AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 39, Number 3, 2023

DOI: 10.1089/aid.2022.0094



Open camera or QR reader and scan code to access this article and other resources online.

# Highly Ambiguous HIV-1 Pol Positions Encoding Multiple Amino Acids Usually Result from Antiviral or Immune Selection Pressure

Kaiming Tao,[1] Soo-Yon Rhee,[1] Philip L. Tzou,[1] Susan P. Holmes,[2] and Robert W. Shafer[1]



**Abstract**

HIV-1 *pol* nucleotide ambiguities encoding amino acid mixtures occur commonly during population-based genotypic drug resistance testing. However, few studies have addressed the validity of sequences with fully ambiguous codons (FACs) containing codons translatable to more than four amino acids. We identified 839 published HIV-1 *pol* sequences with 846 FACs at 131 positions and determined their distribution relative to 215 HLA-associated *pol* positions (HAPs) and 84 drug-resistance positions. Among HIV-1 reverse transcriptase (RT) and protease sequences from antiretroviral therapy (ART)-naive and -experienced persons, there was a strong correlation between the likelihood a position was a FAC and that it was an HAP (Spearman's correlation coefficient rho >0.40; $p < 1e\text{-}6$). Among HIV-1 RT sequences from ART-experienced persons, there was a correlation between the likelihood that a position was a FAC and that it was a drug-resistance position (rho = 0.2; $p = 8e\text{-}4$). In the context of population-based genotypic resistance testing, FACs usually result from antiviral or immune selection pressure.

**Keywords:** antiretroviral therapies, immunogenetics, host evolution

The presence of multiple variants in a sample undergoing population-based dideoxyterminator sequencing results in more than one electrophoretic peak at the same position (i.e., nucleotide ambiguities). Although most nucleotide ambiguities are synonymous, mixtures that result in multiple amino acids are common, particularly at sites under immune or antiviral selection pressure.[1,2] Because the cell-mediated immune response is host specific, amino acids at cytotoxic T lymphocyte (CTL) epitopes frequently alternate between amino acids as HIV-1 is transmitted to hosts with different HLA molecules.[3]

Likewise, the emergence of HIV-1 drug-resistance (HIVDR) mutations during antiviral therapy and their replacement after therapy often transitions through a stage in which multiple amino acids are detected.[4,5] The presence of up to four amino acids at the same position occurs commonly when an HIVDR mutation requires a double nucleotide change. However, there are few studies of plasma virus sequences containing codons translatable into more than four amino acids.

We searched the Stanford HIV Drug Resistance Database (HIVDB) for group M HIV-1 protease, reverse transcriptase (RT), and integrase plasma virus population-based sequences from persons with known drug class treatment histories**.** For RT, we analyzed the first 300 amino acid positions. For each gene, antiretroviral therapy (ART)-experienced sequences were defined as those from persons receiving a drug targeting the gene. For persons containing multiple sequences, we selected the earliest sequence.

Mutations were defined as amino acid differences from the consensus subtype B reference sequence. Mixtures were defined as positions at which the translation of each codon in a sequence yielded ≥2 amino acids. Positions containing codons with multiple twofold ambiguous International Union of Pure and Applied Chemistry nucleotides (KMRSWY) translatable to ≥5 amino acids were classified as fully ambiguous codons (FACs). Hypermutated sequences or those containing more than 8 protease, 15 RT, or 10 integrase unusual mutations, defined as having a global prevalence <0.01%, were excluded.[6]

[1]Division of Infectious Diseases, Department of Medicine and [2]Department of Statistics, Stanford University, Stanford, California, USA.

HLA-associated positions (HAPs) were defined as positions containing amino acids reported in one or more of seven studies to be significantly correlated in a phylogenetic context with an HLA type.[7–13] All seven studies analyzed the first 300 RT positions and the complete protease and integrase. For each position, we determined the number of studies reporting that the position contained an HLA-associated amino acid.

