## Flow cytometry

Samples of fresh, uncoagulated whole blood (100 μL) were incubated with the antibody mix for 20 minutes at room temperature in the dark. Next, Immunoprep Reagent System (Beckman Coulter, Galway, Ireland) was added to fix and lyse each sample. The fluorescence was evaluated using a Gallios™ Flow Cytometer (Beckman Coulter, Miami, FL), establishing a cut-off of 200,000 cells in the lymphocyte gate for each sample. The Kaluza™ acquisition software (version 1.5; Beckman Coulter, Miami, FL) was used for the analysis of the flow cytometry data. The antibodies used were the following: anti-CD38-APC-Cyanine 5.5 (APC-Cy5.5) and anti-CD3-PO (Pacific Orange) from Invitrogen (Frederick, MD, USA); anti-CD28-PE (Phycoerythrin), anti-CD57-FITC (Fluorescein), anti-CD127-PC7 (Phycoerythrin-Cyanin 7), anti-CD25-PC5 (Phycoerythrin-Cyanin 5.1), anti-CD45RA-ECD (Phycoerythrin-Texas Red X) from Beckman Coulter (Marseille, France); and anti-CD8-PB (Pacific Blue) and anti-CD4-APC-Cy7 (APC-Cyanine 7) from BioLegend (San Diego, CA, USA).

The representative flow cytometer plots and gating strategy are shown in **Supplementary** **Figure 1**. Here, we show the gating strategy to visualize Treg cells. First of all, lymphocytes were gated on a forward scatter/side scatter (FSC-A/SSC-A) dot-plot. Next, lymphocytes were displayed using a CD3/CD4 dot-plot and CD4+ T-cells were selected on a gate of CD4+. These CD4+ T-cells are displayed on a CD127/CD25 dot-plot to visualize Treg cells (CD4+CD127low/-CD25+) in a gate. At the same time, CD4+ T-cells were displayed in a CD45RA histogram, gating CD4+CD45RA- and CD4+CD45RA+ on a CD127/CD25 dot-plot to visualize mTreg and rTreg cells.

**Supplementary Figure 1**. Flow cytometry gating strategy for regulatory CD4+ T-cells subsets in a representative sample.



## Multiplex assay and ELISA

The measure of the plasma biomarkers was performed using a Luminex 200™ analyzer (Luminex Corporation, Austin, TX, United States) with the ProcartaPlexTM multiplex immunoassay (Bender MedSystems GmbH, Vienna, Austria) and following the manufacturer’s specifications. The 28 plasma biomarkers measured by ELISA multiplex were: IL-10, IL-1 receptor antagonist (IL-1RA), IL-4, IFN-γ-inducible protein 10 (IP-10), IL-8 (chemokine (C-X-C motif) ligand 8, CXCL8), monocyte chemoattractant protein-1 (MCP-1)], IL-1β, IL-18, IL-6, tumor necrosis factor-alpha (TNF-α), interferon (IFN)-γ, IL-12p70, IL-2, IL-17A, soluble intercellular cell adhesion molecule 1 (sICAM-1), soluble tumor necrosis factor receptor 1 (sTNFR-1), plasminogen activator inhibitor-1 (PAI-1), D-Dimer, soluble receptor activator of nuclear factor-kappaB ligand (sRANKL), osteoprotegerin (OPG), insulin, leptin, vascular endothelial growth factor A (VEGF-A), soluble receptors for vascular endothelial growth factor (sVEGF-R1), soluble programmed death protein 1 (sPD-1), and soluble programmed cell death ligand 1 (sPD-L1).

Besides, we used four single commercial ELISA kits for biomarkers that were not available by multiplex ELISA: Lipopolysaccharide binding protein (LBP) (R&D Systems, Minneapolis, USA), sCD14 and Fatty acid-binding protein 2 (FABP-2) (Raybiotech, Georgia, USA), and transforming growth factor-beta 1 (TGF-β1; Bender MedSystems GmbH, Vienna, Austria). The lipopolysaccharide (LPS; Hycult Biotech, Uden, The Netherlands) was evaluated by a Limulus amebocyte lysate chromogenic endpoint ELISA.

## RNA extraction, library preparation, and RNA sequencing

Total RNA was extracted from PBMC with the RNeasy Minikit (Qiagen™), according to the manufacturer’s instructions. RNA quantity was evaluated with Nanodrop 2000 and quality was assessed with the 2100 Bioanalyzer RNA NANO assay (Agilent), and only those samples with RNA integrity numbers higher than 7.5 were selected for sequencing. Libraries and sequencing of poly-A RNA were performed at the Centre for Genomic Regulation in Barcelona (Spain). Firstly, libraries were synthesized with the Illumina’s TruSeq Stranded mRNA Sample Prep Kit v2 using 500 nanograms of total RNA following the manufacturer’s protocol. This procedure captures both coding RNA and multiple forms of noncoding polyadenylated RNAs. Ten libraries were multiplexed and pooled for sequencing in the same line, to obtain 25 million reads per sample, on average. Sequencing was performed on an Illumina HiSeq2500, single read, 50nts (1x50).

Raw sequences were analyzed as follow: quality control was carried out before and after trimming with FastQC (v. 0.11.8), adapter sequences were trimmed with Trimmomatic (v. 0.33), alignment was performed with TopHat (v. 2.0.14) using GRCH38 as a reference genome, and the read counts were extracted by HTSeq (v. 0.6.1). The raw RNA data are publicly available at the ArrayExpress repository (EMBL-EBI; <https://www.ebi.ac.uk/>) under the accession number E-MTAB-8249.

Raw sequences were analyzed following the bioinformatics protocol. Quality control was carried out before and after trimming with FastQC (v. 0.11.8), adapter sequences were trimmed with Trimmomatic (v. 0.33), alignment was performed with TopHat (v. 2.0.14) using GRCH38 as a reference genome and the read counts were extracted by HTSeq (v. 0.6.1).

**Supplementary Table 1.** Bioinformatic pipeline to analyze raw sequences from RNA-seq of HIV-infected and healthy controls.

|  |
| --- |
| *1. Filtering step*Settings employed to eliminate adapters and low quality reads.- Software: Trimmomatic- Version: 0.33- Code:java -jar PATH-TO-TRIMMOMATIC/trimmomatic-0.33.jar SE -threads 10 -phred33 PATH-TO-SAMPLES/”sample”.fastq.gz “sample”/”sample”\_filtered.fastq ILLUMINACLIP:all\_PE.fa:2:30:10 SLIDINGWINDOW:4:15 MINLEN:50   |
| *2. Mapping Step*Settings employed for mapping the filtered counts.- Software:Tophat2- Version:2.0.14- Human genome:GRCh38- Code:mkdir -p “sample”; qsub -V -b y -j y -cwd -N TOPHATALIGNMENT -q all.q -pe openmp 10 tophat2 -p 10 -o 026C -G ../../REFERENCES/GRCh38\_refseq.gtf --transcriptome-index ../../REFERENCES/ ../../REFERENCES/hg38.fullAnalysisSet.fa ../02-preprocessing/”sample”/”sample”\_filtered.fastq.gz   |
| *3. Count step*Settings employed to obtain the number of counts per gene and sample- Software:HTSeq- Version:0.6.1- Code:#!/bin/bash#$ -V#$ -b y#$ -j y#$ -cwd#$ -N HTSEQCOUNT#$ -q all.q#$ -t 1-100set -eset -xinfile=../samples\_id.txtin=$(awk "NR==$SGE\_TASK\_ID" $infile)mkdir -p $inhtseq-count -f bam ../04-tophat/$in/"$in"\_accepted\_hits.bam ../../REFERENCES/GRCh38\_refseq.gtf > $in/"$in"\_htseqCount.txt |