Use of a mutation-specific genotyping method to assess for HIV-1 drug resistance in antiretroviral-naïve HIV-1 Subtype C-infected patients in Botswana [version 1; peer review: 1 approved, 1 not approved]

Dorcas Maruapula1,2, Iain J. MacLeod3,4, Sikhulile Moyo1,3, Rosemary Musonda1, Kaelo Seatla1,2, Kesaobaka Molebatsi1,2, Melvin Leteane1,2, Max Essex1,3, Simani Gaseitsiwe1,3, Christopher F. Rowley3,5

1Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana
2University of Botswana, Gaborone, Botswana
3Harvard T.H Chan School of Public Health, Boston, MA, USA
4Aldatu Biosciences, Watertown, MA, USA
5Beth Israel Deaconess Medical Center, Boston, MA, USA

Abstract

Background: HIV-1 drug resistance poses a major threat to the success of antiretroviral therapy. The high costs of available HIV drug resistance assays prohibit their routine usage in resource-limited settings. Pan-degenerate amplification and adaptation (PANDAA), a focused genotyping approach based on quantitative PCR (qPCR), promises a fast and cost-effective way to detect HIV drug resistance mutations (HIVDRMs). Given the high cost of current genotyping methods, we sought to use PANDAA for screening key HIVDRMs in antiretroviral-naïve individuals at codons 103, 106 and 184 of the HIV-1 reverse transcriptase gene. Mutations selected at these positions have been shown to be the most common driver mutations in treatment failure.

Methods: A total of 103 samples from antiretroviral-naïve individuals previously genotyped by Sanger population sequencing were used to assess and verify the performance of PANDAA. PANDAA samples were run on the ABI 7500 Sequence Detection System to genotype the K103N, V106M and M184V HIVDRMs. In addition, the cost per sample and reaction times were compared.

Results: Sanger population sequencing and PANDAA detected K103N mutation in three (2.9%) out of 103 participants. There was no evidence of baseline V106M and M184V mutations observed in our study. To genotype the six HIVDRMs it costs approximately 40 USD using PANDAA, while the reagents cost per test for Sanger population
sequencing is approximately 100 USD per sample. PANDAA was performed quicker compared to Sanger sequencing, 2 hours for PANDAA versus 15 hours for Sanger sequencing.

**Conclusion:** The performance of PANDAA and Sanger population sequencing demonstrated complete concordance. PANDAA could improve patient management by providing quick and relatively cheap access to drug-resistance information.

**Keywords**
HIV-1 drug resistance testing, Assay performance, Pan-degenerate amplification and adaptation
Introduction

HIV remains a major global health problem; currently, 37.9 million adults and children are estimated to be living with HIV with Sub-Saharan Africa being the most severely affected region\(^1\). In Botswana, 380,000 people are estimated to be living with HIV of which 310,713 are on treatment\(^2\). In 2016, Botswana introduced universal HIV treatment to provide ART to all HIV positive individuals regardless of their immune status\(^3\). Combination antiretroviral therapy (cART) has been successful in reducing morbidity and mortality in individuals infected with HIV as well as in prevention of mother-to-child transmission (PMTCT) of HIV\(^4\). Despite the availability of antiretroviral drugs, which inhibits HIV replication and reducing mortality, one public health concern about the widescale rollout of cART is the increase in emergence and transmission of HIV drug resistance\(^5-7\); which has the potential to reduce the efficacy and compromise the success of ART programmes\(^8-9\).

HIV-1 reverse transcriptase mutations introduced into the viral genome contributes to the development of resistance to antiretroviral drugs. Major non-nucleoside reverse transcriptase inhibitor (NNRTI) mutations, such as K103N and V106M, are selected when exposed to nevirapine (NVP), which is still used in low resource settings as part of management. Resistance mutations that develop in patients exposed to the first generation NNRTIs, NVP and efavirenz (EFV) have been shown to confer some cross-resistance to second generation NNRTIs like etravirine and rilpivirine\(^10\). M184V is a major NRTI mutation selected for under tenofovir and lamivudine. Studies have shown that >98% of patients failing first-line NNRTI regimens would have one or more drug resistance mutations, such as K103N, V106M and M184V\(^11\). HIV drug resistance testing is routinely used for clinical care in high-income countries; however, routine HIV drug resistance testing is not available to majority of patients in resource-limited settings due to the high costs of implementation and limited trained manpower. While Sanger sequencing-based methodologies remain the gold standard for mutation detection, the assays are costly and resource-intensive. Thus, it is urgent to use a simple and cheaper detection method for HIV drug resistance. Detecting known specific mutations provides important information that guides patient treatment options. Moreover, utilising point mutation assays reduces the need for several laboratory equipment.

In this study, we compare an HIV genotyping method, pan-degenerate amplification and adaptation (PANDAA), a focused point mutation genotyping assay, with Sanger population based sequencing\(^12\). It is anticipated that PANDAA could serve as an alternative method to rapidly detect HIV-1 drug resistance mutations in HIV patients in Botswana.

Methods

Study population

This was a retrospective study utilizing existing data and stored PCR products from 103 specimens previously genotyped by population sequencing from a previous completed study conducted at Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana: Novel strategy for HIV drug resistance monitoring in developing countries (BHP063 study)\(^13\). Briefly, this study enrolled 234 pregnant women diagnosed with HIV and 188 pre-ART from Infectious Diseases Care Clinics (IDCC) between 2012 and 2015. These participants were enrolled to determine the prevalence of HIV transmission at three different geographical locations in Botswana (Gaborone, Molepolole, Mochudi). In samples collected between 2014 and 2015, the following mutations were detected in the main cohort; K103N, G190A and L90M\(^11\).

