








Dynamic thresholding and tissue dissociation optimization for CITE-seq identifies differential surface protein abundance in metastatic melanoma

Ulrike Lischetti ^{1,5,6}, Aizhan Tastanova ^{2,6}✉, Franziska Singer ^{3,4}, Linda Grob ^{3,4}, Matteo Carrara^{3,4}, Phil F. Cheng², Julia M. Martínez Gómez², Federica Sella ², Veronika Hauerdinger², Christian Beisel ^{1,7} & Mitchell P. Levesque ^{2,7}✉

Multi-omics profiling by CITE-seq bridges the RNA-protein gap in single-cell analysis but has been largely applied to liquid biopsies. Applying CITE-seq to clinically relevant solid biopsies to characterize healthy tissue and the tumor microenvironment is an essential next step in single-cell translational studies. In this study, gating of cell populations based on their transcriptome signatures for use in cell type-specific ridge plots allowed identification of positive antibody signals and setting of manual thresholds. Next, we compare five skin dissociation protocols by taking into account dissociation efficiency, captured cell type heterogeneity and recovered surface proteome. To assess the effect of enzymatic digestion on transcriptome and epitope expression in immune cell populations, we analyze peripheral blood mononuclear cells (PBMCs) with and without dissociation. To further assess the RNA-protein gap, RNA-protein we perform codetection and correlation analyses on thresholded protein values. Finally, in a proof-of-concept study, using protein abundance analysis on selected surface markers in a cohort of healthy skin, primary, and metastatic melanoma we identify CD56 surface marker expression on metastatic melanoma cells, which was further confirmed by multiplex immunohistochemistry. This work provides practical guidelines for processing and analysis of clinically relevant solid tissue biopsies for biomarker discovery.

¹Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland. ²Department of Dermatology, University Hospital Zurich, University of Zurich, Zurich, Switzerland. ³ETH Zurich, NEXUS Personalized Health Technologies, Wagistrasse 18, 8952 Schlieren, Switzerland. ⁴SIB Swiss Institute of Bioinformatics, Zurich, Switzerland. ⁵Present address: Department of Biomedicine, University Hospital Basel, University of Basel, 4031 Basel, Switzerland. ⁶These authors contributed equally: Ulrike Lischetti, Aizhan Tastanova. ⁷These authors jointly supervised this work: Christian Beisel, Mitchell P. Levesque. ✉email: Aizhan.Tastanova@usz.ch; Mitchell.Levesque@usz.ch

The last decade has been marked by immense progress in the field of single-cell RNA sequencing (scRNA-seq)^{1,2}. As an unbiased technology that does not profile a panel of preselected transcripts, it is ideally suited to resolve cellular heterogeneity in complex healthy or diseased tissues. Single-nuclei RNA-sequencing has been proposed for larger cohorts of fresh-frozen samples or where tissue dissociation is hard to achieve, such as for mature interconnected brain tissues^{3–5} but this is not without its drawbacks⁶. Notably, both single-cell and single-nuclei RNA-seq do not address the gap between RNA and protein expression, which can result from technical capture difficulties (lack in transcript capture/drop-out) and biological processes (translational impediments, post-translational influences, RNA degradation kinetics or protein trafficking to and from the cell surface)^{7,8}. Compared to other single-cell proteomics technologies such as flow and mass cytometry, cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) allows simultaneous capture of gene and protein information as well as the implementation of larger antibody panels by virtue of a highly diverse combinatorial DNA sequence space that can be utilized to tag individual antibodies, thereby avoiding previous limitations of spectral overlap or matching isotope selections⁹. Increasing antibody panel size is essential for improved cell type and cell state resolution of highly heterogeneous tissue samples such as clinical specimens as well as for novel biomarker discovery. The same principle of DNA-barcoded antibodies can be applied for sample multiplexing by cell hashing through the use of antibodies targeting ubiquitously expressed epitopes, which reduces experimental costs, allows doublet detection, and minimizes batch effects¹⁰.

Despite their differences in signal detection, CITE-seq, and flow cytometry share the same requirements of careful experimental setups and considerations in panel design. This includes testing antibody panels for detection of the epitopes of interest, working with optimal antibody titers, washing out unbound antibodies after staining, including appropriate controls, and addressing background staining signals¹¹. However, tailoring an antibody panel is easier for biological samples where the cell type composition is known or expected a priori, such as for PBMCs or healthy tissues. In disease states such as cancer, heterogeneous sample types can vary substantially in tumor content and tumor microenvironment composition from, e.g., low to high tumor presence and immune-rich to stroma-rich sample types. Moreover, clinical biopsies are often limited in tissue size, making antibody panel testings practically difficult.

Currently, CITE-seq applications focus mainly on immunophenotyping of liquid^{12–14} and solid biopsies^{15,16}. Opening the field to clinically relevant solid tissue samples comes with the challenge of potential bias in gene expression (GEX)^{17–19} and surface protein expression (SPEX)^{20,21} introduced by dissociation and cell handling processes. Most tissue types require enzymatic or mechanical processing steps. The proteolytic enzymes used to dissociate the extracellular matrix and cell–cell-junctions require incubation at 37 °C for a prolonged time for optimal function, which can cause stress responses in cellular transcriptomes^{17–19}. Hence, whenever the tissue sample allows, cold digestion or mechanical dissociation is preferred to avoid these stress signatures^{22,23}. In addition, the selection of enzyme cocktails, incubation times, and steps will also influence the types of released cells, with specific cell types being either more sensitive to the dissociation process or requiring extended dissociation to be released from the tissue²⁴. Lastly, the proteolytic activity of digestion enzymes can cleave cell surface proteins, which poses a problem for surface protein measurement techniques such as cytometry and CITE-seq. Trypsin and dispase are examples of such enzymes²⁵. In summary, the choice of tissue dissociation

protocol can greatly affect the observed cell type composition, gene expression, and the spectrum of detectable surface proteins making protocol optimization to minimize these effects imperative. This could be in the form of optimizing for shortest and least harsh dissociation protocols that result in high cell release, avoiding surface protein-cleaving enzymes such as trypsin, and testing if the obtained cell composition matches expected ratios.

Protein expression is usually inferred from RNA expression and vice versa. However, while protein concentrations positively correlate with transcript abundances, the association is not strong^{7,26,27}. The correlation coefficient of RNA and protein expression has been reported to be between 0.5 and 0.7. Large differences could be seen depending on whether observations are based on single-cell or population level⁷, with lower correlations on cell level attributed to the noise observed in scRNA-seq data, such as technical dropouts or differences in biological half-lives between RNA and protein²⁸. RNA-protein correlations are also dependent on measuring technologies and their precision and accuracy in feature (RNA or protein) detection as well as on types of data analyses^{29,30}. For example, correlations vary depending on whether they are calculated between or within features. Between-feature correlations take the average of all feature correlations per tissue leading to higher correlations, while within-feature correlations are proposed to be more appropriate as they take feature-specific variation in RNA-protein correlations into account³⁰. Thus, with low protein predictions from RNA levels for many features, simultaneous profiling of both is necessary to identify actionable targets, although this is rarely performed^{31,32}.

Here, we report the performance of CITE-seq on a range of liquid and solid tissue biopsies. We developed an improved bioinformatics technique using cell type-specific ridge plots on which we implemented antibody signal thresholds to account for background staining signals and visualized their variance across experiments. Solid biopsies included healthy skin, for which we established an optimized dissociation protocol and primary and metastatic melanoma samples. Using a PBMC model, we assessed dissociation-specific gene perturbations and epitope loss on immune cell populations. We evaluated RNA-protein codetection and calculated feature-specific correlations on aggregated sample and single-cell level and report varying correlations depending on the profiled feature, the cohort, and the type of analysis (sample or single cell-based). Finally, we demonstrated the potential for biomarker discovery by applying large antibody panels with minimal prior selection of included surface markers. In summary, we highlight the applicability and possible pitfalls of CITE-seq for liquid and solid tissue samples and paired differential gene expression and protein abundance analysis on clinical samples.

Results

Workflow overview. CITE-seq and cell hashing were performed on liquid and solid tissue biopsies (Fig. 1a, b). Experimentally, cells from 17 samples (Supplementary Data 2) were hashed and stained with a panel of 97 antibodies (Supplementary Data 2) covering key as well as exploratory immuno-oncology markers resulting in 57,261 cells after preprocessing and quality control. Biopsies included slow-frozen biobanked PBMCs, healthy skin, primary melanoma, and lymph node (LN) melanoma metastasis samples. For skin biopsies, five different dissociation protocols were compared for optimal yield, viability, and epitope preservation (Supplementary Note 1). Stringent red blood cell lysis, dead cell, and clump removal improved the quality of the final single-cell suspension prior to CITE-seq processing. Computationally, we implemented cell type-specific gating and setting of sample-specific manual thresholds in analogy to FACS analysis allowing us to perform RNA-protein correlation and differential abundance (DA)

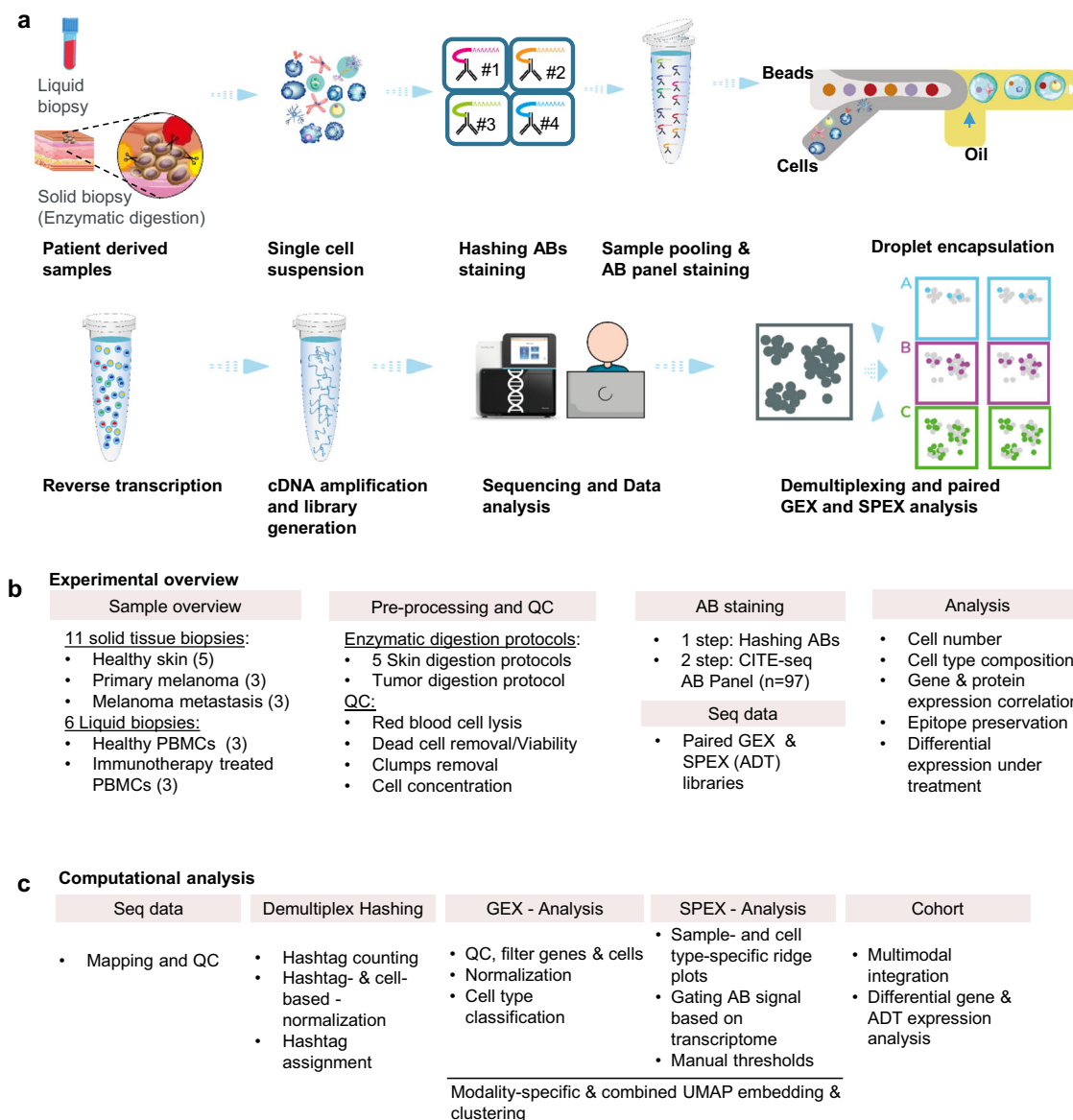


Fig. 1 Study overview. **a** Workflow of sample processing. **b** Experimental overview. **c** Computational analysis. AB Antibody, ADT antibody derived tag, QC Quality control.

analysis (Fig. 1c). Finally, experimental and analytical improvements demonstrated CITE-seq utility for multimodal single-cell profiling and its potential for biomarker discovery.

