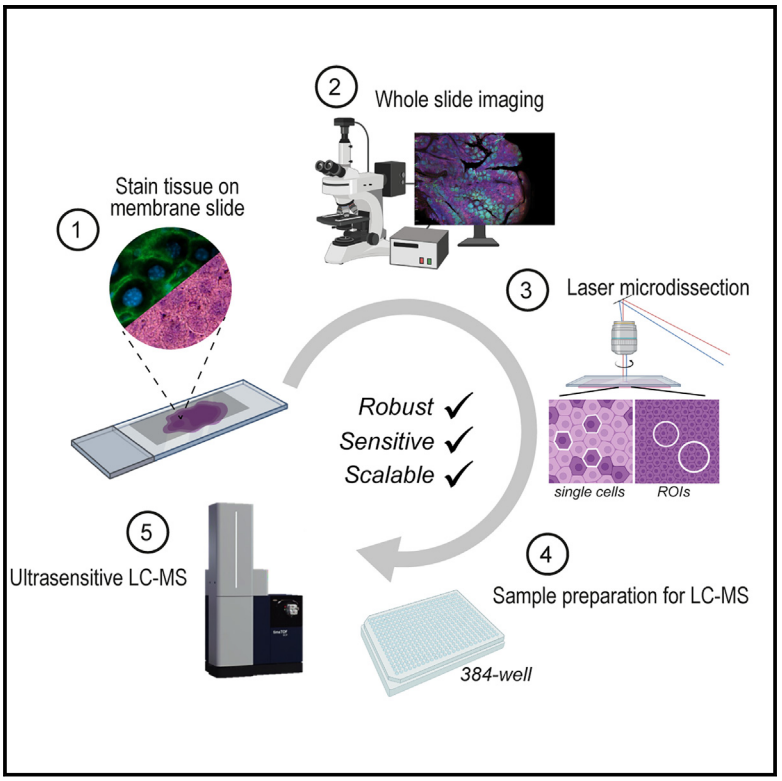
Methods

Cell Systems

A framework for ultra-low-input spatial tissue proteomics

Graphical abstract



Highlights

- Detailed protocol and guidelines for ultra-low-input spatial tissue proteomics
- Single-cell MS-based proteomics data obtained from archival FFPE tissue
- MS signal allows to normalize across tissue types and sample amounts
- Cytokines and transcriptional regulators quantified in tonsil microregions

Authors

Anuar Makhmut, Di Qin, Sonja Fritzsche, Jose Nimo, Janett König, Fabian Coscia

Correspondence

fabian.coscia@mdc-berlin.de

Makhmut et al., 2023, Cell Systems *14*, 1002–1014
November 15, 2023
https://doi.org/10.1016/j.cels.2023.10.003

CellPress

CellPress
OPEN ACCESS

Cell Systems

Methods

A framework for ultra-low-input spatial tissue proteomics

Anuar Makhmut,[1] Di Qin,[1] Sonja Fritzsche,[1] Jose Nimo,[1] Janett König,[1] and Fabian Coscia[1,2,*]
[1]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Spatial Proteomics Group, Berlin, Germany
[2]Lead contact
*Correspondence: fabian.coscia@mdc-berlin.de
https://doi.org/10.1016/j.cels.2023.10.003

SUMMARY

Spatial proteomics combining microscopy-based cell phenotyping with ultrasensitive mass-spectrometry-based proteomics is an emerging and powerful concept to study cell function and heterogeneity in (patho) physiology. However, optimized workflows that preserve morphological information for phenotype discovery and maximize proteome coverage of few or even single cells from laser microdissected tissue are currently lacking. Here, we report a robust and scalable workflow for the proteomic analysis of ultra-low-input archival material. Benchmarking in murine liver resulted in up to 2,000 quantified proteins from single hepatocyte contours and nearly 5,000 proteins from 50-cell regions. Applied to human tonsil, we profiled 146 microregions including T and B lymphocyte niches and quantified cell-type-specific markers, cytokines, and transcription factors. These data also highlighted proteome dynamics within activated germinal centers, illuminating sites undergoing B cell proliferation and somatic hypermutation. This approach has broad implications in biomedicine, including early disease profiling and drug target and biomarker discovery. A record of this paper's transparent peer review process is included in the supplemental information.