For the current study, a convenience sampling method was employed to maximize the number of samples available for analysis and samples collected between 2014 and 2015 were included, provided that the stored sample(s) were still available with sufficient remaining volume for PCR products. At the time of the current study, the first-line ART regimen in Botswana consisted of tenofovir + emtricitabine (or lamivudine) + efavirenz (or NVP).

Ethical considerations

Ethical clearance for the BHP063 study was obtained from the Human Research Development Committee (HRDC) at the Botswana Ministry of Health (Ethics permit number: HPDME 13/18/1 Vol (366)). All participants consented prior to participation in the study.

The current study was approved by the University of Botswana Institute Review Board (IRB) and the Human Research Development Committee at the Botswana Ministry of Health (Ethics permit number: HPDME 13/18/1 Vol (833)) and the need for informed consent was waived since remnant plasma samples were used for this study.

RNA extraction, reverse transcription and PCR amplification

As reported previously\(^13\) EZ1 Advanced XL (Qiagen) automated instrument was used to isolate HIV-1 RNA from 400 µL plasma using the QIAamp viral RNA Mini Kit (Qiagen, Valencia, CA, USA) and eluted in 60 µL of buffer. The HIV protease-reverse transcriptase (PR-RT) region was reverse transcribed to cDNA then amplified by one-step RT-PCR (Roche One Step RNA PCR kit) to generate a 1,569 bp amplicon\(^15\). The primers were used CWF1-LNA2 and CWR1-LNA3 for first round, whereas second-round primers were CWF1-LNA2 and RT20C\(^16\) (Table 1). The amplification conditions were reverse transcription at 50°C for 30 min, initial denaturation at 94°C for 7 min, and then 10 cycles of 94°C for 10 s, 55°C for 30 s and 68°C for 2 min, followed by 35 cycles of 94°C for 10 s, 55.5°C for 30 s and 68°C for 2 min, increasing each cycle by 10 s, with a final elongation step of 5 min at 68°C. The second-round cycling conditions were 98°C for 10 s, then 40 cycles of 98°C for 1 s, 55°C for 5 s and 72°C for 20 s, and then hold at 4°C.

The 1.5 kb generated PCR products were visualized in 1% ethidium bromide stained gels. PCR products were purified using QIAquick Purification (Qiagen, Hilden, Germany), and sequencing clean-up for population sequencing was done using ZR DNA Sequencing Clean up Kit (Zymo, Irvine, CA, USA) according to the manufacture’s protocol\(^13\).
Drug resistance genotyping by population sequencing
Direct population sequencing of the pol gene was previously performed on an ABI 3130xl genetic analyser (Applied Biosystems, Foster City, CA, USA) using BigDye Terminator cycle sequencing kit (Life Technologies, Carlsbad, CA, USA)\(^1\). Sequences were manually edited using Sequencher\(^\circledR\) version 5.0 DNA sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, USA). HIV drug resistance and subtype identification was assessed by entering the generated consensus sequences into the Stanford HIV database\(^14,15\). Sequences were aligned using ClustalW with HXB2 accession number K03455 as a reference from the Los Alamos sequence database. The consensus sequences were then uploaded on the Stanford HIV database.

PANDAA qPCR
The stored pol-derived PCR products were diluted prior to PANDAA focused genotyping.

PANDAA qPCR reactions for detecting drug-resistant point mutations K103N, M184V, V106M were performed on an ABI prism 7500 sequence detection system (Applied Biosystems).

PANDAA is provided as a 10x mix of primers and probes that are specific for each DRM in three triplex qPCR reactions\(^16\). A single target codon is amplified by the PANDAA primers (proprietary properties of Aldatu Biosciences) and the wild-type variants in each patient is detected using a VIC-labelled TaqMan MGB probe, which is differentiated from the resistant variant, which is detected by a FAM-labelled probe (Life Technologies, MA, USA). Components of the PANDAA reaction contained 5 µL buffer (kappa Probe Fast, kappa Biosystems), 1 µL PANDAA probes (VIC labelled wild-type and DRM-specific FAM-labelled) and primer mix (forward and reverse primers), 4 µL template to a final volume of 10 µL. Each sample was performed in triplicate under the following thermal cycling conditions: 98°C for 3 minutes followed by 40 cycles of 95°C for 5 seconds then 60°C for 90 seconds during which fluorescence data were acquired. Each sample was run in triplicate for each DRM. PANDAA primers include locked nucleic acids (LNAs) which increase affinity for their target sequences and contain an adaptor region (ADR) that is matched to the probe-binding site and a pan-degenerate region (PDR) that incorporates degenerate bases in the targeted primer-binding site upstream of the ADR. The principle of PANDAA is shown in Figure 1\(^12\).

The different protocols (K103N, V106M and M184V) were performed separately, each with a corresponding set of standards.

PANDAA data analysis
The threshold was set at 0.02 and using the ABI 7500 software, raw qPCR fluorescence data were exported from Applied Biosystems SDS software to excel and Cq values were corrected for differences in probe-binding efficiencies. All reactions were performed in triplicate, and the mean of the three values was used for calculation.

