Science Translational Medicine
AAAS

Supplementary Materials for

Single-cell transcriptomics identifies prothymosin α restriction of HIV-1 in vivo

Aviva Geretz *et al.*

Corresponding author: Rasmi Thomas, rthomas@hivresearch.org

Sci. Transl. Med. **15**, eadg0873 (2023)
DOI: 10.1126/scitranslmed.adg0873

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

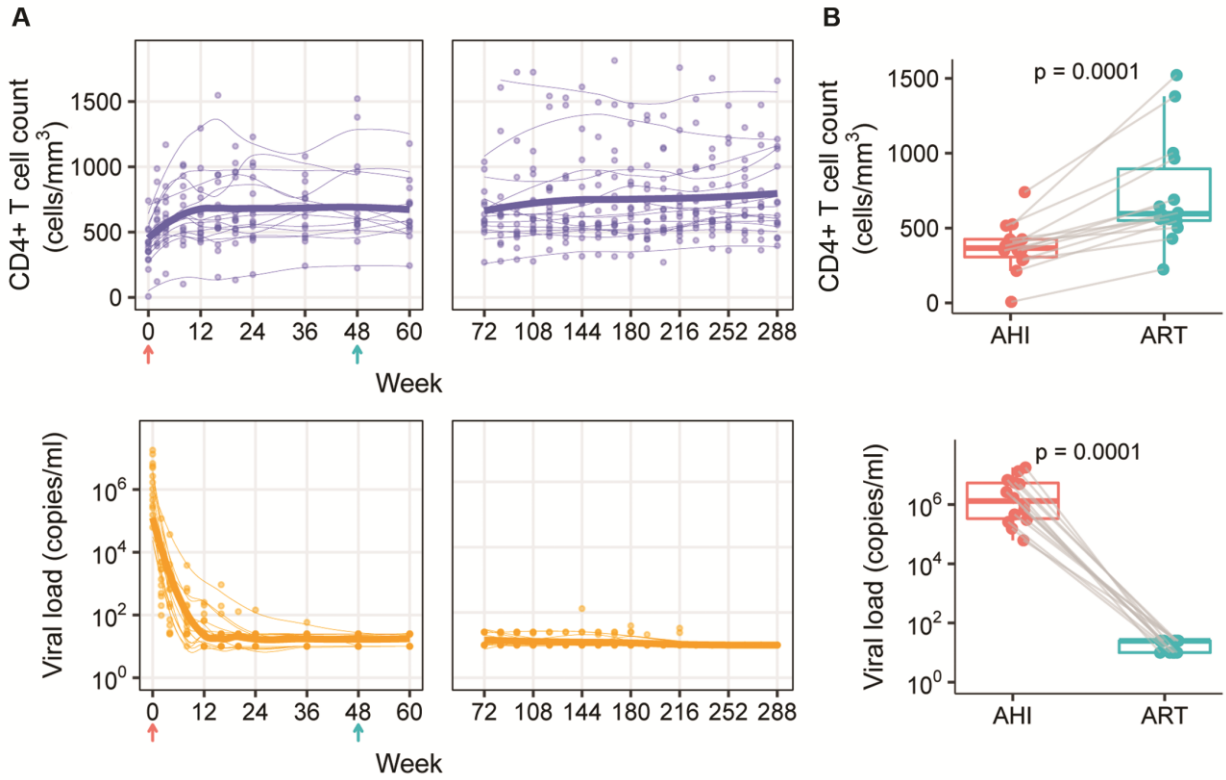

Fig. S1. Participant CD4+ T cell counts and viral load (VL) differ at antiretroviral therapy (ART) and acute HIV infection (AHI) timepoints. (**A**) Absolute CD4+ T cell counts (blue) and VL (orange) for 14 participants over 288 weeks of follow-up. Heavy lines indicate mean values. Participant samples from week 0 (AHI; Fiebig stage III) and week 48 (ART) were used in the study. Participants were all male; Thai; HIV Fiebig staging III; and with subtype CRF01_AE. (**B**) CD4+ T cell counts and HIV VL at weeks 0 (AHI) and 48 after ART initiation (ART) for 14 participants. Significance was determined by Wilcoxon signed rank test.

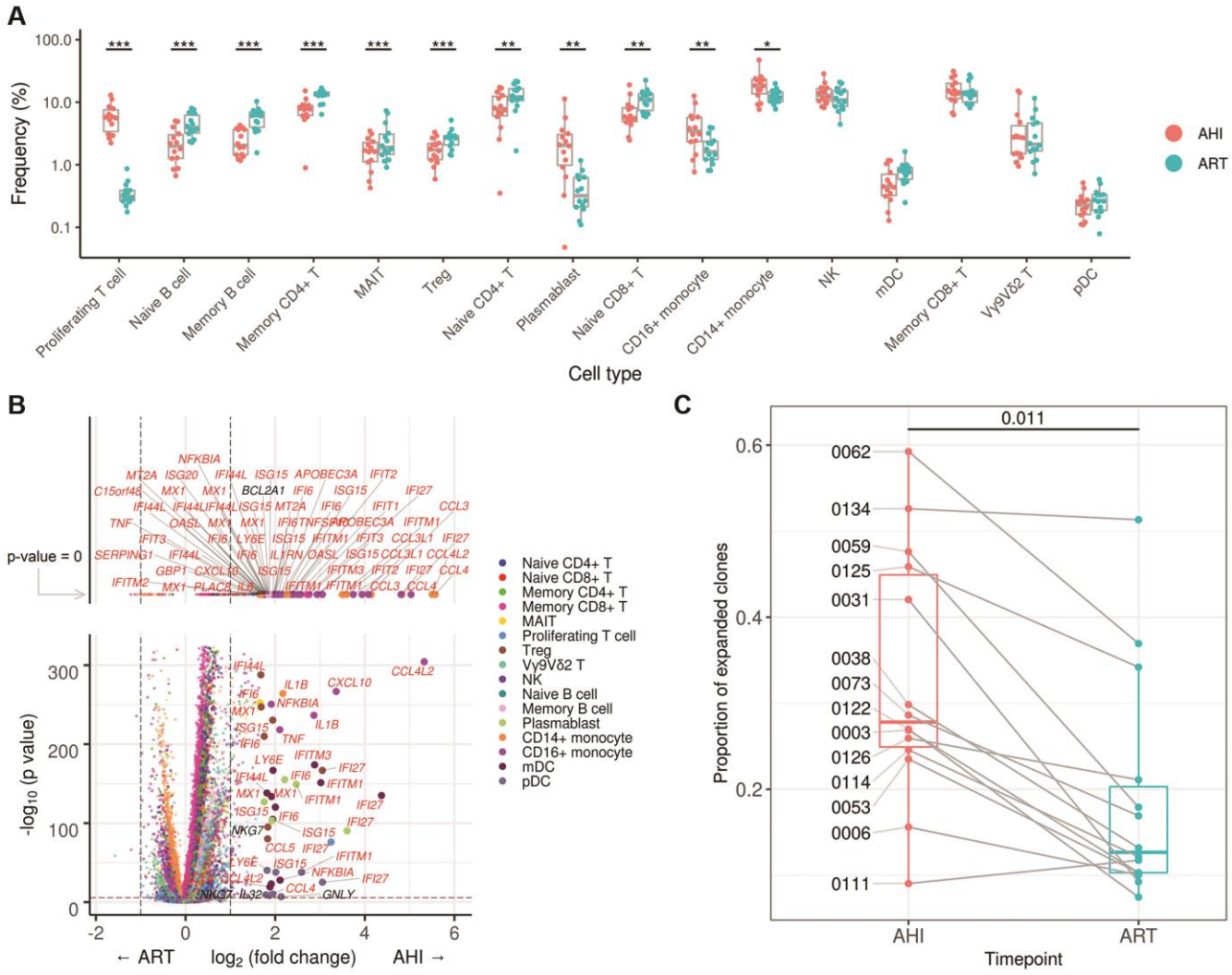

Fig. S2. Difference in cell subset frequencies, gene expression, and T cell receptor (TCR) expansion between the AHI and ART timepoints. (**A**) Population clusters of immune cell subsets differ in frequency between the AHI and ART timepoints. MAIT, mucosal-associated invariant T; Treg, regulatory T; NK, natural killer; mDC, myeloid dendritic cells; pDC, plasmacytoid DC. Asterisks indicate significant results (False discovery rate (FDR)-adjusted *p <0.05, **p <0.01, ***p <0.001). (**B**) The significance and fold change between AHI and ART time points for differentially expressed genes (DEG) in the major immune cell types in peripheral blood mononuclear cell (PBMC) are shown. Top 100 significant genes with the highest fold change are annotated. Red gene symbols indicate the gene is an interferon-stimulated gene (ISG). (**C**) Comparison of TCR expansion including all T cell populations at the AHI compared to ART timepoint. Significance from Wilcoxon signed rank test was implemented for A and C.

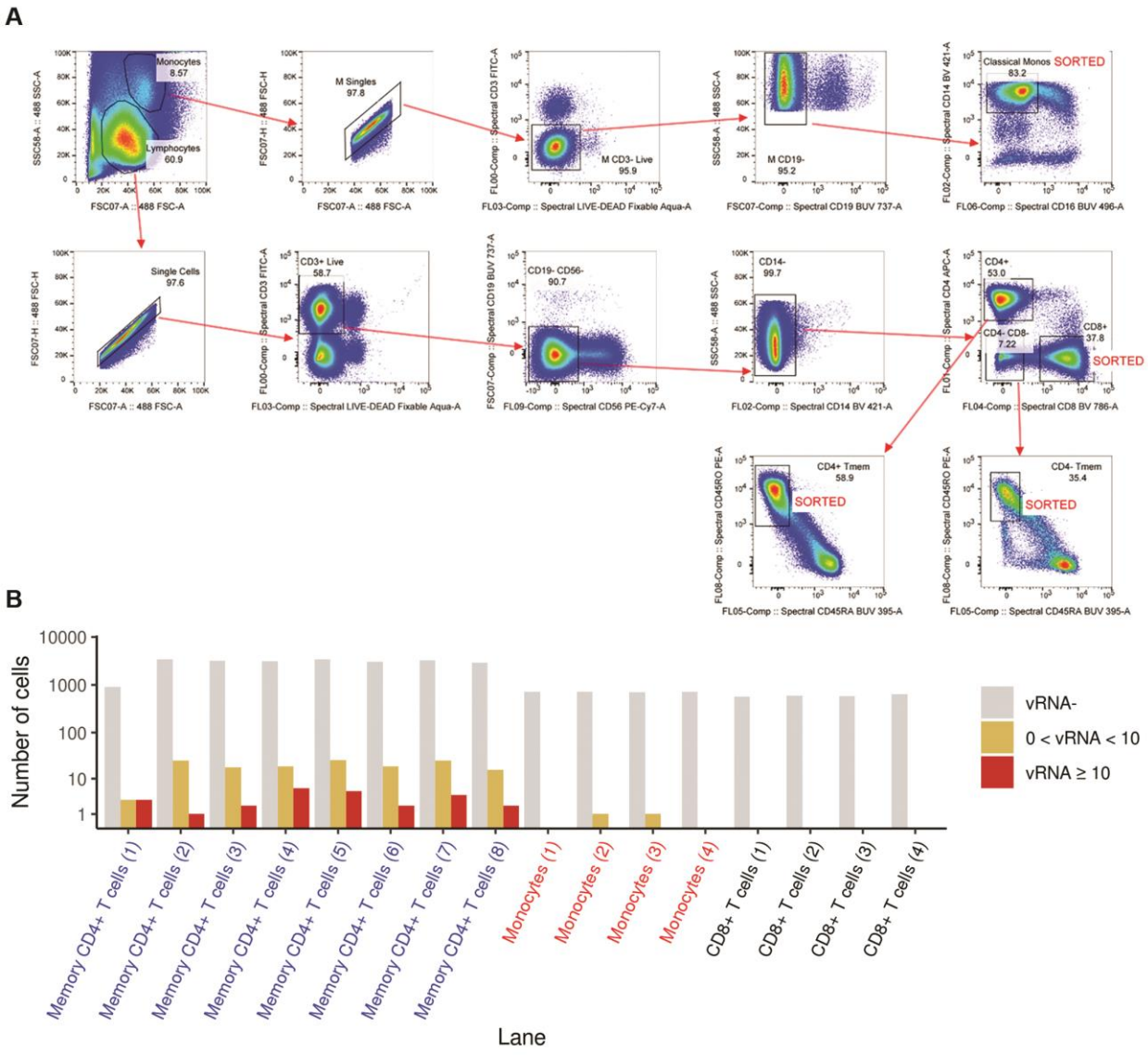

Fig. S3. Cells with high viral RNA (vRNA) using single cell RNA sequencing (scRNA-seq) were only detected in sorted CD4+ T cell populations. Sensitivity of vRNA+ cell detection using the scRNA-seq methodology was assessed in different cell populations including memory CD4+ T cells, monocytes and CD8+ T cells that were isolated by flow cytometry. (**A**) Gating strategy for participant 3 is shown. (**B**) Sorted cell populations loaded per Chromium Chip lane are shown in different font colors. Blue: memory CD4+ T cells; red: monocytes; black: CD8+ T cells.

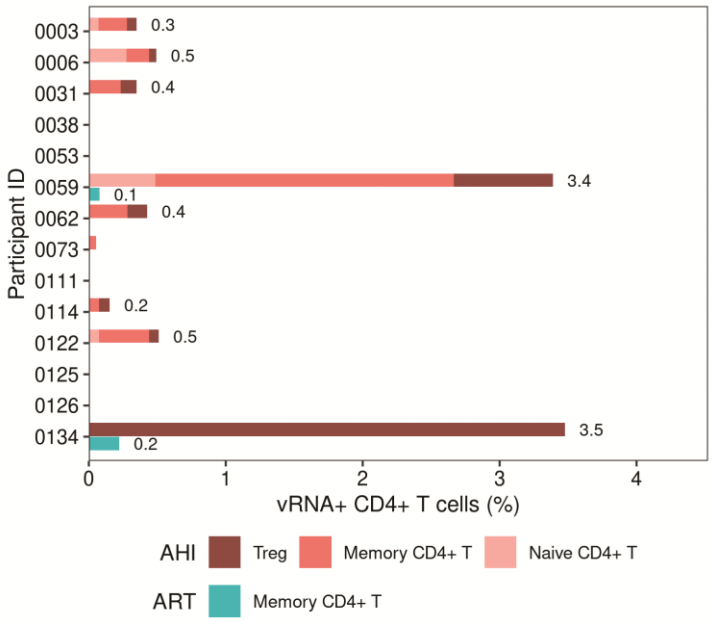

Fig. S4. Distribution of vRNA+ cells in total CD4+ T cell populations. Frequency of vRNA+ CD4+ T cell immune cell subsets at the AHI and ART timepoints.

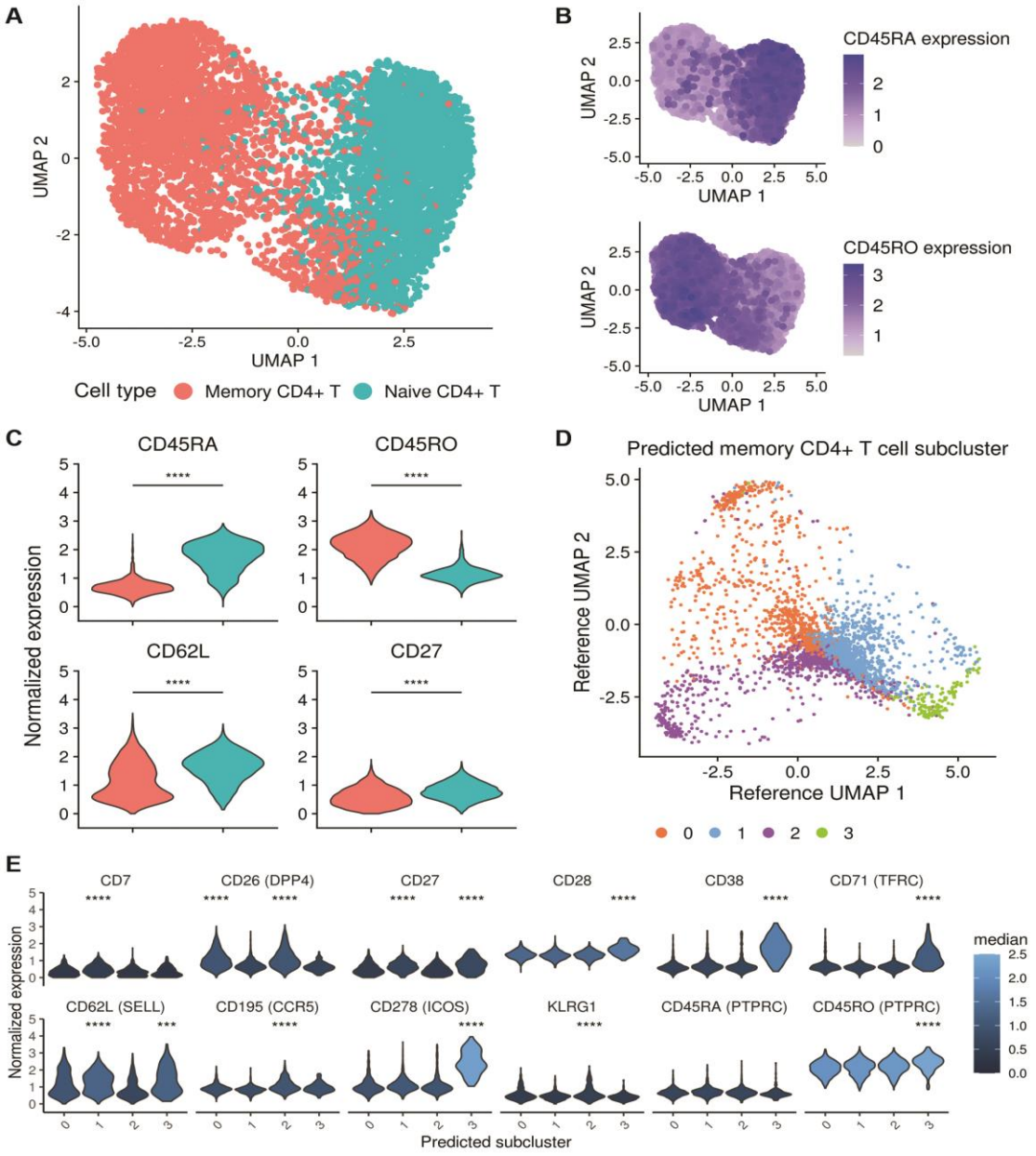

Fig. S5. Surface protein expression from cellular indexing of transcriptomes and epitopes sequencing (CITE-seq) of CD4+ T cells. (**A**) Dimensionality reduction plot of single cells showing CD4+ memory and naive T cells from the 7 participants at AHI. (**B**) Surface expression of CD45RA and RO; darker color indicates higher expression. Memory CD4+ T cells are CD45RO+CD45RA- and naïve cells are CD45RO-CD45RA+. (**C**) Expression distribution of CD45RA, CD45RO, CD62L (SELL), and CD27 in the two CD4+ T cell subsets. Significance from Wilcoxon rank sum test comparing protein expression between the two clusters is shown. (**D**) CD4+ memory T cell gene expression data from the 7 participants was projected into the 0059 dimensionality reduction UMAP (uniform manifold approximation and projection) space to predict subcluster identity. (**E**) Surface protein expression of selected markers from the 7 participants separated by predicted subcluster and colored by median expression of that subcluster. Significance for each predicted subcluster is shown, where Wilcoxon rank sum test was performed comparing that subcluster's expression to the expression of cells in all three other subclusters. P values of $\log_2$ fold change (FC) <0.15 are not shown. *p <0.05, **p <0.01, ***p <0.001, and ****p <0.0001.

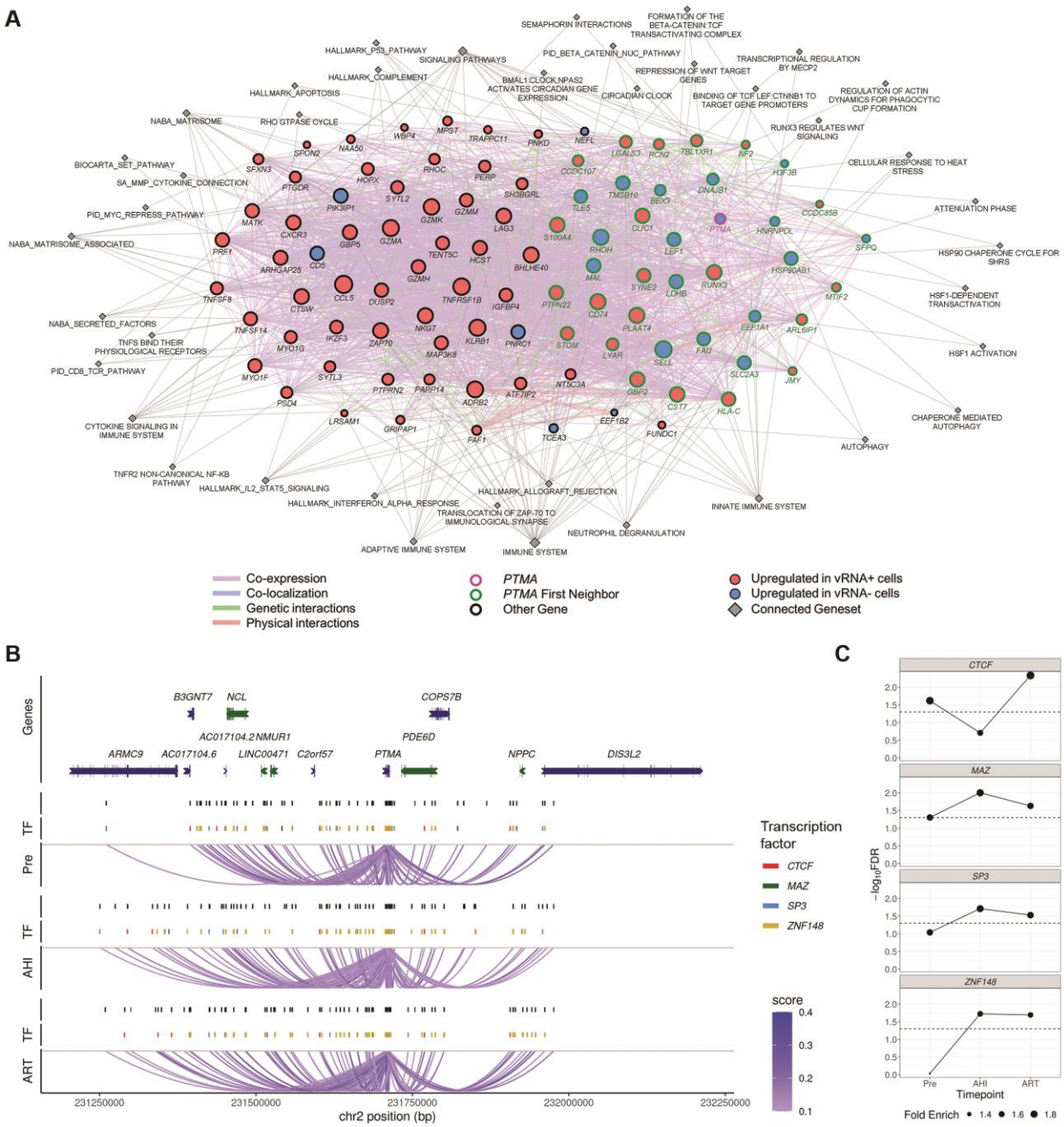

Fig. S6. Transcription factors regulating *PTMA*. (**A**) GeneMania network of the top 100 DEG between vRNA+ and vRNA- cells in participant 0059 shows known connections between genes, as well as the top connected genesets. Genesets and genes are indicated by diamond and circular nodes, respectively. *PTMA* has a magenta outline and first neighbor genes to *PTMA* are outlined in green. Nodes are colored by direction of gene expression: red indicates higher expression in vRNA+ cells, blue in vRNA- cells. Node sizes correspond to the number of connections to other nodes. (**B**) Epigenetic landscape at the *PTMA* locus in memory CD4+ T cells. Gene annotations for the *PTMA* locus are shown on top in blue (forward strand) or green (reverse strand). Chromatin accessibility regions ("peaks"), which represent potential enhancers, are shown as vertical gray lines. Transcription factor (TF) binding site motifs within peaks are shown as red (*CTCF*), green (*MAZ*), blue (*SP3*), and gold (*ZNF148*). Purple arcs indicate co-accessibility of peaks to *PTMA* gene peaks calculated by Cicero at preinfection (Pre), AHI and ART timepoints. (**C**) Motifs enriched in peaks across timepoints. Size indicates fold enrichment of motif within co-accessible peaks. Dashed lines show FDR <0.05 significance threshold.

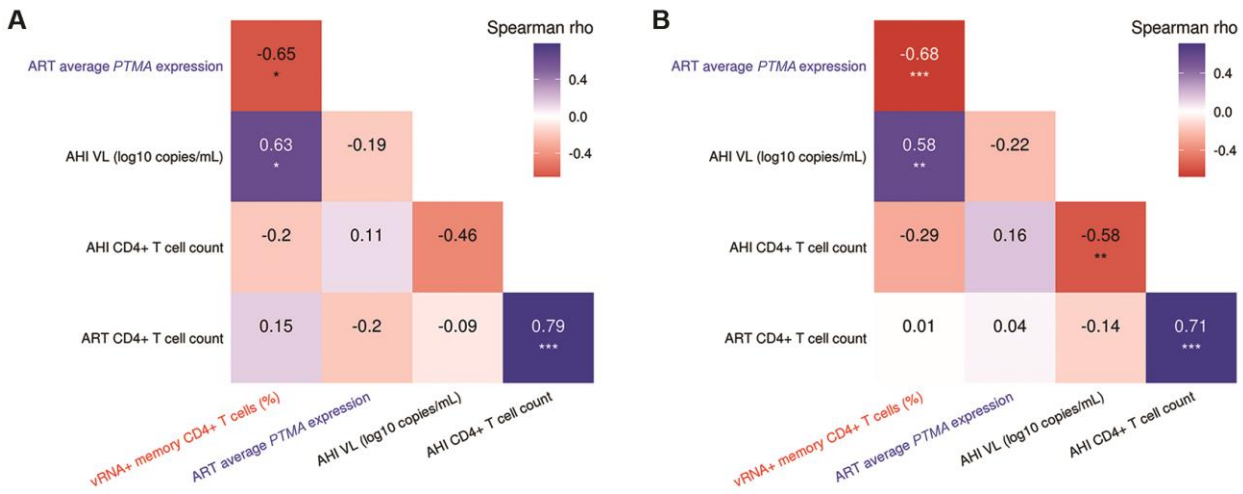

Fig. S7. Correlation of *PTMA* expression at ART with frequency of vRNA+ memory CD4+ T cells is independent of VL and CD4+ T cell counts. (**A** and **B**) Correlation of VL and CD4+ T cell counts with *PTMA* expression at ART in the 14 participant dataset (A) and the 21 participant dataset (B). Viral phenotype and expression of PTMA generated by scRNA-seq is shown in red and blue font, respectively. All clinical parameters are shown in black font.Spearman correlation was performed. *p <0.05, **p <0.01, ***p <0.001, ****p <0.0001.

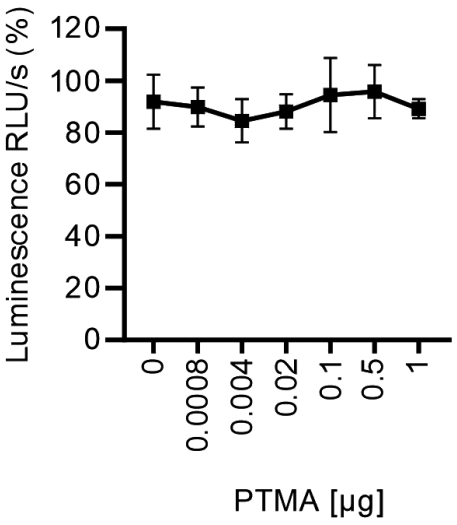

Fig. S8. Overexpression of PTMA is not cytotoxic. HEK293T cells were transfected with the indicated amounts of PTMA expression construct. Two days post-transfection, cells were harvested and cell viability was measured by CellTiter-Glo® Luminescent Cell Viability Assay. Shown are mean values (± SEM; n=2) of luminescence (RLU/s) relative to those obtained for untreated cells (100%).

Table S1. Demographic and clinical characteristics of the participants in this study.

PID	Alias	Cohort	Method	HIV-1 RNA Log_{10} (Copies/ml)(AHI)	CD4+ T cell counts (cells/mm^3) (AHI)	HIV-1 RNA (Copies/ml) (ART)	CD4+ T cell counts (cells/mm^3) (ART)	HIV-1 DNA (cp/million cells) (ART)	Age	Sex	Ethnicitiy	HIV Subtype updated DEC2021	VL/ Fiebig stage	AHI duration (days)[#]
0003	-	RV254	PBMC 10x GEX, TCR	5.5	740	<50	1380	586	28	M	Thai	CRF01_AE	III	20
0006	-	RV254	PBMC 10x GEX, TCR	6.7	525	<50	1522	below LOD	46	M	Thai	CRF01_AE	III	15
0031	-	RV254	PBMC 10x GEX, TCR	6.7	352	<50	690	295	30	M	Thai	CRF01_AE	III	21
0038	-	RV254	PBMC 10x GEX, TCR	5.7	409	<50	503	below LOD	22	M	Thai	CRF01_AE	III	31
0053	-	RV254	PBMC 10x GEX, TCR	4.8	352	<50	602	below LOD	29	M	Thai	CRF01_AE	III	15
0059	1	RV254	PBMC 10x GEX, TCR	7.25	292	<50	562	573	34	M	Thai	CRF01_AE	III	12
0062	-	RV254	PBMC 10x GEX, TCR	6.4	350	<50	590	183	22	M	Thai	CRF01_AE	III	19
0073	-	RV254	PBMC 10x GEX, TCR	5.8	425	<20	964	below LOD	23	M	Thai	CRF01_AE	III	20
0111	-	RV254	PBMC 10x GEX, TCR	5.2	425	<20	559	below LOD	30	M	Thai	CRF01_AE	III	10
0114	-	RV254	PBMC 10x GEX, TCR	6	286	<20	429	below LOD	20	M	Thai	CRF01_AE	III	24
0122	-	RV254	PBMC 10x GEX, TCR	7.1	214	<20	547	300	31	M	Thai	CRF01_AE	III	27
0125	-	RV254	PBMC 10x GEX, TCR	5.4	516	<20	1002	below LOD	20	M	Thai	CRF01_AE	III	21
0126	-	RV254	PBMC 10x GEX, TCR	6.8	381	<20	644	37	22	M	Thai	CRF01_AE/B Recombinant	III	19
0134	-	RV254	PBMC 10x GEX, TCR	6.2	7	25	224	4559	34	M	Thai	CRF01_AE/B Recombinant	III	20
40577	2	RV217	Sorted mem CD4 T cells 10x GEX, TCR	7.58	702	-	-	-	23	M	Thai	CRF01_AE	peak VL	12
40512	3	RV217	Sorted mem CD4 T cells 10x GEX, TCR	6.82	655	-	-	-	18	M	Thai	CRF01_AE	peak VL	15
0164	40774	RV254	PBMC 10x GEX	3.5	1033	<20	977	NA	21	M	Thai	CRF01_AE	III	21
0235	40943	RV254	PBMC 10x GEX	4.4	794	<20	635	NA	29	M	Thai	CRF01_AE	III	7
0366	41105	RV254	PBMC 10x GEX	4.9	712	<20	736	NA	28	M	Thai	CRF01_AE	III	18
0536	41250	RV254	PBMC 10x GEX	5.1	548	<20	1264	NA	42	M	Thai	CRF01_AE	III	12
0394	41139	RV254	PBMC 10x GEX	6.6	656	<20*	**	NA	28	M	Thai	CRF01_AE	III	14
0436	41225	RV254	PBMC 10x GEX	6.8	668	<20	1170	NA	30	M	Thai	CRF01_AE	III	24
0466	41259	RV254	PBMC 10x GEX	7.7	259	<20	684	NA	23	M	Thai	CRF01_AE	III	10

* HIV-1 RNA was not measured at week 48 on ART, but was <20 copies/mL at weeks 36 and 60.
** CD4+ T cell counts were not measured at week 48 on ART, but were 592 and 847 cells/mm^3 at weeks 36 and 60, respectively.
[#] Estimated number of days from HIV exposure to enrollment date prior to ART initiation.

LOD: Below lower limit of detection of assay. All participants with HIV DNA below LOD (10 copies/million cells) were assigned an absolute value of 0.

Supplementary Data Files

Data file S1. Cell counts and percentages of different cell subsets at the AHI and ART timepoints.
Data file S2. Top 100 differentially expressed genes were upregulated at the AHI versus ART timepoints.
Data file S3. DEG between different CD4+ memory T subclusters in participant 0059.
Data file S4. DEG between vRNA+ versus - memory CD4+ T cells in participant 0059.
Data file S5. MAST results using normalized *PTMA* as a continuous variable.
Data file S6. Raw, individual-level data for experiments where n<20.