Genetic Variation in Mother-Child Acute Seroconverter Pairs from Zambia

Federico G. Hoffmann  
*University of Nebraska at Lincoln*

Xiang He  
*CAIDS*

John T. West  
*University of Oklahoma Health Sciences Center, john-west@ouhsc.edu*

Philippe Lemey  
*University of Oxford, philippe.lemey@zoo.ox.ac.uk*

Chipepo Kankasa  
*University of Zambia*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.unl.edu/virologypub](http://digitalcommons.unl.edu/virologypub)

Part of the [Virology Commons](http://digitalcommons.unl.edu/virologypub)


[http://digitalcommons.unl.edu/virologypub/177](http://digitalcommons.unl.edu/virologypub/177)

This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Primary HIV infection is often characterized by ‘flu-like’ symptoms and a high plasma viral load in body fluids [1,2] and has been found to associate with higher rates of heterosexual transmission [3-5]. In spite of the impact of primary HIV-1 infection (PHI) on transmission, very little is known about transmission during PHI, particularly in the context of mother to child transmission (MTCT). HIV-1 MTCT constitutes a major problem in sub-Saharan Africa, where HIV-1 infection is epidemic, and where, until recently, antiretroviral therapy was not widely available as a preventive measure. In the absence of a prophylactic vaccine or universal access to preventive treatment in developing countries, a clearer understanding of MTCT is critical to the development of effective measures to reduce rates of MTCT. We have centered our efforts on identifying cases of MTCT dur-
ing PHI of the mothers in a cohort in Zambia, as part of ongoing efforts to understand transmission and disease progression in subtype C HIV-1 infection.

Most studies of transmission have focused on adult donors who were chronically infected. In this context, only a limited number of viruses within the viral quasispecies population in the donor are transmitted; collectively this suggests that only a few viral genotypes are competent for transmission [6]. This is also true of perinatal MTCT, where a subset of the maternal quasispecies is transmitted to the infant in the majority of cases [7-11], though some studies have documented transmission of multiple genotypes [10,12,13]. A recent study in adults comparing transmission from recently infected and chronically infected patients found that the transmitted virus derived from a single lineage in the donor in all cases, regardless of infection stage of the donor [14]. However nothing has been reported about MTCT transmission during PHI.

For the current study, three mother-infant pairs (MIP) were identified where MTCT presumably occurred shortly after primary infection of the mothers, and longitudinal samples of the viral populations were obtained from mother and infant in each pair. The intention of the study was to characterize the transmitted viruses with relation to the source, and compare this with MTCT from chronically infected mothers. The characterization of the viral populations in these three MIP could also provide a better understanding of HIV-1 subtype C MTCT during PHI. In addition, these MIP provided a unique opportunity to compare longitudinal evolution between viral populations established by closely related viruses in two different hosts.

Methods

Subjects and specimen collection

Archived blood samples were obtained retrospectively from participants in a longitudinal cohort to study HIV-1 and Kaposi sarcoma-associated herpes virus transmission at the University of Zambia Teaching Hospital (Lusaka, Zambia). Informed consent was obtained from all participants. HIV-1 status was determined by two rapid HIV tests followed by an immunofluorescence assay. Infection of infants younger than 18 months of age was confirmed by PCR.

Polymerase chain reaction, gene cloning, and sequencing

Total DNA from cells stored at -80°C was extracted using the Gentra Genome extraction kit (Gentra, Minneapolis, Minnesota, USA). A fragment of env (gp120, V1-V5 region) was amplified by nested PCR as described by Zhang et al. [9] using primers enFl (5’-GATGCAT-GAGG ATATAATCAGTT TATGGGA, positions 6533-6562 of HIV-1 strain HXB2) and enRl (5’-ATTGATGCT-GCG CCCATAGTGC, positions 7828-7806) for the first reaction and enF2 (5’-AGTTTTATGGGACCAAAGCCT AAAGCCATGT, positions 6552-6581) and enR2 (5’-A CTGCTCTTTTTCTCTCACCACCT, positions 7762-7734) for the second. Two independent PCR amplifications were carried out for each sample; products were pooled and cloned into pGEM-T easy vector. Clones were sequenced using vector and internal primers with BigDye 3 chemistry (Applied Biosystems, Foster City, California, USA).

