illumina®

APPLICATION NOTE

Sequencing the HIV genome on the MiSeq™ i100 Series

Accurate detection and characterization of HIV-1

Rapid interrogation of the HIV *pol* gene with Illumina Microbial Amplicon Prep for viral detection and drug resistance detection

Comprehensive HIV genome sequencing with the Viral Surveillance Panel v2 for tracking of viral evolution and detection of > 200 other viral pathogens

Flexible, end-to-end workflow with sequencing on the MiSeq i100 Series and DRAGEN™ secondary analysis

Introduction

Symptoms of acquired immunodeficiency syndrome (AIDS) were first recognized in 1981, and in the following years attributed to infection by human immunodeficiency viruses type 1 and 2 (HIV-1 and HIV-2).[1–4] The HIV pandemic has seen more than 88 million people worldwide infected with HIV and more than 42 million lives lost to the disease.[5] Approximately 1.3 million new infections and 630,000 deaths were reported in 2023 alone.[6]

Advances in disease prevention and the development of effective antiretroviral drugs have largely transitioned the HIV pandemic to a manageable chronic illness, significantly reducing the risk of HIV transmission and disease progression to AIDS.[7–9] However, ongoing, sustained transmission of HIV among vulnerable populations in regions around the world highlights the continued need for HIV surveillance.[8] Molecular surveillance via next-generation sequencing (NGS) enables detection of variants that confer resistance to antiretroviral drugs, tracking of HIV transmission, including transmission of drug-resistant HIV, detection of variants associated with HIV tropism, and tracking of intrahost viral evolution.[10,11] In particular, mutations in the HIV *pol* gene can confer resistance to protease inhibitors (PIs), nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), and integrase strand transfer inhibitors (INSTIs).[12] As such, the HIV *pol* gene is an important target for genotypic detection of antiretroviral drug resistance.

This application note demonstrates HIV detection and characterization in HIV-positive blood plasma and contrived samples using an NGS workflow that integrates Illumina library preparation, sequencing on the MiSeq i100 Series, and DRAGEN secondary analysis (Figure 1).

Methods

Samples

HIV-positive plasma samples were obtained from BIOFLUIDS.com and RNA was extracted using the QIAamp Viral RNA Mini Kit (QIAGEN, Catalog no. 52904), using the Mini Spin procedure, omitting carrier RNA addition to Buffer AVL.* Contrived samples were formulated for testing using purified RNA extracted from virus propagated in peripheral blood mononuclear cells (PBMCs) provided from SeraCare (Table 1). Controls representing four prevalent HIV-1 subtypes were spiked into a background of 10 ng of Human Universal Reference RNA (UHRR) for a total input of 690 viral copies/µl for Illumina Microbial Amplicon Prep and 1176 viral copies/µl for Illumina RNA Prep with Enrichment with the Viral Surveillance Panel v2 to assess performance across differing HIV-1 subtypes belonging to group M (Table 1).[13]

* Sufficient input of viral nucleic acid is critical for good performance. Extraction method considerations such as centrifugation with sucrose (sucrose gradient) or addition of carrier RNA could improve sensitivity, particularly for low-titer samples.[14]

Figure 1: Comprehensive NGS workflow for HIV surveillance
Combine Illumina library preparation using an amplicon-based or target enrichment approach with sequencing on the MiSeq i100 Series and DRAGEN secondary analysis for accurate genomic characterization of HIV.

Table 1: Purified RNAs sourced for subtype controls

SeraCare product name	Material no.	Isolate ID
HIV-1 Purified RNA subtype C	0400-0079	DJ259
HIV-1 Purified RNA subtype B	0400-0078	US1
HIV-1 Purified RNA subtype CRF01-AE	0400-0084	POC30506
HIV-1 Purified RNA subtype CRF02-AG	0400-0076	POC44951

Library preparation

Targeted sequencing of HIV was performed via amplicon sequencing using Illumina Microbial Amplicon Prep (Illumina, Catalog no. 20097857) and by target enrichment using Illumina RNA Prep with Enrichment with the Viral Surveillance Panel v2 (Illumina, Catalog no. 20108081).

