Immune Screening Identifies Novel T Cell Targets Encoded by Antisense Reading Frames of HIV-1

Christoph T. Berger,[a,b] Anuska Llano,[c] Jonathan M. Carlson,[d] Zabrina L. Brumme,[e,f] Mark A. Brockman,[e,f] Samandhy Cedeño,[c] P. Richard Harrigan,[f] Daniel E. Kaufmann,[a,g] David Heckerman,[d] Andreas Meyerhans,[h,i] Christian Brander[c,i,j,k]

Ragon Institute of MGH, MIT and Harvard, Boston, Massachusetts, USA[a]; Translational Immunology, Department of Biomedicine, University Hospital Basel, Basel, Switzerland[b]; IrsiCaixa AIDS Research Institute-HIVACAT, Hospital Germans Trias I Pujol, Badalona, Spain[c]; Microsoft Research, Seattle, Washington, USA[d]; Simon Fraser University, Burnaby, BC, Canada[e]; British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada[f]; University of Montreal Hospital Research Center (CRCHUM), Montreal, QC, Canada[g]; Infection Biology Group, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain[h]; Institucio Catalana de Recerca i Estudis Avancats (ICREA), Barcelona, Spain[i]; University of Vic and Central University of Catalonia, Vic, Catalonia, Spain[j]; Autonomous University of Barcelona, Barcelona, Catalonia, Spain[k]

Cytotoxic-T lymphocyte (CTL) responses to epitopes in alternative HIV reading frames have been reported. However, the extent of CTL responses to putative proteins encoded in antisense reading frames is unknown. Using sequence alignments and computational approaches, we here predict five potential antisense HIV proteins and characterize common CTL responses against them. Results suggest that antisense-derived sequences are commonly transcribed and translated and could encode functional proteins that contain important targets of anti-HIV cellular immunity.

Virus-specific CD8$^+$ cytotoxic-T lymphocyte (CTL) responses are important in the control of HIV infection (1–3). Accumulating evidence indicates that in addition to CTL responses targeting known structural and nonstructural HIV proteins, CTL responses to epitopes in alternative reading frames (ARFs) of HIV can contribute to *in vivo* immune selection pressure (4–6). Importantly, these CTL responses are not limited to ARFs in the sense direction, but may also target antisense-encoded ARFs (5, 6). Antisense transcription of known host genes in human and rodent cells has been extensively described (7–10), and the existence of an antisense protein (ASP) in HIV was first proposed more than 20 years ago (10–12). The extent to which such viral antisense transcripts encode functional proteins remains largely unknown.

Here, we sought to provide further evidence supporting the existence of fully translated ASPs in HIV by assessing the presence of ASP-specific cellular host immune responses. In a first step, 355 unique full-length HIV-1 clade B sequences were retrieved from the Los Alamos database (13) to generate a consensus sequence for all three antisense reading frames. Five sequences starting with a classical methionine translation start codon followed by >150 nucleotides without stop codons and with conservation of >50% (range, 56 to 78%) at all amino acid positions were identified (Fig. 1). The highest conservation scores were obtained for the MAN sequence (entropy score, 0.04; i.e., 99.4% conservation), followed by those for MLK (0.11, 98.2%), MGS (0.18, 97.9%), MPQ (0.16, 97.4%), and the putative MAL protein (0.31, 92.5%). This analysis placed the five ASPs in the range of conservation between the highly conserved HIV p24 (0.21, 99.7%) and the variable Vpu (0.34, 92%) proteins, overall being comparable to other regular reading frame-encoded HIV proteins (14). The ASP sequences ranged from 50 to 179 amino acids in length and included a previously predicted ASP (11). To find immunological evidence that proteins were indeed produced from these antisense open reading frames (ORFs), HIV-infected individuals were tested for T cell responses against corresponding consensus overlapping peptide (OLP) sets by gamma interferon (IFN-γ)-enzyme-linked immunosorbent spot (ELISpot) assay (4) (see Table 1 in the supplemental material). A total of 60 HLA-typed individuals, including 40

