

# Use of a High Resolution Melting (HRM) Assay to Compare *Gag*, *Pol*, and *Env* Diversity in Adults with Different Stages of HIV Infection

Matthew M. Cousins<sup>1</sup>, Oliver Laeyendecker<sup>2,3</sup>, Geetha Beauchamp<sup>4</sup>, Ronald Brookmeyer<sup>5</sup>, William I. Towler<sup>1</sup>, Sarah E. Hudelson<sup>1</sup>, Leila Khaki<sup>1</sup>, Beryl Koblin<sup>6</sup>, Margaret Chesney<sup>7</sup>, Richard D. Moore<sup>2</sup>, Gabor D. Kelen<sup>8</sup>, Thomas Coates<sup>9</sup>, Connie Celum<sup>10</sup>, Susan P. Buchbinder<sup>7,11</sup>, George R. Seage III<sup>12</sup>, Thomas C. Quinn<sup>2,3</sup>, Deborah Donnell<sup>4</sup>, Susan H. Eshleman<sup>1\*</sup>

**1** Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **2** Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **3** Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Baltimore, Maryland, United States of America, **4** Statistical Center for HIV/AIDS Research and Prevention, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **5** Department of Biostatistics, University of California Los Angeles, Los Angeles, California, United States of America, **6** Laboratory of Infectious Disease Prevention, New York Blood Center, New York, New York, United States of America, **7** University of California San Francisco, San Francisco, California, United States of America, **8** Department of Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **9** Program in Global Health, University of California Los Angeles, Los Angeles, California, United States of America, **10** Department of Medicine, University of Washington, Seattle, Washington, United States of America, **11** San Francisco Department of Public Health, San Francisco, California, United States of America, **12** Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America

## Abstract

**Background:** Cross-sectional assessment of HIV incidence relies on laboratory methods to discriminate between recent and non-recent HIV infection. Because HIV diversifies over time in infected individuals, HIV diversity may serve as a biomarker for assessing HIV incidence. We used a high resolution melting (HRM) diversity assay to compare HIV diversity in adults with different stages of HIV infection. This assay provides a single numeric HRM score that reflects the level of genetic diversity of HIV in a sample from an infected individual.

**Methods:** HIV diversity was measured in 203 adults: 20 with acute HIV infection (RNA positive, antibody negative), 116 with recent HIV infection (tested a median of 189 days after a previous negative HIV test, range 14–540 days), and 67 with non-recent HIV infection (HIV infected >2 years). HRM scores were generated for two regions in *gag*, one region in *pol*, and three regions in *env*.

**Results:** Median HRM scores were higher in non-recent infection than in recent infection for all six regions tested. In multivariate models, higher HRM scores in three of the six regions were independently associated with non-recent HIV infection.

**Conclusions:** The HRM diversity assay provides a simple, scalable method for measuring HIV diversity. HRM scores, which reflect the genetic diversity in a viral population, may be useful biomarkers for evaluation of HIV incidence, particularly if multiple regions of the HIV genome are examined.

**Citation:** Cousins MM, Laeyendecker O, Beauchamp G, Brookmeyer R, Towler WI, et al. (2011) Use of a High Resolution Melting (HRM) Assay to Compare *Gag*, *Pol*, and *Env* Diversity in Adults with Different Stages of HIV Infection. PLoS ONE 6(11): e27211. doi:10.1371/journal.pone.0027211

**Editor:** Srikanth Prasad Tripathy, National AIDS Research Institute, India

**Received:** July 20, 2011; **Accepted:** October 11, 2011; **Published:** November 2, 2011

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** This work was supported by: (1) the HIV Network for Prevention Trials (HIVNET) and sponsored by the U.S. National Institute of Allergy and Infectious Diseases (NIAID), the National Institutes of Health (NIH), and Dept. of Health and Human Services (DHHS) (NO1-AI35176; N01-AI35173; N01-AI45200; U01-AI046745), (2) the HIV Prevention Trials Network (HPTN) sponsored by the NIAID, the National Institute on Drug Abuse (NIDA), the National Institute of Mental Health, and the Office of AIDS Research of the NIH and DHHS (U01-AI068613 and UM1-AI068613) (3) NIDA (R01-DA011602), (4) the National Institute of Alcohol Abuse and Alcoholism (R01-AA016893), (5) the Emergency Medicine Foundation (Center of Excellence grant), (6) NIAID (R01-AI095068), and (7) the Division of Intramural Research, NIAID, NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** S.H.E. has given presentations at meetings sponsored by Abbott Diagnostics (distributor of the ViroSeq HIV Genotyping System) and has collaborated with Celera (manufacturer of the ViroSeq HIV Genotyping System) and Abbott Diagnostics on evaluation of HIV-related assays. S.H.E. and W.T. are co-inventors of the HRM diversity assay, and Johns Hopkins University has filed a patent application for this assay with the US-Patent and Trademark Office under the title "Use of a high-resolution melting assay to measure genetic diversity". The inventors may receive royalty payments if the patent is awarded and licensed. This does not alter our adherence to all the PLoS ONE policies on sharing data and materials.

\* E-mail: seshlem@jhmi.edu

## Introduction

Accurate methods for measuring HIV incidence using cross-sectional samples are important for monitoring the HIV epidemic and assessing the efficacy of interventions for HIV prevention [1]. Most cross-sectional HIV incidence studies have been performed using serologic assays, such as the BED capture immunoassay [2]. Serologic HIV incidence assays are based on the premise that the antibody response to HIV matures over time. Unfortunately, some individuals never attain a mature anti-HIV antibody response, and individuals with non-recent HIV infection may be misclassified as recently infected if the antibody response to HIV is blunted by viral suppression or advanced HIV disease [3,4].

Because HIV generally diversifies over time in infected individuals, HIV diversity may serve as a biomarker for assessing HIV incidence [5]. Most studies of HIV diversity have used sequence-based methods to analyze individual HIV variants in infected individuals [6,7,8]. Those studies demonstrate that HIV infection is usually initiated by one or a small number of founder viruses [9]. Over time, rapid viral replication, frequent mutation, and frequent recombination events generate large numbers of distinct viral variants [9,10]. Immune responses to infection, antiretroviral therapy (ART), and other selective pressures drive the diversification and evolution of the viral population [9,11]. Previous reports suggest that *env* sequences are usually homogeneous early in infection [12,13,14], with higher levels of diversity accompanying higher multiplicity of infection [12,13]. After HIV infection is established, *env* diversity usually increases over time and stabilizes or declines in advanced stages of HIV disease [7,14,15]. Genetic diversification in *env* and *gag* may be concordant or discordant over the course of infection [16,17]. Differences in viral diversification in these two regions may reflect different selective pressures targeting *env* and *gag* proteins [16,18].

While sequencing-based studies of HIV diversity have been informative, the cost and effort needed to sequence viral variants (e.g., by cloning, single genome sequencing, or parallel sequencing) make those methods impractical for analysis of HIV incidence in larger cohorts or surveillance studies, which may require analysis of hundreds or thousands of samples. Recent reports have found that the frequency of ambiguous nucleotide calls in population sequencing data may reflect HIV diversity [19,20]. This approach may be useful for assessing HIV diversity using existing sequence databases generated for surveillance of HIV drug resistance. However, patterns of HIV diversity can vary from one genomic region to another, and genetic bottlenecks may occur in some regions during the course of HIV infection. Therefore, discrimination between recent and non-recent HIV infection may require analysis of diversity in more than one region of the HIV genome. For such an approach to be practical for HIV incidence applications, simpler methods are needed for HIV diversity analysis. Heteroduplex mobility assays can be used to analyze HIV diversity without sequencing [14]. In those assays, genetic diversity is quantified by analyzing the mobility pattern of amplified DNA in a gel. Unfortunately, the requirement for gel electrophoresis increases the time and effort needed for analysis and makes heteroduplex mobility assays difficult to scale up for high-throughput analysis.

