Effects of Mutations on Replicative Fitness and Major Histocompatibility Complex Class I Binding Affinity Are Among the Determinants Underlying Cytotoxic-T-Lymphocyte Escape of HIV-1 Gag Epitopes

Yushen Du,[a,b] Tian-Hao Zhang,[c] Lei Dai,[a] Xiaojuan Zheng,[d] Aleksandr M. Gorin,[e] John Oishi,[f] Ting-Ting Wu,[b] Janice M. Yoshizawa,[g] Xinmin Li,[g] Otto O. Yang,[e,h,i] Otoniel Martinez-Maza,[e,f,j] Roger Detels,[f] Ren Sun[a,b,c]

Cancer Institute, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, School of Medicine, Zhejiang University, Hangzhou, China[a]; Department of Molecular and Medical Pharmacology, University of California, Los Angeles, Los Angeles, California, USA[b]; Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California, USA[c]; Key Laboratory of Animal Virology of Ministry of Agriculture, Zhejiang University, Hangzhou, China[d]; Department of Microbiology, Immunology & Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA[e]; Department of Epidemiology, Jonathan and Karin Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California, USA[f]; Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA[g]; Department of Medicine, David Geffen School of Medicine, Division of Infectious Diseases, University of California, Los Angeles, Los Angeles, California, USA[h]; AIDS Healthcare Foundation, Los Angeles, California, USA[i]; Department of Obstetrics and Gynecology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA[j]

ABSTRACT Certain "protective" major histocompatibility complex class I (MHC-I) alleles, such as B*57 and B*27, are associated with long-term control of HIV-1 *in vivo* mediated by the CD8[+] cytotoxic-T-lymphocyte (CTL) response. However, the mechanism of such superior protection is not fully understood. Here we combined high-throughput fitness profiling of mutations in HIV-1 Gag, *in silico* prediction of MHC-peptide binding affinity, and analysis of intraperson virus evolution to systematically compare differences with respect to CTL escape mutations between epitopes targeted by protective MHC-I alleles and those targeted by nonprotective MHC-I alleles. We observed that the effects of mutations on both viral replication and MHC-I binding affinity are among the determinants of CTL escape. Mutations in Gag epitopes presented by protective MHC-I alleles are associated with significantly higher fitness cost and lower reductions in binding affinity with respect to MHC-I. A linear regression model accounting for the effect of mutations on both viral replicative capacity and MHC-I binding can explain the protective efficacy of MHC-I alleles. Finally, we found a consistent pattern in the evolution of Gag epitopes in long-term nonprogressors versus progressors. Overall, our results suggest that certain protective MHC-I alleles allow superior control of HIV-1 by targeting epitopes where mutations typically incur high fitness costs and small reductions in MHC-I binding affinity.

IMPORTANCE Understanding the mechanism of viral control achieved in long-term nonprogressors with protective HLA alleles provides insights for developing functional cure of HIV infection. Through the characterization of CTL escape mutations in infected persons, previous researchers hypothesized that protective alleles target epitopes where escape mutations significantly reduce viral replicative capacity. How-

ever, these studies were usually limited to a few mutations observed *in vivo*. Here we utilized our recently developed high-throughput fitness profiling method to quantitatively measure the fitness of mutations across the entirety of HIV-1 Gag. The data enabled us to integrate the results with *in silico* prediction of MHC-peptide binding affinity and analysis of intraperson virus evolution to systematically determine the differences in CTL escape mutations between epitopes targeted by protective HLA alleles and those targeted by nonprotective HLA alleles. We observed that the effects of Gag epitope mutations on HIV replicative fitness and MHC-I binding affinity are among the major determinants of CTL escape.

KEYWORDS CTL escape, Gag epitopes, HIV-I, high-throughput fitness profiling, MHC binding prediction, intrapatient viral evolution

HIV-1-specific CD8$^+$ cytotoxic T lymphocytes (CTLs) represent the most critical immune response that limits HIV-1 replication *in vivo* (1–3). Their antiviral activity has been demonstrated in laboratory experiments and clinical observations (1, 4–6). HIV-1-specific CTLs from infected persons show robust killing of HIV-1-infected cells *in vitro* (7). In macaque models, *in vivo* depletion of CTLs causes consistent increase of viremia (3). CTLs recognize HIV-1-infected cells through the binding of short, virus-derived peptide epitopes that are presented on the cell surface via major histocompatibility complex class I (MHC-I) molecules. Although CTLs play an essential protective role, they ultimately fail to control HIV replication and to prevent disease progression in most individuals. By mutation, HIV-1 can escape from CTL killing in multiple ways, such as reduction of the binding of viral epitopes to MHC-I, interference with epitope processing, and impairment of binding recognition by T cell receptor (8–11). The pressure for HIV-1 to escape CTL recognition is in fact a major driver of viral evolution at the individual and population levels (12–14).

The genotype of MHC-I alleles determines the HIV-1 epitopes available for presentation and thus the targeting of CTLs (11, 15). Certain "protective" alleles are enriched in long-term nonprogressors (LTNPs) and are associated with superior viral immune control. However, the mechanism behind this protection is not fully understood (15–20). One proposed hypothesis is that CTL escape mutations in epitopes presented by these alleles significantly reduce viral replicative capacity. In support of this hypothesis, the high fitness costs of some CTL escape mutations have been investigated and observed experimentally (21–24). However, previous studies have relied on the identification of epitope mutations in samples from infected persons and characterization of their phenotypes *in vitro* and have revealed only a small number of mutations that survived the selection process. Moreover, as virus fitness *in vivo* depends on both its intrinsic replicative capacity and its ability to evade CTL pressure, protective MHC-I alleles may also target epitopes where evasion of CTL recognition is more difficult to occur. Therefore, a systematic examination of the fitness cost and the effect on CTL escape is necessary to obtain a quantitative comparison of epitopes targeted by protective and nonprotective MHC-I alleles (25–29).

In this study, we integrated multiple approaches, including high-throughput fitness profiling, *in silico* prediction of MHC-peptide binding affinity, and analysis of intraperson virus evolution, to systematically determine the differences between epitopes presented by protective MHC-I alleles and those presented by nonprotective MHC-I alleles with respect to HIV-1 Gag mutations. Mutations in epitopes corresponding to protective MHC-I alleles had higher replicative capacity cost and lower levels of reductions in MHC-I binding affinity. The conclusion was supported with consistent differences observed in Gag sequences from HIV-1-infected LTNPs and progressors.