Cycle quantification (Cq) values were recorded for each sample. Samples were considered positive when the amplification of the mutant was statistically significant with respect to control sample. The percent abundance of the DRM was calculated using E^ΔCq, whereby E is the efficiency of probe-binding, and Cq is the Cq difference between the wild-type and DRM probes, after correcting for variations in probe-binding efficiency.

Reagent cost comparison
The costs of reagents were estimated by according to updated prices. Cost of equipment such as ABI 3130XL and 7500 sequence detection system were not considered as these items of equipment were already available in the laboratory.

**Table 1. Detailed sequences of the primers used for PCR and sequencing**\(^13\).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Target region</th>
</tr>
</thead>
<tbody>
<tr>
<td>CWF1-LNA2</td>
<td>5’+GAA+G+GACACCAAATGAAAGAYTG-3’</td>
<td>2044-2066</td>
</tr>
<tr>
<td>CWR1-LNA3</td>
<td>5’-G+GA+TAC+TTYCCTGTTCAG-3’</td>
<td>3613-3594</td>
</tr>
<tr>
<td>CWF1</td>
<td>5’-GAAGGACACAAATGAAAGAYTG-3’</td>
<td>2044-2066</td>
</tr>
<tr>
<td>CWCS2</td>
<td>5’-AGAACCTCAAGATTTTGGG-3’</td>
<td>2044-2066</td>
</tr>
<tr>
<td>CWCS3</td>
<td>5’-TGCTGGTGCGGTATCC-3’</td>
<td>3145-3129</td>
</tr>
<tr>
<td>CWCS5</td>
<td>5’-TGTTAAAA TTTGATATGTCCAT-3’</td>
<td>3577-3557</td>
</tr>
<tr>
<td>Seq6</td>
<td>5’-CCATCCCTGTGGAAGCACATTA-3’</td>
<td>3008-2987</td>
</tr>
<tr>
<td>Seq2.1-F2</td>
<td>5’-GGCCAGGAATTATTTCTTCAAGGC-3’</td>
<td>2120-2142</td>
</tr>
<tr>
<td>RT20C</td>
<td>5’-CTGCAATCCTAATTCCTCG CTTC-3’</td>
<td>3462-3441</td>
</tr>
</tbody>
</table>

The primers used for Sanger sequencing are those shown in bold.

---

95°C for 5 seconds then 60°C for 90 seconds during which fluorescence data were acquired. Each sample was run in triplicate for each DRM. PANDAA primers include locked nucleic acids (LNAs) which increase affinity for their target sequences and contain an adaptor region (ADR) that is matched to the probe-binding site and a pan-degenerate region (PDR) that incorporates degenerate bases in the targeted primer-binding site upstream of the ADR. The principle of PANDAA is shown in Figure 1\(^12\).

The different protocols (K103N, V106M and M184V) were performed separately, each with a corresponding set of standards.

PANDAA data analysis
The threshold was set at 0.02 and using the ABI 7500 software, raw qPCR fluorescence data were exported from Applied Biosystems SDS software to excel and Cq values were corrected for differences in probe-binding efficiencies. All reactions were performed in triplicate, and the mean of the three values was used for calculation.

Cycle quantification (Cq) values were recorded for each sample. Samples were considered positive when the amplification of the mutant was statistically significant with respect to control sample. The percent abundance of the DRM was calculated using E^ΔCq, whereby E is the efficiency of probe-binding, and Cq is the Cq difference between the wild-type and DRM probes, after correcting for variations in probe-binding efficiency.

Reagent cost comparison
The costs of reagents were estimated by according to updated prices. Cost of equipment such as ABI 3130XL and 7500 sequence detection system were not considered as these items of equipment were already available in the laboratory.

---

95°C for 5 seconds then 60°C for 90 seconds during which fluorescence data were acquired. Each sample was run in triplicate for each DRM. PANDAA primers include locked nucleic acids (LNAs) which increase affinity for their target sequences and contain an adaptor region (ADR) that is matched to the probe-binding site and a pan-degenerate region (PDR) that incorporates degenerate bases in the targeted primer-binding site upstream of the ADR. The principle of PANDAA is shown in Figure 1\(^12\).

The different protocols (K103N, V106M and M184V) were performed separately, each with a corresponding set of standards.

PANDAA data analysis
The threshold was set at 0.02 and using the ABI 7500 software, raw qPCR fluorescence data were exported from Applied Biosystems SDS software to excel and Cq values were corrected for differences in probe-binding efficiencies. All reactions were performed in triplicate, and the mean of the three values was used for calculation.

Cycle quantification (Cq) values were recorded for each sample. Samples were considered positive when the amplification of the mutant was statistically significant with respect to control sample. The percent abundance of the DRM was calculated using E^ΔCq, whereby E is the efficiency of probe-binding, and Cq is the Cq difference between the wild-type and DRM probes, after correcting for variations in probe-binding efficiency.

Reagent cost comparison
The costs of reagents were estimated by according to updated prices. Cost of equipment such as ABI 3130XL and 7500 sequence detection system were not considered as these items of equipment were already available in the laboratory.

