Increased Sequence Coverage through Combined Targeting of Variant and Conserved Epitopes Correlates with Control of HIV Replication

Justine Sunshine, Moon Kim, Jonathan M. Carlson, David Heekerman, Julie Czartoski, Stephen A. Migueles, Janine Maenza, M. Juliana McElrath, James I. Mullins and Nicole Frahm


Updated information and services can be found at:
http://jvi.asm.org/content/88/2/1354

These include:

**SUPPLEMENTAL MATERIAL**
Supplemental material

**REFERENCES**
This article cites 68 articles, 30 of which can be accessed free at: http://jvi.asm.org/content/88/2/1354#ref-list-1

**CONTENT ALERTS**
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
A major challenge in the development of an HIV vaccine is that of contending with the extensive sequence variability found in circulating viruses. Induction of HIV-specific T-cell responses targeting conserved regions and induction of HIV-specific T-cell responses recognizing a high number of epitope variants have both been proposed as strategies to overcome this challenge. We addressed the ability of cytotoxic T lymphocytes from 30 untreated HIV-infected subjects with and without control of virus replication to recognize all clade B Gag sequence variants encoded by at least 5% of the sequences in the Los Alamos National Laboratory HIV database (1,300 peptides) using gamma interferon and interleukin-2 (IFN-γ/IL-2) FluoroSpot analysis. While targeting of conserved regions was similar in the two groups (P = 0.47), we found that subjects with control of virus replication demonstrated marginally lower recognition of Gag epitope variants than subjects with normal progression (P = 0.05). In viremic controllers and progressors, we found variant recognition to be associated with viral load (r = 0.62, P = 0.001). Interestingly, we show that increased overall sequence coverage, defined as the overall proportion of HIV database sequences targeted through the Gag-specific repertoire, is inversely associated with viral load (r = −0.38, P = 0.03). Furthermore, we found that sequence coverage, but not variant recognition, correlated with increased recognition of a panel of clade B HIV founder viruses (r = 0.50, P = 0.004). We propose sequence coverage by HIV Gag-specific immune responses as a possible correlate of protection that may contribute to control of virus replication. Additionally, sequence coverage serves as a valuable measure by which to evaluate the protective potential of future vaccination strategies.
the autologous virus over time (15, 31–33). While greatly informative for our understanding of the dynamics of host-virus interactions, these types of approaches inevitably extend only incrementally our understanding of the potential for HIV-specific CTLs to recognize naturally occurring epitope variants. As a result, the relative contribution that broad epitope variant recognition may play in the control of HIV replication remains unclear.

Here we report a comprehensive evaluation of the cross-reactivity of Gag-specific T-cell responses to frequently found clade B variants in the Los Alamos National Laboratory HIV Sequence Database (HIVDB; http://www.hiv.lanl.gov/) in individuals with and without spontaneous control of viral replication. Contrary to our expectations, we observed extensive epitope variant recognition in progressors, which was associated with viral load. However, we found an inverse correlation between viral load and sequence coverage of frequently occurring variants, suggesting that it is the ability to target the most commonly occurring variants, rather than simply a large number of variants, that contributes to control of viral replication. These findings provide a greater understanding of epitope variant recognition during natural infection and offer important insight for informing the evaluation of variant-inclusive vaccines.

MATERIALS AND METHODS

Study subjects. Thirty HIV-infected participants in the Seattle Long-Term Non-Progressor (34), Natural Progression, Primary Infection (35–37), and NIAID/NHI (38) cohorts were selected for this study. HIV controllers (n = 15) were defined by viral loads < 2,000 RNA copies/ml and CD4+ T-cell counts > 500. Elite controllers (EC) were further defined as having undetectable viral load (<50 RNA copies/ml), whereas viremic controllers (VC) had detectable viral loads of <2,000 RNA copies/ml. Progressors (n = 15) were defined as having median viral loads > 10,000 RNA copies/ml in the prior year. Three of the progressors were identified from the Seattle Primary Infection cohort and therefore had previously assigned publication identification numbers (P11 = 55097, P12 = 59530, and P15 = 75688). All subjects were studied in chronic infection and were antiretroviral therapy naive. The relevant institutional review boards and P15 RNA copies/ml in the prior year. Three of the progressors were identified (i) >55 spot-forming cells (SFC)/million PBMC, (ii) >4 times the average of the contents of at least six negative-control wells from the same experiment, (iii) values > 3 standard deviations above the average of the negative-control well results, and (iv) the presence of at least 5 spots per well.

