare two NGS approaches, a "short read" sequencing protocol using DeepChek® Whole Genome HIV-1 Assay on Illumina, and a "long read" sequencing protocol of HIV-1 *pol* and *env* single-molecule real-time sequencing (SMRT) on Pacific Biosciences (PacBio). We analyzed 16 plasma samples and 13 cellular samples from PWH. HIV-1 whole genome was amplified into five amplicons using DeepChek® Whole Genome HIV-1 Assay and sequenced on an iSeq. 100. In parallel, HIV-1 *pol* and *env* genes were separately amplified and sequenced using PacBio SMRT system with the circular consensus sequencing mode on a Sequel IIe. Concordance rates for determining DRMs with both approaches varied depending on the HIV-1 region, with higher concordance in the integrase region compared to the reverse transcriptase and protease regions. DeepChek® Whole Genome HIV-1 Assay exhibited better sensitivity in HIV-1 RNA sequencing of plasmas with lower viral loads. In cell HIV-1 DNA sequencing, the DeepChek® Whole Genome HIV-1 Assay performed better in *pol* and *env* sequencing but detected more APOBEC-induced DRMs, which can represent defective proviruses. Our findings indicate that both DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing exhibit good performance for subtype determination, detection, and quantification of DRMs of the HIV-1 genome. However, some discrepancies were found in cellular samples, highlighting the challenges of interpreting HIV-1 DNA DRMs.

KEYWORDS
drug-resistance mutation, HIV-1, SMRT sequencing, whole genome sequencing

1 | INTRODUCTION

Despite the success of combination antiretroviral therapy (ART) in treating HIV-1 infection, the risk of emergence of HIV-1 resistance to ART remains a challenge to effective treatment.[1] Therefore, HIV-1 drug-resistance tests are necessary to identify drug-resistance mutations (DRMs). These tests are routinely carried out to enhance the likelihood of treatment success by guiding the selection of ART for individuals newly diagnosed with HIV-1 or experiencing treatment failure. In recent years, novel sequencing approaches capable of identifying minor drug-resistant variants in clinical samples have been developed. A suitable HIV-1 drug-resistance test requires strong performance, including the ability to detect DRMs in samples with low plasma viral load or cell HIV-1 DNA. Analyzing cell DNA through sequencing can offer valuable information for individuals with low or undetectable plasma HIV-1 RNA levels, and when no previous genotype is available.[2,3] However, the interpretation of drug resistance on cell DNA requires consideration of the potential presence of defective, replication-incompetent proviruses.[4,5]

In the past few years, HIV-1 drug resistance tests have been performed by sequencing the protease (PR), the reverse transcriptase (RT), and the integrase (INT) regions of the HIV-1 genome. However, new HIV-1 drugs, including capsid inhibitors,[6] CD4-directed monoclonal antibodies (mAbs),[7] and broadly neutralizing antibodies have been developed.[8] Sequencing the whole HIV-1 genome appears to be necessary to assess the presence or absence of DRMs for all potential therapeutic options.

HIV-1 has elevated rates of reverse transcriptase-related mutation and a rapid viral turnover that conduct to various genetically diverse variants which coexist as a quasispecies.[9,10] During non-suppressive ART, ongoing evolution of the quasispecies seems to be driven by genetic drift, fitness-driven selection, and recombination.[11] Several studies have shown that drug-resistant variants present at low frequencies in people with HIV-1 (PWH) under ART may play a role in the evolution of resistance and treatment failure.[12–15] Minor variants initially present at low levels, could evolve into major DRMs when subjected to selective pressure. NGS represents a powerful tool to identify all drug-resistant variants within the HIV-1 genome, including minor ones. These technologies with enhanced sensitivity can detect subpopulations of drug-resistant variants in clinical samples at very low frequencies and improved our capacity to identify individuals at risk of virological failure compared to the 20% detection threshold utilized in Sanger sequencing.[16–18]

The DeepChek® Whole Genome HIV-1 Assay kit (ABL Diagnostics) has been developed for routine HIV-1 genotyping using both plasma RNA and cell DNA.[19] This assay enables sequencing of the whole HIV-1 genome in five fragments. However, it is performed on an Illumina platform, requiring DNA fragmentation before sequencing. Alternatively, other NGS approaches, such as Single Molecule Real Time (SMRT) sequencing, developed by Pacific Biosciences (PacBio), provide long-read capabilities, enabling the detection of all haplotypes of a quasispecies on the amplified fragments within a sample.

Evaluating the performance of short and long reads NGS technologies in identifying DRMs in people with HIV remains crucial. The aim of our study was to compare two NGS approaches for genotyping plasma HIV-1 RNA and cell HIV-1 DNA: the DeepChek® whole genome HIV-1 genotyping assay on Illumina, and HIV-1 *pol* and *env* PacBio SMRT sequencing.

2 | METHODS

2.1 | Patients and samples

We collected 23 blood samples from 23 PWH, who were being monitored at the Toulouse University Hospital for the determination of HIV genotypic resistance, between February 2022 and April 2024. Plasma HIV-1 RNA load ranged from 4.5 to 6.6 $\log_{10}$ copies/mL in 11 patients and was undetectable in the 12 remaining patients. These latter had cell HIV-1 DNA ranging from <1.4 (undetectable with a routine assay) to 3.3 $\log_{10}$ copies/million cells. To approximate the threshold where results in routine analyses are challenging (around 3 $\log_{10}$ copies/mL), we diluted three plasmas in PBS 7% BSA (Sigma Aldrich), to assess the sensitivity of the sequencing protocols, giving five samples with viral load ranging from 2.6 to 3.8 $\log_{10}$ copies/mL.

Fifty microliters of HIV-1 RNA was extracted from 1000 µL of plasma using the Viral NA Large Volume kit on the MagNA Pure 96 instrument (Roche Diagnostics). One hundred microliters of HIV-1 DNA was extracted from 200 µL of buffy coat (PBMC-enriched blood fraction) using the Viral NA Small Volume Kit on the MagNA Pure 96 instrument (Roche Diagnostics).

