Deep Sequencing Reveals Minor Protease Resistance Mutations in Patients Failing a Protease Inhibitor Regimen

Randall Fisher,a Gert U. van Zyl,a,b Simon A. A. Travers,c Sergei L. Kosakovsky Pond,d Susan Engelbrech,a,b Ben Murrell,e,f Konrad Scheffler,a and Davey Smithfd

Division of Medical Virology, Stellenbosch University, Stellenbosch, South Africa; National Health Laboratory Service, Coastal, Tygerberg, South Africa; South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa; Department of Medicine, Division of Infectious Diseases, University of California, San Diego, San Diego, California; Computer Science Division, Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa; and Biomedical Informatics Research Division, eHealth Research and Innovation Platform, Medical Research Council, Tygerberg, South Africa.

Standard genotypic antiretroviral resistance testing, performed by bulk sequencing, does not readily detect variants that comprise <20% of the circulating HIV-1 RNA population. Nevertheless, it is valuable in selecting an antiretroviral regimen after antiretroviral failure. In patients with poor adherence, resistant variants may not reach this threshold. Therefore, deep sequencing would be potentially valuable for detecting minority resistant variants. We compared bulk sequencing and deep sequencing to detect HIV-1 drug resistance at the time of a second-line protease inhibitor (PI)-based antiretroviral regimen failure. Eligibility criteria were virologic failure (HIV-1 RNA load of >500 copies/ml) of a first-line nonnucleoside reverse transcriptase inhibitor-based regimen, with at least the M184V mutation (lamivudine resistance), and second-line failure of a lopinavir/ritonavir (LPV/r)-based regimen. An amplicon-sequencing approach on the Roche 454 system was used. Six patients with viral loads of >90,000 copies/ml and one patient with a viral load of 520 copies/ml were included. Mutations not detectable by bulk sequencing during first- and second-line failure were detected by deep sequencing during second-line failure. Low-frequency variants (>0.5% of the sequence population) harboring major protease inhibitor resistance mutations were found in 5 of 7 patients despite poor adherence to the LPV/r-based regimen. In patients with intermittent adherence to a boosted PI regimen, deep sequencing may detect minority PI-resistant variants, which likely represent early events in resistance selection. In patients with poor or intermittent adherence, there may be low evolutionary impetus for such variants to reach fixation, explaining the low prevalence of PI resistance.

Standard HIV drug resistance genotypic testing often involves a PCR amplifying the HIV-1 pol coding region, followed by sequencing by electrophoresis of the total sample of circulating HIV RNA species (26); hence, this method is referred to as bulk sequencing. This process has clinical value in the detection of antiretroviral resistance and allows the selection of a new antiretroviral regimen in patients who have experienced failure of their current antiretroviral therapy regimen. However, an important limitation of bulk sequencing is that it is able to reliably detect only viral variants that comprise at least 20% of the circulating viral population; otherwise, the sequence readout may represent only the nucleotide sequence of the predominant variant (13, 17). Nevertheless, if these minor variants harbor resistance-associated mutations, they are likely clinically relevant. For example, in mothers who were exposed to nevirapine (NVP) as part of prevention of mother-to-child transmission (PMTCT), the detection of a minor variant population with resistance mutations such as K103N and Y181C increased the likelihood of future NVP regimen failure (16). In clinical settings where one is interested in the identification of mutations only at a limited number of loci, allele-specific assays are practicable. When one attempts to detect all important minor variant resistance mutations that could have been transmitted or could have been acquired during combination antiretroviral therapy, multiple allele-specific assays would be needed. However, even a large array of allele-specific assays would not be able to detect all mutations (9). Therefore, to fully sample transmitted or acquired variants, a different approach is needed. Next-generation sequencing employs the parallel sequencing of single genomes, which, due to the relative long sequencing read length, has the added advantage over allele-specific assays of being able to detect mutations in the context of a sequence and not just a single locus (23). One approach is the sequencing of PCR amplicons referred to as ultradeep pyrosequencing (UDPS), for example, sequencing on the Roche 454 platform.

