Vertical T cell immunodominance and epitope entropy determine HIV-1 escape

Michael K.P. Liu,[1] Natalie Hawkins,[2] Adam J. Ritchie,[1] Vitaly V. Ganusov,[3] Victoria Whale,[1] Simon Brackenridge,[1] Hui Li,[4] Jeffrey W. Pavlicek,[5] Fangping Cai,[5] Melissa Rose-Abrahams,[6] Florette Treurnicht,[6] Peter Hraber,[7] Catherine Riou,[8] Clive Gray,[8] Guido Ferrari,[5] Rachel Tanner,[1] Li-Hua Ping,[9,10] Jeffrey A. Anderson,[9,10,11] Ronald Swanstrom,[9,10,12] CHAVI Core B, Myron Cohen,[13] Salim S. Abdool Karim,[14] Barton Haynes,[5] Persephone Borrow,[15] Alan S. Perelson,[7,16] George M. Shaw,[4] Beatrice H. Hahn,[4] Carolyn Williamson,[6] Bette T. Korber,[7,16] Feng Gao,[5] Steve Self,[2] Andrew McMichael,[1] and Nilu Goonetilleke[1]

[1]Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom. [2]Statistical Center for HIV/AIDS Research and Prevention (SCHARP), University of Washington, Seattle, Washington, USA. [3]Department of Microbiology, University of Tennessee, Knoxville, Tennessee, USA. [4]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [5]Duke University Medical Research, Duke University, Durham, North Carolina, USA. [6]Division of Medical Virology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. [7]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA. [8]Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. [9]UNC Center for AIDS Research, [10]Lineberger Comprehensive Cancer Center, [11]Division of Infectious Diseases, School of Medicine, [12]Department of Biochemistry and Biophysics, and [13]HIV Prevention Trials Unit, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [14]Centre for the AIDS Programme of Research in South Africa, University of KwaZulu-Natal, Durban, South Africa. [15]Nuffield Department of Clinical Medicine, Division of Experimental Medicine, University of Oxford, Oxford, United Kingdom. [16]Santa Fe Institute, Santa Fe, New Mexico, USA.

HIV-1 accumulates mutations in and around reactive epitopes to escape recognition and killing by CD8$^+$ T cells. Measurements of HIV-1 time to escape should therefore provide information on which parameters are most important for T cell–mediated in vivo control of HIV-1. Primary HIV-1–specific T cell responses were fully mapped in 17 individuals, and the time to virus escape, which ranged from days to years, was measured for each epitope. While higher magnitude of an individual T cell response was associated with more rapid escape, the most significant T cell measure was its relative immunodominance measured in acute infection. This identified subject-level or "vertical" immunodominance as the primary determinant of in vivo CD8$^+$ T cell pressure in HIV-1 infection. Conversely, escape was slowed significantly by lower population variability, or entropy, of the epitope targeted. Immunodominance and epitope entropy combined to explain half of all the variability in time to escape. These data explain how CD8$^+$ T cells can exert significant and sustained HIV-1 pressure even when escape is very slow and that within an individual, the impacts of other T cell factors on HIV-1 escape should be considered in the context of immunodominance.

Introduction

HIV-1–specific CD8$^+$ T cells are central to anti–HIV-1 immunity (1–6). HIV-1–specific CD8$^+$ T cells are first detected prior to peak plasma viremia around 2 to 3 weeks after infection, expand as virus load (VL) declines, and thereafter can be detected throughout chronic infection (3). However, observations that strong HIV-1–specific T cell responses can be detected both in adults who are controlling HIV-1 well as well as in many of those progressing to AIDS (7) suggests that not all specific CD8$^+$ T cell responses are equal and that some control virus better than others.

The parameters that distinguish the more effective CD8$^+$ T cells, particularly in acute HIV-1 infection, are not clear but are likely to be influenced strongly by their epitope specificity. Comprehensive mapping of HIV-1–specific T cell responses is limited by the sequence diversity of HIV-1, with most studies in chronic infection using clade consensus peptide sets that fail to detect a significant proportion of T cell responses (8). Even so, large cross-sectional studies have found that subjects whose CD8$^+$ T cells preferentially target the structural protein Gag are more likely to sustain lower VL in chronic HIV-1 infection (9). The importance of epitope specificity to virus control is also suggested by genetic studies that found lower VL and delayed disease progression associated with the expression of certain HLA alleles, B*57/5801, B*27, B*51, and B*8101 (10–12), and the finding that protection is linked to the number of T cell responses restricted by these HLA types (13).

The specificity of the T cell response is determined by multiple parameters, including HLA, virus sequence, and T cell receptor repertoire (14–16). When more than one epitope is targeted, responding populations expand unevenly, establishing a vertical immunodominance hierarchy within a subject. Vertical T cell immunodominance patterns are highly variable among subjects and change over time, largely as a consequence of the sequence variability of HIV-1 (17–20). This differs from horizontal immunodominance, also described as immunoprevalence (21), in which the same reactive epitope is commonly targeted by individuals of a given HLA type at the population level. Horizontally immunodominant epitopes often are not vertically immunodominant (22, 23). Studies of HIV-1 infection have identified several HLA allotypes that present epitopes that are consistently horizontally

Conflict of interest: The authors have declared that no conflict of interest exists.

Note regarding evaluation of this manuscript: Manuscripts authored by scientists associated with Duke University, The University of North Carolina at Chapel Hill, Duke-NUS, and the Sanford-Burnham Medical Research Institute are handled not by members of the editorial board but rather by the science editors, who consult with selected external editors and reviewers.

Citation for this article: *J Clin Invest.* 2013;123(1):380–393. doi:10.1172/JCI65330.

immunodominant (1), with some, associated with protection (1, 24). While horizontal HIV-1–specific T cell immunodominance has been broadly investigated, until recently, there has been limited information on vertical $CD8^+$ T cell immunodominance hierarchies in HIV-1 infection (13, 20, 25). We propose that the true contribution of epitope specificity and relative immunodominance of $CD8^+$ T cells to control of HIV-1 infection can only be understood when all T cell responses are measured.

In response to $CD8^+$ T cell pressure, HIV-1 mutates to escape recognition (26–29). Virus escape begins almost as soon as $CD8^+$ T cell responses appear, coincident with acute VL decline (19, 20, 26, 29). Complete escape from primary HIV-1–specific $CD8^+$ T cells can occur in only 10 days from first detection of a T cell response (20, 30). Modeling studies, which assumed that higher rates of virus escape are representative of greater killing efficiency by HIV-1–infected cells, calculated that T cells specific for a single epitope could contribute >30% of the killing of infected CD4 T cells per day while virus escape was ongoing (20, 31).

In addition to T cell killing of virus-infected cells, another factor impacting rate of virus escape is the replicative capacity of the emerging mutant virus(es), specifically whether the fitness costs of escape could themselves slow that rate of escape (32). Evidence that T cell selected HIV-1 escape can confer fitness costs to the virus in vivo comes from the consistent reversion of escaped virus when transmitted into patients with different HLA types (33, 34). Documented examples of reversion are limited to a few well-described epitopes, but in vitro studies of particular epitopes have repeatedly shown lower replication following introduction of HIV-1 T cell escape mutations into chimeric or full-length viruses (35–38). Both types of study have focused on epitopes presented by protective HLA, with limited fitness cost information available on the vast majority of epitopes.

Recently, in 5 subjects we observed a correlation between the rank order of escape within patients and the Shannon entropy of the epitope (39). Shannon entropy (40) can be used to measure population diversity at a given sequence position or within a cluster of bases/amino acids, such as an epitope seen by $CD8^+$ T cells. HIV-1 is highly variable at the population level, but some regions of the virus are more conserved (41). Low entropy values correspond to sites that are more conserved (42), and HIV-1 T cell epitopes with lower average entropy escaped more slowly (39). These observations suggested that entropy, which can be calculated for any epitope, can be used as a surrogate measure of the fitness cost of escape.

Here, in a larger cohort, we used epitope entropy in combination with measurements of T cell selection pressure to show that $CD8^+$ T cells can exert significant in vivo pressure on HIV-1, even when the subsequent virus escape is very slow. We examined the complete T cell response to HIV-1 in 17 patients, using peptides matched to both the autologous founder virus sequence and subsequent evolutionary variants. This enabled us to define the vertical immunodominance profiles in each patient and to show clear changes in patterns of immunodominance during the first several months of infection, clearly related to evolutionary pressure on the virus sequence. We found, in multivariate analysis, that the time taken for HIV-1 to escape each of the first T cell responses is determined largely by the combination of the relative immunodominance of the T cell response and the fitness constraints on variability within the epitope (relative risk [RR] = 6.6, $P = 6.4 \times 10^{-8}$). The measure of immunodominance alone explained 39% of the variation in time to virus escape. This highly significant risk ratio means that the role of other factors that include polyfunctionality and T cell exhaustion in influencing escape rate are unlikely to work independently of immunodominance. This study gives insight into how pressure is exerted by $CD8^+$ T cells on HIV-1 and why it changes rapidly during early infection.

Results

Shifts in T cell immunodominance hierarchies are observed early in HIV-1 infection. We studied a cohort of 17 patients infected acutely with HIV-1. Five subjects, four with protective HLA types (B*57, 5801, and 81), exhibited early virus control, with mean VL <5,000 vRNA copies per ml for up to 2 years following infection (Figures 1, 2, and 3, and Table 1). The others had more typical VL set points, ranging from 5,811 to 575,440 vRNA copies per ml. One subject, CH607, started on antiretroviral therapy (ART) within 2 months of infection; the others did not receive ART during this study, in line with local medical guidelines. The transmitted/founder (T/F) virus in each subject was deduced following single genome amplification (SGA) and sequencing of plasma virus taken in acute infection close to peak viremia (43). Fifteen subjects were infected with a single T/F virus (data not shown). Two subjects, CH159 and CH607, were each infected with 2 very closely related viruses, differing by 2 and 13 nucleotides, respectively. SGA sequencing of virus was repeated in most subjects at frequent intervals in the first 3 months of infection and then less frequently in the ART naive subjects over 1 to 3 years (Figures 1–3). This revealed sites, or clusters of sites, in which nonsynonymous (ns) mutations were selected by immune responses in each subject (Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/JCI65330DS1).

Overlapping peptides matched to the full proteome of the T/F virus(es) in each subject were synthesized, and T cell responses were mapped using ex vivo IFN-γ ELISpot assays. Acute anti–HIV-1 T cell responses were therefore measured without bias to known epitopes; indeed only 43% of all T cell responses mapped to previously described or predicted epitopes (Figures 1–3 and Supplemental Table 1). Our previous studies have found that nearly all T cell responses in patients with acute HIV-1 infection measured in this way were mediated by $CD8^+$ T cells (20, 39).