INTRODUCTION

Cells are the functional units of organs, which fulfill essential physiological tasks in a spatially defined manner to maintain tissue integrity.[1] To analyze cell dynamics in space and time, powerful spatial genomics,[2] epigenomics,[3] transcriptomics,[4–6] and imaging-based proteomics[7,8] methods have been developed to better understand cellular and molecular drivers of health and disease states. As proteins are the biomolecules closest to the cellular phenotype determining cell identity and function,[9,10] spatial proteomics (SP) methods are particularly promising for the study of human (patho)physiology. SP methods with the single-cell resolution are dominated by targeted antibody-based methods such as imaging mass cytometry[11] (IMC) or multiplex immunofluorescence (mIF) imaging,[8,12] where several dozen proteins can be analyzed at (sub)cellular resolution. However, although such methods are well suited for the large-scale screening of cellular phenotypes, they fall far short of capturing the actual complexity of the cellular proteome. It is estimated that single-cell types express more than 10,000 unique proteins,[9] which is complemented by millions of potential proteoforms, including splice variants, post-translational modifications (PTMs), and protein sequence variants.[10,13,14] Liquid chromatography mass spectrometry (LC-MS)-based proteomics in contrast enables the study of proteomes at an unbiased (i.e., untargeted), quantitative, and system-wide level.[9] The combination of both of these complementary proteomic approaches is therefore highly desirable but requires integrated and multimodal pipelines. We recently introduced deep visual proteomics (DVP),[15] a new concept combining imaging-based (fluorescence or bright field) single-cell phenotyping with unbiased MS-based proteomics for global proteome profiling with cell type and spatial resolution. To realize DVP, we developed an automated laser microdissection (LMD) workflow for the streamlined collection of nuclei, cells, or larger regions of interest (ROIs) directly into 96- or 384-well plates, thereby connecting whole-slide imaging and deep-learning-based image analysis[16] with ultrasensitive MS-based proteomics.[17] This allowed the profiling of as little as 100 phenotype-matched cells from archival tissue material while also preserving detailed cell type and spatial information. Further advances in sample preparation and MS acquisition recently pioneered the profiling of single-cell proteome heterogeneity in cryosections of murine liver tissue,[18] emphasizing the strong spatial influence on the hepatocyte-specific proteome. Despite these promising proof-of-concept studies, a systematic evaluation and optimization of all experimental steps of IF microscopy-guided spatial tissue proteomics is still missing. In particular, the analysis of few or even single cells of formalin-fixed and paraffin-embedded (FFPE) tissue collected by LMD has remained elusive and relies on optimized and robust "end-to-end" protocols. The successful development of such integrated workflows could pave the way for a plethora of biomedical applications, including early disease proteome profiling studies directly from archived patient material, where only a few cells can be present.

Here, we describe a scalable, robust, and easy-to-use protocol optimized for the profiling of ultra-low-input archival tissue guided by whole-slide (IF) imaging. After benchmarking in murine liver tissue, we applied our workflow to study the cell-type-resolved proteome of B and T lymphocytes in different spatially defined niches, guided by four-marker whole-slide IF imaging. We finally provide detailed guidelines covering all aspects of our workflow (Methods S1), from antigen retrieval (AR) and staining on LMD membrane slides, over manual or automated sample processing, to MS data acquisition and analysis.

RESULTS

Optimizing LMD-based low-input FFPE tissue proteomics

Light microscopy is an integral part of SP workflows that combine LMD with MS-based proteomics[19–21] (Figure 1A). Although hematoxylin and eosin (H&E) staining protocols are well established for tissue sections mounted on specialized LMD membrane slides[23] (Figure S1A), IF-based protocols, which allow the sensitive detection of a higher number of cell type and functional markers (typically 3–5 per imaging cycle), require additional AR steps for epitope unmasking. The choice of the AR method can not only influence staining and image quality but also impacts LMD membrane integrity, tissue collection efficiency, and potentially proteome coverage of trace sample amounts. The evaluation and optimization of AR and staining prior to MS is therefore important for ultra-low input or even single-cell applications that strictly depend on highly