Analysis of HIV Diversity Using a High-Resolution Melting Assay

William I. Towler,1 Maria M. James,1 Stuart C. Ray,1 Lei Wang,2 Deborah Donnell,2 Anthony Mwatha,2 Laura Guay,1 Clemensia Nakabiito,3 Philippa Musoke,3,4 J. Brooks Jackson,1 and Susan Eshleman1

Abstract

HIV viruses are usually genetically homogeneous shortly after infection, and become more heterogeneous over time. We developed a high-resolution melting (HRM) assay to analyze HIV diversity without sequencing. Plasma samples from the HIVNET 012 trial were obtained from nine Ugandan mother–infant pairs. DNA amplified from the HIV gag region was analyzed to determine the number of degrees over which the DNA melted (HRM score). HRM gag DNA was also cloned and sequenced (50 clones/mother; 20 clones/infant). The median HRM score for infants (4.3, range 4.2–5.3) was higher than that for control plasmids (3.4, range 3.2–3.8, p < 0.001) and lower than that for mothers (5.7, range 4.4–7.7, p = 0.005, exact Wilcoxon rank sum test). The intraclass correlation coefficient reflecting assay reproducibility was 94% (95% CI: 89–98%). HRM scores were also compared to sequenced-based measures of HIV diversity; higher HRM scores were associated with higher genetic diversity (p < 0.001), complexity (p = 0.009), and Shannon entropy (p = 0.022), but not with length variation (p = 0.111). The HRM assay provides a novel, rapid method for assessing HIV diversity without sequencing. This assay could be applied to any region of the HIV genome or to other genetic systems that exhibit DNA diversity.

Introduction

Current studies suggest that one or a few HIV variants usually initiate infection, and that the immune response and other selective forces then drive the evolution of viral variants within an infected person, generating a complex population of related viral quasispecies.1–3 Several factors promote rapid HIV evolution, including large viral population size, rapid viral turnover, lack of proof-reading by HIV reverse transcriptase, and a high rate of genetic recombination.2 Changes in HIV diversity have been associated with different stages of HIV disease.1–4 The rapid evolution and high rate of genetic diversity of HIV viruses also complicate HIV therapy and vaccine development.1–3

HIV diversity is usually studied by analyzing sequences from individual HIV variants using phylogenetic or other methods. Individual viral sequences can be obtained by analyzing HIV after cloning, by single genome sequencing, or by high-throughput sequencing methods, such as “deep” pyrosequencing.5 HIV diversity has also been studied using heteroduplex mobility assays.6 Use of these methods may be limited by cost, effort required, or the complexity of analysis. The availability of a simple, rapid method for quantifying the level of HIV diversity could facilitate studies of HIV transmission and pathogenesis. In this report, we present a novel approach for analysis of HIV diversity that involves analysis of the melting profile of DNA amplified from a defined region of the HIV genome.

Materials and Methods

Source of samples used for analysis

Plasma samples were obtained from Ugandan women and infants enrolled in the HIVNET 012 trial.7,8 In HIVNET 012, HIV-1-infected, antiretroviral drug-naïve women received a single dose of nevirapine (sdNVP) during labor, and their infants received sdNVP shortly after birth to reduce the risk of HIV mother-to-child transmission. We analyzed samples collected from women prior to NVP exposure and samples from their HIV-infected infants collected at 6–8 weeks of age. Samples from nine mother–infant pairs were available for analysis. We also analyzed HIV-1 plasmids as controls.

1Johns Hopkins University School of Medicine, Baltimore, Maryland.
2Fred Hutchinson Cancer Research Center, Seattle, Washington.
3Makerere University–Johns Hopkins University (MUJHU) Research Collaboration, Kampala, Uganda.
4Makerere University School of Medicine, Kampala, Uganda.
Preparation of DNA templates for the HRM assay

DNA templates for the HRM assay were prepared from the plasma samples described above using the ViroSeq HIV-1 Genotyping System v2 (ViroSeq, Celeria, Alameda, CA). Briefly, HIV RNA was extracted from 500 μl of maternal plasma or 100 μl of infant plasma and reverse transcribed. A nonnested polymerase chain reaction (PCR) was used to amplify a PCR product encoding a portion of HIV gag, HIV protease, and a portion of HIV reverse transcriptase. The amplified DNA was purified using spin columns, analyzed by agarose gel electrophoresis, and diluted according to the manufacturer’s instructions. In the ViroSeq system, the diluted PCR products are sequenced to identify antiretroviral drug resistance mutations. In this study, PCR products that remained after genotyping were stored at –80°C and used as templates for the high-resolution melting (HRM) assay.