The detection of minor variants using UDPS has recently been applied in clinical settings by a number of research and clinical studies. For example, minor variants have been shown to predict antiretroviral failure to nonnucleoside reverse transcriptase inhibitor (NNRTI)-based regimens (19), including the detection of etravirine resistance-associated mutations at low frequency (23). Despite the clinical value of detecting minority variants, these methods are subject to sampling error and PCR or sequencing artifacts. For example, the detected frequency of identified minority variants in recently infected patients could either be due to these artifacts or be true mutations induced during viral replication (6). Furthermore, PCR or sequencing artifacts on the UDPS platform may be related to the nucleic acid template, and this could also result in spurious mutation detection, such as K65R in HIV-1 subtype C (24).
Considering these issues, we investigated antiretroviral resistance with bulk sequencing and UDPS among a group of patients receiving a second-line antiretroviral regimen containing lopinavir/ritonavir (LPV/r) who developed virologic failure to this regimen largely as a result of poor adherence (22). Bulk sequencing methods in this setting do not usually detect any protease inhibitor (PI) resistance (15). This observation could be due to the high genetic barrier of the LPV/r regimen. Specifically, variants with a single resistance mutation have a very limited survival advantage and may therefore be out-competed by wild-type variants and remain undetectable by bulk sequencing. In the process of PI resistance becoming fixed in the circulating HIV RNA population, it is likely that minor resistance variants wax and wane with intermittent drug adherence; therefore, more sensitive methods, like UDPS, would be needed to detect these mutations.

MATERIALS AND METHODS

Patient selection, viral load measures, and genotyping testing with bulk sequencing. Patients were eligible for the study under the following conditions: (i) if they had previously received a first-line NNRTI-based regimen on which they had experienced virologic failure, with a viral load of >500 HIV RNA copies/ml (Abbott m2000 Real Time HIV-1 Assay; Abbott Molecular, Inc.), and had genotypic resistance testing by bulk sequencing that identified at least lamivudine (3TC) resistance with an M184V/I mutation (patients who demonstrated only NNRTI resistance were excluded); (ii) if they were switched to a second-line protease inhibitor (LPV/r)-based regimen and again had virologic failure, with a viral load of >500 copies/ml and had genotypic resistance testing by bulk sequencing at the time of second-line failure.

In both episodes of virologic failure, bulk sequencing of HIV-1 pol was performed, as previously described (21). All patients gave informed consent and were included as part of a larger observational study, approved by the Stellenbosch University, Committee for Human Research (N06/05/081).

UDPS. UDPS was performed on a sample contemporaneous to that used for bulk sequencing during second-line failure. HIV RNA was extracted from blood plasma with a NucliSENS easyMAG total nucleic acid extraction system (bioMérieux, Boxtel, The Netherlands). Reverse transcription of HIV RNA to yield cDNA along with the first-round PCR was performed according to a one step Reverse Transcriptase PCR (RT-PCR) protocol (Access RT-PCR System; Promega, Madison, WI), which uses the reverse transcriptase enzyme from the avian myeloblastosis virus (AMV) and DNA polymerase from Thermus flavus and HIV-1 subtype C pol-specific primers (see Table S1 in the supplemental material). The resultant HIV pol fragment product was used as the template to amplify the HIV-1 coding regions of the protease and reverse transcriptase genes in three overlapping fragments using primers adapted from Varghese et al. (23) for optimal detection of South African subtype C HIV-1. Primers consisted of a 454 sequencing adaptor, a patient-specific molecular identifier (MID), and the specific viral sequence for priming (see Table S1). For this second-round PCR, Expand High Fidelity DNA polymerase (Roche Applied Science, Basel, Switzerland) was used for PCR amplification of the target regions.

