Low-frequency HIV-1 drug resistance mutations can be clinically significant but must be interpreted with caution

Jeffrey A. Johnson1* and Anna Maria Geretti2

1Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA; 2Department of Virology, Royal Free Hampstead NHS Trust and UCL Medical School, London, UK

With drug-resistant HIV-1 present in at least 10%–20% of new infections in Western countries and in >60% of patients failing antiretroviral therapy (ART), monitoring HIV-1 drug resistance is becoming increasingly important for assessing its impact on therapeutic measures of virus control and for guiding treatment. The sensitivity limitations of conventional bulk genotyping often lead to an underestimation of the total burden of drug resistance in a patient, as resistant variants escape detection when present at low frequency within the viral quasispecies. Using sensitive resistance testing methods, a few investigators have linked low-frequency mutations to poor treatment outcomes, while other studies have shown no correlation. Understanding the technical limitations of sensitive testing methods and the relevance of the amount of a particular resistance mutation in the context of different ART regimens will help to define the clinical benefit of low-frequency resistance testing. Paramount to interpreting the clinical utility of sensitive testing is evaluating resistance mutations selectively, at biologically significant frequencies, and using methods that have been broadly validated on clinical specimens.

Keywords: HIV, drug resistance, low-frequency mutants, testing sensitivity

Introduction

As the era of antiretroviral therapy (ART) lengthens and more people gain access to treatment, HIV infections involving drug-resistant mutants are becoming more prevalent. The conventional method for HIV drug resistance testing employs bulk population genotyping of the plasma virus swarm in an infected patient. However, the sensitivity limitation of the method is such that variants present in an individual at levels <20%-30% are masked in this analysis (Figure 1).1 Underestimating the burden of resistance in drug-naive individuals can affect optimal ART selection, particularly for persons diagnosed long after becoming infected, when transmitted resistant mutants may have decayed. Undiagnosed mutants may also affect the treatment of ART-experienced individuals by failing to give a full representation of resistance.2 To improve detection of resistant mutants, various molecular assays have been developed that uncover low-frequency variants typically not detectable using conventional genotyping. However, important factors, primarily technological issues concerning the generation of PCR artefacts and the presence of a natural background of mutations within the HIV-1 quasispecies, must be considered when investigating the clinical significance of low-frequency variants. Separating mutations that reflect enrichment through drug selective pressure from mutations that occur naturally at very low levels is important for characterizing the clinical implications of transmitted drug resistance. Using different criteria to define which mutation frequencies are representative of drug-selected variants can confound interpretation of clinical relevance.

Most studies to date have investigated the clinical relevance of low-frequency mutations in drug-naive patients starting first-line ART and in women receiving single-dose nevirapine for the prevention of mother-to-child transmission. Two important observations can be drawn from these studies. The first is related to the lack of consistent evidence supporting a significant effect of low-frequency resistant variants on responses to first-line ART, with reports describing either a detrimental impact3–5 or the lack of a discernible effect6,7 However, the conflicting studies employed different testing methodologies, applied different sensitivity cut-offs for interpretation and tested heterogeneous populations in terms of duration of infection and composition of the first-line ART regimen. As a consequence, the results cannot be easily compared. A second important observation is that in women receiving single-dose nevirapine, the negative impact of low-frequency drug-resistant mutants on responses to first-line ART appears to decline over time.8,9 This supports the hypothesis that a quantitative threshold of selected resistant virus expression (‘mutational load’) is needed to impact clinical outcomes, at least in the context of first-line regimens based on two nucleoside reverse transcriptase inhibitors (NRTIs) and one non-nucleoside reverse transcriptase inhibitor (NNRTI). Of note, comparisons between transmitted
and selected drug resistance should be made with caution, as transmitted variants would seed resistance in a high proportion of the viruses archived, with potentially long-term effects on treatment outcomes. Hence, the extent to which resistant virus is archived may determine its ability to impact subsequent therapy.

