JVI Journals.ASM.org

HIV Control Is Mediated in Part by CD8$^+$ T-Cell Targeting of Specific Epitopes

Florencia Pereyra,[a,b] David Heckerman,[c] Jonathan M. Carlson,[c] Carl Kadie,[d] Damien Z. Soghoian,[a] Daniel Karel,[a] Ariel Goldenthal,[a] Oliver B. Davis,[a] Charles E. DeZiel,[c] Tienho Lin,[c] Jian Peng,[c] Alicja Piechocka,[a] Mary Carrington,[a,e] Bruce D. Walker[a,f]

Ragon Institute of MGH, MIT and Harvard, Boston, Massachusetts, USA[a]; Division of Infectious Diseases, Brigham and Women's Hospital, Boston, Massachusetts, USA[b]; Microsoft Research, Los Angeles, California, USA[c]; Microsoft Research, Redmond, Washington, USA[d]; Cancer and Inflammation Program, Laboratory of Experimental Immunology, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA[e]; Howard Hughes Medical Institute, Chevy Chase, Maryland, USA[f]

ABSTRACT

We investigated the hypothesis that the correlation between the class I HLA types of an individual and whether that individual spontaneously controls HIV-1 is mediated by the targeting of specific epitopes by CD8$^+$ T cells. By measuring gamma interferon enzyme-linked immunosorbent spot (ELISPOT) assay responses to a panel of 257 optimally defined epitopes in 341 untreated HIV-infected persons, including persons who spontaneously control viremia, we found that the correlation between HLA types and control is mediated by the targeting of specific epitopes. Moreover, we performed a graphical model-based analysis that suggested that the targeting of specific epitopes is a cause of such control—that is, some epitopes are protective rather than merely associated with control—and identified eight epitopes that are significantly protective. In addition, we use an *in silico* analysis to identify protein regions where mutations are likely to affect the stability of a protein, and we found that the protective epitopes identified by the ELISPOT analysis correspond almost perfectly to such regions. This *in silico* analysis thus suggests a possible mechanism for control and could be used to identify protective epitopes that are not often targeted in natural infection but that may be potentially useful in a vaccine. Our analyses thus argue for the inclusion (and exclusion) of specific epitopes in an HIV vaccine.

IMPORTANCE

Some individuals naturally control HIV replication in the absence of antiretroviral therapy, and this ability to control is strongly correlated with the HLA class I alleles that they express. Here, in a large-scale experimental study, we provide evidence that this correlation is mediated largely by the targeting of specific CD8$^+$ T-cell epitopes, and we identify eight epitopes that are likely to cause control. In addition, we provide an *in silico* analysis indicating that control occurs because mutations within these epitopes change the stability of the protein structures. This *in silico* analysis also identified additional epitopes that are not typically targeted in natural infection but may lead to control when included in a vaccine, provided that other epitopes that would otherwise distract the immune system from targeting them are excluded from the vaccine.

Control of HIV is associated with the expression of certain "protective" HLA alleles (1–5). Although the mechanisms mediating this effect are not fully understood, evidence points to the induction of virus-specific CD8$^+$ T-lymphocyte (CTL) responses as the primary mechanism of control. In animal studies of AIDS virus infection, depletion of CTL leads to rapid loss of control (6), while in human studies, *in vitro* quantification of the quality of T-cell responses is correlated with HIV control, with effective HIV-specific CTLs characterized by their ability to rapidly proliferate (7), avoid exhaustion (8, 9), and efficiently kill infected CD4$^+$ T cells (7, 10, 11). These observations provided the rationale for the development of CTL vaccines, though such vaccines expressing full-length viral proteins as antigens have failed to induce protection or control in human trials (12, 13).

While CTL responses are generally linked with natural HIV control, an important question is whether control is more likely if the CTL response is targeted against specific viral epitopes. Studies suggest that targeting predominantly the HIV Gag protein is correlated with control (14–17), while targeting Env (14) and Nef (16, 17) is correlated with progression, though the reasons for this differentiation are unclear. One hypothesis is that Gag-related protection is due to the relative abundance of Gag proteins soon after viral entry into a cell (18), suggesting that broad Gag-specific CTL targeting should be the aim of a CTL-based vaccine. Another hypothesis is that targeting specific "protective" epitopes leads to control (19–24) and that Gag just happens to have a high concentration of such epitopes. Recent evidence for this hypothesis includes the observation that targeting specific epitopes in Gag is linked with loss of control (21), while targeting some regions outside Gag is linked with control (25–27). If the latter hypothesis is

Received 11 April 2014 **Accepted** 19 August 2014
Published ahead of print 27 August 2014
Editor: G. Silvestri
Address correspondence to Bruce D. Walker, bwalker@partners.org.
F.P., D.H., and J.M.C. contributed equally to this article.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.01004-14.

doi:10.1128/JVI.01004-14
The authors have paid a fee to allow immediate free access to this article.

correct, then identifying protective epitopes is an important step in the development of a vaccine.

One hypothesis that has emerged is that protective epitopes will be those that are mutationally constrained, meaning that CTL targeting of these epitopes followed by escape through mutation would weaken the virus. Perhaps the simplest way to estimate mutational constraint is to use overall sequence conservation. However, although this is an approach that is being actively pursued (20, 28), studies to date have shown at best very weak correlations between the conservation of the sequences targeted by an individual's CTL response and viral control (25), though there is some evidence that more careful definitions of sequence conservation that account for covariation may improve the link between targeting conserved elements and control (22, 24).

In this study, we directly tested the hypothesis that targeting of specific epitopes mediates the correlation between HLA alleles and control by comparing the responses of untreated viremic controllers and progressors to all reported optimally defined epitopes. This approach is unbiased regarding HLA expression, an important consideration since many epitopes have been reported to be targeted in the context of more than one HLA class I allele (reviewed in reference 4). From these data, we identified specific epitopes correlated with control and used statistical methods to show that targeting of specific epitopes is a likely causal mechanism of HLA-mediated control—that is, that these epitopes are likely protective rather than simply correlated with control. Furthermore, we found that although protective epitopes were not distinguished by high levels of sequence conservation, they were distinguished by the tendency of intraepitopic mutation to destabilize the HIV protein structures. These observations provide insight into the mechanisms of natural control and provide a means by which we can identify protective epitopes that are not often targeted in natural infection but may be potentially useful in a vaccine.

TABLE 1 Baseline characteristics of the cohort

Characteristic	Controllers	Progressors
n		
All	245	96
EC	97	
VC	148	
Male gender, *n* (%)		
All	176 (72)	81 (84)
EC	70 (72)	
VC	106 (72)	
Median age (IQR[a]), yr		
All	45 (39–52)	41 (32–47)
EC	48 (41–54)	
VC	44 (37–50)	
Median disease duration (IQR), yr		
All	8 (3–17)	4 (1–12)
EC	15 (5–21)	
VC	6 (3–14)	
Median no. of $CD4^+$ T cells/mm^3 (IQR)		
All	733 (547–918)	466 (321–614)
EC	836 (657–1149)	
VC	664 (494–819)	
Median no. of HIV-1 copies/ml (IQR)		
All	160 (50–750)	24100 (8115–72625)
EC	50 (48–75)	
VC	448 (192–1404)	
Race, n (%)		
Black		
All	79 (32)	29 (30)
EC	23 (24)	
VC	56 (38)	
White		
All	148 (60)	58 (60)
EC	67 (69)	
VC	81 (55)	
Other		
All	18 (7)	9 (9)
EC	7 (7)	
VC	11 (7)	

[a] IQR, interquartile range.

MATERIALS AND METHODS

Study subjects. Subjects for the first part of the analysis were recruited from outpatient clinics at local Boston hospitals and referred from providers throughout the United States, following institutional review board approval and written informed consent. Controller status was determined based on the median HIV RNA levels over available samples prior to enrollment. Subjects were classified as elite controllers ($n = 98$), viremic controllers ($n = 147$), or progressors ($n = 96$) if their median HIV RNA levels were below the limit of detection, ≤2,000 copies/ml, or >2,000 copies/ml, respectively (Table 1). The limit of detection was <75 copies/ml by branched chain DNA testing (bDNA) or <50 copies/ml by ultrasensitive PCR.

HLA typing. High-resolution class I HLA typing for HLA-A, -B, and -C was performed at the National Cancer Institute, National Institutes of Health, Frederick, MD, using sequence-based typing protocols developed by the 13th International Histocompatibility Workshop (29). DNA for typing was extracted from whole blood using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN) according to the manufacturer's instructions. HLA alleles were grouped at the 2-digit resolution level, with the exception of A*68, B*15, whose allotypes have divergent binding motifs (30), and B*58, B*42, and B*35, which were previously reported to exhibit discordant associations with disease progression (31–34).

Assessment of HIV-specific CD8 T-cell responses. Gamma interferon (IFN-γ) enzyme-linked immunosorbent spot (ELISPOT) assays were performed with previously cryopreserved peripheral blood mononuclear cells (PBMC), using a panel of 257 peptides corresponding to optimal clade B cytotoxic T-lymphocyte epitopes (35). For each subject, all defined CD8 T-cell epitopes were tested irrespective of the known HLA class I allele restriction. PBMC were plated at 100,000 cells per well with peptides at a final concentration of 14 μg/μl in 96-well plates and processed as described previously (36). Medium alone served as a negative control and phytohemagglutinin (PHA) as a positive control. The numbers of specific IFN-γ-secreting T cells were enumerated using an automated ELISPOT reader (Cellular Technology Ltd., Shaker Heights, OH) and expressed as spot-forming cells (SFCs)/10^6 PBMC. Responses were considered positive if they had at least three times the mean number of SFC in the three negative-control wells and >50 SFC/10^6 PBMC (36, 37).