Amplicon sequencing

For Illumina Microbial Amplicon Prep, custom primers, including previously published primers[15] and supplementary primers designed by Illumina scientists to ensure amplification of the HIV *pol* gene for multiple HIV-1 subtypes, were sourced from Integrated DNA Technologies (IDT) and diluted to 10 µM per pool (Table 2). Because the amplicons targeting the *pol* gene overlap in position relative to the HIV-1 genome, primers were pooled into two separate tubes and used for amplicon PCR in two independent reactions. This approach is needed to prevent small amplicon generation and enables ARTIC-style amplicon schemes[16] (Figure 2).

Table 2: Custom primers used for amplification of HIV-1 with Illumina Microbial Amplicon Prep

Primer name	Primer sequence 5′–3′
P1-F v2	AAGGGYTGYTGGAAATGYGG
P1-R v2	CTGTADTTCTGCTAYTAAMTCTTTTGATGG
P1-F	TTGGAAATGTGGAAAGGAAGGAC
P1-R	CTGTATTTCTGCTATTAAGTCTTTTGATGGG
P5-F	GGAATCATTCAAGCACAACCAGA
P5-F v2	TATGCAYTAGGAATYATTCARGCACA
P5-R	TCTCCTGTATGCAGACCCCAATAT
P5-R v2	TCTCCTGTATGCAGACCCCAATATG
Primer name	**Primer sequence 5′–3′**
YL1	AGAACCYCCATTCCTTYGGATGGG
YL1 v2	AAGCATCAAAAGGAACCTCCCTT
YL4	CCTTTGTGTGCTGGTACCCATG
YL4 v2	CCACATGGACAGCAACTATTATG

Supplementary primers designed by Illumina scientists for increased taxonomic coverage are denoted by "v2" in the primer name, supplied by Integrated DNA Technologies.

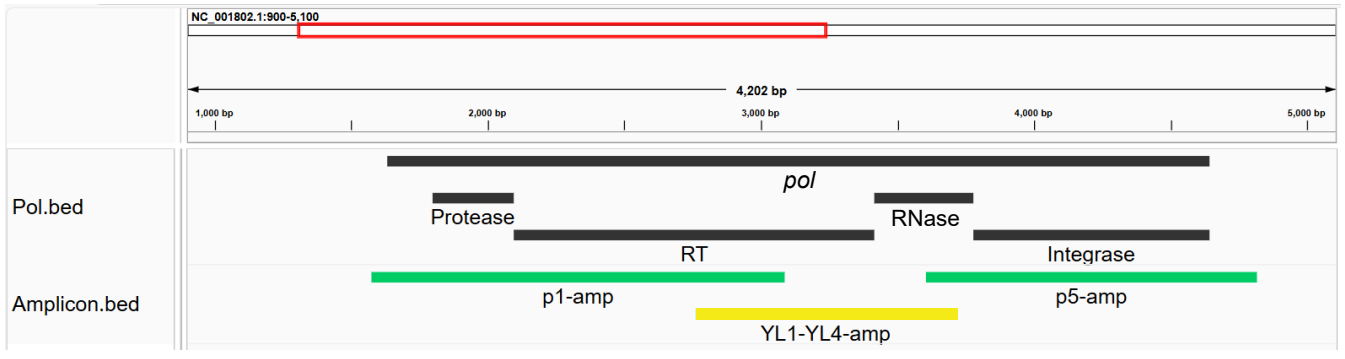


Figure 2: Illumina Microbial Amplicon Prep primer design for HIV-1 *pol* gene
Primers for p1 and p5 amplicons (green) were used in one PCR reaction while primers for the YL1-YL4 amplicon (yellow) were used in a second PCR reaction to cover the *pol* gene (black).

Sequencing-ready libraries were prepared with slight modifications to the Illumina Microbial Amplicon Prep protocol, including a reduction in random hexamer input (2.5 µl instead of 8.5 µl), RNA input (14.5 µl instead of 8.5 µl), and tagmentation reaction amplicon input volumes (10 µl from p1 and p5 amplification reactions, 5 µl from the YL1-YL4 amplification reaction, and 5 µl of DEPC H_20).