We recently developed a rapid assay for HIV diversity based on high resolution melting (HRM) technology [21]. Assays based on HRM of DNA duplexes have been used to detect mutations associated with cancer and genetic diseases; HRM technology is also being developed for analysis of specific mutations in bacterial, viral, and parasitic pathogens [22]. We adapted HRM technology to quantify genetic diversity in HIV [21,23]. The HRM diversity

assay is performed in a 96-well plate format, and each melting procedure takes only a few minutes. The HRM diversity assay provides a single numeric HRM score that reflects the level of HIV diversity in a specific region of the HIV genome, simplifying data analysis. Calculation of the HRM score is straightforward and can be automated using the electronic output of the melting instrument. The HRM diversity assay is highly reproducible, and HRM scores are significantly associated with sequence-based diversity measures such as genetic diversity, genetic complexity, and Shannon entropy [21,23]. In this report, the HRM diversity assay was used to compare, *gag*, *pol*, and *env* diversity in samples from 203 adults with different stages of HIV infection. These data suggest that the HRM diversity assay may be useful for analysis of HIV incidence.

## Methods

### Human subjects (Ethics Statement)

The EXPLORE, HIVNET 001, Johns Hopkins HIV Clinical Cohort (JHHCC), and Johns Hopkins Hospital Emergency Department (JHH ED) studies were conducted according to the ethical standards set forth by the institutional review boards of the participating institutions and the Helsinki Declaration of the World Medical Association; participants provided written, informed consent [24,25,26,27]. The work described in this report involved analysis of stored samples and data from those studies. No participants were recruited or followed in the course of this work. The work described in this report was approved by the Internal Review Board at the Johns Hopkins University School of Medicine.

### Samples used for analysis

Samples were collected from adults in the United States. Acute samples (HIV RNA positive, HIV antibody negative [28], Feibig stage I or II [29],  $n = 20$ ) and recent samples (collected near the time of HIV seroconversion, likely Feibig stage VI [29],  $n = 102$ ) were obtained from men who have sex with men (MSM) in the EXPLORE study [24] (1999–2001, median age: 31 years, range 19–56 years, 66% Caucasian). Additional recent samples (collected near the time of HIV seroconversion, likely Feibig stage VI [29],  $n = 14$ ) were obtained from the HIVNET 001 study [25] (1995–1997) from ten heterosexual intravenous drug users (IDUs, six women and four men) and four heterosexual women who were not IDUs. The median time between collection of recent samples and the last negative HIV test was 189 days (range 14–540) for the EXPLORE study and 165 days (range: 49–216) for the HIVNET 001 study.

Samples from adults with non-recent HIV infection (infected >2 years) were obtained from two sources: (1) The Johns Hopkins HIV Clinical Cohort (JHHCC) study of HIV-infected patients in Baltimore (2002–2008,  $n = 56$ ) [26], and (2) Johns Hopkins Hospital Emergency Department (JHH ED) HIV serosurvey (2001 and 2003,  $n = 11$ ) [27]. Characteristics of the non-recent group ( $n = 67$ , JHHCC and JHH ED) were: 72% men, 36% Caucasian (53 had data for gender or race), median age: 36 years (range: 21–53 years; 60 had data for age), median  $\log_{10}$  HIV viral load: 5.2  $\log_{10}$  copies/ml (range: 3.6–5.9  $\log_{10}$  copies/ml, 55 had data for viral load), and median CD4 cell count: 43 cells/mm<sup>3</sup> (range: 1–388 cells/mm<sup>3</sup>, 57 had data for CD4 cell count).

### Preparation of DNA for HRM analysis

HIV RNA was extracted from plasma or serum using the ViroSeq HIV Genotyping System (Celera, Alameda, CA). HIV DNA used to analyze *gag* and *pol* was prepared using the ViroSeq

system. HIV DNA used to analyze *env* was prepared using the Qiagen OneStep RT-PCR Kit (QIAGEN Inc., Valencia, CA; forward primer: JH35F (5'-TGARGGACAATTGGAGAA-RTGA-3'); reverse primer JH38R (5'-GGTGARTATCCCCTK-CCTAAC-3') [30,31]). PCR products were purified using ExoSAP-IT (United States Biochemical Corporation, Cleveland, OH) and were diluted to approximately 0.5 ng/μl for HRM analysis.

### Preparation of plasmid controls

DNA (*gag-pol* and *env* amplicons) amplified from five recent samples (EXPLORE study) was cloned into the vector, pCR®2.1-TOPO (TOPO TA Cloning® Kit, Invitrogen, Carlsbad, CA). Plasmids were diluted to approximately 5 ng/μl for HRM analysis.

### HRM diversity analysis

The HRM diversity assay was performed as previously described [21]. Six regions of the HIV genome were amplified in the presence of LCGreen® Plus dye (Idaho Technology Inc., Salt Lake City, UT, Table 1, Figure 1). The amplicons were melted using the LightScanner Instrument (Model HR 96, Idaho Technology Inc., Salt Lake City, UT), and release of the dye was quantified as a function of temperature (melting range for *gag* and *pol* amplicons: 68–98°C with a 65°C hold; melting range for *env* amplicons: 60–98°C with a 57°C hold). Melting curves were used to determine HRM scores, as described previously [21], using a 15°C window size. Samples were analyzed in duplicate and the results were averaged; if the difference in the duplicate HRM scores was >0.5, the data were rejected, and the samples were reanalyzed.

### Statistical methods

The Wilcoxon rank sum test was used to compare HRM scores from different sample sets (e.g., non-recent vs. recent) for each region. Outlier values were defined as greater than the third quartile + (1.5 x interquartile range [IQR]) or less than the first quartile – (1.5 x IQR). Extreme values were defined as values greater than the third quartile + (3 x IQR) or less than the first

quartile – (3 x IQR). Logistic regression was used to assess HRM scores for recent and non-recent infection for all six regions both separately and jointly. The Wilcoxon rank sum test was used to compare HRM scores for adults in different subgroups (e.g., with and without ART). Analyses were performed using SAS software version 9.2 (Cary, North Carolina).

### Results

The HRM diversity assay was used to analyze six regions in the HIV genome: two regions in HIV *gag* (GAG1 and GAG2), one region in HIV *pol* (POL), and three regions in HIV *env* (ENV1, ENV2, and ENV3; Figure 1, Table 1). HRM scores for each region were all less than 5.2 for control plasmids (Figure 2). Region-specific differences in HRM scores for the plasmids most likely reflect differences in the length and melting domain characteristics of the amplicons [32].