Drug resistance mutations are present below conventional sequence detection

In the absence of drug pressure, drug-resistant variants will fade out of circulation as they become overgrown by more fit wild-type HIV-1. Mutation decay is an issue in the surveillance of transmitted drug resistance because most people newly diagnosed with HIV-1 infection became infected many months or years earlier. Hence, there has been sufficient time for some transmitted mutations to decay, making it difficult to identify resistance that may have established reservoirs at the time of primary infection.

Resistant variants fade at different rates largely as a result of the fitness cost of the mutations they carry. High-fitness-cost mutations that severely impact replicative capacity, such as K65R and M184V in reverse transcriptase, typically cause the variant to decay very rapidly, while low-fitness-cost mutations, such as M41L and K103N, may persist for years in the absence of drug exposure. However, the varying prevalence of low-frequency transmitted resistance mutations is not explainable by fitness cost alone. Some low-fitness-cost mutations, e.g. M41L, seem to persist at minority levels for quite some time, while others such as K103N persist at higher levels, then appear to reach a point where they revert more rapidly than expected. Thus, even though K103N is one of the most prevalent transmitted mutations detected by bulk genotyping, its low-frequency prevalence does not appear to be significantly greater than that of higher-fitness-cost mutations. Moreover, other mutations that are also thought to confer low fitness costs, e.g. Y181C, are not as prevalent at bulk-detectable levels as is K103N, yet low-frequency populations may add substantially to their prevalence (Figure 2). It is unknown whether this reflects resistant variants that are driven to lower levels by population pressure but are able to adapt to their impaired fitness and sustain a low-frequency niche.

Sensitive resistance testing methods have revealed that there is substantial transmitted drug resistance at frequencies below what is detectable by conventional bulk genotyping. In areas where ART is widely available and transmission risk is high, as much as 30%–50% of transmitted resistance mutations can be missed by conventional sequencing (Figure 2). Obtaining a more comprehensive assessment of transmitted drug resistance that takes into account low-frequency variants may help guide treatment in individual patients, while also serving a public health purpose. Because resistance mutations serve as genetic markers, sensitive drug resistance screening can help to identify transmission networks and populations on which to focus prevention measures. Moreover, primary infections involving multiple or high-fitness-cost mutations may reflect transmission from persons who have received treatment and know their infection status. Those resistant viruses might be especially difficult to detect without sensitive screening because their expression could be weak in ART-naive recipients. Thus, sensitive

Drug resistance mutations are present below conventional sequence detection

In the absence of drug pressure, drug-resistant variants will fade out of circulation as they become overgrown by more fit wild-type HIV-1. Mutation decay is an issue in the surveillance of transmitted drug resistance because most people newly diagnosed with HIV-1 infection became infected many months or years earlier. Hence, there has been sufficient time for some transmitted mutations to decay, making it difficult to identify resistance that may have established reservoirs at the time of primary infection.

Resistant variants fade at different rates largely as a result of the fitness cost of the mutations they carry. High-fitness-cost mutations that severely impact replicative capacity, such as K65R and M184V in reverse transcriptase, typically cause the variant to decay very rapidly, while low-fitness-cost mutations, such as M41L and K103N, may persist for years in the absence of drug exposure. However, the varying prevalence of low-frequency transmitted resistance mutations is not explainable by fitness cost alone. Some low-fitness-cost mutations, e.g. M41L, seem to persist at minority levels for quite some time, while others such as K103N persist at higher levels, then appear to reach a point where they revert more rapidly than expected. Thus, even though K103N is one of the most prevalent transmitted mutations detected by bulk genotyping, its low-frequency prevalence does not appear to be significantly greater than that of higher-fitness-cost mutations. Moreover, other mutations that are also thought to confer low fitness costs, e.g. Y181C, are not as prevalent at bulk-detectable levels as is K103N, yet low-frequency populations may add substantially to their prevalence (Figure 2). It is unknown whether this reflects resistant variants that are driven to lower levels by population pressure but are able to adapt to their impaired fitness and sustain a low-frequency niche.