Dynamic thresholding for the removal of background noise improves cell surface protein signal detection. Taking advantage of single-cell resolution and gene expression-based cell type annotation, we first implemented major cell type-specific protein abundance ridge plots for each sample’s SPEX analysis to identify the protein feature abundance. This strategy is visualized on healthy PBMCs (Supplementary Fig. 1a). In contrast to all-cells protein abundance distribution plots where a marker is plotted for all cells within the sample (Supplementary Fig. 1b), protein abundance visualized on cell type-specific ridge plots allowed for marker detection on individual cell populations (Supplementary Fig. 1c). Furthermore, manual thresholds were set to reduce background noise and increase confidence in antibody signal detection (Supplementary Data 1, Supplementary Data 2) as well as to capture positive antibody signals from cell populations present at a lower frequency (Supplementary Fig. 1). This

approach allowed to confirm protein detection on expected cell types as for CD2 on lymphocytes and rare cell types such as CD16 on natural killer (NK) cells or CD123 on plasmacytoid dendritic cells (Supplementary Fig. 1d and e). Finally, we visualized threshold variation for the top 30 antibodies with the largest variance across experiments underscoring that thresholds vary across experiments and cannot be easily automated or transferred from one experiment to another (Supplementary Fig. 2, Supplementary Data 1). In summary, manual thresholds need to be adapted for each experiment and processed samples, but allow for background signal removal in CITE-seq antibody analysis.

Because of the improved antibody signal detection, we applied the manual thresholds on all samples throughout this study.

Optimization of healthy skin dissociation protocols. The main objective of the study was to establish CITE-seq for skin and primary cutaneous melanoma samples. Skin is a collagen-rich tissue that is difficult to dissociate, requiring several enzymatic digestion steps³³. Therefore, we set out to optimize a dissociation protocol suitable for CITE-seq applications. The following five

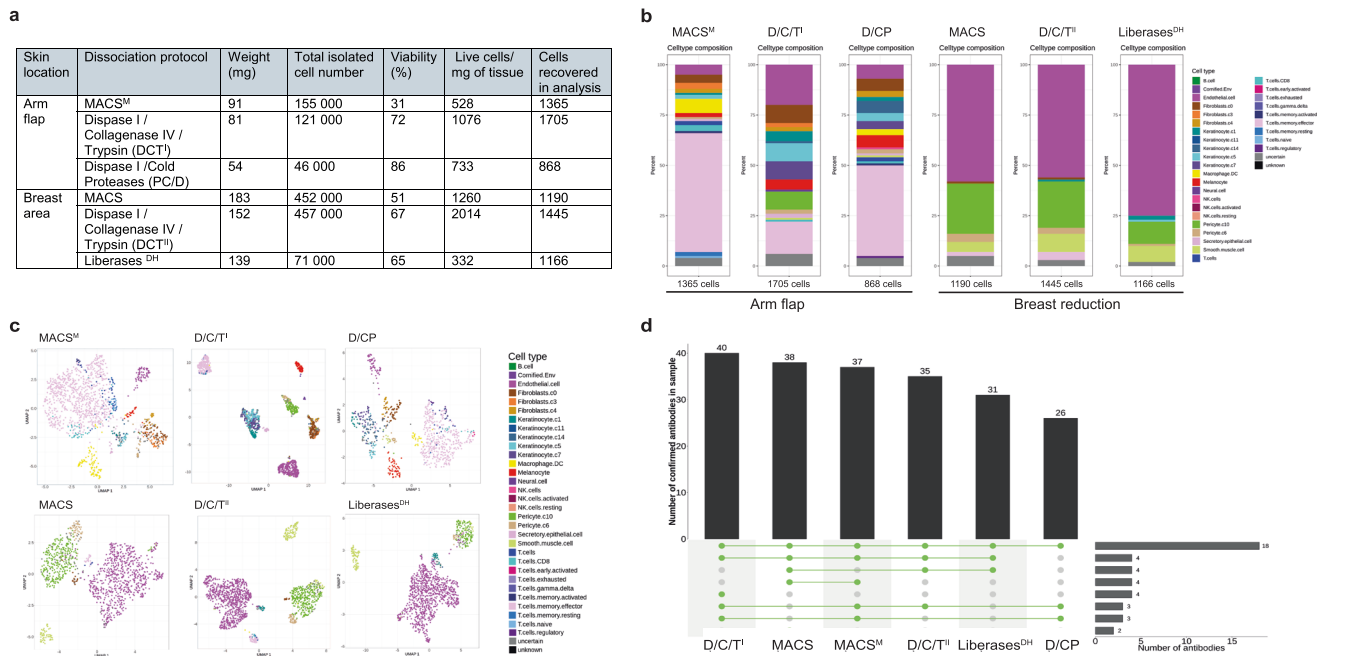


Fig. 2 Protocol optimization for healthy skin dissociation. **a** Table summarizing dissociation efficiency of tested protocols ($n = 2$, biologically independent samples). **b** Cell type composition bar plots, and **c** UMAP plots, of each tested protocol and biopsy type. **d** UpSet plot showing number of detected antibodies per tested protocol.

protocols were tested on surplus healthy skin material from surgery (arm and breast areas, Fig. 2, Supplementary Data 2): MACS skin dissociation kit with and without mechanical dissociation using the gentleMACS Octo Dissociator (MACS^M and MACS, respectively), three-step tissue dissociation protocols consisting of consecutive Dispase I, Collagenase IV and Trypsin with EDTA (0.25%) (D/C/T), a combination of Dispase I and cold-active protease (D/CP) and Liberases^{DH}. In both experiments, the D/C/T protocol had the highest dissociation efficiency, defined as the number of isolated viable cells per mg of tissue with 2–6 fold more released cells (Fig. 2a, Supplementary Data 2). Moreover, this protocol yielded the highest number of cells in the analysis and was able to capture a heterogeneous cell type composition including keratinocytes, melanocytes (in the UV-exposed arm flap skin sample), fibroblasts, and immune cells (Fig. 2b, Supplementary Data 1) with clear cell type clustering (Fig. 2c). In comparison, the MACS protocol-treated sample did not capture the pericyte population; however, it contained a high number of T cells compared to D/C/T and D/CP. Furthermore, the D/C/T protocol yielded high numbers of detected antibodies with 40 and 35 antibodies confirmed in arm flap and breast tissues, respectively (Fig. 2d, Supplementary Data 1, Supplementary Data 2), indicating it retained many epitopes. The selected D/C/T (SkinD) protocol was then applied to a cohort of three healthy skin biopsies from the head and neck area (Fig. 3). The three samples shared similarly diverse cell type compositions including keratinocytes, melanocytes, pericytes, immune cells, and endothelial cells (Fig. 3a, Supplementary Data 1). Combined sample UMAP visualization showed clear clustering into different cell types (Fig. 3b), while the clustering was not affected or biased by hashing or sample origin (Fig. 3c). Finally, cell lineage protein (Fig. 3d, Supplementary Data 1) and gene (Fig. 3e, Supplementary Data 1) marker expression confirmed the cell identities.

A PBMC model reveals the effects of enzymatic treatment on gene expression and epitope preservation. CITE-seq is mainly used on liquid biopsies, which do not require extensive tissue

dissociation steps that could affect cell surface protein presentation. In order to investigate the impact of enzymatic treatments on gene expression and epitope preservation on major immune cell populations, we tested the optimized skin dissociation protocol (SkinD, D/C/T) and a well working solid soft tumor dissociation (TumorD, Supplementary Note 1) protocol³⁴ on PBMC aliquots of three healthy donors against an untreated control (Fig. 4a). Individually treated aliquots of each patient were hashtag-antibody stained, washed and pooled, followed by staining with the panel comprising 97 antibodies (Supplementary Data 2). Stained pools were washed and processed for single-cell sequencing. UMAP visualization of all samples (three donors, each with three treatment conditions) showed robust clustering into B cell, T cell, and myeloid/monocyte populations (Fig. 4b). Major cell types from each donor and treatment condition were pooled and selected protein marker expression was plotted on non-thresholded and thresholded protein data, revealing the degree of background signal in the data and highlighting the impact of setting manual thresholds (Supplementary Fig. 3a and b), see also Grob et al.³⁵. Next, antibody signal detection was assessed between different treatment conditions. Out of 97 tested antibodies, 39 were detected across all samples (Fig. 4c, Supplementary Data 1, Supplementary Data 2); untreated PBMCs showed the highest number with up to 61 detected antibody markers. In addition, three antibodies were detected in all conditions, except for the Donor 2 untreated sample (Fig. 4c, Supplementary Data 1, Supplementary Data 2). Looking at treatment-specific effects in individual samples, the monocyte population clustered separately from untreated ones in both dissociation protocols, while B and T cell populations were mainly unaffected (Fig. 4d, Supplementary Fig. 3c and d). To identify digestion-sensitive surface proteins on major cell types, we performed protein differential abundance (DA) analysis based on normalized counts using the Seurat function findMarkers. Loss of surface protein expression was observed under both digestion protocols in all cell types (Fig. 4e). Specifically, we observed loss of CD4, CD8, CD27, CD335 on T cells, CD335, CD31, CD49f, CD62L, CD69 on monocytes, CD21, CD81,

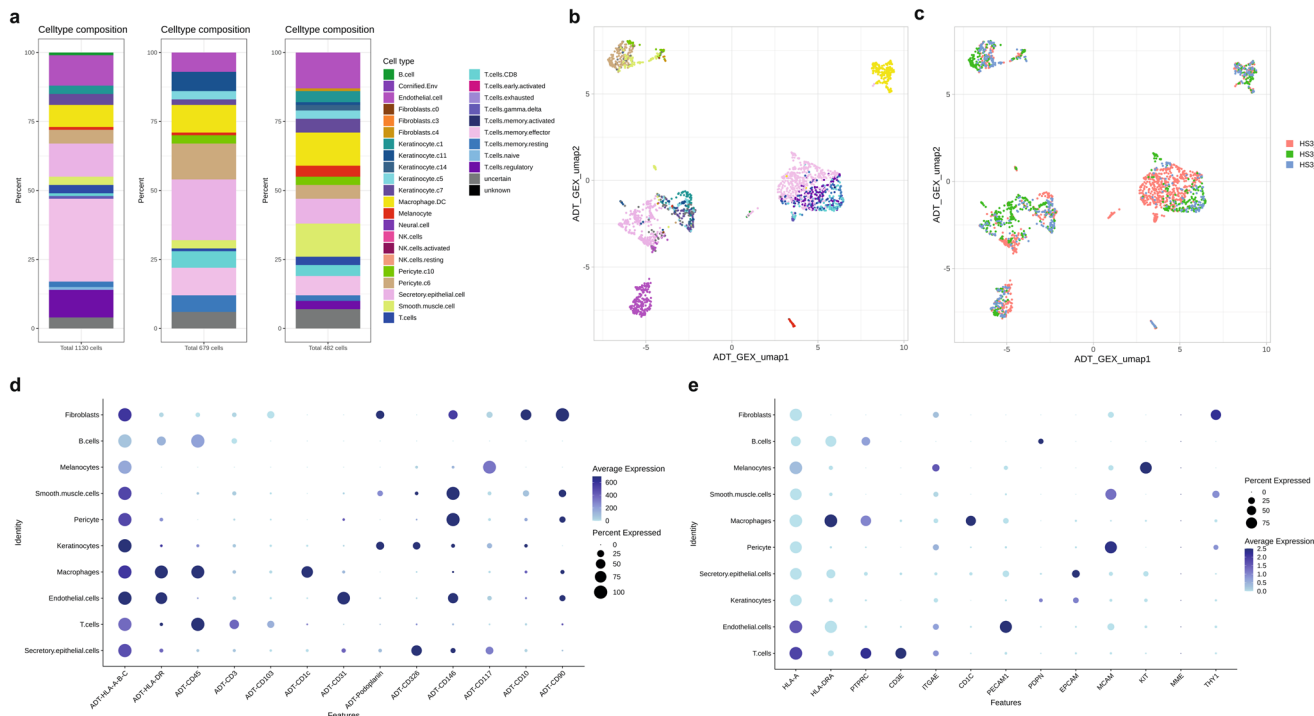


Fig. 3 CITE-seq performance on healthy skin biopsies. **a** Cell type composition bar plots per sample; left, HS3_1; middle, HS3_2; right, HS3_4 (n = 3, biologically independent samples). **b** UMAP plots with cell type, and **c** sample information. **d** Protein and **e** gene lineage marker expression dot plots per major cell type.