Target enrichment

Illumina RNA Prep with Enrichment with the Viral Surveillance Panel v2 targets the entire genome of HIV-1 and HIV-2, providing whole-genome sequencing (WGS) of HIV and ~200 additional viruses. Sequencing-ready libraries were prepared following the manufacturer's protocol.

Sequencing

Prepared libraries were sequenced on the MiSeq i100 Plus System using a 25M flow cell with a run configuration of 2 × 150 bp.

Data analysis

After sequencing was complete, data were normalized to 0.5M clusters/1M paired-end (PE) reads and 1M clusters/2M PE reads for Illumina Microbial Amplicon Prep and Illumina RNA Prep with Enrichment with Viral Surveillance Panel v2, respectively, using the FASTQ Toolkit App. Normalized data were analyzed using the DRAGEN Microbial Amplicon and DRAGEN Microbial Enrichment Plus apps in the cloud in BaseSpace™ Sequence Hub. These apps can also be accessed onboard the MiSeq i100 Plus System. To determine drug resistance information, normalized FASTQs were used as input into the publicly available Stanford database.[17]

The DRAGEN Microbial Enrichment Plus app performs a reference-guided assembly; therefore, the correct panel and associated set of genome reference sequences of viruses targeted by the panel is necessary. Therefore, the Viral Surveillance Panel v2 must be selected from the enrichment panel dropdown menu in the app. Despite the large and diverse set of HIV-1 reference genomes used in the selected application workflow for detection and alignment, it is possible that a more representative reference genome for a given sample is available and more appropriate to use for analysis. Thus an iterative approach was taken, re-running either the DRAGEN Microbial Enrichment Plus or DRAGEN Microbial Amplicon workflows (Figure 3).

