TITLE
HIV Resistance Testing Using Point-of-Care Dynamic Solid-Phase Melt Analysis

PROJECT SUMMARY
The increasing prevalence of HIV-1 drug resistance (DR) is a threat to HIV-1 treatment and prevention in the low- and middle-income countries hardest hit by the HIV-1 pandemic. An inexpensive genotypic resistance test (GRT) that could be performed during a clinic visit, heretofore called a point-of-care (POC) assay, would enable HIV-1 care providers to make informed treatment decisions for antiretroviral (ARV)-naïve patients starting therapy and for patients with virological failure (VF) on therapy. Through the delivery of immediately actionable information upon the detection of VF, a POC GRT would benefit treatment programs by reducing treatment delivery costs and would benefit patients by reducing travel costs and work absenteeism, and by increasing the likelihood treated patients achieve and maintain suppressed virus levels.

An assay that detected the DR mutations (DRMs) K65R/N, K103N/S, V106A/M, Y181C/I/V, M184V/I and G190A/S would be highly sensitive for detecting acquired DR on an NRTI/NNRTI-containing regimen and for detecting transmitted DR that would compromise a first-line NRTI/NNRTI-containing regimen.

The main challenge in developing a POC GRT is the genetic variability at and surrounding each DRM position. We describe preliminary experiments using a novel solid-phase nucleic acid melt-analysis platform that demonstrate a high sensitivity and specificity for detecting multiple variants at the DRM position 103. We describe plans to collaborate with the biotechnology company InSilixa (Sunnyvale, CA) to develop the required solid-phase fluorogenic probe sets and a multiplex asymmetric RT PCR protocol to detect the 35 most common codons at the DRM positions 65, 103, 106, 181, 184, and 190. The probe sets and PCR protocol will be optimized and validated using a dynamic microarray setup predictive of probe performance on InSilixa's complementary metal-oxide semiconductor (CMOS) biochip-based HYRDRA-1K POC platform.

If the aims of this proposal are met, we will extend our collaboration with InSilixa to (i) evaluate sample processing modules to separate plasma from blood, concentrate virus, and extract RNA that can be delivered to the fluidic chamber of the HYDRA-1K POC platform; (ii) test a wide range of clinical samples; (iii) transfer the optimized probe sets onto the HYDRA-1K POC platform; (iv) determine whether the presence of probe hybridization can be used to define a reproducible threshold for detecting HIV-1; and (v) extend the assay to include protease inhibitor and integrase inhibitor DRMs. Developing a POC HIV-1 GRT will also address the broader challenge of detecting mutations in rapidly evolving epidemic viruses whether they are DRMs or critical vaccine-escape or gain-of-function mutations.
SPECIFIC AIMS

The increasing prevalence of acquired and transmitted HIV-1 drug resistance (DR) is a threat to HIV-1 treatment and prevention in the low- and middle-income countries (LMICs) hardest hit by the HIV-1 pandemic. An inexpensive genotypic resistance test (GRT) that could be performed during a clinic visit, heretofore called a point-of-care (POC) assay, would enable HIV-1 care providers to make informed treatment decisions for antiretroviral (ARV)-naive patients starting therapy and for patients developing virological failure (VF) while on therapy. The need for such testing is increasing as LMICs scale up virus load (VL) testing to meet the UNAIDS target of 90% virological suppression for those on ARVs. A POC test would provide access to GRT in LMICs where it is largely not available and avoid the logistical challenges associated with centralized GRT where it is available. If coupled with VL testing, it would allow for same day treatment decisions upon the detection of VF.

The most commonly used first-line ARV regimen worldwide includes the NRTIs tenofovir (TDF) + 3TC/FTC (XTC) in combination with the NNRTI efavirenz (EFV). During the next 10 years, dolutegravir (DTG) will eventually be used more frequently than EFV\(^1,2\); however, EFV is projected to be the companion ARV for at least 40% of first-line regimens through 2025\(^5\). EFV is usually more effective than PIs for initial ARV therapy\(^3,4\) and it will remain an essential ARV in regions where generic DTG cannot be procured and in patients for which DTG has not been well studied such as pregnant women, children, and persons treated for TB\(^1,2\). For these reasons, the detection of both NRTI and NNRTI DRMs should be prioritized in a POC GRT platform.

The main challenge in developing a POC GRT is the genetic variability at and surrounding each DR mutation (DRM) position. In this proposal, we present preliminary data from multiplexed experiments using a solid-phase nucleic acid melt-curve analysis platform showing that melt-curve analysis is sensitive and specific for distinguishing between multiple variants at the same position despite surrounding variability. We describe plans to extend our collaboration with the biotechnology company InSilixa (Sunnyvale, CA) to develop an assay to detect variants at the NRTI DRM positions 65 and 184 and the NNRTI DRM positions 103, 106, 181, and 190. The accurate detection of such variants would make the assay 90 to 95% sensitive for detecting intermediate / high-level acquired TDF or XTC resistance and 90 to 95% sensitive for detecting intermediate / high-level acquired EFV resistance in patients receiving TDF/XTC/EFV. Such an assay would also be 90 to 95% sensitive for detecting intermediate / high-level transmitted TDF or XTC resistance and 90 to 95% sensitive for detecting intermediate / high-level transmitted EFV resistance.