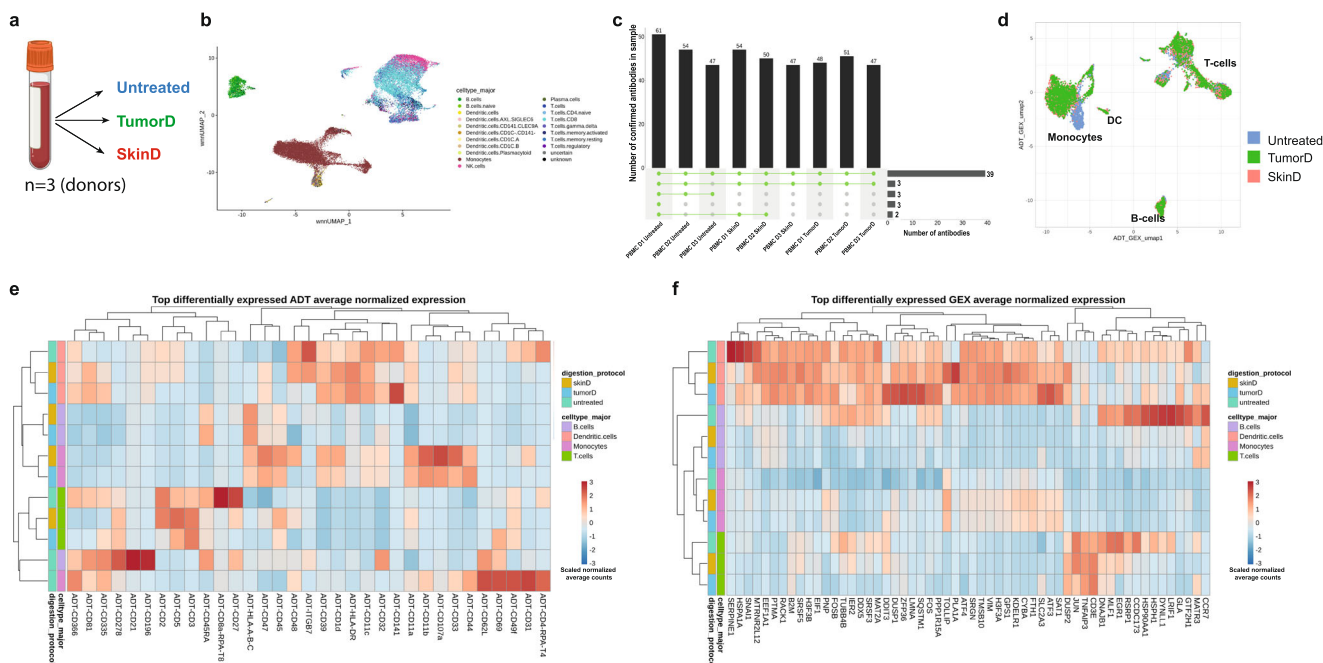


Fig. 4 PBMC model revealing the influence of enzymatic dissociation on gene expression and epitope preservation. **a** Experiment design: healthy PBMC from three healthy donors were aliquoted each in three parts that were left untreated at 37 °C for 1 h or underwent SkinD or TumorD enzymatic incubation at 37 °C for 1 h. **b** Integrated UMAP visualization of all 9 samples (3 donors, 3 treatment conditions) with cell type information. **c** UpSet plot showing confirmed antibody detection across all samples. **d** Representative UMAP plot from Donor 3 showing protocol-specific influence on clustering of different cell types. **e** Differential SPEX and **f** GEX heatmap by sample and major cell type aggregated over three donors showing top differentially expressed features across treatment conditions.

CD196, CD278, CD335, CD336 on B cells as well as ITGB7, CD48, CD31, CD49f, and CD141 on dendritic cells. CD4 and CD336 loss was observed on T cells, dendritic cells, and monocytes, while CD45RA expression was decreased on both B and

T cells (Fig. 4e). Moreover, the DA analysis between treatment conditions by cell type showed upregulation of various surface proteins (Fig. 4e). T cells showed an increase in CD3, CD5, and CD69 under either one or both of the digestion protocols. Cell

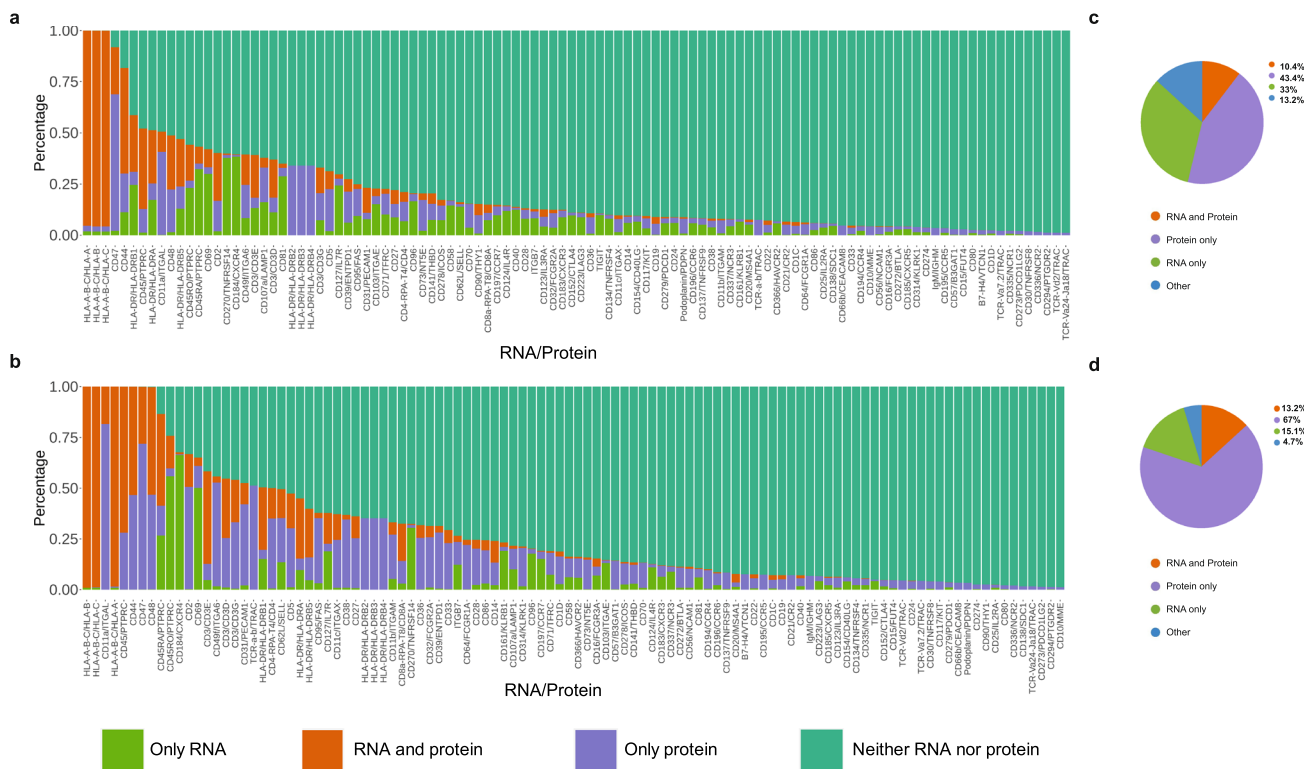


Fig. 5 RNA and protein codetection plots. **a** Feature codetection waterfall plot of solid tissue biopsy cohort including five skin, three primary, and three metastatic melanoma samples ($n = 11$, biologically independent samples) and of **b** liquid biopsy cohort including PBMCs from three healthy donors (no digestion) and three immunotherapy-treated melanoma patients samples ($n = 6$, biologically independent samples). Plots show the percentages of cells with “RNA only”, “RNA and protein”, “protein only”, or “neither RNA nor protein” detection category. **c** Percentage of feature pairs in the solid and **d** liquid cohort that were predominantly assigned to one category based on a threshold of $>50\%$ of cells. Feature pairs where the threshold was not reached for either of the three categories were assigned to “other”.

surface proteins associated with antigen presentation (CD1d, HLA-A-B-C), adhesion and migration (CD11a, CD11b, CD44, CD47), immune regulation (CD39, CD107a) as well as activation (CD11c, CD32, CD33, CD48) were upregulated on monocytes (Fig. 4e). Differential gene expression (DE) between conditions showed upregulation of signatures of innate immune activation (*SLC2A3*, *TOLLIP*, *SATI*, *FOSB*), autophagy (*SQSTM1*), stress response and apoptosis (*DDIT3*, *DUSP1*) as well as down-regulation of genes involved in heat shock response (*HSPA1A*, *HSP90AA1*, *HSPH1*, *DNAJB1*) in enzymatically treated cells compared to untreated (Fig. 4f). Thus, the PBMC model comparing untreated versus enzymatically treated conditions informed on dissociation-sensitive genes and identified not only loss of certain surface epitopes but also increased protein presentation by the enzymatic digestion.

RNA and surface protein codetection. Most differential mRNA expression studies assume the expression of the corresponding protein product, and some studies provide parallel protein stainings from paired formalin-fixed paraffin-embedded (FFPE) slides to demonstrate the translation of mRNA transcript into a protein product. Here, we assessed the codetection of RNA-protein pairs for the CITE-seq antibody panel ($n = 97$) in PBMCs from three healthy donors and three immunotherapy-treated melanoma patient samples (liquid biopsies, $n = 6$, no digestion applied) as well as in five healthy skin, three primary melanoma, and three lymph node metastatic melanoma samples (solid biopsies, $n = 11$, Supplementary Fig. 4, Supplementary Data 2). We analyzed codetection in RNA-protein detection in each cell of the solid and liquid biopsy cohorts by looking at four conditions

(Fig. 5a and b, respectively; RNA-protein pairs sorted by decreasing codetection, Supplementary Data 1, Supplementary Fig. 4; alphabetically listed RNA-protein pairs): (i) detection of RNA only (minimal count of 1 transcript per cell), (ii) detection of protein only (signal above manually defined threshold per cell; Supplementary Fig. 1, Supplementary Data 1), (iii) detection of both RNA and protein, and (iv) absence of both RNA and protein. Next, for both solid and liquid biopsy cohorts, we assessed to which category (“RNA only”, “protein only”, “RNA and protein”) each feature was predominantly assigned ($>50\%$ of cells per category) and calculated the percentages of these categories by cohort (Fig. 5c, d, Supplementary Data 1). The first 15 markers that showed the highest codetection in “RNA and protein” expressions were shared between solid and liquid biopsies (left side of the waterfall plots, Fig. 5a, b, Supplementary Data 1). We found that with 43.4% and 67.0%, most of the markers were predominantly detected only at the protein level, while 33.0% and 15.1% markers were detected at the RNA level only, and to the lowest extent with 10.4% and 13.2% markers were detected on RNA and protein levels in solid and liquid biopsy cohorts, respectively (Fig. 5c, d, Supplementary Data 1). “Protein only” and paired “RNA and protein” were detected more often in liquid biopsies compared to solid biopsies. In both biopsy cohorts, overlap in RNA and protein detection was observed in cell surface markers important for regular cell functions such as major histocompatibility antigens (HLA-A/B/C, HLA-DR), cell lineage markers (CD45, CD3, CD8a), cell surface glycoproteins (CD44) and immunoglobulins (CD48). CD45RO and CD45RA are isoforms used to distinguish between different naive and memory subpopulations, which could not be distinguished at the RNA