We analyzed six data sets comprising three genes (protease, RT, and integrase) in two treatment contexts (ART-naive and ART-experienced). HIVDR positions were defined as 84 positions containing an amino acid assigned a mutation penalty score by the HIVDB HIVDR interpretation program.[14] Spearman's correlation coefficients were used to determine the association between the likelihood a position was a FAC and an HAP. For ART-experienced sequences, Spearman's correlation coefficients were used to determine the association between the likelihood a position was a FAC and an HIVDR position.

We analyzed 137,498 protease sequences from PI-naive persons and 39,507 from PI-experienced persons; 120,019 RT sequences from RT inhibitor (RTI)-naive persons and 67,393 from RTI-experienced persons; 19,588 integrase sequences from integrase strand transfer inhibitor (INSTI)-naive persons and 2,486 from INSTI-experienced persons. The combined sequences included the following group M subtypes and circulating recombinant forms (CRFs): A (6.2%), B (50.5%), C (14.5%), D (1.9%), F (1.8%), G (2.0%), CRF01_AE (11.8%), CRF02_AG (4.7%), and other CRFs or unique recombinant forms (6.6%).

The proportion of sequences containing ≥1 FAC was 0.10%, 0.29%, and 0.18% in ART-naive protease, RT, and

Table 1. Positions with Fully Ambiguous Codons and Number of Studies Reporting the Codon to Be at an HLA-Associated Position: Sequences from Antiretroviral Therapy-Naive Persons