RESULTS

High-throughput fitness profiling of HIV-1 Gag mutations. We have previously demonstrated the feasibility of using quantitative high-throughput genetics to systematically evaluate the fitness effects of point mutations in HIV-1, HCV and influenza virus
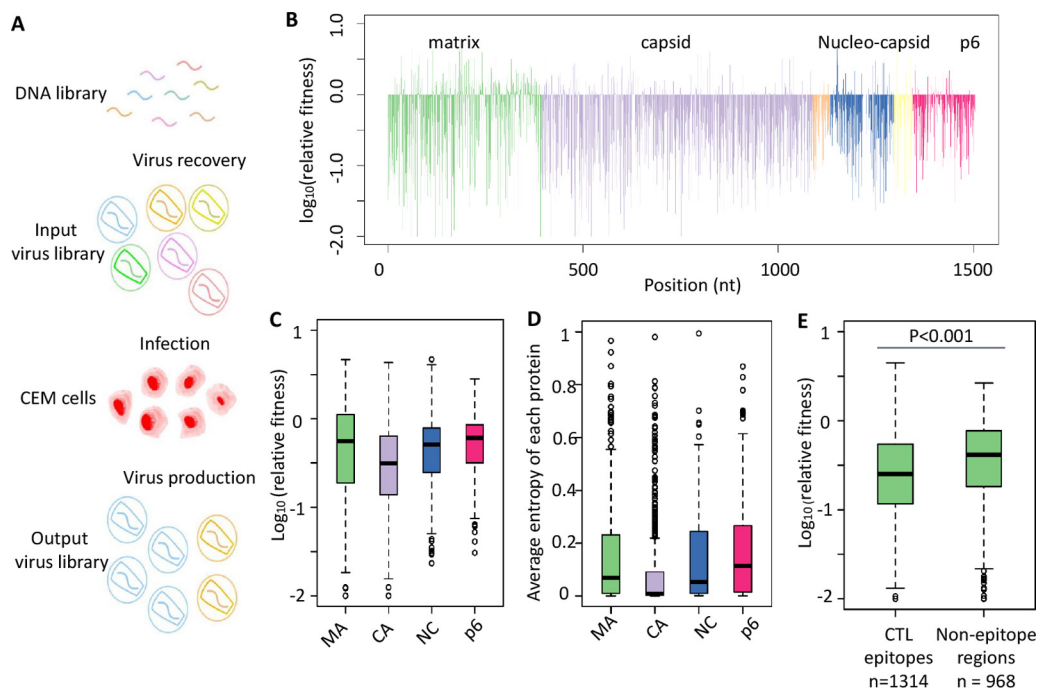
FIG 1 Quantitative high-throughput fitness profiling of HIV-I Gag polyprotein. (A) Experimental design of high-throughput fitness profiling of the HIV-1 Gag region. (B) Relative fitness score of each point mutation in Gag. Each Gag protein is labeled by a different color. MAb, matrix; CA, capsid; SP1, spacer peptide 1; NC, nucleocapsid; SP2, spacer peptide 2; p6, p6 protein; nt, nucleotide. (C) Average relative fitness scores of missense mutations in each Gag protein are shown using a box plot. (D) The average entropy of each Gag protein was calculated based on naturally occurred variants in the HIV sequence database at Los Alamos National Laboratory. (E) Relative fitness scores of mutations within or outside CTL epitope regions were compared. CTL epitopes were defined according to the 2013 update of best-characterized epitopes from the Los Alamos Database. A total of 1,314 mutations within CTL epitopes and a total of 968 mutations outside CTL epitopes were calculated.

(25–27, 29). In this study, we generated plasmid libraries of single nucleotide mutations in the Gag region of HIV-1 molecular clone NL4-3 using error-prone PCR mutagenesis. The corresponding virus libraries were reconstituted in 293T cells by transfecting the plasmid libraries followed by two successive passages of 6 days each in a human leukemic T cell line (CEM) (Fig. 1A). Relative fitness (RF) scores, representing the replicative capacities of individual mutants, were calculated as the ratio of the frequency in the library after the two passages to the frequency in the input viral library (Fig. 1B; see also Table S1 in the supplemental material). The mutant library covered 74% (3,340/4,509) of all possible single nucleotide mutations and 27% (2,788/10,020) of single amino acid mutations in Gag. The clear separation of the RF scores between synonymous mutations and missense mutations suggested efficient selection of viable versus nonviable mutants in the passaging process (see Fig. S1A in the supplemental material). We further quantified the effect of missense mutations on replicative capacity and the fraction of lethal mutations for four major proteins encoded in the Gag region. These four proteins displayed various levels of mutation tolerability (Fig. 1C; see also Fig. S1B). For example, the fitness costs of capsid mutations were significantly higher than those seen with all other proteins ($P < 0.001$, $P < 0.001$, and $P < 0.001$ compared with matrix, nucleocapsid, and p6, respectively; two-tailed Wilcoxon rank sum test). Capsid also had more lethal mutations than other proteins ($P = 0.007$, $P < 0.001$, and $P < 0.001$ compared with matrix, nucleocapsid, and p6, respectively; two-tailed Fisher exact test). Around 20% of the missense mutations on capsid were lethal for viral replication in our profiling (30). The fitness effects of mutations in individual Gag protein correlated well with genetic diversity in naturally occurring sequences in the Los Alamos National Laboratory HIV Sequence Database (https://www.hiv.lanl.gov/content/index) (Fig. 1D; see also Fig. S1C and D).

As Gag-specific CTLs are important for viral control (1, 21, 31), we examined whether mutations in previously characterized CTL epitopes carry higher fitness cost than in other regions. CTL epitopes were defined according to the 2013 update of best-characterized epitopes from the Los Alamos National Laboratory HIV Immunology Database (https://www.hiv.lanl.gov/content/index). Our fitness profiling of Gag covered 1,314 missense mutations in CTL epitopes and 968 missense mutations in nonepitope regions. Mutations in epitopes had higher fitness cost than those in nonepitope regions (Fig. 1E; two-tailed t test, $P < 0.001$), suggesting that Gag-specific CTLs target regions of the proteins with lower tolerance of mutations. Moreover, 70% (50/72) of the epitopes were located on capsid, which was the most conserved protein in the Gag region. Overall, our comprehensive data support the idea that the intolerance of mutations in Gag epitopes might be one of the reasons for better viral control by Gag-specific CTLs.

Systematic evaluation of effects of Gag epitope mutations on MHC-I binding affinity. One of the proposed mechanisms whereby HIV-1 can escape CTL killing is via epitope mutations that reduce their binding affinity to MHC-I (8, 32). Several data-driven computation programs have been developed to predict affinity of peptide binding to specific MHC-I molecules (33–35) and thereby allow identification of epitope variants that facilitate escape from CTL recognition. NetMHC is the state-of-art predictor and is based on an artificial neural network (34, 35), achieving up to 80% correlation with experimental data (36). We used netMHC-4.0 to estimate the effects of all single amino acid mutations in Gag CTL epitopes (compared to NL4-3 as the index sequence) on the binding affinity (dissociation constant [K_d]) to MHC-I (Table S2). A total of 62 epitopes were included in the analysis (excluding epitopes with predicted $K_d > 10$ μM).

Peptides bind to MHC-I primarily through anchor residues, which are usually located at position 2 and the C terminus (9th or 11th residue according to the length of peptide) of the peptide. We thus examined the effect of mutations at different positions within an epitope. For each position on an epitope, the missense mutations across all 62 epitopes were examined. A total of 11,580 mutations were included in the analysis. As expected, the greatest drop of binding affinity was observed at anchor residues (2nd, 9th, and 11th residue; $P < 0.001$ when comparing each of the anchor residues with other residues; two-tailed Wilcoxon rank sum test) Fig. 2A, validating the accuracy of using netMHC4.0 to predict binding affinity.

Next, we compared the effects of mutations on MHC-I binding affinity across different Gag epitopes (Fig. 2B; see also Fig. S2A). Notably, different Gag epitopes showed various profiles of changes of binding affinity caused by single amino acid mutations. For example, most mutations in epitopes targeted by HLA B*2705 remained robust, with fewer mutations leading to increased K_d (i.e., lower binding affinity), suggesting that evasion of CTL recognition via reduced binding to MHC-I is uncommon; in contrast, epitopes targeted by HLA A*0207 were more sensitive to mutations. By comparing the *in silico* binding affinity prediction with the fitness profiling of mutations in each epitope, we further evaluated the relationship between MHC-I binding affinity and viral replicative capacity (RF scores). A weak but significant negative correlation was observed (Fig. S2B, rho $= -0.093$, $P = 0.003$), which may reflect a tradeoff between viral fitness and the reduction of MHC-I binding affinity for escape (32).

Impacts of Gag epitope mutations on viral replication and MHC-I binding for epitopes presented by protective versus nonprotective MHC-I alleles. The replicative capacity of HIV-1 and its ability to escape CTL recognition are two possible determinants of viral replication *in vivo*. To explore the mechanism of superior viral control observed in individuals with protective MHC-I alleles, we combined fitness profiling of mutations and prediction on MHC-I binding affinity to examine the difference between the CTL epitopes targeted by protective MHC-I alleles and those targeted by nonprotective MHC-I alleles.