In order not to waste sequencing coverage on short fragments such as primers or primer-dimers, amplified products were separated on a 0.8% (wt/vol) agarose gel by electrophoresis. DNA in the gel was stained using 1% (vol/vol) GelStar (Cambrex), which allowed the visualization of cDNA under an amber-colored screen and blue transilluminator backlight. This staining method was chosen rather than ethidium bromide and UV light transillumination in order to prevent DNA damage and mutations. PCR products of the correct sizes were excised from the gel and purified using a Wizard GEL and PCR purification kit (Promega). The purified products were then reconstituted in nuclease-free water before being quantified and purity checked by UV spectroscopy. Purified PCR products were then sent to Inqaba Biotech, Pretoria, South Africa, for UDPS. At this facility, equimolar pooling of the three fragments for each patient was performed, followed by emulsion PCR and pyrosequencing on a Titanium FLX (454) Sequencer (Roche).

Data analysis and bioinformatics. Generated sequences were demultiplexed based on the barcodes present in the patient-specific primers. Using the CLC Bio Genomics Workbench (CLC Bio, Denmark), the data set was “cleaned” by excluding bases with a base call confidence below 95%. Regions with poor quality reads were trimmed. Once cleaned, the forward and reverse UDPS reads for each coding region were combined before being aligned to HIV subtype C consensus sequences, obtained from the Los Alamos HIV Sequence Database (available from http://www.hiv.lanl.gov/), for the respective genes (i.e., the two overlapping reverse transcriptase fragments were aligned to the subtype C reverse transcriptase gene consensus sequence, and the protease sequences were aligned to the subtype C protease gene consensus sequence). Sequence alignments were obtained using Segminer, bioinformatics software designed specifically for HIV minority variant detection using the Roche 454 UDPS sequencing platform (1), and minority HIV-1 resistant variants were identified using a threshold of 0.5%. Mutation loci were identified according to The HIV Drug Resistance Database Mirror in Southern Africa (available from http://bioafrica.mrc.ac.za/hivdb).

Posthoc estimation of lower limit of detection for deep sequencing. The lower limit of detection is dependent on the sampling error and background error rate (deep sequencing error). This background error was estimated in a prior publication to occur at 0.21% ± 0.08% across HIV polymerase sequences (6) and has also been determined experimentally by sequencing monoclonal populations (8). The mean substitution rates for the protease gene, the cooccurrence of mutations on the same read (linkage), and phylogenetic analyses were investigated with the use of the online molecular sequence analysis server, Datamonkey (http://www.datamonkey.org/) (5, 14).

RESULTS

Patient demographics and testing. The study included seven patients who had a mean age of 38 years. Their median time on first-line antiretroviral therapy was 12 months, with a documented virologic failure of first-line therapy of at least 2 months. Their median time on second-line therapy was 16 months, with documented virologic failure duration of at least 10 months. Patient demographic details and clinical information are given in Table 1. Bulk sequencing results were available at the time of first- and second-line failure, and UDPS results were available only at the time of second-line failure.

Deep sequencing coverage. The UDPS resulted in a total of 155,074 sequence reads, with an average of ~9,400 reads per patient sample. The coverage was therefore ~4,700 reads in both the forward and reverse directions. Average read length was 278 bases. As the reverse transcriptase A fragment could not be amplified in the case of patient 3, the coverage was 0 for that middle fragment.

Estimated sampling error. Specimens, excluding specimen 3, had viral loads of >90,000 copies/ml. The input in the extraction reaction mixture was 1 ml, diluted into 25 µl, of which 10 µl was used in the RT-PCR. Based on a reverse transcriptase efficiency of 20 to 50%, the expected number of viruses sampled was thus ~7,200. Assuming a simple binomial distribution, any minor variant occurring at a frequency of as low as 0.6% in the sample would be detected in at least 88% of cases at a frequency of >0.5% in such a sample. Therefore, with a high viral input, the expected sampling error is low, which could contribute to true detection of minor variants.
Two major PI resistance mutations (M46I and V82S) were detected. All of these major variants occurred at a frequency of 0.5 to 1.0% during first-line or second-line failure. The mean substitution rate, and rates of protease mutations (MPRMs), namely, M46V, I54T, V82A, and N88(S/D), were detected at lower levels than the expected threshold of detection in patient samples. However, bulk sequencing detected higher levels of PI resistance-associated mutations than UDPS, making up 45% of the sequences. However, bulk sequencing detected only 6% of sequences generated by UDPS. For specimen 3, which has a viral load of 520 copies/ml (repeated value, 630 copies/ml on the same day), an expected 42 copies were sampled resulting in a 19% probability of detecting a minor variant if present at 0.5%; however, variants occurring at a frequency of 0.1% would be detected in at least 5% of cases.