---

95°C for 5 seconds then 60°C for 90 seconds during which fluorescence data were acquired. Each sample was run in triplicate for each DRM. PANDAA primers include locked nucleic acids (LNAs) which increase affinity for their target sequences and contain an adaptor region (ADR) that is matched to the probe-binding site and a pan-degenerate region (PDR) that incorporates degenerate bases in the targeted primer-binding site upstream of the ADR. The principle of PANDAA is shown in Figure 1\(^12\).

The different protocols (K103N, V106M and M184V) were performed separately, each with a corresponding set of standards.

PANDAA data analysis
The threshold was set at 0.02 and using the ABI 7500 software, raw qPCR fluorescence data were exported from Applied Biosystems SDS software to excel and Cq values were corrected for differences in probe-binding efficiencies. All reactions were performed in triplicate, and the mean of the three values was used for calculation.

Cycle quantification (Cq) values were recorded for each sample. Samples were considered positive when the amplification of the mutant was statistically significant with respect to control sample. The percent abundance of the DRM was calculated using E^ΔCq, whereby E is the efficiency of probe-binding, and Cq is the Cq difference between the wild-type and DRM probes, after correcting for variations in probe-binding efficiency.

Reagent cost comparison
The costs of reagents were estimated by according to updated prices. Cost of equipment such as ABI 3130XL and 7500 sequence detection system were not considered as these items of equipment were already available in the laboratory.
Reaction time
To establish the total time to perform each method, we considered the total time to perform genotyping method and interpretation of results.

Concordance statistics
Agreements between PANDAA and Sanger population sequencing were calculated using Cohen’s kappa coefficient. The Mann-Whitney U-test was used to test for differences in CD4 counts and viral loads between the groups with drug resistance mutations and those without drug resistance mutations. Two-sided tests were used and a p-value less than 0.05 implied statistically significant differences. All statistical analysis was carried out using R version 3.5.1, other than R², which was calculated using the linear regression function in Microsoft Excel.

Results
Characteristics of participants
All participants were female. The median age was 28 (Q1; Q3: 24; 32) years (Table 2).

Performance of PANDAA
The lower detection limit was determined by analysing serial dilutions. A linear standard curve generated from 10-fold dilution was obtained as shown in Figure 2.

PANDAA showed reproducible results when 1:1 mix of wild-type and DRM templates over a range of copy numbers tested in triplicate. The correlation of each mutant detected by PANDAA correlated with expected mutant as shown in Figure 3.

Quantification of drug resistance of patient samples by PANDAA
PANDAA was completed on patient-derived amplicons of 103 ARV naïve individuals for the K103N, V106M and M184V DRMs using PANDAA. PANDAA identified the presence of K103N in three samples. The three samples with K103N were the same samples that Sanger sequencing detected. Only wild-type sequences at codons 106 and 184 of the RT could be identified by both PANDAA and population sequencing. There was a complete concordance between population sequencing and PANDAA assay as PANDAA qPCR confirmed the presence of HIV drug-resistant mutations as identified by population-based sequencing as shown in Table 3.

Differences in CD4 counts and viral loads between the groups with drug resistance mutations and those without drug resistance mutations are shown in Figure 4A and Figure 4B.

Cost and time analysis of each reaction
The cost of genotyping six drug resistance mutations per patient using PANDAA is 40 USD and Sanger population sequencing
Table 2. Characteristics of participants.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (Q1, Q3) years</td>
<td>28 (24, 32)</td>
</tr>
<tr>
<td>CD4+ T cell count, median (Q1,Q3) (cells/uL)</td>
<td>331 (207.5, 495.5)</td>
</tr>
<tr>
<td>HIV-1 RNA copies, median (Q1, Q3), log_{10} copies/ml (Q1,Q3)</td>
<td>4.1 (3.49, 4.55)</td>
</tr>
</tbody>
</table>

Figure 2. Standard curves generated from ten-fold serial dilutions. Correlation coefficients ($r^2$) were higher than 99.4%.
is estimated at 100 USD per sample. The turnaround time for PANDAA and Sanger sequencing is approximately 2 hours and 24 hours, respectively (Table 4).

**Discussion**

Here, we show that the results of PANDAA are comparable to those produced by Sanger population sequencing. Our study provides baseline data of PANDAA performance and has added an insight that monitoring HIV drug resistance mutations is possible using PANDAA. Having protocols in place for detecting HIV drug resistance mutations using fast and low-cost platforms is important for guiding future diagnostic and patient management, thereby achieving WHO goal of eliminating HIV by 2020.

**Figure 3. Measured mutant correlated with expected mutant.** (A) K103N: R^2=0.99339. (B) M184V: R^2=0.99472.

<table>
<thead>
<tr>
<th>Assay</th>
<th>PANDAA</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>K103N</td>
<td>Sanger</td>
<td>Yes</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>M184V</td>
<td>Sanger</td>
<td>Yes</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>0</td>
<td>103</td>
</tr>
<tr>
<td>V106M</td>
<td>Sanger</td>
<td>Yes</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>0</td>
<td>103</td>
</tr>
</tbody>
</table>
Table 4. Comparison of sequencing cost and time required for PANDAA and Sanger population sequencing.