**Specific Aim 1:** To develop parsimonious probe sets capable of genotyping >99% of HIV-1 group M variants at 6 DRM positions including the NRTI DRM positions 65 and 184 and the NNRTI DRM positions 103, 106, 181, and 190. We will analyze a database of global HIV-1 sequences and perform in silico hybridization simulations to select probe candidates predicted to capture the 10 most common codons at NRTI DRM positions 65 and 184 and the 25 most common codons at NNRTI DRM positions 103, 106, 181, and 190. We will then perform in vitro experiments using a solid-phase, highly multiplexed, nucleic acid melt-curve analysis prototype platform to test, iteratively optimize and validate probe sets for detecting mutations in a wide variety of synthesized 80-mer targets representing HIV-1 sequence diversity. This prototype platform is predictive of probe performance on InSilixa’s integrated complementary metal-oxide semiconductor (CMOS) biosensor array HYRDRA-1K POC platform. At each position, we will determine the sensitivity and specificity of the assay for all codon variants in pure and mixed DNA populations.

**Specific Aim 2:** To develop a multiplex asymmetric RT PCR protocol that will generate sufficient amounts of single-stranded DNA for expedited microarray hybridization and detection. We will develop a multiplex asymmetric PCR protocol to amplify the 3 regions encompassing codon 65, codons 103 and 106, and codons 181, 184, and 190. The protocol will be optimized to yield \(>10\) nM of single-stranded DNA complementary to the probes created in Specific Aim 1 using culture supernatants and pedigreed plasma virus samples containing known numbers of HIV-1 copies.

If the aims of this proposal are met, we will plan future work to (i) evaluate sample processing modules to separate plasma from blood, concentrate virus, and extract RNA that can be delivered to the fluidic chamber of InSilixa’s HYDRA-1K POC platform; (ii) test a wide range of clinical samples; (iii) transfer the optimized probe sets onto the HYDRA-1K POC platform; (iv) determine whether the presence or absence of probe hybridization can be used to define a reproducible threshold for HIV-1 detection (i.e., a qualitative measure of VL); and (v) expand the assay to include protease inhibitor and integrase inhibitor DRMs.
INTRODUCTION TO A REVISED APPLICATION

We appreciate the reviewers’ comments on our original application, including recognition of potential significance of the work and the strength of the team and research environment. Here we describe three major revisions to the proposal (items 1 – 3) and address the previous reviewers’ concerns (items 4a – 4e).

(1) Additional preliminary studies. We evaluated 3 probe sets (18 probes containing 6 codons x 3 flanking regions) in 81 experiments to detect the reported codons at the NNRTI drug resistance (DR) position K103: AAA (K), AAG (K), AAC (N), AAT (N), AGA (R), AGC (S). In 54 experiments, we tested 18 targets (6 codons x 3 flanking regions) in triplicate, and in 27 experiments, we tested AAC targets diluted to 50%, 25%, and 10%. We attained 98% sensitivity and 99% specificity at detecting the correct codon when present at a level ≥25%.

(2) Incorporation of additional DRM positions. Although further studies are required to optimize detection of K103 variants including those present at levels of 10%, our preliminary data encouraged us to expand our work from 16 codons at 3 RT DR mutation (DRM) positions to 35 codons at the 6 RT DRM positions with the highest sensitivity for detecting acquired and transmitted DR to the NRTIs TDF + 3TC (or FTC) and the NNRTI EFV.

(3) Additional co-investigator. Dr. Benjamin Pinsky, Associate Professor of Pathology, has been added to the project. Dr. Pinsky has developed multiple PCR assays for the diagnosis of viral pathogens and has been collaborating with Dr. Shafer, Dr. Mazarei, and Dr. Hassibi on this project.

(4a) Detecting different codons at the same position (Reviewer 1). (i) “... different codons for WT could lead to differential melting curve. If so, relying on differential melting curve to deduce WT vs mutant in the setting of highly variable HIV sequences is problematic...”; (ii) “Melting curve analysis cannot determine exact DRMs compared with other genotypic approaches”; (iii) “…detectable delta Tms can also occur within WT AAs coded by 2 or more codons (e.g. AAA vs. AAG at position 103).” Preliminary data described above and in detail in the proposal show the reliable detection of 6 codons at RT position 103, demonstrating that melt-curve analysis reliably discriminates between multiple codons at the same position.

(4b) Flanking region variability (Reviewer 1). (i) “Since other DRMs may be present in the vicinity of the targeted position ... it is unclear the probe sets (which will be designed based on WT alignments) ... would be able to hybridize to targets that have multiple DRMs present in the probe binding sites.” (ii) “Figure 1 is based on delta Tm between WT and mutant with identical surrounding sequences. What if melting curves are different when surrounding sequences are variable ...?” Probe sets were actually designed using sequences from treated and untreated persons. We show melt curve analysis is highly sensitive and specific for distinguishing between multiple codons despite surrounding genetic variability. In our analyses, we excluded hybridization data from targets and probes with identical flanking sequences.