The HRM diversity assay was used to analyze samples from 189 adults with different stages of HIV infection (20 with acute HIV infection [EXPLORE study], 102 with recent HIV infection [EXPLORE study], and 67 with non-recent HIV infection [JHHCC and JHU ED], see Methods). ART use and low CD4 cell count have been associated with misclassification of individuals with non-recent infection as recently infected using serologic HIV incidence assays [3,4]. To test whether this type of misclassification would also complicate use of the HRM diversity assay for HIV incidence testing, the non-recent sample set was selected to include a high proportion of “challenge” samples; 30 (44.8%) of the 67 non-recent samples were from adults on ART and 32 (48%) of the 67 non-recent samples were from adults with advanced HIV disease (CD4 cell counts <50 cells/mm<sup>3</sup>).

The HRM scores obtained for adults with acute HIV infection were all low (<6.3). The median HRM scores for those samples were significantly higher than the median scores for control plasmids for all of the regions except GAG2 (Figure 2). The highest median HRM score for acute samples was obtained for the ENV2 region which includes the immunodominant region (IDR) cluster I of HIV gp41 (Figure 1) [33]. The median HRM scores obtained for adults with recent HIV infection (seroconversion samples) were significantly higher than the median scores obtained

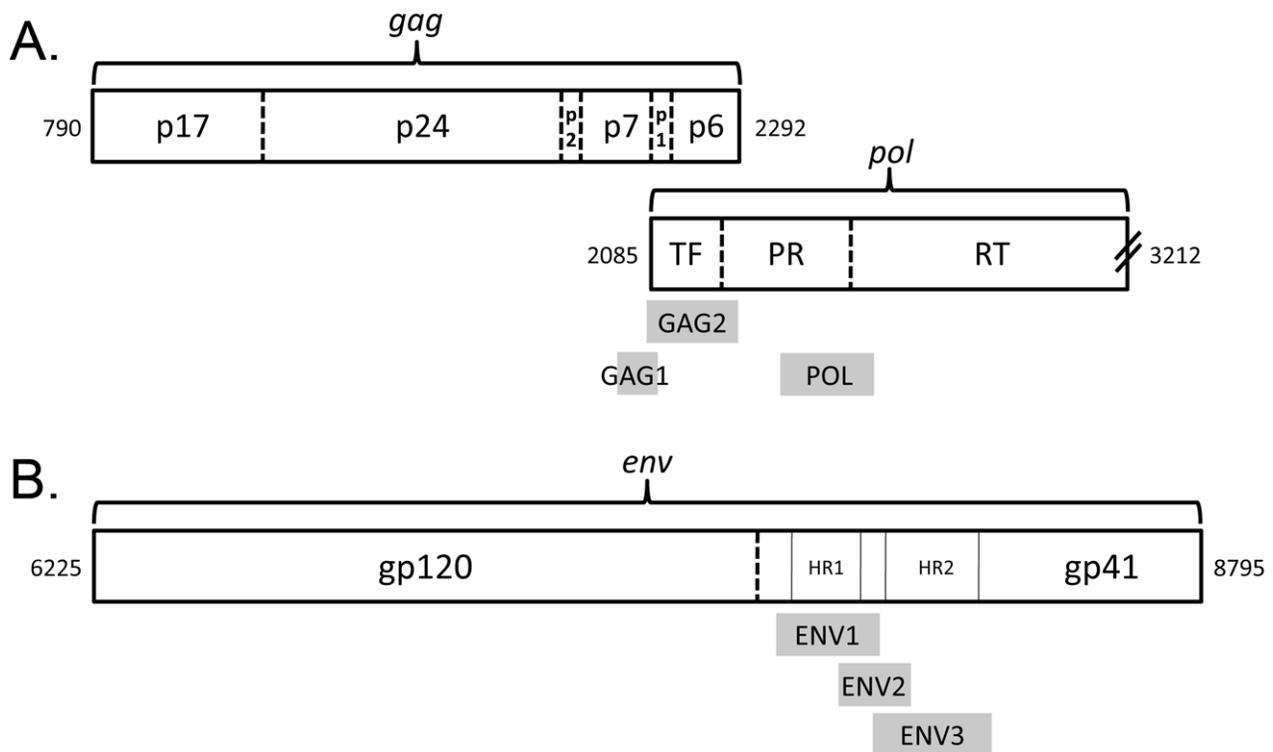
**Table 1.** Regions of the HIV genome analyzed using the high resolution melting (HRM) diversity assay.

Region analyzed <sup>a</sup>	Corresponding region in HXB2 <sup>a</sup>	Sequences of primers used to produce amplicons for HRM diversity analysis <sup>b</sup>	Amplicon size (bp)
GAG1	1998 – 2097	Forward: 5'- AAATTGCAGGGCCCTAGGAA Reverse: 5'- TTTCCCTAAAAAATTAGCCTGTCT	100
GAG2	2068 – 2278	Forward: 5'- ACTGAGAGACAGGCTAATTTTTAG Reverse: 5'- GGTCGTTGCCAAAGAGTGATTTG	211
POL	2373 – 2597	Forward: 5'- AAATGGAAACCAAAAATGATAG Reverse: 5'- CATTCTGGCTTTAATTTTACTG	225
ENV1	7798 – 8036	Forward: 5'- CAGCAGGWAGCACKATGGG Reverse: 5'- GCARATGWGYTTTCCAGAGCADCC	239
ENV2	7950 – 8119	Forward: 5'- CTYCAGRCAAGARTCYTGCC Reverse: 5'- TCCCAYTSCAKCCARGTC	170
ENV3	8016 – 8299	Forward: 5'- TGCTCTGGAARWCATYTGCC Reverse: 5'- AARCTCTACTATCATTATRA	284

<sup>a</sup>See Figure 1.

<sup>b</sup>Mixed nucleotides were present at some positions (W: A/T; K: G/T; R: A/G; Y: C/T; D: A/G/T; S: G/C).

doi:10.1371/journal.pone.0027211.t001



**Figure 1. Regions of the HIV genome analyzed using the HRM diversity assay.** The relevant regions of the HIV genome are shown (adapted from [39]). Numbers at the ends of each genomic segment correspond to coordinates in HXB2 (Genbank accession number: K03455). The amplicons used for HRM diversity analysis (regions analyzed) are indicated by shaded boxes. (A) *Gag* and *pol* amplicons: The GAG1 amplicon includes a portion of the coding regions for *gag* p7 and *gag* p1. The GAG2 amplicon includes the coding regions for *gag* p1 and *gag* p6 and extends into the coding region for HIV protease (PR); this amplicon also corresponds to the transframe (TF) protein. The POL amplicon spans the junction between the coding regions of HIV protease and HIV reverse transcriptase (RT). (B) *Env* amplicons: The ENV1 amplicon includes the coding region for heptad repeat 1 (HR1) of gp41, as well as portions of the coding regions to either side of HR1. The ENV3 amplicon includes the coding region for heptad repeat 2 (HR2), as well as portions of the coding regions to either side of HR2. The ENV2 amplicon includes the coding region for immunodominant region (IDR) cluster I of gp41, as well as portions of the coding regions for HR1 and HR2 [33]. doi:10.1371/journal.pone.0027211.g001