<table>
<thead>
<tr>
<th>Sequencing Method</th>
<th>Instrument used for sequence detection</th>
<th>Cost estimate per test (USD)</th>
<th>Number of DRMs per test</th>
<th>Processing steps required</th>
<th>Total time required for sequencing only (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PANDAA</td>
<td>ABI 7500 sequence detection system</td>
<td>40</td>
<td>6</td>
<td>✓ ✓ × × ✓</td>
<td>2 hours</td>
</tr>
<tr>
<td>Sanger Population</td>
<td>ABI 3130XL genetic analyser</td>
<td>70–100</td>
<td>All PI and RT</td>
<td>✓ × ✓ ✓ ✓</td>
<td>&gt;15 hours</td>
</tr>
</tbody>
</table>

The table shows instruments used by each sequencing method, cost per test, steps required to process and time required for sequencing. ✓ Means required, × means not required.

When the duration of each method was compared, the results showed that PANDAA required the shortest time for genotyping and had the lowest cost, when compared to Sanger sequencing. It is important to note that PANDAA cost 40 USD for six relevant drug resistance mutations, thus making it cost effective. Sanger sequencing is the widely used and validated method and it is still a relevant platform to use; however, using PANDAA to detect key drug resistance mutations will reduce the cost, especially in this test-and-treat era, thereby enabling quicker results to patients.

Our study had small number of positive samples used to compare the results; however, PANDAA was shown to produce concordant results with sanger sequencing. PANDAA can be
considered to rapidly detect drug resistance mutations at a cheaper cost. In addition, PANDAA kit is more cost-effective, and after preparation genotyping results can be obtained in less than two hours.

Botswana has recently introduced universal HIV therapy; however, additional patients are likely to develop drug resistance and transmit these drug-resistant HIV strains to their uninfected partners. As more patients will be receiving ART in Botswana, there is a need to consider investing in fast, low-cost assays to detect mutations associated with drug resistance.

In Botswana, NVP is still used as prophylaxis for prevention of mother-to-child transmission of HIV (PMTCT). Common drug resistance mutations associated with resistance to NVP include K103N (AAA/G to AAC/T) and V106M.19 The key M184V (ATG to GTG) mutation in HIV-1 RT is associated with high-level resistance to the lamivudine (3TC) and emtricitabine (FTC)19; however, M184V has been shown to rapidly decay in the absence of treatment as a result of its impact on viral fitness30. HIVDR testing is important to clinicians for patient management, however the cost of reagents and equipment maintenance for resistance testing is the biggest obstacles in resource-limited settings.

In this study, we used PANDAA, to screen for NRTI and NNRTI drug-resistant viruses in 103 newly diagnosed HIV-infected pregnant women from the BHP063 cohort and compared the PANDAA results to those obtained by Sanger based population sequencing. Standard curves generated proved PANDAA to accurately differentiated mutants from wild type. In one hundred and three samples included in our study, the use of PANDAA assay enabled detection of K103N in 3 antiretroviral naïve individuals. Both PANDAA and Sanger sequencing did not detect any mutations at codons 106 and 184 in the HIV strains from this cohort. This study provides insights on the performance of PANDAA, a simple method that utilises primers and probes on any available real-time qPCR platform to detect key HIV drug resistance mutations.

The data generated by our study confirm the ability of PANDAA to detect K103N HIV drug resistance mutation as a point mutation assay, and these data correspond to Sanger sequencing data. The results generated from the use of PANDAA provide evidence that this assay represents an alternative strategy for rapid, specific detection of mutations of interest. At the time the samples were collected for this study, 2014-2015, the standard of care for treatment of HIV infection in Botswana was a regimen that included tenofovir, emtricitabine, and efavirenz co-formulated into one pill, Atripla, taken once a day. By using PANDAA, we targeted the most likely mutations to develop to these medications in HIV-1 subtype C, the M184V, K103N and V106M mutations in reverse transcriptase, a targeted and cost-effective approach to genotyping is possible21.

Our study had some limitations. Firstly, we only examined the most common relevant resistant mutations, V106M, K103N and M184V of the reverse transcriptase; therefore, there was a limited number of positive mutations available for analysis. Secondly, the samples used were from ART naïve patients and not exposed to ART leading to under-representation of drug resistant mutations. Another limitation in this study was the lack of samples with M184V and V106M, making it difficult to draw a conclusion on the performance of PANDAA in detecting V106M and M184V. There is a need for further studies utilising cohort with positive known mutations. The applicability of this assay can be demonstrated further by testing a larger number of samples with known mutations. Nevertheless, we have shown that it is possible to genotype HIV drug resistance mutations in HIV naïve individuals using PANDAA.

**Conclusion**

Our findings proved the potential use of PANDAA assay for testing drug resistance mutations in resource-limited settings. This study demonstrates that applying this cost-effective assay to samples from treatment-naïve individuals where background HIV drug resistance may be increasing can provide valuable insight into baseline resistance and allow for decisions to be made to ensure the best prospect of successful HIV treatment. PANDAA holds the same promise for detection of HIV DRM in patients failing ART, although the current study did not include any participants with known ART exposure. Given the simplicity and cost-effectiveness of PANDAA, it can be performed in any laboratory with real-time PCR capability and its principle could be easily adapted to other clinically relevant point mutations. Overall, the comparative results indicate that PANDAA assay provides similar results with Sanger population sequencing at a much lower cost.

**Data availability**

Sequence data generated in this study has been deposited with NCBI GenBank under sequential accession numbers MT908833–MT908846 and MT919428–MT919516.


