Prevalence of Genotypic Resistance to Antiretroviral Drugs in Treatment-Naive Youths Infected with Diverse HIV Type 1 Subtypes and Recombinant Forms in Dar es Salaam, Tanzania

F. Mosha,1 W. Urassa,2 S. Aboud,2 E. Lyamuya,2 E. Sandstrom,3 H. Bredell,4 and C. Williamson4

Abstract
As human immunodeficiency virus (HIV) diversity may have an impact on both vaccine efficacy and drug resistance, it is important to have knowledge of circulating genetic variants. With widespread use of antiretroviral (ARV) drugs in Africa, one of the major potential challenges is the risk of emergence of ARV drug-resistant HIV strains. This study aimed to determine the circulating HIV subtypes and recombinant forms, as well as the prevalence of ARV drug resistance mutations, among 75 treatment-naive HIV-infected youths in Dar es Salaam, Tanzania. Gag (n = 48), partial pol (n = 44), and partial env (n = 35) sequencing was performed; all three regions were sequenced in 26 samples. Evidence of infection with recombinant viruses was found in 12 (46%) participants; AC recombinants were the most commonly detected and they were identified in six (23%) participants. Of individuals infected with nonrecombinant strains, subtype A was most commonly detected in seven (27%) participants, followed by subtype C detected in six (23%) participants and subtype D detected in one (4%) participant. Among the pol sequences from 44 individuals, three (7%) had resistance to nucleoside reverse transcriptase (RT) inhibitors and four (9%) had nonnucleoside RT inhibitor resistance mutations. Of these, three (7%) individuals were infected with viruses with cross-resistance mutations to both classes of RT inhibitors. These resistant mutations were all associated with drugs currently used in first-line therapy and in the prevention of vertical transmission. This high prevalence of resistance mutations is of considerable concern in apparently drug-naive populations as it may result in treatment failure and the spread of ARV-resistant strains.

Introduction
Tanzania is among the countries that have been greatly affected by human immunodeficiency virus (HIV) infection. A population-based survey conducted in 2007/2008 revealed an overall national prevalence of 6%, with a slightly higher prevalence among females (7%) compared to males (5%), with the age group of 20–49 years particularly affected.1 Although antiretroviral therapy (ART) has been available in Tanzania through the private sector since mid-1990, widespread access became available only after 2003, when a National Care and Treatment Plan (NCTP) was developed that recommended the provision of ART to all patients with CD4 T cell counts below 200 cells/μL.2 Prior to the NCTP implementation there was an era of treatment anarchy utilizing monotherapy, dual therapy, and less potent combination therapy with suboptimal dosages, thus creating a breeding ground for HIV drug resistance. The emergence of HIV resistance to antiretroviral (ARV) drugs renders the drugs ineffective in suppressing viral replication and is therefore a major threat to public health programs.3 Currently, the most widely used first-line drugs in Tanzania are a fixed dose combination of the nucleoside reverse transcriptase inhibitors (NRTIs; stavudine and lamivudine), plus the nonnucleoside reverse transcriptase inhibitors (NNRTIs; nevirapine), or combinations including zidovudine (NRTI) or efavirenz (NNRTI). Second-line combinations including protease inhibitors are rarely used in Tanzania. The NCTP recommend nevirapine and zidovudine for prevention of mother-to-child transmission.2 The introduction of ARV

1Field Epidemiology and Laboratory Training Programme, Ministry of Health and Social Welfare, Dar es Salaam, Tanzania.
2Department of Microbiology and Immunology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania.
3Department of Clinical Science and Education, Sodersjukhuset and Karolinska Institutet, Stockholm, Sweden.
4Institute of Infectious Disease and Molecular Medicine, Division of Medical Virology, University of Cape Town, Cape Town, South Africa.
drugs reinforces the need to introduce a surveillance system to monitor the emergence of new ARV drug-resistant variants. Genotypic testing among ARV drug-naive patients is therefore important in clinical settings for the detection of baseline mutations prior to the initiation of ART. In this study we aimed to establish the baseline prevalence of ARV drug resistance in drug-naive youths.

