**Abstract**

**Introduction:** HIV genetic diversity and envelope (env) tropism, V3 loop CCR5 (R5) or CXCR4 (X4) co-receptor use, are associated with HIV pathogenesis and transmission (Tx). Co-receptor use was estimated from env V3 loop sequences clones by a Subtype C position specific scoring matrix from 18 pregnant women and 4 of their infants. Viral diversity in env RNA from plasma (PL) and cervix (Cx) was estimated as nucleoside ambiguity and intra-clonal ambiguity from 18 women in late pregnancy. Infant infection was determined by DNA-PCR at 2, 6 and 24 weeks.

**Results:** The 18 pregnant women had median CD4 364 cells/mm$^3$, Pl and Cx virus load of 4.0 log$_{10}$ copies/ml and 3.5 log$_{10}$ copies/ml, respectively. Seven of 18 women transmitted HIV and 4 infants’ env were sequenced. Eight of the 18 women (44%) had X4 tropic clones or X4/R5 tropism whilst 10 had exclusively R5 virus by cPSSM. Of 4 infected infants’ virus 3 were R5, and one X4 tropic. Nucleoside ambiguity and diversity to estimate duration of infection, and X4/R5 Dual tropism were associated with MTCT.

**Importance:** Genotypic detection of X4 and dual tropic X4/R5 virus among subtype C infected women demonstrate high levels of diversity in association with transmission. Low cost, rapid genotyping may identify risk of transmission and guide the appropriate use of R5 entry inhibitors to prevent mother to child transmission in subtype C infection in Southern Africa.

**INTRODUCTION**

Although CCR5 (R5) tropic viruses predominate in early HIV infection [1,2], CXCR4 tropic (X4) and dual tropic viruses (X4/R5) may emerge in 40-50% [3] with increasing duration of infection [4]. The emergence of X4 and dual tropic virus is associated with more rapid CD4+T cell decline and disease progression compared to exclusively R5 tropic virus ([5-10]. The transmission of HIV from mother to infants is associated with maternal virus load, genetic diversity and chemokine receptor tropism. HIV-1 envelope sequence diversity and dual tropic virus pose challenges to the development of vaccines and antiretroviral therapy (ART) to prevent R5 co-receptor attachment, viral entry and infection.

In subtype C infection, the virus which accounts for nearly 50% of HIV infection worldwide, R5 viruses have been associated with sexual and vertical transmission. X4 tropic viruses are more often seen to emerge after exposure to ART [11-13] and when anti-retroviral treatment is used to prevent mother-to-child-transmission (pMTCT) [14]. A longitudinal study of pregnant women with HIV subtypes A, B, C and D demonstrated a high frequency of switches from R5 to X4 between the first and third trimesters of pregnancy [14]. Although infants are usually infected by maternal R5 tropic strains despite the presence of X4 tropism [15-18], the acquisition of X4 variants among infants [19-23] is associated with lower birth-weight and may portend a poor prognosis compared to infants with R5 virus [20].

Diversity of HIV envelope sequences, a consequence of host immune selection, and chronicity of infection may be measured by the rate of synonymous versus non-synonymous mutations (ds/dn) [24] or as the pair wise genetic distances [25,26]. Intra-individual evolution and genetic diversity has been demonstrated by multiple amplicon cloning and single genome sequencing (SGS) [Tamura-Nei] [27]. Viral diversity within an individual has been observed to increase with duration of infection [28-30] and the nucleoside ambiguity of HIV pol and gag are proportional to the duration of infection [31,32].

The extent of genetic variation between genital and plasma compartments is controversial. Diversity between cervical vaginal virus and plasma was first described in the V3 envelope region among recently infected Kenyan women [33]. This was further explored by Overbaugh et al., 1996, who showed that in some cases, mucosal variants of HIV-1 V1-V3 envelope in the
cervix arose from minor variants in blood [34]. Genital tract and plasma viral diversity have been further explored among women with sub-type C infection demonstrating significantly higher genital virus load compared to subtype B infected women [35].

The relationship between HIV-1 envelope sequences from the plasma and cervical vaginal compartments and the intra-individual diversity of subtype C HIV-1 was explored among 18 subtype C infected pregnant women entering an early Zidovudine prevention trial in Harare. High levels of diversity, R5/X4 dual tropic strains in maternal plasma and cervical vaginal compartments present challenges in the prevention of HIV-1 transmission of subtype C HIV infection in Southern Africa.