Flow cytometry. Whole PBMC or expanded T-cell lines were stimulated with specified optimal peptides directly or with B-cell lines pulsed with the specified optimal peptides (2 μg/ml), as previously described (38). Briefly, stimulation was performed at 37°C for 6 h in R10 (RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 10 mM HEPES, 100 U/ml penicillin, and 100 μg/ml strep-

tomycin). To detect CD107a mobilization, a predetermined amount of anti-human CD107a phycoerythrin (PE)-Cy5 (clone H4A3; BD Biosciences) was added to the medium. An unstimulated (medium-only) sample served as a negative control. At 30 min into the stimulation, the transport inhibitors brefeldin A (Sigma) and monensin (BD Biosciences) were added to facilitate detection of T-cell responses. Following stimulation, cells were surface stained with an amine-reactive viability dye and antibodies against CD3, CD4, and CD8. Cells were then washed, permeabilized (Cytofix/Cytoperm; BD Biosciences), and stained with antibodies against IFN-γ and/or tumor necrosis factor alpha (TNF-α). Data were collected on a BD LSR II cytometer.

Predictive model and statistical analysis. L1-regularized logistic regression (39) was used to predict controller status (HIV controllers versus chronic progressors) based on class I HLA alleles alone, epitope-specific $CD8^+$ T-cell targeting alone, and the combination of both sets of observations. The prediction accuracy of one model relative to another was determined by comparing the areas under the receiver operating characteristic (ROC) curves generated using 10-fold cross validation to ensure out-of-sample predictions, while the Lasso penalty for each training set was determined using nested 10-fold cross validation. Significant differences in the areas under the curves (AUCs) were determined by estimating the variance of the AUC differences using a bootstrap approach (40).

To identify HLA class I restrictions for each epitope, we assessed the significance of association between the presence of targeting that epitope and the presence of an allele using Fischer's exact test. To limit the effects of HLA linkage disequilibrium, we used a decision tree approach similar to that described previously (41): for each epitope, we identified the HLA that was most significantly associated with targeting that epitope, then removed all individuals who expressed that HLA, and repeated the procedure until no HLA was associated with targeting at a P value of <0.05 by Fisher's exact test. To estimate q values, we took all contingency tables of all HLA-epitope pairs resulting from the decision tree and computed q values using the exact Fisher's exact test P value distributions of the tables (42).

Directed acyclic graphical models were used to investigate the direction of cause and effect (43). Each node in a graphical model corresponds to a random variable, while arcs correspond to assertions of cause and effect. In a given graphical model, an L1-regularized logistic regression model was constructed for each non-HLA variable/node with regressor variables being those of the parent variables (nodes) in the graph. To measure how well a model fit the data, we determined, for each sample (individual) in the data, the out-of-sample log likelihood summed over each regression node, using 10-fold cross validation. To compare these fits for two models, we tested for differences between these summed log likelihoods using a Wilcoxon matched-pairs signed-rank test.

***In silico* analysis of protein structural stability.** To estimate the expected amino acid distribution at a given site, we estimated the thermodynamic stability changes caused by each of the 20 amino acids using the FoldX software package (http://foldx.crg.es/) and a six-step procedure, as previously described (44). Briefly, the structures of p17 (Protein Data Bank [PDB] code 2GOL) (45), p24 hexamer (PDB 3H4E) (46), protease dimer (PDB 3IXO) (47), reverse transcriptase (RT) (PDB 1DLO) (48), integrase (PDB 1BIS) (49), Tat (PDB 3MI9) (50), Rev (PDB 3LPH) (51), Nef (PDB 1EFN) (52), and gp120 (PDB 3JWD) (53) were mutated to the NL4-3 reference strain background, and the FoldX optimization procedure and probability-based rotamer libraries were used (54) to remove steric clashes and other estimation errors and to reconstruct missing side chain atoms. Then, the absolute change $|ddG|$ in the Gibbs free energy was estimated using the FoldX software for each of the 20 amino acids. The expected frequency $E[f_{ij}]$ of amino acid j at site i was then estimated using the Boltzmann distribution as the normalized negative exponential:

$$p_{ij} = E[f_{ij}] = \frac{\exp(-|ddG_{ij}|)}{\sum_{k=1}^{20} \exp(-|ddG_{ik}|)}$$

The expected entropy $E[H_i]$ of site i was calculated as $E[H_i] = -\sum_{j=1}^{20} p_{ij}\ln(p_{ij})$. The structural entropy of an epitope is reported as the average of $E[H_i]$ over all sites i in the epitope. Standard sequence entropies were calculated in an identical manner, with $E[f_{ij}]$ being estimated as the observed frequency of amino acid j at site i in a previously published alignment of full sequences sampled from approximately 1,800 chronically clade B-infected, antiretroviral (ARV)-naive individuals from North American and Australian cohorts (55).

RESULTS

We evaluated the relationship between epitope targeting specificity and viral control in 341 persons with chronic untreated HIV infection and high-resolution HLA typing, including 245 HIV controllers (56), who maintain viremia at less than 75 RNA copies/ml (elite controllers [EC]) or 2,000 RNA copies/ml (viremic controllers [VC]) in the absence of antiretroviral therapy (median viral load of 160 RNA copies/ml), and 96 chronic progressors (median viral load of 23,800 RNA copies/ml) (Table 1). In particular, peripheral blood mononuclear cells (PBMC) from all 341 subjects were tested for responses to HIV epitopes listed in the Los Alamos Immunology Database that are considered to be optimally defined in terms of epitope length and restricting class I allele ($n = 257$) (35) using an IFN-γ ELISPOT assay. Responses were scored as either positive or negative based on predetermined criteria (57). All individuals were tested for responses to all epitopes. Consistent with previous work (41, 58), we found many undocumented responses to epitopes among individuals who did not express the HLA alleles reported to restrict those epitopes. Although it is possible some of these undocumented responses are $CD4^+$ specific, our experimental setup was unlikely to detect such responses, as CD4 T-cell-induced IFN-γ responses typically require longer exposure time and CD8 T-cell depletion to be measurable (59, 60). Nevertheless, to confirm $CD8^+$ specificity, we performed intracellular cytokine secretion assays to randomly selected responses among five individuals with the highest number of undocumented responses (40 to 50); CD4 T-cell responses did not appear to meaningfully contribute to our findings by flow cytometry (see Fig. S1 in the supplemental material). The frequency of responses from all individuals is shown in Table S1 in the supplemental material.

Epitope targeting mediates the association between HLA and control. To examine the relative contributions of HLA class I alleles and targeting of T-cell epitopes to viral control, we used L1-regularized logistic regression (39) to predict viral control based on these features. To increase power, we grouped together EC and VC and modeled the HLA alleles at the HLA type (formerly known as "two-digit level"), with the exceptions of A*68 and B*15, whose subtypes span multiple supertypes (30), and B*35 (33), B*42 (61), and B*58 (31), whose subtypes have differential effects on control. Specifically, we used groups Px and PY for B*35 and modeled the remaining exceptions at the four-digit level.

Using receiver operator characteristic (ROC) curves to evaluate predictive performance (Fig. 1), we found that the combination of HLA class I alleles and targeting of T-cell epitopes better predicted viral control than HLA class I alleles alone ($P = 0.013$). In contrast, HLA class I alleles did not improve prediction of control beyond that of epitope targeting alone ($P = 0.95$). Thus, within our cohort, epitope targeting mediates the correlation between HLA expression and control. Similar results were observed when we used high-resolution HLA subtype information (Fig. 1, right), and the ROC curves were not improved when HLA homozygosity, HLA frequency, total epitope breadth, or proportion

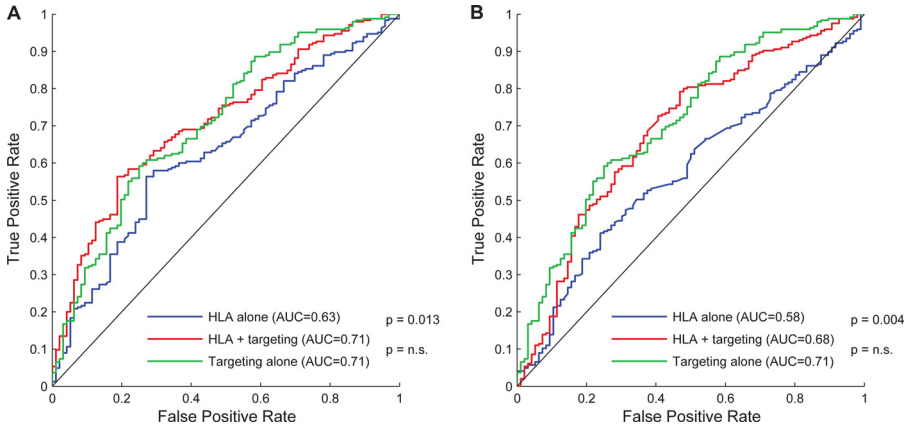

FIG 1 Ability of HLA types and epitope targeting to predict control. Shown are receiver operating characteristic (ROC) curves, plotting the true-positive rate versus false-positive rate of L1-regularized logistic regression models for the prediction of viral control on 341 subjects with chronic HIV infection. Models were trained using a mix of 2-digit and 4-digit resolution (left), as described in the text, and using all 4-digit types (right). The diagonal line represents uninformed predictions. The curve for predictions based on HLA restriction and epitope targeting deviates significantly from the curve for predictions based on HLA restriction alone, indicating significantly improved predictions when epitope targeting data are incorporated. Adding HLA typing information to epitope targeting does not improve predictions. *P* values are estimated using bootstrap estimates of the standard errors of the areas under the curves (AUC) and test adjacent curves as indicated.

of responses in Gag was included in the model (data not shown). Also, epitope targeting could not be used to predict EC versus VC, justifying our decision to group these individuals into a single controller group.