(4c) Analysis (Reviewer 1): “The InSilixa technology was not well described, and thus it’s not clear how results from the proposed studies would lead to the final POC product.” Innovation now better describes the technology. We also outline a path to a product in the Specific Aims and Research Plan: “(i) evaluate sample processing modules to separate plasma from blood, concentrate virus, and extract RNA that can be delivered to the fluidic chamber of the HYDRA-1K biochip; (ii) test a wide range of clinical samples; (iii) transfer the optimized probe sets onto the InSilixa CMOS chip-based HYDRA-1K POC platform; (iv) determine whether the presence or absence of probe hybridization can be used to define a reproducible threshold for HIV-1 detection; and (v) expand the assay to include protease inhibitor and integrase inhibitor DRMs.”

(3d) Sample processing (Reviewers 1, 2, and 3): (i) “...sample prep and sample processing were not adequately addressed.” (ii) “... scant attention is paid to sample processing.” (iii) “While this application has excellent and detailed discussion of resistance variation and probe selection, it lacks discussion on sample prep and how it will be performed at POC.” We have added the following to Specific Aim 2: “Our eventual approach to sample processing will involve using disposable cartridges pre-loaded with reagents for sample lysis and RNA extraction/purification based on established magnetic particle chemistry. Such a cartridge is being developed by InSilixa in collaboration with researchers specializing in sample processing for POC applications.” Sample processing is a generic challenge not specific for HIV. This proposal therefore focuses on the HIV-specific challenge of detecting mutations despite surrounding genetic variability.

3e. Virus load and PCR contamination (Reviewer 3): “It would be more valuable if resistance testing could be multiplexed with a quantitative multi-clade HIV assay for viral load … complexities and limitations for LLLOD … are not discussed … conventional PCR is not considered appropriate for POC testing due to complexity and contamination issues.” We clarify that if the aims are successfully addressed, “the presence or absence of probe hybridization can be used to define a reproducible threshold for HIV-1 detection (i.e., a qualitative measure of VL).” Regarding PCR complexity and contamination, we note in the Research Plan (first paragraph): “The closed-tube format leverages the biochip’s capability of performing both PCR and melting without the need for a wash step and the risk of PCR contamination.”
RESEARCH STRATEGY

(a) SIGNIFICANCE

The Joint UN Program on HIV/AIDS aims for 90% of HIV-infected persons to be diagnosed, for 90% of those diagnosed to be receiving antiretroviral (ARV) therapy (ART), and for 90% of those on ART to be virologically suppressed as part of an ambitious public health program to end the AIDS epidemic by 2030\(^5\). Although successful implementation of these recommendations will reduce HIV-1 incidence, the dramatic increase in ART from \(\sim\)20 million to \(\sim\)32 million persons will increase the prevalence of acquired drug resistance (DR) in treated persons and transmitted DR in newly infected persons. Indeed, between 10% and 15% of persons receiving a first-line NRTI/NNRTI regimen develop virological failure (VF) during their first year of therapy and most persons with VF acquire NRTI- and/or NNRTI-resistance\(^6-12\). Moreover, as the number of persons with acquired DR has increased so has the proportion of newly infected persons with transmitted DR\(^2,5,13\).

Recent surveillance studies have reported that pretreatment NNRTI resistance (defined as either transmitted NNRTI resistance or NNRTI resistance in people with prior ART exposure initiating or reinitiating first-line ARV therapy) is \(\geq\)10% in many countries\(^2,14-18\). As a result, the WHO recommends that alternative first-line regimens including dolutegravir (DTG) + tenofovir (TDF) + 3TC/FTC (XTC) be considered for all treatment initiators in regions with pretreatment NNRTI resistance \(\geq\)10%. In regions where it is not feasible to change first-line ART for all initiators, the WHO recommends considering pretreatment genotypic resistance testing (GRT) to identify which patients should receive an alternate first-line regimen. However, in LMICs, the resources and capacity to perform GRT for patient management are limited or concentrated in a few central laboratories.

An inexpensive GRT that could be performed during a clinic visit, heretofore called a point-of-care (POC) assay, would expand access to GRT and avoid the challenges and delays associated with centralized GRT. Even in the context of a public health approach to ART, where a limited number of standard regimens are available, a reliable POC GRT would enable HIV care providers to make informed treatment decisions for ART-naïve patients starting ART and for patients with VF on an initial NRTI/NNRTI regimen. A POC GRT would be particularly useful in conjunction with a POC virus load (VL) test. Through the delivery of immediately actionable information upon the detection of VF, a POC GRT would benefit treatment programs by reducing treatment delivery costs and would benefit patients by reducing travel costs and work absenteeism and by increasing their likelihood of achieving and maintaining virus suppression\(^19,20\).