Quantification of proviral DNA burden

Real-time PCR was used to quantify the number of copies of HIV-1 in each sample with TaqMan chemistry (Applied Biosystem). The reaction targeted a sequence of the 5’ viral long terminal repeat as described by Yun et al. [15]. The reaction contained 400 nmol/l of forward and reverse primer, and 200 nmol/l of FAM-labeled probe. Standard curves for HIV-1 and human β-globin were generated following the protocols standardized in our laboratory [16].

Sequence analyses

Sequence alignment was carried out on the translated amino acid sequence in ClustalW [17], as implemented in MEGA version 3.1 [18]. Maximum likelihood, minimum evolution, neighbor-joining, and Bayesian phylogenetic analyses were used to explore genealogical relationships among mother and infant viral sequences. Minimum evolution and neighbor-joining analyses were performed in MEGA using the Kimura 2-parameter genetic distance, and support for the nodes was evaluated with bootstrap. Maximum likelihood searches were conducted in Treefinder version May 2006 [19] using an independent GTR + r model for each codon position. Bayesian analyses were carried out using MrBayes version 3.04 [20], running four chains for 5 x 10⁶ generations, collecting trees every 103 generations. After checking for convergence, 2,500 trees were used to build a majority rule consensus.

Variations in genetic diversity, genetic divergence, the number and location of putative N-linked glycosylation sites (PNGS), and the length of the V1-V5 fragment were analyzed. Viral genetic diversity was estimated as the average nucleotide difference between sequences within a contemporaneous set, and genetic divergence was calculated as the average genetic distance to the earliest viral population collected for each patient. The number and location of PNGS were estimated using N-GlycoSite from the Los Alamos National Laboratory (Los Alamos, New Mexico, USA).

The instantaneous rates of nonsynonymous (dN) and synonymous (dS) substitutions were compared to evaluate the role of natural selection. Estimates of dN, dS, and dN/dS for each time point were obtained in Datamon-
key using the Fixed Effects Likelihood procedure [21]. Absolute rates of nonsynonymous [E(N)] and synonymous [E(S)] substitutions between mothers and the corresponding infants were compared using a codon-based extension of the Bayesian relaxed clock model implemented in BEAST (by A. J. Drummond and A. Rambaut, available from http://beast.bio.ed.ac.uk/), which incorporates sampling dates and phylogenetic uncertainty into the reconstruction procedure. HyPhy was used to perform the codon model analyses on a subsample of the posterior genealogies. E(N) and E(S) rates were summarized for internal branches only, since external branches could be biased by mutational load [22,23].

Results

Study subjects

Three MIP infected by HIV-1 at similar time points were identified during the follow-up period of the longitudinal cohort. The HIV-1 status of the three MIP at birth was negative (serology and PCR). They were followed longitudinally from delivery, and their follow-up specimens were serologically tested for HIV-1. The first positive sample for each MIP was found to be from the same visit in mother and child. The patients seroconverted 4 to 18 months after birth, and the first positive HIV-1 tests were 4 and 6 months after a negative HIV-1 test (Table 1). The first positive DNA PCR coincided with the first seropositive sample in all cases (Table 1) except infant 2953, where no specimen was available for nucleic acid extraction at the 11-month seroconversion time point. Viral load was not measured because properly frozen plasma was not available. As a surrogate for virus load, viral burden in peripheral blood mononuclear cells (PBMC) was determined using real-time PCR (Table 1). The first positive sample was found to have the highest viral burden in the three mothers. This provided an opportunity to determine whether patterns of variation associated with transmission during maternal PHI are similar to those observed in transmission during chronic infection. Furthermore, as mother and infant within a MIP were infected with very similar viruses, it was possible to compare molecular evolution of maternal and infant viral populations.