While implementation of treatment programs is essential for people already infected, vaccines are considered to be one of the best ways to ultimately control the epidemic. Tanzania has been involved in preparations for HIV vaccine trials for nearly a decade and in 2006 started the first phase I/II trial involving 60 healthy volunteers. These vaccines were designed on immunogens from subtypes A, B, C, and D. As the genetic differences between the candidate vaccine and the circulating HIV subtypes may have an impact on the efficacy of the vaccine there is a need to obtain information on HIV genotypes in Tanzania. The study also aimed to provide insight into the design of appropriate vaccine candidates for the region through an investigation of circulating HIV subtypes among infected youths in Dar es Salaam, Tanzania.

Materials and Methods

Study population

A cross-sectional study was conducted on ART-naive HIV-infected youths attending the Infectious Disease Centre (IDC) voluntary and counseling clinic in Dar es Salaam, Tanzania, from July 2004 to June 2005. Youths aged 13–25 years were eligible for participation in the study if they reported they had never been on any ART.

Laboratory evaluation

Whole blood was collected in EDTA vacutainer tubes. CD4 T-lymphocyte counts were enumerated using FACSCalibur (BD, USA) as described previously. Whole blood, separated by centrifugation at 4000 revolutions per minute for 20 min at −15°C, and plasma were aliquoted and stored at −80°C until the time for testing.

HIV genotyping

HIV RNA was extracted from plasma using the QIAamp viral RNA extraction kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. cDNA synthesis was performed using the Thermoscript RT-PCR System (Life Technologies, USA) according to the manufacturer’s instructions with the primer IN3 for the pol region, Gag D Rev for the gag region, and ED33 for the env region. A nested polymerase chain reaction (PCR) was done in all three regions.

The pol region

During the first-round PCR, the DNA was amplified with primers IN3 and G25REV producing a 2400-base pair fragment. Primers AV150 and PolM4 were used during the second-round PCR producing a 1770-base pair fragment. Denaturation was done at 94°C for 2 min, with a cycle sequence of 10×(94°C for 10 s, 50°C for 30 s, and 68°C for 2 min), 25×(94°C for 15 s, 50°C for 30 s, and 68°C for 2 min + 20 s for each successive cycle), and then the final elongation step at 68°C for 7 min.

The gag region

During the first-round PCR, the DNA was amplified with primers Gag D Rev and Gag D Fwd producing a 1777-base pair fragment. Three sets of second-round PCRs were performed using the following primers: Gag A Fwd and Rev, Gag B Fwd and Rev, and Gag C Fwd and Rev. Denaturation was done at 94°C for 2 min, with a cycle sequence of 3×(94°C for 1 min, 55°C for 1 min, and 72°C for 1 min), 32×(94°C for 15 s, 55°C for 45 s, and 72°C for 1 min), and then the final elongation step at 72°C for 7 min.

The env region

The primers ED33 and ED31 were used for the first- and second-round PCR. Denaturation was done at 94°C for 2 min, and a cycle sequence of 30×(94°C for 10 s, 50°C for 30 s, and 72°C for 2 min), then the final elongation step at 72°C for 7 min.

Sequencing and sequence analysis

The QIA quick PCR purification kit (Qiagen, Hilden, Germany) was used to purify the nested PCR products and quantified by agarose gel visualization. The ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit was used for sequencing together with the following primers: primers polM0, polM1, polM4, polM1REV, and polM86 for the pol region, primers Gag A Fwd, Gag A Rev, Gag B Fwd, Gag B Rev, Gag C Fwd, and Gag C Rev for the gag region, and primers ED33 and ED31 for the env region.