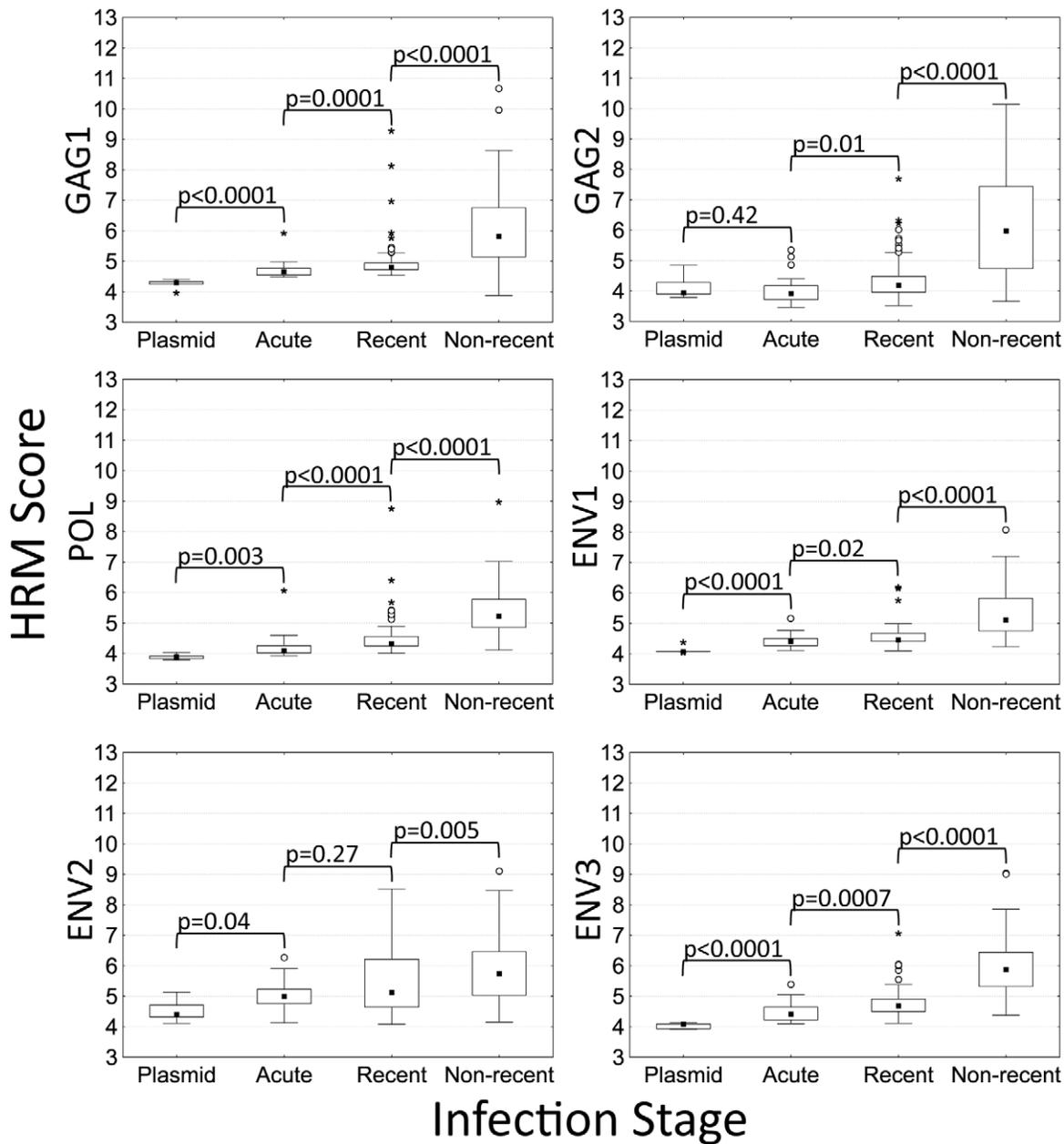
for adults with acute infection for all of the regions except ENV2 (Figure 2).

In all six regions analyzed, the median HRM scores for adults with non-recent infection (infected >2 years) were significantly higher than the median scores for adults with recent HIV infection ( $P=0.005$  for ENV2,  $P<0.0001$  for the other regions, Figure 2). Higher HRM scores in each region were associated with non-recent infection in logistic regression models (compared to recent HIV infection,  $P<0.02$  for ENV2,  $P<0.0001$  for all other regions; Table 2). In a multivariate logistic regression model, HRM scores in three regions were independently associated with non-recent infection (GAG2:  $P<0.04$ , ENV1:  $P<0.004$ , and ENV2:  $P<0.0004$ ; Table 2). Correlation plots for HRM scores in these three regions in adults with acute, recent, and non-recent infection are shown in Figure 3. In general, acute and recent infection samples had low HRM scores in all three regions, while non-recent infection samples had higher HRM scores (e.g., >6) in at least one of the three regions (Figures 3 and 4).

HRM scores for adults with non-recent infection who had CD4 cell counts above vs. below 50 cells/mm<sup>3</sup> were not significantly different for five of the six regions analyzed. For the POL region, the median HRM score was 5.6 for adults with CD4 cell counts <50 cells/mm<sup>3</sup> vs. 5.1 for adults with CD4 cell counts >50 cells/mm<sup>3</sup> ( $P=0.006$ , Wilcoxon test). In Figure 2, we show the distribution of the HRM scores obtained for each group of

participants. A total of 402 HRM scores were included in the analysis of participants with non-recent infection (67 samples, 6 regions each). Six (1.5%) of the 402 HRM scores were outlier or extreme values (see Methods); all six of those scores were from adults who had CD4 cell counts <50 cells/mm<sup>3</sup>. As shown in Figure 2, all of those outlier / extreme values are high values, indicating the presence of viral populations that had high levels of diversity in the regions analyzed. Note that all of the HRM scores (including the outlier / extreme HRM scores) were included in the analysis. The unusually high levels of viral diversity in some adults with very low CD4 cell counts enhances our ability to discriminate between adults with recent HIV infection and adults with non-recent HIV infection, including those with advanced HIV disease. There was no significant association between antiretroviral drug use and HRM score for any of the six regions (not shown).

In the analysis described above, adults with recent infection were all MSM; three (3%) of those men reported using intravenous drugs in the six months before HIV seroconversion. In contrast, the non-recent group included both men and women, most of whom likely acquired HIV infection through intravenous drug use. Some studies indicate that the number of HIV variants present very early in HIV infection varies among different risk groups [12,13,34,35]. Therefore, differences in HRM scores between these two groups (recent and non-recent) may have reflected their different demographic characteristics and risk factors for HIV



**Figure 2. HRM scores for plasmid controls and samples from adults with different stages of HIV disease.** The box and whisker plots show the distribution of HRM scores for six regions in the HIV genome in control plasmids ( $n=5$ ; Subtype B) and in adults with acute ( $n=20$ ), recent ( $n=102$ ), and non-recent ( $n=67$ ) HIV infection (see text). For each column, the median (closed square), interquartile range (IQR, box), lower inner fence (first quartile [Q1] - [1.5 X IQR]) and upper inner fence (third quartile [Q3] + [1.5 X IQR]), whiskers, outliers (greater than [Q3] + [1.5 X IQR], open circle) and extremes (greater than [Q3] + [3 x IQR], asterisk) are shown. doi:10.1371/journal.pone.0027211.g002

acquisition, rather than the duration of HIV infection. To address this possibility, we tested additional samples from women and heterosexual men, most of whom were IDUs (HIVNET 001 cohort, see Methods). In the GAG1, GAG2, POL, and ENV3 regions, HRM scores for men in the EXPLORE cohort and the adults in the HIVNET 001 cohort were not significantly different; HRM scores for two regions (ENV1 and ENV2) were slightly lower for adults in the HIVNET 001 cohort (Table 3). This indicates that the lower HRM scores observed in the recent group from EXPLORE compared to the non-recent group (from the JHHCC and JHU ED serosurvey) most likely reflect differences in

HIV diversity in recent vs. non-recent HIV infection rather than differences in the demographic characteristics and risk factors of the specific cohorts tested.