HRM assay

A region of HIV $gag$ was amplified from the PCR products prepared in the ViroSeq system (see above). Each 10 μl HRM amplification reaction included 1 μl of a 1:10 dilution of ViroSeq PCR products or 5 ng of plasmid control (template DNA). 0.2 μM forward and reverse primers, and 1× LightScanner Master Mix amplification buffer (Idaho Technologies, Salt Lake City, UT), which contains Taq polymerase and LCGreen Plus dye, which is incorporated into the amplified PCR products. The primer sequences were forward: HRM-3F: 5'-CTGAAACAGCTATGACCA), 5'-Sequencing Buffer (Applied Biosystems), and 1 mM dNTPs. HIV $gag$ sequences were obtained for 50 clones from each maternal sample and 20 clones from each infant sample.

Sequence-based analysis of HIV diversity

Phylogenetic methods were used to confirm that the sequences from clones derived from each mother-infant pair grouped together (50 sequences from each maternal sample, 20 sequences from each infant sample, see above), without evidence of a sample mix-up. Sequences from each plasma sample were aligned using MegAlign (DNASTar, Madison, WI) and manually edited to remove and/or align gaps in the sequences. The sequences from maternal samples ($n = 450$) and from infant samples ($n = 180$) were then aligned to a reference alignment (hiv.lanl.gov, subtype reference alignment version 2008) using the Profile Alignment method implemented in Clustal X2, with minimal manual editing to preserve highly conserved motifs. After alignment, sequences were trimmed to shared 5’ and 3’ termini. Sequence length was determined for each sequence as the number of nongap residues between those shared termini. For a given set of sequences (i.e., the set of clonal sequences from each sample), complexity was calculated as the number of unique sequence patterns $n$ divided by the number of sequences $N$ in that set (i.e., $n/N$), counting insertions or deletions as differences. Normalized Shannon entropy,$^{10,11}$ was calculated in a manner similar to complexity, except that for the $n$ unique sequence patterns (individually represented $i$) observed at a frequency $p_i$ among $N$ sequences, the normalized Shannon entropy was calculated as $S = -(1/ \log N) \sum_{i=1}^{n} p_i \log p_i$. Therefore, entropy takes into account the proportion or frequency of unique sequences in the set. For example, if a set of 50 sequences has two unique sequences at frequencies of 1/50 and 49/50, and another set of 50 sequences has two unique sequences, each at a frequency of 25/50, the complexity of both sequence sets would be the same (2/50, or 4%), but the entropy would be higher for the second set ($0.025$ vs. $0.177$). Diversity was calculated for each set of sequences after removing sites containing gaps; diversity was calculated using Mega (version 4.1, http://www.megasoftware.net) to apply the maximum composite likelihood distance model with rate variation among sites estimated using the gamma distribution (with alpha parameter 1.0).

Statistical analysis

The Wilcoxon rank sum test was used to compare median HRM scores between different sample sets (e.g., infants, mothers, control plasmids). Pearson correlation coefficients were calculated to describe the linear relationship between results from the HRM assay and other sequence-based measures of HIV diversity, and to evaluate the dependence between the HRM scores from the same mother-infant pair. To evaluate assay reproducibility, each sample was analyzed four times over the course of a year. The intraclass correlation coefficient, defined as a proportion of the total variance due to between-subject variation, for the HRM score was calculated.
as a measure of assay reproducibility. To assess the correlation between HRM score and HIV RNA viral load, subjects were categorized to either a high or a low viral load group by their log10 transformed viral load using a cutoff value of 5. The Wilcoxon rank sum test was used to compare HRM scores between the two groups. All statistical analyses were performed using SAS version 9.1.3 on the SunOS 5.9 platform.

Human subjects consent

Guidelines of the U.S. Department of Health and Human Services and the authors’ institutions were followed in the conduct of this research. Informed consent was obtained from all women for participation in the HIVNET 012 trial. Approval for this research was obtained from the Institutional Review Boards in Uganda and at the Johns Hopkins University School of Medicine.