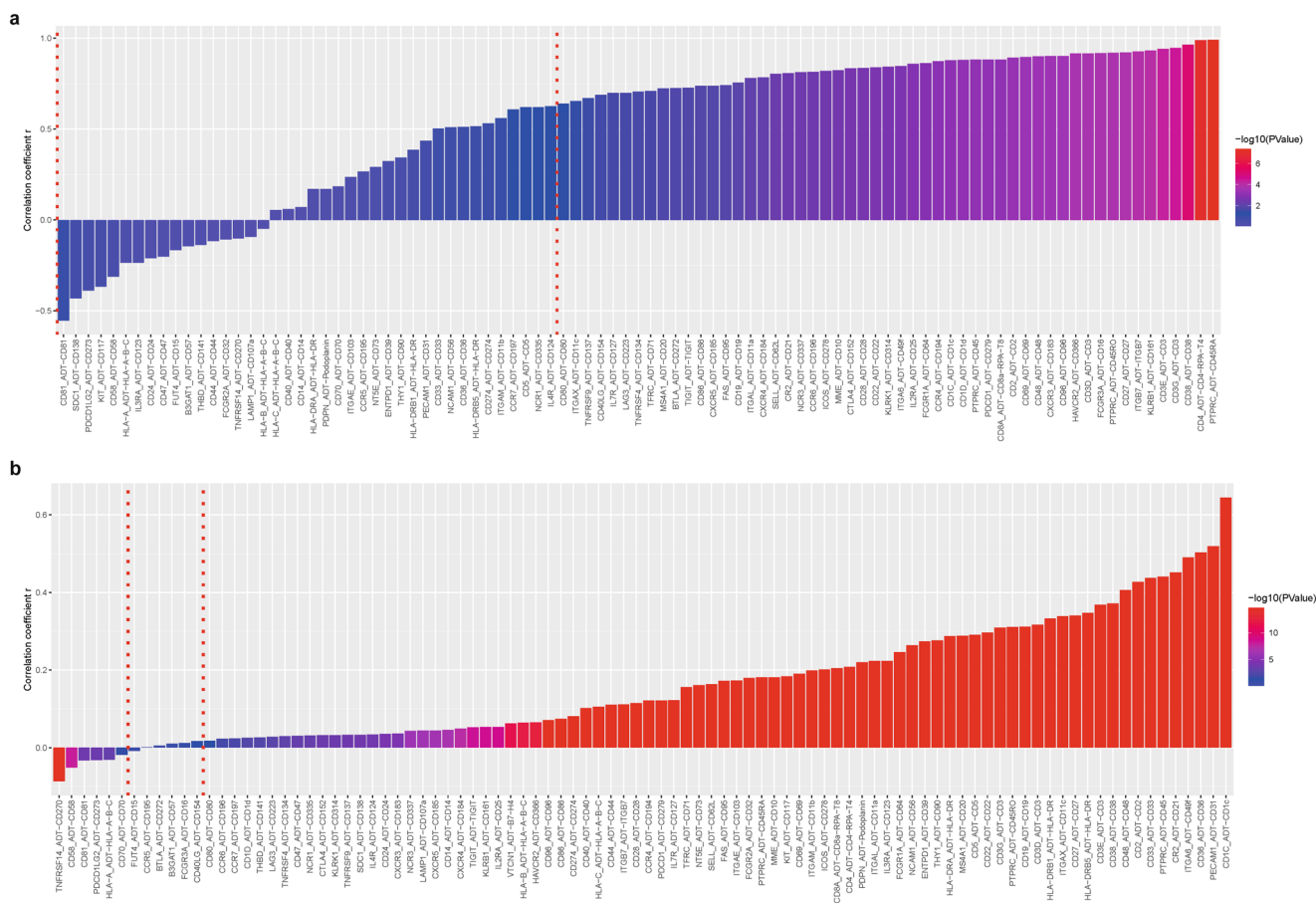


Fig. 6 RNA and protein expression correlation. Pearson correlation coefficients of 93 RNA-protein pairs on aggregated **a** sample level ($n = 11$, biologically independent samples) and **b** cell level ($n = 12,665$ single cells) for the solid tissue biopsy cohort. The Pearson correlation coefficient between each RNA-protein pair is shown along with its related p -value. The significance threshold is set to 0.05 and indicated as a red, dotted vertical line. Pairs were excluded if either one or both members of the pair were not detected.

level (*PTPRC*) without full-length transcript information. Higher CD45RA expression was observed in the liquid biopsy cohort, as expected in peripheral blood with a higher fraction of naïve cells³⁶; whereas, in solid biopsies, CD45RA and CD45RO were detected on the same level indicating that immune cells found in the tissues are more activated and/or differentiated. Markers that showed “RNA only” expression included chemokine receptors and activation markers such as *CXCR3* (CD183), *CXCR4* (CD184), *IL4R* (CD124), *IL7R* (CD137), *CCR7* (CD197), *LAG3* (CD223), and *TIGIT*. Markers predominantly detected at the “protein only” level included among others CD10, CD39, CD36, CD56, B7-H4, CD195, and CD70 with missing GEX information indicating low RNA expression or difficulties in transcript capture. In both liquid and solid biopsy samples, “protein only” expression was detected in cases where no single RNA directly translates into the epitope such as various types of T cell receptor epitopes (TCR- α/β , TCRV $\delta 2$, TCR-V $\alpha 7.2$, TCR-V $\alpha 24$ -Ja18) and CD57 (glycoepitope). In the case of the B cell immunoglobulin M (IgM) isotype, the immunoglobulin heavy constant mu gene (*IGHM*) could not be detected because the gene is located on the 3-prime end of the B cell receptor mRNA molecule³⁷, which was not covered with our 5-prime sequencing approach. The decreased protein detection in the solid cohort, seen as a higher fraction of “RNA only” detection, was attributed to enzymatic digestion-associated loss of epitopes as shown in the PBMC model (Fig. 4e). In summary, the assessment of RNA-protein codetection highlighted similarities and differences in feature detection in liquid and solid biopsy cohorts.

RNA and protein expression correlation. Next, we asked how well RNA and protein expression is correlated and calculated RNA-protein correlation coefficients on a sample level (Fig. 6a, Supplementary Fig. 5a, Supplementary Data 1) and a cell level (Fig. 6b, Supplementary Fig. 5b, Supplementary Data 1) for the solid biopsy (Fig. 6, Supplementary Data 1) and the liquid biopsy cohorts (Supplementary Fig. 5). Only RNA-protein pairs were included for which both features were detected. More RNA-protein pairs with significant positive correlations ($p < 0.05$) were detected in the solid cohort compared to the liquid cohort on a sample (57% vs. 25%) and cell (88% vs. 80%) level reflecting broader feature detection due to a higher cell type diversity. In addition, the higher number of solid samples provided better statistical power compared to the liquid biopsy cohort. Strong differences were observed depending on whether correlations were computed on a sample (population) or single-cell level. While only 57% (53 out of 93 included pairs) of RNA-protein pairs reached positive significance on a sample level in the solid cohort, the correlation of these was high (mean $r = 0.83$). On a single-cell level, 88% (82 out of 93 pairs) of pairs were significantly positively correlated in the solid cohort; however, these displayed an overall lower correlation (mean $r = 0.17$). Only three features (3% of all pairs) showed a correlation coefficient above 0.5 and six pairs (6% of all pairs) were significantly anti-correlated, albeit with coefficients equal to or lower than 0.09, indicating that noise or other small fluctuations can reach significance with a large number of observations. The same was observed for significantly positively correlated pairs in the liquid cohort with

mean $r = 0.92$ on a sample and mean $r = 0.24$ on a cell level. In summary, this analysis provides a list of highly correlated features that could guide antibody or gene panel design in future studies.

RNA and protein abundance analysis of PBMCs. Application of CITE-seq to DA and DE analysis was assessed on a PBMC cohort consisting of three healthy donors and three immunotherapy-treated (IT) melanoma patients (Supplementary Data 2). UMAP visualization showed cell clustering according to distinct cell types, and not according to sample origin (Fig. 7a). IT-treated patients' PBMCs showed higher numbers of detected antibodies, which included CD161, CD137, CD272, and CD335 markers (Fig. 7b, Supplementary Data 1, Supplementary Data 2). We next performed DA/DE analysis between PBMCs from healthy donors and IT-treated melanoma patients on SPEX (Fig. 7c) and GEX (Fig. 7d) levels, respectively. Because the thresholding introduces changes in count distribution, DA was computed on non-thresholded data, however, thresholds were applied in the final results of the visualization of the heatmaps (see methods). DE analysis on GEX level detected strong upregulation of activation, checkpoint, and memory markers such as *CTLA4*, *TIGIT*, and *CD27* on IT-treated T cells, a finding that could not be verified on SPEX level due to above reported poor protein detection of these markers. The absence of CTLA-4 protein expression could be explained by anti-CTLA-4 treatment in case the epitope is already occupied by a therapeutic antibody. However, one of the IT-treated patients received anti-PD-L1 treatment only (Supplementary Data 2). Therefore, to validate the CITE-seq results, the same infusion of PBMCs from three patients that were processed for CITE-seq were selected for an orthogonal experiment. To test variability in antibody clone performance, in addition to the clone included in the CITE-seq panel (BNI3), clone L3D10 was selected based on reports from other studies³⁸. Interestingly, although CTLA-4 is a surface receptor, both instructions from the manufacturer and previous publications on CTLA-4 expression on T-cells³⁹ utilize prior cell stimulation and intracellular staining to detect the protein expression. Therefore, PBMCs were first stimulated for 4 hours with PMA/ionomycin, followed by extracellular and intracellular staining. Cells were gated as lymphocytes/single cells/CD3+/CD4+CD8- and gates for CTLA-4 were adjusted according to the matching isotype control (Supplementary Fig. 6a). CTLA-4 protein expression was only observed when both stimulation and intracellular staining was performed (Supplementary Fig. 6b and c). In addition, compared to the L3D10 clone, lower CTLA-4 detection was observed using the BNI3 clone (Supplementary Fig. 6b and c).

Overall, GEX and SPEX DE analysis provided different results as a consequence of feature limitations in SPEX space and different capture strengths of both types of measured molecules.

Proof-of-concept CITE-seq application on a melanoma cohort for biomarker discovery: differential protein abundance and validation. As a proof-of-concept study of CITE-seq applicability for biomarker discovery, we next performed differential protein abundance analysis on healthy skin melanocytes, primary melanoma, and metastatic melanoma cells and visualized protein and corresponding RNA expression (Fig. 8a, b, Supplementary Data 1). Various markers associated with immune evasion (CD47 and CD274)^{40,41}, invasion and metastasis initiation (CD44, CD49f, CD81, and HLA-DR)^{42,43} were found on metastatic melanoma cells. HLA-DR was detected at low, medium, and high levels in melanocytes, primary melanoma, and metastatic melanoma cells, respectively, although these biopsies contained BRAF and not the NRAS mutation, which is associated with higher HLA-DR expression^{43,44}. CD107a abundance, a marker for

organelles of endosomal-lysosomal lineage and proxy for vesicle secretion⁴⁵, was detected on a larger fraction of healthy melanocytes and reduced on primary and metastatic melanoma cells (Fig. 8a, Supplementary Data 1, Supplementary Data 2). CD117 expression was detected on healthy melanocytes and metastatic melanoma cells but absent on primary melanoma cells. Metastatic melanoma cells showed CD81 expression, a marker associated with metastatic progression⁴⁶ and high expression of CD49f. In addition, protein expression was observed for CD56 (neural cell adhesion molecule 1 (*NCAM1*), involved in cell-cell adhesion), CD71 (transferrin receptor 1)⁴⁷ and CD73 (ecto-5'-nucleotidase responsible for extracellular adenosine production with immunosuppressive function)^{48,49}. CD56 (*NCAM1*) as well as CD274 (*PD-L1*) were only detected in metastatic melanoma cells and primarily on a protein-only level (Fig. 8a, b) consistent with the overall detection pattern for these markers in liquid and solid cohorts (Fig. 5a, b, Supplementary Fig. 4). While CD274 (*PD-L1*) is a known immune checkpoint inhibitor^{50,51}, less is reported on CD56 expression on metastatic melanoma cells⁵². To validate and spatially resolve the expression of CD56, a paired FFPE section of the metastatic melanoma LN was analyzed using Akoya multiplex immunohistochemistry (Fig. 8c). Quantification of CD56/MLANA double-positive melanoma cells, MLANA+ melanoma cells, CD68+ macrophages, and CD8+ T-cells in five regions of interest (ROI) (Fig. 8d, e, Supplementary Data 1) showed that on average 13% of all MLANA+ cells (3,171 cells) were CD56/MLANA double-positive (430 cells).

Discussion

Application of antibody panels for multiplexed protein detection requires careful tailoring of optimal working concentrations, which is typically established by titration experiments. For assays in which antibodies are conjugated to fluorophores (flow cytometry) or isotopes (mass cytometry), compensation steps for each conjugate are performed to separate the true positive signal from the background staining. In addition, the number of antibodies in the panels is limited by the number of unique conjugates due to technological limitations to resolve the signals. Oligo-conjugated antibody barcoding technology such as CITE-seq is based on the unique DNA sequence and due to its immense combinatorial sequence space offers a virtually limitless number of antibodies to be included in the panel. When possible, reducing staining concentrations was proposed to decrease antibody-related sequencing costs⁵³. Titration of a large antibody panel is costly, time-consuming, and not always possible when using clinical samples due to limited tissue material, which is a potential limitation when applying CITE-seq on patient-derived biopsies. In this study, we tested a panel of 97 antibodies without prior titration and achieved good signal-to-noise ratios allowing us to identify positive antibody signals. The background staining was resolved using transcriptome-based gating and plotting cell type-specific protein abundance ridge plots. First, the ridge plots visualize individual cell types, which allows antibody signal detection on rare cell populations that otherwise would have been lost on all-cells protein abundance distribution plots. Further, the background staining was removed by applying dynamic manual thresholds to display only true positive antibody signals. We also demonstrated that these thresholds vary between experiments and cannot be easily transferred from one dataset to another. With wider application of CITE-seq analysis, new approaches are being developed that aim to calculate the background staining from empty droplets as well as by simultaneous staining with non-human species directed antibodies^{54,55}; however, manual examination of each protein marker would still be necessary to correct for possible inaccuracies.

