CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Research Repository

Murdoch
UNIVERSITY

MURDOCH RESEARCH REPOSITORY

This is the author's final version of the work, as accepted for publication following peer review but without the publisher's layout or pagination.
The definitive version is available at

http://dx.doi.org/10.1128/JVI.00147-14

Chikata, T., Carlson, J.M., Tamura, Y., Borghan, M.A., Naruto, T., Hashimoto, M., Murakoshi, H., Le, A.Q., Mallal, S., John, M., Gatanaga, H., Oka, S., Brumme, Z.L. and Takiguchi, M. (2014) Host-Specific adaptation of HIV-1 Subtype B in the Japanese population.
Journal of Virology, 88 (9). pp. 4764-4775.

http://researchrepository.murdoch.edu.au/21949/

Copyright: © 2014 American Society for Microbiology
It is posted here for your personal use. No further distribution is permitted.

JVI Accepts, published online ahead of print on 12 February 2014
J. Virol. doi:10.1128/JVI.00147-14
Copyright © 2014, American Society for Microbiology. All Rights Reserved.

JVI Accepts published online ahead of print

Host-specific adaptation of HIV-1 subtype B in the Japanese population

Takayuki Chikata[1], Jonathan M. Carlson[2], Yoshiko Tamura[1], Mohamed Ali Borghan[1,3], Takuya Naruto[1], Masao Hashimoto[1], Hayato Murakoshi[1], Anh Q. Le[4], Simon Mallal[5], Mina John[5], Hiroyuki Gatanaga[1,6], Shinichi Oka[1,6], Zabrina L. Brumme[4,7], and Masafumi Takiguchi[1]

[1]Center for AIDS Research, Kumamoto University, 2-2-1 Honjo, Chuo-ku, Kumamoto, Japan; [2]eScience Group, Microsoft Research, Los Angeles, CA, USA; [3]Department of Biological Sciences and Chemistry, College of Arts and Sciences, University of Nizwa, Nizwa, Oman; [4]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada; [5]Institute for Immunology & Infectious Diseases, Murdoch University, Murdoch, WA, Australia; [6]AIDS Clinical Center, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, Tokyo, Japan; [7]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada

Running title: HLA-associated polymorphisms in Japanese

Correspondence: Masafumi Takiguchi
Center for AIDS Research, Kumamoto University
2-2-1 Honjo, Chuo-ku, Kumamoto 860-0811, Japan
e-mail:masafumi@kumamoto-u.ac.jp
Tel: +81 96 373 6529, Fax: +81 96 373 6532

Word count: Text 7,021 words, Abstract 249 words

JVI Accepts published online ahead of print

26 **Abstract**

27

28 The extent to which HIV-1 clade B strains exhibit population-specific adaptations to host
29 HLA alleles remains incompletely known, in part due to incomplete characterization of
30 HLA-associated HIV-1 polymorphisms (HLA-APs) in different global populations.
31 Moreover, it remains unknown to what extent the same HLA alleles may drive significantly
32 different escape pathways across populations. As the Japanese population exhibits
33 distinctive HLA class I allele distributions, comparative analysis of HLA-APs between
34 HIV-1 clade B-infected Japanese and non-Asian cohorts could shed light on these questions.
35 However HLA-APs remain incompletely mapped in Japan. In a cohort of 430
36 treatment-naïve Japanese with chronic HIV-1 clade B infection, we identified 284
37 HLA-APs in Gag, Pol and Nef using phylogenetically-corrected methods. The number of
38 HLA-associated substitutions in Pol, notably those restricted by HLA-B*52:01, was weakly
39 inversely correlated with plasma viral load (pVL), suggesting that the transmission and
40 persistence of B*52:01-driven Pol mutations could modulate pVL. Differential selection of
41 HLA-APs between HLA subtype members, including those differing only with respect to
42 substitutions outside the peptide-binding groove, was observed, meriting further
43 investigation as to their mechanisms of selection. Notably, two-thirds of HLA-AP identified
44 in Japan had not been reported in previous studies of predominantly Caucasian cohorts, and
45 were attributable to HLA alleles unique to, or enriched in, Japan. We also identified 71
46 cases where the same HLA allele drove significantly different escape pathways between
47 Japan versus predominantly Caucasian cohorts. Our results underscore the distinct global
48 evolution of HIV-1 clade B as a result of host population-specific cellular immune
49 pressures.

JVI Accepts published online ahead of print

50 **Importance section**

51

52 CTL escape mutations in HIV-1 are broadly predictable based on the HLA class I alleles
53 expressed by the host. Because HLA allele distributions differ among worldwide
54 populations, the pattern and diversity of HLA-associated escape mutations are likely to be
55 somewhat distinct to each race and region. HLA-associated polymorphisms (HLA-APs) in
56 HIV-1 have previously been identified at the population level in European, North American,
57 Australian and African cohorts, however, large-scale analyses of HIV-1 clade B-specific
58 HLA-APs in Asians are lacking. Differential intra-clade HIV-1 adaptation to global
59 populations can be investigated via comparative analyses of HLA-associated
60 polymorphisms across ethnic groups, but such studies are rare. Here, we identify HLA-APs
61 in a large Japanese HIV-1 clade B cohort using phylogeneticaly-informed methods and
62 observe that the majority of them had not been previously characterized in predominantly
63 Caucasian populations. Results highlight HIV's unique adaptation to cellular immune
64 pressures imposed by different global populations.

JVI Accepts published online ahead of print

65 **Introduction**

66

67 HIV CTL escape occurs in a manner that is highly reproducible in context of the HLA
68 class I alleles expressed by the host (1-8). By extension, HIV sequences circulating in a
69 given host population will exhibit polymorphisms that reflect the HLA allele distribution of
70 that population (9). Because HLA class I allele distributions differ among racial and ethnic
71 groups worldwide (10), the pattern and diversity of HLA-associated escape mutations is
72 also likely to be somewhat distinct to each race and region. Numerous population-based
73 studies identifying HLA-associated polymorphisms (HLA-APs) have been conducted in
74 European, North American, Australian, and African cohorts (2, 6, 8). However, comparably
75 fewer have been undertaken in Asian cohorts, where HIV-1 prevalence is also substantial
76 (11). Since Asian populations differ in their HLA allele distributions from the cohorts
77 previously studied, it is important to identify and analyze HLA-APs to achieve a better
78 understanding of HIV-1 pathogenesis in Asia and to inform future HIV vaccine design
79 efforts targeted to these populations. The Japanese epidemic is somewhat unique in Asia.
80 While clades A/E and C predominate in many Asian countries (12-14), the Japanese HIV-1
81 epidemic comprises 80% clade B infections (12). As such, the analysis of Japanese cohorts
82 also provides the opportunity to undertake comparative analyses of HLA-APs between
83 Asian and non-Asian populations infected with HIV clade B.
84 Previous studies have investigated differential HLA-driven HIV evolution across human
85 populations. For example, a study of HLA-specific adaptations in HIV Pol in a Mexican
86 cohort identified "unique" HLA-APs in this population that were not present in an
87 international cohort from Canada, the USA and Australia, even though both cohorts
88 harbored HIV clade B (15). Most of the unique Mexican HLA-APs were restricted by HLA
89 alleles particular to this population (*e.g.* HLA-B*39) but that were underrepresented or
90 absent in the international cohort (15). This study therefore illustrates population-specific
91 HIV adaptation in its most intuitive manifestation: where distinctive HLA-associated

JVI Accepts published online ahead of print

92 polymorphisms are observed in a population due to the presence (or comparatively higher
93 frequency) of an HLA allele in this population compared to another.

94 What remains unknown however, is the extent to which the same HLA allele may drive
95 divergent escape pathways in different human populations. Two critical features are
96 required to address this question. First, the identification of HLA-AP must be undertaken at
97 the HLA subtype level. This is because the majority (>60%) of HLA-associated
98 polymorphisms are best defined at the subtype-level (16) – even for closely related HLA
99 subtype members that present the same or similar peptide epitopes (16, 17, 18 , 19).
100 Comparative studies undertaken at allele-level (two-digit) resolution cannot disentangle
101 whether population-specific HLA-AP are attributable to differential HLA subtype
102 distributions between cohorts, or whether they are "true" cases where the same HLA
103 subtype drives different escape pathways across populations. Indeed, a study investigating
104 >500 Americans with chronic HIV-1 clade B infection observed distinct patterns of
105 HLA-APs among White, Black, and Hispanic individuals that were likely attributable to the
106 differential distribution of closely-related HLA subtypes among these groups (18) rather
107 than true differential escape. The present study is therefore undertaken at subtype-level
108 resolution. Secondly, the identification of population-specific escape pathways driven by
109 the same HLA allele requires a method to do so. Here, we adapt phylogenetically-corrected
110 statistical methods originally developed to assess differential escape among related HLA
111 subtypes (17) and apply them to investigate differential escape across host populations.

112 The present study is divided into two parts, each with a specific major objective. Our
113 first objective was to identify and characterize HLA-AP in HIV-1 Gag, Pol, and Nef
114 proteins in a cohort of 430 chronically clade B-infected Japanese individuals using
115 phylogenetically-informed approaches (20), and to investigate their associations with
116 clinical parameters (CD4+ T cell count and plasma viral load). Importantly, HLA
117 genotyping (and thus HLA-AP identification) was undertaken at subtype-level resolution,
118 allowing us to analyze the effect of genetic differences among closely-related HLA

JVI Accepts published online ahead of print

119 subtypes on the selection of HLA-APs in the Japanese cohort as part of this objective. Our
120 second major objective was to perform a comparative analysis of HLA-AP identified in
121 Japan to those identified in a large international (Canada/USA/Australia) cohort of
122 antiretroviral-naïve, chronically clade B infected, predominantly Caucasian individuals. As
123 expected, a substantial proportion of Japanese HLA-AP were restricted by alleles unique to
124 (or highly enriched in) Japan compared to the non-Asian cohort. Notably we also observed
125 numerous cases where the same HLA allele drove significantly different – sometimes
126 opposing – escape pathways in these two populations. Our results highlight HIV's unique
127 adaptation to cellular immune pressures imposed by different global populations.

