**Supplementary Material for**

# Early immunologic and virologic predictors of clinical HIV-1 disease progression

Yolanda D. Mahnke, Kaimei Song, Mariana M. Sauer, Martha C. Nason, Maria Teresa M. Giret, Karina I. Carvalho, Priscilla R. Costa, Mario Roederer, Esper G. Kallás

## Supplementary Materials and Methods

### Patients

Patients were recruited from individuals seeking free, anonymous testing services for sexually transmitted diseases in São Paulo, Brazil. Those that tested positive for HIV-1 were further tested using the Serologic Testing Algorithm for HIV Recent Seroconversion (STAHRS) and enrolled if they tested negative for the desensitized ELISA, which identifies early HIV infection [1]. HIV-1 Clade B infection was established by full genome sequencing [2, 3]. After a 2-year follow-up, patients were divided into three cohorts according to their clinical status (Table 1): Fast Progressors required to receive antiretrovirals (ARVs) within the first 2 years, while Slow Progressors had similar $CD4^+$ and $CD8^+$ T-cell counts and PVL at entry but no requirement for treatment with ARVs, and Controllers maintained PVL of <2000 copies/ml (i.e. $\log_{10}$ 3.3). Blood samples were collected and PBMC frozen within 1 year of diagnosis. ARVs were initiated when $CD4^+$ T-cell counts were repeatedly <300 cells/µl, following Brazilian HIV treatment guidelines [4]. HIV-1-infected PBMC samples from a single leukapheresis were kindly provided by J Casazza (Immunology Laboratory, NIAID, NIH) and included in every experiment as an internal control. Healthy Donor PBMC and plasma samples were collected in São Paulo, Brazil. Written informed consent was obtained from all donors and the study approved by the institutional review boards of participating sites. This study was approved and annually reviewed by the Institutional Review Board of the Federal University of São Paulo (#362/00), and by the Institutional Review Board of the University of São Paulo (#0458/08), ensuring the compliance with all guidelines for human experimentation set down by the two universities, as well as the Department of Human and Health Services.

Data on gender, age, ethnicity, mode of transmission, and presence of symptoms were noted. $CD4^+$ and $CD8^+$ T-cell counts, as well as plasma HIV-1 RNA copies/ml were recorded at all visits. $CD4^+$ and $CD8^+$ T-cell counts were determined by flow cytometry using fluorescenctly conjugated monoclonal antibodies specific for CD3, CD4 and CD8 (Kit TriTest, BD Biosciences, San Diego, California, USA). The plasma RNA measurements were performed using the Amplicor HIV-1 Monitor test, version 1.5 (Roche Diagnostics, Indianapolis, IN, USA), until January 2007, and was subsequently replaced by the Versant – bDNA (branched DNA) HIV-1 RNA 3.0 Assay (Bayer Health Care LLC Tarrytown, NY).

### CCR5 polymorphism

Genomic DNA samples were extracted from 300 µl of buffy coat using a QIAamp Blood Kit (QIAGEN Inc, CA), according to the manufacturer's instructions. The presence of the CCR5Δ32 allele was determined by polymerase chain reaction (PCR) and subsequent gel electrophoresis as previously described [5]. Briefly, the 20 µl reaction mix contained 0.375 mM of both forward (5'–TCAAAAAGAAGGTCTTCACACC–3') and reverse (5'–AGCCCAGAAGAGAAAATAAACAATC–3') primers, 20 ng genomic DNA, 1 unit of AmpliTaq Gold (Roche Molecular Systems), and 23.4 mM

dNTPs (Roche) in a 1.5 mM $MgCl_2$ buffer. After incubating the mix for 10 minutes at 94ºC, 40 cycles were performed with 45 seconds of melting at 94ºC, annealing for 45 seconds at 58ºC, and extending for 45 seconds at 72ºC. Subsequently, amplified products were separated by electrophoresis in a 3% agarose gel for 40 minutes at 110mV and visualized with ethidium bromide under ultraviolet light. The expected PCR product size was 241bp and 209 bp for the wildtype and CCR5Δ32 alleles, respectively.