efficient tissue collection and protein extraction, as well as near-lossless sample preparation protocols. In other words, the most sensitive mass spectrometer can only be as good as the upstream sample preparation workflow that delivers these trace peptide amounts to the instrument. We therefore first evaluated different AR methods for their general compatibility with whole-slide imaging on LMD slides, efficient LMD, and ultrasensitive MS-based proteomics. We tested two common protocols based on heat-induced epitope retrieval (HIER) or proteolytic (i.e., pepsin-induced epitope retrieval [PIER]) epitope retrieval and immunofluorescently stained murine liver FFPE tissue with an antibody against a ubiquitously expressed plasma membrane marker (Na/K-ATPase). We chose murine liver as benchmarking tissue as hepatocytes make up 60%–80% of the liver cell mass,[24] which allowed us to obtain consistent results from serial, homogeneous tissue slices, and repeated sample collections over the course of our experiments. 5-μm-thick liver sections were mounted on glass membrane (polyethylene naphthalate [PEN] or polyphenylene sulfide [PPS]) or metal frame (PPS) slides, de-paraffinized, and subjected to two common AR protocols (STAR Methods) prior to antibody staining, IF microscopy, LMD, and proteomics (Figure 1B). We included H&E stains for comparison, which do not rely on additional AR after de-paraffinization, allowing us to directly investigate the impact of PIER and HIER on our proteomic results. PIER was fully compatible with both LMD slide types (glass or metal frame), facilitating efficient LMD, but generally came at the cost of higher background staining for the tested antibody, compared with HIER (Figure S1A). HIER occasionally resulted in membrane distortion of the glass-type slides toward the label end of the slide (Figure S1B), which can impede efficient LMD collection in this area if not prevented by the addition of glycerol.[25] We alternatively recommend tissue mounting on the side of the slide, which is distant from the label. Frame slides on the contrary are more suitable for diverse biological sample types (i.e., tissue or cell culture) but are suboptimal for the collection of directly connected contours, for example, for gridded sampling schemes as used for the nanoPOTS approach,[20] as this can ultimately lead to loss of overall membrane integrity. Irrespective of the choice of AR (HIER or PIER), staining technique (H&E or IF), or LMD slide type (PEN or PPS), proteomics results from three different liver tissue amounts were highly consistent (Figures 1C–1E). Tissue samples were processed in 384-well low-binding plates using an MS-compatible, organic solvent-based protocol, which included 60-min heating at 95°C for efficient formalin decrosslinking,[26] sequential Lys-C and trypsin digestion and miniaturized solid-phase extraction. Compared with our previous DVP protocol, we optimized our workflow for lower microliter volumes (1–2 μL) to minimize peptide loss from surface adsorption while still being pipette-able with standard laboratory equipment. At the same time, this also allowed the integration of robotic sample preparation workflows (STAR Methods). In addition, we used an optimized 15-min active nano-LC gradient (Figures S1C–S1E) in combination with an optimal window design dia-PASEF[27] method on a trapped-ion mobility spectrometry (TIMS) mass spectrometer (Bruker timsTOF SCP) for improved sensitivity and sample throughput. Using DIA-NN,[28] we quantified more than 9,000 precursors and 1,700 unique proteins from small tissue regions of approximately one to two hepatocytes (~7,500 μm^3, Figures 1C, S1F, and S1G). From 50-cell samples (50,000 μm^2, 5 μm thick) of H&E, HIER, and PIER-derived samples, 4,000 proteins were consistently quantified with excellent quantitative reproducibility (Pearson r = 0.98, Figure 1D). Furthermore, we found a 95% overlap of protein identifications across AR methods (Figure 1F), as well as nearly identical cellular compartment proportions of the underlying proteomes (Figure 1G), and consistent with a deep liver proteome study of primary cells.[22] We conclude that common FFPE tissue preparation methods for fluorescence microscopy are fully compatible with ultra-low input MS-based proteomics. However, depending on the concrete application, glass membrane or metal frame slides offer unique advantages and disadvantages, which we summarized in Figure 1H to provide general guidelines for LMD proteomics beginners. These data also include input from two additional LMD expert labs and together with our detailed protocol description (Methods S1; supplemental information), they are intended to support the selection of the right tissue preparation and sample collection strategy for diverse SP applications.