JVI Accepts published online ahead of print

128 **Materials and Methods**

129

130 **Ethics statement**

131 This study was approved as a part of "the study of immunological and virological
132 analysis in HIV-1 infection (#540)" by the ethics committee for epidemiology and general
133 study in the faculty of life science in Kumamoto University and the National Center for
134 Global Health and Medicine. All studied individuals were adults. Written informed consent
135 was obtained from all studied individuals according to the Declaration of Helsinki.

136

137 **Subjects**

138 Four-hundred-thirty treatment-naïve Japanese individuals with chronic HIV-1 clade B
139 infection were enrolled in the National Center for Global Health and Medicine (NCGHM)
140 from 2008 to 2011. HLA alleles of these individuals were determined at the 4-digit level by
141 a probe-based sequence-specific oligonucleotide (SSO) typing method (HLA Laboratory,
142 Kyoto, Japan). The median CD4+ T cell count (CD4 count) and plasma viral loads (pVL) at
143 the first visit to NCGHM were 321 cells/μl (IQR: 190 to 440 cells/μl) and 25,000 copies/ml
144 (IQR: 6,800 to 98,000 copies/ml), respectively.

145 HLA-associated polymorphisms derived from the International HIV Adaptation
146 Collaborative (IHAC) cohort, comprising 1,888 treatment-naïve individuals with chronic
147 clade B infection from Canada, USA, and Western Australia (16), identified using identical
148 methods, were used for comparison. The IHAC cohort comprises predominantly Caucasian
149 individuals, with Asians making up less 5% of the total. The median CD4 count in IHAC
150 was 260 cells/μl (IQR: 110 to 418 cells/μl).

151

152 **RT-PCR and sequencing of plasma HIV RNA**

153 HIV-1 viral RNA was extracted from plasma samples using either a QIAamp MinElute
154 virus spin kit (Qiagen, Valencia, CA) or an EZ1 Virus Mini Kit v2.0 (Qiagen, Valencia, CA).

JVI Accepts published online ahead of print

155 Reverse transcription was performed using random hexamers with the SuperScript III
156 First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). HIV-1 Gag, Pol, and
157 Nef genes were amplified from cDNA by nested PCR using Taq DNA polymerase
158 (Promega, Fitchburg, WI) and 10 primer pairs that were designed based on the clade B
159 strain. For subjects with a viral load below 1000 copies/ml, RT-PCR was performed with
160 region-specific primers using the SuperScript III One-Step RT-PCR System with Platinum
161 Taq kit (Invitrogen, Carlsbad, CA). The 1st round PCR product was then used in the 2nd
162 round PCR amplification using Taq DNA polymerase (Promega, Fitchburg, WI) and the 10
163 primer pairs. The 2nd round PCR product was purified by using the ExoSAP-IT reagent
164 containing Exonuclease I and alkaline phosphatase (GE Healthcare, Buckinghamshire).
165 Gag, Pol, and Nef sequences were determined by using the BigDye Terminator v3.1 cycle
166 sequencing kit (Applied Biosystems, Carlsbad, CA) and an ABI 3500 genetic analyzer
167 (Applied Biosystems, Carlsbad, CA). Sequencing reactions were performed in both 5' and
168 3' directions to yield a minimum of bidirectional coverage of all regions. Sequence data
169 was then aligned by using SeqScape software (Applied Biosystems, Carlsbad, CA) based
170 on the HXB2 reference sequence (K03455). Accession numbers are AB873205 to
171 AB873601 (Gag), AB873908 to AB874270 (Pol), and AB873602 to AB873907 (Nef).
172
173 **Identification of HLA-associated polymorphisms**
174 HLA class I associated polymorphisms in HIV-1 (HLA-AP) can be identified in large,
175 cross-sectional linked datasets of host (HLA) and HIV genotypes using statistical
176 association approaches that identify viral polymorphisms significantly over- or
177 under-represented in individuals harboring a specific HLA class I allele (1, 2, 4, 16-18, 21).
178 HLA-APs that are over-represented in individuals harboring the relevant HLA are
179 commonly referred to as "adapted" forms, while those under-represented in individuals
180 harboring the relevant HLA are referred to as "nonadapted" forms (2, 18). As such,
181 "nonadapted" and "adapted" forms can be conceptualized to represent the

JVI Accepts published online ahead of print

182 "immunologically susceptible" and "escape mutant" forms, respectively, for the specific
183 HLA allele in question at that HIV codon position. Statistical association approaches for the
184 identification of HLA-AP also correct for the confounding influences of viral phylogeny,
185 HIV codon covariation and linkage disequilibrium between HLA class I alleles (2, 16, 17,
186 21).

187 Associations between HLA class I alleles and HIV-1 amino acid polymorphisms in the
188 Japanese and IHAC datasets were identified using a published phylogenetically-corrected
189 logistic regression model that corrects for HLA linkage disequilibrium, HIV phylogeny, and
190 HIV codon covariation as potential confounders (17, 20). Briefly, maximum likelihood
191 phylogenetic trees were constructed using Gag, Pol and Nef sequences (one tree per gene),
192 and a model of conditional adaptation was inferred for each observed amino acid at each
193 codon. Amino acids are assumed to evolve independently along the phylogeny, until the
194 tree tips (representing the present host). In each host, HLA-mediated selection and HIV
195 amino acid covariation are directly modeled using weighted logistic regression, in which
196 the individual's HLA repertoire and covarying HIV amino acids are used as binary
197 predictors and the bias is determined by the possible transmitted sequences as inferred by
198 the phylogeny (17). To identify which factors (HLA and/or HIV covariation) contribute to
199 selection pressure, we employ a forward selection procedure where the most significant
200 association is iteratively added to the model, with p-values computed using the likelihood
201 ratio test. We performed post-hoc filtering of the resulting HLA-associated polymorphisms
202 list, restricting our output to instances in which at least 10 individuals carried the allele or
203 polymorphism and at least 10 individuals did not carry the allele or polymorphism.
204 Multiple tests were accounted for using q-values, the p-value analog of the false discovery
205 rate (FDR) (22). The FDR is the expected proportion of false positives among results
206 deemed significant at a given threshold; for example, at $q < 0.2$, we expect 20% of
207 identified associations to be false positives. In the analyses identifying HLA-associated
208 polymorphisms (HLA-AP), significance threshold of $q < 0.2$ was employed.

JVI Accepts published online ahead of print

209

210 **Statistical analysis**

211 Correlations between the total number of HLA-associated substitutions in each
212 individual and clinical parameters (pVL and CD4 count) were performed using Spearman's
213 correlation. To count the total number of HLA-associated substitutions within a given
214 HIV-1 sequence, we first identified all HIV-1 sites within that sequence identified as being
215 associated with any HLA allele. The specific residue at each site was counted as
216 "HLA-associated" if it matched any HLA-associated adapted form, or any residue other
217 than a nonadapted form identified at that position. The HLA alleles expressed by the
218 individual were not considered (unless specifically stated) – rather, our goal was to
219 enumerate the number of HLA-AP associated with any HLA allele in each viral sequence.
220 In analyses where host HLA alleles were not considered, HIV sites harboring residues that
221 simultaneously represented a nonadapted and Adapted form associated with different HLA
222 alleles were excluded from consideration.

223

224 **Detection of differential escape between closely-related HLA alleles, and between**
225 **cohorts.**

226 Two types of differential escape were investigated. First, we investigated differential
227 escape between closely-related HLA class I alleles, defined here as (four-digit) HLA
228 subtype members belonging to the same (two-digit) allele group, in the Japanese cohort.
229 Specifically, seven HLA allele groups (A*02, A*26, B*15, B*40, C*03, C*08, and C*14)
230 for which a minimum of two subtype members were represented in the Japanese cohort,
231 were investigated. For example, the HLA-A*02 allele group featured subtypes A*02:01,
232 A*02:06 and A*02:07, while the A*26 allele group featured subtypes A*26:01 and
233 A*26:03. For each allele group, we took the union of all HLA-AP identified for all subtype
234 members of the group. Then, in a pairwise manner, we compared their strengths of
235 selection between all HLA subtype members using a previously-described

JVI Accepts published online ahead of print

236 phylogenetically-corrected interaction test (17). In this analysis, thresholds of $p < 0.05$, $q <$
237 0.2 were used to define significance.

238 Second, we investigated differential HLA-driven escape pathways between Japanese and
239 IHAC cohorts. As outlined in the introduction, HLA-AP identified in human populations
240 will differ to some extent due to the presence (or enrichment) of certain HLA alleles in one
241 population versus another. However in this analysis we were specifically interested in
242 identifying cases where the same HLA allele drove significantly different escape pathways
243 in the two cohorts. To do this, we took the union of all HLA-AP identified in Japan and
244 IHAC cohorts, that were restricted by HLA subtypes observed a minimum of 10 times in
245 both cohorts. We then compared the strength of selection of each HLA-AP in a pairwise
246 manner, between cohorts. The statistical methods used investigate differential escape
247 between Japanese vs. IHAC cohorts are similar to those used to investigate differential
248 escape between HLA subtype members (17), with some modifications as follows. Briefly, a
249 phylogenetically-corrected logistic regression model is constructed using a single HLA
250 allele as a predictor. Using a likelihood ratio test, we then compare this model to a more
251 expressive one that includes an additional interaction term that is 1 if the individual
252 expresses the HLA allele and is in the IHAC cohort, or 0 otherwise. In this way, we can
253 obtain a p-value testing the hypothesis that selection is the same in both cohorts (null
254 hypothesis) or whether selection differs across cohorts (alternative hypothesis). In contrast
255 to the HLA-AP analyses described thus far, the present one does not feature corrections for
256 HLA linkage disequilibrium or HIV codon covariation – and therefore will yield odds ratios
257 of association and p-values that differ slightly from the original cohort-specific values. In
258 the inter-cohort differential escape analysis, significance was defined as $p < 0.01$, $q < 0.05$.