METHODS AND MATERIALS

Study population and viral isolation

Eighteen HIV-infected pregnant women were recruited between January 1999 and June 2001 to an open-label prospective study of maternal short-course Zidovudine to prevent MTCT of HIV sponsored by the Swedish International Development Agency (SIDA) in Chitungwiza, Zimbabwe. Whole blood samples and cervical-vaginal swabs were collected at 36 weeks gestation, before receipt of Zidovudine. Plasma and cervix viral loads were measured using the Roche Amplicor 1.5 (Roche Molecular Systems, Pleasanton, CA) as per manufacturer’s instruction. Ethical review was obtained from the Institutional review board at the University of Zimbabwe and the panel for human subjects’ protection at Stanford University School of Medicine.

RNA extraction and cDNA synthesis

Viral particles were concentrated from 500µl of plasma by centrifugation at 12,000 g at -4°C for 30 min. 300µl of the supernatant was discarded and the remaining 200µl was used to resuspend the viral pellet. Viral RNA was extracted from the concentrated plasma using the QIAampViral RNA Mini kit (Qiagen). Random hexamers and Super Script III (Thermo Fischer) were used to generate cDNA from 9µl of purified viral RNA.

Polymerase chain reaction amplification

Netsted PCR was used to amplify the C2-V5 region of the envelope gene. Platinum Taq and was used for both 1st and 2nd round PCRs. The primers for 1st round PCR were, forward KK1-GCACAG envelope gene. The 2nd round primers were, forward ES7-CTGTTAAATGGCAGTCTAGC (positions 7002-7021) and reverse ED12-AGTGCTTCCTGCTGCTCCCAAGAACCCAAG (position 7647-7667). The second round PCR products were separated by running on a 1% agarose gel at 100V for 40 minutes.

Sequencing and phylogenetic analysis

The PCR products were purified using the Qiaquick PCR purification kit and the DNA sequenced using two bi-direction primers covering the C2-V5 (~700 nucleotides). Sequencing was performed on a 3130 xl genetic analyzer (Applied Biosystems Inc, Forster City, CA, USA). Sequence assembly was done in Sequencer, alignment and manual editing in MEGA (version 5; Molecular Evolutionary Genetics Analysis) and Geneious R7 (Biomatters). Subtype C position specific scoring matrix http://fortinbrus.su/qgbin/fssm/fssm.pl), using C-PSSM predicted co-receptor usage. Neighbor joining phylogenetic trees of the C2-V5 (~700 nucleotides) of env with 1000 bootstrap replicates were inferred with Phylip package with 2-Kimura parameters and support values for the branching pattern were calculated from 1000 bootstrap replicates, and the trees analyzed for relatedness between infant, cervical and plasma sequences.

Clonal analysis

The second round PCR products from the plasma and cervical virus of the 7 transmitting (Tx) women were done into E. coli TOP10 cells expanded in LB broth and the inserts were sequenced. Two hundred and thirteen (213) C2-V5 env clones, a mean of 30 clones from each woman; 115 from plasma and 98 from cervical env RNA were sequenced. To estimate viral diversity, C2-V5 consensus sequences from plasma and cervix env were analyzed.

Compartmental sequence diversity by population and clonal analysis

Among 11 women an intra-individual diversity between PI and Cx consensus sequences was calculated (Fourment and Gibbs, 2006) as “patristic distance”, the sum of branch lengths that link two terminal nodes of envelope C2-V5 (Figure 3). V3 Nucleoside ambiguity was calculated as the percent positions with ambiguous bases in the V3 loop from consensus PI and Cx sequences from 11 women whose infants were not infected. A V3 intra-clonal nucleoside ambiguity score was calculated as the percent of the 111 nucleotide positions with variant nucleotides from at least 5 clones from each compartment among the 7 transmitting women.

Statistical analysis

Statistical tests were conducted using the Graph Pad Prism software version 6. Non-parametric t-test was conducted to determine P-values, and values less than or equal to 0.05 were considered significant.