Subtyping and transmission of viral env

For each patient, viral env DNA samples were obtained from three to five different time points and 25 to 33 env clones were sequenced per sample, to give a total of 79 to 152 env clones per patient. All the clones sequenced correspond to HIV-1 subtype C, the most prevalent subtype in Zambia. In all three MIP, viral sequences from the mother and the corresponding infant were placed in monophyletic clades. Viral isolates from these MIP were readily distinguishable from other subtype C isolates from the Los Alamos HIV database or from our laboratory. The close affinity of mother and infant viral sequences provided evidence of an epidemiological link between the HIV-1 viral populations in each MIP (supplementary materials available online).

Table 1. Summary of clinical information.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Age (years) a</th>
<th>Sampling times (month) b</th>
<th>Seroconversion month c</th>
<th>PCR positive month d</th>
<th>Clinical information</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP334</td>
<td>26</td>
<td>0, 4, 12, 18</td>
<td>4</td>
<td>4</td>
<td>Weight: 49, 47, 48, 45, 48 kg at 2, 4, 6, 12, 18 months respectively; malaria at 6 and 18 months Papulée non-specific rash on lower limb at 4 months</td>
</tr>
<tr>
<td>Mother</td>
<td>Infant, male</td>
<td>0, 4, 6, 12, 18</td>
<td>4</td>
<td>4</td>
<td>Fever and headache at 2 and 18 months; weight loss at 18 months Sexually transmitted infection: chancroid at 2 months; vaginal discharge at 18 months Existing pulmonary TB; accepted anti-TB treatment from 30 months; malnourished-looking child</td>
</tr>
<tr>
<td>MIP2660</td>
<td>21</td>
<td>0, 6, 12, 18, 24, 30</td>
<td>18</td>
<td>18</td>
<td>Weight: 55, 57, 74, 69, 72, 65, 71 kg at 2, 6, 11, 18, 24, 30, and 39 months, respectively; abdominal pains at 11 months Antimalaria treatment at 5 and 11 months; chicken pox at 18 months; cough at 24 months; fever at 24 and 39 months</td>
</tr>
<tr>
<td>Mother</td>
<td>Infant, male</td>
<td>0, 6, 12, 18, 24, 30, 36</td>
<td>18</td>
<td>18</td>
<td>Weight: 55, 57, 74, 69, 72, 65, 71 kg at 2, 6, 11, 18, 24, 30, and 39 months, respectively; abdominal pains at 11 months Antimalaria treatment at 5 and 11 months; chicken pox at 18 months; cough at 24 months; fever at 24 and 39 months</td>
</tr>
</tbody>
</table>

TB, tuberculosis.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Age (years) a</th>
<th>Sampling times (month) b</th>
<th>Seroconversion month c</th>
<th>PCR positive month d</th>
<th>Clinical information</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP2953</td>
<td>22</td>
<td>0, 2, 5, 11, 18, 24, 30, 39</td>
<td>11</td>
<td>11</td>
<td>Weight: 55, 57, 74, 69, 72, 65, 71 kg at 2, 6, 11, 18, 24, 30, and 39 months, respectively; abdominal pains at 11 months Antimalaria treatment at 5 and 11 months; chicken pox at 18 months; cough at 24 months; fever at 24 and 39 months</td>
</tr>
<tr>
<td>Mother</td>
<td>Infant, female</td>
<td>0, 2, 5, 11, 18, 24, 30, 39</td>
<td>11</td>
<td>18 d</td>
<td>Weight: 55, 57, 74, 69, 72, 65, 71 kg at 2, 6, 11, 18, 24, 30, and 39 months, respectively; abdominal pains at 11 months Antimalaria treatment at 5 and 11 months; chicken pox at 18 months; cough at 24 months; fever at 24 and 39 months</td>
</tr>
</tbody>
</table>

a All infants were breastfed up to 18 months of age.

b Age of the mother in years at the time of delivery.

c Times are expressed in months after infant birth.

d No specimen was available for nucleic acid extraction at the 11-month time point.
Analyses of maternal and infant viral sequences from early time points showed that donor and recipient sequences did not form mutually exclusive groups in any of the three MIP, indicative of the passage of multiple viral lineages from the mother to the child. This lack of separation can be seen in the phylogenetic trees in Figure 1, where early mother and infant sequences within a MIP are intermixed in the base of the trees in all cases. In addition to the transmission of multiple viral genotypes, there was no loss of genetic diversity associated with transmission in these three MIP. The transmission of multiple viral lineages was corroborated by significant results of the approximately unbiased topology test suggested by Shimodaira [24], as implemented in Treefinder [19]. These analyses were restricted to the sequences from the earliest time points.