| *Protease* | | | *RT* | | | *RT (continued)* | | | *Integrase* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] |
| 63 | 64 (0.047) | 3 | 207 | 70 (0.058) | 7 | 294 | 1 (0.002) | 2 | 125 | 5 (0.026) | 2 |
| 37 | 27 (0.020) | 4 | 245 | 46 (0.052) | 4 | 32 | 1 (0.001) | 1 | 119 | 4 (0.020) | 4 |
| 72 | 11 (0.008) | 2 | 123 | 39 (0.033) | 5 | 169 | 1 (0.001) | 1 | 194 | 4 (0.020) | 0 |
| 12 | 9 (0.007) | 4 | 39 | 36 (0.031) | 5 | 249 | 1 (0.001) | 1 | 112 | 3 (0.015) | 3 |
| 67 | 7 (0.005) | 1 | 211 | 20 (0.017) | 4 | 297 | 1 (0.002) | 1 | 95* | 3 (0.015) | 1 |
| 19 | 5 (0.004) | 5 | 135 | 15 (0.013) | 6 | 3 | 1 (0.001) | 0 | 124 | 2 (0.010) | 5 |
| 41 | 4 (0.003) | 3 | 35 | 11 (0.010) | 4 | 20 | 1 (0.001) | 0 | 11 | 1 (0.005) | 6 |
| 36 | 3 (0.002) | 4 | 6 | 10 (0.010) | 2 | 67* | 1 (0.001) | 0 | 163* | 1 (0.005) | 3 |
| 69 | 2 (0.001) | 2 | 173 | 9 (0.008) | 4 | 91 | 1 (0.001) | 0 | 39 | 1 (0.005) | 2 |
| 71 | 2 (0.001) | 2 | 142 | 8 (0.007) | 3 | 146 | 1 (0.001) | 0 | 50 | 1 (0.005) | 2 |
| 61 | 2 (0.001) | 1 | 174 | 8 (0.007) | 3 | 168 | 1 (0.001) | 0 | 134 | 1 (0.005) | 1 |
| 35 | 1 (0.001) | 5 | 162 | 7 (0.006) | 5 | 188* | 1 (0.001) | 0 | 277 | 1 (0.005) | 1 |
| 60 | 1 (0.001) | 2 | 248 | 6 (0.008) | 4 | 213 | 1 (0.001) | 0 | 9 | 1 (0.005) | 0 |
| 21 | 1 (0.001) | 0 | 4 | 5 (0.005) | 4 | 218 | 1 (0.001) | 0 | 38 | 1 (0.005) | 0 |
| 55 | 1 (0.001) | 0 | 11 | 4 (0.004) | 3 | 221* | 1 (0.001) | 0 | 55 | 1 (0.005) | 0 |
| 66 | 1 (0.001) | 0 | 200 | 4 (0.003) | 3 | 226 | 1 (0.001) | 0 | 70 | 1 (0.005) | 0 |
| 86 | 1 (0.001) | 0 | 250 | 4 (0.006) | 3 | 229 | 1 (0.001) | 0 | 96 | 1 (0.005) | 0 |
| 10* | 0 (0.000) | 4 | 122 | 3 (0.003) | 5 | 272 | 0 (0.000) | 6 | 99 | 1 (0.005) | 0 |
| 14 | 0 (0.000) | 4 | 215* | 3 (0.003) | 0 | 278 | 0 (0.000) | 6 | 160 | 1 (0.005) | 0 |
| 93 | 0 (0.000) | 4 | 43 | 2 (0.002) | 4 | 275 | 0 (0.000) | 5 | 239 | 1 (0.005) | 0 |
| 16 | 0 (0.000) | 3 | 177 | 2 (0.002) | 4 | 277 | 0 (0.000) | 5 | 31 | 0 (0.000) | 6 |
| 64 | 0 (0.000) | 3 | 165 | 2 (0.002) | 3 | 102 | 0 (0.000) | 4 | 218 | 0 (0.000) | 6 |
| 82* | 0 (0.000) | 3 | 204 | 2 (0.002) | 3 | 138* | 0 (0.000) | 4 | 45 | 0 (0.000) | 5 |
| 13 | 0 (0.000) | 2 | 60 | 2 (0.002) | 2 | 178 | 0 (0.000) | 4 | 167 | 0 (0.000) | 5 |
| 15 | 0 (0.000) | 2 | 145 | 2 (0.002) | 2 | 281 | 0 (0.000) | 4 | 188 | 0 (0.000) | 5 |
| 20* | 0 (0.000) | 2 | 179* | 2 (0.002) | 2 | 48 | 0 (0.000) | 3 | 219 | 0 (0.000) | 5 |
| 39 | 0 (0.000) | 2 | 134 | 2 (0.002) | 1 | 121 | 0 (0.000) | 3 | 10 | 0 (0.000) | 4 |
| 45 | 0 (0.000) | 2 | 228 | 2 (0.002) | 1 | 158 | 0 (0.000) | 3 | 14 | 0 (0.000) | 4 |
| 62 | 0 (0.000) | 2 | 69* | 2 (0.002) | 0 | 166 | 0 (0.000) | 3 | 32 | 0 (0.000) | 4 |
| 70 | 0 (0.000) | 2 | 132 | 2 (0.002) | 0 | 243 | 0 (0.000) | 3 | 34 | 0 (0.000) | 4 |
| 74* | 0 (0.000) | 2 | 148 | 2 (0.002) | 0 | 276 | 0 (0.000) | 3 | 72 | 0 (0.000) | 4 |
| 77 | 0 (0.000) | 2 | 40 | 1 (0.001) | 3 | 279 | 0 (0.000) | 3 | 136 | 0 (0.000) | 4 |
| | | | 196 | 1 (0.001) | 3 | 286 | 0 (0.000) | 3 | 203 | 0 (0.000) | 4 |
| | | | 36 | 1 (0.001) | 2 | | | | 206 | 0 (0.000) | 4 |
| HAP (rho = 0.47, $p$ = 4.7e-07)[c] | | | HAP (rho = 0.46, $p$ < 1e-16)[c] | | | | | | HAP (rho = 0.16, $p$ = .004)[c] | | |