## Discussion

We used a novel HRM diversity assay to compare HIV diversity in adults with different stages of HIV disease. Adults with acute HIV infection had uniformly low HRM scores (low levels of HIV diversity). However, the median HRM scores for those individuals were significantly higher than those obtained for plasmid controls

**Table 2.** Comparison of HRM scores from adults with recent vs. non-recent HIV infection\*.

Region analyzed <sup>a</sup>	Univariate model		Multivariate model	
	Odds ratio (95% CI)	P value	Odds ratio <sup>b</sup> (95% CI)	P value
GAG1	<b>0.20 (0.10–0.34)</b>	<b>&lt;0.001</b>	0.60 (0.29–1.18)	0.15
GAG2	<b>0.27 (0.17–0.40)</b>	<b>&lt;0.001</b>	<b>0.58 (0.32–0.95)</b>	<b>0.04</b>
POL	<b>0.10 (0.05–0.20)</b>	<b>&lt;0.001</b>	0.60 (0.25–1.54)	0.33
ENV1	<b>0.03 (0.006–0.08)</b>	<b>&lt;0.001</b>	<b>0.17 (0.04–0.52)</b>	<b>0.004</b>
ENV2	<b>0.72 (0.54–0.95)</b>	<b>0.02</b>	0.92 (0.56–1.51)	0.73
ENV3	<b>0.07 (0.03–0.15)</b>	<b>&lt;0.001</b>	<b>0.20 (0.07–0.46)</b>	<b>&lt;0.001</b>

\*Abbreviations: HRM: high resolution melting. CI: confidence intervals. Recent samples were collected a median of 189 days after a negative HIV test (range: 14–540 days); non-recent samples were from individuals infected >2 years (see Methods). Results from adults with acute HIV infection were not included in this analysis. Statistically significant values are shown in bold text.

<sup>a</sup>See Figure 1.

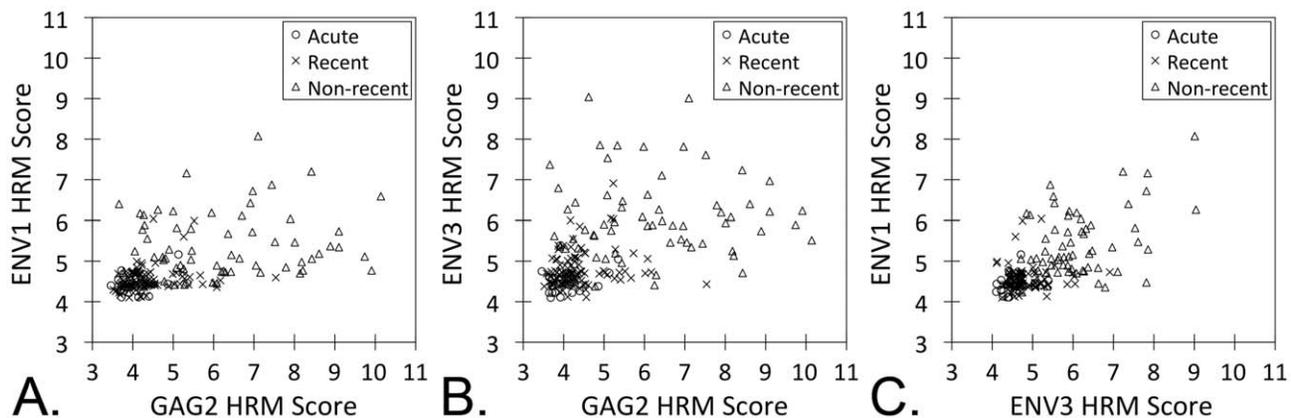
<sup>b</sup>Based on multivariate logistic regression adjusting for all regions simultaneously.

doi:10.1371/journal.pone.0027211.t002

for all regions except GAG2. This indicates that the HRM diversity assay can detect a low level of HIV diversity very early in HIV infection. In adults with acute HIV infection, the highest median HRM score was obtained for the ENV2 region, which contains IDR cluster I of gp41. HRM scores for all regions except for ENV2 were significantly higher in seropositive adults with recent HIV infection than in adults with acute HIV infection.

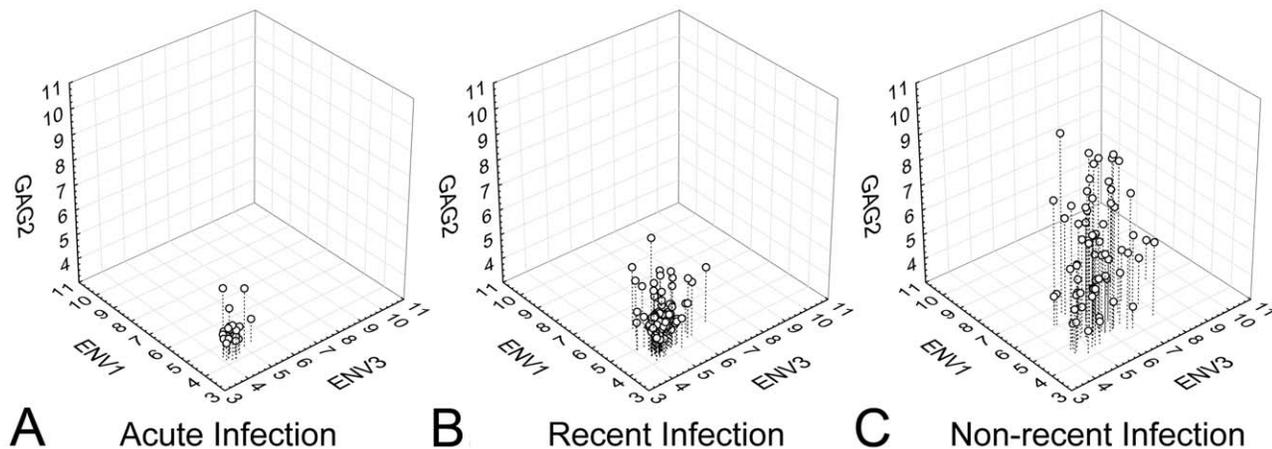
In all six regions analyzed, we found significantly higher levels of HIV diversity in adults who were infected for at least two years (non-recent group) than in adults near the time of HIV seroconversion (recent group). An important finding of this study was that viruses from individuals with non-recent infection often exhibited low diversity in at least one of the regions analyzed; in each region, there was some overlap in the HRM scores from adults with recent and non-recent infection. A similar finding was reported in a study that used a heteroduplex mobility assay to compare HIV diversity in the V3-V5 region of HIV *env* in adults with likely recent vs. likely non-recent infection (classified using a detuned enzyme immunoassay strategy) [36]. These findings and ours indicate that diversity-based measures (obtained using the HRM diversity assay or another method for viral diversity analysis) are not likely to be useful for HIV incidence analysis if they rely on analysis of a single genomic region. Our study extends the previous report [36] by comparing diversity in multiple genomic regions

and by using sample sets from individuals with known recent and known non-recent infection. This expanded analysis revealed that individuals with non-recent HIV infection rarely had low diversity in all regions that we examined. Furthermore, multivariate logistic regression showed that higher HRM scores in three regions (GAG2, ENV1, and ENV3) were independently associated with non-recent HIV infection. Independent diversification in different regions of the HIV genome is likely to reflect several factors, including: (1) different selective forces act on different HIV gene products, inducing diversification of different genomic regions, and (2) the very high frequency of genetic recombination of HIV lowers genetic linkage of different subgenomic regions. Our findings suggest that HIV diversity may be a useful biomarker for HIV incidence determination, provided that multiple HIV genomic regions are analyzed. The HRM diversity assay is simpler and less expensive than many other laboratory approaches used to measure HIV diversity, and it is particularly well-suited to multi-region analysis. The HRM diversity assay can be used to measure diversity in any RNA or DNA sample, including HIV RNA and proviral DNA. For HIV incidence applications, it makes most sense to analyze the actively replicating pool (e.g., plasma HIV RNA) rather than proviral DNA, which is likely to include archived sequences from viruses that were circulating earlier in infection.