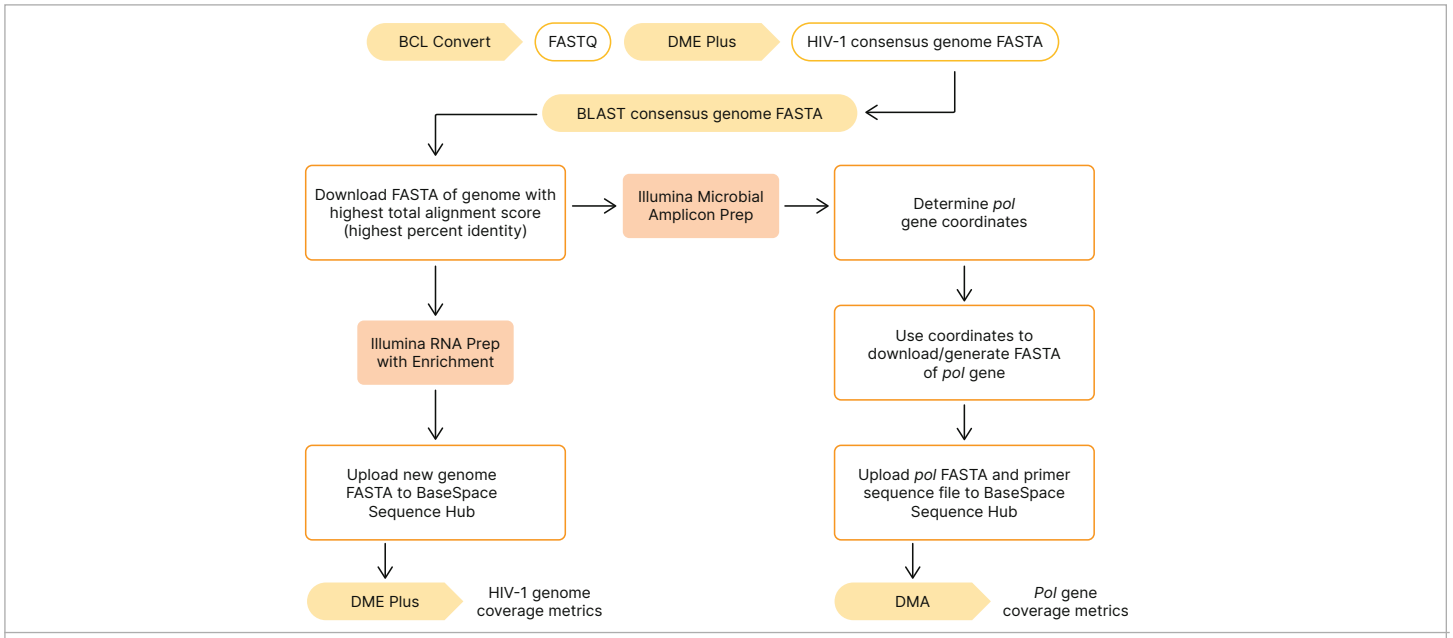


Figure 3: DRAGEN Microbial Enrichment Plus/DRAGEN Microbial Amplicon iterative analysis workflow
In addition to coverage metrics, variant calling and consensus sequence generation are also output in the DRAGEN Microbial Amplicon (DMA) and DRAGEN Microbial Enrichment Plus (DME plus) workflows.

Results

HIV-1 *pol* gene coverage with Illumina Microbial Amplicon Prep

Analysis of Illumina Microbial Amplicon Prep sequencing data from blood plasma sample using the DRAGEN Microbial Enrichment Plus app resulted in detection of HIV-1 and generation of a consensus HIV-1 FASTA file that was subsequently used for BLAST analysis to determine a representative accession FASTA file. To attain *pol* gene–specific metrics, the Gene Cutter tool from the Los Alamos National Laboratory HIV Sequence Database was used to parse the HIV-1 genome sequence to just the *pol* gene region. Alignment of the *pol* output to a reference sequence using the DRAGEN Microbial Amplicon app showed 100% coverage across the *pol* gene (Figure 4).

For contrived samples, DRAGEN Microbial Enrichment Plus analysis detected HIV-1 in all Illumina Microbial Amplicon Prep library technical replicates. Re-analysis with DRAGEN Microbial Amplicon was done for all samples using the iterative analysis workflow. DRAGEN Microbial Amplicon analysis demonstrates full *pol* gene coverage for each contrived sample tested (Figure 5).

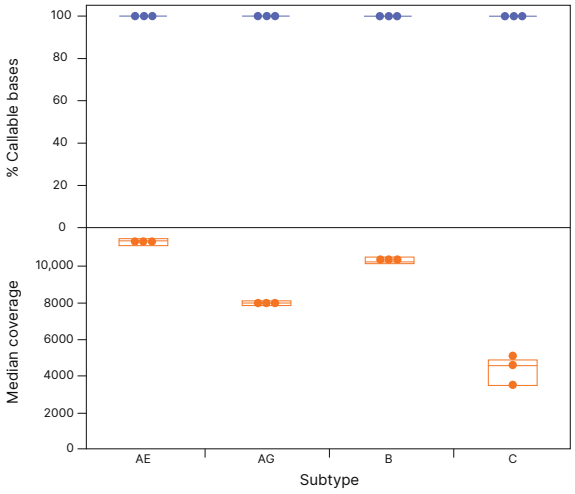


Figure 5: Coverage metrics for HIV-1 subtypes with Illumina Microbial Amplicon Prep
Percent callable bases and median depth of coverage for the *pol* gene. Results show that the selected primer pool amplifies HIV-1 strains representing diverse subtypes.