[a]The FAC percentage indicates the number of times that a position had a FAC divided by the total number of sequences.
[b]Number of studies showing that the position is an HAP in sequences from ART-naive persons.
[c]Spearman's rank correlation coefficients (rho) are shown for the association between FACs and HAPs in protease, RT, and integrase. HIVDR positions are indicated with an asterisk. Because of space limitations not all of the HAPs are shown (Supplementary File S3).
HAP, HLA-associated position; HIVDR, HIV-1 drug-resistance; Pos, position; RT, reverse transcriptase.

integrase, and 0.18%, 0.35%, and 0.36% in ART-experienced protease, RT, and integrase, respectively. Of the 839 sequences with a FAC, 7 (0.8%) had two FACs. Overall, there were 249 studies containing a sequence with a FAC, including 70 studies containing ≥1,000 sequences. The median proportion of sequences with a FAC in these 70 studies was 0.2% (range: 0.06%–1.2%).

In protease, 34 of 99 positions had an HAP in ≥1 study including 26 for which two to five studies found the position to be an HAP (Supplementary File S1). In RT, 90 of the first 300 positions had an HAP in ≥1 study including 54 for which two to seven studies found the position to be an HAP (Supplementary File S1). In integrase, 91 positions had an HAP in ≥1 study including 57 for which two to six studies found the position to be an HAP (Supplementary File S1).

Tables 1 and 2 list each of the FACs and most of the HAPs in the three ART-naive and three ART-experienced sequence sets, respectively. In the ART-naive and -experienced protease and RT sequence sets, there was a strong correlation between the number of times a position had a FAC and the number of times it was reported to be an HAP (Table 1): ART-naive protease (rho = 0.47, $p = 4.7e-07$) and RT (rho = 0.46; $p < e-16$); ART-experienced protease (rho = 0.44, $p = 2.1e-6$) and RT (rho = 0.41; $p = 2.1e-14$).



Table 2. Positions with Fully Ambiguous Codons and Number of Studies Reporting the Codon to Be at an HLA-Associated Position: Sequences from Antiretroviral Therapy-Experienced Persons