We first examined several well-characterized epitopes presented by the protective HLA alleles B*57 (KF11 [position 162 to 172] and TW10 [position 240 to 249]) and B*27 (KK10 [position 263 to 272]) and observed that mutations in these epitopes led to
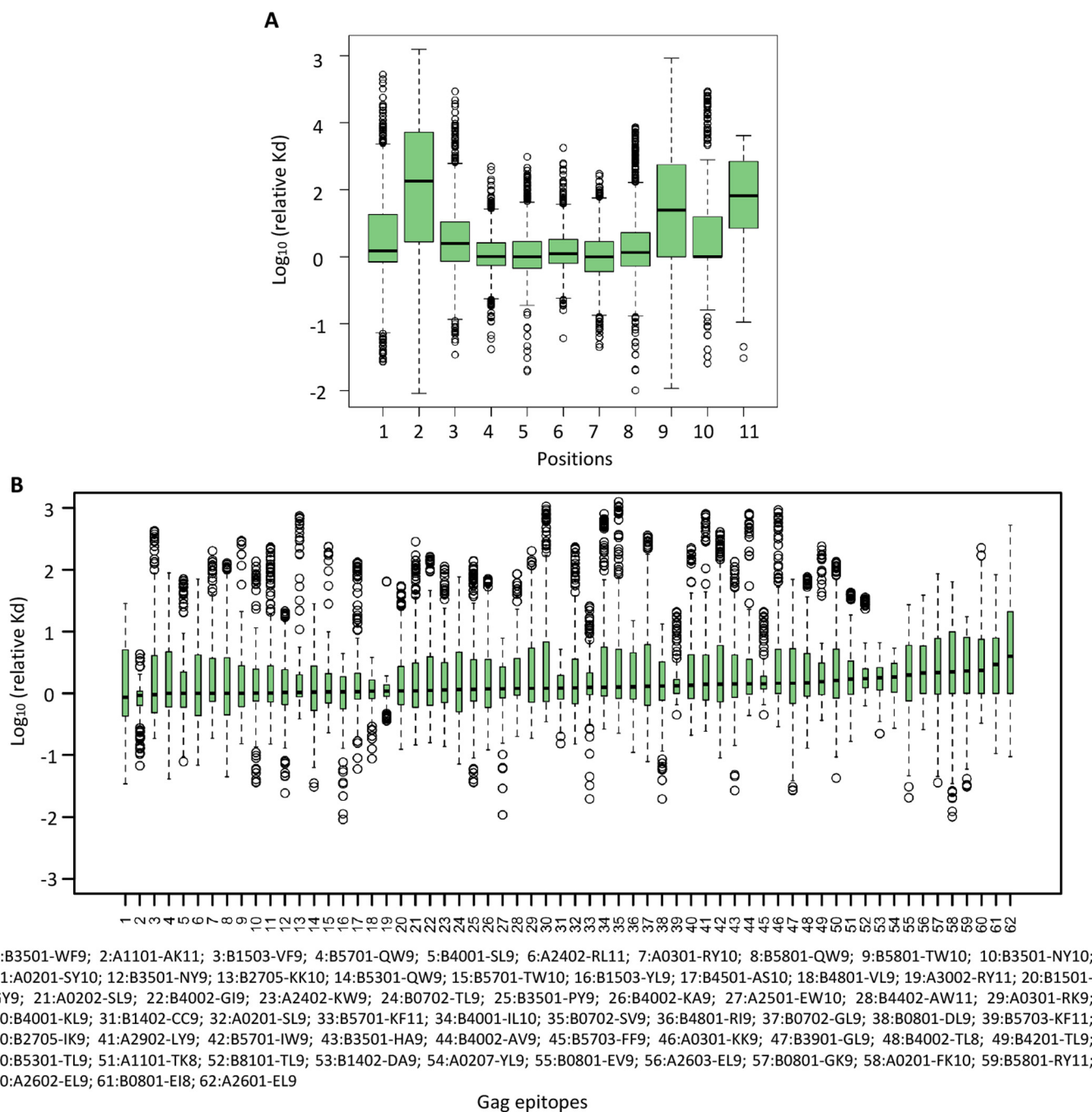
A

B

1:B3501-WF9; 2:A1101-AK11; 3:B1503-VF9; 4:B5701-QW9; 5:B4001-SL9; 6:A2402-RL11; 7:A0301-RY10; 8:B5801-QW9; 9:B5801-TW10; 10:B3501-NY10; 11:A0201-SY10; 12:B3501-NY9; 13:B2705-KK10; 14:B5301-QW9; 15:B5701-TW10; 16:B1503-YL9; 17:B4501-AS10; 18:B4801-VL9; 19:A3002-RY11; 20:B1501-GY9; 21:A0202-SL9; 22:B4002-GI9; 23:A2402-KW9; 24:B0702-TL9; 25:B3501-PY9; 26:B4002-KA9; 27:A2501-EW10; 28:B4402-AW11; 29:A0301-RK9; 30:B4001-KL9; 31:B1402-CC9; 32:A0201-SL9; 33:B5701-KF11; 34:B4001-IL10; 35:B0702-SV9; 36:B4801-RI9; 37:B0702-GL9; 38:B0801-DL9; 39:B5703-KF11; 40:B2705-IK9; 41:A2902-LY9; 42:B5701-IW9; 43:B3501-HA9; 44:B4002-AV9; 45:B5703-FF9; 46:A0301-KK9; 47:B3901-GL9; 48:B4002-TL8; 49:B4201-TL9; 50:B5301-TL9; 51:A1101-TK8; 52:B8101-TL9; 53:B1402-DA9; 54:A0207-YL9; 55:B0801-EV9; 56:A2603-EL9; 57:B0801-GK9; 58:A0201-FK10; 59:B5801-RY11; 60:A2602-EL9; 61:B0801-EI8; 62:A2601-EL9

Gag epitopes

FIG 2 Systematic evaluation of effects of mutations on MHC-I binding affinity. (A) Effects on MHC-I binding affinity of single amino acid mutations at different positions within an epitope. The binding affinity of mutated epitopes was predicted by the use of netMHC-4.0. (B) The distribution of predicted changes in MHC-I binding affinity caused by single amino acid mutations across CTL epitopes in Gag (~171 to ~209 mutations for each epitope, depending on the length).