Gag sequencing. A total of 10 to 25 gag sequences were obtained from plasma RNA from five VC and five progressors. For 8/10 study subjects, sequencing was performed at the same visit date used in the immunological assays. Because of plasma availability, clones from VC C12 and C15 were sequenced from plasma at the next available visit dates (2 and 3 months later, respectively).

HIV-1 RNA was purified from plasma using a QIAamp Viral RNA MiniKit (Qiagen, Valencia, CA). HIV-1-specific PCR was performed to amplify HIV-1 gag products with the use of endpoint-limiting dilution to obtain single templates. The HIV-1 gag gene (1,590-nucleotide [nt] fragment; the complete coding region of Gag and 103 nt of that of protease [PR]) was amplified by nested PCR with the following primer sets. The first round PCR, nt 683 to 705 of HIV-1_HXB2 (F683; 5′-CTC TCG CAC GAG GAC TCG GCT TG-3′) and nt 2484 to 2511 of HIV-1_HXB2 (StepGR_1.0_2511; 5′-TTC CAA TTA TGT TGA CAG GTG TAG GTC-3′) were used. For the second round, nt 762 to 786 of HIV-1_HXB2 (StepGR_2.0_2403; 5′-TTC CAA TTA TGT TGA CAG GTG TAG GTC C-3′) were used. The first round of PCR used Advantage 2 polymerase (Clontech, Mountain View, CA) and the following conditions: 95°C for 2 min, 35 cycles of 95°C for 30 s, 58°C for 30 s, and 68°C for 2 min, and a final extension step at 68°C for 10 min. The second round of PCR used Biolase Taq (Bioline, Taunton, MA) and the following conditions: 95°C for 2 min, 35 cycles of 95°C for 30 s, 58°C for 30 s, and 68°C for 2 min, and a final extension step at 68°C for 10 min. PCR products were visualized using a QIAxcel DNA Fast Analysis kit (Qiagen), and PCR-positive samples were purified using a NucleoSpin Gel and PCR Clean-Up kit (Clontech) before being submitted directly for dideoxynucleotide chain termination (Sanger) sequencing. Chromatograms were assembled and edited using Geneious 5.6.5 (Biomatters Ltd., New Zealand). HIV populations from these sequenced 10 individuals were confirmed to represent clade B viruses (data not shown).

Data analysis. The breadth of Gag-specific responses was calculated conservatively based on epitopic regions, which means that responses for up to 4 subsequent 11mer peptides (i.e., representing a potential shared 8mer) were counted as one epitopic region. Dually functional responses were determined by an IFN-γ plus IL-2 response to at least one 11mer peptide within a targeted epitopic region. The magnitude of the response to an epitopic region (in SFC/million PBMC) was defined as the highest response to a single peptide observed within the targeted epitopic region.

Variant recognition was calculated as the proportion of variants eliciting a response among all the 11mer variants tested in the targeted epitopic region, including overlapping peptide responses (see Table S2 in the supplemental material). For example, if a response is directed toward an epitopic region containing two overlapping peptides, each represented by four variants (i.e., 8 variants in total), and for both overlapping peptides two variants are recognized (i.e., a total of 4 variants recognized), variant recognition is calculated to be at 50%. Responses in which only a single variant was tested (i.e., highly conserved, with no other variant
present in ≥5% of sequences] were excluded from the variant recognition analysis as no variants were tested. Sequence coverage of frequently occurring Gag variants was calculated per epitopic region as the sum [frequency of targeted 11mers] over the sum [frequency of tested 11mers], where “frequency” refers to the proportion of the 11mer sequence in the clade B sequence alignment used to design the peptide set (see Table S2 in the supplemental material). We chose to normalize sequence coverage for each epitopic region by the sum [frequency of tested 11mers] to account for the fact that many epitopic regions had overlapping 11mers. Without normalization, the sum [frequency of targeted 11mers] would often be over 100%. In all targeted epitopic regions identified in this study, we found that the median value for the sum [frequency of tested 11mers] (i.e., the denominator in Table S2 in the supplemental material) that we used to normalize the coverage for each response was 81%. This highlights that although our peptide set does not cover all possible variations in clade B Gag sequences, the peptide design ensured that we were able to test for responses to a high proportion of circulating Gag variants.

