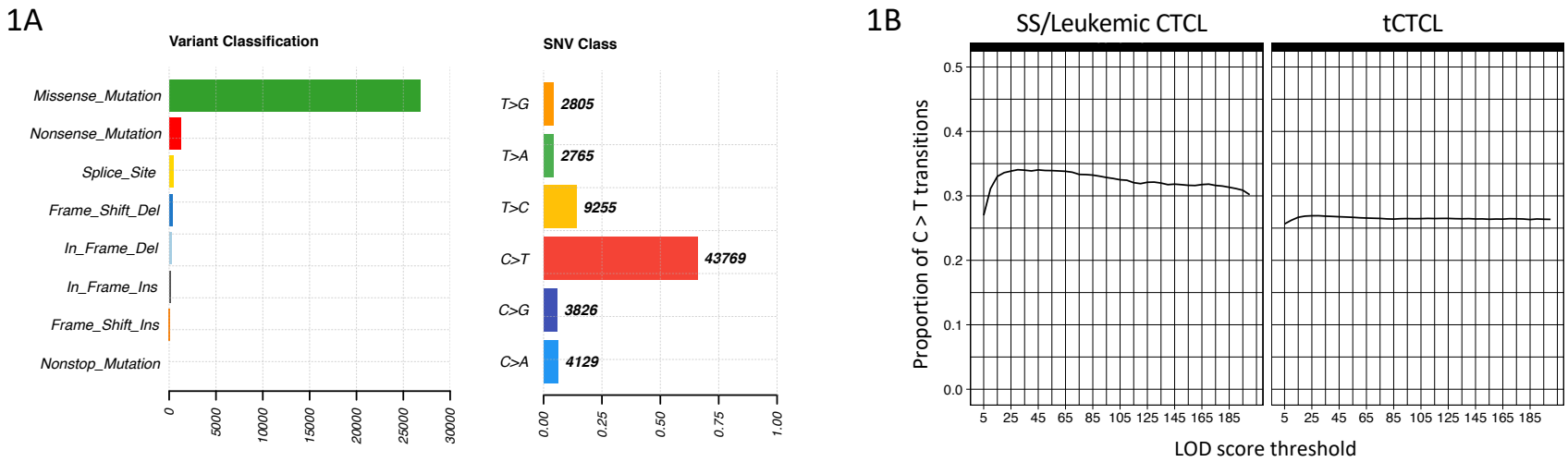
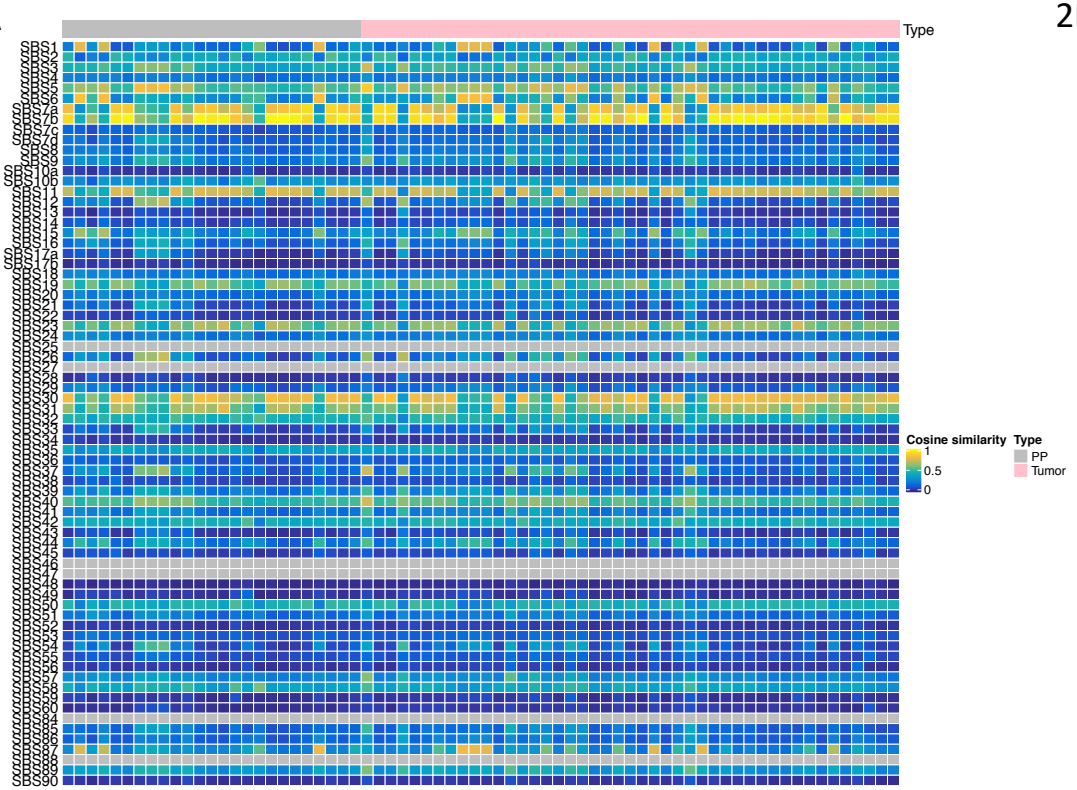
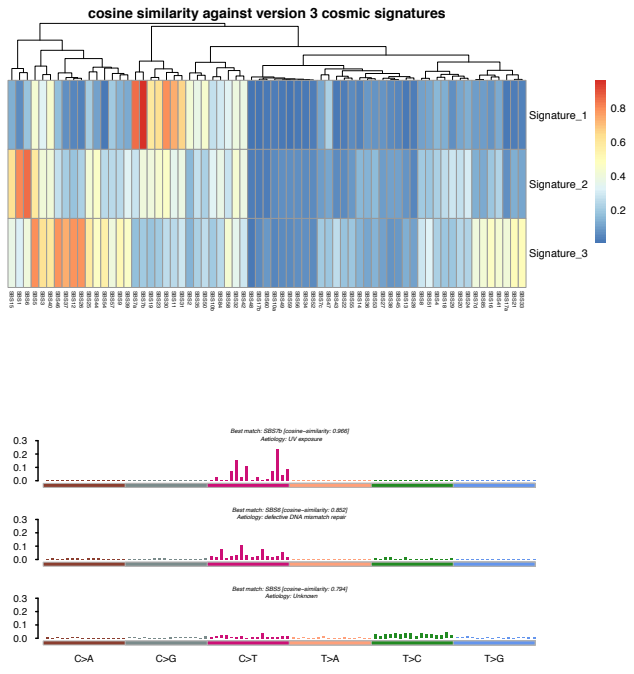


Supplementary Fig. S1. WES mutation variant classification and distribution of the proportion of C>T/G>A transitions by log odds (LOD) scores. A. Somatic single nucleotide variants (SSNVs) and small indels calling from WES by comparing TT and PP samples to matched normal germline (n=41 patients) or panel of normal (n=13 patients). The predominant mutations are missense mutations (97.1%), and the predominant SNV class is C>T transitions (65.8%). B. The distribution of the proportion of C>T/G>A transitions (y-axis) and LOD scores of mutation calls (x-axis). LOD scores from Mutect outputs are used to estimate sequencing quality, with increasing LOD score indicating higher sample quality (Hu. et. al, Nature Communications 2019). SS/leukemic CTCL36 (peripheral blood samples) and tCTCL samples (PP and TT, FFPE samples) show comparable LOD scores and sequencing quality.

2A

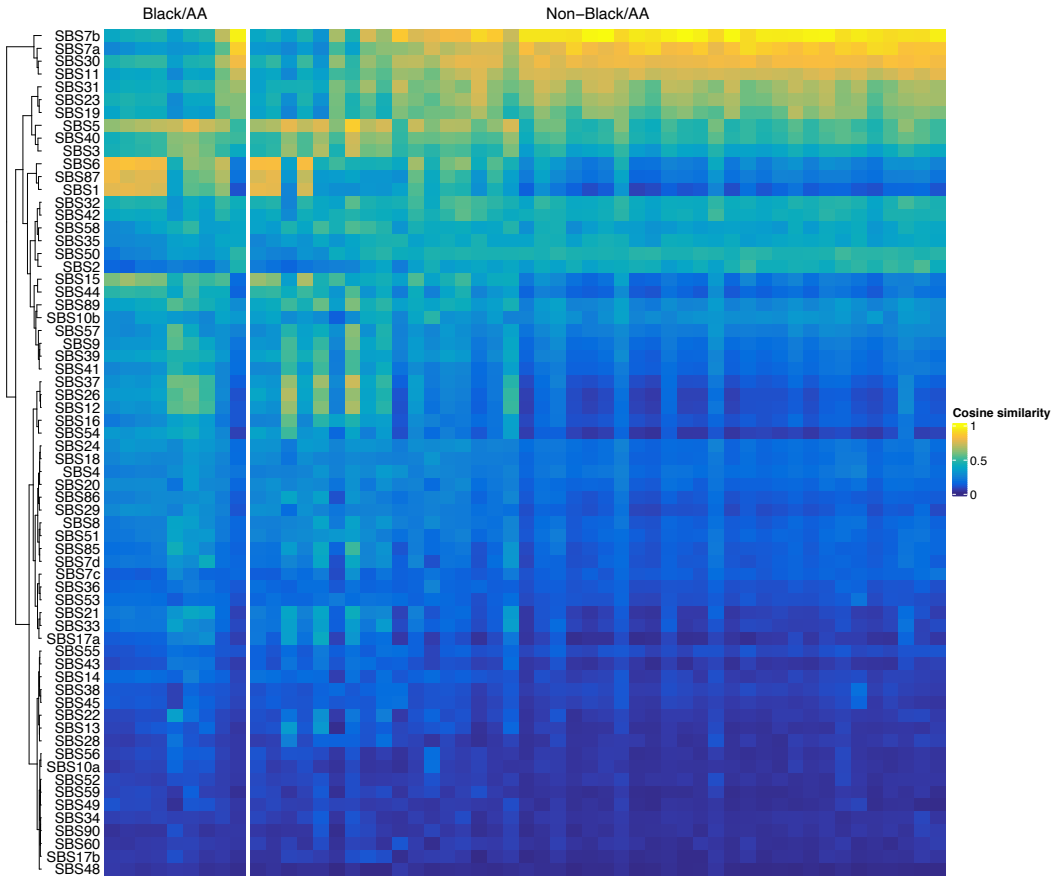


2B

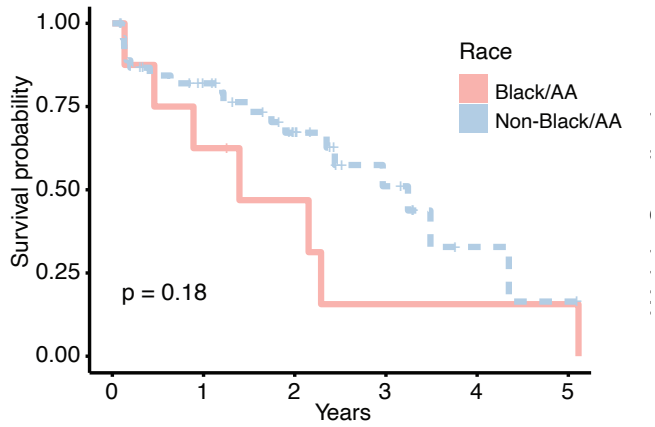


Supplementary Fig. S2. Mutational processes that contribute to the genomic landscape of tCTCL. A. Plot represents contribution of COSMIC v3.2 mutational signatures in each tissue sample (n= 70 samples from tCTCL patients, TT in pink, PP in gray) using MutationalPatterns, with highest cosine similarity to signatures SBS7a, SBS7b and SBS30. Cosine similarity is color coded to the right. B. At the cohort level, non-negative matrix factorization (NMF) methods using Maftools revealed best match to COSMIC v3.2 mutational signatures SBS7b, SBS6 and SBS5.

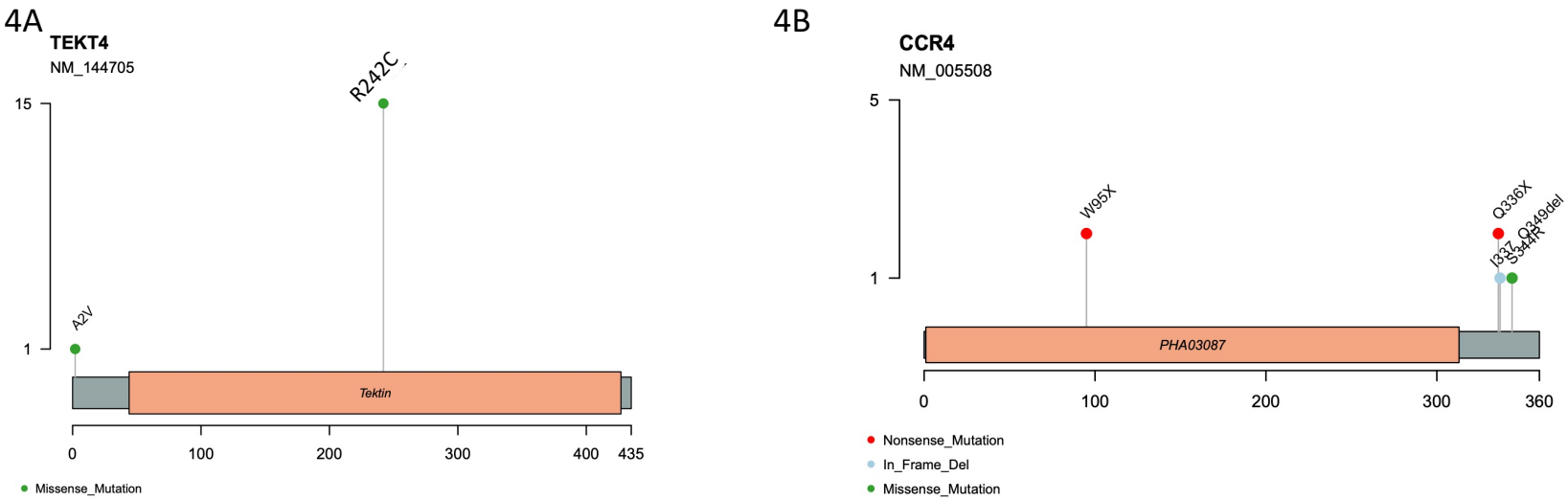
3A



3B

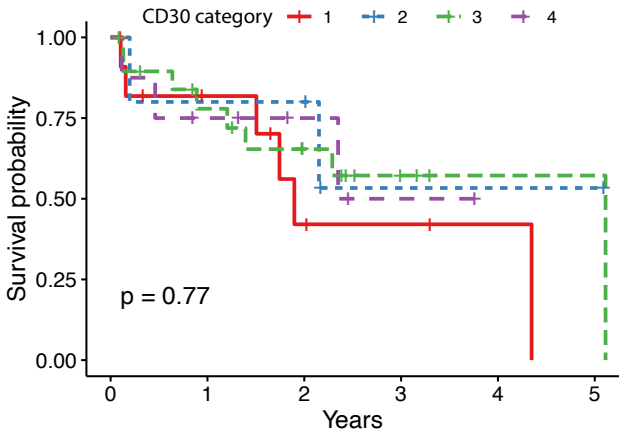


Supplementary Fig. S3. A. MutationalPatterns analysis. Heatmap of cosine similarities between the mutational profile of each sample and COSMIC v3.2 mutational signature in Black/AA vs non-Black/AA patients and ranked by SBS7a-d. B. Survival probability (y-axis) of Black/AA patients (pink) versus non-Black/AA patients (light blue) from time of LCT (years, x-axis).



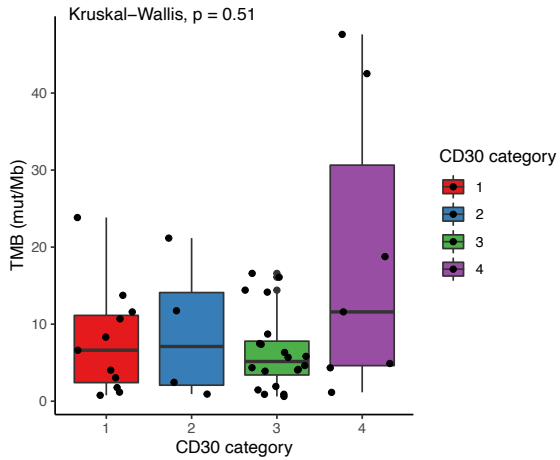
Supplementary Fig. S4. Mutations in representative predicted driver genes by dNdScv and MutSigCV. Schematic of somatic mutations in TEKT4 (A) and CCR4 (B).

5A

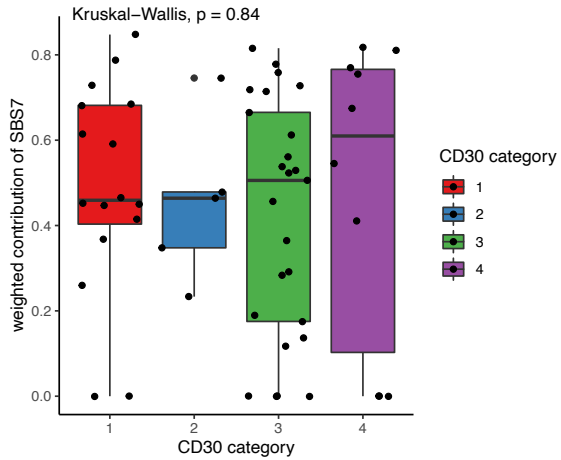


CD30%	Expression Category
1% or less	1
>1% to <10%	2
10-50%	3
50% or higher	4

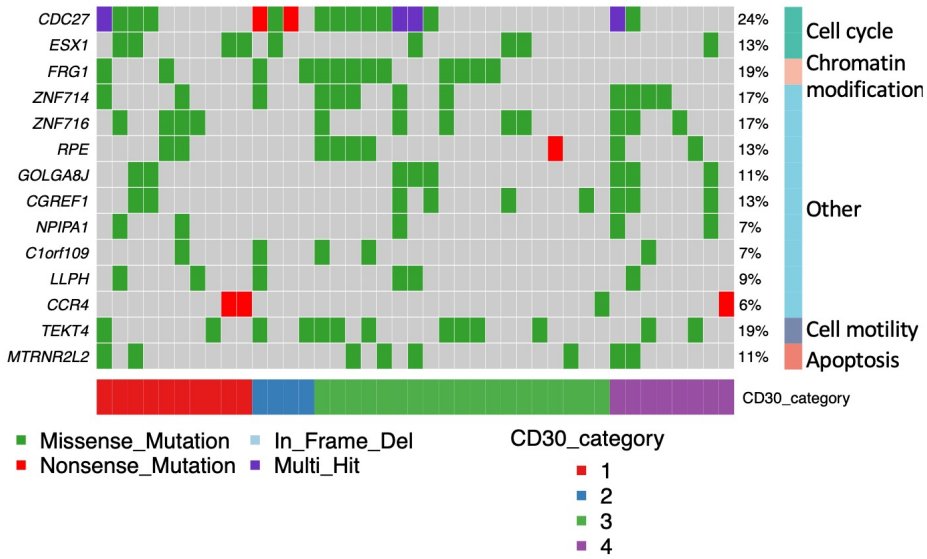
5B



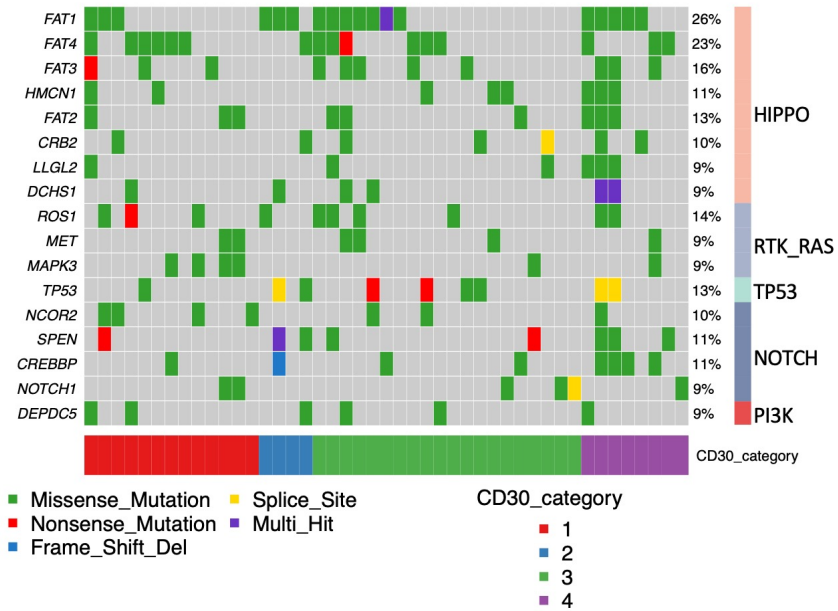
5C



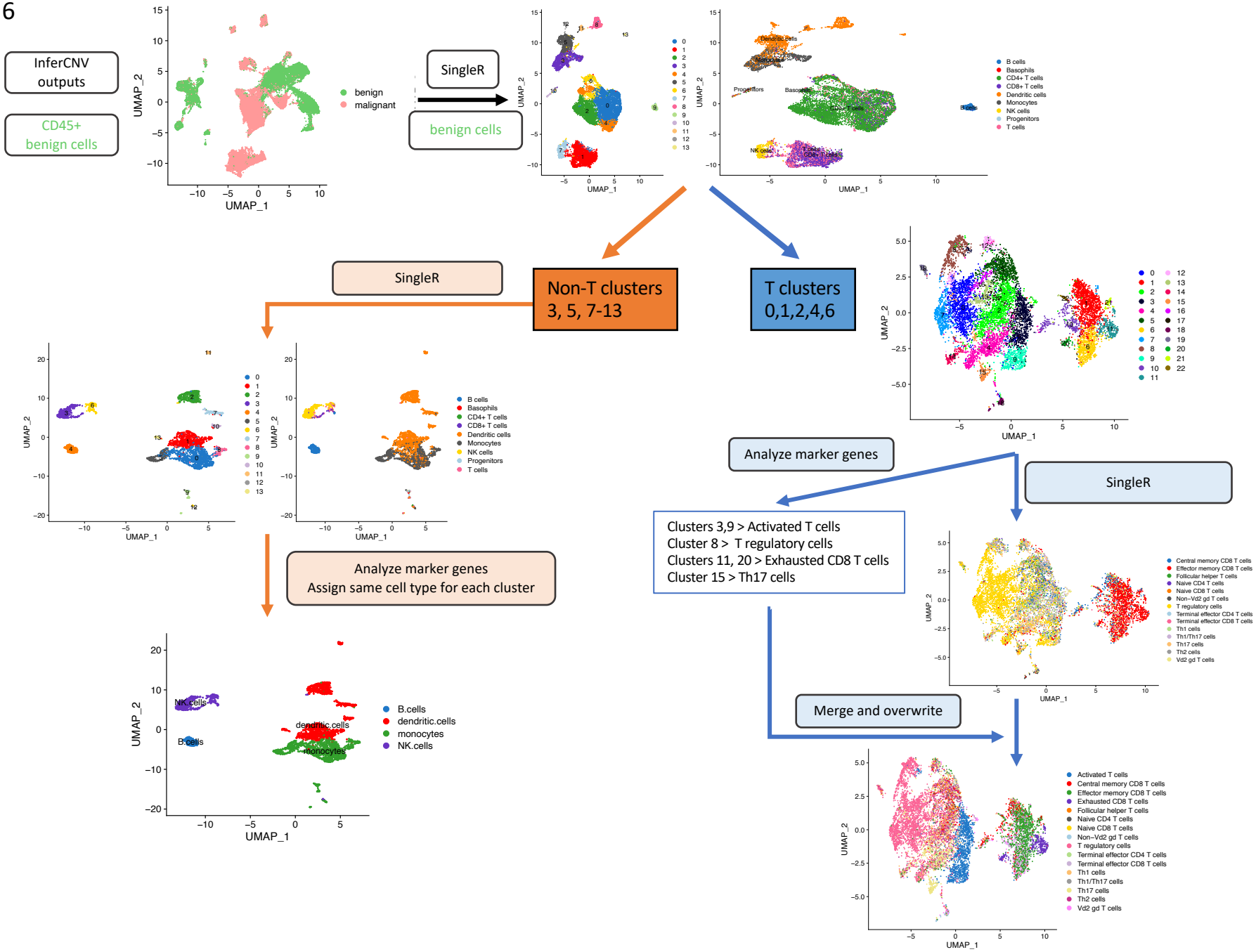
5D



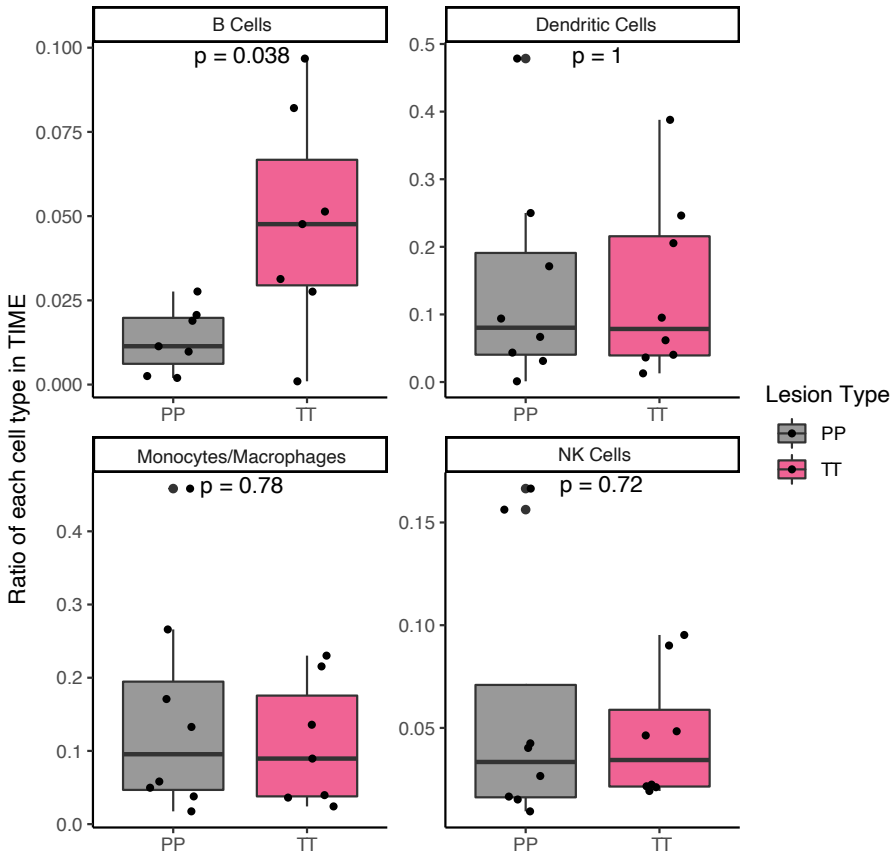
5E



Supplementary Fig. S5. WES genomic data stratified by CD30 expression. CD30 expression categories, 1: ≤1% (red), 2: >1% - <10% (blue), 3: 10-50% (green), 4: ≥50% (purple). A. Overall survival probability (y-axis) against time from onset of LCT to time of death or last follow up (in years, x-axis). B. TMB (mut/MB) in y-axis; CD30 expression categories (x-axis). C. Weighted contribution of SBS7 (UV signature) in y-axis; CD30 expression categories (x-axis). D. Oncoplot of predicted driver genes by dNdScv and MutsigCV (n=70 samples) and E. most recurrently mutated pathway genes (n=70 samples) in tCTCL. CD30 expression categories color coded at bottom.



Supplementary Fig. S6. Workflow for the characterization of benign immune cell types in the tCTCL TIME. In brief, benign immune cells (benign T-cells from inferCNV output and all remaining CD45+ immune cells) were clustered (Methods - data processing) and divided into T-cell clusters versus non-T cell clusters first based on SingleR. Results from SingleR annotation were further validated by cross-referencing with marker genes (Supplementary table S10). For non-T cell clusters (clusters 3, 5, 7-13), extraction of differentially expressed marker genes was performed with the FindAllMarkers function of Seurat, and the clusters were confirmed orthogonally by manually curated marker genes (Supplementary Table S10) and manually annotated/confirmed as B cells, dendritic cells, macrophage/monocytes, and NK cells. To further refine the benign T-cell subtypes (clusters 0, 1, 2, 4, 6), we performed clustering of benign T-cells and identified clusters of activated T cells (cluster 3, 9), T regulatory cells (cluster 8), exhausted CD8+ T cells (cluster 11, 20), and Th17 cells (cluster 15) based on differentially expressed genes in the T-cell clusters and known marker genes for these cell types. We then used SingleR to determine the remaining T-cell subtypes.



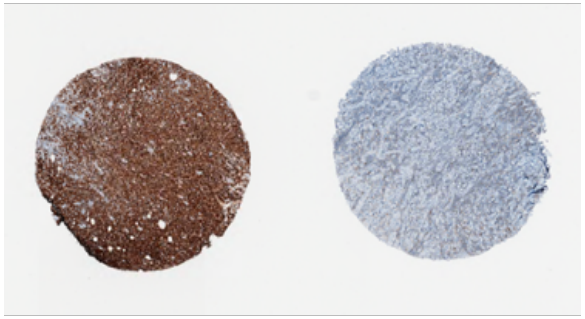
Supplementary Fig. S7. Characterization of non-T cell immune composition in tCTCL TIME. Boxplots comparing the proportion of B-cells (top left), dendritic cells (top right), macrophages (bottom left) and NK cells (bottom right) (y-axis) in PP vs. transformed tumors (x-axis) in the scRNAseq dataset.

Cell Type

- Malignant T cells in TT
- Malignant T cells in PP
- Naive CD4 T cells
- Follicular helper T cells
- Terminal effector CD4 T cells
- Activated T cells
- T regulatory cells
- Th1 cells
- Th1/Th17 cells
- Th17 cells
- Th2 cells

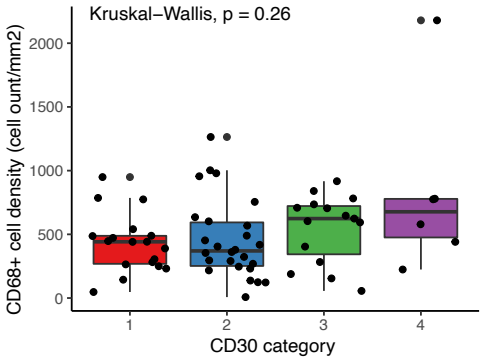
Supplementary Fig. S8. Kinetics of 55-gene malignant T-cell oncogenic program in tCTCL (Fig. 4C) plotted along pseudotime (x-axis). Cell types dotted in colors (e.g. malignant T-cells in PP – grey, malignant T-cells in TT – magenta, benign T-cells in remaining color schemes). Gene expression level based on scRNAseq data (y-axis). Supplementary Table S13 and Fig. 4D.

9A

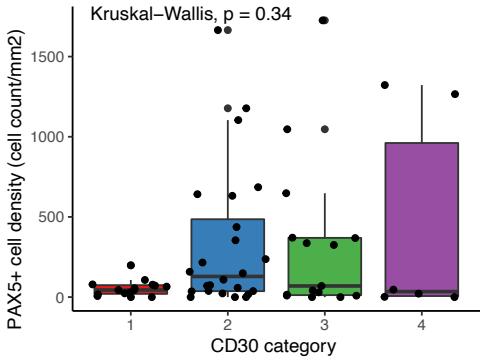


CD30%	Expression Category
1% or less	1
>1% to <10%	2
10-50%	3
50% or higher	4

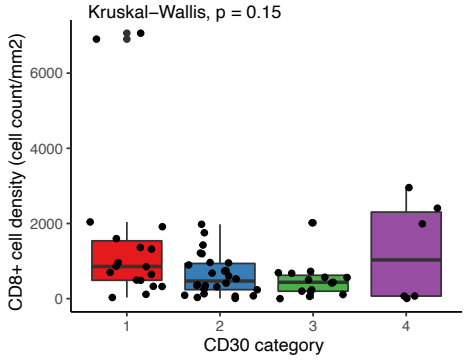
9B



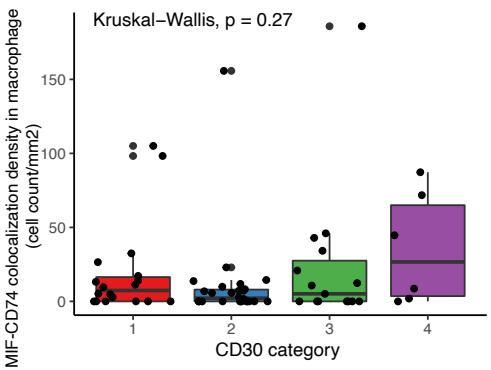
9C



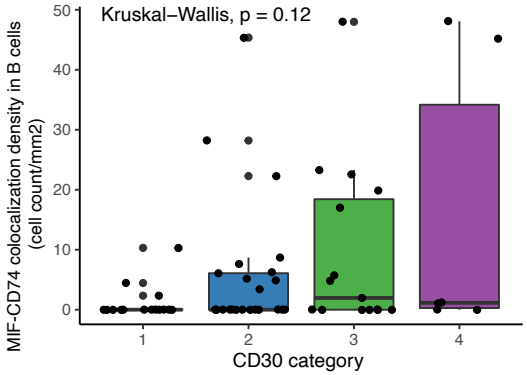
9D



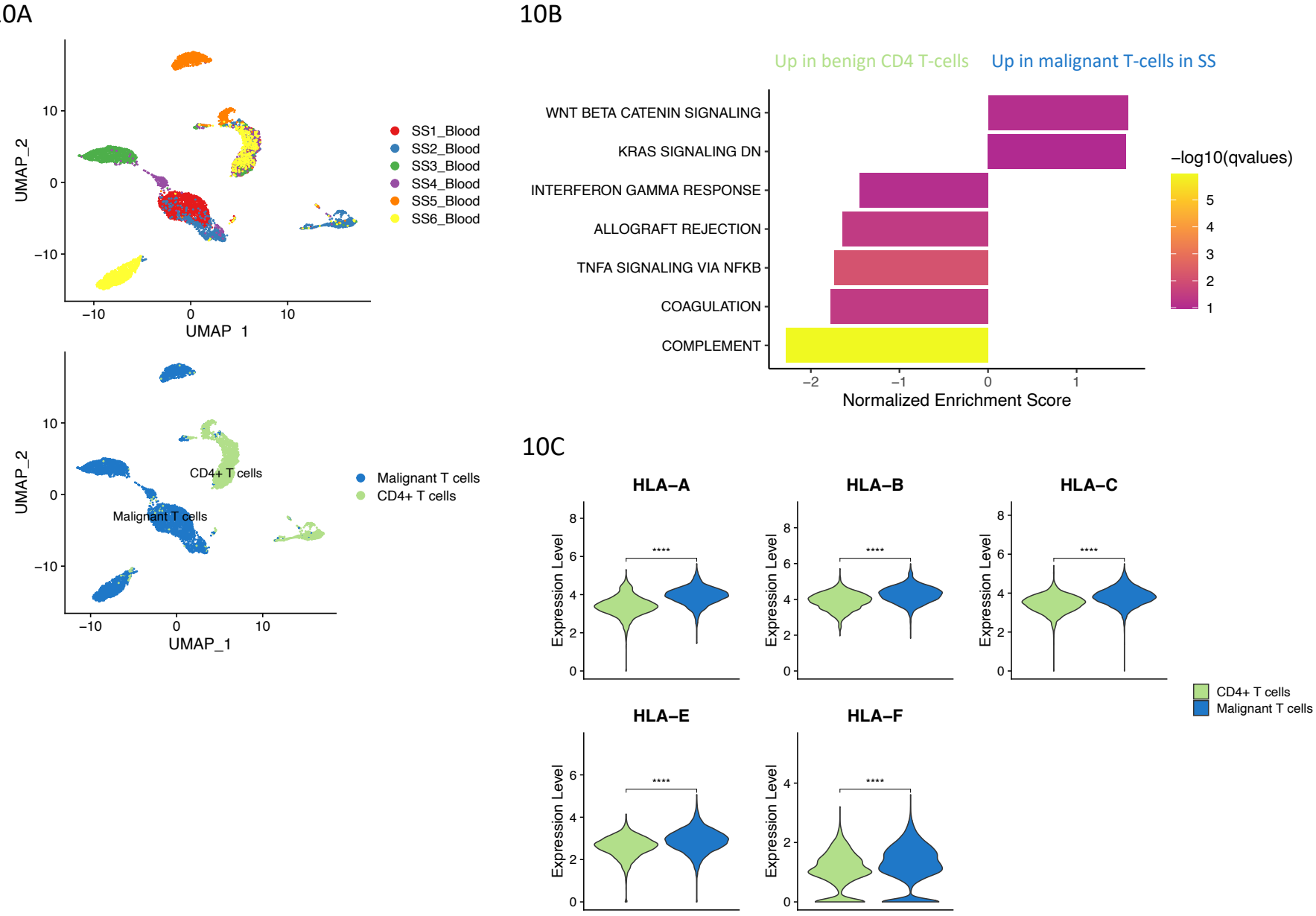
9E



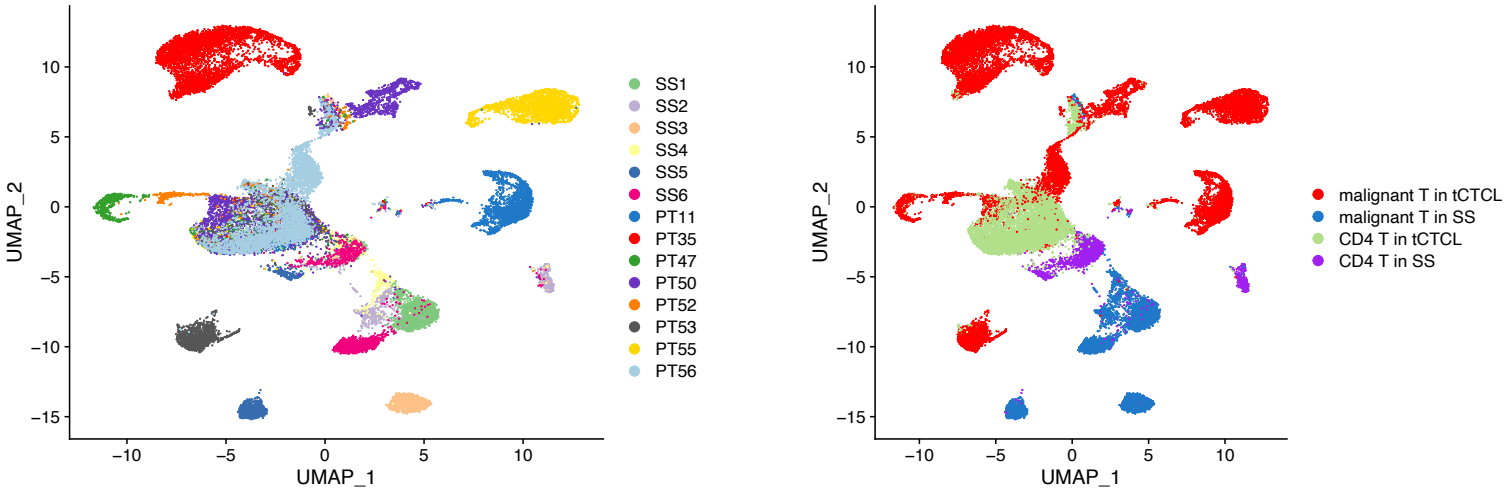
9F



Supplementary Fig. S9. Analysis of tCTCL TIME by CD30 expression. CD30 expression categories, 1: ≤1% (red), 2: >1% - <10% (blue), 3: 10-50% (green), 4: ≥50% (purple). **A.** Select images of TMA cores stained with anti-CD30 antibodies by immunohistochemistry. Density of **B.** CD68+ macrophages, **C.** PAX5+ B-cells, **D.** CD8+ T-cells, **E.** colocalization of MIF in malignant T-cells and CD74 in macrophages, and **F.** colocalization of MIF in malignant T-cells and CD74 in B-cells across 4 categories of CD30 expression.



Supplementary Fig. S10. Charting the malignant T-cell signature in SS by scRNAseq (Herrera cohort; ECCITE- seq). **A.** UMAP of single-cell profile of blood samples from 6 SS patients (SS1 to SS6, top) colored by malignant cell status (bottom; Supplementary Methods), with clear separation of malignant T-cell clusters (blue) from benign CD4 T-cells (green). **B.** GSEA comparing malignant T-cells in SS versus benign CD4+ T-cells. Significant enrichment of genes in Wnt-beta catenin and KRAS signaling pathways and downregulation of genes in the down-regulation of IFN-g, TNF-a and complement pathways. Normalized enrichment score (NES, x-axis) reflects the extent of enrichment and allows comparison across gene sets. Listed pathways are ranked by their NES and colored by their significance. C. Violin plots of distribution of HLA-A, B, C, E, F gene expression in malignant T-cells (blue) and benign CD4 T-cells from SS1-SS6 patients, with benign CD4 T-cells showing lower MHC-I expression in this cohort. **** denotes p<0.001.



Supplementary Fig. S11. (Left) UMAP of malignant T-cells and benign CD4 T-cells from SS patients (SS1 to SS6, blood samples, Herrera cohort) and 8 tCTCL patients (PT11, 35, 47, 50, 52, 53, 55, 56; PP + TT). (Right) Malignant T-cells from tCTCL patients (red, PP+TT) and SS patients (blue). Benign CD4 T-cells from tCTCL patients (green) and SS patients (purple).