Next Generation Sequencing Reveals a High Prevalence of HIV-1 Minority Variants and an Expanded Drug Resistance Profile Among Individuals Initiated to Antiretroviral Therapy in a Resource-constrained Setting

Maria Nannyonjo
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Jonah Omooja (Jonah.Omooja@mrcuganda.org)
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Daniel Lule Bugembe
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Nicholas Bbosa
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Sandra Lunkuse
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Stella Esther Nabiye
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Faridah Nassolo
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Hamidah Namagembe
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Andrew Abaasa
Medical Research Council/Uganda Virus Research Institute & London School of Hygiene and Tropical Medicine, Uganda Research Unit

Anne Kazibwe
Research Article

**Keywords:** HIV-1, Minority HIV-1 drug resistance, treatment outcomes, HIV-1 ART, ART-naïve, Uganda, Sanger sequencing and Next-generation sequencing

**Posted Date:** March 29th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-2741155/v1

**License:** ☋ This work is licensed under a Creative Commons Attribution 4.0 International License. 
[Read Full License]
Abstract

Introduction

Because next-generation sequencing (NGS) can detect minority mutations that have been linked to treatment failure but are missed by population Sanger sequencing, it has the potential to enhance HIV treatment monitoring. Comparing NGS based on the Illumina platform to Sanger sequencing, we assessed the clinical importance of NGS in HIV-1 medication resistance testing.

Methods

In this retrospective case-control study, 167 people with matched Sanger sequencing data underwent HIV genotypic analysis using an Illumina-Miseq. These included 122 time-matched controls from the same cohort who had viral suppression at 12 months, and 45 patients with virologic failure at 12 months.

Results

NGS identified all major HIV drug resistance mutations detected by Sanger sequencing and revealed additional major mutations M184V and K65R that increased the resistance profile to antiretroviral therapy used in this cohort. Abacavir's HIV drug resistance score increased 60-fold, Zidovudine's by 25-fold and to Emtricitabine/Lamivudine by 90-fold. Overall, 108/167 (64.7%) of our subjects had minority DRMs at baseline. K70E, M184V, Y115F, and K70R were among the NRTI minority SDRMs discovered, whereas K103N, Y181C, and K101E were NNRTI minority SDRMs. PI minority SDRMs were also detected in 4 individuals. Being female ($p = 0.005$) and having a CD4 $< 250$ cells/mL ($p = 0.029$) were associated with minority mutations. Minority surveillance DRMs expanded the HIV drug resistance profiles of individuals. A higher frequency of baseline minority mutations correlated with a higher viral load count at end point ($p < 0.005$).

Conclusions

NGS identified pre-treatment minority variants linked to increased viral load count and enhanced resistance to NNRTIs and NRTIs, and it detected more major mutations than Sanger sequencing. Being female and having low CD4 count were associated with presence of minority mutations. NGS could be used to create drug resistance profiles for people receiving HIV-1 ART, allowing clinicians to use both major and minor mutation profiles to inform treatment choices and so increase the effectiveness of the currently available antiretroviral medication. This may be crucial if the UNAIDS 95-95-95 targets are to be met and if we are to eliminate HIV/AIDS as a public health issue by 2030.

Introduction
HIV, the causative agent of AIDS still remains a global public health concern with the UNAIDS estimating that 38.4 million individuals across the globe were living with HIV in 2021\[1\] The WHO has a commitment to end this public health threat by 2030 through its 95-95-95 goals that aim to have 95% of individuals living with HIV to know their sero-status; 95% of those tested having access to HIV treatment and of which 95% of them achieving virologic suppression by 2030 [2]. Achieving these ambitious targets requires concerted efforts in HIV/AIDS management strategies to ensure that treatment is not only accessible, but its outcomes are sufficiently monitored.

HIV antiretroviral therapy (ART) has remarkably improved HIV/AIDS treatment outcomes at both individual and population level by reducing HIV-associated morbidity, mortality, and transmission, especially when virological suppression is attained. These crucial outcomes have inspired a massive global roll-out of ART and initiatives like test and treat. However, despite the gains registered so far from the rapid scale up of ART, the emergence and spread of HIV drug resistance [3] associated with virological failure (VF), are hampering the efforts in controlling the HIV epidemic. For this reason, drug resistance testing has become a standard procedure for the clinical management of the HIV/AIDS patients. Population sequencing using the Sanger method is used worldwide to test for drug resistance variants (Metzker, 2010). However, the Sanger method cannot detect minority drug resistant variants with a frequency below 20% of the viral quasispecies [4–6]. Selection of minority variants due to ART drug pressure can make them the predominant variants [7].