In 2015, we reported that an assay that included exactly 6 DRMs including K65R, K103N, V106M, Y181C, M184V, and G190A was \(\sim\)99% sensitive for detecting intermediate / high-level acquired NRTI or NNRTI resistance to any ARV (including AZT and NVP) in patients with VF on a first-line WHO-recommended NRTI/NNRTI regimen and \(\sim\)60% sensitive for detecting intermediate / high-level transmitted NRTI or NNRTI resistance to any ARV\(^13,21\). In an analysis specific for a first-line TDF/XTC/EFV-containing regimen\(^22\), an assay that includes each of the most common variants at these 6 positions (K65R/N and M184V/I for the NRTIs; K103N/S, V106A/M, Y181C/I/V, and G190A/S for the NNRTIs) would be 90 to 95% sensitive for detecting intermediate / high-level acquired resistance to TDF or XTC and 90 to 95% sensitive for detecting intermediate / high-level acquired EFV resistance. The assay would also be 90 to 95% sensitive for detecting intermediate / high-level transmitted resistance to TDF or XTC and 90 to 95% sensitive for intermediate / high-level transmitted EFV resistance.

This proposal presents preliminary data demonstrating the reliability of a novel solid-phase nucleic acid melt-analysis platform for distinguishing between multiple codons at a single HIV-1 DR position despite surrounding nucleic acid variability. We outline a plan to build on these data to detect the 10 most common codons at NRTI DRM positions 65 and 184 and the 25 most common codons at NNRTI DRM positions 103, 106, 181, and 190.

(b) INNOVATION

The development of a POC HIV-1 GRT addresses the challenge of detecting mutations in rapidly evolving epidemic viruses whether they are DRMs or critical vaccine-escape or gain-of-function mutations. The main challenge in developing such an assay is the genetic variability of HIV-1 at and surrounding each DRM. Several published HIV-1 DRM point-mutation assays – the oligonucleotide ligation assay (OLA) developed by the Frenkel laboratory\(^23,24\), the multiplex allele-specific primer extension assay developed by the CDC\(^25,26\), and the PANDAA assay developed by Aldatu\(^27\) – employ a variety of approaches to address sequence variability such as using different primer sets for different geographic regions and using an additional enzymatic reaction (e.g., ligation and sequence extension) to improve specificity. However, each involves multiple steps and
requires benchtop equipment and trained laboratory technicians. A recent study showed that melt curve analysis using TaqMan probes can also accurately detect DR variants at codons 103, 181, and 184. However, this approach cannot be highly multiplexed and also requires benchtop equipment and trained laboratory technicians\textsuperscript{26}. A phenotypic approach to detecting RTI resistance avoids the problem of sequence variability but it relies on extracellular RT inhibition and is not applicable to other drug classes\textsuperscript{29}.

This proposal addresses two of the main challenges to developing a closed-tube GRT that can be performed on a single POC device: (1) Creating parsimonious probe sets capable of hybridizing to the vast majority of HIV-1 variants and detecting each of the reported codons at 6 key DRM positions and (2) Developing a multiplex asymmetric PCR assay comprising 3 primer pairs capable of generating sufficient single-stranded DNA (ssDNA) for hybridization and detection. To address these challenges, we are collaborating with the biotechnology company, InSilixa, which has developed a dynamic microarray platform that offers a closed tube solution to nucleic acid amplification; capture hybridization; and mutation detection by solid-phase melt-curve analysis. Combined with a disposable sample processing cartridge, GRT using this platform would be affordable, sensitive, specific, rapid, user-friendly, and deliverable to remote locations\textsuperscript{30}.

The InSilixa technology merges semiconductor-based integrated circuits with nucleic acid detection to create fully integrated complementary metal-oxide semiconductor (CMOS) biosensors that detect hybridization events at individual features\textsuperscript{31}. The current version of this technology, the HYDRA-1K biochip, comprises a 32 x 32 microarray with 1024 features for highly multiplexed melt-analysis. Every feature includes a high-performance fluorescence photo-sensor, a low-noise and high dynamic range sensor circuitry, a thermo-cycling device, and a digital signal analysis component. Synthetic oligonucleotide capture probes are covalently attached to each feature. Viral RNA is RT-PCR amplified in a closed tube fluidic reaction chamber that surmounts the 32 x 32 feature array. The resulting amplicons then hybridize to their complementary capture probes in the same reaction chamber. On-chip solid-phase melt analysis then discriminates between variants in the hybridized amplicons. It is the integration of signal detection into the CMOS chip that obviates the need for a fluorescent microscope and thus makes this technology POC-compatible. Stanford Professor Emeritus Gary Schoolnik, InSilixa Chief Medical Officer, has used the HYDRA-1K technology to develop a POC assay capable of detecting 88 nucleotide variants associated with DR in 6 \textit{M. tuberculosis} genes (manuscript in preparation). The HYDRA-1K biochip has a cost-of-goods of $15 per assay if chips are mass produced and a potential cost of ~$26 per assay when an integrated RNA-extraction sample processing cartridge is included.