2.2 | Plasma HIV-1 RNA assay and cell HIV-1 DNA assay

Plasma HIV-1 RNA (pol and LTR regions) was quantified by real-time TMA on a Panther instrument using the Aptima® HIV-1 Quant Dx (Hologic).[20] Cell HIV-1 DNA (LTR region) was quantified by real-time PCR (LightCycler 480 system, Roche) from DNA extracted from PBMCs using the Generic HIV DNA Cell kit (Biocentric) as previously described.[21]

2.3 | DeepChek® whole genome HIV-1 assay

Extracted RNA or DNA was amplified using the DeepChek® whole genome HIV-1 assay kit (Ref 170B) (ABL Diagnostics). Briefly, five amplicons were amplified to cover the whole genome, following manufacturer's instructions (Figure 1). Each PCR product was controlled using the 5200 Fragment Analyzer (Agilent) and quantified using the Quantifluor dsDNA HS Assay kit (Promega). If the amplified fragments were insufficient, a nested PCR reaction was performed. PCR products were purified using clean DNA magnetic beads (ABL diagnostics) and pooled. The appropriate volume of the pool of amplicons was adjusted to 3 ng/µL. Subsequently, libraries were prepared using the DeepChek® NGS Library preparation kit (Ref 116B) (ABL Diagnostics) according to manufacturer's instructions. The libraries were qualified

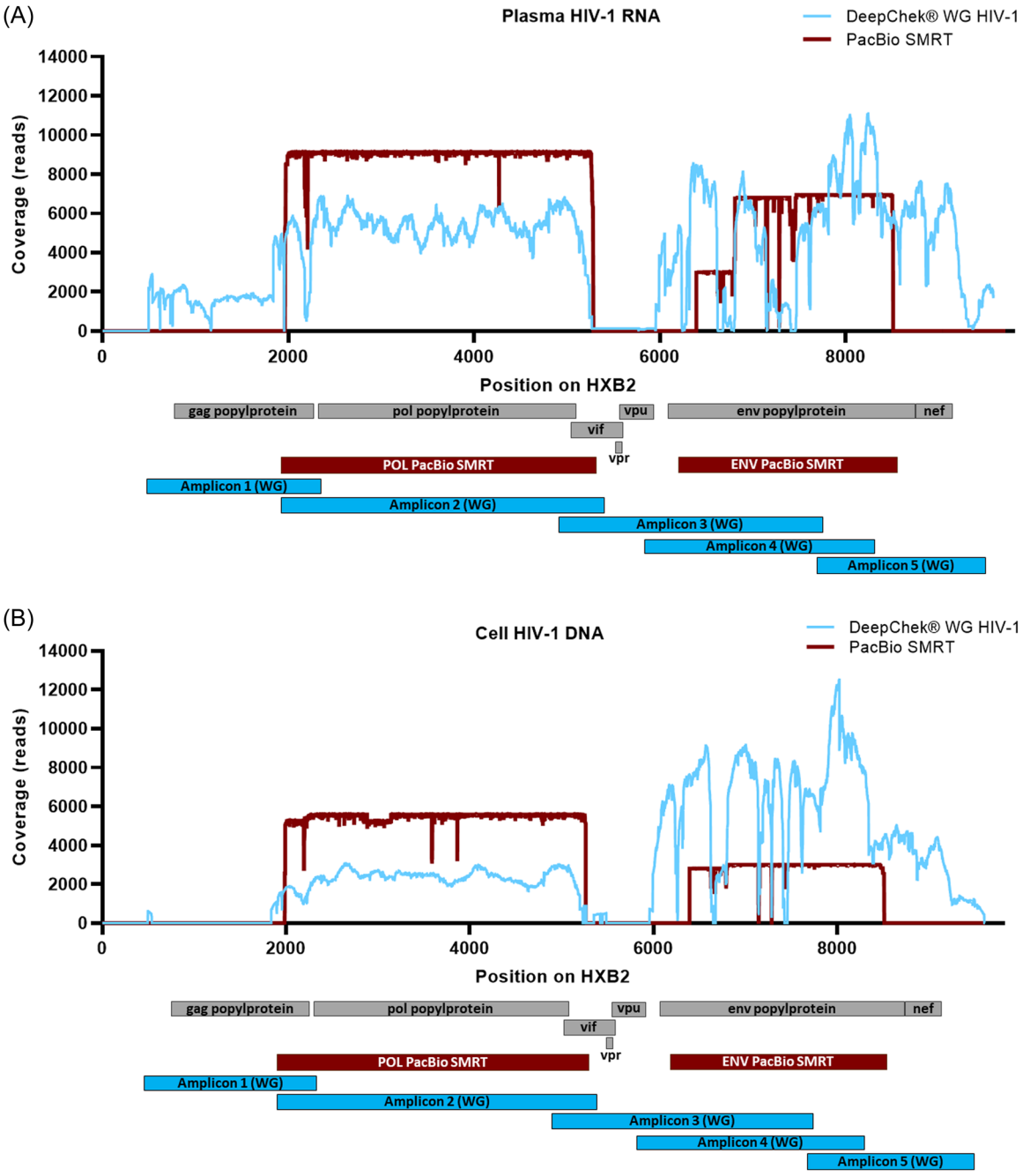

FIGURE 1 Median coverage along HIV-1 genome (reference HXB2) of plasma HIV-1 RNA (A) and cell HIV-1 DNA (B) sequences, and localization of the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT amplicons for HIV drug resistance. Gray, HIV-1 proteins; blue, DeepChek® Whole Genome HIV-1 assay and red, PacBio SMRT. PacBio, Pacific Biosciences; SMRT, single-molecule real-time sequencing.

using the 5200 Fragment Analyzer (Agilent) and quantified using the Qubit DSDNA HS Assay kit (Thermo Fisher Scientific). The resulting libraries were sequenced on an iSeq. 100 sequencer (Illumina) with a run time of 19 h. HIV libraries were subjected to paired-end sequencing (2 × 150 bp read length, forward and reverse). Sequencing analyzers' outputs (sequences) were analyzed using tailored bioinformatics pipelines from DeepChek® software as previously described by Mohamed et al.[19] Briefly, sequences were aligned against the HIV-1 reference (HXB2) using BWA (bioinformatics pipeline for Illumina outputs) and split by regions. HIV-1 subtype was determined by BLAST. Variants calling and filtering were performed by an expert system. The minimum coverage depth was 100 reads to analyze the results. The 2022 International AIDS Society list of mutations was used to identify resistance-associated mutations, and the interpretation of antiretroviral drug resistance was done by selecting the ANRS resistance algorithm (2023, v.34, available at http://www.hivfrenchresistance.org) and the Stanford HIV Drug Resistance Database interpretation system (v.9.5, available at https://hivdb.stanford.edu) on Deepchek® software.

TABLE 1 HIV-1 RNA *pol* and HIV-1 RNA *env* sequencing success and failure using the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT Sequencing on plasma samples for HIV-1 drug resistance genotyping.

			Pol region		*Env* region	
Sample ID	Subtype	Viral load (log_{10} copies/mL)	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT
1	B	5.9				
2	F1	6.6				
3	CRF02-AG	5.8				
4	CRF06	6.4				
5	CRF02-AG	5.6				
6	CRF02-AG	5.0				
7	B	5.5				
8	CRF06	5.0				
9	CRF02-AG	5.8				
10	CRF02-AG	4.5				
11	CRF02-AG	5.5				
12 (#1 1/100)	B	3.8				
13 (#1 1/1000)	B	2.8				
14 (#2 1/1000)	F1	3.6				
15 (#2 1/10000)	F1	2.6				
16 (#3 1/500)	CRF02-AG	3.1				

Note: Samples from 12 to 16 correspond to diluted plasma from samples 1, 2, and 3. Green, sequencing success; orange, partially covered; red, sequencing failure.

Two thresholds (≥5% and ≥20%) were selected for reporting. APOBEC-induced DRMs were detected using the HIV Drug Resistance Database algorithm (2021, available at https://hivdb.stanford.edu/page/apobecs/).

2.4 | PacBio SMRT sequencing

The *pol* and the *env* regions were separately amplified with a one-step RT-PCR from the extracted RNA or DNA (Figure 1). A nested PCR reaction with barcoded primers was then performed on each PCR products, as previously described.[22] Details of the PCR conditions and primers are provided in the supplementary methods. Each nested product was controlled using the 5200 Fragment Analyzer (Agilent) and quantified using the Quantifluor dsDNA HS Assay kit (Promega). PCR products were purified using AMPure PB magnetic beads (Pacific Biosciences). Barcoded amplicon libraries were prepared using the SMRT Bell Express Template Prep kit 3.0 (Pacific BioSciences) and sequenced by SMRT on a Sequel IIe System sequencer (Pacific BioSciences) with a 4-h pre-extension and 15-h movie time. HiFi reads were generated on the instrument using PacBio circular consensus sequencing (CCS) mode.

HiFi reads were demultiplexed with lima (v.2.2.0, https://github.com/PacificBiosciences/barcoding) and filtered with a minimum threshold on their quality (Q30) and their length (*env* 2000 bp, *pol* 3000 bp). The minimum coverage depth was 200 reads to analyze the results. The resulting reads were mapped against the HIV-1 reference (HXB2), and variant calling was performed with Juliet (v1.12.0, PacBio https://github.com/PacificBiosciences/minorseq). Concurrently, the filtered reads were clustered by pbAA to generate the haplotypes (PacBio tool, v.0.1.3 https://github.com/PacificBiosciences/pbAA). Chimeric and low frequency (≤2%) haplotypes were discarded. The HIV-1 subtype of each sample was determined using a maximum likelihood phylogenetic analysis processed by IQ-TREE (v.2.0.3),[23] with the best-fit model automatically selected by Model Finder (TVM + F + R9) and 1000 bootstraps (ultra-fast bootstrap), using the 2017 curated alignment from online HIV databases as the reference set of sequences. The absence of inter-sample contamination was verified using a homemade script and the phylogenetic tree. The drug resistance of each haplotype was compiled by a homemade script. We used the 2022 International AIDS Society list of mutations to identify resistance-associated mutations, and the interpretation of antiretroviral drug resistance was done using the ANRS resistance algorithm (2023, v.34, available at http://www.hivfrenchresistance.org) and the Stanford HIV Drug Resistance Database interpretation system (v.9.5, available at https://hivdb.stanford.edu). Two thresholds (≥5% and ≥20%) were selected for reporting. APOBEC-induced DRMs were detected using the HIV Drug Resistance Database algorithm (2021, available at https://hivdb.stanford.edu/page/apobecs/).

TABLE 2 Comparison of detection of HIV-1 drug-resistance mutations in RNA genotypes (cut-off 5%) using DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing.

		PR[b]		RT[c]		INT[d]		ENV[e]	
Sample ID	Viral load[a]	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT
1	5.9	33V (99%), 62V (99%)	33V + 62V (100%)	None	None	None	None	**43K (23%)**	**43K (68%)**
2	6.6	10I (99%), 20M (98%), 36I (99%), 89M (99%)	10I + 20M + 36I + 89M (100%)	None	None	None	None	**375M (99%)**	**375M (99%)**
3	5.8	20I (99%), 36I (99%), 63P (99%), 69K (99%), 89M (99%)	20I + 36I + 63P + 69K + 89M (100%)	103R (95%), 106I (99%)	103R + 106I (100%)	None	***138K (13%)***	None	None
4	6.4	16E (98%), 20I (99%), 36I (98%), 69K (99%), 77I (6%), 89M (99%)	16E + 20I + 36I + 69K + 89M (95%) 16E + 20I + 36I + 69K + 89M + 77I (5%)	None	None	None	None	None	None
5	5.6	20I (99%), 36I (70%), 64L (99%), 69K (99%), *82A (5%)*, 89M (99%)	20I + 64L + 69K + 89M + 36I (65%) 20I + 64L + 69K + 89M (35%)	None	None	None	None	None	None
6	5.0	**10V (99%)**, **16E (99%)**, 20R (70%), *20T (29%)*, 36I (97%), 69K (99%), 89I (99%)	**10V** + **16E** + 20R + 36I + 69K + 89I (100%)	None	***65E (5%)***, *101E (5%)*	None	None	None	None
7	5.5	60E (99%), 63P (99%), 71T (12%), 77I (25%)	63P + 60E (64%) 63P + 71T (14%) 63P + 60E + 77I (9%) 63P + 71T + 77I (8%) 63P (5%)	None	None	None	None	None	None
8	5.0	16E (99%), *20I (18%)*, 36I (87%), 69K (99%), 89I (100%)	16E + 36I + 69K + 89I (100%)	None	None	None	None	None	None
9	5.8	20I (99%), 36I (100%), 69K (99%), 89M (99%)	20I + 36I + 69K + 89M (100%)	None	None	None	None	None	None
10	4.5	*10V (8%)*, 20I (100%), 36I (99%), 69K (99%), *77I (11%)*, 89M (99%)	20I + 36I + 69K + 89M (100%)	108I (100%)	108I (100%)	None	None	None	None
11	5.5	20I (100%), 36I (100%), 63P (99%), 69K (99%), 89M (99%)	20I + 36I + 63P + 69K + 89M (100%)	None	None	None	None	None	None
12 (#1 1/100)	3.8	33V (99%), 62V (100%)	33V + 62V (100%)	None	None	None	None	**43K (17%)**	**43K (100%)**
13 (#1 1/1000)	2.8	33V (100%), 62V (100%)	x	None	x	None	x	**43K (31%)**	x

(Continues)

TABLE 2 (Continued)

		PR[b]		RT[c]		INT[d]		ENV[e]	
Sample ID	Viral load[a]	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT
14 (#2 1/1000)	3.6	10I (99%), 20 M (98%), 36I (100%), 89M (100%)	10I + 20M + 36I + 89M (100%)	***138Q (9%)***	None	None	None	x	x
15 (#2 1/10000)	2.6	10I (99%), 20M (99%), 36I (100%), 89M (100%)	10I + 20 M + 36I + 89 M (100%)	None	None	None	None	375M (99%)	x
16 (#3 1/500)	3.1	20I (100%), 36I (100%), 63P (99%), 69K (99%), 89M (100%)	x	106I (100%)	x	None	x	x	None

[x]genotyping failure, bold: mutation with drug resistance, italic: difference between the two approaches, + haplotype covariations.
[a]log_{10} copies/mL.
[b]protease.
[c]reverse transcriptase.
[d]integrase.
[e]envelope.

The mean error rate of amplification and sequencing was performed using HIV-1 clones giving a result of 4.52×10^{-6} nucleotides for the *pol* region and 1.85×10^{-6} nucleotides for the *env* region (supplementary methods). The accuracy and reproducibility of the method were validated by amplifying and sequencing the *env* region of HIV-1 clone mixtures with theoretical numbers of haplotypes and frequencies of 0%, 5%, 10%, 20%, 50%, 75%, and 100% (Supplementary Methods).

2.5 | Statistical analysis

Correlations between frequencies of DRMs quantified by both NGS approaches were estimated by calculating Spearman's rank correlation coefficients. All tests were two-sided, and p values < 0.05 were considered statistically significant. Statistical analyses and Graphs were performed with GraphPad Prism 9.0. The concordance rate of DRMs detection was calculated as the number of samples where the same DRMs were detected by both approaches, divided by the total number of samples successfully sequenced by both methods.

3 | RESULTS

3.1 | Plasma HIV-1 RNA

3.1.1 | Limit of detection

Firstly, we compared the viral load thresholds required to achieve sufficient coverage with the two approaches. For plasma HIV-1 RNA sequenced, we observed a median coverage of 4957 (IQR, 1734–5987) reads and 9114 (IQR, 6541–9131) reads for the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing, respectively (Figure 1A).

For the 11 plasmas, both approaches achieved a 100% success rate in sequencing the *pol* region. When considering lower viral loads obtained after plasma dilutions, the DeepChek® Whole Genome HIV-1 Assay successfully sequenced the *pol* region of all diluted plasma samples (ranging from 2.5 to 3.8 log_{10} copies/mL) whereas the PacBio SMRT sequencing approach experienced failures in two out of five samples (Table 1).

Concerning *env* sequencing on HIV-1 RNA, all 11 plasma samples were successfully sequenced. However, for the five diluted plasma samples, two were successfully sequenced using the PacBio SMRT sequencing approach, and four using the DeepChek® Whole Genome HIV-1 Assay (including one that was only partially covered) (Table 1).

3.1.2 | HIV-1 subtype determination and HIV-1 drug-resistance mutation

HIV-1 subtype prediction was performed using the *pol* sequences obtained with each NGS approach. We found a 100% concordance for the determination of plasma HIV-1 RNA subtypes. Results of

HIV-1 subtype prediction using the *env* sequences are available in the supplementary results.

Then, we compared the results of HIV-1 drug resistance by examining mutations of interest. In HIV-1 RNA sequencing, we observed an overall good correlation between the two approaches. Using a 20% threshold, we found 100% concordance for the presence of DRMs in the RT, integrase, and envelope region, and 93% in the protease region. Using a 5% threshold, we found 86% concordance for the presence of mutations of interest in the RT region, 93% in the integrase region, and 100% in the envelope region. Although a lower concordance rate of 71% was noted for the protease region, these discrepancies did not impact the interpretation regarding sensitivity to HIV-1 drug resistance (Table 2, Figure 2). When examining the percentage of mutations, disparities were observed for the 43 K

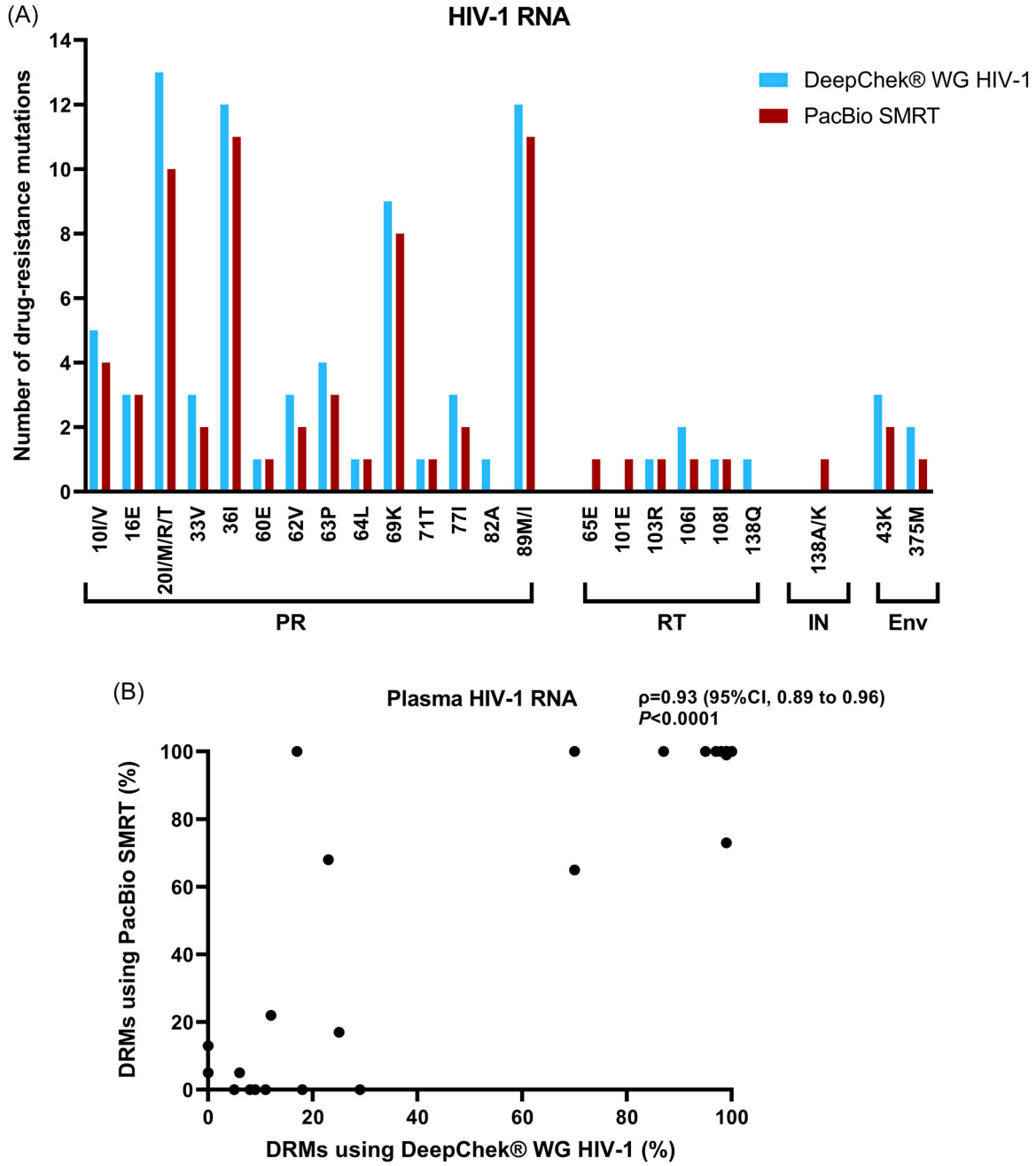

FIGURE 2 (A) Positions and number of drug-resistance mutations identified in the HIV-1 polymerase (PR, RT, IN) and envelope (env) region from 16 plasma samples using DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing. Each bar represents the number of mutations at one position of the protease (PR), reverse transcriptase (RT), integrase (IN), or envelope (Env) region identified by the indicated sequencing method from 16 plasma samples. (B) Comparison of DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing for the quantifying drug-resistance mutations in plasma HIV-1 RNA samples. Each point represents one of the 74 mutations quantified by both NGS approaches (5% threshold). Spearman's correlation coefficients (ρ) with 95% confidence intervals and p values are shown.

TABLE 3 HIV-1 DNA *pol* and HIV-1 DNA *env* sequencing success and failure using the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT Sequencing on cellular samples for HIV-1 drug resistance genotyping.

			Pol region		*Env* region	
Sample ID	Subtype	Viral load (log_{10} copies/10^6 PBMCs)	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT
1	B	<1.4				
2	B	1.5				
3	B	1.4				
4	B	3.3				
5	B	2.3				
6	B	2.8				
7	CRF02-AG	3.0				
8	B	2.9				
9	B	2.6				
10	B	1.7				
11	B	2.3				
12	CRF02-AG	2.8				

Note: Green, sequencing success; orange, partially covered; red, sequencing failure.

mutation in the *env* gene between the two approaches. However, in general, the two approaches showed similar results, even detecting minority variants (e.g., <10% for some) in some samples (e.g. plasma sample number 4), and the frequencies of the 74 DRMs quantified by both NGS approaches were well correlated (5% threshold, Spearman's correlation coefficients $\rho = 0.93$, $p < 0.0001$) (Figure 2B).

3.2 | Cell HIV-1 DNA

3.2.1 | Limit of detection

For cell HIV-1 DNA, we observed a median coverage of 2653 (IQR, 2179–5792) reads and 5420 (IQR, 3000–5582) reads for the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing, respectively (Figure 1B). We compared the ability of the two approaches in sequencing HIV-1 *pol* and HIV-1 *env* on cell HIV-1 DNA. We found that DeepChek® Whole Genome HIV-1 Assay achieved 92% and 83% of success for *pol* and *env* sequencing, respectively, whereas PacBio SMRT sequencing approach achieved 83% and 58% of success for *pol* and *env* sequencing, respectively (Table 3).

3.2.2 | HIV-1 subtype determination and HIV-1 drug-resistance mutations

Using the *pol* sequences, we determined cell HIV-1 DNA subtypes and found a 100% concordance of subtype determination between the two approaches. Results of HIV-1 subtype prediction using the *env* sequences are available in the supplementary results.

Concerning HIV-1 DNA DRMS, using a 20% threshold, we found 100% concordance for DRMs detection in the protease, integrase and envelope region, and 56% in the integrase region.

Using a 5% threshold, the two approaches exhibited lower concordance regarding the identification of variants at positions of interest in the protease, RT, integrase, and envelope regions (Table 4, Figure 3A). The DeepChek® Whole Genome HIV-1 Assay detected more APOBEC-induced DRMs, whereas these mutations were less frequently found using PacBio SMRT sequencing, specifically in the RT region. However, overall frequencies of the 59 DRMs quantified by both NGS approaches were well correlated (5% threshold, Spearman's correlation coefficients $\rho = 0.84$, $p < 0.0001$) (Figure 3B).

4 | DISCUSSION

The present study provides insights into the relative performance of two NGS approaches. Our findings suggest that the DeepChek® Whole Genome HIV-1 Assay on Illumina and HIV-1 *pol* and *env* SMRT sequencing on Sequel IIe demonstrate good performance. The DeepChek® assay shows slightly better sensitivity in HIV-1 RNA sequencing of plasmas with lower viral loads, and HIV-1 DNA sequencing. Concordance rates for determining drug resistance varied depending on the HIV-1 region, with higher concordance in the integrase region compared to the RT and protease regions.

The presence of low-frequency drug-resistant variants may affect the response to ART. Studies have demonstrated that for NNRTIs and entry inhibitors these low-frequency drug-resistant variants could impact the effectiveness of ART.[14,24–26] However, no

TABLE 4 Comparison of detection of HIV-1 drug-resistance mutations in DNA genotypes (cut-off 5%), using the DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing.

		PR[b]		RT[c]		INT[d]		ENV[e]	
Sample ID	Viral load[a]	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT	DeepChek® WG HIV-1	PacBio SMRT
1	<1.4	*16E (15%)*, 36I (100%)	36I (99%)	179I (99%), ***184I (17%)***,[f] ***230I (19%)***[f]	179I (100%)	None	None	x	x
2	1.5	63P (53%), *77I (19%)*	63P (99%)	***184I (27%)***,[f] ***230I (27%)***[f]	None	none	None	None	None
3	1.4	x	62V + 63P + 77I (100%)	x	None	x		None	x
4	3.3	63P (98%), 77I (42%),	x	69D (6%), **230I (6%)**[f]	x	**74I (99%)**	x	**44M (99%)** **375I (10%)**	***36S (11%)*** **44M (71%)** **44M + 375I (17%)**
5	2.3	71V (99%)	*36V* + 71V (12%) 71V (27%)	*none*	***41L + 67N + 184V + 210W + 215Y (12%)***[f]	None	None	None	***36S + 375T (5%)***
6	2.8	63P (95%), 77I (98%)	63P + 77I (100%)	***184I (6%)***[f]	None	None	None	None	x
7	3.0	10I (13%), 16E (97%), 69K (99%), 89M (99%)	10I + 16E + 69K + 89M (20%) 16E + 69K + 89M (80%)	***90I (4%)***, ***98S (14%)***, **103N (84%)**, ***138G (19%)***, ***184I (8%)***,[f] ***184V (24%)***, **225H (24%)**, ***230I (9%)***[f]	**103N (77%), 225H (9%)**	***138K (5%)***, **163R (7%)**[f]	**163R (8%)**[f]	x	**36S (12%)**
8	2.9	63P (97%), 77I (99%)	x	**215Y (37%)**, **230I (19%)**[f]	x	None	x	**375I (62%)**, **434I (99%)**	**375I + 434I (78%)** **434I (22%)**
9	2.6	36I (74%), 62V (97%), 63P (96%), 71V (97%)	36I + 62V + 71V (15%) 62V + 63P + 71V (34%) 36I + 62V + 63P + 71V (51%)	None	***184I (34%)***[f]	None	None	**426L (12%)**	x
10	1.7	None	None	None	***67N + 70R + 215Y + 219E (74%)*** ***215Y (13%)***[f]	None	None	**36S (21%)**	**36S (12%)**
11	2.3	63P (87%), 77I (31%)	63P (63%) 63P + 77I (29%)	None	None	None	None	None	x
12	2.8	36I (100%), 69K (98%), 89M (99%)	36I + 69K + 89M (100%)	None	*90I (14%)*	**157Q (95%)**	**157Q (100%)**	***434I (6%)***	None

[x]genotyping failure, bold: mutation with drug resistance, italic: difference between the two approaches, + haplotype covariations.

[a]log_{10} copies/10^6 PBMCs.

[b]protease.

[c]reverse transcriptase.

[d]integrase.

[e]envelope.

[f]the following drug-resistance mutations could reflect APOBEC activity.

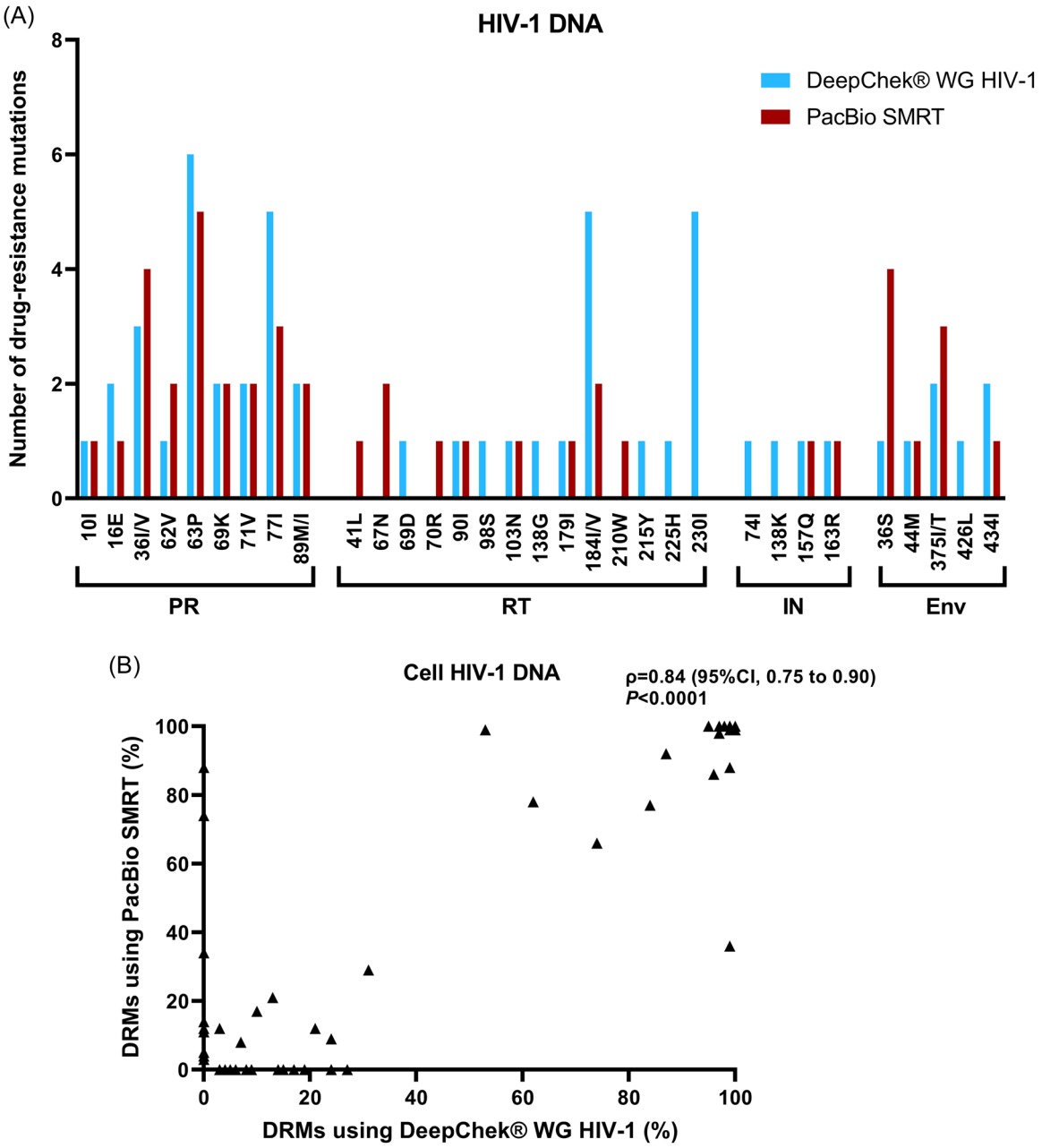

FIGURE 3 (A) Positions and number of drug-resistance mutations identified in the HIV-1 polymerase (PR, RT, IN) and envelope (env) region from 12 cellular samples using DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing. Each bar represents the number of mutations at one position of the protease (PR), reverse transcriptase (RT), integrase (IN), or envelope (Env) region identified by the indicated sequencing method from 12 cellular samples. (B) Comparison of DeepChek® Whole Genome HIV-1 Assay and PacBio SMRT sequencing for the quantifying drug-resistance mutations in cellular HIV-1 DNA samples. Each point represents one of the 59 mutations quantified by both NGS approaches (5% threshold). Spearman's correlation coefficients (ρ) with 95% confidence intervals and p values are shown.

studies have yet found the impact of these low-frequency drug-resistant variants on the effectiveness of NRTIs, PIs, or integrase inhibitors.[27,28]

For plasma HIV-1 RNA, with a 5% cut-off we observed concordance rates between the two NGS approaches of 71%, 86%, 93%, and 100% for the protease, RT, integrase, and envelope regions, respectively, based on the presence of mutations of interest. When examining the value of the percentages, some discrepancies were notably found in the *env* region known for its genetic variability. However, in some other samples and regions, comparable mutation percentages were observed between the two approaches, even for minority variants (<10%). In analyzing diluted plasma samples with the DeepChek® Whole Genome HIV-1 Assay, we detected the presence of mutation 138Q at 9%. Interestingly, this mutation was absent in the original plasma sample or its other dilutions, and was not identified using PacBio SMRT. While inadequate coverage in these regions could have accounted for the disparities, our analysis revealed sufficient coverage.

Despite the inherent differences between the two NGS platforms (short and long reads), the regions targeted by the initial PCR for both methods are relatively similar. The *pol* PacBio SMRT amplicon closely aligns with the amplicon 2 of the ABL Deepchek WG kit, and the *env* PacBio SMRT amplicon is quite close to the amplicon 4 of the ABL Deepchek WG kit. However, the ABL Deepchek WG assay is designed so that parts of the HIV-1 envelope are also covered by amplicons 3 and 5 that could contribute to a better coverage of the envelope region.

Both approaches take 3 days from extraction to sequencing results. A strength of the DeepChek® Whole Genome HIV-1 Assay, is its kit-based approach which reduces the risks of technical errors, and the Deepchek® software that enables automated bioinformatics analyses from the fastq files. The PacBio SMRT requires the development of a bioinformatic pipeline and specific computing resources, however it allows haplotyping analysis. DeepChek® Whole Genome HIV-1 Assay on Illumina is a short-reads sequencing method, involving amplicon fragmentation before sequencing. In contrast, PacBio SMRT sequencing is a long-reads approach that allows detection of mutations on the same genomic fragment. The lower sensitivity of PacBio SMRT could be due to the long reads approach requiring optimization. However, the long reads sequencing allows the identification of DRMs occurring on the same haplotype, providing information on HIV-1 drug-resistance at a quasispecies level instead of focusing on the level of individual mutations. Mutations detected on the same haplotype or on distinct haplotypes could have a different impact on ARV drug activity. Transmitted drug-resistant variants can acquire new mutations and the synergistic effect of certain mutations on a single haplotype can contribute to the development of drug resistance.[29–31]

A previous study has demonstrated the importance of selecting an appropriate threshold for HIV-1 DNA NGS to achieve a balance between sensitivity and the presence of defective proviruses resulting from APOBEC activity.[32] In our study, using a 5% threshold, we identified mutations of interest in the *pol* region that can arise from G-to-A mutations induced by APOBEC in 6 out of 12 (50%) cellular samples using the DeepChek® Whole Genome HIV-1 Assay, and in 4 out of 12 (33%) cellular samples using PacBio SMRT sequencing. However, when we increased the threshold to 10%, the proportion of APOBEC-induced mutations of interest detected dropped: 3 out of 12 (25%) with both approaches. This discrepancy between the two NGS approaches does not appear to be related to the filtering used during sequencing analyses as we retrospectively examined at the raw sequence data obtained with the PacBio SMRT sequencing, and did not find more APOBEC-induced DRMs. Our findings suggest that this observation may likely be due to the first PCR that is influenced by the low viral load and Poisson distribution rules conducted to the amplification of certain haplotypes. These discrepancies highlight the challenges of sequencing and interpreting HIV-1 DNA in cellular samples, which contain an archived reservoir with both intact and defective proviruses.

Limitations of our study include the small number of samples sequenced and the few HIV-1 subtypes tested. However, our data demonstrate that both NGS approaches can successfully sequence recombinants subtypes such as CRF02-AG and CRF06. Further testing is necessary to validate the performance of these methods on additional non-B subtypes.

In conclusion, our study presented the comparison of two NGS approaches, one "short read" and one "long read" sequencing protocol, for sequencing plasma HIV-1 RNA and cell HIV-1 DNA. Our findings indicate good performance with both approaches for the detection and quantification of DRMs in the *pol* and *env* regions of the HIV-1 genome. Further studies are needed to analyze the relevance of APOBEC mutations detected in cell HIV-1 DNA sequencing for the interpretation of DRMs.

AUTHOR CONTRIBUTIONS

Camille Vellas designed the study, collected the data, and wrote the original draft. Camille Vellas and Jacques Izopet designed the study. Jacques Izopet supervised the study, reviewed, and edited the manuscript. Amira Doudou, Sofiane Mohamed, and Noémie Ranger contributed to the experiments. Stéphanie Raymond participated in collecting and analyzing the data. Nicolas Jeanne, Justine Latour, and Sofia Demmou performed bio-informatics analyses. Dimitri Gonzalez contributed to the conception of the study. Pierre Delobel helped to provide clinical samples. All authors contributed to read and approved the version of the submitted manuscript.

ACKNOWLEDGMENTS

The authors acknowledge ABLD for providing DeepChek® whole genome HIV-1 assay and NGS library preparation kits.

CONFLICT OF INTEREST STATEMENT

Amira Doudou, Sofiane Mohamed, and Dimitri Gonzalez are employees of the ABLD group.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

According to the French law (Loi Jardé), anonymous retrospective studies do not require institutional review board approval.

ORCID

Camille Vellas http://orcid.org/0000-0002-8637-6695
Sofiane Mohamed https://orcid.org/0009-0002-4741-1432

REFERENCES

1. HIV drug resistance report 2019. Accessed February 7, 2024. https://www.who.int/publications-detail-redirect/WHO-CDS-HIV-19.21
2. Allavena C, Rodallec A, Leplat A, et al. Interest of proviral HIV-1 DNA genotypic resistance testing in virologically suppressed patients candidate for maintenance therapy. *J Virol Methods*. 2018;251:106-110. doi:10.1016/j.jviromet.2017.10.016
3. Armenia D, Zaccarelli M, Borghi V, et al. Resistance detected in PBMCs predicts virological rebound in HIV-1 suppressed patients

switching treatment. *J Clin Virol*. 2018;104:61-64. doi:10.1016/j.jcv.2018.04.001
4. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*. 2003;424(6944):99-103. doi:10.1038/nature01709
5. Noguera-Julian M, Cozzi-Lepri A, Di Giallonardo F, et al. Contribution of APOBEC3G/F activity to the development of low-abundance drug-resistant human immunodeficiency virus type 1 variants. *Clin Microbiol Infect*. 2016;22(2):191-200. doi:10.1016/j.cmi.2015.10.004
6. Segal-Maurer S, DeJesus E, Stellbrink HJ, et al. Capsid inhibition with lenacapavir in multidrug-resistant HIV-1 infection. *N Engl J Med*. 2022;386(19):1793-1803. doi:10.1056/NEJMoa2115542
7. Emu B, Fessel J, Schrader S, et al. Phase 3 study of ibalizumab for multidrug-resistant HIV-1. *N Engl J Med*. 2018;379(7):645-654. doi:10.1056/NEJMoa1711460
8. Gaebler C, Nogueira L, Stoffel E, et al. Prolonged viral suppression with anti-HIV-1 antibody therapy. *Nature*. 2022;606(7913):368-374. doi:10.1038/s41586-022-04597-1
9. Eigen M. Viral quasispecies. *Sci Am*. 1993;269(1):42-49. doi:10.1038/scientificamerican0793-42
10. Zanini F, Brodin J, Thebo L, et al. Population genomics of intrapatient HIV-1 evolution. *eLife*. 2015;4:e11282. doi:10.7554/eLife.11282
11. Kemp SA, Charles OJ, Derache A, et al. HIV-1 evolutionary dynamics under nonsuppressive antiretroviral therapy. *mBio*. 2022;13(3):e0026922. doi:10.1128/mbio.00269-22
12. Dykes C, Najjar J, Bosch RJ, et al. Detection of drug-resistant minority variants of HIV-1 during virologic failure of indinavir, lamivudine, and zidovudine. *J Infect Dis*. 2004;189(6):1091-1096. doi:10.1086/382033
13. Halvas EK, Wiegand A, Boltz VF, et al. Low frequency nonnucleoside reverse-transcriptase inhibitor-resistant variants contribute to failure of efavirenz-containing regimens in treatment- experienced patients. *J Infect Dis*. 2010;201(5):672-680. doi:10.1086/650542
14. Li JZ. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA*. 2011;305(13):1327-1335. doi:10.1001/jama.2011.375
15. Li JZ, Kuritzkes DR. Clinical implications of HIV-1 minority variants. *Clin Infect Dis*. 2013;56(11):1667-1674. doi:10.1093/cid/cit125
16. Inzaule SC, Hamers RL, Noguera-Julian M, et al. Clinically relevant thresholds for ultrasensitive HIV drug resistance testing: a multi-country nested case-control study. *Lancet HIV*. 2018;5(11):e638-e646. doi:10.1016/S2352-3018(18)30177-2
17. Simen BB, Simons JF, Hullsiek KH, et al. Low-abundance drug-resistant viral variants in chronically HIV-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes. *J Infect Dis*. 2009;199(5):693-701. doi:10.1086/596736
18. Boltz VF, Zheng Y, Lockman S, et al. Role of low-frequency HIV-1 variants in failure of nevirapine-containing antiviral therapy in women previously exposed to single-dose nevirapine. *Proc Natl Acad Sci*. 2011;108(22):9202-9207. doi:10.1073/pnas.1105688108
19. Mohamed S, Boulmé R, Sayada C. From capillary electrophoresis to deep sequencing: an improved HIV-1 drug resistance assessment solution using in vitro diagnostic (IVD) assays and software. *Viruses*. 2023;15(2):571. doi:10.3390/v15020571
20. Sauné K, Raymond S, Boineau J, Pasquier C, Izopet J. Detection and quantification of HIV-1 RNA with a fully automated transcription-mediated-amplification assay. *J Clin Virol*. 2016;84:70-73. doi:10.1016/j.jcv.2016.09.002
21. Véronique AF, Marie-Laure C, Stéphane B, et al. LTR real-time PCR for HIV-1 DNA quantitation in blood cells for early diagnosis in infants born to seropositive mothers treated in HAART area (ANRS CO 01). *J Med Virol*. 2009;81(2):217-223. doi:10.1002/jmv.21390
22. Raymond S, Jeanne N, Vellas C, et al. HIV-1 genotypic resistance testing using single molecule real-time sequencing. *J Clin Virol*. 2024;174:105717. doi:10.1016/j.jcv.2024.105717
23. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268-274. doi:10.1093/molbev/msu300
24. Bansal V, Metzner KJ, Niederöst B, et al. Minority K65R variants and early failure of antiretroviral therapy in HIV-1-infected Eritrean immigrant. *Emerging Infect Dis*. 2011;17(10):1966-1968. doi:10.3201/eid1710.110592
25. Raymond S, Nicot F, Pallier C, et al Impact of human immunodeficiency virus type 1 minority variants on the virus response to a rilpivirine-based first-line regimen. *Clin Infect Dis*. 2018;66(10):1588-1594. doi:10.1093/cid/cix1070
26. Swenson LC, Mo T, Dong WWY, et al. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. *J Infect Dis*. 2011;203(2):237-245. doi:10.1093/infdis/jiq030
27. Metzner KJ, Rauch P, von Wyl V, et al. Efficient suppression of minority drug-resistant HIV type 1 (HIV-1) variants present at primary HIV-1 infection by ritonavir-boosted protease inhibitor-containing antiretroviral therapy. *J Infect Dis*. 2010;201(7):1063-1071. doi:10.1086/651136
28. Mbunkah HA, Bertagnolio S, Hamers RL, et al. Low-abundance drug-resistant HIV-1 variants in antiretroviral drug-naive individuals: a systematic review of detection methods, prevalence, and clinical impact. *J Infect Dis*. 2020;221(10):1584-1597. doi:10.1093/infdis/jiz650
29. Vingerhoets J, Azijn H, Fransen E, et al. TMC125 displays a high genetic barrier to the development of resistance: evidence from in vitro selection experiments. *J Virol*. 2005;79(20):12773-12782. doi:10.1128/JVI.79.20.12773-12782.2005
30. Feng M, Sachs NA, Xu M, et al. Doravirine suppresses common nonnucleoside reverse transcriptase inhibitor-associated mutants at clinically relevant concentrations. *Antimicrob Agents Chemother*. 2016;60(4):2241-2247. doi:10.1128/AAC.02650-15
31. Boltz VF, Shao W, Bale MJ, et al. Linked dual-class HIV resistance mutations are associated with treatment failure. *JCI Insight*. 2019;4(19). doi:10.1172/jci.insight.130118
32. Armenia D, Santoro MM, Bellocchi MC, et al. Viral resistance burden and APOBEC editing correlate with virological response in heavily treatment-experienced people living with multi-drug resistant HIV. *Int J Antimicro Ag*. 2022;59(1):106492. doi:10.1016/j.ijantimicag.2021.106492

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vellas C, Doudou A, Mohamed S, et al. Comparison of short-read and long-read next-generation sequencing technologies for determining HIV-1 drug resistance. *J Med Virol*. 2024;96:e29951. doi:10.1002/jmv.29951