Interest in minority HIV-1 drug resistant variants is driven by the development of more sensitive and precise assays that can detect and quantify minority variants in large genetically complex populations of intra-host viruses. This is currently possible by use of next generation sequencing (NGS) which are available for the detection of minority drug resistant variants [8]. The NGS platforms for deep sequencing, metagenomics and whole-genome sequencing have become invaluable for identifying and characterising viral pathogens, studying viral variations, enabling their accurate classification, and identification of viral genetic markers that correlate with virulence. Therefore, there is a suitable avenue for surveillance, prevention and control, as well as design of therapies for viral infections [9–12].

NGS platforms have been widely used to assess HIV viral diversity and minority drug resistance mutations. The NGS assays have a high sensitivity and offer a platform for high-throughput sequencing, enabling NGS to detect HIV minority variants that constitute as low as 0.05–20% of the HIV viral population [13–15]. Minority drug-resistance mutations (DRMs) are of clinical relevance since they can cause treatment failure in individuals initiated on HIV antiretroviral drugs [16–19]. Minority variants can reduce the efficacy of HIV drugs, for example, etravirine [20] and other NNRTI-based first-line regimens [21].

Most HIV drug resistance studies done in Uganda have looked at major drug resistance mutations [22–24] with reliance on Sanger sequencing for HIV-drug-resistance (HIVDR) genotyping. Although NGS promises to revolutionise HIV genotyping, its relevance in the clinical context has not been widely investigated. In Uganda, a few studies have used NGS in HIVDR genotyping: using a deep sequencing,
Kyeyune et al [25] analysed individuals who were failing on ART that had no detectable DRMs by Sanger sequencing, and reported the existence of minority mutations that were associated with virologic failure and drug resistance. Recently, Ayitewala et al [26] reported that the NGS-based in-house assay could be utilized for clinical HIVDR. In the current study, we assessed the performance of the Illumina Miseq next generation sequencing and compared the results to those previously obtained from Sanger sequencing[24]. We also determined the prevalence of minority mutations in HIV-1 infected adults initiating ART and at 12 months post-ART as well as assessed their clinical relevance in the same Ugandan cohort. This study analysed samples from a previous study in which we used a WHO HIV drug resistance survey protocol to assess acquired HIVDR using Sanger sequencing, among individuals initiated on ART at three treatment centres from 2012 to 2013 [24]. In that survey, 20.9% of individuals had major HIVDR, with the most prevalent mutations being M184V, Y181C, K65R as well as thymidine analogue mutations. Baseline viral load (VL) > 100,000 copies and CD4 count < 250 cells/µl were independent predictors of HIVDR [24].

Materials And Methods

Study design

This was a case-control study of HIV-1 ART-naïve individuals that were initiated on ART and followed for 12 months after ART initiation. The study was nested in a parent observational cohort study conducted by the Ministry of Health, Uganda to profile pre-treatment HIV-1 drug resistance in three treatment centres of Masaka, Nsambya and Mbale. A detailed description of the study site, study population and sample collection in the parent study were published earlier on [24]. In this study, we used this cohort’s samples to assess HIV-1 minority variants. In the final analysis, 45 cases and 122 controls were analysed, altogether 167 participants. In terms of samples, 45 samples at baseline and 45 samples at 12 months (time of virological failure) were retrieved for the cases, while 122 samples at baseline were retrieved for the controls. The study therefore analysed 212 samples. The dependent variable was the output of NGS in terms of the prevalence of minority drug resistance mutations and additional major mutations undetected by Sanger sequencing, and the clinical relevance of these mutations. A minority drug resistance mutation refers to any nucleotide change in the HIV-1 genome that changes the amino acid sequence of the wild-type virus, and this alteration is present in < 20% of the viral population. This study analysed the following factors to determine how they were associated with minority variants: Age, sex, viral load, drug regimen, marital status, CD4 counts and the HIV/AIDS WHO stage.

Laboratory methods

RNA extraction, amplification, and detection of the pol HIV gene

A total of 212 samples (45 samples before ART initiation, 45 samples after ART initiation and 122 controls) were extracted and amplified as earlier described [24]. The sanger sequencing assay was
Library preparation and NGS for the detection and quantification of minority resistant variants

