

METHODOLOGY

Open Access



Next generation sequencing based in-house HIV genotyping method: validation report

Alisen Ayitewala^{*} , Isaac Ssewanyana and Charles Kiyaga

Abstract

Background: HIV genotyping has had a significant impact on the care and treatment of HIV/AIDS. At a clinical level, the test guides physicians on the choice of treatment regimens. At the surveillance level, it informs policy on consolidated treatment guidelines and microbial resistance control strategies. Until recently, the conventional test has utilized the Sanger sequencing (SS) method. Unlike Next Generation Sequencing (NGS), SS is limited by low data throughput and the inability of detecting low abundant drug-resistant variants. NGS can improve sensitivity and quantitatively identify low-abundance variants; in addition, it has the potential to improve efficiency as well as lowering costs when samples are batched. Despite the NGS benefits, its utilization in clinical drug resistance profiling is faced with mixed reactions. These are largely based on a lack of a consensus regarding the quality control strategy. Nonetheless, transitional views suggest validating the method against the gold-standard SS. Therefore, we present a validation report of an NGS-based in-house HIV genotyping method against the SS method in Uganda.

Results: Since there were no established proficiency test panels for NGS-based HIV genotyping, 15 clinical plasma samples for routine care were utilized. The use of clinical samples allowed for accuracy and precision studies. The workflow involved four main steps; viral RNA extraction, targeted amplicon generation, amplicon sequencing and data analysis. Accuracy of 98% with an average percentage error of 3% was reported for the NGS based assay against the SS platform demonstrating similar performance. The coefficient of variation (CV) findings for both the inter-run and inter-personnel precision showed no variability ($CV \leq 0\%$) at the relative abundance of $\geq 20\%$. For both inter-run and inter-personnel, a variation that affected the precision was observed at 1% frequency. Overall, for all the frequencies, CV registered a small range of (0–2%).

Conclusion: The NGS-based in-house HIV genotyping method fulfilled the minimum requirements that support its utilization for drug resistance profiling in a clinical setting of a low-income country. For more inclusive quality control studies, well-characterized wet panels need to be established.

Keywords: HIV Genotyping, Next Generation Sequencing, Sanger Sequencing, Uganda

Background

Globally, antiretroviral therapy campaigns have led to a tremendous reduction in morbidity and mortality [1]. Despite this, the risk of virologic failure increases with the emergence and potential transmission of drug-resistant variants which threaten the UNAIDS 95-95-95 goals

to control the HIV epidemic worldwide by 2030 [2–4]. Fortunately, interventions such as HIV genotyping that monitor HIV-1 drug resistance (HIVDR) and surveillance of transmitted drug resistance serve a critical role in the fight against HIV/AIDS [5–7].

HIV Genotyping assays are diverse given the dynamic nature of the technique and driven by a vast range of technologies [8]. Currently, commercial assays such as ViroSeq are available in developed countries. However, because of the cost and varying sensitivity across HIV-1

*Correspondence: ayitewala@gmail.com

National Health Laboratories and Diagnostic Services, Central Public Health Laboratories, Ministry of Health, PO Box 7272, Kampala, Uganda



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

subtypes, these are hardly utilized in resource-limited settings such as Uganda [9, 10]. Instead, most laboratories develop in-house assays that are affordable and designed to genotype HIV-1 subtypes and circulating recombinant forms (CRFs) that are predominant within their localities.

Conventional HIV genotyping using Sanger sequencing techniques serves as the mainstay for clinical HIVDR testing. Although Sanger sequencing has been commonly applied as the “gold standard” for a while, there are some intrinsic limitations with this technology. As opposed to Next Generation Sequencing, the technique has low data throughput and limited capacity to detect variants below 20% intra-host frequency of the quasi-species [11, 12]. Literature suggests that low-abundance HIVDR variants could have a relevant clinical impact and that their detection could benefit treatment management [13–15]. However, a clinically significant threshold is yet to be defined amidst the promising benefits of the new NGS technologies.

Although NGS chemistries differ, all platforms are characterized by high-output, clonal, and parallel sequencing [16]. These outperform conventional Sanger sequencing in scalability, sensitivity, and quantitative detection of minority resistance variants [17]. The sensitive methodology can accurately profile the protease, reverse transcriptase, integrase, and maturation inhibitors, as well as HIV-1 coreceptor tropism in a single run. Despite this, NGS-based assays have been primarily limited to research settings and are rarely used in clinical settings, especially in low- and middle-income countries. This can be attributed to the assay not being standardized, which is necessary for accreditation by regulatory agencies and the lack of an appropriate validation and performance assessment platform for NGS-based HIV genotyping.

Therefore, considering the principles of molecular assay validations and benchmark to the available guiding documents [18], here we present a validation report for the NGS-based assay. The validation report was used to assess the HIVDR assay of the Central Public Health Laboratories (CPHL) which is the national reference laboratory for specialized tests. The report was also used as a benchmark for the NGS assay for Joint Clinical Research Centre (JCRC) which is a WHO-accredited HIVDR laboratory and part of the HIVDR surveillance network.

Results

Accuracy

Accuracy was demonstrated by sequencing 10 samples at frequencies above 20% which is the threshold for clinical interpretation. Accuracy of 98% with an average percentage error of 3% was reported for the NGS assay

against the Sanger sequencing platform demonstrating similar performance, Table 1. Considering the linear range of mutation detection, no significant difference ($r=0.99$, $p=2.5$ Pearson coefficient correlation, Fig. 1) was observed between the NGS platform and the gold standard.

Clinical interpretation of drug resistance mutation profiles is based on frequencies 20% and above. As expected, the linear range of the number of DRMs detection showed a significant reduction in correlation when the data indicated a significant difference ($r=0.7$, $p<0.05$ Pearson coefficient correlation, Fig. 1) between the two platforms <20% frequencies. This is attributed to the sequencing depth capacity of the NGS platform as opposed to Sanger sequencing.