JVI Accepts published online ahead of print

259 **Results**

260

261 **Identification of HLA-associated polymorphisms in chronically HIV-1 clade**
262 **B-infected Japanese individuals.**

263 The first objective of our study was to identify and characterize HLA-AP in Japan, a
264 unique population in terms of its HLA class I distribution and predominantly HIV clade B
265 epidemic. Towards this end, we analyzed linked HIV/HLA genotypes from 430
266 antiretroviral therapy-naive Japanese individuals chronically infected with HIV-1 clade B.
267 A total of 78 unique HLA class I alleles, defined at subtype-level (4-digit) resolution, were
268 observed in our cohort (**Fig. S1**) at frequencies consistent with the published literature (23).
269 Of these, 37 (including 9 HLA-A, 17 HLA-B, and 11 HLA-C alleles) were observed in at
270 least 10 individuals, and thus were included in the statistical analysis of HLA-APs (see
271 methods). Amplification and sequencing of HIV-1 Gag, Pol without the transframe (TF)
272 protein, and Nef was successful for 397 (92.3%), 363 (84.4%), and 306 (71.2%) individuals,
273 respectively. As described in the methods, HLA-APs within these three genes were
274 identified using a phylogenetically corrected logistic regression model which corrects for
275 the confounding effects of viral phylogeny, HIV-1 codon covariation and linkage
276 disequilibrium between host HLA class I alleles (16, 17, 20). A false discovery rate
277 (q-value) approach was employed to address multiple tests.

278 At a threshold of $q < 0.2$, a total of 284 HLA-APs, comprising 143 adapted and 141
279 non-adapted associations, were identified in Gag (N=94), Pol (N=86), and Nef (N=104
280 associations) (**Fig. 1 and Table S1**). HLA-APs were more frequently detected in Nef
281 (occurring at 45 of 206 codons; 21.8%) compared to Gag (51 of 500 of codons; 10.2%) or
282 Pol (51 of 947 codons; 5.1%). Although HLA class I allele frequencies in Japan are
283 somewhat distinct globally, the distribution of HLA-AP across HIV-1 proteins was
284 consistent with that reported in previous studies of other populations infected with clades B
285 or C (1, 2, 6, 7, 16). Broken down by HLA locus, the numbers of HLA-A-, HLA-B-, and

JVI Accepts published online ahead of print

286 HLA-C-associated polymorphisms were 78, 140, and 66, respectively, numbers that were
287 also consistent with previous reports from Caucasian and African cohorts that HLA-B
288 alleles restrict more associations than HLA-A or HLA-C alleles (1, 6, 18).
289

290 **Correlation between the total number of HLA-associated substitutions and clinical**
291 **parameters in Japanese individuals.**

292 We next wished to investigate the relationship between the presence of HLA-associated
293 substitutions in each gene and patient HIV-1 plasma viral load (pVL) and CD4+ T cell
294 count (CD4 count) in the Japanese cohort. As described in the methods, substitutions within
295 a given HIV-1 sequence were counted as "HLA-associated" if they had been identified as
296 being associated with any HLA class I allele in our study, regardless of the HLA alleles
297 expressed by the patient. For example, Gag-9S is a HLA-B*15:01-associated nonadapted
298 polymorphism (**Fig. 1 and Table S1**); as such, any amino acid other than "S" at codon 9
299 was counted as an HLA-associated substitution. Similarly, Gag-123G is an
300 HLA-C*01:02-associated adapted polymorphism (but no specific nonadapted forms,
301 restricted by C*01:02 or others, were identified at this position); as such, any sequence
302 harboring "G" at codon 123 was counted as having an HLA-associated substitution at this
303 site.

304 A weak yet statistically significant inverse correlation was observed between pVL and
305 the total number of HLA-associated substitutions in Pol (Spearman's R = −0.11; p = 0.04)
306 (**Fig. 2A**). However, no such correlations were observed for Gag (Spearman's R= −0.056, p
307 = 0.3) or Nef (Spearman's R= −0.029, p = 0.6) (**Fig. 2A**). Moreover, no significant
308 correlations were observed between the total number of HLA-associated substitutions in
309 any HIV protein, and CD4 count (**Fig. 2A**). Though the overall association is weak, results
310 raise the intriguing hypothesis that selection of certain HLA-driven substitutions in Pol
311 could modulate VL in the Japanese population.

312 We next wondered whether the observed correlation between Pol polymorphisms and

JVI Accepts published online ahead of print

313 lower pVL could be attributed to polymorphisms restricted by HLA alleles that are
314 protective in Japanese populations. HLA-B*67:01 and the HLA-B*52:01-HLA-C*12:02
315 haplotype are examples of such protective alleles (24). As such, we investigated whether
316 they could play a role in the observed pVL correlation. No HLA-B*67:01-associated
317 substitution was identified in Pol, whereas four HLA-B*52:01-associated and one
318 HLA-C*12:02-associated substitutions were detected in this protein (**Table S1**). Exclusion
319 of the single HLA-C*12:02-associated substitution from analysis did not affect the
320 relationship between the number of HLA-associated substitutions in Pol and pVL (**data not**
321 **shown**). In contrast, exclusion of the four HLA-B*52:01-associated Pol substitutions
322 substantially weakened the overall relationship between the number of HLA-associated Pol
323 substitutions and pVL (Spearman's $R = -0.057$; $p = 0.3$) (**Fig. 2B**). Similarly, specific
324 consideration of only HLA-B*52:01-associated Pol substitutions revealed a highly
325 significant inverse correlation with pVL (Spearman's $R = -0.18$; $p = 0.0007$) (**Fig. 2C**) that
326 represented the strongest such relationship detected in Pol for common HLA alleles
327 observed in our cohort (**Fig. S2**). We therefore reasoned that B*52:01-restricted
328 substitutions were likely to be critical mediators of the observed pVL effect.

329 Finally, stratification of B*52:01-associated Pol substitutions by host B*52:01
330 expression revealed that the inverse correlation with pVL remained strongly detectable in
331 HLA-B*52:01^{-} individuals (Spearman's $R = -0.18$; $p = 0.003$), but not in HLA-B*52:01^{+}
332 individuals (Spearman's $R = 0.026$; $p = 0.8$) (**Fig. 2D**). We interpret our observations to
333 suggest that HLA-B*52:01-restricted Pol substitutions possess fitness costs that manifest
334 themselves in terms of lower pVL upon transmission to, and persistence in, HLA-B*52:01^{-}
335 individuals. In contrast, no such pVL effects are detectable in B*52:01^{+} individuals, likely
336 because the fitness costs of these substitutions are outweighed by the advantages conferred
337 by immune escape.

338

339 **Differential escape between HLA subtypes in Japanese individuals.**

JVI Accepts published online ahead of print

340 Our final goal in characterizing HLA-AP in Japan was to investigate the extent of
341 differential escape between closely-related HLA subtypes. In particular, we hypothesized
342 that HLA subtype members differing with respect to amino acids located within in the
343 peptide-binding groove of the HLA molecule may differ with respect to the nature (or
344 binding affinity) of the specific HIV epitopes presented (25-28), and therefore that they
345 may exhibit differential escape pathways. In contrast, HLA subtype members that differ
346 with respect to amino acids located outside the peptide-binding groove may be more likely
347 to present the same epitopes (29-31), and therefore will generally exhibit less evidence for
348 differential escape between them. Of the 284 HLA-AP identified in our cohort, 128 were
349 restricted by HLA allele groups (A*02, A*26, B*15, B*40, C*03, C*08, and C*14)
350 containing two or more subtype members (**Table S1**). For five of these allele groups (A*02,
351 A*26, B*15, B*40, and C*08), subtype members differed by substitutions within the
352 peptide-binding groove (**Fig. S3**), supporting them as potential candidates for differential
353 HLA-AP selection. In contrast, members of the C*03 and C*14 subtypes differed by
354 substitutions outside the peptide-binding groove (**Fig. S3),** suggesting that their epitope
355 repertoire (and thus escape pathways) would be more similar to one another.

356 We began by simply comparing HLA-AP identified in context of the different HLA
357 subtypes. As expected, viral polymorphisms associated with HLA subtype members
358 differing within their peptide-binding grooves appeared to be quite specific to each HLA
359 subtype (**Fig. S3A, S3B, S3C, S3D, and S3F**). Surprisingly however, viral polymorphisms
360 associated with HLA subtype members differing only with respect to amino acids located
361 outside their peptide-binding grooves also appeared to be quite specific to each HLA
362 subtype (**Fig. S3E and S3G**). For example, HLA-C*03:03 and C*03:04, which differ only
363 by substitutions at position 91 that have no contact with the groove (29-31), were
364 associated with a total of 11 HLA-APs, none of which appeared to be shared (**Fig. S2E**).
365 Similarly, HLA-C*14:02 and C*14:03, that differ only by a substitution at position 21
366 located outside of the floor of the peptide-binding groove (**Fig. S2G**), shared only 10 of the

JVI Accepts published online ahead of print

367 24 HLA-APs identified between them.

368 However, qualitative comparisons of HLA-AP meeting a specific significance threshold,
369 such as those described above, are not statistically robust (since individual associations may
370 fail to meet the threshold and thus not be detected, or variations in allele frequency may
371 limit power to detect associations). Thus, to explicitly investigate whether the above
372 examples represent statistically significant instances of differential escape between subtype
373 members we applied a phylogenetically-corrected interaction test to compare their strengths
374 of selection between subtypes (17). For each HLA allele group, we took the union of all
375 HLA-AP identified for all subtype members, and compared their strength of selection
376 between all subtype members in a pairwise manner. Representative examples of our results
377 are shown in **Fig. 3**. For example, HLA-A*26:01 or A*26:03 differ with respect to
378 substitutions at amino acids 74, 76 and 77, located within the peptide-binding groove of the
379 of the HLA molecule (**Fig. S3B**). A total of 10 HLA-APs, located at 8 HIV codons, were
380 originally identified as associated with either HLA-A*26:01 or A*26:03 (**Fig. S3B**).
381 Although qualitatively, all 10 HLA-APs appear to be differentially selected by
382 HLA-A*26:01 or A*26:03 (**Fig. S3B**), the phylogenetically-corrected interaction test
383 revealed only 3 of them (located at Pol residues 276 and 551, and Nef residue 85) to be
384 significantly differentially selected in terms of their natural logarithm of the odds ratios of
385 association ($p < 0.05$, $q < 0.2$) (**Fig. 3A**). Surprisingly, significant differential escape was
386 also observed between subtype members that differed only with respect to substitutions
387 outside of their peptide-binding grooves: 3 of 9 (33.3%) sites restricted by HLA-C*03
388 allele group members and 5 of 14 (35.7%) sites restricted by C*14 allele group members
389 similarly exhibited statistically significant evidence of differential selection (**Fig. 3B and**
390 **3C**).