chronically HIV-infected subjects and 20 HIV-uninfected controls, were tested. Responses to the ASP-derived OLPs were detected exclusively in HIV-infected persons (15 of 40 [37.5%] showed reactivities to at least one OLP) and targeted all five putative ASPs. Most (n = 22 [30%] and n = 24 [36%], respectively, of 66 responses in total) were detected against the previously undescribed antisense-encoded MGS protein within the +1 antisense reading frame (ASRF) of Pol and the previously reported MPQ ASP within the +1 ASRF of Env (11) (Fig. 2A and B). The remaining three putative ASPs, MAN, MLF, and MAL, were less frequently targeted, with three, six, and nine responses, respectively (see Fig. S1 in the supplemental material). Responses to the two most commonly targeted OLPs, M27 (PLMGGAYIAFPTSC HMFI within MPQ ASP, targeted by 7/40 HIV-infected individuals) and M49 (SVNCFTSLVWAPLILAYF, within MGS ASP, targeted by 11/40 subjects), were further characterized (Fig. 2B to D). Peptide titration analyses (Fig. 2C) were performed to (i) determine the functional avidity of the reactive T cells and (ii) rule out that the responses were artifacts of unspecific reactivities due to high peptide concentrations used in the screening assays (15, 16). The data show the peptide concentrations that elicited the half-maximum response (SD 50%) that were consistent with what
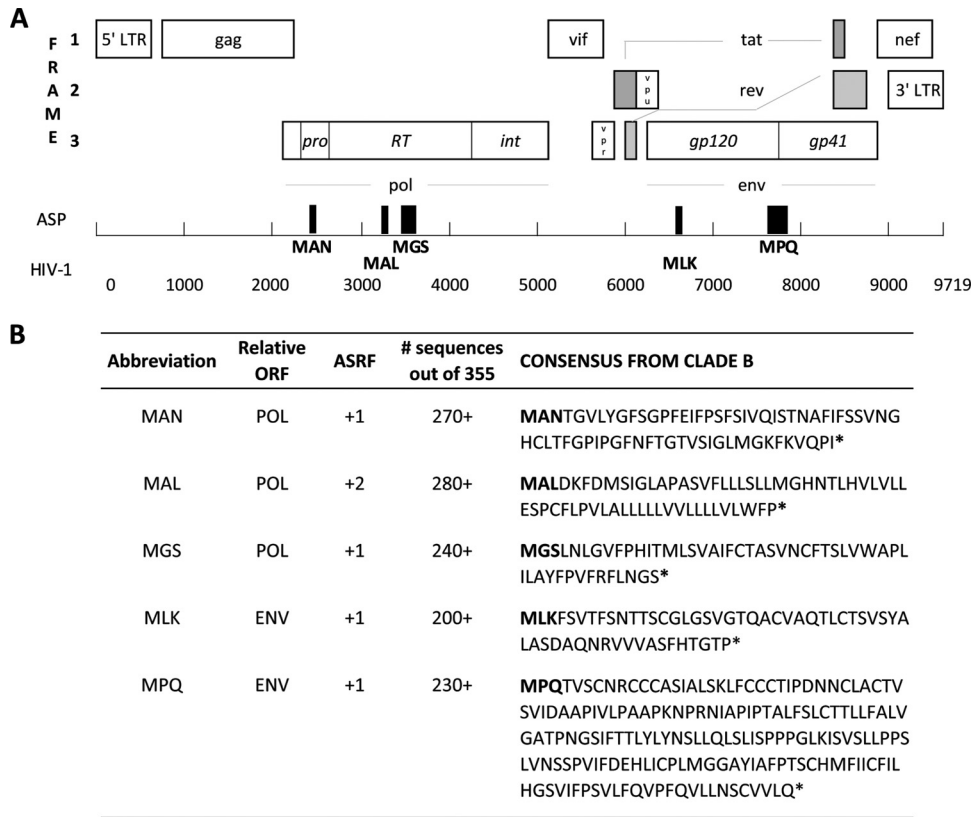
FIG 1 Sequences and localization of putative antisense proteins (ASP) in the HIV genome. (A) The full-length HIV genome is shown with structural and accessory proteins encoded in the three regular open reading frames (ORFs). Black boxes depict the location of the five predicted ASPs relative to the regular ORFs. (B) The consensus amino acid sequences of the five predicted proteins are shown. Their antisense reading frames (ASRFs) and the frequencies at which they were found among 355 HIV isolates are indicated. *, stop codon.