| *Protease* | | | *RT* | | | *RT (continued)* | | | *Integrase* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] | *Pos* | *FAC* n *(%)*[a] | *No. of studies*[b] |
| 63 | 16 (0.041) | 3 | 207 | 26 (0.039) | 7 | 166 | 1 (0.002) | 3 | 194 | 6 (0.246) | 0 |
| 37 | 12 (0.030) | 4 | 40 | 24 (0.036) | 3 | 174 | 1 (0.001) | 3 | 134 | 2 (0.080) | 1 |
| 20* | 9 (0.023) | 2 | 123 | 23 (0.034) | 5 | 204 | 1 (0.001) | 3 | 125 | 1 (0.040) | 2 |
| 67 | 6 (0.015) | 1 | 245 | 13 (0.024) | 4 | 145 | 1 (0.001) | 2 | 11 | 0 (0.000) | 6 |
| 82* | 4 (0.010) | 3 | 190* | 13 (0.019) | 0 | 203 | 1 (0.001) | 2 | 31 | 0 (0.000) | 6 |
| 72 | 3 (0.008) | 2 | 200 | 9 (0.013) | 3 | 294 | 1 (0.001) | 2 | 218 | 0 (0.000) | 6 |
| 79 | 3 (0.008) | 0 | 39 | 8 (0.013) | 5 | 32 | 1 (0.001) | 1 | 45 | 0 (0.000) | 5 |
| 19 | 2 (0.005) | 5 | 101* | 6 (0.009) | 0 | 88 | 1 (0.001) | 1 | 124 | 0 (0.000) | 5 |
| 12 | 2 (0.005) | 4 | 135 | 5 (0.007) | 6 | 103* | 1 (0.001) | 1 | 167 | 0 (0.000) | 5 |
| 36 | 2 (0.005) | 4 | 138* | 5 (0.007) | 4 | 106* | 1 (0.001) | 1 | 188 | 0 (0.000) | 5 |
| 18 | 2 (0.005) | 0 | 142 | 5 (0.007) | 3 | 111 | 1 (0.002) | 1 | 219 | 0 (0.000) | 5 |
| 84* | 1 (0.003) | 0 | 179* | 5 (0.007) | 2 | 134 | 1 (0.001) | 1 | 10 | 0 (0.000) | 4 |
| 10* | 1 (0.003) | 4 | 69* | 5 (0.007) | 0 | 159 | 1 (0.001) | 1 | 14 | 0 (0.000) | 4 |
| 41 | 1 (0.003) | 3 | 215* | 5 (0.007) | 0 | 237 | 1 (0.001) | 1 | 32 | 0 (0.000) | 4 |
| 45 | 1 (0.003) | 2 | 122 | 4 (0.006) | 5 | 249 | 1 (0.001) | 1 | 34 | 0 (0.000) | 4 |
| 60 | 1 (0.003) | 2 | 68 | 4 (0.006) | 1 | 19 | 1 (0.001) | 0 | 72 | 0 (0.000) | 4 |
| 70 | 1 (0.003) | 2 | 70* | 4 (0.006) | 0 | 22 | 1 (0.001) | 0 | 119 | 0 (0.000) | 4 |
| 7 | 1 (0.003) | 0 | 75* | 4 (0.006) | 0 | 38 | 1 (0.002) | 0 | 136 | 0 (0.000) | 4 |
| 38 | 1 (0.003) | 0 | 188* | 4 (0.006) | 0 | 46 | 1 (0.003) | 0 | 203 | 0 (0.000) | 4 |
| 48* | 1 (0.003) | 0 | 162 | 3 (0.004) | 5 | 118 | 1 (0.002) | 0 | 206 | 0 (0.000) | 4 |
| 35 | 0 (0.000) | 5 | 43 | 3 (0.004) | 4 | 130 | 1 (0.002) | 0 | 28 | 0 (0.000) | 3 |
| 14 | 0 (0.000) | 4 | 177 | 3 (0.004) | 4 | 144 | 1 (0.002) | 0 | 74* | 0 (0.000) | 3 |
| 93 | 0 (0.000) | 4 | 248 | 3 (0.004) | 4 | 151* | 1 (0.001) | 0 | 91 | 0 (0.000) | 3 |
| 16 | 0 (0.000) | 3 | 139 | 3 (0.007) | 1 | 157 | 1 (0.001) | 0 | 101 | 0 (0.000) | 3 |
| 64 | 0 (0.000) | 3 | 201 | 3 (0.004) | 1 | 185 | 1 (0.001) | 0 | 112 | 0 (0.000) | 3 |
| 13 | 0 (0.000) | 2 | 67* | 3 (0.004) | 0 | 195 | 1 (0.001) | 0 | 163* | 0 (0.000) | 3 |
| 15 | 0 (0.000) | 2 | 35 | 2 (0.005) | 4 | 220 | 1 (0.001) | 0 | 193 | 0 (0.000) | 3 |
| 39 | 0 (0.000) | 2 | 250 | 2 (0.004) | 3 | 242 | 1 (0.001) | 0 | 207 | 0 (0.000) | 3 |
| 62 | 0 (0.000) | 2 | 6 | 2 (0.003) | 2 | 260 | 1 (0.002) | 0 | 208 | 0 (0.000) | 3 |
| 69 | 0 (0.000) | 2 | 148 | 2 (0.005) | 0 | 272 | 0 (0.000) | 6 | 220 | 0 (0.000) | 3 |
| 71 | 0 (0.000) | 2 | 219* | 2 (0.003) | 0 | 278 | 0 (0.000) | 6 | 256 | 0 (0.000) | 3 |
| 74* | 0 (0.000) | 2 | 102 | 1 (0.001) | 4 | 275 | 0 (0.000) | 5 | 265 | 0 (0.000) | 3 |
| 77 | 0 (0.000) | 2 | 173 | 1 (0.001) | 4 | 277 | 0 (0.000) | 5 | 269 | 0 (0.000) | 3 |
| | | | 211 | 1 (0.001) | 4 | 4 | 0 (0.000) | 4 | | | |
| | | | 11 | 1 (0.001) | 3 | 178 | 0 (0.000) | 4 | | | |
| | | | 165 | 1 (0.001) | 3 | 281 | 0 (0.000) | 4 | | | |
| HAP (rho = 0.44, $p = 2.1e-06$)[c] HIVDR positions (rho = −0.002, $p = .5$)[c] | | | HAP (rho = 0.41, $p = 2.1e-14$)[c] HIVDR positions (rho = 0.18, $p = .0008$)[c] | | | | | | HAP (rho = 0.062, $p = .15$)[c] HIVDR positions (rho = −0.031, $p = .7$)[c] | | |

