

## Supplemental Methods

**Patient samples and cell isolation** Peripheral blood from CTCL patients and healthy donors was collected in lithium heparin tubes at the Yale Cancer Center after obtaining written informed consent and following all regulations in accordance with the Yale Human Investigational Review Board. Peripheral blood mononuclear cells (PBMCs) were immediately separated from whole blood by Ficoll density gradient. Total CD4<sup>+</sup> T cells were purified via either MACS CD4<sup>+</sup> negative selection kit (Miltenyi Biotec) or RosetteSep™ human CD4<sup>+</sup> T cell enrichment cocktail kit (STEMCELL Technologies) under the manufacturer's protocol. Malignant CD4<sup>+</sup> T cells were purified via MACS CD4<sup>+</sup> negative selection kit supplemented with biotin-conjugated anti-CD26 and/or anti-CD7 plus anti-biotin microbeads, based on the phenotype of aberrant T cells identified by clinical flow cytometry.

**Cell culture and stimulation** The culture medium used in this study is RPMI1640 (Invitrogen) supplemented with 20% heat-inactivated human AB serum (Corning, 35060112), penicillin-streptomycin (100 U/mL, Invitrogen), L-glutamine (2 mM, Invitrogen), HEPES (10 mM, Invitrogen), sodium pyruvate (1 mM, Invitrogen), 1% MEM NEAA (Invitrogen), 50 mM beta-mercaptoethanol (Bio-Rad) and the following R&D Systems cytokines: human recombinant IL-2 (10 ng/ml), IL-15 (10 ng/ml) and IL-7 (5 ng/ml). Purified malignant CD4<sup>+</sup> T-cells were cultured at  $1 \times 10^6$ /ml with no stimulation for 4 days or were stimulated either immediately following isolation or after a 4-day rest. Stimulation consisted of 5 mg/ml plate-bound anti-CD3 + 1 mg/ml soluble anti-CD28 for 2 days followed by washing and a 2 day expansion in cytokine containing medium.

**Flow cytometry** Cells were blocked with human TrueStain FcX (Biolegend) and stained with fixable viability dye (ThermoFisher Scientific, Fixable Viability Dye) for 10 minutes on ice. Then without wash, fluorescent-conjugated antibody mixtures targeting indicated surface markers (CD3, CD4, CD26, CD7, CD82, all from Biolegend) were added for 15 minutes on ice. Cells were then washed, fixed with 1% paraformaldehyde in PBS, and held at 4°C. The CellTrace violet cell proliferation kit (ThermoFisher Scientific) was used for cell proliferation studies. Phospho-protein staining followed the published protocol<sup>1</sup> with some modifications. In brief, cells were starved overnight in culture medium without cytokines or serum, but containing 1% BSA, followed by cytokine cocktail (IL-2, IL-7, and IL15) stimulation for 30 min. Cells were then fixed in 1.5% formaldehyde for 10 minutes at room temperature by directly adding 16% formaldehyde stock into cultured cells. Fixed cells were permeabilized with ice-cold methanol for 30 minutes on ice. Permeabilized cells were stained with primary anti-phospho-protein antibody pJAK2, pSTAT5, pJAK3-FITC, pAKT1-PE/Cy7, pPI3K-PE (all from ThermoFisher Scientific) for 30 minutes on ice, washed, then incubated, if needed, with fluorescent-conjugated secondary antibody (Alexa488-Donkey anti-Rabbit, Jackson Immunoresearch) for 30 minutes on ice. Cell apoptosis was examined by Annexin V Apoptosis detection kit (eBioscience, 88-8007-72) following the manufacturer's protocol. All data was collected on CytoFlex LX (Beckman Coulter) and then analyzed with FlowJo 10.6.2 software. Paired t test was used for statistical comparisons.

**JAK inhibitor drug screening** Malignant CTCL cells were isolated from peripheral blood and cultured overnight in cytokine supplemented medium as described above. Drug screening was performed at the Yale Center for Molecular Discovery where 6000 cells/well

were dispensed into 384-well plates (Multidrop Combi, ThermoFisher Scientific) and cultured for 72 hours following addition of various concentrations of a JAK 2-selective inhibitor (Pacritinib, Entrectinib or Fedratinib), JAK 2/3 inhibitor AT9283, JAK 1/2 inhibitor (Ruxolitinib, Baricitinib, Filgotinib, Mometinib or Upadacitinib), or pan-JAK 1/2/3 inhibitor (Peficitinib, Tofacitinib, Delgocitinib or Cerdulatinib). JAK inhibitors were nanodroplet dispensed in three-fold dilutions (Echo 550, Labcyte). Positive and negative controls were 10% and 0.2% DMSO, respectively. Cell viability was measured using CellTiter-Glo (Promega) with luminescence measurements taken on a Synergy Neo2 plate reader (BioTek Instruments). Mean and standard deviation of positive and negative control wells were used to quantify signal-to-background and Z' values for each screening plate to ensure assay robustness. Drug data were normalized to the mean values of negative control (set as 0% effect) and positive control (set as 100% effect) wells using Microsoft Excel, then 50% inhibitory concentrations (IC50s) were calculated using GraphPad Prism (version 8.4.3). Nonparametric Mann-Whitney U Test was used for statistical comparison among groups.

**Single cell RNA library preparation and sequencing** Libraries were prepared by the Yale Center for Genome Analysis using Chromium Next GEM Single Cell 5' Library and Gel Bead Kit V2 (1000263), Chromium Single Cell Human TCR Amplification Kit (1000252), Dual Index Kit TT Set A (1000215), and Chromium Next GEM Chip K Single Cell Kit (1000286) according to the manufacturer's instructions. Following quality control (QC) assessment, libraries were subjected to Novaseq 6000 (Illumina) sequencing with 50,000 read pairs per cell for the Single Cell 5' Library and 10,000 read pairs per cell for the TCR library.

**Single cell RNA data processing and QC** We collected scRNAseq and scTCRseq data from 11 CTCL patient samples as well as three samples from healthy donors as controls. For each sample, raw 10X sequencing data was processed into a UMI count matrix by aligning the reads to the GRCh38 genome following the standard 10X Cell Ranger software. The downstream data analysis was performed using Seurat V3 R package<sup>2</sup>. In the quality control (QC) analysis, poor quality cells with  $nFeature\_RNA \geq 250$ ,  $nCount\_RNA \geq 500$ ,  $\log_{10}GenePerUMI > 0.8$ , and with mitochondrial gene percentages over 20% were excluded. Genes with zero count numbers were removed. In addition, TCR genes were removed to prevent clustering bias caused by the contribution of variable V(D)J transcripts in major variable components.

**Single sample analysis** The feature-barcode matrix for every single sample was first normalized and scaled with default settings in Seurat. Then the 2000 top variable genes were identified, which served as the input to principal component analysis (PCA) for dimensionality reduction. We applied Louvain algorithm for clustering and retained the 20 leading principal components as an input for further visualization. The UMAP embedding<sup>3</sup> was used to visualize the single cells on a two-dimensional space with a perplexity of 100. scTCRseq data were annotated with the V(D)J module of the Cell Ranger pipeline (10x Genomics, version 3.0.1). Then, the TCR data was added into corresponding Seurat objects as the metadata. Subsequently, the share nearest neighbor (SNN) graph was constructed by calculating the Jaccard index between each cell and its 20-nearest neighbors, which was then used for cell clustering based on Louvain algorithm (with a resolution of 0.01). The cluster with the largest proportion of cells belonging to the most abundant clonotype (indexed by clonotype-1) was identified as CTCL T-cells, while the cluster with a minimum of

clonotype-1 occupation was identified as normal T-cells. The remaining cluster with a medium number of clonotype-1 T-cells was inferred to be transitional T-cells after PHATE<sup>4,5</sup> was used to infer the trajectory of T-cell population evolution. Directionality of evolution was determined based on the interrogation of clonotype composition across clusters, which was then revalidated based on the changing pattern of single nucleotide variation and copy number variation, as described in detail below.

**Integration analysis** The Seurat integration strategy<sup>2</sup> was employed to identify common anchors for combining different samples and reduce technical variations across libraries. The integrated data matrix was then processed following the same workflow as described in the above single sample analysis. T-cells were stratified into three categories: normal T-cells, transitional T-cells, and CTCL T-cells, based on their cluster-wise clonotype compositions (cell numbers in each category for each patient are shown in Supplemental Table S3). Non-parametric Wilcoxon rank sum test was used to identify gene markers showing differential expression patterns from control T-cells to CTCL T-cells, which were visualized by heatmaps and violin plots. We then used the Enrichr<sup>6,7</sup> to perform pathway analysis and enrichment analysis with The Molecular Signatures Pathway Database (MsigDB\_Hallmark\_2020) to discover enriched biological pathways based on the list of identified DEGs.

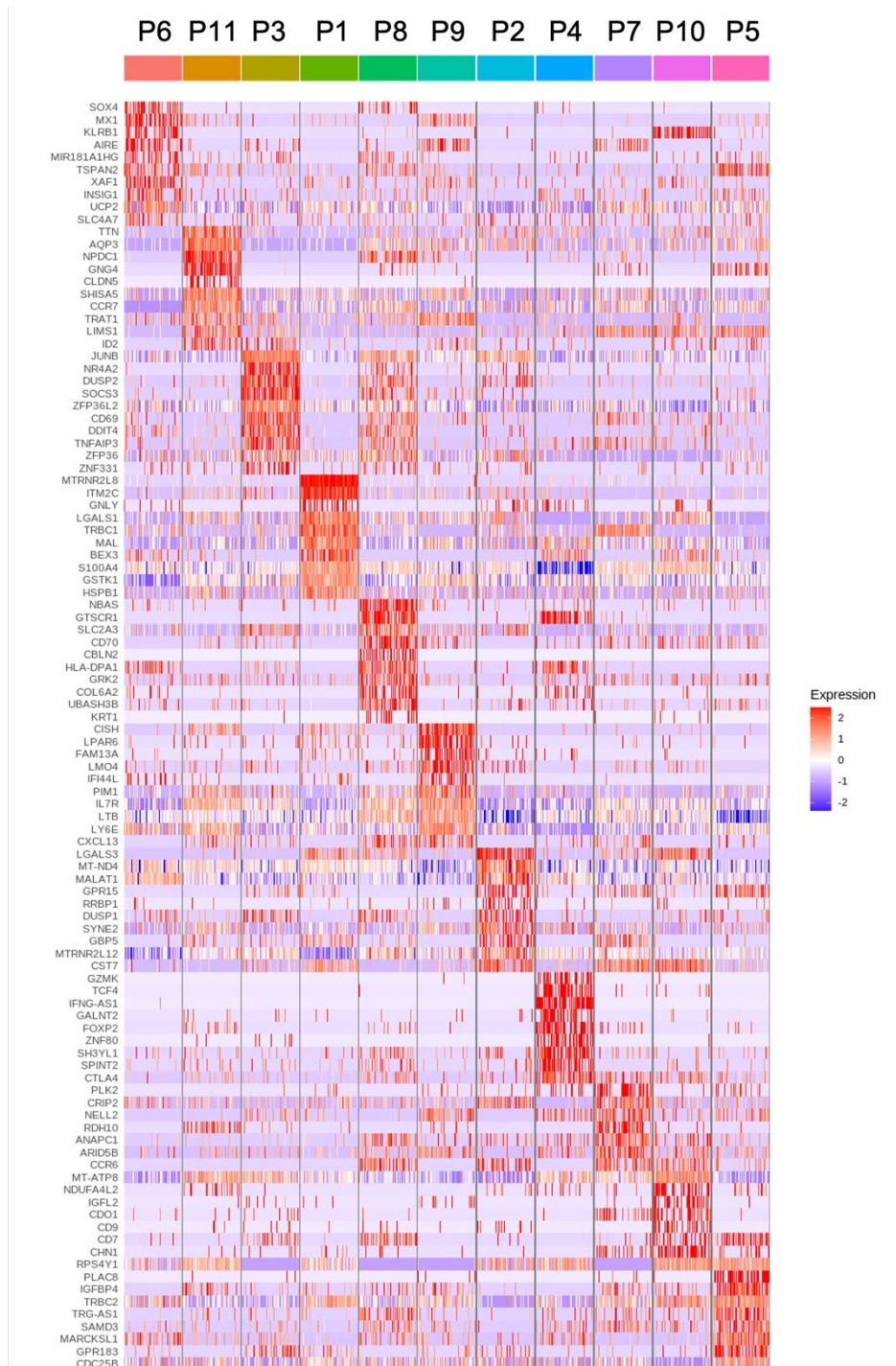
**Single nucleotide variation analysis** CellSNP mode 1<sup>8</sup> was leveraged to infer single nucleotide variation (SSNV) per cell for all samples. Briefly, a list of SSNVs, the common human variants from the 1000 Genomes Project, were compiled at the single cell level from the bam file of each sample, enabling us to track the change in number of SSNVs among different T-cell populations. SSNVs found in normal T-cells were considered as normal genomic variations or polymorphisms of the individual and hence used for background filtering. The remaining SSNVs in transitional T-cells and/or CTCL cells were profiled and served as the input to define cellular SSNV level, represented by the number of SSNVs found in each cell. The SSNV level was overlaid onto each corresponding PHATE embedding to depict its dynamical change across T-cell evolution.

**Copy number variation analysis** The R package InferCNV<sup>9</sup> was employed to infer copy number variation (CNV) at a single-cell level, as described in the Supplemental Methods. Specifically, using the transcriptome profiles of normal T-cells as background, we applied HMM to predict the status of detected SCNVs by assigning them into three states: amplification in CTCL, neutral, or deletion in CTCL. For each gene, we counted the number of samples that contained SCNVs located within the corresponding gene region and visualized the distribution of SCNVs across all chromosomes based on a combined track plot. Gene signatures which were previously reported to be associated with CTCL regulation were highlighted.

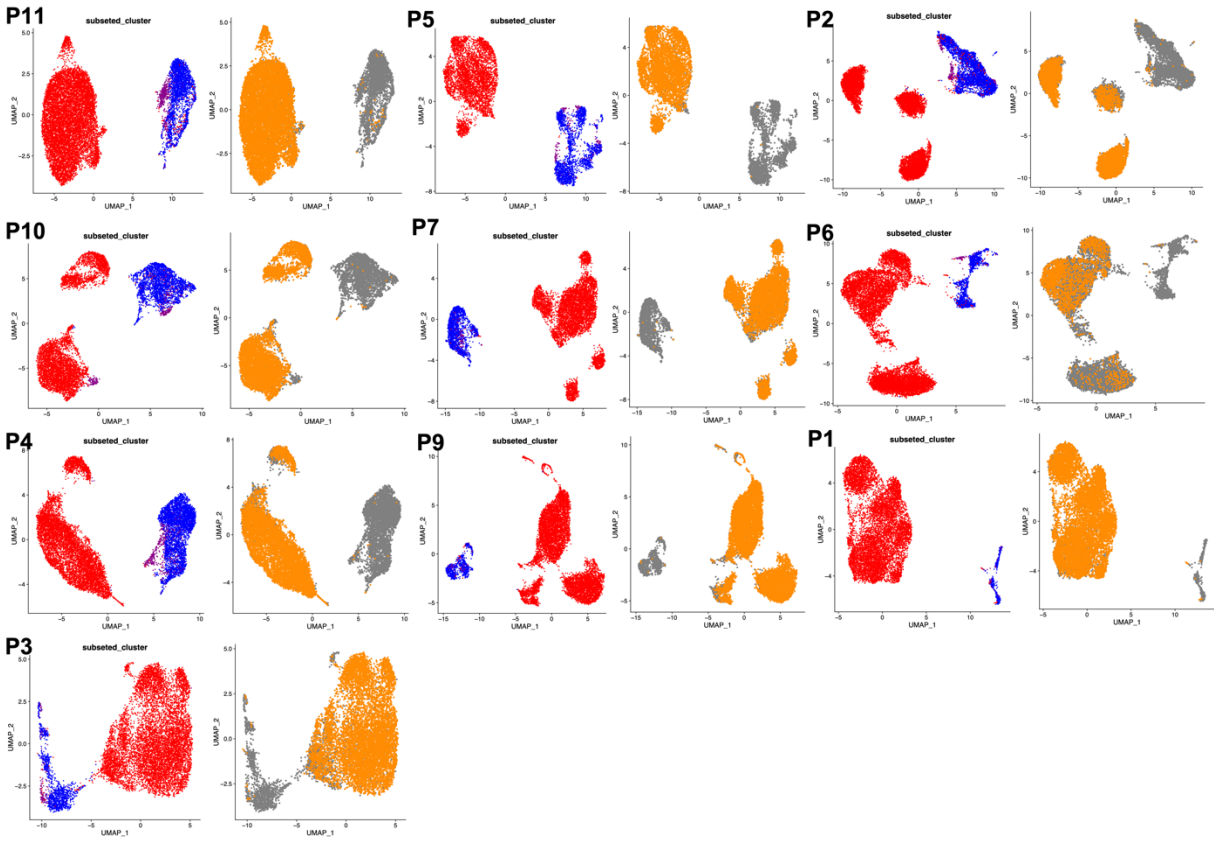
**CD82 knockout via Crispr-Cas9 system** Purified malignant T cells from the peripheral blood of patients were cultured for 3 days in the culture medium with cytokine cocktail mentioned above, and then CD82 gene knockout was performed following the IDT protocol (Electroporation of primary human CD34+ hematopoietic stem and progenitor cells) with some modifications, as described in the Supplemental Methods. In brief, the top two selected crRNAs of human CD82 (IDT pre-designed and recommended) at 1 to 1 ratio or one negative control crRNA (IDT 1072544) were annealed with the universal tracrRNA (IDT

1072532) to form the guide RNA duplex, which was then incubated with Cas9 protein (IDT 1081058) at room temperature for 20 minutes to generate ribonucleoproteins (RNPs). The CD82 Negative Control (NC) utilizes a safe-targeting guide to allow for proper control of toxicity from on-target DNA damage. The safe-targeting guide sequence used was GTGAGCTAGCACCAATTA and was obtained from IDT<sup>10</sup>. For electroporation,  $1-2 \times 10^7$  malignant T cells were resuspended in P3 buffer (Lonza V4XP-3024) plus electroporation enhancer (IDT 1075915) and transferred into a 100 ml nucleocuvette. Electroporation was performed using the Lonza 4D-Nucleofector X with program E0-115. For optimal cell recovery, 1 ml pre-warmed complete culture medium was added to the nucleocuvette immediately post electroporation and then transferred into a cell culture incubator for 2 hours resting before adding cytokines into the medium. 2-3 days post electroporation, cells were stained with violet proliferation dye as described, and seeded into anti-CD3 and anti-CD28 coated plates for cell activation for two days. The knockout efficiency of CD82, phosphorylation of potential downstream signaling molecules and proliferation of malignant T cells were detected via flow cytometry.

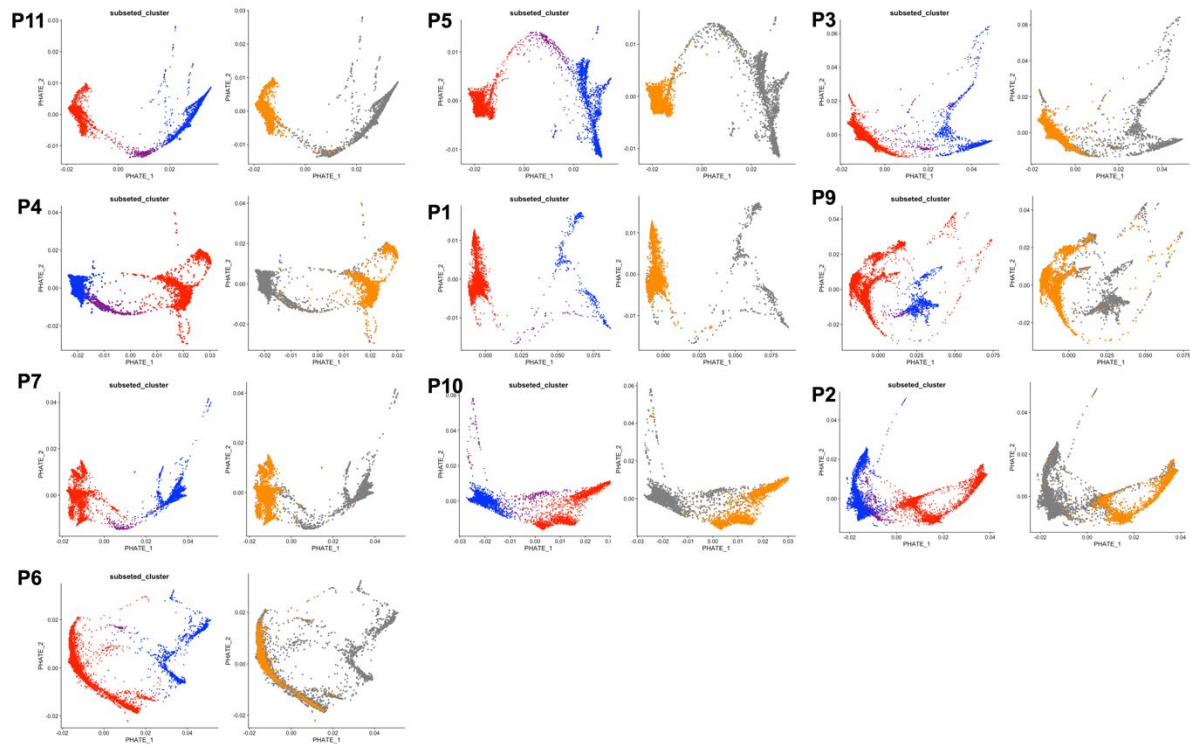
Visual Abstract: The visual abstract was created with BioRender.com.



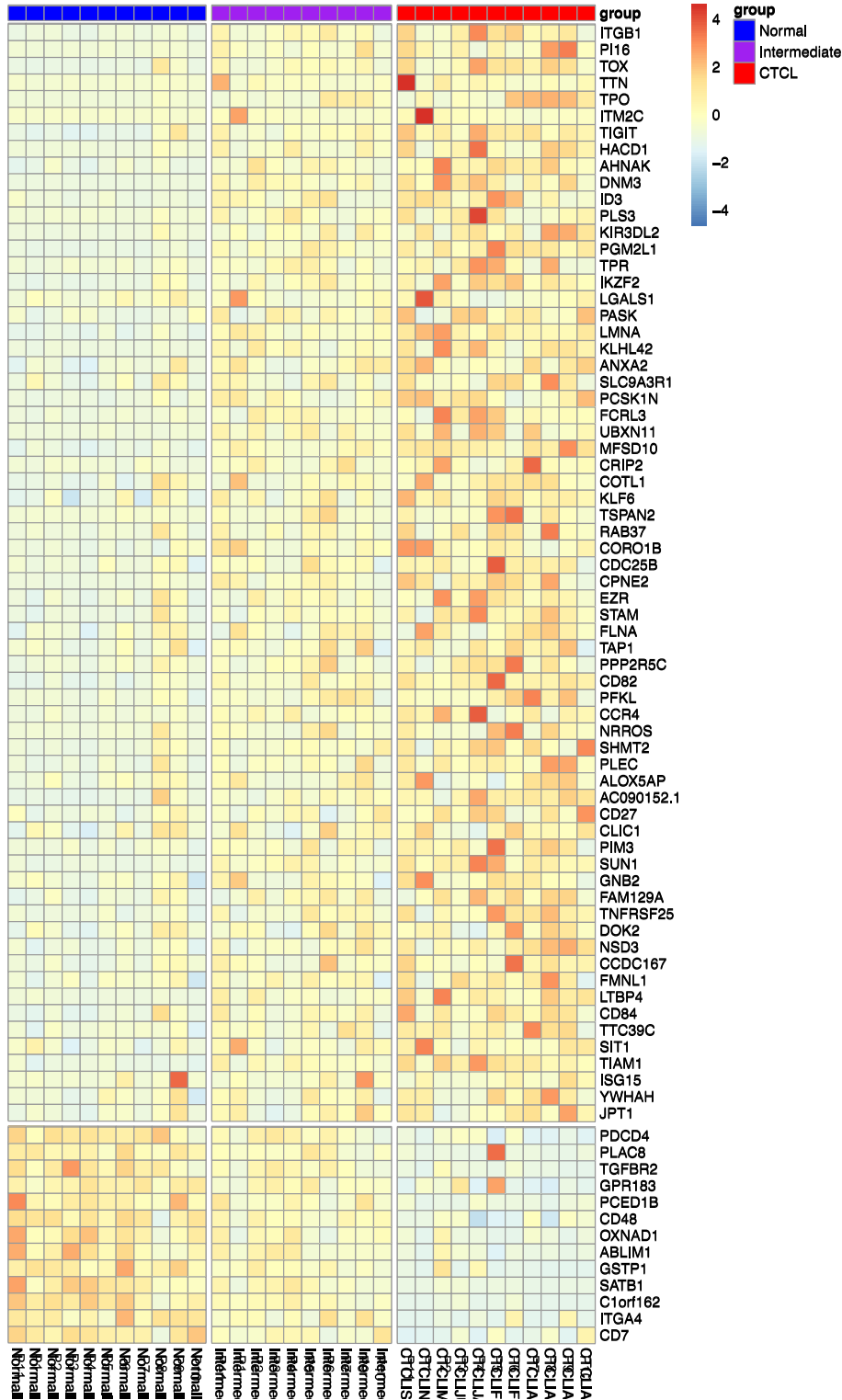
**Supplemental Figure S1. Inter-patient CTCL heterogeneity revealed by paired scRNAseq and scTCRseq.** Purified CD4+ T cells from the peripheral blood of 11 CTCL patients (P1-P11) underwent 5' scRNAseq and paired scTCRseq. Gene expression of cells taken only from annotated CTCL clusters in the unbiased PCA were then compared, and a heatmap displaying the top 10 differentially expressed genes is presented.



**Supplemental Figure S2. CTCL clonal annotation consistently reveals CTCL clonotype positive cells within the non-CTCL CD4+ T cell cluster.** Side by side UMAP plots of scRNAseq PCA (left) and CTCL scTCRseq clonal annotation (right) for each patient (P1-P11) are presented. CTCL cells (red) cluster separately from non-CTCL CD4+ T cells (blue + purple). The dominant TCR clonotype (yellow) highlighted over the same UMAP reveals CTCL clonotype positive cells within the non-CTCL CD4+ T cell cluster of each patient.

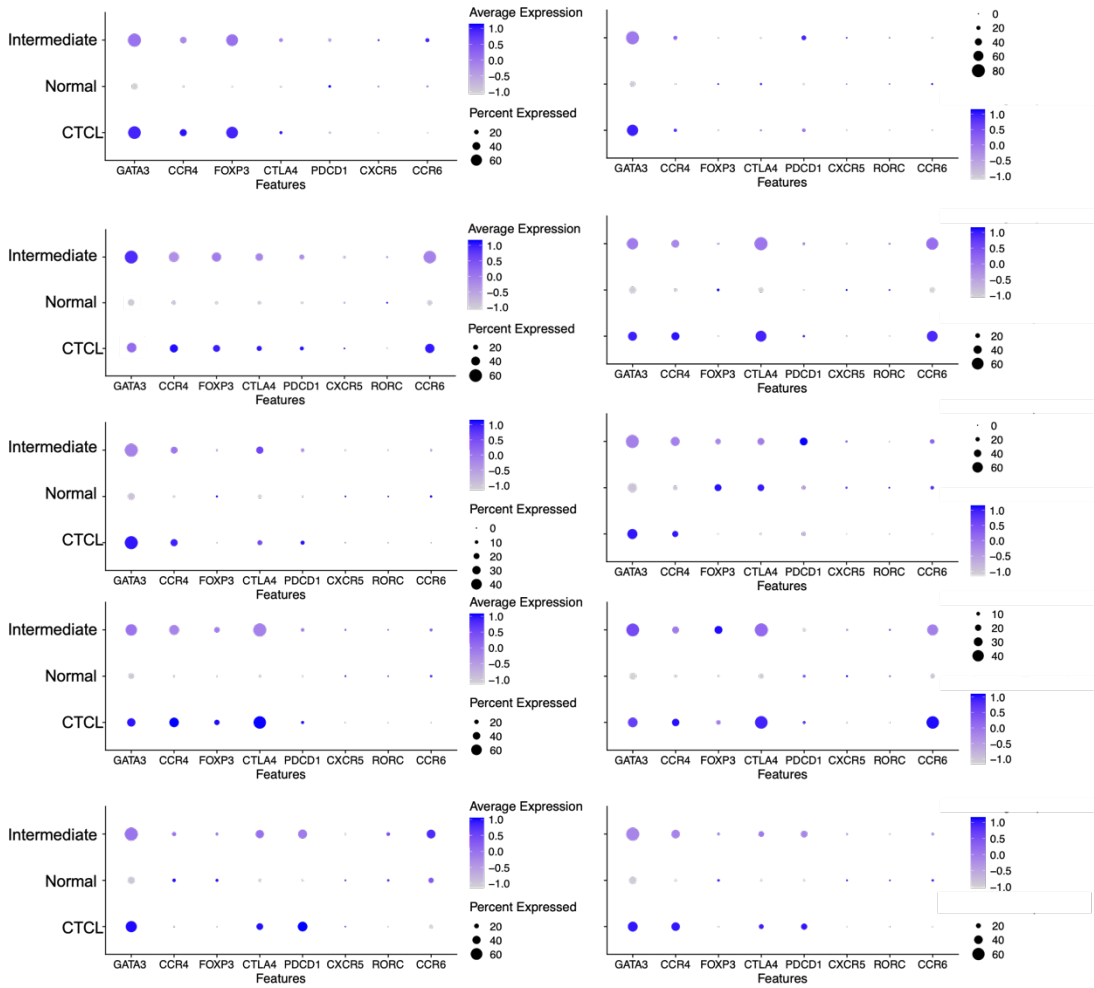


**Supplemental Figure S3. PHATE mapping reveals intermediate clonotype positive cells.** Side by side PHATE plots for each patient (P1-P11) reveal cell trajectory from normal CD4+ T cells (blue) to an intermediate population (purple) to malignant CTCL (red). The dominant TCR clonotype (yellow) highlighted over the same PHATE map reveals CTCL clonotype positive cells within the intermediate population.

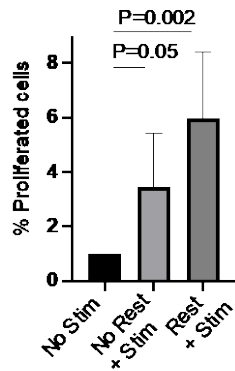


**Supplemental Figure S4. Heatmap of expression levels of the DEGs reported in Figure 3A for each patient separately.** The mean expression values for individual samples grouped by cell subset as normal, intermediate, or CTCL reveals heterogeneity between individuals, but with the trajectory of expression preserved in the transition from normal to intermediate to CTCL.

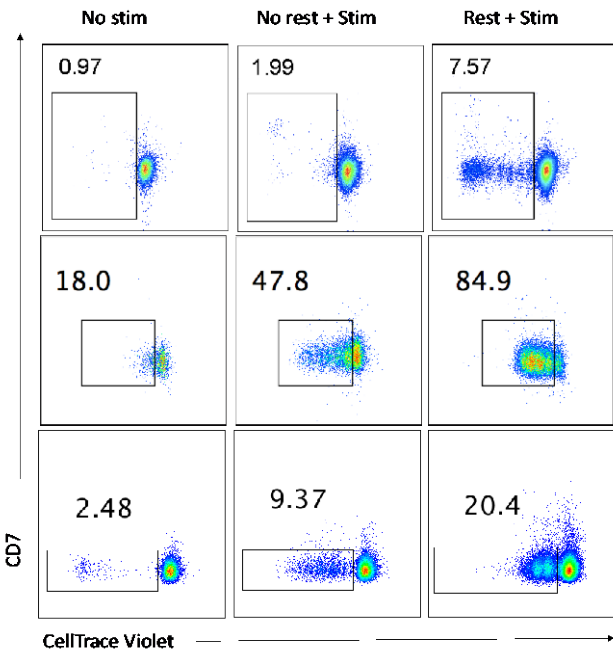




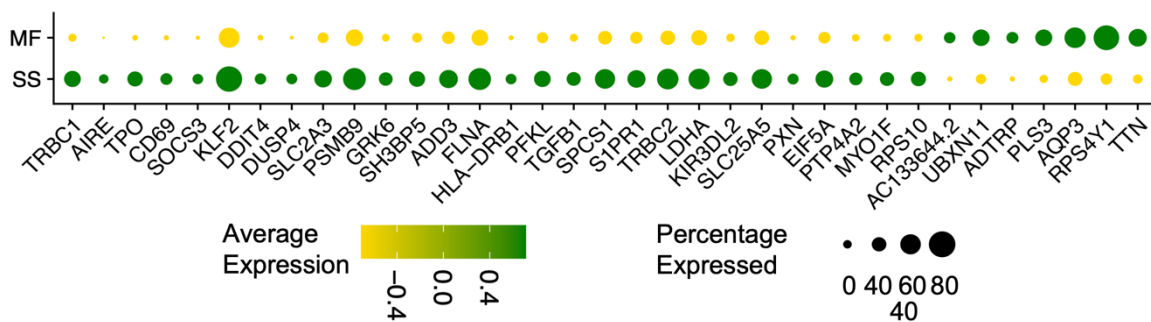
**Supplemental Figure S5.** DotPlot of each patient sample showing T-helper (Th) cell subset hallmark gene expression among three different categorized groups: CTCL, Normal and Intermediate.



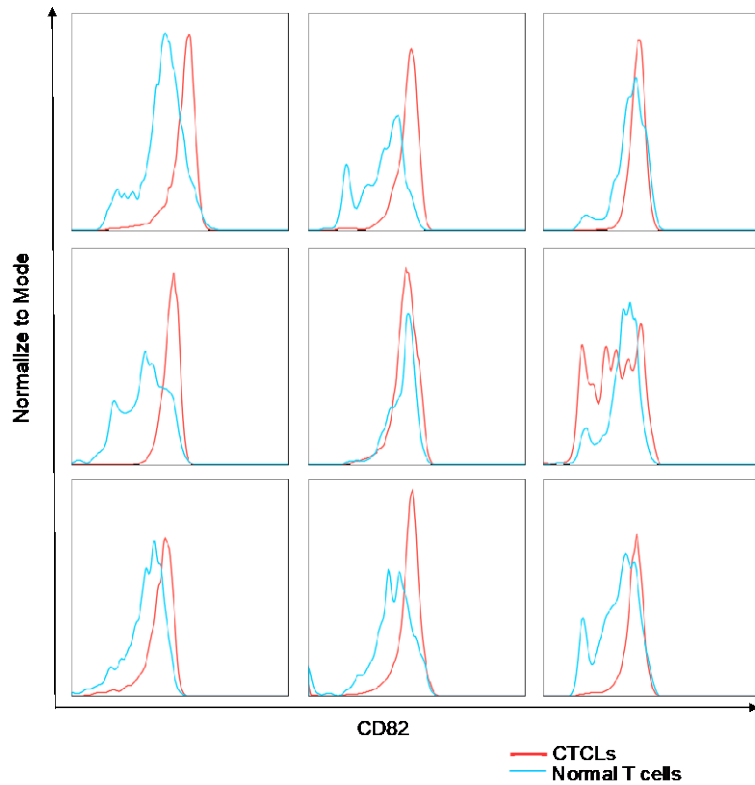
**Supplemental Figure S6.** The percentage of CTCLs that proliferate in response to TCR engagement is significantly increased following a period of rest in vitro. The data from Figure 4D is presented as Mean  $\pm$  SE (N=3 patient samples, unpaired t-test).



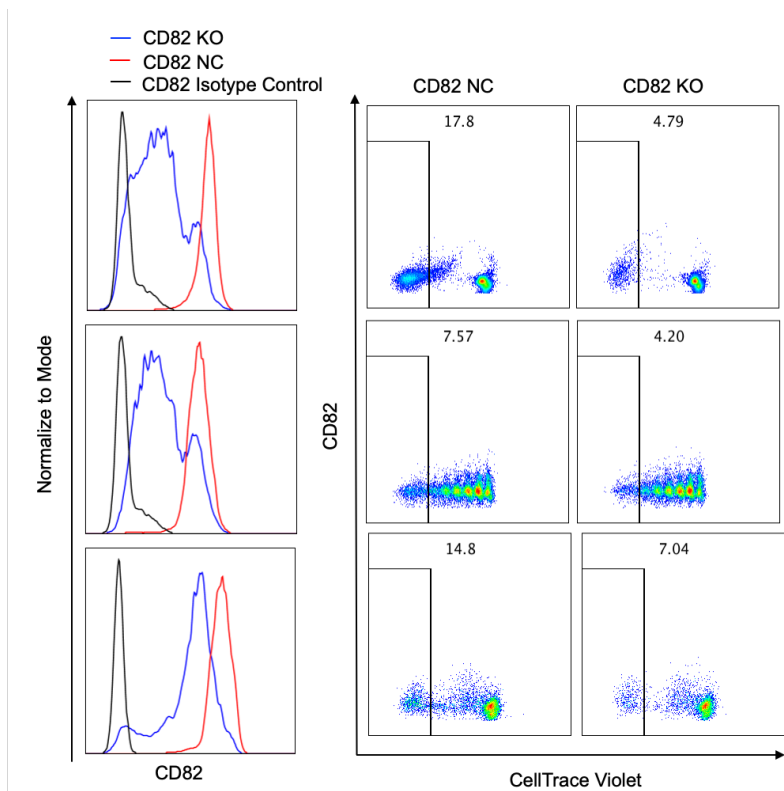
**Supplemental Figure S7. The percentage of CTCLs that proliferate in response to TCR engagement is significantly increased following a period of rest in vitro.** Flow cytometry data for three individual patient samples are shown (corresponding to Figure 4D). For two patients (top two rows) malignant cells were isolated and gated as CD4+7-; for the third patient (bottom row) the malignant cells were isolated as CD4+7- then further gated on TRBV+ cells for the flow cytometric analysis. Stimulation conditions are described in the methods and the Figure 4 legend.



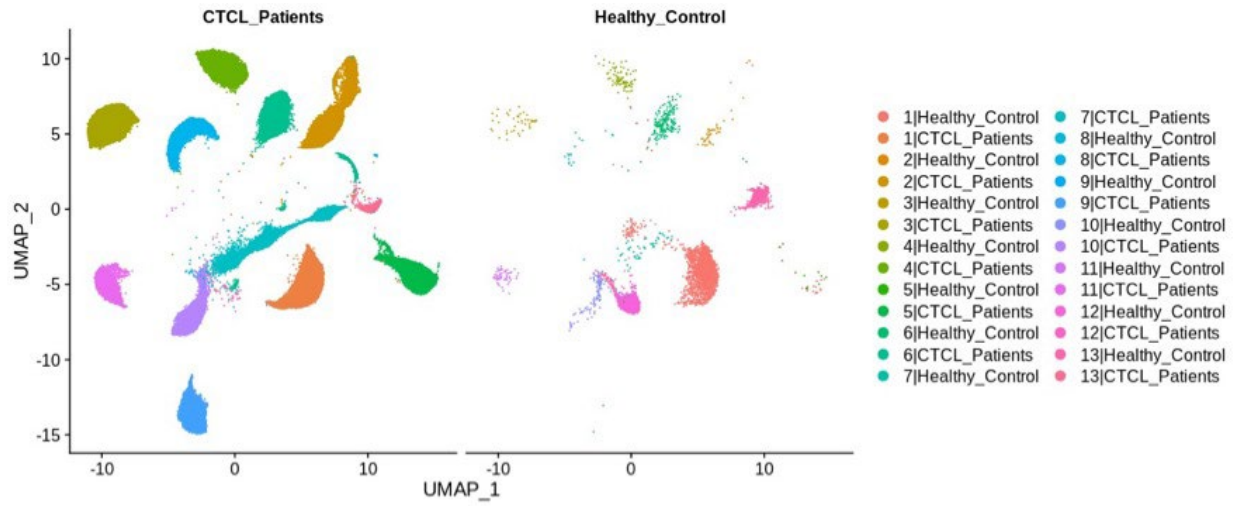
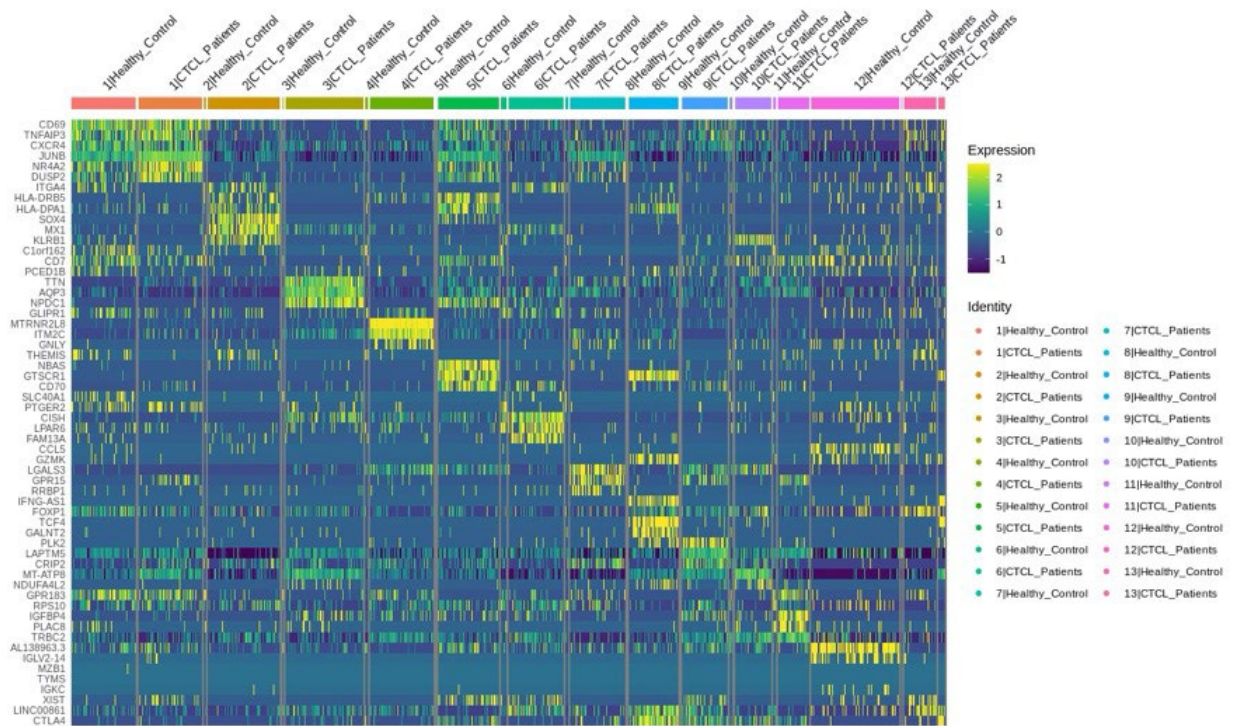
**Supplemental Figure S8. Differentially expressed genes in CTCL cells from patients diagnosed with MF vs SS.** Purified CD4+ T cells from the peripheral blood of 11 CTCL patients (P1-P11) underwent 5' scRNAseq and paired scTCRseq. Gene expression of cells taken only from annotated CTCL clusters in the unbiased PCA were then compared between patients diagnosed with MF vs SS. A dot plot of average gene expression is presented.

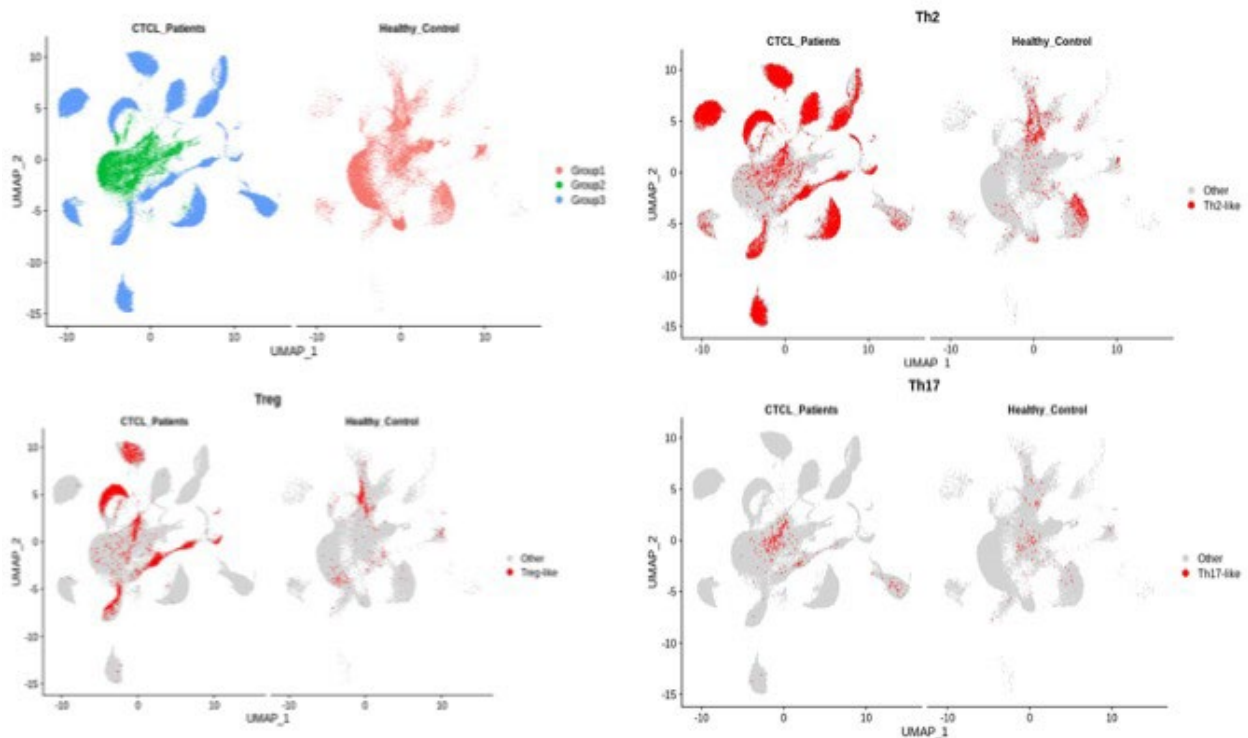


**Supplemental Figure S9. CD82 protein expression by CTCL cells vs normal CD4<sup>+</sup> T-cells from the same patient.** The flow cytometry data for each patient corresponding to Figure 5B is shown. PBMC from 9 CTCL patients were stained with CD82, CD3, CD4, CD7, CD26. Histograms show the level of CD82 expression on the normal CD4<sup>+</sup> T cell subset (blue) and the CTCL subset (red).



**Supplemental Figure S10. CD82 expression level correlates with proliferative capacity.** CD82-NC (negative control) and CD82-KO (knockout) cells were prepared from purified malignant CTCL cells from N=4 patient samples. Flow cytometry was used to assess the level of CD82 expression (left) and proliferative capacity based on CellTrace Violet dilution following CD3+CD28 stimulation (right). This figure depicts each individual experiment shown in Figure 5C-5E.

**A****B**

**C**

**Supplemental Figure S11. Analysis of normal T-cells that cluster in the same position as patient CTCL cells. (A)** Side-by-side view of the integrated UMAP of CTCL malignant cells from 11 CTCL patients (left) and overlaid only those normal T cells from healthy controls that appear in the same cluster positions as the CTCL cells (right) along with a heatmap **(B)** created using the Seurat FindConservedMarkers function to display the extent of shared similarly expressed genes in those overlaid subpopulations (indicated by the same cluster index). Based on CD4+ T-cell subtype annotation **(C)**, we did not find that the overlaid cell subpopulations from healthy donors belong to any particular T-cell subtypes.

**Supplemental Table S1.** Enriched signaling pathways with MsigDB Hallmark 2020 database via Enrichr<sup>6,7</sup> analysis.

Signaling pathways	gene ratio	Adjusted P-value	Genes
IL-2/STAT5 Signaling	24/199	3.80E-08	XBP1;AHNAK;PTGER2;CAPG;SLC2A3;TRAF1;IKZF2;TNFRSF1B;HOPX;BATF;TIAM1;NT5E;CCND3;KLF6;CCND2;SOCS1;SELL;IL2RB;TNFSF10;CTLA4;CD48;LTB;CCR4;PLEC
Interferon Gamma Response	23/200	1.40E-07	GZMA;STAT3;MX1;HLA-B;NLR5;TAP1;ISG15;HLA-A;SRI;PSMB8;PSMB9;SOCS3;MT2A;RNF213;SOCS1;CCL5;IL2RB;TNFSF10;STAT4;IRF7;TXNIP;HLA-DRB1;LY6E
Myc Targets V1	22/200	4.98E-07	NPM1;RPL34;RPS5;RPL22;RPLP0;RPS6;RPL6;PSMA7;PPM1G;EEF1B2;COPS5;PSMB3;RPS3;RPL14;SRSF2;PGK1;AP3S1;PABPC1;RPL18;HNRNPA1;RPS10;KPNB1
Inflammatory Response	21/200	1.80E-06	PTGER4;IFITM1;CD82;SEMA4D;CD70;PTGER2;RHOG;EMP3;SRI;ATP2B1;TNFRSF1B;KLF6;SELL;GPR183;CCL5;IL2RB;TNFSF10;IRF7;CCR7;CD48;LY6E
TNF-alpha Signaling via NF-kB	20/200	6.42E-06	DUSP4;PTGER4;PPP1R15A;PLK2;TAP1;SLC2A3;TRAF1;ATP2B1;SAT1;KLF2;EIF1;NR4A2;SOCS3;ZFP36;KLF6;GPR183;CCL5;JUNB;IER2;BIRC3
Apoptosis	17/161	1.75E-05	BCAP31;SATB1;PRF1;TAP1;HSPB1;SAT1;CDC25B;LGALS3;CCND2;LMNA;TNFSF10;RARA;PDCD4;TXNIP;TSPO;SPTAN1;BIRC3
mTORC1 Signaling	19/200	1.95E-05	PPP1R15A;XBP1;TPI1;SHMT2;INSIG1;RPN1;ITGB2;SLA;SLC2A3;SLC9A3R1;PFKL;DDX39A;COPS5;PPA1;PRDX1;PGK1;CALR;GAPDH;CTSC
p53 Pathway	18/200	5.99E-05	PPP1R15A;CD82;PLK2;PVT1;TAP1;SAT1;LDHB;CCND3;CCND2;SOCS1;RPL36;TXNIP;STOM;S100A4;RPL18;RALGDS;RPS12;S100A10
Interferon Alpha Response	12/97	6.78E-05	IFITM1;SELL;OAS1;MX1;IRF7;TAP1;HLA-C;TXNIP;ISG15;PSMB8;LY6E;PSMB9
Complement	16/200	6.13E-04	CPM;GZMA;RHOG;GATA3;GNAI2;PSMB9;PDP1;DPP4;GZMK;LGALS3;PPP4C;GNB2;CCL5;IRF7;CTSC;HSPA1A
KRAS Signaling Up	14/200	0.005362659	PPP1R15A;MAP3K1;SATB1;ITGB2;TRAF1;TNFRSF1B;ETS1;HDAC9;PSMB8;PCSK1N;CCND2;KCNN4;TRIB2;BIRC3
Reactive Oxygen Species Pathway	6/49	0.007917961	PRDX2;ABCC1;PRDX1;MBP;LSP1;JUNB
TGF-beta Signaling	6/54	0.012155542	PPP1R15A;FN1A;ID3;JUNB;RHOA;PPP1CA
IL-6/JAK/STAT3 Signaling	6/87	0.03046721	SOCS3;SOCS1;ITGA4;STAT3;LTB;CXCL13;TNFRSF1B
UV Response Up	10/158	0.035046335	DNAJB1;CCND3;RPN1;TAP1;TMBIM6;ARRB2;CDO1;AQP3;JUNB;ATP6V1F

**Supplemental Table S2.** Top 10 mutated genes (SSNVs) in the intermediate cells shared by  $\geq 50\%$  of CTCL patients analyzed (N=10), as detected by CellSNP analysis of scRNAseq data.

Gene	% patients
DNM3	70%
CCDC47	60%
AC016831.7	60%
RPTOR	50%
TOX	50%
SMYD3	50%
SKAP1	50%
TSHZ2	50%
HCG27	50%
MUC20-OT1	50%

**Supplemental Table S3.** Number of single cells used for analysis after filtering out low-quality cells.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Normal	393	3941	1234	3820	2077	1280	1608	37	855	2436	2539
Transitional	66	495	192	512	222	86	170	0	80	449	730
CTCL	9297	8124	9290	7867	4600	10968	6509	8688	8233	5467	10579

## Supplemental References

1. Krutzik PO, Nolan GP. Intracellular phospho-protein staining techniques for flow cytometry: monitoring single cell signaling events. *Cytometry A*. 2003;55(2):61-70.
2. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177(7):1888-1902 e1821.
3. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2018.
4. Moon KR, van Dijk D, Wang Z, et al. Author Correction: Visualizing structure and transitions in high-dimensional biological data. *Nat Biotechnol*. 2020;38(1):108.
5. Moon KR, van Dijk D, Wang Z, et al. Visualizing structure and transitions in high-dimensional biological data. *Nat Biotechnol*. 2019;37(12):1482-1492.
6. Xie Z, Bailey A, Kuleshov MV, et al. Gene Set Knowledge Discovery with Enrichr. *Curr Protoc*. 2021;1(3):e90.
7. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14:128.
8. Huang X, Huang Y. Cellsnp-lite: an efficient tool for genotyping single cells. *Bioinformatics*. 2021.
9. Tickle T TI, Georgescu C, Brown M, Haas B. inferCNV of the Trinity CTAT Project. <https://github.com/broadinstitute/inferCNV>. 2019.
10. Morgens DW, Wainberg M, Boyle EA, et al. Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat Commun*. 2017;8:15178.