blood®
Regular Article

LYMPHOID NEOPLASIA

Multimodal single-cell analysis of cutaneous T-cell lymphoma reveals distinct subclonal tissue-dependent signatures

Alberto Herrera,[1,*] Anthony Cheng,[2,3,*] Eleni P. Mimitou,[4] Angelina Seffens,[1,5] Dean George,[6] Michal Bar-Natan,[1,7] Adriana Heguy,[1,8] Kelly V. Ruggles,[9] Jose U. Scher,[10] Kenneth Hymes,[11] Jo-Ann Latkowski,[12] Niels Ødum,[13] Marshall E. Kadin,[6] Zhengqing Ouyang,[3] Larisa J. Geskin,[14] Peter Smibert,[4] Terkild B. Buus,[1,13] and Sergei B. Koralov[1]

[1]Department of Pathology, New York University School of Medicine, New York, NY; [2]Department of Genetic and Genome Sciences, University of Connecticut School of Medicine, Farmington, CT; [3]Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA; [4]Technology Innovation Laboratory, New York Genome Center, New York, NY; [5]Columbia University Vagelos College of Physicians and Surgeons, New York, NY; [6]Department of Dermatology, Boston University and Roger Williams Medical Center, Brown University, Providence, RI; [7]Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY; [8]Genome Technology Center, New York University School of Medicine, New York, NY; [9]Division of Translational Medicine, [10]Division of Rheumatology, Department of Medicine, [11]Division of Hematology/Oncology, and [12]Department of Dermatology, New York University School of Medicine, New York, NY; [13]LEO Foundation Skin Immunology Research Center, Department of Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark; [14]Department of Dermatology, Columbia University, New York, NY

KEY POINTS

- **There is striking subclonal molecular heterogeneity within clonal malignant T-cell populations in the skin and blood of leukemic CTCL.**
- **Tissue microenvironment influences the transcriptional state of malignant T cells, likely contributing to evolution of malignant clones.**

Cutaneous T-cell lymphoma (CTCL) is a heterogeneous group of mature T-cell neoplasms characterized by the accumulation of clonal malignant CD4+ T cells in the skin. The most common variant of CTCL, mycosis fungoides (MF), is confined to the skin in early stages but can be accompanied by extracutaneous dissemination of malignant T cells to the blood and lymph nodes in advanced stages of disease. Sézary syndrome (SS), a leukemic form of disease, is characterized by significant blood involvement. Little is known about the transcriptional and genomic relationship between skin- and blood-residing malignant T cells in CTCL. To identify and interrogate malignant clones in matched skin and blood from patients with leukemic MF and SS, we combine T-cell receptor clonotyping with quantification of gene expression and cell surface markers at the single cell level. Our data reveal clonal evolution at a transcriptional and genetic level within the malignant populations of individual patients. We highlight highly consistent transcriptional signatures delineating skin- and blood-derived malignant T cells. Analysis of these 2 populations suggests that environmental cues, along with genetic aberrations, contribute to transcriptional profiles of malignant T cells. Our findings indicate that the skin microenvironment in CTCL promotes a transcriptional response supporting rapid malignant expansion, as opposed to the quiescent state observed in the blood, potentially influencing efficacy of therapies. These results provide insight into tissue-specific characteristics of cancerous cells and underscore the need to address the patients' individual malignant profiles at the time of therapy to eliminate all subclones.

Introduction

Insights into the molecular progression of cutaneous T-cell lymphoma (CTCL) have revealed a significant degree of heterogeneity in affected tissues between different patients and disease stages and prospective intraindividual variability.[1–10] In addition, single-cell profiling of advanced stage tumor biopsies and circulating Sézary syndrome (SS) cells has further highlighted substantial inter- and intratumoral variability.[11,12] While these approaches have provided important insights into the genomic instability and the transcriptional heterogeneity of CTCL, there is a gap in the understanding of disease progression between skin and blood, and little is known about the degree to which malignant T cells in the blood resemble malignant T cells in lesional skin within leukemic disease.[13–15] In the present study, we interrogate multiple modalities of information in single cells from matched skin and blood samples of patients with leukemic disease and healthy controls within a single workflow. By using expanded CRISPR-compatible cellular indexing of transcriptomes and epitopes by sequencing (ECCITE-seq),[16] we sought to compare the molecular profile of malignant clones residing in the skin and circulation of these patients across gene and protein expression modalities. We apply inferred copy number variation (CNV) and phylogenetic analysis to examine subclonal heterogeneity to gain insights into the evolution and the relationship of malignant clones across tissues.

Materials and methods

Clinical samples

Patient and control samples were collected at New York University Langone Health Medical Center, Columbia University Irving Medical Center, and Roger Williams Medical Center in accordance with protocols approved by the New York University School of Medicine Institutional Review Board and Bellevue Facility Research Review Committee (IRB#i15-01162), Columbia University Irving Medical Center Institutional Review Board (IRB#AAAQ8751), Roger William Medical Center Protocol Review Committee (IRB#15-520-99), and Rhode Island Hospital Institutional Review Board (IRB#4159-16). Patients with CTCL were diagnosed according to ISCL/EORTC classification criteria, and all patients included in the study had detectable leukemic involvement and staging of mycosis fungoides (MF) stage IV or SS. Most patients included were not undergoing any topical or systemic therapies for at least 2 weeks before the time of biospecimen collection. A summary of the clinical notes, including past therapies for each patient, is included in supplemental Table 1, available on the *Blood* Web site. After written informed consent was obtained, skin biopsy and peripheral blood samples were harvested. All authors had access to de-identified patient clinical data.

Harvesting of cells from clinical tissue

Peripheral blood mononuclear cells (PBMCs) were isolated from the blood of patients and healthy controls by gradient centrifugation using Ficoll-Paque PLUS (GE Healthcare) and Sepmate-50 tubes (Stemcell). Buffy coat PBMCs were collected and washed twice with phosphate-buffered saline (PBS) 2% fetal bovine serum (FBS). Skin biopsies were digested in 0.25 mg/mL of Liberase TL and 0.5 mg/mL DNase (Sigma) diluted in RPMI and incubated for 1.5 hours at 37°C. Skin dissociated cells were filtered through a 70-μM cell strainer. All cells from single-cell suspensions were cryopreserved in freezing medium (40% RPMI 1640, 50% FBS and 10% dimethyl sulfoxide). Cryopreserved PBMCs and skin-dissociated cells were thawed for 1 to 2 minutes in a 37°C water bath, washed twice in warm PBS 2% FBS, and resuspended in complete media (RPMI 1640 supplemented with 10% FBS and 2 mM L-Glut) on the day of the experiments. Biopsies from 2 patients did not yield sufficient T cells to be included in analysis.

Oligo-conjugated antibodies

Antibodies used for cell hashing and for our ECCITE-seq antibody-derived tags (ADT) panel were covalently and irreversibly conjugated to barcode oligos by iEDDA-click chemistry ("home conjugated") as previously described[16–18] or purchased as TotalSeq-C reagents (BioLegend). Concentrations and clones used for all conjugated antibodies for staining are listed in supplemental Table 4.

ECCITE-seq protocol compatible with 10x genomics instruments

Cells were stained with barcoded antibodies and loaded onto the 10x Chromium single-cell workflow (10X Genomics) as previously described[16] with a few modifications. Approximately 1 million to 3 million cells per sample were resuspended in CITE-seq staining buffer (2% BSA, 0.01% Tween in PBS) and incubated for 10 minutes with Fc receptor block (TruStain FcX, BioLegend and FcR blocking reagent, Miltenyi) to prevent antibody binding to Fc receptors. Cells were then incubated with hashing and ECCITE-seq surface panel antibodies for 30 minutes at 4°C. After staining, cells were washed 3 times in PBS containing 2% BSA and 0.01% Tween, followed by centrifugation (300 ×g for 5 minutes at 4°C) and supernatant aspiration. After the final wash, cells were resuspended in PBS and filtered through 40-μm cell strainers.

Stained cells from each sample were pooled and loaded into the 10x Chromium Single Cell Immune Profiling workflow according to the manufacturer's instructions with a few modifications as previously described.[16] Additional modifications included 1. Addition of 2 pmol of TotalSeq-C BioLegend protein-tag additive (CTCGTGGGCTCGGAGATGTGTATAAGAGACG) during cDNA amplification PCR for pools stained with TotalSeq-C antibodies and 2. N7NY_xx indexing primers, as described by Mimitou et al,[16] were used along with SI-PCR primer to make libraries for TotalSeq-C antibodies. Libraries were pooled and sequenced on Illumina NovaSeq 6000, MiSeq, or HiSeq 2500.

Single-cell data processing

Alignment, quantification, and demultiplexing FASTQ files from the 10x libraries were processed using the count module of Cell Ranger pipeline, version 3.0.1 (10x Genomics). FASTQ files from the T-cell receptor (TCR) α/β and TCR γ/δ libraries were processed using the vdj module of Cell Ranger pipeline, version 3.0.1 (10x Genomics). For ADT libraries and hashtag oligos (HTO) libraries, the tag quantification pipeline version 1.3.2, available at https://github.com/Hoohm/CITE-seq-Count, was used to tally the counts from reads. A Hamming distance of 1 was used as criteria for read sequence matches to preassigned barcode sequences. All reads from the 5' mRNA, TCR α/β, and TCR γ/δ libraries were aligned to the GRCh38 reference. Raw UMI gene expression (GEX), TCR α/β, TCR γ/δ clonotype, ADT, and HTO count matrices were obtained. Samples were demultiplexed by their HTO using Seurat HTODemux with positive quantile parameter set to 0.999.

Single-cell data analysis Removal of cell doublets Scrublet[19] was used to detect and remove intrasample cell duplicates. We parameterized the model using a 7% expected doublet rate, 30 principal components (default), and a minimum gene variability of the 85th percentile (default). We performed manual thresholding by using a doublet score of 0.2.

Seurat normalization The count matrices of all modalities were loaded into a Seurat v3 object (Seurat version 3.0.2). For GEX data, we normalized the count data using LogNormalize and scaled the resulting normalized values using ScaleData. For all other modalities (ADT, HTO, TCR α/β, TCR γ/δ), we normalized the count data using centered-log ratio (margin = 2) and scaled the resulting normalized values using ScaleData.

Multimodal WNN analysis We applied weighted-nearest neighbor (WNN) analysis on our ECCITE-seq data, enabling the integrative analysis of GEX and ADT modalities in the same cells as previously described.[20] k-Nearest neighbors graphs (k = 20) were first constructed, and the neighbors of each cell in each modality were identified. For graph construction, the first 30 principal components (PCs) of the log-normalized values from the GEX modality and the first 10 PCs of the centered-log ratio values from the ADT modality were used. The dimensionally reduced molecular profiles of the k-nearest neighbors of each cell were

then averaged and compared with the actual values to obtain the residuals. This procedure was done in a within-modality (RNA-RNA, protein-protein) and cross-modality (RNA-protein, protein-RNA) manner. Next, for each of these residuals, they were converted into affinity-based similarities using Uniform Manifold Approximation and Projection (UMAPs) locally adaptive kernel. Finally, the ratios of the within-modality and cross-modality affinities were softmax transformed to obtain the cell-specific modality weights. The weighted similarity then formed the basis of the construction of the WNN graph. Using the integrated WNN graph, we constructed *t*-SNE visualizations and performed modularity clustering from it.

TCR analysis TCR α/β reads attributed to each cell barcode were grouped, assembled into a single contig, and then had their V, D, and J segments annotated. Only clonotypes with productive and full-length contigs were retained for analysis. TCR α/β and TCR γ/δ clonotype count matrices were obtained and appended alongside GEX and ADT modalities for simultaneous analysis.

CNV inference InferCNV version 1.2.1 was applied for analysis of individual patient samples using malignant cells (defined by their expanded TCR β clonotype, distinctive clustering, and GEX patterns) (supplemental Figures 3 and 4) from both the blood and skin as the targets, and nonmalignant CD4+ T cells were used as the reference. For subcluster analysis, the analysis_mode parameter was set to "subclusters." Default parameters (cutoff = 0.1, cluster_by_groups = TRUE, denoise = TRUE, HMM = TRUE) were used for these analyses.

Phylogenetic analysis Using the de-noised CNV inference hidden Markov model results, we built copy number profiles for each patient at 500-Kb resolution. Neighbor-joining trees of malignant T cells were then constructed and rooted with an individual nonmalignant T cell using R package ape version 5.3. Bootstrap analysis of the NJ tree was performed using 100 replicates to estimate the confidence of the clades. Visualization of the tree as a phylogram was done using R package ggtree version 2.0.1.

Pseudotime trajectory analysis Pseudotime trajectory analysis was performed on each patient with CTCL with matched blood and skin samples (SS1 to SS4 and MF stage IV) using Monocle 2 version 2.14.[21] Briefly, moderately expressed genes identified using detectGenes (min_expr = 0.1) and at least 20 cells expressing the genes were used for the analysis. PCA was performed, and the top 20 components were used for *t*-SNE visualization. Ordering genes were defined based on DifferentialGeneTest on the tissue of origin and used for trajectory construction (DDRTree)[21,22] and pseudotime analysis. We rooted the state based on the CNV subtype that has the shortest branch length as compared with nonmalignant CD4+ T cells.

Gene-set analysis Two gene sets—resting CD4+ T cells and activated CD4+ T cells upon CD3/CD28 stimulation in human tissues—were retrieved from Szabo et al.[23] We used Seurat function ScoreModules to score each malignant cell from both blood and skin on each gene set.

Code availability Code for the pipeline to pre-process ECCITE-seq data are available on GitHub: http://github.com/ouyang-lab/CTCL. All downstream analysis will be made available on request.

Results

Clonal subpopulations in skin and blood

We used ECCITE-seq to examine clonal malignant T cells and survey other leukocytes in the blood and skin of 5 patients with leukemic MF and SS and in the blood of 2 additional patients with SS (supplemental Table 1). While sequencing-based capture of surface epitopes recapitulates distribution of markers used for fluorescence -activated cell sorting analysis of malignant T cells in CTCL (supplemental Figure 1), ECCITE-seq enables a far more detailed analysis of the samples with information on >50 surface epitopes, single cell transcriptome and V(D)J repertoire.[16] As expected, we detected clonally expanded malignant T cells only in CTCL patient-derived specimens (Figure 1A-B; supplemental Figures 2-4). Differences in malignant burden and in the distribution of other leukocyte populations were observed between individual patients (Figure 1B; supplemental Figure 2; supplemental Tables 2 and 3). The malignant T cells within each patient were defined based on their expanded TCR β CDR3 and on distinct GEX and ADT profiles clustering them away from polyclonal nonmalignant T cells in both tissues (Figure 1A; supplemental Figure 2). Malignant GEX profiles included dysregulated expression of hallmark genes frequently affected in leukemic MF and SS, including *SATB1* , *DPP4* (CD26), *CD7*, *TOX*, *PSL3*, *HDAC9*, *DNM3*, *CDO1*, *KIR3DL2*, and *GATA3* (supplemental Figures 3 and 4). Additionally, the malignant T cells in all patients with CTCL with matched lesional skin and blood shared identical TCR clonotypes between tissues, indicating distinct clonal involvement across tissues in late-stage CTCL.

We then reclustered our CTCL samples, only including the malignant T cells, and focused on their transcriptional profiles. Our analysis detected transcriptional heterogeneity at the single-cell level among malignant T cells from the same patient, allowing for grouping into distinct clusters (Figure 1C-D; supplemental Figure 5). Interestingly, when malignant T cells from skin and blood were analyzed together, most of the subclonal transcriptional heterogeneity was observed in the blood, whereas the skin-derived malignant T cells mainly constituted a single subcluster (Figure 1C-D; supplemental Figure 5). This was consistent across all patients investigated; together, these results exposed a high degree of transcriptional heterogeneity in blood-derived malignant T cells of multiple patients with CTCL and a unique transcriptional bias generally grouping skin-derived malignant T cells in their own subcluster.

CNVs identify malignant subclones

SS is known to encompass a significant degree of genetic instability across patients, portrayed by an abundance of CNVs.[1–3,24,25] To determine if the intraclonal malignant transcriptional heterogeneity in our CTCL cohort reflects the presence of distinct malignant subclones, we inferred CNVs at the single-cell level in clonal malignant T cells from single -cell RNA-sequencing (scRNA-seq) data as has been previously demonstrated in other tumor types[26] (Figure 2A-C; supplemental Figure 5). This technique (InferCNV) was previously used in the analysis of uveal melanoma to ascertain clonal structure in tumors and reveal noncanonical large-scale CNVs with potential implications in tumor progression.[27] CNV inference methods use GEX data to find broad, contiguous sections of consistent copy number changes. We took advantage of this method to identify broad CNV events to distinguish major subclonal malignant populations.

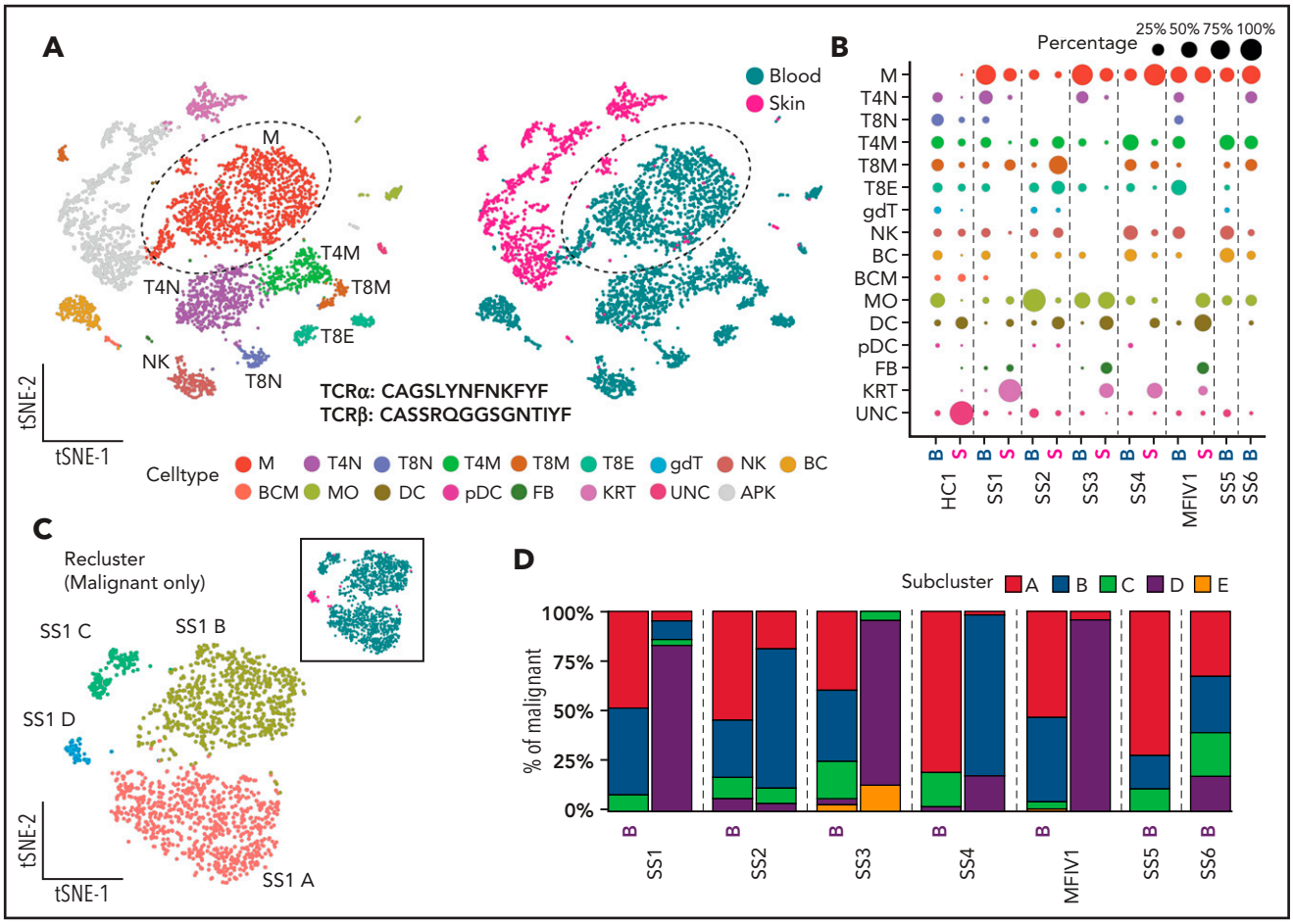

Figure 1. ECCITE-seq captures transcriptional, surface protein and clonotypic information from multiplexed healthy and CTCL tissues as well as malignant heterogeneity in disease. (A) Representative *t*-SNE plots of single cells from PBMCs and skin dissociated tissue of a CTCL patient (SS1) clustered and visualized by GEX and ADT expression using WNN analysis and colored by cluster (left) or tissue of origin (right). (B) Dot plot showing the percentage of populations making up the skin and blood of a healthy control (HC1), 5 CTCL patients (SS1-MF stage IV1), and blood from 2 CTCL patients (SS6 and SS7). Apoptotic keratinocytes from skin tissue were filtered out. (C) Representative t-SNE plot depicting transcriptional heterogeneity within the malignant T cells from matched blood and skin defined by their common expanded clonotype in a CTCL patient. (D) Bar plots showing the subclonal composition based on transcriptional differences in malignant T cells from 7 CTCL patients (5 of whom have matched blood and skin). APK, apoptotic keratinocytes; BC, B cells; BCM, memory B cells; DC, dendritic cells; FB, fibroblasts; gdT, γ/δ T cells; KRT, keratinocytes; M, malignant T cells; MO, monocytes/macrophages; NK, natural killer cells; pDC, plasmacytoid dendritic cells; T4M, memory CD4+ T cells; T4N, naïve CD4+ T cells; T8E, effector CD8+ T cells; T8M, memory CD8+ T cells; T8N, naïve CD8+ T cells; UNC, unclassified.

We corroborated our results using an alternative method of CNV profiling from scRNA-seq (copy number karyotyping of aneuploid tumors)[28] (supplemental Figure 6). One limitation of these algorithms involves possible false discovery of CNV patterns if significant GEX differences occur between the chosen reference and target cells in the absence of genetic aberrations. We reinforced our confidence in inferring CNV events from scRNA-seq data by applying the method on scRNA-seq from a patient with SS not otherwise included in the study and comparing with matched whole-exome sequencing and methylation microarray datasets. This comparison revealed high concordance in the CNVs detected by scRNA-seq, whole-exome sequencing, and methylation microarray, albeit with low sensitivity for small-scale aberrations, as would be expected (supplemental Figure 7). To independently validate our inferred CNVs, we also measured loss of heterozygosity stemming from CNV events in our individual patients with CTCL (supplemental Figure 8). Our analysis of skin and blood revealed multiple subclonal populations within each patient's malignant T cells, defined by varying degrees of acquired CNVs in both tissues (Figure 2A-C). Each patient also contained a unique combination of CNVs within their malignant subclones, reinforcing the wide range of genetic aberrations found in late-stage CTCL tissues. Similar CNV profiles are often shared between individual malignant subclones across the tissues, suggesting that genetic heterogeneity alone does not dictate tissue localization of malignant cells. Notably, the CNV defined subclones in 5 out of 7 patients (SS1-SS4 and SS6) showed a moderate degree of concordance (adjusted Rand index: 0.28-0.50) with the previously transcriptionally defined clonal subclusters (Figure 2D; supplemental Figure 5) supporting InferCNV's predictions. Although the other 2 patients (MF stage IV1 and SS6) showed a lower degree of alignment (adjusted Rand index: 0.08-0.10), most cells from each defined state still visually aligned with each other.

Because CTCL is thought to arise from skin effector/memory T cells and progress to leukemic involvement in a subgroup of patients,[13,29,30] we sought to examine whether the evolutionary relationship between CNV-defined subclones in each patient can identify the tissue of origin of malignant cells and

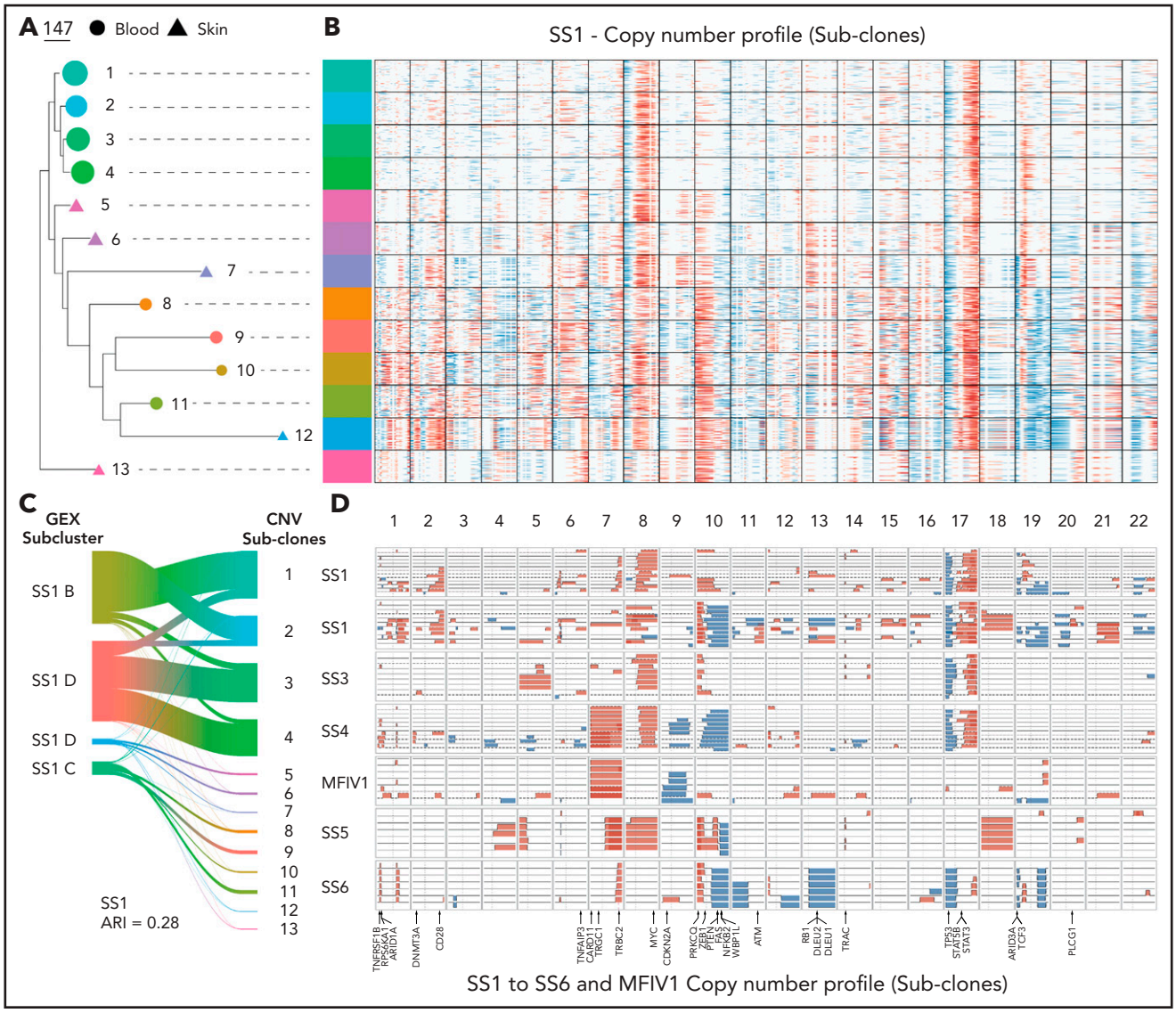

Figure 2. Copy number variations (CNVs) identify distinct malignant subclones phylogenetically intermixed between CTCL tissues. (A) Subclonal populations are colored based on inferred CNV analysis . Predicted regions of copy number gain (in red) or loss (in blue) highlighted across the genomic position (columns) and single cells (rows) divided into their subclonal branches from phylogenetic analysis.(B) Plot showing gain (in red) and loss (in blue) events in CNV defined subclones from the skin- and blood-derived clonal malignant T cells from 5 patients (SS1-MF stage IV1) and blood from 2 patients (SS6 and SS7). Skin derived CNV states are shown as dashed lines and blood derived CNV states as solid lines. CNV states with less than 1**%** frequency in their tissue or based on less than 5 cells were filtered out. (C) Representative phylogram of the malignant skin- and blood-derived T cells of a CTCL patient (SS1) bootstrapped 100 times and rooted using their non-malignant polyclonal T cells to determine malignant evolution of CNV defined clusters. Branch length depicts level of CNV burden. (D) Representative river plot showing connections between GEX based malignant subclusters and CNV defined malignant subclones in a CTCL patient (SS1).

provide a map of clonal evolution in leukemic disease. To this end, we applied a phylogenetic approach in which we constructed neighbor-joining trees as well as minimum evolution and average linkage distance–based phylograms for each patient's malignant T cells, rooted by the polyclonal nonmalignant T cells from the same patient (Figure 2C; supplemental Figures 5, 9, and 10). We observed that in most of the matched samples, there was no clear monodirectional phylogenetic relationship between skin- and blood-derived subclones to suggest a tissue of origin for disease, with distinct trajectories of clonal evolution among individual patients. Our analysis showed branches of clonal evolution leading to advanced skin and blood subclones in each patient. Advanced subclones also shared several common intermediate-late branching points regardless of tissue distribution, indicating that closely related subclones could reside in either affected tissue. Interestingly, only the patient diagnosed with MF stage IV1, who did not meet clinical criteria for SS diagnosis, showed a more clear distinction in subclonal trajectories between tissues. This complex and variable phylogenetic landscape based on CNV accumulations reveals a snapshot of the development of multiple related but unique subclones in the skin and blood of patients with SS, with both tissues frequently sharing similar branches of evolution. Collectively, these data highlight the highly heterogeneous nature of late-stage CTCL at the subclonal level and highlight a clinically relevant challenge presented by the emergence of distinct subclonal populations across multiple tissues.

Tissue-dependent transcriptional signatures

In contrast to their blood-derived counterparts, skin-derived malignant T cells from our patients with CTCL showed a strong tendency to group either as a single main subcluster (SS1, SS3-MF stage IV1) or as a separate island of cells away from circulating cells (SS2) when analyzed by their transcriptional profile (Figure 1C; supplemental Figure 5). To assess the transcriptional relationship between malignant skin- and blood-derived T cells from these patients, we applied a transcriptional trajectory analysis of matched samples as previously described[21] (Figure 3A,C; supplemental Figure 11). This analysis revealed transcriptional trajectories that generally started with subclones having the fewest CNVs and ended with subclones containing more extensive CNVs. We also observed a strong transcriptional bias, with skewing of all skin-derived subclones, toward the end of the branches that started at blood-derived subclone trajectories. This is in

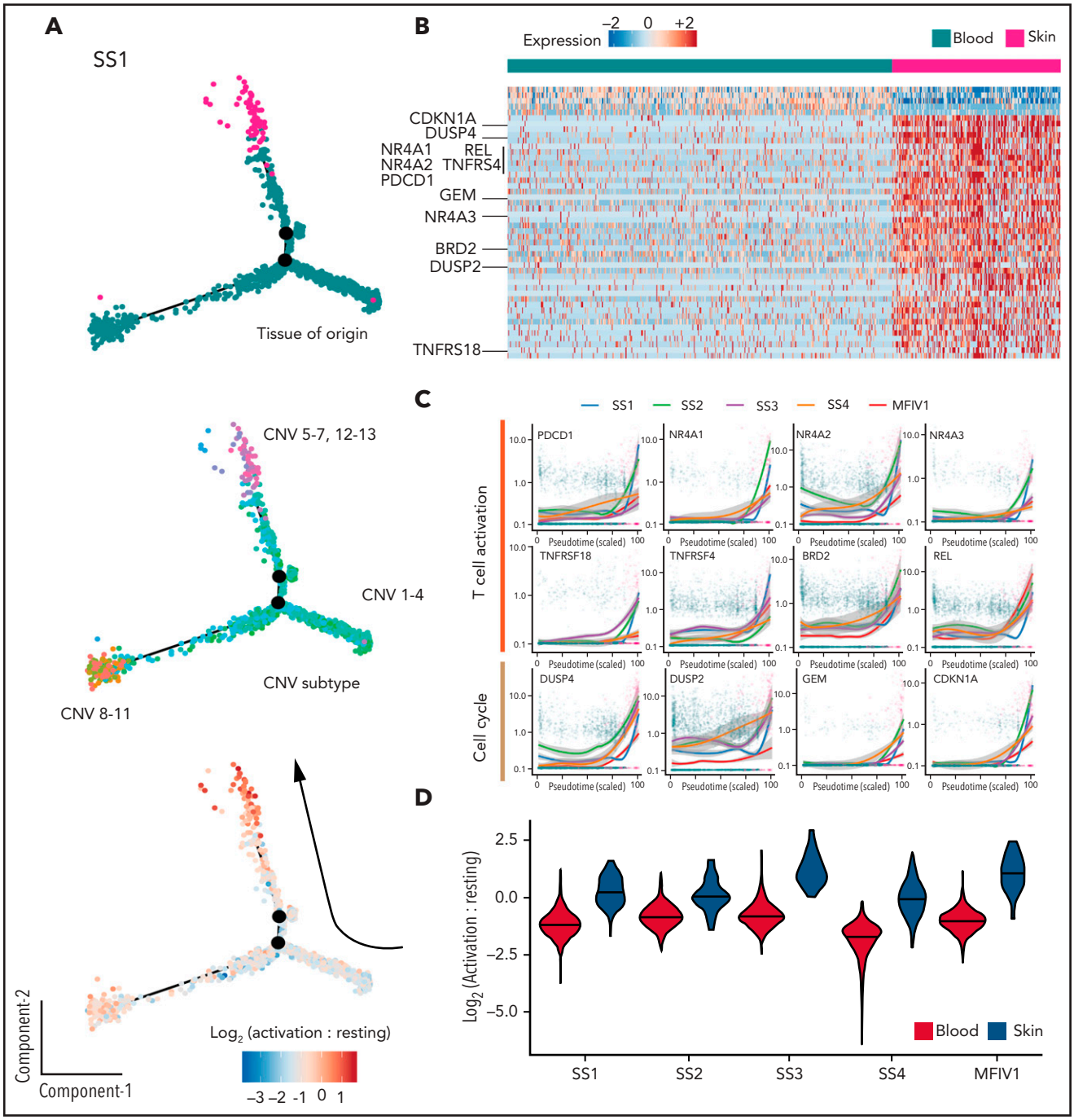

Figure 3. Transcriptional comparison of skin- and blood-derived malignant T cells in CTCL. (A) Representative transcriptional trajectory analysis of malignant skin- and blood-derived T cells from a CTCL patient (SS1). Colored by tissue of origin (top), inferred CNV subclones (middle) and activation to resting score ratio (bottom). (B) Pairwise comparison of differentially expressed genes between skin- and blood-derived malignant T cells of 5 CTCL patients (SS1-MF stage IV1). (C) Plots showing expression of selected "T-cell activation" and "cell-cycle" genes in malignant T cells found in circulation and skin. (D) Violin plots showing T-cell activation to resting score ratios, across malignant T cells from skin and blood of 5 CTCL patients (SS1-MF stage IV1).

contrast with our previous observations using phylogenetic analysis of the inferred CNV subclones, in which we found an intermixed distribution among skin and blood subclonal branches in regards to their path of evolution and their common ancestors (Figure 2B; supplemental Figure 5). These results suggested the presence of a strong tissue-dependent transcriptional bias in malignant T cells from the skin, effectively masking subclonal differences.

To determine the underlying mechanism behind this transcriptional bias observed in skin-derived malignant T cells, we performed paired transcriptional comparisons between matched skin- and blood-derived malignant T cells from 5 of our patients with CTCL (Figure 3B-C; supplemental Figure 12). This comparison revealed a consistent and strong upregulation of several T cell activation-, TCR ligation-, and mitogen-induced transcripts as well as regulators of cell cycle in malignant T cells from skin as compared with their blood-derived counterparts. These genes included *PDCD1* (PD-1), *NR4A1*, *NR4A2*, *NR4A3*, *TNFRSF18* (GITR), *TNFRSF4* (OX-40R), *BRD2*, *REL*, *DUSP2*, *DUSP4*, *GEM,* and *CDKN1A* among others. Additionally, upregulated transcripts in blood-derived malignant T cells compared with their skin-derived counterparts included *KLF2*, an important factor in maintaining T-cell quiescence,[31] and *TCF7* and *SELL* (CD62L), known markers of resting T cells.[32,33] Because PD-1 and CD62L were included in our ADT surface protein panel (supplemental Table 4), we examined whether the tissue-dependent expression of these proteins was also evident in individual malignant T cells (supplemental Figure 13). Our results revealed significantly higher levels of surface PD-1 expression in skin-derived malignant T cells compared with blood counterparts in 4 of 5 patients with CTCL (SS1-4) and significantly higher levels of CD62L expression in blood-derived malignant T cells compared with their skin counterparts in 4 of 5 patients with leukemic CTCL (SS2-5) in line with tissue-specific distribution of activation signature.

The abundance of differentially expressed transcripts relating to differences in activation status between malignant T cells depending on their tissue location prompted us to investigate whether independent "T-cell activation status" gene sets were enriched in malignant T cells from skin and blood (Figure 3A,D; supplemental Figure 11).[23] Notably, skin-derived malignant T cells had a consistently higher T-cell activation score compared with blood-derived malignant T cells, whereas the opposite was true for the resting T-cell score. It resulted in a markedly increased activation-to-resting-score ratio in malignant T cells from skin as compared with their blood-derived counterparts in each patient (Figure 3D). Importantly, this activated signature is unique to the lesional skin microenvironment of CTCL, because there was no difference in the activation to resting score ratio between skin-derived CD4+ T cells from the nonlesional skin of a healthy control and a patient with psoriasis compared with their blood counterparts (supplemental Figure 14).

To explore whether this chronically activated phenotype of malignant T cells in the skin tumor microenvironment coincides with a high proliferative potential, we applied a cell cycle scoring analysis using canonical markers of G1, S, and G2/M as previously described[34] and superimposed them on the transcriptional trajectories of malignant T cells (supplemental Figure 15). Our results revealed that the frequency of highly proliferative malignant T cells in the skin of patients with CTCL was higher than that found in circulation. These results are consistent with observed activation of malignant T cells in the skin and indicate that the skin tumor microenvironment provides signals conducive to frequent malignant T-cell activation and expansion. Of note, regressing out tissue-dependent transcriptional signatures allowed malignant subclones, based on inferred CNV profiles from skin, to cluster together with their most phylogenetically closely related subclones in blood (supplemental Figure 16). Overall, these results show consistent transcriptional signatures delineating malignant subclones from skin and blood, highlighting the importance of the tumor microenvironment in shaping the transcriptional programs that contribute to distinct functional responses.

Discussion

SS is known to be a heterogeneous disease, with different patients demonstrating highly variable transcriptional profiles,[6,11,35–37] but not much is known about the relationship between malignant T cells from the skin and blood in these patients. Our findings reveal the presence of a distinct T-cell clone identified in both the matched skin and blood of patients with MF stage IV and patients with SS with multiple CNV-defined subclonal populations in both tissues. In accordance with previous studies,[1,3,38] many of these subclones displayed recurrent CNVs, particularly involving gains in 8q (containing *MYC* , 5 of 7 patients) and 17q (5 of 7 patients) and losses in 10q (containing *PTEN*, 4 of 7 patients) and 17p (containing TP53, 5 of 7 patients) that could be found in both tissues. This broad assortment of CNVs was accompanied by a high degree of transcriptional heterogeneity between clonal malignant T cells across patients (supplemental Figure 17). The presence of a highly branching subclonal evolution regardless of tissue was also reflected in our phylogenetic analysis, in which we saw no clear linear relationship between skin- and blood-derived subclones. GEX analysis using T-helper (Th) subtype differentiation markers between tissues also revealed preferential expression of Th2 and Th17 signature transcripts across malignant T cells, consistent with previous reports on Th identity of malignant cells in CTCL[39–41] (supplemental Figure 18). These results highlight a close genetic relationship between subclonal populations in circulation and skin, frequently sharing branches of clonal evolution, and suggest that migration between tissues is continuous, not a one-time event. While one of our patients (MF stage IV1) showed clear phylogenetically related subclones based on tissue distribution, that individual was the only patient not considered to have SS based on lack of clinical criteria, suggesting that recirculation of evolved malignant clones may be a characteristic of SS. This highly diverse landscape at the subclonal level provides insight regarding the refractory nature of advanced-stage CTCL,[42] in which malignant subpopulations can be resistant to treatment,[6] and highlights the need for a personalized, multipronged therapeutic approach wherein the complete landscape of subclonal diversity in each patient should be considered. Because most patients with SS are treated with sequential single-agent therapies,[30] our data provide additional rationale behind a polytherapeutic approach.

Our differential GEX analysis revealed robust differences between malignant T cells from blood and skin and highlighted marked enrichment of transcripts related to activation status and cell cycle progression. This indicates that the tumor microenvironment in CTCL can dictate the transcriptional program of malignant T cells. Several lines of evidence have pointed to the critical role of the

skin microenvironment in the pathogenesis of CTCL, a site of interaction between skin pathobionts and immune cells and as a site of chronic activation and proinflammatory signaling that promotes the selection of malignant CD4+ T cells.[11,43–50] A report by Cristofoletti et al[51] also proposes that skin-derived T cells in SS have a higher proliferative index compared with their blood-derived counterparts. Together with the highly consistent differences we observe in expression of activation- and mitogen-induced transcripts in malignant T cells found in the skin and blood of patients, these findings indicate that the skin microenvironment of CTCL promotes a transcriptional landscape driving sustained activation that can lead to rapid malignant expansion.

Because treatment failure is a common occurrence in patients with SS,[42] it is imperative to improve our understanding of the mechanisms that promote disease heterogeneity, specifically the emergence of multiple malignant subclones. Our results suggest that the skin microenvironment can greatly influence the transcriptional state of malignant T cells and may promote evolution of malignant subclones that can readily disseminate through the circulation. Our study provides a foundation for further understanding the molecular drivers that orchestrate tissue distribution in the skin and blood of patients with advanced-stage CTCL at the single-cell level and makes the case for extending this multimodal approach to interrogate the key drivers of both early skin-confined and progressive disease.

Acknowledgments

The authors thank New York University Genome Technology Center for technical assistance and support.

S.B.K. was supported by the National Institutes of Health, National Heart, Lung, and Blood Institute R01 grant (HL-125816), LEO Foundation grant (LF-OC-20-000351), New York University Cancer Center pilot grant (P30CA016087), the Judith and Stewart Colton Center for Autoimmunity pilot grant, and a grant from the Drs. Martin and Dorothy Spatz Foundation. A.S. was supported by HHMI Medical Scholar Fellowship. P.S. was supported by the National Institutes of Health, National Human Genome Research Institute (R21HG009748) and the Chan Zuckerberg Initiative (HCA-A-1704-01895). T.B.B. and N.Ø. were supported by the Danish Cancer Society (Kræftens Bekæmpelse), the Danish Council for Independent Research (Danmarks Frie Forskningsfond), and the LEO Foundation. M.E.K. was supported by a grant from the Drs. Martin and Dorothy Spatz Foundation.

Authorship

Contribution: A. Herrera and S.B.K. initiated the project and were responsible for experiment design with input from E.P.M., N.Ø., A. Heguy, and P.S., E.P.M. and A. Herrera performed the majority of experiments; A.S., J.U.S., L.J.G., M.E.K., D.G., M.B., K.H., and J.L. were directly involved in patient recruitment, sample collection and study design; A.C. performed the bioinformatics work under the supervision of Z.O. with input from A. Herrera, K.V.R., T.B.B., and S.B.K; and A. Herrera, A.C., T.B.B., and S.B.K. wrote the paper with all authors providing input on the manuscript.

Conflict-of-interest disclosure: P.S. is listed a co-inventor on a patent application related to ECCITE-seq (US provisional patent application 62/515-180). The remaining authors declare no competing financial interests.

ORCID profiles: A.H., 0000-0003-4189-9051; A.C., 0000-0002-0778-8238; E.P.M., 0000-0001-9737-6394; A.S., 0000-0001-5703-5472; K.V.R., 0000-0002-0152-0863; J.U.S., 0000-0002-1072-6994; N.Ø., 0000-0003-3135-5624; Z.O., 0000-0003-2842-8503; P.S., 0000-0003-0772-1647; T.B.B., 0000-0001-7180-6384; S.B.K., 0000-0002-4843-3791.

Correspondence: Sergei B. Koralov, Department of Pathology, New York University School of Medicine, 550 First Ave, MSB 531A, New York, NY 10016; e-mail: sergei.koralov@nyumc.org; and Terkild B. Buus, Department of Immunology and Microbiology, University of Copenhagen, Blegdamsvej 3B, Bldg. 07-12-62, DK-2200 Copenhagen, Denmark; e-mail: terkild.buus@sund.ku.dk.

Footnotes

Submitted 30 September 2020; accepted 3 April 2021; prepublished online on *Blood* First Edition 7 July 2021. DOI 10.1182/blood.2020009346.

*A.H. and A.C. contributed equally to this work.

The ECCITE-seq data reported in this article have been deposited in the Gene Expression Omnibus database (accession number GSE171811). All other data are available upon reasonable request. Please contact sergei.koralov@nyumc.org.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

REFERENCES

1. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat Genet.* 2015;47(12):1426-1434.
2. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet.* 2015;47(9):1011-1019.
3. Vermeer MH, van Doorn R, Dijkman R, et al. Novel and highly recurrent chromosomal alterations in Sézary syndrome. *Cancer Res.* 2008;68(8):2689-2698.
4. Iżykowska K, Przybylski GK, Gand C, et al. Genetic rearrangements result in altered gene expression and novel fusion transcripts in Sézary syndrome. *Oncotarget.* 2017;8(24):39627-39639.
5. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in cutaneous T-cell lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *OncoImmunology.* 2017;6(5):e1306618.
6. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115-2126.
7. Vidulich KA, Talpur R, Bassett RL, Duvic M. Overall survival in erythrodermic cutaneous T-cell lymphoma: an analysis of prognostic factors in a cohort of patients with erythrodermic cutaneous T-cell lymphoma. *Int J Dermatol.* 2009;48(3):243-252.
8. Iyer A, Hennessey D, O'Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175-1184.
9. Iyer A, Hennessey D, O'Keefe S, et al. Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma. *Blood Adv.* 2020;4(11):2489-2500.
10. Laharanne E, Chevret E, Idrissi Y, et al. CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. *Mod Pathol.* 2010;23(4):547-558.
11. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-cell lymphocyte heterogeneity in advanced cutaneous T-cell lymphoma skin tumors. *Clin Cancer Res.* 2019;25(14):4443-4454.
12. Borcherding N, Voigt AP, Liu V, Link BK, Zhang W, Jabbari A. single-cell profiling of cutaneous T-cell lymphoma reveals underlying heterogeneity associated with disease progression. *Clin Cancer Res.* 2019;25(10):2996-3005.
13. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic

rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767-771.

14. Clark RA, Watanabe R, Teague JE, et al. Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients. Science Translational Medicine. 2012;4(117):117ra117.

15. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. *Br J Dermatol*. 2002; 147(3):464-475.

16. Mimitou EP, Cheng A, Montalbano A, et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods*. 2019;16(5):409-412.

17. van Buggenum JA, Gerlach JP, Eising S, et al. A covalent and cleavable antibody-DNA conjugation strategy for sensitive protein detection via immuno-PCR. *Sci Rep*. 2016; 6(1):22675.

18. Stoeckius M, Zheng S, Houck-Loomis B, et al. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*. 2018; 19(1):224.

19. Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst*. 2019;8(4):281-291.e9.

20. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587.e29.

21. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979-982.

22. Mao Q, Wang L, Tsang IW, Sun Y. Principal graph and structure learning based on reversed graph embedding. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(11): 2227-2241.

23. Szabo PA, Levitin HM, Miron M, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun*. 2019;10(1):4706.

24. Iżykowska K, Przybylski GK. Genetic alterations in Sezary syndrome. *Leuk Lymphoma*. 2011;52(5):745-753.

25. van Doorn R, van Kester MS, Dijkman R, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. *Blood*. 2009;113(1):127-136.

26. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014;344(6190):1396-1401.

27. Durante MA, Rodriguez DA, Kurtenbach S, et al. Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat Commun*. 2020;11(1):496.

28. Gao R, Bai S, Henderson YC, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat Biotechnol*. 2021;39(5):599-608.

29. Kim EJ, Hess S, Richardson SK, et al. Immunopathogenesis and therapy of cutaneous T cell lymphoma. *J Clin Invest*. 2005;115(4):798-812.

30. Mazzeo E, Rubino L, Buglione M, et al. The current management of mycosis fungoides and Sézary syndrome and the role of radiotherapy: principles and indications. *Rep Pract Oncol Radiother*. 2013;19(2):77-91.

31. Preston GC, Feijoo-Carnero C, Schurch N, Cowling VH, Cantrell DA. The impact of KLF2 modulation on the transcriptional program and function of CD8 T cells. *PLoS One*. 2013; 8(10):e77537.

32. Willinger T, Freeman T, Herbert M, Hasegawa H, McMichael AJ, Callan MF. Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. *J Immunol*. 2006;176(3):1439-1446.

33. Yang S, Liu F, Wang QJ, Rosenberg SA, Morgan RA. The shedding of CD62L (L-selectin) regulates the acquisition of lytic activity in human tumor reactive T lymphocytes. *PLoS One*. 2011;6(7):e22560.

34. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352(6282):189-196.

35. Dulmage BO, Geskin LJ. Lessons learned from gene expression profiling of cutaneous T-cell lymphoma. *Br J Dermatol*. 2013;169(6): 1188-1197.

36. Wu CH, Yang CY, Wang L, et al. Cutaneous T-cell lymphoma PDX drug screening platform identifies cooperation between inhibitions of PI3Kα/δ and HDAC. *J Invest Dermatol*. 2021;141(2):364-373.

37. Poglio S, Prochazkova-Carlotti M, Cherrier F, et al. Xenograft and cell culture models of Sezary syndrome reveal cell of origin diversity and subclonal heterogeneity. *Leukemia*. 2021;35(6):1696–1709.

38. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat Genet*. 2015;47(12):1465-1470.

39. Guenova E, Watanabe R, Teague JE, et al. TH2 cytokines from malignant cells suppress TH1 responses and enforce a global TH2 bias in leukemic cutaneous T-cell lymphoma. *Clin Cancer Res*. 2013;19(14):3755-3763.

40. Gibson HM, Mishra A, Chan DV, Hake TS, Porcu P, Wong HK. Impaired proteasome function activates GATA3 in T cells and upregulates CTLA-4: relevance for Sézary syndrome. *J Invest Dermatol*. 2013;133(1): 249-257.

41. Fanok MH, Sun A, Fogli LK, et al. Role of dysregulated cytokine signaling and bacterial triggers in the pathogenesis of cutaneous T-cell lymphoma. *J Invest Dermatol*. 2018; 138(5):1116-1125.

42. Alberti-Violetti S, Talpur R, Schlichte M, Sui D, Duvic M. Advanced-stage mycosis fungoides and Sézary syndrome: survival and response to treatment. *Clin Lymphoma Myeloma Leuk*. 2015;15(6):e105-e112.

43. Krejsgaard T, Lindahl LM, Mongan NP, et al. Malignant inflammation in cutaneous T-cell lymphoma: a hostile takeover. *Semin Immunopathol*. 2017;39(3):269-282.

44. Blümel E, Willerslev-Olsen A, Gluud M, et al. Staphylococcal alpha-toxin tilts the balance between malignant and non-malignant $CD4^+$ T cells in cutaneous T-cell lymphoma. *OncoImmunology*. 2019;8(11):e1641387-e1641387.

45. Posner LE, Fossieck BE Jr, Eddy JL, Bunn PA Jr. Septicemic complications of the cutaneous T-cell lymphomas. *Am J Med*. 1981;71(2): 210-216.

46. Axelrod PI, Lorber B, Vonderheid EC. Infections complicating mycosis fungoides and Sézary syndrome. *JAMA*. 1992;267(10): 1354-1358.

47. Emge DA, Bassett RL, Duvic M, Huen AO. Methicillin-resistant *Staphylococcus aureus* (MRSA) is an important pathogen in erythrodermic cutaneous T-cell lymphoma (CTCL) patients. *Arch Dermatol Res*. 2020; 312(4):283-288.

48. Willerslev-Olsen A, Buus TB, Nastasi C, et al. *Staphylococcus aureus* enterotoxins induce FOXP3 in neoplastic T cells in Sézary syndrome. *Blood Cancer J*. 2020;10(5):57.

49. Lindahl LM, Willerslev-Olsen A, Gjerdrum LMR, et al. Antibiotics inhibit tumor and disease activity in cutaneous T-cell lymphoma. *Blood*. 2019;134(13):1072-1083.

50. Shalabi D, Bistline A, Alpdogan O, et al. Immune evasion and current immunotherapy strategies in mycosis fungoides (MF) and Sézary syndrome (SS). *Linchuang Zhongliuxue Zazhi*. 2019;8(1):11.

51. Cristofoletti C, Bresin A, Picozza M, et al. Blood and skin-derived Sezary cells: differences in proliferation-index, activation of PI3K/AKT/mTORC1 pathway and its prognostic relevance. *Leukemia*. 2019;33(5): 1231-1242.