[a]The FAC percentage indicates the number of times that a position had a FAC divided by the total number of sequences.
[b]Number of studies showing that the codon is at an HAP in sequences from ART-experienced persons.
[c]Spearman's rank correlation coefficients (rho) are shown for the association between FACs and HAPs in protease, RT, and integrase. HIVDR positions are indicated with an asterisk. Because of space limitations not all of the HAPs are shown (Supplementary File S2).

There was a weak correlation between FACs and HAPs in the ART-naive integrase data set (rho = 0.16; $p = 4e\text{-}3$).

For the 846 FACs occurring at an HAP, 51.0% of the nonconsensus amino acids were among the amino acids present in the corresponding HAP. The presence of two or three amino acids at a position can result in a codon that when translated contains more mutations than are simultaneously present in a sample. Thus, some of the amino acids in FACs may represent transitional amino acids that are no longer present in a sample. Supplementary File S3 shows the specific amino acids shared by FACs and HAPs.

In the ART-experienced data sets, there was a significant positive correlation between FACs and HIVDR positions in the RT data set (rho = 0.20; $p = 8e\text{-}4$) but not in the protease and integrase data sets (Table 2). In RT, there were 14 HIVDR positions containing one or more FACs, including positions 190, 101, 138, 179, 69, 215, 70, 75, 188, 67, 219, 103, 106, and 151. At each of these positions except 103 and 151, ≥3 different HIVDR mutations have been reported. In the ART-experienced protease data set, there were five HIVDR positions with FACs and in integrase, there were no HIVDR positions with FACs.

Of the complete set of 846 FACs, 90.3% occurred at an HAP (83.2%), an HIVDR position (3.9%), or a position that was both an HAP and HIVDR position (3.2%). For RT and protease, but not for integrase, FACs were overwhelmingly likely to result from immune or antiviral selection pressure. We were unable to explain the markedly reduced correlation of integrase FACs with HAPs and HIVDR positions. Although the number of integrase sequences was about one-eighth the number of protease and RT sequences, these sequences contained a proportion of FACs similar to RT. Moreover, similar to RT and protease, integrase also contains many HAPs and HIVDR positions.

FACs occurred more often at HAPs than at HIVDR positions because the number of CTL escape positions was greater than the number of HIVDR positions and because HIVDR positions are usually highly conserved. We correlated FACs with HAPs rather than specific CTL epitopes because CTL escape can result from mutations upstream, downstream, and at different positions within an epitope.[15]

Our analysis excluded sequences with poor sequence quality. Had we not excluded these sequences, it is possible that a greater proportion of FACs may have resulted from technical artifact rather than from immune or antiviral selection pressure. Nonetheless, our data suggest that in the context of genotypic resistance testing, the presence of an HIVDR mutation within a FAC should be reported to HIV care providers.

## Authors' Contributions

K.T., S.-Y.R., and P.L.T. performed the analyses in the study. S.P.H. assisted with statistical analyses. R.W.S. wrote the article. K.T. and S.P.H. reviewed the article and provided edits.

## Author Disclosure Statement

R.W.S. has served on advisory boards for Gilead Sciences and GlaxoSmithKline.