We further investigated which epitopes were individually correlated with viral control by using L1-regularized logistic regression to predict control based on HLA alleles and epitope targeting. We constructed 1,000 bootstrap replicates of the data and measured the frequency with which each epitope appeared in the model. The targeting of eight epitopes was positively and significantly associated with virus control (bootstrap frequency, >0.5): Gag KF11, Gag QW11, Gag AS10, Env RR11, Env RW9, Nef HW9, Pol GK9, and Rev SL9 (Table 2). In contrast, the targeting of 21 epitopes was significantly associated with chronic progression (Table 3). Note that while we use 0.5 for our bootstrap frequency threshold for significance, there is no general standard (see Tables 2 and 3 and Table S1 in the supplemental material for specific values).

TABLE 2 Epitope-specific $CD8^+$ T-cell epitopes associated with HIV control

Sequence	Protein	Name	Coordinates[a]		Effect size[b]	Bootstrap frequency[c]
			Start	Stop		
KAFSPEVIPMF	Gag (p24)	KF11	30	40	2.49	0.946
RLRDLLLIVTR	Env (gp41)	RR11	259	269	4.06	0.754
GIPHPAGLK	Pol (RT)	GK9	93	101	5.26	0.746
HTQGYFPDW	Nef	HW9	116	124	4.86	0.743
AEAMSQVTNS	Gag (p2)	AS10	1	10	4.30	0.638
SAEPVPLQL	Rev	SL9	67	75	2.57	0.626
QAISPRTLNAW	Gag (p24)	QW11	13	23	1.21	0.599
RIKQIINMW	Env (gp120)	RW9	419	427	2.90	0.513

[a] HXB2 coordinates with respect to protein domains (p24, RT, gp41, and gp120). See Table S1 in the supplemental material for coordinates with respect to polyproteins (Gag, Pol, Env).
[b] The natural logarithm of the odds ratio that an individual will be a progressor or not given that he or she targets this epitope. Parameters are estimated using a nonregularized logistic regression model including all HLAs and epitopes.
[c] The fraction of bootstrap samples in which the sequence is selected for inclusion in the predictive model.

When we repeated this analysis with exclusion of individuals who expressed the well-established protective B*57 and B*58:01 alleles, seven epitopes were positively and significantly associated with control, suggesting that the predictive effect of epitopes extends beyond the well-characterized B*57/58:01 alleles. These included two new epitopes (Env RI10 and Gag GL8, both of which were just below the threshold of significance in the initial analysis), while three previously associated epitopes lost their association (Gag KF11, Nef HW9, and Rev SL9).

Protective epitope targeting causes control, while nonprotective epitope targeting is the result of loss of control. One explanation of the predictive ability of targeting beyond HLA alleles is that targeting is a cause of viral control. Another explanation is that increased virus load resulting from lack of control leads to the targeting of additional epitopes (62). To explore the relative likelihood of these two explanations, we considered how well alternative statistical models uniquely consistent with each explanation predicted the data. Specifically, we first defined two groups of epitopes: those with a positive and negative trending association with viral control, where trends were determined using a (nonregularized) univariate logistic regression model with HLA alleles as covariates. We then fit each group of epitopes to two different graphical models: model A, a model in which HLA class I alleles determine viral control and this viral control, along with HLA restrictions, determines epitope targeting (Fig. 2A); and model B, a model in which HLA class I alleles determine epitope targeting and epitope targeting determines viral control (Fig. 2B). Model A is consistent with lack of control causing targeting, whereas model B is consistent with targeting causing control (see Materials and Methods). Out-of-sample prediction accuracy using 10-fold cross validation was then measured for epitope targeting and control. For the group of positively associated epitopes, model B produced better predictions than did model A according to a Wilcoxon matched-pairs signed-rank test ($P = 0.05$). In contrast, for the group of negatively associated epitopes, model A produced the

TABLE 3 Epitope-specific $CD8^+$ T-cell targeting associated with HIV chronic progression

Sequence	Protein	Name	Coordinates[a]		Effect size[b]	Bootstrap frequency[c]
			Start	Stop		
RPLVTIKI	Pol (Pr)	RI8	8	15	7.12	0.971
QELKNSAVSL	Env (gp41)	QL10	294	303	5.70	0.893
KYCWNLLQY	Env (gp41)	KQY9	283	291	3.94	0.855
RPMTYKAAV	Nef	RV9	77	85	6.71	0.836
IKLEPVHGVY	Pol (RT)	IY10	309	318	4.41	0.759
KQNPDIVIYQY	Pol (RT)	KY11	173	183	2.97	0.679
RAEQASQEV	Gag (p24)	RV9	173	181	3.13	0.677
VPLRPMTY	Nef	VY8	74	81	3.67	0.666
SFEPIPIHY	Env (gp120)	SY9	209	217	4.36	0.645
KAYETEVHNVW	Env (gp120)	KW11	59	69	2.57	0.635
LVGKLNWASQIY	Pol (RT)	LY12	260	271	3.27	0.614
LRSLFLFSY	Env(gp41)	LY9	249	257	3.13	0.596
SLYNTVATL	Gag (p17)	SL9	77	85	2.37	0.588
THLEGKIIL	Pol (Int)	TL9	66	74	4.14	0.564
IQQEFGIPY	Pol (Int)	IY9	135	143	3.18	0.548
ILKEPVHGV	Pol (RT)	IV9	309	317	1.63	0.539
GAETFYVDGA	Pol (RT)	GA10	436	445	2.76	0.527
VPVWKEATTTL	Env (gp120)	VL11	42	52	4.08	0.524
DTVLEEWNL	Pol (Pr)	DWL9	30	38	3.80	0.514
YTPGPGIRY	Nef	YY9	127	135	0.74	0.511
RRQDILDLWVY	Nef	RY11	105	115	2.04	0.505

[a] HXB2 coordinates with respect to protein domains (p24, RT, gp41, and gp120). See Table S1 in the supplemental material for coordinates with respect to polyproteins (Gag, Pol, Env).
[b] The natural logarithm of the odds ratio that an individual will be a progressor or not given that he or she targets this epitope. Parameters are estimated using a nonregularized logistic regression model including all HLAs and epitopes.
[c] The fraction of bootstrap samples in which the sequence is selected for inclusion in the predictive model.

better predictions ($P = 0.005$). These results suggest that targeting of the positively associated epitopes is a cause of viral control, whereas targeting of negatively associated epitopes is a causal effect of the lack of viral control. The latter observation is further supported by our observation that individuals with higher viral load target more epitopes and at higher intensity (see Fig. S2 in the supplemental material), consistent with a model in which a higher viral load leads to a high cell surface viral peptide concentration that stimulates further responses (similar observations with respect to variant recognition within Gag have been reported previously [62]). Note that these results suggest but do not prove cause and effect, as the models we examined were by no means exhaustive. For example, our models did not include latent variables. Furthermore, because we have sufficient statistical power to test only the positively and negatively correlated groups as a whole, our analysis does not rule out the possibility, e.g., that some negatively correlated epitopes cause lack of control. Here we refer to epitopes that are significantly associated with control as either protective (if the association is positive) or nonprotective (if the association is negative) epitopes.

Protective epitopes are distinguished by protein structures that constrain immune escape. Having uncovered the possibility of a cause-and-effect relationship between targeting specific epitopes and control, we next explored a possible mechanism to explain why some epitopes cause control and others do not. We hypothesized that a protective epitope is one where immune escape substantially weakens the virus. We further hypothesized that an escape at a given position in the protein would weaken the virus if that mutation altered the stability of the viral protein, as can be estimated *in silico* through a physics simulation of the interactions among the atoms in the protein (see Materials and Methods). We performed these simulations for every protein with an available high-resolution crystal structural (portions of p17, p24, protease, RT, integrase, Rev, Tat, and gp41) and from these results constructed a measure of protection that we call structural entropy. In particular, for each position in the protein, we mapped the predicted change in protein stability (energy) due to an amino acid substitution to the probability that the amino acid would be observed in the population of circulating viruses: the greater the change in stability, the lower the probability. We did so for each possible amino acid substitution (including wild type) at a given position, yielding a probability distribution over the 20 amino acids. Finally, we defined the structural entropy of a site to be the Shannon entropy of this probability distribution and the structural entropy of an epitope to be the average entropy over all sites in that epitope. Thus, the structural entropy of an epitope can be viewed as a hypothesized surrogate for how readily the virus can mutate in that region without harm to the virus. Note that structural entropy should not be confused with classical sequence entropy that is based on the observed frequencies of amino acids in a given population (see Discussion).

The structural entropies completely segregated protective from nonprotective epitopes ($P = 0.001$ by logistic regression, with covariates for each protein domain) (Fig. 3), with the exception of the protective Env RW9 epitope, which spans important protein-protein interaction interfaces that mediate coreceptor (63), CD4 (64, 65), and gp41 (63) binding. Results remained significant at a P value of <0.02 over all data sets resulting from the removal of one or two epitopes. Furthermore, the three protective epitopes that lay outside solved protein structures were similarly consistent with mutational constraint. In particular, gp41-RR10 lies on the N terminus of the lentiviral lytic peptide 2 (LLP-2) domain. Site-directed mutagenesis of position 1 of the epitope, which is associated with escape from A*31 and A*74 (55), has been shown to dramatically reduce cell-cell fusion and associated cell

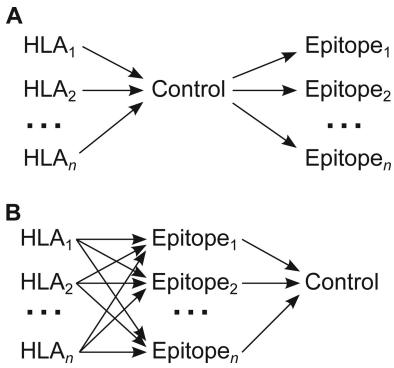

FIG 2 Graphical models used to investigate the direction of cause and effect between targeting and viral control. Nodes correspond to variables, and an arc from one node to another reflects the assumption that the first node is a cause of the second. Thus, model A reflects the assumptions that HLA types are causes of control and that HLA and (lack of) control are causes of epitope targeting. (Not shown in the figure but present in the model are additional arcs from HLA nodes to epitope nodes corresponding to the known influence of HLA on epitope binding.) Model B reflects the assumptions that HLA types are causes of targeting of specific epitopes, which in turn are causes of control.