Precision

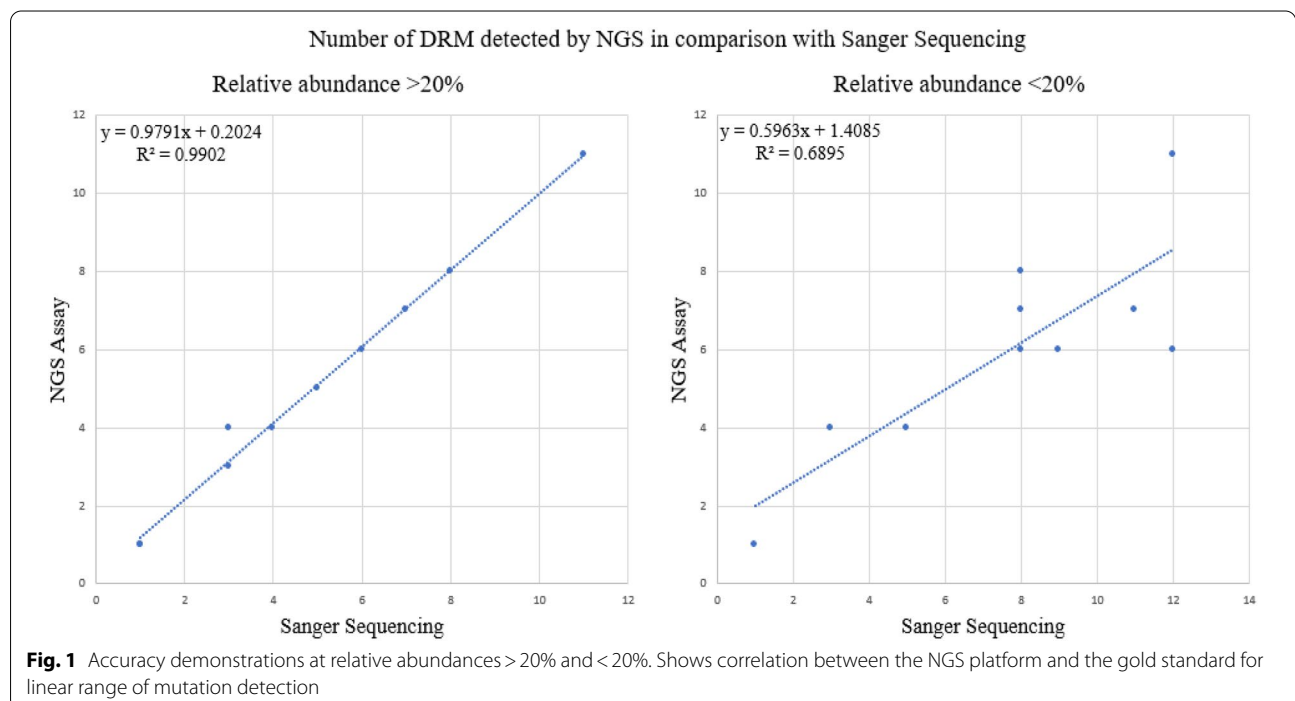
Repeatability (inter-run) was demonstrated by processing and sequencing five samples 5 times using the same conditions in 1 week. Reproducibility (inter-personnel) was demonstrated by processing and sequencing five samples by five laboratory technologists in 1 week. The NGS assay passed the precision demonstrations at frequencies of above 20%, discrepancies occurred at very low frequencies of about 1%. As expected, the NGS assay detected similar mutation profiles at 20% frequency which is the threshold for clinical interpretation of drug resistance mutation profiles, Fig. 2. Despite the difference at a low cut off of 1%, the NGS assay detected comparable similar mutation profiles at frequencies lower than 20%. The difference can be attributed to; inter-personnel and experimental errors i.e., pipetting errors, Fig. 2.

The coefficient of variation (CV) was used to determine the inter-run and inter-personnel precision, Table 2. The findings showed no variability ($CV \leq 0\%$) at the relative abundance of 20% for both the inter-run and inter-personnel precision. For both inter-run and inter-personnel, the variation that affected the precision was observed at 1% frequency. Overall, for all the frequencies, CV registered a small range of (0%-2%).

Precision demonstrations were further interrogated for the most prevalent mutations that are, M184V and K103N. Figure 3 shows the frequencies of M184V across the precision trials for the five samples. M184V is one of the major and most common mutations among the NRTI class of drugs. M184V causes high-level in vitro and in vivo resistance to lamivudine (3TC) and emtricitabine (FTC) and low-level resistance to Didanosine and Abacavir (ABC). As expected, frequencies >20% were relatively uniform across all the runs for every sample. However, at frequencies <2%, M184V was not visible for some runs. Since the frequencies are too low, the difference can be attributed to; inter-personnel

Table 1 Drug resistance mutation profiles for accuracy of NGS against Sanger Sequencing