## Funding Information

K.T., S.-Y.R., P.L.T., S.P.H., and R.W.S. were funded in part by a grant from the National Institute of Allergy and Infectious Diseases (NIAID), R24AI136618.

## Supplementary Material

Supplementary File S1
Supplementary File S2
Supplementary File S3

## References

1. Poon AFY, Pond SLK, Bennett P, et al. Adaptation to human populations is revealed by within-host polymorphisms in HIV-1 and hepatitis C virus. PLoS Pathog 2007; 3(3):e45; doi: 10.1371/journal.ppat.0030045
2. Poon AFY, Kosakovsky Pond SL, Richman DD, et al. Mapping protease inhibitor resistance to human immunodeficiency virus type 1 sequence polymorphisms within patients. J Virol 2007;81(24):13598–13607; doi: 10.1128/JVI.01570-07
3. Delport W, Scheffler K, Seoighe C. Frequent toggling between alternative amino acids is driven by selection in HIV-1. PLoS Pathog 2008;4(12):e1000242; doi: 10.1371/journal.ppat.1000242
4. Wei X, Ghosh SK, Taylor ME, et al. Viral dynamics in human immunodeficiency virus type 1 infection. Nature 1995;373(6510):117–122; doi: 10.1038/373117a0
5. Devereux HL, Youle M, Johnson MA, et al. Rapid decline in detectability of HIV-1 drug resistance mutations after stopping therapy. AIDS 1999;13(18):F123–F127.
6. Tzou PL, Kosakovsky Pond SL, Avila-Rios S, et al. Analysis of unusual and signature APOBEC-mutations in HIV-1 pol next-generation sequences. Mor O. ed. PLoS One 2020;15(2):e0225352; doi: 10.1371/journal.pone.0225352
7. Brumme ZL, John M, Carlson JM, et al. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. Nixon DF. ed. PLoS One 2009;4(8):e6687; doi: 10.1371/journal.pone.0006687
8. Carlson JM, Listgarten J, Pfeifer N, et al. Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. J Virol 2012;86(9):5230–5243; doi: 10.1128/JVI.06728-11
9. Carlson JM, Schaefer M, Monaco DC, et al. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science 2014;345(6193):1254031; doi: 10.1126/science.1254031
10. Chikata T, Carlson JM, Tamura Y, et al. Host-specific adaptation of HIV-1 subtype B in the Japanese population. J Virol 2014;88(9):4764–4775; doi: 10.1128/JVI.00147-14
11. Rousseau CM, Daniels MG, Carlson JM, et al. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype C proteome: Immune escape and viral load. J Virol 2008;82(13):6434–6446; doi: 10.1128/JVI.02455-07
12. Van Tran G, Chikata T, Carlson JM, et al. A strong association of human leukocyte antigen-associated Pol and Gag mutations with clinical parameters in HIV-1 subtype A/E infection. AIDS 2016;30(5):681–689; doi: 10.1097/QAD.0000000000000969

13. Kinloch NN, Lee GQ, Carlson JM, et al. Genotypic and mechanistic characterization of subtype-specific HIV adaptation to host cellular immunity. J Virol 2019;93(1): e01502–e01518; doi: 10.1128/JVI.01502-18
14. Paredes R, Tzou PL, van Zyl G, et al. Collaborative update of a rule-based expert system for HIV-1 genotypic resistance test interpretation. PLoS One 2017;12(7):e0181357; doi: 10.1371/journal.pone.0181357
15. Carlson JM, Le AQ, Shahid A, et al. HIV-1 adaptation to HLA: A window into virus–host immune interactions. Trends Microbiol 2015;23(4):212–224; doi: 10.1016/j.tim .2014.12.008

Address correspondence to:
*Robert W. Shafer*
*Division of Infectious Diseases*
*Department of Medicine*
*Stanford University*
*240 Pasteur Drive*
*Stanford, CA 94305*
*USA*

*E-mail:* rshafer@stanford.edu