Autologous gag sequences (whole-protein and viral portions corresponding to the targeted epitopic region) were aligned using the Muscle algorithm within Seaview v4.4.0 (39). The Le and Gascuel (LG) substitution model (40)-corrected pairwise diversity and divergence from clade B consensus were calculated using DIVEIN (41). Analyses were restricted to only those epitopic regions in which variant recognition was assessed. To generate the panel of founder viruses, we obtained previously generated sequences isolated from the first time point in early acute infection (median, 7 days post-onset of symptoms) from eight individuals (14). For each individual, a consensus sequence was generated from the first time point sequences using Seaview v4.4.0 and was used to represent the founder virus.

Statistical analysis. For each study subject, the median variant recognition or median epitope coverage was calculated across all epitopes and used in comparisons between groups using a nonparametric Mann-Whitney test. Differences between pairwise distances between groups were assessed using a nonparametric Mann-Whitney test (see Table S2 in the supplemental material) that we used to normalize the coverage for each response was 81%. This highlights that although our peptide set does not cover all possible variations in clade B Gag sequences, the peptide design ensured that we were able to test for responses to a high proportion of circulating Gag variants.

Autologous gag sequences (whole-protein and viral portions corresponding to the targeted epitopic region) were aligned using the Muscle algorithm within Seaview v4.4.0 (39). The Le and Gascuel (LG) substitution model (40)-corrected pairwise diversity and divergence from clade B consensus were calculated using DIVEIN (41). Analyses were restricted to only those epitopic regions in which variant recognition was assessed. To generate the panel of founder viruses, we obtained previously generated sequences isolated from the first time point in early acute infection (median, 7 days post-onset of symptoms) from eight individuals (14). For each individual, a consensus sequence was generated from the first time point sequences using Seaview v4.4.0 and was used to represent the founder virus.

Statistical analysis. For each study subject, the median variant recognition or median epitope coverage was calculated across all epitopes and used in comparisons between groups using a nonparametric Mann-Whitney test. Differences between pairwise distances between groups were assessed using a nonparametric Mann-Whitney test. Spearman rank testing was used to assess all correlations. GraphPad Prism X was used for all statistical analyses.

RESULTS

Significant differences in the quality but not the quantity of Gag-specific responses between controllers and progressors. To assess whether broad recognition of Gag epitope variants was associated with control of viral replication, we screened for Gag-specific T-cell responses in 15 HIV controllers and 15 HIV progressors. The controller cohort was comprised of seven elite control. Given that half of our controllers possessed B*57 and/or B*27 alleles, we did observe more responders among controllers to residues 162 to 172 (P = 0.46) and residues 263 to 272 (P = 0.65), corresponding to the immunodominant B*57 KF11 and B*27 KK10 epitopes, respectively. However, these differences were not significant.

Higher epitope variant recognition observed in progressors. As the results determined for breadth and the targeting of Gag residues were comparable, we next determined if there were differences in the number of distinct sequence variances recognized (“variant recognition”) as shown by Gag-specific T-cell responses between the controller and progressor groups. The peptide set used in this study was designed to include any variant present in at least 5% of clade B Gag sequences from the HIVDB, allowing us to comprehensively identify responses to frequently occurring Gag variants. For each targeted epitope, we calculated variant recogni-
Cross-reactivity capability (9, 10, 45). We did not detect a significant difference between viral recognition and viral load. When we examined all study subjects, we observed a trend toward a correlation between epitope variant recognition and viral load (r = 0.33, P = 0.07; Fig. 4A). Exclusion of individuals with undetectable viral load from the analysis revealed a significant correlation between epitope variant recognition and viral load (r = 0.62, P = 0.001; Fig. 4B). We found no significant association between response magnitude and viral load (r = 0.24, P = 0.18; data not shown).