| Sample ID | Sequencing platform | RT region | | PR region | |
|-----------|---------------------|-----------------------------------|--|------------------------|------------|
| | | NRTI mutations | NNRTI mutations | Major mutations | Accessory |
| DR-968-19 | Sanger Sequencing | K70R, M184V, K219Q | K103N, H221Y, M230L, L234I | None | None |
| | NGS | K70R, M184V, K219Q, D67G | K103N, H221Y, M230L, L234I | None | None |
| DR-969-19 | Sanger Sequencing | None | K103KNRS | None | None |
| | NGS | None | K103KNRS | None | None |
| DR-970-19 | Sanger Sequencing | M41L, D67N, M184V, T215F | A98G, G190A | N88S | V32A, L33F |
| | NGS | M41L, D67N, M184V, T215FIS | A98G, G190A | N88S | V32A, L33F |
| DR-971-19 | Sanger Sequencing | M41L, K65R, M184V, K219N | L100I, K103N | None | None |
| | NGS | M41L, K65R, M184V, K219N, T215FIS | L100I, K103N, G190A, A98G | N88S | V32A, L33F |
| DR-972-19 | Sanger Sequencing | D67N | K103N, K238T, Y188N | None | None |
| | NGS | D67GNS | K103N, K238T | None | None |
| DR-975-19 | Sanger Sequencing | M41L, M184V, L210W, T215F | A98G | M46I, I50V, I54V, V82A | L10F, L33F |
| | NGS | M41L, M184V, L210W, T215F | A98G, V108I | M46I, I50V, I54V, V82A | L10F, L33F |
| DR-976-19 | Sanger Sequencing | M184V | K101E, G190A, E138A | None | None |
| | NGS | K70R, M184V, K219Q | K101E, G190A, K103N, H221Y, M230L, L234I | None | None |
| DR-979-19 | Sanger Sequencing | T69DN, K70R, M184V | G190S | None | None |
| | NGS | T69DN, K70R, M184V | G190S | None | G48R |
| DR-980-19 | Sanger Sequencing | M184V, T215Y | K101E, E138A, G190A | None | None |
| | NGS | M184V, T215NSY | K101E, E138A, G190A, H221Y, P225H | M46I, L76V, I84V | Q58E |
| DR-981-19 | Sanger Sequencing | E44D | None | None | None |
| | NGS | E44D, K65E, M184IV | A98G, K103N, P225H | I54V, V82A | None |



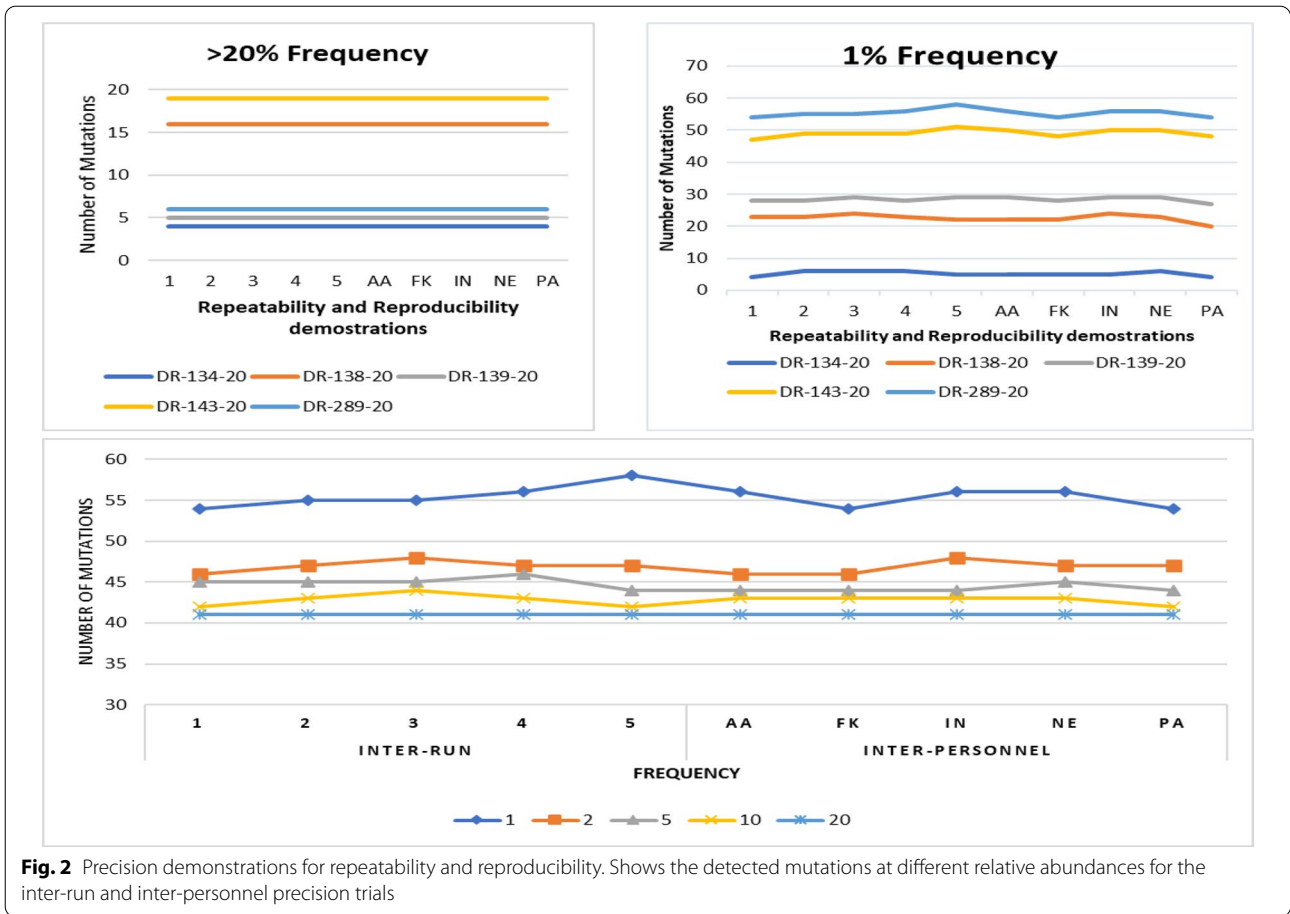


Fig. 2 Precision demonstrations for repeatability and reproducibility. Shows the detected mutations at different relative abundances for the inter-run and inter-personnel precision trials

Table 2 Coefficient of Variation for inter-run and inter-personnel precision

| | Coefficient of Variation, CV % | | | | |
|-----------------|--------------------------------------|----|----|----|----|
| | DRM Relative abundance (% Frequency) | | | | |
| | 1 | 2 | 5 | 10 | 20 |
| Precision | | | | | |
| Inter-Run | 2% | 1% | 1% | 2% | 0% |
| Inter-Personnel | 2% | 2% | 1% | 1% | 0% |

and experimental errors i.e., pipetting errors. Figure 3 also shows the frequencies of K103N across the precision runs for four samples. K103N is one of the major and most common mutations among the NNRTI class of drugs. K103N is a non-polymorphic mutation that causes high-level reductions in Nevirapine and Efavirenz susceptibility. As expected, the frequencies > 20% were relatively uniform across all the runs for every sample. The slight variations are negligible and could be attributed to experimental errors. The fifth sample had no K103N mutation in all the demonstrations.