We performed primer ID illumina/MiSeq sequencing based on methods earlier published [28]. The amplified PCR product of HIV-1 Pol gene (1.3kb) from the protease (PR) and reverse transcriptase (RT) regions was cleaned using the Qiagen purification kit (Qiagen, German) according to the manufacturer's instructions. We used 10 µl of the cleaned PCR amplicons for quantitation using the Qubit fluorometer (Invitrogen Thermoscientific) and the qubit ds DNA HS assay kit according to the manufacturer's instructions. We prepared Sequencing libraries using the Nextera XT DNA library preparation kit (Illumina, SanDieogo, CA, USA) according to the manufacturer's protocol but after diluting the PCR products to 0.2ng/µl. Library preparation included fragmentation based on transposon technology, then a PCR step incorporating dual indexes to the fragments and simultaneously tags the DNA with adaptor sequences. To ensure equal library representation during sequencing, library normalization was done using the Nextera library normalization kit to obtain a 10-12pM library with inserts that are 500–1000 base pairs. We diluted and pooled libraries prior to sequencing in a MiSeq Illumina platform. We used denatured Phix control (20pM) from the phix kit that was spiked at 20% in the pooled amplicons as a control.

In total three miseq runs were done. All the raw Miseq data obtained in FAST Q format was processed using HyDRA web (https://hydra.canada.ca), a free pipeline for NGS-based HIVDR data analysis. HIVDR mutations detected above a 1% frequency were reported based on the default HyDRA Web Mutation Database, which is a combination of the Stanford 2015 list of HIV-1 drug resistance mutations (http://hivdb.stanford.edu), with added annotations from the WHO 2009 list of mutations for surveillance of transmitted HIV drug resistance. HIVDR mutations were reported for reverse-transcriptase and protease using the Stanford classification designations (Taylor et al., 2019).

Data collection, processing, and statistical analysis

Laboratory data for this study was collected between 2018 and December 2020. All Data were managed in Excel sheets, cleaned, and transferred to STATA version 15 (Stata Corp LP, Texas, USA) for statistical analysis. Medians (interquartile ranges) were used for description of continuous variables. For categorical variables, proportions, frequencies, and percentages were used. Independent variables were compared between cases and controls using Chi-Squares, and the Kruskal-Wallis test was used for comparison between groups, with only variables with p < 0.05 reported as statistically significant. Because of the skewed distribution, viral load was transformed on log base 10 scale. The Hydra generated reads were reported as estimated frequencies for both minority and major mutations.

The Illumina MiseQ platform generated FASTQ texts from nucleotide sequences along with matching quality scores. We analysed the FASTQ files using HyDra [29], an online-based pipeline. The Hydra outputs included consensus sequences alongside an amino acid variant format (AAVF) file. The AAVF file had a summary of the amino acid variation translated from the NGS read pileup across the analysed
region of the HIV genome [30]. The AAVF files were uploaded to the Stanford University HIVDR Database for HIVDR profiling.

**Ethical considerations**

Ethical approval was obtained from the Uganda Virus Research Institute (UVRI) Research and Ethics Committee and the Uganda National Council for Science and Technology (Ref number: GC/127/15/12/203). All study subjects consented to the use of their samples for genetic studies and research.

**Results**

**Study Cohort Demographics, Sampling and Clinical Characteristics**

We analysed 167 HIV-1 infected adults with a median age of 32.5 years (IQR: 26.5–39.5) who were initiated on ART. After 12 months on ART, 45 (16.9%) individuals had failed to achieve virologic suppression, defined here as VL ≤ 1000 copies/mL.

The 45 individuals with VF provided two samples each, one at baseline and one at the time of VF (12 months post-ART initiation), yielding 90 samples from cases that were used for NGS. Of the 122 individuals who achieved virologic suppression after 12 months, we successfully retrieved and performed NGS and Hydra analysis of 122 respective baseline samples (Fig. 1). The sequencing outputs were obtained in form of FASTA files and CSV files.

**Comparison of Sanger Sequencing and NGS at baseline**

Of the 45 individuals with VF, 6 (13.3%) had major SDRMs based on Sanger sequencing. Next generation sequencing detected all the major SDRMs in the same individuals (100%) as earlier detected by Sanger sequencing at baseline. However, NGS detected additional SDRMs to both NNRTIs and NRTIs (Table 1). In participant NSA1100042, Sanger sequencing did not identify any NRTI SDRMs while high throughput NGS identified M184V and K65R, moreover as major mutations with frequencies above 20% of the viral population.
Table 1
Comparison of SDRMs for cases by Sanger and NGS outputs

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Mutations to:</th>
<th>Sanger</th>
<th>NGS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NRTIs</td>
<td>M184V; E44D</td>
<td>M184V; E44D</td>
</tr>
<tr>
<td>MSK3300054</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NNRTIs</td>
<td>Y181C</td>
<td>Y181C; K103N; G190A</td>
</tr>
<tr>
<td>NSA1100042</td>
<td>NRTIs</td>
<td>None</td>
<td>M184V; K65R</td>
</tr>
<tr>
<td></td>
<td>NNRTIs</td>
<td>K103N</td>
<td>K103N;</td>
</tr>
<tr>
<td>MSK3300124</td>
<td>NRTI</td>
<td>T215S</td>
<td>T215S</td>
</tr>
</tbody>
</table>

Also, among controls, NGS was able to detect additional 24 major SDRMs to NRTI mutations and 24 major SDRMs to NNRTIs missed out by Sanger sequencing. Those SDRMs were distributed among 15/122 (12.3%) of the controls (Fig. 2). NGS provided greater sensitivity due to deep sequencing.