T cell responses were first detected around peak viremia and were typically narrow (median of 3 epitopes, range 1–11). The magnitude of individual T cell responses measured by ELISpot was used to rank T responses and define immunodominance hierarchies within subjects. Different hierarchies were observed, coincident with VL decline from peak viremia (Figures 1–3). The dominant or most highly ranked T cell responses were mostly focused on the Nef, Env, and Gag proteins while Pol was targeted infrequently (Figure 4A) (44). The strong, early Nef-specific T cell responses (Figure 4B) were consistent with previous observations (25, 45). In as little as a week, shifts in T cell immunodominance hierarchies were observed, often following rapid decline of the first T cell responses as virus escaped (Figures 1–3). By 6 months following infection, additional HIV-1–specific T cell responses had emerged in most subjects (16 ART naive subjects: median number of epitopes targeted at 6 months = 6, range 2–16) (17–20). In the 7 subjects expressing HLA B*57/*5801 or B*8101, T cell responses restricted by these protective HLA alleles were dominant within weeks of infection but were often not the first dominant response (Figures 1–3 and Supplemental Table 1).

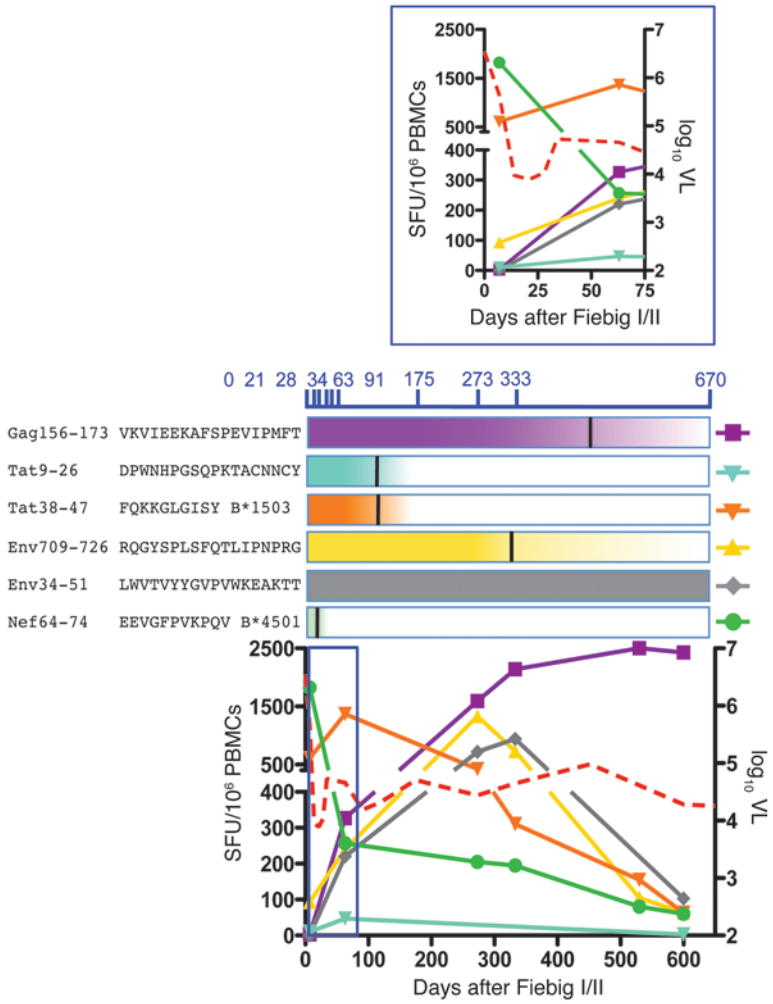

Figure 1
Shifting immunodominance of CD8+ T cell responses induced in early HIV-1 infection. Subject CH131 was detected in acute HIV-1 infection (Fiebig stages I/II), just prior to peak viremia. HIV-1–specific CD8+ T cell responses in PBMCs were mapped using ex vivo IFN-γ ELISpot against autologous, overlapping peptides spanning the subject's T/F virus (solid colored lines). Distinct CD8+ T cell immunodominance was observed as early as peak viremia (red dotted line). See inset for greater detail. Above the panel, the blue text indicates the days from Fiebig stages I/II when SGA sequencing was performed. Sequence at day 0 was used to deduce the T/F and subsequent sequencing used to track emerging ns mutations in the virus population. Below this text, colored bars represent the changes in the epitope sequence of the corresponding (color-matched) T cell response, with black vertical lines showing days to 50% T/F virus sequence. As the virus escapes, T cell responses often decline, after which new or subdominant T cell responses become immunodominant. The HXB2 location of each reactive peptide epitope and, if known, the HLA restriction is identified to the left of the bar. A positive T cell response required a background subtracted response of >30 SFU per million PBMCs and >4× background.

HIV-1 escape from CD8+ T cells occurs at very different rates – from days to years. We focused on HIV-1 escape from the first T cell responses detected within 50 days of estimated time Fiebig stages I/II. Fiebig categories, which range from Fiebig stage I–VI, classify stages of acute HIV-1 infection in accordance with reactivity in diagnostic assays (46). Fiebig stages I/II, when subjects are HIV-1 vRNA positive but seronegative, broadly corresponds to acute VL ramp up shortly before peak viremia is observed (47). The following 50-day phase occurs coincident with acute plasma VL decline and early onset of VL set point (ref. 3 and Figures 1–3). This window is likely to be crucial in determining subsequent VL control. Frequent blood sampling in this period enabled detailed measurement of T cell responses to T/F and mutant viruses, T cell immunodominance hierarchies, and virus sequence changes. We would note that while high VLs were measured in the first 50 days from Fiebig stages I/II (Figures 1–3), our weekly sampling schedule was insufficient to track the rapidly changing acute VL kinetics and confidently calculate peak viremia in our subjects. Similarly, while acute VL decline was observed in most subjects (Figures 1–3), without a peak viremia measurement, we were unable to calculate VL slopes. Our analyses were therefore focused on understanding how HIV-1–specific T cell responses in this critical acute stage of HIV-1 infection exert antiviral pressure and select escape mutants.

Virus escape from CD8+ T cells was first identified by ns changes in reactive T cell epitopes over time. In most cases, escape was then examined experimentally by comparing peptides containing the mutant amino acid sequences with their corresponding reactive T/F peptide. T cell recognition of the mutant peptides was generally impaired compared with peptides matching the T/F sequence (Supplemental Table 1). Also, 3 examples were found of mutations adjacent to an invariant T cell epitope that coincided with loss of the T cell response; these mutations probably impaired antigen processing (AgP) (ref. 48 and Supplemental Table 1). Limited blood volumes precluded experimental confirmation of these putative AgP escapes, so subsequent analyses are presented with and without their inclusion.

Extending earlier results in 3 of these subjects (20), early rapid escape from the first CD8+ T cells induced following HIV-1 infection was observed in one or more epitopes in all subjects (Supplemental Figure 1). Overall, in 15 subjects, T cell escape occurred in the first 50 days from Fiebig stages I/II and in 2 subjects within the first 70 days (Supplemental Table 1). In contrast with previous observations (49), rapid virus escape was not more common in epitopes restricted by protective HLA allotypes (Figures 1–3 and Supplemental Table 1). Our observations showed clearly that when comprehensive T cell mapping is performed, rapid HIV-1 escape was measured in T cell epitopes restricted by a broad range of HLA