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

**Acknowledgements**

The authors would like to acknowledge the study participants, principal investigator, and study coordinator from the Novel strategy completed study. We would like to also extend our acknowledgements to University of Botswana and the Botswana Harvard HIV Reference Laboratory for their support and contribution to the success of the study. We thank Aldatu Biosciences for their support throughout the experiment.
Diana B. Dickinson
Independence Surgery, Gaborone, Botswana

This is a very relevant, clearly presented and well done paper, readily reproducible, particularly in the context of a country, Botswana, that has very strong guidelines such that there is a standard universal first line regimen for both public and private facilities. This study shows us that a cost effective and rapid way is possible to demonstrate any pre-existing drug resistant mutations in all patients pre-initiation of treatment. We would have liked to have been doing that for years, but the cost of universal genotyping using Sanger sequencing was too much for us to do routinely.

Fortunately, as this study shows, our level of transmitted drug resistance in the untreated population is low, however with the PANDAA assay as demonstrated by this study, routine screening for important mutations would be possible, still keeping to a budget, enabling rapid response if relevant mutations are detected.

Of course, now our first line has changed to an integrase inhibitor regimen, the exact point mutations would have to be changed appropriately, but as the investigators point out this is easily done.

It would also be useful to use this modified PANDAA in the rare cases of treatment failure expected, to give a rapid overview on important mutations to craft a salvage regimen while awaiting the full Sanger sequence to fine tune that regimen.

I hope that this study can be used for our national guidelines to focus more easily and effectively on specific mutations expected to be revealed by our local PANDAA modifications.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes
Are sufficient details of methods and analysis provided to allow replication by others?  
Yes

If applicable, is the statistical analysis and its interpretation appropriate?  
Yes

Are all the source data underlying the results available to ensure full reproducibility?  
Yes

Are the conclusions drawn adequately supported by the results?  
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical HIV specialist with a specific interest in resistance and complicated or failing patients

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Methods:
1. The ABI 7500 system is not a sequence detection system BUT a real time PCR system.

Results:
1. The costing comparison is between detecting six mutations using PANDAA vs Sanger sequencing. However, the study evaluated the detection of only 3 mutations?

2. The time comparison only focuses on the real time PCR vs sequencing, it would be more objective to add the first half of the work, RNA extraction to PCR.

Introduction:
1. "HIV-1 reverse transcriptase mutations introduced into the viral genome contributes to the development of resistance to antiretroviral drugs" - Mutations in the PR, Integrase genes also contribute to ART resistance.

2. The statement that, "Major non-nucleoside reverse transcriptase inhibitor (NNRTI) mutations, such as K103N and V106M, are selected when exposed to nevirapine (NVP), which is still used in low resource settings as part of management." omits the role of EFV in the selection of those two mutations and EFV is also used in RLS. It's the drug that was in use at the time of collection of the samples used in this study.

3. "Studies have shown that >98% of patients failing first-line NNRTI regimens would have one or more drug resistance mutations, such as K103N, V106M and M184V" is highly inaccurate. I believe the statement originated from Rhee et al. (2015) which implied that the set six mutations (K65R, M184V, K103N, V016M, Y181C, and G190A) was 98% sensitive in detecting HIVDR in treatment failure in RLS at that time, taking into consideration the regimens that were in use. Furthermore, if one mentions "studies" it is expected that they provide multiple references or a single systematic review as the reference.

4. "Moreover, utilising point mutation assays reduces the need for several laboratory equipment". This statement may be misleading, especially in reference to the PANDAA assay. The major difference between PANDAA and Sequencing is the method used in the detection of the mutations. One uses real-PCR to detect specific point mutations and the other uses sequencing. The extraction, reverse transcription and PCR are similar. When doing such comparisons, one should be cognoscente of the fact that there are now low cost sequencing platforms that may be even cheaper that some of the real time PCR platforms on the market.

Methods:
1. It's not clear which other samples, other than the Gaborone ANC 2014-2015 (K103N, n=4) samples were used for this study. Please clarify this. This is key to understanding the analysis that was done, the results and their interpretation.

2. The investigators used amplicons that were generated during the sequencing process for the PANDAA evaluation. If this is correct, there is no need to provide details of RNA extraction, reverse transcription, PCR and PCR product clean up. This would have been reported already in the primary paper.

3. When reporting target regions for primers (Table 1), its important to indicate the reference,
for example HXB2 if that is the reference you used.

4. Did the investigators do any sequencing for this paper, if not, there is no need to provide details of the method. Reference should be made to the primary paper.

5. "All reactions were performed in triplicate, and the mean of the three values was used for calculation", isn't it better to use median rather than mean?

6. For cost comparison, the investigators should indicate the cost inputs.

7. The analysis of the differences in CD4 counts and viral loads between the groups with drug resistance and those without, seems to be out of place. It's not indicated anyway in the objectives of the study. Nothing about this is even mentioned in the introduction. The power of this analysis is further diminished by the small number of samples with HIVDR mutations (3) in this study.

**Results:**

1. The investigators mention that they did an analysis to determine the "lower limit of detection" for PANDAA but the result is not indicated.

2. Turnaround time (Table 4). For the statement that the turn around time for PANDAA is 2 hrs and that of Sanger is 24 hrs to be objective, one needs to highlight that this time is just for the real time PCR vs Sequencing, excluding RNA extraction, Reverse transcription, PCR, and PCR clean up. They should also provide the sequencing protocol as well as the sequencing clean up protocol. In general the sequencing PCR should not take more that 1.5 hrs, the clean up 30 minutes and the sequencing electrophoresis, not more than 1.5 hours per injection. If you have a 16 capillary array, your turn around time will be much shorter compared to the 8 capillary array. So, its important to indicate all this information so as to determine how the 15 hrs came about.