Results

We developed and evaluated an HIV diversity assay based on analysis of the melting of DNA duplexes. Using the LightScanner instrument, the gag amplicons from a sample were heated, causing the DNA duplexes to melt and release the fluorescent dye that was incorporated into the amplicons during PCR (Fig. 1A). By determining the slope of the fluorescence curve and inverting the curve (multiplying by −1), we generated a melting curve for each sample (Fig. 1B, −d[fluorescence]/dT). The left and right margins of the melting curve were marked, and the distance between these two margins was defined as the HRM score (Fig. 1B). Figure 1C shows the melting curves for six control plasmids. The median HRM score for the plasmids was 3.4 (range 3.2–3.8). Because the plasmid templates are clonal, any diversity in the gag amplicons from these samples is likely to reflect errors introduced during PCR amplification.

We next used the HRM assay to analyze HIV from nine HIV-infected mother–infant pairs (Table 1 and Fig. 2). The HRM scores for each woman and her infant were independent of one another (for the paired samples, Pearson correlation r = 0.204, p = 0.599). The median HRM score for the nine infants (4.3, range 4.2–5.3) was higher than that for control plasmids (p < 0.0001) and was lower than that for the nine mothers (5.7, range 4.4–7.7, p = 0.005, exact Wilcoxon rank sum test). Viral load data were available for 17 of the 18 samples; we found no association between HRM score and number of copies of HIV RNA analyzed (p = 0.081 for all 18 samples; p = 0.21 for 9 maternal samples only, exact Wilcoxon rank sum test).

Reproducibility of the HRM assay was assessed by analyzing the 18 samples four times over the course of a year. For this analysis, DNA templates for the HRM assay (PCR products produced in the ViroSeq system) were stored at −80°C. For each run, samples were thawed and the HRM amplification and data analysis were repeated. Reproducibility of the HRM assay was high [intraclass correlation coefficient: 94% (95% CI: 89%, 98%), see Materials and Methods].

We next compared results from the HRM assay to results obtained using sequence-based measures of HIV diversity. For this analysis, HIV gag amplicons produced in the HRM assay were cloned and sequenced (50 clones for each maternal sample, 20 clones for each infant sample). Sequences were analyzed for genetic diversity, complexity, Shannon entropy, and length variation (see Materials and Methods, Table 1). Higher HRM scores were associated with higher genetic diversity (r = 0.76, p < 0.001), complexity (r = 0.59, p = 0.009), and Shannon entropy (r = 0.54, 0.022), but not with length variation (r = 0.39, p = 0.111).

Discussion

The HRM assay described in this study provides a rapid, high-throughput method for quantifying genetic diversity without sequencing. This method differs from the AmpliCot method, which has been used to analyze human genomic diversity.12 In AmpliCot, genetic diversity is determined by measuring the hybridization kinetics of DNA duplex formation. In contrast, the HRM assay uses the range of melting temperatures of DNA duplexes to measure HIV diversity. The HRM assay can detect low levels of sequence diversity (e.g.,
Table 1. HRM Scores and Sequence-Based Measures of HIV Diversity*