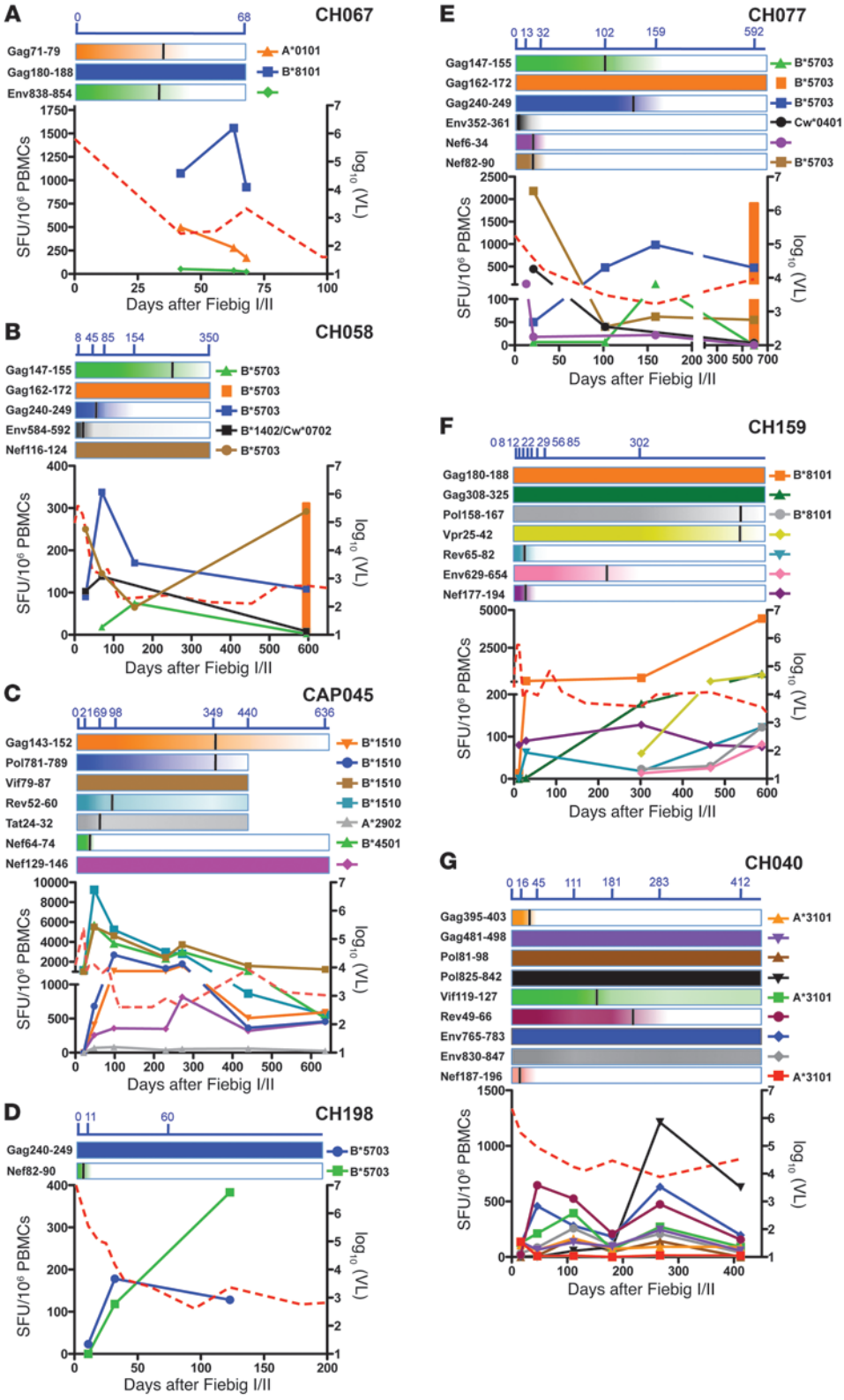

Figure 2
HIV-1–specific CD8 T cell kinetics. T cell kinetics in subjects (**A**) CH067, (**B**) CH058, (**C**) CAP045, (**D**) CH198, (**E**) CH077, (**F**) CH159, and (**G**) CH040 with VL set points from 40–14,791 copies/ml (red dotted line). HIV-1–specific CD8+ T cell responses in PBMCs were mapped using ex vivo IFN-γ ELISpot against autologous, overlapping peptides spanning the subject's T/F virus (solid colored lines). Above each panel, the blue text indicates the days from Fiebig stages I/II when SGA sequencing was performed. Below this text, colored bars represent the changes in the epitope sequence of the corresponding (color-matched) T cell response, with black vertical lines showing days to 50% T/F virus sequence. The HXB2 location of each reactive peptide is identified to the left of the bar, and HLA restriction (if known) is identified to the right. A positive T cell response required a background subtracted response of >30 SFU per million PBMCs and >4× background.

allotypes, not limited to HLA allotypes associated with lower VLs. Mathematical modeling (20) showed that, within each subject, individual T cell responses killed up to 50% of infected CD4+ T cell targets per day (Supplemental Table 1). By shortening the life span of productively infected cells in vivo, acutely responding CD8+ T cells exerted substantial antiviral pressure in all patients studied.

In addition to T cells selecting rapid escape, there were an equal number of T cell responses specific for epitopes that varied very little during the first 50 days of observation (Figure 4C and Supplemental Figure 1). However, most of these epitopes escaped in the first year following infection, 70% within the first 6 months (Figure 4C). Only 14% of T epitopes remained invariant throughout the whole period of observation (sequencing 350–1,187 days from Fiebig stages I/II). Therefore, the large majority of HIV-1 epitopes recognized by CD8+ T cells escape, but, within the same patient, the time taken ranges from days to years (Figure 4C). Estimates, using existing mathematical models (31), of killing by HIV-1–specific CD8+ T cells selecting slow escape were much lower than for the rapid escapes (Supplemental Table 1). Among the slow-escaping epitopes, several were restricted by protective HLA alleles (Supplemental Table 1). Given the strong association of these HLA types with lower VL, we would predict that these T cell responses exert

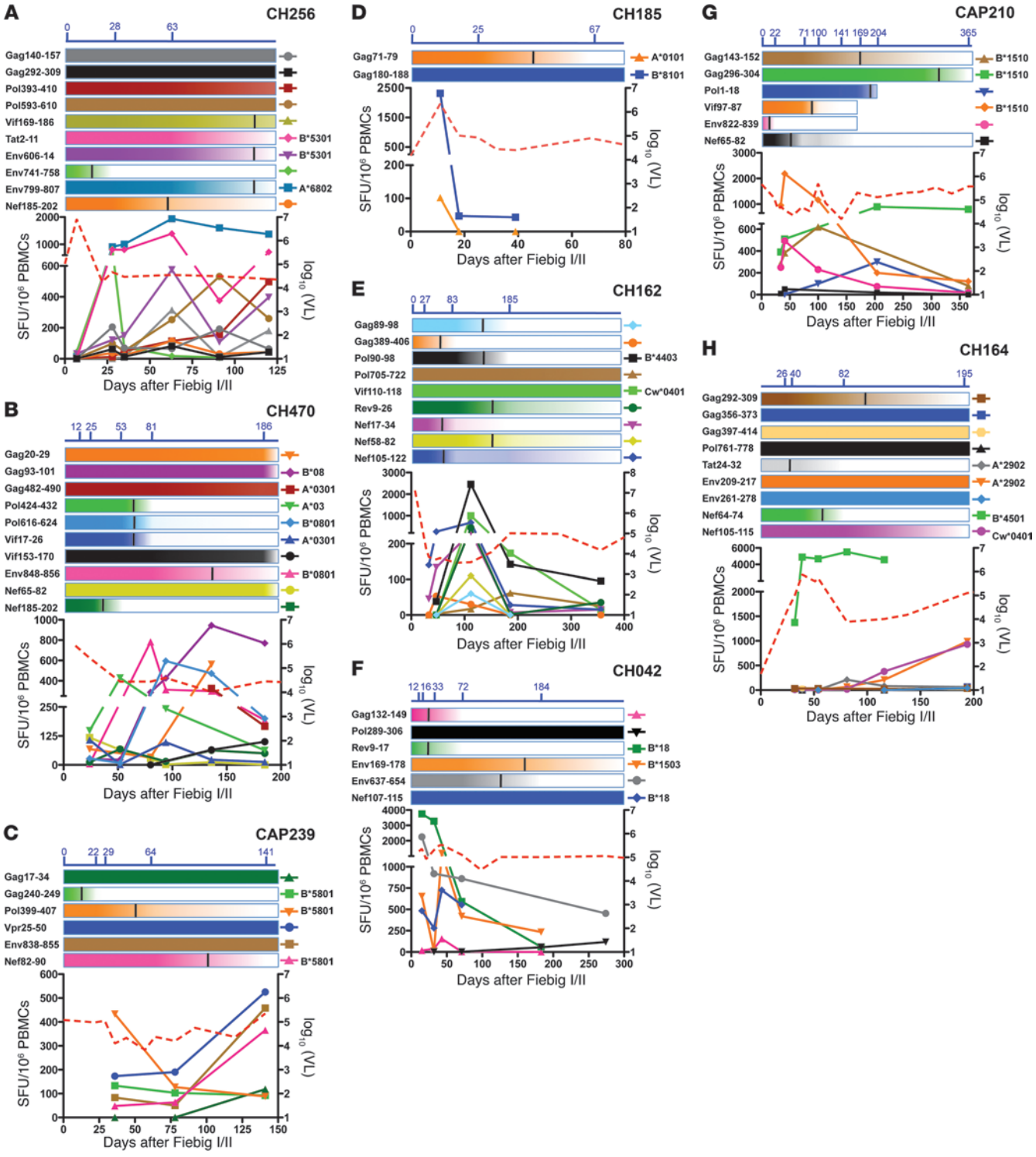

Figure 3

HIV-1–specific CD8 T cell kinetics continued. T cell kinetics in subjects (**A**) CH256, (**B**) CH470, (**C**) CAP239, (**D**) CH185, (**E**) CH162, (**F**) CH042, (**G**) CAP210, and (**H**) CH164, with VL set points from 18,621–575,440 copies per ml (red dotted lines). HIV-1–specific $CD8^+$ T cell responses in PBMCs were mapped using ex vivo IFN-γ ELISpot against autologous, overlapping peptides spanning the subject's T/F virus (solid colored lines). Above each panel, the blue text indicates the days from Fiebig stages I/II when SGA sequencing was performed. Below this text, colored bars represent the changes in the epitope sequence of the corresponding (color-matched) T cell response, with black vertical lines showing days to 50% T/F virus sequence. The HXB2 location of each reactive peptide is identified to the left of the bar and HLA restriction (if known) is identified to the right. A positive T cell response required a background subtracted response of >30 SFU per million PBMCs and more than 4 times background. Subject CH607 went onto ARV approximately 5 weeks after Fiebig stages I/II. Therefore, T cell kinetics were not followed.

Table 1
Demographic and clinical data from study subjects

Subject identifier	Gender (M/F)	Country[A]	Age at first positive HIV-1 vRNA (yr)	HIV-1 subtype	Median CD4 count[B]	Virus set point[C] (copies vRNA/ml) (cells/μl)	HLA type
CH067	F	S. Africa	32	C	1,557	40	A*01:01,A*33:01, B*53:01,**B*81**:01, Cw*04:01,Cw*18:01
CH058	M	USA	23	B	861	234	A*01:01,A*23:01, B*14:02,**B*57**:01-04, Cw*07:01,Cw*08:02
CAP045	F	S. Africa	41	C	1,046	871	A*23:01,04, A*29:02,03, B*15:10,B*45:01, Cw*16:01,Cw*16:02
CH198	M	S. Africa	33	C	894	1,175	A*03:01,A*74:01, B*08:01,**B*57**:01-04, Cw*07:01,Cw*07:02
CH077	M	USA	23	B	756	3,631	A*02:05,A*02:05, B*53:01,**B*57**:01-04, Cw*04:01,Cw*18:01
CH159	F	Malawi	18	C	440	8,511	A*23:01,A*6802, **B*81**:01,B*18:01, Cw*07:04,Cw*18:01
CH040	M	USA	56	B	986	14,791	A*31:01,A*02:01, B*44:02,B*40:01, Cw*03:02,Cw*05:01
CH256	F	Malawi	19	C	601	18,621	A*33:01,A*68:01, B*14:01,B*53:01, Cw*04:01,Cw*08:02
CH131	M	Malawi	20	C	300	22,909	A*29:01-04,A*23:01, B*45:01,B*15:03, Cw*02:02,Cw*06:02
CH470	M	USA	17	B	470	23,710	A*03:01,A*01:01, B*08:01,B*49:01, Cw*07:01,Cw*07:01
CAP239	F	S. Africa	44	C	792	26,915	A*01:01,A*29:02,03, B*42:01,**B*58:01**, Cw*06:02,Cw*07:02,03
CH185	F	S. Africa	20	C	309	40,738	A*01:01,A*74:01, B*15:03,**B*81**:01, Cw*02:02,Cw*18:01
CH162	M	S. Africa	42	C	376	114,815	A*34:02,A*30:01, B*44:03,B*42:02, Cw*04:01,Cw*17:01
CH042	M	S. Africa	38	C	349	128,825	A*23:01,A*30:02, B*15:03,B*18:01, Cw*02:02,Cw*07:04
CAP210	F	S. Africa	43	C	306	295,121	A*68:02,A*68:02, B*15:10,B*15:10, Cw*03:04,Cw*03:04
CH164	M	S. Africa	27	C	447	575,440	A*02:01,A*29:01-04, B*44:03,B*45:01, Cw*07:01,Cw*16:01
CH607	M	USA	18	B	984[D]	–[D]	A*03:01,A*01:01, B*08:01,B*49:01, Cw*07:01,Cw*07:01

[A]All African subjects are black Africans. The USA subjects, CH040, CH470, and CH607, are of mixed European descent, and subjects CH058 and CH077 are African American. [B]CD4 count calculated from available time points (minimum 2) from visits between 6 months to 1 to 3 years following subject screening. [C]HIV-1 viral load set point calculated as described in Fellay et al. (2). Briefly, data were manually inspected, and outliers and time points in either the acute stage of infection or following progression were excluded. Set point was the average of all remaining results. [D]Subject CH607 received ART from approximately 5 weeks from Fiebig stages I/II. Bolded underlined text identifies HLA alleles associated with lower VL set points and or delayed progression to HIV-1–related disease. F, female; M, male; S. Africa, South Africa.

significant in vivo immune pressure. In the light of our previous observation of little functional variability in acutely responding T cells, irrespective of rate of virus escape (39), this suggests that slow escape does not equate to poor ongoing immune pressure (32). We therefore asked whether other factors impacted time to escape.