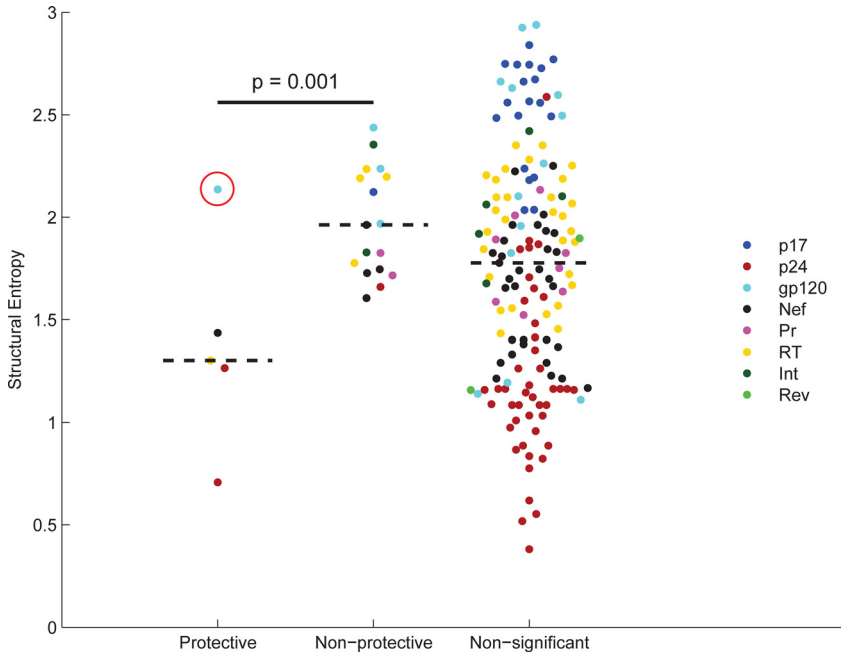

FIG 3 Structural entropy for protective, nonprotective, and nonsignificant epitopes. The structural entropy of an epitope can be viewed as an *in silico* estimate for how readily the virus can mutate in that region without harming the virus. Each point corresponds to a specific epitope from a protein (indicated by color). Epitopes are grouped as being protective, nonprotective, or nonsignificant based on the analysis of the ELISPOT data. There is a significant difference between the structural entropies of the protective and nonprotective epitopes ($P = 0.001$ using a logistic regression test in which the dependent variable is protective versus nonprotective and the independent variables are the structural entropy and the protein domain in which the epitope resides). The circled data point corresponds to the envelope epitope RW9, which spans important protein-protein interaction interfaces that mediate coreceptor, CD4, and gp41 binding. The reported *P* value includes this point.

death of $CD4^+$ T cells (66). Interestingly, Rev-SL9 is immediately upstream of gp41-RR10 and also overlaps the gp41 LLP-2 domain, though in a different reading frame. Finally, Gag-AS10 lies on the proteolytic cleavage site for the capsid-spacer 1 (p24-p2) boundary, consisting of p2 positions 1 to 10; cleavage appears to be largely mediated by the interaction of the protease and cleavage site structures (for a review, see reference 67), and mutations within the AS10 epitope have been shown to disrupt p2 cleavage and drastically reduce viral infectivity (68). These striking results suggest that a decrease in viral fitness due to escape after targeting is a possible causal mechanism for control.

Finally, we note that many of the epitopes tested have low structural entropies but were neither significantly protective nor nonprotective (Fig. 3). These epitopes may be protective but lower on the immunodominance hierarchy and hence not targeted frequently enough to show significant protection. Indeed, 35 out of 49 of these epitopes were targeted in fewer than 10% of our subjects (compared to an average targeting frequency of 23% for significant protective epitopes) and thus are good candidates for protective epitopes.

A majority of noncontrollers are capable of targeting protective epitopes. Although our analysis identified only eight protective epitopes with high enough targeting frequencies to be detected, 215 of 461 (47%) responses to these epitopes were in individuals who lacked the published restricting HLA allele, consistent with previous reports in which a subset of these protective epitopes were frequently targeted by controllers who did not express the published restricting allele (41). Such a high rate of undocumented targeting suggests that a vaccine including only these epitopes may cover a substantial proportion of the population. To further explore this possibility, we identified HLA-epitope associations using an iterative Fisher exact test that corrects for linkage disequilibrium (see Table S2 in the supplemental material and Materials and Methods), using a *q* value of <0.05 as a threshold to define HLA alleles associated with each epitope. Overall, 74/96 noncontrollers (77%) expressed at least one HLA allele that was linked with the presentation of a protective epitope, suggesting that a majority of individuals have the capacity to target these epitopes. Importantly, 340/341 individuals (99.7%) in the study express alleles that were statistically linked to nonprotective epitopes, highlighting the potential importance of vaccinating for specific protective epitopes and the exclusion of nonprotective epitopes. Although precise HLA-epitope restrictions will require further *in vitro* confirmation, these results highlight the potentially broad coverage that may be achieved by a vaccine that focuses on a small number of protective epitopes.

DISCUSSION

By comprehensively analyzing the IFN-γ ELISPOT response profiles of 245 natural controllers, compared to those of 96 progressors, against all known optimally defined epitopes, we obtained compelling evidence that a primary mechanism of control associated with HLA types is mediated by the targeting of specific epitopes, providing a broad justification for $CD8^+$ T-cell based vaccines and identifying specific epitopes that are primary candidates for such vaccines. The set of identified protective epitopes covered multiple proteins, including the Env protein, and was characterized by a heightened tendency for intraepitopic variation

to result in altered protein stability. These observations argue for the inclusion of specific epitopes in a vaccine, rather than whole viral proteins that may result in unfocused and unhelpful targeting, suggest that a hallmark of protective epitopes is a reduced capacity to escape the immune response through viral mutation, and provide an approach to identify additional protective epitopes.

Although the preponderance of evidence has pointed toward $CD8^+$ T-cell targeting as a causal mechanism behind HLA-mediated control (reviewed in reference 4), it has been difficult to prove this hypothesis. The most compelling evidence in humans has come from longitudinal studies, in which viral mutation led to loss of targeting of the immunodominant epitopes by the protective B27 (69–72), B57 (73), or B81 (74) HLA allele and was temporally followed by loss of control, though it is difficult to know if these results generalize to other alleles that are not characterized by such highly immunodominant targeting. In cross-sectional studies, one study provided some evidence that the targeting of a select set of epitopes in early infection improved prediction of CD4 decline beyond HLA typing information alone (21), while another study used a regression model that identified numerous overlapping peptide (OLP) targets as being predictive of reduced viral load in addition to HLA typing information (25), each suggesting that relative control among individuals expressing similar alleles can be better predicted given epitope targeting information. Here, we extended these analyses by examining all reported optimal epitopes in a large cohort of persons with chronic infection and through two formal statistical approaches. First, we showed that epitope-specific targeting mediates the correlation between HLA and control in this cohort, as epitope targeting significantly improves prediction of viremic control compared to HLA typing alone, while HLA type information adds no predictive value beyond that of epitope targeting. Second, an additional analysis revealed that a model in which HLA control was mediated through targeting of epitopes positively associated with control was significantly better at explaining the data than a model in which HLA alleles directly predicted control while targeting was a function of HLA restriction and control. Importantly, we observed the opposite result for epitopes negatively associated with control: targeting of such epitopes in general appeared to be the result of loss of control, while lack of control was associated with a greater breadth and higher intensity of epitope targeting. These results thus provide new evidence that a subset of epitopes are causally protective, including epitopes outside Gag, and are thus primary vaccine candidates, though future longitudinal studies in early infection cohorts are warranted to confirm these results.

If certain epitopes are causally protective and not simply markers of the presence of specific allelic types or haplotypes, then certain characteristics are likely to be common among protective epitopes. A natural potential mechanism involves viral immune escape. If CTL targeting is effective at limiting viral reproduction, then there will be significant selection pressure that will favor viral variants that ameliorate the immune response by, for example, disrupting epitope processing (75–77), HLA binding (70, 78), and/or TcR recognition (79–82). The importance of such immune escape is suggested by the extent to which specific HIV polymorphisms are correlated with the expression of specific HLA alleles (55, 83–85), as well as the observation that roughly a third of observed mutations in early infection are consistent with escape or reversion (86). Furthermore, protective alleles are distinguished by a tendency to select for escapes that abrogate HLA binding (55) and by a relatively high *in vitro* replicative cost of escape (24, 44), indicating that selection pressure is sufficiently high—and the potential sites of escape are sufficiently limited—to warrant the selection of otherwise unfavorable viral variants. An attractive hypothesis has therefore emerged: protective epitopes will be those for which escape options are constrained.