**Prevalence of minority drug resistance among participants at baseline**

Of the 167 (45 controls and 122 cases) participants with NGS results available, pre-treatment minority DRM s were present in 108 (64.7%) of them. Specifically, we detected surveillance pre-treatment minority SDRMs in 43/167 (25.7%). For the 45 cases before ART-initiation, minority DRMs with a frequency range of 1.09–14.12%, were present in 29/45 (64.4%) of our participants (cases). Minority SDRMs existed in 11/45 (24.4%) of these cases. There were 11 individuals with minority SDRMs to NRTIs and 9 individuals with SDRMs to NNRTIs. The most frequent NRTI minority SDRM was which was present in 2 individuals. Other NRTIs detected include M184V, D67E, T215C, among others. Also, K103N and G190A, each of which were detected in two individuals, were the most common NNRTI minority SDRMs. Four individuals had PI minority SDRMs, such as V82A and M46I, among others (Fig. 3).

At baseline, minority mutations with a frequency range of 1.01–19.9% were detected, of which minority SDRMs were present in 32/122 = 26.2%) of controls. Controls had minority SDRMs that cut across all the three drug classes, i.e., NNRTIs, NRTIs and PIs. We observed that 17/122 (13.9%) of the controls had minority SDRMs to NNRTIs, 13/122 (10.7%) had minority SDRMs to NRTIs and 8/122 (6.6%) had minority SDRMs to PIs (Fig. 4).

**Prevalence of minority drug resistance variants among participants experiencing virological failure 12 months after ART initiation**
After 12 months on ART, 33/45 (73.3%) of the cases had minority DRMs, however, only 13/45 (28.9%) individuals harboured minority SDRMs to NNRTIs while 12/45 (26.7%) had minority SDRMs to NRTIs. The profiles of the mutation patterns of the cases at 12 months are summarised in Fig. 5.

**Factors associated with the presence of minority drug resistance**

We analysed the sequences generated and linked the variants to the demographic and clinical characteristics of the individuals they were obtained from. Females had more minority variants at baseline compared to males (51.8% vs 31.1%; $p = 0.005$) and a similar trend was observed after 12 months on ART, although there was no statistical significance at 12 months (Table 2).

Minority drug resistance mutations were significantly associated with CD4 count. Individuals with a CD4 count less than 250 cells/mL were more likely to harbour minority HIV-1 variants compared to individuals whose CD4 count was more than 250 cells/mL ($p = 0.029$). High mean viral load was associated with presence of minority variants before ART initiation ($p = 0.021$).
Table 2
Factors associated with presence of minority variants among individuals