T cell response magnitude, immunodominance rank, and breadth are significantly associated with time to virus escape. We examined the parameters that contribute to time to virus escape, anticipating that this approach would give us a measure of the antiviral pressure exerted by $CD8^+$ T cells. Given the contribution of T cell immunodominance hierarchies in antiviral immunity in mouse models (14), we first sought correlations between T cell magnitude and relative immunodominance and time to virus escape in each epitope (see Methods for definition of time to escape). Data were examined in terms of "survival" of the founder T cell epitope, fitting Cox proportional hazards (CPH) models to calculate RR of escape in nonstratified and subject-stratified data. Significance tests examined the null hypothesis that the covariate being tested had no impact on escape, i.e., RR = 1.

RR calculations found that increases in the $\log_{10}$ of the absolute magnitude (M) of individual T cell responses significantly increased the risk of virus escape (RR = 3.3, $P = 0.03$; Table 2), consistent with our previous rank order observations in a smaller data set (39). Our larger and more detailed data sets enabled us to express data for each patient as a percentage magnitude (%M) of the total T cell response measured by IFN-γ ELISpot at the given time point. %M

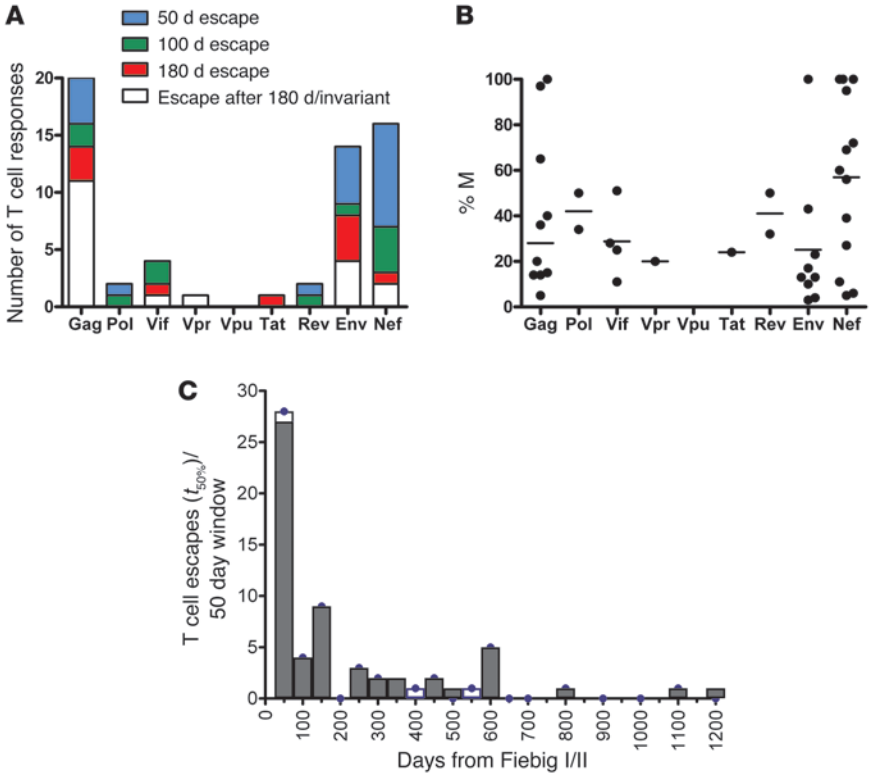

Figure 4
The first HIV-1–specific CD8+ T cells target all HIV-1 proteins but are largely immunodominant against Nef, Gag, and Envelope. (**A**) The total number and the proportion of mapped T cell epitopes found in each HIV-1 protein that underwent escape in the first 50 (blue), 100 (green), and 180 (red) days from Fiebig stages I/II. (**B**) The magnitude of each mapped T cell response was calculated as a percentage of the total response measured in that individual. Horizontal bars represent median values; each symbol represents an individual response. (**C**) Histogram binning time to escape for each epitope across subjects in 50-day windows. The white blocks in columns indicate data adjusted to include putative AgP mutations (n = 3). Note that the median %M in **B** was higher for Nef than other proteins but was not significantly different than Gag, Vif, and Env proteins (P > 0.05, 1-way ANOVA Friedman test).

was our operational measure of immunodominance, providing both the proportional contribution of each T cell response and its rank in its immunodominance hierarchy in acute HIV-1 infection. Importantly, calculation of %M in this study enabled comparisons across all subjects not examined in our previous study (39). The impact of %M of T cell responses measured in acute infection on non-subject stratified data were highly significant (RR = 3.6, $P = 5.5 \times 10^{-5}$; Table 2). The correlation improved for both M and %M when probable AgP mutations were included (%M: RR = 5, $P = 7.5 \times 10^{-7}$; Table 2). Data resolved into tertiles are shown in survival plots in Figure 5, A and B. These plots illustrate that the greatest effect of M and %M on time to escape came from data falling into the highest tertile, i.e., the highest magnitude (>320 SFU/10^6 PBMCs) and the most dominant T cell responses (>33%). The greater the M and the higher the position of the T cell response in the immunodominance hierarchy, the more rapidly virus escape was selected.

The breadth of the HIV-1–specific T cell response measured in acute infection also had a significant effect on RR of virus escape (RR = 0.85, P = 0.0027). Given that breadth is a subject-level measure, RR analyses, again using CPH tests, were limited to non-subject stratified data (Table 2 and Figure 5C). Consistent with previous predictions from mathematical modeling (31), higher T cell breadth in acute HIV-1 infection (n > 3) correlated with a slower rate of escape from individual T cell responses within a subject.

Shannon entropy of the epitope influences the rate of virus escape. In our data set, some epitopes (e.g., B*81-restricted Gag 180-188 TL9 in subjects, CH067, CH159, and CH185; Figure 2, A and F; Figure 3D; and Supplemental Table 1), despite becoming immunodominant very early in infection, selected escape slowly over months to years. This suggested that while overall dominant T cell responses were the most effective at recognizing and killing virus-infected cells, virus-intrinsic factors were retarding the ability of mutant viruses to emerge. In this much larger data set than our previous analysis (39), we again used Shannon entropy (40) as a measure of the variability of each epitope at the population level to assess associations with time to escape. Our mapping studies had identified T cell epitopes to the level of a reactive 18-mer peptide. In more than a half of cases, we went further and experimentally identified the optimal T cell epitope (Figures 1–3 and Supplemental Table 1). We initially limited studies of entropy and time to escape to optimal epitopes. A significant correlation between Shannon entropy values for the optimal epitopes and time to escape was observed in CPH models; the more conserved the epitope at a population level, the slower the escape (n = 34, RR = 6.6, P = 0.017). Notably, mapped epitopes restricted by protective HLA alleles B*57/5801 and B*8101 had lower entropies compared with epitopes restricted by HLA alleles not associated with protective effects on VL (P = 0.03, Mann-Whitney 1 tailed; Supplemental Table 1 and Supplemental Figure 2).

The analysis was then expanded by calculating entropies for reactive peptides in which the optimal 8- to 11-mers T cell epitope had not been experimentally defined (see Methods and Supplemental Figure 3). Analysis of the larger data set found similarly that increased $\log_{10}$-transformed entropy scores increased the rate of virus escape (n = 59, RR = 4, P = 0.0099; Table 2). Tertile analysis showed that the lowest entropy tertile, i.e., the most conserved epitopes (≤0.194 nats), escaped markedly more slowly (Figure 5D and Supplemental Table 2).

Multivariate analyses — T cell pressure and epitope entropy are independent predictors of escape. The univariate analyses above showed that T cell immunodominance (%M), M, breadth of the T cell response, and epitope entropy were all significantly

Table 2
Univariate tests[A] reporting RR of T cell response covariates, M, %M, T cell breadth, and entropy on time to T cell escape[B] in epitopes[C] with and without putative AgP[D]

Cov	Analysis	TTE, unstratified	TTE, subject stratified	TTE + AgP, unstratified	TTE + AgP, subject stratified
log_{10}(M)	RR/cov *P*	1.8/0.04	3.3/0.04	1.9/0.01	4.9/0.009
	(RR 95% CI)	(1, 3)	(1.1, 10)	(1.2, 3.3)	(1.5, 16)
	Model *P* value	**0.042**	**0.03**	**0.015**	**0.0052**
	n all, *n* events[E]	59, 49	59, 49	59, 52	59, 52
log_{10}(%M)	RR/cov *P*	$3.6/1 \times 10^{-4}$	3.2/0.04	$5/4 \times 10^{-6}$	4/0.02
	(RR 95% CI)	(1.9, 6.8)	(1, 9.8)	(2.5, 9, 7)	(1, 3, 13)
	Model *P* value	$\mathbf{5.5 \times 10^{-5}}$	**0.036**	$\mathbf{7.5 \times 10^{-7}}$	**0.012**
	n all, *n* events	57, 47	57, 47	57, 50	57, 50
Breadth[F]	RR/cov *P*	0.87	–	0.85	–
	(RR 95% CI)	(0.77, 0.98)		(0.76, 0.96)	
	Model *P* value	**0.01**		**0.0027**	
	n all, *n* events	59, 49		59, 52	
log_{10}(entropy)	RR/cov *P*	4/0.01	13/0.005	3.4/0.02	12/0.005
	(RR 95% CI)	(1.4, 12)	(2.2, 75)	(1.2, 9.5)	(2.1, 69)
	Model *P* value	**0.0099**	**0.0024**	**0.019**	**0.0025**
	n all, *n* events	59, 49	59, 49	59, 52	59, 52