<table>
<thead>
<tr>
<th>Subject</th>
<th>HRM score</th>
<th>Log_{10} HIV RNA</th>
<th>Genetic diversity (%)</th>
<th>Median length (bp)</th>
<th>Length range (bp)</th>
<th>Length diff (bp)</th>
<th>Complexity</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-1</td>
<td>6.8</td>
<td>4.48</td>
<td>3.80</td>
<td>161</td>
<td>159–164</td>
<td>5</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>M-2</td>
<td>5.7</td>
<td>5.82</td>
<td>3.57</td>
<td>177</td>
<td>157–184</td>
<td>27</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>M-3</td>
<td>4.4</td>
<td>5.37</td>
<td>2.87</td>
<td>163</td>
<td>152–167</td>
<td>15</td>
<td>0.80</td>
<td>0.92</td>
</tr>
<tr>
<td>M-4</td>
<td>6.3</td>
<td>4.20</td>
<td>3.51</td>
<td>158</td>
<td>155–162</td>
<td>7</td>
<td>0.78</td>
<td>0.91</td>
</tr>
<tr>
<td>M-5</td>
<td>7.7</td>
<td>4.70</td>
<td>4.50</td>
<td>158</td>
<td>150–164</td>
<td>14</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>M-6</td>
<td>5.4</td>
<td>4.27</td>
<td>1.69</td>
<td>158</td>
<td>155–181</td>
<td>26</td>
<td>0.92</td>
<td>0.97</td>
</tr>
<tr>
<td>M-7</td>
<td>6.1</td>
<td>5.41</td>
<td>5.85</td>
<td>158</td>
<td>153–161</td>
<td>8</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>M-8</td>
<td>4.8</td>
<td>5.76</td>
<td>2.10</td>
<td>162</td>
<td>158–165</td>
<td>7</td>
<td>0.90</td>
<td>0.96</td>
</tr>
<tr>
<td>M-9</td>
<td>4.8</td>
<td>4.45</td>
<td>2.27</td>
<td>162</td>
<td>159–164</td>
<td>5</td>
<td>0.86</td>
<td>0.94</td>
</tr>
<tr>
<td>I-1</td>
<td>4.8</td>
<td>6.12</td>
<td>0.10</td>
<td>161</td>
<td>161</td>
<td>0</td>
<td>0.35</td>
<td>0.58</td>
</tr>
<tr>
<td>I-2</td>
<td>4.8</td>
<td>5.18</td>
<td>1.60</td>
<td>177</td>
<td>174–180</td>
<td>6</td>
<td>0.80</td>
<td>0.89</td>
</tr>
<tr>
<td>I-3</td>
<td>4.2</td>
<td>6.37</td>
<td>0.30</td>
<td>163</td>
<td>161–163</td>
<td>2</td>
<td>0.45</td>
<td>0.66</td>
</tr>
<tr>
<td>I-4</td>
<td>4.3</td>
<td>5.78</td>
<td>0.31</td>
<td>158</td>
<td>158</td>
<td>0</td>
<td>0.25</td>
<td>0.26</td>
</tr>
<tr>
<td>I-5</td>
<td>4.3</td>
<td>4.69</td>
<td>0.52</td>
<td>158</td>
<td>157–158</td>
<td>1</td>
<td>0.50</td>
<td>0.63</td>
</tr>
<tr>
<td>I-6</td>
<td>5.1</td>
<td>4.65</td>
<td>0.60</td>
<td>157.5</td>
<td>157–158</td>
<td>1</td>
<td>0.45</td>
<td>0.59</td>
</tr>
<tr>
<td>I-7</td>
<td>5.3</td>
<td>5.80</td>
<td>0.20</td>
<td>157</td>
<td>157</td>
<td>0</td>
<td>0.25</td>
<td>0.34</td>
</tr>
<tr>
<td>I-8</td>
<td>4.3</td>
<td>NA</td>
<td>0.10</td>
<td>162</td>
<td>162</td>
<td>0</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>I-9</td>
<td>4.2</td>
<td>5.50</td>
<td>0.51</td>
<td>161</td>
<td>160–161</td>
<td>1</td>
<td>0.55</td>
<td>0.65</td>
</tr>
</tbody>
</table>

*aPlasma HIV from nine Ugandan mothers (M-1 to M-9) and their infants (I-1 to I-9) was analyzed in the HRM assay (see Materials and Methods). HRM scores are defined as the number of degrees centigrade over which melting of the DNA amplicons occurred (Fig. 2B). HIV gag amplicons produced in the HRM assay were cloned. Sequencing was performed for 50 clones from each maternal sample and 20 clones from each infant sample. For each sample, the sequences were used to determine HIV genetic diversity (%), the median length of the gag amplicons (base pairs, bp), the difference in length from the longest to the shortest gag region sequence (length diff, bp), complexity, and Shannon entropy (see Materials and Methods).

bNA, not available; insufficient sample for analysis.

FIG. 2. HRM data from nine mother–infant pairs. Melting curves for the nine mother–infant pairs are shown (MI-1 to MI-9). In each panel, the melting curve from the mother’s sample is shown in red and the melting curve from the corresponding infant is shown in blue.
the level of HIV diversity in newly infected infants) and provides a simple measure of diversity, the HRM score, which is significantly associated with sequence-based diversity measures. The stability of DNA duplexes in a complex population is likely to be influenced by the number and type of nucleotide differences and insertions/deletions, as well as the proximity of these sequence differences to the ends of the duplex and to each other. For these reasons, the HRM score is a more complex measure of diversity than sequence-based approaches that are based on relatively simple algorithms, such as the frequency of nucleotide differences in a sequence set (genetic diversity) or the frequency of unique sequences in a sequence set (complexity).