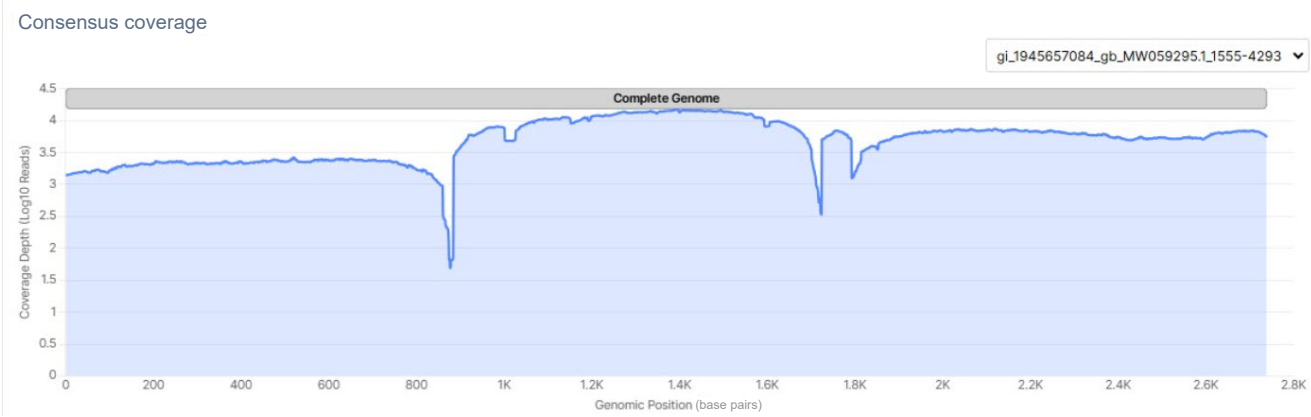


Figure 4: Genome coverage of HIV-1 *pol* region with Illumina Microbial Amplicon Prep
Genome coverage plot for a representative plasma sample analyzed with the DRAGEN Microbial Enrichment Plus app. Results show full coverage (100%) across the *pol* gene.

HIV-1 genome coverage with Illumina RNA Prep with Enrichment

DRAGEN Microbial Enrichment Plus analysis detected HIV-1 in some unenriched and all Viral Surveillance Panel v2–enriched libraries at varying viral copy inputs (Figure 6). Initial DRAGEN Microbial Enrichment Plus analysis produced coverage plots for HIV-1 detected samples. For one representative plasma sample, 95.27% of the HIV-1 genome was covered (≥ 1× depth) with a median depth of 757.5× and 114,150 aligned reads (Figure 7A). The consensus genome sequence from this analysis was analyzed using BLAST and the Nucleotide database. The highest scoring HIV-1 genome was used as input into the iterative DRAGEN Microbial Enrichment Plus analysis workflow. This resulted in greater coverage metrics with 99.11% of the genome covered (≥ 1× depth) (Figure 7B).

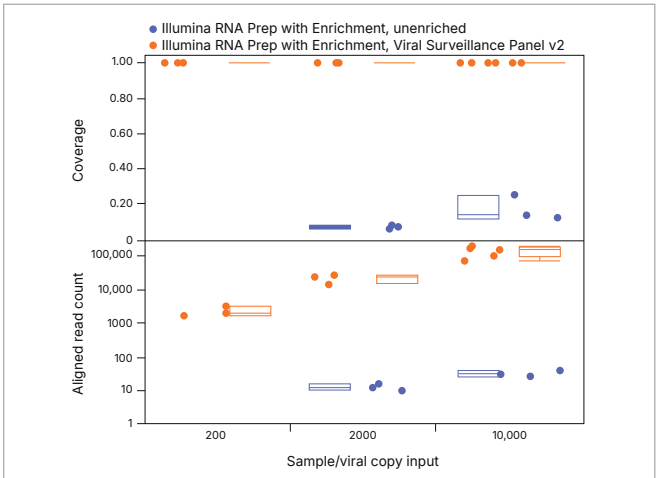
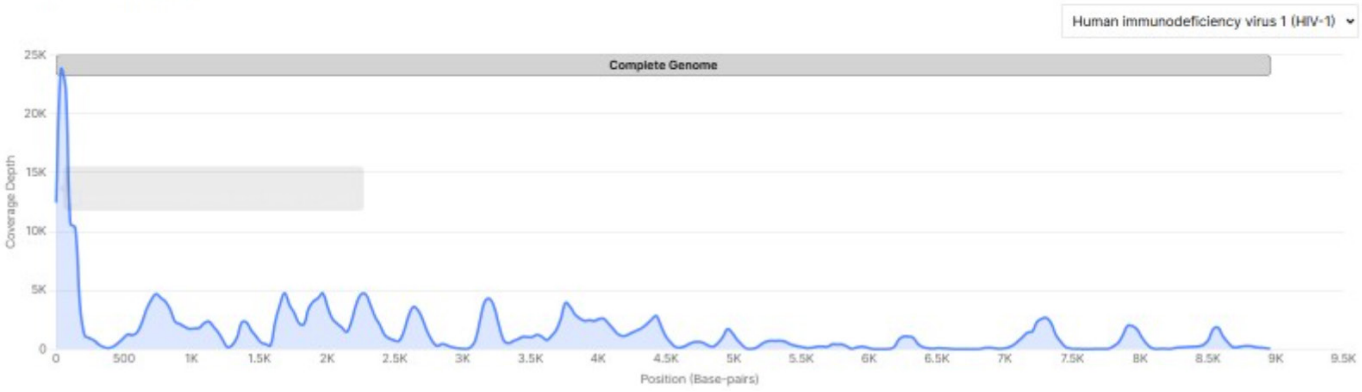