[A]Cox proportional hazards (CPH) tests. [B]Time to escape (TTE) measured in days from first detection of T cell response (detected in first 50 days from Fiebig stages I/II) to time to ≤50% transmitted virus within epitope. [C]17 subjects; total epitopes analyzed = 59. [D]Time to escape measured in days from first detection of T cell response to time to ≤50% transmitted virus, where mutation emerges within proximity of T cell epitope, occurring coincident with decline of T cell response. [E]"*n* all" refers to the total number of reactive epitopes. "*n* events" refers to the difference in total number of reactive epitopes and invariant epitopes (see Methods) over study window. *n* events = 49 increasing to 52 when AgP mutations were considered. [F]Breadth is a subject level measure. cov, covariates. Bold text highlights model *P* values. RR, relative risk; M, magnitude; %M, percentage magnitude; AgP, putative antigen processing mutation.

correlated with the time to virus escape in acute HIV-1. Other subject-specific parameters were also tested but did not impact on the RR of virus escape; these included subject, gender, and age; continent of origin; infecting HIV-1 virus clade; and the virus set point (Supplemental Tables 3 and 4). Epitope-specific parameters were also analyzed. Neither expression of a protective HLA allele (B*57, B*58, B*81) nor Gag specificity of the T cell responses independently contributed to time to escape (Supplemental Tables 3 and 4).

Our large comprehensive data set provided us with a unique opportunity to next apply multivariate analysis to define the relative contributions of M, %M, T cell breadth, and entropy to time to escape. This approach produced a RR (and *P* value) for each covariate and then assessed whether combining covariates produced a better "fit" when explaining time to escape. Two basic models were developed in which either M or %M were added to a model containing entropy. Both M and %M significantly improved models of entropy (Table 3), and the overall *P* values of these models were highly significant (Table 3, bold text). The best fit was observed in the model of %M and entropy; the RR of each covariate was 6.6, producing a combined model *P* value of 6.4×10^{-8} in non-subject stratified data. We next accommodated extra variation in time to escape using multiple imputation methods (Supplemental Table 5). The results were the same as those from CPH models, suggesting that variation due to periodic sampling and differences in the number of SGA sequences available at different visits did not explain the associations observed. We also explored whether the effects between covariates were more than additive. No interaction was detected between M or %M and epitope entropy, consistent with the hypothesis that the T cells exert pressure on the T/F virus, whereas epitope entropy relates more to the rate at which viral escape mutants emerge (Table 3).

Two additional more complex models, adding %M or T cell breadth to a model containing M and entropy, were developed. Adding %M as a covariate significantly improved the non-subject stratified models of M and entropy ($P < 0.0001$; Table 3) without any interaction between %M and M ($P > 0.05$; Table 3). This observation was important as it showed that the position of a T cell response in the immunodominance hierarchy is independent of its M and independently selects for more rapid virus escape, advancing our previous observations (39). In the next model, T cell breadth also significantly improved a model of M and entropy ($P = 0.02$; Table 3), with increased breadth slowing the emergence of viral escape mutants within a subject. As could be predicted, subjects whose first HIV-1 T cell responses included a highly dominant (higher %M) T cell response had an overall lower total T cell breadth ($r = -0.59$, $P = 0.01$, Spearman 2 tailed). Given that these 2 parameters were related, T cell breadth did not significantly add to multivariate models containing %M (data not shown). In summary, after rigorous statistical testing, %M and epitope entropy were identified as the major in vivo factors determining time to escape in acute HIV-1 infection.

Last, we calculated a coefficient of explained randomness (50) of the time to virus escape, which is analogous to explained variation (r^2) for linear regression models. We observed that nearly half of the variation in time to escape was accounted for by just

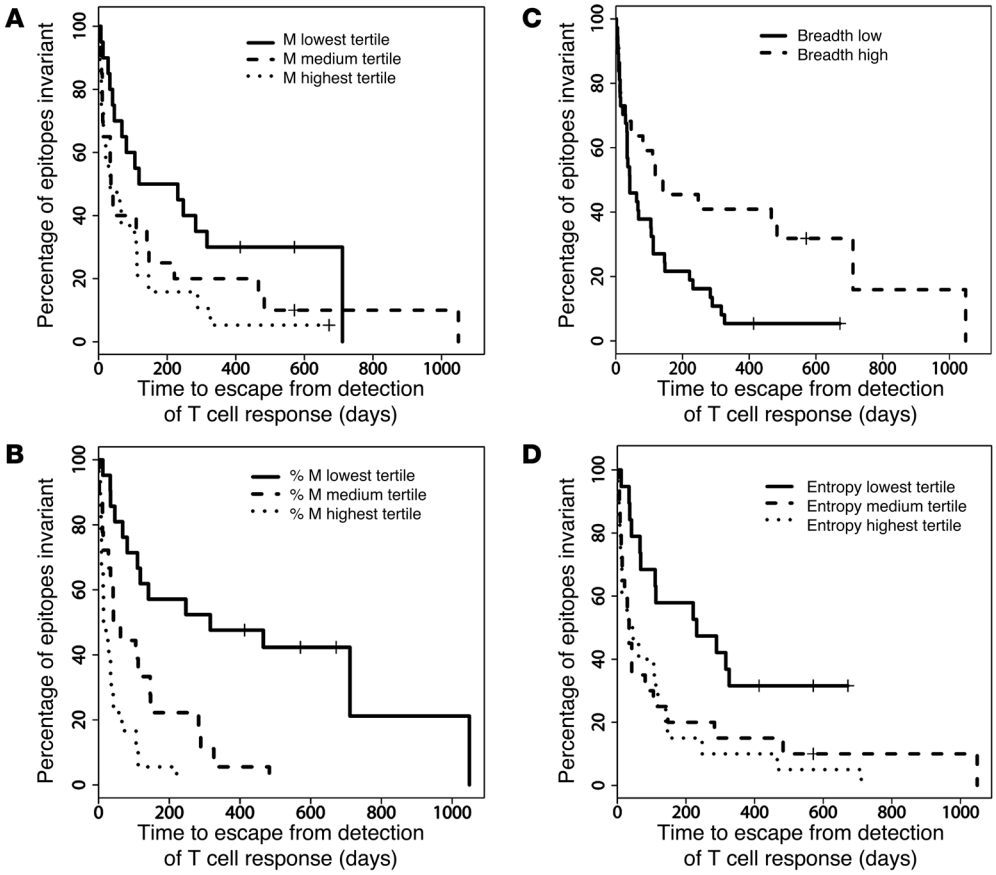

Figure 5
M, immunodominance (%M), breadth, and the epitope entropy of the HIV-1–specific T cell response all contribute to time to escape. Univariate correlations of time to escape with (**A**) magnitude (M), (**B**) %M of the T cell response, and (**C**) T cell breadth and (**D**) entropy of the epitope targeted by the T cell response. T cell data are restricted to the first HIV-1–specific T cell responses detected using IFN-γ ELISpot detected within 50 days of Fiebig stages I/II. Time to virus escape was calculated from the time of first detection of the T cell response and time to 50% change in SGA sequences compared with the T/F sequence. Data were resolved into tertiles in **A**, **B**, and **D** and median in **C**. See Supplemental Table 2 for risk ratios and *P* values using CPH models.

2 factors, %M and epitope entropy (Table 4). Individually, %M and entropy informed 39% and 10% of the variation, respectively. Notably, as discussed in univariate analysis above, neither protective HLA allele nor Gag restriction explained any of the variation in time to virus escape.

Epitope entropy is an incomplete surrogate for virus fitness. While entropy correlated highly significantly with time to escape, independently it informed only approximately 10% of the variation in time to escape (Table 4). This was surprising given that virus fitness should be a major determinant for the ability of mutants to emerge. We wondered whether the contribution of entropy to virus escape represented a lower estimate for in vivo fitness impact. We examined patient CAP239, who generated a subdominant HLA B*5801-restricted T cell response that targeted the classical Gag 240–249 epitope (TW10), which has a low entropy (lowest tertile in our data set, see Supplemental Table 1). Escape within this epitope occurred within days (usual escape mutant, 242 threonine to asparagine mutation [T242N]), compared with slower escape in this same epitope observed in other subjects (CH058, CH077; Supplemental Table 1). Neither the relative immunodominance of the TW10 T cell response in patient CAP239 (15% CAP239 vs. 17% CH058) nor the M of response was higher than in subject CH058, which could have explained more rapid escape (Supplemental Table 1). Further examination of the T/F virus in this subject showed that it already contained upstream Gag mutations, H219Q and I223V (Supplemental Table 1). These mutations are known to occur more frequently in B*57/58 progressors than in nonprogressors and to partially compensate for the fitness cost of the T242N escape within the TW10 epitope (51). This likely facilitated the more rapid escape observed in this patient, rendering this T cell response less effective than in other B*57/5801 patients; none of whom had these compensatory mutations in their T/F virus (Supplemental Table 1). This exception suggests that epitope entropy is a partial surrogate for fitness and emphasizes the strength of the relationship between "true" in vivo fitness and time to virus escape.

Discussion

Vertical T cell immunodominance and epitope entropy combine to explain nearly half of the variation in the time to HIV-1 escape from primary $CD8^+$ T cells. This means that within an individual, dominant HIV-1–specific T cells must kill or suppress virus-infected cells more efficiently than subdominant T cells, but the emergence of mutants in some regions of the HIV-1 proteome is constrained, significantly slowing the rate of escape despite sustained T cell pressure. Importantly, these observations were made in acute infection in individuals infected with a single T/F virus when HIV-1–specific $CD8^+$ T cells are capable of recognizing and killing all circulating HIV-1–infected cells. Note that these observations also hold if nonlytic mechanisms of selection pressure are attributed to HIV-1–specific $CD8^+$ T cells (52). Combined with our previous observations showing no difference in polyfunctional profiles of dominant versus subdominant T cells (39), our data suggest that $CD8^+$ T cells can sustain strong T cell pressure in vivo even when the result is slow to no virus escape.