It has therefore been proposed that protective epitopes will be those in regions with low sequence diversity (20, 25, 28), as low sequence diversity generally implies that most variants reduce the replicative fitness of the virus enough to sustain purifying selection. Importantly, targeting of conserved regions has been correlated with a lower viral load (25), though we found no difference in sequence entropy between our protective and nonprotective epitopes (see Fig. S3 in the supplemental material). Furthermore, when we used both structural and sequence entropy to predict whether a significant epitope was protective or nonprotective, structural entropy remained significant ($P = 0.004$), while sequence entropy was not ($P = 0.85$). Although there were only five protective epitopes wholly contained within solved structures, the fact that each structure that contains a protective epitope also contains at least one nonprotective epitope allows for some robust comparisons. Indeed, when we repeated the analysis on all 264 combinations resulting from the removal of one or two significant epitopes, the maximum P value observed for structural entropy (conditioned on sequence entropy) was 0.03, while the minimum P value observed for sequence entropy (conditioned on structural entropy) was 0.59. One possible explanation for this difference may be the extreme phenotype we employed: the overall range in effect size in one study was roughly one log in viral load (25), whereas we are attempting to distinguish a more extreme controller phenotype. More simply, the lack of signal for sequence entropy in our data may reflect the smaller sample size of our data set, which is consistent with the weak correlation observed by Mothe and colleagues ($|r| < 0.15$) (25).

The weakness of the previously reported correlation, coupled with the lack of association in our data set, is somewhat surprising and suggests that sequence conservation is at best a weak marker for epitope protectiveness. One explanation is that while sequence conservation suggests constraints on viral evolution, lack of conservation does not rule out such constraints. Indeed, *in vivo*, the virus must balance the replicative fitness costs of a mutation with the fitness gained by abrogating the immune response. For a sufficiently strong immune response, it may well benefit the virus to select for escape despite the incursion of replicative fitness costs. The net result will be an increase in overall sequence diversity, driven primarily by the frequency of the HLA in the population. Furthermore, even if mutation is impossible for the virus in isolation, compensatory mutations in other sites may ameliorate the replicative fitness costs (87, 88). Such events will increase apparent sequence diversity but may still confer a protective advantage for certain epitopes: even though replicative fitness can be compensated, the requirement for multiple mutations increases the expected time until escape occurs and may provide a crucial window in early infection during which the immune system can effectively control the virus. A corollary to high sequence diversity resulting from effective epitope targeting is that high sequence conservation may simply be a marker for a typically ineffective immune re-

sponse that does not mount sufficient selection pressure to drive viral diversification.

Our measure of structural entropy addresses these potential discrepancies between the protective ability of an epitope and how well the epitope sequence is conserved. In particular, an epitope with low structural entropy is one for which the majority of mutations are expected to alter protein stability and thus have low replicative fitness (44). Although several of our epitopes were not located in available crystal structures, among those that were, structural entropy distinguished epitopes deemed protective by our ELISPOT data from those deemed nonprotective. The lone exception was a gp120 epitope that spans several protein-protein interaction domains (63–65) not measurable by the available structure but consistent with a constrained ability of the virus to mutate in this region. The ability of structural entropy to distinguish protective from nonprotective epitopes corroborates the causal role that targeting protective epitopes plays in natural viral control and is consistent with other data showing the existence of higher-order structural constraints on viral evolution, where targeting of these regions is associated with immune control (22).

One drawback of looking at natural T-cell targeting is that targeting of subdominant epitopes is, by definition, observed infrequently, making it difficult to identify subdominant epitopes that may be protective. This difficulty is particularly present among individuals who have a low level of circulating virus and therefore a low level of circulating T cells of the effector phenotype, which is what is typically measured on these assays (89). Our results indicate that mutational constraints in general, and structural entropy in particular, constitute a feature that is useful in identifying potentially subdominant protective epitopes for further study. Indeed, an important recent result indicates that vaccination against specific epitopes can substantially alter immunodominance hierarchies (90), even if administered after the onset of infection (91), validating the approach of designing a prophylactic or therapeutic vaccine to elicit such (hypothetical) protective subdominant targeting.

An alternative to our approach of measuring responses to known epitopes would have been to perform an overlapping peptide (OLP) scan of the entire HIV proteome. Limitations of this approach include the reduction of specificity, as each OLP may contain more than one epitope, and an increase in false-negative rates due to the oversize peptides and required processing (23, 62), requiring additional follow-up to identify the protective epitopes. Two advantages of the OLP approach are that it provides an unbiased scan of HIV peptides and covers the entire proteome. Nevertheless, our *in silico* measure of structural entropy provides the opportunity to scan the entire proteome and identify additional peptide sequences for testing.

Our identification of specific protective and nonprotective epitopes argues for the inclusion (and exclusion) of specific epitopes in a vaccine, an approach that is currently in development by several groups (20, 25, 28), and argues against expression of entire viral proteins. Finally, our observation that protective epitopes are distinguished by low structural entropy broadly confirms the hypothesis that protective epitopes are those for which viral escape is constrained and provides a means to identify additional subdominant epitopes that are potentially protective and thus candidates for inclusion in a vaccine.

ACKNOWLEDGMENTS

This study was supported in part by the International HIV Controllers Study (IHCS), funded by the Bill and Melinda Gates Foundation (F.P. and B.D.W.), the AIDS Healthcare Foundation (F.P.), and the Harvard University Center for AIDS Research (CFAR), an NIH-funded Center for AIDS Research (P30 AI060354), which is supported by the following NIH Co-Funding and Participating Institutes and Centers: NIAID, NCI, NICHD, NHLBI, NIDA, NIMH, NIA, FIC, and OAR. Members of IHCS can be found at http://ragoninstitute.org/hivcontrollers/health-professionals/study-members/the-international-hiv-controllers-study-members/. We also gratefully acknowledge support from the Mark and Lisa Schwartz Foundation. This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under contract no. HHSN261200800001E. This research was supported in part by the Intramural Research Program of the NIH, Frederick National Laboratory, Center for Cancer Research.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

David Heckerman, Jonathan M. Carlson, Carl Kadie, and Charles E. DeZiel were employed by Microsoft while performing this research.

REFERENCES

1. **Kaslow RA, Carrington M, Apple R, Park L, Muñoz A, Saah AJ, Goedert JJ, Winkler C, O'Brien SJ, Rinaldo C, Detels R, Blattner W, Phair J, Erlich H, Mann DL.** 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat. Med. **2:**405–411. http://dx.doi.org/10.1038/nm0496-405.
2. **Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K, O'Brien SJ.** 1999. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science **283:**1748–1752. http://dx.doi.org/10.1126/science.283.5408.1748.
3. **Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal SA, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB.** 2007. A whole-genome association study of major determinants for host control of HIV-1. Science **317:**944–947. http://dx.doi.org/10.1126/science.1143767.
4. **Goulder PJR, Walker BD.** 2012. HIV and HLA class I: an evolving relationship. Immunity **37:**426–440. http://dx.doi.org/10.1016/j.immuni.2012.09.005.
5. **Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PIW, Walker BD, Ripke S, Brumme CJ, Pulit SL, Carrington M, Kadie CM, Carlson JM, Heckerman D, Graham RR, Plenge RM, Deeks SG, Gianniny L, Crawford G, Sullivan J, Gonzalez E, Davies L, Camargo A, Moore JM, Beattie N, Gupta S, Crenshaw A, Burtt NP, Guiducci C, Gupta N, Gao X, Qi Y, Yuki Y, Piechocka-Trocha A, Cutrell E, Rosenberg R, Moss KL, Lemay P, O'Leary J, Schaefer T, Verma P, Toth I, Block B, Baker B, Rothchild A, Lian J, Proudfoot J, Alvino DML, Vine S, Addo MM, Allen TM, et al.** 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science **330:**1551–1557. http://dx.doi.org/10.1126/science.1195271.
6. **Schmitz JE, 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. http://dx.doi.org/10.1126/science.283.5403.857.
7. **Migueles SA, Laborico AC, Shupert WL, Sabbaghian MS, Rabin R, Hallahan CW, Van Baarle D, Kostense S, Miedema F, McLaughlin M, Ehler L, Metcalf J, Liu S, Connors M.** 2002. HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. Nat. Immunol. **3:**1061–1068. http://dx.doi.org/10.1038/ni845.
8. **Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, Reddy S, Mackey EW, Miller JD, Leslie AJ, DePierres C, Mncube Z, Duraiswamy J, Zhu B, Eichbaum Q, Altfeld M, Wherry EJ, Coovadia HM, Goulder PJR, Klenerman P, Ahmed R, Freeman GJ, Walker BD.**

2006. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature **443:**350–354. http://dx.doi.org/10.1038/nature05115.

