**Supplemental information**

**Figure S1. Study participant characteristics comparing immediate and delayed antiretroviral therapy,** Related to Figure 1.

**Figure S2. Batch effect removal by fastMNN, gene expression differences across clusters, and modules identified by WGCNA to differentiate HIV-1 viremic conditions from HIV-1 viral suppression conditions.** Related to Figure 1.

**Figure S3. Differences related to immediate or deferred initiation of ART.** Related to Figure 1, Figure 2, and Figure 3.

**Figure S4. HIV-1-specific antigen responses.** Related to Figure 3.

**Figure S5. HIV-1 RNA mapping and expression level.** Related to Figure 4.

**Figure S6. Clone phenotypes and the size of HIV-1 RNA$^{+}$ clones vs HIV-1 RNA$^{-}$ clones.** Related to Figure 5 and Figure 6.

**Figure S7. Machine learning identification of cellular markers necessary and sufficient to distinguish HIV-1 RNA$^{+}$ T cell clones from HIV-1 RNA$^{-}$ T cell clones by scRFE and flow cytometry validation for HIV-1 p24$^{+}$ cells.** Related to Figure 7.

Related to Figure 7.

**Table S1 Clinical demographics.** Related to Figure 1.

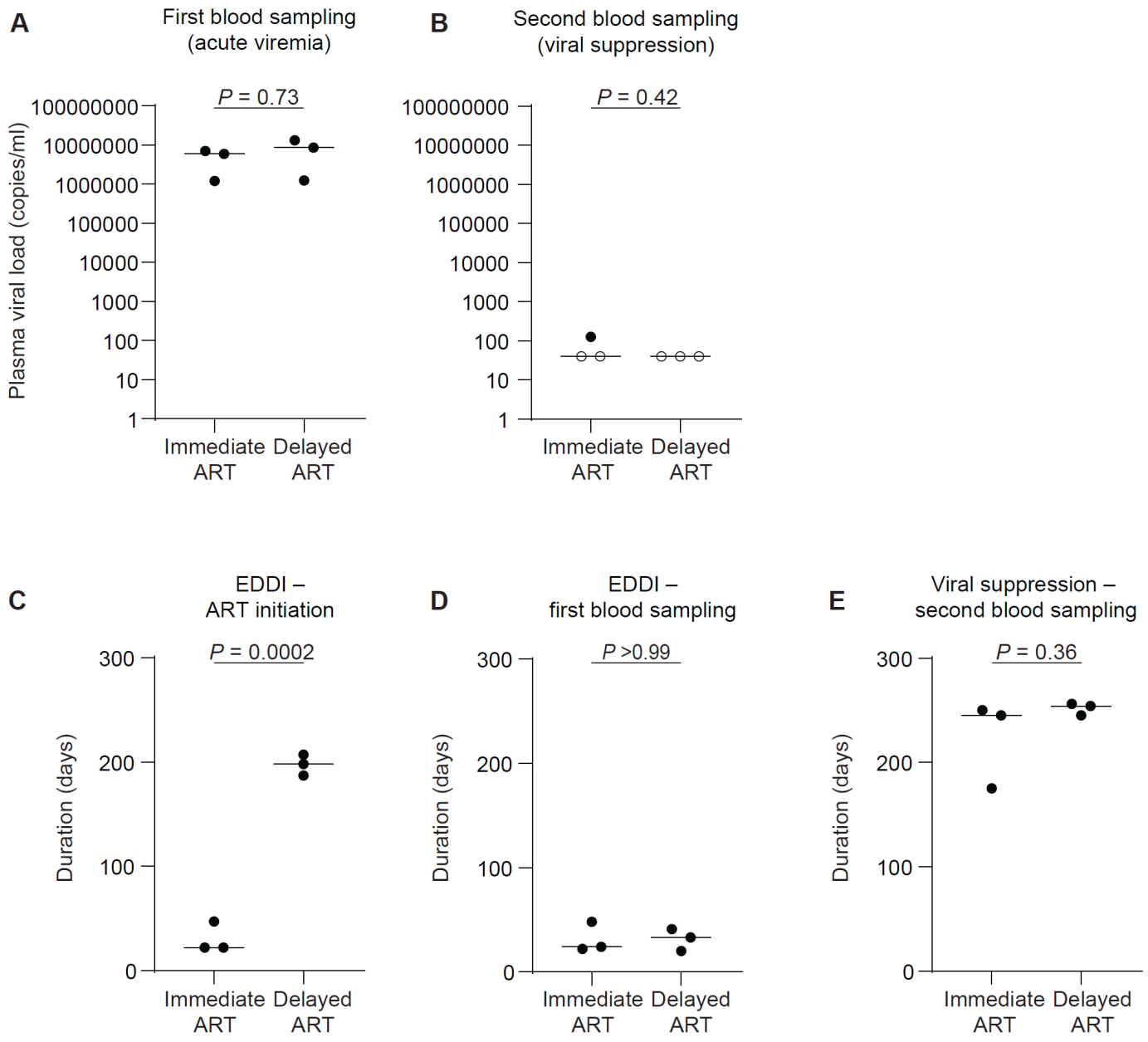



**Figure S1. Study participant characteristics comparing immediate and delayed antiretroviral therapy,** Related to Figure 1. Plasma viral loads during the **A,** viremia timepoint and after **B,** viral suppression is achieved. Number of days between the estimated date of infection and the **C,** initiation of antiretroviral therapy or **D,** the time of first blood sampling. E, The duration of viral suppression. *P* values were determined by two-tailed paired Student's *t*-test. EDDI, estimated date of detectable infection.

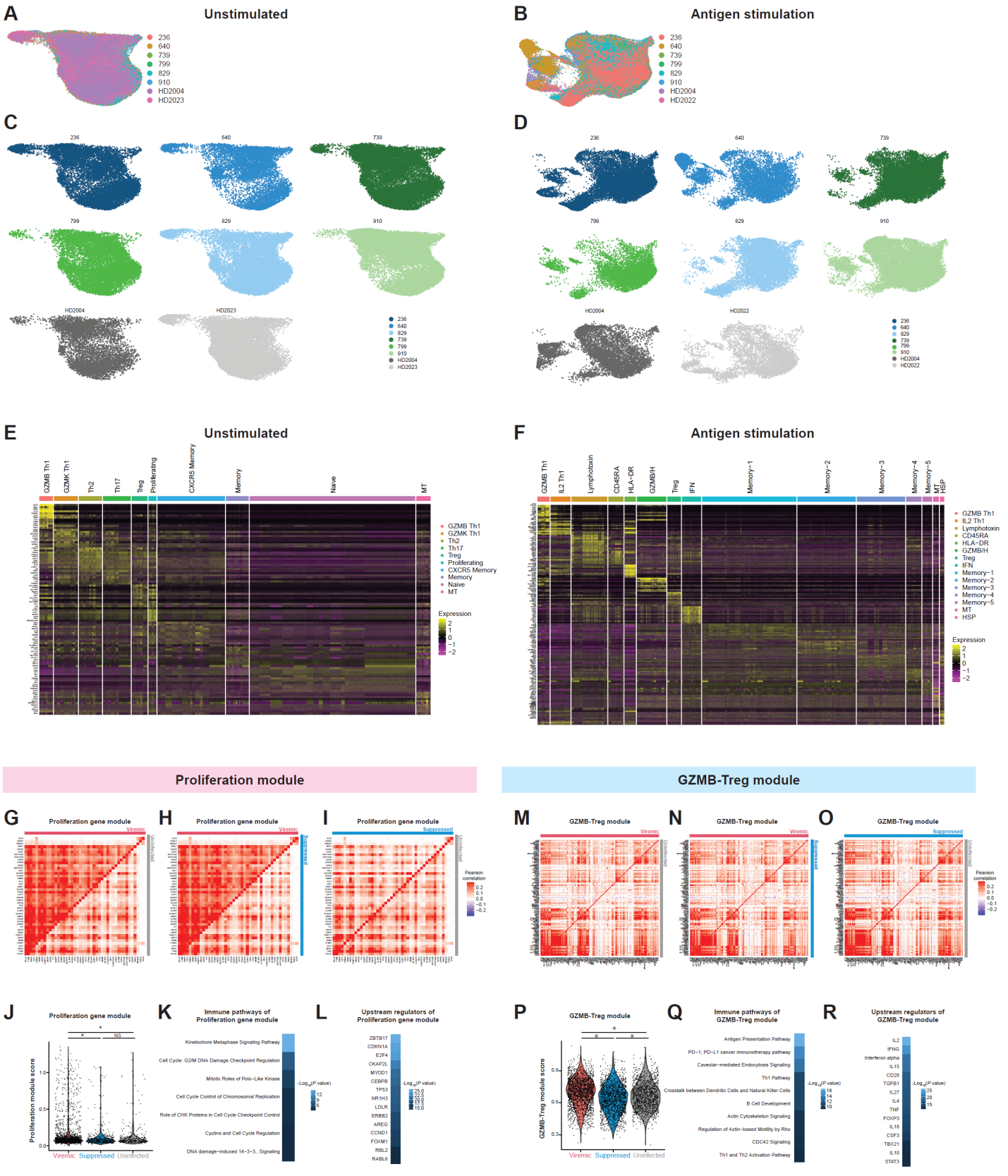



**Figure S2. Batch effect removal by FastMNN, Gene expression differences across clusters, and Modules identified by WGCNA to differentiate HIV-1 viremic conditions from HIV-1 viral suppression conditions.** Related to Figure 1.

UMAP plots showing cells colored by the participant in unstimulated condition (**A**) and antigen stimulated condition (**B**). Heatmaps showing the top 10 genes for each cluster in unstimulated condition (**C, E**) and antigen-stimulated condition (**D, F**). Of the WGCNA

defined modules, two separated specifically viremia and viral suppression. The proliferation defined module was induced by viremia relative to **G,** uninfected and **H,** virally suppressed individuals, but not in **I**, suppressed individuals relative to uninfected individuals. **J,** Increased expression of this module is restricted to proliferating cells in the viremic condition. Ingenuity pathway analysis (IPA) showing the **K,** immune pathways and **L,** upstream regulators of the proliferation gene module. The GZMB-Treg defined module was not induced by viremia relative to **M,** uninfected or **N,** virally suppressed individuals, and not in **I,** suppressed individuals relative to uninfected individuals. **O,** Increased expression of this module is restricted to Treg cells in the viremic condition relative to suppressed and uninfected, and in suppressed relative to uninfected. IPA analysis showing the **P,** immune pathways and **L,** upstream regulators of the proliferation gene module. **J** and **P**, * denotes $P < 0.05$ by Wilcoxon rank-sum test. **K, L, Q,** and **R**, Fisher's exact test with Benjamini-Hochberg procedure multiple hypothesis correction.

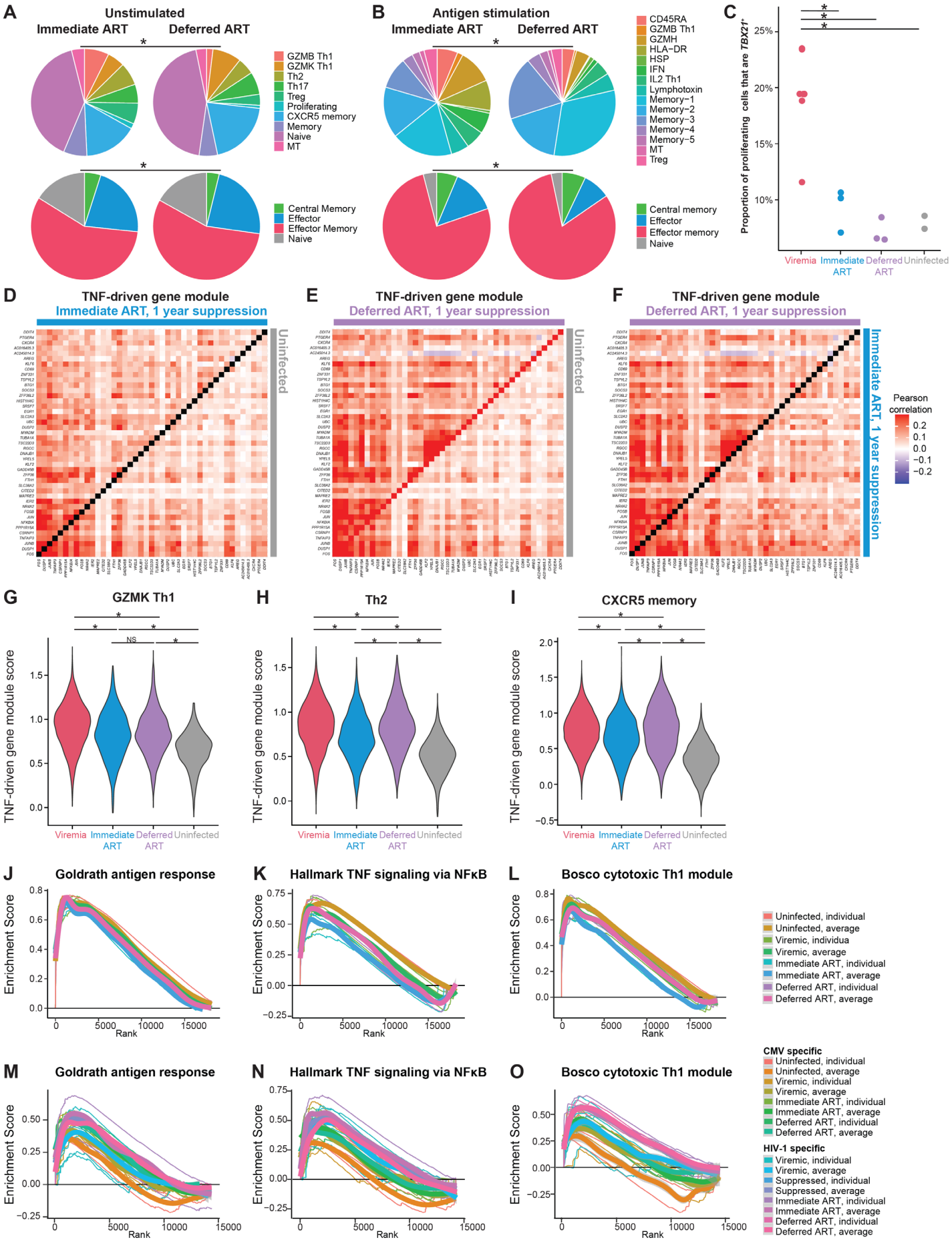
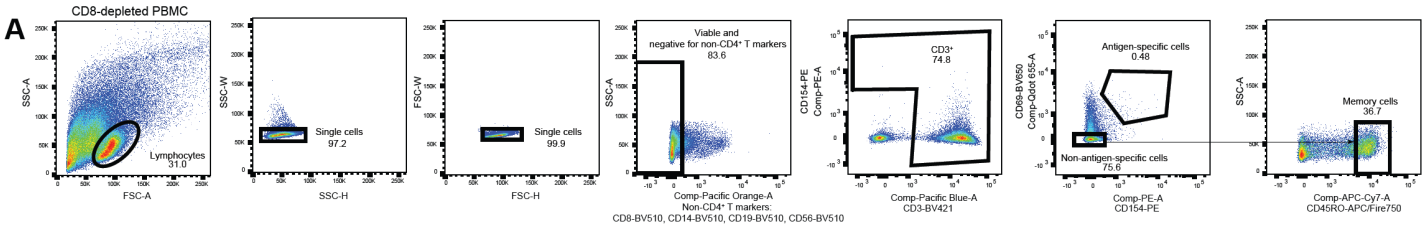
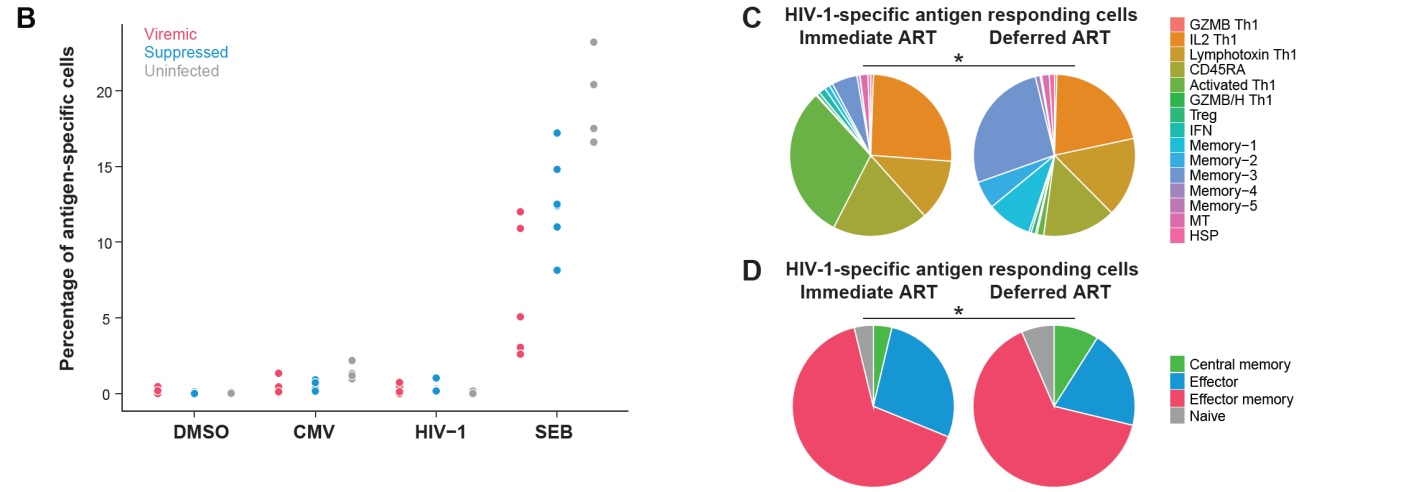
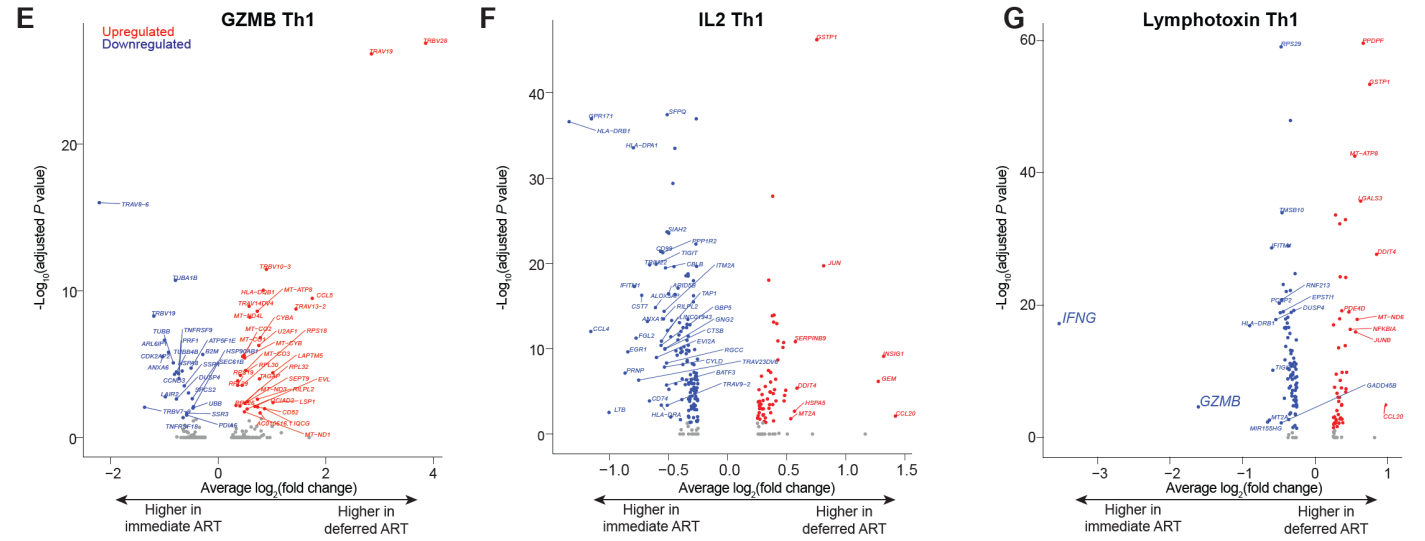

A
Unstimulated
Immediate ART
Deferred ART
GZMB Th1
GZMK Th1
Th2
Th17
Treg
Proliferating
CXCR5 memory
Memory
Naive
MT
Central Memory
Effector
Effector Memory
Naive
B
Antigen stimulation
Immediate ART
Deferred ART
CD45RA
GZMB Th1
GZMH
HLA−DR
HSP
IFN
IL2 Th1
Lymphotoxin
Memory−1
Memory−2
Memory−3
Memory−4
Memory−5
MT
Treg
Central memory
Effector
Effector memory
Naive
C
Proportion of proliferating cells that are TBX21+
25%
20%
15%
10%
Viremia
Immediate ART
Deferred ART
Uninfected
D
TNF-driven gene module
Immediate ART, 1 year suppression
Uninfected
E
TNF-driven gene module
Deferred ART, 1 year suppression
Uninfected
F
TNF-driven gene module
Deferred ART, 1 year suppression
Immediate ART, 1 year suppression
Pearson correlation
0.2
0.1
0.0
−0.1
−0.2
G
GZMK Th1
TNF-driven gene module score
NS
Viremia
Immediate ART
Deferred ART
Uninfected
H
Th2
TNF-driven gene module score
Viremia
Immediate ART
Deferred ART
Uninfected
I
CXCR5 memory
TNF-driven gene module score
Viremia
Immediate ART
Deferred ART
Uninfected
J
Goldrath antigen response
Enrichment Score
Rank
K
Hallmark TNF signaling via NFκB
Enrichment Score
Rank
L
Bosco cytotoxic Th1 module
Enrichment Score
Rank
Uninfected, individual
Uninfected, average
Viremic, individua
Viremic, average
Immediate ART, individual
Immediate ART, average
Deferred ART, individual
Deferred ART, average
M
Goldrath antigen response
Enrichment Score
Rank
N
Hallmark TNF signaling via NFκB
Enrichment Score
Rank
O
Bosco cytotoxic Th1 module
Enrichment Score
Rank
CMV specific
Uninfected, individual
Uninfected, average
Viremic, individual
Viremic, average
Immediate ART, individual
Immediate ART, average
Deferred ART, individual
Deferred ART, average
HIV-1 specific
Viremic, individual
Viremic, average
Suppressed, individual
Suppressed, average
Immediate ART, individual
Immediate ART, average
Deferred ART, individual
Deferred ART, average

**Figure S3. Differences related to immediate or deferred initiation of ART.** Related to Figure 1, Figure 2, and Figure 3.

The cluster composition and memory phenotypes as defined by surface protein expression of **A,** unstimulated and **B,** antigen stimulated conditions comparting immediate and delayed antiretroviral therapy initiation. **C,** Proportion of cells in the proliferating cell cluster expressing *TBX21* (encoding Th1 transcription factor Tbet). The module in of Figure 1F–J showing a modest induction in **D,** immediate ART, and an increased induction in deferred ART relative to **E,** uninfected and **F,** immediate ART conditions. This TNF driven module has equivalent expression in immediate and deferred ART conditions in **G,** GZMK Th1, but increased expression in deferred ART conditions in **H,** Th2 and **I,** CXCR5 memory cells. Gene sets identified by GSEA correlating with clone size in unstimulated conditions include: **J,** Goldrath antigen response, **K,** Hallmark TNF signaling via NFκB, and **L,** Bosco cytotoxic Th1 module. Gene sets identified by GSEA correlating with clone size in antigen stimulated conditions include: **M,** Goldrath antigen response, **N,** Hallmark TNF signaling via NFκB, and **O,** Bosco cytotoxic Th1 module. **A** and **B,** $P < 0.05$, Fisher's exact test. **C, G, H,** $P < 0.05$, Wilcoxon rank-sum test.


A
CD8-depleted PBMC
SSC-A
FSC-A
Lymphocytes
31.0
SSC-W
SSC-H
Single cells
97.2
FSC-W
FSC-H
Single cells
99.9
SSC-A
Viable and
negative for non-CD4+ T markers
83.6
Comp-Pacific Orange-A
Non-CD4+ T markers:
CD8-BV510, CD14-BV510, CD19-BV510, CD56-BV510
CD154-PE
Comp-PE-A
CD3+
74.8
Comp-Pacific Blue-A
CD3-BV421
CD69-BV650
Comp-Qdot 655-A
Antigen-specific cells
0.48
Non-antigen-specific cells
75.6
Comp-PE-A
CD154-PE
SSC-A
Memory cells
36.7
Comp-APC-Cy7-A
CD45RO-APC/Fire750


B
Viremic
Suppressed
Uninfected
Percentage of antigen-specific cells
DMSO
CMV
HIV−1
SEB
C
HIV-1-specific antigen responding cells
Immediate ART
Deferred ART
*
GZMB Th1
IL2 Th1
Lymphotoxin Th1
CD45RA
Activated Th1
GZMB/H Th1
Treg
IFN
Memory−1
Memory−2
Memory−3
Memory−4
Memory−5
MT
HSP
D
HIV-1-specific antigen responding cells
Immediate ART
Deferred ART
*
Central memory
Effector
Effector memory
Naive


E
GZMB Th1
Upregulated
Downregulated
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART
F
IL2 Th1
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART
G
Lymphotoxin Th1
IFNG
GZMB
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART

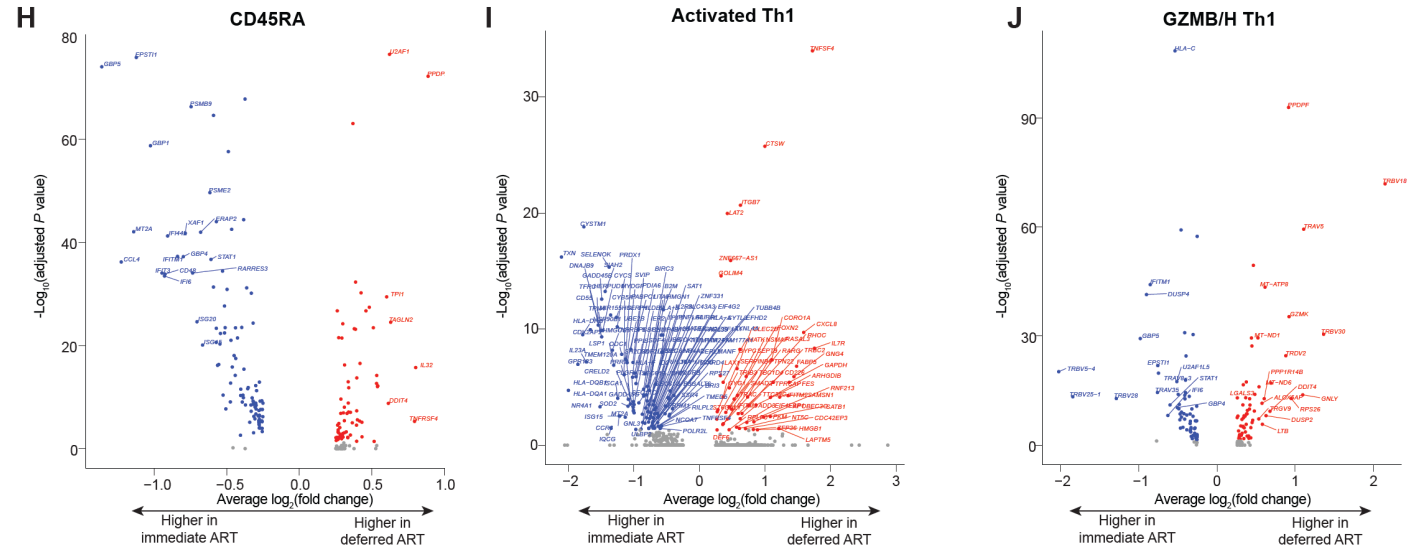

H
CD45RA
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART
I
Activated Th1
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART
J
GZMB/H Th1
-Log10(adjusted P value)
Average log2(fold change)
Higher in immediate ART
Higher in deferred ART

**Figure S4. HIV-1-specific antigen responses.** Related to Figure 3.

**A**, $CD8^+$ T cells were depleted from PBMC using magnetic depletion before antigen stimulation. CD8-depleted PBMC were gated on flow cytometry to select lymphocytes, single cells, cells that are not $CD4^+$ T cells (CD8, CD14, CD19, CD56 negative), and CD3 positive. Antigen-specific cells were defined by CD69 and CD154 double positive expression. Memory cells were defined as CD69 and CD154 double negative and CD45RO positive populations. Representative plots from participant 910 HIV-1 in antigen stimulated condition were shown. **B**, The frequency of CMV-specific and HIV-1-specific cells. SEB, Staphylococcal enterotoxin B. **C,** Cluster distribution and **D,** Memory distribution differences between suppressed immediate and suppressed delayed individuals. Differential expressional analysis comparing deferred suppression to immediate suppression cells which are HIV-1specific within **E,** GZMB Th1, **F,** IL2 Th1, **G,** Lymphotoxin Th1, **H,** CD45RA, **I,** Activated Th1, and **J,** GZMB/H Th1 cell clusters. **C** and **D**, *, $P < 0.05$, Fisher's exact test. **E–J,** Wilcoxon Rank-Sum test with Benjamini-Hochberg Procedure multiple hypothesis correction.

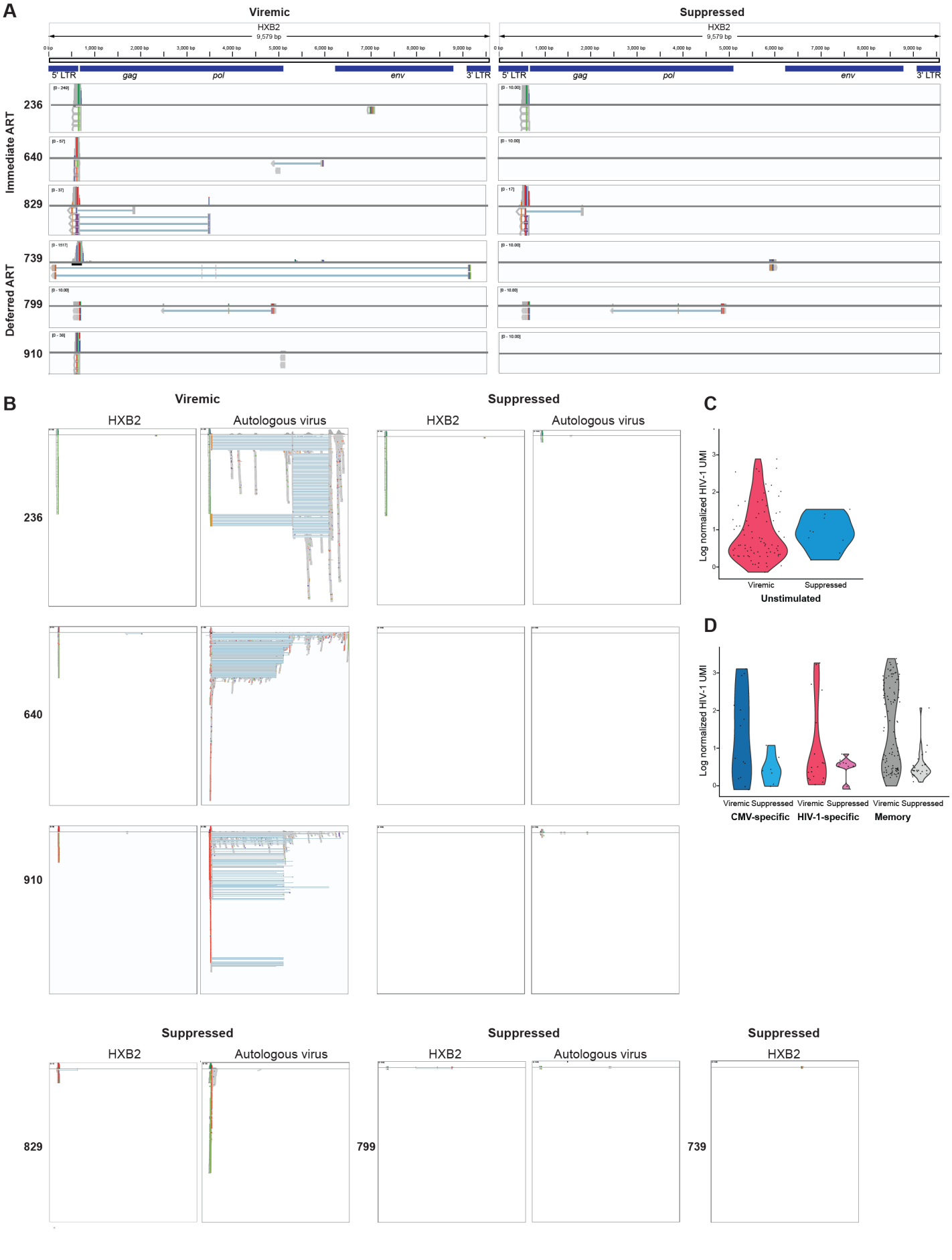

A
Viremic
Suppressed
HXB2
9,579 bp
5' LTR
gag
pol
env
3' LTR
Immediate ART
Deferred ART
236
640
829
739
799
910
B
Viremic
Suppressed
HXB2
Autologous virus
C
Log normalized HIV-1 UMI
Viremic
Suppressed
Unstimulated
D
Log normalized HIV-1 UMI
Viremic Suppressed
CMV-specific
HIV-1-specific
Memory

**Figure S5. HIV-1 RNA mapping and expression level.** Related to Figure 4.

Integrative genomics viewer (IGV) plots of **A,** all six participants to HXB2 both during viremia and suppression and **B,** alignments of select participants comparing HXB2 alignments to alignments to infected individual autologous viral sequences. Violin plot showing the level of HIV-1 RNA expression in HIV-1 RNA$^{+}$ cells in the unstimulated condition during **C,** viremia and after viral suppression and **D,** in antigen stimulated condition during viremia and after viral suppression.

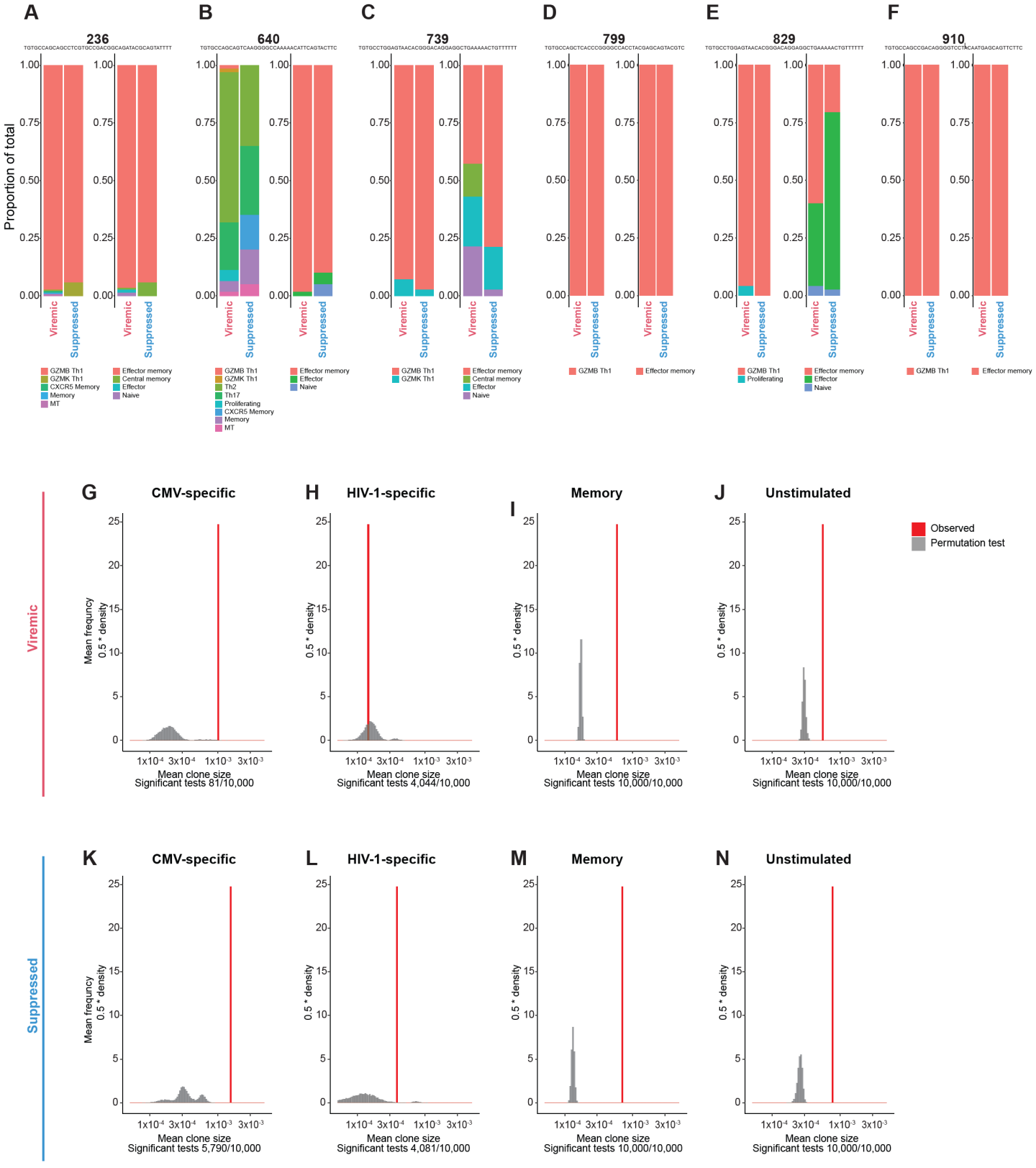



**Figure S6. Clone phenotypes and the size of HIV-1 RNA$^+$ clones vs HIV-1 RNA$^-$ clones.** Related to Figure 5 and Figure 6. **A–F,** stacked bar graphs showing the differences in cluster membership and memory phenotype between clone members in viremia and suppression in six example clones. Histograms showing the permutation mean clone size (grey) and the observed HIV-1 RNA$^+$ clone mean size (red) for **A,** CMV-specific clones during viremia, **B,** HIV-specific clones during viremia, **C,** clones for the sorted memory

condition during viremia, **D,** clones from the unstimulated condition during viremia, **E,** CMV-specific clones during viral suppression, **F,** HIV-specific clones during viral suppression, **G,** clones for the sorted memory condition during viral suppression, and **H,** clones from the unstimulated condition during viral suppression, $P < 0.05$ in a one-sided (greater) Wilcoxon rank sum test was considered significant**.**

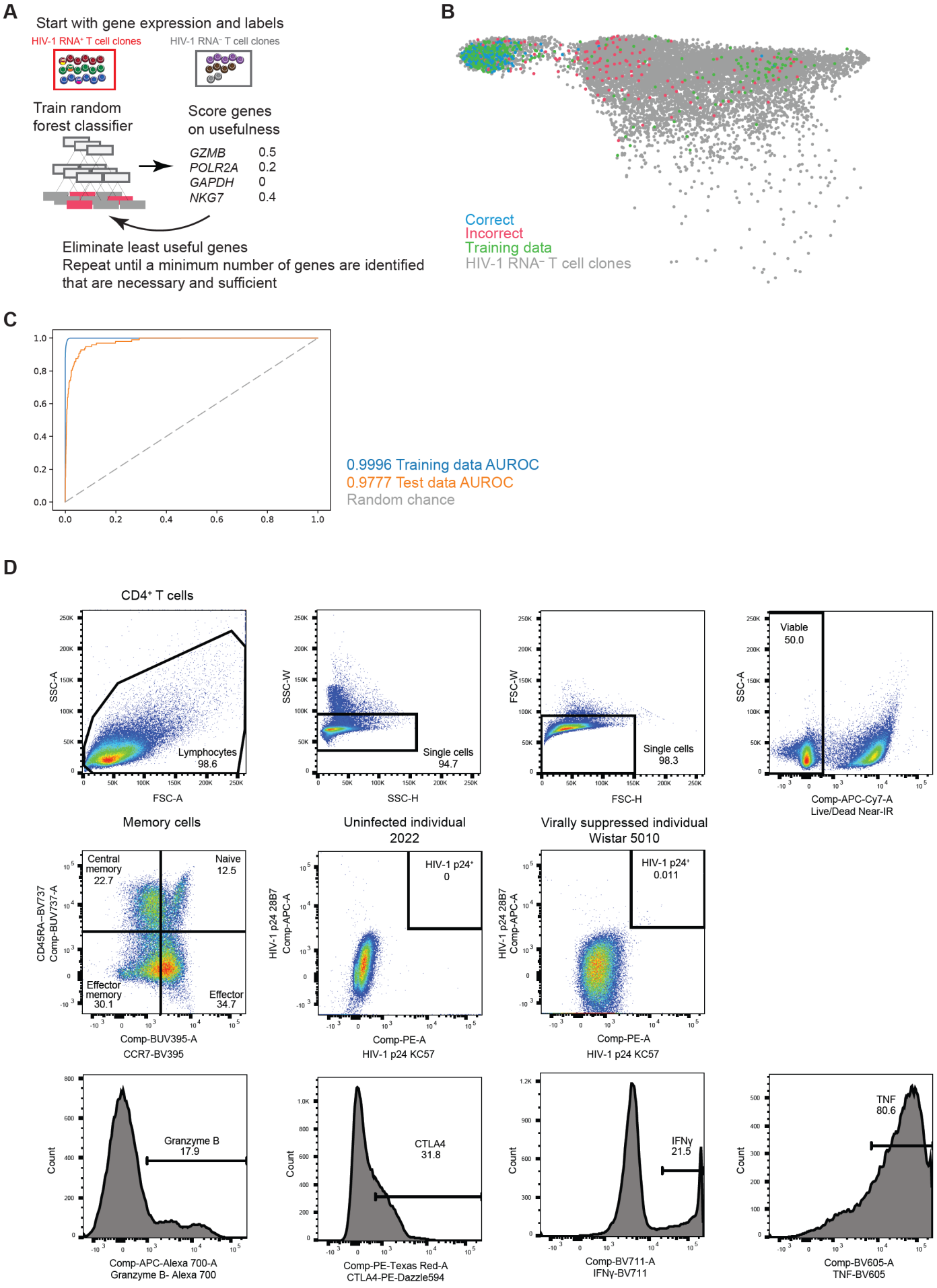

A
Start with gene expression and labels
HIV-1 RNA+ T cell clones
HIV-1 RNA- T cell clones
Train random forest classifier
Score genes on usefulness
GZMB 0.5
POLR2A 0.2
GAPDH 0
NKG7 0.4
Eliminate least useful genes
Repeat until a minimum number of genes are identified that are necessary and sufficient
B
Correct
Incorrect
Training data
HIV-1 RNA- T cell clones
C
0.9996 Training data AUROC
0.9777 Test data AUROC
Random chance
D
CD4+ T cells
SSC-A
FSC-A
Lymphocytes 98.6
SSC-W
SSC-H
Single cells 94.7
FSC-W
FSC-H
Single cells 98.3
Viable 50.0
Comp-APC-Cy7-A Live/Dead Near-IR
Memory cells
CD45RA--BV737 Comp-BUV737-A
Central memory 22.7
Naive 12.5
Effector memory 30.1
Effector 34.7
Comp-BUV395-A CCR7-BV395
Uninfected individual 2022
HIV-1 p24 28B7 Comp-APC-A
HIV-1 p24+ 0
Comp-PE-A HIV-1 p24 KC57
Virally suppressed individual Wistar 5010
HIV-1 p24+ 0.011
Comp-PE-A HIV-1 p24 KC57
Count
Granzyme B 17.9
Comp-APC-Alexa 700-A Granzyme B- Alexa 700
CTLA4 31.8
Comp-PE-Texas Red-A CTLA4-PE-Dazzle594
IFNγ 21.5
Comp-BV711-A IFNγ-BV711
TNF 80.6
Comp-BV605-A TNF-BV605

**Figure S7. Machine learning identification of cellular markers necessary and sufficient to distinguish HIV-1 RNA$^+$ T cell clones from HIV-1 RNA$^-$ T cell clones by scRFE and flow cytometry validation for HIV-1 p24$^+$ cells.** Related to Figure 7.

**A**, Schematic representation of the process employed for scRFE to identify a subset of genes that were necessary and sufficient to differentiate HIV-1 RNA$^+$ T cell clones from HIV-1 RNA$^-$ T cell clones. **B**, UMAP plots of clonal cells from HIV-1-infected individuals in the unstimulated condition showing training data (green), cells in the test data correctly identified as HIV-1 RNA$^+$ T cell clone (blue), and cells in the test data incorrectly identified as HIV-1 RNA$^-$ T cell clone (red). **C**, Receiver operator curve (ROC) of model performance on the training dataset (blue) and the test dataset (orange). **D,** CD4$^+$ T cells were isolated from PBMC using magnetic negative selection. CD4$^+$ T cells were gated on flow cytometry to select lymphocytes, single cells, and viable cells (Live/Dead near-IR negative). Flow cytometric gating of memory populations (defined by CCR7 and CD45RA expression), HIV-1 p24, granzyme B, CTLA4, IFNγ, and TNF. Gating for p24$^+$ cells in an uninfected individual (2022) and in an HIV-1-infected individual (Wistar 5010) was shown.

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

| Participant ID | Study arm | Age | Sex | Ethnicity | ART | Peak viral load (copies/ml) | Duration since EDDI at the viremic time point (days) | Duration from EDDI to ART (days) | Duration from EDDI to suppressed time point (days) | Duration of viral suppression at the viral suppression time point (days) | CD4 count at the viremic time point (/µl) | CD4 count at the viral suppression time point (/µl) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sabes | | | | | | | | | | | | |
| 236 | Immediate | 22 | M | Hispanic | FTC/TDF/EFV | 7,117,757 | 24 | 22 | 431 | 384 | 625 | 712 |
| 640 | Immediate | 26 | M | Hispanic | EVG/c/TDF/FTC | 6,023,064 | 22 | 22 | 336 | 172 | 586 | 772 |
| 829 | Immediate | 21 | M | Hispanic | EFV/TDF/FTC | 1,205,138 | 48 | 47 | 329 | 245 | 388 | 796 |
| 739 | Delay | 23 | M | Hispanic | EFV/TDF/3TC | 8,604,710 | 41 | 207 | 342 | 256 | 467 | 821 |
| 799 | Delay | 31 | M | Hispanic | EFV/TDF/3TC | 13,313,300 | 20 | 187 | 329 | 245 | 1,203 | 1,464 |
| 910 | Delay | 24 | M | Hispanic | EFV/AZT/3TC | 1,257,674 | 33 | 198 | 338 | 254 | 375 | 371 |
| Wistar | | | | | | | | | | | | |
| 5004 | Long-term ART | 46 | M | White | ABC/DTG/3TC | NA | NA | NA | NA | 3,003 | NA | 737 |
| 5005 | Long-term ART | 41 | M | AA | TDF, ABC/DTG/3TC | NA | NA | NA | NA | 1,399 | NA | 498 |
| 5007 | Long-term ART | 40 | M | AA | EVG/c/FTC/TAF | NA | NA | NA | NA | 5,480 | NA | 784 |
| 5008 | Long-term ART | 35 | M | AA | BIC/FTC/TAF | NA | NA | NA | NA | 1,235 | NA | 797 |
| 5009 | Long-term ART | 62 | M | AA | BIC/FTC/TAF | NA | NA | NA | NA | 2,052 | NA | 583 |
| 5010 | Long-term ART | 53 | M | AA | BIC/FTC/TAF | NA | NA | NA | NA | 1,234 | NA | 569 |
| 5011 | Long-term ART | 43 | M | AA | DTG/3TC | NA | NA | NA | NA | 1,444 | NA | 667 |
| 5012 | Long-term ART | 33 | M | AA | EVG/c/FTC/TAF | NA | NA | NA | NA | 2,189 | NA | 1,322 |
| Uninfected | | | | | | | | | | | | |
| 2004 | Uninfected | 55 | M | AA | NA | NA | NA | NA | NA | NA | NA | NA |
| 2022 | Uninfected | 59 | M | AA | NA | NA | NA | NA | NA | NA | NA | NA |
| 2023 | Uninfected | 47 | M | AA | NA | NA | NA | NA | NA | NA | NA | NA |

*AA African American; 3TC, lamivudine; ABC, abacavir; AZT, azidothymidine; BIC, bictegravir; /c, cobicistat; DTG, dolutegravir; EFV, efavirenz; EVG, elvitegravir; FTC, emtricitabine; TAF, tenofovir alafenamide; TDF, tenofovir disoproxil; EDDI, estimated dates of detectable infection; NA, not available.