A summary, provided in Table 2, compares UDPS results (mutations present at ≥5%) with those of bulk sequencing. For the protease gene coding region, no major variants were detected by UDPS. For the coding region of the reverse transcriptase gene, the M184V/I mutation was detected by UDPS in three patients. Additionally, during first-line failure, the only thymidine analogue mutation (TAM) that was detected by UDPS was M41L in patient 7. During second-line failure, no TAM was detected by bulk sequencing; however, UDPS detected K65N in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N in four of six patients in the second line. During first-line failure, no TAM was detected by bulk sequencing. This was expected as all patients were receiving lamivudine in their first-line regimen, which selects for this mutation, and the persistence of this mutation was an entry criterion for this study. In all seven patients, M184V/I was no longer detectable by bulk sequencing from the subset of reads which spanned both positions. In another case, UDPS detected T215I, which is a common resistance mutation used to distinguish between the second and third PI resistance mutations (20, 24). For mutations associated with NNRTI resistance, results from UDPS differed in some patients. In one case, UDPS detected K65R mutation. This is a common resistance mutation to be associated with NRTI resistance, the M184V/I mutation was detected by UDPS in three patients. Additionally, during first-line failure, the only thymidine analogue mutation (TAM) that was detected by UDPS was M41L in patient 7. During second-line failure, no TAM was detected by bulk sequencing; however, UDPS detected K65N in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25). Nevertheless, UDPS detected D67N, which was detected in four of six patients in the second regimen not containing lamivudine and by the fact that the M184V/I mutation is a less fit variant in the absence of lamivudine selection pressure (20, 25).
### TABLE 2 Resistance detected with bulk sequencing during first-line (bulk sequencing) and second-line (bulk sequencing and UDPS) failure