Figure 6: Increased coverage with Illumina RNA Prep with Enrichment relative to unenriched libraries
HIV-1 genome coverage for subtype B contrived sample. Low viral reads detected in unenriched libraries demonstrate the use of amplification or enrichment.

A. Reference coverage—initial analysis



B. Reference coverage—iterative analysis workflow

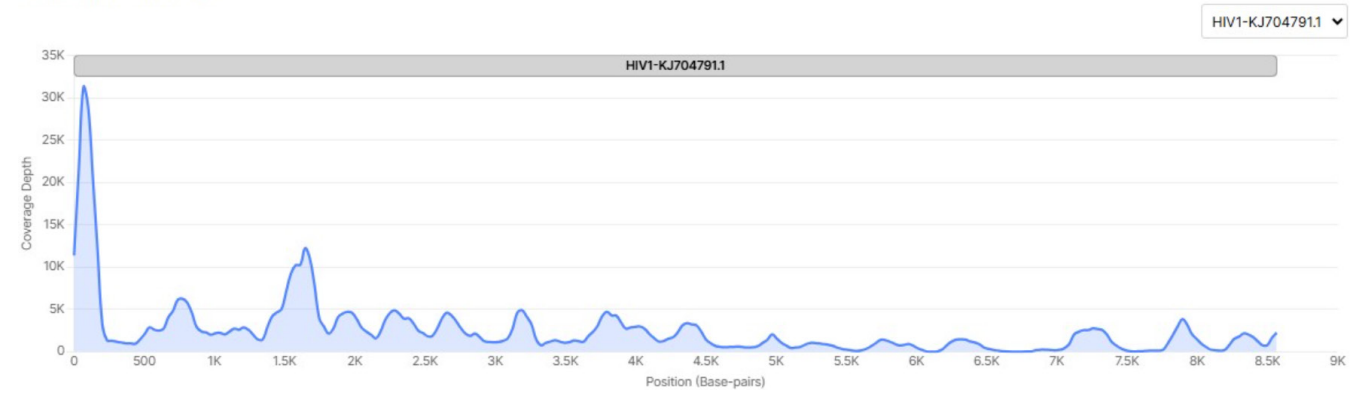


Figure 7: Genome coverage of HIV-1 with Illumina RNA Prep with Enrichment
Genome coverage plots for a representative blood plasma sample (A) analyzed with the DRAGEN Microbial Enrichment Plus app and (B) analyzed with the DRAGEN Microbial Enrichment Plus iterative workflow. Results show that the iterative workflow improved HIV-1 genome coverage from 95.27% to 99.11%.