**Detection of GB virus type C RNA**

The procedure for the detection of GBV-C RNA was performed as previously described [6, 7]. Viral RNA was extracted from 140 µl plasma samples using the QIAamp Viral RNA Mini Kit (QIAGEN Inc, CA) according to the manufacturer's instructions. 5 µl of the extracted RNA was diluted in a mix containing 150 ng of Random Primer (Pharmacia Biotech, Sweden) and 10 mmol/l dNTPs (Invitrogen Inc., Carlsbad, California, USA); the solution was kept at 65°C for 5 minutes. Following the addition of 200 U SuperScript III Reverse Transcriptase (Invitrogen Inc, CA), complementary DNA (cDNA) was synthesized at 25°C for 5 minutes, 50°C for 60 minutes and 70°C for 15 minutes in a buffer solution containing 10 U ribonuclease inhibitor (Invitrogen Inc, CA) at a final volume of 20 µl.

A 344 bp fragment of the 5' non-coding region was amplified by nested RT-PCR using the following primers: 5'–AGGTGGTGGATGGGTGAT–3' outer sense primer; 5'–TGGTAGGTCGTAAATCCCGGT–3' inner sense primer; 5'–GGAGCTGGGTGGCCCCATGCAT–3' inner antisense primer; 5–TGCCACCCGCCCTCACCCGAA–3' outer antisense primer [8, 9]. 40 amplification cycles were performed for both the first and second PCR rounds as follows: 30 seconds each at 94°C, 50°C, and 72°C for the first round, and 30 seconds each at 94°C, 60°C, and 72°C for the second round. After amplification, 5 µl of the PCR product was analyzed by electrophoresis on a 2% agarose gel.

**Sample preparation and Ag-stimulation**

Cryopreserved PBMC samples were thawed in pre-warmed RPMI 1640, 10% FCS, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin (all from Gibco; this medium will hereafter be referred to as RPMI complete), in the presence of 20 µg/ml benzonase nuclease (Novagen). A minimum of $1x10^6$ live cells were immediately analyzed for T-cell subsets and activation markers by polychromatic flow cytometry. For functional assays, cells were rested overnight in RPMI complete at 37ºC, 5% $CO_2$. Cells were harvested the next morning, washed in pre-warmed RPMI complete and 1-$2x10^6$ PBMC stimulated for 6 hrs in 200 µl with 2.5 µg/ml peptide pools for either Gag, Env, Nef, Pol, or Tat, Rev, Vif, and Vpr (TRVV), or left unstimulated (costimulation only control). Stimulation cultures contained monensin, Brefeldin A, and mAb to CD49d and $CD28^{PE\text{-}Cy5}$ (all from BD Biosciences). Healthy Donor PBMC were also stimulated with SEB (Sigma) to serve as a positive control.

**Flow Cytometry**

The following reagents were used for ex vivo: $CCR5^{PE\text{-}Cy7}$ (clone 2D7/CCR5), $CD25^{PE\text{-}Cy5}$ (clone M-A251), $Ki67^{FITC}$ (clone B56) (BD Pharmingen), $CCR7^{Ax680}$ (clone 150503), $CD3^{QD800}$ (clone OKT3), $CD4^{QD655}$ (clone M-T477), $CD8^{QD855}$ (clone RPA-T8), $CD14^{Pacific\ Blue}$ (clone M5E2), $CD19^{Pacific\ Blue}$ (clone HIB19), $CD27^{APC\text{-}Ax700}$ (clone 1A4LDG), $CD38^{PE}$ (clone OKT10), $CD45RO^{QD545}$ (clone UCHL1), $CD57^{QD705}$ (clone NK-1) (conjugated in-house), and Granzyme $B^{APC}$ (clone GB12) (Caltag).