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Reverse transcriptase</th>
<th>Protease</th>
<th>Mutation(s) during first-line NNRTI failure detected by bulk sequencing</th>
<th>Bulk sequencing</th>
<th>Mutation(s) during second-line PI failure detected by:&lt;sup&gt;a&lt;/sup&gt;</th>
<th>UDPS (frequency [%])</th>
<th>NNRTI</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A62V, M184V, V106I, Y181C, H221Y</td>
<td>M46I, L89M, I93L</td>
<td>None</td>
<td>M36I, L89M, I93L</td>
<td>K65R (1.1), D67N (0.9), D67E (0.9), K219R (0.7)</td>
<td>V90I (0.8), A98E (0.9), K101E (5.9), K103R (44.7), K103N (5.1), K103E (3.2), V179I (0.5), L89M (0.6), F227L (0.8), F227S (0.6), K238R (0.5)</td>
<td>V90I (0.8), K101R (0.7), K103R (2.2), Y181C (1.9), F227S (0.5)</td>
<td>L23P (0.5), M36I (98.7), L89I (99.1), I93L (99.1)</td>
</tr>
<tr>
<td>2</td>
<td>M184V, V106M</td>
<td>M36I, L63P, L89M, I93L</td>
<td>None</td>
<td>M36I, L63P, L89M, I93L</td>
<td>K65R (3.8), D67N (8.4), F77L (0.7), M184V (8.0), L210S (0.7), T215 (0.8)</td>
<td>V90I (0.8), K101R (0.7), K103R (2.2), Y181C (1.9), F227S (0.5)</td>
<td>K101E (0.6), K101R (0.6), P225T (1.5), F227L (0.6), K238T (3.3), K238R (0.7)</td>
<td>M46V (0.7), F53I (0.6), F53S (0.5), K20R (73.5), M36I (67.5), M36I (31.6), D60E (59.3), L63P (38.0), L89M (79.0), I93L (99.7)</td>
</tr>
<tr>
<td>3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>M184V, V90I, K103N, Y181C</td>
<td>K20R, M36I, L63P, L89M, I93L</td>
<td>None</td>
<td>K20R, M36I, L63P, L89M, I93L</td>
<td>V118A (0.5)<em>, K219E (0.5)</em></td>
<td>V179I (5.8)</td>
<td>K101E (0.6), K101R (0.6), P225T (1.5), F227L (0.6), K238T (3.3), K238R (0.7)</td>
<td>M36I (98.6), L89M (99.6), I93L (99.8)</td>
</tr>
<tr>
<td>4</td>
<td>M184V, V106I, Y181C, H221Y</td>
<td>K20R, M36I, D60E, L89M, I93L</td>
<td>None</td>
<td>K20R, M36I, D60E, L89M, I93L</td>
<td>K65R (2.7), D67N (1.5), F116S (0.5), M184V (2.6)</td>
<td>K101E (0.6), K101R (0.6), P225T (1.5), F227L (0.6), K238T (3.3), K238R (0.7)</td>
<td>K101E (0.6), K101R (0.6), P225T (1.5), F227L (0.6), K238T (3.3), K238R (0.7)</td>
<td>M46V (0.7), F53I (0.6), F53S (0.5), K20R (73.5), M36I (67.5), M36I (31.6), D60E (59.3), L63P (38.0), L89M (79.0), I93L (99.7)</td>
</tr>
<tr>
<td>5</td>
<td>M184V</td>
<td>M36I, L63P, I93L</td>
<td>K103N</td>
<td>M36L, L63P, I93L</td>
<td>K65R (1.6), K65E (0.5), D67N (0.5), M184V (3.1)</td>
<td>K101E (1.3), K101R (0.6), K103N (55.2), V179D (1.0), P225T (1.0), F227L (0.6), F227S (0.6), K238T (2.9)</td>
<td>K101E (1.3), K101R (0.6), K103N (55.2), V179D (1.0), P225T (1.0), F227L (0.6), F227S (0.6), K238T (2.9)</td>
<td>M36I (98.6), L89M (99.7), I93L (99.7)</td>
</tr>
<tr>
<td>6</td>
<td>M184V, K103N</td>
<td>D60E, L63P, I93L</td>
<td>None</td>
<td>D60E, L63P, L93L</td>
<td>K65R (1.0), K65E (0.6), K219E (0.5)</td>
<td>K103E (0.8), G190E (0.7)</td>
<td>K103E (0.8), G190E (0.7)</td>
<td>V82A (0.6), K20R (34.5), M36I (28.1), D60E (97.9), I62V (0.7), L63P (81.2), L93L (99.6)</td>
</tr>
<tr>
<td>7</td>
<td>M41L, K65R, V75I, M184V, K103R, V179D</td>
<td>M36I, L63S, T74S, I93L</td>
<td>V179D</td>
<td>M36I, L63S, T74S, I93L</td>
<td>K65R (2.2), T215A (0.7), K219R (0.7), K219E (0.5)</td>
<td>V90I (1.0), V179D (6.1)</td>
<td>V90I (1.0), V179D (6.1)</td>
<td>M36I (93.2), D60E (8.1), L63S (90.7), L63P (9.0), T74S (90.2), I93L (99.6)</td>
</tr>
</tbody>
</table>

<sup>a</sup> For the NNRTI failure episode, NRTI mutations are in roman, and NNRTI mutations are in italics.

<sup>b</sup> For UDPS mutations, major PI resistance mutations are shown in bold, accessory mutations are in italics, and other amino acid variants at PI resistance loci are in roman type.

<sup>c</sup> Asterisks indicate the detection of minor variants below the predicted threshold, based on the sample input (viral load of 520 copies/ml).
variants harboring drug resistance that were not detected by bulk sequences at the time of first-line failure are JN638092, JN638100, JN638126, JN638158, JN638164, JN638183, and JN638219 and at the time of second-line failure are JN700927 to JN700933.

Sequence data. The GenBank accession numbers of HIV-1 pol sequences at the time of first-line failure are JN638092, JN638100, JN638126, JN638158, JN638164, JN638183, and JN638219 and at the time of second-line failure are JN700927 to JN700933.

