HIV-1 subtype C envelope (Env) region amino acid length polymorphism and glycosylation sites variation; correlations with markers of disease progression among 7 antiretroviral therapy naïve heterosexuals

Duri K.1, Kristiansen K. I.2, Mapingure M. P.3, Chirenje M. Z.4 Stray-Pedersen B.5, Müller F.2

1Department of Immunology, University of Zimbabwe College of Health Sciences, Harare, Zimbabwe.
2Division of Microbiology, Rikshospitalet, Oslo University Hospital and University of Oslo, Oslo, Norway.
3Department of Community Medicine, University of Zimbabwe College of Health Sciences, Harare, Zimbabwe.
4Department of Obstetrics and Gynecology, University of Zimbabwe College of Health Sciences, Harare, Zimbabwe.
5Division of Women and Children, Rikshospitalet, Oslo University Hospital and University of Oslo, Oslo, Norway.

Accepted 17 January, 2013

HIV env gp120 characteristics associated with disease progression have so far shown conflicting results, yet such information may provide insight into HIV-1 vaccine design. Env C2V5 characteristics associated with immunological and virological markers of disease progression were assessed among 7 HIV-1 subtype C infected but antiretroviral therapy naïve heterosexuals. Amino acid sequence lengths and potential N-glycosylation sites (PNGs) were counted and regression analysis was used to assess associations with plasma viral load and CD4 count. Each unit increase in PNGs or amino acid sequence length within C2V5 region was associated with a 340000 or 110000 copies/ml decrease in viral load, p=0.001 or 0.005, respectively. Viral load suppressions of 720000 and 2 000000 copies/ml were observed for each unit increase in PNGs within the C3 and C4 sub-regions; p=0.035 and <0.001, respectively. Each unit increase in amino acid sequence length within the V3 region correlated with a 129 cell increase in CD4 count. Increases in PNGs within the C3 and C4 sub-regions could be central in viral replication whilst increase in amino acid sequence lengths within the V3 sub-region may be essential in immunological recovery. However, bigger longitudinal studies are necessary to substantiate these findings.

Key words: HIV-1 gp120 env C2V5, potential N glycosylation sites, amino acid length polymorphism, disease progression.

INTRODUCTION

Structural plasticity of the viral envelope glycans has been shown to potentially influence the capacity of HIV-1 to replicate (Yoo et al., 2010; Poumbourios et al., 2003). Any modification in the glycosylation profile particularly the potential N-glycosylation site (PNGs) has been shown to affect viral susceptibility to host immune responses and HIV/AIDS disease progression (Sagar et al., 2006; Ganeshan et al., 1997; Hill et al., 2004). Studies have shown increased amino acid sequence lengths and PNGs throughout chronic HIV-1 infection which declines with disease progression (Bunnik et al., 2008; Kitirinos et al., 2003; Palmer et al., 1996; Wang et al., 2000; Archary et al., 2010). Contrarily, some studies have not observed this association (Abrahams et al., 2009; Hughes et al., 1997; Chohan et al., 2005). Thus, env characteristics contributing to disease progression remain unclear yet such information is important for HIV-1 vaccine design and development and in this regard the importance of such studies in different populations may not be over-emphasized.

Most of the currently available knowledge on HIV env gp120 characteristics associated with differential disease
progression is biased towards subtype B variants, focusing mainly on the V1/V2 sub-regions (Curlin et al., 2010). Extrapolation of such findings to subtype C and other HIV-1 gp120 sub-regions other than V1/V2 regions remains to be substantiated. There is a paucity of such data with respect to V3 including V4 and virtually nothing for subtype C V5 sub-region which contributes about 64% of all worldwide HIV-1 infections (Curlin et al., 2010). Our previous studies have shown that at least in HIV/AIDS pediatric patients the C4 sub-region’s PNGs may be influential in viral replications whilst V3 region amino acid sequence length polymorphism could be key in host immunological surveillance. Further authentication is necessary to ascertain whether such findings could be extrapolated to the adult population. Consequently, we sought to determine the relationship between the entire C2V5 env gp120 region including individual sub-regions viral characteristics and disease progression measured as a function of CD4+ T cell loss and increase HIV-1 subtype C plasma RNA load among seven antiretroviral naïve heterosexuals.