The HRM primers used in this study were designed for amplification of several HIV subtypes, including A, B, C, and D; we have used the HRM assay to analyze HIV from >400 individuals infected with these HIV subtypes (data not shown). The region we selected for analysis within HIV gag is poorly conserved across HIV strains and often contains both point mutations and insertions/deletions. This part of HIV gag is likely to evolve more slowly during HIV infection than the other regions, such as HIV env, but is still of interest since it contains epitopes that are targets for cytotoxic T lymphocytes. Recent studies using the HRM assay show that HRM scores based on analysis of the HIV gag p6 region vary with the stage of HIV disease in adults (acute < recent < chronic and AIDS). Because the portion of the gag region analyzed in this study is positioned immediately upstream of HIV protease, it is often included in amplicons produced in HIV genotyping assays used for antiretroviral drug resistance testing. Therefore, DNA remaining after genotyping can be used as a template for gag amplification in the HRM assay, which may be an advantage if primary samples are no longer available.

In this study, the median HRM score (reflecting HIV diversity in the gag region) was higher for women than infants. The low HRM scores and low sequenced-based diversity measures obtained for infants are consistent with the fact that the infants were recently HIV infected. There is a wider range of variability in the HRM scores and sequence-based diversity measures for the women. Some of the women may have been recently infected or may have less diverse viral populations for other reasons. We considered whether the viral populations in the infants may have been reduced by NVP exposure (genetic bottlenecking); however, the HRM scores of six infants in HIVNET 012 who were not NVP exposed (mother and infant received placebo) were similar to those of the NVP-exposed infants (data not shown). We also considered whether variations in HIV viral load may have influenced the analysis (e.g., sampling error in low viral load samples). This does not appear to be the case; there was no association between HRM score and viral load. Based on the viral load of the samples, the amount of plasma used for testing (500 µl for women, 100 µl for infants), and the amount of extracted HIV RNA used for amplification (1/5 of the RNA from each sample), we calculated that all of the samples tested had at least 889 copies of HIV RNA analyzed.

The HRM assay described in this report may facilitate studies of the complexity of HIV populations and the evolution of HIV diversity during infection. We are currently using the HRM assay to evaluate HIV diversity and evolution in infants, children, and adults, and to investigate the relationship between HIV diversity and other factors, such as HIV subtype and route of HIV infection. This study evaluated HIV diversity in the gag region of HIV in plasma samples, but the HRM assay could be applied to any region of the HIV genome, to HIV in other sample types, to other pathogens that can occur as mixtures of genetically related variants, and to other genetic systems that exhibit DNA diversity.

Acknowledgments

We are indebted to Prof. Francis Mmiro for improving the health of women and infants with HIV infection. Prof. Mmiro was the Ugandan Principal Investigator for the HIVNET 012 trial. Sadly, he died before this study was performed. We also thank the HIVNET 012 study team, the women and infants who participated in the HIVNET 012 study, and the staff in Uganda and at Johns Hopkins University who assisted with sample processing. We also thank Jason McKinney, Luke Stewart, and Justin Geiger (Idaho Technologies) for helpful discussions.

We acknowledge the following sources of support: HIV Network for Prevention Trials (HIVNET), sponsored by the U.S. National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), through contract N01-AI-035173 with Family Health International, contract N01-AI-045200 with the Fred Hutchinson Cancer Research Center, and contract N01-AI-035173-417 with Johns Hopkins University.

HIV Prevention Trials Network (HPTN), sponsored by NIAID, Eunice Kennedy Shriver National Institutes of Child Health and Human Development (NICHD), National Institute on Drug Abuse (NIDA), National Institute of Mental Health (NIMH), and the Office of AIDS Research, of the NIH, DHHS (U01-AI-046745, U01-AI-048054, and U01-AI-068613).

International Maternal Pediatric and Adolescent AIDS Clinical Trials Group (U01-AI-068632, NIAID, NICHD).

R01 DA024565 (NIH, NIDA).

Author Disclosure Statement

None of the authors has a commercial or other association that might pose a conflict of interest with the following except: Dr. Susan Eshleman and Mr. William Towler are co-inventors of the High-Resolution Melting (HRM) assay and Johns Hopkins University is planning to file a patent application with the U.S.-Patent and Trademark Office. The inventors may receive royalty payments if the patent is awarded and licensed.

References


Address correspondence to:
Susan H. Eshleman
Johns Hopkins University School of Medicine
Ross Bldg., Room 646
720 Rutland Ave.
Baltimore, Maryland 21205
E-mail: seshlem@jhmi.edu