It is possible to perform real-time PCR on the HYDRA-1K platform\textsuperscript{32,33} providing the possibility of performing both GRT and VL testing on the same platform. However, two POC VL tests are CE-IVD approved and others are in development\textsuperscript{34,35}. This proposal aims to develop a GRT, which has not received attention from industry and yet will be increasingly needed as ART coverage and VL testing increase in LMICs\textsuperscript{20}. However, if the aims of this proposal are successfully addressed, the presence or absence of probe hybridization can be used to define a reproducible threshold for HIV-1 detection (i.e., a qualitative measure of VL).

\textbf{(c) RESEARCH PLAN}

This proposal addresses the two main challenges to developing a closed-tube GRT that can be performed on a POC device. The first is to create probe sets capable of hybridizing to the vast majority of HIV-1 group M variants and of distinguishing between the most commonly reported codons at 6 key DRM positions. Part of this challenge involves determining the sensitivity and specificity for detecting codons present in mixed populations. The second is to develop a multiplex asymmetric PCR assay to generate sufficient single-stranded DNA (ssDNA) for capture hybridization for the 3 regions encompassing codon 65, codons 103 and 106, and codons 181, 184, and 190. This challenge constitutes its own aim because it is a nontrivial essential step for performing GRT in a closed-tube format. The closed-tube format leverages the biochip’s capability of performing both PCR and melting without the need for a wash step and the risk of PCR contamination.

\textbf{Specific Aim 1: To develop parsimonious probe sets capable of genotyping >99% of HIV-1 group M variants at 6 DRM positions including the NRTI DRM positions 65 and 184 and the NNRTI DRM positions 103, 106, 181, and 190.} \textbf{Overview:} Table 1 lists 35 wild type and mutant codons at positions 65, 103, 106, 181, 184 and 190 that account for ~99.5% of all reported codons at these positions in QC-filtered plasma virus sequences from LMICs (n=63,000 persons) and upper-income countries (n=69,000 persons) of whom 54\% were ARV-naïve, 30\% were ARV-experienced, and 16\% had uncertain ARV histories (“HIV-1 Genetic Variability for Drug Resistance” section of the Stanford HIV Drug Resistance Database). Experiments for this aim are designed to create parsimonious probe sets capable of detecting each of these 35 codons and
determining the sensitivity and specificity for each in a mixed population. We will use hybridization simulations to identify capture probe sets predicted to hybridize >99% of sequence variants surrounding each codon and to distinguish between each codon at these 6 positions. We will then test all probe sets in vitro on a solid-phase, highly multiplexed, nucleic acid melt-curve analysis platform using 80-mer targets representing diverse viral variants and containing these 35 codons. In Specific Aim 2, we will evaluate probe hybridization and discrimination following multiplexed asymmetric RT-PCR of HIV-1 RNA from diverse patient samples and culture supernatants.

Preliminary studies: We performed a series of experiments to determine the feasibility of microarray-based hybridization and melt-curve analysis to distinguish between the 6 reported codons at RT position 103: AAA, AAG, AAC, AAT, AGA, AGC. We used VisualOMP (DNA Software, MI) to identify parsimonious probe sets such that ≥1 member of one probe set was predicted to have a Tm ≥55°C to >99% of LMIC RT sequences and such that mismatching codons have a ΔTm ≥2°C. This analysis yielded 3 sets of 6 probes. Each probe set had a different 31 bp flanking sequence (Table 2). Each probe in a set was complementary to one of the 6 codon 103 variants. Flanking sequence of probe set 1 was complementary to the most common sequence surrounding codon 103. Flanking sequences of probe sets 2 and 3 were common variants identified from a weighted codon 103 region alignment.

To evaluate the performance of the 3 probe sets comprising 18 probes in vitro, we synthesized each probe and covalently printed it in triplicate onto a functionalized 1x3" silicon substrate slide. We then synthesized 18 80-mer Cy3-labeled targets that exactly matched one of the probes in its 34-bp probe-binding region but differed from the remaining probes either at its central codon location or flanking sequence. Targets were diluted to 100 nM and, in separate experiments, were injected onto the capped slide. We then used a hybridization/melting protocol involving (i) hybridization for 1 hour at 45°C; (ii) washing of unhybridized target; and (iii) a 20-minute melting period as temperature was raised to 90°C during which serial images were taken every 10 seconds.

We performed 81 preliminary experiments: (i) Each of the 18 targets (6 codons x 3 flanking sequences) was tested in triplicate (n=54 experiments); and (ii) Each of the 3 targets with AAC was diluted to 50%, 25%, and 10% (with targets containing AAA) and also tested in triplicate (n=27 experiments). Figures 1A 1D show melt curves for 4 of the 54 experiments with a single undiluted target and Figures 2A and 2B show melt curves for 2 of the 27 experiments in which 80-mer targets with AAC were diluted with AAA targets.