L74I was the only mutation observed for the Integrase Strand Transfer Inhibitors (INSTI). Figure 4 shows the frequencies across the reproducibility and repeatability runs. L74I is a polymorphic accessory mutation commonly selected by each of the INSTI. Although a slight difference of ~1% was registered, the assay performed as expected. The difference can be attributed to inter-run and inter-personnel bias.

Overall, the average coverage (sequencing depth) per nucleotide position for precision demonstrations was 2795 as also illustrated in Fig. 5.

Discussion

HIV genotyping is increasingly picking pace in Uganda despite resource limitations. This highly specialized molecular test has been a preserve for research purposes for decades which is attributed to operational costs. Mobilization of resources for public health comes at a time when HIV genotyping is advancing technologically. The conventional technology that is currently utilized in the country is based on the Sanger sequencing method. The method is faced with limitations such as low data output and the inability to detect mutations of the viral

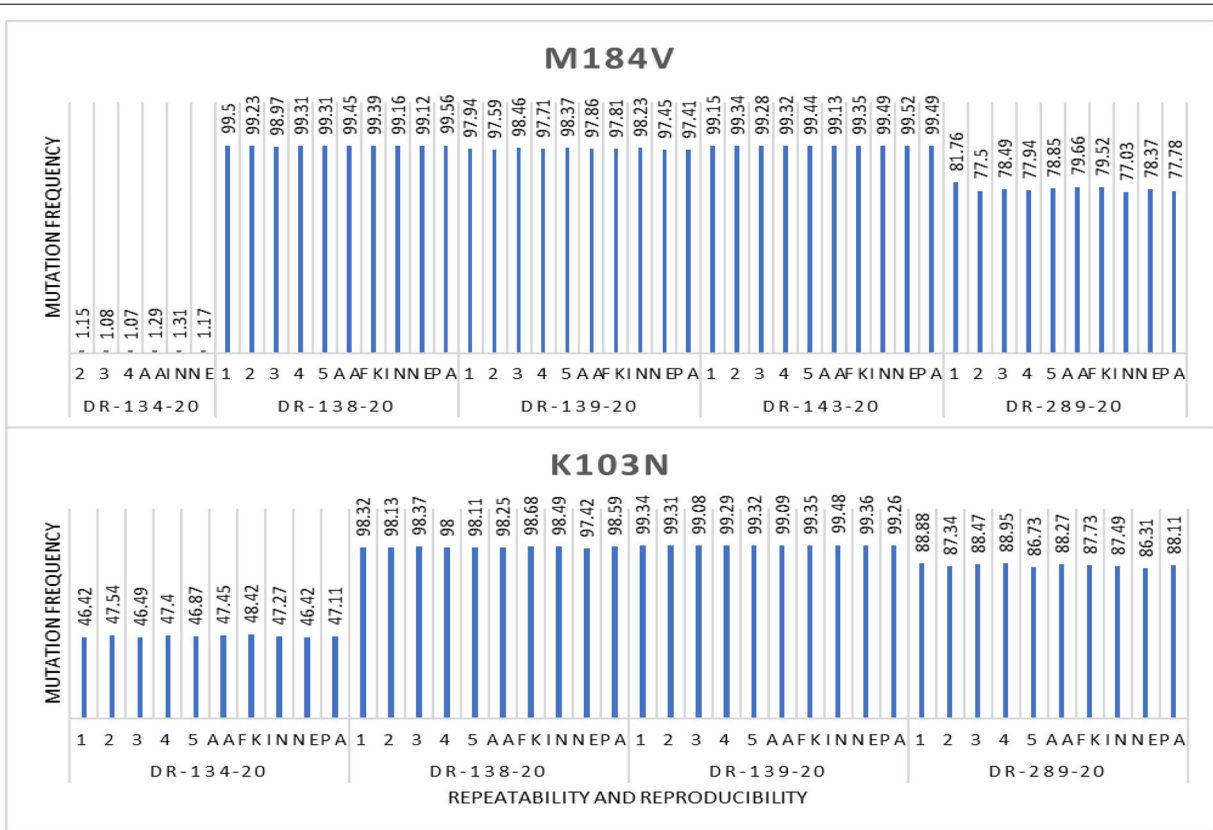


Fig. 3 Precision demonstrations for M184V and K103N mutations. Shows the mutation frequencies of common mutations M184V and K103N across the inter-run and inter-personnel demonstrations