The panel used for the detection of cytokine production after in vitro stimulation has been published [10], and used the following Abs: $CD3^{APC\text{-}Cy7}$ (clone SK7), $CD27^{PE\text{-}Cy7}$ (clone M-T271), $CD28^{PE\text{-}Cy5}$ (clone CD28.2), IFN-$\gamma^{APC}$ (clone B27), (BD Pharmingen), $CCR7^{Ax594}$ (clone 150503), $CD8^{QD855}$ (clone RPA-T8), $CD14^{Pacific\ Blue}$ (clone M5E2), $CD19^{Pacific\ Blue}$ (clone HIB19), $CD45RO^{QD545}$ (clone UCHL1), $CD57^{QD705}$ (clone NK-1), $TNF^{Ax594}$ (clone MAb11) (conjugated in-house), $CD4^{QD605}$

(clone M-T477) (Molecular Probes, Invitrogen), $CD127^{PE}$ (clone R34.34) (Immunotech Coulter), and $IL\text{-}2^{Ax488}$ (clone MQ1-17H12) (ReaMetrix). $PD\text{-}1^{biot}$ (clone MIH4) was revealed using $SA^{QD655}$ (Invitrogen). In both panels, dead cells were detected with the Violet Amine Reactive Viability Dye (Molecular Probes, Invitrogen). For intracellular staining, cells were treated with BD Cytofix/Cytoperm Permeabilization Solutions (BD Biosciences). Data were acquired on a FACSAria II cell sorter or an LSR II (both from BD Biosciences) and analyzed using FlowJo Version 9.0.1 (Tree Star), Pestle (from M.R.), and Spice Version 5.1.

**Cell-associated viral load (CAVL)**

Four sub-populations of $CD4^+$ T-cells were sorted on a FACSAria II cell sorter (BD Biosciences): Naïve ($CCR7^+$ $CD45RO^-$), non-naïve $CCR5^+$ $CD45RO^+$, non-naïve $CCR5^-$ $CCR7^+$ $CD45RO^+$, and non-naïve $CCR5^-$ $CCR7^-$. The cells were pelleted by centrifugation and, after aspiration of the supernatant, cryopreserved at -20ºC. Quantitative PCR (qPCR) was performed as described elsewhere [11] using previously described primers and probes [12]. Briefly, thawed pellets were lysed in Proteinase K buffer, and duplicates tested for *albumin* and HIV *gag* DNA on a 7500 Real Time PCR System (Applied Biosystems). *Albumin* and HIV *gag* DNA were quantified using standard curves obtained from PCR reactions with serial dilutions of plasmid DNA and corresponding duplicates averaged. Samples were excluded when ≤100 cells were measured by PCR. The adjusted number of HIV *gag* copies per cell was calculated by adding one-half of the lowest measured HIV *gag* value to each averaged HIV *gag* value before dividing by half the corresponding *albumin* copies measured (since there are 2 *albumin* copies per cell) plus 1 [13]. This provided an estimate of the HIV gag copies per cell even for those samples were none were measureable due to low input cell numbers and/or very low cell-associated viral load. Total CAVL was calculated by multiplying the viral burden in each subset by the frequency of that subset and adding the results for the four $CD4^+$ T-cell subsets per patient.

**Sample analysis and modeling**

The median fluorescence intensity (MFI) of $CD38^+$ cells was calibrated using the corresponding measurement in the experiment-matched internal control sample.

Only those samples with >10 $cytokine^+$ cells were included in the phenotypic analyses of cytokine-producing cells. Responses measured following stimulation with Env, Gag, Nef, Pol, and TRVV peptide pools were summed for the analysis of total HIV-1-specific response magnitude and cytokine pattern, but averaged for phenotypic investigation of $cytokine^+$ cells.

**Statistical analysis**

CAVL data were analyzed in Prism (GraphPad) using the Wilcoxon signed rank test for matched samples, and Mann-Whitney test for comparisons between groups. For flow cytometry data, groups were compared using the Wilcoxon-Rank test. Data in pie charts were evaluated using a permutation test based on 10,000 permutations in Spice Version 5.1 [14].