Assay for drug resistance mutations

Normalized FASTQs from both Illumina Microbial Amplicon Prep and Illumina RNA Prep with Enrichment assays for subtype B, 10K viral copy input, were analyzed using the Stanford HIV Drug Resistance Database to assess presence of DRMs with default parameters (minimum read depth threshold ≥ 50, nucleotide mixture threshold ≤ 2%, mutation detection threshold ≥ 10%). No major or minor PI, NRTI/NNRTI, or INSTI mutations were detected in data generated from either assay. Other mutations (n = 24) predicted by the HIV-1 genome sequence provided by SeraCare were detected using both assays (Figure 8).

Summary

The MiSeq i100 Series combined with high-quality library prep using both Illumina Microbial Amplicon Prep and Illumina RNA Prep with Enrichment with the Viral Surveillance Panel v2 demonstrated targeted sequencing capabilities for HIV-1 detection and drug resistance profiling. While Illumina Microbial Amplicon Prep provided deeper coverage of the HIV *pol* gene with a simpler workflow, Illumina RNA Prep with Enrichment with the Viral Surveillance Panel v2 enrichment yielded full genome coverage of HIV strains. This application note demonstrates that the MiSeq i100 Series is part of a flexible, end-to-end NGS workflow for HIV characterization that can be tailored to user needs.

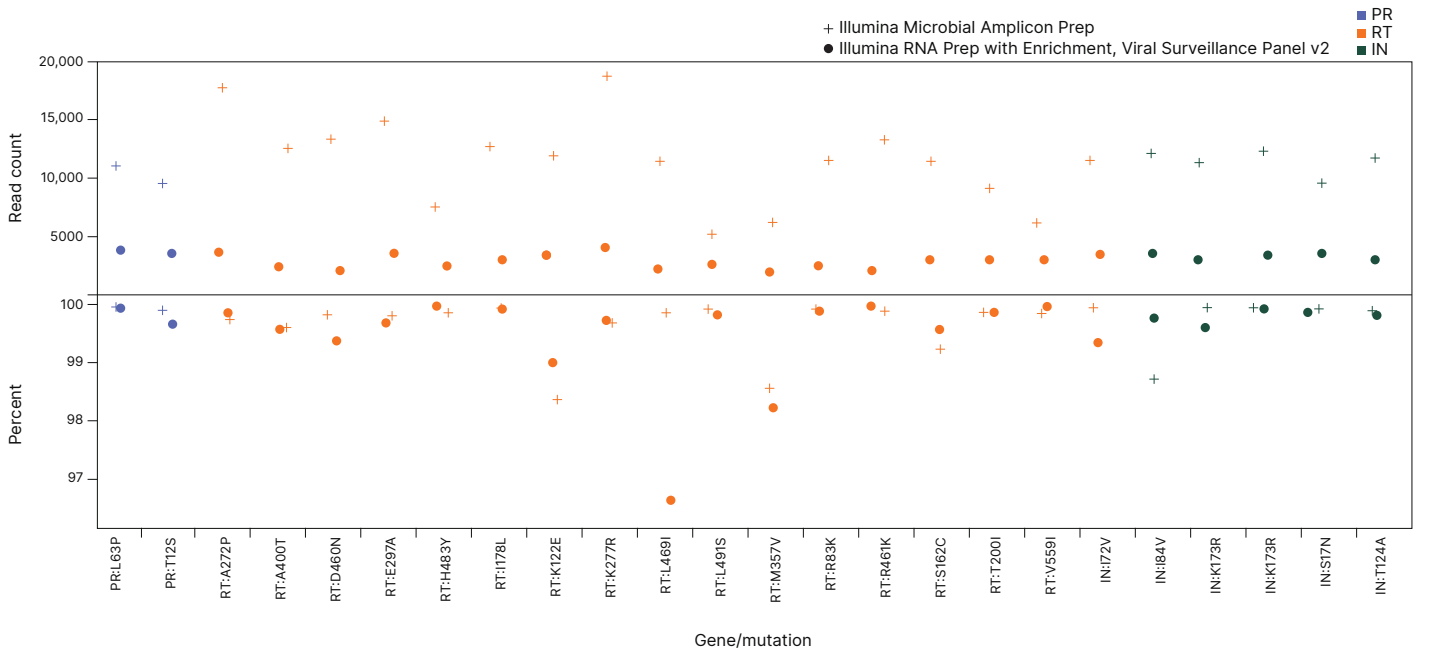


Figure 8: DRM detection with Illumina Microbial Amplicon Prep and Illumina RNA Prep with Enrichment
Mutations detected in the *pol* gene and their associated relative frequency and read count for a contrived sample containing subtype B. Notation on the X-axis refers to Gene: Reference Amino Acid: Mutated Amino Acid.