9. **Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, Bessette B, Boulassel M-R, Delwart E, Sepulveda H, Balderas RS, Routy J-P, Haddad EK, Sekaly R-P.** 2006. Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. Nat. Med. **12:**1198–1202. http://dx.doi.org/10.1038/nm1482.
10. **Sáez-Cirión A, Lacabaratz C, Lambotte O, Versmisse P, Urrutia A, Boufassa F, Barré-Sinoussi F, Delfraissy J-F, Sinet M, Pancino G, Venet A.** 2007. HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. Proc. Natl. Acad. Sci. U. S. A. **104:**6776–6781. http://dx.doi.org/10.1073/pnas.0611244104.
11. **Chen H, Ndhlovu ZM, Liu D, Porter LC, Fang JW, Darko S, Brockman MA, Miura T, Brumme ZL, Schneidewind A, Piechocka-Trocha A, Cesa KT, Sela J, Cung TD, Toth I, Pereyra F, Yu XG, Douek DC, Kaufmann DE, Allen TM, Walker BD.** 2012. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. Nat. Immunol. **13:**691–700. http://dx.doi.org/10.1038/ni.2342.
12. **Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D, Gilbert PB, Lama JR, Marmor M, Del Rio C, McElrath MJ, Casimiro DR, Gottesdiener KM, Chodakewitz JA, Corey L, Robertson MN.** 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet **372:**1881–1893. http://dx.doi.org/10.1016/S0140-6736(08)61591-3.
13. **Hammer SM, Sobieszczyk ME, Janes H, Karuna ST, Mulligan MJ, Grove D, Koblin BA, Buchbinder SP, Keefer MC, Tomaras GD, Frahm N, Hural J, Anude C, Graham BS, Enama ME, Adams E, DeJesus E, Novak RM, Frank I, Bentley C, Ramirez S, Fu R, Koup RA, Mascola JR, Nabel GJ, Montefiori DC, Kublin J, McElrath MJ, Corey L, Gilbert PB.** 2013. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. N. Engl. J. Med. **369:**2083–2092. http://dx.doi.org/10.1056/NEJMoa1310566.
14. **Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, Reddy S, de Pierres C, Mncube Z, Mkhwanazi N, Bishop K, van der Stok M, Nair K, Khan N, Crawford H, Payne R, Leslie AJ, Prado J, Prendergast A, Frater J, McCarthy N, Brander C, Learn GH, Nickle D, Rousseau C, Coovadia H, Mullins JI, Heckerman D, Walker BD, Goulder PJR.** 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat. Med. **13:**46–53. http://dx.doi.org/10.1038/nm1520.
15. **Rolland M, Heckerman D, Deng W, Rousseau CM, Coovadia H, Bishop K, Goulder PJR, Walker BD, Brander C, Mullins JI.** 2008. Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. PLoS One **3:**e1424. http://dx.doi.org/10.1371/journal.pone.0001424.
16. **Huang K-HG, Goedhals D, Carlson JM, Brockman MA, Mishra S, Brumme ZL, Hickling S, Tang CSW, Miura T, Seebregts C, Heckerman D, Ndung'u T, Walker BD, Klenerman P, Steyn D, Goulder PJR, Phillips R, van Vuuren C, Frater J.** 2011. Progression to AIDS in South Africa is associated with both reverting and compensatory viral mutations. PLoS One **6:**e19018. http://dx.doi.org/10.1371/journal.pone.0019018.
17. **Zuñiga R, Lucchetti A, Galvan P, Sanchez C, Hernandez A, Sanchez H, Frahm N, Linde CH, Hewitt HS, Altfeld M, Allen TM, Walker BD, Korber BT, Leitner T, Sanchez J, Brander C, Sanchez S, Hildebrand W.** 2006. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. J. Virol. **80:**3122–3125. http://dx.doi.org/10.1128/JVI.80.6.3122-3125.2006.
18. **Sacha JB, Chung C, Rakasz EG, Spencer SP, Jonas AK, Bean AT, Lee W, Burwitz BJ, Stephany JJ, Loffredo JT, Allison DB, Adnan S, Hoji A, Wilson NA, Friedrich TC, Lifson JD, Yang OO, Watkins DI.** 2007. Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. J. Immunol. **178:**2746–2754. http://dx.doi.org/10.4049/jimmunol.178.5.2746.
19. **Heckerman DE, Frahm N, Pereyra F, Dubey S, Geraghty DE, Carlson JM, Robertson MN, McElrath MJ, Casimiro DR, Walker BD.** 2009. Vaccine-induced targeting of epitopes associated with spontaneous control of HIV viral replication is assocaited with lower set-point viral loads in HIV-infected participants from the STEP trial, p 95. *In* 16th Conf. Retroviruses Opportun. Infect., Montreal, Canada.
20. **Rolland M, Nickle DC, Mullins JI.** 2007. HIV-1 group M conserved elements vaccine. PLoS Pathog. **3:**e157. http://dx.doi.org/10.1371/journal.ppat.0030157.
21. **Dinges WL, Richardt J, Friedrich D, Jalbert E, Liu Y, Stevens CE, Maenza J, Collier AC, Geraghty DE, Smith J, Moodie Z, Mullins JI, McElrath MJ, Horton H.** 2010. Virus-specific CD8+ T-cell responses better define HIV disease progression than HLA genotype. J. Virol. **84:**4461–4468. http://dx.doi.org/10.1128/JVI.02438-09.
22. **Dahirel V, Shekhar K, Pereyra F, Miura T, Artyomov M, Talsania S, Allen TM, Altfeld M, Carrington M, Irvine DJ, Walker BD, Chakraborty AK.** 2011. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. Proc. Natl. Acad. Sci. U. S. A. **108:**11530–11535. http://dx.doi.org/10.1073/pnas.1105315108.
23. **Mothe B, Llano A, Ibarrondo J, Zamarreño J, Schiaulini M, Miranda C, Ruiz-Riol M, Berger CT, Herrero MJ, Palou E, Plana M, Rolland M, Khatri A, Heckerman D, Pereyra F, Walker BD, Weiner D, Paredes R, Clotet B, Felber BK, Pavlakis GN, Mullins JI, Brander C.** 2012. CTL responses of high functional avidity and broad variant cross-reactivity are associated with HIV control. PLoS One **7:**e29717. http://dx.doi.org/10.1371/journal.pone.0029717.
24. **Ferguson AL, Mann JK, Omarjee S, Ndung'u T, Walker BD, Chakraborty AK.** 2013. Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. Immunity **38:**606–617. http://dx.doi.org/10.1016/j.immuni.2012.11.022.
25. **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. http://dx.doi.org/10.1186/1479-5876-9-208.
26. **Navis M, Schellens IMM, van Swieten P, Borghans JAM, Miedema F, Kootstra NA, van Baarle D, Schuitemaker H.** 2008. A nonprogressive clinical course in HIV-infected individuals expressing human leukocyte antigen B57/5801 is associated with preserved CD8+ T lymphocyte responsiveness to the HW9 epitope in Nef. J. Infect. Dis. **197:**871–879. http://dx.doi.org/10.1086/528695.
27. **Adland E, Carlson JM, Paioni P, Kløverpris H, Shapiro R, Ogwu A, Riddell L, Luzzi G, Chen F, Balachandran T, Heckerman D, Stryhn A, Edwards A, Ndung'u T, Walker BD, Buus S, Goulder P, Matthews PC.** 2013. Nef-specific CD8+ T cell responses contribute to HIV-1 immune control. PLoS One **8:**e73117. http://dx.doi.org/10.1371/journal.pone.0073117.
28. **Létourneau S, Im E-J, Mashishi T, Brereton C, Bridgeman A, Yang H, Dorrell L, Dong T, Korber B, McMichael AJ, Hanke T.** 2007. Design and pre-clinical evaluation of a universal HIV-1 vaccine. PLoS One **2:**e984. http://dx.doi.org/10.1371/journal.pone.0000984.
29. **Hansen JA (ed).** 2005. Immunobiology of the human MHC: 13th International Histocompatibility Workshop and Congress. IHWG Press, Seattle, WA.
30. **Sidney J, Peters B, Frahm N, Brander C, Sette A.** 2008. HLA class I supertypes: a revised and updated classification. BMC Immunol. **9:**1. http://dx.doi.org/10.1186/1471-2172-9-1.
31. **Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, Rathnavalu P, Moore C, Pfafferott KJ, Hilton L, Zimbwa P, Moore S, Allen T, Brander C, Addo MM, Altfeld M, James I, Mallal SA, Bunce M, Barber LD, Szinger J, Day C, Klenerman P, Mullins JI, Korber B, Coovadia HM, Walker BD, Goulder PJR.** 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature **432:**769–775. http://dx.doi.org/10.1038/nature03113.
32. **Kloverpris HN, Stryhn A, Harndahl M, van der Stok M, Payne RP, Matthews PC, Chen F, Riddell L, Walker BD, Ndung'u T, Buus S, Goulder PJR.** 2012. HLA-B*57 micropolymorphism shapes HLA allele-specific epitope immunogenicity, selection pressure, and HIV immune control. J. Virol. **86:**919–929. http://dx.doi.org/10.1128/JVI.06150-11.
33. **Gao X, Nelson GW, Karacki P, Martin MP, Phair J, Kaslow R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, O'Brien SJ, Carrington M.** 2001. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. N. Engl. J. Med. **344:**1668–1675. http://dx.doi.org/10.1056/NEJM200105313442203.
34. **Kløverpris HN, Harndahl M, Leslie AJ, Carlson JM, Ismail N, van der Stok M, Huang K-HG, Chen F, Riddell L, Steyn D, Goedhals D, van Vuuren C, Frater J, Walker BD, Carrington M, Ndung'u T, Buus S, Goulder PJR.** 2012. HIV control through a single nucleotide on the

HLA-B locus. J. Virol. **86:**11493–11500. http://dx.doi.org/10.1128/JVI.01020-12.