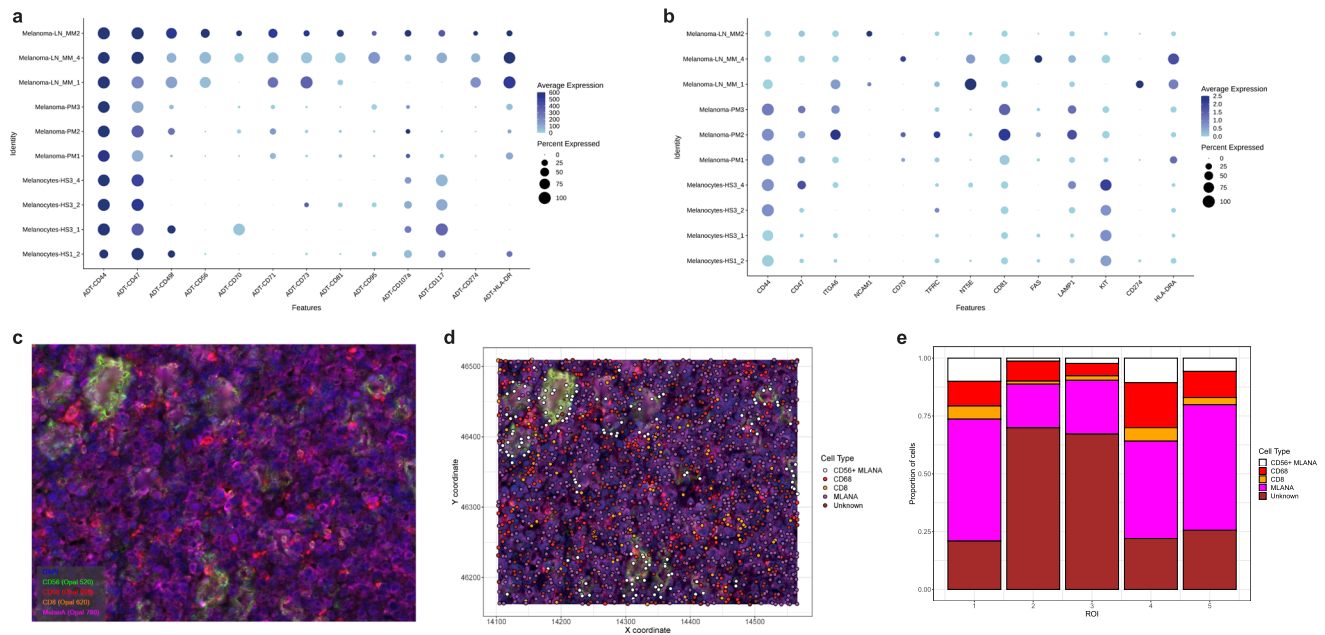


Fig. 8 Proof of concept CITE-seq application for biomarker discovery on a cohort of solid tissues. **a** Protein and **b** gene marker expression dot plot per cell type and sample. Melanocytes from healthy skin samples ($n = 4$, biologically independent samples, dissociated with D/C/T protocol—the skin biopsy from breast reduction was excluded from this analysis because of lack of melanocyte presence), primary melanoma cells from primary melanoma samples ($n = 3$, biologically independent samples), and metastatic melanoma cells from the metastatic lymph node samples ($n = 3$, biologically independent samples). **c** Multiplex immunohistochemistry staining of a matched metastatic melanoma lymph node tissue sample showing CD56 signal on MLANA-positive melanoma cells. **d** Analysis of a representative region of interest (ROI) quantifying single cells positive for: CD8 (T cells), CD68 (macrophages), MLANA (melanoma cells), and DAPI (nuclei staining). **e** Cell type composition bar plots from five analyzed ROIs.

Surface protein detection on solid tissues is more complicated due to the pre-processing into single-cell suspensions using mechanical or enzymatic digestion that introduces bias into cellular composition as well as affects the transcriptome and the surface proteome of the cells⁵⁶. Dissociation of solid tissues using proteolytic enzymes is the most reliable method to obtain single-cell suspensions that capture the heterogeneity of the original tissue and optimization is required to obtain heterogeneous cell populations and preserve cell viability²⁵. Immune cell release can be achieved by shorter treatment times (~15 min), while stromal cells require prolonged enzymatic treatment due to complex cell–cell and cell–matrix connections⁵⁷. Although prolonged dissociation is often necessary to extract most cell types from tissues, it also introduces bias in gene expression by upregulating early stress⁵⁷ and heat shock signatures²³ and can cause artificial loss of surface protein expression²¹. To investigate preservation of surface epitopes on immune cells, we treated healthy PBMCs with skin (SkinD) and tumor dissociation (TumorD) protocols and compared the RNA and protein expression patterns against an untreated control. In agreement with previous reports, enzymatically treated PBMCs showed loss of CD4, CD8, CD21, and CD62L in our CITE-seq analysis in samples treated by both SkinD and TumorD protocols^{20,21}. CD141 signal was reduced in SkinD-treated dendritic cells, consistent with trypsin sensitivity as previously reported⁵⁸. Interestingly, differential gene and surface protein expression analysis showed that lymphocytes were most affected by enzymatic digestion compared to other cell populations with loss of dissociation-sensitive epitopes (such as CD4, CD8, CD21). The monocyte population showed prominent gains in many cell surface markers such as CD107a (degranulation and activation), CD39 (metabolic activation), CD1d (antigen presentation) and CD33 (macrophage differentiation) indicating functional sensitivity to enzymatic exposure. This is in line with previous reports of transcriptomic and proteomic changes of

microglia in enzymatically digested mouse brain tissues compared to mechanical dissociation²² and collagen-induced chemotaxis of human blood monocytes⁵⁹.

Global single-cell transcriptome profiling of primary tissues from various tumors and non-malignant samples revealed intrinsic cell type heterogeneity and greatly improved our understanding of disease mechanisms and drug actions in multiple clinical fields^{60–63}. However, scRNA-seq studies are limited only to the description of the transcriptome, while protein expression is either inferred from gene expression or orthogonally validated from paired FFPE samples⁶². Single-cell analysis pipelines are rapidly advancing into technologies that can integrate multiple profiling modalities, and recent progress in sequencing-based technologies allows simultaneous capture of cell surface protein and gene expression from a single cell^{9,63,64}, thereby linking genes and their protein products and capturing a more realistic functional state. In clinical settings, CITE-seq was applied to monitor responses to therapy and relapse by Ibrutinib (irreversible Bruton Tyrosine Kinase inhibitor) treatment of chronic lymphocytic leukemia (CLL)¹⁴. The analysis revealed that CD3 is well identified by both protein-labeling as well as by gene expression, whereas CD69 or CD19 were mainly detected by the cell surface protein labeling. CITE-seq analysis showed a decrease in CD69 expression (prognostic factor) correlated to clinical progression in CLL at three months post-treatment, and subsequent increase at month 27 correlating with disease progression¹⁴. A study that applied CITE-seq to myeloid populations within human and mouse glioblastoma samples reported RNA–protein codetection for many markers; other features, however, were differentially expressed on protein level only⁶⁵. In this study, we analyzed the detection intersection in 97 RNA–protein pairs, reported predominant detection (more than 50% of cells with at least one observation), and calculated their correlation. We corroborate previous observations of varying degrees of

feature detection, provide calculations on feature correlation and list features in each category that can serve as a lookup table for future investigators. The high amount of features that could only be well detected in protein space highlights the additional value CITE-seq can provide considering clinically relevant readouts. Nevertheless, markers such as *TIGIT* and *LAG3* were better detected at the RNA level showing neither technology alone is superior in feature detection, thus highlighting the added value of combining both omics arms. Furthermore, transcript isoforms are not covered with the short read sequencing data. As noted by Liu et al.²⁹, the possibility that protein levels may be compared to splice isoforms that do not contain the respective protein sequence may distort the RNA-protein correlation. An example of this is the detection of CD45RA and CD45RO splice isoforms of CD45 on immune cells, which could be detected with isoform-specific antibodies. Due to the lack of full-transcript sequence coverage, they could not be detected at the transcript level. Coupling long-read sequencing technologies allowing for isoform detection^{66,67} with surface protein capture would be the next logical step to further gain biological insights. More “protein only” and “RNA and protein” expression was observed in liquid biopsies (three healthy and three immunotherapy-treated PBMCs) or in solid biopsies that were digested with a milder enzymatic blend (metastatic melanoma in LN) indicating digestion-associated loss of cell surface proteins. The absence of positive antibody signals could also be influenced by other factors such as suboptimal antibody clone performance (e.g., CD4 and CD8 clones SK3 and SK8, performed better than clones RPA-T4 and RPA-T8, respectively). An additional limitation of the CITE-seq method might be due to how liquid biopsy is processed prior to single cell droplet generation. Usually, to have minimal impact on the transcriptome of cells derived from patients, samples are processed as quickly as possible. However, for certain surface proteins, such as CTLA-4, it has been described that surface receptors reside on the cell surface only for a short time, and are internalized in the absence of a ligand, which means that at any given time point, there is very low surface expression of CTLA-4⁶⁸. In vitro stimulation is necessary for reliable detection of CTLA-4 expression, which is not feasible if fast sample processing is required. Therefore, surface protein expression of such markers might be overlooked during CITE-seq profiling. Moreover, our analysis compared well with RNA-protein detections in an independent study comprised of glioblastoma-isolated immune cells⁶⁵, which showed higher protein detection for CD36, podoplanin, CD64, CD49f, CD86, CD15, CD56, CD90, CD11a and CD11b, high RNA and protein codetection for CD44 and higher RNA detection for *CXCR4* (CD184). This was consistent with our classifications of predominantly “protein only”, “RNA and protein”, and “RNA only”, indicating that technological and biological effects influencing RNA-protein detection are preserved across studies and experiments, thus supporting a wider application of CITE-seq in basic and translational research. CITE-seq applications on liquid biopsies enabled the identification of cell populations present at low frequencies, such as antigen-presenting cells (CD1c+ and CD141+ cells) or naïve and memory phenotypes by discriminating expression of CD45RA and CD45RO⁶⁹. RNA and protein correlation analysis was calculated on a sample and cell level with more significantly positively correlated pairs found in the solid cohort due to higher cell diversity and sample numbers. Cell-based correlations provided more significantly correlated pairs resulting from a higher number of observations, however, had lower correlation coefficients. The generalized low correlation in single-cell observations reflected most likely differences in transcript detection, cell-to-cell variability, transcriptional noise, and potentially poor antibody binding and is in line with previously reported differences in

correlations on sample (population) or single-cell levels⁷. This indicates that to mitigate these effects and to achieve higher confidence in feature detection, RNA-protein correlations should be investigated at least on a cluster or cell type instead of single-cell level.

Studies that used CITE-seq to analyze solid tissues so far have only focused on subtyping the immune milieu of various organs such as the brain^{15,65}, kidney⁷⁰, breast¹⁶, and skin⁷¹. Besides immune cells, the tumor microenvironment (TME) is composed of various cell types such as endothelial cells, pericytes, fibroblasts, and neural cells. As has been shown, the non-malignant components of the TME can also contribute to tumorigenesis, progression, and metastasis^{72,73}. Therefore, understanding the cell-cell interactions as well as profiling the functional states of various cell types within the environment of the solid tissue samples is as important. Commonly, in single-cell protein profiling techniques such as flow cytometry or CyTOF, antibody panels are pre-selected and prioritized based on previous knowledge, and the combination of established and exploratory markers is technologically limited. For biomarker discovery in CITE-seq, in addition to well-established antibodies, we added a panel of exploratory markers and identified known as well as less described metastatic melanoma markers. After analysis of the panel, we compared the expression of the selected surface markers on healthy melanocytes, as well as primary and metastatic melanoma cells. Among markers identified on metastatic melanoma cells, protein expression did not correlate with the corresponding gene expression in almost 50% of all markers (6 out of 13—CD56, CD70, CD71, CD81, CD107a, and CD274) showing that application of CITE-seq extended antibody panels can reveal unusual surface markers that would otherwise be missed based on mere RNA profiling. Previously, CD56 expression was shown by immunohistochemistry staining in various tumors with neuroendocrine differentiation⁷⁴, including melanoma^{52,75,76}. CD56 is involved in cell-cell and cell-matrix interactions during development and differentiation and was implicated in homophilic and heterophilic interactions⁷⁷. *NCAM1* gene expression did not correlate with CD56 protein abundance in the cell type-specific analysis. CD56 was furthermore predominantly found on “protein only” level in both liquid and solid cohorts and in an independent study⁶⁵. To address the loss of spatial resolution and validate CD56 expression identified by CITE-seq, we performed multiplex immunohistochemistry coupled with advanced multispectral imaging. We profiled CD56 expression in a paired histological section of metastatic melanoma LN and identified clusters of MLANA-positive cells expressing CD56 that were organized in circular shapes. Previous studies have highlighted the possible role of CD56 as a predictive biomarker⁷⁸, druggable target⁷⁹, and its role in NK-cell mediated immunity⁸⁰. In primary melanoma cultures, increased expression of CD56 was identified in cells showing statistically significant higher capability to cross the in vitro blood-brain-barrier model⁸¹ and all cultures established from brain metastases of melanoma showed CD56 expression⁸², suggesting its role in invasion and its potential importance as a biomarker in metastatic melanoma.