Drug-resistant mutations were classified according to the International AIDS Society recommendations, with accessory mutations identified utilizing the Stanford database. Sequences were aligned using Clustal X, manually edited with BioEdit, and neighbor-joining phylogenetic trees generated using Mega-4.0.0 software. The REGA subtyping tool and Genetic algorithm subtype prediction (GASPRED) were used for the screening of recombination patterns and subtyping of relevant fragments. The sequences were analyzed for recombination using bootscanning methods. All sequences that showed evidence of recombination as well as those that could not be subtyped using REGA were screened using the Recombination Detection Program (RIP) to identify possible breakpoints. Epi Info Version 3.5.1 (2008) was used for the statistical analysis. Representative sequences of HIV-1 group M subtypes from the Los Alamos HIV-1 database were used to construct phylogenetic trees and to evaluate bootstrap robustness values. Individuals were classified as infected with recombinant viruses if there was evidence of intersubtype recombination within the pol, gag, or env sequences, or if there was discordance in subtype classification between regions sequenced.

Ethical considerations

The ethical committee of the Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania approved the study protocol. Pre- and post-HIV test counseling was done for all study participants. Written informed consents were obtained from all participants prior to HIV testing and inclusion into the study. Individuals found to be HIV infected were managed at the IDC according to the existing treatment guidelines in Tanzania. All samples collected were...
ARV RESISTANCE MUTATIONS IN TANZANIA 3

sequenced. However, the sequences contained a region belonging to another subtype; this region was not of sufficient length to classify.

Table 1. Subtype and Recombinant Form DISTRIBUTION AMONG THE SEQUENCED SAMPLES IN DAR ES SALAAM, TANZANIA DURING 2005

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Env region</th>
<th>Gag region</th>
<th>Pol region</th>
<th>Env/gag/pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtype A</td>
<td>16 (46%)</td>
<td>12 (25%)</td>
<td>16 (36%)</td>
<td>7 (27%)</td>
</tr>
<tr>
<td>Subtype C</td>
<td>6 (17%)</td>
<td>13 (27%)</td>
<td>16 (36%)</td>
<td>6 (23%)</td>
</tr>
<tr>
<td>Subtype D</td>
<td>1 (3%)</td>
<td>7 (15%)</td>
<td>6 (14%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Unclassified</td>
<td>12 (34%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>C-likec</td>
<td>0 (0%)</td>
<td>2 (4%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>D-likec</td>
<td>0 (0%)</td>
<td>2 (4%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Subtype AD</td>
<td>0 (0%)</td>
<td>2 (4%)</td>
<td>2 (5%)</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Subtype AC</td>
<td>0 (0%)</td>
<td>8 (17%)</td>
<td>1 (2%)</td>
<td>2 (6%)</td>
</tr>
<tr>
<td>Subtype ACD</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Subtype AB</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (2%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Subtype CD</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (2%)</td>
<td>3 (12%)</td>
</tr>
<tr>
<td>CRF10_CD</td>
<td>0 (0%)</td>
<td>2 (4%)</td>
<td>1 (2%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>48</td>
<td>44</td>
<td>26</td>
</tr>
</tbody>
</table>

*Intersubtype recombination within the pol, gag, or env sequences, or if there was discordance in subtype classification between regions sequenced.

**Sequence did not group within a particular subtype and no definitive recombinant was identified.

The majority of the sequence belonged to a particular subtype, however, the sequences contained a region belonging to another subtype; this region was not of sufficient length to classify.

given new identification numbers and were delinked from the patient’s identifiers.

Results

The study population consisted of 75 youth with a mean age of 22 years (range 13–25), of which 17 (23%) were males and 58 (77%) were females. Sixty-two (83%) were single, 12 (16%) were married, and one (1%) was divorced. The median CD4 T cell count was 228 cells/μL. Forty-four (59%) samples were sequenced in the pol region, 35 (47%) in the env region, and 26 (35%) samples were sequenced in all three regions (Table 1).

Of the 26 participant samples that were sequenced in all three regions, 12 (46%) were found to be infected with unique recombinant viruses including six (23%) AC recombinants, three (12%) CD, and one (4%) ACD (Table 1).