greater loss of viral fitness than in the epitopes presented by the nonprotective HLA allele A*02 (SL9 [position 77 to 83]) (Fig. 3A) (37, 38). Furthermore, we ranked MHC-I alleles by the ratio of their prevalence in HIV controllers to their prevalence in progressors based on The International HIV Controllers (TIHIVC) study (Table S3) (2). The top 5 MHC-I alleles with available binding affinity predictions (B*5701, B*2705, B*1402, B*2501, and B*5801) were assigned as "protective," while the bottom 5 (A*3002, B*0702, A*2902, B*3501, and B*4001) were assigned as "nonprotective." We observed that the mutations in Gag epitopes presented by protective MHC-I alleles showed significantly greater loss of RF scores than the mutations presented by nonprotective alleles (P =
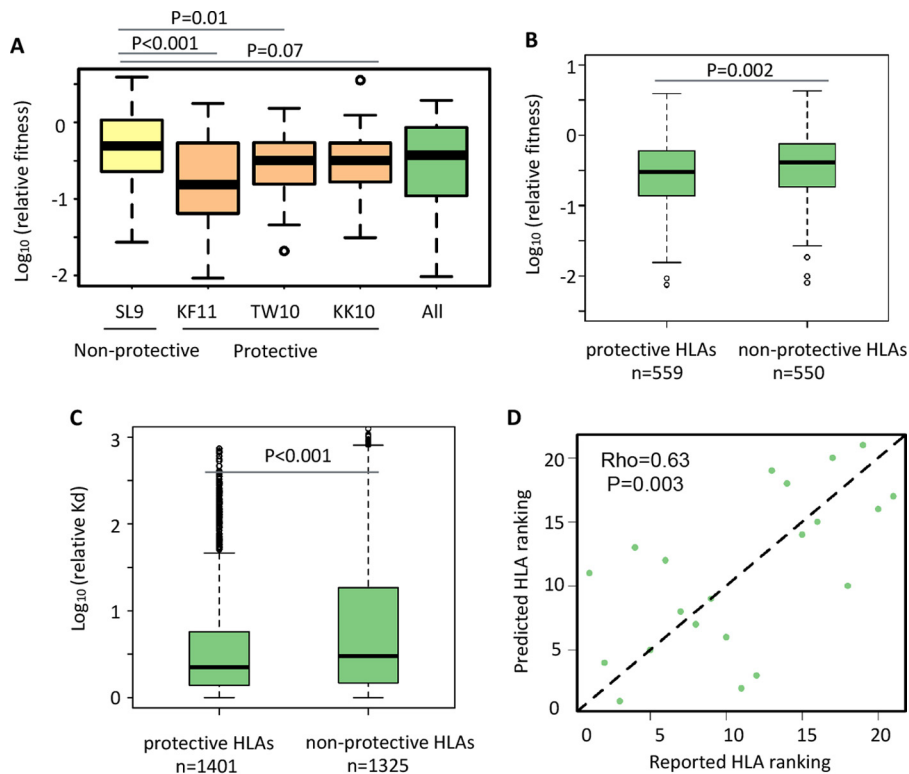
FIG 3 Systematic analysis of differences between protective and nonprotective MHC-I alleles. (A) Relative fitness scores of mutations in indicated epitopes. Epitopes shown as examples are SL9 (Gag positions 77 to 85), KF11 (Gag positions 163 to 173), TW10 (Gag positions 240 to 249), and KK10 (Gag positions 263 to 272). (B) Relative fitness scores of mutations in epitopes presented by protective or nonprotective MHC-I alleles. A total of 559 mutations in the epitopes targeted by protective MHC-I alleles and a total of 550 mutations targeted by nonprotective MHC-I alleles were analyzed. (C) MHC-I binding affinity changes of mutations in epitopes presented by protective or nonprotective MHC-I alleles were compared. A total of 1,401 mutations in the epitopes targeted by protective MHC-I alleles and a total of 1,325 mutations in the epitopes targeted by nonprotective MHC-I alleles were analyzed. (D) The correlation between predicted ranking of MHC-I protectiveness and ranking calculated based on TIHIVC study data. Predicted ranking was fitted by a linear regression model using the average effect of mutations on relative fitness scores and MHC-I binding affinity as variables (Spearman correlation, 0.63; $P = 0.003$).

0.004; two-tailed t test) (Fig. 3B). These results suggest that protective MHC-I alleles promote targeting HIV-1 epitopes that are less tolerant of mutations.

Next, we examined the effect of mutations on the MHC-I binding affinity of epitopes presented by protective versus nonprotective MHC-I alleles. Interestingly, single amino acid mutations in epitopes presented by protective alleles showed lower levels of reduction in binding affinity to MHC-I ($P < 0.001$; two-tailed t test) (Fig. 3C). Overall, these data suggest that the epitopes presented by protective MHC-I alleles have two important properties: higher fitness costs and less abrogation of MHC-I binding caused by mutations in the epitope sequences. These results were robust if considering B*2705 and B*5701 only as protective MHC-I alleles and reached statistical significance in the TIHIVC study (Fig. S3).

We further examined whether these two properties of CTL epitopes (fitness cost and MHC-I binding affinity) might explain the protective effect of different MHC-I alleles. We first tested if a single property is sufficient to explain the protectiveness of MHC-I. Average relative fitness scores alone were correlated only weakly with the ranking of protectiveness of MHC-I alleles, while the effects of mutations on MHC-I binding did not show a significant correlation with the ranking of their protective effect (Fig. S4). Then, we examined whether combining the two properties can better explain the protectiveness. Indeed, the ranking of the MHC-I types fitted by a linear regression model that includes both properties as predictor variables was significantly correlated with the

actual ranking (Spearman rank rho = 0.63, $P = 0.003$) (Fig. 3D). Consistent with the comparison between protective and nonprotective MHC-I alleles (Fig. 3B and C), this finding suggests that these two variables could largely account for their influence on immune containment of HIV-1.

Mutations in HIV-1 Gag observed in intraperson viral evolution. Finally, we examined the evolution of HIV-1 epitopes in 4 progressors and 4 long-term nonprogressors (LTNPs) with chronic HIV-1 infection. Samples were collected from the Multicenter AIDS Cohort Study (MACS) (Table S4) with matched CD4 cell percentages at the first time point. All subjects were antiretroviral therapy (ART) naïve; thus, the major selective pressure for viral evolution was imposed by the host immune system. Progressors proceeded to AIDS-related death, while LTNPs maintained stable (drop of less than 10%) blood $CD4^+$ T cell levels (Table S4) over 4 years of observation after enrollment in the cohort. Proviral DNA was extracted from 10 million peripheral blood mononuclear cells (PBMC), and the entire *gag* region (1,500 bp) was amplified and subjected to deep sequencing (Materials and Methods and Fig. S5A). The consensus *gag* sequences examined at enrollment and 4 years later displayed the expected phylogenetic clustering of sequences in each individual (Fig. S5B). We also reconstructed full-length (1,500-bp) viral haplotypes by the use of PredictHaplo. The frequency of mutations calculated from reconstructed haplotypes was highly correlated with the frequency of mutations in the raw data, indicating that the reconstruction of haplotypes was reliable (Fig. S4C). For both groups, we observed an accumulation of mutations in specific HLA epitopes (example shown in Fig. 4A) and a minor increase in the Shannon entropy of epitope regions at the second time point (Fig. S4D), suggesting that these epitope regions might be under the control of positive selection to escape CTL restriction in the infected individuals.

We next examined the effect of the observed mutations on viral replicative fitness and CTL escape. Mutations on virus in each individual were determined by comparing the sequences seen at time point 2 to the consensus sequence at time point 1. The fitness effects of these mutations were extracted from our profiling data. For CTL epitope regions, we found that the mutations observed in LTNPs had slightly higher fitness cost than those in progressors (Fig. 4B), although the results did not reach statistical significance. Additionally, the CTL epitope mutations observed in progressors resulted in a more substantial drop in the binding affinity to MHC-I than the mutations in LTNPs (Fig. 4C). Finally, we extracted the possible MHC-I escape mutations from each individual by comparing the binding affinity of mutations with the global consensus defined in the database at Los Alamos National Laboratory (39). The escape mutations associated with LTNPs showed significantly lower replicative fitness than those associated with progressors ($P = 0.01$; two-tailed Wilcoxon tank sum test) (Fig. 4D). Thus, we observed the same trend as that seen in our previous analysis, where LTNPs are linked with higher replicative fitness cost and lower drops in binding affinity to MHC-I introduced by mutations in corresponding epitopes.

DISCUSSION

In the present study, by combining high-throughput fitness profiling and *in silico* prediction of MHC-peptide binding affinity, we observed that epitopes targeted by protective MHC-I alleles have two important properties: mutations in these epitopes are more deleterious with respect to viral replication and have a smaller effect on evasion of CTL recognition than those targeted by nonprotective MHC-I alleles. Collectively, these two properties can largely account for the superior viral control conferred by protective MHC-I alleles. We observed a similar trend in virus populations in infected individuals, wherein escape mutations in LTNPs were associated with higher fitness costs and smaller effects on HLA binding affinity than were seen in escape mutations in progressors.