A scalable FFPE tissue proteomics workflow allowing single-cell analysis

Having established that common AR and staining methods are fully compatible with LMD and ultra-low-input tissue proteomics, we next assessed the scalability, robustness, and minimally required sample amount of our workflow. Homogeneous areas of murine liver tissue were collected by LMD into 384-well

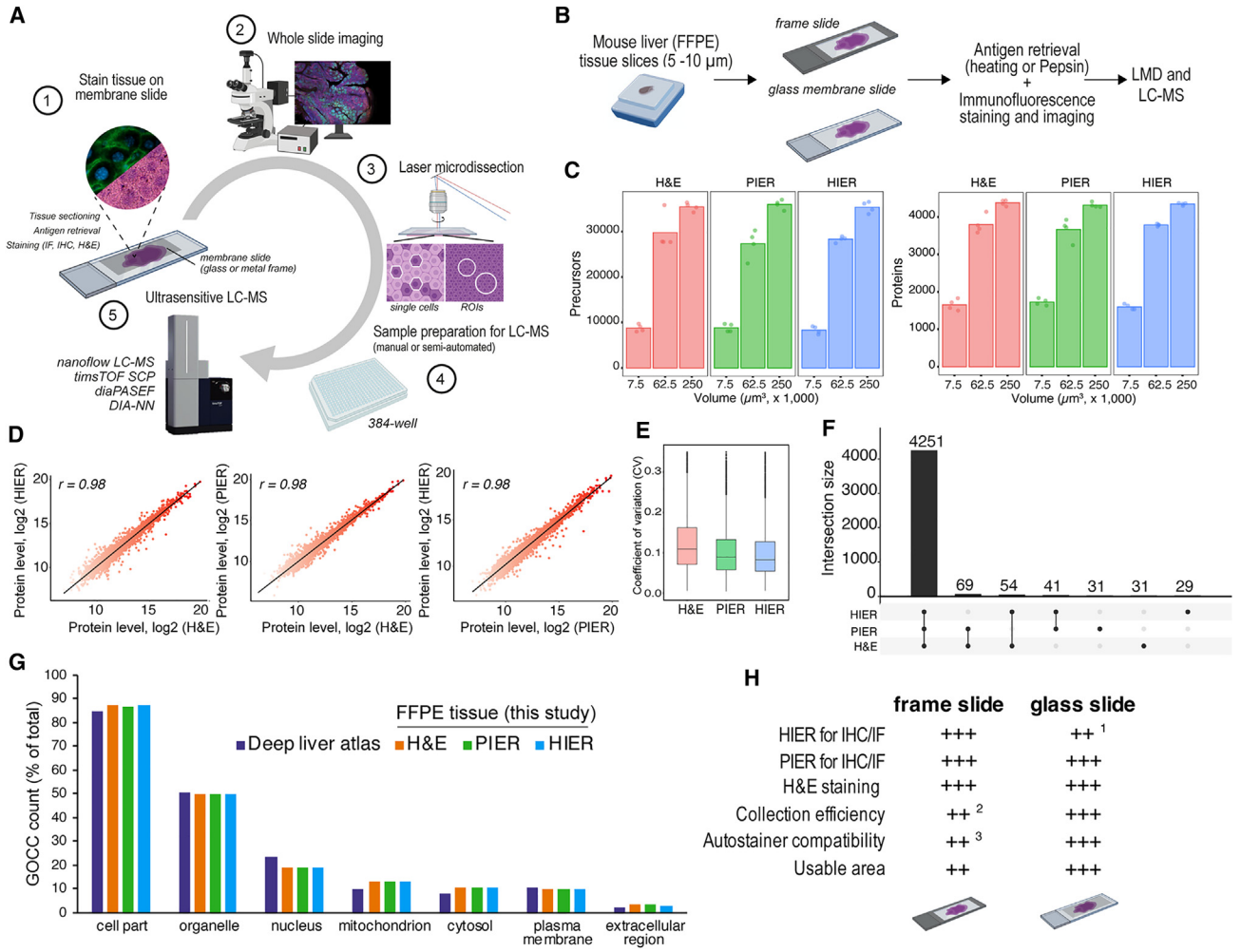


Figure 1. Optimizing laser microdissection-based low-input FFPE tissue proteomics

(A) Overview of the spatial tissue proteomics workflow.

(B) Tissue preparation strategy for laser microdissection-based proteomics benchmarking experiments using specialized LMD slides (metal frame or glass).

(C) Precursor and protein identifications from tissue samples processed with different staining and antigen retrieval methods. Areas of 1,562, 12,500, and 50,000 μm^2 of a 5-μm thick section were laser microdissected. Averages are shown from quadruplicate measurements.

(D) Proteome correlations (Pearson's *r*) of HIER, PIER, and H&E-based tissue samples.

(E) Boxplots showing coefficients of variations (CVs) of triplicate proteome measurements from the three antigen retrieval methods. Related to (D).

(F) Upset plot showing common and exclusive proteins for different antigen retrieval strategies based on 50,000 μm^2 tissue samples.

(G) Protein identifications from major cellular compartments ("cytosol," "nucleus," "plasma membrane," and "extracellular region," Gene Ontology Cellular Component [GOCC]) from H&E, PIER, and HIER-treated mouse liver tissues. Percentages are the number of quantified proteins per compartment over all quantified proteins in the corresponding sample. For comparison, a deep mouse liver proteome dataset was included[22] based on non-fixed, primary cells.

(H) Summary of the laser-microdissection optimizations for low-input proteomics. Three labs assessed the applicability of glass and frame slides and rated each category with moderate (+), good (++), and excellent (+++). The average score is shown from all three ratings. [1]HIER can cause membrane distortion of glass-type slides. [2]Frame slides are more problematic for grid-based sampling schemes, the collection of many closely connected contours can lead to loss of overall membrane integrity. [3]Tested autostainers: Ventana (Roche) supported frame and glass slides, DAKO (Agilent) system supported glass slides. (A) and (B) created with Biorender.

low-binding plates, ranging from single hepatocytes (600 μm^2, 5 μm thick) to approximately 100 cells (100,000 μm^2, Figures 2A, 2B, S1F, and S1G). Proteomic results showed a linear increase in MS2 intensity, precursor, and protein identifications from "low" to "high" tissue amounts, which, as expected, was inversely correlated with the median coefficient of variations (CV) of protein quantifications calculated from triplicate measurements of adjacent regions (Figures 2C–2E). However, even single hepatocyte contours featured low median CVs of 20% for the 1,500–2,000 quantified proteins per contour (Figures 2E and S2A–S2C), which we and others previously only achieved from many thousands of cells collected from FFPE tissue.[30,31] It is noteworthy that CVs also included true biological variation from known spatially defined hepatocyte heterogeneity,[18,32] letting us conclude that reproducible single-cell proteome analysis from FFPE tissue is achievable based on our