<table>
<thead>
<tr>
<th>Variable</th>
<th>Before ART initiation</th>
<th>12 months after ART initiation</th>
<th>p-value</th>
<th>Before ART initiation</th>
<th>12 months after ART initiation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minority</td>
<td>Not minority</td>
<td></td>
<td>Minority</td>
<td>Not minority</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>79 (43.4)</td>
<td>103 (56.6)</td>
<td></td>
<td>28 (24.4)</td>
<td>87 (75.7)</td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>0.028</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>52 (50.5)</td>
<td>51 (49.5)</td>
<td></td>
<td>na</td>
<td>Na</td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>27 (34.2)</td>
<td>52 (65.8)</td>
<td></td>
<td>28 (24.4)</td>
<td>87 (75.7)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.005</td>
<td>0.128</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>56 (51.8)</td>
<td>52 (48.2)</td>
<td></td>
<td>20 (29.4)</td>
<td>48 (70.6)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>23 (31.1)</td>
<td>51 (68.9)</td>
<td></td>
<td>8 (17.0)</td>
<td>39 (83.0)</td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td>0.104</td>
<td>0.070</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>20 (44.4)</td>
<td>25 (55.6)</td>
<td></td>
<td>Hi6 (50.0)</td>
<td>6 (50.0)</td>
<td></td>
</tr>
<tr>
<td>25–34</td>
<td>21 (33.3)</td>
<td>42 (66.7)</td>
<td></td>
<td>9 (17.7)</td>
<td>42 (82.3)</td>
<td></td>
</tr>
<tr>
<td>35+</td>
<td>38 (51.3)</td>
<td>36 (48.7)</td>
<td></td>
<td>13 (25.0)</td>
<td>39 (75.0)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>0.319</td>
<td>0.145</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary/None</td>
<td>45 (46.9)</td>
<td>51 (53.1)</td>
<td></td>
<td>16 (30.8)</td>
<td>36 (69.2)</td>
<td></td>
</tr>
<tr>
<td>Secondary+</td>
<td>34 (39.5)</td>
<td>52 (60.5)</td>
<td></td>
<td>12 (19.1)</td>
<td>51 (80.9)</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td>0.053</td>
<td>0.357</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>49 (50.0)</td>
<td>49 (50.0)</td>
<td></td>
<td>16 (28.1)</td>
<td>41 (71.9)</td>
<td></td>
</tr>
<tr>
<td>Single*</td>
<td>30 (35.7)</td>
<td>54 (64.3)</td>
<td></td>
<td>12 (20.7)</td>
<td>46 (79.3)</td>
<td></td>
</tr>
<tr>
<td>WHO staging</td>
<td>0.457</td>
<td>0.491£</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>12 (35.3)</td>
<td>22 (64.7)</td>
<td></td>
<td>3 (25.0)</td>
<td>9 (75.0)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>46 (43.8)</td>
<td>59 (56.2)</td>
<td></td>
<td>20 (29.0)</td>
<td>49 (71.0)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>12 (42.9)</td>
<td>16 (57.1)</td>
<td></td>
<td>2 (16.7)</td>
<td>10 (83.3)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>9 (60.0)</td>
<td>6 (40.0)</td>
<td></td>
<td>3 (13.6)</td>
<td>19 (86.4)</td>
<td></td>
</tr>
<tr>
<td>CD4</td>
<td>0.128</td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 250</td>
<td>50 (39.7)</td>
<td>76 (60.3)</td>
<td></td>
<td>22 (22.7)</td>
<td>75 (77.3)</td>
<td></td>
</tr>
<tr>
<td>250+</td>
<td>29 (51.8)</td>
<td>27 (48.2)</td>
<td></td>
<td>6 (33.3)</td>
<td>12 (66.7)</td>
<td></td>
</tr>
</tbody>
</table>
### Variable

<table>
<thead>
<tr>
<th>First line ART regimen</th>
<th>Before ART initiation</th>
<th>12 months after ART initiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZT/3TC/EFV</td>
<td>-</td>
<td>1 (33.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (66.7)</td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>-</td>
<td>3 (27.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 (72.7)</td>
</tr>
<tr>
<td>TDF/3TC/EFV</td>
<td>-</td>
<td>3 (14.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18 (85.7)</td>
</tr>
<tr>
<td>TDF/3TC/NVP</td>
<td>-</td>
<td>8 (27.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21 (72.4)</td>
</tr>
<tr>
<td>TDF/FTC/EFV</td>
<td>-</td>
<td>1 (33.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (66.7)</td>
</tr>
<tr>
<td>TDF/FTC/NVP</td>
<td>-</td>
<td>12 (25.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36 (75.0)</td>
</tr>
</tbody>
</table>

| Log10 viral load       | Mean (SD)             | 0.021                         |
|------------------------|                       | 5.5                            |
|                        |                       | 5.5                            |
|                        |                       | 0.922                         |

*Single; never married or ever married (separated/divorced), £-Fisher’s exact p, SD-Standard deviation, na-not applicable

**Clinical relevance of Next generation sequencing in HIV treatment monitoring**

To evaluate the clinical relevance of next generation sequencing, we assessed the output of NGS in terms of minority mutations and additional major mutations that had not been detected by Sanger sequencing. We examined if the mutations detected by only NGS had any clinical implications on treatment outcomes of individuals with them. The presence of minority SDRMs detected by NGS increased the resistance profiles of the individuals with them to both NRTI and NNRTI drugs in use and those that were not being used in this cohort (see the bolded genotypic sensitivity score (GSS) records on Tables 3 and 4). For instance, individual NSA1100042 had increased resistance to all the NRTI drugs when NGS detected M184V and K65R minority NRTI mutations. HIV drug resistance score to ABC increased 60-fold, resistance to AZT by 25-fold and to FTC/3TC by 90-fold (Table 3).
Table 3  
Comparison of effect of NRTI mutations based on Sanger sequencing and NGS at baseline