Limitations of CITE-seq analysis include the loss of surface epitopes due to enzymatic dissociation, capture of only pre-selected surface targets and loss of spatial resolution. These can be addressed by complementary spatial profiling, which is required to validate the findings and identify the biologically relevant spatial interactions. A limitation of our CITE-seq analysis using manual thresholds could be further optimized by using automated thresholding as recently proposed⁵⁵, while general limitations of antibodies for protein detection in regard to restricted panels, binding specificity and cross-specificity should eventually

be overcome by further developments in single-cell proteomics using mass spectrometry such as reported by SCoPE-MS⁸³.

In this study, we demonstrated that CITE-seq can be applied to clinically relevant solid tissue biopsies and addressed common limitations that arise from enzymatic digestion. We demonstrated an optimization of the enzymatic digestion protocol for a complex solid tissue such as healthy skin and presented a simple method of applying thresholds on cell type-specific ridge plots to utilize large antibody panels for CITE-seq without prior titration steps. Finally, we applied the CITE-seq method for biomarker discovery, identified a potentially druggable surface protein CD56^{84,85} on metastatic melanoma cells, and validated its expression on a matched tissue. CITE-seq represents a powerful profiling modality that by virtue of its quantitative and qualitative dual transcriptome and surface proteome readouts allows global unbiased cell type composition analysis, RNA-protein correlations quantification as well as simultaneous DE and DA analysis, on RNA and protein levels, respectively.

Methods

Human primary tissue and live-cell biobanking. Human healthy skin, primary and melanoma metastasis samples, as well as PBMCs from consenting patients treated with immunotherapy were obtained from the Dermatology Biobank, University Hospital Zurich (BASEC Nr.2017-00494). Experimental and clinical information are summarized in Supplementary Data 2. After collection, tissue sample biopsies and surgical material were live-cell biobanked. Briefly, collected tissue samples were cut into small pieces of a maximum 2 × 2 mm and up to four pieces were placed into one cryovial (Sarstedt, cat. no. 72.380.002) containing 1 mL of freezing medium consisting of 90% FBS (Biowest, cat. no. S006420E01, batch no. S169419181H), 10% DMSO (Sigma, cat. no. 102148154, filter-sterilized using 0.22 µm filter (Steriflip, Milipore, cat. no. SCGP00525)), slow-frozen in pre-chilled CoolCell® Containers (Corning, cat. no. 432001) and stored in -80 °C fridges within 30 min after collection.

PBMCs isolation: Buffy coats from healthy donors were obtained from Blutspende Zurich (Kunden No. 6561, Biobank project). PBMCs were isolated using a protocol described elsewhere (PBMC isolation and cryopreservation⁸⁶) with minor modifications. Briefly, buffy coats diluted with PBS (Ca²⁺/Mg²⁺-free, Gibco, cat. no.10010-015) were layered onto HISTOPAQUE®-1077 (Sigma, 10771-500 mL) for healthy donor 1 and Ficoll-Paque-1084 (Cytiva, cat. no.17-5446-52) for healthy donors 2-3. Gradient centrifugation was performed at 760 g for 20 min with the brakes OFF. After collecting the mononuclear cell layer, cells were washed four times with PBS and spun down at 350 × g for 8 min at room temperature. Cell count and viability were accessed on Luna-FL™ cell counter (Dual Fluorescence Cell counter, Logos Biosystems Inc., cat. no. L1001) using acridine orange propidium iodide (AOPI, Logos Biosystems Inc., cat. no. F23001) live/dead staining. Up to 10⁷ PBMCs were cryopreserved as described above in the live-cell biobanking section.

Processing of cryopreserved solid tissue and enzymatic dissociation. Live-cell biobanked tissue samples were quickly thawed in a water bath set to 37 °C, resuspended in 10 mL of ice-cold RPMI (Sigma, cat. no. R0883) with 0.04% BSA (Sigma-Aldrich, cat. no. A7906) and incubated for 10 min on ice to allow DMSO to diffuse from the tissue. Samples were spun down at 300 × g for 5 min and cut into small pieces with a scalpel. This step was performed to increase contact between the enzymes and the total surface area of the tissues in a small amount of enzyme mixture. Small tissue pieces were placed in the optimal enzymatic mixture and incubated for 30 min to 3 h at 37 °C with continuous rotation on a MACSmix tube rotator (Miltenyi Biotec, cat. no. 130-090-753) and trituration every 15 min using wide-bore pipet tips (Supplementary Note 1).

After incubation with the respective enzymatic mixture, digested tissue was sequentially filtered through 100 µm (Falcon, cat. no. 352360) and 35 µm cell strainers (Falcon, blue capped FACS tubes - cat. no. 352235). For samples with viability below 70% and when cell numbers allowed (>10⁵ cells total), apoptotic and dead cells were removed by immunomagnetic cell separation using the Annexin Dead Cell Removal Kit (StemCell Technologies, cat. no. 17899) and EasySep™ Magnet (StemCell Technologies, cat. no. 18000). If the cell pellet appeared red, red blood cell (RBC) lysis was performed following the manufacturer's instructions (Roche, cat. no. 11814389001). Briefly, the cell pellet was resuspended in 500 µL of RBC lysis buffer and incubated at room temperature for 3 min. Next, cells were washed with a resuspension buffer (PBS with 0.04% BSA), spun down and resuspended in a resuspension buffer. If the cell pellet appeared still red, RBC lysis was repeated. Cell number and viability were assessed on a Cellometer K2 Image Cytometer (Nexcelom Bioscience, cat. no. Cellometer K2) using ViaStain AOPI Staining Solution (Nexcelom Bioscience, cat. no. CS2-0106-5mL) and PD100 cell counting slides (Nexcelom Bioscience, cat. no. CHT4-PD100-003).

Optimization of healthy skin dissociation - experiment outline: we selected slow-frozen healthy skin biopsies of sufficient size that allowed for testing of three dissociation protocols on the same sample. In short, tissues were thawed and DMSO washed out. Biopsies were cut into three pieces and weighed on a precision scale. Each piece underwent a separate dissociation protocol as described in Supplementary Note 1 and Fig. 2a. Dead cell removal was performed whenever necessary and possible, and suspensions were stained with hashing antibodies whenever possible. Finally, samples were pooled and stained with a CITE-seq antibody cocktail.

Enzymatic digestion effect on PBMCs - experiment outline: three vials of PBMCs from healthy donors were thawed, counted, and split into three equal parts. Aliquot 1 was kept in PBS with 1% BSA and incubated for one hour at 37 °C (untreated), aliquot 2 was treated with SkinD for one hour at 37 °C while aliquot 3 was treated with TumorD protocol for one hour at 37 °C. Processing for antibody staining was performed as described below in the "Antibody staining" section. Donors 2 and 3 were processed on a separate day from Donor 1.

Oligo-labeled antibody panels and staining. Oligo-labeled antibody panels were ordered from the BioLegend TotalSeq-C product line compatible with 10x Genomics 5P V(D)J immune profiling kits and concentrations were chosen according to manufacturer instructions for each antibody. The panel consisting of 97 antibodies (Supplementary Data 2) was pooled in a labeling buffer (PBS with 1% BSA). Antibody staining was performed similarly to the demonstrated 10x Genomics protocol Cell Surface Protein Labeling for Single Cell RNA Sequencing Protocols (https://assets.ctfassets.net/an68im79xiti/6p0emleLO8bsxinEbKgcFf/275a5752f4e4347f75a1f649bd824463/CG000149_DemonstratedProtocol_CellSurfaceProteinLabeling_RevB.pdf), omitting dextran sulfate staining as recommended by BioLegend (<https://www.biolegend.com/en-us/protocols/totalseq-b-or-c-with-10x-feature-barcoding-technology>). After Fc receptor blocking, samples were first stained with hashing antibodies (BioLegend, see Supplementary Data 2) in ~100 µL volume for 30 min on ice. Afterwards, cells were washed three times with labeling buffer (PBS with 1% BSA) before incubation in a final pool of oligo-barcoded BioLegend TotalSeq-C antibodies in 100 µL volume for another 30 min. Finally, cells were washed three times and resuspended in a resuspension buffer (PBS with 0.04% BSA). Cell number and viability were profiled as described in the above section (Processing of cryopreserved tissue and enzymatic digestion) and optimal cell concentrations were set according to 10x Genomics protocols (700-1200 cells/µL).

Single-cell droplet generation and processing. Stained cell suspensions were loaded and processed using the 10x Genomics Chromium platform with the 5P V(D)J immune profiling kit on 10x Genomics Chromium Single Cell Controller (10x Genomics, PN-120263). Hashed samples were super-loaded with 9,000 to 20,000 cells per lane (see Supplementary Data 2). GEX and SPEX libraries were amplified and sequenced on the Illumina NextSeq 500 or NovaSeq 6000 platform at recommended sequencing depth (20,000-50,000 reads/cell for GEX libraries, and >7000 reads/cell for SPEX libraries).

Multiplex immunohistochemistry staining and multispectral imaging (Akoya Vector Polaris). Histological slides were stained with primary (at 1:100 concentration) and opal antibody pairs on a Leica Bond RXm following manufacturer instructions supplied from Akoya (Table 1). Prior to scanning, the bottom of the slides was cleaned with 70% ethanol. Multispectral slide scanning was performed on the Vectra Polaris at ×40 magnification. Regions of interest (ROIs) were annotated with PhenoChart and spectral unmixing and cell segmentation were performed with inForm 2.4.9. Raw QTIFF images were exported as TIFF files.

PBMCs stimulation and flow cytometry. Live-frozen PBMCs from three patients receiving immunotherapy (Supplementary Data 2) were thawed by drop-wise resuspension in media (RPMI 1640 (Sigma-Aldrich, cat. no. R0883) supplemented with 5 nM L-glutamine (Gibco, Thermo Scientific, cat. no. 25030-024), 1 mM sodium pyruvate (Sigma-Aldrich, cat. no. S8636), 10% heat-inactivated fetal bovine serum (Biowest, cat. no. S181H) and 1% Pen-Strep (Gibco, Thermo Scientific, cat. no. 15140-122)) and rested on ice for 10 min. Cells were counted with AOPI (Logos Biosystems, cat. no. F23002) and viability was found to be over 90%. Cells were seeded in round bottom 96 well plates in media at 0.5 Mio cells/well and incubated with or without cell stimulation/protein export inhibitor cocktail (eBioscience, cat. no. 00-4975-93) for 4 h at 37 °C and 5% CO₂. Next, cells were washed with PBS with 2% FCS, and stained with anti-human CD3 (APC, BioLegend, cat. no. 317318, stock concentration 50 µg/ml, dilution 1:50, final concentration 1 µg/ml), anti-human CD4 (PerCP, BioLegend, cat. no. 317432, stock concentration 200 µg/ml, dilution 1:50, final concentration 5 µg/ml), anti-human CD8 (PE-Cy7, BioLegend, cat. no. 344712, stock concentration 100 µg/ml, dilution 1:100, final concentration 1 µg/ml), anti-human CTLA-4 (Biolegend, clone BNI3, APC/Fire 750, cat. no. 369627, stock concentration 25 µg/ml, dilution 1:20, final concentration 1.25 µg/ml) or Mouse IgG2a, κ Isotype control (APC/Fire 750, Biolegend, cat. no. 400283), anti-human CTLA-4 (PE, Biolegend, clone L3D10, cat. no. 349905, stock concentration 200 µg/ml, dilution 1:20, final concentration 10 µg/ml) or PE Mouse IgG1, κ Isotype control (PE, Biolegend, cat. no. 400113). Samples were analyzed on

Table 1 Antibodies and Opal reagents for mIHC.

Primary	Opal
MelanA (NBPI-30151, NovusBio)	Opal 780 (NEL871001KT)
CD68 (ab213363, Abcam)	Opal 690 (NEL871001KT)
CD56 (ab220360, Abcam)	Opal 520 (NEL871001KT)
CD8 (ab4055, Abcam)	Opal 620 (NEL871001KT)

a LSRFortessa™ Cell Analyzer (BD Biosciences). Isotype controls were used at the same concentrations as the matched antibodies. Compensation was performed with UltraComp eBeads™ (Invitrogen, cat. no. 01-2222-42). The flow cytometry data were analyzed in FlowJo™ v10.0.8.

Data analysis. All CITE-seq data analysis was embedded in a workflow for single-cell GEX and SPEX analysis as well as hashing analysis³⁵. In the following, we describe in brief the main steps of the sequencing data analysis.

Preprocessing and hashing analysis: Raw reads were mapped to the GRCh38 reference genome using 10x Genomics Cell Ranger 3.1.0 to infer read counts per gene per cell. GEX and SPEX libraries were processed independently, and the Cell Ranger “-force-cells” option was set to initially retrieve the number of loaded cells per experiment. For samples sequenced on the NovaSeq platform, index-hopping removal using a method developed by Griffiths et al. was performed⁸⁷. Next, hashed samples were demultiplexed by first applying CITE-seq count in order to count hashtags per cell and subsequent normalization and hashtag assignment using Seurat⁸⁸.