CellPress
OPEN ACCESS

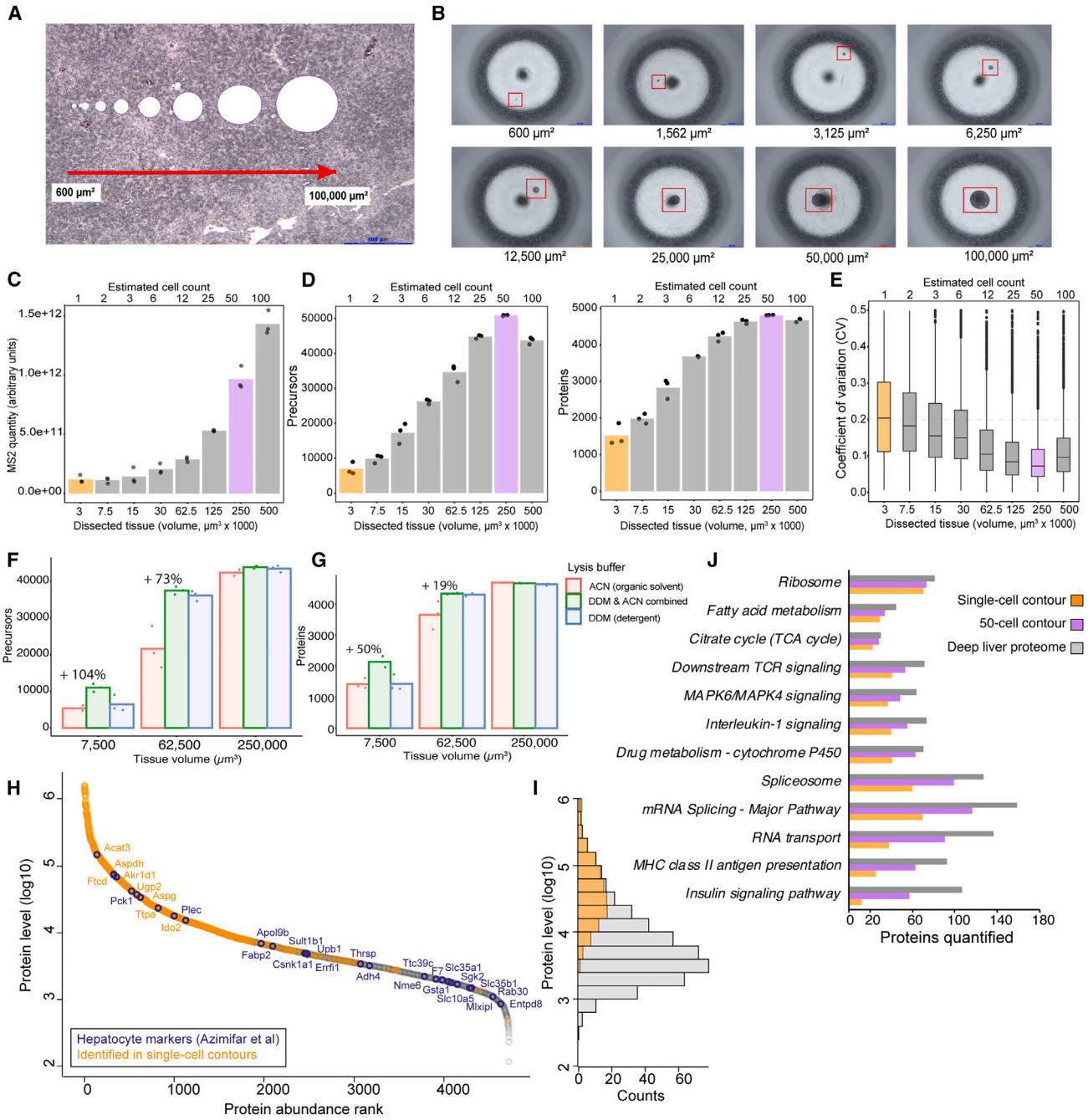


Figure 2. A scalable FFPE tissue proteomics workflow allowing single-cell analysis

(A) H&E-stained mouse liver tissue section (5 µm). Collected LMD tissue samples are shown from 600 up to 100,000 µm^2. Single hepatocyte contour isolation (average size 600 µm^2) was guided by the immunofluorescence signal of the plasma membrane marker Na/K-ATPase and the DAPI signal. Larger regions of interest (ROIs) were isolated based on circular contours of pre-defined sizes ranging from 1,562 to 10,000 µm^2 (2–100 cells), respectively. Hepatocyte volumes of 4,000–6,000 µm^3[29] were used for cell count estimations of circular ROIs. Scale bars, 400 µm.

(B) Tissue inspection after collection into a 384-well low-binding plate. Scale bars, 400 µm.

(C and D) Average MS2 quantities (C), number of identified precursors (D, left), and proteins (D, right) of the collected tissue samples are shown. For all amounts, three replicates were collected and measured. The tissue optimum is highlighted in purple and single-cell contours in orange.

(E) Boxplots showing the CVs of protein quantification across different tissue areas. CVs were calculated from triplicates of non-log-transformed data. The boxplots define the range of the data (whiskers), 25th and 75th percentiles (box), and medians (solid line).

(F and G) Impact of different lysis buffers on proteome coverage. Tissues of 7,500, 62,500, and 250,000 µm^3 of a 5-µm thick mouse liver section were profiled using three different methods based on organic solvent only (ACN), DDM only (no ACN), or combined (DDM and ACN) protocol. For all amounts, three replicates were collected and measured.