**DISCUSSION**

Similar to results in a previous study that used UDPS (7), some drug resistant HIV-1 variants previously detectable at the time of regimen failure became undetectable after removal of selective drug pressure (7). Nevertheless, UDPS was able to detect minority variants harboring drug resistance that were not detected by bulk sequencing. Many factors could contribute to the detection of minor variants: (i) resistant viral populations that decayed below the threshold of detection with bulk sequencing in the absence of drug pressure (e.g., nonnucleoside reverse transcriptase [NNRT] mutations and M184V); (ii) minor variants emerging on the current regimen; (iii) genetic hitchhiking of mutations which occur in cis with mutations that have unrelated survival advantages (immune escape or mutations that could contribute to resistance of the second regimen); (iv) sampling or technical artifact.

The detection of K65R in all six patients where it was sequenced is likely an artifact of UDPS due to PCR error which is template dependent: codon usage differs among subtypes in the KKK-encoding nucleotide sequence (codons 64 to 66 of the reverse transcriptase gene), which for subtype C is either AAA-AAG-AAG or AAA-AAG-AAA and for subtype B is either AAG-AAA-AAA or AAG-AAA-AAG. The use of high-fidelity enzymes could prevent this error (24). As a high-fidelity enzyme was used in the second-round PCR, this is probably an artifact of the first-round PCR.

UDPS was able to detect minor PI-resistant variants in five of seven patients, whereas bulk sequencing did not detect any resistance. Therefore, UDPS may be better able to identify the development of PI resistance before it becomes fixed in the circulating viral population. The presence of these minor variants harboring PI resistance may be indicative of selection pressure for PI resistance; however, competition with wild-type strains may have prevented these resistant variants from predominating as the lopinavir levels were previously reported to be low in this patient group (22). Selection due to low drug levels could have been inadequate for emerging resistant variants to predominate (2). Stochastically, in patients without prior PI resistance, as in our study, most minor resistant variants would initially be single-mutation variants and would thus represent a first step in the evolution of clinically significant resistance but would need to acquire additional mutations (either de novo mutations or through recombination) before clinical resistance due to multiple resistance mutation variants would emerge (10). Nevertheless, there was some evidence of possible epistatic interactions of minor resistant variants: the linkage of one major PI mutation (I54V) with the accessory mutation L63P in patient 3 and two accessory mutations, D60E and L63P in patient 7.

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>PI mutation(s) (frequency [%])</th>
<th>Mean nucleotide substitution rate (%)</th>
<th>Linkage of mutations (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I54T (0.5), M36I (99.5), L63P (0.9), L89M (99.3), I93L (99.7)</td>
<td>2.8</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>L23P (0.5), M36I (98.7), L63P (99.6), L89M (99.5), I93L (99.1)</td>
<td>2.2</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>I54T (0.7)*; I84V (0.9)*; K20R (98.3), M36I (98.6), L62V (1.2), L63P (2.0), A71T (1.3), L89M (99.6), I93L (99.8)</td>
<td>3</td>
<td>I54V and L63P (0.0015)</td>
</tr>
<tr>
<td>4</td>
<td>M46V (0.7), F35L (0.6), F35S (0.5), K20R (73.5), M36I (67.5), M36L (31.6), D60E (59.3), L63P (38.0), L89M (79.0), I93L (99.7)</td>
<td>5.2</td>
<td>M36L and L63P (0.001)</td>
</tr>
<tr>
<td>5</td>
<td>F53L (0.7), N88S (0.7), N88D (0.6), K20R (0.5), M36I (3.0), M36L (96.5), I93L (99.7)</td>
<td>2.7</td>
<td>None</td>
</tr>
<tr>
<td>6</td>
<td>V82A (0.6), K20R (34.5), M36I (28.1), D60E (97.9), I62V (0.7), L63P (81.2), I93L (99.6)</td>
<td>5.7</td>
<td>None</td>
</tr>
<tr>
<td>7</td>
<td>M36I (93.2), D60E (81.1), L63S (90.7), L63P (9.0), T74S (90.2), I93L (99.6)</td>
<td>2.9</td>
<td>D60E and L63P (0.001)</td>
</tr>
</tbody>
</table>

*Since the viral load of patient 3 was low (520 copies/ml), there was a low probability of detecting minor variants. However, the linkage of I54V and L63P suggests that the detected minor variants may be true minor populations.