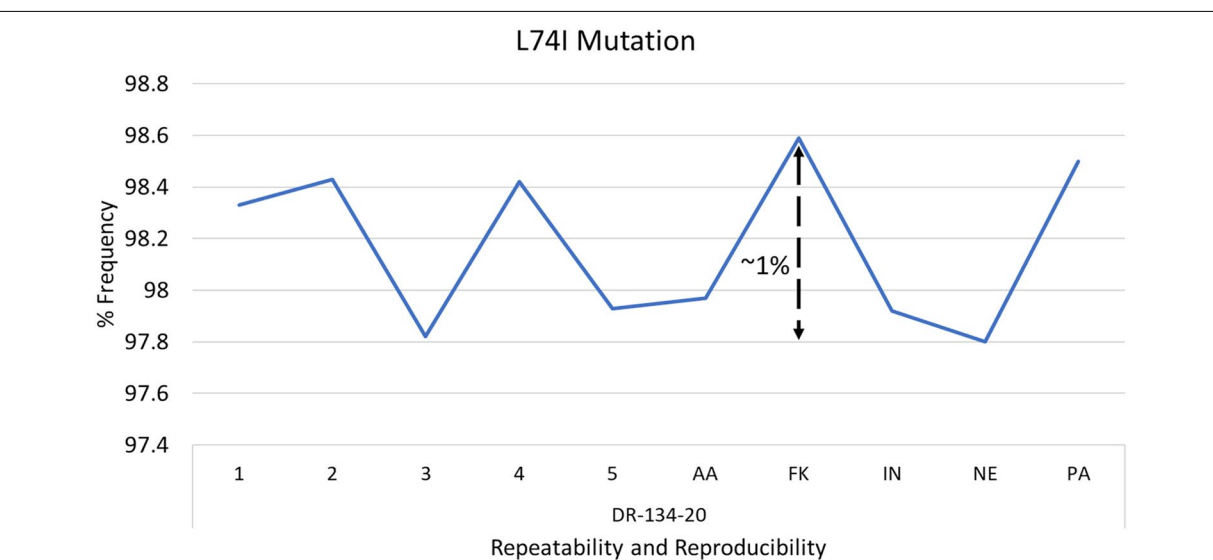
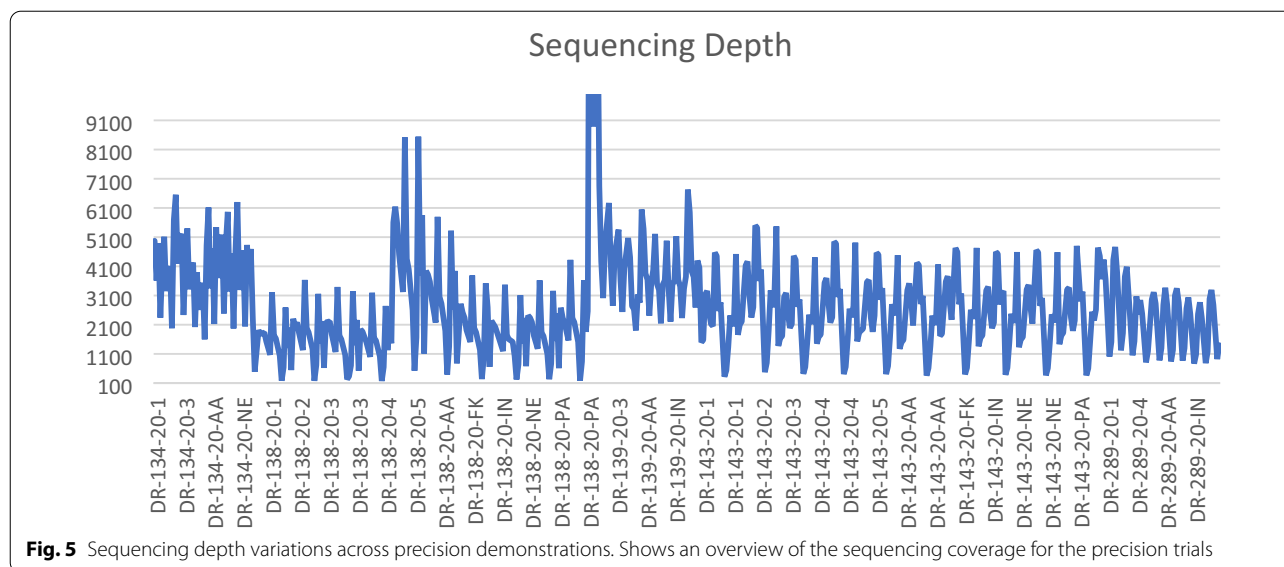


Fig. 4 Precision demonstration for L74I mutation. Shows the percentage difference of frequencies across precision trials for the integrase inhibitor mutation L74I



minority populations. Despite this, SS still serves as a gold standard primarily because the NGS technologies lack a standardized validation strategy and a consensus is yet to be reached on the relevance of the rich data output.

Nevertheless, CPHL acquired an NGS platform for use in clinical diagnostics. Therefore, in this report, we present a validation report of the in-house assay based on NGS technologies. The assay was validated on two criteria; accuracy and precision. Accuracy of 98% was reported for the NGS assay against the SS assay. The calculations were set at 20% as the mutation frequency threshold given that SS assays can arguably detect mutations in approximately 15% of the viral population variants. Qualitatively all mutations reported on SS were also reported on the NGS platform, Table 1. This was expected as NGS is documented to have a higher sequencing depth as compared to SS. The assay can massively generate data without compromising on quality. A significant difference was observed when the threshold was set below 20%, the findings actualize the literature that the NGS platforms can detect mutations occurring at low frequencies as opposed to the SS platforms [17]. Although the NGS platform performed as expected against SS, this approach oversimplifies the complexity and richness of NGS HIVDR data that can report on a highly diverse population of HIV. Therefore, well-validated External Quality Assurance (EQA) strategies that check the NGS platform remain to be established.

The coefficient of variation was used to study precision. The report documented no variation ($CV \leq 0\%$) at the relative abundance of 20%. However, slight variations that affected precision were noted at relative abundances below 20%, but more prominent at 1%. This could

be attributed to the sequence error rate of $\sim 1\%$ and the increased likelihood of cross-contamination and sampling/PCR biases [19]. Although, minority mutations at frequencies $< 0.5\%$ could result in treatment failure [20], the clinical relevance of variants below 5% is still a question of debate [21]. However, if the variants below 5% are proven relevant, the validation process and quality assurance strategies could have to leverage the use of unique molecular identifiers (UMIs) to determine the accuracy of NGS assays at such lower frequencies [22].