(legend continued on next page)

workflow. We attribute the excellent quantitative reproducibility to the combination of optimized tissue preparation, low-volume sample processing, and highest-sensitivity MS acquisition and analysis using an optimal window dia-PASEF scheme (STAR Methods). Precursor and protein identifications peaked for 50–100 cell regions (25,000–50,000 μm^2), where we quantified close to 50,000 precursors and 5,000 proteins (Table S1). Interestingly, further increasing sample amounts resulted in lower precursor identifications and higher median CVs, indicative of sample overloading (Figures 2D and 2E). Although this phenomenon could possibly be balanced out using longer LC gradients and adjusted trypsin amounts, thereby further improving proteome coverage, we note that our optimized 15-min active nanoflow gradient provides an excellent compromise between single-cell sensitivity, high proteome coverage for 50–100 cell samples, and reasonable sample throughput of around 30–40 samples per day. Additionally, we tested the applicability of our protocol in combination with the Bruker timsTOF Pro2, which shows a roughly 4- to 5-fold lower total ion current (TIC) compared with the SCP instrument.[17] For the lowest tissue amounts measured (7,500 μm^3 samples, 1–2 hepatocytes[29,33]), close to 1,000 proteins could still be quantified with high quantitative reproducibility and a somewhat similar proteome coverage was observed for the 50-cell samples (Figures S2D–S2F). We next tested if we could further increase proteome coverage of single FFPE tissue contours through the use of MS-compatible detergents. We assessed the integration of n-Dodecyl-β-D-maltoside (DDM), which was previously shown to improve proteome coverage of low-input FFPE tissue, in particular when applied at high temperatures.[34] Combining our acetonitrile (ACN) based protocol (10% final concentration) with DDM (0.1% final concentration), which likewise included 60-min controlled heating at 95°C for efficient formalin de-crosslinking, proteome coverage of the single-cell samples improved by 104% and 50% for precursor and protein identifications, respectively (Figures 2F and 2G). For 62,500 (~12 cells) or 250,000 μm^3 (~50 cells) samples, proteome coverage was more similar between protocols. As DDM is not removed during peptide clean-up steps, its accumulation on the analytical column can compromise chromatographic performance over time, which can be prevented by additional high organic solvent washing. Our organic-solvent-based protocol performed equally well for 50-cell contours, thus offering an excellent and cleaner alternative to the DDM/ACN combination for "higher" tissue amounts.

The unprecedented depth of our single FFPE hepatocyte contours, reproducibly quantifying up to 2,000 proteins depending on tissue thickness (Figure S2C) and with high data completeness (89% complete; Figure S2G), encouraged us to further explore these single-cell tissue proteome data. Protein levels generally showed a high degree of concordance with higher loading amounts (Figure S2H) and included many known hepatocyte-specific markers distributed over a dynamic range of approximately three orders of magnitude (Figures 2H and 2I). Proteins involved in "housekeeping" cell functions were quantified at a similar depth of analysis as the 50-cell samples, despite the approximately 33-fold lower total MS2 signal, and, remarkably, compared with the deep mouse liver study covering more than 10,000 proteins[22] (Figure 2J). For example, we quantified 70 of 81 ribosomal proteins, 29 of 44 involved in fatty acid metabolism and 22 of 30 tricarboxylic acid (TCA) cycle-related proteins (Table S2). Interestingly, although housekeeping functions such as ribosomal proteins were more stably expressed, cell-metabolism-related proteins showed higher variation in the single-cell contours compared with the 50-cell contours, likely due to proteome averaging (Figure S2I). This is in line with a recent liver study revealing strong metabolic differences in single hepatocytes along the liver zonation axis.[18] For lower abundant pathways, for example, insulin signaling or major histocompatibility complex (MHC)-2 antigen presentation (12 vs. 107 and 25 vs. 92 detected proteins, respectively) (Figure 2J), a higher discordance in detected proteins was observed between the low and high tissue amounts, making the comprehensive single-cell analysis of these pathways only achievable with further improved workflow sensitivity, or alternatively, by pooling phenotype-matched cells, as we conceptualized recently.[15] We conclude that reproducible single-cell-based FFPE tissue profiling is achievable using optimized sample processing and ultrasensitive LC-MS workflows revealing important insights into cell identity and function. The scalability of our workflow also enabled the spatially resolved quantification of ~5,000 proteins from 50-cell regions, capturing a substantial fraction of the cell-type-specific proteome, thereby complementing single-cell-based analyses.

Optimizing sample input across tissue and cell types

Based on our tissue dilution experiment in the murine liver, we empirically determined the optimal tissue amount for the highest proteome coverage of small tissue areas using a 15-min active nanoflow gradient combined with dia-PASEF on the Bruker timsTOF SCP. We next addressed how this liver optimum translated to other tissue and cell types and how the recorded MS readout could be exploited to normalize sample loading from tissue to tissue or cell type to cell type. Such adjustments are of particular importance for ultra-low sample amounts, which are not amenable to peptide concentration measurements routinely used prior to MS bulk analysis. In addition, various factors can affect the TIC derived from different tissue specimens, including sample-related sources of variability such as tissue archival time, which can affect the retrieval of lower abundant proteins,[35] or biologically due to protein abundance differences across tissue and cell types.[36] In any case, normalizing and adjusting sample amounts is hence particularly important for low input tissue

(H) Dynamic range of protein abundance for 50-cell (250,000 μm^3) contours. Proteins identified in single-cell contour samples are highlighted, as well as hepatocyte-specific markers.[22] A minimum of two quantified values per quadruplicate measurement was required for the single-contour samples.
(I) Histogram of $\log_{10}$ protein intensities obtained from 600 and 50,000 μm^2 samples. Proteins identified in single cells cover 2–3 orders of magnitude from the top abundant fraction of the liver proteome.
(J) Reactome and KEGG pathway coverage for single-cell and 50-cell samples. Values show the number of proteins quantified per pathway. For comparison, a deep (>10,000 proteins) mouse liver dataset[22] was included. A minimum of two quantified values from quadruplicate measurements was required for single cells and three values from triplicates for the 50-cell samples.

CellPress
OPEN ACCESS

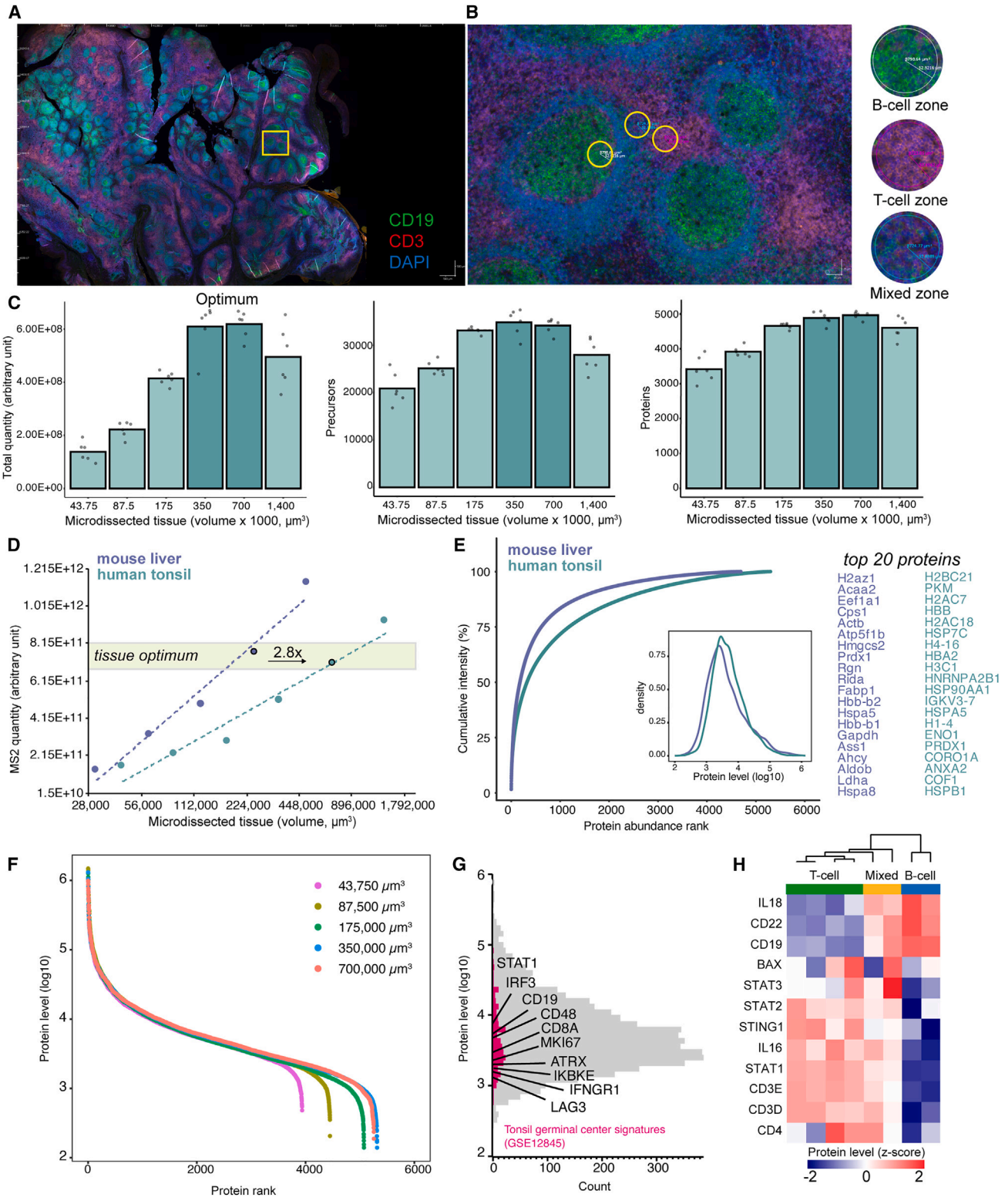


Figure 3. Optimizing sample input across tissue and cell types

(A) Immunofluorescence whole-slide image of a 5-µm-thick tonsil tissue section stained for CD3 (T cells), CD19 (B cells), and DNA (DAPI). Scale bars, 500 µm.

(B) Magnifications of exemplary B and T cell enriched regions used for laser microdissection and proteomics profiling. Scale bars, 25 µm.