**Discussion:**

1. This evaluation or comparison would have benefited more from a sample with higher proportions of the different mutations evaluated.


3. Turn around times need to be assessed in the context of the workflow in patient management. For example, if a patient has to go home and come and collect results after 48hrs, a difference of 13 hrs will not add value for the patient. However, if PANDAA enabled same day delivery of results, then, it would make a significant difference in patient management. Unfortunately the authors only focused on the time post PCR, which gives an illusion that it takes 2 hrs to get results for PANDAA. A proper assessment would have to be from RNA extraction to genotyping result.

4. The statement that, in Botswana, NVP is still being used as prophylaxis for prevention of mother to child transmission of HIV, is mostly likely in accurate. Please provide a reference like the current HIV management guidelines for Botswana.
5. The authors properly highlight the limitations of this study, some which have been highlighted above. However these are very strong limitations which will significantly reduce the impact of this paper.

References
PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular Virology, Bioinformatics, Molecular Diagnostics, HIV

We confirm that we have read this submission and believe that we have an appropriate level of expertise to state that we do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 13 Apr 2021

Dorcas Maruapula, Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana

Authors’ Response to Reviewers’ Comments:

Manuscript Title: Use of a mutation-specific genotyping method to assess for HIV-1 drug resistance in antiretroviral-naive HIV-1 Subtype C-infected patients in Botswana

We thank the reviewers for their valuable comments, suggestions and insights that have
helped us improve our manuscript. We have revised the manuscript by clearly discussing the limitations of the method in the Discussion section. We hope the reviewers will find our revised manuscript improved and suitable for publication.

Please find below a point-by-point response to the reviewers comments.

Sincerely,
D.M

Abstract:

Background:

Reviewer's comments: The combination of mutations evaluated in this study is suboptimal in the context of the current ISTI based regimens. In the context of a TLD regimen, a panel of mutations that includes at least the K65R mutation for TDF in addition to the M184V mutation in the current panel would be much more useful.

Authors' response: We thank the reviewer for the suggestions. The mutations used here are important although they are not all inclusive. Future work can focus on other mutations that were not included here. It is most likely that the list of important mutations to be evaluated will always change and more mutations added to the list.

Reviewer's comments: The statement that, "Mutations selected at these positions have been shown to be the most common driver mutations in treatment failure" is no longer true in the context of the current regimens. This is a historical statement and it should be clearly stated. This will allow the evaluation of mutations that are important to the current regimens.

Authors' response: We are thankful to the reviewers for the suggestions. We have indicated that the mutations selected were some of the most common driver mutations for treatment failure at the time the study was conducted and for NNRTI based ART regimen. We have also included more literature which shows that baseline NNRTI resistance are linked to poor response to first-line dolutegravir, therefore it is still important to look at NNRTI resistance mutations (MJ Siedner et al, 2020).

Methods:

1. Reviewer's comment: The ABI 7500 system is not a sequence detection system BUT a real time PCR system.

2. Author's response: We thank the reviewer for the observation. We have corrected the wording in the new manuscript.

Results:

1. Reviewer's comment: The costing comparison is between detecting six mutations using PANDAA vs Sanger sequencing. However, the study evaluated the detection of only 3 mutations?

Author's response: We have now clearly stated the costing comparison in the revised manuscript.

1. Reviewer's comment: The time comparison only focuses on the real time PCR vs sequencing, it would be more objective to add the first half of the work, RNA extraction to PCR.

Author's response: We thank the reviewer for the suggestion. We have added the first half
of the work, RNA extraction to PCR into the manuscript.

Introduction:

1. **Reviewer's comment**: "HIV-1 reverse transcriptase mutations introduced into the viral genome contributes to the development of resistance to antiretroviral drugs" - Mutations in the PR, Integrase genes also contribute to ART resistance.

**Author's response**: We thank the reviewer for the comment. We have corrected the sentence.

2. **Reviewer's comment**: The statement that, "Major non-nucleoside reverse transcriptase inhibitor (NNRTI) mutations, such as K103N and V106M, are selected when exposed to nevirapine (NVP), which is still used in low resource settings as part of management." omits the role of EFV in the selection of those two mutations and EFV is also used in RLS. It's the drug that was in use at the time of collection of the samples used in this study.

**Author's response**: We thank the reviewer for the observation. We have improved on the wording of the manuscript.

3. **Reviewer's comment**: "Studies have shown that >98% of patients failing first-line NNRTI regimens would have one or more drug resistance mutations, such as K103N, V106M and M184V" is highly inaccurate. I believe the statement originated from Rhee et al. (2015) which implied that the set six mutations (K65R, M184V, K103N, V016M, Y181C, and G190A) was 98% sensitive in detecting HIVDR in treatment failure in RLS at that time, taking into consideration the regimens that were is use. Furthermore, if one mentions "studies" it is expected that they provide multiple references or a single systematic review as the reference.