Interestingly, when we segregated our controller cohort into elite controllers and viremic controllers, we found that progressors recognized significantly more variants only than viremic controllers (Kruskal-Wallis [KW] test P = 0.047, Dunn’s posttest P < 0.05 for VC versus progressors; Fig. 4C). To further explore the relationship between viral load and variant recognition, we re-

**TABLE 1 Subject characteristics**

<table>
<thead>
<tr>
<th>Study subject</th>
<th>Status</th>
<th>CD4 count</th>
<th>Viral load</th>
<th>HLA class I designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controllers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>EC</td>
<td>809</td>
<td>&lt;50</td>
<td>A<em>03/A</em>31/B<em>27/B</em>57/C<em>02/C</em>12</td>
</tr>
<tr>
<td>2</td>
<td>EC</td>
<td>1,211</td>
<td>&lt;50</td>
<td>A<em>02/A</em>03/B<em>13/C</em>18/C<em>06/C</em>07</td>
</tr>
<tr>
<td>3</td>
<td>EC</td>
<td>540</td>
<td>&lt;50</td>
<td>A<em>02/A</em>11/B<em>35/B</em>57/C<em>04/C</em>06</td>
</tr>
<tr>
<td>4</td>
<td>EC</td>
<td>603</td>
<td>&lt;50</td>
<td>A<em>29/ND/B</em>13/B<em>49/C</em>06/C*07</td>
</tr>
<tr>
<td>5</td>
<td>EC</td>
<td>643</td>
<td>&lt;50</td>
<td>A<em>24/A</em>31/B<em>27/B</em>57/C<em>02/C</em>02</td>
</tr>
<tr>
<td>6</td>
<td>EC</td>
<td>1,389</td>
<td>&lt;50</td>
<td>A<em>11/A</em>31/B<em>15/B</em>15/C<em>03/C</em>04</td>
</tr>
<tr>
<td>7</td>
<td>EC</td>
<td>1,916</td>
<td>&lt;50</td>
<td>A<em>01/A</em>24/B<em>08/C</em>35/C<em>04/C</em>07</td>
</tr>
<tr>
<td>8</td>
<td>VC</td>
<td>523</td>
<td>246</td>
<td>A<em>01/A</em>26/B<em>49/B</em>57/C<em>06/C</em>07</td>
</tr>
<tr>
<td>9</td>
<td>VC</td>
<td>784</td>
<td>125</td>
<td>A<em>24/A</em>25/B<em>39/B</em>57/C<em>06/C</em>12</td>
</tr>
<tr>
<td>10</td>
<td>VC</td>
<td>580</td>
<td>432</td>
<td>A<em>29/A</em>32/B<em>44/C</em>05/C*16</td>
</tr>
<tr>
<td>11</td>
<td>VC</td>
<td>747</td>
<td>250</td>
<td>A<em>01/A</em>03/B<em>27/B</em>57/C<em>02/C</em>06</td>
</tr>
<tr>
<td>12</td>
<td>VC</td>
<td>570</td>
<td>386</td>
<td>A<em>02/A</em>03/B<em>18/B</em>07/C<em>07/C</em>12</td>
</tr>
<tr>
<td>13</td>
<td>VC</td>
<td>702</td>
<td>163</td>
<td>A<em>01/A</em>02/B<em>38/B</em>57/C<em>06/C</em>12</td>
</tr>
<tr>
<td>14</td>
<td>VC</td>
<td>790</td>
<td>191</td>
<td>A<em>11/A</em>26/B<em>35/B</em>58/C<em>04/C</em>14</td>
</tr>
<tr>
<td>15</td>
<td>VC</td>
<td>1,181</td>
<td>430</td>
<td>A<em>03/A</em>68/B<em>13/B</em>56/C<em>06/C</em>07</td>
</tr>
<tr>
<td>Median</td>
<td>747</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progressors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>TP</td>
<td>349</td>
<td>22,730</td>
<td>A<em>01/A</em>01/B<em>08/B</em>08/C<em>07/C</em>07</td>
</tr>
<tr>
<td>2</td>
<td>TP</td>
<td>550</td>
<td>286,000</td>
<td>ND</td>
</tr>
<tr>
<td>3</td>
<td>DP</td>
<td>441</td>
<td>78,400</td>
<td>A<em>01/A</em>02/B<em>15/B</em>57/C<em>03/C</em>06</td>
</tr>
<tr>
<td>4</td>
<td>TP</td>
<td>522</td>
<td>39,550</td>
<td>A<em>03/A</em>24/B<em>35/B</em>44/C<em>02/C</em>04</td>
</tr>
<tr>
<td>5</td>
<td>TP</td>
<td>510</td>
<td>60,000</td>
<td>A<em>02/A</em>03/B<em>07/B</em>44/C<em>05/C</em>07</td>
</tr>
<tr>
<td>6</td>
<td>TP</td>
<td>175</td>
<td>115,240</td>
<td>A<em>24/A</em>25/B<em>15/B</em>58/C<em>04/C</em>07</td>
</tr>
<tr>
<td>7</td>
<td>TP</td>
<td>440</td>
<td>622,800</td>
<td>A<em>01/A</em>01/B<em>08/B</em>40/C<em>03/C</em>07</td>
</tr>
<tr>
<td>8</td>
<td>TP</td>
<td>371</td>
<td>217,200</td>
<td>A<em>11/A</em>31/B<em>08/B</em>07/C<em>07/C</em>07</td>
</tr>
<tr>
<td>9</td>
<td>TP</td>
<td>407</td>
<td>3,455</td>
<td>A<em>02/A</em>33/B<em>35/B</em>44/C<em>04/C</em>16</td>
</tr>
<tr>
<td>10</td>
<td>TP</td>
<td>340</td>
<td>23,750</td>
<td>ND</td>
</tr>
<tr>
<td>11</td>
<td>TP</td>
<td>317</td>
<td>8,606</td>
<td>A<em>23/A</em>30/B<em>08/B</em>42/C<em>07/C</em>17</td>
</tr>
<tr>
<td>12</td>
<td>TP</td>
<td>448</td>
<td>153,286</td>
<td>A<em>01/A</em>01/B<em>08/B</em>08/C<em>07/C</em>07</td>
</tr>
<tr>
<td>13</td>
<td>TP</td>
<td>575</td>
<td>24,224</td>
<td>A<em>03/A</em>30/B<em>35/B</em>57/C<em>04/C</em>07</td>
</tr>
<tr>
<td>14</td>
<td>TP</td>
<td>416</td>
<td>23,279</td>
<td>A<em>68/A</em>74/B<em>14/B</em>57/C<em>07/C</em>08</td>
</tr>
<tr>
<td>15</td>
<td>TP</td>
<td>500</td>
<td>9,333</td>
<td>A<em>02/A</em>32/B<em>08/B</em>55/C<em>03/C</em>07</td>
</tr>
<tr>
<td>Median</td>
<td>440</td>
<td>39,550</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* VC, viremic controller; EC, elite controller; TP, typical progressor; DP, delayed progressor. DP refers to an individual who previously was a controller and later progressed.