391 To compare whether the extent of differential escape between HLA subtype members
392 varied between HLA allele groups that differed with respect to substitutions within or
393 outside the binding groove, we asked whether the extent of differential escape between

JVI Accepts published online ahead of print

394 subtype members of the former group (comprising A*02, A*26, B*15, B*40, and C*08)
395 differed compared to the latter group (comprising HLA-C*03 and C*14). Overall, we found
396 no significant differences in the proportion of differential escape between them (34.8% for
397 HLA-C*03/C*14 subtypes compared to 36.8% for subtypes of all other HLA alleles, p =
398 0.5) (**Table S2**). This intriguing result suggests that variations outside the HLA binding
399 groove may contribute as much to differential escape as variations within the binding
400 groove.
401
402 **Comparison of HLA-APs between Japanese and non-Asian individuals chronically**
403 **infected with HIV-1 clade B.**
404 Our second objective was to investigate HLA-AP identified in Japan versus those
405 previously identified in non-Asian cohorts infected with HIV clade B. The comparison
406 cohort in this analysis was the International HIV Adaptation Collaborative (IHAC) cohort,
407 comprising 1,888 antiretroviral-naïve individuals with chronic clade B infection in Canada,
408 the USA, and Australia (of which <5% of cohort participants are Asian) (16).
409 HLA-AP will differ to some extent between human populations due to the presence (or
410 enrichment) of certain HLA alleles in one population versus another. Indeed, HLA allele
411 frequencies differed markedly between the Japan and IHAC cohorts (**Fig. S1**). As such, we
412 begin with a qualitative comparison of HLA-AP between them. We begin with a simple
413 positional analysis. In the Japanese cohort, HLA-APs were observed at a total of 147 codon
414 positions in Gag, Pol, and Nef (**Fig. 4**). Of these, 117 (79.6%) were also associated with at
415 least one HLA allele in IHAC. In contrast, the remaining 30 positions (including 16, 7 and
416 7 in Gag, Pol, and Nef, respectively) that harbored HLA associations in Japan were not
417 associated with any HLA alleles in IHAC (**Fig. 4**). That 30/147 (20.4%) of HIV codons
418 exhibited evidence of HLA-driven selection in Japan but not IHAC already strongly
419 suggests that HIV is evolving under population-specific selection pressures in Japan
420 compared to other regions.

JVI Accepts published online ahead of print

421 Next we compared HLA-AP over HIV position and specific HLA restriction. Of the 284
422 HLA-APs identified in Japan, 188 (66.2%) were not reported in IHAC. As expected, a
423 substantial portion of these (46 of 188, 24.5%) were associated with 8 HLA subtypes
424 (A*26:03, B*40:06, B*54:01, B*55:02, B*59:01, B*67:01, C*08:03, and C*14:03)
425 common in Japan but essentially absent (<1% frequency) in IHAC. Others were likely
426 attributable to alleles observed at much higher frequencies in Japan compared to IHAC: for
427 example, an additional 27.1% were associated with HLA alleles present in both cohorts, but
428 whose frequencies were at least fourfold higher in Japan compared to IHAC. Overall,
429 results suggest that HLA-APs identified in Japan are quite distinctive, in large part
430 reflecting the unique HLA allele distribution in the Japanese population.

431 We also wished to investigate the existence of differential HLA-associated escape
432 pathways between the two populations, that are not attributable to HLA frequency
433 differences between them – in other words, cases where the same HLA subtypes drive
434 significantly different escape pathways in Japan versus IHAC cohorts. This required the
435 application of statistical tests (see methods and below). Specifically, we first identified a list
436 of N=551 HLA-AP in HIV Gag, Pol and Nef, which represented the union of all HLA-AP
437 identified in either Japan or IHAC for which both the viral polymorphism and the
438 restricting HLA allele were observed in a minimum of 10 individuals per cohort (not
439 shown). The latter criteria were employed in order to achieve some minimal statistical
440 power to compare strengths of individual associations between cohorts. It is important to
441 emphasize that these criteria would by definition exclude HLA alleles (and/or viral
442 polymorphisms) present in one cohort but essentially absent in the other (as we would have
443 no power, and in fact no rationale, to test whether their strengths of selection were
444 statistically significantly different between cohorts).

445 For each HLA-AP, we calculated its natural logarithm of the odds ratio (lnOR) of
446 association in each cohort – a measure that can be interpreted as an estimate of the strength
447 of selection exerted by the HLA allele on that particular HIV codon, in that cohort. We then

JVI Accepts published online ahead of print

448 applied a phylogenetically-corrected interaction test (17) to assess whether these lnORs of
449 selection were significantly different in the Japanese versus the IHAC cohorts. In these
450 analyses, statistical significance was defined as $p < 0.01$ and $q < 0.05$.
451 Overall, 71 (of 551, 12.8%) HLA-APs originally identified in either Japan or IHAC
452 cohorts exhibited significantly different strengths of selection between the two populations
453 (**Figure 5** and **Table S3**). The HLA-B*44:03-associated 125H substitution in Nef serves as
454 an example of how to interpret these data. The lnOR of this association is 1.73 in Japan
455 (with a cohort-specific p-value of 3.26×10^{-6}) versus 0.42 for IHAC (with a cohort-specific
456 p-value of 0.36). Both lnORs are positive, indicating that 125H is positively associated with
457 B*44:03 in both cohorts, but the higher lnOR in Japan indicates that the strength of
458 selection of Nef-125H by B*44:03 is greater in Japan compared to IHAC (indeed, the
459 cohort-specific p-values reveal that this association is significant in Japan but not IHAC).
460 Finally, the p- and q-values for the intercohort comparison ($p = 1.02 \times 10^{-6}$ and $q = 1.19 \times$
461 10^{-4}; **Table S3**) confirm that the strength of selection of Nef-125H by B*44:03 is
462 significantly greater in Japan compared to IHAC. Importantly, this difference is not simply
463 attributable to intercohort differences in B*44:03 frequency (which is comparable between
464 populations; **Fig S1**).
465 In addition to the HLA-B*44:03-associated 125H polymorphism in Nef, we identified
466 21 other HLA-AP whose strengths of selection were significantly greater in Japan
467 compared to IHAC, yielding a total of 22 (of 71; 31.0%) HLA-APs in this category.
468 Conversely, 39 (of 71; 54.9%) differentially-selected HLA-AP exhibited strengths of
469 selection that were greater in IHAC compared to Japan. The HLA-A*26:01-associated
470 889S substitution in Pol serves as an example. The lnOR of this association is −0.18 in
471 Japan (with a cohort-specific p-value of 0.3) versus −1.17 for IHAC (with a cohort-specific
472 p-value of 7.92×10^{-9}). Both lnORs are negative, indicating that 889S is negatively
473 associated with A*26:01 in both cohorts, but the more negative value for IHAC indicates
474 that this association is stronger in IHAC compared to Japan. Finally, the p- and q-values for

JVI Accepts published online ahead of print

475 the intercohort comparison (p = 1.15 × 10^{-4} and q = 4.48 × 10^{-3}; Table S3) confirm that of
476 the strength of the negative association between Pol-889S by A*26:01 is significantly
477 greater in IHAC compared to Japan.

478 Strikingly, the remaining 10 (of 71; 14.1%) differentially-selected HLA-APs displayed
479 diametrically opposed directions of selection between the cohorts (defined here as lnORs of
480 association that were positive in one cohort but negative in the other, where the
481 cohort-specific p-values were <0.05 in both cases) (**Fig. 5**). The HLA-B*44:03-associated
482 120F substitution in Nef serves as an example. The lnOR of this association is 1.44 in Japan
483 (with a cohort-specific p-value of 2.03 × 10^{-4}), indicating that HLA-B*44:03 is
484 significantly positively associated with 120F in Japan. In contrast, the lnOR of this
485 association is −0.69 in IHAC (with a cohort-specific p-value of 9.50 × 10^{-3}), indicating that
486 HLA-B*44:03 is significantly negatively associated with 120F in Japan. The p- and
487 q-values for the intercohort comparison (p = 2.15 × 10^{-8} and q = 3.75 × 10^{-6}; **Table S3**)
488 confirm that the opposing direction of selection of Nef-120F by B*44:03 between Japanese
489 and IHAC cohorts is a statistically significant observation.

490 Of interest, the 71 HLA-APs identified as being under significantly different selection
491 between Japan and IHAC cohorts were differentially distributed across HLA loci and HIV
492 proteins (**Fig. 6A and 6B**). Specifically, HLA-A-associated polymorphisms that were
493 significantly differentially selected across cohorts were most abundant in Gag, followed by
494 Pol and Nef, whereas differentially-selected HLA-B-associated and HLA-C-associated
495 polymorphisms were most numerous in Nef, followed by Pol and Gag. Taken together,
496 results support the existence of HLA class I alleles that drive significantly different HIV
497 escape pathways in global populations infected with the same viral clade. The uneven
498 distribution of the locations of these differentially-selected polymorphisms across HLA loci
499 and HIV regions raise the intriguing hypothesis that Gag and Pol/Nef may differentially
500 evolve under selection pressures dominated by HLA-A versus HLA-B/C allele-restricted
501 immune responses, respectively.