**Figure 3. Relationship between HRM scores for the ENV1, ENV2, and GAG2 regions.** Scatter plots are shown for HRM scores for adults with acute, recent, and non-recent infection: (A) ENV1 vs. GAG2, (B) ENV3 vs. GAG2, and (C) ENV1 vs. ENV3.

doi:10.1371/journal.pone.0027211.g003



**Figure 4. HRM diversity scores for GAG2, ENV1 and ENV3 plotted in 3 dimensions.** High HRM scores for these regions were determined to be independently associated with non-recent infection. The use of data from multiple regions in tandem demonstrates that HRM scores are generally compact in acute infection (A) with a slight increase in the distribution of data in recent infection (B), and wide dispersion of the data in non-recent infection (C).

doi:10.1371/journal.pone.0027211.g004

Advanced HIV disease is associated with misclassification of individuals with non-recent infection as recently infected using serologic incidence assays [3]. Our results indicate that this is not likely to confound the use of the HRM diversity assay for HIV incidence testing. In this study, almost half (48%) of the samples in the non-recent group were from individuals with CD4 cell counts  $<50$  cells/mm<sup>3</sup>. In the non-recent group, HRM scores were not significantly different among adults with CD4 cell counts above vs. below 50 cells/mm<sup>3</sup>, and all of the unusually high HRM scores in this group (outlier and extreme values) were from adults with CD4 cell counts  $<50$  cells/mm<sup>3</sup>. These data show that advanced HIV disease is not associated with misclassification using the HRM diversity assay, and suggest that the HRM diversity assay may be

useful for identifying samples from adults with advanced HIV disease who were misclassified as recently infected using serologic incidence assays.

Viral suppression is also associated with misclassification of individuals with non-recent infection as recently infected using serologic incidence assays [3]. In this study, samples from adults on ART had detectable HIV RNA (to permit amplification of HIV RNA for analysis); it is not known whether those individuals were non-adherent to their treatment regimens or were failing ART. Because the HRM diversity assay uses different primer pairs for amplification of various regions of the HIV genome, and because the primers are designed to bind to relatively conserved sequences, sequence differences in HIV samples are not likely to impair binding of all of the relevant primer pairs. Therefore, amplification failure for all primer pairs is likely to indicate low viral load. We do not feel that it is necessary to screen samples for viral load prior to testing with the HRM diversity assay. If desired, a viral load assay could be used to confirm viral suppression in samples with multi-region amplification failure. We recognize that recently-infected individuals who have very low viral loads could be misclassified as non-recent if virologic suppression is used as a biomarker for non-recent infection. However, because infected individuals are not likely to have natural or ART-induced viral suppression early in infection, this type of misclassification should be infrequent and should have very little impact on HIV incidence estimates.

For samples that do amplify, we recognize that there is a potential to underestimate diversity when fewer copies of HIV RNA are used in the analysis. However, in a previous study [23], we demonstrated that results from the HRM diversity assay were not significantly affected by differences in sample volume (0.1 vs. 0.5 ml), HIV viral load (range: 2,000 to 50,000 copies/ml), or the number of HIV RNA copies used to prepare DNA templates for amplification (range: 100 to 5,000 copies of HIV RNA). Those results support the use of the HRM diversity assay for analysis of clinical samples with variable viral loads.

The HRM diversity assay provides data that is likely to be independent of data from serologic incidence tests. Therefore, use of the HRM diversity assay in combination with serologic testing is likely to improve the precision of multi-assay algorithms for HIV incidence, lowering misclassification rates. Figure 5A shows an example of an existing multi-assay algorithm that combines four

**Table 3. Comparison of HRM scores from recently infected individuals from the HIVNET 001 cohort and the EXPLORE cohort\*.**

Region analyzed <sup>a</sup>	HRM Score <sup>b</sup>		P value <sup>c</sup>
	HIVNET 001 (n = 14)	EXPLORE (n = 102)	
GAG1	4.8 (4.7, 4.9)	4.8 (4.7, 4.9)	0.95
GAG2	4.4 (4.1, 4.7)	4.2 (4.0, 4.5)	0.13
POL	4.3 (4.1, 4.4)	4.3 (4.3, 4.6)	0.24
ENV1	4.4 (4.3, 4.4)	4.5 (4.4, 4.7)	0.02
ENV2	4.6 (4.4, 4.8)	5.1 (4.6, 6.2)	0.01
ENV3	4.7 (4.4, 5.6)	4.7 (4.5, 4.9)	0.55

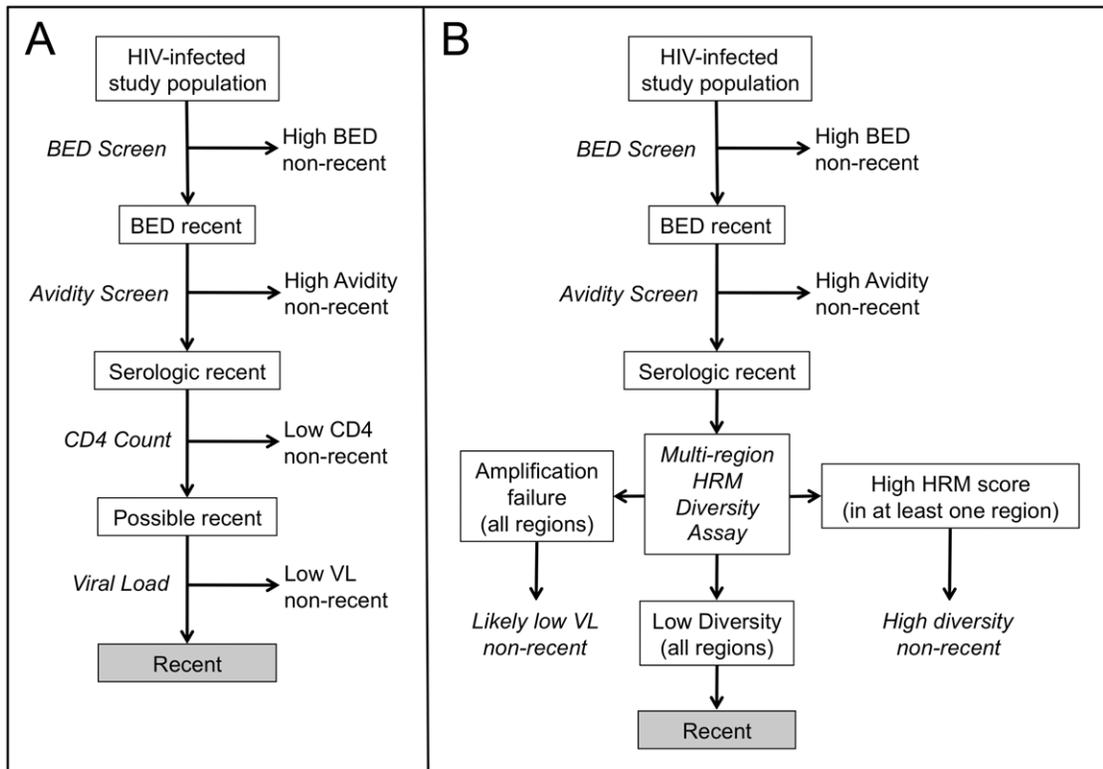
\*Abbreviations: HRM: high resolution melting. CI: confidence intervals.