Sensitive resistance testing methods have revealed that there is substantial transmitted drug resistance at frequencies below what is detectable by conventional bulk genotyping. In areas where ART is widely available and transmission risk is high, as much as 30%–50% of transmitted resistance mutations can be missed by conventional sequencing (Figure 2). Obtaining a more comprehensive assessment of transmitted drug resistance that takes into account low-frequency variants may help guide treatment in individual patients, while also serving a public health purpose. Because resistance mutations serve as genetic markers, sensitive drug resistance screening can help to identify transmission networks and populations on which to focus prevention measures. Moreover, primary infections involving multiple or high-fitness-cost mutations may reflect transmission from persons who have received treatment and know their infection status. Those resistant viruses might be especially difficult to detect without sensitive screening because their expression could be weak in ART-naive recipients. Thus, sensitive
resistance testing can be a powerful tool to help both monitor the epidemic and highlight prevention needs. Finally, resistance decay is not only a concern for new infections. For treated patients, changing drug regimens can suppress resistant variants while new strains emerge. Identifying prior resistance that may be persisting at low levels and current resistance that has yet to emerge at high frequency could benefit the patient long term by guiding optimization of therapy.

**Sensitive drug resistance testing in drug-naive patients requires careful interpretation**

Several investigators have developed molecular assays to identify low-frequency resistance mutations, and each method has its advantages and disadvantages. Some of the reported methods include oligoligase assay (OLA), LigAMP, allele-specific or mutation-specific real-time PCR, parallel allele-specific sequencing (PASS), ultra-deep sequencing and single-genome sequencing (SGS), each with differing levels of technical complexity and manipulation. Some assays have the sensitivity to detect 1-in-10000 to 1-in-100 000 cloned mutant viruses in a wild-type background. Whatever the method used, when screening for transmitted drug resistance for individual patient care or population-based surveys, two key factors need to be considered: (i) the natural background prevalence of the mutation in the HIV-1 quasispecies; and (ii) the frequency of PCR artefacts if amplifying the resistance-associated codon sequence for mutation detection. Because some testing methods are sufficiently sensitive to detect just a few mutant copies in a wild-type background, defining the mutation background is critical for distinguishing variants present above threshold levels that signify a prior enrichment by drug selection.

What has not been elucidated for the observed ‘natural’ mutation frequencies is whether PCR amplification, particularly in poly(A) tracts or in sequences adjacent to GC-rich structures, artificially introduces mutations in a manner similar to what can occur in these regions during in vivo replication. An approach taken by Johnson et al. sought to address both issues by evaluating specimens that had been collected from HIV-1-infected patients in the early 1980s, before the introduction of antiretrovirals. The use of clinical samples dating back to the pre-ART era eliminated the possibility of detecting transmitted resistance and allowed for measurements of ‘background’ assay reactivity reflecting the natural quasispecies. Furthermore, any reactivity due to artefactual mutations generated by PCR amplification was accounted for by setting the interpretative cut-off for each mutation above the spectrum of reactivity seen in all early HIV specimens tested. This exercise demonstrated that mutations vary in background frequency, with some of the lower-fitness-cost mutations appearing to sustain higher natural levels than variants carrying more deleterious resistance mutations. Johnson et al. found also that even assays that exhibited very low background reactivity on cloned specimens occasionally detected mutations at frequencies as high as 0.3% in cross-sectional testing of the early HIV samples from the pre-ART era.

Mutations associated with defective variants in the natural quasispecies swarm or that are generated artefacts would not be expected to impart a clinical impact. Indeed, when Johnson et al. re-analysed a previously reported case–control study using detection cut-offs for transmitted K103N lower than the 0.9% established on pre-ART era HIV specimens, the previously detected association between low-frequency K103N and poor outcome was no longer significant (P calculated using Fisher’s exact test).