a Data represent viral loads on the day of sample collection.

b Data represent CD4 counts within 4 weeks of sample collection.

c Viral loads of 10,000 copies/ml compared to that determined for the prior year.

d B*27 and B*57 alleles are shown in bold. ND, not determined.
peated the variant recognition assessment in five progressors who started antiretroviral therapy during the study period. Following 6 months of treatment, we observed a trend toward reduced epitope variant recognition in all five progressors ($P = 0.06$; Fig. 4D). Collectively, these data suggest that while viral load can explain the differential levels of variant recognition observed between viremic controllers and progressors, there are likely alternative mechanisms responsible for driving and maintaining the high levels of variant recognition observed in elite controllers.

**Greater diversity in Gag epitopes found in progressors.** Persistence of viral antigen has been shown to promote proliferation and maintain antigen-specific CTL populations during chronic viral infection (46). Therefore, the high levels of variant-specific responses observed in progressive infection may be maintained due to the presence of highly diverse Gag variants in the viral population. To address the association between viral diversity and epitope variant recognition, we directly sequenced 10 to 25 individual Gag genes from plasma RNA from five viremic controllers and five progressors. We observed a significant correlation between viral load and mean pairwise distance ($r = 0.75, P = 0.02$; Fig. 5A). Diversity of the entire Gag protein was marginally higher in progressors than in controllers ($P = 0.05$, data not shown), even though controllers were on average infected for significantly longer (median, 20 years) than progressors (4 years, $P = 0.008$; data not shown), and diversity usually increases with duration of infection (47). This effect was more pronounced when examining pairwise distance only in viral regions corresponding to targeted epitopes ($P = 0.0005$; Fig. 5B).