Figure 1. Melt-curve experiments using 80-mer targets with nucleotides encoding 4 variants at position 103 in the context of flanking sequence 1: the wild type variants AAA (K) and AAG (K) and the DRMs AAC (N) and AAT (N). The 3rd replicate from each experiment is shown.

| Table 1. Observed DRMs (in bold) and wild type variants at 6 RT DRM positions |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| K65 | AAA(K), AAG(K), AGA(R), AGG(R), AAT(N), AAC(N) |
| K103 | AAA(K), AAG(K), AGA(R), AAC(N), AAT(N), AGC(S) |
| V106 | GTA(V), GTG(V), ATA(I), ATG(M), GCA(A) |
| Y181 | TAT(Y), TAC(Y), TGT(C), TGC(C), ATT(I), GTT(V) |
| M184 | ATG(M), GTG(V), GTA(V), ATA(I) |
| G190 | GGA(G), GGG(G), GGC(G), GCA(A), GCG(A), GCC(A), AGC(S), AGT(S) |

| Table 2. Probe sets used to distinguish between 6 codons at RT 103 |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Probe Set | 1 | 2 | 3 |
| Probe Set 1 | TAC | TAC | CAT |
| Probe Set 2 | TGT | TAC | TGG |
| Probe Set 3 | TTA | TTT | TTA |
| Each probe set has 6 variants TTT, GTT, CTT, ATT, TCT, and GCT (indicated by NYT) complementary to AAA, AAC, AAG, AAT, AGA, and AGC. |
For each experiment, we computed 3 \( T_{m_{\text{max}}} \)'s defined as the highest \( T_m \) within a probe set and 18 \( \Delta T_{m_{\text{max}}} \)'s defined as the difference between the \( T_{m_{\text{max}}} \) for a probe set and the \( T_m \) of a probe within that set (Table 3). We then calculated a mean \( \Delta T_{m_{\text{max}}} \) for the 2 probe sets that did not contain an exact match with the 80-mer target. Thus, if the target had the codon AAT and flanking sequence 1, the \( \Delta T_{m_{\text{max}}} \)'s were the mean of the 2 probe sets with flanking sequences 2 and 3.

Figure 3 summarizes the 486 mean \( \Delta T_{m_{\text{max}}} \)'s as defined in Table 3 (81 experiments x 6 codons with each result representing the mean \( \Delta T_{m_{\text{max}}} \) from the 2 non-exactly matching probe sets) according to the \% of a codon in experimental targets (100\% for undiluted targets; 50\%, 25\%, 10\% for diluted AAC targets; and 50\%, 75\%, 90\% for AAA targets used to dilute AAC). These experiments demonstrate little overlap between the \( \Delta T_{m_{\text{max}}} \)'s generated by probes not complementary to a target's codon (i.e., median of the mean \( \Delta T_{m_{\text{max}}} \)'s = 6.1\text{oC}) compared with probes complementary to a target codon present at \( \geq 25\% \) (median mean \( \Delta T_{m_{\text{max}}} < 0.3\text{oC} \)). Overall, a mean \( \Delta T_{m_{\text{max}}} \) of \( \leq 1.5\text{oC} \) was 98\% sensitive and 99\% specific for detecting variants present at a level \( \geq 25\% \) (Figure 3 ROC curve insert). The \( \Delta T_{m_{\text{max}}} \)'s for codons present at 10\% (median mean \( \Delta T_{m_{\text{max}}} = 2.3\text{oC} \)) were significantly lower than the \( \Delta T_{m_{\text{max}}} \)'s for absent codons. However, additional probes and analytical approaches are needed to consistently detect mutations present at levels of 10\% without increasing the risk of false positive results.

**Table 3. Metrics used to assess the results of a melting-curve experiment**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_m )</td>
<td>Temperature at which 50% of the target is hybridized to a probe.</td>
</tr>
<tr>
<td>( T_{m_{\text{max}}} )</td>
<td>Highest ( T_m ) within a set of the 6 probes (one for each K103 codon) with the same flanking sequence</td>
</tr>
<tr>
<td>( \Delta T_{m_{\text{max}}} )</td>
<td>Difference between the ( T_m ) of an individual probe and the ( T_{m_{\text{max}}} ) within the set of probes sharing the same flanking sequence. By definition, at least 1 probe in a set will have a ( \Delta T_{m_{\text{max}}} ) of 0\text{oC}.</td>
</tr>
<tr>
<td>Mean ( \Delta T_{m_{\text{max}}} )</td>
<td>Mean ( \Delta T_{m_{\text{max}}} ) for the 2 probe sets that do not contain an exact match with the 80-mer target. With unknown targets, the mean ( \Delta T_{m_{\text{max}}} ) would be the mean of all probe sets for a DRM position.</td>
</tr>
</tbody>
</table>