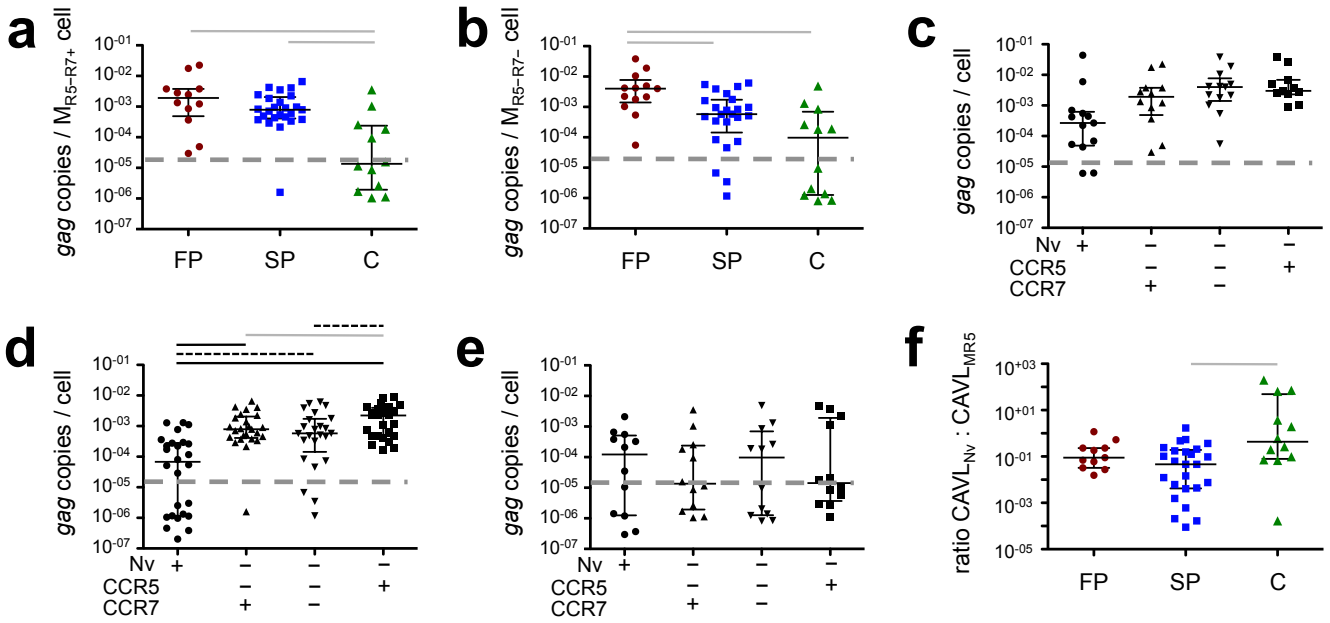
Classification tree models were built using a recursive partitioning algorithm, implemented in the Rpart function in the statistical software R (version 2.14.1), with a minimum node size of 15 and a complexity parameter of 0.01. Any missing values in the predictors were handled by the software via the surrogate splits method.

In order to assess the predictive accuracy of our classification tree, we performed leave-one-out cross-validation (LOOCV). To estimate prediction accuracy using this method, one observation at a time is excluded from the data, and the model refit, using all the original steps. Once the model was refit, a prediction was created for the excluded observation, based on this modified model. The process was then repeated for all the observations in the dataset, yielding a vector of predictions approximating