We next determined the relationship between the epitope variants that were recognized and the viral variants present in the autologous population. If the observed variant-specific T cells were effective at eliminating infected cells, we would not expect to see recognized epitope variants present in the plasma viral population. Within targeted epitopic regions, viremic controllers had more viral variants present in the plasma that were tested but not recognized than progressors ($P = 0.008$; Fig. 5C), suggesting greater evidence for escape in viremic controllers. In both controllers and progressors, we found targeted epitopic regions for which autologous viral variants were not contained in our peptide set ($P = 0.34$; Fig. 5C). These autologous variants represent uncommon sequences present in less than 5% of clade B sequences in the HIVDB. Although our data show that these regions are targeted, we are unable to address whether these autologous variants are recognized. Intriguingly, viral populations in progressors tended to more often match the recognized epitope variants than those in viremic controllers ($P = 0.07$; Fig. 5C).

**Coverage of frequently occurring Gag variants is associated with control of viral replication.** An alternative approach to investigating the capacity of CTLs to cope with HIV sequence diversity is to evaluate recognized variants relative to the frequency with which they occur in circulating clade B sequences. For example, if an individual recognizes 3/5 tested variants (60%) but each of these variants is present in only 5% of circulating sequences, then the actual sequence coverage afforded through that variant recognition is quite poor (i.e., 5% each or 15% total). To assess the level of coverage in a given epitopic region, we summed the frequency of recognized 11mers versus the summed frequency of all tested 11mers in that region (see Table S2 in the supplemental material). Although our peptide set did not test every possible clade B Gag variant in the database, its design ensured that we were evaluating coverage of all frequently occurring Gag variants. For each targeted epitopic region, including both conserved and variable responses, we determined the level of sequence coverage through the recognized variants (see Fig. S7A and B). Controllers and progressors had comparable levels of overall sequence coverage through the Gag-specific repertoire ($P = 0.29$); however, elite controllers trended toward greater sequence coverage than both viremic controllers and progressors ($P = 0.09$; Fig. 6A). Interestingly, although viremic controllers recognized significantly fewer variants than progressors (KW test $P = 0.38$; Fig. 6A), coverage levels were equivalent in the two groups, indicating that the number of variants recognized does not necessarily translate into broad sequence coverage. We found a significant inverse correlation between sequence coverage and viral load ($r = -0.38, P = 0.03$; Fig. 6B). There was no correlation between sequence coverage and response magnitude ($r = -0.01, P = 0.95$; data not shown).

**Increased sequence coverage is correlated with recognition of more founder viruses.** Several vaccination strategies have designed multivalent immunogens with hopes of inducing T-cell responses that provide broad sequence coverage (19, 20, 24). Intuitively, increased sequence coverage should increase the number of potentially infecting sequences that would be recognized. To
explore how the sequence coverage measured in this study can inform future vaccine evaluation, we analyzed all 347 responses relative to how often they matched a panel of eight founder viruses. This panel is a collection of clade B viruses isolated from recently infected individuals (median, 7 days post-onset of symptoms) in the same demographic population as the majority of participants in this study (14). For each targeted epitopic region, we determined what proportion of the founder viruses would be recognized given the epitope specificity of that individual’s T-cell response (see Fig. S8A and B in the supplemental material). While we did not see a correlation between recognition of founder viruses and median epitope variant recognition ($r = 0.22$, $P = 0.23$; Fig. 6C), we did find a significant correlation between the number of founder viruses that were recognized and overall sequence coverage ($r = 0.50$, $P = 0.004$ Fig. 6D), suggesting that for vaccine-induced responses, a greater protective potential may be achieved

FIG 2 Controllers and progressors target conserved and variable regions of Gag. Data represent the number of controllers and progressors that target each amino acid residue in Gag p17 (top), p24 (middle), and p15 (bottom panel), as well as the conservation score for each residue (gray line). *, $P < 0.05$.