RESULTS

Clinical characteristics of 18 study subjects

Among 18 HIV+ pregnant women participating in an early study of pMTCT with Zidovudine in Zimbabwe, the median CD4 count was 364 cells/mm³ (IQR 279- 522.25) and the median PI and Cx virus load was 4.0 log10 copies/ml (IQR 2.88-3.74) and 3.5 log10 copies/ml (IQR 2.89-3.74), respectively (Table 1a,1b). All virus sequences were classified as subtype C HIV-1 save for one subtype B sequence that was identified from the Cx virus (subject 3030).
**Table 1a:** Characteristics of 11 non-transmitters’ population sequences.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Diversity V3 (%) (111 Nts)</th>
<th>Diversity V1-V5 (~700 Nts)</th>
<th>CD4</th>
<th>Plasma Tropism</th>
<th>Cervical Tropism</th>
<th>Infant transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V3 Nucleoside Ambiguity¹</td>
<td>Patristic distance (V1-V5)</td>
<td>per cu mm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3048</td>
<td>0</td>
<td>0.008</td>
<td>723</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3054</td>
<td>0</td>
<td>0.054</td>
<td>567</td>
<td>X4</td>
<td>X4</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3031</td>
<td>0.901</td>
<td>0.041</td>
<td>375</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3058</td>
<td>0.901</td>
<td>0.004</td>
<td>636</td>
<td>X4</td>
<td>X4</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3032</td>
<td>2.703</td>
<td>0.019</td>
<td>317</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3036</td>
<td>3.604</td>
<td>0.059</td>
<td>133</td>
<td>X4</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3040</td>
<td>3.604</td>
<td>0.023</td>
<td>542</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3034</td>
<td>4.505</td>
<td>0.009</td>
<td>11</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3037</td>
<td>4.505</td>
<td>0.031</td>
<td>33</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
<tr>
<td>3030</td>
<td>6.306</td>
<td>0.309</td>
<td>520</td>
<td>X4</td>
<td>Infant negative</td>
<td></td>
</tr>
<tr>
<td>3028</td>
<td>9.009</td>
<td>0.023</td>
<td>303</td>
<td>R5</td>
<td>R5</td>
<td>Infant negative</td>
</tr>
</tbody>
</table>

¹ V3 Nucleoside Ambiguity defined as number of positions with ambiguous bases in the V3 cervical and plasma population sequences, expressed as a percentage of the total number of nucleotides in the V3 loop (111).

**Table 1b:** Characteristics of 7 transmitting women and 4 infants.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Clones</th>
<th>V3 Intra-clonal Ambiguity²</th>
<th>Diversity V1-V5 (~700 Nts)</th>
<th>CD4</th>
<th>Plasma Tropism</th>
<th>CV Tropism</th>
<th>mode of infant transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>3022</td>
<td>Y</td>
<td>0.901</td>
<td>n/a</td>
<td>452</td>
<td>X4</td>
<td>X4</td>
<td>LL/P⁴</td>
</tr>
<tr>
<td>3039</td>
<td>Y</td>
<td>13.51</td>
<td>n/a</td>
<td>278</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>LP⁵</td>
</tr>
<tr>
<td>3003</td>
<td>Y</td>
<td>1.802</td>
<td>n/a</td>
<td>474</td>
<td>R5</td>
<td>R5</td>
<td>IP/EPP⁴</td>
</tr>
<tr>
<td>3053*</td>
<td>Y</td>
<td>13.604</td>
<td>n/a</td>
<td>523</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IP/EPP⁵</td>
</tr>
<tr>
<td>4053*</td>
<td>N</td>
<td>0</td>
<td>n/a</td>
<td>523</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IP/EPP⁵</td>
</tr>
<tr>
<td>3062a</td>
<td>Y</td>
<td>5.405</td>
<td>n/a</td>
<td>265</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IU</td>
</tr>
<tr>
<td>4062a</td>
<td>N</td>
<td>0</td>
<td>n/a</td>
<td>523</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IU</td>
</tr>
<tr>
<td>3056e</td>
<td>Y</td>
<td>n/a¹</td>
<td>n/a</td>
<td>282</td>
<td>R5</td>
<td>R5</td>
<td>IU</td>
</tr>
<tr>
<td>4056e</td>
<td>N</td>
<td>0</td>
<td>n/a</td>
<td>282</td>
<td>R5</td>
<td>R5</td>
<td>IU</td>
</tr>
<tr>
<td>3038e</td>
<td>Y</td>
<td>21.622</td>
<td>n/a</td>
<td>523</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IU</td>
</tr>
<tr>
<td>4038e</td>
<td>N</td>
<td>0</td>
<td>n/a</td>
<td>523</td>
<td>X4/R5</td>
<td>X4/R5</td>
<td>IU</td>
</tr>
</tbody>
</table>

* ø φ †§ Infant-Mother pairs
1. V3 Nucleoside Ambiguity defined as number of positions with ambiguous bases in the V3 cervical and plasma population sequences, expressed as a percentage of the total number of nucleotides in the V3 loop (111).
2. V3 intra-clonal ambiguity defined as the number of positions with one or more nucleotide variant in the V3 loops of plasma and cervical vaginal clones of 7 subjects, expressed as a percentage of total number of nucleotides in the V3 loop.
3. Subject 3056 excluded because of insufficient number of clones generated in cervical-vaginal compartment.
4. Infant virus genotyped from Peripheral blood mononuclear cells.
5. Infant virus not available.