Learn more →

MiSeq i100 Series

Illumina Microbial Amplicon Prep

Viral Surveillance Panel v2

References

1. Greene WC. A history of AIDS: looking back to see ahead [published correction appears in *Eur J Immunol*. 2008 Jan;38(1):309]. *Eur J Immunol*. 2007;37 Suppl 1:S94-S102. doi:10.1002/eji.200737441
2. Barré-Sinoussi F, Chermann JC, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 1983;220(4599):868-871. doi:10.1126/science.6189183
3. Gallo RC, Salahuddin SZ, Popovic M, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*. 1984;224(4648):500-503. doi:10.1126/science.6200936
4. Popovic M, Sarngadharan MG, Read E, Gallo RC. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*. 1984;224(4648):497-500. doi:10.1126/science.6200935
5. World Health Organization. The Global Health Observatory HIV. who.int/data/gho/data/themes/hiv-aids. Published July 2024. Accessed May 20, 2025.
6. World Health Organization. HIV data and statistics. who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics. Published July 2024. Accessed May 20, 2025.
7. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*. 2011;1(1):a006841. doi:10.1101/cshperspect.a006841
8. Deeks SG, Overbaugh J, Phillips A, Buchbinder S. HIV infection. *Nat Rev Dis Primers*. 2015;1:15035. Published 2015 Oct 1. doi:10.1038/nrdp.2015.35
9. Swinkels HM, Nguyen AD, Gulick PG. HIV and AIDS. [Updated 2024 Jul 27]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-.
10. Yu F, Wen Y, Wang J, et al. The Transmission and Evolution of HIV-1 Quasispecies within One Couple: a Follow-up Study based on Next-Generation Sequencing. *Sci Rep*. 2018;8(1):1404. Published 2018 Jan 23. doi:10.1038/s41598-018-19783-3
11. Ouyang F, Yuan D, Zhai W, Liu S, Zhou Y, Yang H. HIV-1 Drug Resistance Detected by Next-Generation Sequencing among ART-Naïve Individuals: A Systematic Review and Meta-Analysis. *Viruses*. 2024;16(2):239. Published 2024 Feb 2. doi:10.3390/v16020239
12. Jones LR, Moretti F, Calvo AY, et al. Drug resistance mutations in HIV pol sequences from Argentinean patients under antiretroviral treatment: subtype, gender, and age issues. *AIDS Res Hum Retroviruses*. 2012;28(8):949-955. doi:10.1089/AID.2011.0287
13. Hemelaar J, Elangovan R, Yun J, et al. Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis [published correction appears in *Lancet Infect Dis*. 2020 Mar;20(3):e27. doi: 10.1016/S1473-3099(19)30747-9.]. *Lancet Infect Dis*. 2019;19(2):143-155. doi:10.1016/S1473-3099(18)30647-9
14. Chaillon A, Gianella S, Dellicour S, et al. HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J Clin Invest*. 2020;130(4):1699-1712. doi:10.1172/JCI134815
15. Winters MA, Coolley KL, Girard YA, et al. A 6-basepair insert in the reverse transcriptase gene of human immunodeficiency virus type 1 confers resistance to multiple nucleoside inhibitors. *J Clin Invest*. 1998;102(10):1769-1775. doi:10.1172/JCI4948
16. Quick J, Grubaugh ND, Pullan ST, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*. 2017;12(6):1261-1276. doi:10.1038/nprot.2017.066
17. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res*. 2003;31(1):298-303. doi:10.1093/nar/gkg100

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-03571 v1.0