<sup>a</sup>See Figure 1.

<sup>b</sup>HRM scores were obtained for samples collected from adults with recent HIV infection. Samples from the HIVNET 001 cohort were obtained from women and from men who likely acquired HIV through intravenous drug use. Samples from the EXPLORE cohort were obtained from men who likely acquired HIV infection through sexual exposure with other men. The median number of days between collection of the sample used for testing and the individual's last negative HIV tests was 165 days (range: 49–216) for the HIVNET 001 cohort and 189 days (range: 14–540) for the EXPLORE cohort. Median HRM scores and the first and third quartile HRM scores (in parentheses) are shown.

<sup>c</sup>Wilcoxon Rank sum test.

doi:10.1371/journal.pone.0027211.t003



**Figure 5. Use of the HRM diversity assay as part of a multi-assay algorithm for HIV incidence determination.** Panel A shows one example of a multi-assay algorithm developed for HIV incidence determination. In this algorithm, samples from HIV-infected individuals are first tested using the BED-CEIA assay, using a high assay cutoff to indicate non-recent HIV infection (BED screen). Samples that are below the BED screen cutoff (BED recent samples) are then tested using a second serologic assay, such as one based on antibody avidity (avidity screen). Samples that are below the cutoff for the second serologic assay are considered to be “serologic recent” samples. Samples with low CD4 cell count test results are then excluded as non-recent (note that CD4 cell count test results are usually obtained for all HIV-infected individuals at the time of sample collection). Finally, samples that are not excluded based on CD4 cell count are tested using a viral load assay, and samples with low viral loads are excluded as non-recent. The remaining samples are characterized as recent for the purpose of estimating HIV incidence. Panel B shows an alternative multi-assay algorithm that incorporates the HRM diversity assay. In this algorithm, samples that are characterized as serologic recent based on two assays (BED screen and an avidity screen) are tested with a multi-region HRM diversity assay. Samples that have a high HRM score in at least one of the regions tested are excluded as non-recent. Samples that fail to amplify in all regions tested are also excluded as non-recent, based on the assumption that they have low viral loads; this could be confirmed with a viral load assay. Samples that have low HRM scores in all regions tested are characterized as recent for the purpose of estimating HIV incidence.  
doi:10.1371/journal.pone.0027211.g005

assays for HIV incidence determination: a BED screening assay and an avidity screening assay (using a high cut-off for recent HIV infection for both assays), CD4 cell count, and viral load [23,37,38]. While CD4 cell count is a useful biomarker for reducing misclassification, inclusion of CD4 cell count data in incidence algorithms presents certain logistical challenges. First, because CD4 cell counts must be obtained in real-time (before other HIV incidence testing has been performed), CD4 cell count testing must be performed for all HIV-infected individuals evaluated, rather than the smaller subset who appear to be recently infected based on serologic testing. Second, many valuable sample sets from clinical trials and surveillance studies include only stored serum or plasma. Unless CD4 cell counts were obtained at the time of sample collection, it is not possible to assess incidence in those sample sets using an algorithm that includes CD4 cell count data.

We are exploring whether inclusion of the HRM diversity assay as part of a multi-assay algorithm will eliminate the need for CD4 cell count data, allowing all of the testing to be performed using a single plasma or serum sample. Figure 5B shows a possible alternative multi-assay algorithm that incorporates the HRM diversity assay. To reduce the cost and effort needed for analysis, it

would be most effective to screen samples for recent infection using serologic HIV incidence assays (e.g., BED and avidity screens). The subset of samples that are classified as “recent” based on serologic testing could then be tested using the HRM diversity assay to improve the precision of HIV incidence estimates. With this analytic plan, the number of samples that would require HRM diversity analysis would be relatively small. Therefore, the cost of the HRM diversity assay would not greatly impact the overall cost of the incidence assessment. Our preliminary data suggest that HIV from recently-infected individuals usually lacks diversity across the HIV genome, while HIV from individuals with non-recent infection is genetically diverse in one or more genomic regions. Studies are underway to identify genomic regions and assay cutoffs in each region that optimally discriminate between recent and non-recent infection. Once those parameters are set, it might be possible to use a simple approach in which samples are characterized as non-recent if they have a high HRM score in at least one of several regions tested.

In summary, this study provides proof of principle that HIV diversity can be used as a biomarker to distinguish between adults with recent vs. non-recent HIV infection. Further studies are needed to evaluate the performance of multi-assay algorithms for

HIV incidence determination that include the HRM diversity assay.

## Acknowledgments

The authors would like to thank the participants and the study teams of the EXPLORE and HIVNET 001 studies and patients who provided samples for the JHHCC Study and the JHH ED serosurvey. The authors thank Judy Shanan for management of the JHH ED serosurvey project.

**Disclaimer.** The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the National Institutes of Health (NIH). Use of trade names is for identification purposes only and does not constitute endorsement by the NIH.

**Previous Presentation of Data.** Part of this work was presented at the 17th Conf. on Retroviruses and Opportunistic Infections in San

Francisco, CA, February, 2010 (Towler WI, et al. Abstract #267), at the HIV Prevention Trials Annual Meeting in Washington, DC, June, 2011, and at the WHO Working Group on HIV Incidence Assays Annual Meeting in Atlanta, GA, August, 2011.

## Author Contributions

Conceived and designed the experiments: MMC OL WIT SHE. Performed the experiments: MMC WIT SEH LK. Analyzed the data: MMC SHE GB RB DD. Contributed reagents/materials/analysis tools: BK MC RDM GDK TC CC SPB GRS. Wrote the paper: MMC RB DD SHE. All authors contributed to manuscript preparation: MMC OL GB RB WIT SEH LK BK MC RDM GDK TC CC SPB GRS TCQ DD SHE.