Table 3
Multivariate analysis[A] reporting RR of T cell response covariates M and %M and entropy on time to T cell escape[B] of epitopes[C] with and without putative AgP[D]

Model	Cov for each model	Analysis	TTE, unstratified	TTE, subject stratified	TTE + AgP, unstratified	TTE + AgP, subject stratified
M/entropy	log_{10}(M)	RR/cov *P* (RR 95% CI)	1.8/0.02 (1.1, 3.1)	4.9/0.01 (1.4, 17)	2/0.007 (1.2, 3.3)	9.1/0.003 (2.1, 39)
	log_{10}(entropy)	RR/cov *P* (RR 95% CI) Model *P* value *P* value adding M to entropy Interaction term[E] M & entropy	4.4/0.006 (1.5, 13) **0.003** 0.03 0.80	22/0.003 (2.9, 171) **3×10^{-4}** 0.01 0.30	3.8/0.01 (1.3, 11) **0.0021** 0.01 0.65	32/0.002 (3.6, 288) **2.8×10^{-5}** 0.00 0.43
%M/entropy	log_{10}(%M)	RR/cov *P* (RR 95% CI)	5.01/0.00004 (2.32, 10.81)	4.7/0.02 (1.3, 17)	$6.6/2 \times 10^{-6}$ (3, 15)	6.8/0.007 (1.7, 27)
	log_{10}(entropy)	RR/cov *P* (RR 95% CI) Model *P* value *P* value adding %M to entropy Interaction to %M & entropy	8/18/0.002 (2.22–30.17) **1.7×10^{-6}** 0.001 0.38	22/0.004 (2.6, 187) **0.00057** 0.01 0.12	6.6/0.004 (1.8, 24) **6.4×10^{-8}** <0.001 0.72	25/0.004 (2.8, 232) **0.00018** 0.002 0.11
M/entropy/%M	log_{10}(M)	RR/cov *P* (RR 95% CI)	0.96/0.9 (0.54, 1.7)	–[F]	0.94, 0.8 (0.53, 1.6)	–
	log_{10}(entropy)	RR/cov *P* (RR 95% CI)	8.2/0.002 (2.2, 30)		6.5/0.004 (1.8, 2.4)	
	log_{10}(%M)	RR/cov *P* (RR 95% CI) Model *P* value *P* value adding %M to M/entropy Interaction term M & %M	$5.1/1 \times 10^{-4}$ (2.2, 12) **7×10^{-6}** 0.00009 0.08		$6.8/8 \times 10^{-6}$ (3, 16) **2.9×10^{-7}** <0.00001 0.15	
M/entropy/breadth	log_{10}(M)	RR/cov *P* (RR 95% CI)	1.5/0.1 (0.88, 2.6)	–[G]	1.6/0.08 (0.95, 2.7)	–
	log_{10}(entropy)	RR/cov *P* (RR 95% CI)	5/0.005 (1.6, 15)		4.1/0.01 (1.4, 12)	
	Breadth	RR/cov *P* (RR 95% CI) Model *P* value *P* value adding breadth to M/entropy	0.87/0.05 (0.76, 1) **0.0011** 0.036		0.86/0.03 (0.76, 0.99) **0.00048** 0.019	

[A]CPH test. [B]Time to escape measured in days from first detection of T cell response (detected in first 50 days from Fiebig stages I/II) to time to ≤50% transmitted virus within reactive epitope. [C]17 subjects; total epitopes analyzed = 59. [D]Time to escape measured in days from first detection of T cell response to time to ≤50% transmitted virus, where mutation emerges within proximity of T cell epitope, occurring coincident with decline of T cell response. [E]"Interaction term" examines whether the effect of the listed covariates is more than additive. [F]%M, which ranks the magnitude of individual T cell responses, is redundant in models that (a) also include M and (b) have data stratified by subject. –, not applicable. [G]Breadth is a subject level measure. Bold text highlights model *P* values.

These results are distinct from, and advance on, previous papers (e.g., refs. 13, 24). This is the first study to our knowledge to comprehensively measure and analyze vertical immunodominant T cell hierarchies in HIV-1 infection and to relate these to measurements of longitudinal immune escape. While, our previous work described acute and rapid virus escape, we did not examine or demonstrate the contribution of immunodominance or T cell breadth to escape (20, 39). The design of this study ensured that all HIV-1–specific T cell responses induced in subjects were quantified and considered for analysis; indeed, almost 60% of our mapped T cell responses had not been described previously. Our results, because they are independent of HLA

Table 4
%M and epitope entropy combine to explain nearly half of the of the variation in time to escape of epitopes[A]

CPH model[B]	Cov		TTE, unstratified	TTE + AgP[C], unstratified
Univariate	$\log_{10}$(%M)	%(ρ^2_n)[D]	29%	39%
	Breadth	%(ρ^2_n)	13%	26%
	$\log_{10}$(entropy)	%(ρ^2_n)	13%	10%
Multivariate	$\log_{10}$(%M)/ $\log_{10}$(entropy)	%(ρ^2_n)	43%	48%

[A]17 subjects; epitopes analyzed = 57; invariant epitopes/ censored events =10. [B]CPH test. [C]Time to escape measured in days from first detection of T cell response to time to ≤50% transmitted virus, where mutation emerges within proximity of T cell epitope, occurring coincident with decline of T cell response. [D]ρ^2_n proportion of total variance (= 1) in CPH test explained by listed covariate (50), expressed as a percentage.

restriction, are not biased to known T cell epitopes nor are they based on theoretical escape sequence footprints (49). Moreover, we showed that infecting HIV-1 clade, country of origin, and gender did not impact time to escape. Our results are therefore broadly generalizable to all subjects infected with HIV-1, whether or not they express a protective HLA allele. We show that the immune selection pressure exerted on HIV-1 by an individual T cell response must be considered in the context of the other T cell responses induced at that stage of infection. T cells targeting the same epitope in different individuals will contribute more or less to the total T cell immune pressure, dependent on whether the response is dominant or subdominant. This may reflect competition between different T cell populations, each seeing different epitopes on the same infected cell.

It should be noted that we define immunodominance very precisely as the magnitude of the response relative to the total response within each patient at a given time point. This differs from many previous studies in which a response has been called "immunodominant" because it occurs repeatedly in patients with a given HLA type (e.g., B*57 or B*27), without recognizing that it may not be the strongest response at several time points (24, 49). This is clearly illustrated for the responses of patients CH058 and CH077 to the TW10 immunoprevalent epitope, which was only immunodominant in the true sense at some time points and not in the first weeks of infection (Figure 2, B and E). Given than both T cell mapping approaches afford valuable though different data, we suggest in the future that vertical immunodominance studies be clearly distinguished from horizontal measures across patients that are more accurately investigating immunoprevalence.

This work is distinct from virus challenge models in mice, which, while showing protection associated with certain T cell immunodominance hierarchies, use invariant challenge viruses such as lymphocytic choriomeningitis virus (53, 54). Those studies highlighted the contribution of certain T cell immunodominance hierarchies to virus control but did not explore how different T cell immunodominance hierarchies combine to exert in vivo pressure on the infecting virus. Furthermore, the constancy of the infecting virus in these earlier animal experiments generated immunodominance hierarchies that were sustained over time, in marked contrast to the studies described here for HIV-1 (17).

In our study, subjects are aligned by stage of infection (days from Fiebig stages I/II) and therefore exclude the impact of both differential VL, which can influence T cell function and shifting T cell immunodominance patterns. These factors can confound studies of chronic infection and even studies of early HIV-1 infection. Here, deduction of the T/F virus(es) in each subject and subsequent virus sequencing were combined with detailed T cell kinetics data to produce by far the most accurate virus escape measurements to date. We show that escape from acute T cell responses within the same subjects can range from days to years. We went further and applied error modeling to these escape measurements to confirm the robustness of our results. Finally, we used multivariate analysis to demonstrate conclusively that nearly half of the variability in time to escape, which is an in vivo measure of T cell pressure, can be explained by just 2 measurements, immunodominance and entropy. Significantly, measurements of vertical immunodominance alone accounted for 39% of the variability in time to escape. The strength of this association shows clearly that vertical immunodominance hierarchies in HIV-1 are a major determinant of time to virus escape. Given that immunodominance is likely a compound consequence of many other variables, it will be important to investigate how factors such as polyfunctionality (55) and exhaustion profiles (56), lytic (57)/nonlytic control (52), and TCR repertoire (58–61) contribute to virus escape.

We were unable to quantify the contribution of these first HIV-1–specific T cell responses to decline in VL in acute infection, primarily because of insufficiently frequent VL data in these subjects preventing calculation of VL slopes. While it has been shown in SIV challenge of macaques that certain T cell immunodominance hierarchies are associated with VL control, we predict HIV-1 control in human infection will be more complex (62, 63). In macaques, standardized SIV challenge strains have been used across experiments; this differs from HIV-1 infection for which every subject's infection is founded by a different virus (64). Differences in T/F virus sequence could impact virus replicative fitness. Epidemiological studies suggest that intrinsic properties of the T/F virus account for as much as 50% of the variation in VL set point (65, 66). Consistently, in our study, examination of T/F virus sequence outside reactive T cell epitopes found that subjects with lower acute VL set points contained footprints of previous virus escape in B57/5801, B81, and B51 epitopes (Supplemental Table 1 and refs. 67, 68). These escape variants have been associated with lower in vitro replication, suggesting the transmitted viruses of these subjects had poorer replicative fitness, which contributed to the lower set point VL observed in some subjects (37, 64–68). Studies are ongoing in an extended patient set to examine associations of T cell responses with VL over the first 6 months of HIV-1 infection. The changing patterns of T cell immunodominance during this period will exert variable pressure, which will have to be incorporated into a model.

These results have practical implications for T cell vaccine design. Although nothing can be done to influence the nature of the actual T/F virus, thereafter vaccination could influence events profoundly. Our data show that the fitness loss to HIV-1 of some mutations is so great that the virus will endure ongoing killing or viral suppression by immunodominant $CD8^+$ T cells, often for years. In our data set, in primary infection most of the immunodominant T cell responses targeted high entropy regions of HIV-1, often in Nef and Env, and escape was very rapid. Low entropy epitopes were also targeted, but these responses were typically subdominant. We argue that the host T cell response in natural infection, which targets both low and high entropy epitopes, with the latter more likely to be immunodominant early on, explains partly the failure to clear the virus. Our data support vaccine immunogens that focus on conserved regions of HIV-1 but possibly excluding the Pol protein, which was poorly immunogenic in acute HIV-1 infection. This approach would ensure, regardless of the immunodominance hierarchy established in individual vaccinees, that all T cell responses would be focused on conserved regions of HIV-1, minimizing rapid, early escape. Induction of broad T cell responses is particularly important in HIV-1 to maximize the likelihood of vaccine-induced T cell response recognizing highly variable infecting viruses. However, our observations demonstrated that increased T cell breadth often occurs at the cost of immunodominance, because immunodominant T cell responses targeted to high entropy epitopes would out-compete subdominant responses that recognize lower entropy epitopes, resulting in rapid virus escape and less sustained T cell pressure. This suggests that, in some subjects, high T cell breadth may be detrimental to HIV-1 control. Notably, successful induction of broad T cell responses against a conserved HIV-1 immunogen would satisfy both the need to maximize T cell detection of incoming viruses and to overcome any detrimental effect of breadth on immunodominance, because, even if an epitope escapes, the next T cell response would also likely focus on a low entropy epitope.