Of the gag and pol sequences, 16 (33%) and 6 (13%) were recombinant viruses, respectively. Of the recombinant viruses identified in gag sequences, only two (4%) belonged to a previously described circulating recombinant form (CRF10_CD) and the remainders were unique recombinants. Other identified recombinants in the gag region were AD (4%), AC (17%), C-like (4%), and D-like (4%). Sequences were classified as C- or D-like when the majority of the sequence belonged to a particular subtype together with a region belonging to another subtype, but this region was not of sufficient length to classify. The identified recombinants in the pol regions were AD (5%), AC (2%), AB (2%), CD (2%), and one (2%) CRF10_CD while 12 (34%) env region sequences were unclassified (Table 1 and Fig. 1) as they did not reliably group within a particular subtype and no definitive recombinant was identified.

Among the participants infected with viruses where no intersubtype recombination was detected, three dominant subtypes were identified: subtypes A, C, and D. Subtype A

![FIG. 1. Neighbor-joining phylogenetic tree analysis of pol sequences from drug-naive youth in Dar es Salaam, Tanzania during 2005. Bootstrap values over 70% are illustrated. The bar equals 0.5% DNA distance.](image-url)
was the dominant subtype identified in 16 (46%) env sequences compared to only 12 (25%) and 16 (36%) gag and pol sequences, respectively. Subtype C was the most common subtype identified in both gag and pol, and was found in 13 (27%) and 16 (36%) sequences, respectively, compared to 6 (17%) env sequences. Subtype D had the lowest prevalence and was detected in one (3%), seven (15%), and six (14%) of env, gag, and pol sequences, respectively. Among the 26 samples that were sequenced in all three regions, seven (27%) sequences clustered with subtype A, six (23%) with subtype C, and one (4%) sequence with subtype D (Table 1 and Fig. 1).

Among 44 pol sequences that were analyzed for ARV resistance mutations, four (9%) had resistance to at least one of the ARV groups (Table 2). The ARV resistance mutations were grouped as mutations conferring resistance to protease inhibitors, NRTIs, and NNRTIs.

Three (7%) individuals, all females, harbored viruses with mutations conferring resistance to both classes of RT inhibitors, in which all were resistant to first-line ARV in the country; two were resistant to stavudine, lamivudine, nevirapine, and efavirenz and one to lamivudine, nevirapine, and efavirenz (Table 2).

Discussion

Characterization of circulating HIV subtypes is important in HIV candidate vaccine development and testing. In addition, baseline data on ARV drug resistance are important in the surveillance of the emergence of HIV drug resistance in ART programs. This study aimed to identify circulating HIV subtypes and recombinant forms together with mutations associated with ARV drug resistance in treatment-naive HIV-infected youths. We selected young people as proxy to “new infection.” However, the low mean CD4 T cell counts suggest that some of these individuals were not recently infected but could not access early HIV testing.

Characterization of circulating HIV subtypes was performed by analysis of genome sequences of HIV-1 strains. Overall, the highest percentage of samples clustered with subtype A and C with AC being the most common recombinant. A high frequency (at least 52%) of unique recombinant viruses is characteristic of Tanzanian cohorts and presumably is a result of dual infection with more than one subtype. These findings complement and are consistent with existing molecular epidemiological data in Tanzania, which indicated the presence of subtypes A, C, D, and their recombinants. Despite the fact that the samples were collected as long ago as 2004–2005, these findings are useful for monitoring the trends of ARV drug resistance in Tanzania.

Of concern was the frequent detection of resistance to the first-line ARV in Tanzania, which is Triomune (stavudine, lamivudine, and nevirapine), in individuals who reported that they had not taken ARVs. This ARV resistance pattern was not limited to a particular subtype but was observed more in female patients than males, which could be due to the fact that the clinic receives more female patients than male patients. Polymorphisms detected in RT and protease sequences were mainly mutations that have been found in the consensus sequence of non-B subtypes. However, this could not be verified in this study. The PI mutations observed were minor mutations and could be possibly naturally occurring polymorphisms and may not have clinical relevance.

Few patients were sequenced in all three regions based on successful pol region sequencing and PCR failures could not be repeated because of the limited available budget. It would be of value to extend this study to a larger number of participants to determine if the subtype affects the development of drug resistance. In addition, future studies would need to differentiate more accurately individuals who had been exposed to drugs compared to infection with drug resistance viruses. In this study investigators depended on patient self-reporting on ARV drug exposure and some participants may not have been truly ARV naive.

