Viral adaptation to immune selection pressure by HLA class I–restricted CTL responses targeting epitopes in HIV frameshift sequences

Christoph T. Berger,[1] Jonathan M. Carlson,[2] Chanson J. Brumme,[1] Kari L. Hartman,[1] Zabrina L. Brumme,[1,3,4] Leah M. Henry,[1] Pamela C. Rosato,[1] Alicja Piechocka-Trocha,[1] Mark A. Brockman,[1,3,4] P. Richard Harrigan,[3,5] David Heckerman,[2] Daniel E. Kaufmann[1], and Christian Brander[1,6,7]

[1]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard, Boston, MA 02129
[2]Microsoft Research, Seattle, WA 98033
[3]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC V6B 5S8, Canada
[4]Simon Fraser University, Burnaby, BC V5A 4Y7, Canada
[5]Division of AIDS, University of British Columbia, Vancouver, BC V6Z 1Y6, Canada
[6]Institucio Catalana de Recerca i Estudis Avancats, 08010 Barcelona, Spain
[7]Irsicaixa HIV Research Institute–HIVACAT, Hospital Germans Trias i Pujol, Badalona, 08916 Barcelona, Spain

CD8⁺ cytotoxic T lymphocyte (CTL)–mediated immune responses to HIV contribute to viral control in vivo. Epitopes encoded by alternative reading frame (ARF) peptides may be targeted by CTLs as well, but their frequency and in vivo relevance are unknown. Using host genetic (human leukocyte antigen [HLA]) and plasma viral sequence information from 765 HIV-infected subjects, we identified 64 statistically significant ($q < 0.2$) associations between specific HLA alleles and sequence polymorphisms in alternate reading frames of gag, pol, and nef that did not affect the regular frame protein sequence. Peptides spanning the top 20 HLA-associated imprints were used to test for ex vivo immune responses in 85 HIV-infected subjects and showed responses to 10 of these ARF peptides. The most frequent response recognized an HLA-A*03–restricted +2 frame–encoded epitope containing a unique A*03-associated polymorphism at position 6. Epitope-specific CTLs efficiently inhibited viral replication in vitro when viruses containing the wild-type sequence but not the observed polymorphism were tested. Mutating alternative internal start codons abrogated the CTL-mediated inhibition of viral replication. These data indicate that responses to ARF-encoded HIV epitopes are induced during natural infection, can contribute to viral control in vivo, and drive viral evolution on a population level.

CORRESPONDENCE
Christian Brander:
cbrander@irsicaixa.es

Abbreviations used: ARF, alternative reading frame; BCL, B cell line; OLP, overlapping peptide; ORF, open reading frame; SD50%, concentration indicating half-maximal response; SFC, spot-forming cell; TCL, T cell line.

HLA class I–restricted CD8⁺ CTL responses have been associated with viral containment during acute infection and are thought to contribute to the establishment of viral setpoint in the chronic stages of infection (Borrow et al., 1994; Koup et al., 1994; Jin et al., 1999; Schmitz et al., 1999; Brander et al., 2006; Northfield et al., 2007; Streeck et al., 2009). Aside from CTL epitopes located in proteins encoded by HIV open reading frames (ORFs), recent data point toward the existence of epitopes in protein sequences encoded by frameshifted HIV (Cardinaud et al., 2004) and SIV nucleotide sequences (Maness et al., 2007). Such alternative reading frame (ARF)–encoded epitopes have been reported in a variety of human diseases, including influenza infection, malignancies, and autoimmunity (Bullock and Eisenlohr, 1996; Wang et al., 1996; Bullock et al., 1997; Rimoldi et al., 2000; Probst-Kepper et al., 2001; Saeterdal et al., 2001; Saulquin et al., 2002), but they may occur at particularly high frequency in HIV and other retroviral pathogens that show frequent nucleotide deletions or insertions in their genome. Early studies in mouse

The Rockefeller University Press $30.00
J. Exp. Med. Vol. 207 No. 1 61-75
www.jem.org/cgi/doi/10.1084/jem.20091808

The Journal of Experimental Medicine

AIDS and the recent reports in HIV and SIV infections suggest that immune responses mounted against such frame-shifted epitopes could potentially contribute to viral control in vivo (Mayrand et al., 1998; Ho and Green, 2006). Indeed, CTLs targeting an HLA-B*07–restricted ARF epitope, first identified in HLA-B*07 transgenic mice, were able to recognize HIV-infected cells and kill peptide-pulsed antigen-presenting cells (Cardinaud et al., 2004). Moreover, ARF epitope–specific CTLs inhibited SIV replication in autologous infected cells in vitro and appeared to select for viral escape variants that abrogated the epitope binding to Mamu-B*17 in vivo (Yant et al., 2006; Maness et al., 2007).

The analysis of immune escape mutations in HIV-infected individuals expressing specific HLA class I alleles has allowed for the identification of HLA class I allele–specific "footprints," which reflect the accumulation of particular HIV sequence polymorphisms in individuals carrying or lacking a specific HLA class I allele (Moore et al., 2002; Leslie et al., 2005; Bhattacharya et al., 2007; Rousseau et al., 2008; Brumme et al., 2009). The presence of such HLA footprints is thought to represent direct evidence for in vivo immune selection pressure on the virus by the host cellular immune response. Importantly, these footprints have also facilitated the identification of novel CTL epitopes in HIV coding sequences, although the presentation of individual epitopes on several HLA alleles may complicate the interpretation (Bhattacharya et al., 2007; Frahm et al., 2007).

Little is known about the potential in vivo relevance of ARF epitopes and how responses against such targets could affect viral evolution. To address this, we assessed the presence and frequency of HLA-associated viral polymorphisms that were statistically significant in one or more ARFs but not in the primary ORF (e.g., synonymous changes) in a large cohort of 765 untreated, chronically HIV-infected individuals (Brumme et al., 2007, 2008). Our data indicate that cellular immune responses to ARF-encoded sequences are not infrequent in HIV-infected individuals and that these responses can inhibit viral replication in vitro. Their in vivo relevance is further reflected by the identification of HLA class I allele–specific associations with particular viral polymorphisms that did not affect the amino acid sequence in the original reading frame yet represent effective escape from ARF epitope–specific CTL populations.

RESULTS

Population-level HLA-associated viral polymorphisms in ARF sequences indicate CTL-mediated immune pressure

To identify ARF-directed T cell responses and to assess population-wide immune selection pressure mediated by these responses, HIV gag, pol, and nef sequences from a cohort of 765 individuals were analyzed for HLA allele–associated sequence polymorphisms in ARFs (Brumme et al., 2007, 2008). All nucleotide sequences were initially translated in the putative +1 and +2 reading frames and analyzed for HLA class I allele–associated viral polymorphisms, as previously described (Carlson et al., 2008). A correction for multiple comparisons using a q-value approach (Storey and Tibshirani, 2003) was applied, and associations with a false discovery rate of 20% or lower ($q < 0.2$) were included. A total of 64 HLA-associated viral polymorphisms unique to the ARF were identified in gag (29 associations), pol (27 associations), and nef (8 associations). Of these, 12, 18, and 7 were derived from +1 shifted reading frames and 17, 9, and 1 were derived from +2 shifted sequences, respectively (Table I). It is important to note that the HLA-associated polymorphisms listed in Table I were not reflected in the coding sequence of the HIV protein, and these associations are unique to the +1 and/or +2 ARF only. Indeed, our method also identified HLA-associated polymorphisms in the coding sequence, which resulted in (statistically significant) nonsynonymous substitutions in the +1 and +2 frames, but these were excluded from further analysis as we conservatively assumed that these imprints were driven by immune responses to the known HIV proteins. Furthermore, all HLA-associated polymorphisms in ARF that had an HLA association with any HLA allele within ±2 amino acids in the corresponding primary coding sequence were also excluded from analysis (for a complete list of imprints see Table S1).

According to a recent publication that introduces a consistent nomenclature of HLA-associated viral polymorphisms (Brumme et al., 2009), the associations identified in this study are referred to as adapted and nonadapted forms. Adapted forms (also commonly called escape variants) are amino acids significantly enriched in the presence of the HLA allele in question (and vice versa). Nonadapted forms (also commonly called wild-type or susceptible forms) are amino acids significantly depleted in the presence of the HLA allele in question (and vice versa).

Peptides spanning regions of ARF sequence polymorphisms are targeted by HIV-infected individuals

To investigate if the observed HLA-associated polymorphisms could in fact be the result of viral adaptation to CTL-mediated immune pressure on alternative-frame HIV peptides, in vitro T cell responses to ARF peptides spanning the region of the associations were assessed. Peptides were synthesized to either contain the identified imprint in the center of two overlapping peptide (OLP) sequences (between 13 and 20 amino acids in length, and overlapping by 9–10 amino acids) or as short (9–12 amino acids long) predicted CTL epitopes using a recently developed epitope prediction algorithm (Heckerman et al., 2007), and were then tested by ex vivo IFN-γ ELISPOT assay using PBMCs from untreated, chronically HIV-infected individuals. Peptides included 9 sequences from the +1 ORF and 12 from the +2 ORF, covering 10, 8, and 2 regions of imprints in gag, pol, and nef, respectively (Table II). Responses were observed in 21 out of 85 HIV-infected individuals and 0 out of 32 HIV-negative individuals tested ($P = 0.001$ using Fisher's exact test), indicating that responses are HIV specific (Fig. 1 A). Of note, there was no statistical difference in HLA class I allele representation between the HIV-infected and -uninfected individuals, indicating that

the HIV-negative group was a suitable control population (Tables S2 and S3).

Among the 21 subjects with responses, 27 peptide-specific activities were seen with a median magnitude of 140 spot-forming cells (SFCs)/10^6 PBMCs (range: 50–820 SFCs/10^6 PBMCs; Fig. 1 B). Responses in seven individuals were re-confirmed by intracellular cytokine staining using flow cytometry and were in all cases mediated by CD8 T cells (not depicted). The frequency of individuals with detectable responses (24.7%) is comparable to that seen to first-frame responses to HIV protease, Vpu, Rev, Tat, and Vif (11–40% of responders, based on protein-specific peptide sets ranging from 9 to 24 OLPs). These proteins are consistently among the less frequently targeted HIV proteins (Frahm et al., 2004; Zuñiga et al., 2006) that elicit responses of comparable magnitude (all $P > 0.05$; Zuñiga et al., 2006). Thus, although of lower magnitude and frequency than responses to dominant targets such as HIV Gag, RT, and Nef, responses to ARF were well above stringent assay backgrounds, also reflected by the fact that no spurious responses were seen in HIV-uninfected subjects (Fig. 1 A). Interestingly, the responses to the tested predicted ARF epitopes and longer ARF peptides did not occur exclusively in the individuals expressing the HLA allele for which an imprint was present in that region. Although these situations were less frequent (19 out of the 27 responses detected were observed in subjects expressing either the imprint allele [$n = 8$] or at least an HLA supertype–matched allele [$n = 11$]), 8 responses were unexpected based on the HLA type of the tested individual. The recently described extensive HLA promiscuity even among optimally defined CTL epitopes and the fact that the in vitro testing was based on longer peptide sequences that may contain multiple epitopes presented by different HLA alleles likely explain the detection of these responses in the absence of the imprint HLA alleles (Goulder et al., 2000; Frahm et al., 2002; Frahm et al., 2007). Among the reactive peptides, three regions (four peptides) were targeted by >1 out of the 21 reactive individuals (Fig. 1 C). In particular, the GQQRSTLERTSKASLER (GR17 peptide) sequence was recognized by 10 individuals of whom all but 1 expressed either HLA-A*03 or an HLA-A*03 supertype allele (Table III; Sidney et al., 2008). The median magnitude of the response was 185 SFCs/10^6 PBMCs (range: 70–640 SFCs/10^6 PBMCs). This peptide spanned the site of the strongest imprint identified in this study ($P = 1.78 \times 10^{-8}$; $q = 0.000048$), which was indeed associated with an adapted polymorphism in the

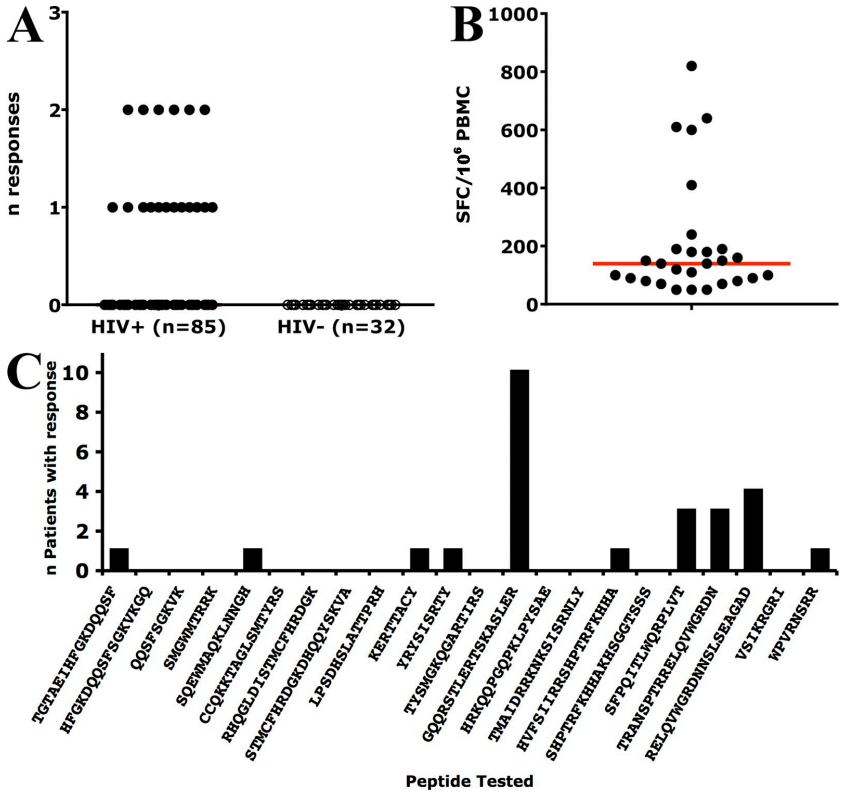

Figure 1. Magnitude and breadth of CTL responses to frameshift-derived peptides. (A–C) 85 HIV-infected individuals and 32 HIV-negative controls were tested for CTL responses in ex vivo IFN-γ ELISPOT using fresh PBMCs. 21 out of 85 patients displayed a CTL response to at least one of the peptides tested. Peptides were exclusively recognized in HIV-infected individuals. (A) The number of peptides targeted by each individual tested is shown (A). Magnitude of responses in the IFN-γ ELISPOT is expressed as SFCs/10^6 PBMCs. The horizontal bar represents the median magnitude (B). The frequency and sequence of the responses to individual peptides are shown. All peptides tested and the respective number of patients mounting a CTL response are indicated. The 17mer GQQRSTLERTSKASLER (GR17) was recognized by 10 out of the 85 individuals tested (C).

Table I. HLA-associated viral polymorphisms in the +1 and +2 ARFs

RF	Protein	Association	HLA	Pos HXB2	aa	p-value	q-value	Dist
+1	Gag	Adapted	C05	12	T	4.76×10^{-4}	1.91×10^{-1}	14
+1	Gag	Nonadapted	A03	45	Q	1.47×10^{-4}	1.02×10^{-1}	13
+1	Gag	Adapted	A29	86	F	1.07×10^{-4}	7.94×10^{-2}	5
+1	Gag	Adapted	B40	99	R	3.42×10^{-4}	1.73×10^{-1}	4
+1	Gag	Adapted	B08	132	T	3.84×10^{-4}	1.73×10^{-1}	5
+1	Gag	Nonadapted	A23	138	S	2.04×10^{-4}	1.23×10^{-1}	8
+1	Gag	Adapted	C06	168	S	8.89×10^{-5}	6.79×10^{-2}	20
+1	Gag	Adapted	B57	253	T	4.41×10^{-4}	1.85×10^{-1}	11
+1	Gag	Nonadapted	B57	253	I	4.43×10^{-4}	1.85×10^{-1}	11
+1	Gag	Adapted	B48	321	Y	3.08×10^{-4}	1.64×10^{-1}	9
+1	Gag	Nonadapted	B48	321	C	3.08×10^{-4}	1.64×10^{-1}	9
+1	Gag	Nonadapted	B52	499	H	4.07×10^{-4}	1.76×10^{-1}	17
+1	pol (pro)	Adapted	A01	73	A	4.26×10^{-6}	3.55×10^{-3}	10
+1	pol (RT)	Adapted	B07	25	P	2.87×10^{-4}	1.18×10^{-1}	10
+1	pol (RT)	Adapted	A01	53	K	2.17×10^{-4}	9.2×10^{-2}	18
+1	pol (RT)	Adapted	B46	151	K	1.49×10^{-4}	7.66×10^{-2}	7
+1	pol (RT)	Nonadapted	B46	151	R	1.53×10^{-4}	7.66×10^{-2}	7
+1	pol (RT)	Nonadapted	B40	258	R	5.1×10^{-4}	1.74×10^{-1}	13
+1	pol (RT)	Adapted	B40	258	K	5.1×10^{-4}	1.74×10^{-1}	13
+1	pol (RT)	Adapted	B58	271	M	8.2×10^{-5}	4.93×10^{-2}	6
+1	pol (RT)	Adapted	B14	305	R	5.63×10^{-4}	1.84×10^{-1}	8
+1	pol (RT)	Nonadapted	B14	305	K	5.63×10^{-4}	1.84×10^{-1}	8
+1	pol (RT)	Adapted	C06	326	*	5.75×10^{-4}	1.86×10^{-1}	3
+1	pol (RT)	Nonadapted	B27	356	E	4.39×10^{-4}	1.58×10^{-1}	11
+1	pol (int)	Adapted	B44	153	L	1.15×10^{-4}	6.49×10^{-2}	3
+1	pol (int)	Adapted	A03	240	N	2.4×10^{-4}	10^{-1}	6
+1	pol (int)	Nonadapted	A03	240	S	3.22×10^{-4}	1.26×10^{-1}	6
+1	pol (int)	Nonadapted	A03	241	F	1.78×10^{-8}	3.55×10^{-5}	7
+1	pol (int)	Adapted	A03	241	S	3.06×10^{-5}	2.03×10^{-2}	7
+1	pol (int)	Nonadapted	A23	241	F	3.56×10^{-4}	1.33×10^{-1}	7
+1	Nef	Nonadapted	B13	24	S	4.34×10^{-4}	1.02×10^{-1}	3
+1	Nef	Adapted	C05	28	M	6.29×10^{-4}	1.3×10^{-1}	4
+1	Nef	Adapted	A02	28	P	9.83×10^{-4}	1.74×10^{-1}	4
+1	Nef	Nonadapted	B18	58	*	4.28×10^{-4}	1.02×10^{-1}	4
+1	Nef	Nonadapted	B44	146	Y	8.49×10^{-4}	1.6×10^{-1}	3
+1	Nef	Adapted	C15	158	K	3.01×10^{-4}	7.77×10^{-2}	3
+1	Nef	Adapted	A24	177	K	1.41×10^{-4}	3.98×10^{-2}	3
+2	Gag	Adapted	C05	12	L	4.76×10^{-4}	1.8×10^{-1}	14
+2	Gag	Nonadapted	A03	45	S	5.19×10^{-5}	4.08×10^{-2}	13
+2	Gag	Nonadapted	C06	53	S	5.29×10^{-5}	4.08×10^{-2}	5
+2	Gag	Adapted	B40	99	G	3.42×10^{-4}	1.51×10^{-1}	4
+2	Gag	Adapted	A68	123	R	9.03×10^{-6}	1.16×10^{-2}	4
+2	Gag	Nonadapted	A23	138	P	1.82×10^{-4}	1.04×10^{-1}	8
+2	Gag	Adapted	C06	168	H	8.89×10^{-5}	6.22×10^{-2}	20
+2	Gag	Nonadapted	B49	183	F	1.47×10^{-4}	9.18×10^{-2}	28
+2	Gag	Adapted	B57	253	P	4.61×10^{-4}	1.8×10^{-1}	11
+2	Gag	Nonadapted	B57	253	S	4.61×10^{-4}	1.8×10^{-1}	11
+2	Gag	Adapted	C02	272	S	2.31×10^{-4}	1.15×10^{-1}	4
+2	Gag	Adapted	B48	298	S	5.09×10^{-4}	1.9×10^{-1}	4

Table I. HLA-associated viral polymorphisms in the +1 and +2 ARFs *(Continued)*

RF	Protein	Association	HLA	Pos HXB2	aa	p-value	q-value	Dist
+2	Gag	Adapted	B48	321	I	2.32×10^{-4}	1.15×10^{-1}	9
+2	Gag	Nonadapted	B53	465	G	1.18×10^{-5}	1.29×10^{-2}	12
+2	Gag	Adapted	B52	499	A	1.95×10^{-5}	1.88×10^{-2}	17
+2	Gag	Nonadapted	B51	499	T	7.49×10^{-5}	5.41×10^{-2}	17
+2	Gag	Nonadapted	B52	499	T	9.38×10^{-5}	6.37×10^{-2}	17
+2	pol (pro)	Adapted	A01	73	H	4.26×10^{-6}	6.86×10^{-3}	10
+2	pol (pro)	Nonadapted	A01	73	Y	2.61×10^{-4}	1.63×10^{-1}	10
+2	pol (RT)	Adapted	B07	25	L	2.87×10^{-4}	1.75×10^{-1}	10
+2	pol (RT)	Adapted	A01	53	K	2.18×10^{-4}	1.49×10^{-1}	18
+2	pol (RT)	Adapted	B58	271	C	6.64×10^{-5}	6.5×10^{-2}	6
+2	pol (int)	Adapted	A03	240	T	2.4×10^{-4}	1.53×10^{-1}	6
+2	pol (int)	Nonadapted	A03	240	A	2.41×10^{-4}	1.53×10^{-1}	6
+2	pol (int)	Nonadapted	A03	241	S	1.78×10^{-8}	4.84×10^{-5}	7
+2	pol (int)	Adapted	A03	241	P	3.08×10^{-5}	3.24×10^{-2}	7
+2	Nef	Adapted	B40	157	*	5.17×10^{-4}	1.25×10^{-1}	4

RF indicates the reading frame relative to the regular coding frame of the respective protein. For associations, adapted forms are amino acids significantly enriched in the presence of the HLA allele in question (and vice versa), whereas nonadapted forms (also commonly called wild-type or susceptible forms) are amino acids significantly depleted in the presence of the HLA allele in question (and vice versa). HLA indicates the HLA allele for which the association was observed. Pos HXB2 indicates the amino acid position in the corresponding HXB2 primary reading frame sequence (Leitner et al., 2005). aa indicates the amino acid for which the HLA-associated polymorphism was observed. dist indicates the distance (in amino acids) to the next HLA-associated viral polymorphism (to any HLA allele) in the primary ORF. Asterisks indicate stop codons.

context of HLA-A*03. As indicated in Table II, individuals expressing the HLA-A*03 allele were significantly less likely to have a serine residue at position 241 in the +2 frameshifted integrase sequence compared with HLA-A*03–negative individuals. The HLA types of all patients with a response to GR17 are summarized in Table III. To rule out that the observed HLA-A*03 imprint originated from immune pressure on epitopes derived from other reading frames (i.e., the original Pol sequence or its +1 ARF), responses in HLA-A*03–expressing individuals to these regions were assessed. Overall, only one HLA-A*03–expressing individual mounted a response to the corresponding peptide in ARF +1. In addition, in listings from the Los Alamos HIV immunology database and in datasets from two clade B–infected cohorts with a total of >600 comprehensively tested individuals, no evidence emerges that the corresponding region in the original frame (i.e., HIV integrase 241) contains an HLA-A*03 (supertype)–restricted epitope (unpublished data; Frahm et al., 2004; Zuñiga et al., 2006). Similarly, available population-based studies of HLA footprinting provide no evidence of an A*03-associated polymorphism in the coding sequence (Rousseau et al., 2008; Brumme et al., 2009). Collectively, these data suggest that responses to the peptide encoded by the +2 ARF, not the integrase coding sequence peptide, were driving the observed imprint.

Identification of the optimal HLA-A*03 ARF epitope and its associated escape mutation

To further describe the recognized epitope within the GR17 peptide and to evaluate whether the observed imprint in ARF +2 was affecting functional CTL recognition, standard epitope mapping analyses using serial truncations of the GR17 peptide were performed (Fig. 2). Binding motif analysis suggested the presence of two potentially A*03-restricted epitopes, RTSKASLER (RR9) and RSTLERTSK (RK9; Fig. 2 B). As shown in Fig. 2 C, the optimal epitope mapped to the 9mer sequence RR9. The highly statistically significant HLA-A*03–associated ARF polymorphism observed at the population level mapped to position 6 (underlined) of this epitope (RTSKAPLER), suggesting abrogation of TCR recognition as a possible mechanism of escape (Shibata et al., 1992; Rudolph et al., 2006). Subsequent HLA restriction analysis using partly HLA-matched APCs and RR9-specific T cell lines (TCLs) confirmed the restriction of this epitope by HLA-A*03 (Fig. 2 D). To test whether sequence mutation within the RR9 epitope affected CTL recognition, the most common variants (minimal frequency of 2% in the sequences of the 765 individuals included in the population-based viral polymorphism analysis) were synthesized and tested for their ability to elicit T cell activity in RR9 responders (Fig. 3). The three tested variants were more common in A*03-positive individuals, with two of them producing a significant HLA-associated imprint in the initial analysis: RTSKAPLER (A*03-RPR9; $P = 3.08 \times 10^{-5}$; $q = 0.0324$), RTSKTSLER (A*03-RTR9; $P = 2.4 \times 10^{-4}$; $q = 0.153$), and RTSKAALER (A*03-RAR9; Fig. 3 A). Recognition of these epitope variants was then tested in the context of five HLA-A*03–expressing subjects. Neither of the subjects tested recognized the A*03-RPR9 or the A*03-RTR9. In contrast, two of the HLA-A*03–positive and the one HLA-A*03–negative but HLA-A*74–expressing individual mounted a detectable response to the A*03-RAR9 variant. Thus, the amino acid change at position 6 from a serine

to an alanine (which did not appear as a common variant in the imprint analysis; Table S1) is only variably associated with escape from recognition in the HLA-A3 context (Fig. 3 B).

A*03-RR9–specific CTLs inhibit viral replication in vitro

To demonstrate the expression and presentation of RR9 in infected cells and that the observed polymorphisms were effective immune escape variants, we engineered multiple viral constructs and assessed them in in vitro inhibition assays. We introduced different point mutations on the backbone of a laboratory strain (NL4-3), including (a) the clade B consensus variant encoding a serine (underlined) at residue 6 (RTSKASLER; RR9), (b) the most common, and putative escape form, in A*03-positive individuals encoding a proline (underlined) at that position (RTSKAPLER; RPR9), and (c) a variant containing an amino acid change at one of the HLA-binding anchors at position 9 (underlined; RTSKASLEG; RG9) to cause the loss of binding capacity to the A*0301 allele. We anticipated that the consensus variant (RR9) should be recognized by a peptide-specific T cell clone derived from an A*03-positive individual, whereas the RPR9 would escape from recognition. Similarly, the RG9 variant should be less well recognized because of reduced HLA binding. Importantly, all viral variants were designed to be synonymous in the corresponding regular reading frame. To assess viral inhibition, $CD4^+$ T cells were infected with the different viral constructs and coincubated with effector cells, and viral replication was monitored for 7 d of culture. To avoid outgrowth of autologous virus, $CD4^+$ T cells from an elite controller were used for these experiments. An A*03-RR9–specific $CD8^+$ CTL clone (Fig. S1) from the same subject was used as an effector cell population and added at different E/T ratios to the infected $CD4^+$ cells (Fig. 4). Even at an E/T ratio of 1:10, >80% of the viral replication was inhibited at day 7 as compared with infection of target cells alone when virus containing the clade B consensus sequence was used. In contrast, when using the variant virus containing the proline variant (RPR9) or the variant with the change at the anchor position that should abrogate binding of the epitope to HLA-A*03 (RG9), no relevant inhibition of viral replication was observed. Together with the variant recognition data derived from experiments with PBMCs (Fig. 3 B), the results obtained in the viral replication inhibition assay strongly suggest that the A*03-RR9 epitope is indeed expressed, processed, and presented in HIV-infected cells in sufficient amounts to sensitize these cells for CTL-mediated killing.

Alternative start codon usage may reflect a mechanism of frameshift epitope expression in infected cells

Different mechanisms leading to the generation of CTL epitopes encoded by ARFs have been described in the literature (for review see Mayrand and Green, 1998). In particular, the presence of upstream alternative splice sites and alternative start codons has been suggested to potentially facilitate the

Table II. HLA imprints and regions tested

RF	Protein	HLA	aa	Pos	Association	p-value	q-value	Dist	Region of imprint
+1	gag	B52	H	499	Nonadapted	4.07×10^{-4}	1.76×10^{-1}	17	*LPSDHSLATTPR**H***
+1	pol (RT)	B07	P	124	Adapted	2.87×10^{-4}	1.18×10^{-1}	10	*SQEWMAQKLNNG**H***
+1	pol (RT)	A03	F	142	Nonadapted	1.78×10^{-8}	3.55×10^{-5}	7	GTAEIHFGKDQQS**F**SGKVKGQ
+1	pol (RT)	B46	K	149	Adapted	1.49×10^{-4}	7.66×10^{-2}	7	RHQGLDISTMCFH**R**DGKDHQQYSKVA
+1	pol (RT)	B40	R	258	Nonadapted	5.1×10^{-4}	1.74×10^{-1}	13	*CCQKKTAGLSMTY**R**S*
+1	pol (Int)	A03	N	240	Adapted	2.4×10^{-4}	10^{-1}	6	TGTAEIHFGKDQQ**S**FSGKVKGQ
+1	pol (Int)	A03	S	240	Nonadapted	3.22×10^{-4}	1.26×10^{-1}	6	TGTAEIHFGKDQQ**S**FSGKVKGQ
+1	nef	A24	K	177	Adapted	1.41×10^{-4}	3.98×10^{-2}	3	*ASMGWMTR**R**KKC*
+1	nef	C15	K	158	Adapted	3.01×10^{-4}	7.77×10^{-2}	3	*KRPM**Q**ERTTACYTL*
+2	gag	C05	L	12	Nonadapted	4.76×10^{-4}	1.8×10^{-1}	14	GCESVSIKRGR**I**R*
+2	gag	A03	S	45	Nonadapted	5.19×10^{-5}	4.08×10^{-2}	13	TYSMGKQGARTIR**S***
+2	gag	C06	S	53	Nonadapted	5.29×10^{-5}	4.08×10^{-2}	5	SWPVRN**S**RRL
+2	gag	A68	R	123	Adapted	9.03×10^{-6}	1.16×10^{-2}	4	*HR**K**QQPGQPKLPYSAE
+2	gag	B49	F	183	Nonadapted	1.47×10^{-4}	9.18×10^{-2}	28	HVFSIIRRSHPTR**F**KHHAKHSGGTSSS
+2	gag	B51	T	499	Nonadapted	7.49×10^{-5}	5.41×10^{-2}	17	SFPQITLWQRPLV**T***
+2	gag	B52	T	499	Nonadapted	9.38×10^{-5}	6.37×10^{-2}	17	SFPQITLWQRPLV**T***
+2	gag	B53	G	465	Nonadapted	1.18×10^{-5}	1.29×10^{-2}	12	TRANSPTRRELQV**W**GRDNNSLSEAGAD
+2	pol (Pro)	A01	H	73	Nonadapted	4.26×10^{-6}	6.86×10^{-3}	10	*S**Y**R**Y**SISRT**Y**TCQHNWK
+2	pol (RT)	B07	L	25	Nonadapted	2.87×10^{-4}	1.75×10^{-1}	10	*TMA**I**DRRKNKSISRNLY
+2	pol (Int)	A03	S	241	Nonadapted	1.78×10^{-8}	4.84×10^{-5}	7	GQQRSTLERTSKA**S**LER*

HLA indicates the HLA allele for which the association was observed. aa indicates the amino acid for which the HLA-associated polymorphism was observed. Pos indicates the amino acid position in the corresponding HXB2 primary reading frame sequence (Leitner et al., 2005). For associations, adapted forms are amino acids significantly enriched in the presence of the HLA allele in question (and vice versa), whereas nonadapted forms are amino acids significantly depleted in the presence of the HLA allele in question (and vice versa). Dist indicates the distance (in amino acids) to the next HLA-associated viral polymorphism (to any HLA allele) in the primary ORF. Region of imprint shows the sequences surrounding the imprint (bold). In cases where predicted epitopes were tested, the respective sequence is underlined. *, stop codon.

expression of ARF sequences (Bullock and Eisenlohr, 1996; Mayrand and Green, 1998; Schwab et al., 2004). However, none of these mechanisms that could enable or increase the expression of such frameshifted sequences has been analyzed in HIV/SIV in detail.

Because a known splice acceptor site, SA1, is located right upstream of the RR9 epitope (Arrigo et al., 1990), we first hypothesized that HIV might be making mistakes when splicing mRNA using this acceptor site. We isolated intracellular HIV mRNA from T cells infected with NL4-3 and amplified putative splice variants by nested RT-PCR using forward primers just upstream of Gag and reverse primers just downstream of HIV integrase. During PCR amplification a short (50-s) elongation time was used to selectively amplify splice variants but exclude amplification of genomic RNA. PCR amplicons were cloned into *Escherichia coli*, and 96 clones were sequenced and analyzed for evidence of alternative splice variants. However, all 96 sequences were identical and exhibited the expected splicing pattern (unpublished data), suggesting that alternative splicing is not the major mechanism of production of the RR9 ARF peptide.

We then investigated if initiation of translation at a start codon in the ARF upstream of the epitope could be a mechanism whereby the ARF peptide is produced in vivo. When the ARF +2 sequence was screened for potential start codons (AUG, methionine), the closest methionine was found >100 amino acids upstream of the RR9 epitope. However, the potential reading frame after this methionine was not open, as there were five stop codons between the methionine and the RR9 epitope that would abrogate expression. When searching for alternative (nonmethionine) start codons, a leucine residue just 2 amino acid residues upstream of the RR9 epitope (Fig. 2 A) was identified and found to be extremely conserved in the British Columbia cohort (not depicted). Leucine has previously been described to be able to act as an alternative start codon (Schwab et al., 2004). To investigate if the RR9 epitope would indeed be processed and presented from such a short precursor protein and whether this leucine residue in fact acted as an alternative start codon, we designed specific viral variants on the backbone of an NL4-3 virus containing the RR9 epitope. The leucine residue was changed either to a regular start codon (methionine; L→M) or an amino acid that is not described as an alternative start codon (valine; L→V). We hypothesized that if the leucine indeed acts as an alternative start codon for this epitope, the change to a regular start codon would increase its expression and hence its inhibition by a specific T cell clone. On the other hand, the variant with the valine should not be expressed at all and, therefore, we expected to see no inhibition of viral replication at all. These mutations were again carefully designed to be synonymous in the regular reading frame and, therefore, not alter the sequence integrity of the original reading frame (Fig. 5 A). Using these two viral variants in viral inhibition assays and comparing viral inhibition to the consensus variant, we demonstrated that the L→M variant was strongly inhibited (Fig. 5 B). Interestingly, this inhibition exceeded the suppression seen with the virus carrying the leucine (0.9 vs. 3.1 log at an E/T ratio of 1:5 at day 7; Fig. 5 D) and, therefore, suggests that the RR9 epitope is more highly expressed in the context of a regular start codon. On the other hand, the L→V variant was, as expected, not inhibited at all (Fig. 5 C). Collectively, these data indicate that expression of the alternative leucine is in fact a possible mechanism for RR9 epitope expression in HIV-infected cells.

DISCUSSION

RNA viruses with relatively small genomes, such as HIV, human T lymphocyte virus, hepatitis B virus, and hepatitis C virus, have found ways to increase the genetic information of their genome by encoding essential proteins in different forward or backward reading frames (Jacks et al., 1988; Xu et al., 2001; Gaudray et al., 2002). Through single or double nucleotide deletion or insertions and by the presence of alternative start codons, alternative splice sites, or possibly other mechanisms, highly variable pathogens such as RNA-based HIV

Table III. HLA types and clinical characteristics of individuals recognizing the GQQRSTLERTSKASLER (GR17) peptide

Patient	HLA-A	HLA-B	HLA-Cw	SFCs/10^6 PBMCs	CD4 count	VL	Clinical course
1	0201/1101	0702/1401	0702/0802	190	393	6,270	Chronic untreated
2	**0301/0301**	0702/5101	0702/1402	70	147	165,995	Chronic untreated
3	0201/**0301**	3501/4501	ND /1601	140	371	216,000	Chronic untreated
4	**0301**/2501	1801/3501	0401/1203	410	286	5,322	Chronic untreated
5	2601/7401	3701/4901	0602/0701	240	109	7,310	Chronic untreated
6	2902/3201	4402/4403	0501/1601	610	582	8,450	Chronic untreated
7	0201/**0301**	5101/5101	0303/1502	640	791	74	Controller
8	**0301**/3001	4201/5802	0602/1700	110	344	2,650	Chronic untreated
9	0101/3301	0801/1402	0701/0802	80	540	156,000	Chronic untreated
10	**0301**/1101	0702/4402	0501/0702	180	553	13,100	Chronic untreated

SFCs/10^6 PBMCs indicates the magnitude of IFN-γ ELISPOT responses. CD4 counts are indicated in cells per microliter. VL indicates the viral load in RNA copies per milliliter. Clinical course indicates if a patient is chronically HIV infected and untreated or is a spontaneous controller of HIV infection in the absence of any HIV-specific treatment. A*03 alleles are bolded to highlight which patients were expressing A*03. ND, not done.

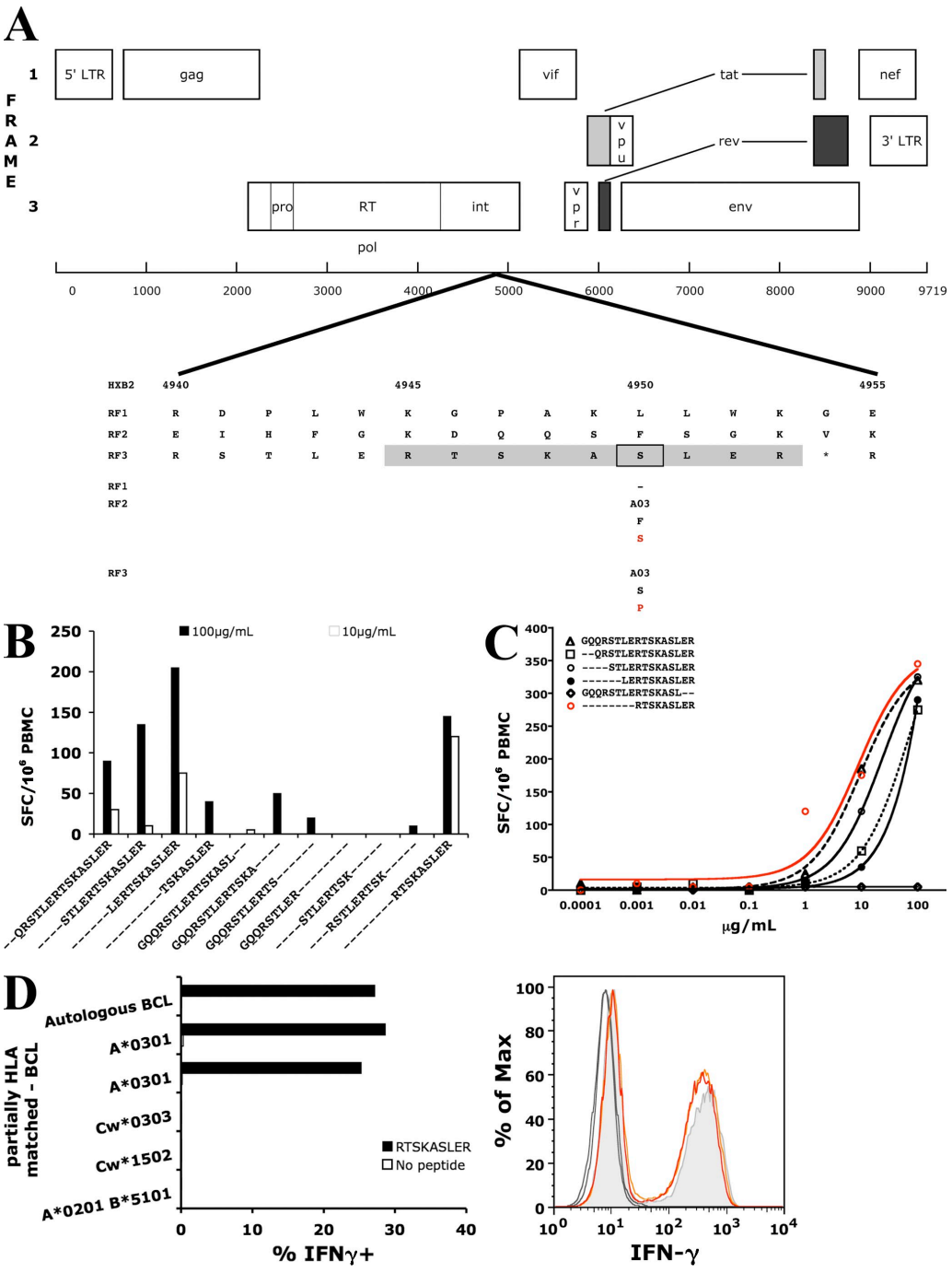

Figure 2. Epitope mapping of A*03–RTSKASLER. HIV sequencing and analysis for HLA-associated polymorphisms revealed that viruses isolated from A*03-positive patients show a common S⟶P change at position 241 in the +2 ARF relative to integrase. The amino acid sequence of this region in all three reading frames (RF1–3) and the location relative to the protein-encoding HIV genome (HXB2) are displayed. An A*03-associated imprint was also observed in the +1 ARF at the same position (A*03 is associated with an S instead of F at that position). 10 out of 85 individuals tested mounted a response to the region in the +2 ARF versus only 1 individual to the +1 ARF region. The shaded sequence indicates the newly identified A*03 epitope (A). The optimal epitope length of the GQQRSTLERTSKASLER (GR17) peptide was determined using sequential peptide truncations. IFN-γ ELISPOT responses at 100 and 10 µg/ml are displayed (B). Functional avidity was determined using serial 10-fold dilutions of the peptides ranging from 100 to 0.0001 µg/ml. The shortest peptide with the lowest SD50% was considered the optimal epitope (RTSKASLER; RR9; C). Partly HLA-matched BCLs were used for HLA restriction analyses and responses were determined with intracellular cytokine staining. Bars represent the percentage of IFN-γ–positive $CD8^+$ CTLs (left), and the histogram shows IFN-γ positivity of $CD3^+$ $CD8^+$ gated cells (right). RR9 peptide–loaded BCLs only matched in the A*03 allele (orange and red lines) coincubated with an RR9-specific effector TCL resulted in an IFN-γ response comparable to incubation with peptide-loaded autologous BCLs (gray-shaded area). Peptide-loaded BCLs matched in the other alleles did not induce IFN-γ production (gray lines; D). One out of at least three independent experiments is shown.

and hepatitis C virus may consistently produce frameshifted mRNA sequences, giving rise to alternative viral antigens that may be processed and presented by host HLA molecules. Indeed, two recent reports, one in HIV and one in SIV, indicate that antigens resulting from ARF translation can serve as a source of CTL epitopes against which the host can mount a cellular immune response (Cardinaud et al., 2004; Maness et al., 2007). The present study expanded on these observations and addressed whether and how strongly frameshifted sequences could be under immune selection pressure on a population level. The data also document that common HLA allele–driven epitope variants can be effective CTL escape mutations and can affect CTL recognition of virally infected cells. Finally, a potential mechanism for out-of-frame epitope expression, namely the use of an alternative start codon, is also supported by our data. Collectively, the data provide for the first time evidence that ARF-based epitope sequences cannot only induce immune responses comparable in breadth to those observed against regulatory HIV proteins, but that these HLA-restricted responses specific for an ARF-encoded epitope are also able to inhibit viral replication in vivo. Furthermore, we demonstrate that HLA-restricted immune pressure exerted on ARF peptides is apparently strong enough to drive HIV evolution on a population level. Indeed, within the studied HLA-A*03–restricted ARF peptide, we document and functionally validate an escape mutation occurring at position 6 of this ARF epitope, which does not affect the coding sequence of HIV integrase. The results are thus consistent with adoptive CTL transfer experiments in mouse retroviral infection and provide human experimental data in support of recent observations in the SIV model (Ho and Green, 2006; Maness et al., 2007).

Since the first report of HLA footprints in HIV sequences (Moore et al., 2002), the bioinformatic approaches to identify these imprints have been refined and take HIV founder sequence effects, HLA linkage disequilibrium, and HIV sequence covariation into account (Bhattacharya et al., 2007; Carlson et al., 2008; Brumme et al., 2009). Using such a recently developed refined approach, we identified 64 discrete HLA-driven imprints across the HIV gag, pol, and nef spanning ARF +1 and +2 sequences (Carlson et al., 2008). Their distribution across the three proteins (Table I) is shifted as more ARF imprints were observed in gag and pol regions than in nef, whereas first-frame imprints were more common in Nef (Brumme et al., 2009). For the present study, where we focused on the 20 best-supported imprints for peptide synthesis, this led us to test more gag-ARF and pol-ARF than nef-ARF imprints for functional responses (Table II). Again, compared with analyses in the regular coding frame, this selection is unlikely to have reduced or overestimated the breadth of detectable responses, as the imprints reported in the first frame were located with comparable frequency in previously defined epitopes across the three proteins tested (unpublished data). Also, compared with response rates to regular coding frame–based screenings (i.e., traditional OLP-based tests), the observed frequency of responses to ARF peptides is comparable to the frequency at which regulatory HIV proteins and the HIV protease protein are being targeted (Frahm et al., 2004; Zuñiga et al., 2006). The magnitude of responses is, however, clearly below the level that has been observed for regular HIV protein–specific responses (Frahm et al., 2004; Zuñiga et al., 2006). As ARF-derived epitopes are probably expressed and presented at a lower density than some of the presumably more abundantly produced immunodominant Gag-derived epitopes, a lower magnitude of the CTL response would not be surprising. In addition, responses with median frequencies as reported in this study are comparable to infections other than HIV, which is known to induce particularly strong responses (Bihl et al., 2005). The frequency of responses is also similar to what Cardinaud et al. (2004) have reported for a single HLA-B*07–restricted response in highly active antiretroviral therapy–treated individuals. Their HLA-B*07–restricted response was, however, not included in our screening, as no HLA-B*07–associated imprint in the vicinity of the described ARF epitope was identified, possibly either because of a lack of statistical

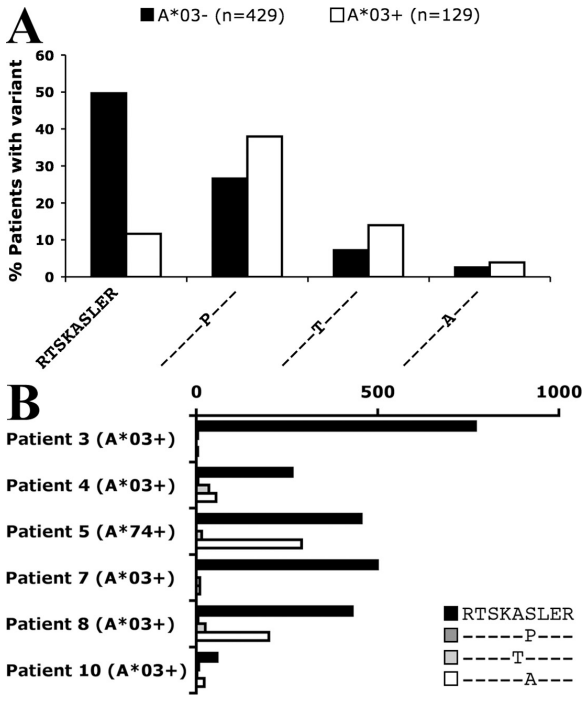

Figure 3. Commonly occurring HIV variants can escape recognition. (A and B) Viral variants of the A*03-RR9 epitope occurring with >2% frequency in the British Columbia cohort were identified. Frequency of the analyzed variants in A*03-negative (black bars) and -positive (white bars) individuals are displayed (A). Recognition of the consensus sequence peptide and the three variants was assessed in A*03-RR9 responders. All five A*03-positive individuals tested recognized the consensus sequence (black bars) but not the variants with the S→P change at amino acid position 6 (RTSKAPLER) or the A→S change at position 5 (RTSKTSLER) of the epitope, suggesting that these variants are escape forms (gray bars). In contrast, two of the A*03-positive and one A*74-positive individuals recognized the S→A variant (RTSKAALER; white bar) as well (B).

power in our cohort for this allele or because of inconsistent or diverse escape patterns from this response.

A frequent limitation in HLA imprint studies is that effective CTL escape by the observed mutation is assumed rather than experimentally demonstrated. In this study, we assessed variant recognition, both on a peptide basis as well as by expressing variant viral sequences. As the data demonstrate, the common escape variant in position 6 of the RR9 epitope abrogated CTL recognition and control of viral replication. Effective CTL escape by such a mutation that is enriched in individuals expressing the restricting HLA-A*03 allele lends further strong support to the in vivo relevance of this mutation and the effectiveness of the CTL immune pressure against the RR9 epitope.

HLA footprint analyses can only provide a partial picture of the host immune response to HIV, and it is difficult to judge the in vivo relevance of ARF-derived responses or to estimate its total contribution to the global immune response to HIV from such analyses. Aside from simple statistical considerations, such as insufficient power because of low allele frequency, fixation or near-to-fixation of escape variants in the population, and differential kinetics of reverting mutations, the current analysis is likely to underestimate the true breadth of responses because of several additional reasons. A basic issue is that frameshifts or single nucleotide insertion in the viral genome may occur repetitively, and a final epitope sequence could contain sequences of more than one reading frame. Although some positions in the genome may be more or less frequently affected by frameshifts, it is thus possible that epitopes exist that are encoded by an essentially unpredictably vast array of random mosaic epitope sequences composed of more than one reading frame. The identification of such potential epitopes based on HLA imprints is difficult, as all combinations of possible frameshifts in the

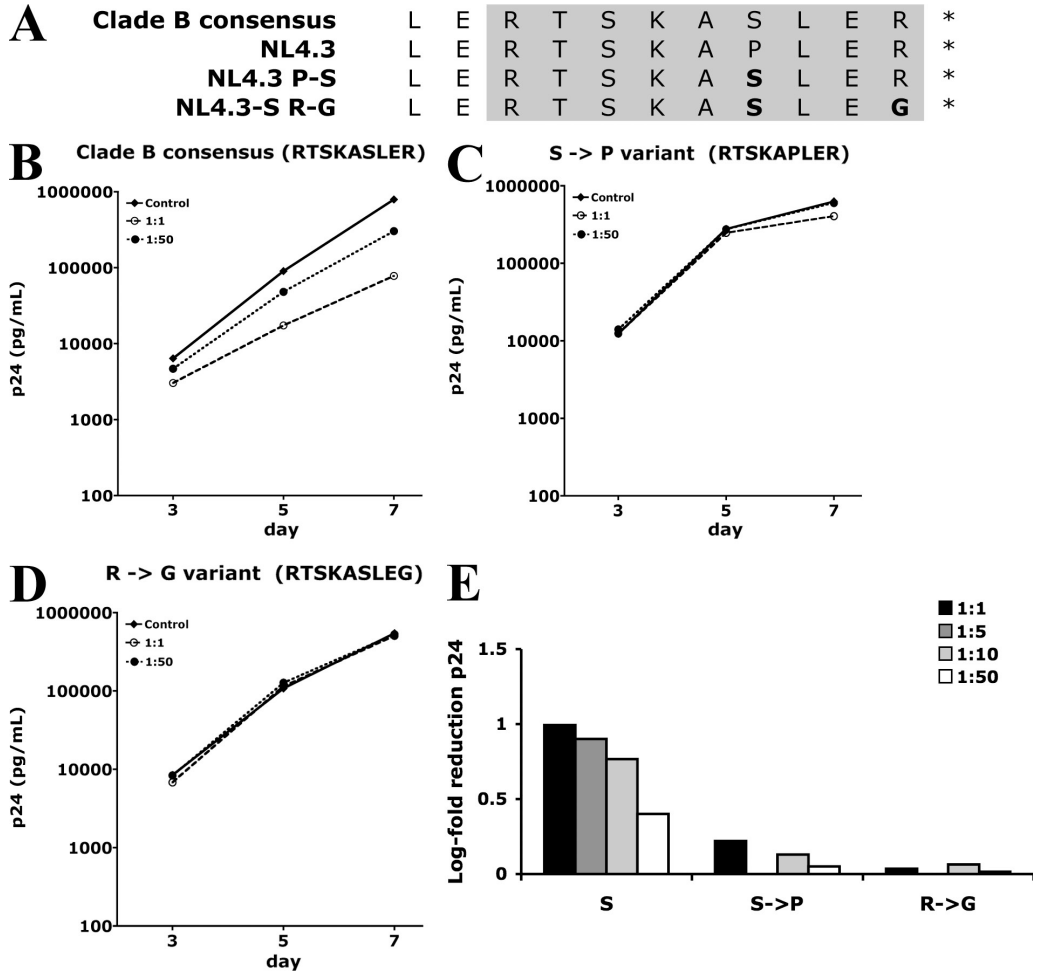

Figure 4. A*03–RR9–specific CTLs inhibit viral replication. (A–E) Different virus variants were designed on an NL4-3 backbone (A) using site-directed mutagenesis and were tested in viral inhibition assays. CD4+ T cells were infected and co-cultured with an RR9-specific CD8+ clone from the same subject at various E/T ratios over 7 d. Viral inhibition was determined by p24 measurement in supernatants at days 3, 5, and 7. The consensus (B) but not the S→P variant (C) was strongly inhibited. Similarly, an R→G mutation at position 9, potentially disrupting MHC class I binding to A*03, was not inhibited (D). The inhibition of viral replication of the consensus variant was dependent on the E/T ratio (E). One representative out of three independent experiments is shown. Inhibition assays were set up in triplicates.

vicinity of the imprint would need to be synthesized as test peptides. Despite this potential complexity, however, our data also indicate that HIV-infected cells seem to produce and present ARF-encoded peptides consistently enough across individuals to result in population-based immune pressure reflected as population-wide HLA footprints. Further documentation of how consistently HIV-infected cells may present ARF-encoded peptide sequences may also come from epitope elution studies using HIV-infected cells that may help to identify epitopes with partial sequence matches in different reading frames (Hickman et al., 2003).

Apart from statistical power to identify specific HLA allele–associated imprints, ex vivo immune analyses will only reveal responses if they are present in an appreciable number of subjects. Although some of the most immunodominant known HIV epitopes are targeted by >50% of individuals expressing the respective restricting HLA allele, many responses are found less frequently and are only detected if the test cohorts are massively expanded so that even infrequent alleles are represented several times (Bihl et al., 2006). In addition, HLA imprints in ARF that caused a nonsynonymous change in the first, coding frame were excluded in the present study, assuming that the observed change was driven by immune selection pressure against an epitope in the original HIV protein. Although this may usually be the case, it is a conservative assumption and may have underestimated the true frequency of ARF-based CTL targets. On the other hand, it cannot be excluded that specific codon usage and preferred RNA structure for amino acids in the first frame may have driven some changes in the ARF while being synonymous in the regular coding frame.

The in vivo relevance of ARF-encoded epitopes will also depend on the likelihood of and the mechanisms involved in their production in virally infected cells. Data from a recent study suggest the possibility of ribosomal slippage being a potential important mechanism given the RNA structure of HIV (Watts et al., 2009). However, to date, it still remains unclear to what level ARF-derived epitopes in HIV/SIV are a degradation product of functional proteins or RNA sequences encoded by the frameshift or antisense sequence (Miller, 1988; Landry et al., 2007), or are rather simply defective ribosomal products that result from accidental frameshift and are rapidly degraded (Yewdell and Nicchitta, 2006). The data shown in this study do not resolve this point but indicate that the use of alternative start codons could likely help the production of ARF epitopes, as the replacement of leucine

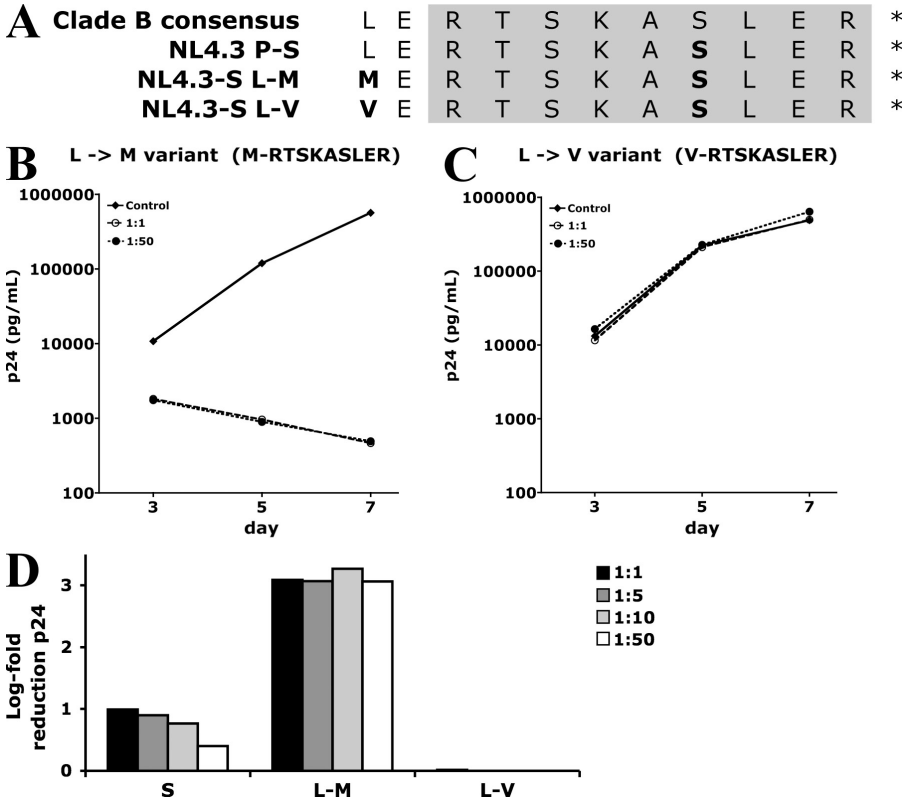

Figure 5. Alternative start codon usage might be a mechanism for frameshift epitope expression of A*03-RR9. (A–D) To address if an upstream leucine could figure as a alternative start codon and therefore regulate expression of A*03-RR9, viral variants carrying a regular start codon (AUG, methionine) or lacking a potential start codon (valine instead of leucine) at this position were created (A). The inhibition of the variants by an RR9-specific $CD8^+$ clone was assessed and compared with the consensus variant with the leucine. The L→V variant (C) was no longer inhibited, whereas the L→M variant (B) was more strongly inhibited than the L variant (−3 vs. −1 log at an E/T ratio of 1:10 at day 7), pointing at a potential mechanism for the expression of A*03-RR9 (D). One representative out of three independent experiments is shown. Inhibition assays were set up in triplicates.

with valine showed an abrogated suppression of viral replication and, presumably, RR9 presentation. Although the substitution of leucine to methionine caused significantly improved suppression of viral replication, the lower level expression of the leucine is probably sufficient to sensitize infected cells for CTL recognition. This is in line with data by Sykulev et al. (1996) showing that a single peptide–MHC complex on a target cell might be sufficient to elicit a cytolytic T cell response. Also, as the relative magnitude of responses and the relative abundance of presented epitopes do not necessarily correlate with in vivo protection (Gallimore et al., 1998), the expression level of the RR9 epitope and the magnitude of the response are likely not the only factors that determine its in vivo relevance. Regardless of this open question, the present study is the first demonstration of population-wide viral adaptation to immune selection pressure exerted by CTL responses targeting HLA class I–restricted epitopes encoded by HIV frameshifted sequences. These responses lead to effective CTL escape variants in vivo and were able to inhibit viral replication in vitro. Thus, they show all the characteristics that are commonly seen as important attributes of an effective HIV-specific CTL immunity and may be considered for inclusion in HIV vaccine design.

MATERIALS AND METHODS

Ethical approval. The study was approved by the respective local ethics boards (Partners Institutional Review Board and Providence Health Care/University of British Columbia Research Ethics Board).

Virus sequencing and HLA typing. Plasma RNA from 765 treatment-naive, chronic, predominantly HIV clade B–infected (97.5%) subjects enrolled in the British Columbia HOMER Study was isolated and sequenced as previously described (Hogg et al., 2001). In brief, HIV RNA was extracted from a pretherapy plasma sample and HIV gag, protease, RT (codons 1–400), integrase, and nef were amplified using nested RT-PCR and bulk sequenced on an automated DNA sequencer (ABI 3700; Applied Biosystems). For all genes, data on at least 500 individuals were available (gag, $n = 567$; protease/RT, $n = 532$; integrase, $n = 582$; and nef, $n = 686$). HLA class I typing was performed by sequence-based typing as previously described (Brumme et al., 2007, 2008). Accession numbers of the sequences are listed in the respective publications (Brumme et al., 2007, 2008).

Identification of HLA-associated amino acid variation in HIV genomes. Gag, Pol, and Nef sequences were aligned to HXB2 and translated into the frameshifted reading frames +1 and +2. HLA-associated polymorphisms in the frameshifted sequences were identified using previously published, phylogenetically informed methods, which additionally correct for HLA linkage disequilibrium as well as HIV codon covariation (Carlson et al., 2008). Statistical significance is reported using q-values, which estimate the false discovery rate for each comparison (Storey and Tibshirani, 2003). We used a threshold of $q \leq 0.2$, meaning that we expect a false-positive proportion of 20% among identified associations. To avoid identifying false-positive HLA associations in ARFs that could reflect immune pressure on the regular ORF, all HLA associations in the +1 and +2 ORFs were excluded when they were (a) associated with a nonsynonymous change at the same position in the first frame, (b) had an imprint for any other allele within ±2 amino acids in the original reading frame, or (c) had a known optimal epitope in the original reading frame that was restricted by the same HLA allele for which the imprint was observed for the ARF. In addition, all regions with imprints in the +1 and +2 ORFs that had a length of <10 amino acids between two stop codons were excluded from further analysis because of a low probability that these epitopes were in fact expressed all the way in the same frame and would hence not be discovered by a peptide screen.

Assessment of T cell responses to frameshift-derived epitopes. To test for immune responses to ARF epitopes, 85 untreated chronically HIV-infected individuals were recruited from hospitals in the Boston area. 32 HIV-negative individuals were included as controls. From all individuals, fresh PBMCs were isolated from whole blood using density gradient centrifugation and used in direct ex vivo ELISPOT assays as previously described (Frahm et al., 2004). In brief, 100,000 fresh PBMCs were incubated with peptides in pools of four peptides at a final concentration of 28 μg/ml in 96-well polyvinylidene difluoride plates (Millipore) precoated with the anti–IFN-γ mAb 1-D1K (Mabtech) according to the manufacturer's instructions. A positive (PHA) and at least four negative (no peptide) controls were included for each individual tested. ELISPOT results were expressed as SFCs/10^6 input cells. The threshold for a positive response was defined as at least 50 SFCs/10^6 input cells and responses exceeding three times the mean of negative wells. Peptides were synthesized as listed in Table II. Specific epitopes with an HLA-binding motif for the observed imprint were predicted using epitope prediction algorithms (available at http://atom.research.microsoft.com/bio/epipred.aspx [Heckerman et al., 2007] and www.hiv.lanl.gov/content/immunology). In brief, for each HLA-associated polymorphism, a consensus sequence was constructed from sequences that contained the predicted nonadapted form of the association, and Epipred was used to predict the likeliest epitope (9–11mer) within a window extending ±13 amino acids from the observed HLA-associated polymorphism or until the first consensus stop codon. In cases where the imprint was not included in the predicted epitope, OLPs of 17–18 amino acids in length were designed and included in the screening (Fig. 1 C).

Determination of the functional avidity and HLA restriction analyses. The optimal epitope length was determined using sequential peptide truncations in serial 10-fold dilutions ranging from 100 to 0.0001 μg/ml (Frahm et al., 2007). Freshly isolated or frozen, unexpanded PBMCs were used in standard IFN-γ ELISPOT assays. The functional avidity, defined as the antigen dose needed to elicit a concentration indicating half-maximal response (SD50%), was determined. The shortest peptide with the lowest SD50% was considered the optimal epitope (Frahm et al., 2008). HLA restriction was determined using standard ELISPOT assays and epitope-specific TCLs. TCLs were used for restriction analysis after 10–14 d of peptide-specific in vitro expansion. EBV-transformed, partly HLA-matched B cell lines (BCLs) were used as APCs for restriction analysis, as previously described (Goulder et al., 2001). BCLs were incubated for 1 h with the optimal peptides at 20 μg/ml, followed by extensive washing. BCLs without peptide were used as negative controls. Peptide-specific TCLs (E/T ratio of 1:1) and 2 μg/ml of anti-CD28 and -CD49d co-stimulatory antibodies (BD) were added and incubated for 6 h at 37°C. For the last 5 h of incubation, 10 μg/ml Brefeldin A was added. Cells were fixed and permeabilized using Cytofix/Cytoperm (BD) according to the manufacturer's instructions, and the percentage of responding cells was determined using anti-CD3–PerCP, anti-CD4–PE, anti-CD8–allophycocyanin, and anti–IFN-γ–FITC antibodies (all from BD).

Viral variants expressing escape mutations and artificial start codons in ARFs. Viruses harboring single or multiple point mutations in ARFs within and/or surrounding the A*03-RR9 epitope were created on an NL4-3 backbone using site-directed mutagenesis (QuikChange XL; Agilent Technologies) according to the manufacturer's instructions. The viral variants generated and used in in vitro inhibition assays are summarized in Fig. 4 A and Fig. 5 A. Importantly, mutagenesis primers were designed to only cause synonymous changes in the primary ORF, meaning that the engineered point mutations did not disrupt the regular reading frame. NL4-3 contains the sequence of the escape variant (RTSKAPLER). Therefore, we first created a virus variant that carries the clade B consensus sequence. Based on this P→S variant, the other viruses were created. The following primer

sequences were used: P→S variant, 5′-CCAGTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAAGGGGCAGTAG-3′ and 5′-CTACTGCCCCTTCACCTTTCCAGAGAAGCTTTGCTGGTCCTTTCCAAACTGG-3′; R→G variant, 5′-CCAGTTTGGAAAGGACCAGCAAAGCTTCTCTGGAGGGTGAAGGGGCAGTAGTAATACAAG-3′ and 5′-CTTGTATTACTACTGCCCCTTCACCCTTCCAGAGAAGCTTTGCTGGTCCTTTCCAAACTGG-3′; L→M variant, 5′-TTATTACAGGGACAGCAGAGATCCAGTATGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGT-3′ and 5′-ACCTTTCCAGAGAAGCTTTGCTGGTCCTTTCCATACTGGATCTCTGCTGTCCCTGTAATAA-3′; and L→V variant, 5′-TTTATTACAGGGACAGCAGAGATCCAGTGTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGT-3′ and 5′-ACCTTTCCAGAGAAGCTTTGCTGGTCCTTTCCACACTGGATCTCTGCTGTCCCTGTAATAAA-3′.

In vitro inhibition of viral replication. Viral inhibition assays were performed as described previously (Yang et al., 1996). In brief, autologous $CD4^+$ T cells were used as target cells after expansion from PBMCs using 1 μg/ml of a bispecific anti-CD3/8 mAb in R10-50 media (RPMI 1640 media supplemented with 10% fetal calf serum, 1% L-glutamine, and 50 U/ml IL-2) for 5 d (Chen et al., 2009). Cells were used without further CD4 separation for infection with the respective viruses. Infection was performed by incubating $CD4^+$ T cells for 4 h with the different HIV stocks at a multiplicity of infection of 0.01 followed by extensive washing. 100,000 target cells were added per well of a round-bottom 96-well plate and incubated with an A*03-RR9–specific $CD8^+$ clone at E/T ratios of 1:1, 1:5, 1:10, and 1:50. Cells were co-cultured for 7 d in 200 μl R10-50. At days 3, 5, and 7, 100 μl of the supernatant was replaced by fresh media. p24 was quantitated by a standard quantitative ELISA (commercial kit; Dupont) according to the manufacturer's instructions and as described previously (Brander et al., 1999; Jones et al., 2003; Chen et al., 2009).

Epitope-specific T cell clones were generated from a peptide-specific TCL by selection of epitope-specific $CD8^+$ T cells using an IFN-γ capture assay according to the manufacturer's instructions (Miltenyi Biotec) followed by live single-cell sorting of $CD3^+$ $CD8^+$ $CD14^-$ $CD19^-$ IFN-γ^+ cells on a FACSAria (BD) using a 70-μm nozzle at 70 psi. The single-sorted cells were expanded on irradiated feeder cells with 0.5 μg/ml of an anti-CD3 antibody (12F6; Jones et al., 2003) in R10-50. Flowcytometric polyfunctionality stainings (as described in Bihl et al., 2007) showed that the cells produced IFN-γ (>85%), TNF (>95%), and MIP-1β (>99%) and expressed CD107a (>98%; Fig. S1).

Online supplemental material. Table S1 provides the complete list of individual HLA associations with HIV sequence polymorphisms with a q-value <0.2, independent of their respective location in relation to associations seen in the primary ORF. Tables S2 and S3 provide HLA types of the HIV-infected and -uninfected individuals, respectively, included in the study. Fig. S1 shows the intracellular cytokine staining of an A*03-RR9–specific $CD8^+$ T cell clone derived from single-cell sorting in response to stimulation with A*03-RR9 peptide. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20091808/DC1.

We thank Dr. A. Khatri for peptide synthesis and J. Sela for excellent laboratory assistance. C.T. Berger thanks C. Hess for ongoing support.

The work was supported by a grant from the National Institutes of Health (R01 A1-067077 to C. Brander and D.E. Kaufmann). C.T. Berger is supported by a postdoctoral fellowship from the Swiss National Science Foundation (PBBSP3-123141) and by the Freiwillige Akademische Gesellschaft Basel. Z.L. Brumme was supported by a postdoctoral fellowship and, currently, a New Investigator award from the Canadian Institutes for Health Research. C. Brander is an Institucio Catalana de Recerca i Estudis Avancats Senior Research Professor at the Irsicaixa AIDS Research Institute.

The authors have no conflicting financial interests.

Submitted: 19 August 2009
Accepted: 11 December 2009

REFERENCES

Arrigo, S.J., S. Weitsman, J.A. Zack, and I.S. Chen. 1990. Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. *J. Virol.* 64:4585–4588.

Bhattacharya, T., M. Daniels, D. Heckerman, B. Foley, N. Frahm, C. Kadie, J. Carlson, K. Yusim, B. McMahon, B. Gaschen, et al. 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science.* 315:1583–1586. doi:10.1126/science.1131528

Bihl, F., N. Frahm, L. Di Giammarino, J. Sidney, M. John, K. Yusim, T. Woodberry, K. Sango, H.S. Hewitt, L. Henry, et al. 2006. Impact of HLA-B alleles, epitope binding affinity, functional avidity, and viral coinfection on the immunodominance of virus-specific CTL responses. *J. Immunol.* 176:4094–4101.

Bihl, F., M. Narayan, J.V. Chisholm III, L.M. Henry, T.J. Suscovich, E.E. Brown, T.M. Welzel, D.E. Kaufmann, T.M. Zaman, S. Dollard, et al. 2007. Lytic and latent antigens of the human gammaherpesviruses Kaposi's sarcoma-associated herpesvirus and Epstein-Barr virus induce T-cell responses with similar functional properties and memory phenotypes. *J. Virol.* 81:4904–4908.

Bihl, F.K., E. Loggi, J.V. Chisholm III, H.S. Hewitt, L.M. Henry, C. Linde, T.J. Suscovich, J.T. Wong, N. Frahm, P. Andreone, and C. Brander. 2005. Simultaneous assessment of cytotoxic T lymphocyte responses against multiple viral infections by combined usage of optimal epitope matrices, anti- CD3 mAb T-cell expansion and "RecycleSpot". *J. Transl. Med.* 3:20. doi:10.1186/1479-5876-3-20

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

Brander, C., P.J.R. Goulder, K. Luzuriaga, O.O. Yang, K.E. Hartman, N.G. Jones, B.D. Walker, and S.A. Kalams. 1999. Persistent HIV-1-specific CTL clonal expansion despite high viral burden post in utero HIV-1 infection. *J. Immunol.* 162:4796–4800.

Brander, C., N. Frahm, and B.D. Walker. 2006. The challenges of host and viral diversity in HIV vaccine design. *Curr. Opin. Immunol.* 18:430–437. doi:10.1016/j.coi.2006.05.012

Brumme, Z.L., C.J. Brumme, D. Heckerman, B.T. Korber, M. Daniels, J. Carlson, C. Kadie, T. Bhattacharya, C. Chui, J. Szinger, et al. 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. *PLoS Pathog.* 3:e94. doi:10.1371/journal.ppat.0030094

Brumme, Z.L., I. Tao, S. Szeto, C.J. Brumme, J.M. Carlson, D. Chan, C. Kadie, N. Frahm, C. Brander, B. Walker, et al. 2008. Human leukocyte antigen-specific polymorphisms in HIV-1 Gag and their association with viral load in chronic untreated infection. *AIDS.* 22:1277–1286. doi:10.1097/QAD.0b013e3283021a8c

Brumme, Z.L., M. John, J.M. Carlson, C.J. Brumme, D. Chan, M.A. Brockman, L.C. Swenson, I. Tao, S. Szeto, P. Rosato, et al. 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS One.* 4:e6687. doi:10.1371/journal.pone.0006687

Bullock, T.N., and L.C. Eisenlohr. 1996. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *J. Exp. Med.* 184:1319–1329. doi:10.1084/jem.184.4.1319

Bullock, T.N., A.E. Patterson, L.L. Franlin, E. Notidis, and L.C. Eisenlohr. 1997. Initiation codon scanthrough versus termination codon readthrough demonstrates strong potential for major histocompatibility complex class I–restricted cryptic epitope expression. *J. Exp. Med.* 186:1051–1058. doi:10.1084/jem.186.7.1051

Cardinaud, S., A. Moris, M. Février, P.S. Rohrlich, L. Weiss, P. Langlade-Demoyen, F.A. Lemonnier, O. Schwartz, and A. Habel. 2004. Identification of cryptic MHC I–restricted epitopes encoded by HIV-1 alternative reading frames. *J. Exp. Med.* 199:1053–1063. doi:10.1084/jem.20031869

Carlson, J.M., Z.L. Brumme, C.M. Rousseau, C.J. Brumme, P. Matthews, C. Kadie, J.I. Mullins, B.D. Walker, P.R. Harrigan, P.J. Goulder, and D. Heckerman. 2008. Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. *PLOS Comput. Biol.* 4:e1000225. doi:10.1371/journal.pcbi.1000225

Chen, H., A. Piechocka-Trocha, T. Miura, M.A. Brockman, B.D. Julg, B.M. Baker, A.C. Rothchild, B.L. Block, A. Schneidewind, T. Koibuchi, et al. 2009. Differential neutralization of human immunodeficiency virus (HIV) replication in autologous CD4 T cells by HIV-specific cytotoxic T lymphocytes. *J. Virol.* 83:3138–3149. doi:10.1128/JVI.02073-08

Frahm, N., P.J.R. Goulder, and C. Brander. 2002. Total assessment of HIV-specific CTL responses: epitope clustering, processing preferences, and the impact of HIV sequence heterogeneity. *In* HIV Molecular Immunology Compendium. B.T. Korber, C. Brander, B.F. Haynes, R. Koup, C. Kuiken, J.P. Moore, B.D. Walker, and D.I. Watkins, editors. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM. I-1–22.

Frahm, N., B.T. Korber, C.M. Adams, J.J. Szinger, R. Draenert, M.M. Addo, M.E. Feeney, K. Yusim, K. Sango, N.V. Brown, et al. 2004. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J. Virol.* 78:2187–2200. doi:10.1128/JVI.78.5.2187-2200.2004

Frahm, N., K. Yusim, T.J. Suscovich, S. Adams, J. Sidney, P. Hraber, H.S. Hewitt, C.H. Linde, D.G. Kavanagh, T. Woodberry, et al. 2007. Extensive HLA class I allele promiscuity among viral CTL epitopes. *Eur. J. Immunol.* 37:2419–2433. doi:10.1002/eji.200737365

Frahm, N., B. Baker, and C. Brander. 2008. Identification and optimal definition of HIV-derived cytotoxic T-lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution. *In* HIV Molecular Immunology 2008. B.T. Korber, C. Brander, B.F. Haynes, R. Koup, J.P. Moore, B.D. Walker, and D.I. Watkins, editors. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM. 3–24.

Gallimore, A., T. Dumrese, H. Hengartner, R.M. Zinkernagel, and H.G. Rammensee. 1998. Protective immunity does not correlate with the hierarchy of virus-specific cytotoxic T cell responses to naturally processed peptides. *J. Exp. Med.* 187:1647–1657. doi:10.1084/jem.187.10.1647

Gaudray, G., F. Gachon, J. Basbous, M. Biard-Piechaczyk, C. Devaux, and J.M. Mesnard. 2002. The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. *J. Virol.* 76:12813–12822. doi:10.1128/JVI.76.24.12813-12822.2002

Goulder, P.J., C. Brander, K. Annamalai, N. Mngqundaniso, U. Govender, Y. Tang, S. He, K.E. Hartman, C.A. O'Callaghan, G.S. Ogg, et al. 2000. Differential narrow focusing of immunodominant human immunodeficiency virus gag-specific cytotoxic T-lymphocyte responses in infected African and caucasoid adults and children. *J. Virol.* 74:5679–5690. doi:10.1128/JVI.74.12.5679-5690.2000

Goulder, P.J., M.M. Addo, M.A. Altfeld, E.S. Rosenberg, Y. Tang, U. Govender, N. Mngqundaniso, K. Annamalai, T.U. Vogel, M. Hammond, et al. 2001. Rapid definition of five novel HLA-A*3002-restricted human immunodeficiency virus-specific cytotoxic T-lymphocyte epitopes by elispot and intracellular cytokine staining assays. *J. Virol.* 75:1339–1347. doi:10.1128/JVI.75.3.1339-1347.2001

Heckerman, D., C. Kadie, and J. Listgarten. 2007. Leveraging information across HLA alleles/supertypes improves epitope prediction. *J. Comput. Biol.* 14:736–746. doi:10.1089/cmb.2007.R013

Hickman, H.D., A.D. Luis, W. Bardet, R. Buchli, C.L. Battson, M.H. Shearer, K.W. Jackson, R.C. Kennedy, and W.H. Hildebrand. 2003. Cutting edge: class I presentation of host peptides following HIV infection. *J. Immunol.* 171:22–26.

Ho, O., and W.R. Green. 2006. Cytolytic CD8+ T cells directed against a cryptic epitope derived from a retroviral alternative reading frame confer disease protection. *J. Immunol.* 176:2470–2475.

Hogg, R.S., B. Yip, K.J. Chan, E. Wood, K.J. Craib, M.V. O'Shaughnessy, and J.S. Montaner. 2001. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA.* 286:2568–2577. doi:10.1001/jama.286.20.2568

Jacks, T., M.D. Power, F.R. Masiarz, P.A. Luciw, P.J. Barr, and H.E. Varmus. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature.* 331:280–283. doi:10.1038/331280a0

Jin, X., D.E. Bauer, S.E. Tuttleton, S. Lewin, A. Gettie, J. Blanchard, C.E. Irwin, J.T. Safrit, J. Mittler, L. Weinberger, et al. 1999. Dramatic rise in plasma viremia after $CD8^+$ T cell depletion in simian immunodeficiency virus–infected macaques. *J. Exp. Med.* 189:991–998. doi:10.1084/jem.189.6.991

Jones, N., D. Agrawal, M. Elrefaei, A. Hanson, V. Novitsky, J.T. Wong, and H. Cao. 2003. Evaluation of antigen-specific responses using in vitro enriched T cells. *J. Immunol. Methods.* 274:139–147. doi:10.1016/S0022-1759(02)00510-0

Koup, R.A., J.T. Safrit, Y. Cao, C.A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, and D.D. Ho. 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J. Virol.* 68:4650–4655.

Landry, S., M. Halin, S. Lefort, B. Audet, C. Vaquero, J.M. Mesnard, and B. Barbeau. 2007. Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology.* 4:71. doi:10.1186/1742-4690-4-71

Leitner, T., B. Foley, B. Hahn, P. Marx, F. McCutchan, J. Mellors, S. Wolinksy, and B. Korber, editors. 2005. HIV Sequence Compendium 2005. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. 698 pp.

Leslie, A., D. Kavanagh, I. Honeyborne, K. Pfafferott, C. Edwards, T. Pillay, L. Hilton, C. Thobakgale, D. Ramduth, R. Draenert, et al. 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J. Exp. Med.* 201:891–902. doi:10.1084/jem.20041455

Maness, N.J., L.E. Valentine, G.E. May, J. Reed, S.M. Piaskowski, T. Soma, J. Furlott, E.G. Rakasz, T.C. Friedrich, D.A. Price, et al. 2007. AIDS virus–specific $CD8^+$ T lymphocytes against an immunodominant cryptic epitope select for viral escape. *J. Exp. Med.* 204:2505–2512. doi:10.1084/jem.20071261

Mayrand, S.M., and W.R. Green. 1998. Non-traditionally derived CTL epitopes: exceptions that prove the rules? *Immunol. Today.* 19:551–556. doi:10.1016/S0167-5699(98)01342-5

Mayrand, S.M., D.A. Schwarz, and W.R. Green. 1998. An alternative translational reading frame encodes an immunodominant retroviral CTL determinant expressed by an immunodeficiency-causing retrovirus. *J. Immunol.* 160:39–50.

Miller, R.H. 1988. Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. *Science.* 239:1420–1422. doi:10.1126/science.3347840

Moore, C.B., M. John, I.R. James, F.T. Christiansen, C.S. Witt, and S.A. Mallal. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science.* 296:1439–1443. doi:10.1126/science.1069660

Northfield, J.W., C.P. Loo, J.D. Barbour, G. Spotts, F.M. Hecht, P. Klenerman, D.F. Nixon, and J. Michaëlsson. 2007. Human immunodeficiency virus type 1 (HIV-1)-specific CD8+ T(EMRA) cells in early infection are linked to control of HIV-1 viremia and predict the subsequent viral load set point. *J. Virol.* 81:5759–5765. doi:10.1128/JVI.00045-07

Probst-Kepper, M., V. Stroobant, R. Kridel, B. Gaugler, C. Landry, F. Brasseur, J.P. Cosyns, B. Weynand, T. Boon, and B.J. Van den Eynde. 2001. An alternative open reading frame of the human macrophage colony-stimulating factor gene is independently translated and codes for an antigenic peptide of 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes. *J. Exp. Med.* 193:1189–1198. doi:10.1084/jem.193.10.1189

Rimoldi, D., V. Rubio-Godoy, V. Dutoit, D. Lienard, S. Salvi, P. Guillaume, D. Speiser, E. Stockert, G. Spagnoli, C. Servis, et al. 2000. Efficient simultaneous presentation of NY-ESO-1/LAGE-1 primary and nonprimary open reading frame-derived CTL epitopes in melanoma. *J. Immunol.* 165:7253–7261.

Rousseau, C.M., M.G. Daniels, J.M. Carlson, C. Kadie, H. Crawford, A. Prendergast, P. Matthews, R. Payne, M. Rolland, D.N. Raugi, et al. 2008. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. *J. Virol.* 82:6434–6446. doi:10.1128/JVI.02455-07

Rudolph, M.G., R.L. Stanfield, and I.A. Wilson. 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* 24:419–466. doi:10.1146/annurev.immunol.23.021704.115658

Saeterdal, I., J. Bjørheim, K. Lislerud, M.K. Gjertsen, I.K. Bukholm, O.C. Olsen, J.M. Nesland, J.A. Eriksen, M. Møller, A. Lindblom, and G. Gaudernack. 2001. Frameshift-mutation-derived peptides as tumor-specific

antigens in inherited and spontaneous colorectal cancer. *Proc. Natl. Acad. Sci. USA.* 98:13255–13260. doi:10.1073/pnas.231326898

Saulquin, X., E. Scotet, L. Trautmann, M.A. Peyrat, F. Halary, M. Bonneville, and E. Houssaint. 2002. +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J. Exp. Med.* 195:353–358. doi:10.1084/jem.20011399

Schmitz, J.E., M.J. Kuroda, S. Santra, V.G. Sasseville, M.A. Simon, M.A. Lifton, P. Racz, K. Tenner-Racz, M. Dalesandro, B.J. Scallon, et al. 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science.* 283:857–860.

Schwab, S.R., J.A. Shugart, T. Horng, S. Malarkannan, and N. Shastri. 2004. Unanticipated antigens: translation initiation at CUG with leucine. *PLoS Biol.* 2:e366. doi:10.1371/journal.pbio.0020366

Shibata, K., M. Imarai, G.M. van Bleek, S. Joyce, and S.G. Nathenson. 1992. Vesicular stomatitis virus antigenic octapeptide N52-59 is anchored into the groove of the H-2Kb molecule by the side chains of three amino acids and the main-chain atoms of the amino terminus. *Proc. Natl. Acad. Sci. USA.* 89:3135–3139. doi:10.1073/pnas.89.7.3135

Sidney, J., B. Peters, N. Frahm, C. Brander, and A. Sette. 2008. HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 9:1. doi:10.1186/1471-2172-9-1

Storey, J.D., and R. Tibshirani. 2003. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. USA.* 100:9440–9445. doi:10.1073/pnas.1530509100

Streeck, H., J.S. Jolin, Y. Qi, B. Yassine-Diab, R.C. Johnson, D.S. Kwon, M.M. Addo, C. Brumme, J.P. Routy, S. Little, et al. 2009. 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.* 83:7641–7648. doi:10.1128/JVI.00182-09

Sykulev, Y., M. Joo, I. Vturina, T.J. Tsomides, and H.N. Eisen. 1996. Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. *Immunity.* 4:565–571. doi:10.1016/S1074-7613(00)80483-5

Wang, R.F., M.R. Parkhurst, Y. Kawakami, P.F. Robbins, and S.A. Rosenberg. 1996. Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. *J. Exp. Med.* 183:1131–1140. doi:10.1084/jem.183.3.1131

Watts, J.M., K.K. Dang, R.J. Gorelick, C.W. Leonard, J.W. Bess Jr., R. Swanstrom, C.L. Burch, and K.M. Weeks. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature.* 460:711–716. doi:10.1038/nature08237

Xu, Z., J. Choi, T.S. Yen, W. Lu, A. Strohecker, S. Govindarajan, D. Chien, M.J. Selby, and J. Ou. 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J.* 20:3840–3848. doi:10.1093/emboj/20.14.3840

Yang, O.O., S.A. Kalams, M. Rosenzweig, A. Trocha, N. Jones, M. Koziel, B.D. Walker, and R.P. Johnson. 1996. Efficient lysis of human immunodeficiency virus type 1-infected cells by cytotoxic T lymphocytes. *J. Virol.* 70:5799–5806.

Yant, L.J., T.C. Friedrich, R.C. Johnson, G.E. May, N.J. Maness, A.M. Enz, J.D. Lifson, D.H. O'Connor, M. Carrington, and D.I. Watkins. 2006. The high-frequency major histocompatibility complex class I allele Mamu-B*17 is associated with control of simian immunodeficiency virus SIVmac239 replication. *J. Virol.* 80:5074–5077. doi:10.1128/JVI.80.10.5074-5077.2006

Yewdell, J.W., and C.V. Nicchitta. 2006. The DRiP hypothesis decennial: support, controversy, refinement and extension. *Trends Immunol.* 27:368–373. doi:10.1016/j.it.2006.06.008

Zuñiga, R., A. Lucchetti, P. Galvan, S. Sanchez, C. Sanchez, A. Hernandez, H. Sanchez, N. Frahm, C.H. Linde, H.S. Hewitt, et al. 2006. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. *J. Virol.* 80:3122–3125. doi:10.1128/JVI.80.6.3122-3125.2006