would be expected for responses to epitopes in standard ORFs. To further demonstrate the potential HLA class I restriction of the MPQ-M27 and MGS-M49 responses, flow cytometry-based HLA restriction analyses were performed (4), identifying HLA-B*35:01 and HLA-C*04:01 as potential restricting elements and showing that responses were mediated by CD8+ T cells (Fig. 2D and E). While the data suggest that these immunodominant antisense-encoded epitopes are targeted by CD8 T cells, we cannot exclude that some of the observed responses against antisense-encoded OLPs may be CD4 T cell mediated.

To further support the notion that ASPs were expressed and can be targeted by host T cells, we subjected the five ASP sequences to a CTL epitope prediction analysis, using the "Epipred" algorithm trained on well-characterized CD8+ T cell epitopes from the Immune Epitope Database (IEDB) and the Los Alamos HIV database (17). Epipred predicts epitopes based on scores determined by the probability of binding to respective HLA molecules, also considering the presence of specific epitope-flanking regions. For instance, a prediction score of 0.98 indicates a 98% probability that the predicted epitope indeed can be presented and recognized by a T cell. Based on the HLA type of each screened individual and OLP pair, we determined the best epitope (i.e., the epitope with the highest prediction score) for each response by testing all 8 to 11 mers embedded within the OLP sequence against all six HLA class I alleles of the responding subject. As expected, several individuals who responded to the same OLP likely did so through the same

HLA-epitope pair. Interestingly, the median posterior probability score for the best-predicted epitopes was significantly higher for those OLPs that were in fact targeted in the ELISpot screens compared to the nontargeted OLPs (P = 0.004) (Fig. 3A). Similarly, the distribution of best-predicted epitopes per OLP and subject showed an enrichment of epitopes with elevated prediction scores in targeted OLPs compared to those in nontargeted OLPs (Fig. 3B), thus supporting a direct link between high epitope prediction score and ELISpot reactivity. These findings are further validated by the Epipred predictions for the two most immunodominant OLPs, M49 and M27, when not only the best-, but all, predicted epitopes were considered. In the case of M49, Epipred predicted CFTSLVWAPLI (CI-11) as an HLA-C*04-restricted epitope with a prediction score of 0.90 (Fig. 3C) and NCFTSLVWAPL (NL-11) as a putative HLA-C*01 and C*07 epitope (posterior probabilities of 87% and 83%, respectively), in line with the fact that 8 out of the 11 M49 responders expressed at least one of these HLA class I alleles. With lower probability, Epiprep also predicted a putative HLA-B*35:01-restricted epitope, APLILAYFPVF (AF-11; score of 0.51), still in line with three of the M49 responders expressing HLA-B*35:01. Similarly, in the case of OLP M27, Epipred predicted AYIAFPTSCHM (AM-11) as being HLA-C*0401 restricted, which is consistent with three out of seven responders expressing this allele. The detection of additional OLP-specific responses in individuals not expressing the predicted HLA alleles is in agreement with the possible presence of additional CTL
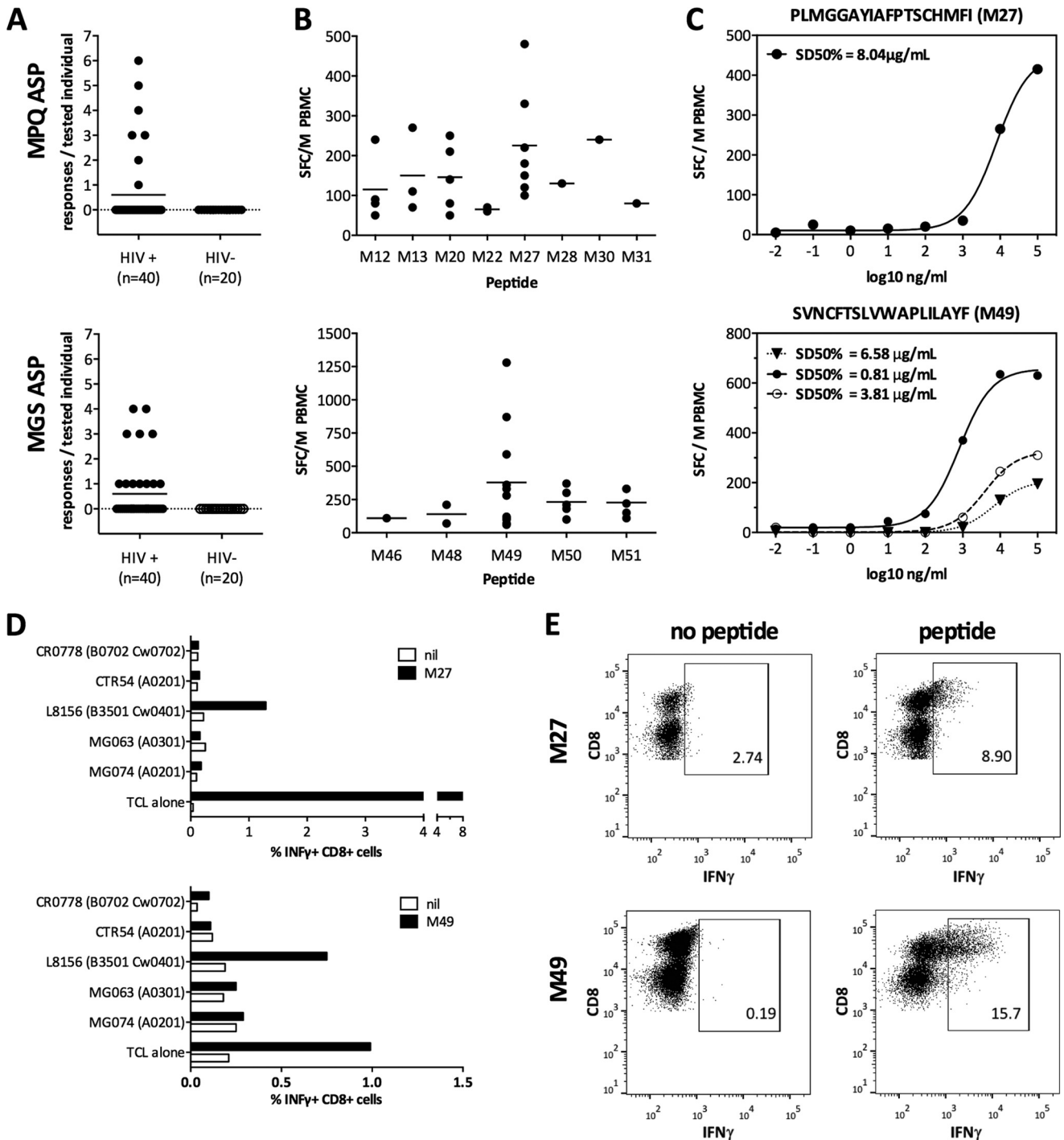
FIG 2 CTL responses to antisense proteins are common and HIV specific. Peripheral blood mononuclear cells (PBMCs) were tested in ELISpot assays for reactivity against overlapping peptides at a peptide concentration of 28 μg/ml as described previously (4). Responses of >50 spot-forming cells per million (SFC/M) PBMCs were considered positive. (A) The breadths (i.e., number of responses per tested individual) of the two most frequently targeted ASPs, MPQ (upper panel) and MGS (lower panel), are displayed. None of the HIV-negative controls showed a response to any of the peptides tested. (B) Magnitudes (SFC/M PBMCs) of the positive responses targeting MPQ (upper panel) or MGS (lower panel) are indicated. (C) Peptide titrations (ranging from 100 μg/ml to 0.001 pg/ml) were performed on unexpanded PBMCs to determine the functional avidity (SD 50%) of responses targeting M27 (upper panel) and M49 (lower panel). (D) Partially HLA-matched B cell lines were peptide pulsed (M27, upper panel; M49, lower panel) and incubated with peptide-specific expanded T cells from HIV-infected individuals. Black bars indicate the IFN-γ response as assessed by flow cytometry; white bars indicate the background response (media alone). The indicated HLA alleles represent alleles matched between antigen-presenting cells and the T cell line (TCL) reactive with the tested peptide. Values are expressed as percent IFN-γ$^+$ CD8$^+$ T cells after stimulation for 5 h. (E) Intracellular cytokine staining following incubation for 5 h with M27 (upper panel) or M49 (lower panel), with peptides directly added to the T cell lines, confirmed the phenotype of the responding cells as CD8$^+$ T cells. Panels show peptide-specific T cell lines stimulated with media alone (left) versus with peptide (right). The percent IFN-γ$^+$ cells is indicated. Experiments for panels D and E were performed at different time points, resulting in different maximal responses.

OLP#	HLA	Predicted Epitope	Epitope Score	n with HLA and OLP response
M49	**C0401**	**CFTSLVWAPLI**	**0.90**	**4 / 11**
M2	B3501	GPIFEIFPSF	0.90	1 / 1
M27	**C0401**	**AYIAFPTSCHM**	**0.87**	**3 / 7**
M49	**C0102**	**NCFTSLVWAPL**	**0.87**	**1 / 11**
M49	**B0801**	**TSLVWAPL**	**0.74**	**2 / 11**
M46	C0401	VFPHITML	0.74	1 / 1
M27	**C0501**	**AYIAFPTSCHM**	**0.72**	**1 / 7**
M28	C0701	TSCHMFIICF	0.69	1 / 1
M20	C0401	IFTTLYLYNSL	0.67	2 / 5
M49	**C0304**	**CFTSLVWAPLI**	**0.64**	**1 / 11**
M48	C0401	IFCTASVNCF	0.62	2 / 2
M49	**C0701**	**TSLVWAPL**	**0.60**	**1 / 11**
M49	**A0201**	**SLVWAPLI**	**0.59**	**1 / 11**
M45	A0201	LLLVVLLLLV	0.59	1 / 1
M49	**A0217**	**SLVWAPLI**	**0.58**	**1 / 11**
M27	**C0602**	**PLMGGAYIAF**	**0.56**	**1 / 7**
M40	B0801	APASVFLLL	0.55	1 / 3
M27	**B5703**	**AYIAFPTSCHM**	**0.53**	**1 / 7**
M40	C0401	VFLLLSLLM	0.52	1 / 3
M50	B3501	APLILAYFPVF	0.51	3 / 5
M51	C0401	YFPVFRFL	0.48	2 / 5
M44	C0401	CFLPVLALLL	0.48	2 / 3
M20	A3002	IFTTLYLY	0.47	1 / 5
M34	C0802	QACVAQTL	0.47	1 / 1
M20	B5703	IFTTLYLY	0.45	1 / 5
M31	C0401	FQVPFQVLL	0.42	1 / 1
M6	B5703	LTFGPSIPGF	0.42	1 / 2

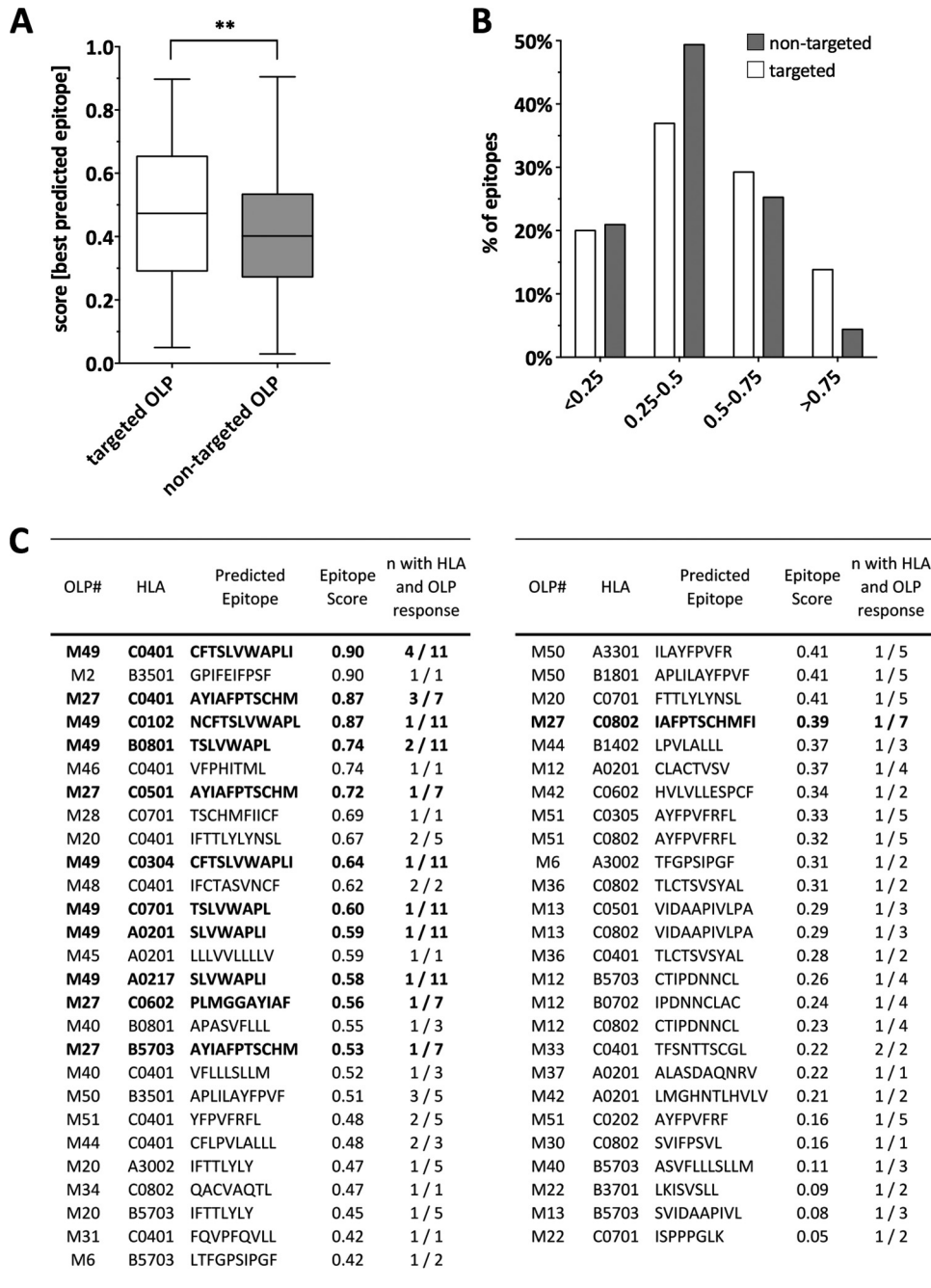| OLP#	HLA	Predicted Epitope	Epitope Score	n with HLA and OLP response
M50	A3301	ILAYFPVFR	0.41	1 / 5
M50	B1801	APLILAYFPVF	0.41	1 / 5
M20	C0701	FTTLYLYNSL	0.41	1 / 5
M27	**C0802**	**IAFPTSCHMFI**	**0.39**	**1 / 7**
M44	B1402	LPVLALLL	0.37	1 / 3
M12	A0201	CLACTVSV	0.37	1 / 4
M42	C0602	HVLVLLESPCF	0.34	1 / 2
M51	C0305	AYFPVFRFL	0.33	1 / 5
M51	C0802	AYFPVFRFL	0.32	1 / 5
M6	A3002	TFGPSIPGF	0.31	1 / 2
M36	C0802	TLCTSVSYAL	0.31	1 / 2
M13	C0501	VIDAAPIVLPA	0.29	1 / 3
M13	C0802	VIDAAPIVLPA	0.29	1 / 3
M36	C0401	TLCTSVSYAL	0.28	1 / 2
M12	B5703	CTIPDNNCL	0.26	1 / 4
M12	B0702	IPDNNCLAC	0.24	1 / 4
M12	C0802	CTIPDNNCL	0.23	1 / 4
M33	C0401	TFSNTTSCGL	0.22	2 / 2
M37	A0201	ALASDAQNRV	0.22	1 / 1
M42	A0201	LMGHNTLHVLV	0.21	1 / 2
M51	C0202	AYFPVFRF	0.16	1 / 5
M30	C0802	SVIFPSVL	0.16	1 / 1
M40	B5703	ASVFLLLSLLM	0.11	1 / 3
M22	B3701	LKISVSLL	0.09	1 / 2
M13	B5703	SVIDAAPIVL	0.08	1 / 3
M22	C0701	ISPPPGLK	0.05	1 / 2

FIG 3 Epitope predictions are consistent with experimental data and support high probability of extensive CTL targeting of antisense-protein-derived epitopes. (A) The posterior probability scores of the best-predicted epitopes for each overlapping peptide (OLP) and subject's HLA were compared between OLPs that were (open bars) or were not (gray bars) targeted in HIV-infected individuals. Each best epitope-HLA allele combination was counted separately for both targeted and nontargeted epitopes. (B) Distribution of predicted epitopes based on their prediction scores stratified by targeted and nontargeted OLPs. The percentage of predicted epitopes in each prediction score quartile is shown, with results for epitopes located in targeted OLPs shown as open bars and results for epitopes in nontargeted OLPs shown as gray bars. (C) The best-predicted epitope sequences for each HLA-OLP pair are displayed, and the number of individuals with the respective HLA-OLP combination is indicated. The high scores for predicted HLA restrictions of M27 and M49 (in bold) match the experimentally determined HLA restriction (Fig. 2D).

epitopes in this OLP sequence or epitopes with a wide HLA promiscuity (18). Overall, the epitope prediction analyses provide further evidence that OLP responses were mediated by HLA-I-restricted T cells and provide specific candidate epitopes that may drive these responses.

Together, the present data identify five potential ASP sequences in HIV which are relatively frequently targeted by T cells in HIV-infected subjects. Thus, these data (i) strongly support the expression and translation of these proteins in HIV infection and (ii) identify additional T cell targets that may contribute to host

immune control of HIV. Relative to their length and the frequency whereby they are targeted, responses to these ASPs compare well with observed reactivities to epitopes encoded in forward cryptic HIV reading frames (4) and response rates to well-studied regulatory HIV proteins, such as Tat, Vif, and Rev (19). The response rates are also comparable to activities reported by Champiat et al., who identified responses to ASPs that partially overlapped with four of our ASP sequences (6). Together, the present study corroborates earlier reports of responses to epitopes in alternative HIV reading frames and suggests that HIV may express several additional, relatively conserved proteins that may be under evolutionary constraints and thus exert hitherto-unknown functions.

ACKNOWLEDGMENTS

REFERENCES

1. Schmitz J, Kuroda M, Santra S, Sasseville V, Simon M, Lifton M, Racz P, Tenner-Racz K, Dalesandro M, Scallon B, Ghrayeb J, Forman M, Montefiori D, Rieber E, Letvin N, Reimann K. 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.

2. Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, Borkowsky W, Farthing C, Ho DD. 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.

3. Streeck H, Jolin JS, Qi Y, Yassine-Diab B, Johnson RC, Kwon DS, Addo MM, Brumme C, Routy JP, Little S, Jessen HK, Kelleher AD, Hecht FM, Sekaly RP, Rosenberg ES, Walker BD, Carrington M, Altfeld M. 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. http://dx.doi.org/10.1128/JVI.00182-09.

4. Berger CT, Carlson JM, Brumme CJ, Hartman KL, Brumme ZL, Henry LM, Rosato PC, Piechocka-Trocha A, Brockman MA, Harrigan PR, Heckerman D, Kaufmann DE, Brander C. 2010. Viral adaptation to immune selection pressure by HLA class I-restricted CTL responses targeting epitopes in HIV frameshift sequences. J Exp Med 207:61–75. http://dx.doi.org/10.1084/jem.20091808.

5. Bansal A, Carlson J, Yan J, Akinsiku OT, Schaefer M, Sabbaj S, Bet A, Levy DN, Heath S, Tang J, Kaslow RA, Walker BD, Ndung'u T, Goulder PJ, Heckerman D, Hunter E, Goepfert PA. 2010. CD8 T cell response and evolutionary pressure to HIV-1 cryptic epitopes derived from antisense transcription. J Exp Med 207:51–59. http://dx.doi.org/10.1084/jem.20092060.

6. Champiat S, Andre R, Raposo S, Maness NJ, Lehman JL, Purtell SE, Hasenkrug A, Miller JC, Dean H, Koff WC, Hong CM, Martin JN, Deeks SG, Spotts G, Pilcher CD, Hecht FM, Kallas E, Garrison K, Nixon D. 2012. Influence of HAART on alternative reading frame immune responses over the course of HIV-1 infection. PLoS One 7:e39311. http://dx.doi.org/10.1371/journal.pone.0039311.

7. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. 2008. The antisense transcriptomes of human cells. Science 322:1855–1857. http://dx.doi.org/10.1126/science.1163853.

8. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C. 2005. Antisense transcription in the mammalian transcriptome. Science 309:1564–1566. http://dx.doi.org/10.1126/science.1112009.

9. Lavorgna G, Dahary D, Lehner B, Sorek R, Sanderson CM, Casari G. 2004. In search of antisense. Trends Biochem Sci 29:88–94. http://dx.doi.org/10.1016/j.tibs.2003.12.002.

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

11. Miller RH. 1988. Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. Science 239:1420–1422. http://dx.doi.org/10.1126/science.3347840.

12. Vanhée-Brossollet C, Thoreau H, Serpente N, D'Auriol L, Levy JP, Vaquero C. 1995. A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals. Virology 206:196–202. http://dx.doi.org/10.1016/S0042-6822(95)80034-4.

13. Leitner T, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, Wolinsky S, Korber B (ed). 2005. HIV sequence compendium 2005. LA-UR 06-0680. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. http://hiv.lanl.gov.

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

15. Karlsson RK, Jennes W, Page-Shaferc K, Nixon D, Shacklett BL. 2004. Poorly soluble peptides can mimic authentic ELISPOT responses. J Immunol Methods 285:89–92. http://dx.doi.org/10.1016/j.jim.2003.11.013.

16. Berger CT, Frahm N, Price DA, Mothe B, Ghebremichael M, Hartman KL, Henry LM, Brenchley JM, Ruff LE, Venturi V, Pereyra F, Sidney J, Sette A, Douek DC, Walker BD, Kaufmann DE, Brander C. 2011. High-functional-avidity cytotoxic T lymphocyte responses to HLA-B-restricted Gag-derived epitopes associated with relative HIV control. J Virol 85:9334–9345. http://dx.doi.org/10.1128/JVI.00460-11.

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

18. 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 S, 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.

19. Frahm N, Korber BT, Adams CM, Szinger JJ, Draenert R, Addo MM, Feeney ME, Yusim K, Sango K, Brown NV, SenGupta D, Piechocka-Trocha A, Simonis T, Marincola FM, Wurcel AG, Stone DR, Russell CJ, Adolf P, Cohen D, Roach T, StJohn A, Khatri A, Davis K, Mullins J, Goulder PJ, Walker BD, Brander C. 2004. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. J Virol 78:2187–2200. http://dx.doi.org/10.1128/JVI.78.5.2187-2200.2004.