Science Immunology
AAAS

Supplementary Materials for

Integrated BATF transcriptional network regulates suppressive intratumoral regulatory T cells

Feng Shan *et al.*

Corresponding author: Dario A.A. Vignali, dvignali@pitt.edu

Sci. Immunol. **8**, eadf6717 (2023)
DOI: 10.1126/sciimmunol.adf6717

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Blood and tissue sample processing

Peripheral blood mononuclear cells (PBMC) were isolated from whole blood by density gradient centrifugation in Ficoll/Hypaque for 20 minutes at 400 xg with the brake off. Carryover red blood cells were lysed with BD Pharm Lyse, and samples were then resuspended in staining buffer (phosphate buffered saline [PBS] with 0.1% azide, 10 mM HEPES, and 2% FBS). Single-cell suspensions from tonsil tissues were generated by mechanical disruption followed by enzymatic digestion with 50 µg/mL of Liberase DL (Roche) in 5 mL of serum-free Roswell Park Memorial Institute media (RPMI) for 15 minutes at 37°C. After initial isolation from tissue, cells were passed through a 100 µm filter and spun at 500xg for 10 minutes to yield single-cell suspensions. Cells were stained by incubation with PerCP-Cy5.5-conjugated mouse anti-human CD4 (clone: RPA-T4; BioLegend), BV421-conjugated mouse anti-human CD8 (clone: RPA-T8; BioLegend), APC-conjugated mouse anti-human CD127 (clone: A019D5; BioLegend) and BV650-conjugated mouse anti-human CD25 (clone: BC96; BioLegend) at a 1:100 dilution for 15 minutes at 4°C, spun down at 500xg for 5 minutes, and incubated in PBS with 1:4000 eFluor 780 viability dye (eBioscience) for 15 minutes at 4°C. Cells were then resuspended in staining buffer, and sorted for live CD4+ T_{conv} ($CD4^{+}CD25^{-}CD127^{high}$) and T_{regs} ($CD4^{+}CD25^{high}CD127^{low}$) by fluorescence-activated cell sorting (FACS) on the Sony MA900 at the Hillman Cancer Center Cytometry Facility. Data S6 summarizes the detailed information of antibodies used for cell sorting.

Single-cell RNA-seq library preparation and sequencing

Immediately following sorting, T_{conv} and T_{regs} were centrifuged for 5 min at 500 xg and were resuspended in PBS with 0.04% BSA. Cells were then counted using the Cellometer Auto2000 (Nexcelom) and loaded into the 10X Controller (10X Genomics) targeting a recovery of 2000 cells per sample. The RNA capture, barcoding, cDNA and library preparation were performed according to the manufacturer's recommendations. 10x libraries were pooled and sequenced on

either a NextSeq500 at the Health Sciences Sequencing Core at Children's Hospital of Pittsburgh or on a NovaSeq6000 at the UPMC Genome Core. Data S1 summarizes the final cell number per participant analyzed in the study after QC.

T_{reg} isolation and expansion

Human umbilical cord samples were collected from the umbilical vein immediately after vaginal delivery by Obstetric Specimen Procurement Unit at UPMC Magee-Womens Research Institute. PBMC isolation was followed same procedures as described above. T_{regs} from cord blood PBMC were enriched with the EasySep™ Human $CD4^+CD127^{low}CD25^+$ Regulatory T Cell Isolation Kit (Stemcell Technologies). Pre-enriched T_{regs} were purified again by staining with PerCP-Cy5.5-conjugated mouse anti-human CD4 (clone: RPA-T4; BioLegend), BV421-conjugated mouse anti-human CD8 (clone: RPA-T8; BioLegend), APC-conjugated mouse anti-human CD127 (clone: A019D5; BioLegend) and BV650-conjugated mouse anti-human CD25 (clone: BC96; BioLegend) and eFluor 780 viability dye (eBioscience). Live $CD4^+CD25^{high}CD127^{low}$ T_{regs} were isolated by FACS on the Sony MA900 at the Hillman Cancer Center Cytometry Facility. Freshly isolated T_{regs} were cultured in complete Roswell Park Memorial Institute (cRPMI: RPMI-1640 (Sigma) supplemented with 5 mM HEPES (Gibco), 2 mM glutamine (Gibco), 50 μg ml−1 penicillin/streptomycin (Gibco), 5 mM nonessential amino acids (Gibco), 5 mM sodium pyruvate (Gibco) and 10% FBS (Atlanta Biologicals)). T_{regs} were expanded for seven days in 24-well tissue culture plates (Corning Costar) coated overnight with 1 μg/ml αCD3 (eBioscience, clone OKT3) in cRPMI supplemented with 2 μg/ml αCD28 (eBioscience, clone CD28.2) and 200 UI/ml recombinant human IL-2 (PeproTech). 1 ml cRPMI was removed and replenished with 1ml αCD28 and human IL-2 cRPMI every two days. Data S6 summarizes the detailed information of antibodies used for cell sorting and T ell activation.

In vitro human T_{reg} culture with TCR-stimulation in hypoxia

T_{regs} from cord blood were isolated as shown above. T_{regs} were then activated at 20,000 cells per 96-well round-bottomed plates with an equivalent number of magnetic beads coated with anti-CD3/anti-CD28 in the presence of 1000 U/ml IL-2 in 200 µl of complete RPMI. After 24-hour activation, anti-CD3/anti-CD28 Dynabeads were magnetically removed and T_{regs} were washed and split into four conditions at 20,000 cells per well. T_{regs} were expanded with IL-2 only ('acute' TCR stimulation) or co-cultured with 200,000 anti-CD3/anti-CD28 Dynabeads and IL-2 ('continuous' TCR stimulation), with the medium changed regularly to prevent nutrient depletion. T_{regs} under acute or continuous TCR activation were cultured in either normoxic (18.6% O_2) or hypoxic (1.5% O_2) gas atmosphere conditions for 10 days. T_{regs} were split in half every 48 hours with fresh medium in presence of 1000 U/ml IL-2. The number of Dynabeads were kept consistent per well as 10:1 bead/cell ratio.

Cas9 RNP assembly and electroporation

100 µM crRNA (IDT) and 100 µM trans-activating RNA (tracrRNA) (IDT) were mixed in a 1:1 ratio and incubated for 30 min at 37 °C to generate 50 µM crRNA–tracrRNA CRISPR duplex. 50 µM TrueCut Cas9 Protein v2 (Thermo Fisher) was mixed with the crRNA–tracrRNA duplex and incubated for 15 min at 37 °C to generate 25 µM Cas9 ribonucleoprotein (RNP). 1×10^6 ex vivo expanded human T_{regs} were pelleted and resuspended in 20 µl of P3 buffer. Then, 2 µl of 25 µM Cas9 RNP and 1 µl of 100 µM electroporation enhancer (IDT) were added directly to the cells and the entire volume was transferred to a 96-well reaction cuvette (Lonza). T_{regs} were electroporated using program EH-115 on the Amaxa 4D-Nucleofector (Lonza). Next, 80 µl of pre-warmed cRPMI was immediately added to each well after electroporation and the cells were allowed to recover for 15 min at 37 °C. CRISPR-edited T_{regs} were then transferred to 24-well tissue culture plates with cRPMI supplemented with 200 UI ml−1 human IL-2 and rested for 48 hours at 37 °C. CRISPR-edited T_{regs} and scrambled control T_{regs} were then activated for 72 hours with TCR-stimulation.

Cells were activated in 24-well tissue culture plates (Corning Costar) coated overnight with 1 µg/ml αCD3 (eBioscience, clone OKT3) in cRPMI supplemented with 2 µg/ml αCD28 (eBioscience, clone CD28.2) and 200 UI/ml recombinant human IL-2 (PeproTech). Data S7 summarizes the sequences of guides used in the experiments.

In vitro microsuppression assay

Naïve $CD8^+$ T cells were isolated by EasySep Human Naïve CD8+ T Cell Isolation Kit (Stemcell) and antigen-presenting cells (APCs) were isolated by FACS using Beckman Coulter MoFlo Astrios from healthy donor PBMC. Isolated naïve CD8+ T cells were then labeled with 5 µM CellTrace™ Violet (CTV) Cell Proliferation Kit (Invitrogen) for 10 min at 37 °C. After 72-hour TCR-restimulation, CRISPR-treated T_{regs} were stained with anti-human CD4, anti-human CD25 and anti-human CD127 antibodies with eFluor 780 viability dye for further purification on the Sony MA900. 2,000 CRISPR-edited T_{regs} in 50 ml cRPMI were seeded in the first column of a round-bottom 96-well plate (Costar). A serial two-fold dilution of T_{regs} was conducted for nine columns. 2,000 APCs and 2,000 CTV-labeled naïve $CD8^+$ T cells in 50ml cRPMI were added separately to all wells to reach the T_{regs}:$CD8^+$ T cell ratio from 1:2 to 1:1000. 50 ml cRPMI supplemented with anti-CD3 (2 µg/ml) (eBioscience, clone OKT3) were then added to all wells. No-proliferation control and no-suppression control were included, either composed of CTV labeled $CD8^+$ T cells alone or with APC. Cells were cultured for 5 days at 37°C with 5% CO_2. Then cells were spun down at 500 xg for 5 minutes and stained with anti-human CD4 antibody, anti-human CD8 antibody and eFluor 660 viability dye analysis flow cytometry analysis. Experiments were repeated with five independent replicates.

Surface and intracellular antibody staining

CRISPR-edited T_{regs}, scrambled controls and unperturbed controls were resuspended in staining buffer and labelled with antibodies at 1:100 for 25 mins at 4 °C, followed by viability staining using

fixable dye (eFluor 780 viability dye (Invitrogen) at 1:4000 or Zombie NIR (BioLegend) at 1:2000 ratio. Cells were then spun down at 500 xg for 5 minutes and washed with PBS. Foxp3/Transcription Factor Staining Buffer (eBioscience) was added into cells for 60 mins at room temperature for cell fixation. Cells were then washed by permeabilization buffer (eBioscience) and labelled with antibodies for 60 minutes at room temperature for intracellular staining. Cells were washed with permeabilization buffer and FACS buffer. BD LSRFortessa II flow cytometer or Cytek Aurora were used for acquiring flow cytometry readouts and FlowJo V10 was used for data analysis.

Bulk RNA sequencing

CRISPR edited T_{regs} and scrambled controls were double sorted (purity >99.5%) directly into individual wells of a 96-well plate containing 2μl lysis buffer (0.2% Triton X-100 with RNase inhibitor at 2Uμl −1). The plate was spun down at 2,000 xg for 2min and reverse transcription was then performed. For reverse transcription, a mixture (2μl) of reverse-transcription primers and dNTP (1mM) was added to each well followed by incubation at 70°C for 3 min. Denatured templates and reverse-transcription primers were quickly spun down and 6μl reverse-transcription master mix containing MgCl2 (9mM), first-strand buffer (5×), Superscript II reverse transcriptase (10Uμl −1), DTT (5mM), Betaine (1M), RNase inhibitors (1Uμl −1) and template switch oligo (1μM) were added followed by incubation on a thermocycler with the following parameters: 42°C for 90min, ten cycles of 50/42°C and 70°C for 15min, hold at 4°C. An addition of 15-cycle complementary DNA amplification was performed following cDNA synthesis by the KAPA Hot Start II High-Fidelity DNA Polymerase. The amplified cDNA was purified using Ampure XP beads (at 0.6:1 bead to cDNA ratio) and eluted with 17.5μl elution buffer. cDNA size (peak at roughly 1.5–2kb) was verified with TapeStation5000 and quantified by the Qubit (ThermoFisher). Sequencing libraries were prepared from 1ng cDNA using the Nextera XT DNA Library Prep kit (Illumina FC-131-1096), following the manufacturer's instructions. cDNA Libraries were quantified

by the KAPA library quantification kit (KAPA KK4854) and size (peak at ~400bp) was verified on a TapeStation1000. Ten diluted libraries (2nM) were pooled and sequenced with the NextSeq 500/550 High Output v.2 kit using 75bp single-end reads.

Bulk RNAseq data analysis

We used the R package DESeq2 (1.40.1) for downstream bulk RNA-seq analysis (*66*). The batch effects between donors were corrected by ComBat-seq (v3.36.0) (*67*). Genes were filtered as not expressed if their read count less than 10. Differential gene analysis was done by DESeq function in DESeq2 (1.40.1). Log2 fold changes were added into the analysis by lfcShrink function in DESeq2 (1.40.1). Gene set enrichment analysis was performed by using R package clusterProfiler (4.8.1) (*68*).

scRNAseq data integration and clustering

We used the R package Seurat v4 for downstream scRNA-seq analysis. Cells were filtered that have over 4000 features and more than 20% mitochondrial counts according to Seurat QC suggestions. The dataset was split by patient and normalized using SCTransform v1 (*23*). The most variable genes in each patient were identified by the SelectIntegrationFeatures function. The top 2,000 variable genes overlapped across groups were selected for integration. Then we scaled the integrated dataset to remove confounding sources of variation by regressing out the percentage of mitochondrial genes expressed per cell and number of genes per cell. Further, we used principal component analysis (PCA) to dimensionally reduce the dataset to 30 dimensions and chose 1-20 PCs that explained the most variance in the dataset for visualization and cell clustering. We next performed Louvain graph-based clustering in Seurat to identify the clusters and applied a non-linear dimensional reduction method Uniform Manifold Approximation Embedding (UMAP) to visualize the dataset. T_{conv} and T_{regs} were identified based on the original

paper and previous sorting results. Cell types and cell origins were used to label the cells in the UMAP embedding.

Differential expression analysis and gene set enrichment analysis

We next used FindAllMarkers function in Seurat to run differential expression (DE) analysis between different groups of T cell subpopulations depending on the comparison. We used a Wilcoxon rank-sum test and only detected gene markers that expressed at least 25% of cells in either of the two groups of cells. The top DE genes in each subpopulation were ranked based on the log2 fold change. We then performed the gene set enrichment analysis to T_{regs} in the dataset by using the R package singleseqgset as previously described (*19*). Gene sets were extracted from the MSigDB Database (C7 Immunology Gene Sets, Reactome subset of canonical pathways, KEGG subset of canonical pathways and GO: Gene Ontology gene sets). The log2 fold change in gene expression was calculated for all genes across T_{reg} clusters and used as input for a variance inflation corrected Wilcoxon rank sum test to calculate whether the gene sets were upregulated in a concerted manner within a cluster.

Gene regulatory network reconstruction

To reconstruct the gene regulatory networks (GRN) within T_{reg} subpopulations, we used SCENIC to infer regulons with normalized data matrices as input. The regulons were generated and the regulon activities were calculated following the pySCENIC (0.11.2) pipeline (*46*). We identified the differentially activated regulons in each T_{reg} subpopulation by the Wilcoxon rank sum test against all the other cells. Regulons were considered as enriched by the adjusted false discovery rate (FDR) p-value < 0.05 and ranked in the heatmap based on log2 fold change. To identify GRNs central to TNFR+ T_{reg} subpopulation, we combined the target genes and regulators in the top 5 regulons ranked by log2 fold change and used normalized data metrices as input for the causal mixed graphical modeling (MGM) (*42*). We used MGM to build a skeleton network, then Fast

Causal Inference MAX (FCI-MAX) to refine the network and determine correlation (*69*). Edges represent partial correlation with p-value < 0.05 and FDR < 0.1. Signaling pathways enriched in the GRN and combined regulon networks were analyzed by Reactome (*70*).

Pseudotime analysis

To infer T_{reg} developmental trajectory, we applied the RNA Velocity algorithm to infer pseudotime for each T_{reg} (*26, 71*). The unspliced UMIs for each gene in each cell were counted by the python package velocyto (version 0.17.15) and the spliced gene expression matrices was exported from Seurat project. The differentiation trajectory and pseudotime inference were performed by scanpy (version 1.6.1) and scVelo (v0.2.4) (*26*). Filtering and normalization were applied to both unspliced and spliced count matrices. A k nearest-neighbor graph (k=30) was built using the top 30 principal components. RNA velocity was then estimated by scv.tl.recover_dynamics function in velocyto and followed by pseudotime inference. To infer the developmental trajectory of T_{regs} from the online published datasets, we applied the Slingshot (*72*) algorithm to infer pseudotime for each T_{reg}. Briefly, we took the top three principal components from the PCA generated during cell clustering procedure (as described above) and normalized the dataset for input into the R package Slingshot. We used the pre-defined clustering results to identify the global lineage structure by a cluster-based minimal spanning tree (MST), fitted principal curves to describe each lineage and finally aligned cells on a pseudotime trajectory.

Survival analysis using The Cancer Genome Atlas

To determine whether the enrichment of T_{reg} subpopulations was associated with survival outcomes of patients with cancer, we utilized bulk RNA-seq data of patients with HNSCC, NSCLC and melanoma available through TCGA. Similarly, bulk RNA-seq data of normal lung and normal skin (sun exposed) were obtained from Genotype-Tissue Expression Project (GTEx). The enrichment score for each T_{reg} subpopulation from each patient was calculated as previously

described (*73, 74*). Briefly, we derived genes sets that were reflective of the cell populations of interest and determined an enrichment score for each patient in the single-cell RNAseq dataset. The gene set was defined by using the top 200 genes that differentially expressed in each T_{reg} subpopulation ranked by log2 fold change. Enrichment scores were calculated by AddModuleScore function in Seurat v4. The enrichment scores of each patient in the single-cell RNAseq dataset were used as a reference profile to deconvolute of bulk RNA-Seq data from TCGA and GTEx by CIBERSORTx (*30*). We stratified patients into signature-high versus -low groups based on their enrichment score by surv_cutpoint function in R package survminer. We then performed monovariate and multi-variable analyses based on patients' progression-free survival.

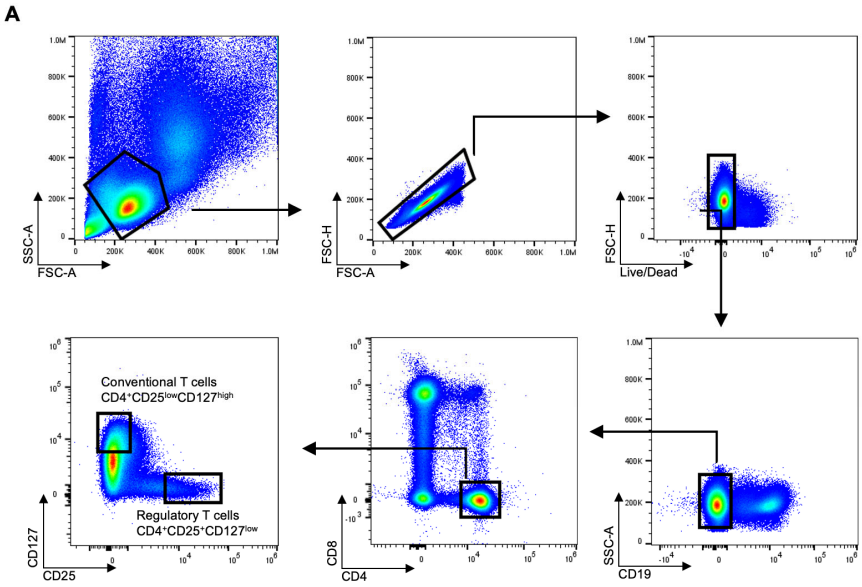


Fig. S1 Gating strategy of sorting T_{regs} and T_{conv} (**A**) Representative gating strategy of sorting T_{regs} and T_{conv} in TIL and PBMC from patients with HNSCC, tonsils from patients with tonsilitis and HD PBMC.

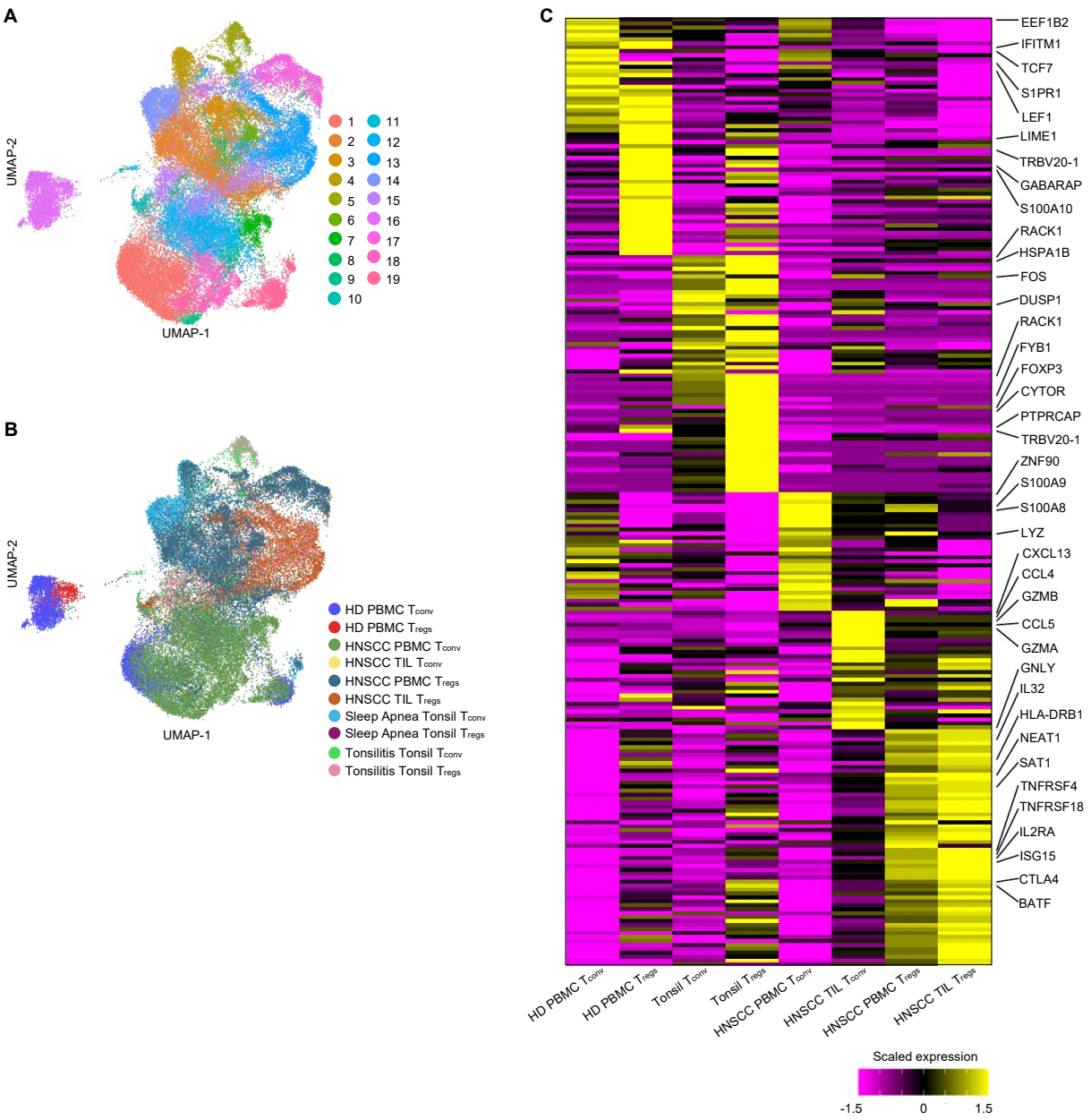


Fig. S2 Unbiased clustering showed distinct patterns for CD4⁺ T cells across cell origins. (**A**) UMAP embedding and clustering of 51,195 CD4+ T cells from matched TIL and PBMC from patients with HNSCC (n=26), inflamed tonsil tissues from patients with tonsilitis (n=11) and healthy donor PBMC (n= 10). UMAP embedding of 51,195 CD4⁺ T cells were color coded by cluster identity. **(B)** UMAP embedding and clustering of 51,195 CD4⁺ T cells from matched TIL and PBMC from patients with HNSCC (n=26), tonsil tissues from patients with tonsilitis (n=5),

tonsil tissues from patients with sleep apnea (n=6), and healthy donor PBMC (n= 10). UMAP embedding of 51,195 $CD4^+$ T cells were color coded by cell type and cell origin. (**C**) Relative expression of genes differentially expressed in each subset of $CD4^+$ T cells grouped by cell type and cell origin on a heatmap. Top 5 DE genes ranked by log2 fold changed were labelled in the heatmap. The gene expression is scaled by transforming the expression in each population to zero mean and unit standard deviation is shown in the heatmap.

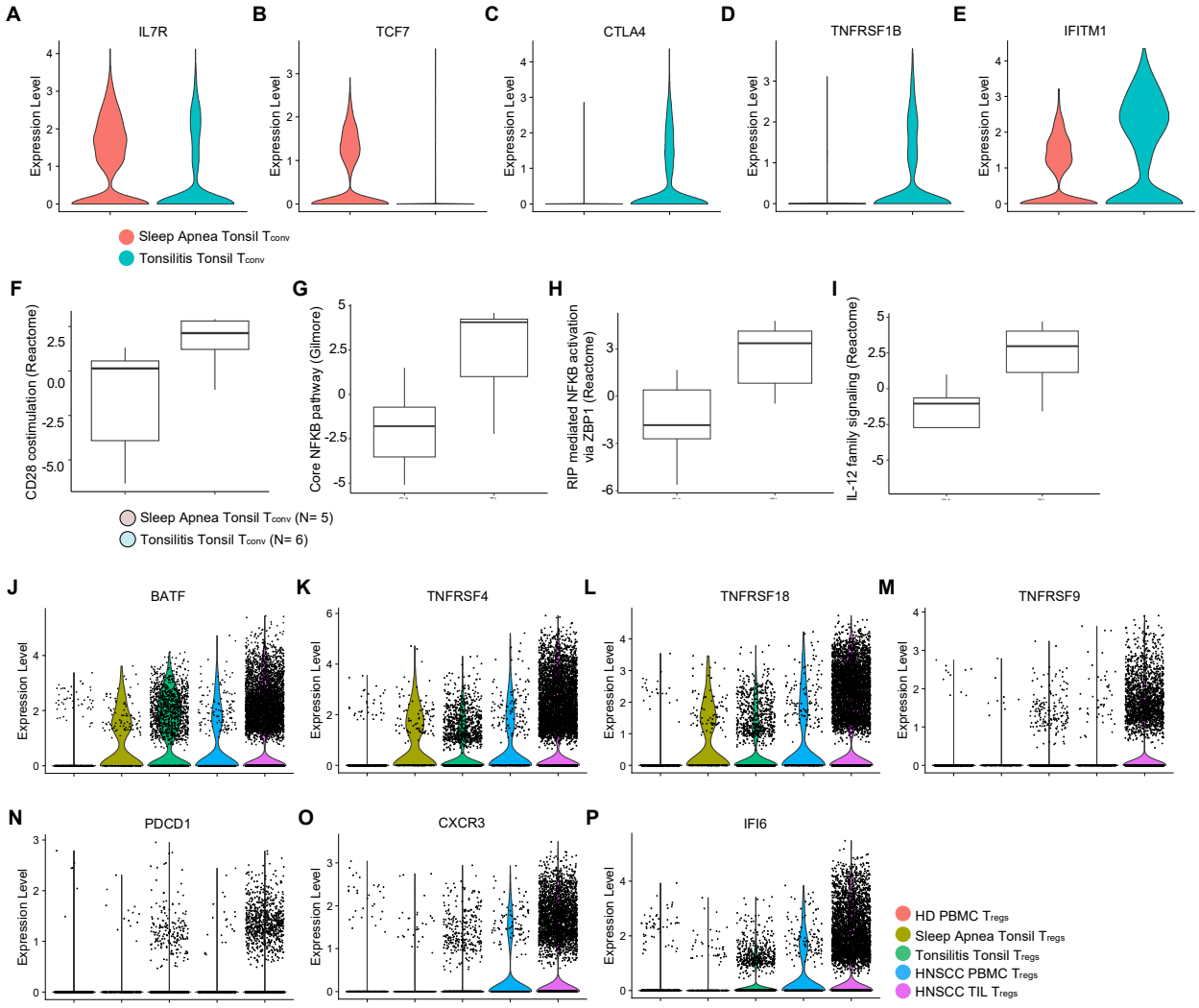


Fig. S3 Tonsil CD4⁺ T cell comparisons between sleep apnea and tonsilitis. (**A-E**) Gene expressions of *IL7R*, *TCF7*, *CTLA4*, *TNFRSF1B* and *IFITM1* in CD4⁺ T_{conv} from patients with sleep apnea and tonsilitis are shown in violin plots. UMAP embedding and clustering of 51,195 CD4+ T cells from matched TIL and PBMC from patients with HNSCC (n=26), inflamed tonsil tissues from patients with tonsilitis (n=11) and healthy donor PBMC (n= 10). UMAP embedding of 51,195 CD4⁺ T cells were color coded by cluster identity. **(F-I)** Bar plots showing the gene set enrichments of in CD4⁺ T_{conv} from patients with sleep apnea and tonsilitis. (**J-P**) Gene expressions of selected markers in tonsil T_{regs} from patients with sleep apnea and tonsilitis are shown in violin plots. Gene expression is normalized by log transformation.

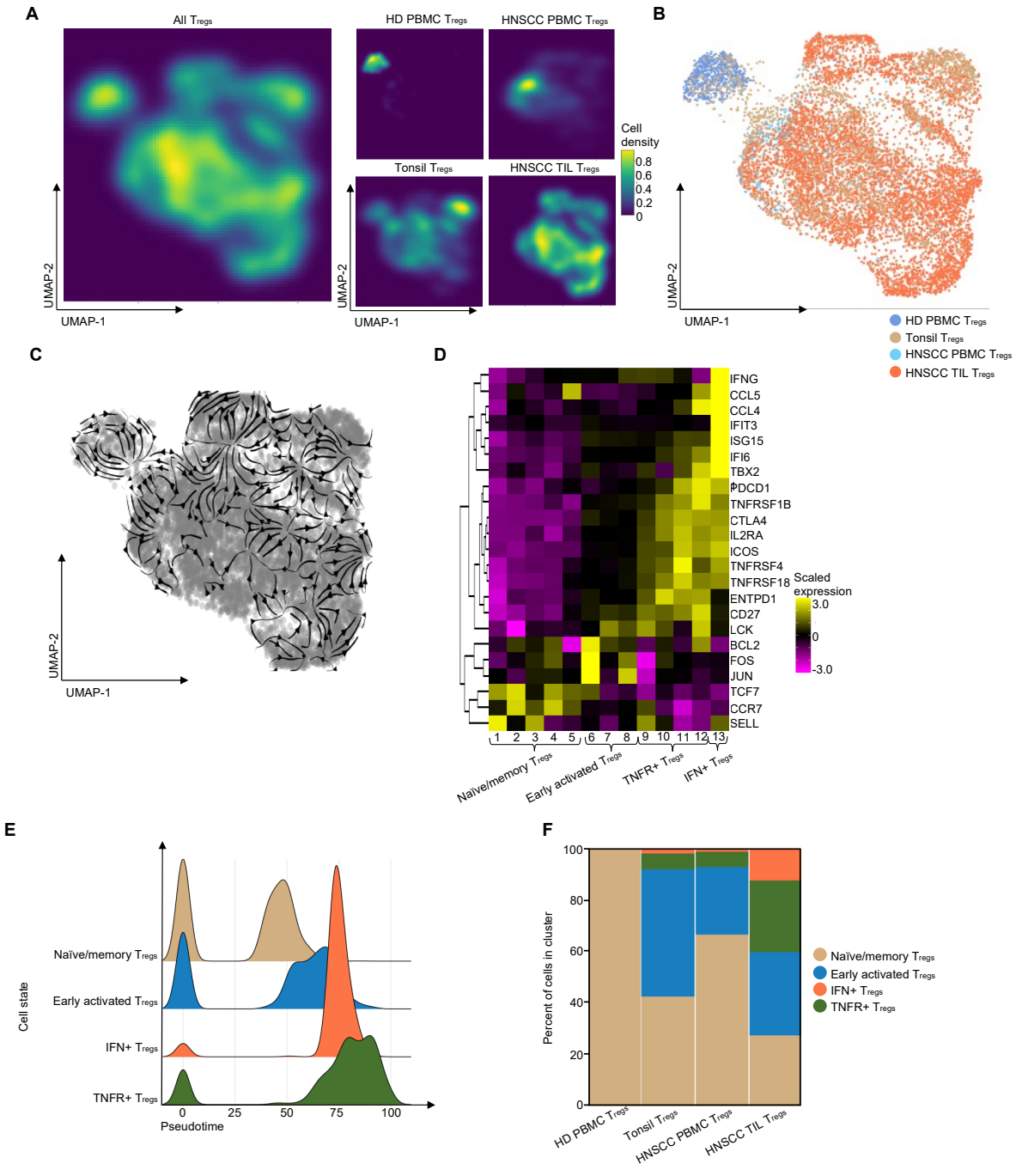


Fig. S4. Intratumoral T_{regs} were characterized by cell state, and their expression follows differentiation trajectory in the HNSCC TME. (**A**) 2D galaxy plots of all T_{regs} were grouped by cell origins. The 2D galaxy plots were colored by percentage of cell density. The percentage was calculated by using the total number of cells in the plot as the parent population. (**B**) UMAP

embedding all T_{regs} were color coded by cell type and cell origin. (**C**) Differentiation trajectory inferred from RNA Velocity was visualized on a 2D UMAP embedding. (**D**) Relative expression of selective canonical marker genes associated with cell state were visualized on a heatmap. Cluster 1-5 are associated with naïve/memory T_{reg} phenotypes. Cluster 6-8 have early activated T_{regs} phenotype and cluster 9-13 exhibit activated T_{reg} phenotype. Cluster 9-12 can be distinguished from cluster 13 based on TNFR member genes, IFN-response genes and Th-1 like expression signatures. The gene expression is scaled by row and shown in the heatmap. (**E**) Cell distribution at each T_{reg} cell state on the imputed pseudotime was visualized by a ridge plot. Y-axis shows the number of cells in each cluster by the imputed pseudotime and indicates the abundance of cells on the pseudotime trajectory. (**F**) Percentage of T_{regs} at each cell state in each Treg origin was visualized by a stacked plot. T_{regs} were colored based on the cell state annotation results.

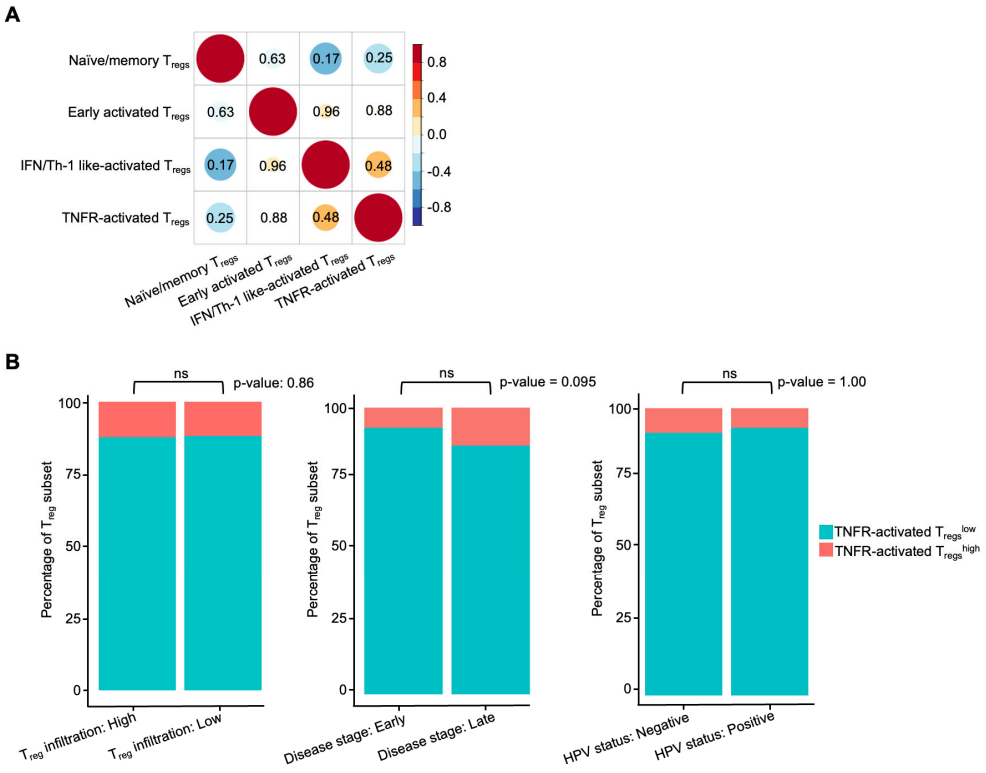


Fig. S5. Correlation analysis of each T_{reg} subpopulation and multivariate analysis of TNFR-activated T_{regs}. (**A**) Pearson correlation analysis between each T_{reg} subpopulation was shown on a dot plot. The enrichment of T_{reg} subpopulations were not correlated with each other. The dot size was scaled based on the correlation coefficient score. Positive correlation coefficient scores were colored in red and negative correlation coefficient scores were colored in blue. P-values were calculated by t test and displayed on each dot. (**B**) The abundance of TNFR-activated T_{regs} stratified by level of overall T_{reg} infiltration in the HNSCC TME, disease stage of patients with HNSCC and HPV status of patients with HNSCC were shown in bar plots. P values were calculated by Fisher's exact test. ns: not significant.

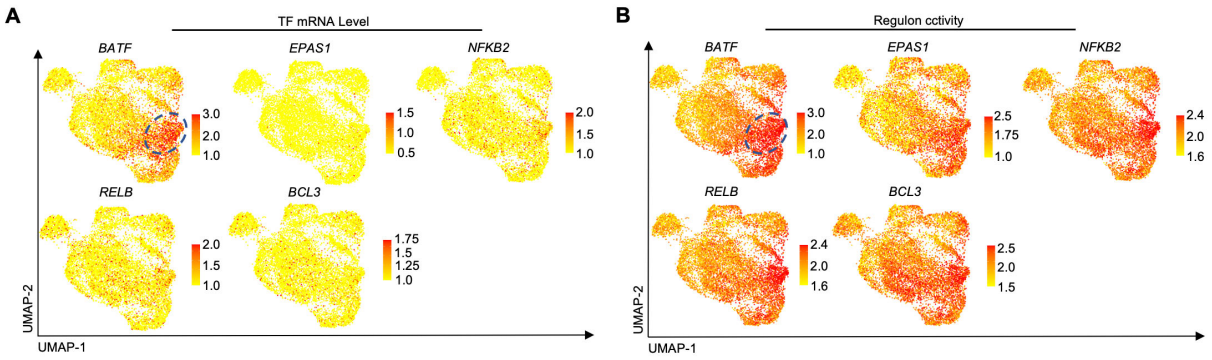


Fig. S6. Regulons inferred by SCENIC enabled to recover TF activity in T_{regs}. (A) Relative mRNA expression level of selective TFs was shown by 2D UMAP embeddings. (**B**) Activity score of regulons inferred from SCENIC was plotted by 2D UMAP embeddings. The Blue dash circle indicated the cluster of TNFR-activated T_{regs}. SCENIC inference recovered TF activity that had low level of mRNA expression in the scRNA-seq data. Yellow indicates low expression and red indicates high expression in the plots.

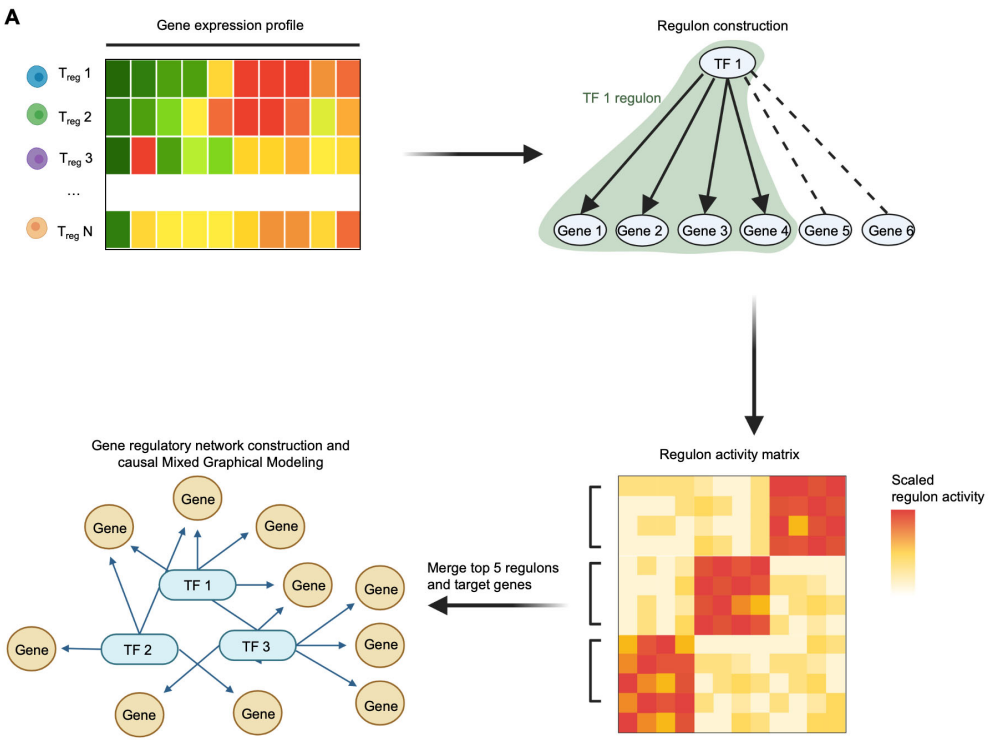


Fig. S7. Reconstruction of gene regulatory networks of T_{regs}. (**A**) Schematics of gene regulatory network construction. Regulons were generated by using SCENIC pipeline. The differential enriched regulons in each T_{reg} subpopulation were identified by non-parametric Wilcoxon rank sum test. Regulons were considered as significantly enriched by the adjusted false discovery rate (FDR) p-value < 0.05. The top 5 regulons ranked by log2 fold changed were selected and integrated into gene regulatory network (GRN). The GRN of TNFR-activated T_{regs} was constructed by causal mixed graphical modeling and Fast Causal Inference MAX algorithms.

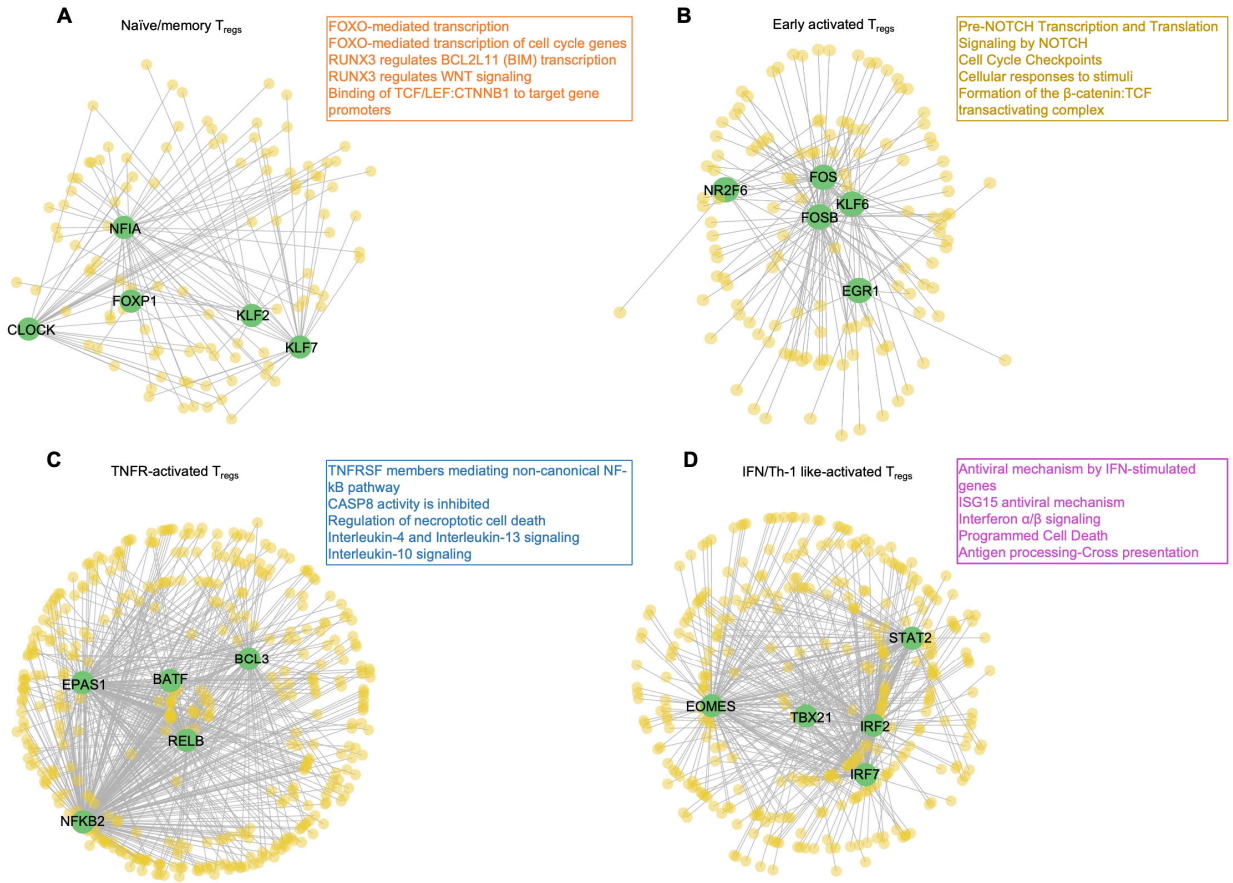


Fig. S8. Gene regulatory networks reveal distinct transcriptional program of T_{reg} subpopulations. (**A**) The gene regulatory network reconstructed in Naïve/memory T_{regs} was visualized. FOXO-mediated pathways, RUNX-WNT axis and TGFβ pathways were highly enriched in Naïve/memory T_{regs} by Reactome signaling pathway analysis and highlighted in the text box. (**B**) The gene regulatory network reconstructed in Early-activated T_{regs} was visualized. NOTCH mediated pathways, NGF-stimulated transcription and β-catenin:TCF transcription signaling were highly enriched in Early-activated T_{regs} by Reactome signaling pathway analysis and highlighted in the text box. (**C**) The gene regulatory network reconstructed in TNFR+ T_{regs} was visualized. Signaling pathways associated with IL-3/IL13, IL-10, IL-21 were highly enriched in TNFR+ T_{regs} by Reactome signaling pathway analysis and highlighted in the text box. (**D**) The gene regulatory network reconstructed in IFN+ T_{regs} was visualized. Signaling pathways related

to type I/II interferon responses and antigen-presentation were highly enriched in IFN+ T_{regs} by Reactome signaling pathway analysis and highlighted in the text box. Transcriptional regulators within each gene regulatory network were colored in green and the downstream target genes were colored in yellow.

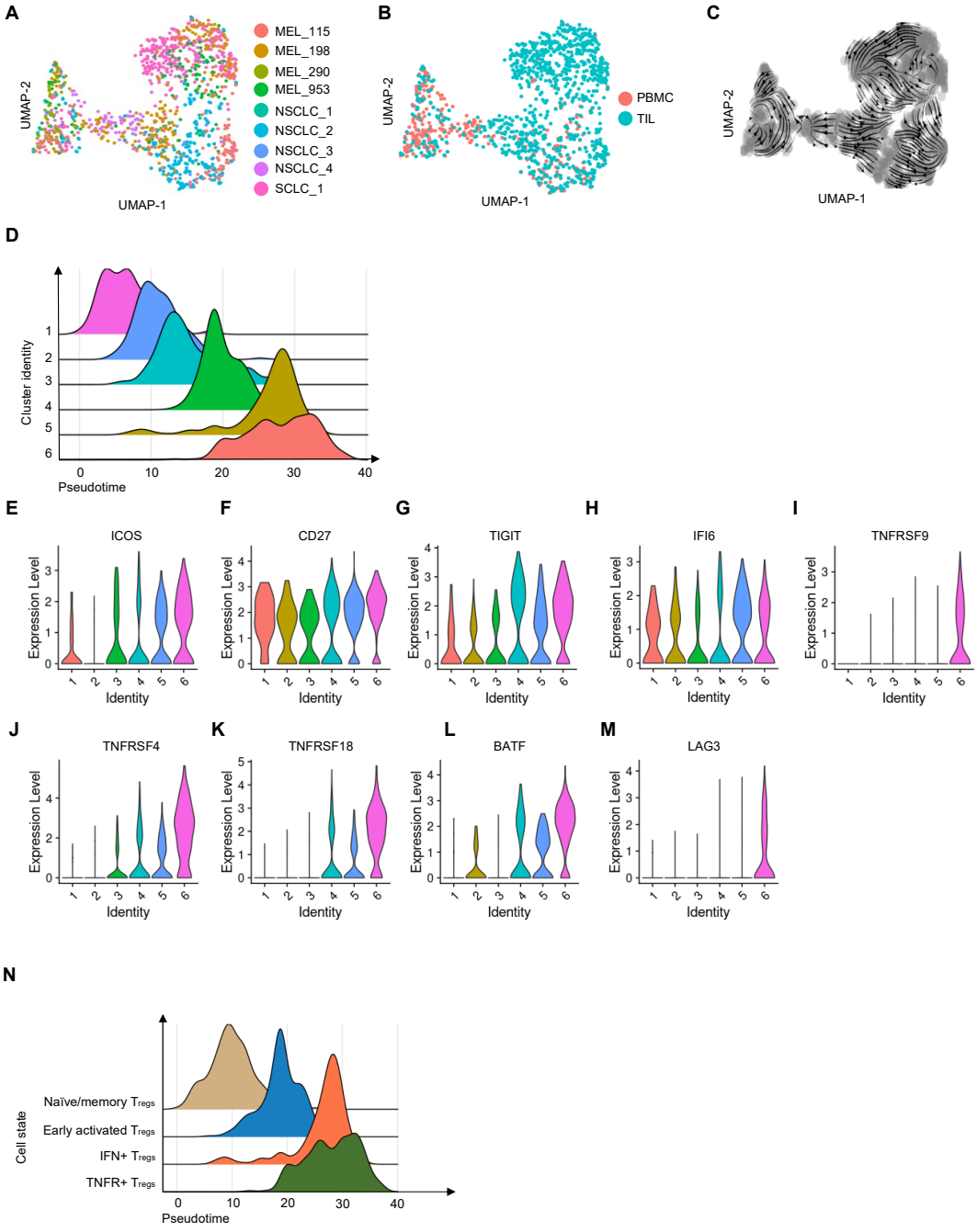


Fig. S9. TNFRSF-activated T_{regs} are highly enriched in lung cancer and melanoma (**A**) The distribution of T_{regs} from lung cancer and melanoma was visualized by a UMAP embedding. (**B**) T_{regs} from TIL and PBMC were separated by the unbiased clustering and visualized on a UMAP embedding. Y-axis shows the number of cells in each cluster by the imputed pseudotime and indicates the abundance of cells on the pseudotime trajectory. (**C-D**) Differentiation trajectory and

pseudotime ordering of clusters imputed by RNA Velocity were visualized in a Ridge plot. (**E-M**) Gene expression of selected markers associated with cell states are shown in violin plots. Log normalized expression of each gene was visualized in the violin plot. (**N**) The pseudotime ordering of T_{regs} imputed by RNA Velocity were grouped by cell state and visualized in a Ridge plot. Y-axis shows the number of cells in each cluster by the imputed pseudotime and indicates the abundance of cells on the pseudotime trajectory.

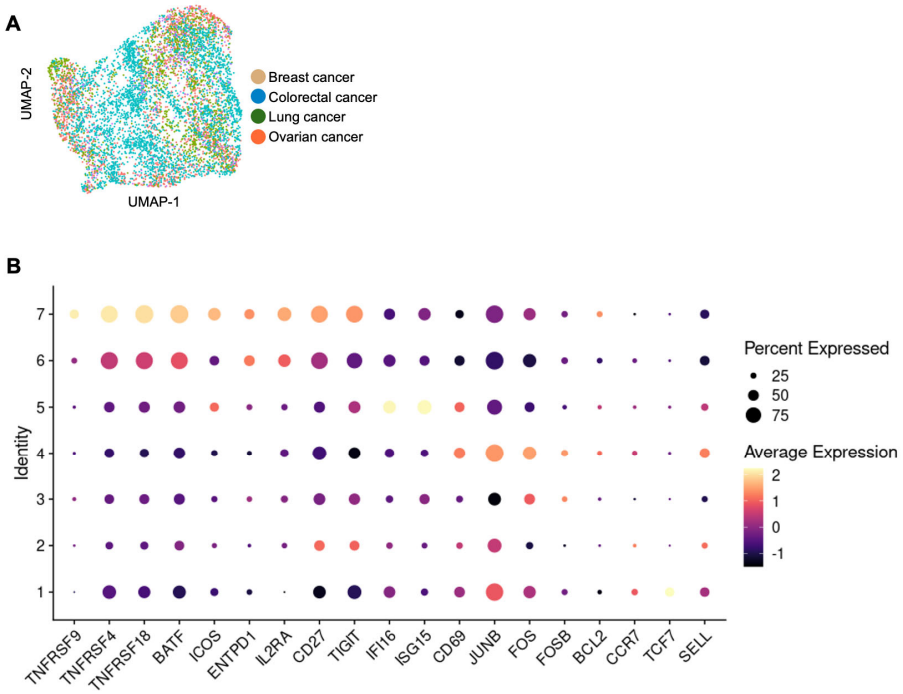
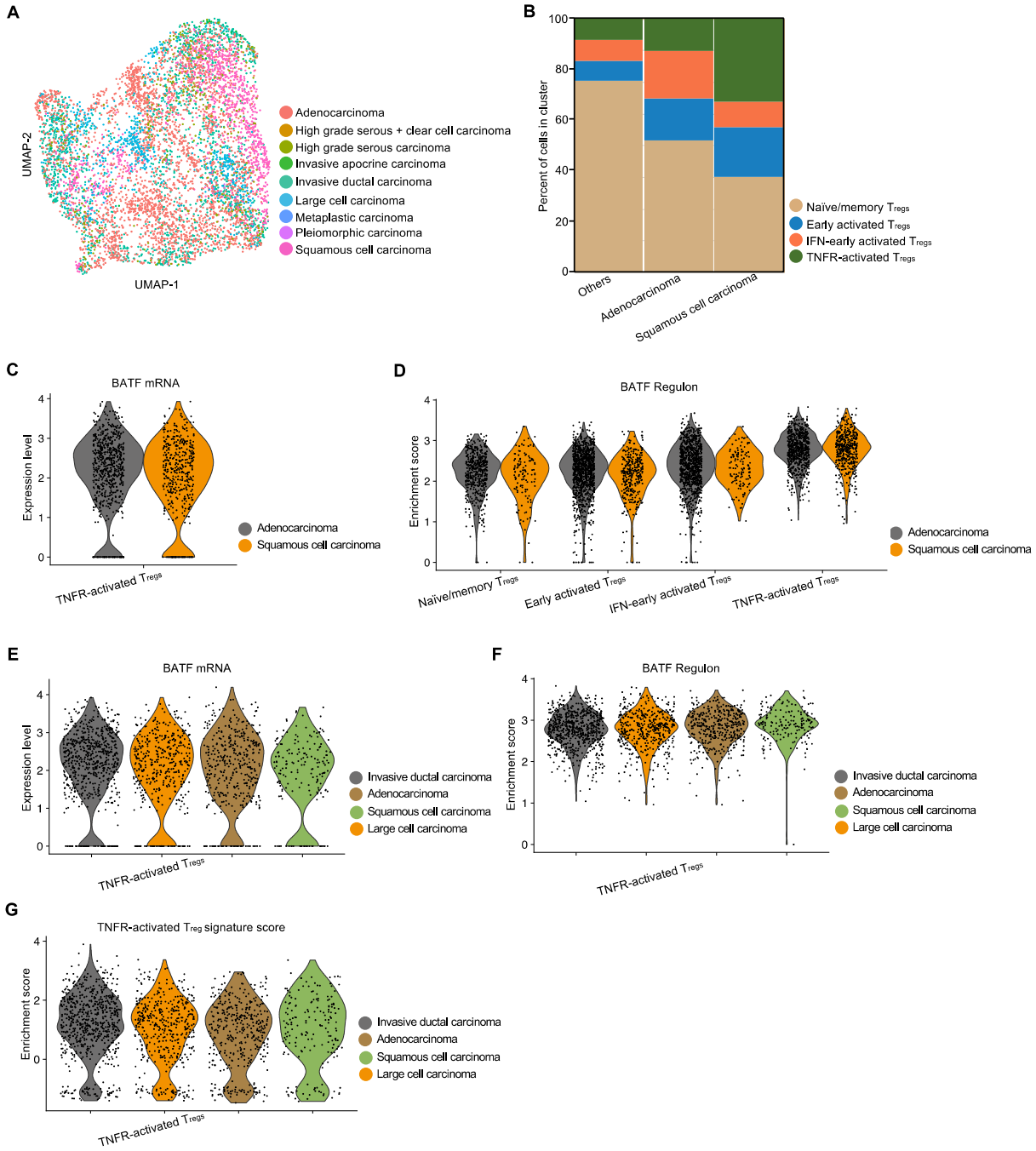


Fig. S10. TNFR-activated T_{regs} were highly enriched in solid TMEs. (**A**) A UMAP embedding of 7,045 T_{regs} in TIL from patients with ovarian cancer (n=5), lung cancer (n=8), breast cancer (n=14) and colorectal cancer (n=7). 7 clusters were identified by unbiased clustering. (**B**) Gene expression of selected markers associated with cell states and the percentage of cells in each cluster expressing marker genes are shown in dot plots.



was visualized by a stacked plot. T_{regs} were colored based on the cell state annotation results. (**C**)

The expression of BATF in TNFR-activated T_{regs} from adenocarcinoma and squamous cell carcinoma was shown in violin plot. (**D**) The enrichment score of BATF regulon in T_{regs} at different cell states from adenocarcinoma and squamous cell carcinoma was shown in violin plot. (**E**) The expression of BATF in TNFR-activated T_{regs} from invasive ductal carcinoma, adenocarcinoma, squamous cell carcinoma and large cell carcinoma was shown in violin plot. (**F**) The enrichment score of BATF regulon in TNFR-activated T_{regs} from invasive ductal carcinoma, adenocarcinoma, squamous cell carcinoma and large cell carcinoma was shown in violin plot. (**G**) The enrichment score of TNFR-activated T_{reg} signature in TNFR-activated T_{regs} from invasive ductal carcinoma, adenocarcinoma, squamous cell carcinoma and large cell carcinoma was shown in violin plot.

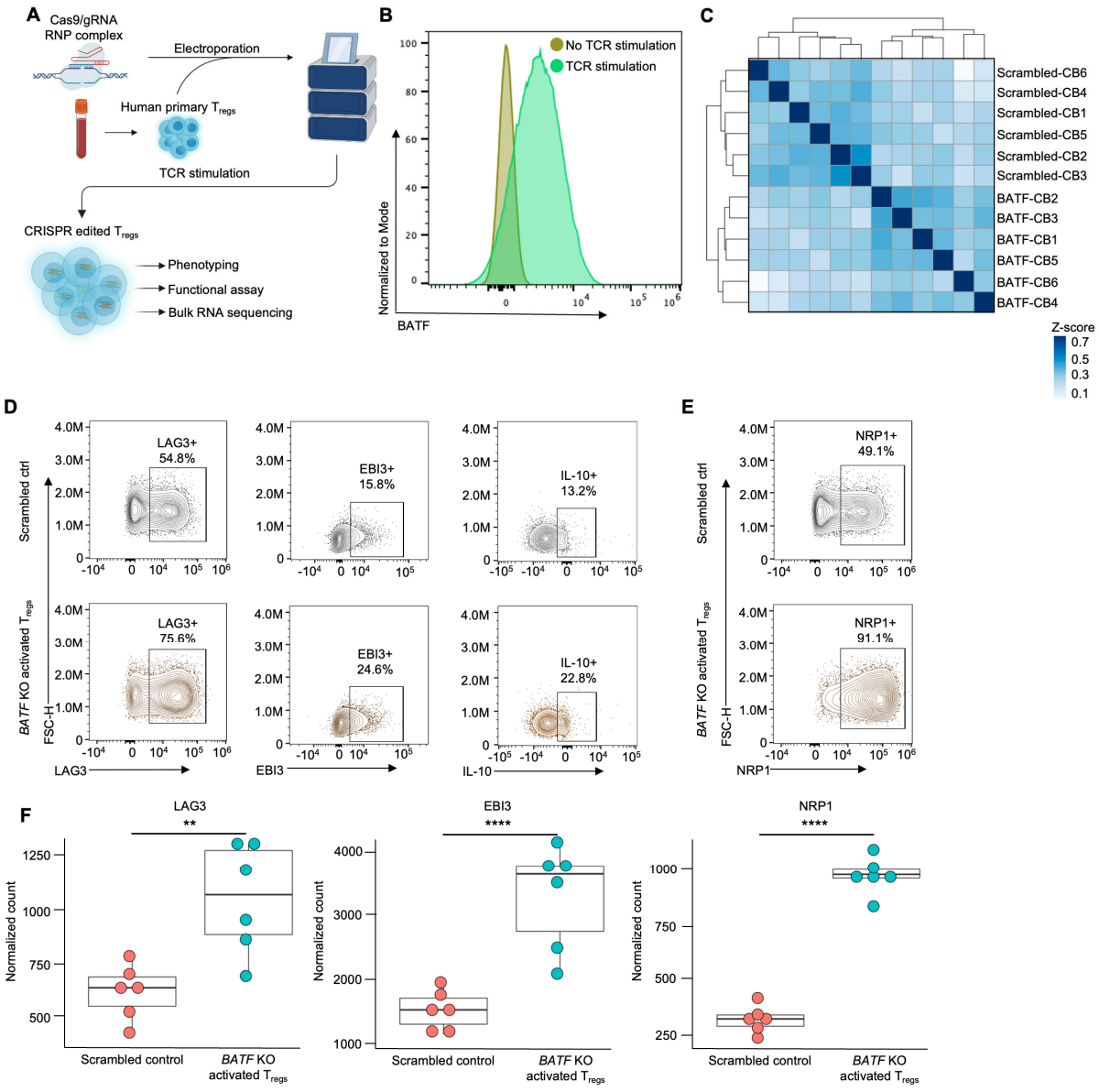


Fig. S12. CRISPR-Cas9 RNP KO approach successfully altered BATF expression level and enabled to examine the effects of BATF disruption in human primary activated T_{regs}. (**A**) Schematic workflow of Cas9 RNP knockout on human primary T_{regs}. Human primary T_{regs} were isolated from human umbilical cord blood, ex vivo cultured and expanded. Cas9 RNPs targeting BATF and nontargeting scrambled control RNP were electroporated into T_{regs}. After three-day TCR stimulation, immunophenotyping on protein expression, microsuppression assay for T_{reg}

functional test and RNA sequencing were done with paired CRISPR edited T_{regs}. (**B**) Represented flow cytometry plot showing BATF is upregulated in human primary T_{regs} after three-day TCR-stimulation compared to ex vivo T_{reg}. (**C**) The hierarchical clustering of six paired CRISPR-treated T_{regs} indicated that T_{regs} were grouped based on knockout condition with restricted batch effect from individual replicates. (**D**) The expression of LAG3, EBI3 and IL-10 in *BATF* KO activated T_{regs} and scrambled control were visualized by represented flow cytometry plots from five individual replicates. (**E**) The expression of NRP1on *BATF* KO activated T_{regs} and scrambled control were visualized by represented flow cytometry plots from five individual replicates. (**F**) Dot plots showing the expression levels of *LAG-3*, *EBI3* and *NRP1* in *BATF* KO activated T_{regs} (n=8) and scrambled controls from the same donor in bulk RNA-sequencing data. Each dot indicates an individual replicate. The FDR p-values were calculated by Wald test.

P-values: **p< 0.01; ****p< 0.0001.

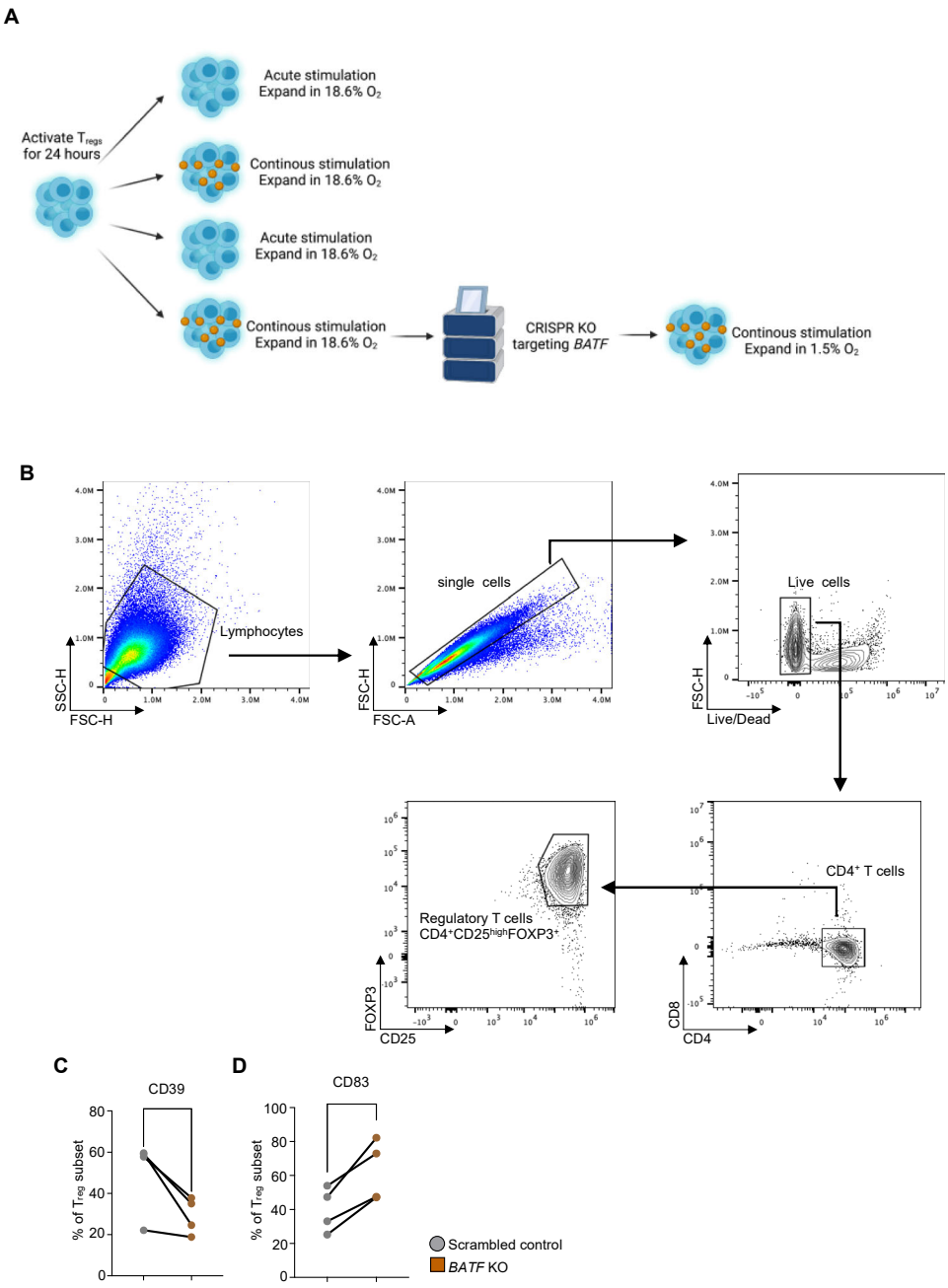


Fig. S13. Human T_{regs} cultured with continuous TCR stimulation under hypoxia in conjunction with CRISPR-Cas9 RNP KO targeting BATF (**A**) Schematic workflow of in vitro culture system for human T_{regs} with continuous TCR stimulation under hypoxia. T_{regs} from cord blood PBMC were isolated and then activated at 20,000 cells per 96-well round-bottomed plates with an equivalent number of magnetic beads coated with anti-CD3/anti-CD28 in the presence of 1000 U/ml IL-2 in 200 µl of complete RPMI. After 24-hour activation, anti-CD3/anti-CD28

Dynabeads were magnetically removed and T_{regs} were washed and split into four conditions at 20,000 cells per well: “acute stimulation” in normoxia (18.6% O_2); “continuous stimulation” in normoxia; “acute stimulation” in hypoxia (1.5% O_2) and “continuous stimulation in hypoxia. T_{regs} were cultured in four conditions for 10 days and electroporated for BATF CRISPR KO. After 48-hour rest in complete RPMI in the presence of IL-2, *BATF* KO T_{regs} and scrambled control Tregs were restimulated by anti-CD3/anti-CD28 Dynabeads in hypoxia for additional 48 hours. (**B**) Representative gating strategy for BATF KO activated T_{regs} and scrambled controls after 3-day TCR-stimulation post electroporation. (**C and D)** Knockout of *BATF* in T_{regs} with continuous stimulation in hypoxia (n= 4) resulted in decreased expression of CD39 and increased expression of CD83. Samples from the same donor were connected by solid lines. Each dot indicates an individual replicate. Data were analyzed by a ratio paired t test.

P-values: *: p <= 0.05; **: p <= 0.01.

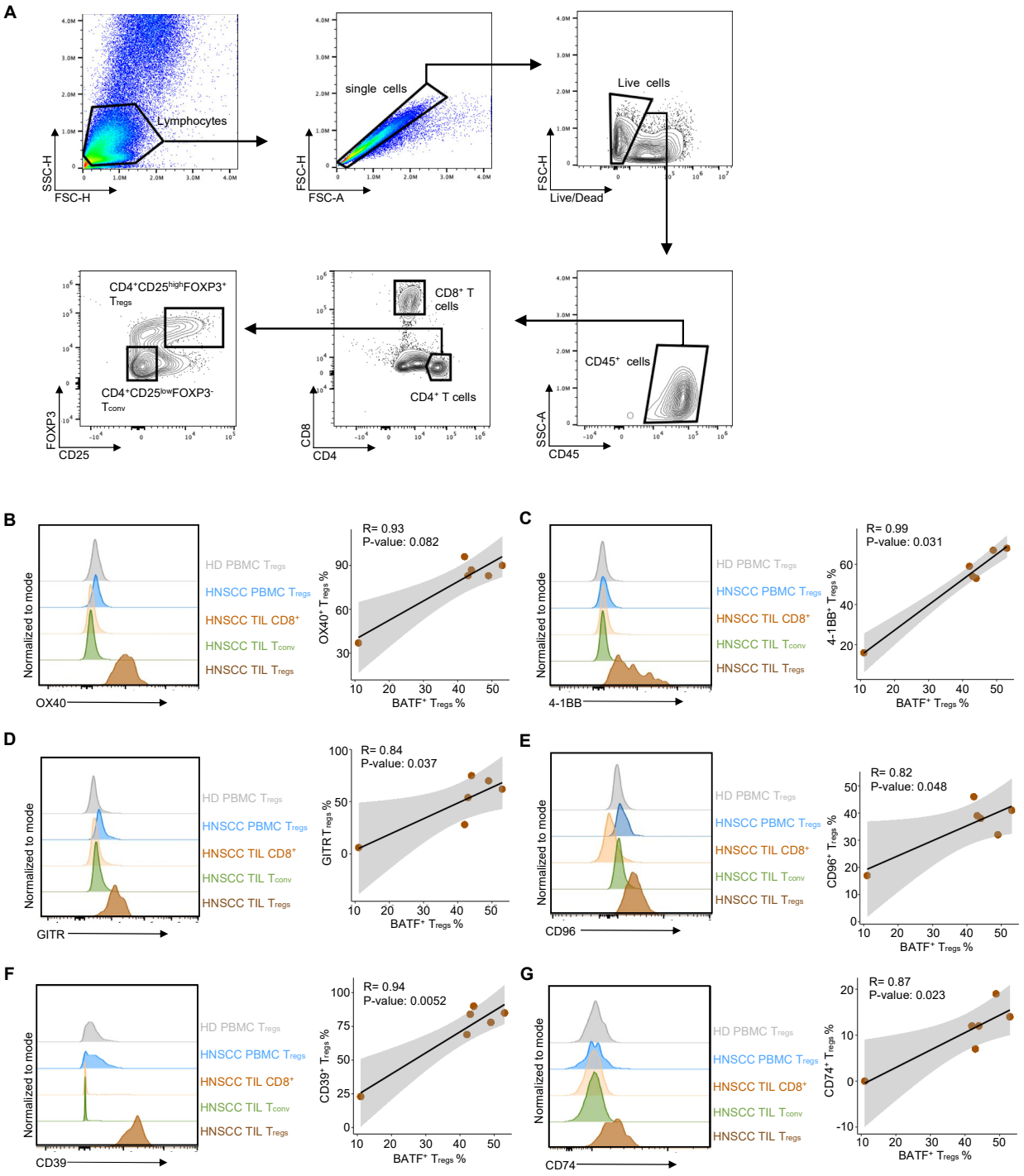


Fig. S14. BATF co-expresses with several surface markers on HNSCC TIL T_{regs}. (**A**) Representative gating strategy of T_{regs} in TIL and PBMC from patients with HNSCC, HD PBMC, and T_{conv} and $CD8^+$ T cells in TIL from patients with HNSCC. (**B-G**) Representative flow plot of

OX40, 41BB, GITR, CD96, CD96 and CD74 expression in TIL and PBMC from patients with HNSCC and HD PBMC. Normalized to mode scales as a percentage of the maximum count. The correlation between the percentage of HNSCC TIL Tregs expressing BATF and marker genes were shown in scatter plots. The correlation coefficient between the expression BATF and marker genes were calculated by Pearson correlation and p value were calculated by t test. Each dot indicates an individual replicate.