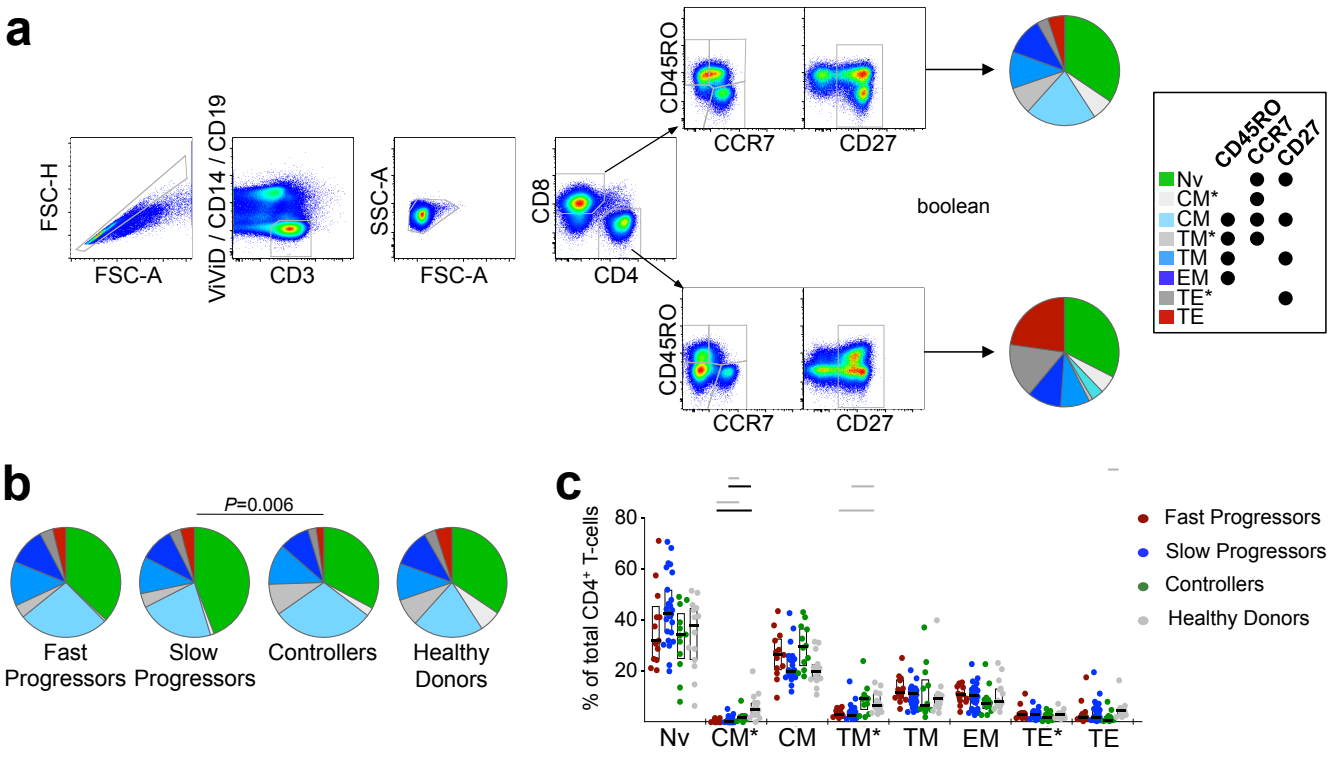
the predictions that would be obtained from testing the model on an independent data set. Finally, a permutation test was used to calculate a p-value for the predictive power of the classification tree, based on the LOOCV predictions.