CTL escape mutations with highly predictable patterns are frequently observed in HIV-infected individuals (8, 9, 12, 15, 17, 33, 40–42). Previous studies have revealed that escape mutations in epitopes presented by protective MHC-I alleles often result in a
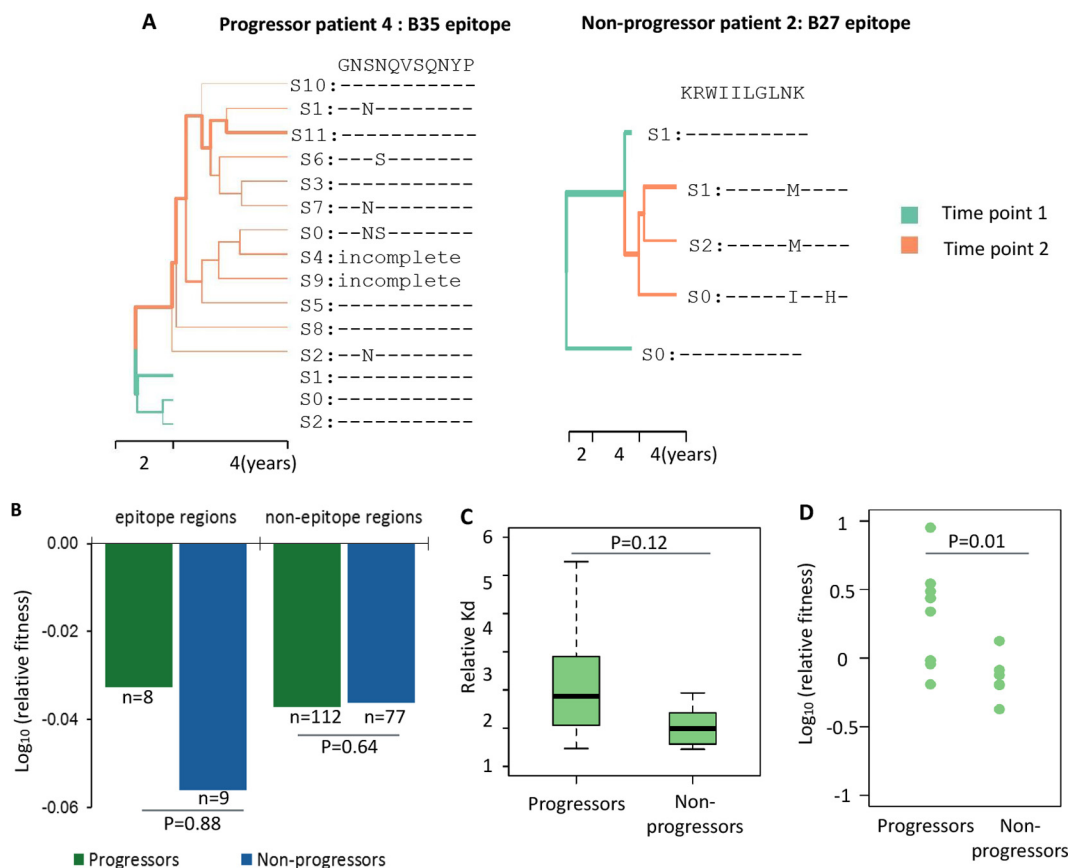
FIG 4 HIV intraperson evolution in LTNPs and progressors. (A) Representative phylogenetic trees of virus haplotypes of LTNPs and progressors. Viral haplotypes were assembled by PredictHaplo. Maximum clade credibility (MCC) trees were constructed by BEAST. The mutations in a representative HLA epitope are labeled. The width of branches is proportional to the abundance of the corresponding haplotypes in the population. The colors green and orange represent sampling time point 1 and time point 2. (B) Relative fitness scores of naturally aroused variants in both groups. (C) Predicated effects of epitope mutations in samples on MHC-I binding affinity. For each individual, the epitopes were selected based on the corresponding HLA serotypes. Epitope sequences from all reconstructed viral haplotypes (11 from progressors and 6 from LTNP) were included in the analysis. (D) Relative fitness scores of possible MHC-1 escape mutations in both groups. Possible MHC-I escape mutations from each individual were defined as the ones with levels of binding affinity lower than those seen with the global consensus sequences.

high cost with respect to replicative capacity or require the preexistence of compensatory mutations (11, 32, 40, 43). However, those studies were usually restricted to studying a few mutations observed *in vivo*. In this work, we overcame the restrictions represented by the limited sampling with a systematic unbiased fitness profiling of Gag mutations. Consistent with previous observations, our comprehensive fitness data and integrated analyses support the notion that protective MHC-I alleles result in viral epitopes with lower mutational tolerance. Reduced HIV replication capacity resulting from deleterious mutations is associated with reduced rates of CD4 decline and disease progression (11, 16, 44). Thus, the high fitness cost of CTL escape mutations can in part explain the superior viral control observed in individuals with protective MHC-I alleles.

The viral growth that occurs under conditions of CTL selective pressure depends not only on intrinsic replication capacity but also on the ability to escape CTL recognition. Reduction of MHC-I binding affinity through the activity of epitope mutations is another contributor to CTL escape. Based on the calculations performed using Net-MHC4.0, we noted that mutations in epitopes presented by protective MHC-I alleles had a smaller effect on MHC-I binding than those presented by nonprotective MHC-I. Furthermore, we found that incorporating the effect of mutations on MHC-I binding into a linear regression model substantially improved the prediction of the protectiveness of MHC-I alleles, in comparison with the model that accounted only for fitness costs.

There are a few caveats pertaining the current study. First, we used random mutagenesis to introduce mutations in the Gag region. Although we optimized the mutation rate to approach 1 mutation per fragment, it is possible that there were multiple mutations in single viral clones, which might bias the fitness score of some mutations. Second, our mutant library was limited to single nucleotide mutations in the NL4-3 virus backbone; thus, we were unable to assess the fitness effect of compensatory mutations or the effect of these mutations on different genetic backgrounds (i.e., different strains of HIV-1). Third, we did not generate every possible amino acid variant at every position in Gag since our mutagenesis method usually introduces single nucleotide changes, thus limiting the diversity of amino acid changes at each position. Fourth, although drops in MHC-I binding affinity are considered to represent a major mechanism of HIV CTL escape (8), other factors such as intracellular epitope processing and recognition of T cell receptor may also contribute to CTL escape. Lastly, the sample size of infected individuals without any antiviral treatment was small for our intraperson viral evolution analysis, which limited the statistical power of our comparisons between LTNPs and progressors in terms of fitness cost and MHC binding. More samples are required for the further validation of our results. However, despite these limitations, we have performed a most comprehensive profiling of variants within Gag, which enables us to integrate with prediction of MHC binding affinity and obtain verification with viral sequence analyses of sequential samples from infected individuals.

Understanding the mechanism of viral control achieved in LTNPs with protective MHC-I alleles provides insights for developing functional cure and T cell-mediated vaccine against HIV (16, 44, 45). Several T cell vaccine strategies focus on using evolutionarily conserved regions in HIV genome as immunogens, with the promise that escape mutations in the conserved regions will incur higher fitness cost (46–50). However, many studies have documented that conserved regions are not necessarily essential for viral fitness, although there is some correlation (27, 29). Our systematic investigations of single amino acid mutations in Gag will more precisely pinpoint the sequences that are essential for viral replication and that are less likely to escape CTL, thereby aiding the rational design of immunogens for vaccine development.

MATERIALS AND METHODS

Construction of high-density mutant libraries for the HIV Gag region. To generate mutant plasmid libraries, we divided the entire *gag* genes in replication-competent proviral plasmid NL43 into 3 fragments, each spanning position 790 to position 1419, position 1419 to position 1995, and position 1995 to position 2292 on HXB2 coordinates. Mutations were randomly introduced into each fragment by performing error-prone PCR using Mutazyme II DNA polymerase (Stratagene). Mutated segments were then ligated back into the proviral backbone. The ligated products were electroporated into high-efficiency MegaX DH10B T1R electrocompetent cells (Invitrogen). Clones ($n = 5$ to 10) were randomly picked and analyzed by Sanger sequencing to estimate the mutation rate (approximately 1 mutation/clone). Approximately 50,000 bacterial colonies were collected for each of the three small libraries.