JVI Accepts published online ahead of print

502 **Discussion**

503

504 The present study comprised two major objectives, both of which are novel in terms of
505 populations studied and/or analytical methods used. First, we characterized HLA-AP in
506 HIV-1 clade B Gag, Pol and Nef and their relationship with clinical parameters in a large
507 Japanese cohort. Second, we compared HLA-AP in Japanese versus non-Asian populations
508 infected with HIV clade B, to identify population-specific differences in their selection. In
509 particular, we wished to identify HLA-AP that are unique to Japan by virtue of the
510 distinctive HLA distribution in this population, as well as cases where the same HLA allele
511 drives divergent escape pathways in Japan vs. non-Asian populations.
512 This study is the first to identify HLA-APs in HIV-1's structural and functional genes in
513 Japanese populations. Only one previous study investigated HLA-AP in HIV-1 clade
514 B-infected Asians (11): this study comprised 231 Chinese individuals infected during a
515 narrow-source outbreak, and identified 141 HLA-associated polymorphisms at two-digit
516 resolution. Our study differs from the latter with respect to cohort size, HLA genetics of the
517 host population, HLA typing resolution and epidemiology of the epidemic. Using
518 phylogenetically-informed approaches, we identified 284 HLA-APs within HIV-1 Gag, Pol
519 and Nef in our cohort, supporting a strong influence of population-specific, HLA-driven
520 immune pressures in shaping HIV-1 evolution in Japan. In contrast to a previous study
521 undertaken in a predominantly Caucasian population that observed approximately one-half
522 of the total number of Gag HLA-APs to be located within or flanking reported CTL
523 epitopes (3), the majority of HLA-APs identified in the present study were not located
524 nearby reported CTL epitopes. This discrepancy may be due to the limited number of
525 Asian-specific HLA-restricted CTL epitopes identified to date, underscoring the need for
526 further epitope discovery in these populations.
527 This study revealed differential frequencies of HLA-APs across HIV genes in the
528 Japanese population. Consistent with previous studies of HLA-AP in HIV clade B (2, 16,

JVI Accepts published online ahead of print

529 18), HLA-APs were more frequently detected in Nef than in Gag and Pol. Also consistent
530 with previous observations in Caucasian, African, Chinese, and Mexican populations (1, 6,
531 11, 15, 18), the number of HLA-B-associated polymorphisms in our cohort was higher than
532 that of HLA-A- or HLA-C-associated ones, further supporting a dominant role of HLA-B
533 in HIV evolution (32). An interesting feature of the Japanese population is that
534 approximately 70% of individuals carry HLA-A*24:02 (23). Despite sufficient statistical
535 power to detect HLA-A*24:02-associated polymorphisms in our cohort, we identified only
536 9 of these, 6 of which were located in epitopes identified by our group (33-35). A possible
537 explanation for the relatively low number of A*24:02-associated polymorphisms in Japan is
538 that they have accumulated over time in circulating sequences such that they are no longer
539 significantly enriched among persons expressing HLA-A*24:02. Further analysis of
540 mutations selected by HLA-A*24:02-restricted CTLs should clarify the mechanism
541 whereby high frequency HLA alleles influence the formation of HIV-1 polymorphisms.
542 Protective HLA alleles such as HLA-B*57, B*58, and B*27, select Gag mutations
543 affecting viral replication in Caucasians and Africans (36-41) that may also provide some
544 clinical benefit if they are transmitted to hosts lacking these alleles (42, 43). HLA-B*57,
545 B*58, and B*27 are not present at appreciable frequencies in Japan (23). It is therefore
546 perhaps unsurprising that no correlations between HLA-associated substitutions in Gag and
547 HIV clinical parameters were observed in our cohort. In contrast, we observed a weak but
548 significant inverse correlation between the frequency of HLA-APs in Pol and plasma viral
549 load, which appeared to be driven by polymorphisms selected by HLA-B*52:01, an allele
550 identified as protective in Japan (24). Upon further stratification by HLA-B*52:01
551 expression, the inverse correlation between VL and the total number of B*52:01-associated
552 Pol substitutions was maintained in HLA-B*52:01^{-}, but not in HLA-B*52:01^{+} individuals.
553 Taken together, these findings suggest that transmitted B*52:01-associated polymorphisms
554 could reduce viral fitness in a dose-dependent manner, though further studies would be
555 required to assess this. In addition, these substitutions were not located within or nearby

JVI Accepts published online ahead of print

556 known B*52:01-restricted epitopes. Thus, further research would be required to identify
557 these epitopes and elucidate their mechanisms of escape.

558 Many previous studies of HLA-APs were performed at 2-digit HLA resolution (1-4, 6).
559 Here, we performed HLA genotyping at 4-digit resolution, which allowed us to investigate
560 differential escape between closely-related HLA subtypes in the Japanese cohort. Nearly
561 one half of the HLA-AP identified in Japan were restricted by HLA allele groups
562 containing two or more subtype members (A*02, A*26, B*15, B*40, C*03, C*08, and
563 C*14). For five of these groups (A*02, A*26, B*15, B*40, and C*08), subtype members
564 differed by substitutions within the peptide-binding groove, while for the remaining two
565 groups (HLA-C*03 and -C*14), subtype members differed by substitutions located outside
566 the peptide binding groove. Reasoning that amino acid differences located within the
567 peptide-binding groove could modulate the nature or presentation of CTL epitopes, we
568 hypothesized that the former group would generally exhibit distinct HLA-AP between
569 subtype members, while the latter would generally exhibit similar or identical HLA-AP.
570 However, we were surprised to observe substantial evidence for differential HLA-AP
571 selection between closely-related HLA subtypes regardless of whether they differed in
572 sequence within or outside the peptide binding groove. Significantly differential HLA-AP
573 selection was observed at 3 of 9 HLA-C*03 associated sites and 5 of 14 HLA-C*14
574 associated sites (**Fig. 3**), proportions that were not significantly lower than the frequency of
575 differential selection between subtypes that differed in their peptide-binding groove.

576 This observation raised several hypotheses. HLA polymorphic sites outside the
577 peptide-binding groove may indirectly influence the binding groove conformation, thus
578 altering HLA-peptide interactions and/or T cell recognition. Another possibility is selection
579 by NK cells, as KIR may recognize sites outside the peptide-binding groove. Indeed,
580 KIR3DL1 bind to the loop including position 91 of HLA-B*57:01 (44). However it is not
581 clear whether KIR2DLs, which are receptors for HLA-C, can bind to the loop outside the
582 peptide-binding groove of HLA-C molecules. A recent study showed that HLA-C antigens

JVI Accepts published online ahead of print

583 are expressed at different levels on the cell surface, even among HLA-C subtypes (45).
584 This study also observed a strong positive correlation between HLA-C expression level and
585 the strength of HLA-C-mediated selection pressure conferred on HIV. Differential
586 expression levels of these HLA-C subtype members in Japanese populations thus provide
587 another potential explanation for this observation, for future follow-up.

588 Our second objective was to investigate differential HLA-AP between Japanese and
589 non-Asian cohorts infected with HIV clade B. Here, the IHAC cohort (comprising clade
590 B-infected Canadians, Americans and Australians) was used as a comparison group (16).
591 HLA-AP identified in human populations will differ to some extent due to
592 population-specific HLA distributions, yielding population-specific HLA-AP driven by
593 HLA alleles present in one population but not another (15). Indeed, two-thirds of the
594 HLA-APs identified in Japan had not previously been identified due to the presence of the
595 restricting HLA alleles in Japan, but its absence (or far lower prevalence) in IHAC.

596 What remains unknown however, is the extent to which the same HLA allele may drive
597 significantly different escape pathways in different human populations. To this end, we
598 applied novel phylogenetically-corrected statistical approaches to assess the extent to which
599 HLA-AP identified in either Japan or IHAC, that were restricted by HLA alleles present in
600 both populations, exhibited significantly different strengths of selection between them. Of
601 the 551 HLA-AP investigated, 71 (12.9%) were significantly differentially selected in
602 Japan versus IHAC at a stringent statistical threshold of $q < 0.05$. Of these 71, 31%
603 exhibited significantly greater strengths of selection in Japan compared to IHAC whereas
604 55% exhibited greater strengths of selection in IHAC compared to Japan. Surprisingly, the
605 remaining 14% displayed diametrically opposed selection pathways in the two cohorts
606 (where an HIV polymorphism represented the “adapted” form associated with a given allele
607 in one cohort, but the “nonadapted” form associated with the same allele in the other
608 cohort). It is important to emphasize that these significantly different pathways of HLA-AP
609 selection are not simply attributable to differences in HLA frequency between the cohorts.

JVI Accepts published online ahead of print

610 We feel that these are intriguing observations that merit further study. Nevertheless we
611 propose the following potential interpretations. Firstly, these differences could be explained
612 by functional differences in HIV-1 specific T cells elicited between Japanese and Caucasian
613 cohorts, possibly as a result of differences in host genetics (for example in the genes that
614 encode the T-cell receptor and/or modulate their expression). Such differences may
615 influence the structure of the T-cell receptor(s) and thus the quality, quantity and/or makeup
616 of the HIV-1 specific T cell repertoire, thus influencing the specific escape mutations
617 selected in context of peptide-bound HLA. Further analysis of HIV-1 specific T cells
618 driving the selection of these mutants in both cohorts is therefore warranted. It is also
619 important to note that the intercohort HLA-AP comparisons, unlike previous analyses, did
620 not correct for HLA linkage disequilibrium (LD) or HIV codon covariation. Although both
621 Japan and IHAC cohorts feature HIV clade B infections, intra-clade differences in the viral
622 backbone could also influence differential escape via epistatic effects. In-depth analyses of
623 intercohort differences in HIV codon covariation relationships are also therefore warranted.
624 Intercohort differences in HLA linkage disequilibrium are another possible contributor.
625 Finally, the differentially-selected HLA-AP between cohorts appeared to be unevenly
626 distributed by HLA locus: while HLA-A associated polymorphisms exhibiting differential
627 selection between cohorts were more abundant in Gag compared to other proteins, HLA-B
628 and HLA-C associated polymorphisms exhibiting differential selection between cohorts
629 tended to be more abundant in Nef. This suggests that inter-cohort differential HLA-APs
630 across HIV proteins may be arising as a result of cellular immune pressures exerted by
631 distinct HLA class I loci, though this also requires further study.