Mother to child transmission

Of the 7 mothers who transmitted HIV to their infants, 3 had both X4 and R5 virus co-circulating virus, 3 had only R5 virus, and 1 had only X4. All 4 infants were infected with virus that was closely related to maternal PL virus. For each of the 7 transmitters, neighbor joining trees with 1000 bootstrap replicates for Cx and PL clones and infant sequences (envelope V1-V5; ~700 nucleotides) were used to estimate intra-subject diversity from an average of 17PL and 15 Cx clones per subject (Figure 2).

Diversity

The “V3 intra-clonal ambiguity” scores among 6/7 women (Subject 3056 excluded because only one Cx clone was obtained) were determined as the percent of variant nucleotides in the PL and Cx clones. The median V3 intra-clonal ambiguity score was 4.5% (IQR 3.6% - 7.7%) (Table 1). Transmitters had V3 intra-clonal ambiguity scores ranging from 0.90% - 21.62%, whereas non-transmitters had lower V3 diversity scores ranging from 0% - 9.009% (Table 1). Among transmitters, higher V3 diversity scores were associated with dual tropism, whereas lower V3...
among the 18 HIV-infected women studied at 36 weeks of gestational age, 7/18 of their infants were infected. The 11 non-transmitters had either X4 (3) or R5 (8) monotypic virus. Despite maternal Zidovudine prophylaxis at 36 weeks, 7/18 infants acquired HIV infection; 2 were apparently infected in-utero, 3 intra-partum and 2 through breastfeeding. Eleven of the 18 infants remained HIV DNA negative on follow-up to 48 weeks.

Compartmentalization of consensus sequences from plasma and cervix

A neighbor joining tree with 1000 bootstrap replicates was generated using population PL andCxenvsequences from 11 women (Figure 4). PL and Cx population sequences were closely related within women, with the exception of one (subject 3030) who displaying a subtype B virus in the cervical compartment. Diversity between the viruses in different anatomical compartments as measured using population envelope sequences from 10 PL and Cx pairs demonstrated intra-individual patristic distances from 0.004 to 0.059 between compartments (excluding subject 3030) (Table 1a, Figure 4). Among 7 transmitters, anatomic compartmentalization of PL and Cx virus (defined as distinct clusters with bootstrap values > 70%) was observed among 3 women (3022, 3003 and 3053) all of whom had monotypic virus. There was no evidence for compartmentalization of virus from women with dual tropic viruses (3039,3062,3038) (supplemental Figures).

Temporal association of nucleoside ambiguity

Nucleoside ambiguity scores and viral diversity are associated with duration of infection. Nucleoside ambiguity in V3 envelope and time since infection were plotted as Nucleoside ambiguity scores (n=11) from the population sequences and V3 intra-clonal ambiguity scores (n=7) were used to estimate duration of infection. The fraction of ambiguous nucleotides over time was plotted as 0.2% per year as described by Kouyos et al., as an estimated time since infection from the with V3 nucleoside and intra-clonal ambiguity scores for 18 women and 4 infants [31]. The relationship between time and nucleoside ambiguity is displayed graphically and assessed by regression analysis. Dual tropic viruses (X4/R5) were associated with higher diversity scores, and were associated with greater than 8 years of infection (Figure 5). The diversity score of R5 or X4 monotypic viruses were consistent with earlier infection compared with patients with both R5 and X4 (dual-mixed virus) (Figure 5).

DISCUSSION

Among Subtype C infected women, before the era of ART, X4 tropic strains were found among 8/18 (44%) of pregnant women in Zimbabwe. Despite the provision of Zidovudine at 36 weeks, 7/18 of these women transmitted HIV to their infants. Transmission was associated with an increased diversity in env sequence and dual tropic R5 and X4 virus. Although transmission of HIV subtype C infection has been strongly associated with R5 tropism [36, 37], X4 and X4/R5 dual tropic viruses were found more often than in most studies from Southern Africa [11,16,20,37-45]. The higher frequency of X4 variants observed could be due to the small sample size? although the presence of X4 tropic variants is consistent with recently reported evidence of the evolution of co-receptor utilization from CCR5 to CXCR4 during pregnancy and in children [14-18].