## References

- Brookmeyer R (2010) Measuring the HIV/AIDS epidemic: approaches and challenges. *Epidemiol Rev* 32: 26–37.
- Dobbs T, Kennedy S, Pau CP, McDougal JS, Parekh BS (2004) Performance characteristics of the immunoglobulin G-capture BED-enzyme immunoassay, an assay to detect recent human immunodeficiency virus type 1 seroconversion. *J Clin Microbiol* 42: 2623–2628.
- Murphy G, Parry JV (2008) Assays for the detection of recent infections with human immunodeficiency virus type 1. *Euro Surveill* 13.
- Guy R, Gold J, Calleja JM, Kim AA, Parekh B, et al. (2009) Accuracy of serological assays for detection of recent infection with HIV and estimation of population incidence: a systematic review. *Lancet Infect Dis* 9: 747–759.
- Busch MP, Pilcher CD, Mastro TD, Kaldor J, Vercauteren G, et al. (2010) Beyond detuning: 10 years of progress and new challenges in the development and application of assays for HIV incidence estimation. *AIDS* 24: 2763–2771.
- Vrancken B, Lequime S, Theys K, Lemey P (2010) Covering all bases in HIV research: unveiling a hidden world of viral evolution. *AIDS Rev* 12: 89–102.
- Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, et al. (1999) Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 73: 10489–10502.
- Kearney M, Palmer S, Maldarelli F, Shao W, Polis MA, et al. (2008) Frequent polymorphism at drug resistance sites in HIV-1 protease and reverse transcriptase. *AIDS* 22: 497–501.
- Keele BF (2010) Identifying and characterizing recently transmitted viruses. *Curr Opin HIV AIDS* 5: 327–334.
- Tebit DM, Nankya I, Arts EJ, Gao Y (2007) HIV diversity, recombination and disease progression: how does fitness "fit" into the puzzle? *AIDS Rev* 9: 75–87.
- Richman DD, Little SJ, Smith DM, Wrin T, Petropoulos C, et al. (2004) HIV evolution and escape. *Trans Am Clin Climatol Assoc* 115: 289–303.
- Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, et al. (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* 105: 7552–7557.
- Li H, Bar KJ, Wang S, Decker JM, Chen Y, et al. (2010) High multiplicity infection by HIV-1 in men who have sex with men. *PLoS Pathog* 6: e1000890.
- Delwart EL, Pan H, Sheppard HW, Wolpert D, Neumann AU, et al. (1997) Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. *J Virol* 71: 7498–7508.
- Learn GH, Muthui D, Brodie SJ, Zhu T, Diem K, et al. (2002) Virus population homogenization following acute human immunodeficiency virus type 1 infection. *J Virol* 76: 11953–11959.
- Piantadosi A, Chohan B, Panteleeff D, Baeten JM, Mandaliya K, et al. (2009) HIV-1 evolution in gag and env is highly correlated but exhibits different relationships with viral load and the immune response. *AIDS* 23: 579–587.
- Liu S, Xing H, He X, Xin R, Zhang Y, et al. (2008) Dynamic analysis of genetic diversity of gag and env regions of HIV-1 CRF07\_BC recombinant in intravenous drug users in Xinjiang Uygur Autonomous Region, China. *Arch Virol* 153: 1233–1240.
- Frahm N, Brander C (2007) HIV viral diversity and escape from cellular immunity. *Curr Infect Dis Rep* 9: 161–166.
- Kouyos RD, von Wyl V, Yerly S, Boni J, Rieder P, et al. (2011) Ambiguous nucleotide calls from population-based sequencing of HIV-1 are a marker for viral diversity and the age of infection. *Clin Infect Dis* 52: 532–539.
- Wilson E, Shao W, Brooks J, Dewar R, Kearney M, et al. (2011) New bioinformatic algorithm to identify recent HIV-1 infection. 18th Conf on Retroviruses and Opportunistic Infections, Boston, MA, Abstract # 1057.
- Towler WI, James MM, Ray SC, Wang L, Donnell D, et al. (2010) Analysis of HIV diversity using a high-resolution melting assay. *AIDS Res Hum Retroviruses* 26: 913–918.
- Montgomery JL, Sanford LN, Wittwer CT (2010) High-resolution DNA melting analysis in clinical research and diagnostics. *Expert Rev Mol Diagn* 10: 219–240.
- James MM, Wang L, Musoke P, Donnell D, Fogel J, et al. (2011) Association of HIV diversity and survival in HIV-infected Ugandan infants. *PLoS ONE* 6: e18642.
- Koblin B, Chesney M, Coates T (2004) Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 364: 41–50.
- Celum C L, Buchbinder S P, Donnell D, Douglas J M, Mayer K, et al. (2001) Early human immunodeficiency virus (HIV) infection in the HIV network for prevention trials vaccine preparedness cohort: risk behaviors, symptoms, and early plasma and genital tract virus load. *J Infect Dis* 183: 23–35.
- Moore RD (1998) Understanding the clinical and economic outcomes of HIV therapy: the Johns Hopkins HIV clinical practice cohort. *J Acquir Immune Defic Syndr Hum Retrovirol* 17 Suppl 1: S38–41.
- Laeyendecker O, Rothman RE, Henson C, Horne BJ, Ketlogetswe KS, et al. (2008) The effect of viral suppression on cross-sectional incidence testing in the Johns Hopkins hospital emergency department. *J Acquir Immune Defic Syndr* 48: 211–215.
- Eshleman SH, Khaki L, Laeyendecker O, Piwowar-Manning E, Johnson-Lewis L, et al. (2009) Detection of individuals with acute HIV-1 infection using the ARCHITECT HIV Ag/Ab Combo assay. *J Acquir Immune Defic Syndr* 52: 121–124.
- Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, et al. (2003) Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS* 17: 1871–1879.
- Swanson P, Devare SG, Hackett J, Jr. (2003) Molecular characterization of 39 HIV isolates representing group M (subtypes A-G) and group O: sequence analysis of gag p24, pol integrase, and env gp41. *AIDS Res Hum Retroviruses* 19: 625–629.
- Badreddine S, Smith K, van Zyl H, Bodelle P, Yamaguchi J, et al. (2007) Identification and characterization of HIV type 1 subtypes present in the Kingdom of Saudi Arabia: high level of genetic diversity found. *AIDS Res Hum Retroviruses* 23: 667–674.
- Reed GH, Wittwer CT (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. *Clin Chem* 50: 1748–1754.
- Dorn J, Masciotra S, Yang C, Downing R, Biryahwaho B, et al. (2000) Analysis of genetic variability within the immunodominant epitopes of envelope gp41 from human immunodeficiency virus type 1 (HIV-1) group M and its impact on HIV-1 antibody detection. *J Clin Microbiol* 38: 773–780.
- Bar KJ, Li H, Chamberland A, Tremblay C, Routy JP, et al. (2010) Wide variation in the multiplicity of HIV-1 infection among injection drug users. *J Virol* 84: 6241–6247.
- Long EM, Martin HL, Jr., Kreiss JK, Rainwater SM, Lavreys L, et al. (2000) Gender differences in HIV-1 diversity at time of infection. *Nat Med* 6: 71–75.
- Machado D, Delwart E, Diaz R, de Oliveira C, Alves K, et al. (2002) Use of the sensitive/less-sensitive (detuned) EIA strategy for targeting genetic analysis of HIV-1 to recently infected blood donors. *AIDS* 16: 113–119.
- Laeyendecker O, Oliver A, Astemborski J, Owen M, Kirk G, et al. (2010) Improved precision of cross-sectional HIV incidence testing using a multi-assay algorithm that includes BED and an avidity assay with modified assay cut-offs. 17th Conf on Retroviruses and Opportunistic Infections, San Francisco, CA, Abstract # 935.
- Mastro TD, Kim AA, Hallett T, Rehle T, Welte A, et al. (2010) Estimating HIV incidence in populations using tests for recent infection: issues, challenges and the way forward. *J HIV AIDS Surveill Epidemiol* 2: 1–14.
- Kuiken C, Foley B, Leitner T, Apetrei C, Hahn B, et al. (2010) HIV Sequence Compendium 2010. Los AlamosNM: Los Alamos National Laboratory, Theoretical Biology and Biophysics.