In MIP 834, the first sample positive by serology and PCR was collected 4 months after birth (Table 1). Genetic diversity in the initial viral populations was lower than 0.5% for both mother (0.41%) and infant (0.44%), implying a typical limited diversity infection from the original source, presumably a heterosexual contact for the mother (Table 2). No significant differences between the earliest maternal and infant viral populations were observed in the median number of PNSG, in the median V1-V5 length, or in estimates of dN/dS. For MIP 2660, the first positive sample for mother and infant was collected 18 months after birth. This MIP had the highest genetic diversity among the samples from the first virus-positive visit, which was 0.74% in the infant and 0.63% in the mother (Table 2). There were no significant differences in median number of PNSG or median V1-V5 lengths between mother and infant. Estimates of dN/dS for these populations were two-fold larger in the infant relative to the mother. For the third pair, MIP 2953, the first positive HIV-1 test was at 11 months after birth. However, the infant PBMC specimen from this visit was not available for DNA analysis and the estimate of genetic diversity in the maternal population at 11 months was 0.31%, the lowest among all samples tested (Table 2). A PBMC sample from infant 2953 was available at 18 months, which was the infant sample closest to the time of seroconversion. Therefore, this sample was used for genetic comparisons with the corresponding maternal sample at 18 months. Differences in genetic diversity, dN/dS, median number of PNSG, and median V1-V5 length were not significant (Table 2).

Longitudinal variation of viral env
In each of the MIP studied, there was an increase in genetic diversity with time (Table 2). The lowest genetic diversity was observed in viral populations from earliest time points, where genetic diversity ranged from 0.3 to 1.2% (Table 2). Conversely, the highest viral genetic diversity was observed in populations collected towards the end of the study, where genetic distance ranged from 0.8 to 2.1% (Table 2). Parallel to the accumulation of genetic diversity within a patient, estimates of genetic divergence, which measure overall genetic changes relative to the initial viral population, also increased with time, as did the distance between contemporaneous viral samples from mother and infant within a MIP.

Changes in the number of PNSG and V1-V5 length have been associated with a trade-off between transmission competence and resistance to immune challenges [6]. To evaluate this hypothesis, changes in the number of PNSG and V1-V5 length were examined. In the infants, median PNSG number decreased in infant 834, and increased in infants 2660 and 2953. Similarly, V1-V5 length decreased in infant 834 but increased in infants 2660 and 2953. For the mothers, the median number of PNSG decreased in
mother 834 but remained constant for mothers 2660 and 2953. With regard to V1-V5 length, it remained constant in mother 834, decreased in mother 2953, but oscillated in mother 2660. Therefore, variation in V1-V5 length was different between mother and the corresponding infant in all cases. Overall, there was no correlation between PNGS numbers or V1-V5 length with time for either infants or their mothers (Table 2).

A visual inspection of nonsynonymous changes over time showed that amino acid replacements were not evenly distributed along the Env fragment sequenced (Figure 2). Nonsynonymous changes accumulated in V1-V2, the 3’ end of C2, V3, the 5’ end of C3, V4, and in the 3’ half of C4 and V5. This coincides with the pattern of longitudinal variation observed in other infants from the same cohort [9]. Nonsynonymous changes accumulated in similar regions of env in both mother and infant, and a subset of codon positions were inferred to be under positive selection in both mother and infant (indicated by as terisks in Figure 2).