Patients 4 and 6 have two phylogenetically distinct populations in all three regions. In the case of patient 4, the linkage of M36L and L63P could therefore not be used as evidence of coselection under therapy pressure.

I54V was detected at only 0.1% but significantly cooccurred with L63P on the same sequence reads.
and L63P, in patient 7 could suggest early steps in the evolutionary pathway under selective pressure. In the case of patient 3, the viral load had come down to 520 copies/ml from 11,000 copies/ml 2 months before, probably as a result of improved adherence and increased selection pressure.

In patients with low levels or intermittent adherence, the viral populations spend short periods under conditions that select for PI-resistant variants, but these are interspersed with long periods where wild-type virus is favored as single PI resistance mutation variants are less fit in the absence of drug pressure. This explains the low probability of acquiring multiple PI resistance mutation variants that would result in clinical resistance. It also explains the absence of resistance as detected by bulk sequencing in most patients with intermittent or low levels of adherence on boosted PI regimens (3, 10). It is, however, difficult to study this evolutionary process as it would require frequent sampling and powerful analyses to study the pathways to multiple resistance, disentangling recombination from sequential resistance evolution—especially as these would be low probability events, as evident from the low prevalence of resistance in patients on boosted PI regimens (11). To study the evolution of these minor variants would therefore require the retrospective analysis of a very large cohort of patients on PI therapy who had archived specimens available during subsequent periods of virologic failure.

Despite the intriguing findings presented here, this study has a number of limitations that should be acknowledged. First, specimens were available for UDPs only at the time of second-regimen failure. It would have been informative also if UDPs results could have been performed at baseline and at first-regimen failure. Second, the mean sequencing read length was only 278 bases, which reduced coverage distal to the primers. A longer read length may have allowed more uniform coverage and better minor variant discrimination across the whole of the target sequence. Third, patient 3 had a low viral load, which may have increased the possibility of sampling error, with the result that minor variants that could have been present were not detected or that minority variants were overrepresented. Also for patient 3, we were unable to amplify the reverse transcriptase A fragment and therefore could not investigate the presence of mutations in this fragment. Fourth, although HIV RNA levels were high for six of the seven patients, we did not quantitate the HIV cDNA input into UDPs: a low input could have led to disproportionate sampling of variants and may have increased the minority variant threshold. Fifth, deep sequencing is likely not sensitive enough to detect all archived mutations, and inherent PCR and sequencing artifacts may compromise its specificity in some instances, such as K65R (24). Lastly, since we do not have follow-up data on the patients after deep sequencing, we could not determine the clinical significance of the deep sequencing results.

In patients with intermittent adherence on a boosted PI regimen, deep sequencing may detect the emergence of minority PI-resistant variants, which likely represent singly resistant variants, a first step in the evolution of resistance. However, in patients with poor or intermittent adherence, conditions may not favor the further evolution of resistance, which may explain the low prevalence of boosted PI resistance in these patients. In summary, for patients without prior resistance tests and who fail a second antiretroviral regimen, UDPs may provide a valuable tool to study the evolution of resistance.

ACKNOWLEDGMENTS

This study was supported by a South African Department of Health Comprehensive Care and Treatment Programme (CCMT) grant, the South African Polymyelitis Research Foundation, and the National Health Laboratory Research Trust (South Africa). B.M. is supported by a European grant (SANTE/2007/147-790) from the European Commission. Computational analyses were supported in part by the BEAST Core of the University of California, San Diego, Center for AIDS Research (AI36214). D.S. is supported by the National Institutes of Health under grants AI69432, AI043638, MH62512, MH08352, AI077304, AI36214, AI047745, AI74621, and AI080353 and the James B. Pendleton Charitable Trust.

REFERENCES