GEX analysis: After hashtag-based demultiplexing of samples, GEX data of each sample was analyzed using the scAmp workflow⁸⁹. In brief, UMI counts were quality controlled and cells and genes filtered to remove known contaminants: cells where over 50% of the reads mapped to mitochondrial genes and cells with fewer than 400 different expressed genes were removed, as well as non protein-coding genes and genes that were expressed in <20 cells. The 400 DEG was chosen as the default threshold and is in place to remove cells with low gene diversity, as those “cells” have an increased risk of being artifacts. Subsequently, counts were normalized and corrected for cell-cycle effects and library size using scran⁹⁰. Similar cells were grouped based on unsupervised clustering using Phenograph⁹¹ and an automated cell type classification was performed independently for each cell⁹². Cell type annotation was performed using lists of cell type defining highly expressed genes from previous publications as follows (Supplementary Data 2): for lymph node melanoma metastasis and healthy skin samples, gene sets from Tirosh et al.⁹³ and Tabib et al.⁹⁴ were used, respectively. For primary cutaneous melanoma samples, gene sets from Tirosh et al.⁹³ were enriched with keratinocyte gene sets from Tabib et al.⁹⁴. Healthy PBMC data was annotated using gene lists from Zhang et al.⁹⁵ and Newman et al.⁹⁶ and enriched with dendritic cell gene sets from Villani et al.⁹⁷. The label “unknown” was assigned if for a particular cell none of the compared cell type marker lists showed sufficient similarity to the gene expression profile of the cell. In contrast, a cell was labeled “uncertain”, if the expression profile of the cell is sufficiently similar to more than one marker list (i.e., the *p*-values returned by the Mann-Whitney *U*-Test performed by the scROSHI cell typing method⁹² are close to the most similar cell type and the second most similar cell type).

SPEX processing: preprocessing, thresholding and visualization: Initial SPEX counts per cell were determined using 10x Genomics Cell Ranger 3.1.0 (Hohhm/CITE-seq-Count: 1.4.2 - <https://zenodo.org/record/2590196#.YWcNFmQza2x>). Raw counts were log-transformed and visualized in a cell type-specific expression ridge plot to allow manual threshold definition. Similar to, e.g., FACS experiments, the observed raw counts contain background noise, which can be removed using manually selected thresholds (for a manually set threshold overview, refer to Supplementary Data 1). Based on the thresholds, only cells with a SPEX count exceeding the threshold were determined as positive for the respective antibody (Supplementary Data 2). GEX and SPEX counts were combined to calculate a UMAP embedding that displays both GEX and SPEX-based effects. In addition, we performed a clustering based on combined GEX and SPEX counts using BREM-SC⁹⁸. To inform on cell type-dependent SPEX counts, ridge plot visualization was performed on a per cell type level.

Cohort integration: Samples originally part of the same hashing experiment are assumed to show no sequencing or antibody staining batch effects, as they have been pooled together before processing. Thus, when grouping samples from the same hashing experiment, no additional batch correction was performed. Instead, after grouping the individual samples, the original GEX and SPEX counts from the single samples are visualized on the combined UMAP.

When integrating samples across hashing experiments the top 3000 genes observed as variables in most samples were used as anchors; additional batch correction was applied using the Seurat CCA method⁹⁹ and multimodal integration was performed using the Seurat WNN method⁸⁸.

Differential gene expression and protein abundance analysis: To compare the effect of different treatment protocols on PBMC samples (untreated versus SkinD versus TumorD) a differential gene expression and surface protein abundance analysis was performed on the integrated samples. To avoid any bias due to sample

composition differences across the compared groups, the DE analysis was performed by major cell types as follows: Per cell type subset, expression levels were compared using Seurat FindMarkers with a significance level of $FDR < 0.0001$ and $|\logFC| > 3$ for GEX data and $FDR < 0.0001$ and $|\logFC| > 0.5$ for SPEX. Specifically for SPEX data, the differential expression analysis is performed on un-thresholded counts, as existing tools are not built to be aware of thresholding; the final results are visualized however throughout the manuscript on the thresholded counts (Supplementary Data 1) to reduce the noise and highlight expression differences (Supplementary Fig. 1c and d), see also Grob et al.³⁵.

GEX and SPEX codetection analysis: To compare the detection of GEX and SPEX counts, the average of the fraction of cells was calculated when (i) positive for both RNA and associated surface protein, (ii) negative for both RNA and surface protein, (iii) positive for RNA but negative for surface protein, and (iv) negative for RNA but positive for surface protein. A cell was identified as “positive” for a RNA if the UMI count was greater than zero. A cell was identified as “positive” for a surface protein if the UMI count was greater than the predefined threshold (see Supplementary Data 1). The comparison was performed across six liquid biopsy samples (three healthy and three immunotherapy-treated PBMCs) and across all solid biopsy samples (five healthy skin samples, three primary melanomas and three metastatic melanomas from the lymph node). For the three categories (“RNA only”, “protein only” and “RNA and protein”), feature counts were converted into percentages. Features belonged predominantly to one category, if >50% of cells were assigned to one category. In case the threshold was not reached for a single category, the feature was labeled as “other” (Supplementary Data 1).

RNA-protein correlation analysis: We calculated RNA-protein correlation coefficients on aggregated sample and cell levels: *Aggregated sample level:* In brief, counts were aggregated by samples and a matrix of counts for all RNAs and proteins (antibody-derived oligos) was created for each sample. RNAs or proteins without counts were excluded from the analysis. Next, RNA-protein pairs were filtered, keeping only the relevant ones as shown in the RNA-protein match lookup table (Supplementary Data 2). The Pearson correlation was computed on the matrices, obtaining both the correlation coefficient and the significance level for every possible pairing. *Cell level:* the expression of RNA and protein abundance was extracted from each cell, the Pearson correlation was computed using each cell as a separate observation and followed by multiple testing correction using the Bonferroni method on the *p*-values.

The results were visualized in the form of barplots for all detected RNA-protein pairs of the sorted correlation coefficients, with color coding for the *p*-values: gray-to-blue gradient representing *p*-values below threshold; blue-to-red gradient representing *p*-values above thresholds; the red dotted lines show the threshold ($p < 0.05$) for both the positive and anti-correlated pairs. Pairs were positively correlated if the expression of both RNA and protein in the available observations showed a tendency to change in the same direction, i.e. a higher observed expression of one feature in the pair corresponded to a higher observed expression of the other feature. Anti-correlation was assigned to RNA-protein pairs if the change in expression of one feature was opposite to the changes observed in the other feature, i.e., a higher observed expression in either the RNA or the protein corresponding to a lower observed expression of the other feature.

Single-cell spatial image analysis: Spatial expression analysis using Giotto (version 2.0.0.957) was performed as follows. Cell segmentation data output from inForm image processing was imported via the createGiottoObject function. The giotto object was further filtered using filterGiotto with the following argument thresholds: expression_threshold = 1, feat_det_in_min_cells = 3, min_det_feats_per_cell = 1. Normalization was applied using normalizeGiotto with the default scale factor of 6000, log_norm = FALSE, library_size_norm = FALSE, scale_feats = FALSE, and scale_cells = TRUE. Dimensionality reduction was performed with runPCA then runUMAP. Clustering was performed with doLeidenCluster with a resolution of 0.2 and 100 iterations. Clusters with similar expression patterns were merged and a final number of 5 clusters was chosen. Clusters were named CD56+/MLANA+, CD68, CD8, MLANA, and Unknown based on the average expression of the markers. Cell identities were mapped back to the original image using ggplot2.

Statistics and reproducibility. In the following, we provide a summary of the statistical analysis performed throughout this manuscript. Further details are available as part of the methods.

GEX analysis included counts normalization, library size, and cell-cycle effect correction using scran⁹⁰. Cell type assignment was based on the *p*-values returned by the Mann-Whitney *U*-Test performed by the scROSHI cell typing method⁹². Combined GEX and SPEX counts clustering was performed using BREM-SC⁹⁸. Cohort integration included additional batch correction using the Seurat CCA method⁹⁹ and multimodal integration using the Seurat WNN method⁸⁸. Differential expression analysis was performed using Seurat FindMarkers with a significance level of $FDR < 0.0001$ and $|\logFC| > 3$ for GEX data and $FDR < 0.0001$ and $|\logFC| > 0.5$ for SPEX. Specifically for SPEX data, the differential expression analysis is performed on un-thresholded counts, as existing tools are not built to be aware of thresholding. RNA-protein correlation was performed by calculating the Pearson correlation on the count matrices, obtaining both the correlation coefficient and the significance level for every possible pairing. The significance threshold was set to $p < 0.05$.

Patients' samples were used for healthy skin, primary melanoma, metastatic melanoma, healthy and immunotherapy-treated PBMCs to reach statistical power in our investigations. The explorative nature of the study did not permit estimating effect sizes prior to the analysis. Healthy skin samples were pre-processed, divided into equal parts, and randomly assigned to one of the digestion protocols. All the CITE-seq samples passed the QC of the 10x Cell Ranger data analysis pipeline. In flow cytometry experiment, $n = 3$ represents biologically independent PBMCs samples.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The scRNA-seq data generated in this study has been deposited to the European Genome-phenome Archive (EGA) database under accession code EGAS00001005849. Source data is available from Supplementary Data 1.

Code availability

The scAmp workflow⁸⁹ for GEX analysis is available on github: https://github.com/ETH-NEXUS/scAmp_single_cell_RNA, as well as the gExcite workflow³⁵ for GEX, SPEX, and hashing analysis: https://github.com/ETH-NEXUS/gExcite_pipeline. The code used for the downstream analysis is available on GitLab: https://gitlab.ethz.ch/nexuscbu/cite-seq_method_paper. The custom scripts used for the paper are on Zenodo¹⁰⁰.

Received: 25 July 2022; Accepted: 26 July 2023;