**Supplementary References**

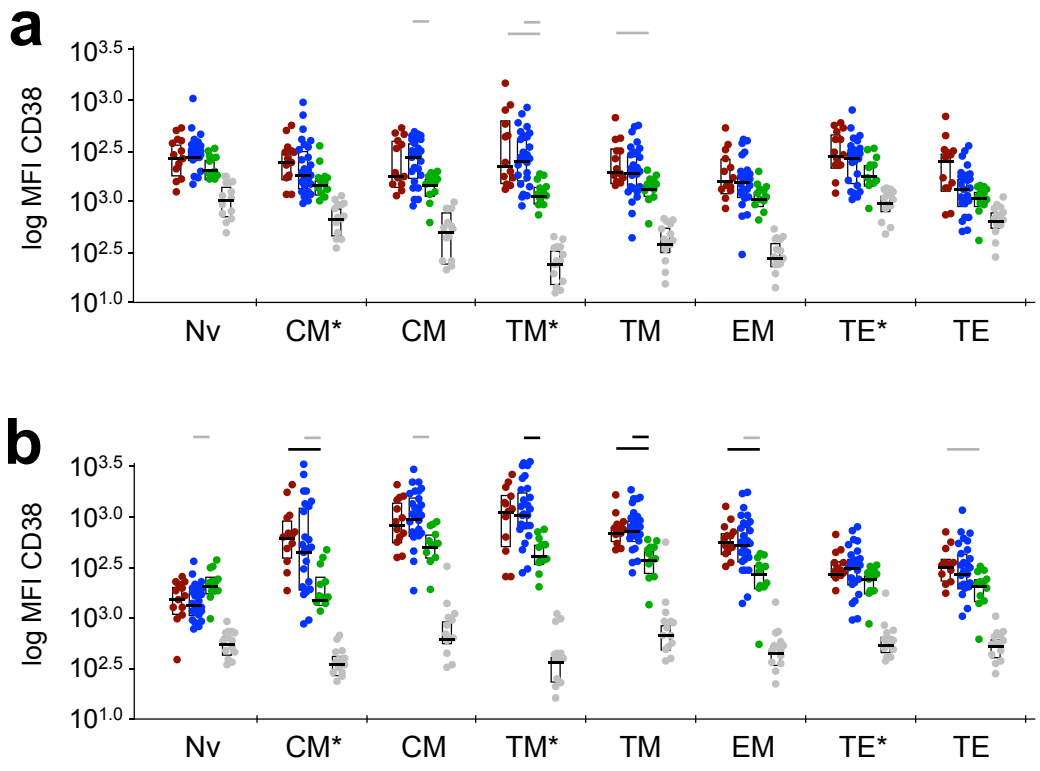

1. Kallas EG, Bassichetto KC, Oliveira SM, Goldenberg I, Bortoloto R, Moreno DM*, et al.* Establishment of the serologic testing algorithm for recent human immunodeficiency virus (HIV) seroconversion (STARHS) strategy in the city of Sao Paulo, Brazil. *Braz J Infect Dis* 2004,**8**:399-406.
2. Sa-Filho D, Kallas EG, Sanabani S, Sabino E, Sucupira MC, Sanchez-Rosa AC*, et al.* Characterization of the full-length human immunodeficiency virus-1 genome from recently infected subjects in Brazil. *AIDS Res Hum Retroviruses* 2007,**23**:1087-1094.
3. Sanabani SS, Pastena ER, da Costa AC, Martinez VP, Kleine-Neto W, de Oliveira AC*, et al.* Characterization of partial and near full-length genomes of HIV-1 strains sampled from recently infected individuals in Sao Paulo, Brazil. *PLoS One* 2011,**6**:e25869.
4. http://www.aids.gov.br/sites/default/files/consensoAdulto005c_2008montado.pdf. Accessed: 14 July 2012.
5. Munerato P, Azevedo ML, Sucupira MC, Pardini R, Pinto GH, Catroxo M*, et al.* Frequency of polymorphisms of genes coding for HIV-1 co-receptors CCR5 and CCR2 in a Brazilian population. *Braz J Infect Dis* 2003,**7**:236-240.
6. Giret MT, Miraglia JL, Sucupira MC, Nishiya A, Levi JE, Diaz RS*, et al.* Prevalence, incidence density, and genotype distribution of GB virus C infection in a cohort of recently HIV-1-infected subjects in Sao Paulo, Brazil. *PLoS One* 2011,**6**:e18407.
7. Maidana-Giret MT, Silva TM, Sauer MM, Tomiyama H, Levi JE, Bassichetto KC*, et al.* GB virus type C infection modulates T-cell activation independently of HIV-1 viral load. *AIDS* 2009,**23**:2277-2287.
8. Jarvis LM, Davidson F, Hanley JP, Yap PL, Ludlam CA, Simmonds P. Infection with hepatitis G virus among recipients of plasma products. *Lancet* 1996,**348**:1352-1355.
9. Tucker TJ, Smuts H, Eickhaus P, Robson SC, Kirsch RE. Molecular characterization of the 5' non-coding region of South African GBV-C/HGV isolates: major deletion and evidence for a fourth genotype. *J Med Virol* 1999,**59**:52-59.
10. Mahnke YD, Roederer M. OMIP-001: Quality and phenotype of Ag-responsive human T-cells. *Cytometry A* 2010,**77**:819-820.
11. Mattapallil JJ, Douek DC, Hill B, Nishimura Y, Martin M, Roederer M. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* 2005,**434**:1093-1097.
12. Douek DC, Brenchley JM, Betts MR, Ambrozak DR, Hill BJ, Okamoto Y*, et al.* HIV preferentially infects HIV-specific CD4+ T cells. *Nature* 2002,**417**:95-98.
13. Ganesan A, Chattopadhyay PK, Brodie TM, Qin J, Gu W, Mascola JR*, et al.* Immunologic and virologic events in early HIV infection predict subsequent rate of progression. *J Infect Dis* 2010,**201**:272-284.
14. Roederer M, Nozzi JL, Nason MC. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytometry A* 2011,**79**:167-174.