632 Nevertheless, the present study confirms of the existence of population-specific HIV-1
633 adaptations that are attributable to the unique HLA allele distributions of that population
634 (15). We additionally provide evidence of population-specific HIV adaptation to
635 HLA-restricted immune responses that cannot be explained by differential HLA frequencies
636 alone: cases where the same HLA allele drives significantly different, sometimes opposing,

637 escape pathways in different host populations. Taken together, results support differential
638 HIV-1 adaptation to human populations worldwide, that might be driven by multiple host
639 and viral mechanisms.

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

640 **Acknowledgements**

641

642 The authors thank Dr. Madoka Koyanagi and Rie Maruyama for collection of samples
643 from the patients, Mari Hasegawa and Sayaka Nagata for technical assistance, Sachiko
644 Sakai for her secretarial assistance, and Kyle Cobarrubias for assistance in constructing
645 Figure 1.

646 TC is a JSPS Research Fellow. AQL is the recipient of a Frederick Banting and Charles
647 Best Masters award from the Canadian Institutes for Health Research (CIHR). ZLB is the
648 recipient of a New Investigator Award from the CIHR and a Scholar Award from the
649 Michael Smith Foundation for Health Research (MSFHR). This research was supported by
650 the Global COE program "Global Education and Research Center Aiming at the control of
651 AIDS," launched as a project commissioned by the Ministry of Education, Science, Sports,
652 and Culture, Japan; and by Grants-in Aid for AIDS Research from the Ministry of Health,
653 Labour, and Welfare, Japan.

654

655 **Disclosure**

656

657 The authors have no financial conflicts of interest.

JVI Accepts published online ahead of print

658 **References**

659

660 1. **Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson J,**
661 **Kadie C, Bhattacharya T, Chui C, Szinger J, Mo T, Hogg RS, Montaner JS,**
662 **Frahm N, Brander C, Walker BD, Harrigan PR.** 2007. Evidence of differential
663 HLA class I-mediated viral evolution in functional and accessory/regulatory genes
664 of HIV-1. PLoS Pathog. **3:**e94.
665 2. **Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA,**
666 **Swenson LC, Tao I, Szeto S, Rosato P, Sela J, Kadie CM, Frahm N, Brander C,**
667 **Haas DW, Riddler SA, Haubrich R, Walker BD, Harrigan PR, Heckerman D,**
668 **Mallal S.** 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag,
669 Pol and Nef proteins. PLoS One **4:**e6687.
670 3. **Brumme ZL, Tao I, Szeto S, Brumme CJ, Carlson JM, Chan D, Kadie C,**
671 **Frahm N, Brander C, Walker B, Heckerman D, Harrigan PR.** 2008. Human
672 leukocyte antigen-specific polymorphisms in HIV-1 Gag and their association with
673 viral load in chronic untreated infection. AIDS **22:**1277-1286.
674 4. **Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA.** 2002.
675 Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population
676 level. Science **296:**1439-1443.
677 5. **Rolland M, Carlson JM, Manocheewa S, Swain JV, Lanxon-Cookson E, Deng**
678 **W, Rousseau CM, Raugi DN, Learn GH, Maust BS, Coovadia H, Ndung'u T,**
679 **Goulder PJ, Walker BD, Brander C, Heckerman DE, Mullins JI.** 2010.
680 Amino-acid co-variation in HIV-1 Gag subtype C: HLA-mediated selection
681 pressure and compensatory dynamics. PLoS One **5**.
682 6. **Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, Prendergast A,**
683 **Matthews P, Payne R, Rolland M, Raugi DN, Maust BS, Learn GH, Nickle DC,**
684 **Coovadia H, Ndung'u T, Frahm N, Brander C, Walker BD, Goulder PJ,**

JVI Accepts published online ahead of print

685 **Bhattacharya T, Heckerman DE, Korber BT, Mullins JI.** 2008. HLA class
686 I-driven evolution of human immunodeficiency virus type 1 subtype c proteome:
687 immune escape and viral load. J. Virol. **82:**6434-6446.
688 7. **Wang YE, Li B, Carlson JM, Streeck H, Gladden AD, Goodman R,**
689 **Schneidewind A, Power KA, Toth I, Frahm N, Alter G, Brander C, Carrington**
690 **M, Walker BD, Altfeld M, Heckerman D, Allen TM.** 2009. Protective HLA class
691 I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral
692 escape mutations located in highly conserved regions of human immunodeficiency
693 virus type 1. J. Virol. **83:**1845-1855.
694 8. **Huang KH, Goedhals D, Carlson JM, Brockman MA, Mishra S, Brumme ZL,**
695 **Hickling S, Tang CS, Miura T, Seebregts C, Heckerman D, Ndung'u T, Walker**
696 **B, Klenerman P, Steyn D, Goulder P, Phillips R, Bloemfontein-Oxford**
697 **Collaborative G, van Vuuren C, Frater J.** 2011. Progression to AIDS in South
698 Africa is associated with both reverting and compensatory viral mutations. PLoS
699 One **6:**e19018.
700 9. **Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M,**
701 **Gatanaga H, Fujiwara M, Hachiya A, Koizumi H, Kuse N, Oka S, Duda A,**
702 **Prendergast A, Crawford H, Leslie A, Brumme Z, Brumme C, Allen T,**
703 **Brander C, Kaslow R, Tang J, Hunter E, Allen S, Mulenga J, Branch S, Roach**
704 **T, John M, Mallal S, Ogwu A, Shapiro R, Prado JG, Fidler S, Weber J, Pybus**
705 **OG, Klenerman P, Ndung'u T, Phillips R, Heckerman D, Harrigan PR,**
706 **Walker BD, Takiguchi M, Goulder P.** 2009. Adaptation of HIV-1 to human
707 leukocyte antigen class I. Nature **458:**641-645.
708 10. **Solberg OD, Mack SJ, Lancaster AK, Single RM, Tsai Y, Sanchez-Mazas A,**
709 **Thomson G.** 2008. Balancing selection and heterogeneity across the classical
710 human leukocyte antigen loci: a meta-analytic review of 497 population studies.
711 Hum. Immunol. **69:**443-464.

JVI Accepts published online ahead of print

712 11. **Dong T, Zhang Y, Xu KY, Yan H, James I, Peng Y, Blais ME, Gaudieri S,**
713 **Chen X, Lun W, Wu H, Qu WY, Rostron T, Li N, Mao Y, Mallal S, Xu X,**
714 **McMichael A, John M, Rowland-Jones SL.** 2011. Extensive HLA-driven viral
715 diversity following a narrow-source HIV-1 outbreak in rural China. Blood
716 **118:**98-106.
717 12. **Lau KA, Wang B, Saksena NK.** 2007. Emerging trends of HIV epidemiology in
718 Asia. AIDS reviews **9:**218-229.
719 13. **Rodrigo C, Rajapakse S.** 2009. Current Status of HIV/AIDS in South Asia. J.
720 Glob. Infect. Dis. **1:**93-101.
721 14. **Ruxrungtham K, Brown T, Phanuphak P.** 2004. HIV/AIDS in Asia. Lancet
722 **364:**69-82.
723 15. **Avila-Rios S, Ormsby CE, Carlson JM, Valenzuela-Ponce H, Blanco-Heredia J,**
724 **Garrido-Rodriguez D, Garcia-Morales C, Heckerman D, Brumme ZL, Mallal**
725 **S, John M, Espinosa E, Reyes-Teran G.** 2009. Unique features of HLA-mediated
726 HIV evolution in a Mexican cohort: a comparative study. Retrovirology **6:**72.
727 16. **Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ,**
728 **Chui CK, Cotton LA, Knapp DJ, Riddler SA, Haubrich R, Nelson G, Pfeifer N,**
729 **Deziel CE, Heckerman D, Apps R, Carrington M, Mallal S, Harrigan PR, John**
730 **M, Brumme ZL, International HIVAC.** 2012. Correlates of protective cellular
731 immunity revealed by analysis of population-level immune escape pathways in
732 HIV-1. J. Virol. **86:**13202-13216.
733 17. **Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker BD, Ndung'u T,**
734 **Shapiro R, Frater J, Brumme ZL, Goulder PJ, Heckerman D.** 2012.
735 Widespread impact of HLA restriction on immune control and escape pathways of
736 HIV-1. J. Virol. **86:**5230-5243.
737 18. **John M, Heckerman D, James I, Park LP, Carlson JM, Chopra A, Gaudieri S,**
738 **Nolan D, Haas DW, Riddler SA, Haubrich R, Mallal S.** 2010. Adaptive

JVI Accepts published online ahead of print

739 interactions between HLA and HIV-1: highly divergent selection imposed by HLA
740 class I molecules with common supertype motifs. J. Immunol. **184:**4368-4377.
741 19. **Leslie A, Price DA, Mkhize P, Bishop K, Rathod A, Day C, Crawford H,**
742 **Honeyborne I, Asher TE, Luzzi G, Edwards A, Rousseau CM, Mullins JI,**
743 **Tudor-Williams G, Novelli V, Brander C, Douek DC, Kiepiela P, Walker BD,**
744 **Goulder PJ.** 2006. Differential selection pressure exerted on HIV by CTL targeting
745 identical epitopes but restricted by distinct HLA alleles from the same HLA
746 supertype. J. Immunol. **177:**4699-4708.
747 20. **Carlson JM, Brumme ZL, Rousseau CM, Brumme CJ, Matthews P, Kadie C,**
748 **Mullins JI, Walker BD, Harrigan PR, Goulder PJ, Heckerman D.** 2008.
749 Phylogenetic dependency networks: inferring patterns of CTL escape and codon
750 covariation in HIV-1 Gag. PLoS Comput. Biol. **4:**e1000225.
751 21. **Carlson J, Kadie C, Mallal S, Heckerman D.** 2007. Leveraging hierarchical
752 population structure in discrete association studies. PLoS One **2:**e591.
753 22. **Storey JD, Tibshirani R.** 2003. Statistical significance for genomewide studies.
754 Proc. Natl. Acad. Sci. U. S. A. **100:**9440-9445.
755 23. **Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, Kashiwase K, Kikkawa E,**
756 **Kulski JK, Satake M, Inoko H.** 2005. High-throughput DNA typing of HLA-A,
757 -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese
758 population. Immunogenetics **57:**717-729.
759 24. **Naruto T, Gatanaga H, Nelson G, Sakai K, Carrington M, Oka S, Takiguchi M.**
760 2012. HLA class I-mediated control of HIV-1 in the Japanese population, in which
761 the protective HLA-B*57 and HLA-B*27 alleles are absent. J. Virol.
762 **86:**10870-10872.
763 25. **Sette A, Sidney J.** 1999. Nine major HLA class I supertypes account for the vast
764 preponderance of HLA-A and -B polymorphism. Immunogenetics **50:**201-212.
765 26. **Reche PA, Reinherz EL.** 2007. Definition of MHC supertypes through clustering

JVI Accepts published online ahead of print

766 of MHC peptide-binding repertoires. Methods Mol. Biol. **409:**163-173.