The observed prevalence of 9% ARV drug-resistant mutations in drug-naive patients in the current study is in contrast to the <5% prevalence reported previously among ARV-naive antenatal clinic attendees. However, a worldwide surveillance program found the prevalence of ARV resistance among treatment-naive populations in Africa to be 5.5%. Minor mutations were also present in the protease gene, reducing the “genetic barrier” to drug resistance by increasing the selection of resistant strains. These mutations are associated with resistance to each of the PIs when present with other mutations. Interestingly, resistance mutations in the protease inhibitor genes were also observed in two different studies in Tanzania and in other developing countries on ARV-naive antenatal clinic attendees. In the Tanzania treatment regimen, PIs are used only as a second-line treatment in the few patients who are recognized to have clinical treatment failures as indicated by clinicians within NCTP or outside the program when they have access to these drugs.

Given this level of resistance in this population, further work would be needed to determine if resistance is a consequence of short-term exposure during pregnancy or if in fact

### Table 2. Mutations Conferring Resistance to NRTIs and NNRTIs among Sequenced Samples in Dar es Salaam, Tanzania during 2005

<table>
<thead>
<tr>
<th>PID</th>
<th>Subtype</th>
<th>NRTI mutations</th>
<th>NNRTI mutations</th>
<th>Gender</th>
<th>CD4 count (cells/μL)</th>
<th>Drug resistance to first-line therapya</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>AC</td>
<td>M184V, T215F</td>
<td>K103N, Y181C</td>
<td>Female</td>
<td>83</td>
<td>Lamivudine/stavudine/nevirapine/efavirenz</td>
</tr>
<tr>
<td>17</td>
<td>D</td>
<td>M184V</td>
<td>K103N</td>
<td>Female</td>
<td>175</td>
<td>Nevirapine/efavirenz/lamivudine</td>
</tr>
<tr>
<td>33</td>
<td>AC</td>
<td>Y181C</td>
<td></td>
<td>Male</td>
<td>39</td>
<td>Nevirapine/efavirenz</td>
</tr>
<tr>
<td>66</td>
<td>C</td>
<td>K65R</td>
<td>Y181C</td>
<td>Female</td>
<td>374</td>
<td>Nevirapine/efavirenz/stavudine/lamivudine</td>
</tr>
</tbody>
</table>

*First line therapy: nucleoside reverse transcriptase inhibitors (NRTIs; stavudine and lamivudine); plus the nonnucleoside reverse transcriptase inhibitors (NNRTIs; nevirapine), or combinations including zidovudine (NRTI) or efavirenz (NNRTI). The NCTP recommends nevirapine and zidovudine for prevention of mother-to-child transmission.*
these individuals had accessed ARVs. It would be of public health relevance to perform an epidemiological ARV drug resistance surveillance in drug-naive population groups to evaluate these preliminary findings and determine the frequency of transmitted resistance viruses. In addition, it will help to track the emergence of new ARV HIV-resistant variants. Together, this information will be used to guide development of ART policy guidelines. ARV resistance monitoring will help avoid the unnecessary cost of using ineffective treatments. Thus, monitoring of the emergence of drug-resistant HIV in populations starting and using ARV drugs would be a useful approach for planning an effective ART program in Tanzania. It is concluded that the existence of ARV drug-resistant mutations and the great diversity of HIV subtypes and circulating recombinant forms have implications for treatment and vaccination strategies. The observed mutations within the pol region are of considerable concern because it may increase the development and spread of ARV-resistant strains.

Acknowledgments

We would like to thank the study participants for consenting to participate in the study, the clinic staff at IDC for assisting with specimen collection, the technical staff in the Department of Microbiology and Immunology, MUHAS, for assisting with the processing of specimens, and the technical staff at the University of Cape Town for assisting with genotyping. The study was supported by European Union Grant ICA4-CT-2002-10036, HIVIS.

Author Disclosure Statement

No competing financial interests exist.

References