METHODS

Study population and procedures

Study population was part of the parents of HIV-1 infected families that participated in the phylogenetic study labeled 205, 366, 375 and 567. In one of the families, 366 the father figure was missing as he was working in another regional country. Consent was initially obtained from the four pregnant mothers of each family 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.

None had received antiretroviral therapy at the time of sample collection. Three of the four mothers were in monogamous marriage except for mother 375 who was in a polygamous relationship. All had at least 7 years in school and were of low socio-economic status. Mothers were generally 5 years younger than their male counterparts. Sexual partners did not know when and how they got infected. However, mode of HIV-1 acquisition was most likely heterosexually as none mentioned any history of blood transfusion, drug abuse, or homosexuality except for mother, 366, who had a history of blood transfusion. One lifetime sexual partner was generally reported except for mother 375 who reported two. All mothers had negative results for the rapid plasma reagin (RPR) syphilis test but two of the four mothers were Herpes simplex virus type 2 (HSV-2) positive. Mothers 375 and 567 reported itchy genitals as previously described (Duri et al., 2012).

Similar recruitment and procedures were followed as previously described for the mothers (Gumbo et al., 2011). Nucleic acid extraction, PCR amplification, cloning and DNA sequencing methods for the HIV-1 env gp120 C2V5 region were done on plasma samples as previously described and so was HIV-1 subtype determination (Duri et al., 2011). CD4 counts and viral load were also determined as previously described (Duri et al., 2010). Briefly PCR amplification was done using first round amplified an 800 bp fragment using outer sense and antisense primers, 50-GTCAGCACAGTACATGTACAT-30 and 50-GGCCCATATGCTTCTGGTGC-30, respectively. Secondary PCR amplified an approximately 535-bp env gene fragment using inner sense and antisense primers 50-AATATGYACACATGAATTARGCCA-30 and 50-GAGGGCCATACATTGCCA-30, respectively. Amplicons were cloned in an nitrogen TA TOPO vector. An average of 4 randomly selected colonies was sequenced using the big dye terminator standard protocol.

Sequence and data analysis

Clustal X program was used to align sequences with manual editing. C2V5 amplified region was divided into 3 constant sub-regions; C2, C3, C4 and three hypervariable V3, V4 and V5 sub-regions corresponding to HXB2 6948-7109, HXB2 7218-7376, HXB2 7479-7556 and HXB2 7110-7217, HXB2 7377-7478 and HXB2 7557-7637. Regressions were conducted assessing the number of amino acid lengths or PNGs as a function of viral load in copies/ml and CD4 count. 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 as previously described (Derdeyn et al., 2004).

Data were entered and analysed using Stata version 10. The number of PNGs and including sequence length of C2V5 and sub-regions were determined. Regression analysis was used to assess the association between PNGs, amino acid sequence lengths of C2V5 region/ sub-regions and viral load or CD4 %. The Student t-test was used to compare mean log_{10} viral load after stratification by gender. 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...
Table 1. Relationship between PNGs, Amino acid lengths and viral load in adults

<table>
<thead>
<tr>
<th>HIV-1 Env</th>
<th>PNGs</th>
<th>Amino acid length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regression coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>C2</td>
<td>7430</td>
<td>0.991</td>
</tr>
<tr>
<td>C3</td>
<td>-71362</td>
<td>0.035</td>
</tr>
<tr>
<td>C4</td>
<td>-1923261</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>V3</td>
<td>529042</td>
<td>0.257</td>
</tr>
<tr>
<td>V4</td>
<td>-4129115</td>
<td>0.090</td>
</tr>
<tr>
<td>V5</td>
<td>543968</td>
<td>0.128</td>
</tr>
<tr>
<td>C2V5</td>
<td>-342787</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 2. Relationship between PNGs, amino acid length and CD4 count in adults.

<table>
<thead>
<tr>
<th>HIV-1 Env</th>
<th>PNGs</th>
<th>Amino acid length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regression coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>C2</td>
<td>+24.32</td>
<td>0.746</td>
</tr>
<tr>
<td>C3</td>
<td>+99.714</td>
<td>0.005</td>
</tr>
<tr>
<td>C4</td>
<td>+114.89</td>
<td>0.007</td>
</tr>
<tr>
<td>V3</td>
<td>+0.900</td>
<td>0.02</td>
</tr>
<tr>
<td>V4</td>
<td>+44.502</td>
<td>0.092</td>
</tr>
<tr>
<td>V5</td>
<td>+26.343</td>
<td>0.504</td>
</tr>
<tr>
<td>C2V5</td>
<td>+33.125</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Council of Zimbabwe (MRCZ) and the Ethical Review Committee in Norway. Written consent to participate in the HIV-1 genetic research study was obtained from the parent(s) who also consented on behalf of their minors. Study participants were free to discontinue at any time without any prejudice. Parents also consented to have their blood samples be used for other future HIV related research studies.

