SUPPLEMENTARY MATERIALS AND METHODS

Targeted single-cell RNA sequencing

Three 1-cm punch biopsies were performed on one patient (who was enrolled in the study NCT00177268 (University of Pittsburgh Institutional Review Board number 0411029)) at three different time points (1 week later and 4 weeks later from the first one), and whole-skin dissociation was performed using the Miltenyi Biotec Human Whole Skin Dissociation Kit on the same day as the biopsy. Targeted single-cell RNA sequencing was performed using the BD Rhapsody Single-Cell Analysis System (BD Biosciences, Franklin Lakes, NJ) using the BD Rhapsody Human Immune Response Panel, TCR panel, and a custom panel of 61 genes. A total of 996 cells were recovered from those three biopsies, but after quality control and double depletion, we analyzed 912 cells for Uniform Manifold Approximation and Projection and reference mapping.

The custom panel of 61 genes

The custom panel of 61 genes included the following:

1. Tumor suppressor: *FAT1*, *GLI3*, *PTEN*, *RASA2*, *RB1*, and *TP53*;
2. Apoptosis regulation: *BAG4*, *BCL7C*, *CAAP1*, *CAPN12*, *DAD1*, *FAS*, *ITPR1*, *TRAF2*, and *USP28*;
3. Cell cycle progression and survival: *ADRA1A*, *BUB3*, *FIGLA*, *PTPN6*, *RNF123*, *TBRG4*, and *ZEB1*;
4. Cell motility, adhesion, and contact: *ANK3*, *BAIAP2*, *EPHA4*, *FLNB*, *FRAS1*, *ICAM2*, *ITPR2*, *PLS3*, *RAMP3*, and *TACR3*;
5. T-cell receptor recombination and signaling: *CARD11*, *GATA3*, *PKHD1L1*, *RAG2*, and signal activator and transducer of transcription 5B gene *STAT5B*;
6. Proto-oncogene: *CBLB*, *JUNB*, and *TWIST1*;
7. Metabolism: *ALDH1A1*, *CPN2*, *GLUD2*, *MTMR2*, *MTMR7*, and *NAV3*;
8. mRNA regulation and decay: *DCPS* and *PABPC3*;
9. Epigenetic modification: *DNMT3A*, *HDAC2*, *NCOR1*, and *SUZ12*;
10. Membrane stabilization: *EHD1*, *SGMS1*, and *TMEM244*;
11. DNA synthesis and chromatin assembly: *FLAD1* and *TOX*;
12. Tyrosine kinase signaling: *PLCG1*, *RIPK2*, and *TYK2*; and
13. Protein degradation: *PSMD3*.

Data analysis

RNA sequencing and TCR sequencing files for our patient with hypopigmented mycosis fungoides were preprocessed using the BD Targeted Multiplex Rhapsody Analysis Pipeline (version 1.9-beta). Sequencing reads were demultiplexed and mapped onto the reference genome (GRCh38) using the Seven Bridges Genomic Platform (BD Rhapsody Bioinformatics), and a gene expression matrix was output. The 10X Genomics gene expression matrices were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus for two classic mycosis fungoides samples (GSM5280111 and GSM5047045), one normal skin sample (GSM4450726), three atopic dermatitis samples (accession number GSE153760), three psoriasis samples (accession number GSE162183), and six vitiligo samples (accession number GSE203262) (Gao et al., 2021; Gaydosik et al., 2020; Jonak et al., 2021; Rindler et al., 2021; Rojansky et al., 2020; Shiu et al., 2022).

Before analysis, the classic mycosis fungoides and normal skin whole-transcriptome single-cell RNA-sequencing datasets were reduced to the targeted sequencing panel used to sequence our hypopigmented mycosis fungoides dataset. The Seurat package (version 4.0.6) in R (version 2022.07.1, Build 554) was then used to analyze the single-cell RNA-sequencing data. TCR-sequencing data were used to identify cells expressing clonal TCR, and single-cell RNA-sequencing data without TCR sequencing were used for all other analyses. Because raw TCR CDR3 sequencing data were not available in the publicly deposited classic mycosis fungoides datasets, we relied on identical TCRα variant chain and TCRβ variant chain genotypes to determine clonality. The scDblFinder package (version 3.15) was used to identify and remove predicted doublets from each dataset with default settings for 10x Genomics datasets and expected doublet rate = 5% for the hypopigmented mycosis fungoides dataset, as recommended by the BD Rhapsody handbook (McGinnis et al., 2019). To further identify a population of intact cells, cells expressing <40 unique genes in each dataset were removed. Individual datasets were scaled and log normalized, and 2,000 highly variable features were identified. Principal components analysis was performed using the highly variable features, followed by unsupervised clustering using the Louvain algorithm and 20 dimensions. For cluster identification, a resolution of 0.75 was used for GSM5280111, and a resolution of 0.5 was used for all other datasets. Clusters were then visualized in Uniform Manifold Approximation and Projection with 20 dimensions and default settings otherwise. Differentially expressed gene markers for each cluster were identified using the FindAllMarkers function. Clusters were identified on the basis of TCR clonality and manual analysis of differentially expressed genes for cell type markers.

Reference mapping was performed in Seurat using the workflow recommended by the package's authors (Stuart et al., 2019). Briefly, the reference and query datasets were normalized, transfer anchors were found between the datasets, and the data were scaled. Transfer anchors were found using the FindTransferAnchors function, and the reference and query datasets were integrated using the MapQuery function. Through these functions, transfer anchors were used to spatially group cells with similar expression profiles between the two datasets and predict the identity of cells on the basis of their proximity to reference cells with a known identity. Cells with a malignant cell-type prediction score >0.75 were called malignant. Area under the curve plots were produced using the cell-type prediction score generated by the MapQuery function and the pROC package (version 1.18.0).

SUPPLEMENTARY REFERENCES

Gao Y, Yao X, Zhai Y, Li L, Li H, Sun X, et al. Single cell transcriptional zonation of human psoriasis skin identifies an alternative immunoregulatory axis conducted by skin resident cells. Cell Death Dis 2021;12:450.

Gaydosik AM, Queen DS, Trager MH, Akilov OE, Geskin LJ, Fuschiotti P. Genome-wide transcriptome analysis of the STAT6-regulated

genes in advanced-stage cutaneous T-cell lymphoma. Blood 2020;136:1748–59.

Jonak C, Alkon N, Rindler K, Rojahn TB, Shaw LE, Porkert S, et al. Single-cell RNA sequencing profiling in a patient with discordant primary cutaneous B-cell and T-cell lymphoma reveals micromilieu-driven immune skewing. Br J Dermatol 2021;185:1013–25.

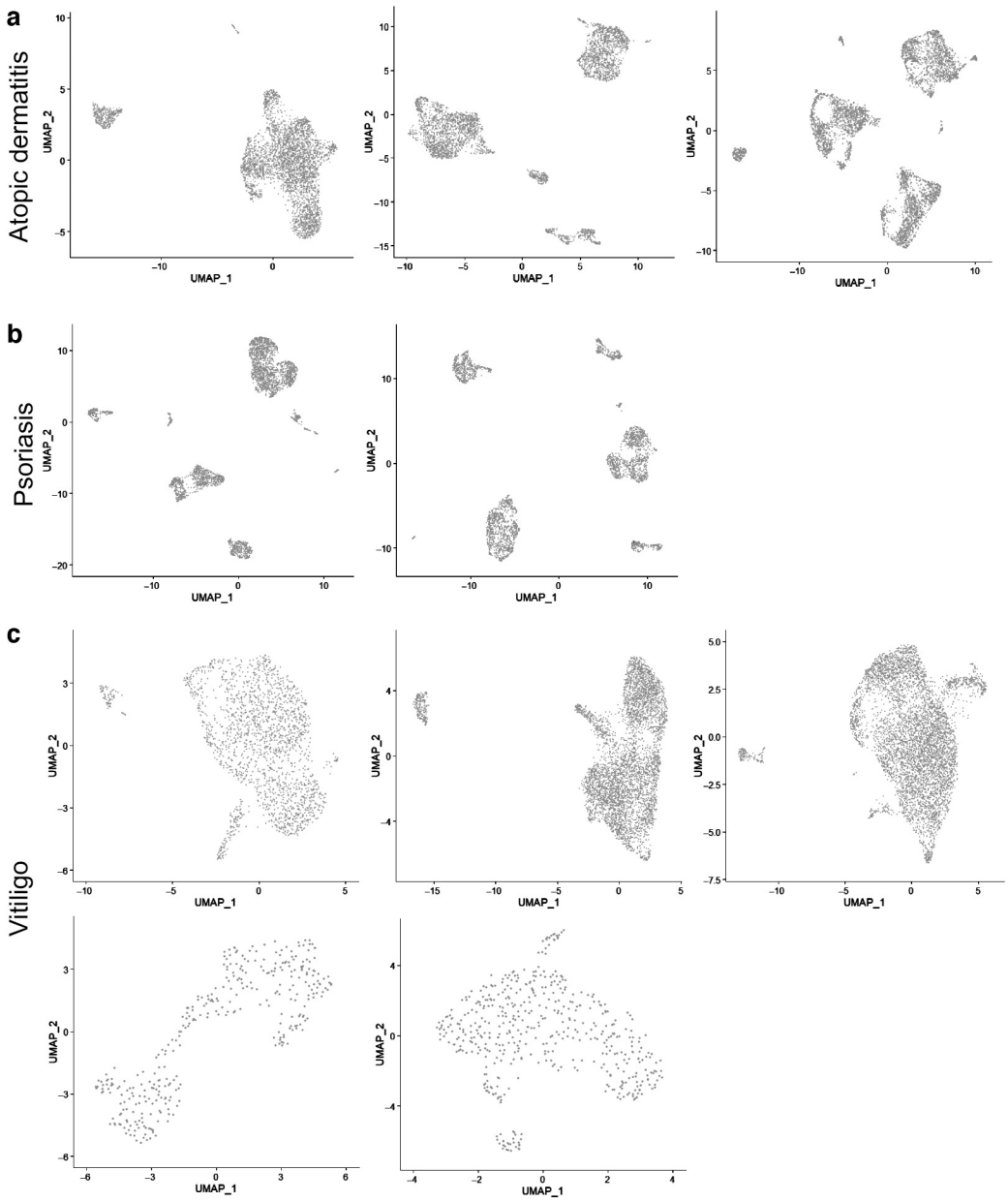
McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst 2019;8:329–337. e4.

Rindler K, Bauer WM, Jonak C, Wielscher M, Shaw LE, Rojahn TB, et al. Single-cell RNA sequencing reveals tissue compartment-specific plasticity of mycosis fungoides tumor cells. Front Immunol 2021;12:666935.

Rojansky R, Fernandez-Pol S, Wang E, Rieger KE, Novoa RA, Zehnder JL, et al. Cutaneous T-cell lymphomas with pathogenic somatic mutations and absence of detectable clonal T-cell receptor gene rearrangement: two case reports. Diagn Pathol 2020;15:122.

Shiu J, Zhang L, Lentsch G, Flesher JL, Jin S, Polleys C, et al. Multimodal analyses of vitiligo skin identify tissue characteristics of stable disease. JCI Insight 2022;7: e154585.

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. Cell 2019;177:1888–902. e21.



Supplementary Figure S1. The reference mapping model applied to nonmalignant inflammatory skin diseases shows good model specificity. The reference mapping model was applied to samples from (**a**) three patients with atopic dermatitis, (**b**) two patients with psoriasis, and (**c**) five patients with vitiligo. UMAP, Uniform Manifold Approximation and Projection.