We acknowledge the limitation of using clinical samples as opposed to well-characterized wet panels. Because of this, Sensitivity, Specificity and the Lower limit of detection (LOD) were not studied. Sensitivity determines the probability of the assay to detect mutations when present. Specificity determines the probability of the assay not detecting mutations when absent. LOD determines the lowest viral load at which the assay can still be able to detect all the mutations present. The true picture of the above-mentioned parameters could not be achieved with clinical samples consisting of viral RNA of variable quality, quantity and variant diversity. However, even for accuracy and precision, a possibility for bias could have been introduced due to the error-prone reverse transcription PCR step and sequencing error [23]

Conclusion

The NGS assay presents opportunities to revolutionize the field of HIV Genotyping. However, these are under-exploited partly because, their relevance has not been well studied in the context of clinical management. This has left such a powerful tool to be oversimplified when compared with inferior assays such as SS. In this report,

we validated the NGS assay against the SS assay. The NGS assay proved fit for HIVDR testing in a clinical setting of a low resource country. However, with the advance in research, a well elaborate quality control strategy needs to be established.

Methods

In the absence of established proficiency test panels for NGS-based HIVDR testing, 15 clinical plasma samples for routine care were utilized as shown in Table 3. These were collected for routine HIV-1 drug resistance testing as the standard of care for HIV-infected individuals failing treatment. The selected samples and size represented the predominantly circulating variants of HIV-1 in Uganda, this was sufficient for accuracy and precision demonstrations. However, the size was also pre-determined by the availability of resources. Briefly in Fig. 6, Viral RNA was extracted using Qiagen QIAmp Viral RNA kit [24].

RT-PCR was done using a Superscript III One-Step system with Platinum Taq [25]. Amplification of the target genes in the pol region was done using two (2) pairs of external and nested primers; External primers (Pair 1 covering the protease gene and reverse transcriptase spanning between positions 2058 and 3529; (forward, GAAAGACTGCACTGAAAGACAGGC), (reverse, GCT ATTAAGTCTTTTGATGGGTCAT), Pair 2 covering the integrase gene spanning between positions 3694 and 5455; (forward, TATGGGGAAAGACTCCTAAATTTA), (reverse, GTCCTGCTTGATATTCACACC)). Nested primers (Pair 1 covering the protease gene and reverse transcriptase spanning between positions 2155 and 3323;

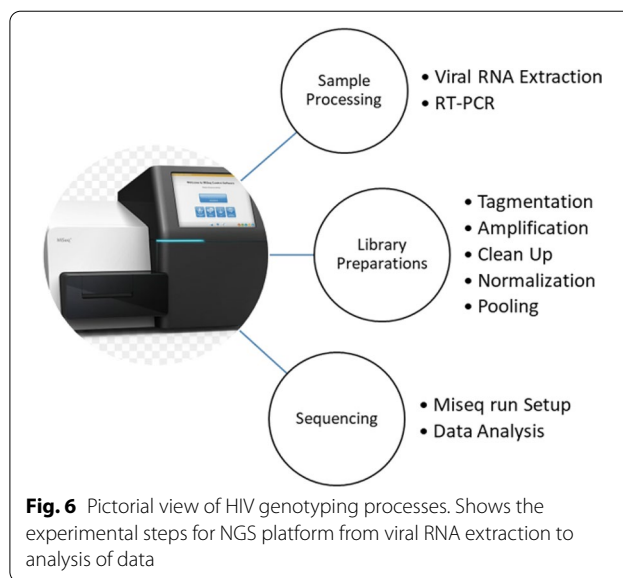


Fig. 6 Pictorial view of HIV genotyping processes. Shows the experimental steps for NGS platform from viral RNA extraction to analysis of data

(forward, ACAGCCCCACCAGCAGAG), (reverse, CTG TATATCATTGACAGTCCAGCT), Pair 2 covering the integrase gene spanning between positions 4022 and 5258; (forward, AGAAGTAAACATAGTAACAGACTC ACA) (reverse, TGCAGACCCCAATATGTTCTA)) [26]. The primer set allows for the robustness of the assay with the ability to genotype all circulating HIV-1 subtypes in the country, Table 3.

Amplicons generated using the aforementioned nested primers were used to prepare NGS libraries using the Nextera XT DNA library preparation kit [27]. The amplicons for the protease and reverse transcriptase region were generated separately from the integrase region, however, these were combined before library preparation at proportional concentrations for sequencing in the same pooled run. Each pooled library had 25 samples; these were run using the Miseq V2 reagent kit on the Miseq system [28] at CPHL. Sanger sequencing was done using BigDye Terminator v3.1 cycle sequencing kit [29] on the ABI genetic analyzer 3730xl [30] at JCRC. To achieve double coverage, two sets were used for forward and reverse primers, in addition to the aforementioned nested primers, the following were also added; for protease and reverse transcriptase region: forward (CTG TACCAGTAAAATTAAGCCAGG), reverse (TCT TCTGTCAATGGCCATTGTTTA), for integrase region: reverse (TGCAGACCCCAATATGTTCTA). Base-calling parameters were set to end at PCR stop, the quality threshold was set to assign Ns to bases with quality values (QVs) less than 15, peak heights were also considered to reduce background noise.

Table 3 Experimental sample Characteristics

| Sample ID | HIV-1 subtypes | Demonstrations | Sequenced Genes |
|-----------|----------------|----------------|-----------------|
| DR-968-19 | D | Accuracy | RT, PR |
| DR-969-19 | D | Accuracy | RT, PR |
| DR-970-19 | A | Accuracy | RT, PR |
| DR-971-19 | D | Accuracy | RT, PR |
| DR-972-19 | A | Accuracy | RT, PR |
| DR-975-19 | A | Accuracy | RT, PR |
| DR-976-19 | B | Accuracy | RT, PR |
| DR-979-19 | B | Accuracy | RT, PR |
| DR-980-19 | D | Accuracy | RT, PR |
| DR-981-19 | C | Accuracy | RT, PR |
| DR-134-20 | A | Precision | RT, PR, INT |
| DR-138-20 | C | Precision | RT, PR, INT |
| DR-139-20 | D | Precision | RT, PR, INT |
| DR-143-20 | D | Precision | RT, PR, INT |
| DR-289-20 | A | Precision | RT, PR, INT |

Data analysis

Data generated was in two formats; FASTA and FASTQ depending on the sequencing platform. FASTA formats are created from chromatograms generated by the Sanger sequencing platforms. Raw chromatograms were analyzed using a web-based software RECall as recommended by World Health Organization (WHO) for laboratories in the HIVDR network [31]. The RECall is an automated base-calling software [32]. This enabled us to do quality checks on; raw chromatogram quality, single-stranded coverage, sequence length, stop codon, excessive mixtures, bad insertions, ambiguous nucleotide/amino acids, APOBEC mutations, atypical mutations and genetic distance. Among the outputs of the RECall analysis included; a consensus sequence, an excel file with the mutation list, a CSV file with the susceptibility scores and pdf files including an image of a neighbour-joining tree.

FASTQ formats were generated from the NGS platforms. These are text-based formats for nucleotide sequences and their corresponding quality scores. Analysis of the FASTQ files was done using an online-based pipeline, Hydra [33]. Among the outputs from Hydra included; a consensus sequence and an Amino Acid Variant Format (AAVF) file. The AAVF file report provides a compact summary of the amino acid variation obtained by conceptual translation of the NGS read pileup across the examined region of the HIV genome [34]. The AAVF files for the samples were then uploaded to the Stanford University HIV Drug Resistance Database [35] for drug resistance profiling.

Criteria for assessment of the HIVDR NGS based home-brewed assay

Accuracy is defined as the closeness of the measurements to the gold standard. Accuracy was demonstrated by sequencing 10 samples on both platforms. For each sample, DRMs detected by SS were compared with those detected by NGS and their number was plotted on scatter graphs. Accuracy was calculated in percentages by comparing the number of DRMs detected by NGS with the standard SS. An average percentage for the 10 samples was established as the accuracy. Precision is defined as how close the agreement of the outcomes is between repeated measurements. This was categorized into Repeatability -the variation arising when all efforts are made to keep conditions constant by using the same instrument and operator and repeating the measurements during a short period. Reproducibility -the variation arising using the same measurement process among different instruments or operators over longer periods. Repeatability was demonstrated by processing and sequencing five samples 5 times using the same

conditions in 1 week. Reproducibility was demonstrated by processing and sequencing five samples by five laboratory personnel in 1 week designated by the initials AA, FK, IN, NE, PA.

Abbreviations

APOBEC: Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like gene; AAVF: Amino Acid Variant Format; NGS: Next Generation Sequencing; SS: Sanger Sequencing; UNHLS: Uganda National Health Laboratory Services; JCRC: Joint Clinical Research Center; PCR: Polymerase Chain Reaction; NNRTI: Non-nucleoside reverse transcriptase Inhibitors; EQA: External Quality Assessment; INSTI: Integrase Strand Transfer Inhibitors; DRM: Drug Resistance Mutations; HIVDR: HIV drug resistance.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12981-021-00390-8>.

Additional file 1. Table showing the detailed precision report from the Miseq system.

Acknowledgements

We are thankful to the Drug Resistance Team of JCRC and CPHL for the support they gave in the laboratory.

Authors' contributions

This work was carried out in collaboration between all authors. Alisen Aytewala (AA) designed the protocols and drafted the methodology. Charles Kiyaga (CK) and Isaac Ssewanyana (IS) performed the literature search and data review. AA wrote the first draft of the manuscript and managed manuscript revisions. All authors read and approved the final manuscript.

Funding

All expenses incurred during the validation process were covered by Uganda National Health Laboratory Services.

Availability of data and materials

All data generated or analyzed during this study are included in this published article [and its Additional file 1].

Declarations

Ethics approval and consent to participate

This study data was obtained following permission to develop an HIV genotyping protocol. The need for approval was waived by the research ethics committee of Uganda National Health Laboratory Services, Kampala Uganda.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 1 April 2021 Accepted: 17 September 2021

Published online: 02 October 2021

References

1. AR Zolopa 2010 The evolution of HIV treatment guidelines: current state-of-the-art of ART *Antiviral Res* 85 241 244
2. SG Deeks 2003 Treatment of antiretroviral-drug-resistant HIV-1 infection *Lancet* 362 2002 2011
3. JW Tang D Pillay 2004 Transmission of HIV-1 drug resistance *J Clin Virol* 30 1 10

4. UNAIDS. (2021). Global AIDS Strategy 2021–2026 —End Inequalities. End AIDS. UNIADS. Retrieved from: https://www.unaids.org/sites/default/files/media_asset/global-AIDS-strategy-2021-2026_en.pdf
5. P Rojas Sánchez A Holguín 2014 Drug resistance in the HIV-1-infected paediatric population worldwide: a systematic review *J Antimicrob Chemother* 69 2032 2042
6. VA Johnson V Calvez HF Günthard R Paredes D Pillay RW Shafer AM Wensing DD Richman 2014 2014 Update of the drug resistance mutations in HIV-1 *Top Antivir Med* 22 642 650
7. D Ssemwanga RW Lihana C Ugoji AL Abimiku J Nkengasong P Dakum N Ndembu 2015 Update on HIV-1 acquired and transmitted drug resistance in Africa *AIDS Rev* 17 3 20
8. S Ávila-Ríos N Parkin R Swanstrom R Paredes R Shafer H Ji R Kantor 2020 Next-generation sequencing for HIV drug resistance testing: laboratory, clinical, and implementation considerations *Viruses* 12 617
9. LL Jagodzinski JD Cooley M Weber NL Michael 2003 Performance characteristics of human immunodeficiency virus type 1 (HIV-1) genotyping systems in sequence-based analysis of subtypes other than HIV-1 subtype B *J Clin Microbiol* 41 998 1003
10. S Saravanan M Vidya P Balakrishnan N Kumarasamy SS Solomon S Solomon R Kantor D Katzenstein B Ramratnam KH Mayer 2009 Evaluation of two human immunodeficiency virus-1 genotyping systems: ViroSeq 2.0 and an in-house method *J Virol Methods* 159 211 216
11. N Stella-Ascariz JR Arribas R Paredes JZ Li 2017 The role of HIV-1 drug-resistant minority variants in treatment failure *J Infect Dis* 216 S847 S850
12. M Casadellà R Paredes 2017 Deep sequencing for HIV-1 clinical management *Virus Res* 239 69 81
13. RG Fisher DM Smith B Murrell R Slabbert BM Kirby C Edson MF Cotton RH Haubrich SL Pond GU Zyl Van 2015 Next generation sequencing improves detection of drug resistance mutations in infants after PMTCT failure *J Clin Virol* 62 48 53
14. SC Inzaule RL Hamers M Noguera-Julian M Casadellà M Parera C Kityo K Steegen D Nanche B Clotet TF Wit de R Paredes PanAfrican Studies to Evaluate Resistance 2018 Clinically relevant thresholds for ultrasensitive HIV drug resistance testing: a multi-country nested case-control study *Lancet* 5 e638 e646
15. A Derache CC Iwuji K Baisley S Danaviah AG Marcelin V Calvez T Oliveira de F Dabis K Porter D Pillay 2019 Impact of Next-generation sequencing defined human immunodeficiency virus pretreatment drug resistance on virological outcomes in the ANRS 12249 treatment-as-prevention trial *Clin Infect Dis* 69 207 214
16. PL Tzou P Ariyaratne V Varghese C Lee E Rakhmanaliev C Villy M Yee K Tan G Michel BA Pinsky RW Shafer 2018 Comparison of an in vitro diagnostic next-generation sequencing assay with sanger sequencing for hiv-1 genotypic resistance testing *J Clin Microbiol* 56 e00105 e118
17. RM Gibson CL Schmotzer ME Quiñones-Mateu 2014 Next-generation sequencing to help monitor patients infected with hiv: ready for clinical use? *Curr Infect Dis Rep* 16 401
18. RJ Hutchins KL Phan A Saboor JD Miller A Muehlenbachs CDC NGS Quality Workgroup 2019 Practical guidance to implementing quality management systems in public health laboratories performing next-generation sequencing: personnel, equipment, and process management (Phase 1) *J Clin Microbiol* 57 e00261 e319
19. MG Becker D Liang B Cooper Y Le T Taylor ER Lee S Wu P Sandstrom H Ji 2020 Development and application of performance assessment criteria for next-generation sequencing-based hiv drug resistance assays *Viruses* 12 627
20. JZ Li DR Kuritzkes 2013 Clinical implications of HIV-1 minority variants *Clin Infect Dis* 56 1667 1674
21. HA Mbunkah S Bertagnolio RL Hamers G Hunt S Inzaule TF Rinke Wit De R Paredes NT Parkin MR Jordan KJ Metzner WHO HIVResNet Working Group 2020 Low-abundance drug-resistant HIV-1 variants in antiretroviral drug-naïve individuals: a systematic review of detection methods, prevalence, and clinical impact *J Infect Dis* 221 10 1584 1597
22. R Kou H Lam H Duan L Ye N Jongkam W Chen S Zhang S Li 2016 Benefits and challenges with applying unique molecular identifiers in next generation sequencing to detect low frequency mutations *PLoS ONE* 11 e0146638
23. JR Kugelman MR Wiley ER Nagle D Reyes BP Pfeffer JH Kuhn M Sanchez-Lockhart GF Palacios 2017 Error baseline rates of five sample preparation methods used to characterize RNA virus populations *PLoS ONE* 12 e0171333
24. Qiagen. QIAamp Viral RNA Mini Handbook [Online]. <https://www.qiagen.com/us/resources/download.aspx?id=c80685c0-4103-49ea-aa72-8989420e3018&lang=en>. Accessed 09 Jan 2021.
25. ThermoFisher. SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase Reference Guide [Online]. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/superscriptIII_onestepRTPCR_man.pdf. Accessed 09 Jan 2021.
26. C Kuiken B Korber RW Shafer 2003 HIV sequence databases *AIDS Rev* 5 1 52 61
27. Illumina. Nextera XT DNA Library Prep Kit Reference Guide [Online]. https://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html. Accessed 09 Jan 2021.
28. Miseq Systems Guide [Online]. https://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-system-guide-for-miseq-reporter-1000000061014-00.pdf. Accessed 09 Jan 2021.
29. ThermoFisher. BigDye™ Terminator v3.1 Cycle Sequencing Kit [Online]. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0015666-BigDyeTerminator_V3_1_CycleSeq_QR.pdf. Accessed 09 Jan 2021.
30. Biosystems, Applied. 3730xl DNA Analyzer User Manual [Online]. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/cms_039278.pdf. Accessed 09 Jan 2021.
31. WHO. WHO HIVResNet HIV Drug Resistance Laboratory Operational Framework [Online]. <https://apps.who.int/iris/rest/bitstreams/1312780/retrieve>. Accessed 09 Jan 2021.
32. CK Woods CJ Brumme TF Liu CK Chui AL Chu B Wynhoven TA Hall C Trevino RW Shafer PR Harrigan 2012 Automating HIV drug resistance genotyping with RECall, a freely accessible sequence analysis tool *J Clin Microbiol* 50 1936 1942
33. Hydra. Hydra Web [Online]. Government of Canada. <https://hydra.canada.ca/pages/home?lang=en-CA>. Accessed 09 Jan 2021.
34. H Ji E Enns CJ Brumme N Parkin M Howison ER Lee R Capina E Marinier S Avila-Rios P Sandstrom G Domselaar Van 2018 Bioinformatic data processing pipelines in support of next-generation sequencing-based HIV drug resistance testing: the Winnipeg Consensus *J Int AIDS Soc* 21 e25193
35. Stanford University. HIV Drug Resistance Database [Online]. <https://hivdb.stanford.edu/hivdb/by-reads/>. Accessed 09 Jan 2021

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.