HIV-1 subtype C envelope C2V5 characteristics; associations with markers of disease progression among 6 slowing progressing pediatric patients

Kerina Duri1*, Felicity Zvanyadza Gumbo2, Knut Ivan Kristiansen1, Munyaradzi Paul Mapingure3, Simba Rusakaniko4, Fredrik Müller3 and Babill Stray-Pedersen6
*Correspondence: tkduri@yahoo.co.uk

1Department of Immunology, University of Zimbabwe.
2Department of Pediatrics and Child Health, University of Zimbabwe.
3Department of Microbiology, University of Oslo and Oslo University Hospital, Rikshospitalet, Oslo, Norway.
4Letten Foundation Research Centre, Harare, Zimbabwe.
5Department of Community Medicine, University of Zimbabwe. 
6Division of Women and Children, Oslo University Hospital, Rikshospitalet and Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

Abstract
Background: HIV-1 gp120 envelope variable regions potential N-glycosylation sites (PNGs) and amino acid length polymorphisms have been shown to play pivotal roles in disease progression in spite of other studies demonstrating contradictory results.

Methods: Regression analysis was used to assess the association between changes in PNGs or amino acid sequence lengths of C2V5 region/sub-regions and disease progression as a function of HIV-1 RNA load or CD4% among antiretroviral therapy naïve subtype C infected but slowly disease progressing children.

Results: Unit increases in amino acid sequence lengths within the V3 region was associated with a 5 unit increase in CD4%, p=0.010. Unit increases in PNGs within the C2V5 region, V3 and V5 sub-regions were associated with 3, 9 and 8 unit increases in CD4%; p=0.041, 0.040 and 0.02, respectively. Interestingly, a unit PNGs increase within the C4 sub-region correlated with a 1.2 million copies/ml decrease in viral load, p=0.009.

Conclusion: C4 sub-region PNGs may be influential in viral replications whilst amino acid length polymorphism within the V3 regions could be key in host immunological surveillance. However, bigger studies with more clones per sample are warranted to substantiate these preliminary findings and whether such observations are also true for the adult the population.

Keywords: HIV subtype C, C2V5 sequence lengths, potential N glycosylation sites, children

Introduction
Disease progression in HIV infection demonstrates inter-individual variation due to the complex interplay between environmental, host genetic, immunologic and viral factors. Rapid HIV disease progression has been related to high maternal viral load, ineffectiveness of the cellular immune control, heterozygosity of host restrictive genes associated with delayed progression, timing of transmission including immaturity of the adaptive immune response of infants including viral characteristics [1-3]. Viral characteristics associated with slow HIV disease progression in pediatric patients remain poorly described and more over the little that is currently known comes from subtype B studies. Characterisation of HIV-1 env gene from slow progressors is not only important in understanding HIV-1 pathogenesis but is also fundamental in providing information on viral correlates of protection essential for effective vaccine design and development.

In the absence of antiretroviral therapy, most HIV-1 infected infants in Sub Saharan Africa progress to AIDS or death within two years of life [4]. Irrespective of the enhanced immune activation due to the heavy parasitic burden aggravated by nutritional deficiencies, recent studies have shown that a quarter of HIV-infected antiretroviral naive African infants are “slow progressors”, presenting for care for the first time during adolescence [5-7]. In Zimbabwe, an estimated 120000 children are living with HIV of which 3.4% are long term survivors of vertical transmission [8,9]. Despite being infected by the same mother, studies have shown that siblings have shown marked differences in the rate of disease progression [10] with in utero infected infants progressing more rapidly [11,12]. Children infected post-natally through breastfeeding have been shown to be slow progressors [13,14]. Studies have shown that the route and timing of infection alone are not predictive of disease progression in infected infants [15]. HIV-1 gp120 variable regions exhibit an unusual pattern with encoded amino acids (asparagine, serine, and threonine) that leads to the creation of new N-linked glycosylation sites, which helps the virus to escape from the immune pressure [16]. Potential N glycosylation sites (PNGs) glycans and amino acid length polymorphism have been shown to play essential roles in disease progression.
with changes potentially affecting the capacity of the virus to replicate [17-19]. Studies have shown that at birth all HIV infected infants’ envelope amino acid sequences are initially of the same length as the maternal ones with amino acid sequence length polymorphism observed from six months postpartum [20]. However, other studies have observed no clear trends in changes in neither the envelope PNGs nor median amino acid lengths with disease progression [20,21]. In that view we sought to determine HIV-1 env gp120 sequence characteristics associated with pediatric disease progression assessed as a function of laboratory markers of CD4 percentage and HIV-1 RNA load.