35. **Frahm N, Baker B, Brander C.** 2008. Identification and optimal definition of HIV-derived cytotoxic T lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution, p 3–24. *In* Korber BT, Brander C, Haynes BF, Koup RA, Moore JP, Walker BD, Watkins DI (ed), HIV molecular immunology 2008. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM.
36. **Addo MM, Yu XG, Rathod A, Cohen D, Eldridge RL, Strick D, Johnston MN, Corcoran C, Wurcel AG, Fitzpatrick CA, Feeney ME, Rodriguez WR, Basgoz N, Draenert R, Stone DR, Brander C, Goulder PJR, Rosenberg ES, Altfeld M, Walker BD, Irol JV.** 2003. Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses but no correlation to viral load. J. Virol. **77:**2081–2092. http://dx.doi.org/10.1128/JVI.77.3.2081-2092.2003.
37. **Feeney ME, Roosevelt KA, Tang Y, Pfafferott KJ, McIntosh K, Burchett SK, Mao C, Walker BD, Goulder PJR.** 2003. Comprehensive screening reveals strong and broadly directed human immunodeficiency virus type 1-specific CD8 responses in perinatally infected children. J. Virol. **77:**7492–7501. http://dx.doi.org/10.1128/JVI.77.13.7492-7501.2003.
38. **Soghoian DZ, Jessen H, Flanders M, Sierra-Davidson K, Cutler S, Pertel T, Ranasinghe S, Lindqvist M, Davis I, Lane K, Rychert J, Rosenberg ES, Piechocka-Trocha A, Brass AL, Brenchley JM, Walker BD, Streeck H.** 2012. HIV-specific cytolytic CD4 T cell responses during acute HIV infection predict disease outcome. Sci. Transl. Med. **4:**123ra25. http://dx.doi.org/10.1126/scitranslmed.3003165.
39. **Andrew G, Gao J.** 2007. Scalable training of L1-regularized log-linear models, p 33–40. *In* Proceedings of the 24th International Conference on Machine Learning. ACM Press, New York, NY.
40. **Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M.** 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics **12:**77. http://dx.doi.org/10.1186/1471-2105-12-77.
41. **Frahm N, Yusim K, Suscovich TJ, Adams S, Sidney J, Hraber P, Hewitt HS, Linde CH, Kavanagh DG, Woodberry T, Henry LM, Faircloth K, Listgarten J, Kadie C, Jojic N, Sango K, Brown NV, Pae E, Zaman MT, Bihl F, Khatri A, John M, Mallal SA, Marincola FM, Walker BD, Sette A, Heckerman D, Korber BT, Brander C.** 2007. Extensive HLA class I allele promiscuity among viral CTL epitopes. Eur. J. Immunol. **37:**2419–2433. http://dx.doi.org/10.1002/eji.200737365.
42. **Carlson JM, Heckerman D.** 2009. Estimating false discovery rates for contingency tables. Microsoft Res. MSR-TR-2009-53.
43. **Spirtes P, Glymour CN, Scheines R.** 2000. Causation, prediction, and search. MIT Press, Cambridge, MA.
44. **Boutwell CL, Carlson JM, Lin T-H, 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. http://dx.doi.org/10.1128/JVI.03233-12.
45. **Kelly BN, Howard BR, Wang H, Robinson H, Sundquist WI, Hill CP.** 2006. Implications for viral capsid assembly from crystal structures of HIV-1 Gag(1-278) and CA(N)(133-278). Biochemistry **45:**11257–11266. http://dx.doi.org/10.1021/bi060927x.
46. **Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M.** 2009. X-ray structures of the hexameric building block of the HIV capsid. Cell **137:**1282–1292. http://dx.doi.org/10.1016/j.cell.2009.04.063.
47. **Robbins AH, Coman RM, Bracho-Sanchez E, Fernandez MA, Gilliland CT, Li M, Agbandje-McKenna M, Wlodawer A, Dunn BM, McKenna R.** 2010. Structure of the unbound form of HIV-1 subtype A protease: comparison with unbound forms of proteases from other HIV subtypes. Acta Crystallogr. D Biol. Crystallogr. **66:**233–242. http://dx.doi.org/10.1107/S0907444909054298.
48. **Hsiou Y, Ding J, Das K, Clark AD, Hughes SH, Arnold E.** 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure **4:**853–860. http://dx.doi.org/10.1016/S0969-2126(96)00091-3.
49. **Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR.** 1998. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proc. Natl. Acad. Sci. U. S. A. **95:**9150–9154. http://dx.doi.org/10.1073/pnas.95.16.9150.
50. **Tahirov TH, Babayeva ND, Varzavand K, Cooper JJ, Sedore SC, Price DH.** 2010. Crystal structure of HIV-1 Tat complexed with human P-TEFb. Nature **465:**747–751. http://dx.doi.org/10.1038/nature09131.
51. **Daugherty MD, Liu B, Frankel AD.** 2010. Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. Nat. Struct. Mol. Biol. **17:**1337–1342. http://dx.doi.org/10.1038/nsmb.1902.
52. **Lee CH, Saksela K, Mirza UA, Chait BT, Kuriyan J.** 1996. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell **85:**931–942. http://dx.doi.org/10.1016/S0092-8674(00)81276-3.
53. **Pancera M, Majeed S, Ban Y-EA, Chen L, Huang C, Kong L, Kwon Do Y, Stuckey J, Zhou T, Robinson JE, Schief WR, Sodroski J, Wyatt R, Kwong PD.** 2010. Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. Proc. Natl. Acad. Sci. U. S. A. **107:**1166–1171. http://dx.doi.org/10.1073/pnas.0911004107.
54. **Guerois R, Nielsen JE, Serrano L.** 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J. Mol. Biol. **320:**369–387. http://dx.doi.org/10.1016/S0022-2836(02)00442-4.
55. **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.** 2012. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. J. Virol. **86:**13202–13216. http://dx.doi.org/10.1128/JVI.01998-12.
56. **Deeks SG, Walker BD.** 2007. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. Immunity **27:**406–416. http://dx.doi.org/10.1016/j.immuni.2007.08.010.
57. **Frahm N, Kiepiela P, Adams S, Linde CH, Hewitt HS, Sango K, Feeney ME, Addo MM, Lichterfeld M, Lahaie MP, Pae E, Wurcel AG, Roach T, St John MA, Altfeld M, Marincola FM, Moore C, Mallal SA, Carrington M, Heckerman D, Allen TM, Mullins JI, Korber BT, Goulder PJR, Walker BD, Brander C.** 2006. Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes. Nat. Immunol. **7:**173–178. http://dx.doi.org/10.1038/ni1281.
58. **Sabbaj S, Bansal A, Ritter GD, Perkins C, Edwards BH, Gough E, Tang J, Szinger JJ, Korber B, Wilson CM, Kaslow, RA, Mulligan MJ, Goepfert PA.** 2003. Cross-reactive CD8+ T cell epitopes identified in US adolescent minorities. J. Acquir. Immune Defic. Syndr. **33:**426–438. http://dx.doi.org/10.1097/00126334-200308010-00003.
59. **Kaufmann DE, Bailey PM, Sidney J, Wagner B, Norris PJ, Johnston MN, Cosimi LA, Addo MM, Lichterfeld M, Altfeld M, Frahm N, Brander C, Sette A, Walker BD, Rosenberg ES.** 2004. Comprehensive analysis of human immunodeficiency virus type 1-specific CD4 responses reveals marked immunodominance of gag and nef and the presence of broadly recognized peptides. J. Virol. **78:**4463–4477. http://dx.doi.org/10.1128/JVI.78.9.4463-4477.2004.
60. **Ranasinghe S, Flanders M, Cutler S, Soghoian DZ, Ghebremichael M, Davis I, Lindqvist M, Pereyra F, Walker BD, Heckerman D, Streeck H.** 2012. HIV-specific CD4 T cell responses to different viral proteins have discordant associations with viral load and clinical outcome. J. Virol. **86:**277–283. http://dx.doi.org/10.1128/JVI.05577-11.
61. **Kløverpris HN, Harndahl M, Leslie AJ, Carlson JM, Ismail N, van der Stok M, Huang K-HG, Chen F, Riddell L, Steyn D, Goedhals D, van Vuuren C, Frater J, Walker BD, Carrington M, Ndung'u T, Buus S, Goulder P.** 2012. HIV control through a single nucleotide on the HLA-B locus. J. Virol. **86:**11493–11500. http://dx.doi.org/10.1128/JVI.01020-12.
62. **Sunshine J, Kim M, Carlson JM, Heckerman D, Czartoski J, Migueles a, Maenza J, McElrath MJ, Mullins JI, Frahm N.** 2014. Increased sequence coverage through combined targeting of variant and conserved epitopes correlates with control of HIV replication. J. Virol. **88:**1354–1365. http://dx.doi.org/10.1128/JVI.02361-13.
63. **Korber B, Gnanakaran S.** 2009. The implications of patterns in HIV diversity for neutralizing antibody induction and susceptibility. Curr. Opin. HIV AIDS **4:**408–417. http://dx.doi.org/10.1097/COH.0b013e32832f129e.
64. **Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA.** 1998. Structure of an HIV gp120 envelope glycoprotein in complex

with the CD4 receptor and a neutralizing human antibody. Nature **393:** 648–659. http://dx.doi.org/10.1038/31405.

65. **Zhou T, Georgiev I, Wu X, Yang Z-Y, Dai K, Finzi A, Kwon Do Y, Scheid JF, Shi W, Xu L, Yang Y, Zhu J, Nussenzweig MC, Sodroski J, Shapiro L, Nabel GJ, Mascola JR, Kwong PD.** 2010. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science **329:**811–817. http://dx.doi.org/10.1126/science.1192819.
66. **Kalia V, Sarkar S, Gupta P, Montelaro C, Montelaro RC.** 2003. Rational site-directed mutations of the LLP-1 and LLP-2 lentivirus lytic peptide domains in the intracytoplasmic tail of human immunodeficiency virus type 1 gp41 indicate common functions in cell-cell fusion but distinct roles in virion envelope incorporation. J. Virol. **77:**3634–3646. http://dx.doi.org/10.1128/JVI.77.6.3634-3646.2003.
67. **Lee S-K, Potempa M, Swanstrom R.** 2012. The choreography of HIV-1 proteolytic processing and virion assembly. J. Biol. Chem. **287:**40867–40874. http://dx.doi.org/10.1074/jbc.R112.399444.
68. **Checkley MA, Luttge BG, Soheilian F, Nagashima K, Freed EO.** 2010. The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. Virology **400:**137–144. http://dx.doi.org/10.1016/j.virol.2010.01.028.
69. **Goulder PJR, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA, Giangrande P, Luzzi G, Morgana B, Edwards A, McMichael AJ, Rowland-Jones S.** 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat. Med. **3:**212–217. http://dx.doi.org/10.1038/nm0297-212.
70. **Kelleher AD, Long C, Holmes EC, Allen RL, Wilson J, Conlon C, Workman C, Shaunak S, Olson K, Goulder PJR, Brander C, Ogg G, Sullivan JS, Dyer W, Jones I, McMichael AJ, Rowland-Jones S, Phillips RE.** 2001. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. J. Exp. Med. **193:**375–386. http://dx.doi.org/10.1084/jem.193.3.375.
71. **Feeney ME, Tang Y, Roosevelt KA, Leslie AJ, Mcintosh K, Karthas N, Walker BD, Goulder PJR.** 2004. Immune escape precedes breakthrough human immunodeficiency virus type 1 viremia and broadening of the cytotoxic T-lymphocyte response in an HLA-B27-positive long-term-nonprogressing child. J. Virol. **78:**8927–8930. http://dx.doi.org/10.1128/JVI.78.16.8927-8930.2004.
72. **Ammaranond P, van Bockel DJ, Petoumenos K, McMurchie M, Finlayson R, Middleton MG, Davenport MP, Venturi V, Suzuki K, Gelgor L, Kaldor JM, Cooper DA, Kelleher AD.** 2011. HIV immune escape at an immunodominant epitope in HLA-B*27-positive individuals predicts viral load outcome. J. Immunol. **186:**479–488. http://dx.doi.org/10.4049/jimmunol.0903227.
73. **Crawford H, Lumm W, Leslie AJ, Schaefer M, Boeras D, Prado JG, Tang J, Farmer P, Ndung'u T, Lakhi S, Gilmour J, Goepfert P, Walker BD, Kaslow R, Mulenga J, Allen S, Goulder PJR, Hunter E.** 2009. Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. J. Exp. Med. **206:** 909–921. http://dx.doi.org/10.1084/jem.20081984.
74. **Ntale RS, Chopera DR, Ngandu NK, Assis de Rosa D, Zembe L, Gamieldien H, Mlotshwa M, Werner L, Woodman Z, Mlisana K, Abdool Karim S, Gray CM, Williamson C.** 2012. Temporal association of HLA-B*81:01 and B*39:10 mediated HIV-1 p24 sequence evolution with disease progression. J. Virol. **86:**12013–24. http://dx.doi.org/10.1128/JVI.00539-12.
75. **Draenert R, Brander C, Yu XG, Altfeld M, Verrill CL, Feeney ME, Walker BD, Goulder PJR.** 2004. Impact of intrapeptide epitope location on CD8 T cell recognition: implications for design of overlapping peptide panels. AIDS **18:**871. http://dx.doi.org/10.1097/00002030-200404090-00004.
76. **Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, Kojima A, Nagai Y, Iwamoto A, Matsuda Z, Ariyoshi K.** 2004. Impaired processing and presentation of cytotoxic-T-lymphocyte (CTL) epitopes are major escape mechanisms from CTL immune pressure in human immunodeficiency virus type 1 infection. J. Virol. **78:**1324–1332. http://dx.doi.org/10.1128/JVI.78.3.1324-1332.2004.
77. **Zhang SC, Martin E, Shimada M, Godfrey SB, Fricke J, Locastro S, Lai NY, Liebesny P, Carlson JM, Brumme CJ, Ogbechie OA, Chen H, Walker BD, Brumme ZL, Kavanagh DG, Le Gall S.** 2012. Aminopeptidase substrate preference affects HIV epitope presentation and predicts immune escape patterns in HIV-infected individuals. J. Immunol. **188:** 5924–5934. http://dx.doi.org/10.4049/jimmunol.1200219.
78. **Ammaranond P, Zaunders J, Satchell C, van Bockel D, Cooper DA, Kelleher AD.** 2005. A new variant cytotoxic T lymphocyte escape mutation in HLA-B27-positive individuals infected with HIV type 1. AIDS Res. Hum. Retroviruses **21:**395–397. http://dx.doi.org/10.1089/aid.2005.21.395.
79. **Cale EM, Bazick HS, Rianprakaisang TA, Alam SM, Letvin NL.** 2011. Mutations in a dominant Nef epitope of simian immunodeficiency virus diminish TCR:epitope peptide affinity but not epitope peptide: MHC class I binding. J. Immunol. **187:**3300–3313. http://dx.doi.org/10.4049/jimmunol.1101080.
80. **Theodossis A, Guillonneau C, Welland A, Ely LK, Clements CS, Williamson NA, Webb AI, Wilce Ja Mulder RJ, Dunstone MA, Doherty PC, McCluskey J, Purcell AW, Turner SJ, Rossjohn J.** 2010. Constraints within major histocompatibility complex class I restricted peptides: presentation and consequences for T-cell recognition. Proc. Natl. Acad. Sci. U. S. A. **107:**5534–5539. http://dx.doi.org/10.1073/pnas.1000032107.
81. **Iversen AKN, Stewart-Jones G, Learn GH, Christie N, Sylvester-Hviid C, Armitage AE, Kaul R, Beattie T, Lee JK, Li Y, Chotiyarnwong P, Dong T, Xu X, Luscher MA, MacDonald K, Ullum H, Klarlund-Pedersen B, Skinhøj P, Fugger L, Buus S, Mullins JI, Jones EY, van der Merwe PA, McMichael AJ.** 2006. Conflicting selective forces affect T cell receptor contacts in an immunodominant human immunodeficiency virus epitope. Nat. Immunol. **7:**179–189. http://dx.doi.org/10.1038/ni1298.
82. **Iglesias MC, Almeida JR, Fastenackels S, van Bockel DJ, Hashimoto M, Venturi V, Gostick E, Urrutia A, Wooldridge L, Clement M, Gras S, Wilmann PG, Autran B, Moris A, Rossjohn J, Davenport MP, Takiguchi M, Brander C, Douek DC, Kelleher AD, Price DA, Appay V.** 2011. Escape from highly effective public CD8+ T-cell clonotypes by HIV. Blood **118:**2138–2149. http://dx.doi.org/10.1182/blood-2011-01-328781.
83. **Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker BD, Ndung'u T, Shapiro R, Frater J, Brumme ZL, Goulder PJR, Heckerman D.** 2012. Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. J. Virol. **86:**5230–5243. http://dx.doi.org/10.1128/JVI.06728-11.
84. **Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson JM, Kadie C, Bhattacharya T, Chui C, Szinger J, Mo T, Hogg RS, Montaner JSG, Frahm N, Brander C, Walker BD, Harrigan PR.** 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. PLoS Pathog. **3:**e94. http://dx.doi.org/10.1371/journal.ppat.0030094.
85. **Bartha I, Carlson JM, Brumme CJ, McLaren PJ, Brumme ZL, John M, Haas DW, Martinez-Picado J, Dalmau J, López-Galíndez C, Casado C, Rauch A, Günthard HF, Bernasconi E, Vernazza P, Klimkait T, Yerly S, O'Brien SJ, Listgarten J, Pfeifer N, Lippert C, Fusi N, Kutalik Z, Allen TM, Müller V, Harrigan PR, Heckerman D, Telenti A, Fellay J.** 2013. A genome-to-genome analysis of associations between human genetic variation, HIV-1 sequence diversity, and viral control. eLife **2:**e01123. http://dx.doi.org/10.7554/eLife.01123.
86. **Brumme ZL, Brumme CJ, Carlson JM, Streeck H, John M, Eichbaum Q, Block BL, Baker B, Kadie C, Markowitz M, Jessen H, Kelleher AD, Rosenberg E, Kaldor J, Yuki Y, Carrington M, Allen TM, Mallal SA, Altfeld M, Heckerman D, Walker BD.** 2008. 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. **82:**9216–9227. http://dx.doi.org/10.1128/JVI.01041-08.
87. **Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, Desouza I, Ryvkin F, Derdeyn CA, Allen S, Hunter E, Mulenga J, Goepfert PA, Walker BD, Allen TM.** 2007. 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. **81:**12608–12618. http://dx.doi.org/10.1128/JVI.01369-07.
88. **Schneidewind A, Brumme ZL, Brumme CJ, Power KA, Reyor LL, O'Sullivan K, Gladden A, Hempel U, Kuntzen T, Wang YE, Oniangue-Ndza C, Jessen H, Markowitz M, Rosenberg ES, Sékaly R-P, Kelleher AD, Walker BD, Allen TM.** 2009. Transmission and long-term stability of compensated CD8 escape mutations. J. Virol. **83:**3993–3997. http://dx.doi.org/10.1128/JVI.01108-08.
89. **Ndhlovu ZM, Proudfoot J, Cesa K, Alvino DM, McMullen A, Vine S, Stampouloglou E, Piechocka-Trocha A, Walker BD, Pereyra F.** 2012. Elite controllers with low to absent effector CD8+ T cell responses maintain highly functional, broadly directed central memory responses. J. Virol. **86:**6959–6969. http://dx.doi.org/10.1128/JVI.00531-12.

90. **Borthwick N, Ahmed T, Ondondo B, Hayes P, Rose A, Ebrahimsa U, Hayton E-J, Black A, Bridgeman A, Rosario M, Hill AV, Berrie E, Moyle S, Frahm N, Cox J, Colloca S, Nicosia A, Gilmour J, McMichael AJ, Dorrell L, Hanke T.** 2014. Vaccine-elicited human T cells recognizing conserved protein regions inhibit HIV-1. Mol. Ther. **22:**464–475. http://dx.doi.org/10.1038/mt.2013.248.

91. **Kloverpris H, Karlsson I, Bonde J, Thorn M, Vinner L, Pedersen AE, Hentze JL, Andresen BS, Svane IM, Gerstoft J, Kronborg G, Fomsgaard A.** 2009. Induction of novel CD8+ T-cell responses during chronic untreated HIV-1 infection by immunization with subdominant cytotoxic T-lymphocyte epitopes. AIDS **23:**1329–1340. http://dx.doi.org/10.1097/QAD.0b013e32832d9b00.