Published online: 10 August 2023

References

- Hedlund, E. & Deng, Q. Single-cell RNA sequencing: technical advancements and biological applications. *Mol. Asp. Med.* **59**, 36–46 (2018).
- Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
- Slyper, M. et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792–802 (2020).
- Narayanan, A. et al. Nuclei isolation from fresh frozen brain tumors for single-nucleus RNA-seq and ATAC-seq. *J. Vis. Exp.* <https://doi.org/10.3791/61542> (2020).
- Krishnaswami, S. R. et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499–524 (2016).
- Thrupp, N. et al. Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. *Cell Rep.* **32**, 108189 (2020).
- Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232 (2012).
- Payne, S. H. The utility of protein and mRNA correlation. *Trends Biochem. Sci.* **40**, 1–3 (2015).
- Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
- Stoeckius, M. et al. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
- Laskowski, T. J., Hazen, A. L., Collazo, R. S. & Haviland, D. Rigor and reproducibility of cytometry practices for immuno-oncology: a multifaceted challenge. *Cytom. A* **97**, 116–125 (2020).
- Vafadarnejad, E. et al. Dynamics of cardiac neutrophil diversity in murine myocardial infarction. *Circ. Res.* **127**, e232–e249 (2020).
- Granja, J. M. et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat. Biotechnol.* **37**, 1458–1465 (2019).
- Cadot, S. et al. Longitudinal CITE-Seq profiling of chronic lymphocytic leukemia during ibrutinib treatment: evolution of leukemic and immune cells at relapse. *Biomark. Res.* **8**, 72 (2020).
- Golomb, S. M. et al. Multi-modal single-cell analysis reveals brain immune landscape plasticity during aging and gut microbiota dysbiosis. *Cell Rep.* **33**, 108438 (2020).
- Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* **53**, 1334–1347 (2021).
- van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
- Vieira Braga, F. A. & Miragaia, R. J. Tissue handling and dissociation for single-cell RNA-seq. *Methods Mol. Biol.* **1979**, 9–21 (2019).
- Adam, M., Potter, A. S. & Potter, S. S. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development* **144**, 3625–3632 (2017).
- Autengruber, A., Gereke, M., Hansen, G., Hennig, C. & Bruder, D. Impact of enzymatic tissue disintegration on the level of surface molecule expression and immune cell function. *Eur. J. Microbiol. Immunol.* **2**, 112–120 (2012).
- Abuzakouk, M., Feighery, C. & O'Farrelly, C. Collagenase and dispase enzymes disrupt lymphocyte surface molecules. *J. Immunol. Methods* **194**, 211–216 (1996).
- Mattei, D. et al. Enzymatic dissociation induces transcriptional and proteotype bias in brain cell populations. *Int. J. Mol. Sci.* **21**, 7944 (2020).
- O'Flanagan, C. H. et al. Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. *Genome Biol.* **20**, 210 (2019).
- Denisenko, E. et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* **21**, 130 (2020).
- Reichard, A. et al. Best practices for preparing a single cell suspension from solid tissues for flow cytometry. *Cytom. A* **95**, 219–226 (2019).
- Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E. M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* **25**, 117–124 (2007).
- Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **583**, 3966–3973 (2009).
- Darmanis, S. et al. Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell Rep.* **14**, 380–389 (2016).
- Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550 (2016).
- Fortelny, N., Overall, C. M., Pavlidis, P. & Freue, G. V. C. Can we predict protein from mRNA levels? *Nature* **547**, E19–E20 (2017).
- Gubin, M. M. et al. High-dimensional analysis delineates myeloid and lymphoid compartment remodeling during successful immune-checkpoint cancer therapy. *Cell* **175**, 1443 (2018).
- Goswami, S. et al. Immune profiling of human tumors identifies CD73 as a combinatorial target in glioblastoma. *Nat. Med.* **26**, 39–46 (2020).
- Jian, L., Cao, Y. & Zou, Y. Dermal-epidermal separation by enzyme. *Methods Mol. Biol.* **2109**, 27–30 (2020).
- Restivo, G. et al. Live slow-frozen human tumor tissues viable for 2D, 3D, ex vivo cultures and single-cell RNAseq. *Commun. Biol.* **5**, 1144 (2022).
- Grob, L. et al. gExcite - A start-to-end framework for single-cell gene expression, hashing, and antibody analysis. *Bioinformatics* **39**, btad329 (2023).
- Wong, M. T. et al. A high-dimensional analysis of human T cell diversity reveals tissue-specific trafficking and cytokine signatures. *Immunity* **45**, 442–456 (2016).
- Menzel, U. et al. Comprehensive evaluation and optimization of amplicon library preparation methods for high-throughput antibody sequencing. *PLoS ONE* **9**, e96727 (2014).
- Garcia-Perez, J. E. et al. CTLA4 message reflects pathway disruption in monogenic disorders and under therapeutic blockade. *Front. Immunol.* **10**, 998 (2019).
- Zhang, Y. et al. Hijacking antibody-induced CTLA-4 lysosomal degradation for safer and more effective cancer immunotherapy. *Cell Res.* **29**, 609–627 (2019).
- Ngo, M. et al. Antibody therapy targeting CD47 and CD271 effectively suppresses melanoma metastasis in patient-derived xenografts. *Cell Rep.* **16**, 1701–1716 (2016).
- Massi, D. et al. PD-L1 marks a subset of melanomas with a shorter overall survival and distinct genetic and morphological characteristics. *Ann. Oncol.* **25**, 2433–2442 (2014).
- Lee, N. et al. Melanoma stem cells and metastasis: mimicking hematopoietic cell trafficking? *Lab. Invest.* **94**, 13–30 (2014).
- Costantini, F. & Barbieri, G. The HLA-DR mediated signalling increases the migration and invasion of melanoma cells, the expression and lipid raft recruitment of adhesion receptors, PD-L1 and signal transduction proteins. *Cell. Signal.* **36**, 189–203 (2017).
- Johnson, D. B. et al. Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. *Nat. Commun.* **7**, 10582 (2016).
- Kushimoto, T. et al. A model for melanosome biogenesis based on the purification and analysis of early melanosomes. *Proc. Natl Acad. Sci. USA* **98**, 10698–10703 (2001).
- Hong, I.-K. et al. The tetraspanin CD81 protein increases melanoma cell motility by up-regulating metalloproteinase MT1-MMP expression through the pro-oncogenic Akt-dependent Sp1 activation signaling pathways. *J. Biol. Chem.* **289**, 15691–15704 (2014).
- Shen, Y. et al. Transferrin receptor 1 in cancer: a new sight for cancer therapy. *Am. J. Cancer Res.* **8**, 916–931 (2018).
- Ghalamfarsa, G. et al. CD73 as a potential opportunity for cancer immunotherapy. *Expert Opin. Ther. Targets* **23**, 127–142 (2019).

49. Antonioli, L., Yegutkin, G. G., Pacher, P., Blandizzi, C. & Haskó, G. Anti-CD73 in cancer immunotherapy: awakening new opportunities. *Trends Cancer* **2**, 95–109 (2016).
50. Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and its ligands in tolerance and immunity. *Annu. Rev. Immunol.* **26**, 677–704 (2008).
51. Alsab, H. O. et al. PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome. *Front. Pharmacol.* **8**, 561 (2017).
52. Juhlin, C. C., Zedenius, J. & Haglund, F. Metastatic malignant melanoma with neuroendocrine differentiation: a case report and review of the literature. *J. Med. Case Rep.* **14**, 44 (2020).
53. Buus, T. B. et al. Improving oligo-conjugated antibody signal in multimodal single-cell analysis. *eLife* **10**, e61973 (2021).
54. Gayoso, A. et al. Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 (2021).
55. Mulé, M. P., Martins, A. J. & Tsang, J. S. Normalizing and denoising protein expression data from droplet-based single cell profiling. *Nat. Commun.* **13**, 2099 (2022).
56. Hosseini, V. et al. A mechanical non-enzymatic method for isolation of mouse embryonic fibroblasts. *Mol. Biol. Rep.* **47**, 8881–8890 (2020).
57. Waise, S. et al. An optimised tissue disaggregation and data processing pipeline for characterising fibroblast phenotypes using single-cell RNA sequencing. *Sci. Rep.* **9**, 9580 (2019).
58. Botting, R. A. et al. Phenotypic and functional consequences of different isolation protocols on skin mononuclear phagocytes. *J. Leukoc. Biol.* **101**, 1393–1403 (2017).
59. Postlethwaite, A. E. & Kang, A. H. Collagen-and collagen peptide-induced chemotaxis of human blood monocytes. *J. Exp. Med.* **143**, 1299–1307 (1976).
60. Ramelyte, E. et al. Oncolytic virotherapy-mediated anti-tumor response: a single-cell perspective. *Cancer Cell* **39**, 394–406.e4 (2021).
61. Kalluri, R. et al. Single-Cell Analyses in the Multi-omics Era. *Cancer Cell* **38**, 9–10 (2020).
62. Durante, M. A. et al. Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat. Commun.* **11**, 496 (2020).
63. Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).
64. Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).
65. Pombo Antunes, A. R. et al. Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. *Nat. Neurosci.* **24**, 595–610 (2021).
66. Al'Khafaji, A. M. et al. High-throughput RNA isoform sequencing using programmed cDNA concatenation. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-023-01815-7> (2023).
67. Dondi, A. et al. Detection of isoforms and genomic alterations by high-throughput full-length single-cell RNA sequencing for personalized oncology. Preprint at *BioRxiv* <https://doi.org/10.1101/2022.12.12.520051> (2022).
68. Khailaie, S. et al. Characterization of CTLA4 trafficking and implications for its function. *Biophys. J.* **115**, 1330–1343 (2018).
69. Mair, F. et al. A targeted multi-omic analysis approach measures protein expression and low-abundance transcripts on the single-cell level. *Cell Rep.* **31**, 107499 (2020).
70. Zhao, Y., Panzer, U., Bonn, S. & Krebs, C. F. Single-cell biology to decode the immune cellular composition of kidney inflammation. *Cell Tissue Res.* **385**, 435–443 (2021).
71. Liu, Y. et al. Classification of human chronic inflammatory skin disease based on single-cell immune profiling. *Sci. Immunol.* **7**, eabl9165 (2022).
72. Hinshaw, D. C. & Shevde, L. A. The tumor microenvironment innately modulates cancer progression. *Cancer Res.* **79**, 4557–4566 (2019).
73. Guo, T., Li, W. & Cai, X. Applications of single-cell omics to dissect tumor microenvironment. *Front. Genet.* **11**, 548719 (2020).
74. Weissferdt, A. & Moran, C. A. Neuroendocrine differentiation in thymic carcinomas: a diagnostic pitfall: an immunohistochemical analysis of 27 cases. *Am. J. Clin. Pathol.* **145**, 393–400 (2016).
75. Katerji, H., Childs, J. M., Bratton, L. E., Peyre, C. G. & Huber, A. R. Primary esophageal melanoma with aberrant CD56 expression: a potential diagnostic pitfall. *Case Rep. Pathol.* **2017**, 9052637 (2017).
76. Ronen, S., Czaja, R. C., Ronen, N., Pantazis, C. G. & Iczkowski, K. A. Small cell variant of metastatic melanoma: a mimicker of lymphoblastic leukemia/lymphoma. *Dermatopathology* **6**, 231–236 (2019).
77. Van Acker, H. H., Capsomidis, A., Smits, E. L. & Van Tendeloo, V. F. CD56 in the immune system: more than a marker for cytotoxicity? *Front. Immunol.* **8**, 892 (2017).
78. Cottini, F. et al. Redefining CD56 as a biomarker and therapeutic target in Multiple Myeloma. *Mol. Cancer Res.* <https://doi.org/10.1158/1541-7786.MCR-21-0828> (2022).
79. Barilà, G., Rizzi, R., Zambello, R. & Musto, P. Drug conjugated and bispecific antibodies for multiple myeloma: improving immunotherapies off the shelf. *Pharmaceuticals* **14**, 40 (2021).
80. Taouk, G. et al. CD56 expression in breast cancer induces sensitivity to natural killer-mediated cytotoxicity by enhancing the formation of cytotoxic immunological synapse. *Sci. Rep.* **9**, 8756 (2019).
81. Rizzo, A. et al. Melanoma cells homing to the brain: an in vitro model. *Biomed. Res. Int.* **2015**, 476069 (2015).
82. Geertsen, R., Zenklusen, R., Kamarashev, J., Burg, G. & Dummer, R. Inverse regulation of neuronal cellular adhesion molecule (NCAM) by IFN-gamma in melanoma cell cultures established from CNS lesions. *Int. J. Cancer* **83**, 135–140 (1999).
83. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCOPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).
84. Feng, Y. et al. Differential killing of CD56-expressing cells by drug-conjugated human antibodies targeting membrane-distal and membrane-proximal non-overlapping epitopes. *MAbs* **8**, 799–810 (2016).
85. Leiendecker, L., Jung, P. S. & Obenaus, A. C. Targeting CD56 with an antibody-drug conjugate in Merkel cell carcinoma. *Br. J. Dermatol.* **186**, 209–210 (2022).
86. PBMC isolation and cryopreservation - <https://www.infinity.inserm.fr/wp-content/uploads/2018/01/PBMC-isolation-and-cryopreservation.pdf>.
87. Griffiths, J. A., Richard, A. C., Bach, K., Lun, A. T. L. & Marioni, J. C. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat. Commun.* **9**, 2667 (2018).
88. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
89. Bertolini, A. et al. scAmpI-A versatile pipeline for single-cell RNA-seq analysis from basics to clinics. *PLoS Comput. Biol.* **18**, e1010097 (2022).
90. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
91. Levine, J. H. et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
92. Prummer, M. et al. scROSHI: robust supervised hierarchical identification of single cells. *NAR Genom. Bioinform.* **5**, lqad058 (2023).
93. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
94. Tabib, T., Morse, C., Wang, T., Chen, W. & Lafyatis, R. SFRP2/DPP4 and FMO1/LSP1 define major fibroblast populations in human skin. *J. Invest. Dermatol.* **138**, 802–810 (2018).
95. Zhang, X. et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* **47**, D721–D728 (2019).
96. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
97. Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
98. Wang, X. et al. BREM-SC: a Bayesian random effects mixture model for joint clustering single cell multi-omics data. *Nucleic Acids Res.* **48**, 5814–5824 (2020).
99. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
100. Lischetti, U. et al. Dynamic thresholding and tissue dissociation optimization for CITE-seq identifies differential surface protein abundance in metastatic melanoma. *Zenodo* <https://doi.org/10.5281/zenodo.8124136> (2023).

Acknowledgements

We thank Dr. Gaetana Restivo and SKINTEGRITY.CH consortium for providing healthy skin samples and Dr. Corinne Stoffel for help with sample processing. We wish to thank Ina Nissen and Elodie Burcklen from the Genomics Facility Basel of the University of Basel and the ETHZ Department of Biosystems Science and Engineering for technical support with sample sequencing. This work was funded by the Swiss Cancer Research foundation (KFS-4459-02-2018) and by the grant #2017-510 of the Strategic Focal Area “Personalized Health and Related Technologies (PHRT)” of the ETH Domain.

Author contributions

M.P.L., C.B., A.T., and U.L. designed the study. A.T. and U.L. designed and performed experiments. U.L. performed sequencing. F.S., L.G., and M.C. performed bioinformatic analysis. P.F.C., Fed.S., and A.T. performed mIHC staining and analysis. V.H. and A.T. designed and analyzed flow cytometry experiments. J.M.M.G. and A.T. selected a clinical cohort for CITE-seq. U.L. and A.T. analyzed data and drafted the manuscript. All authors contributed to manuscript writing and editing. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The use of clinical material for research purposes was approved by the corresponding cantonal ethic commissions (Zürich: (Biobank): EK-687 and 800, KEK 2017-00688, Basel: EKBB 2019-00816) and informed consent was obtained for all human primary material.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05182-6>.

Correspondence and requests for materials should be addressed to Aizhan Tastanova or Mitchell P. Levesque.

Peer review information *Communications Biology* thanks Naveen Ramalingam and the other anonymous reviewer for their contribution to the peer review of this work. Primary Handling Editor: Joao Valente.

Reprints and permission 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) 2023