<table>
<thead>
<tr>
<th>Id</th>
<th>Mutations by Sanger</th>
<th>Mutations by NGS</th>
<th>GSS of drugs by Sanger</th>
<th>GSS of drugs by NGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSK3300054</td>
<td>E44D; M184V</td>
<td>E44D; M184V</td>
<td>ABC (15); AZT (1); D4T (-10); FTC (60); 3TC (60); TDF (-10)</td>
<td>ABC (15); AZT (1); D4T (-10); FTC (60); 3TC (60); TDF (-10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSK3300124</td>
<td>T215S</td>
<td>T215S</td>
<td>ABC (15); AZT (20); D4T (20); FTC (0); 3TC (0); TDF (5)</td>
<td>ABC (15); AZT (20); D4T (20); FTC (0); 3TC (0); TDF (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSA1100042</td>
<td>None</td>
<td>M184V; K65R</td>
<td>ABC (60); AZT (-25); D4T (50); FTC (90); 3TC (90); TDF (50)</td>
<td>ABC (60); AZT (-25); D4T (50); FTC (90); 3TC (90); TDF (50)</td>
</tr>
</tbody>
</table>

Similarly, the detection of minority NNRTI mutations K103N and G190A in the sequences of individual MSK3300054 increased resistance to DOR from 10 to 40, to EFV from 30 to 135, to ETR from 30 to 50, NVP from 60 to 180 and RPV from 45 to 70 (Table 4).
### Table 4
Comparison of effect of NNRTI mutations on drug resistance genotypic sensitivity scores at baseline

<table>
<thead>
<tr>
<th>Id</th>
<th>Mutations by sanger</th>
<th>GSS of drugs</th>
<th>Mutations by NGS</th>
<th>GSS of drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSK3300054</td>
<td>Y181C</td>
<td>DOR (10); EFV (30); ETR (30); NVP (60); RPV (45)</td>
<td>Y181C; K103N; G190A</td>
<td>DOR (40); EFV (135); ETR (50); NVP (180); RPV (70)</td>
</tr>
<tr>
<td>MSK3300124</td>
<td>None</td>
<td>DOR (0); EFV (0); ETR (0); NVP (0); RPV (0)</td>
<td>None</td>
<td>DOR (0); EFV (0); ETR (0); NVP (0); RPV (0)</td>
</tr>
<tr>
<td>NSA1100042</td>
<td>K103N</td>
<td>DOR (0); EFV (60); ETR (0); NVP (60); RPV (0)</td>
<td>K103N</td>
<td>DOR (0); EFV (60); ETR (0); NVP (60); RPV (0)</td>
</tr>
</tbody>
</table>

In this study, NGS baseline minority mutation results were crucial in predicting treatment outcomes of individuals after 12 months on ART. The presence of minority mutations at baseline was associated with higher viral loads after 12 months. For all samples (both controls and cases), the higher the frequency of minority mutations at baseline, the higher the viral load count at 12 months (Fig. 6A). One would probably hypothesise that this could have been due to variation in viral loads between the cases and controls at baseline, however, our analysis showed no significant difference between the count of viral load of cases and controls at baseline (Kruskal-Wallis test; p = 0.221; Fig. 6B).

Even when cases at baseline were grouped into those with major and those with minority mutations, we noted that the higher the frequency of minority mutations, the higher the endpoint viral load count (Fig. 7).

**Discussion**
There is still a dearth of evidence with regards to the real clinical relevance of next generation sequencing for HIV treatment monitoring in resource-constrained settings. This study assessed the clinical relevance of next generation sequencing in monitoring HIV treatment outcomes. We used NGS to determine the prevalence of minority mutations and major mutations missed out by Sanger sequencing and their clinical relevance. Our NGS assay detected pre-treatment minority variants among 64.7% of our study participants. This is similar prevalence to that obtained by Clutter et al. [31] where they detected minority DRMs in 60% of the study participants. The prevalence reported in this study is however lower than the 80% prevalence obtained in a Malawian study [19]. The current study provides sufficient evidence that NGS can detect minority mutations that are often missed out by the most widely used population Sanger sequencing. This finding concurs with the reports of several studies that have credited NGS with the ability to detect low-frequency HIV drug resistance variants [19, 25, 31–34].

HIV-1 minority variants (also referred to as low-frequency mutations) though undetectable to the conventional population Sanger sequencing, are said to be clinically crucial as they have been associated with virological failure in individuals initiated on ART [25, 35]. Some studies have documented that specific HIV-1 variants are clinically significant at a level as low as 1% of the viral population. These minority variants can replicate quickly and become the major viral population due to selective pressure of ART drugs, resulting in treatment failure [34]. We detected the most common NRTI minority mutations as M184V/I, Y115F and D67E, while the most prevalent NNRTI minority mutations included K103N, Y181C and G190A. These findings mirror those that were reported in Malawi among individuals on an ART regimen composed of NRTIs and one NNRTI [19]. In the current study, NGS sequencing detected PI mutations that had not been detected by Sanger sequencing. The common PIs detected by NGS included M46I/L, D30N, I47V AND V82A. Zhou et al [19] reported the detection of PIs which Sanger sequencing did not reveal. Since the PIs were not yet being used in this population at the time of sample collection, the PI mutations detected by NGS here could be natural polymorphs of HIV-1. It is however concerning as natural polymorphisms of M46L have been shown to have a replicative advantage for HIV-1 subtype B [36]. As Uganda has included the use of PIs for second line regimens, the presence of PI minority mutations should be closely monitored to prevent exacerbation of drug resistance to this class of salvage therapy.