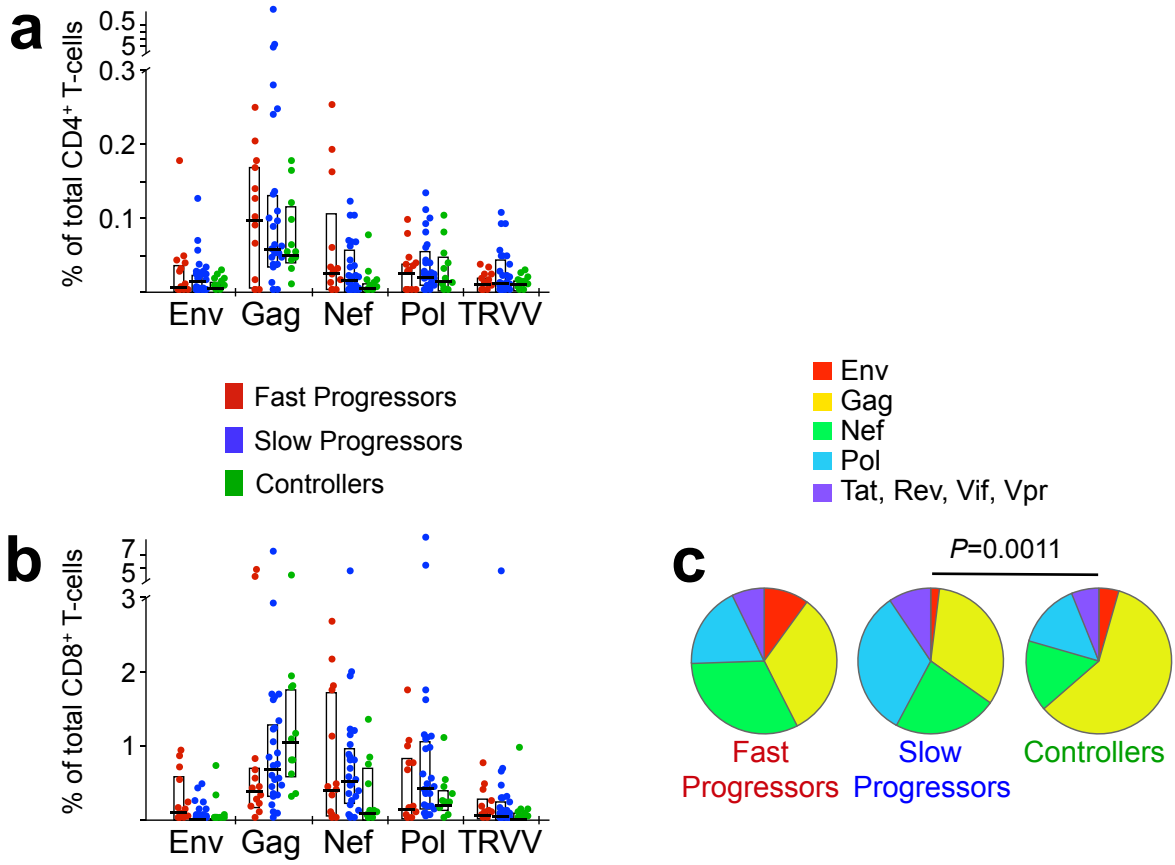

**Supplementary Fig. 1. Cell-associated viral load (CAVL).** CAVL in $M_{R5-R7+}$ (**a**) and $M_{R5-R7-}$ $CD4^+$ T-cells (**b**) are shown for Fast Progressors (FP), Slow Progressors (SP) and Controllers (C). CAVL is illustrated for all four sorted subpopulations in Fast Progressors (**c**), Slow Progressors (**d**) and Controllers (**e**). Bars represent medians and standard error means. Dotted lines indicate the threshold of detection for this assay; measurements that fall below this line had no detectable HIV *gag*, and the distribution of these values provide a conservative estimate of the CAVL in these samples. Statistically significant differences are indicated by lines above the graphs: solid grey, $P<0.01$; dotted black, $P<0.001$; solid black, $P<0.0001$.



**Supplementary Fig. 2. CD4$^+$ T-cell subpopulations.** Gating scheme for determination of T-cell differentiation subsets (**a**). Pies (**b**) and bars (**c**) illustrate the differentiation of CD4$^+$ T-cells in the 3 HIV-1 infected study groups and healthy donors, as defined by expression patterns of CD45RO, CCR7 and CD27. Nv, naive; CM, central memory; TM, transitional memory; EM, effector memory; TE, terminal effector; CM*, TM*, TE* are phenotypically-defined populations that are not described in the literature, but that arise by this gating scheme; their activation phenotype and cytokine potential most closely resembles that of $T_{CM}$, $T_{TM}$ and $T_{TE}$, respectively, hence their nomenclature. Bars indicate medians and boxes interquartile ranges. *P*-values below 0.01 are reported for pies. For the bar chart, statistically significant differences are indicated by color-coded bars: grey, $P<0.01$; black, $P<0.001$.




**Supplementary Fig. 3. MFI of CD38 on T-cell subsets.