767 27. **Sidney J, Southwood S, Sette A.** 2005. Classification of A1- and A24-supertype
768 molecules by analysis of their MHC-peptide binding repertoires. Immunogenetics
769 **57:**393-408.

770 28. **Sidney J, Peters B, Frahm N, Brander C, Sette A.** 2008. HLA class I supertypes:
771 a revised and updated classification. BMC Immunol. **9:**1.

772 29. **Kostyu DD, Hannick LI, Traweek JL, Ghanayem M, Heilpern D, Dawson DV.**
773 1997. HLA class I polymorphism: structure and function and still questions. Hum.
774 Immunol. **57:**1-18.

775 30. **Boyington JC, Motyka SA, Schuck P, Brooks AG, Sun PD.** 2000. Crystal
776 structure of an NK cell immunoglobulin-like receptor in complex with its class I
777 MHC ligand. Nature **405:**537-543.

778 31. **Fan QR, Wiley DC.** 1999. Structure of human histocompatibility leukocyte antigen
779 (HLA)-Cw4, a ligand for the KIR2D natural killer cell inhibitory receptor. J. Exp.
780 Med. **190:**113-123.

781 32. **Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S,**
782 **Rathnavalu P, Moore C, Pfafferott KJ, Hilton L, Zimbwa P, Moore S, Allen T,**
783 **Brander C, Addo MM, Altfeld M, James I, Mallal S, Bunce M, Barber LD,**
784 **Szinger J, Day C, Klenerman P, Mullins J, Korber B, Coovadia HM, Walker**
785 **BD, Goulder PJ.** 2004. Dominant influence of HLA-B in mediating the potential
786 co-evolution of HIV and HLA. Nature **432:**769-775.

787 33. **Fujiwara M, Tanuma J, Koizumi H, Kawashima Y, Honda K,**
788 **Mastuoka-Aizawa S, Dohki S, Oka S, Takiguchi M.** 2008. Different abilities of
789 escape mutant-specific cytotoxic T cells to suppress replication of escape mutant
790 and wild-type human immunodeficiency virus type 1 in new hosts. J. Virol.
791 **82:**138-147.

792 34. **Koizumi H, Iwatani T, Tanuma J, Fujiwara M, Izumi T, Oka S, Takiguchi M.**

JVI Accepts published online ahead of print

793 2009. Escape mutation selected by Gag28-36-specific cytotoxic T cells in
794 HLA-A*2402-positive HIV-1-infected donors. Microbes and infection / Institut
795 Pasteur **11:**198-204.
796 35. **Ikeda-Moore Y, Tomiyama H, Miwa K, Oka S, Iwamoto A, Kaneko Y,**
797 **Takiguchi M.** 1997. Identification and characterization of multiple
798 HLA-A24-restricted HIV-1 CTL epitopes: strong epitopes are derived from V
799 regions of HIV-1. J. Immunol. **159:**6242-6252.
800 36. **Altfeld M, Kalife ET, Qi Y, Streeck H, Lichterfeld M, Johnston MN, Burgett N,**
801 **Swartz ME, Yang A, Alter G, Yu XG, Meier A, Rockstroh JK, Allen TM,**
802 **Jessen H, Rosenberg ES, Carrington M, Walker BD.** 2006. HLA Alleles
803 Associated with Delayed Progression to AIDS Contribute Strongly to the Initial
804 CD8(+) T Cell Response against HIV-1. PLoS Med. **3:**e403.
805 37. **Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, Desouza I,**
806 **Ryvkin F, Derdeyn CA, Allen S, Hunter E, Mulenga J, Goepfert PA, Walker**
807 **BD, Allen TM.** 2007. Escape and compensation from early HLA-B57-mediated
808 cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag
809 alter capsid interactions with cyclophilin A. J. Virol. **81:**12608-12618.
810 38. **Gao X, Bashirova A, Iversen AK, Phair J, Goedert JJ, Buchbinder S, Hoots K,**
811 **Vlahov D, Altfeld M, O'Brien SJ, Carrington M.** 2005. AIDS restriction HLA
812 allotypes target distinct intervals of HIV-1 pathogenesis. Nat. Med. **11:**1290-1292.
813 39. **Kelleher AD, Long C, Holmes EC, Allen RL, Wilson J, Conlon C, Workman C,**
814 **Shaunak S, Olson K, Goulder P, Brander C, Ogg G, Sullivan JS, Dyer W,**
815 **Jones I, McMichael AJ, Rowland-Jones S, Phillips RE.** 2001. Clustered
816 mutations in HIV-1 gag are consistently required for escape from
817 HLA-B27-restricted cytotoxic T lymphocyte responses. J. Exp. Med. **193:**375-386.
818 40. **Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, Feeney M, Tang Y,**
819 **Holmes EC, Allen T, Prado JG, Altfeld M, Brander C, Dixon C, Ramduth D,**

JVI Accepts published online ahead of print

820 **Jeena P, Thomas SA, St John A, Roach TA, Kupfer B, Luzzi G, Edwards A,**
821 **Taylor G, Lyall H, Tudor-Williams G, Novelli V, Martinez-Picado J, Kiepiela**
822 **P, Walker BD, Goulder PJ.** 2004. HIV evolution: CTL escape mutation and
823 reversion after transmission. Nat. Med. **10:**282-289.
824 41. **Schneidewind A, Brockman MA, Yang R, Adam RI, Li B, Le Gall S, Rinaldo**
825 **CR, Craggs SL, Allgaier RL, Power KA, Kuntzen T, Tung CS, LaBute MX,**
826 **Mueller SM, Harrer T, McMichael AJ, Goulder PJ, Aiken C, Brander C,**
827 **Kelleher AD, Allen TM.** 2007. Escape from the dominant HLA-B27-restricted
828 cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in
829 human immunodeficiency virus type 1 replication. J. Virol. **81:**12382-12393.
830 42. **Chopera DR, Woodman Z, Mlisana K, Mlotshwa M, Martin DP, Seoighe C,**
831 **Treurnicht F, de Rosa DA, Hide W, Karim SA, Gray CM, Williamson C, Team**
832 **CS.** 2008. Transmission of HIV-1 CTL escape variants provides HLA-mismatched
833 recipients with a survival advantage. PLoS Pathog. **4:**e1000033.
834 43. **Goepfert PA, Lumm W, Farmer P, Matthews P, Prendergast A, Carlson JM,**
835 **Derdeyn CA, Tang J, Kaslow RA, Bansal A, Yusim K, Heckerman D, Mulenga**
836 **J, Allen S, Goulder PJ, Hunter E.** 2008. Transmission of HIV-1 Gag immune
837 escape mutations is associated with reduced viral load in linked recipients. J. Exp.
838 Med. **205:**1009-1017.
839 44. **Vivian JP, Duncan RC, Berry R, O'Connor GM, Reid HH, Beddoe T, Gras S,**
840 **Saunders PM, Olshina MA, Widjaja JM, Harpur CM, Lin J, Maloveste SM,**
841 **Price DA, Lafont BA, McVicar DW, Clements CS, Brooks AG, Rossjohn J.**
842 2011. Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of
843 human leukocyte antigen B. Nature **479:**401-405.
844 45. **Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, Yuki Y, Del Prete GQ,**
845 **Goulder P, Brumme ZL, Brumme CJ, John M, Mallal S, Nelson G, Bosch R,**
846 **Heckerman D, Stein JL, Soderberg KA, Moody MA, Denny TN, Zeng X, Fang**

847 **J, Moffett A, Lifson JD, Goedert JJ, Buchbinder S, Kirk GD, Fellay J,**
848 **McLaren P, Deeks SG, Pereyra F, Walker B, Michael NL, Weintrob A,**
849 **Wolinsky S, Liao W, Carrington M.** 2013. Influence of HLA-C expression level
850 on HIV control. Science **340:**87-91.
851

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

852 **Figure legends**

853

854 **FIG 1 Escape Map of HLA-associated polymorphisms for Gag, Pol, and Nef.**

855 Escape maps indicate the locations, specific residues and HLA restrictions of
856 HLA-associated polymorphisms (all $q < 0.2$). The global HIV-1 clade B consensus amino
857 acid sequence is used as a reference. Shaded vertical bars separate blocks of 10 amino acids.
858 "adapted" amino acids (those significantly overrepresented in the presence of a given HLA
859 allele) are red. "nonadapted" amino acids (those significantly underrepresented in the
860 presence of a given HLA allele) are blue. Polymorphisms associated with the same HLA
861 allele that occur in proximity to one another are grouped together in yellow boxes. A list of
862 all HLA-associated polymorphisms is provided in Table S1.

863

864 **FIG 2 Correlation between HLA-associated substitutions in Gag, Pol, and Nef and**
865 **viral load or CD4 count.**

866 The total number of HLA-associated substitutions in each subject's Gag, Pol, and Nef
867 sequence was counted (see methods). (A) Correlation between the number of
868 HLA-associated substitutions in Gag, Pol or Nef and pVL or CD4 count, (B) Correlation
869 between pVL and the number of HLA-associated substitutions in Pol, with
870 HLA-B*52:01-associated substitutions excluded (C) Correlation between pVL and the
871 number of HLA-B*52:01-associated substitutions in Pol (all patients). (D) Correlation
872 between the number of HLA-B*52:01 associated substitutions in Pol in
873 HLA-B*52:01-positive individuals (left panel) and HLA-B*52:01-negative individuals
874 (right panel). Analyses were performed using Spearman's correlation. Linear regression
875 lines are included in the plots.