**Implications of sensitive resistance testing for clinical practice**

From a general perspective, multiple factors are likely to influence the impact of resistant mutants on the outcomes of ART, including the frequency of the mutant in the quasispecies, but also its phenotypic and fitness effects, co-existence of other mutations, the mutational load, the level of adherence and its phenotypic and fitness effects, co-existence of other mutations, the mutational load, the level of adherence and the composition of the ART regimen. There has been significant progress in recent years in developing an improved understanding of the interplay between these multiple determinants, but much remains to be clarified.

Recent studies have indicated that when using a consistent approach in terms of selection of the study population and use of interpretative cut-offs, low-frequency resistant mutants negatively affect responses to first-line ART based on NNRTIs in combination with NRTIs. In a case–control study by Johnson et al., 7/95 (7%) persons who experienced virological failure on first-line efavirenz-based ART had low-frequency drug resistance mutations at baseline; however, low-frequency resistance was found in only 2/221 (0.9%) treatment successes (Fisher’s exact test, P = 0.0038). In a study by Geretti et al., the NNRTI mutation K103N was detected at low frequency (interpretative cut-off
0.9%) in 4/18 patients who experienced virological failure on first-line NNRTI-based ART and 0/75 patients who maintained virological suppression on similar regimens, and low-frequency K103N was significantly associated with the odds of virological failure (Fisher’s exact test \( P = 0.001 \)). The consistency of the findings in these two studies can be explained by the use of identical real-time PCR methodology and interpretative cut-offs, and a common focus on NNRTI mutations in patients receiving NNRTI-based ART. A preliminary analysis of a study by Goodman et al.\(^2\) using a different real-time PCR methodology similarly found that detection of K103N at a frequency >2% (or 2000 copies/mL) but below the detection limit of bulk genotyping was significantly associated with virological failure of first-line NNRTI-based ART. A fourth study from Simen et al.\(^3\) employed ultra-deep sequencing with a lower limit of detection of 1% to assess resistance in patients starting ART within the Flexible Initial Retrovirus Suppressive Therapies (FIRST) study, which compared three initial ART strategies. Virological failure was significantly associated with the presence of NNRTI resistance detected by both bulk genotyping and ultra-deep sequencing [hazard ratio (HR) 12.40; 95% confidence interval (CI) 3.41–45.10] and by resistance detected only by ultra-deep sequencing (HR 2.50; 95% CI 1.17–5.36), albeit the association was less strong for the latter. Taken together, these findings clearly indicate that NNRTI-based regimens are vulnerable to the presence of transmitted drug resistance, that NNRTI-resistant variants may retain their impact even if present at low frequency within the quasispecies, and that the threshold frequency for clinical significance, at least in the case of K103N, appears to be in the range 0.9%–2%.

There is less evidence on the clinical impact of mutations conferring resistance to the NRTIs, including M184V and K65R, which confer resistance to emtricitabine/tenofovir, one of the most common NRTI backbones in current clinical use. It may be anticipated that the two mutations negatively affect NNRTI-based regimens but have more limited impact on regimens based on rilronavir-boosted protease inhibitors. Even in the context of NNRTI-based regimens, the fitness cost of the two mutations may lead to a less evident impact on virological suppression than that seen with K103N. Data are needed to support this hypothesis.