Another important implication that comes from these data is that measuring T cell responses to pools of HIV-1 peptides is not adequate to test immunogenicity of HIV-1 T cell vaccines. Better predictions of vaccine efficacy could be made by mapping of individual vaccine-induced T cell responses to define immunodominance hierarchies, breadth, and epitope entropy. This approach would enable identification of those vaccinees who have dominant T cell responses targeting low entropy regions of HIV-1 and who therefore have the greatest likelihood of exerting significant and sustained T cell pressure in the first weeks of subsequent HIV-1 infection.

The observation of virus escape in HIV-1, first described over 20 years ago (28), has been the focus of numerous studies investigating escape to understand in vivo immune pressure exerted by T cells on HIV-1 (69). In this study, we have explained virus escape, showing that almost half of all the variability of time to escape is determined by just 2 factors — immunodominance and epitope entropy. This study therefore simplifies the field. Factors, such as polyfunctionality, T cell exhaustion, and T cell receptor hierarchies, that have been previously proposed as independent determinants of in vivo T cell pressure in HIV-1 could play roles but more likely only when they work in the context of vertical T cell immunodominance.

Methods

Patients and samples. Acute infection was defined as a positive for HIV-1 vRNA in plasma but negative or equivocal for HIV-1 serology at a screening visit. Patients were bled at multiple time points over 1 to 3 years. See Table 1 and Supplemental Table 1 for the demographical and clinical data for each patient. VL set point was determined using an average of VLs within a "set point window" (2). Patients were aligned using days from Fiebig stages I/II. Days after Fiebig stages I/II were estimated for a subset of subjects who were first detected slightly later in Fiebig stages III and IV (Fiebig stage III = 7 days after Fiebig stages I/II; Fiebig stage IV = 11 days after Fiebig stages I/II; ref. 46).

Virus sequencing. RNA extraction, complementary DNA synthesis, and SGA have been described previously (70, 71). The T/F virus(es) was deduced in the overlapping 5′ and 3′ halves for each subject from plasma collected in Fiebig stages I–IV. For the bulk of subjects (15 out of 17), serial SGA sequencing was performed approximately 4, 8, 24, and 60 weeks following subject screening. For 2 subjects, more limited screening (1 time point within 8 weeks of screening) was performed because of either early ART treatment (subject CH607) or strong and sustained VL control <40 copies vRNA per ml (subject CH067), which prevented successful amplification of single virus genomes. For 9 out of 17 subjects, samples were also available for additional sequencing at time points between 2 and 3 years from screening.

T cell mapping. 18-mer peptides overlapping by 10 amino acids were synthesized (Sigma-Genosys or MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine) to match the transmitted sequence of each subject. Approximately 400 peptides for each patient (including >200 that were unique to that patient) were arranged into pools in a matrix format using the Peptide Portal program (SCHARP), adapting code from the Deconvolute this! program (72). Each peptide was repeated 4 times in the matrix peptide plate. Cryopreserved PBMCs were thawed and rested for 2 hours before being placed in the ELISpot plates (Merck, Millipore) at 1×10^5 cells per well. Antigens in the peptide plates were mixed with 1:1 with PBMCs in the ELISpot plate to a final concentration of 2 or 5 μg/ml and incubated for 20 hours at 37°C, 5% CO_2. Coating, development (MabTech), and reading of ELISpot plates (AID Reader) has been described previously (73). At this stage, 30 SFU per 10^6 PBMCs and >3× background threshold of positivity was applied to peptide pools to maximize sensitivity. Data from positive pools were then deconvoluted (72) to identify candidate epitope peptides. Putative positive 18-mer peptide–specific T cell responses were confirmed in triplicate in a follow-up IFN-γ ELISpot with 1×10^5 cells per well and a peptide concentration of 2 μg/ml with 4 to 6 negative control wells (media only) and at least 1 positive control well (10 μg/ml PHA; Sigma-Aldrich) together. Positive T cell responses were defined as ≥30 SFU per million, more than 4 times background. Zero values were not accepted in any replicate of antigen-stimulated wells.

LANL ELF software was then used to identify HLA-matched known and predicted CTL epitopes (http://www.hiv.lanl.gov/content/sequence/ELF/epitope_analyzer.html). These 9- to 11-mer peptides were then synthesized and tested using ELISpot, applying the same positivity criteria described above. Optimal epitopes were defined for 58% of reactive 18-mer peptides (see Supplemental Table 1). For univariate and multivariate analysis, data from the experimentally confirmed optimal epitope was used in preference to corresponding 18-mer data.

HLA typing. HLA typing (Weatherall Institute of Molecular Medicine) was performed using the sequence-specific primer method adapted from ref. 74 that uses allele-specific primer combination in PCR amplification to provide absolute HLA resolution to 2 digits and high-probability resolution to 4 digits.

Time to escape calculations. Peptides and substitutions were aligned relative to HXB2 standards (http://www.hiv.lanl.gov/content/sequence/QUICK_ALIGN/QuickAlign.html). Virus escape within the reactive epitope was defined as ns mutations occurring with ≥50% change in SGA sequences compared to the founder T/F virus sequence. ns mutant epitopes elicited an attenuated response (see Supplemental Table 1) when tested in ex vivo IFN-γ ELISpots (cells permitting) in comparison to the T/F epitope; otherwise an attenuated response was assumed. An AgP mutation was identified when a decline in magnitude of the T cell response was observed subsequent to the emergence of ns mutations occurring in close proximity (≤20 amino acids 5′ or 3′) to the reactive epitope.

Time to escape was calculated as the time in days from first detection of the selecting T cell response to virus escape in the reactive epitope. This latter time point was determined using the gradient of the trend line (Excel, Microsoft) between (a) the last visit when the T/F sequence was found in >50% of SGA sequences generated and (b) the subsequent time point when the proportion of T/F sequences within the reactive peptide fell below 50%. These data were used for all CPH tests. The term "invariant" epitope indicates that the percentage of T/F sequences corresponding to the reactive peptide remained >50% across all sequencing time points generated from a minimum sequencing window of ≥200 days from subject's initial screening visit.

Entropy calculations. A Shannon entropy score was calculated for all reactive peptides in this study (40). Entropy was calculated for experimentally confirmed optimal epitopes or for predicted 9-mers epitopes where the optimal was not known (see Supplemental Figure 3 for details).

For each epitope, the LANL sequence database was used to identify all known sequences that belong to the same HIV-1 clade that infected the responding participant. Data were limited to one randomly selected sequence per patient, and sequences defined as "problematic" by the LANL database search tool were omitted. Between 516 and 43,538 sequences were analyzed per epitope. To focus the analysis on peptides as they appear to T cell receptors, all gaps within sequences were removed. Sequences were aligned and gaps included at either terminus to maximize similarity between sequences. Stop codons, and any sequence downstream, were stripped and replaced with gaps. Gaps were treated as a separate symbol. The Shannon entropy of each position in the epitope measured in nats was then calculated with the following formula ($-\sum^{n}_{i=1} p_i \log_e p_i$), where p_i stands for the frequency of a given amino acid at position i. The mean of these scores was used as the epitope entropy.

Rules were defined to *predict* the location of 9-mer T cell epitopes in reactive 18-mer peptides. Predicted 9-mers were centered on residues where ns changes from the T/F emerged over time and mean entropy calculated. Where no mutations emerged, the peptide was split into nonoverlapping 9-mers, and the average of entropies for each 9-mer was calculated.

Statistics. The relationship between several covariates and an outcome variable, "time to escape" (see above), with and without 3 AgP mutations was examined. In order to estimate RRs, CPH models were fitted to data. A univariate model for each covariate was fitted to data, and those significantly associated with time to escape ($P \leq 0.05$) were considered in multivariate models. Both nonstratified and subject-stratified models were fitted. To better characterize the relative effect sizes across covariates, each covariate was also binned into tertiles and entered as a categorical covariate into the regression model, with the first tertile serving as the reference group. *P* values of less than 0.05 were considered significant. To determine how much of the variation in time to escape was determined by %M, breadth, and entropy, we applied the O'Quigley model ρ^2_n (50), which is a modified version of the generalized linear R^2 equation (75) that replaces total number of observations with the number of noncensored observations.

Study approval. Subjects gave full written informed consent to enroll into the acute infection arms of the CHAVI 001 cohort (OXTREC 3306, Oxford) and Centre for AIDS Programme Research in South Africa (CAPRISA) 002 (Research Ethics Committee, Faculty of Health Sciences, UCT REC no. 025/2004, Capetown). Approval for these experiments has also been given by OXTREC 3306, Oxford.

Acknowledgments

We thank CHAVI and Duke management, particularly Kelly Soderberg, Jennifer Kirchnerr, and Marybeth McCauley for study coordination; A. Williams, C. Margaret, C. deBoer, S.C. Heeti, and R. Thomas of SCHARP for database support; J. Roberts for administrative support; T. Rostron for HLA typing; and K. diGleria and Z. Yu for peptide synthesis. This work was supported by the Center for HIV/AIDS Vaccine Immunology grant AI067854. Additional support came from the Medical Research Council Human Immunology Unit, the NIH Research Oxford Biomedical Research Centre. P. Borrow and A.J. McMichael are Jenner Institute Investigators.

Received for publication July 3, 2012, and accepted in revised form October 5, 2012.

Address correspondence to: Nilu Goonetilleke, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DS, United Kingdom. Phone: 44.1865.222.145; Fax: 44.1865.222.502; E-mail: nilu.goonetilleke@ndm.ox.ac.uk.

