Supporting Information

Catalfamo *et al.* 10.1073/pnas.0810032105

SI Text

Datasets for Analysis. A series of analyses were carried out on the study cohort identified above (Table 1 summarizes the different subsets of the cohort that were involved in the different analyses). 341 patients and 373 controls underwent a total of 5240 and 1453 evaluations respectively. Patients coinfected with hepatitis B or C, or on IL-2 or IFN-α treatment were eliminated from the analyses yielding a final number of 284 patients, contributing to a total of 3725 observations. For our initial analyses a subset of samples was created. Each of the 3725 samples from the 284 patients was categorized into 1 of 5 categories (LowCD4-LowVL, HighCD4-LowVL, LowCD4-HighVL, HighCD4-HighVL and other). Samples in the other category were eliminated. Patients whose samples fell into multiple categories were assigned to the category with the fewest patients. This left 161 unique patients with each patient contribution a variable number of samples to a single category. The 373 samples from the healthy donors were divided based on CD4 T cell count into 3 categories (LowCD4: $<$ 500 cell/μl, HighCD4: $>$800 cell/μl and other). Samples in the other category were eliminated. This created a dataset of 276 unique subjects with each subject contributing a variable number of samples to each of the 2 categories. For the cytokine analysis, the same algorithm as above was applied to all patients for whom there were a minimum of 3 serum samples at a given time point. For each patient, a single sample was randomly extracted from a single time point for which at least 3 serum samples were available and the serum used for the cytokine assay.

Statistical Methods. For our initial analyses we examined the effects of $CD4^+$ T cells and viral load on the proliferation of $CD4^+$ and $CD8^+$ T cells (CD4B and CD8B respectively) in the 4 subsets of patients defined by CD4 counts $>$800 or $<$300 cells per microliter and HIV RNA levels $>$10,000 or $<$50 copies per milliliter and the 2 subsets of normal volunteers defined by CD4 counts $>$800 or $<$500 cells per microliter. In the datasets used for these initial analyses (Table S1 - D1.S and D2.S) the median and the variance of the median of these samples in each group for CD4B and CD8B were calculated from 1000 within cluster resamples (1, 2). The within cluster resampling procedure was also used for computing the *p*-values for the comparisons among the different cohorts, using Wilcoxon tests (2).

To explore the association of CD4 and CD8 proliferation with CD4 counts and viral load, a multivariate analysis was performed with mixed-effects linear models to handle the multiple samples from all patients. In these models, log transformed CD4 and CD8 proliferation, and the log transformed viral load were used. To assess the predictivity of a linear model, an R-squared averaged over within cluster resamples was reported (1, 2). An ANCOVA (analysis of covariance) type analysis was used to compare CD4 and CD8 T cell proliferation between the HIV infected group with suppressed viral load (sub group of patients with VL $<$50 copies per milliliter) and normal volunteers, adjusting for CD4 cell count. In the subset of patients selected for the cytokine analysis, an ANCOVA analysis was used to evaluate the effects of viral load level (high level versus low level, i.e., treating the viral load as a binary variable) on IL-7 adjusting for CD4 cell count. In addition, a multivariate analysis was used to study the association of CD4 and CD8 T cell proliferation with IL-7, CD4 T cell count and log-transformed viral load, and the association of IL-7, IFN-γ and TNF-α with CD4 T cell count and viral load. The significance of paired differences between CD4 and CD8 proliferation was determined by Wilcoxon signed rank test, and the differences between 2 independent groups by the Wilcoxon rank sum test.

1. Hoffman EB, S.P.a.W.C. (2001) Within Cluster Resampling Biometrika 88(40 1121–1134. *Biometrika* 88:1121–1134.

2. Follmann D, Proschan M, Leifer E. (2003) Multiple outputation: Inference for complex clustered data by averaging analyses from independent data. *Biometrics* 59:420–429.

Table S1. Description of the datasets used for each table and figure

Data set	Subjects, no.	Samples, no.	Characteristics	Key Variables	Analysis
D1	284	3,725	HIV+ Multiple samples per subject	CD4B, CD8B CD4 T cell count, Viral load	Table S2 Fig. 2*A* Fig. 2*B*
D1.s	161	997	HIV+ Subset of subjects from D1 dataset assigned to each of the 4 categories (CD4/VL)	CD4B, CD8B, CD4 T cell count, Viral load	Fig. 1*A*
D2	373	1,453	HIV− Multiple samples per subject	CD4B, CD8B CD4 T cell count	Fig. 2*A*
D2.s	276	912	HIV− Subset of subjects from D2 dataset assigned to each of the 2 categories (CD4)	CD4B, CD8B CD4 T cell count	Fig. 1*B*
D3	102	102	HIV+ Random extraction of one sample per subject from D1 dataset samples containing ≥ 3 serum vials (for cytokine assay)	CD4B, CD8B CD4 T cell count, Viral load, Gamma cytokines	Fig. 3*A* Fig. 3*B*
D4	28	28	HIV+ Samples from D3 with VL > 50	CD4B, CD8B CD4 T cell count, Viral load, Gamma cytokines	Table S3 Table S4
D5	34	34	HIV+ One sample per subject	p-stat5, CD4 T cell count, Viral load	Fig. 4
D6	18	18	HIV− One sample per subject	p-stat5, CD4 T cell count,	Fig. 4

D1, 1 subject eliminated; D2, 3 subjects eliminated due to inability to measure levels of proliferation (below the limits of the assay).

Table S2. Multivariate analysis of the relationship between cytokines IL-7, IFN-γ and TNF-α and the covariates CD4 count and viral load over the 28 subjects with VL > 50 copies/ml, $n = 28$

			Regresion coefficient (95% CL)		
Cytokine, pg/ml	Analysis	R^2 (P)	CD4, (cells per microliter)$^{-1}$	Log VL, (copies per milliliter)$^{-1}$	P, univariate vs bivariate
IL-7	Bivariate	0.489 (<0.001)	−0.0133 (−0.020, −0.006)	4.2995 (−0.229, 8.828)	
	Univariate	0.4182 (<0.001)	−0.0156 (−0.023, −0.009)		0.075
	Univariate	0.2188 (0.012)		7.1203 (1.948, 12.292)	0.001
IFN-γ	Bivariate	0.2728 (0.019)	−0.001 (−0.002, 0.0005)	1.0116 (0.087, 1.936)	
	Univariate	0.139 (0.051)	−0.0015 (−0.003, −0.0001)		0.042
	Univariate	0.2206 (0.012)		1.2237 (0.340, 2.108)	0.193
TNF-α	Bivariate	0.0483 ($P = 0.539$)	0.0117 (−0.022, 0.045)	−6.5321 (−27.604, 14.540)	
	Univariate	0.0342 ($P = 0.346$)	0.0152 (−0.016, 0.046)		0.549
	Univariate	0.0302 ($P = 0.376$)		−9.0192 (−28.667, 10.628)	0.497

Table S3. Multivariate analysis of the relationship between CD4 and CD8 T cell proliferation and the covariates IL-7, CD4 counts and viral load in 28 patients with VL >50 copies per milliliter

	Regression R^2			
Output	IL-7 (pg/mL)	CD4 (cells per microliter)$^{-1}$	log VL (copies per milliliter)$^{-1}$	*P* (Univariate vs. multivariate)
CD4 B				
Multivariate		0.378		
Univariate	0.125			0.017
Univariate		0.327		0.389
Univariate			0.147	0.022
CD8 B				
Multivariate		0.292		
Univariate	0.102			0.058
Univariate		0.016		0.020
Univariate			0.269	0.688