FIG 3 Higher variant recognition in progressors. Epitope variant recognition is measured as the number of 11mers recognized per number of 11mers tested in a targeted epitopic region. For each study subject, the median percentages of variant recognition were calculated across all targeted epitopic regions and compared between controllers and progressors ($P = 0.05$). EC, black circles; VC, open circles; progressors, gray squares.
by targeting the most commonly occurring variants rather than simply a large number of variants.

**DISCUSSION**

Our report provides the first investigation into the ability to recognize all frequently occurring Gag variants in subjects with and without spontaneous control of HIV replication. In agreement with a recent study by Mothe et al., the shorter peptide length, tight overlap, and vast sequence representation in our peptide set enabled us to detect a larger breadth of Gag-specific responses and that the targeting of the Gag protein by controllers is remarkably similar to the targeting of the Gag protein by progressors (11). We found that recognition of epitope variants was widespread in both controlled and progressive HIV infections. Interestingly, while responses to conserved epitopes and high numbers of HIV variants have each previously been associated with immune control of HIV infection (9, 11, 12, 48), our study identified the overall sequence coverage provided through the Gag-specific repertoire as a superior correlate of immune responses associated with virologic control.

Cross-reactivity has been proposed as an essential feature of T-cell receptors (TCR) (49, 50). Recognition of naturally occurring Gag variants can result either from a T cell expressing a TCR that is exceptionally flexible or through several different clonotypes, each recognizing one epitope variant or a range of epitope variants. As we identified responses by cytokine secretion using PBMC, our assay was unable to distinguish between these two possibilities. In viremic individuals, we found that contemporaneous viremia was associated with the extent of epitope variant recognition. Investigations into virus-host dynamics during acute HIV infection have shown that de novo responses to viral variants develop throughout infection (15, 31, 32, 51). Therefore, as different TCRs have differential abilities to tolerate variation within epitopes, sequential de novo responses associated with high viremia have the potential to increase the breadth and depth of responses, resulting in the extensive epitope variant recognition observed in progressive infection. Interestingly, we observed greater evidence for escape (i.e., fixation of a viral variant that was tested but not recognized) in epitopes targeted by viremic controllers than in those targeted by progressors. While these responses appear to be effective at driving escape, the lower level of variant recognition in viremic controllers may reflect that these individuals make fewer de novo responses to viral variants, possibly due to the weak potential of low viremia to prime new responses or to selection of escape mutations that abrogate epitope-HLA binding (52).

Constant antigenic stimulation has been shown to be necessary for long-term persistence of virus-specific CTLs during chronic viral infection (31, 46). The lack of IL-2 production in the majority of responses to epitope variants, the association between response magnitude and variant recognition, and the contraction of variant recognition upon ARV-mediated suppression of viral load support a model in which the maintenance of variant-specific responses is antigen dependent. Increasingly diverse viral quasispecies can also provide more opportunities for different TCRs to engage with a recognized antigen to stimulate T-cell proliferation. However, persistent antigenic stimulation can also promote T-cell exhaustion and functional impairment (53–57). Exhausted T cells characteristically exhibit weak cytotoxic potential and loss of poly-
functionality but maintain the ability to secrete IFN-γ (57). In the viral populations of progressors and, to a lesser extent, of viremic controllers, we found evidence for ineffective variant-specific responses, as many viral variants were present in the population despite being recognized. As this is a cross-sectional study, we cannot rule out the possibility that the viral populations were in the process of escaping (in the cases in which a recognized variant was a minor variant) or eventually would escape (in the cases in which a recognized variant was a major variant) in response to immune pressure. Given that the majority of responses to epitope variants detected in this analysis were monofunctional IFN-γ-positive responses, the ineffectiveness of these responses may have been a consequence of exhaustion. Our results are consistent with previous reports that have shown the ability to recognize peptide variants as measured by IFN-γ secretion does not directly translate into an ability to inhibit replication of viruses harboring recognized variants (58, 59). Taken together, our data suggest that a perpetual cycle exists in viremic individuals by which the combination of high viral loads and diversity can drive the development and persistence of variant-specific responses while simultaneously rendering these responses ineffective at fully suppressing viral replication, ultimately permitting the emergence of new viral variants.