**Figure 3. Distribution of \( \Delta T_{m_{\text{max}}} \) according to \% of codon in the experimental targets.**

Continued probe design, testing, and optimization: For the remaining 29 codons at the NRTI positions 65 and 184 and the NNRTI positions 106, 181, and 190, we will identify, test, and iteratively optimize capture probe sets using in silico and in vitro analyses as described in our preliminary studies. We will perform hybridization simulations to identify parsimonious sets of 30 to 35 bp probes predicted to have a \( T_m \geq 55\text{oC} \) and a \( \Delta T_m \geq 2\text{oC} \) for mismatched codons. Probe sets will be printed onto a silicon substrate slide and evaluated as described above using synthesized 80-mer targets representing diverse viral variants. We plan to test 5 to 10 targets per codon (therefore, several additional targets per codon will be required for position 103). Targets will be highly diverse differing from the probes' \(~30\) bp flanking regions by up to 4-5 bp as it is unusual for \( pol \) sequences to differ from each other by >5 bp over a \(~30\) bp span.)
As we test the initially designed probe sets, we will generate data that will enable us to improve the accuracy at detecting codons with divergent flanking sequences and to improve the sensitivity for minority variant codons. A combination of probe sets will be considered successful if undiluted targets hybridize most strongly (i.e., ΔTm,max = 0) to the complimentary probe in each probe set and if the combination accurately identifies minority variants at levels down to 10% to 25%. As codons differ in their hybridization kinetics, ROC curves will be generated to identify threshold mean ΔTm,max’s for each codon.

If the initial combination of printed probe sets does not unequivocally identify majority variants or is not highly sensitive and specific for identifying minority variants at levels between 10% and 25% (approximating Sanger sequencing), we will design additional probe sets as each probe set provides independent information as to whether a target sequence contains a particular codon. Additional approaches will include designing probes with (i) variant bases in their flanking regions such as diaminopurine, which increases Tm when base-paired with T³⁷ and the “universal base” deoxyinosine³⁸ to improve hybridization to divergent sequences; (ii) diaminopurine in codon binding regions to improve the detection of codons that bind weakly; and (iii) locked nucleic acid (LNAs), which though more expensive can be useful in special cases to improve hybridization efficiency and sequence discrimination depending on their placement³⁹.

**Analysis:** In our preliminary studies, we discriminated between codons based entirely upon the ΔTm,max’s of each probe. In subsequent studies, we will attempt to improve the ability to detect minority variants by using two additional metrics: (i) AUC, the differences in the area under two melt curves and (ii) ΔTm,’ the difference between the peaks of the derivative of two melt curves. In addition, the difference between ΔTm,max (or AUC and ΔTm,’) for the probes in a set with the 2nd and 3rd highest Tm may be helpful for distinguishing minority variants from absent codons because it is highly unusual for >2 variants to be present at the same RT position at levels >5% even by deep sequencing⁴⁰,⁴¹. Indeed, Figure 2B shows that targets with an AAA and AAC mixture hybridize much more strongly to probes with these codons than to the third best binding probe.

The experiments outlined above will determine how many different DRMs can be detected on a 1024-feature microarray on which each probe set is printed in triplicate. For example, if an average of 4 probe sets per codon are required, then 420 (35 codons x 4 probe sets x 3 replicates) of the microarray’s 1024 features would be needed. This would leave enough features for (i) gridding, process-management, and signal normalization (~25 features); (ii) potentially additional RT mutations to further increase sensitivity and specificity for detecting transmitted and acquired NRTI and/or NNRTI resistance; and (iii) eventually protease- and/or integrase-inhibitor DRMs. Should new chip versions with >1024 features be developed this would further improve the prospects for detecting multiple additional DRMs.

**Rigor and reproducibility:** All experiments will be done using a microarray setup that has been validated by InSilixa to predict performance on the CMOS chip-based HYDRA-1K POC platform. HPLC-purified probes and fluorophore-labelled targets will be purchased from established vendors. Signal intensities from serial microarray images are permanently stored and the R scripts we use for analysis will be available upon request.

**Specific Aim 2: To develop a multiplex asymmetric RT-PCR protocol that will generate sufficient amounts of single-stranded DNA for microarray hybridization. Overview:** To develop an eventual “sample-in answer-out” assay, it is necessary to develop a closed tube PCR protocol to generate sufficient amounts of ssDNA encompassing codons 65, codons 103 and 106, and codons 181, 184, and 190 from samples with ≥1000 HIV-1 RNA copies. Separate PCRs for each region are needed because probe binding on InSilixa’s melt-curve analysis platform requires amplicons of ~100 to 200 bp. The first step in this process is to identify ranked sets of primer pairs through in silico hybridization simulations. The second step is to calculate the efficiency of candidate primer pairs by generating and analyzing standard curves. Each primer pair will then be used in an asymmetric PCR and the resulting ssDNA will be quantified using a droplet digital PCR approach developed at InSilixa. Finally selected primers will be evaluated in a multiplex asymmetric PCR protocol using pedigreed culture supernatants and plasma virus samples containing known virus levels.