RESULTS

HIV-1 gp120 env PNGs and amino acid lengths correlations with viral load

From the plasma samples of seven adults, four females and 3 males, a total 50 sequences were cloned of which a total of 22 unique clones were analysed. There were no statistically significant relationships between PNGs along the C2, V3, V4 or V5 sub-regions and viral load. However, each unit increase in PNGs within the entire C2V5 region, including C3 and C4 sub-regions was associated with about 340000, 720000 and 2 000000 HIV-1 RNA copies/mL decrease; p=0.001, 0.035 and <0.001, respectively (Table 1).

Amino acid sequence length increases within the entire C2V5 region correlated negatively with viral load. Each unit increase in amino acid length correlated with an 110000 RNA copies per ml decrease in viral load, p=0.005. Statistically significant negative viral load correlations of over 2 million and 200000 copies per ml were also observed for each unit increase in amino acid sequence length within the V3 and V4 sub-regions, p=<0.001 and 0.012, respectively.

HIV-1 gp120 env PNGs and amino acid lengths correlations with CD4 count

Contrary to the previously observed negative association with viral load, positive correlations were observed between PNGs within the C2V5 entire region including V3 region and CD4 counts.

For a unit PNGs increase within the V3, C2V5, (sub) regions, the CD4 count increased by 1 and 34, cell(s); p=0.02 and0.005, respectively. Interestingly, PNGs increases within the constant regions C3 and C4 were associated with the most dramatic increases in CD4 counts; 100 and 115 cells, p=0.005 and 0.007, respectively. Unit increase in amino acid sequence length within the V3 sub-region was associated with 129 cells increase in CD4 count. However, there was a negative correlation of amino acid sequence length and CD4 count with regard to the V5 region with each unit increase in amino acid associated with 39 CD4 cells decrease per/µl (Table 2).
DISCUSSION

Similar to our previous pediatric study, increases in PNGs or amino acid lengths within the C4 and V3 sub-regions were associated with dramatic increases in CD4 count and decrease in plasma viral RNA. Unlike in children, in addition to the C4 region, the C3 region and the entire C2V5 PNGs variation were also implicated in viral control in adults. Likewise, some gender differences with respect to virus replication were also observed with females generally having on average higher copies of viral load relative to their male counterparts. It is not very clear why such differences were observed. Since it was a small sample this observation could be due to chance hence a bigger studies are required in order to sustain this conclusion.

An increase in amino acid sequence length within V3 region had the most negative effect on virus replication. This is expected as the amino acid sequence elongation has been shown to be associated with adaptation as measured by improved immunological markers such as increase in CD4 count. A relatively recent meta-analysis study of over 5000 HIV-1 env gene sequences publicly available from the Los Alamos database assessing envelope sub-regions length variation and disease progression has demonstrated lots of data for subtypes A, B with virtually nothing for HIV-1 subtype C regions V4 and V5 (Curlin et al., 2010). Our study is unique in the sense that it reported on subtype C V4 and V5 sub-regions sequence lengths including glycans associations with markers of disease progression. These preliminary results of the V5 region are suggestive of contradictive or inconclusive observations regarding amino acid length correlations and CD4 counts. Characterization of glycosylation at the sequence level may not adequately address the issue of disease progression as carbohydrates are structurally much more complex molecules. Thus more innovative computer based framework of bioinformatics tools should complement such studies. Bigger studies with more clones per sample are warranted to substantiate these interesting preliminary findings.

ACKNOWLEDGEMENTS

We gratefully acknowledge the women and infants who participated in this study and the study support staff. Special mention goes to the Letten Foundation and Professor Letten herself for funding this study.

REFERENCES


Gumbo FZ, Kandawasvika GQ, Duri K et al.: Reduced HIV transmission at subsequent pregnancy in a resource-poor setting. Trop Doct 2011.


Kitrinos KM, Hoffman NG, Nelson JA, Swanstrom R:


