

Table S1. Epitopes tested in this study.

<u>HLA^a</u>	<u>Epitope sequence^b</u>	<u>HIV-1 protein</u>	<u>Database frequency^d</u>
A*0101 (A1)	RRGWEVLKY	gp41 (787–795)	0.147
A*0201 (A2)	VIYQYMDDL	RT (179–187)	0.856
A*0201 (A2)	ILKEPVHGV	RT (309–317)	0.773
A*0201 (A2)	ALVEICTEM	RT (33–41)	0.711
A*0201 (A2)	AHIRLQQL	Vpr (59–67)	0.591
A*0201 (A2)	SLYNTVATL	Gag p17 (77–85)	0.318
A*0201 (A2)	VLEWRFD SRL	Nef (180–189)	0.144
A*0201 (A2)	PLTFGW CYKL	Nef (136–145)	0.058
A*0201 (A2)	RGPGR AFVTI	gp160 (311–320)	0.010
A*0201 (A2)	SLLNATDIAV	gp160 (813–822)	0.000
A*0301 (A3)	TVYYGVPVWK	gp160 (37–46)	0.924
A*0301 (A3)	KIRLRPGGK	Gag p17 (18–26)	0.813
A*0301 (A3)	QVPLRPMTYK	Nef (73–82)	0.770
A*0301 (A3)	ALVEICTEMEK	RT (33–43)	0.691
A*0301 (A3)	RLRPGGKKKY	Gag p17 (20–29)	0.646
A*0301 (A3)	RLRPGGKKK	Gag p17 (20–28)	0.646
A*0301 (A3)	AIFQSSMTK	RT (158–166)	0.546
A*0301 (A3)	RLRDLLIVTR	gp160 (770–780)	0.371
A*1101 (A11)	IYQEPFKNLK	RT (341–350)	0.794
A*1101 (A11)	PLRPMTYK	Nef (75–82)	0.782
A*1101 (A11)	ACQGVGGPGHK	Gag p24 (217–227)	0.722
A*1101 (A11)	AIFQSSMTK	RT (158–166)	0.546
A*1101 (A11)	TLYCVHQRI	Gag p17 (84–92)	0.293
A*1101 (A11)	AVDLSHFLK	Nef (84–92)	0.288
A*2301 (A23)	RYLKDQQLL	gp41 (585–593)	0.497
A*2301 (A23)	LRSFLFS	gp41 (760–767)	0.091
A*2402 (A24)	LFCASDAKAY	gp160 (52–61)	0.848
A*2402 (A24)	RYPLTFGW	Nef (134–141)	0.547
A*2402 (A24)	RYLKDQQLL	gp160 (585–593)	0.497
A*2402 (A24)	KYKCLKHIVW	Gag p17 (28–36)	0.268
A*2402 (A24)	RDYVDRFFKTL	Gag p24 (162–172)	0.005
A*2501 (A25)	ETINEEAAEW	Gag p24 (71–80)	0.869
A*2501 (A25)	QAISPRTLNAW	Gag p24 (13–23)	0.540
A*2601 (A26)	EVIPMFSAL	Gag p24 (35–43)	0.828
A*2902 (A29)	SFEPIPIHY	gp120 (209–217)	0.751
A*3002 (A30)	RSLYNTVATLY	Gag p17 (76–86)	0.177
A*3101 (A31)	RLRDLLIVTR	gp160 (770–780)	0.371
A*3201 (A32)	RIKQIINMW	gp160 (419–427)	0.477
A*3201 (A32)	PIQKETWETW	RT (392–401)	0.330
A*33	RYPLTFGW	Nef (134–141)	0.547
B*0702 (B7)	TPQDLNTML	Gag p24 (48–56)	0.975

B*0702 (B7)	SPRTLNAWV	Gag p24 (16–24)	0.924
B*0702 (B7)	RPNNNTRKSI	gp160 (298–307)	0.421
B*0702 (B7)	IPRRIRQGL	gp160 (843–851)	0.340
B*0702 (B7)	TPGPGVRYPL	Nef (128–137)	0.193
B*0702 (B7)	RPMTYKAAL	Nef (77–85)	0.091
B*0702 (B7)	FPVTPQVPLR	Nef (68–77)	0.066
B*0702 (B7)	TPQVPLRPM	Nef (71–79)	0.066
B*0801 (B8)	GPKVKQWPL	RT (18–26)	0.959
B*0801 (B8)	DCKTILKAL	Gag p24 (197–205)	0.874
B*0801 (B8)	EIYKRWII	Gag p24 (128–135)	0.864
B*0801 (B8)	FLKEKGGL	Nef (90–97)	0.778
B*0801 (B8)	QGLERALL	gp160 (849–856)	0.538
B*0801 (B8)	YLKDQQLL	gp160 (586–593)	0.497
B*0801 (B8)	GGKKKYKL	Gag p17 (24–31)	0.278
B*0801 (B8)	ELRSLYNTV	Gag p17 (74–82)	0.197
B*0801 (B8)	WPTVRERM	Nef (13–20)	0.189
B*0801 (B8)	RVKEKYQHL	gp160 (2–10)	0.020
B*1402 (B14)	DRFYKTLRA	Gag p24 (166–174)	0.899
B*1402 (B14)	ERYLKDQQL	gp160 (584–592)	0.513
B*1501 (B62)	LVGKLNWASQIY	RT (260–271)	0.969
B*1501 (B62)	GLNKIVRMV	Gag p24 (137–145)	0.944
B*1501 (B62)	IKLEPVHGVY	RT(309–318)	0.773
B*1501 (B62)	TQGYFPDWQNY	Nef (117–127)	0.757
B*1501 (B62)	RMRRAEPAA	Nef (19–27)	0.251
B*1501 (B62)	AAVDLSHFL	Nef (83–91)	0.206
B*15	SFNCGGEFF	gp120 (375–383)	0.716
B*1503 (B72)	RKAKIIRDY	INT (263–271)	0.701
B*1503 (B72)	WRFDSRLAF	Nef (183–191)	0.272
B*1801 (B18)	FRDYVDIFYK	Gag p24 (161–170)	0.919
B*1801 (B18)	YPLTFGWYCY	Nef (135–143)	0.058
B*2705 (B27)	IRLRPGGKK	Gag p17 (19–27)	0.869
B*2705 (B27)	KRWIILGLNK	Gag p24 (131–140)	0.823
B*2705 (B27)	GRRGWEALKY	gp160 (786–795)	0.386
B*2705 (B27)	RRQDILDWI	Nef (105–114)	0.107
B*3501 (B35)	TVLDVGDAY	RT (107–115)	0.959
B*3501 (B35)	WASRELERF	Gag p17 (36–44)	0.949
B*3501 (B35)	VPLRPMTY	Nef (74–81)	0.831
B*3501 (B35)	VPVWKEATTTL	gp160 (42–52)	0.695
B*3501 (B35)	NPDIVIYQY	RT (175–183)	0.660
B*3501 (B35)	DPNPQEVVL	gp160 (78–86)	0.381
B*3501 (B35)	TAVPWNASW	gp160 (606–614)	0.289
B*3501 (B35)	VPLDEDFRKY	RT (118–127)	0.165
B*3501 (B35)	PPIPVGDIY	Gag p24 (122–130)	0.086
B*3501 (B35)	NSSKVSQNY	Gag p17 (124–132)	0.020
B*3501 (B35)	HPDIVIYQY	RT (175–183)	0.000
B*3801 (B38)	MHEDIISLW	gp120 (101–110)	0.787
B*3901 (B39)	GHQAAMQML	Gag p24 (61–69)	0.944
B*4001 (B60)	SEGATPQDL	Gag p24 (44–52)	0.955
B*4001 (B60)	IEELRQHLL	RT (202–210)	0.701

B*4001 (B60)	KEKGGLEGL	Nef (92–100)	0.679
B*4001 (B60)	QELKNSAVSL	gp160 (805–814)	0.563
B*4001 (B60)	AAVDLSHFL	Nef (83–91)	0.206
B*4001 (B60)	IEIKDTKEAL	Gag p17 (92–101)	0.167
B*4002 (B61)	CTERQANFL	Gag p15 (63–71)	0.601
B*4201 (B42)	TPQDLNTML	Gag p24 (48–56)	0.975
B*4201 (B42)	TPGPGVRYPL	Nef (128–137)	0.193
B*4201 (B42)	YPGIKVRQL	RT (271–279)	0.124
B*4402 (B44)	RDYVDRFYKTL	Gag p24 (162–172)	0.889
B*4402 (B44)	AEQASQDVKNW	Gag p24 (174–184)	0.278
B*4402 (B44)	AENLWVTVY	gp160 (31–40)	0.061
B*5101 (B51)	EKEGKISKI	RT (42–50)	0.866
B*5101 (B51)	TAFTIPSI	RT (128–135)	0.722
B*5101 (B51)	EAVRHFPRI	Vpr (29–37)	0.333
B*5101 (B51)	RAIEAQQML	gp160 (557–565)	0.000
B*5701 (B57)	KAFSPEVI	p24 (30–37)	0.970
B*5701 (B57)	KAFSPEVIPMF	Gag p24 (30–40)	0.970
B*5701 (B57)	KTAVQMAVF	INT (173–181)	0.907
B*5701 (B57)	YFPDWQNYT	Nef (120–128)	0.778
B*5701 (B57)	TSTLQEIQGW	Gag p24 (108–117)	0.727
B*5701 (B57)	ISPRTLNAW	Gag p24 (15–23)	0.697
B*5701 (B57)	HTQGYFPDWQ	Nef (116–125)	0.683
B*5701 (B57)	IVLPEKDSW	RT (244–252)	0.619
B*5701 (B57)	QASQEVKNW	Gag p24 (176–184)	0.520
B*5801 (B58)	TSTLQEIQGW	Gag p24 (108–117)	0.727
B*5801 (B58)	TSTVEEQIQW	Gag p24 (108–117)	0.000
Cw5	SAEPVPLQL	Rev (61–75)	0.206
Cw7	DILDLWIY	Nef (106–115)	0.107
Cw8	AAVDLSHFL	Nef (83–91)	0.206
Cw*1203 (Cw12)	CCFHCQVC	Tat (30–37)	0.776

^a HLA restriction. For cases in which the genotype is defined, the serotype is shown in parentheses.

^b 120 defined optimal epitopes [27, 28] were analyzed in 45 subjects in this study. Conserved epitopes were arbitrarily defined as belonging to the upper quartile of the tested epitopes in terms of database frequency (conserved in $\geq 78.68\%$ of the sequences), while epitopes in the lower three quartiles were considered to be variable epitopes.

^c HXB2 amino acid numerations are used.

^d Frequency of the epitopic sequence in HIV-1 subtype B in the Los Alamos HIV Sequence Database in year 2005 [29].

Table S2. Correlation of CTL responses with viral load and CD4⁺ T cell counts ^a.

Infection stage	VL/CD4 ⁺ T cell counts	CTL responses	Epitopes (location/type)	Spearman r	<i>p</i>		
Primary infection	log ₁₀ VL ^b	Number of recognized epitopes	All epitopes	0.1320	0.4568		
			Gag	-0.3390	0.0499		
			Pol	-0.0589	0.7407		
			Aux	0.0725	0.6835		
			Env	0.3786	0.0273		
			Nef	0.2413	0.1693		
			Variable	0.1715	0.3322		
			Conserved	-0.0025	0.9887		
		Fraction of recognized epitopes	All epitopes	0.0608	0.7329		
			Variable	0.0496	0.7804		
			Conserved	-0.0335	0.8506		
		Total magnitude of CTL responses	All epitopes	0.3092	0.0752		
			Gag	-0.2151	0.2218		
			Pol	-0.0940	0.5968		
			Aux	0.1539	0.3850		
			Env	0.4380	0.0096		
			Nef	0.1424	0.4219		
			Variable	0.4010	0.0188		
			Conserved	0.0446	0.8023		
		Mean magnitude of CTL responses	All epitopes				
			Variable	0.3412	0.0483		
			Conserved	0.0495	0.7811		
		CD4 ⁺ T cell counts		Number of recognized epitopes	All epitopes	0.0385	0.8291
					Gag	0.0081	0.9636
Pol	0.0160				0.9283		
Aux	0.1280				0.4705		
Env	0.0203				0.9092		
Nef	-0.0229				0.8978		
Variable	0.1032				0.5614		
Conserved	-0.1288				0.4678		
Fraction of recognized epitopes	All epitopes			0.0878	0.6214		
	Variable			0.1433	0.4190		
	Conserved			-0.1253	0.4801		
Total magnitude of CTL responses	All epitopes			-0.1285	0.4688		
	Gag			0.0075	0.9666		
	Pol			0.0361	0.8394		
	Aux			0.0843	0.6357		
	Env			-0.0788	0.6579		
	Nef			-0.1238	0.4854		
	Variable			-0.1775	0.3153		
	Conserved			-0.1272	0.4734		
Mean magnitude of CTL responses	All epitopes			-0.1661	0.3478		
	Variable			-0.3022	0.0824		
	Conserved			-0.1449	0.4137		

Chronic infection ^b	log ₁₀ VL	Number of recognized epitopes	All epitopes	-0.1266	0.6403	
			Gag	-0.2967	0.2645	
			Pol	0.1736	0.5203	
			Aux	Analysis 1 ^c	-0.3027	0.2544
				Analysis 2 ^c	-0.4976	0.0498
			Env	0.1111	0.6820	
			Nef	0.1386	0.6087	
			Variable	0.0030	0.9912	
			Conserved	-0.2585	0.3338	
		Fraction of recognized epitopes	All epitopes	-0.3976	0.1272	
			Variable	-0.2024	0.4523	
			Conserved	-0.4019	0.1228	
		Total magnitude of CTL responses	All epitopes	-0.0324	0.9035	
			Gag	Analysis 1 ^c	-0.3587	0.1725
				Analysis 2 ^c	-0.5130	0.0421
			Pol	0.2108	0.4332	
			Aux	-0.2579	0.3349	
			Env	0.1057	0.6969	
			Nef	0.1703	0.5283	
			Variable	0.0133	0.9612	
			Conserved	-0.0983	0.7174	
		Mean magnitude of CTL responses	All epitopes	-0.1015	0.7083	
			Variable	0.0721	0.7907	
			Conserved	-0.0747	0.7834	
CD4 ⁺ T cell counts	Number of recognized epitopes	All epitopes	-0.0734	0.7871		
		Gag	0.3647	0.1649		
		Pol	-0.3006	0.2579		
		Aux	0.1396	0.6061		
		Env	-0.5398	0.0309		
		Nef	-0.2093	0.4366		
		Variable	-0.1576	0.56		
		Conserved	0.2475	0.3553		
		Fraction of recognized epitopes	All epitopes	0.1722	0.5237	
	Variable		0.0207	0.9395		
	Conserved		0.3485	0.1859		
	Total magnitude of CTL responses	All epitopes	-0.0765	0.7783		
		Gag	0.2242	0.4039		
		Pol	-0.2755	0.3017		
		Aux	0.1436	0.5956		
		Env	-0.4884	0.0594		
		Nef	-0.1134	0.6757		
		Variable	-0.0765	0.7783		
		Conserved	0.0346	0.8989		
		Mean magnitude of CTL responses	All epitopes	0.1235	0.6485	
	Variable		0.0265	0.9225		
	Conserved		0.0063	0.9816		

^a Correlations were determined by Spearman's rank coefficient, and significant correlations (p value >0.05) are highlighted gray.

^b VL, viral load.

^c Chronic CTL responses were measured in 16 subjects: at a single time point in 7 subjects and at two time points in 9 subjects. Two parallel analyses were conducted. Analysis 1 used data from the former seven subjects and the data of the first time points from the latter nine subjects, and Analysis 2 used data from the former seven subjects and the data of the second time points from the latter nine subjects. Results from both analyses were generally similar. Hence, only the results from Analysis 1 are reported, unless there was disagreement between the two analyses.