The multiplex PCR protocol will be first performed using benchtop PCR machines and its products will be individually assessed in TaqMan assays using real-time qPCR. Once we identify a protocol that produces sufficient ssDNA product, we will evaluate its performance on our microarray setup. For this step, we will include the fluorophore on our PCR primers and will continue to include a wash step to remove unhybridized PCR product. In future studies, our probes will contain fluorophores and our PCR products will contain
quenchers so that PCR, hybridization, and melt-curve analysis can be performed in a single tube without a wash step. The increased cost of probes with fluorophores precludes their use for this project.

**Preliminary in silico studies and singleplex PCR evaluation:** We used PrimerDesign-M, a multiple-alignment based multiple-primer design tool for walking across variable genomes to identify primers predicted to amplify diverse HIV-1 strains encompassing each region of interest. We then used VisualOMP to select primers taking into account HIV-1 RNA secondary structure and sequence variability. We identified 18 primer pairs yielding amplicons ranging in size between 91 and 213 bp and evaluated their efficiencies in qPCRs using serial dilutions (10 to 100,000 copies/µL) of one of three 301 bp DS-DNA blocks encompassing each region of interest. Six primer pairs (2 pairs per region of interest) with the highest efficiency and r-squared values have been identified for developing the multiplex asymmetric RT-PCR protocol. We will apply InSilixa’s droplet digital PCR approach to quantifying ssDNA generated in each singleplex asymmetric PCR. The primer pairs that produce the highest amount of ssDNA product will then be used for multiplex asymmetric RT-PCR.

**Samples and sample processing:** We will use 3 sources of virus to develop the multiplex PCR protocol: (i) 30 cryopreserved plasma HIV-1 samples containing viruses belonging to subtypes A, C, CRF01_AE, and CRF02_AG, of which 8 have the DRMs K103N and/or M184V. The VL of these samples has been determined by commercially available VL assays; (ii) 24 multiply NRTI- and NNRTI-resistant subtype B plasmid clones that contain 17 of the 35 codons at positions 65, 103, 106, 181, 184, and 190; and (ii) ~50 EQAPOL global diversity panel culture supernatants (Duke University) containing diverse HIV-1 subtypes. We will dilute culture supernatants ~100-fold and then adapt a real-time RT-PCR assay to quantify virus in these samples.

Our eventual approach to sample processing will involve using disposable cartridges pre-loaded with reagents for sample lysis and RNA extraction/purification based on established magnetic particle chemistry. Such a cartridge is being developed by InSilixa in collaboration with researchers specializing in sample processing for POC applications. Until such cartridges are available, we plan to use standard benchtop methods for extracting RNA from culture supernatants and plasma samples before submitting it to multiplex asymmetric PCR.

**Multiplex PCR optimization.** To evaluate the performance of a multiplex PCR protocol, we will make serial 10-fold dilutions of PCR product beginning with 1/100, 1/1000, or further after 20, 25, 30, and 35 PCR cycles. We will then quantify each of the amplicons at each cycle in separate PCRs for each primer pair using real-time TaqMan qPCR to characterize multiplex PCR efficiency. In addition, absolute quantification of ssDNA will be repeated for the multiplex PCR. PCR products will also be injected onto our dynamic microarray and undergo hybridization for up to 1 hour at 45°C followed by washing and melt-analysis to determine whether sufficient ssDNA PCR product was produced by the multiplexed protocol. The multiplex asymmetric PCR protocol will be iteratively optimized to produce sufficient amounts of ssDNA PCR product in both singleplex and multiplex PCRs. Additional strategies for optimizing asymmetric PCR protocols may be required such as designing limiting primers with a higher T_m than the excess primer.

These experiments will provide feedback on the primers and conditions optimal for producing sufficient single-stranded product during singleplex and multiplex RT-PCRs. We will analyze the relationship between specific RT-PCR primers, PCR conditions and the amounts of ssDNA product determined by time to microarray probe hybridization. We expect that it will be possible to determine a limit of detection for the final optimized protocol such that for a sample with a known RNA copy number, we will be able to predict the probability that sufficient ssDNA will hybridize all three probes. As noted above, our target for this purpose is ≥1000 HIV-1 RNA copies.

**Rigor and reproducibility:** Our multiplex asymmetric PCR protocol will be validated using TaqMan assays, droplet digital PCR, and our microarray setup. We will publish the final protocol in a peer-reviewed journal.

**Future studies.** If the aims of this proposal are met, we will plan future collaborations with InSilixa that will (i) evaluate sample processing modules to separate plasma from blood, concentrate virus, and extract RNA that can be delivered to the fluidic chamber of the HYDRA-1K biochip; (ii) test a wide range of clinical samples; (iii) transfer the optimized probe sets onto the InSilixa CMOS chip-based HYDRA-1K POC platform; (iv) determine whether the presence or absence of probe hybridization can be used to define a reproducible threshold for detecting HIV-1 (i.e., a qualitative measure of VL); and (v) expand the assay to include protease inhibitor and integrase inhibitor DRMs.