**Methods**

**Study population and procedures**

Four index children (older/ first sibling) and two index child’s siblings (younger or second siblings) from HIV-1 infected families labeled 205, 366, 375 and 567 constituted the study population. The index child in this study was defined as the first child to be recruited into our study. Two families, 205 and 567 had two subsequent biological children, the first and second siblings. All the children were HIV-1 infected and none had received antiretroviral therapy at the time of sample collection. Consent was obtained from the respective pregnant mothers of each of the four families participating in the national PMTCT programme in peri urban Harare mother and child clinic who were known to be HIV-1 positive at 36 weeks gestations. Similar recruitment and procedures were followed as previously described for the mothers and infants [22]. Despite being encouraged to exclusively breastfeed during the first six months of life, all the infants were exposed to breast milk for at least nine months. First siblings’ samples were collected at 60±10 months of age as there were insufficient sample volumes from their respective first HIV positive samples. The first available HIV-1 positive sample was genotyped for the second siblings at about 15±3 months. CD4 counts, viral load determination, nucleic acid extraction, PCR amplification, cloning and DNA sequencing methods for the HIV-1 env gp120 C2V5 region were done as previously described and so was subtype determination [23].

**Data analysis**

The 520 base pair nucleotide sequences were translated to amino acid sequence using the Gene Doc program. The amino acid sequences in their Fasta formats were entered into a glycosylation analysis site: http://www.hiv.lanl.gov/content/hiv-db/GLYCOSITE/glycosite, where PNGs along the sequences were marked and counted. Numbers of PNGs of C2V5 region and sub-regions were enumerated. Median and range values of sequence length and PNGs were calculated for child as previously described [24]. An average of two clones was done per sample. Sequences were aligned, manually edited and separated into C2, V3, C3, V4, C4 and V5 sub-regions. Sequence sub-region lengths were determined and PNGs counted. Data were entered and analysed using Stata version 10. The number of PNGs and including sequence length polymorphisms of C2V5 and sub-regions were determined for each child. Regression analysis was used to investigate the association between PNGs or amino acid sequence lengths polymorphism of the C2V5 region/ sub-regions and viral load or CD4 %. Student t-test was used to compare mean log_{10} viral load between the male and female children. Tests of statistical significance included the 95% confidence interval (CI) of relative risks; two sided p values of less than 0.05 were considered statistically significant.

**Ethical consideration**

The study was approved by the Medical Research Council of Zimbabwe (MRCZ) and the Ethical Review Committee of Norway. Written consent to participate in the research study was obtained from the mothers on behalf of the minors. There was freedom to discontinue participating in the study at any given time without any prejudice. Mothers also consented to usage of blood samples of their children in future other HIV related researches.

**Results**

**Demography, clinical presentation**

Of the six children the 4 first siblings were about 50 months old whilst the two second siblings were about 20 months old. They were equally balanced in respect to gender but presented difference clinical abnormalities ranging from lymphadenopathy to Mycobacterium tuberculosis (TB) infection. See **Table 1**. Almost all of the parameters analyzed varied substantially from child to child, including CD4 count, CD4%, and viral RNA load. The CD4% of all the children

**Table 1. Siblings’ demographics, virological, immunological parameters and clinical abnormalities.**

<table>
<thead>
<tr>
<th>Sibling</th>
<th>Age (Years)</th>
<th>Sex</th>
<th>CD4 count</th>
<th>CD4 %</th>
<th>Viral load Copies/mL</th>
<th>Abnormal clinical signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>205 1st</td>
<td>5.3</td>
<td>M</td>
<td>3170</td>
<td>23</td>
<td>100110</td>
<td>none</td>
</tr>
<tr>
<td>205 2nd</td>
<td>2</td>
<td>F</td>
<td>512</td>
<td>41.2</td>
<td>254676</td>
<td>underweight</td>
</tr>
<tr>
<td>366 1st</td>
<td>3.9</td>
<td>F</td>
<td>197</td>
<td>25</td>
<td>138628</td>
<td>Scabies, dysentery, TB</td>
</tr>
<tr>
<td>375 1st</td>
<td>3.7</td>
<td>F</td>
<td>3156</td>
<td>30</td>
<td>3232925</td>
<td>lymphadenopathy</td>
</tr>
<tr>
<td>567 1st</td>
<td>4.5</td>
<td>M</td>
<td>1849</td>
<td>19.5</td>
<td>1736</td>
<td>T. capitis</td>
</tr>
<tr>
<td>567 2nd</td>
<td>1.5</td>
<td>M</td>
<td>1588</td>
<td>21.7</td>
<td>353418</td>
<td>kwashikor</td>
</tr>
</tbody>
</table>