876

877 **FIG 3 Polymorphic positions in HLA class I molecules and differential escape between**
878 **pairs of HLA subtypes.**

JVI Accepts published online ahead of print

879 In each ribbon diagram depicting the HLA-peptide-binding groove, the locations of
880 residues differing among subtype members of the (A) HLA-A*26, (B) HLA-C*03, and (C)
881 HLA-C*14 allele groups are highlighted in red and labeled with their locations and amino
882 acids. HLA-AP comparisons between subtype members are shown in the corresponding
883 plot below. Horizontal bars represent the natural logarithm of the odds ratio (lnOR), with
884 colors indicating the restricting allele. Infinite lnORs are set to values of ±4. Boldface type
885 indicates HLA-AP whose strengths of selection are statistically significantly different
886 between the two subtype members ($p < 0.05$, $q < 0.2$).
887

888 **FIG 4 Location of HLA-associated sites common to HIV-1 clade B-infected Japanese**
889 **and Caucasian cohorts, and those unique to Japan.**

890 The locations of all HLA-APs in Gag (500 codons), Pol (1,003 codons), and Nef (206
891 codons) are illustrated. The residues in the Pol transframe protein (TF) were not analyzed in
892 IHAC and are thus excluded (grey bar). Blue squares identify codons that harbored at least
893 one HLA-AP in both Japanese and IHAC cohorts. Red squares indicate codons that
894 harbored HLA-AP in Japan, but that were not associated with any HLA alleles in IHAC.
895

896 **FIG 5 HLA-AP displaying significantly different strengths of selection between**
897 **Japanese and IHAC cohorts.**

898 A phylogenetically-corrected interaction test was used to compare the natural logarithm
899 of the odds ratio (lnOR) of selection of HLA-APs in the Japanese cohort versus the IHAC
900 cohort. Comparisons with a $p < 0.01$ and $q < 0.05$ are reported. Bars represent the lnOR.
901 Infinite lnORs are set to values of ±4. Boldface type indicates HLA-AP that display
902 diametrically opposed directions of selection between the cohorts (defined here as lnORs of
903 association that were positive in one cohort but negative in the other, where the
904 cohort-specific p-values were <0.05 in both cases). The complete list of all comparisons
905 with $p < 0.05$ is available in Table S3 in the supplemental material.

JVI Accepts published online ahead of print

906

907 **FIG 6 HLA-AP identified as being under differential strength of selection in Japanese**
908 **and IHAC cohorts.**

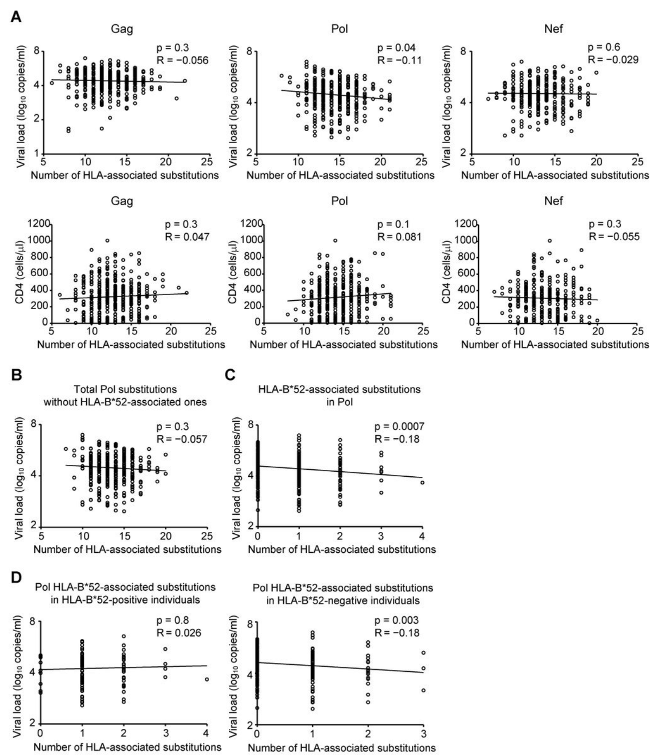
909 At a threshold of $p < 0.01$, $q < 0.05$, a total of 71 HLA-APs were identified as being
910 under significantly different strengths of selection in Japanese and IHAC cohorts. The
911 restricting HLA alleles and their HIV-1 protein locations are shown in (A). The number of
912 differentially-selected HLA-AP, broken down by HLA locus and HIV-1 protein, is shown
913 in (B).

JVI Accepts published online ahead of print

Figure 1

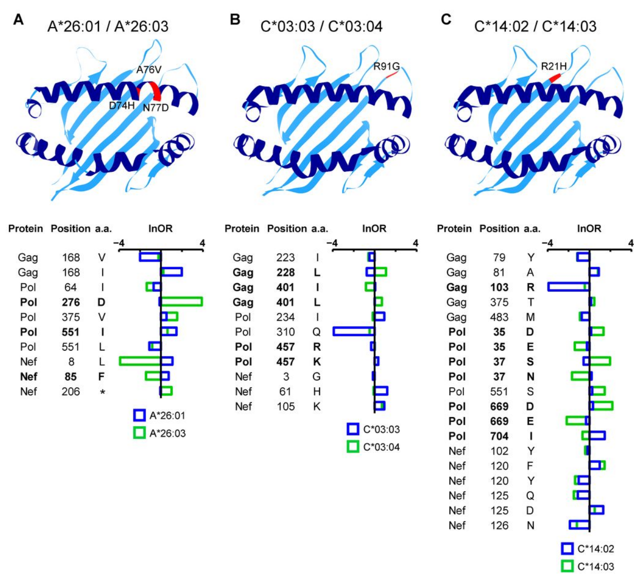
JVI Accepts published online ahead of print

Figure 2



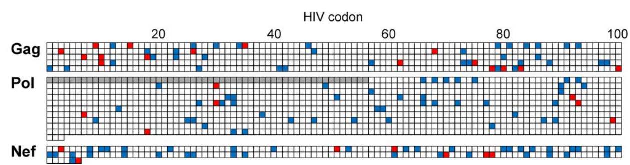
JVI Accepts published online ahead of print

Figure 3



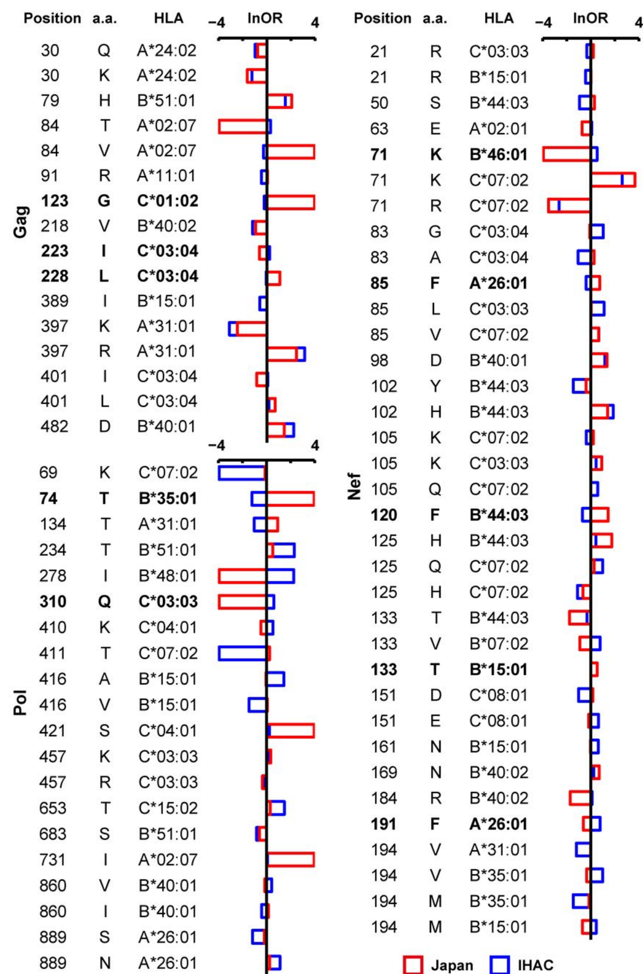
JVI Accepts published online ahead of print

Figure 4



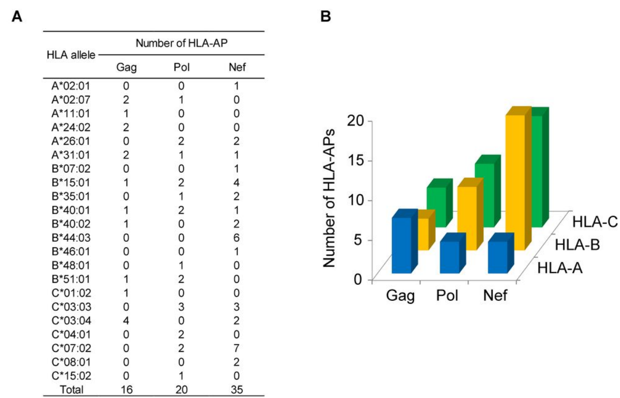
JVI Accepts published online ahead of print

Figure 5



JVI Accepts published online ahead of print

Figure 6



A

HLA allele	Number of HLA-AP		
	Gag	Pol	Nef
A*02:01	0	0	1
A*02:07	2	1	0
A*11:01	1	0	0
A*24:02	2	0	0
A*26:01	0	2	2
A*31:01	2	1	1
B*07:02	0	0	1
B*15:01	1	2	4
B*35:01	0	1	2
B*40:01	1	2	1
B*40:02	1	0	2
B*44:03	0	0	6
B*46:01	0	0	1
B*48:01	0	1	0
B*51:01	1	2	0
C*01:02	1	0	0
C*03:03	0	3	3
C*03:04	4	0	2
C*04:01	0	2	0
C*07:02	0	2	7
C*08:01	0	0	2
C*15:02	0	1	0
Total	16	20	35