**Author's response**: We thank the reviewer for the suggestions. We have added more references.

4. **Reviewer's comment**: Moreover, utilising point mutation assays reduces the need for several laboratory equipment. This statement may be misleading, especially in reference to the PANDAA assay. The major difference between PANDAA and Sequencing is the method used in the detection of the mutations. One uses real-PCR to detect specific point mutations and the other uses sequencing. The extraction, reverse transcription and PCR are similar. When doing such comparisons, one should be cogniscente of the fact that there are now low cost sequencing platforms that may be even cheaper that some of the real time PCR platforms on the market.

**Author's response**: We thank the reviewer for the comment. We have corrected the wording of the sentence.

Methods:

1. **Reviewer's comment**: It's not clear which other samples, other than the Gaborone ANC 2014-2015 (K103N, n=4) samples were used for this study. Please clarify this. This is key to understanding the analysis that was done, the results and their interpretation.

**Author's response**: We thank the reviewer for the comment. We have now clearly stated the samples.

2. **Reviewer's comment**: The investigators used ampicons that were generated during the sequencing process for the PANDAA evaluation. If this is correct, there is no need to
provide details of RNA extraction, reverse transcription, PCR and PCR product clean up. This would have been reported already in the primary paper.

Author's response: We thank the reviewer for the comment. We have highlighted in the manuscript that the amplicons generated were reported elsewhere.

1. Reviewer's comment: When reporting target regions for primers (Table 1), it's important to indicate the reference, for example HXB2 if that is the reference you used.

Author's response: We thank the reviewer for the comment. We indicated the reference in the revised manuscript.

1. Reviewer's comment: Did the investigators do any sequencing for this paper, if not, there is no need to provide details of the method. Reference should be made to the primary paper.

Author's response: We thank the reviewer for the suggestions. We will update the reference of the primary paper.

1. Reviewer's comment: "All reactions were performed in triplicate, and the mean of the three values was used for calculation", isn't it better to use median rather than mean?

Author's response: We thank the reviewers for the suggestion. To enable comparisons with previous studies, we also report means.

1. Reviewer's comment: For cost comparison, the investigators should indicate the cost inputs.

Author's response: We thank the reviewer for the suggestion. We have clearly indicated the cost inputs for cost comparison.

1. Reviewer's comment: The analysis of the differences in CD4 counts and viral loads between the groups with drug resistance and those without, seems to be out of place. It's not indicated anyway in the objectives of the study. Nothing about this is even mentioned in the introduction. The power of this analysis is further diminished by the small number of samples with HIVDR mutations (3) in this study.

Author's response: We thank the reviewers for the comments. We have included the above point as a limitation in the revised manuscript.

Results:

1. Reviewer's comment: The investigators mention that they did an analysis to determine the “lower limit of detection” for PANDAA but the result is not indicated.

Author's response: We thank the reviewer for the observation. We have corrected the above text in the manuscript.

1. Reviewer's comment: Turnaround time (Table 4). For the statement that the turn around time for PANDAA is 2 hrs and that of Sanger is 24 hrs to be objective, one needs to
highlight that this time is just for the real time PCR vs Sequencing, excluding RNA extraction, Reverse transcription, PCR, and PCR clean up. They should also provide the sequencing protocol as well as the sequencing clean up protocol. In general the sequencing PCR should not take more than 1.5 hrs, the clean up 30 minutes and the sequencing electrophoresis, not more than 1.5 hours per injection. If you have a 16 capillary array, your turn around time will be much shorter compared to the 8 capillary array. So, its important to indicate all this information so as to determine how the 15 hrs came about.

**Author's response:** We thank the reviewer for the suggestions. We have highlighted the above comment in the manuscript.

**Discussion:**

1. **Reviewer's comment:** This evaluation or comparison would have benefited more from a sample with higher proportions of the different mutations evaluated.

**Author's response:** We completely agree with the reviewers and we have clarified this point in the Discussion as a limitation. We had to work with the samples that we had at the time the study was conducted.

1. **Reviewer's comment:** The WHO 90-90-90 goals target the elimination of HIV as a public health threat by 2030 NOT 2020.

**Author's response:** We agree with reviewer, and we have now modified this sentence accordingly.

1. **Reviewer's comment:** Turn around times need to be assessed in the context of the workflow in patient management. For example, if a patient has to go home and come and collect results after 48hrs, a difference of 13 hrs will not add value for the patient. However, if PANDAA enabled same day delivery of results, then, it would make a significant difference in patient management. Unfortunately the authors only focused on the time post PCR, which gives an illusion that it takes 2 hrs to get results for PANDAA. A proper assessment would have to be from RNA extraction to genotyping result.

**Author's response:** A proper assessment was addressed in the revised manuscript.

1. **Reviewer's comment:** The statement that, in Botswana, NVP is still being used as prophylaxis for prevention of mother to child transmission of HIV, is mostly likely in accurate. Please provide a reference like the current HIV management guidelines for Botswana.

**Author's response:** We have modified the manuscript with citations to the statement.

1. **Reviewer's comment:** The authors properly highlight the limitations of this study, some which have been highlighted above. However these are very strong limitations which will significantly reduce the impact of this paper.

**Author's response:** We agree with the reviewer that our study has some limitations that have been highlighted in the manuscript. We however think that this is important work on the use of the groundbreaking PANDAA assay in an HIV-1 subtype C endemic setting. Future work will build on this project as more HDR mutation positions are covered by the assay.
We appreciate your valuable comments and we will address all these in the revised manuscript.

**Competing Interests:** No competing interests were disclosed.