1st: First sibling  
2nd: Second
The correlation between PNGs and viral load or CD4% was significant in children. There were no statistically significant differences in the rate of CD4% decline and subsequent amino acid length polymorphism in all sub-regions except for the V3 region. Interestingly, each unit increase in V3 amino acid length was associated with a 5 units increase in CD4%, p=0.010, see Table 3. A unit increase in amino acid length within the V5 region was associated with 320000 copies increase in HIV-1 RNA load, p=0.003, otherwise no other statistically significant associations were observed in other sub-regions.

Discussion
Slowly progressing HIV-1 infected but treatment naïve pediatric patients afforded us the opportunity to assess the natural immune selection of env gene after years of infection. In this paper we assessed whether immunological progression and virological responses of HIV-1 subtype C correlated with viral env gp120 C2V5 region PNGs and amino acid length polymorphism. The Paediatric HIV-1 RNA load remains high during the first years of life, gradually declining slowly only after development of cytotoxic T Lymphocytes (CTL) response after the age of 6 months [25]. In our study, a unit increase in glycan was associated with decreases in viral load although a different picture was depicted by child 567 first sibling who surprisingly despite the fewer PNGs had the least viral load (1700 copies/mL) who interestingly was slightly over five years of age without treatment. This observation could probably be a case of attenuated HIV variants with low replicative capacity linked to non-progressive disease [26,27]. Against all odds this child had also T. capitis infection which under normal circumstances would be expected to enhance viral replication due to immune activation thereby facilitating a faster progression to AIDS [28,29]. Slow progressing infants typically have virus with a low replicative capacity which may be the influence of viral attenuation the selection of Gag CD8 T-cell escape variants [30].

A contradictory trend was observed for V5 region amino acid length and viral load. This observation is complimenting recent findings which have demonstrated that increased sequence length and glycosylation in in the V5 region may collectively create steric hindrance that lowers binding affinity, thereby increasing resistance to neutralization [31]. A subtype C study of a single Zambian HIV infected but slowly progressing infant showed that neither lengthening of the V1V5 domain nor the acquisition of glycosylation site had much better than the second sibling, the rate of immunological progression was worse for the first siblings who have been infected for relatively longer periods compared to the first siblings. Interestingly, child 567 first sibling’s mean amino acid length was not significantly different from that of other children with viral load in hundred thousands of RNA copies per mL. Even more interesting was the comparison of measured parameters two the 567 sibling pairs of which the first sibling, 567-1 seemed to be doing much better than the second sibling, 567-2 based on the data presented in Table 1. Despite their 3 years age difference they had similar CD4 counts and CD4%.

Table 2. Relationship between PNGs, Amino acid lengths and viral load (copies/mL) in children.

<table>
<thead>
<tr>
<th>Region</th>
<th>PNGs</th>
<th>Amino Acid Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>569544</td>
<td>0.299</td>
</tr>
<tr>
<td>C3</td>
<td>-56432</td>
<td>0.880</td>
</tr>
<tr>
<td>C4</td>
<td>-1151683</td>
<td>0.009</td>
</tr>
<tr>
<td>V3</td>
<td>-483628</td>
<td>0.347</td>
</tr>
<tr>
<td>V4</td>
<td>158145</td>
<td>0.689</td>
</tr>
<tr>
<td>V5</td>
<td>298808</td>
<td>0.479</td>
</tr>
<tr>
<td>C2V5</td>
<td>-65563</td>
<td>0.689</td>
</tr>
</tbody>
</table>

Table 3. Relationship between PNGs, amino acid length and CD4 % in Children.