While Sanger sequencing did not detect any variants among controls at baseline, the Illumina Miseq NGS platform intriguingly detected major surveillance DRMs among 15/122 (12.3%) of the controls. This suggests that conventional population Sanger sequencing possibly misses some major HIV-1 SDRMs. This observation has serious ramifications on the HIV treatment at individual and population level. It means that even if Sanger sequencing had been used at baseline to inform treatment regimen for those individuals, 15 of them would have been started on already failing regimens. Kyeyune et al [25] previously observed that some adhering individuals continued to fail on ART and yet Sanger sequencing revealed no HIV-1 drug resistance mutations. This suggests that Sanger sequencing could have possibly missed out on some HIV mutations (probably minority mutations with clinical relevance), therefore, points to the need to use more robust deep sequencing platforms to enhance HIV treatment monitoring and guide treatment decisions. In addition, the NGS platform also revealed minority DRMS among 79/122 (64.8%)
of the controls in this study. These mutations spanned across all the three drug classes of NRTIs, NNRTIs and PIs. The minority DRMs observed here are probably a mixture of transmitted drug resistance mutations that faded to frequencies undetectable by Sanger sequencing and de novo mutations arising from poor incorporation and high error rate of HIV-1 transcriptase enzyme [37]. The NGS analysis of samples of controls that were collected at baseline revealed the presence of minority variants. Without drug pressure, resistant virus populations (variants) are selected against as they are outmatched by wild type variants which have more efficient replication ability [38]. This results in resistant variants having low frequencies (below 20%) which cannot be detected by population sanger sequencing, but are detected by NGS [38, 39].

Not only did NGS show all the mutations earlier identified by Sanger sequencing, but it also identified additional mutations that have a substantial impact on treatment outcomes. For participant NSA1100042, NGS was able to expose NRTI mutations M184V and K65R that were not detected by population Sanger sequencing. The presence of those mutations notably increased resistance to all the NRTI drugs. On a similar note, for participant MSK3300054, Sanger genotyping had only identified the NNRTI mutation Y181C. The NGS platform identified additional mutations K103N and G190A that increased the drug resistance scores to the existing NNRTI drugs. Studies have already reported that there is concordance between NGS and Sanger sequencing in terms of identifying mutations identified by Sanger sequencing [4–6, 40]. The findings of this study agree with those of another study in which NGS identified additional mutations (to those identified by Sanger sequencing) that increased resistance to existing NRTI and NNRTI drugs [19]. Another survey that employed NGS in South America concluded that reducing the variant detection threshold to 5% enhanced the identification of virologic failure among HIV-infected individuals [41]. This further affirms to the benefits of NGS, that does deeper and wider sequencing due to the parallel sequencing mechanism of the assay.

In this study, controls were individuals who had viral load of < 1000 copies/mL at 12 months post-ART initiation. In the current Ugandan guidelines on treatment of HIV/AIDS, these would be regarded as virologically suppressing and with a good treatment outcome [42]. These individuals would not be recommended for genotypic testing for HIV drug resistance. However, with NGS analysis, major NNRTI, NRTI and PI drug resistance mutations were detected in 11, 13 and 1 individual (s) respectively. Since VL test results are used to guide on who should be subjected to genotypic resistance testing, the VL ≥1000 copies/mL used to define virologic failure may need to be revised downwards. In developed settings, the VL threshold for virologic failure is more stringent, for example in Europe, a threshold VL of > 50–200 copies/mL defines virologic failure [43]. From our findings, we note that some of the individuals assumed to be virologically suppressed had major DRMs. Some researchers have already observed treatment failure in form of drug resistance among virologically suppressing individuals with a VL < 1000 copies/mL [25, 44], some individuals failing on treatment may be left out of the necessary GRT on the basis of the reasoning that those with a VL < 1000 copies/mL are less likely to have HIV DRMs.

While the actual clinical relevance of minority DRMs remains contested [38–40], other studies [19, 21, 25, 31, 35] have reported association between minority variants with virologic failure. This concurs with our
observation that the higher the frequency of baseline minority mutations the more the viral load count at end point (Kruskal Wallis p < 0.005) when we compared cases and controls.

In the current study, we observed that the presence of minority DRMS increased the genotypic sensitivity score for some NRTI and NNRTI drugs used in this cohort. Similarly, Zhou et al [19] analysed minority HIV resistance in a Malawian cohort and observed that minority mutations increased the resistance levels of HIV to, not only some of the NRTIs and NNRTIs used, but also to future possible salvage regimen of the same of NNRTIs and NRTIs. The implication of this is that minority HIV-1 drug resistance mutations could possibly hamper HIV treatment efforts by reducing the efficacy of the possible future regimen. That could in turn prevent the realisation of the UNAIDS goal to eliminate the public threat posed by HIV by 2030, which requires the achievement of the 95-95-95 targets.

Our study is among a few studies in Uganda that have evaluated the utilization of NGS for HIV drug resistance genotyping, the other having been done by Kyeyune et al.[25] and a recent validation study by Ayitewala et al [26]. This report adds to the limited data on minority HIV-1 drug resistance in our region and provides insights into the potential relevance of NGS in HIV drug resistance testing. We had a sufficient sample size that included both cases and controls which augments our findings. Noteworthy, is also the sequencing data that was available at both baseline and time of virologic failure, which enable comparison of the NGS platform output with that of Sanger sequencing.

However, the findings of this study should be interpreted with the consideration that the samples were collected in 2014 when individuals were on some of the HIV-1 ART drugs that are now not being used. The Ministry of Health has updated its treatment guidelines and some NNRTI drugs like NVP have been discontinued [42]. Although, this study did not evaluate the cost of using the Illumina NGS platform for HIVDR genotyping, we do report that the initial costs of acquiring equipment and reagents are high. However, this could be offset in the long run when large numbers of samples are tested.

**Conclusion**

This study highlights potential clinical benefits of next generation sequencing. NGS detected minority variants that lowered the efficacy of the NNRTI and NRTI drugs used in this setting and were associated with high endpoint viral load count. Further, NGS identified additional major SDRMS that expanded the ART resistance profile of this cohort. The use of NGS could be adopted to generate drug resistance profiles of individuals on HIV-1 ART so that clinicians can use both major and minor mutation profiles to guide treatment decisions and thus enhance the efficacy of the existing antiretroviral therapy. This could be cardinal in the pursuit of the 95-95-95 targets and in efforts to end the HIV/AIDS as a public health threat by 2030.

**Declarations**

**Competing interests**
The authors have no conflicts of interest to declare.

**Author contributions**

M.N. & J.O.: Wrote the original manuscript text, and prepared figures and tables

M.N., J.O. & D.L.: Data curation and analysis


P.K. & D.S.: Funding acquisition; Project administration & Supervision

D.S.; M.N.; J.O.: Conceptualization and Investigation

_all authors reviewed the manuscript._

**Acknowledgements**

We acknowledge the Ugandan Ministry of Health for the technical and logistical support.

**Funding**

This work was funded by the UK Medical Research Council (MRC) and the UK Department for International Development (DFID) that is under the MRC/DFID Concordat Agreement and is also part of the EDCTP2 programme supported by the European Union.

**Data availability**

The datasets generated and analysed during the current study are not publicly available due to restriction policies and data protection policies of the MRC/UVRI & LSHTM but can be made available by the corresponding authors on reasonable request and on approval of the UVRI Research Ethics Committee.

**References**


**Figures**
Figure 1

Study overview
Figure 2

Mutation patterns of Controls at baseline showing major SDRMs detected only by NGS
Figure 3

*Minority drug resistance profiles of cases at baseline*

![Graph showing minority drug resistance profiles of cases at baseline.](image)

Figure 4

*Minority drug resistance mutations of controls at baseline*

![Graph showing minority drug resistance mutations of controls at baseline.](image)
Figure 5

Minority drug resistance mutations of cases at 12 months post-ART initiation
Figure 6

6A: Baseline HIV-1 drug resistance mutations plotted against end-point viral load for both controls and cases show that the higher the frequency of baseline minority mutations the higher the viral load count at end point (Kruskal Wallis \( p<0.005 \)).

6B: Baseline HIV-1 viral loads are similar across time points for the cases and controls (Kruskal Wallis, \( p=0.231 \)).
The test statistic is adjusted for ties. Multiple comparisons were not performed since there were only two test fields.

**Figure 7**

Baseline HIV-1 drug resistance mutations plotted against end point viral load for cases shows that the higher the frequency of minority mutations at baseline, the higher the viral load count at end point (Kruskal Wallis $p=0.011$).