Transfection, viral titer determinations, and passage of HIV mutant libraries. To reconstitute the mutant virus library, approximately 15 million 293T cells were transfected with each mutant plasmid library (one of 3 small libraries; 16 μg) using Lipofectamine 2000 (Life Technologies, Inc.). The cells were rinsed with phosphate-buffered saline (PBS) at 12 to 14 h posttransfection and were maintained in fresh Dulbecco's modified Eagle's medium (DMEM) growth media supplemented with 10% fetal bovine serum (FBS) and 1$\times$ penicillin-streptomycin. The supernatants were harvested at 72 h posttransfection, filtered through a 0.45-μm-pore-size disposable syringe filter (Olympus), and stored at -80°C in small aliquots. The 50% tissue culture infective dose (TCID$_{50}$) of viral supernatants was measured using GHOST 3-X4/R5 indicator cells (gift of Matthew Marsden and Jerry Zack), which were derived from human osteosarcoma cells and stably transfected with the HIV long terminal repeat (LTR) driving a human green fluorescent protein (hGFP) construct (23). To passage each viral mutant library, approximately 30 million CEM T-lymphocyte cells were used for infection at a low multiplicity of infection (MOI = 0.05) and were supplemented with 2 μg/ml Polybrene (Sigma). At ~14 to ~16 h postinfection, cells were centrifuged at 1,000 rpm for 5 min and washed with PBS followed by the addition of fresh RPMI 1640 growth medium. Extracellular viruses were harvested at approximately 6 days postinfection when syncytium formation can be observed in ~60% to ~80% of cells. Two rounds of passaging were performed for each library.

Library preparation for deep sequencing. Viral RNAs were isolated from the viral supernatants using a QIAamp Viral RNA Minikit (Qiagen), treated with DNase I, and reverse transcribed using Superscript III reverse transcriptase (Life Technologies, Inc.). The plasmid mutant libraries or cDNAs from the viral mutant libraries (transfection or infection) were amplified using KOD Hot Start DNA polymerase. The amplified fragments were then ligated with the sequencing adapter, which had three nucleotide

multiplexing identifiers (IDs) to distinguish the different samples. An Illumina HiSeq 2000 PE100 system was used for sequencing.

Sequencing data analysis. A Burrows-Wheeler aligner (BWA) was used to map sequencing reads to reference sequences (18). Paired-end reads were used for error correction. The relative frequency of each mutation was calculated for each condition, and the relative fitness score was calculated as the difference between the frequency in the passaged library and the frequency in the transfected library. To further improve data quality, mutations with a frequency of <0.01% in the transfection library were filtered out, and possible G-A hypermutations were removed (3). Lethal mutations were defined as the mutations with $\log_{10}$(RF score) of less than -0.95. That value represented the distribution peak of nonsense mutations, while only 2.5% of synonymous mutations fell below this cutoff.

Conservation analysis among clinical isolates. A total of 6,097 prealigned HIV-1 subtype B Gag sequences were downloaded from the database at Los Alamos National Laboratory. No filter was applied for sampling time, country, or individuals' information. The Shannon entropy of all residues was calculated by custom scripts, and the resulting data were deposited at https://github.com/Tian-hao/HIV-clinical/.

Prediction of MHC binding affinity using netMHC4.0. MHC-I binding affinity (i.e., K_d increase) was calculated by netMHC4.0. Epitope sequences and MHC-I alleles were paired according to the best-characterized HIV-1 CTL epitopes from the database at Los Alamos National Laboratory (51). The binding affinity change is calculated as the ratio of the K_d value determined for the mutated epitope and that determined for the parental epitope with MHC-1.

Ranking of protectiveness of MHC-I alleles. We defined the protectiveness of MHC-I alleles as the ratio of the HLA allele prevalence in nonprogressors to the HLA allele prevalence in progressors (2). We then ranked the MHC-I types by two different properties of mutations in the targeting epitopes as follows: (i) $\text{rank}_{\text{fitness}}$, representing the average effect on viral replication fitness of missense mutations profiled in our fitness data (i.e., relative fitness score), and (ii) $\text{rank}_{\text{binding}}$, representing the average effect on MHC-I binding affinity of single amino acid substitutions (i.e., increase in K_d). We used these two properties of MHC-I types to fit the ranking of protectiveness $\text{rank}_{\text{protect}}$ using the following linear regression model:

$$\text{rank}_{\text{protect}} = \alpha \cdot \text{rank}_{\text{fitness}} + \beta \cdot \text{rank}_{\text{binding}}$$

The best-fit parameters were $\alpha = 0.51$ and $\beta = 0.35$. The fitted ranking of protectiveness showed a better correlation than predictions based solely on $\text{rank}_{\text{fitness}}$ (see Fig. S3 in the supplemental material), indicating that the effect of mutations on MHC-I binding affinity contributed to the protectiveness of MHC-I alleles.

As a control, we also fitted the ranking of protectiveness using the following alternative model:

$$\text{rank}_{\text{protect}} = \alpha \cdot \text{rank}_{\text{fitness}} + \beta \cdot \text{rank}_{\text{random}}$$

where we assigned a trait ($\text{rank}_{\text{fitness}}$) that was randomly ranked among MHC-I alleles, in addition to the ranking of fitness effects. We found that only 4.3% (43/1,000) randomly ranked traits produced a better fit of $\text{rank}_{\text{protect}}$ than the incorporated model, suggesting that the effect of mutations on MHC-I binding contributes to the protectiveness of MHC-I alleles.

Sequencing and analysis of viral samples from infected individuals. PBMC samples from 4 paired progressors and long-term nonprogressors (LTNPs) were kindly provided from the Multicenter AIDS Cohort Study (MACS). All of these infected individuals were treatment naive and paired according to their CD4 cell percentage at baseline. For each individual, we obtained two PBMC samples at time points 4 years apart, where the first sample was collected at the earliest time point in the MACS cohort. The progressors proceeded to the AIDS phase and died at year 4 during the cohort study, while the LTNPs maintained stable CD4 cell counts. DNAs were extracted from 10 million PBMC from each individual. The entire *gag* region (1,500 bp) was amplified by nested PCR. The gel-purified PCR products were then subjected to random fragmentation by sonication to achieve fragments of 200 to 700 bp. The fragmented libraries were prepared for high-throughput sequencing with an Illumina HiSeq 2000 system.

Viral sequences were mapped onto HIV-1 molecular clone NL4-3. The haplotypes of Gag genes were constructed by the use of PredictHaplo1.0 (52). Consensus sequences were determined. Variations at each nucleotide were identified if a haplotype sequence was found to be different from the consensus sequence. The tree of haplotypes was constructed using Phylip.

A mutation at time point 2 was defined as a single nucleotide polymorphism (SNP) if it was different from the consensus sequence at time point 1. Incomplete reads were filtered out. The epitopes were called for all individuals' HLA serotypes that had targeted epitopes in the table of the best-defined CTL epitopes. Escape mutations were defined by the ones with a level of MHC-I binding affinity lower than that seen with the global consensus sequence. For prediction of MHC-1 binding affinity, the epitope sequences of all reconstructed haplotypes were used as the input sequences. A total of 59 progressor sequences and a total of 47 LTNP sequences were used. All the custom scripts were deposited at https://github.com/Tian-hao/HIV-clinical/.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio .01050-17.

FIG S1, TIF file, 2.6 MB.

FIG S2, TIF file, 2.6 MB.

FIG S3, TIF file, 2.6 MB.
FIG S4, TIF file, 2.6 MB.
FIG S5, TIF file, 2.6 MB.
TABLE S1, XLSX file, 0.1 MB.
TABLE S2, XLSX file, 0.02 MB.
TABLE S3, XLSX file, 0.01 MB.
TABLE S4, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB. 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J Virol 68:6103–6110.

2. Pereyra F, Jia X, McLaren PJ, Kadie CM, Carlson JM, Heckerman D, De PIW, Sullivan J, Gonzalez E, Davies L, Camargo A, Moore JM, Gupta S, Crenshaw A, Burtt NP, Guiducci C, Cutrell E, Rosenberg R, Moss KL, Lemay P, Leary JO, Schaefer T, Verma P, Toth I, Block B, Rothchild A, Lian J, Proudfoot J, Alvino DML, Vine S, Addo MM, Allen TM, Altfeld M, Henn MR, Le Gall S, Streeck H, Walker BD; International HIV Controllers Study. 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 330:1551–1557. https://doi.org/10.1126/science.1195271.