1. Altfeld M, et al. HLA Alleles Associated with delayed progression to AIDS contribute strongly to the initial CD8(+) T cell response against HIV-1. *PLoS Med*. 2006;3(10):e403.
2. Fellay J, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science*. 2007;317(5840):944–947.
3. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. Nature reviews. *Immunology*. 2010;10(1):11–23.
4. Pereyra F, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science*. 2010;330(6010):1551–1557.
5. Barouch DH, et al. Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. *Nature*. 2002;415(6869):335–339.
6. Schmitz JE, et al. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science*. 1999;283(5403):857–860.
7. Draenert R, et al. Persistent recognition of autologous virus by high-avidity CD8 T cells in chronic, progressive human immunodeficiency virus type 1 infection. *J Virol*. 2004;78(2):630–641.
8. Altfeld M, et al. Enhanced detection of human immunodeficiency virus type 1-specific T-cell responses to highly variable regions by using peptides based on autologous virus sequences. *J Virol*. 2003;77(13):7330–7340.
9. Kiepiela P, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med*. 2007;13(1):46–53.
10. Goulder P, Price D, Nowak M, Rowland-Jones S, Phillips R, McMichael A. Co-evolution of human immunodeficiency virus and cytotoxic T-lymphocyte responses. *Immunol Rev*. 1997;159:17–29.
11. Kaslow RA, et al. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nat Med*. 1996;2(4):405–411.
12. Leslie A, et al. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *J Virol*. 2010;84(19):9879–9888.
13. Dinges WL, et al. Virus-specific CD8+ T-cell responses better define HIV disease progression than HLA genotype. *J Virol*. 2010;84(9):4461–4468.
14. Yewdell JW. Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. *Immunity*. 2006;25(4):533–543.
15. Chen W, Anton LC, Bennink JR, Yewdell JW. Dissecting the multifactorial causes of immunodominance in class I-restricted T cell responses to viruses. *Immunity*. 2000;12(1):83–93.
16. Tscharke DC, et al. Identification of poxvirus CD8+ T cell determinants to enable rational design and characterization of smallpox vaccines. *J Exp Med*. 2005;201(1):95–104.
17. Nowak MA, et al. Antigenic oscillations and shifting immunodominance in HIV-1 infections. *Nature*. 1995;375(6532):606–611.
18. Iversen AK, et al. Conflicting selective forces affect T cell receptor contacts in an immunodominant human immunodeficiency virus epitope. *Nat*

Immunol. 2006;7(2):179–189.
19. Jones NA, et al. Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. *J Exp Med*. 2004;200(10):1243–1256.
20. Goonetilleke N, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med*. 2009;206(6):1253–1272.
21. Oseroff C, et al. Dissociation between epitope hierarchy and immunoprevalence in CD8 responses to vaccinia virus western reserve. *J Immunol*. 2008;180(11):7193–7202.
22. Yu XG, et al. Consistent patterns in the development and immunodominance of human immunodeficiency virus type 1 (HIV-1)-specific CD8+ T-cell responses following acute HIV-1 infection. *J Virol*. 2002;76(17):8690–8701.
23. McMichael AJ, Gotch FM, Rothbard J. HLA B37 determines an influenza A virus nucleoprotein epitope recognized by cytotoxic T lymphocytes. *J Exp Med*. 1986;164(5):1397–1406.
24. Streeck H, et al. Human immunodeficiency virus type 1-specific CD8+ T-cell responses during primary infection are major determinants of the viral set point and loss of CD4+ T cells. *J Virol*. 2009;83(15):7641–7648.
25. Turnbull EL, et al. Kinetics of expansion of epitope-specific T cell responses during primary HIV-1 infection. *J Immunol*. 2009;182(11):7131–7145.
26. Borrow P, et al. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat Med*. 1997;3(2):205–211.
27. Goulder PJ, et al. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat Med*. 1997;3(2):212–217.
28. Phillips RE, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature*. 1991;354(6353):453–459.
29. Price DA, et al. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc Natl Acad Sci U S A*. 1997;94(5):1890–1895.
30. Fischer W, et al. Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PloS One*. 2010;5(8):e12303.
31. Ganusov VV, et al. Fitness costs and diversity of the cytotoxic T lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *J Virol*. 2011;85(20):10518–10528.
32. Asquith B, Edwards CT, Lipsitch M, McLean AR. Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. *PLoS Biol*. 2006;4(4):e90.
33. Leslie AJ, et al. HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med*. 2004;10(3):282–289.
34. Kawashima Y, et al. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature*. 2009; 458(7238):641–645.
35. Crawford H, et al. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. *J Virol*. 2007;81(15):8346–8351.
36. Martinez-Picado J, et al. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *J Virol*. 2006;80(7):3617–3623.
37. Miura T, et al. Impaired replication capacity of acute/early viruses in persons who become HIV controllers. *J Virol*. 2010;84(15):7581–7591.
38. Troyer RM, et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog*. 2009;5(4):e1000365.
39. Ferrari G, et al. Relationship between functional profile of HIV-1 specific CD8 T cells and epitope variability with the selection of escape mutants in acute HIV-1 infection. *PLoS Pathog*. 2011;7(2):e1001273.
40. Shannon CE. A mathematical theory of communication. *Bell Sys Tech J*. 1948;27(3):379–423.
41. Gaschen B, et al. Diversity considerations in HIV-1 vaccine selection. *Science*. 2002;296(5577):2354–2360.
42. Allen TM, et al. 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*. 2005;79(21):13239–13249.
43. Keele BF, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A*. 2008;105(21):7552–7557.
44. Altfeld M, et al. Cellular immune responses and viral diversity in individuals treated during acute and early HIV-1 infection. *J Exp Med*. 2001;193(2):169–180.
45. Mlotshwa M, et al. Fluidity of HIV-1-specific T-cell responses during acute and early subtype C HIV-1 infection and associations with early disease progression. *J Virol*. 2010;84(22):12018–12029.
46. Fiebig EW, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS*. 2003;17(13):1871–1879.
47. Fiebig EW, Heldebrant CM, Smith RI, Conrad AJ, Delwart EL, Busch MP. Intermittent low-level viremia in very early primary HIV-1 infection. *J Acquir Immune Defic Syndr*. 2005;39(2):133–137.
48. Draenert R, et al. Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. *J Exp Med*. 2004;199(7):905–915.
49. Brumme ZL, et al. Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection. *J Virol*. 2008;82(18):9216–9227.
50. O'Quigley J, Xu R, Stare J. Explained randomness in proportional hazards models. *Stat Med*. 2005;24(3):479–489.
51. Brockman MA, et al. Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. *J Virol*. 2007;81(22):12608–12618.
52. Elemans M, Seich Al Basatena NK, Klatt NR, Gkekas C, Silvestri G, Asquith B. Why don't CD8+ T cells reduce the lifespan of SIV-infected cells in vivo? *PLoS Comput Biol*. 2011;7(9):e1002200.
53. Gallimore A, Dumrese T, Hengartner H, Zinkernagel RM, Rammensee HG. Protective immunity does not correlate with the hierarchy of virus-specific cytotoxic T cell responses to naturally processed peptides. *J Exp Med*. 1998;187(10):1647–1657.
54. van der Most RG, et al. Analysis of cytotoxic T cell responses to dominant and subdominant epitopes during acute and chronic lymphocytic choriomeningitis virus infection. *J Immunol*. 1996;157(12):5543–5554.
55. Betts MR, et al. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood*. 2006;107(12):4781–4789.
56. Petrovas C, et al. Differential association of programmed death-1 and CD57 with ex vivo survival of CD8+ T cells in HIV infection. *J Immunol*. 2009;183(2):1120–1132.
57. Makedonas G, et al. Perforin and IL-2 upregulation define qualitative differences among highly functional virus-specific human CD8 T cells. *PLoS Pathog*. 2010;6(3):e1000798.
58. Price DA, et al. Public clonotype usage identifies protective Gag-specific CD8+ T cell responses in SIV infection. *J Exp Med*. 2009;206(4):923–936.
59. Mendoza D, et al. HLA B*5701-positive long-term nonprogressors/elite controllers are not distinguished from progressors by the clonal composition of HIV-specific CD8+ T cells. *J Virol*. 2012;86(7):4014–4018.
60. Iglesias MC, et al. Escape from highly effective public CD8+ T-cell clonotypes by HIV. *Blood*. 2011;118(8):2138–2149.
61. Chen H, et al. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nat Immunol*. 2012;13(7):691–700.
62. Hasegawa A, et al. Analysis of TCRalphabeta combinations used by simian immunodeficiency virus-specific CD8+ T cells in rhesus monkeys: implications for CTL immunodominance. *J Immunol*. 2007;178(6):3409–3417.
63. Loffredo JT, et al. Patterns of CD8+ immunodominance may influence the ability of Mamu-B*08-positive macaques to naturally control simian immunodeficiency virus SIVmac239 replication. *J Virol*. 2008;82(4):1723–1738.
64. Loh L, et al. Complexity of the inoculum determines the rate of reversion of SIV Gag CD8 T cell mutant virus and outcome of infection. *PLoS Pathog*. 2009;5(4):e1000378.
65. Hecht FM, et al. HIV RNA level in early infection is predicted by viral load in the transmission source. *AIDS*. 2010;24(7):941–945.
66. Alizon S, et al. Phylogenetic approach reveals that virus genotype largely determines HIV set-point viral load. *PLoS Pathog*. 2010;6(9):e1001123.
67. Goepfert PA, et al. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *J Exp Med*. 2008;205(5):1009–1017.
68. Chopera DR, et al. Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. *PLoS Pathog*. 2008;4(3):e1000033.
69. Goulder PJ, Walker BD. The great escape - AIDS viruses and immune control. *Nat Med*. 1999; 5(11):1233–1235.
70. Salazar-Gonzalez JF, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med*. 2009;206(6):1273–1289.
71. Abrahams MR, et al. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *J Virol*. 2009; 83(8):3556–3567.
72. Roederer M, Koup RA. Optimized determination of T cell epitope responses. *J Immunol Methods*. 2003;274(1–2):221–228.
73. Goonetilleke N, et al. Induction of multifunctional human immunodeficiency virus type 1 (HIV-1)-specific T cells capable of proliferation in healthy subjects by using a prime-boost regimen of DNA- and modified vaccinia virus Ankara-vectored vaccines expressing HIV-1 Gag coupled to CD8+ T-cell epitopes. *J Virol*. 2006;80(10):4717–4728.
74. Bunce M. PCR-sequence-specific primer typing of HLA class I and class II alleles. In: Powis SH, Vaughan RW, eds. *MHC Protocols*. Vol. 210. Totowa, New Jersey, USA: Human Press; 2003:143–172.
75. Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika*. 1991;78(3):691–692.