The majority of women with monotypic either R5 or X4 viruses who did not transmit HIV to their infants appear to be more recently infected compared to pregnant women with diverse R5 and X4 (dual-mixed virus). This is consistent with the emergence of X4 tropic variants from monotypic R5 viruses with the presence of dual tropic virus. Evidence of an R5 to X4 switch in subtype C may be limited by the use of C-PSSM genotyping which assessed tropism based on sequence only from the V3 loop. This may underestimate X4 tropism in that mutations in the V1/V2 region of envelope [46] and changes in the V4 have been shown to complement or stabilize the R5 to X4 co-receptor switching [47]. Genotypic identification of X4 variants using C-PSSM in subtype C infection, late in pregnancy, also provides further evidence of a delayed switch from R5 to dual tropic virus associated with the higher mutational barrier in subtype C [46]. Lin et al recently demonstrated that R5 done from patients with dual mixed virus had greater diversity than those with purely R5 virus [38]. Fit to the linear model of increasing diversity with time for the first 8 years of infection [31], monotypic viruses (either purely X4 or purely R5) are found earlier while viruses with both X4 and R5 variants were associated with immune selection and increased duration of infection.

Sequencing multiple envelope clones provides evidence in some women for compartmentalization, the clustering of sequences of differing tropism in blood and genital secretions, which has been ascribed to tissue specific selective immune pressure, founder effects, or compartmental conditions such as presence of STIs ([33,34,48-53] or the emergence of drug resistance mutations [54]. This increase in diversity in dual
tropic strains is consistent with mechanisms of V3 evolution and tropism as a sequential transition from monotropic virus to dual tropic virus, suggesting more recent infection among women with R5 or X4 monotropic viruses compared to those with both R5 and X4 (dual-mixed virus). Immunologic selection is cited by in a genotypic study of 12 women, in which compartmental differences between blood and genital tract were associated with higher CD4 counts [55]. Similarly, the presence of X4 virus in vaginal secretions was associated with a threshold population of X4 quasi species during the HIV-induced immune decline [56]. Here, in clonal analysis, virus from the transmitting women with monotropic X4 or R5 virus demonstrated compartmental clustering. In contrast, among the transmitting women with dual tropic virus there was no evidence for compartmentalization, suggesting increasing diversity and exchange of circulating viruses across anatomic compartments.

Maternal env diversity and dual X4/R5 tropism was
Figure 4 V1-V5 Envelope Neighbor joining tree of cervical and plasma population sequences and patristic distances identifying the differences between cx and pl.

A neighbor joining phylogenetic tree of the V1-V5 (~700 nucleotides) of envelope region with 1000 bootstrap replicates was created to determine the genetic relationship between for the cervical and plasma consensus sequences for 11 patients for whom consensus sequences were generated. Patristic distance generated using Geneious R7 software and defined as the “sum of lengths of branches that link two nodes in a tree, where those nodes are typically terminal nodes that represent extant gene sequences or species.

Figure 5 V3 loop diversity scores versus time since infection vv
V3 intra-clonal ambiguity and Nucleoside ambiguity scores to explore the relationship of diversity in the V3 loop has with the time since infection.

associated with Mother-to-child-transmission while the 4 infants genotyped were infected with monotropic (1 X4, 3R5) virus. These findings suggest that env (V3) genotype analysis may be a useful tool to focus the effective use of R5 entry inhibitors in the treatment and prevention of MTCT. Population-sequencing approaches to fully understanding the diversity and evolution of HIV are limited compared to next generation sequencing and cloning to identify intra-patient viral diversity. Recent studies using Ultra-deep sequencing have demonstrated low levels (3-6.4%) of X4 virus that may be detected even in primary infection [57]. Further studies of tropism and diversity pregnancy as a risk for transmission are warranted and may provide effective interventions to further reduce MTCT.

REFERENCES


56. Haaland RF, Sullivan ST, Evans-Strickfaden T, Lennox JL, Hart CE. Female genital tract shedding of CXCR4-tropic HIV Type 1 associated with a majority population of CXCR4-tropic HIV Type 1 in blood and declining C34+ cell counts. AIDS Res Hum Retroviruses. 2012; 28: 1524-1532.