3. Schmitz E, Kuroda MJ, Santra S, Sasseville VG, Simon MA, Lifton MA, Racz P, Tenner-Racz K, Dalesandro M, Scallon BJ, Ghrayeb J, Forman MA, Montefiori DC, Rieber EP, Letvin NL, Reimann KA. 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. Science 283:857–860.

4. McMichael AJ, Rowland-Jones SL. 2001. Cellular immune responses to HIV. Nature 410:980–987. https://doi.org/10.1038/35073658.

5. Ogg GS, Jin X, Bonhoeffer S, Dunbar PR, Nowak MA, Monard S, Segal JP, Cao Y, Rowland-Jones SL, Cerundolo V, Hurley A, Markowitz M, Ho DD, Nixon DF, McMichael AJ. 1998. Quantitation of HIV-1—specific cytotoxic T lymphocytes and plasma load of viral RNA. Science 279:2103–2106.

6. Streeck H, Nixon DF. 2010. T cell immunity in acute HIV-1 infection. J Infect Dis 202:S302–S308. https://doi.org/10.1086/655652.

7. Yang OO, Kalams SA, Rosenzweig M, Trocha A, Jones N, Koziel M, Walker BD, Johnson RP. 1996. Efficient lysis of human immunodeficiency virus type 1-infected cells by cytotoxic T lymphocytes. J Virol 70:5799–5806.

8. Lin GG, Scott JG. 2012. HIV escape mutations occur preferentially at HLA binding sites of CD8 T cell epitopes. AIDS 100:130–134.

9. Martin E, Carlson JM, Le AQ, Chopera DR, McGovern R, Rahman MA, Ng C, Jessen H, Kelleher AD, Markowitz M, Allen TM, Milloy MJ, Carrington M, Wainberg MA, Brumme ZL. 2014. Early immune adaptation in HIV-1 revealed by population-level approaches. Retrovirology 11:64. https://doi.org/10.1186/s12977-014-0064-1.

10. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science 296:1439–1443. https://doi.org/10.1126/science.1069660.

11. Kløverpris HN, Leslie A, Goulder P. 2015. Role of HLA adaptation in HIV evolution. Front Immunol 6:665. https://doi.org/10.3389/fimmu.2015.00665.

12. Goulder PJ, Brander C, Tang Y, Tremblay C, Colbert RA, Addo MM, Rosenberg ES, Nguyen T, Allen R, Trocha A, Altfeld M, He S, Bunce M, Funkhouser R, Pelton SI, Burchett SK, McIntosh K, Korber BT, Walker BD. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. Nature 412:334–338. https://doi.org/10.1038/35085576.

13. Cao J, McNevin J, Malhotra U, McElrath MJ. 2003. Evolution of CD8+ T cell immunity and viral escape following acute HIV-1 infection. J Immunol 171:3837–3846. https://doi.org/10.4049/jimmunol.171.7.3837.

14. Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, Desouza I, Feeney ME, Eldridge RL, Maier EL, Kaufmann DE, Lahaie MP, Reyor L, Tanzi G, Johnston MN, Brander C, Draenert R, Rockstroh JK, Jessen H, Rosen-

berg ES, Mallal SA, Walker BD. 2005. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. J Virol 79:13239–13249. https://doi.org/10.1128/JVI.79.21.13239-13249.2005.

15. Carlson JM, Le AQ, Shahid A, Brumme ZL. 2015. HIV-1 adaptation to HLA: a window into virus-host immune interactions. Trends Microbiol 23:212–224. https://doi.org/10.1016/j.tim.2014.12.008.

16. Walker BD. 2007. Elite control of HIV Infection: implications for vaccines and treatment. Top HIV Med 15:134–136.

17. Pereyra F, Heckerman D, Carlson JM, Kadie C, Soghoian DZ, Karel D, Goldenthal A, Davis OB, DeZiel CE, Lin T, Peng J, Piechocka A, Carrington M, Walker BD. 2014. HIV control is mediated in part by CD8+ T-cell targeting of specific epitopes. J Virol 88:12937–12948. https://doi.org/10.1128/JVI.01004-14.

18. McLaren PJ, Coulonges C, Ripke S, van den Berg L, Buchbinder S, Carrington M, Cossarizza A, Dalmau J, Deeks SG, Delaneau O, De Luca A, Goedert JJ, Haas D, Herbeck JT, Kathiresan S, Kirk GD, Lambotte O, Luo M, Mallal S, van Manen D, Martinez-Picado J, Meyer L, Miro JM, Mullins JI, Obel N, O'Brien SJ, Pereyra F, Plummer FA, Poli G, Qi Y, Rucart P, Sandhu MS, Shea PR, Schuitemaker H, Theodorou I, Vannberg F, Veldink J, Walker BD, Weintrob A, Winkler CA, Wolinsky S, Telenti A, Goldstein DB, de Bakker PIW, Zagury JF, Fellay J. 2013. Association study of common genetic variants and HIV-1 acquisition in 6,300 infected cases and 7,200 controls. PLoS Pathog 9:e1003515. https://doi.org/10.1371/journal.ppat.1003515.

19. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, Yuki Y, Del Prete GQ, Goulder P, Brumme ZL, Brumme CJ, John M, Mallal S, Nelson G, Bosch R, Goedert JJ, Buchbinder S, Kirk GD, Fellay J, McLaren P, Deeks SG, Pereyra F, Walker B, Michael NL, Weintrob A, Wolinsky S, Liao W, Carrington M. 2013. Influence of HLA-C expression level on HIV control. Science 340:87–91. https://doi.org/10.1126/science.1232685.

20. van Manen D, van 't Wout AB, Schuitemaker H. 2012. Genome-wide association studies on HIV susceptibility, pathogenesis and pharmacogenomics. Retrovirology 9:70. https://doi.org/10.1186/1742-4690-9-70.

21. Sunshine JE, Larsen BB, Maust B, Casey E, Deng W, Chen L, Westfall DH, Kim M, Zhao H, Ghorai S, Lanxon-Cookson E, Rolland M, Collier AC, Maenza J, Mullins JI, Frahm N. 2015. Fitness-balanced escape determines resolution of dynamic founder virus escape processes in HIV-1 infection. J Virol 89:10303–10318. https://doi.org/10.1128/JVI.01876-15.

22. Boutwell CL, Carlson JM, Lin TH, Seese A, Power KA, Peng J, Tang Y, Brumme ZL, Heckerman D, Schneidewind A, Allen TM. 2013. Frequent and variable cytotoxic-T-lymphocyte escape-associated fitness costs in the human immunodeficiency virus type 1 subtype B Gag proteins. J Virol 87:3952–3965. https://doi.org/10.1128/JVI.03233-12.

23. Schneidewind A, Brockman MA, Yang R, Adam RI, Li B, Le Gall S, Rinaldo CR, Craggs SL, Allgaier RL, Power KA, Kuntzen T, Tung CS, LaBute MX, Mueller SM, Harrer T, McMichael AJ, Goulder PJ, Aiken C, Brander C, Kelleher AD, Allen TM. 2007. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. J Virol 81:12382–12393. https://doi.org/10.1128/JVI.01543-07.

24. Martinez-Picado J, Martínez MA. 2008. HIV-1 reverse transcriptase inhib-

itor resistance mutations and fitness: a view from the clinic and ex vivo. Virus Res 134:104–123. https://doi.org/10.1016/j.virusres.2007.12.021.

25. Al-Mawsawi LQ, Wu NC, Olson CA, Shi VC, Qi H, Zheng X, Wu TT, Sun R. 2014. High-throughput profiling of point mutations across the HIV-1 genome. Retrovirology 11:124. https://doi.org/10.1186/s12977 -014-0124-6.