Elite controllers maintain levels of epitope variant recognition in the absence of antigenic stimulation comparable to those in progressors, suggesting that there may be alternative immunological mechanisms responsible for preserving the ability to recognize epitope variants. T-cell functional avidity has been associated with mediating epitope variant recognition (10, 11, 60). Given that functional avidity increases in the absence of antigen (55), the observed high levels of variant recognition in elite controllers may be due to the maintenance of a high-avidity Gag repertoire. Additionally, it has been reported that compared to typical progressors, B*27 and B*57 elite controllers possess distinct clonotypes that have greater cross-reactivity to epitope variants (45). Although we did not detect a significant difference in the overall variant recognition between individuals with and without B*57 and B*27 alleles, on a per-T-cell basis individual highly cross-reactive clonotypes may have been contributing to the observed levels of variant recognition found in elite controllers.

We found that increased sequence coverage by Gag-specific T-cell responses is associated with lower viral loads in infected subjects. The combination of all variant-specific and conserved responses contributes to overall sequence coverage, highlighting the importance of the collective Gag-specific repertoire in control of viral replication. On a single-epitope level, coverage of frequently occurring variants can lead to the selection of rare mutations to permit escape. Given the weak relationship between sequence conservation and viral fitness costs (61–64), even high sequence coverage of a single epitope may not translate directly to control. However, recent analyses have revealed extensive covariation and networks of interdependent amino acids within the Gag protein (44, 65, 66). Structural and functional constraints limit the tolerance of multiple simultaneous mutations within coevolving Gag sites (44). Additionally, accumulation of mutations within Gag has recently been shown to decrease HIV replicative...
capacity as well as to contribute to reducing the viral load set point when transmitted to a new recipient (67, 68). Therefore, broad sequence coverage afforded through a functionally effective Gag-specific repertoire has the potential to select for multiple rare mutations across the protein. Increasing the overall level of coverage increases the likelihood that the consequences of escape will compromise the viability of the virus and allow for immune control.

Increased sequence coverage through an individual’s Gag-specific repertoire, but not variant recognition, correlated with higher recognition of different founder viruses. Although our study was limited to clade B infections, these results are encouraging for the protective potential of several vaccination strategies designed to elicit responses of broad coverage to global HIV sequences. Polyvalent mosaic vaccines and conserved-element vaccines both utilize multivalent immunogens to induce variant-specific responses that will ideally provide broad coverage either against whole proteins (19) or primarily against conserved regions (24). Our results suggest that the emphasis in the evaluation of these vaccines should be on the combined sequence coverage that can be achieved through all induced responses. However, our results were defined based on only Gag-specific responses, and it remains to be determined whether the protective potential of sequence coverage is applicable to other HIV proteins. Additionally, immunogenicity evaluation of these vaccines may depend on the identification of T-cell functional qualities that correlate with effective variant recognition. Given that dual IFN-γ/IL-2 responses to epitope variants were uncommon in our study, we were unable to evaluate whether the IFN-γ/IL-2 FluoroSpot assay better predicts effective variant recognition than the traditional IFN-γ enzyme-linked immunosorbent spot (ELISpot) assay.

Taking these results together, our report provides a new potential correlate, “sequence coverage,” of protection from progression in HIV infection. Sequence coverage combines previously suggested markers for viral control (variant recognition and targeting of conserved epitopes) into a single measure that is more informative in predicting viral control. The inverse relationship between sequence coverage and viral load, as well as the positive correlation with the recognition of founder viruses, makes sequence coverage a candidate for assessment in novel vaccine strategies that aim at increasing the breadth and depth of HIV-specific immune responses, as well as those designed to focus the immune response on conserved regions.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R21AI078809, T32AI007140, P01AI057005, R01AI047086, and UM1AI068618.

The content of this publication is solely our responsibility and does not necessarily represent the official views of the National Institutes of Health.

We thank all study participants for their participation in this research study as well as the clinic staff for their critical work and dedication. We thank Paula Walsh and Dave Friedrich for technical support and Stephen Voght for editorial assistance.

REFERENCES


Downloaded from http://jvi.asm.org/ on January 8, 2014 by Cornell University Library

Sunshine et al.