<table>
<thead>
<tr>
<th>Region</th>
<th>PNGs</th>
<th>Amino Acid Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>+5.323</td>
<td>0.229</td>
</tr>
<tr>
<td>C3</td>
<td>+4.10</td>
<td>0.171</td>
</tr>
<tr>
<td>C4</td>
<td>-4.326</td>
<td>0.277</td>
</tr>
<tr>
<td>V3</td>
<td>+8.283</td>
<td>0.040</td>
</tr>
<tr>
<td>V4</td>
<td>-0.584</td>
<td>0.858</td>
</tr>
<tr>
<td>V5</td>
<td>+7.621</td>
<td>0.025</td>
</tr>
<tr>
<td>C2V5</td>
<td>+2.590</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Correlation between PNGs and viral load or CD4%
There was no clear relationship between children’s PNGs within neither the C2V5 region nor C2, C3, V3, V4 including V5 sub-regions and viral load. However, each unit increase in PNGs within the C4 sub-region was associated with a 1.2 million copies decrease in viral load, p=0.009. This relationship remained statistically significant even after adjustment for potentially important confounders such as age and gender. For a unit increase in PNGs within the entire C2V5 region including the V5 and V3 sub-regions, the CD4 percentage increased by about 3, 8 and 9 units p=0.041, 0.025 and 0.040 respectively See Table 2.

Correlation between amino acid length polymorphism and viral load or CD4%
There were no statistically significant differences in the rate of CD4% decline and subsequent amino acid length polymorphisms in all sub-regions except for the V3 region. Interestingly each unit increase in V3 amino acid length was associated with a 5 units increase in CD4%, p= 0.010, see Table 3. A unit increase in amino acid length within the V5 region was associated with 320000 copies increase in HIV-1
yet differed by over 2 log10 in viral load. Thus the older child was quite close to meeting the criteria for being considered an elite controller. This is not surprising as disease progression is influenced by many other host genetics factors such as the genotype at the major histocompatibility complex class I loci, polymorphism in CCR5, SDF-1, apolipoprotein B mRNA editing enzyme catalytic polyprotein-like 3G (APOBEC3G), ring finger protein 39 and the zinc ribbon domain-containing genes [32-34]. Moreover other subtype C studies have attributed slower disease progression to altered Nef protein functions [35]. Our children are generally slow progressors as all of them were infected post-natally through breast feeding as similarly described by others [13,14]. This timing of HIV infection could have a critical impact on the anti-HIV immune response as the immune system has been shown to undergo major changes during the postpartum period, shifting away from active acquisition of self-tolerance in utero towards defensive responses to antigens encountered postnatally [36].

The observation of slower HIV disease progression in African patients is in apparent contradiction with the observation that the typical rate of disease progression seems to be faster in Africa than among patients of European descent in industrialized countries [37,38]. Recent findings have shown that the lack of Duffy antigen receptor for chemokines (DARC) on the red blood cells of patients of African descent is associated with slower disease progression [39]. We have no data on DARC expression in our patients, although the DARC negative phenotype is almost universal in sub-Saharan Africa [39]. Given the substantial genetic diversity present in the quasi-species of chronically HIV infected patients the few sequences analyzed per child compounded by lack of longitudinal data are the major limitations. Despite the interesting preliminary findings the overall sample size is also the major challenge of this study. Thus larger studies are warranted with enough statistical power to infer more meaningful associations of the parameters investigated.

Conclusion
Sequencing of the HIV-1 subtype C envelope and assessment of PNs is important as this may lead to identification of functional sites which can shed a little bit more light on correlations immunological and virological relevant sites with important implications in both drug and vaccine designs and development. C4, V3 and V5 regions prospective regions worth further exploration on a bigger prospective cohort.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
DK collected data, carried out the laboratory analysis and drafted the manuscript, GFZ collected data, KKI participated in laboratory analysis, MMP carried out data analysis and interpretation of results, RS supervised data analysis and interpretation of results, designing of the study, MF supervised laboratory analysis, SB participated in designing of the study. All authors read and corrected the final version of the manuscript.

Acknowledgement
We gratefully acknowledge the women and infants who participated in this study and the study support staff. The study was funded by The Letten Foundation, Oslo, Norway. Special mention goes to the Letten Foundation and Professor Letten herself for funding this study.

Publication history
Editor: Qingming Dong, University of Tennessee Health Science Center, USA.
EIC: Preet M Chaudhary, Keck School of Medicine USC, USA.
Received: 15-Feb-2013 Accepted: 29-Jul-2013
Published: 02-Aug-2013

References


Citation: