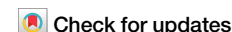


<https://doi.org/10.1038/s41540-024-00474-x>

Spatial cell graph analysis reveals skin tissue organization characteristic for cutaneous T cell lymphoma



Suryadipto Sarkar¹, Anna Möller¹, Anne Hartebrodt¹, Michael Erdmann², Christian Ostalecki², Andreas Baur² & David B. Blumenthal¹ ✉

Cutaneous T-cell lymphomas (CTCLs) are non-Hodgkin lymphomas caused by malignant T cells which migrate to the skin and lead to rash-like lesions which can be difficult to distinguish from inflammatory skin conditions like atopic dermatitis (AD) and psoriasis (PSO). To characterize CTCL in comparison to these differential diagnoses, we carried out multi-antigen imaging on 69 skin tissue samples (21 CTCL, 23 AD, 25 PSO). The resulting protein abundance maps were then analyzed via scoring functions to quantify the heterogeneity of the individual cells' neighborhoods within spatial graphs inferred from the cells' positions in the tissue samples. Our analyses reveal characteristic patterns of skin tissue organization in CTCL as compared to AD and PSO, including a combination of increased local entropy and egophily in T-cell neighborhoods. These results could not only pave the way for high-precision diagnosis of CTCL, but may also facilitate further insights into cellular disease mechanisms.

Cutaneous T-cell lymphomas (CTCL) present in the form of erythematous lesions, eruptions or patches on the skin. These lesions can both clinically and histologically resemble other non-cancerous inflammatory dermatological conditions, including atopic dermatitis (AD) and psoriasis (PSO)^{1–4}, making diagnosis of CTCL challenging. Since CTCL needs to be treated thoroughly early on^{5–9}, this is a major clinical problem. Techniques to reliably distinguish CTCL from the mimicking conditions AD and PSO therefore hold the promise to improve patient care in CTCL.

While several studies have explored the spatial heterogeneity of the tumor microenvironment in CTCL and its potential relevance in prognosis^{10,11}, a systematic quantification of spatial tissue heterogeneity in the context of CTCL does not exist. This prompted us to generate imaging-based spatially resolved protein abundance maps of skin tissue samples from CTCL, AD, and PSO patients treated at the University Hospital Erlangen, using multi-epitope ligand cartography (MELC)^{12,13}. Based on a visual assessment of the images, tissue organization was indeed altered in CTCL compared to AD and PSO samples. To quantify and objectivize this first subjective impression, we used the popular Squidpy package¹⁴ to generate graph representations of our data, where nodes are cells annotated with their cell type and edges encode spatial vicinity.

Since analysis with existing techniques available in Squidpy revealed only few differences, we developed a Python package called SHouT (short for “spatial heterogeneity quantification tool”, available at <https://github.com/bionetslab/SHouT>), which allows to quantify tissue heterogeneity based on spatial cell graphs. Relying on well-established concepts such as Shannon's entropy¹⁵ which have been successfully used for spatial heterogeneity quantification in medical imaging^{16–18} and single-cell omics data^{19–21}, SHouT provides fast and user-friendly implementations of six graph-based scores to characterize tissue organization in spatial omics data. SHouT revealed clear CTCL-specific characteristics of tissue organization, including higher mixing of cells of different types in the vicinity of T cells in CTCL as compared to AD and PSO. Randomization tests based on label permutation and subsampling verified the robustness of these findings.

Results

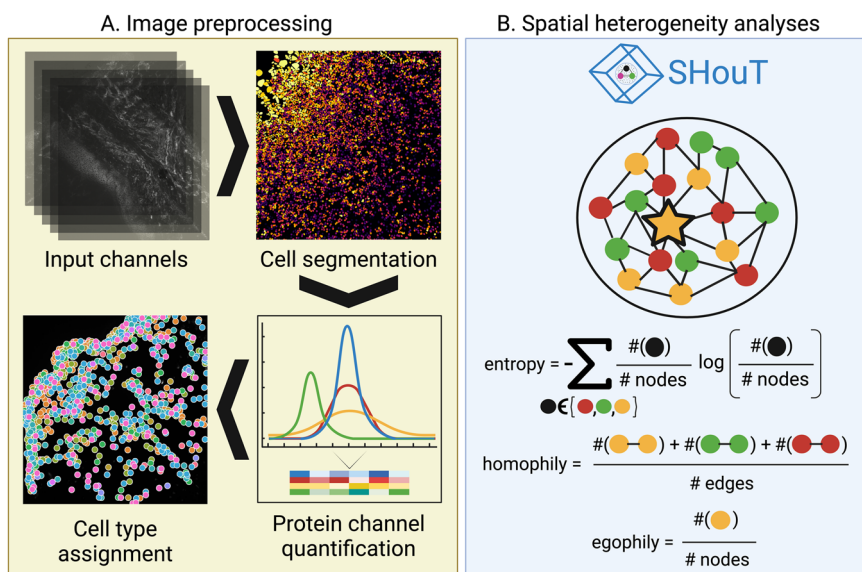
Overview of study design

Figure 1 provides an overview of our work (see “Methods” for details): Using MELC, we generated spatial protein abundance maps for 69 skin tissue samples (21 CTCL, 23 AD, 25 PSO). With this, we obtained images of resolution 2018 × 2018 pixels for at least 35 protein channels per sample, leading to over 140 million pixel values. We then carried out cell

¹Biomedical Network Science Lab, Department Artificial Intelligence in Biomedical Engineering (AIBE), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany. ²Department of Dermatology, Uniklinikum Erlangen, Deutsches Zentrum Immuntherapie (DZI), Comprehensive Cancer Center Erlangen-European Metropolitan Area of Nuremberg (CCC ER-EMN), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany.

✉ e-mail: david.b.blumenthal@fau.de

Fig. 1 | Overview of our analyses. **A** We generated multi-antigen images for skin tissue samples from CTCL, AD, and PSO patients. Subsequently, images were pre-processed via cell segmentation, cell-level protein abundance quantification, and cell-type assignment. **B** We then computed spatial graph representations for all samples, which we analyzed using Squidpy, Leibovici's entropy, modularity, as well as different heterogeneity scores implemented in our Python package SHouT: local and global entropy, local and global homophily, and egophily.



segmentation for all images, using the propidium iodide and CD45 channels as markers for the nucleus and cell membrane, respectively. Upon manual inspection of the segmentation results, one CTCL sample was excluded due to insufficient segmentation quality, leaving us with a cohort of 68 patients in total. Subsequently, adaptive thresholding was used to quantify protein abundances within the individual cells, and cell types were assigned via a rule-based marker gene approach (Supplementary Fig. 1), using skin tissue single-cell RNA sequencing (scRNA-seq) data from the Human Protein Atlas (HPA)²² as reference. We then projected the cell-type labels back on the 2018 × 2018 image grid, and quantified heterogeneity of cell-type co-occurrence patterns in the individual samples via Leibovici's entropy²³ (using the `leibovici` function from the `geontropy` package available at <https://github.com/maxkrysch/geontropy>).

Quantifying spatial heterogeneity of spatial omics data with methods such as Leibovici's entropy that operate directly on the image space is straightforward but has the disadvantage that pixels, i.e. entities without direct biological meaning, are the primary units of analysis. To also enable cell-based spatial heterogeneity quantification, we therefore computed a spatial graph for each sample, where two cells are connected by an edge if they are adjacent to each other in the spatial map. The spatial graphs were then analyzed with Squidpy's `centrality_scores` and `nhood_enrichment` functions, which compute the degree and closeness centralities of the individual cells within the spatial graphs and enrichment of cell-type co-occurrence within one-hop spatial neighborhoods, and by computing the modularity²⁴ of the node partitions induced by the assigned cell types (using the `modularity` function of NetworkX²⁵). We selected these scores because they are readily available in the Python packages Squidpy and NetworkX, which are very widely used in the field of spatial omics analysis. However, it is important to note that none of these scores was originally proposed as a score to distinguish between samples from different conditions (for instance, the primary purpose of the neighborhood enrichment method is to mine spatial omics data for cell communication events).

In addition to relying on these existing tools, we analyzed the cell graphs with our novel Python package SHouT, which we developed with the task of downstream differential analysis in mind. SHouT provides two sample-level scores that quantify overall tissue heterogeneity:

- The edge-based *global homophily* score represents the fraction of edges in the spatial graph that connect cells of the same type. A high score indicates low overall heterogeneity.
- The node-based *global entropy* score represents how well-balanced the numbers of cells of each cell type are throughout the network. A high score indicates high overall heterogeneity.

Moreover, SHouT provides three cell-level scores that quantify tissue heterogeneity within the *r*-hop neighborhood of an individual cell *c* in the spatial graphs (Fig. 1B):

- The edge-based *local homophily* score represents the fraction of edges in the *r*-hop neighborhood of *c* that connect cells of the same type. A high score indicates low heterogeneity in the tissue region surrounding cell *c*.
- The node-based *local entropy* score represents how well-balanced the numbers of cells of each cell type are in the *r*-hop neighborhood of *c*. A high score indicates high heterogeneity in the tissue region surrounding cell *c*.
- The node-based *egophily* score represents the fraction of cells within the *r*-hop neighborhood of *c* that have the same cell type as *c*. A high score indicates low heterogeneity in the tissue region surrounding cell *c*.

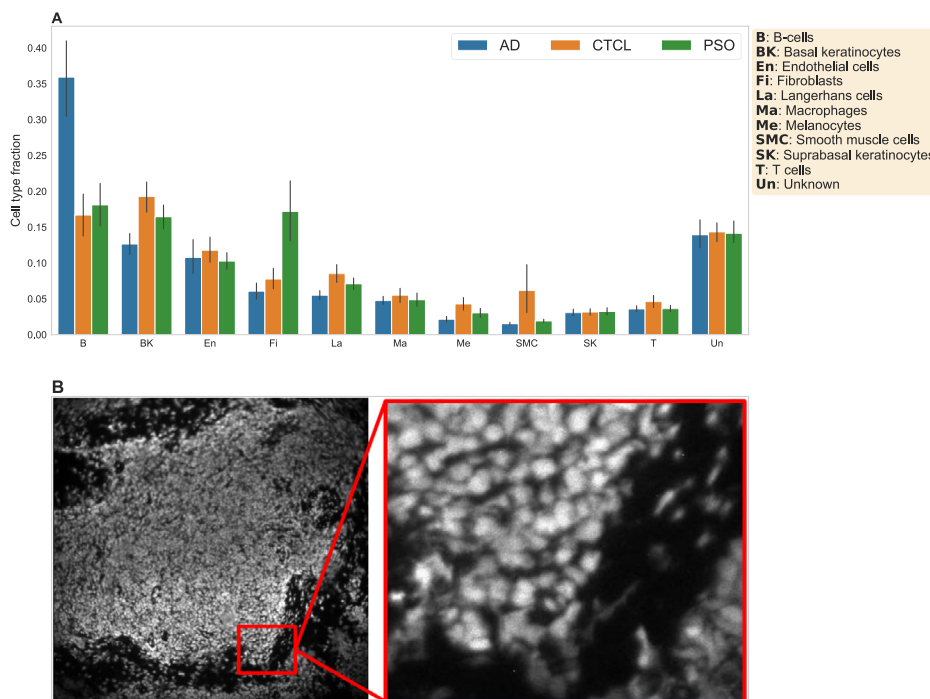
By specifying the radius *r*, the user can select the desired granularity of the cell-level scores. With *r* = 1, only cells in the immediate vicinity of *c* are considered. On the other extreme, setting *r* to the diameter of the spatial graph renders local homophily and entropy equivalent to global homophily and entropy, respectively. Setting *r* to intermediate values allows the quantification of mesoscale patterns in the spatial graphs.

Figure 2 provides a high-level overview of the dataset. Figure 2A shows the average cell-type composition stratified by dataset type (see Supplementary Fig. 2 for underlying distributions of sample-specific cell-type fractions and Supplementary Fig. 3 for a joint t-SNE projection of the protein expression data of all 68 patients, colored by cell type). With the exception of an increased abundance of endothelial cells in AD samples and increased abundance of macrophages in PSO samples, we did not observe strong differences in cell-type composition between the three conditions, emphasizing the need for more in-depth analyses. Figure 2B shows the propidium iodide channel used for cell nucleus identification for one sample.

Heterogeneity analysis reveals CTCL-specific patterns in the neighborhoods of T cells and basal keratinocytes

Figure 3 shows the results of analyzing the spatial graphs with SHouT, Leibovici entropy, modularity, and Squidpy's `centrality_scores` function (see Supplementary Fig. 10 for the results obtained with the `nhood_enrichment` function). All shown *P* values were computed with the two-sided Mann-Whitney *U* (MWU) test and were Bonferroni-corrected with respect to the number of tests per score type (neighborhood enrichment scores: number of condition pairs × number of cell-type pairs; centrality scores, network modularity, Leibovici centrality, and all global

Fig. 2 | Summary of the multi-antigen imaging dataset used for this study. **A** Mean cell-type fractions per condition, across all of the samples. **B** The propidium iodide channel used for cell segmentation for one sample. Left: Compressed image of the entire tissue section. Right: Magnified view representing the actual resolution used for cell segmentation.



SHouT scores: number of condition pairs; local heterogeneity SHouT scores: number of condition pairs \times number of cell types \times number of tested radii (10)).

The two global heterogeneity scores show that overall tissue heterogeneity is highest in CTCL (lowest global homophily, largest global entropy), followed by PSO and AD (Fig. 3A1). This is supported by the results obtained with Leibovici centrality (Fig. 3A2, largest values for CTCL, indicating higher heterogeneity) and network modularity (Fig. 3A3, lowest values for CTCL, indicating decreased clustering of cells of the same type). However, of all four global heterogeneity measures, only the differences in global homophily reach statistical significance for both the comparison between CTCL and PSO and the comparison between CTCL and AD.

When focusing on T cells (Fig. 3B), we observe significantly elevated local entropy (Fig. 3B1) and egophily (Fig. 3B2) scores in CTCL. That is, in skin samples from CTCL patients, T cells tend to cluster together (unsurprisingly, since the tumor cells are T cells) and at the same time are surrounded by tissue that exhibits a higher mixing of cell types than the spatial neighborhoods of T cells in AD or PSO (increased local entropy). Interestingly, the differences are more pronounced and the score distributions are smoother for radius $r = 5$ than for $r = 1$, thereby highlighting the importance of incorporating r -hop neighborhoods into SHouT's local heterogeneity scores. At radius $r = 100$, the score distributions become less smooth, but nonetheless maintain the pronounced differences between the conditions (see Supplementary Figs. 4–6 for additional results with $r = 5$ for all cell types; and Supplementary Fig. 7 for distributions of local SHouT scores across radii $r \in \{1, 2, 3, 4, 5, 10, 20, 50, 100, 500\}$ in T-cell neighborhoods).

In contrast to SHouT, Squidpy's centrality_scores function did not reveal significant differences between the three conditions (Fig. 3B3, see Supplementary Fig. 9 for additional results corresponding to all cell types). Neighborhood enrichment analysis with Squidpy's nhoo-d_enrichment function (Supplementary Fig. 10) did reveal some differences in co-occurrence between T cells and, respectively, macrophages, fibroblasts, and basal keratinocytes. However, the observed differences are much smaller than for local entropy and homophily, do not reach statistical significance after multiple testing correction, and no clear picture emerges that would allow to robustly distinguish CTCL samples from AD and PSO samples based on neighborhood enrichment scores of T cells.

We obtained highly significant differences in SHouT's local heterogeneity scores also for another cell type besides T cells, namely, basal keratinocytes (Fig. 3C, see Supplementary Fig. 8 for distributions of local SHouT scores across radii $r \in \{1, 2, 3, 4, 5, 10, 20, 50, 100, 500\}$ in basal keratinocyte neighborhoods). Like for T cells, local entropy is increased in CTCL in comparison to AD and PSO (Fig. 3C1), i.e., tissue in the vicinity of basal keratinocytes exhibits a higher cell-type mixing in CTCL than in the two other conditions. Moreover, and in contrast to the results for T cells, we simultaneously observe elevated local homophily scores for basal keratinocytes (Fig. 3C2). That is, tissue in the vicinity of basal keratinocytes exhibits a higher co-localization of cells of the same type in CTCL than in PSO or AD (increased local homophily), even though the cell-type heterogeneity is increased (increased local entropy). This shows that the seemingly similar local entropy and local homophily scores indeed quantify distinct properties of spatial tissue organization. As for T cells, the differences between the three conditions are more pronounced for $r = 5$ than for $r = 1$. For larger radii including $r = 100$, the differences remain highly significant, but the score distributions again become noisier than for $r = 5$.

Squidpy's centrality_scores function did not reveal significant differences related to basal keratinocytes (Fig. 3C3). Neighborhood enrichment analysis, however, showed that in CTCL samples, the occurrence of macrophages, melanocytes, smooth muscle cells, suprabasal keratinocytes, and Langerhans cells is increased in the vicinity of basal keratinocytes (Supplementary Fig. 10), although results do not reach statistical significance after multiple testing correction.

The identified CTCL-specific patterns are robust to label permutation and subsampling

We carried out permutation tests to assess the reliability of the identified differences in the local heterogeneity scores with radius $r = 5$ shown in Fig. 3. Specifically, we shuffled the condition labels across all 68 samples and then used the MWU test to compare the local heterogeneity scores between the obtained randomized conditions. We repeated this for 100 iterations, leading to 100 P values per cell type and condition pair. Figure 4 shows the resulting P value distributions (histograms), together with the P values for the original condition labels (red lines). For all eleven combinations of scores and condition pairs where results are significant for the original, unshuffled condition labels (all plots in Fig. 4, except for the one at the bottom left), the P

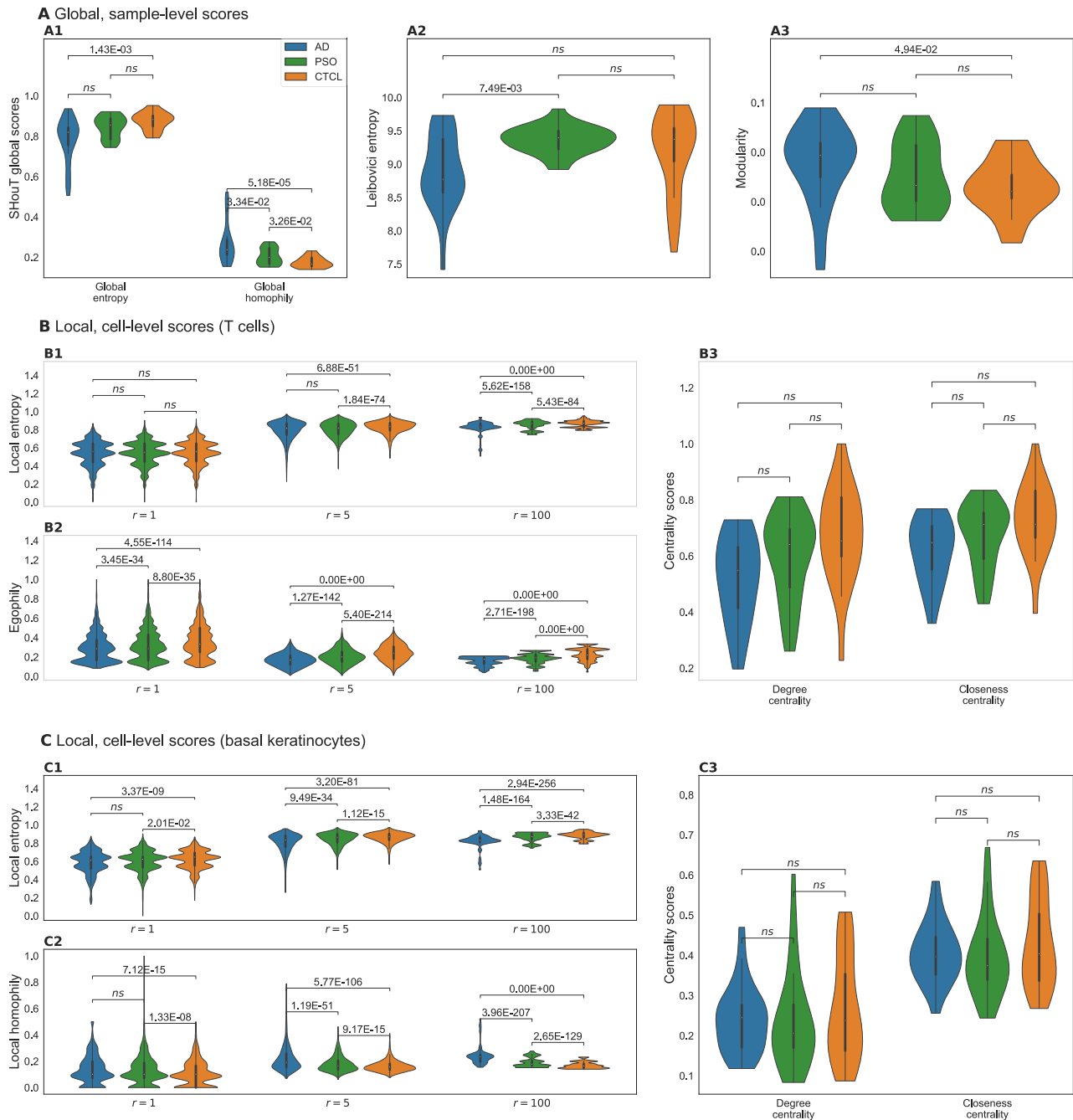


Fig. 3 | Results of spatial analyses of AD, PSO, and CTCL samples with adjusted MWU *P* values. The most important differences between CTCL samples as compared to AD or PSO are: elevated local entropy and egophily scores for neighborhoods of T cells; and elevated local entropy and local homophily scores for

neighborhoods of basal keratinocytes. **A** Distributions of global (i.e., sample-level) scores for the different conditions. **B** Distributions of local (i.e., cell-level) scores for all T cells from samples from the different conditions. **C** Distributions of local scores for all basal keratinocytes from samples from the different conditions.

values obtained for the original condition labels are much smaller than those obtained for the shuffled labels. This indicates that the heterogeneity scores indeed reveal robust differences between CTCL, AD, and PSO.

Moreover, we carried out subsampling to assess if the identified CTCL-specific tissue organization patterns are robust to varying compositions of the CTCL, AD, and PSO cohorts. Specifically, we subsampled 15 samples for each of the three conditions. Subsequently, for each condition pair, we computed MWU *P* values based on the SHouT's local heterogeneity scores (with $r = 5$) for T cells and basal keratinocytes, using only the samples from the subsampled patients. We repeated this 100 times, leading to 100 MWU *P* values for each condition pair, considered cell type, and SHouT score. The resulting *P* value distributions (boxplots) are shown in Fig. 5, together with

the *P* values obtained when using all samples (red lines) and the Bonferroni-corrected significance cutoff $P_{\text{cutoff}} = 0.05 / (\text{number of condition pairs} \times \text{number of cell types} \times \text{number of tested radii})$ (blue lines). For all cases where the original *P* values are significant (red lines to the right of blue lines), all or the vast majority of the *P* values obtained upon subsampling are significant, too. This shows that SHouT reveals CTCL-specific tissue patterns that are robust to subsampling.

SHouT scales to samples with large numbers of cells

To ensure usability, runtime efficiency is an important property of data-centric methods for the analysis of biomedical data. We therefore systematically tested SHouT with respect to its runtime requirements when run on

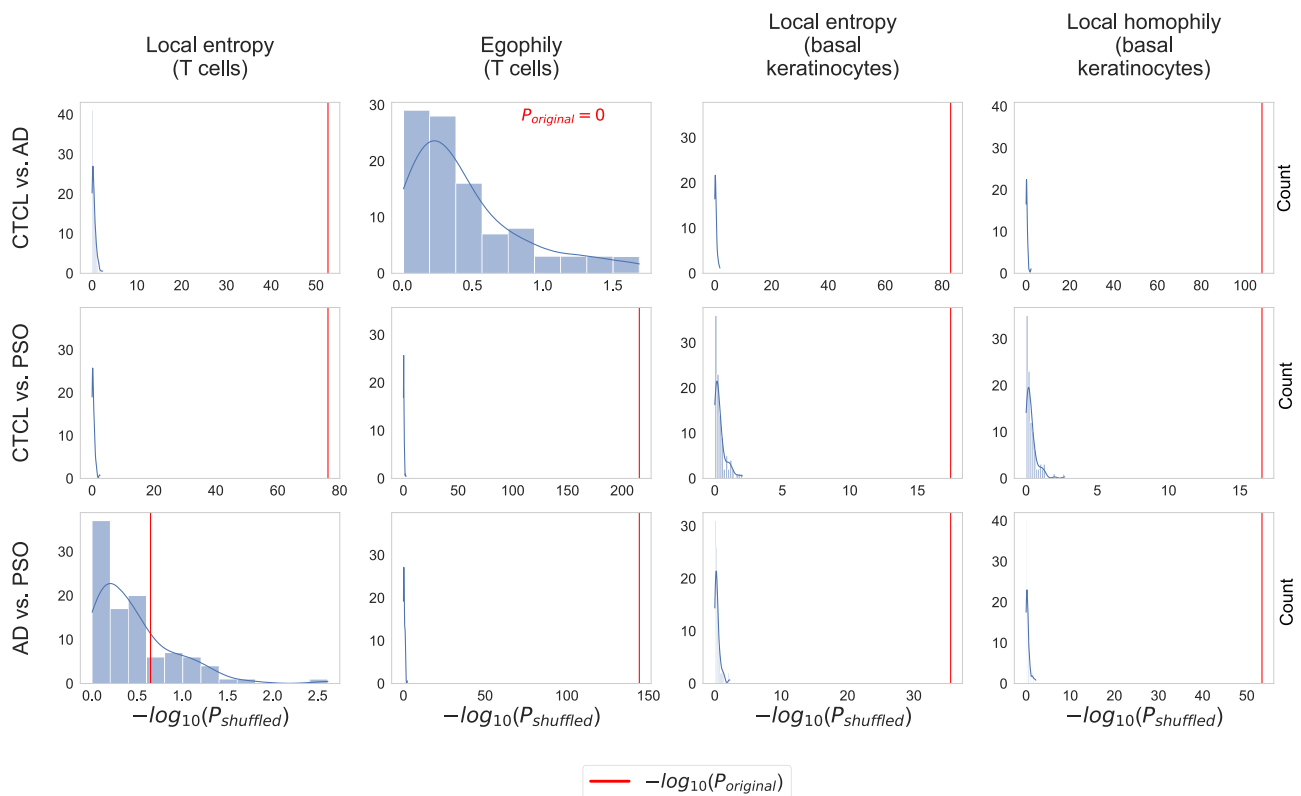


Fig. 4 | Results of permutation tests to assess the robustness of the differences in the local heterogeneity scores with radius $r = 5$ shown in Fig. 3. All P values were computed with the MWU test. Unlike for the P values shown in Fig. 3, we did not

apply Bonferroni correction to adjust for multiple testing (since we are comparing P values, adjusting for multiple testing leaves the results invariant).

samples with varying numbers of cells or when run with varying radii. The results are shown in Fig. 6. We observe that SHouT scales linearly with the numbers of cells per sample, achieving runtimes of little more than a minute even for the samples with the highest cell counts in our dataset (Fig. 6A). Increasing the radius r only marginally increases the runtime, showing that SHouT's heterogeneity scores can be computed efficiently independently of the choice of r (Fig. 6B).

Discussion

Our analyses identified CTCL-specific patterns of tissue organization as compared to PSO and AD in the vicinity of T cells and basal keratinocytes. Since CTCL is a T-cell malignancy, observing characteristic patterns in the vicinity of T cells is not too surprising. In fact, existing studies suggest that malignant T cells and their cross-talk with other cells induce disorganization in the epidermal architecture^{26–28}, which is well aligned with our findings. Also, our results for basal keratinocytes are plausible in the light of the literature. For instance, several studies have identified hyperproliferation and/or de-differentiation of keratinocytes in CTCL^{27,29,30}.

An important aspect of our findings is that the SHouT scores underlying our results are (1) purely quantitative, (2) interpretable by design (each SHouT score has a natural interpretation that can be explained with few sentences), and (3) deterministically computed (computing the SHouT scores does not involve randomized subroutines). Their purely quantitative nature sets our SHouT scores apart from the current standard of care in CTCL diagnosis, which is typically based on (subjective) expert opinions. Their interpretability is a decisive advantage over predictions provided by deep learning models, which are often perceived as black boxes³¹. The fact that SHouT is deterministic ensures that it consistently computes the same results when run several times on equivalent input—unlike many other methods in data-centric biomedicine^{32,33}. Together, these three properties make the SHouT scores ideal ingredients for potential future biomarkers based on spatial omics data, both in the CTCL use case presented here and

beyond. Therefore, we are convinced that SHouT will be of value for the analysis of spatial omics data also beyond this study, complementing existing toolboxes such as Squidpy.

However, several limitations of our study have to be addressed before translation into clinical care may become feasible. First and foremost, the MELC technology used to generate the data underlying this study is highly non-standardized (only few prototypes exist worldwide, two of them at UKER). This had strong implications for the data analysis pipeline used for this study, where we had to resort to customized solutions for almost all pre-processing steps (cell segmentation, protein abundance quantification, cell-type assignment). While we are confident that the methodological choices we made for individual steps in our pipeline are adequate for the data, they still introduce a certain amount of contingency which may distort the results of our analyses in a hard-to-control way. Before translation into clinical care can become an option, it would hence be important to see if our results can be reproduced for spatial omics generated with more established technologies (e.g., MERFISH³⁴ or MIBI³⁵), for which validated pre-processing pipelines are available.

Methods

Sample collection and multi-antigen imaging

In total, 69 skin tissue samples (21 CTCL, 23 AD, 25 PSO) from a total of 27 treated patients (8 CTCL, 7 AD, 12 PSO) were collected at the University Hospital Erlangen. The study has been approved by the Ethics Committee of the Medical Faculty of the Friedrich-Alexander-Universität Erlangen-Nürnberg (approval date: July 5, 2023; approval number: 23-132-B). We used the MELC technology^{12,13,36} to generate multi-antigen imaging data. MELC efficiently combines seamless assimilation of molecular and anatomical information in situ by employing a cyclic process of three steps: (1) protein-specific fluorescent antigen staining, (2) imaging, and (3) photobleaching. The workflow is completely automated, performing multi-antigen imaging on a fixed and mounted tissue or cell sample, without the

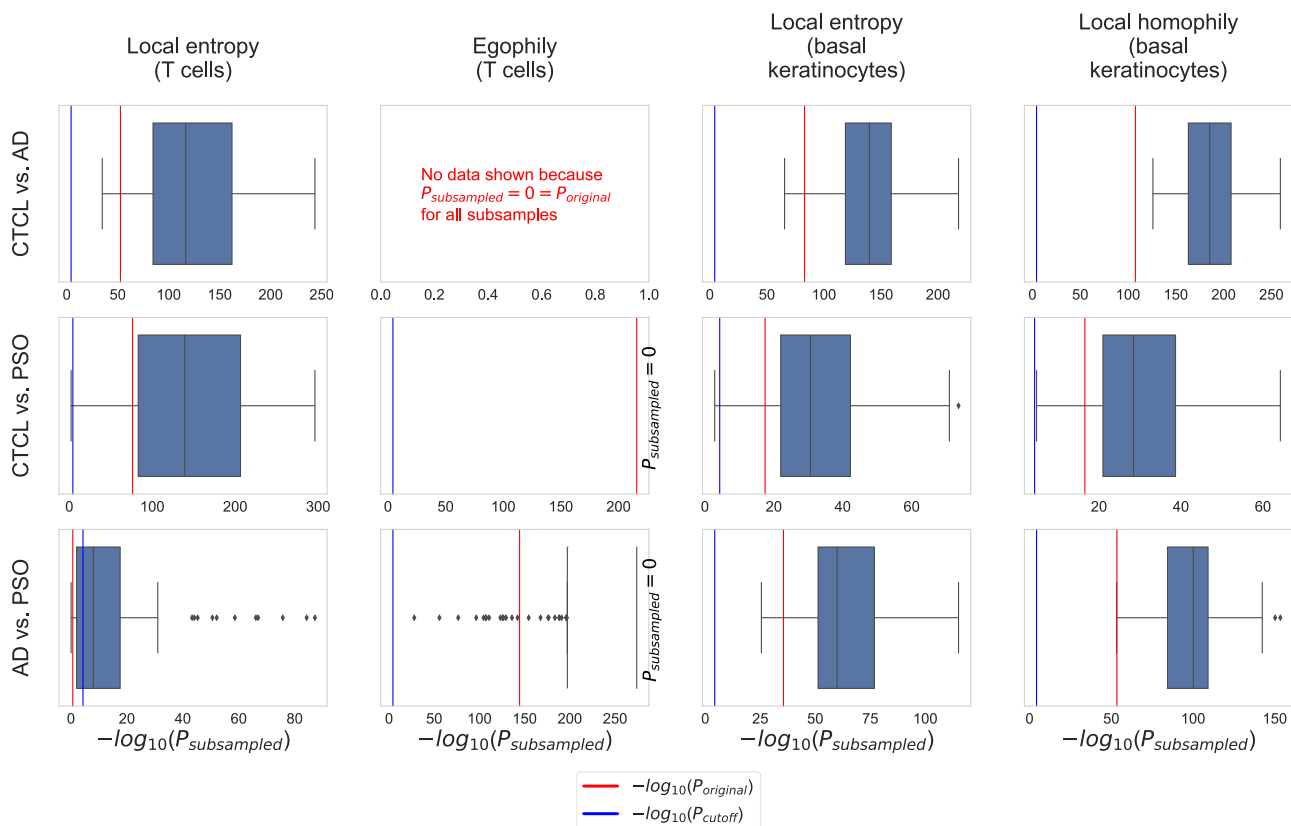


Fig. 5 | Results of subsampling tests to assess the robustness of the differences in the local heterogeneity scores with radius $r = 5$ shown in Fig. 3. All P values were computed with the MWU test. The significance cutoff is Bonferroni-adjusted, i.e.,

$P_{\text{cutoff}} = 0.05 / (\text{number of condition pairs} \times \text{number of cell types} \times \text{number of tested radii}) = 0.05 / (3 \times 11 \times 10)$. P values are non-adjusted.

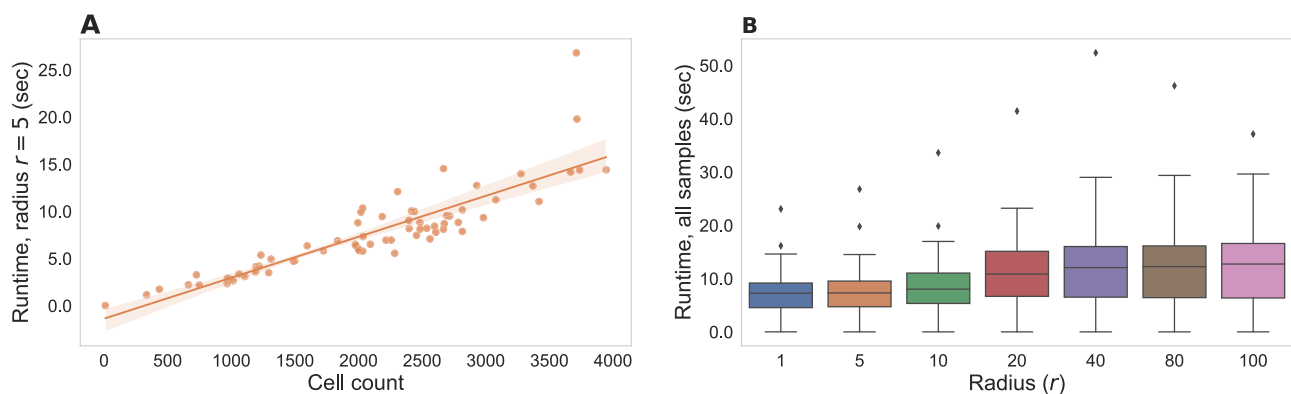


Fig. 6 | Results of scalability tests for our Python package SHouT. Runtime measurements include all subroutines detailed in the “Quantification of spatial tissue heterogeneity with SHouT” subsection of the “Methods”: construction of the spatial neighborhood graph $G = (V, E, \lambda_V)$, computation of the two global scores $H(G)$ and $h(G)$, and computation of the three local scores $H_r(c)$, $h_r(c)$, and $e_r(c)$ for all cells $c \in V$ and a fixed radius r . To ensure a stable execution environment, tests were run on a

Linux compute server with 500 GB of main memory and four AMD EPYC 7402 24-core processors with 1.5 GHz (without using multi-threading). However, SHouT does not require such large-scale resources and can be run on a standard laptop. **A** Runtimes with a fixed radius $r = 5$ for all samples in our dataset. **B** Runtime distributions for varying radii across all samples in our dataset.

requirement of any human involvement. Thanks to this automated workflow, post imaging, the individual protein channels can be mapped into one consolidated tissue map whilst preserving spatial information. In order to assemble a set of antibodies with a strong and specific staining pattern, we screened over 500 antibodies on CTCL tissue. This unbiased approach yielded 36 antibodies (Supplementary Table 1).

Cell segmentation and protein abundance quantification

Owing to the properties of DNA- and histone-binding, propidium iodide is pervasively used to stain cell nuclei in fluorescent microscopy³⁷. We

therefore used the propidium iodide channel for cell nucleus segmentation, relying on a pre-trained model provided by the `stardist.models.StarDist2D` function of the popular cell detection library StarDist^{38,39}.

For cell membrane segmentation, we used the channel for the transmembrane protein tyrosine phosphate (CD45), which is not only present in all nucleated hematopoietic cells but is also one of the most commonly found membrane proteins in such cells⁴⁰. However, owing to the often non-convex shape of cell membranes as opposed to cell nuclei in our dataset, the StarDist model under-segments the cytoplasm. As a workaround, for all nuclei where the StarDist model managed to automatically identify the cell

membrane, we calculated an average ratio between the radii of cells and nuclei. We then used this average ratio to draw a circle around each of the segmented nuclei for which the StarDist model did not identify a cell membrane. If two such circles overlapped, pixels within the intersection were assigned to the closest nucleus, thereby ensuring that all cell segments were pairwise disjoint. All segmentation results were manually inspected by a histopathology expert, and one CTCL sample was excluded due to insufficient segmentation quality.

Transitioning from pixel intensity to cell-level protein abundances was challenging due to the following factors: (1) MELC images are especially susceptible to salt-and-pepper noise. (2) Different protein channels have different intensity levels. (3) Even within the same channel, different regions often have different levels of intensity. Because of these factors, there is no single mapping which would allow to uniformly compute cell-level protein abundance scores based on pixel intensities across all samples and sample regions. We therefore made use of adaptive thresholding which helps circumvent this issue by breaking the image into smaller windows comprising fewer pixels, subsequently computing a different threshold for every window. Specifically, we used OpenCV's (<https://opencv.org/>) `cv2.adaptiveThreshold` function to perform adaptive thresholding. This function binarizes the individual channels by setting each pixel to 1 if its intensity is above the Gaussian weighted mean of its vicinity (here, a window of size 201×201) plus a constant C (here, the standard deviation of the intensity values across the current protein channel). We then transformed the binarized images into floating point protein abundance matrices A by setting the abundance $A(c, p)$ of protein p in cell c to the fraction of positive binarized pixels within the segment corresponding to c .

Cell-type assignment

Since only few exemplars of the MELC imaging system exist, automated cell-type annotation tools for our data do not exist. Moreover, initial tests showed that cell-type annotation tools developed for data generated by other multiplexing platforms are not applicable for our data. We therefore developed a simple rule-based cell-type annotation workflow, making use of single-cell RNA sequencing (scRNA-seq) data specific to skin tissue from HPA as reference. Specifically, we downloaded normalized gene expression values $X(C, g)$, where C denotes cell clusters provided in HPA and g denotes genes. Moreover, we made use of HPA's cell-type annotations $\sigma(C) \in \Sigma$, where Σ is the set of cell types for skin tissue used by HPA. Since, in HPA, there are cell types to which several clusters are assigned, we computed cell-type-specific gene expression values

$$X(t, g) = \left(\sum_{C: \sigma(C)=t} |C| \right)^{-1} \cdot \sum_{C: \sigma(C)=t} |C| \cdot X(C, g) \quad (1)$$

via weighted mean aggregation, for all pairs of cell types t and genes g .

Let \mathcal{C} be the set of all cells contained in any of the samples for any of the three conditions CTCL, PSO, and AD, \mathcal{G} be the set of all genes encoding the measured proteins, and $A(c, g)$ be the abundance of the protein encoded by g in the cell c . Our cell-type assignment workflow iteratively assigns cell types $t^* \in \Sigma$ to a subset $\mathcal{C}_{t^*} \subseteq \mathcal{C}$ of the cells. Upon assigning the cell type t^* to the cells contained in \mathcal{C}_{t^*} , \mathcal{C} and Σ are updated as $\mathcal{C} \leftarrow \mathcal{C} \setminus \mathcal{C}_{t^*}$ and $\Sigma \leftarrow \Sigma \setminus \{t^*\}$, respectively. The process stops when either $\mathcal{C} = \emptyset$, $\Sigma = \emptyset$, or there are no good genes (defined below). In the latter two cases, the remaining cells in \mathcal{C} are assigned the cell-type label "unknown".

To find t^* and \mathcal{C}_{t^*} within one iteration, we computed HPA-based spread scores

$$s(t, g) = X(t, g) - \max\{X(t', g) | t' \in \Sigma \setminus \{t\}\} \quad (2)$$

for all pairs of not yet assigned cell types $t \in \Sigma$ and genes $g \in \mathcal{G}$ and sorted the pairs in decreasing order of $s(t, g)$, leading to a sorted list L . Genes g with large spread scores $s(t, g)$ are potential marker genes for the cell type t , based on the scRNA-seq data in HPA. Next, we fit a bimodal Gaussian mixture

models to the vectors of protein abundances $(A(c, g))_{c \in \mathcal{C}}$ for the not yet assigned cells (using scikit-learn's `GaussianMixture` class) and all genes $g \in \mathcal{G}$ and assessed if the obtained distributions are indeed bimodal. For this, we checked the condition

$$\mu_0(g) + 1.96 \cdot \sigma_0(g) < \mu_1(g) - 1.96 \cdot \sigma_1(g), \quad (3)$$

where $\mu_0(g)$ and $\sigma_0(g)$ are the mean and standard deviation values for the mode with the lower mean and $\mu_1(g)$ and $\sigma_1(g)$ are the mean and standard deviation values for the mode with the higher mean. Genes for which the condition holds are called "good". When we found at least one good gene, we picked the first pair (t^*, g^*) from the sorted list L for which g^* is good. Then, we assigned the cell type t^* to the cells $\mathcal{C}_{t^*} = \{c \in \mathcal{C} | A(c, g^*) > \mu_1(g^*) - 1.96 \cdot \sigma_1(g^*)\}$ and continued with the next iteration of our cell-type assignment protocol. When no good gene was found (for our data, this happened when "granulocytes" was the only cell-type label left in Σ), we stopped the protocol.

The resulting rule-based cell-type assignment tree is shown in Supplementary Fig. 1. Heatmaps showing the corresponding (A) cell-type-averaged gene expression values from HPA and (B) protein abundances from our data, are shown in Supplementary Fig. 11. In the following, we let T denote the set of used cell-type labels (in our case: the label "unknown" and all labels contained in Σ except "granulocytes") and let $\lambda : \mathcal{C} \rightarrow T$ denote the constructed cell-type label function.

Quantification of spatial tissue heterogeneity with SHouT

SHouT starts by computing sample-specific spatial neighborhood graphs $G = (V, E, \lambda_V)$ from the pre-processed imaging data, where $V \subseteq \mathcal{C}$ is the set of cells for the sample under consideration, the set E contains an edge cc' for two cells $c, c' \in V$ if c and c' are spatially adjacent (computed with Squidpy's `spatial_neighbors` function with the parameter `delaunay` set to `True`), and λ_V denotes the restriction of the cell-type label function λ to V . We decided to use Delaunay triangulation to construct the spatial graphs instead of other approaches such as k -nearest neighbor or distance threshold graphs because the latter approaches require essential hyper-parameters (the number of neighbors k and the distance threshold) which are very difficult to select in a principled way. Based on the spatial graphs, SHouT computes two global scores that quantify heterogeneity for the entire graph G and three local, cell-specific scores.

The first global score—global (normalized) entropy—is defined as

$$H(G) = -\log(|T|)^{-1} \cdot \sum_{t \in T} p_G(t) \cdot \log(p_G(t)) \in [0, 1], \quad (4)$$

where $p_G(t) = |\{c \in V | \lambda_V(c) = t\}| / |V|$ is the fraction of cells in V that are of type t . Large values of $H(G)$ indicate that cell-type heterogeneity is high for the sample represented by G . The second global score—global homophily—is defined as the fraction of edges

$$h(G) = |E|^{-1} \cdot \sum_{cc' \in E} [\lambda_V(c) = \lambda_V(c')] \in [0, 1] \quad (5)$$

in the spatial graph G that connect cells of the same type ($[\cdot] : \{\text{True}, \text{False}\} \rightarrow \{0, 1\}$ is the Iverson bracket, i.e., $[\text{True}] = 1$ and $[\text{False}] = 0$). Large values of $h(G)$ indicate that cells tend to be adjacent to cells of the same type in the sample represented by G .

In addition to the four global scores, SHouT provides three local scores to quantify heterogeneity within the r -hop neighborhood $N_r(c) = \{c' \in V | d_G(c, c') \leq r\}$ of an individual cell $c \in V$. Here, r is a hyper-parameter and $d_G : V \times V \rightarrow \mathbb{N}$ is the shortest path distance. The first local score—local (normalized) entropy—is defined as follows:

$$H_r(c) = -\log(|T|)^{-1} \cdot \sum_{t \in T} p_{N_r(c)}(t) \cdot \log(p_{N_r(c)}(t)) \in [0, 1] \quad (6)$$

The only difference to its global counterpart is that the cell-type fractions $p_{N_r(c)}(t) = |\{c \in N_r(c) | \lambda_V(c) = t\}| / |N_r(c)|$ are computed only with respect to the cells contained in the r -hop neighborhood of c . Also the second local score—local homophily—is defined in a similar way as the global version:

$$h_r(c) = |E_{N_r(c)}|^{-1} \cdot \sum_{c' \in E_{N_r(c)}} [\lambda_V(c') = \lambda_V(c'')] \in [0, 1] \quad (7)$$

Here, the difference to the global version is that we only consider the subset of edges $E_{N_r(c)} = \{c'c'' \in E | c', c'' \in N_r(c)\}$ that connect two cells contained in the r -hop neighborhood of c . The last local score—egophily—does not have a global counterpart. It is defined as the fraction

$$e_r(c) = p_{N_r(c)}(\lambda_V(c)) \in [0, 1] \quad (8)$$

of cells within the r -hop neighborhood of c that have the same cell type as c . SHouT provides very efficient vectorized implementations of all heterogeneity scores, relying on SciPy's `sparse.csgraph.shortest_path` function for fast computation of shortest path distances.

Data availability

The MELC data underlying this study are available on Zenodo: <https://doi.org/10.5281/zenodo.11125482>. Pre-clustered, skin-specific scRNA-seq data used as a reference for cell-type assignment are available at https://www.proteinatlas.org/download/rna_single_cell_type_tissue.tsv.zip. The corresponding cluster annotations can be obtained at https://www.proteinatlas.org/download/rna_single_cell_cluster_description.tsv.zip.

Code availability

The code of the SHouT Python package is available at <https://github.com/bionetslab/SHouT>. Code to reproduce the results reported in this paper is available at https://github.com/bionetslab/ctcl_case_study.

Received: 17 May 2024; Accepted: 22 November 2024;
Published online: 02 December 2024

References

1. Semaan, S., Abel, M. K., Raffi, J. & Murase, J. E. A clinician's guide to cutaneous T-cell lymphoma presenting as recalcitrant eczematous dermatitis in adults. *Int. J. Womens. Dermatol.* **7**, 422–427 (2021).
2. Tomasini, C., Novelli, M., Fanoni, D. & Berti, E. F. Erythema multiforme-like lesions in primary cutaneous aggressive cytotoxic epidermotropic CD8+ T-cell lymphoma: a diagnostic and therapeutic challenge. *J. Cutan. Pathol.* **44**, 867–873 (2017).
3. Papadaki, M., Saraki, K., Karagianni, F., Piperi, C. & Papadavid, E. Cutaneous T-cell lymphoma: aetiopathogenesis and current diagnostic and therapeutic developments. *Eur. J. Dermatol.* **30**, 85–102 (2020).
4. Roediger, B. & Schlapbach, C. T cells in the skin: lymphoma and inflammatory skin disease. *J. Allergy Clin. Immunol.* **149**, 1172–1184 (2022).
5. Willemze, R. et al. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. *Blood* **133**, 1703–1714 (2019).
6. Hristov, A. C., Tejasvi, T. & Wilcox, R. Cutaneous T-cell lymphomas: 2021 update on diagnosis, risk-stratification, and management. *Am. J. Hematol.* **96**, 1313–1328 (2021).
7. Franceschi, J. et al. Survival and prognostic factors in patients with aggressive cutaneous T-cell lymphomas. *Acta Derm. Venereol.* **102**, adv00676 (2022).
8. Brunner, P. M., Jonak, C. & Knobler, R. Recent advances in understanding and managing cutaneous T-cell lymphomas. *F1000Res* **9**, 331 (2020).
9. Miyagaki, T. Diagnosis and prognostic stratification of cutaneous lymphoma. *J. Dermatol.* **49**, 210–222 (2022).

10. Kalliaras, E., Belfrage, E., Gullberg, U., Drott, K. & Ek, S. Spatially guided and single cell tools to map the microenvironment in cutaneous T-cell lymphoma. *Cancers* **15**, 2362 (2023).
11. Feng, Y. et al. Spatial transcriptomics reveals heterogeneity of macrophages in the tumor microenvironment of granulomatous slack skin. *J. Pathol.* **261**, 105–119 (2023).
12. Schubert, W. et al. Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat. Biotechnol.* **24**, 1270–1278 (2006).
13. Schubert, W. Topological proteomics, toponomics, MELK-technology. *Adv. Biochem. Eng. Biotechnol.* **83**, 189–209 (2003).
14. Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
15. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
16. Suresh, R. et al. Entropy-based spatial heterogeneity analysis in pathological images for diagnostic applications. In *Data Science for Photonics and Biophotonics*, Vol. 13011 (ed. Bocklitz, T) 130110E (SPIE, 2024). <https://doi.org/10.1117/12.3022363>.
17. Vu, T. et al. FunSpace: a functional and spatial analytic approach to cell imaging data using entropy measures. *PLoS Comput. Biol.* **19**, e1011490 (2023).
18. Weis, C.-A. et al. Histopathology of thymectomy specimens from the MGTX-trial: Entropy analysis as strategy to quantify spatial heterogeneity of lymphoid follicle and fat distribution. *PLoS ONE* **13**, e0197435 (2018).
19. Chen, Z., Yang, Z., Yuan, X., Zhang, X. & Hao, P. scSensitiveGeneDefine: sensitive gene detection in single-cell RNA sequencing data by Shannon entropy. *BMC Bioinform.* **22**, 211 (2021).
20. Dussiau, C. et al. Hematopoietic differentiation is characterized by a transient peak of entropy at a single-cell level. *BMC Biol.* **20**, 60 (2022).
21. Richard, A. et al. Single-cell-based analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. *PLoS Biol.* **14**, e1002585 (2016).
22. Karlsson, M. et al. A single-cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, eabh2169 (2021).
23. Leibovici, D. G. Defining spatial entropy from multivariate distributions of co-occurrences. In *9th International Conference on Spatial Information Theory, Lecture Notes in Computer Science*, Vol. 5756 (eds Hornsby, K.S., Claramunt, C., Denis, M., & Ligozat, G) 392–404 (Springer Berlin Heidelberg, 2009).
24. Newman, M. E. J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **103**, 8577–8582 (2006).
25. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *SciPy 2008* (eds Varoquaux, G., Vaught, T. & Millman, J.) 11–15 (Pasadena, 2008).
26. Thumann, P. et al. Interaction of cutaneous lymphoma cells with reactive T cells and dendritic cells: implications for dendritic cell-based immunotherapy. *Br. J. Dermatol.* **149**, 1128–1142 (2003).
27. Thode, C. et al. Malignant T cells secrete galectins and induce epidermal hyperproliferation and disorganized stratification in a skin model of cutaneous T-cell lymphoma. *J. Invest. Dermatol.* **135**, 238–246 (2015).
28. Stolareenco, V. et al. Cellular interactions and inflammation in the pathogenesis of cutaneous T-cell lymphoma. *Front. Cell. Dev. Biol.* **8**, 851 (2020).
29. Weiss, R. A., Eichner, R. & Sun, T. T. Monoclonal antibody analysis of keratin expression in epidermal diseases: a 48- and 56-kdalton keratin as molecular markers for hyperproliferative keratinocytes. *J. Cell Biol.* **98**, 1397–1406 (1984).
30. Edelson, R. L. Cutaneous T cell lymphoma: the helping hand of dendritic cells. *Ann. N. Y. Acad. Sci.* **941**, 1–11 (2001).
31. Qamar, T. & Bawany, N. Z. Understanding the black-box: towards interpretable and reliable deep learning models. *PeerJ Comput. Sci.* **9**, e1629 (2023).

32. Bernett, J. et al. Robust disease module mining via enumeration of diverse prize-collecting Steiner trees. *Bioinformatics* **38**, 1600–1606 (2022).
 33. Sarkar, S. et al. Online bias-aware disease module mining with ROBUST-Web. *Bioinformatics* **39**, btad345 (2023).
 34. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
 35. Ptacek, J. et al. Multiplexed ion beam imaging (MIBI) for characterization of the tumor microenvironment across tumor types. *Lab. Investig.* **100**, 1111–1123 (2020).
 36. Schubert, W. Exploring molecular networks directly in the cell. *Cytometry A* **69**, 109–112 (2006).
 37. Martin, R. M., Leonhardt, H. & Cardoso, M. C. DNA labeling in living cells. *Cytometry A* **67**, 45–52 (2005).
 38. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell detection with star-convex polygons. In *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*, Vol. 11071 (eds Frangi, A. F., Schnabel, J. A., Davatzikos C. & Alberola-López, C) 265–273 (Springer International Publishing, 2018).
 39. Weigert, M., Schmidt, U., Haase, R., Sugawara, K. & Myers, G. Star-convex polyhedra for 3D object detection and segmentation in microscopy. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Vol. 0, 3655–3662 (IEEE, 2020).
 40. Jung, Y., Wen, L., Altman, A. & Ley, K. CD45 pre-exclusion from the tips of T cell microvilli prior to antigen recognition. *Nat. Commun.* **12**, 3872 (2021).
- analyses. C.O. generated the data. C.O. and M.E. annotated the data. A.M. pre-processed the data. D.B.B., A.B., and A.H. supervised this work. All authors reviewed and approved the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00474-x>.

Correspondence and requests for materials should be addressed to David B. Blumenthal.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Acknowledgements

D.B.B. and A.M. were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—516188180. Figure 1 and Supplementary Fig. 1 were generated with BioRender.com. D.B.B. and A.H. were supported by the German Federal Ministry of Education and Research (BMBF, grant no. 031L0309A).

Author contributions

S.S. and D.B.B. conceived and designed this study, implemented SHouT, analyzed the results, and drafted the manuscript. S.S. carried out the