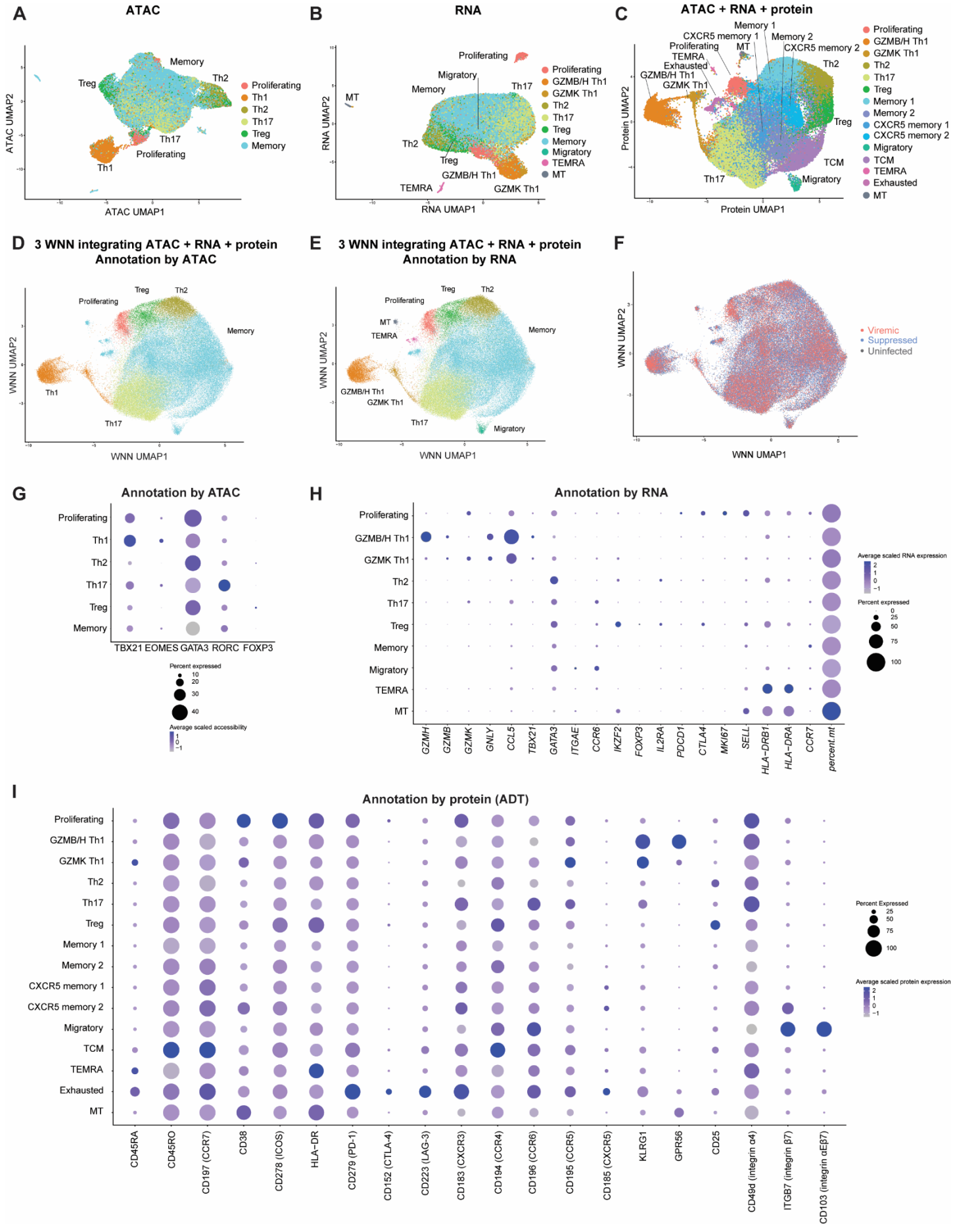
A
ATAC
ATAC UMAP1
ATAC UMAP2
Memory
Treg
Th2
Th17
Proliferating
Th1
Proliferating
Th1
Th2
Th17
Treg
Memory
B
RNA
RNA UMAP1
RNA UMAP2
Proliferating
Migratory
Memory
Th17
MT
Th2
Treg
GZMB/H Th1
TEMRA
GZMK Th1
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory
Migratory
TEMRA
MT
C
ATAC + RNA + protein
Protein UMAP1
Protein UMAP2
Memory 1
CXCR5 memory 1
Proliferating
MT
Memory 2
CXCR5 memory 2
TEMRA
Exhausted
GZMB/H Th1
GZMK Th1
Th2
Treg
TCM
Th17
Migratory
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory 1
Memory 2
CXCR5 memory 1
CXCR5 memory 2
Migratory
TCM
TEMRA
Exhausted
MT
D
3 WNN integrating ATAC + RNA + protein
Annotation by ATAC
WNN UMAP1
WNN UMAP2
Proliferating
Treg
Th2
Memory
Th1
Th17
E
3 WNN integrating ATAC + RNA + protein
Annotation by RNA
WNN UMAP1
WNN UMAP2
Proliferating
Treg
Th2
MT
TEMRA
Memory
GZMB/H Th1
GZMK Th1
Th17
Migratory
F
WNN UMAP1
WNN UMAP2
Viremic
Suppressed
Uninfected
G
Annotation by ATAC
Proliferating
Th1
Th2
Th17
Treg
Memory
TBX21 EOMES GATA3 RORC FOXP3
Percent expressed
10
20
30
40
Average scaled accessibility
1
0
−1
H
Annotation by RNA
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory
Migratory
TEMRA
MT
GZMH GZMB GZMK GNLY CCL5 TBX21 GATA3 ITGAE CCR6 IKZF2 FOXP3 IL2RA PDCD1 CTLA4 MKI67 SELL HLA−DRB1 HLA−DRA CCR7 percent.mt
Average scaled RNA expression
2
1
0
−1
Percent expressed
0
25
50
75
100
I
Annotation by protein (ADT)
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory 1
Memory 2
CXCR5 memory 1
CXCR5 memory 2
Migratory
TCM
TEMRA
Exhausted
MT
CD45RA CD45RO CD197 (CCR7) CD38 CD278 (ICOS) HLA−DR CD279 (PD-1) CD152 (CTLA-4) CD223 (LAG-3) CD183 (CXCR3) CD194 (CCR4) CD196 (CCR6) CD195 (CCR5) CD185 (CXCR5) KLRG1 GPR56 CD25 CD49d (integrin α4) ITGB7 (integrin β7) CD103 (integrin αEβ7)
Percent Expressed
25
50
75
100
Average scaled protein expression
2
1
0
−1

Figure S1. Identities of memory CD4$^+$ T cell subsets were defined by epigenetic, transcriptional, and protein states, related to Figure 1. (**A**) ATAC UMAP projection of memory CD4$^+$ T cells defined by chromatin accessibility profile (ATAC). (**B**) RNA UMAP projection of memory CD4$^+$ T cells defined by transcriptional profile (RNA). (**C**) Protein UMAP projection of memory CD4$^+$ T cells defined by surface protein expression profile. (**D**) WNN UMAP showing cell subsets annotated by chromatin accessibility features. (**E**) WNN UMAP showing cell subsets annotated by RNA features. (**F**) Batch effect removal by Harmony (RNA), reciprocal LSI (ATAC), and reciprocal PCA (protein) showed integration across viremic, suppressed, and uninfected conditions. (**G**–**I**) Key epigenetic, transcriptional, and protein markers used to annotate all 15 computationally identified clusters. (**G**) Average normalized and scaled chromatin accessibility (ATAC) in each cell group. (**H**) Average normalized and scaled RNA expression in each cell group. (**I**) Average CLR-normalized and scaled surface protein expression in each cell group.

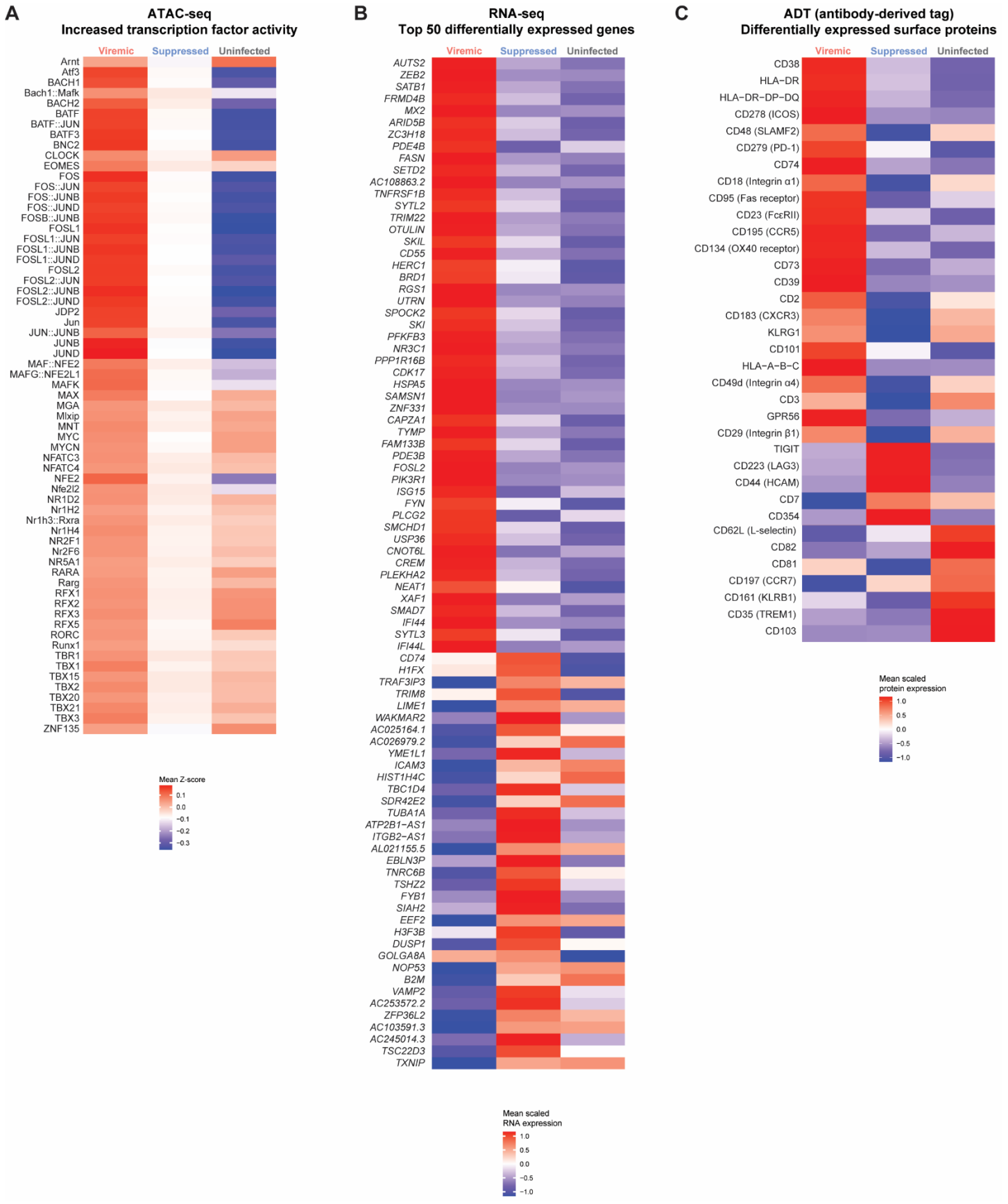
A
ATAC-seq
Increased transcription factor activity
Viremic
Suppressed
Uninfected
Arnt
Atf3
BACH1
Bach1::Mafk
BACH2
BATF
BATF::JUN
BATF3
BNC2
CLOCK
EOMES
FOS
FOS::JUN
FOS::JUNB
FOS::JUND
FOSB::JUNB
FOSL1
FOSL1::JUN
FOSL1::JUNB
FOSL1::JUND
FOSL2
FOSL2::JUN
FOSL2::JUNB
FOSL2::JUND
JDP2
Jun
JUN::JUNB
JUNB
JUND
MAF::NFE2
MAFG::NFE2L1
MAFK
MAX
MGA
Mlxip
MNT
MYC
MYCN
NFATC3
NFATC4
NFE2
Nfe2l2
NR1D2
Nr1H2
Nr1h3::Rxra
Nr1H4
NR2F1
Nr2F6
NR5A1
RARA
Rarg
RFX1
RFX2
RFX3
RFX5
RORC
Runx1
TBR1
TBX1
TBX15
TBX2
TBX20
TBX21
TBX3
ZNF135
Mean Z-score
0.1
0.0
−0.1
−0.2
−0.3
B
RNA-seq
Top 50 differentially expressed genes
Viremic
Suppressed
Uninfected
AUTS2
ZEB2
SATB1
FRMD4B
MX2
ARID5B
ZC3H18
PDE4B
FASN
SETD2
AC108863.2
TNFRSF1B
SYTL2
TRIM22
OTULIN
SKIL
CD55
HERC1
BRD1
RGS1
UTRN
SPOCK2
SKI
PFKFB3
NR3C1
PPP1R16B
CDK17
HSPA5
SAMSN1
ZNF331
CAPZA1
TYMP
FAM133B
PDE3B
FOSL2
PIK3R1
ISG15
FYN
PLCG2
SMCHD1
USP36
CNOT6L
CREM
PLEKHA2
NEAT1
XAF1
SMAD7
IFI44
SYTL3
IFI44L
CD74
H1FX
TRAF3IP3
TRIM8
LIME1
WAKMAR2
AC025164.1
AC026979.2
YME1L1
ICAM3
HIST1H4C
TBC1D4
SDR42E2
TUBA1A
ATP2B1−AS1
ITGB2−AS1
AL021155.5
EBLN3P
TNRC6B
TSHZ2
FYB1
SIAH2
EEF2
H3F3B
DUSP1
GOLGA8A
NOP53
B2M
VAMP2
AC253572.2
ZFP36L2
AC103591.3
AC245014.3
TSC22D3
TXNIP
Mean scaled
RNA expression
1.0
0.5
0.0
−0.5
−1.0
C
ADT (antibody-derived tag)
Differentially expressed surface proteins
Viremic
Suppressed
Uninfected
CD38
HLA−DR
HLA−DR−DP−DQ
CD278 (ICOS)
CD48 (SLAMF2)
CD279 (PD-1)
CD74
CD18 (Integrin α1)
CD95 (Fas receptor)
CD23 (FcεRII)
CD195 (CCR5)
CD134 (OX40 receptor)
CD73
CD39
CD2
CD183 (CXCR3)
KLRG1
CD101
HLA−A−B−C
CD49d (Integrin α4)
CD3
GPR56
CD29 (Integrin β1)
TIGIT
CD223 (LAG3)
CD44 (HCAM)
CD7
CD354
CD62L (L-selectin)
CD82
CD81
CD197 (CCR7)
CD161 (KLRB1)
CD35 (TREM1)
CD103
Mean scaled
protein expression
1.0
0.5
0.0
−0.5
−1.0

Figure S2. Distinct epigenetic, transcriptional, and surface protein profiles of memory CD4$^+$ T cells were captured during viremia and viral suppression, related to Figure 1. (**A**) Differential chromatin accessibility of transcription factor binding motifs, represented in heatmap as comparisons of means in ATAC-seq chromVAR bias-corrected deviations (Z-score) between cell groups. All motifs shown had significantly increased accessibility in the viremic group, by P <0.05 and mean Z-score difference ≥ 0.1. (**B**) Differentially expressed genes, represented in heatmap as comparisons of cell group means in normalized and scaled RNA expression. All genes passed P <0.05, min.pct ≥ 0.1, and $\log_2$FC ≥ 0.15. (**C**) Differentially expressed proteins, represented in heatmap as comparisons of cell group means in DSB-normalized and scaled protein expression. All features passed P <0.05 and $\log_2$FC ≥ 0.25. The mean expression of all protein features shown were also tested to be greater than the mean expression of their specific isotype controls (Z > 2; two-sample Z-test). For all heatmaps, n = 25,778, 56,771, 10,660 memory CD4$^+$ T cells in the viremic, viral suppression, and uninfected conditions, respectively. Statistical significance for heatmaps was determined by Wilcoxon rank-sum test for comparisons between each group and all cells in the other two groups. All P values were FDR-adjusted.

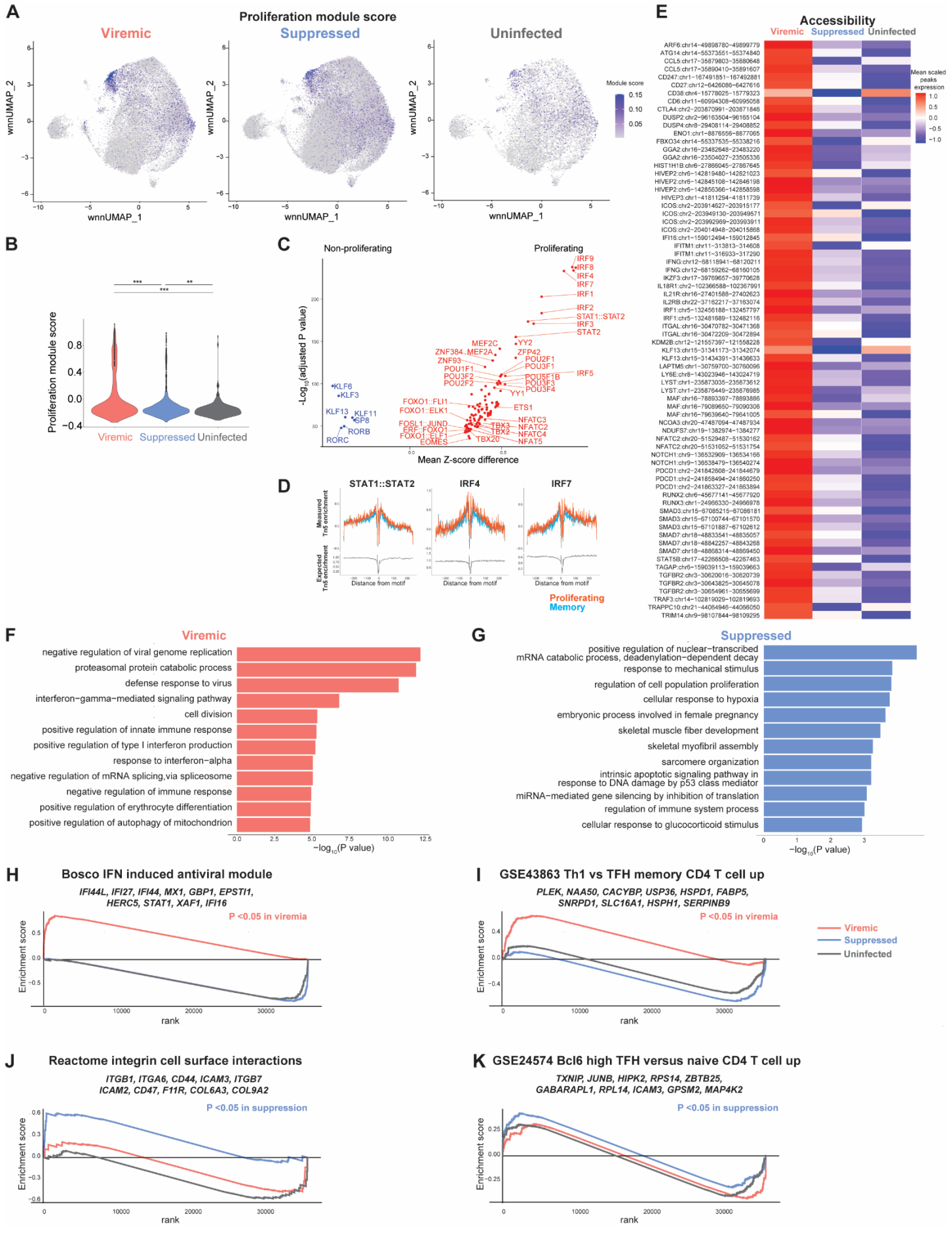
A
Proliferation module score
Viremic
Suppressed
Uninfected
wnnUMAP_1
wnnUMAP_2
Module score
B
Proliferation module score
Viremic
Suppressed
Uninfected
C
Non-proliferating
Proliferating
-Log10(adjusted P value)
Mean Z-score difference
D
STAT1::STAT2
IRF4
IRF7
Measured Tn5 enrichment
Expected Tn5 enrichment
Distance from motif
Proliferating
Memory
E
Accessibility
Viremic
Suppressed
Uninfected
Mean scaled peaks expression
F
Viremic
negative regulation of viral genome replication
proteasomal protein catabolic process
defense response to virus
interferon−gamma−mediated signaling pathway
cell division
positive regulation of innate immune response
positive regulation of type I interferon production
response to interferon−alpha
negative regulation of mRNA splicing,via spliceosome
negative regulation of immune response
positive regulation of erythrocyte differentiation
positive regulation of autophagy of mitochondrion
−log10(P value)
G
Suppressed
positive regulation of nuclear−transcribed mRNA catabolic process, deadenylation−dependent decay
response to mechanical stimulus
regulation of cell population proliferation
cellular response to hypoxia
embryonic process involved in female pregnancy
skeletal muscle fiber development
skeletal myofibril assembly
sarcomere organization
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator
miRNA−mediated gene silencing by inhibition of translation
regulation of immune system process
cellular response to glucocorticoid stimulus
−log10(P value)
H
Bosco IFN induced antiviral module
IFI44L, IFI27, IFI44, MX1, GBP1, EPSTI1, HERC5, STAT1, XAF1, IFI16
P <0.05 in viremia
Enrichment score
rank
I
GSE43863 Th1 vs TFH memory CD4 T cell up
PLEK, NAA50, CACYBP, USP36, HSPD1, FABP5, SNRPD1, SLC16A1, HSPH1, SERPINB9
P <0.05 in viremia
Viremic
Suppressed
Uninfected
Enrichment score
rank
J
Reactome integrin cell surface interactions
ITGB1, ITGA6, CD44, ICAM3, ITGB7 ICAM2, CD47, F11R, COL6A3, COL9A2
P <0.05 in suppression
Enrichment score
rank
K
GSE24574 Bcl6 high TFH versus naive CD4 T cell up
TXNIP, JUNB, HIPK2, RPS14, ZBTB25, GABARAPL1, RPL14, ICAM3, GPSM2, MAP4K2
P <0.05 in suppression
Enrichment score
rank

Figure S3. The epigenetic and transcriptional programs of proliferating cells during HIV-1 infection, related to Figure 1. (**A**–**B**) Proliferation module score in viremia (n = 25,778 cells), viral suppression (56,771 cells), and uninfected conditions (10,660 cells). *P* values were determined by Wilcoxon rank-sum test. (**C**) Volcano plot indicating transcription factor binding motifs having increased chromatin accessibility as measured by mean differences in chromVAR bias-corrected deviations (Z-score) between cells in the proliferating cluster (red, n = 2,727 cells) and cells not in proliferating cluster (blue, n = 90,482 cells) (FDR-adjusted $P < 0.05$ and mean Z-score difference ≥ 0.3). (**D**) Transcription factor footprints of STAT1:STAT2, IRF4, IRF7, built using normalized ATAC reads in chromatin accessible peaks called by MACS3, TF-binding sequence motifs from the JASPAR2022 reference database and reference genome hg38. STAT1:STAT2, IRF4, IRF7 binding motifs had greater observed/expected chromatin accessibility in the proliferating cluster than other clusters (FDR-adjusted $P < 0.05$). (**E**) ATAC peaks (differentially accessible chromatin regions) highly accessible in cells from the proliferating cluster grouped by conditions, represented in heatmap as comparisons of cell group means in normalized and scaled expression of peaks called by MACS3. All features passed $P < 0.05$, min.pct ≥ 0.05, $\log_2$FC ≥ 0.15 with latent variable = total number of in-peak counts. Statistical significance for all heatmaps was determined by Wilcoxon rank-sum test for comparisons between each group and all cells in the other two groups; all *P* values were FDR-adjusted. (**F**–**G**) Top enriched Gene Ontology biological processes (GO:BP) determined using transcriptionally upregulated genes ($P < 0.05$) in the proliferating cluster in viremia (**F**) and in viral suppression (**G**). All Gene Ontology pathways passed $P < 0.01$ by Fisher's exact test. (**H**–**K**) To identify immune pathways correlated with cells in the proliferating cluster in viremia and viral suppression conditions, all 36,601 genes were ranked by $\log_2$ fold change in normalize gene expression between viremic and suppressed cell groups. Significantly enriched pathways were identified using GSEA with a total of 111,635 reference gene sets from H: hallmark, C2: curated gene sets, and C7: immunologic signature gene sets. ** $P < 0.01$, *** $P < 0.001$.

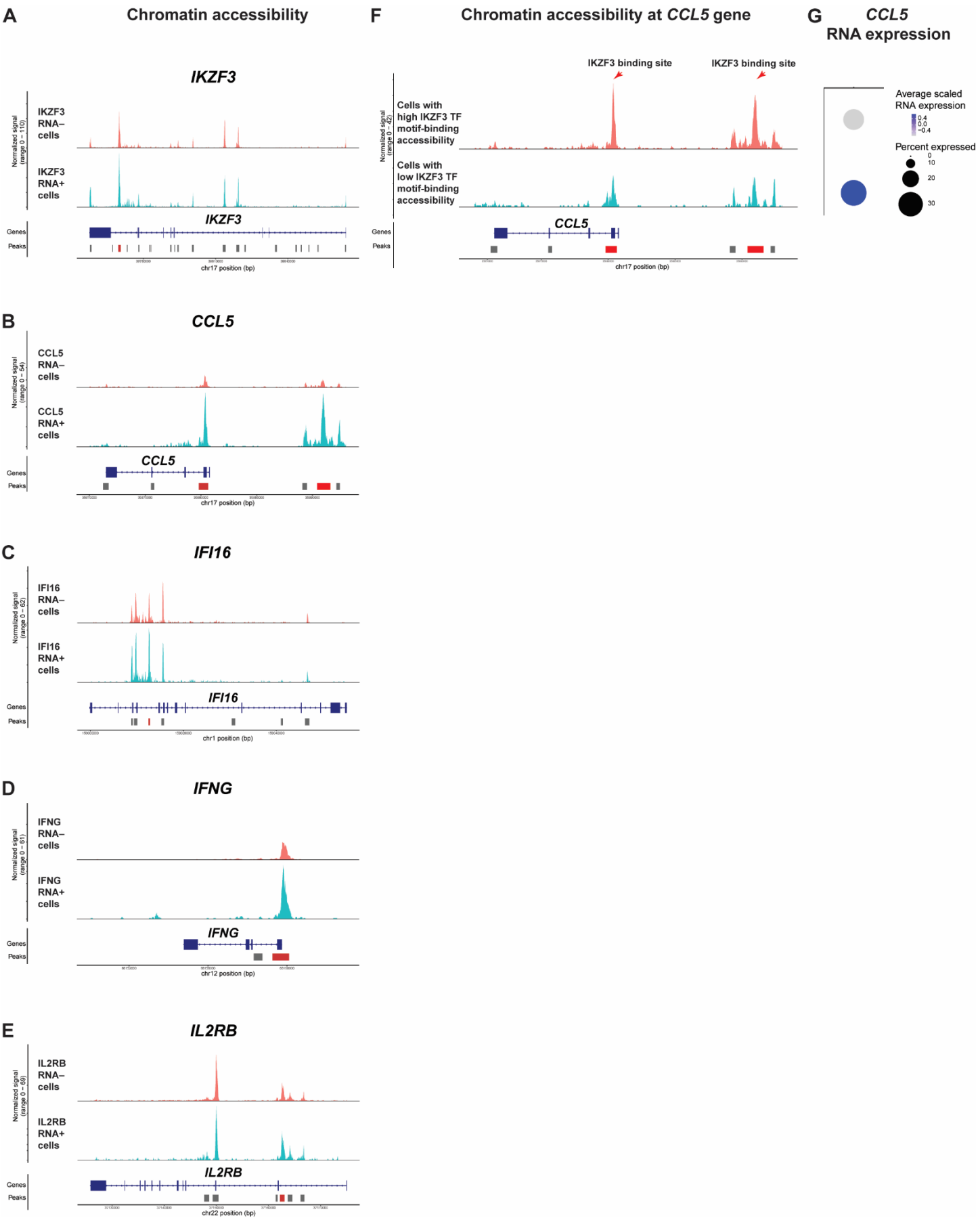
A
Chromatin accessibility
IKZF3
Normalized signal (range 0 – 110)
IKZF3 RNA– cells
IKZF3 RNA+ cells
Genes
Peaks
chr17 position (bp)
B
CCL5
Normalized signal (range 0 – 54)
CCL5 RNA– cells
CCL5 RNA+ cells
chr17 position (bp)
C
IFI16
Normalized signal (range 0 – 62)
IFI16 RNA– cells
IFI16 RNA+ cells
chr1 position (bp)
D
IFNG
Normalized signal (range 0 – 51)
IFNG RNA– cells
IFNG RNA+ cells
chr12 position (bp)
E
IL2RB
Normalized signal (range 0 – 69)
IL2RB RNA– cells
IL2RB RNA+ cells
chr22 position (bp)
F
Chromatin accessibility at CCL5 gene
IKZF3 binding site
IKZF3 binding site
Cells with high IKZF3 TF motif-binding accessibility
Cells with low IKZF3 TF motif-binding accessibility
chr17 position (bp)
G
CCL5 RNA expression
Average scaled RNA expression
Percent expressed

Figure S4. RNA expression of *IKZF3, CCL5*, *IFI16*, *IFNG*, and *IL2RB* are epigenetically regulated, related to Figure 1. (A–E) Increased chromatin accessibility at gene coordinates highlighted in red (corresponding to significantly differentially accessible peaks identified in Figure 1I) in cells in Proliferating cluster that are (**A**) *IKZF3*$^{+}$ versus *IKZF3*$^{–}$, (**B**) *CCL5*$^{+}$ vs *CCL5*-, (**C**) *IFI16*$^{+}$ vs *IFI16*-, (**D**) *IFNG*$^{+}$ vs *IFNG*-, and (**E**) *IL2RB*$^{+}$ vs *IL2RB*-. (**F–G**) Matching increase in chromatin accessibility at differentially accessible peaks (peaks identified in Figure 1I) at *CCL5* locus (**F**) and average normalized and scaled *CCL5* expression (**G**) in proliferating cells having high IKZF3 transcription factor motif-binding accessibility (chromVAR > 0) versus proliferating cells having low IKZF3 transcription factor motif-binding accessibility (chromVAR < 0).

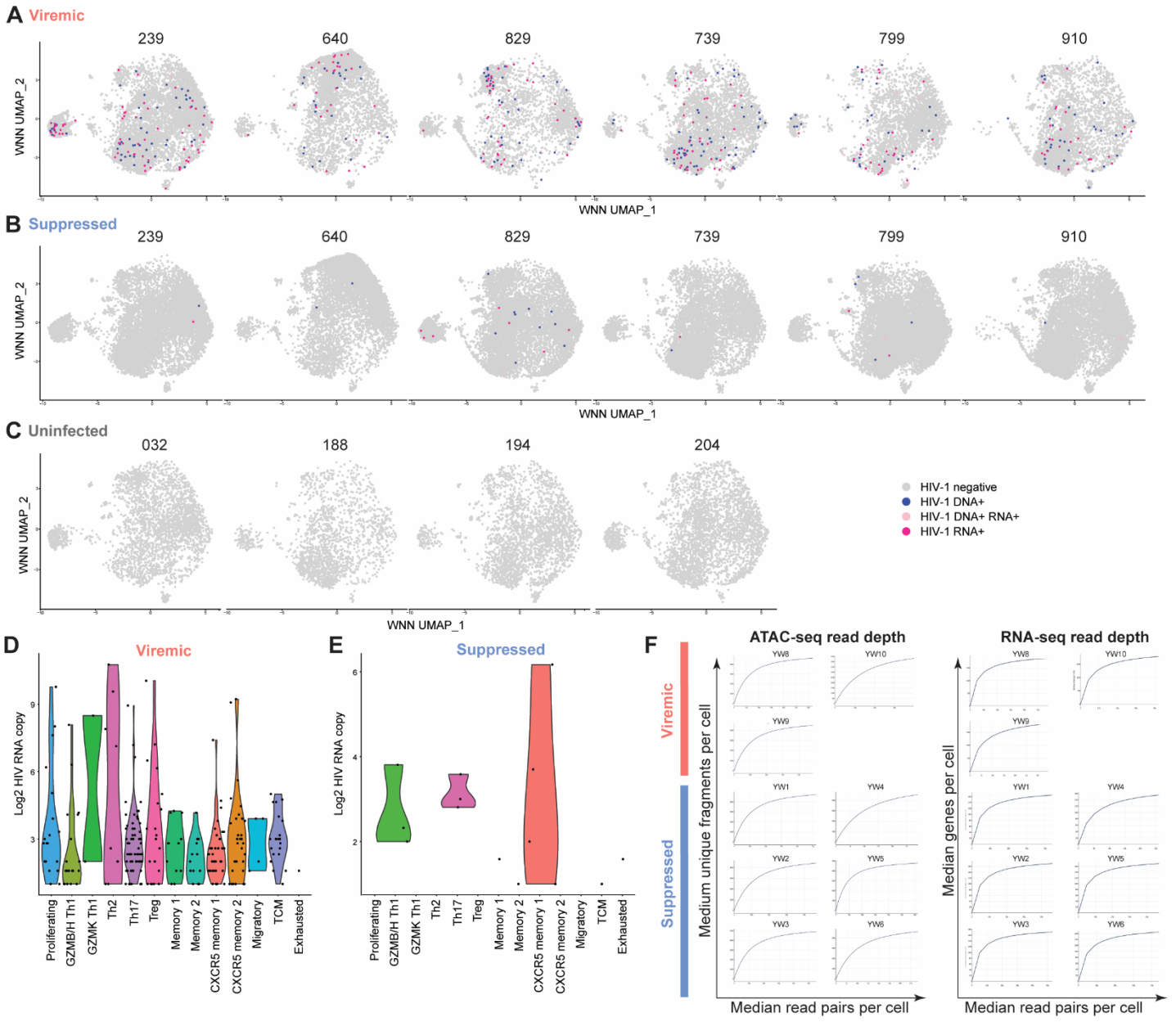
A Viremic
239
640
829
739
799
910
WNN UMAP_2
WNN UMAP_1
B Suppressed
239
640
829
739
799
910
WNN UMAP_2
WNN UMAP_1
C Uninfected
032
188
194
204
WNN UMAP_2
WNN UMAP_1
HIV-1 negative
HIV-1 DNA+
HIV-1 DNA+ RNA+
HIV-1 RNA+
D
Viremic
Log2 HIV RNA copy
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory 1
Memory 2
CXCR5 memory 1
CXCR5 memory 2
Migratory
TCM
Exhausted
E
Suppressed
Log2 HIV RNA copy
F
ATAC-seq read depth
RNA-seq read depth
Viremic
Suppressed
Medium unique fragments per cell
Median genes per cell
Median read pairs per cell
Median read pairs per cell
YW8
YW10
YW9
YW1
YW4
YW2
YW5
YW3
YW6
G HIV-1 DNA mapping perparticipant, viremia
I HIV-1 RNA mapping perparticipant, viremia
H HIV-1 DNA mapping perparticipant, viral suppression
J HIV-1 RNA mapping perparticipant, viral suppression
HXB2
U3
RU5
gag
pol
vif
vpu
tat
rev
nef
R U5
236
640
829
739
799
910

Figure S5. The HIV-1 reservoir is heterogeneous in each participant, related to Figure 2. (**A–C**) Per-participant distribution of HIV-1-infected cells visualized in WNN UMAP projection, in (**A**) viremia, (**B**) viral suppression, and (**C**) uninfected conditions. (**D–E**) Distributions of HIV-1 RNA read copy numbers in HIV-1 RNA^+ cells by cell subset, in (**D**) viremia and (**E**) viral suppression. (**F**) CellRanger sequencing saturation plots per 10x run of ATAC-seq and RNA-seq on viremic and virally suppressed samples. (**G–J**) Integrative genomics viewer (IGV) plots of per-participant HIV-1 DNA^+ reads in viremia (**G**) and in viral suppression (**H**), and per-participant HIV-1 RNA^+ reads in viremia (**I**) and in viral suppression (**J**), mapped to HXB2 reference genome and autologous HIV-1 sequences.

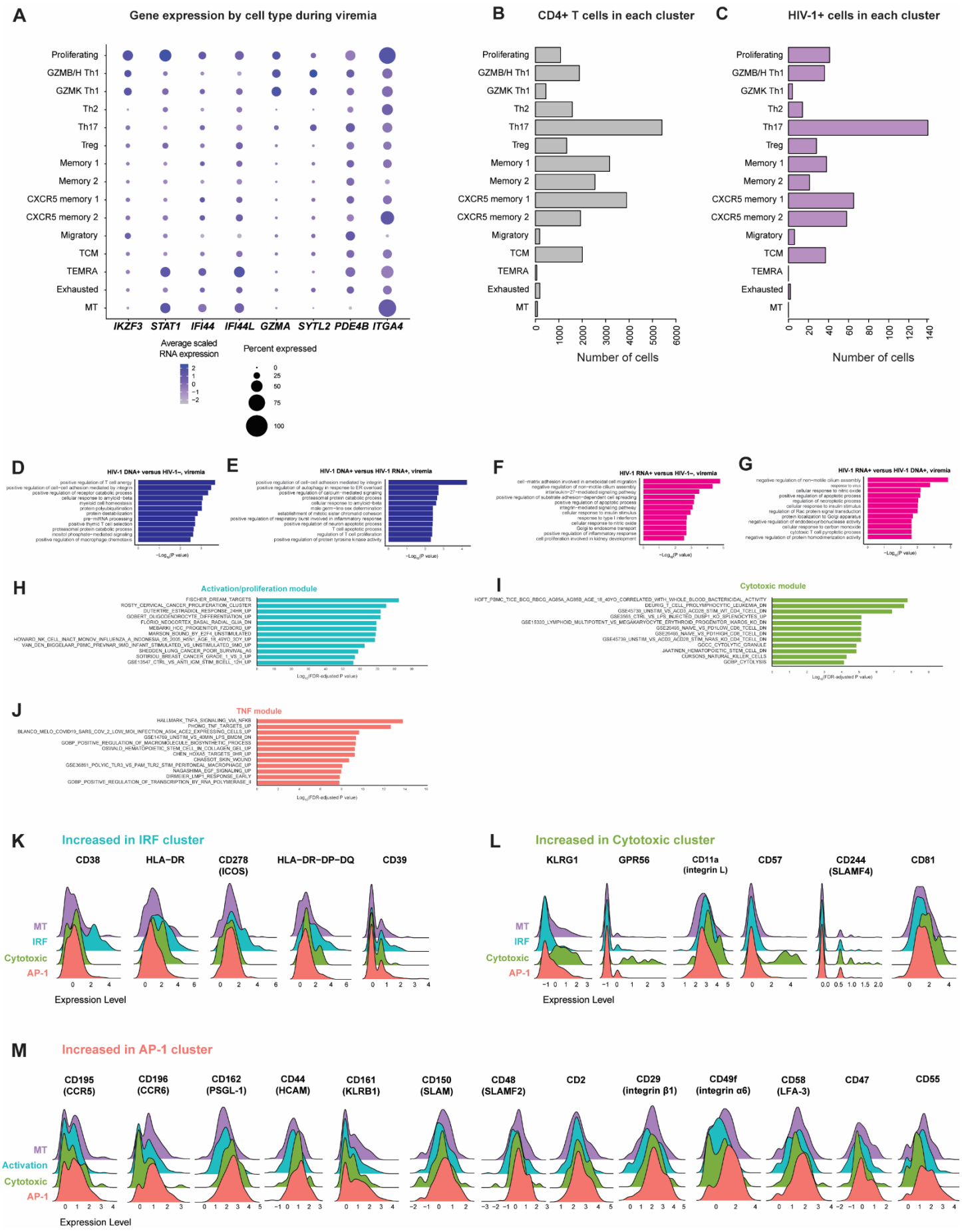
A
Gene expression by cell type during viremia
Proliferating
GZMB/H Th1
GZMK Th1
Th2
Th17
Treg
Memory 1
Memory 2
CXCR5 memory 1
CXCR5 memory 2
Migratory
TCM
TEMRA
Exhausted
MT
IKZF3 STAT1 IFI44 IFI44L GZMA SYTL2 PDE4B ITGA4
Average scaled RNA expression
Percent expressed
B
CD4+ T cells in each cluster
Number of cells
C
HIV-1+ cells in each cluster
Number of cells
D
HIV-1 DNA+ versus HIV-1−, viremia
E
HIV-1 DNA+ versus HIV-1 RNA+, viremia
F
HIV-1 RNA+ versus HIV-1−, viremia
G
HIV-1 RNA+ versus HIV-1 DNA+, viremia
H
Activation/proliferation module
I
Cytotoxic module
J
TNF module
K
Increased in IRF cluster
CD38
HLA−DR
CD278 (ICOS)
HLA−DR−DP−DQ
CD39
MT
IRF
Cytotoxic
AP-1
Expression Level
L
Increased in Cytotoxic cluster
KLRG1
GPR56
CD11a (integrin L)
CD57
CD244 (SLAMF4)
CD81
M
Increased in AP-1 cluster
CD195 (CCR5)
CD196 (CCR6)
CD162 (PSGL-1)
CD44 (HCAM)
CD161 (KLRB1)
CD150 (SLAM)
CD48 (SLAMF2)
CD2
CD29 (integrin β1)
CD49f (integrin α6)
CD58 (LFA-3)
CD47
CD55
Activation

Figure S6. Transcriptional and protein profile of HIV-1-infected cells, related to Figures 4, 5, and 6. (**A**) Per-cell subset average normalized and scaled RNA expression of eight key genes (that were identified in viremic HIV-1-infected cells) in all cells in viremia. (**B**) Total number of cells identified per cell subset in viremia. (**C**) Total number of HIV-1-infected cells identified per cell subset in viremia. (**D–G**) Top enriched Gene Ontology biological processes (GO:BP) determined using genes upregulated in HIV-1 DNA^+ cells (**D**–**E**) and in HIV-1 RNA^+ cells (**F**–**G**). In the four distinct HIV-1-infected cell states, (**H**–**J**) Top enriched Gene Ontology immune pathways (MSigDB H, C2, GO, C7 collections) determined using genes identified in the three WGCNA modules: (**H**) Proliferation gene module (gene list shown in Figure 6G), (**I**) Cytotoxic gene module (gene list shown in Figure 6H), and (**K**) AP-1 gene module (gene list shown in Figure 6I). All hypergeometric $P < 0.05$ were corrected for multiple comparisons using the Benjamini-Hochberg procedure. (**L**–**N**) Ridge plots of DSB-normalized and scaled protein expression distributions for significantly highly expressed surface proteins (shown in Figure 6F; FDR-adjusted $P < 0.05$) identified, including markers highly expressed in the IRF cluster (**L**), Cytotoxic cluster (**M**), and AP-1 cluster (**N**).

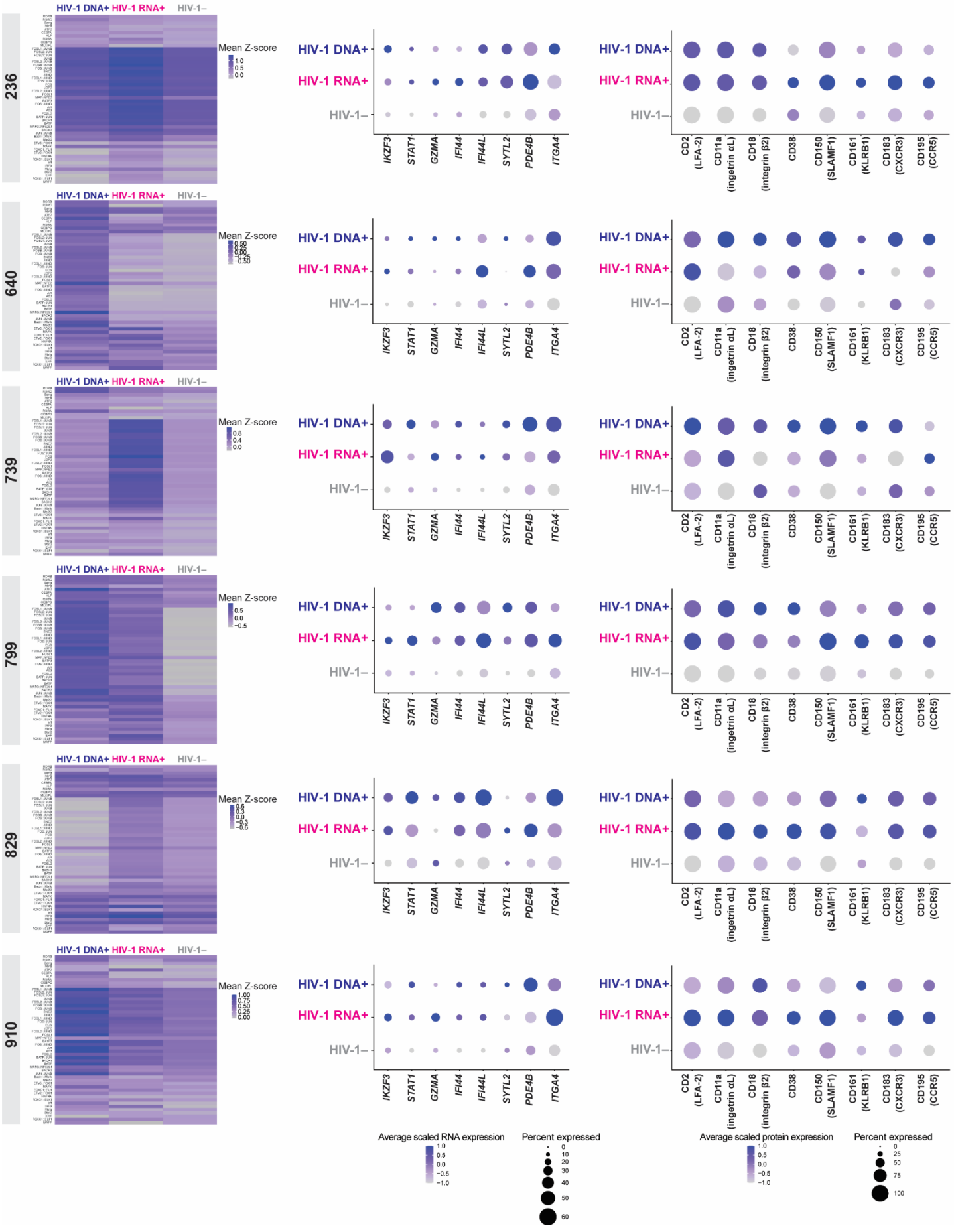
A ATAC-seq (transcription factor)
B RNA-seq (RNA expression)
C ADT (protein expression)
HIV-1 DNA+
HIV-1 RNA+
HIV-1−
Mean Z-score
236
640
739
799
829
910
IKZF3
STAT1
GZMA
IFI44
IFI44L
SYTL2
PDE4B
ITGA4
CD2 (LFA-2)
CD11a (ingetrin αL)
CD18 (integrin β2)
CD38
CD150 (SLAMF1)
CD161 (KLRB1)
CD183 (CXCR3)
CD195 (CCR5)
Average scaled RNA expression
Percent expressed
Average scaled protein expression
Percent expressed

Figure S7. The epigenetic, transcriptional, and protein profiles of the HIV-1 reservoir do not differ among the six HIV-1$^+$ participants, related to Figures 3, 4, and 5. (**A**) Per-participant global chromatin accessibility of transcription factor binding motifs that were identified to have enriched accessibility in pooled viremic HIV-1 DNA$^+$ cells or viremic HIV-1 RNA$^+$ cells (shown in Figure 3F), represented in heatmap as comparisons of group means in chromVAR bias-corrected deviations (Z-score). (**B**) Per-participant average normalized and scaled RNA expression for eight key genes that were determined to be significantly upregulated in both pooled viremic HIV-1 DNA$^+$ cells and HIV-1 RNA$^+$ cells (shown in Figure 4A, 4B). (**C**) Per-participant average CLR-normalized and scaled protein expression for proteins that were determined to be significantly upregulated in both pooled viremic HIV-1 DNA$^+$ cells and HIV-1 RNA$^+$ cells (shown in Figure 5A, 4B). All per-participant comparisons are made between HIV-1 DNA$^+$ cells, HIV-1 RNA$^+$ cells, and HIV-1- cells. *P* values did not reach statistical significance for differential expression between conditions due of the low number of HIV-1$^+$ cells captured per-participant (see Table S3 for infected cell count per participant).

Table S1. Clinical characteristics of study participants, related to Figure 1.

Participant ID	Age	Sex	Ethnicity	ART	Peak viral load (copies/ml)	Duration since EDDI at viremia (days)	Duration of ART during viral suppression (days)	CD4 count at viremia (/µl)	CD4 count at viral suppression (/µl)
People living with HIV-1*									
236	22	M	Hispanic	EFV/FTC/TDF	7,117,757	24	431	625	712
640	26	M	Hispanic	EVG/c/TDF/FTC	6,023,064	22	336	586	772
829	21	M	Hispanic	EFV/TDF/FTC	1,205,138	48	329	388	796
739	23	M	Hispanic	EFV/TDF/3TC	8,604,710	41	342	467	821
799	31	M	Hispanic	EFV/TDF/3TC	13,313,300	20	329	1,203	1,464
910	24	M	Hispanic	EFV/AZT/3TC	1,257,674	33	338	375	371
People living without HIV-1**									
032	52	M	Hispanic						
188	29	M	Hispanic						
194	35	M	Hispanic						
204	28	M	Hispanic						

3TC, lamivudine; AZT, azidothymidine; /c, cobicistat; EFV, efavirenz; EVG, elvitegravir; FTC, emtricitabine; TDF, tenofovir disoproxil; EDDI, estimated dates of detectable infection.

*These are the same study participants in the Sabes cohort described in Collora *et al.*, Immunity 2022

**HIV-1 infection status was determined by a point-of-care third-generation HIV immunoassay described in Lama *et al.*, Clinical Infectious Diseases 2021

Methods S1. DOGMA-seq protocol modifications, related to STAR Methods.

Yulong Wei, Ya-Chi Ho

- Ensure having all required primers prior to starting Step 3 of the multiome protocol.

Step 1–3 of the 10X multiome protocol (10X documentation GC000338 Rev A) should be followed without modification.

PREAMPLIFICATION PCR (Step 4)

Complete step 4 with only these modification

In step 4.1: Add 1 µl of 0.2 µM of ADT and HTO primers to the preamplification mix.

1X	
50 µl	P-A Mix
4 µl	P-A Primer
1 µl	0.2 µM ADT
1 µl	0.2 µM HTO
56 µL	Mastermix per sample

In step 4.3.k Elute beads in 100.5 µl of Buffer EB

In step 4.3.o Recover 100 µl postamplification product

ATAC AMPLIFICATION (Step 5)

In step 5.1 Use only 25 µl of post amplification product for ATAC sample index PCR.
Adjust sample index PCR mix accordingly:

- Amp Mix: 31.3 µl per sample
- Si-PCR Primers: 4.7 µl per sample

In step 5.1.b Add 35.6 µl sample index PCR mix to 25 µl post amplification product for ATAC sample index PCR.

In step 5.1.c Add 1.6 µl of sample index N set A for each well.
Record assignment, pipette mix and centrifuge briefly.

In step 5.2.a Add 37.5 µl SPRI select (0.6X) to each sample and mix.
Incubate for 5 minutes, then place on magnet.
Transfer 92 µl of supernatant to new strip.

In step 5.2.e Add 59.4 µl SPRI select (1.55X) to each sample and mix.

cDNA AMPLIFICATION (Step 5)

Complete without modification.

To complete the CITE/HTO amplification (not in 10X protocol):

Avoid reusing indices to maintain Unique Dual Indexing.

Prepare a PCR reaction mix:

50 µl	2x KAPA mix
2.5 µl	SI-PCR 10 µM – use the indexed ones provided
2.5 µl	RPxx (TSA family) and D7xx (TSA hashing) mix- provided
10 µl	H_2O
65 µl	Total

Note the index used for each sample (both the SI PCR and the RP/D7 index number).

Add 65 µl to 35 µl of pre-amplified sample

Cycle samples according to these conditions:

95℃ 3 min
95℃ 20 sec |
60℃ 30 sec | 11 total cycles
72℃ 20 sec |
72℃ 5 min
4℃ hold

Expected product is 190 bp in size.

Isolate library using the following 2.0X SPRI protocol:

1. Add 200 µl SPRI (2X) to PCR reaction (transfer to 1.5 ml Eppendorf tube if necessary) and mix.
2. Incubate for 10 minutes at room temperature.
3. Place the tube on magnet and wait for solution to clear.
4. Discard the supernatant.
5. Add 400 µl of 80% freshly made ethanol without disturbing the pellet.
6. Wait for 30 seconds.
7. Discard ethanol wash.
8. Centrifuge briefly, place the tube on the magnet, and remove residual ethanol.
9. Resuspend beads in 50 µl water.
10. Add 100 µl SPRI (2X) to PCR reaction (transfer to 1.5ml Eppendorf tube if necessary) and mix.
11. Incubate for 10 minutes at room temperature.
12. Place the tube on the magnet and wait for the solution to clear.
13. Discard the supernatant.
14. Add 400ul of 80% ethanol without disturbing the pellet.
15. Wait for 30 seconds.
16. Discard ethanol wash.
17. Repeat steps 14–16 one additional time for a total of two washes.
18. Centrifuge briefly, place the tube on the magnet, and remove residual ethanol.
19. Air-dry for two minutes (avoid over-drying beads).
20. Remove from magnet and resuspend beads in 20 µl water through vigorous pipetting.
21. Incubate at room temperature for 5 minutes.
22. Place the tube on the magnet. Wait for solution to clear.
23. Transfer clear solution to new PCR tubes.
24. Perform Qubit fluorometric quantification. Dilute 1 µl accordingly. Examine library size and quantity on Bioanalyzer (or equivalent equipment).

DOGMA-seq primer sequences:

Name	Sequence	Notes
ADT	CCTTGGCACCCGAGAATT*C*C	* indicates phosphorothioate
HTO	GTGACTGGAGTTCAGACGTGTGC*T*C	* indicates phosphorothioate
SIPCR 1	AATGATACGGCGACCACCGAGAagtacgagTCTAC ACTCTTTCCCTACACGACGCTC	Dual index common primer for both
SIPCR 2	AATGATACGGCGACCACCGAGAgcatgaccTCTACA CTCTTTCCCTACACGACGCTC	Dual index common primer for both
SIPCR 3	AATGATACGGCGACCACCGAGAtgcatgctTCTACA CTCTTTCCCTACACGACGCTC	Dual index common primer for both
SIPCR 4	AATGATACGGCGACCACCGAGAgactagcaTCTAC ACTCTTTCCCTACACGACGCTC	Dual index common primer for both
RPX1	CAAGCAGAAGACGGCATACGAGATagatcgcaGTG ACTGGAGTTCCTTGGCACCCGAGAATTCCA	Dual index common primer for CITE
RPX2	CAAGCAGAAGACGGCATACGAGATtgacatagGTG ACTGGAGTTCCTTGGCACCCGAGAATTCCA	Dual index common primer for CITE
RPX3	CAAGCAGAAGACGGCATACGAGATgatcagacGTG ACTGGAGTTCCTTGGCACCCGAGAATTCCA	Dual index common primer for CITE
RPX4	CAAGCAGAAGACGGCATACGAGATttcatagcGTGA CTGGAGTTCCTTGGCACCCGAGAATTCCA	Dual index common primer for CITE
D7X1	CAAGCAGAAGACGGCATACGAGATgccatcaaGTG ACTGGAGTTCAGACGTGTGC	Dual index common primer for HTO
D7X2	CAAGCAGAAGACGGCATACGAGATtggcattaGTGA CTGGAGTTCAGACGTGTGC	Dual index common primer for HTO
D7X3	CAAGCAGAAGACGGCATACGAGATggcatacaGTG ACTGGAGTTCAGACGTGTGC	Dual index common primer for HTO
D7X4	CAAGCAGAAGACGGCATACGAGATgcgcagctGTG ACTGGAGTTCAGACGTGTGC	Dual index common primer for HTO