Finally, absolute rates of nonsynonymous and synonymous changes for each individual were characterized, integrating relaxed clock and codon model analysis [23]. Overall estimates of E(N), E(S) and E(N)/E(S) were higher in the infants relative to the mother in the three MIP (Table 3). Differences in E(N)/E(S) were primarily attributable to differences in the estimates of the rate of E(N), suggesting that the selective pressure on the virus was higher in the infants than in the mothers.

Discussion

In the three cases of MTCT during maternal PHI studied here, we observed the transmission of multiple closely related viral lineages, as evidenced by the statistically significant lack of monophy of infant sequences relative to mothers (Figure 1). In addition, we did not detect significant loss in genetic diversity associated with trans-
Fig. 2. Estimated number of nonsynonymous substitutions along the HIV-1 Env V1–V5 fragment sequenced, estimated using Datamony. Results are presented cumulatively for each mother–infant pair. The plot is restricted to the time points for which both mother and infant sequences are available. Numbers on the x axis correspond to the coordinates that define the various loops, V1, V2, V3, V4 and V5, which are shaded. Numbers in boxes identify positions inferred to be under positive selection in one or more of the time points comparisons; bolded numbers identify those positions inferred to be under positive selection in two or more time points; asterisks identify positions inferred to be under positive selection in both mother and infant.
mission, since levels of genetic diversity of the viral populations from the earliest comparable time point were statistically indistinguishable in comparisons between infants and mothers in all MIP. In comparison, only a small fraction of the viral quasispecies from the donor is generally transmitted to the recipient during chronic HIV-1 infection, both in MTCT settings [8,9,10,25], and in transmission among adults [6]. Similar losses in Env diversity upon HIV-1 transmission have been reported in other cohorts consisting of either sexually active adults or mother-to-child constituents [6,26]. Interestingly, our results support a recent study comparing the sexual transmission of HIV-1 subtype B from recent or chronically infected donors, which showed that infection from a chronically infected donor resulted in a loss of genetic diversity in the new host, but this did not occur if transmission was from a recently infected donor [14].

There are several possible explanations for the observed transmission of multiple viral lineages in these three MIP. One would be the presence of a viral population readily adapted for transmissibility in the mother, which is transmitted to the child without further selection. Alternatively, differences might result from the already limited viral sequence diversity observed in the three PHI donors. It would be difficult to differentiate between these two possibilities unless individuals with transmission from patients with a PHI and harboring a genetically diverse HIV-1 population can be identified and studied. A third possibility would be the result of multiple transmissions to the child through breastfeeding prior to our initial sampling time point. Our previous studies of MTCT from chronically infected mothers from this same cohort, and sampled at similar time points, showed that all infants received a restricted viral genotype [9,25]. If the explanation was multiple transmissions via breast milk, we would have observed multiple transmissions in those nine cases as well. However, our analyses cannot discriminate among these three possibilities.

The closely related viral quasispecies from the mother and infant within an MIP accumulated changes in similar positions of the fragment studied, in spite of coming from different hosts and in spite of the observed increase in divergence between them over time, and the differences between their hosts (Figure 2). These amino acid changes accumulated in areas that are highly variable among HIV-1 subtype C sequences, as identified by a previous study of longitudinal variation in infants from the same cohort infected with HIV-1 subtype C [9]. This parallel evolution of the viral populations in mother and infants suggests that the closely related viral quasispecies from mother and child within a MIP share evolutionary constraints. These constraints are also reflected in the analyses of selective pressure, which identified a subset of positions inferred to be under positive selection in both mother and infant within a MIP (balded positions in Figure 2).

Longitudinal studies of HIV-1 have reported that higher levels of selective pressure, measured as dN/dS, and higher levels of diversity are associated with slower rates of disease progression [9,27-31]. The higher dN/dS estimates have been linked to selection for changes in the virus by the host adaptive immune response, which includes neutralizing antibodies, cytolytic T cells against virus-infected cells, and/or the per-cell responses [32,33]. These changes will enable the virus to escape the immune response and are most likely to associate with longer survival times, as expected for the mothers in this study [29,30]. Contrary to our expectations, higher overall E(N)/E(S) estimates were found in the infants relative to the mothers in the three MIP analyzed (Table 3). These estimates suggest that environmental challenges to the virus are stronger in the infant.