Implementing sensitive resistance testing in routine diagnostic practice

Despite the growing body of supportive evidence, implementation of sensitive drug resistance testing in clinical practice remains currently limited. Real-time PCR assays demand stringent conditions to ensure accuracy and are vulnerable to sequence variation, requiring subtype-specific optimization. They target one mutation at a time, meaning that reagent cost and labour can become significant if multiple mutations are targeted. While it may be tempting to adopt a simplified approach to real-time PCR testing, confirmation of findings is essential if results are used to guide clinical practice, or to inform population-based surveys, which adds to the complexity of testing. In the studies by Johnson et al.\(^2\) and Geretti et al.\(^1\) for instance, the presence of mutations detected only by real-time PCR was confirmed by sequencing of both the products of the primary reverse transcription PCR (RT–PCR) and the mutation-positive amplicons derived from the real-time PCR. The significant success of first-line NNRTI-based ART among patients who show no evidence of resistance by conventional genotyping clearly indicates that the issue of low-frequency resistance mutants currently has relevance for a small, albeit significant minority of the treated population. This indicates that cost-effectiveness analyses may be required to drive the adoption of sensitive testing methodologies in routine diagnostic settings. Nonetheless, available evidence clearly indicates that further treatment optimization can be achieved in patients due to start NNRTI-based ART by applying a more sensitive assessment of drug resistance. One possible algorithm would screen patients due to start NNRTI-based ART by sensitive detection of NNRTI-resistant mutants, thus identifying quickly those that require further resistance testing, while avoiding the costs of routine genotyping in others. The same practice could be applied to screening for the more commonly selected mutations (e.g., M184V) predicted to impact standard regimens.

While, presently, the cost of ultra-deep sequencing is prohibitive for large population settings it is an attractive option for the not-too-distant future. However, methodological complexity remains significant, including the need to optimize the reading (including the production of longer sequences) and interpretation of the many sequences obtained for each sample. Nonetheless, allowing a deeper look into quasispecies swarms and expanding identification of resistant variants, sensitive testing will improve our understanding of drug resistance mutation perversiveness and clinical influence. In time, this could help determine future best clinical practices for resistance genotyping.

One interesting observation made with both real-time PCR and ultra-deep sequencing is that patients often have K103N as the sole detectable transmitted variant, in the absence of other mutations. This observation leads to a final consideration related to the future of NNRTI therapy. Second-generation NNRTIs such as etravirine or rilpivirine have an improved genetic barrier and may prove more resilient to the problem of transmitted drug resistance than nevirapine or efavirenz, reducing the need for sensitive testing of particular NNRTI mutations (e.g., K103N). However, other NNRTI mutations could still be detrimental and warrant sensitive screening. Studies are ongoing to address this question.

Conclusions

Drug-resistant variants of HIV-1 at levels not detectable by conventional genotyping are common in areas of prolonged ART utilization. Low-frequency resistance can be a result of transmitted drug resistance that has decayed, or remnants of earlier drug selection in persons previously exposed to ART. For at least one drug class, the NNRTIs, the presence of low-frequency transmitted resistance has been shown to have a negative clinical impact on NNRTI-based ART when NNRTIs were the only other drugs in the combination. Even though ART based on the NNRTI efavirenz plus two NRTIs remains the standard of care for first-line treatment worldwide, there is sufficient evidence to indicate that a small but significant proportion of patients will fail therapy as a result of resistance that escaped detection by conventional genotyping. Interpretation of low-frequency NNRTI resistance, especially K103N, but also Y181C and G190A,
can be done with sufficient confidence based upon available data. While the practicalities of using sensitive methodologies to universally screen drug-naïve persons are cumbersome, they are not insurmountable, particularly if cost and options for streamlined approaches to testing prove beneficial. The purpose and feasibility of the individual resistance detection method will dictate which one will be the most logistically favourable.

The contribution of sensitive resistance testing to clinical practice is contingent on its ability to provide comprehensive resistance information in the realm of what is relevant to available treatment options. Pilot studies offering a reliable real-time PCR targeting K103N could be the first step towards establishing clinical applicability. Further studies are needed to define the impact of low-frequency resistant mutants in treatment-naïve patients starting second-generation NNRTIs and in treatment-experienced patients.

**References**


22. Goodman DD, Margo NA, McColl DJ et al. Pre-existing low-levels of the K103N HIV-1 RT mutation above a threshold are associated with virological failure in treatment-naïve patients undergoing EFV-containing antiretroviral treatment. Antivir Ther 2009; 14: S43.