** The MFI of CD38 was determined for the eight differentiation subsets described in Supplementary Fig. 2a on CD4$^+$ (**a**) and CD8$^+$ T-cells (**b**). No statistically significant differences were observed between Slow and Fast Progressors, while all subsets of all HIV-1$^+$ groups were statistically different from corresponding measurements in healthy donors ($P<0.001$; except CD4$^+$ $T_{TE}$ in Controllers: $P=0.002$). Statistically significant differences between Progressors and Controllers are indicated by lines above the graphs: grey, $P<0.01$; solid black, $P<0.001$.




**Supplementary Fig. 4. Response magnitudes to individual HIV-1 derived peptide pools.** Response magnitudes for individual HIV-1 derived peptide pools generated by $CD4^+$ (**a**) and $CD8^+$ T-cells (**b**). No statistically significant differences were observed in (**a**) and (**b**). Relative proportion of HIV-1-reactive $CD8^+$ T-cells responding to either of the peptide pools (**c**). The statistically significant difference between pie charts of Slow Progressors and Controllers is due to the proportional response to Gag ($P$=0.004).

**Supplementary Table 1. Power of correlates identified in previous studies to predict outcome in the present patient cohort.**

| | **Progressors *vs.* Controllers [p-value]** [a] | **Fast *vs.* Slow Progressors [p-value]** [a] |
|---|---|---|
| *CD4 count* | 0.163 | 0.191 |
| *PVL* | **<0.001** | 0.069 |
| *total CAVL* | **0.007** | **0.001** |
| *%CD38$^+$ CD8$^+$* | **<0.001** | 0.775 |
| *%CD38$^+$ CD4$^+$* | **<0.001** | 0.479 |
| *CD38 MFI CD8$^+$* | **<0.001** | 1.0 |
| *CD38 MFI CD4$^+$* | **0.001** | 1.0 |
| *%Ki67$^+$ CD8$^+$* | **<0.001** | 0.168 |
| *%Ki67$^+$ CD4$^+$* | **<0.001** | 1.0 |
| *%CD127$^+$ CD8$^+$* | 0.989 | 0.761 |
| *%$T_{CM}$ CD4$^+$* | 0.068 | 0.121 |
| *%$T_{NV}$ CD8$^+$* | 0.230 | 0.455 |

[a] evaluated using a Wilcoxon signed-rank test