This is the first comparative study of this type and it suggests that transmission during PHI involves multiple, closely related viral genotypes. In contrast to what was generally observed in nonacute transmission, we did not detect a loss in genetic diversity of the transmitted virus associated with transmission in these three cases of dual acute seroconversion. Even with the potential limitations, our data still suggests that there is no loss of genetic diversity during transmission from the PHI donors, probably because the viral population has recently been selected for transmission in the mothers. Moreover, our results showed that the newly transmitted viruses in

### Table 3. Absolute rates of nonsynonymous and synonymous genetic change estimated for each mother and infant.

<table>
<thead>
<tr>
<th></th>
<th>E(N)</th>
<th>E(S)</th>
<th>E(N)/E(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2660 M</td>
<td>1.20 × 10⁻⁶ (8.98 × 10⁻⁶-1.60 × 10⁻⁵)</td>
<td>4.97 × 10⁻⁴ (3.55 × 10⁻⁴-6.66 × 10⁻⁵)</td>
<td>2.45 (1.73-3.07)</td>
</tr>
<tr>
<td>834 M</td>
<td>5.84 × 10⁻⁴ (3.68 × 10⁻⁴-9.18 × 10⁻⁵)</td>
<td>4.07 × 10⁻⁴ (2.21 × 10⁻⁴-5.89 × 10⁻⁵)</td>
<td>1.46 (0.91-2.19)</td>
</tr>
<tr>
<td>834 M</td>
<td>1.01 × 10⁻³ (7.35 × 10⁻⁴-1.30 × 10⁻³)</td>
<td>8.34 × 10⁻⁴ (6.08 × 10⁻⁴-1.09 × 10⁻³)</td>
<td>1.22 (0.90-1.52)</td>
</tr>
<tr>
<td>2953 M</td>
<td>8.38 × 10⁻⁵ (5.08 × 10⁻⁵-1.23 × 10⁻⁵)</td>
<td>7.96 × 10⁻⁵ (5.94 × 10⁻⁵-1.14 × 10⁻⁵)</td>
<td>1.07 (0.71-1.45)</td>
</tr>
<tr>
<td>2953 M</td>
<td>1.29 × 10⁻³ (9.55 × 10⁻⁴-1.76 × 10⁻³)</td>
<td>4.97 × 10⁻⁴ (3.51 × 10⁻⁴-6.66 × 10⁻⁵)</td>
<td>2.67 (1.64-3.82)</td>
</tr>
<tr>
<td>2953 M</td>
<td>7.62 × 10⁻⁴ (5.94 × 10⁻⁴-9.61 × 10⁻⁴)</td>
<td>3.47 × 10⁻⁴ (2.64 × 10⁻⁴-4.53 × 10⁻⁴)</td>
<td>2.21 (1.73-2.73)</td>
</tr>
</tbody>
</table>

E(N), absolute rate of nonsynonymous substitution; E(S), absolute rate of synonymous substitution; M, mother; I, infant. Estimates (with their 95% highest posterior density intervals) were obtained using the procedure described in [73], excluding terminal branches, using different trees estimated under the relaxed clock procedure as described in [31].
both hosts, mother and infants, have undergone similar patterns of evolution, which probably reflects evolutionary-constaints shared between the mother and infant viruses. A better understanding of the factors underlying viral transmission during the acute phase can potentially lead to the development of better strategies to prevent viral transmission and disease progression.

Acknowledgements
We thank Robert Weldon, Hong Zhang, and Florencia Meyer for helpful discussions.

Sponsorship
This work was supported in part by PHS grants HD39620, RO1 CA75903, Fogarty International Training grant D43 TW01492, and NCRR COBRE grant P20 RR15635 to CW, and by a Postdoctoral Fellowship in Population Biology to FGH from the University of Nebraska. XH was a Fogarty Fellow.

Note: F. G. Hoffmann and X. He contributed equally to this paper.

References