26. Qi H, Olson CA, Wu NC, Ke R, Loverdo C, Chu V, Truong S, Remenyi R, Chen Z, Du Y, Su SY, Al-Mawsawi LQ, Wu TT, Chen SH, Lin CY, Zhong W, Lloyd-Smith JO, Sun R. 2014. A quantitative high-resolution genetic profile rapidly identifies sequence determinants of hepatitis C viral fitness and drug sensitivity. PLoS Pathog 10:e1004064. https://doi.org/ 10.1371/journal.ppat.1004064.

27. Du Y, Wu NC, Jiang L, Zhang T, Gong D, Shu S, Wu TT, Sun R. 2016. Annotating protein functional residues by coupling high-throughput fitness profile and homologous-structure analysis. mBio 7:e0801-16. https://doi.org/10.1128/mBio.01801-16.

28. Haddox HK, Dingens AS, Bloom JD. 2016. Experimental estimation of the effects of all amino-acid mutations to HIV's envelope protein on viral replication in cell culture. PLoS Pathog 12:e1006114. https://doi.org/10 .1371/journal.ppat.1006114.

29. Wu NC, Olson CA, Du Y, Le S, Tran K, Remenyi R, Gong D, Al-Mawsawi LQ, Qi H, Wu TT, Sun R. 2015. Functional constraint profiling of a viral protein reveals discordance of evolutionary conservation and functionality. PLoS Genet 11:e1005310. https://doi.org/10.1371/journal.pgen.1005310.

30. Rihn SJ, Wilson SJ, Loman NJ, Alim M, Bakker SE, Bhella D, Gifford RJ, Rixon FJ, Bieniasz PD. 2013. Extreme genetic fragility of the HIV-1 capsid. PLoS Pathog 9:e1003461. https://doi.org/10.1371/journal.ppat.1003461.

31. Mothe B, Llano A, Ibarrondo J, Daniels M, Miranda C, Zamarreño J, Bach V, Zuniga R, Pérez-Álvarez S, Berger CT, Puertas MC, Martinez-Picado J, Rolland M, Farfan M, Szinger JJ, Hildebrand WH, Yang OO, Sanchez-Merino V, Brumme CJ, Brumme ZL, Heckerman D, Allen TM, Mullins JI, Gómez G, Goulder PJ, Walker BD, Gatell JM, Clotet B, Korber BT, Sanchez J, Brander C. 2011. Definition of the viral targets of protective HIV-1-specific T cell responses. J Transl Med 9:208. https://doi.org/10.1186/ 1479-5876-9-208.

32. Mostowy R, Kouyos RD, Hoof I, Hinkley T, Haddad M, Whitcomb JM, Petropoulos CJ, Keşmir C, Bonhoeffer S. 2012. Estimating the fitness cost of escape from hla presentation in HIV-1 protease and reverse transcriptase. PLoS Comput Biol 8:e1002525. https://doi.org/10.1371/journal.pcbi .1002525.

33. Gutteridge A, Bartlett GJ, Thornton JM. 2003. Using a neural network and spatial clustering to predict the location of active sites in enzymes. J Mol Biol 330:719–734. https://doi.org/10.1016/S0022-2836(03)00515-1.

34. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, Røder G, Peters B, Sette A, Lund O, Buus S. 2007. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. PLoS One 2:e796. https://doi.org/10 .1371/journal.pone.0000796.

35. Andreatta M, Nielsen M. 2016. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32:511–517. https://doi.org/10.1093/bioinformatics/btv639.

36. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. 2008. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. Nucleic Acids Res 36:W509–W512. https://doi.org/10.1093/nar/gkn202.

37. Gorin AM, Du Y, Liu FY, Zhang TH, Ng HL, Hofmann C, Cumberland WG, Sun R, Yang OO. 2017. HIV-1 epitopes presented by MHC class I types associated with superior immune containment of viremia have highly constrained fitness landscapes. PLoS Pathog 13:e1006541. https://doi .org/10.1371/journal.ppat.1006541.

38. Horton H, Frank I, Baydo R, Jalbert E, Penn J, Wilson S, McNevin JP, McSweyn MD, Lee D, Huang Y, De Rosa SC, McElrath MJ. 2006. Preservation of T cell proliferation restricted by protective HLA alleles is critical for immune control of HIV-1 infection. J Immunol 177:7406–7415. https://doi.org/10.4049/jimmunol.177.10.7406.

39. Los Alamos National Laboratory. 2016. Best-defined CTL/CD8+ epitope summary. Los Alamos National Laboratory, Los Alamos, NM. https:// www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary .html.

40. Wang YE, Li B, Carlson JM, Streeck H, Gladden AD, Goodman R, Schneidewind A, Power KA, Toth I, Frahm N, Alter G, Brander C, Carrington M, Walker BD, Altfeld M, Heckerman D, Allen TM. 2009. Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. J Virol 83:1845–1855. https:// doi.org/10.1128/JVI.01061-08.

41. Schellens IMM, Navis M, van Deutekom HWM, Boeser-Nunnink B, Berkhout B, Kootstra N, Miedema F, Keşmir C, Schuitemaker H, van Baarle D, Borghans JAM. 2011. Loss of HIV-1-derived cytotoxic T lymphocyte epitopes restricted by protective HLA-B alleles during the HIV-1 epidemic. AIDS 25:1691–1700. https://doi.org/10.1097/QAD.0b013e32834981b3.

42. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, Swenson LC, Tao I, Szeto S, Rosato P, Sela J, Kadie CM, Frahm N, Brander C, Haas DW, Riddler SA, Haubrich R, Walker BD, Harrigan PR, Heckerman D, Mallal S. 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. PLoS One 4:e6687. https://doi.org/ 10.1371/journal.pone.0006687.

43. Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ, Chui CKS, Cotton LA, Knapp DJHF, Riddler SA, Haubrich R, Nelson G, Pfeifer N, Deziel CE, Heckerman D, Apps R, Carrington M, Mallal S, Harrigan PR, John M, Brumme ZL; International HIV Adaptation Collaborative. 2012. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. J Virol 86:13202–13216. https://doi.org/10.1128/JVI.01998-12.

44. Autran B, Descours B, Avettand-Fenoel V, Rouzioux C. 2011. Elite controllers as a model of functional cure. Curr Opin HIV AIDS 6:181–187. https://doi.org/10.1097/COH.0b013e328345a328.

45. Okulicz JF. 2012. Elite controllers and long-term nonprogressors: models for HIV vaccine development? J AIDS Clin Res 3:1–4.

46. Gray GE, Laher F, Lazarus E, Ensoli B, Corey L. 2016. Approaches to preventative and therapeutic HIV vaccines. Curr Opin Virol 17:104–109. https://doi.org/10.1016/j.coviro.2016.02.010.

47. Streeck H. 2016. Designing optimal HIV-vaccine T-cell responses. Curr Opin HIV AIDS 11:593–600. https://doi.org/10.1097/COH.0000000000000313.

48. Rolland M, Manocheewa S, Swain JV, Lanxon-Cookson EC, Kim M, Westfall DH, Larsen BB, Gilbert PB, Mullins JI. 2013. HIV-1 conserved-element vaccines: relationship between sequence conservation and replicative capacity. J Virol 87:5461–5467. https://doi.org/10.1128/JVI.03033-12.

49. Rolland M, Nickle DC, Mullins JI. 2007. HIV-1 group M conserved elements vaccine. PLoS Pathog 3:e157. https://doi.org/10.1371/journal.ppat .0030157.

50. Allen TM, Altfeld M. 2008. Crippling HIV one mutation at a time. J Exp Med 205:1003–1007. https://doi.org/10.1084/jem.20080569.

51. Llano A, Williams A, Olvera A, Silva-Arrieta S, Brander C. 2013. Best-characterized HIV-1 CTL epitopes: the 2013 update. HIV Mol Immunol 2013:3–25.

52. Prabhakaran S, Rey M, Zagordi O, Beerenwinkel N, Roth V. 2014. HIV haplotype inference using a propagating dirichlet process mixture model. IEEE/ACM Trans Comput Biol Bioinform 11:182–191. https://doi .org/10.1109/TCBB.2013.145.