(legend continued on next page)

proteomics and should be carefully assessed. We hypothesized that a simple normalization strategy using the total excised tissue volume (microdissected area × section thickness) across specimens should be a poor estimator of the optimal sample amount. To test this, we analyzed a second tissue type, the human FFPE tonsil, which as a secondary lymphoid organ is important for the development of immune tolerance and adaptive immune functions and comprises B and T cell subsets, as well as other immune and non-immune-related cell types.[37] In murine liver, 50–100 cell contours (125,000–250,000 μm^3) yielded the highest proteome depth and lowest median protein CVs (Figures 2C–2E and S3A–S3C). Analyzing the same tissue amounts obtained from tonsil, focusing on T and B cell enriched microregions after IF imaging (Figures 3A and 3B), revealed clearly different sampling optima (Figure 3C), despite a similar total number of quantified proteins per experiment (5,000–5,500 proteins; Table S3). Based on the MS2 signal (sum of MS2 quantities of all peaks) and total quantity (sum of MS2 quantities of identified precursors) retrieved from the DIA-NN output, we estimated that murine liver on average resulted in roughly 3-fold (2.8×) higher MS2 signal compared with tonsil (Figure 3D), likely due to different protein abundances in these two organs (Figure 3E). Notably, the MS2 signal derived from the liver reference sample was consistent over the tested time frame of 6 months and independent of section thickness (Figure S3D). In other words, a 5-μm-thick contour of 25,000 μm^2 resulted in an almost identical MS2 signal as a 10-μm-thick contour of 12,500 μm^2. The optimal tissue amount for tonsil was 350,000–700,000 μm^3 instead, allowing us to quantify ~35,000 precursors and ~5,000 proteins in the 15-min dia-PASEF measurement (Figure 3C), and beyond which no further increase in proteome depth was apparent. In fact, we observed lower identification rates beyond this saturation point, similar to our findings in liver (Figures 2C–2E), which we mostly attribute to increased tryptic miscleavage (Figure S3E), as there was no sign of TIMS saturation for the tested tissue amounts. However, further increased peptide amounts could result in precursor-intensity-dependent TIMS cell fragmentation, emphasizing the need to carefully assess and normalize sample loading. This prompted us to test whether these tissue saturation points obtained from the liver and tonsil data could be exploited to predict the best tissue sampling amount from "single-shot" measurements alone. This simulates a scenario at the beginning of a spatial tissue proteomics study, where a priori proteomics information for a new tissue type is lacking. The accurate prediction of the optimal tissue amount could hence save time and resources and maximize data output. To this end, we laser microdissected replicates of small 60,000 μm^3 samples (6,000 μm^2 of a 10 μm slice) of a consecutive tonsil tissue section, such that the obtained intensity was in the linear range of the mass spectrometric signal (>50,000 μm^3 tissue; Figure 3D) and compared the recorded MS2 precursor quantity with our pre-determined optimum (MS signal = 6–8 × 10^{11}). To minimize variability from different extracellular matrix compositions, we focused on homogeneous tissue regions of high cellularity (Figure 3B), similar to the liver titration experiment (Figure S1G). This way, we extrapolated that 450,000–753,000 μm^3 would be the optimal sampling amount for this tonsil tissue, in excellent agreement with our measurements (Figure 3C). We note that this simple normalization strategy is strictly dependent on the employed LC-MS setup and therefore requires empirical data to begin with but generally recommend to perform such "survey" experiments to pinpoint the optimal tissue amount needed for a specific SP research question. In our example, murine liver served as highly consistent reference tissue to predict the ideal tonsil amount; however, other tissue types should be equally suited for such normalization. Moreover, this strategy can prevent TIMS overloading, which can negatively affect proteome coverage and quantification. As a positive byproduct, the acquired raw files of the titration data can also be used to create project-specific refined spectral libraries (STAR Methods), thereby drastically speeding up subsequent DIA-NN searches.

As the size of the total dissected tissue area (or the number of collected single-cell contours per sample), which determines the obtained spatial resolution, is inversely correlated with proteome coverage (Figures 2C and 2D), such survey measurements can also help to find a good balance between these two parameters in the context of the specific research question. To demonstrate this exemplarily, we further analyzed our tonsil proteome data, which in total covered more than 5,000 proteins distributed over four orders of magnitude (Figure 3F). Projecting the different tissue dilution measurements onto the measured dynamic range of protein abundance showed that B or T cell-specific regions of only 8,750 μm^2 (5-μm-thick section) were already sufficient to quantify many key players of immune cell signaling, cell-type-specific markers, cytokines, and even transcription factors (e.g., STAT1, IRF3, IFNGR1, CD8A, CD19, Ki-67, LAG3, interleukin [IL]-16, and IL-18; Figures 3B, 3G, 3H, S3F, and S3G) at a spatial resolution of ~70–100 μm (center-to-center; Figure 3A). Consequently, our pipeline should allow the analysis of spatially resolved proteomes of various B and T cell niches from regions of as little as ~4,000 μm^2 (10-μm-thick section).

(C) Bar plots of MS2 intensities (left), precursors (middle) and proteins quantified (right) obtained from increasing amounts of tonsil tissue. For all amounts, six replicates were collected and measured. Note, the tissue-specific sampling optimum was reached at 350,000–700,000 μm^3, beyond which identifications dropped again.

(D) Comparison of liver and tonsil tissue dilution data. MS2 quantities are plotted against increasing (log_2 transformed) tissue amounts (volume in μm^3). Based on the empirically determined MS2 intensity optimum (MS2 quantity = 6–8E11 for liver and tonsil tissues), 2.8-fold more tonsil tissue was sampled to reach the same MS2 quantity. For tonsil and liver tissue samples, data show averages from minimum six and five replicates per group, respectively.

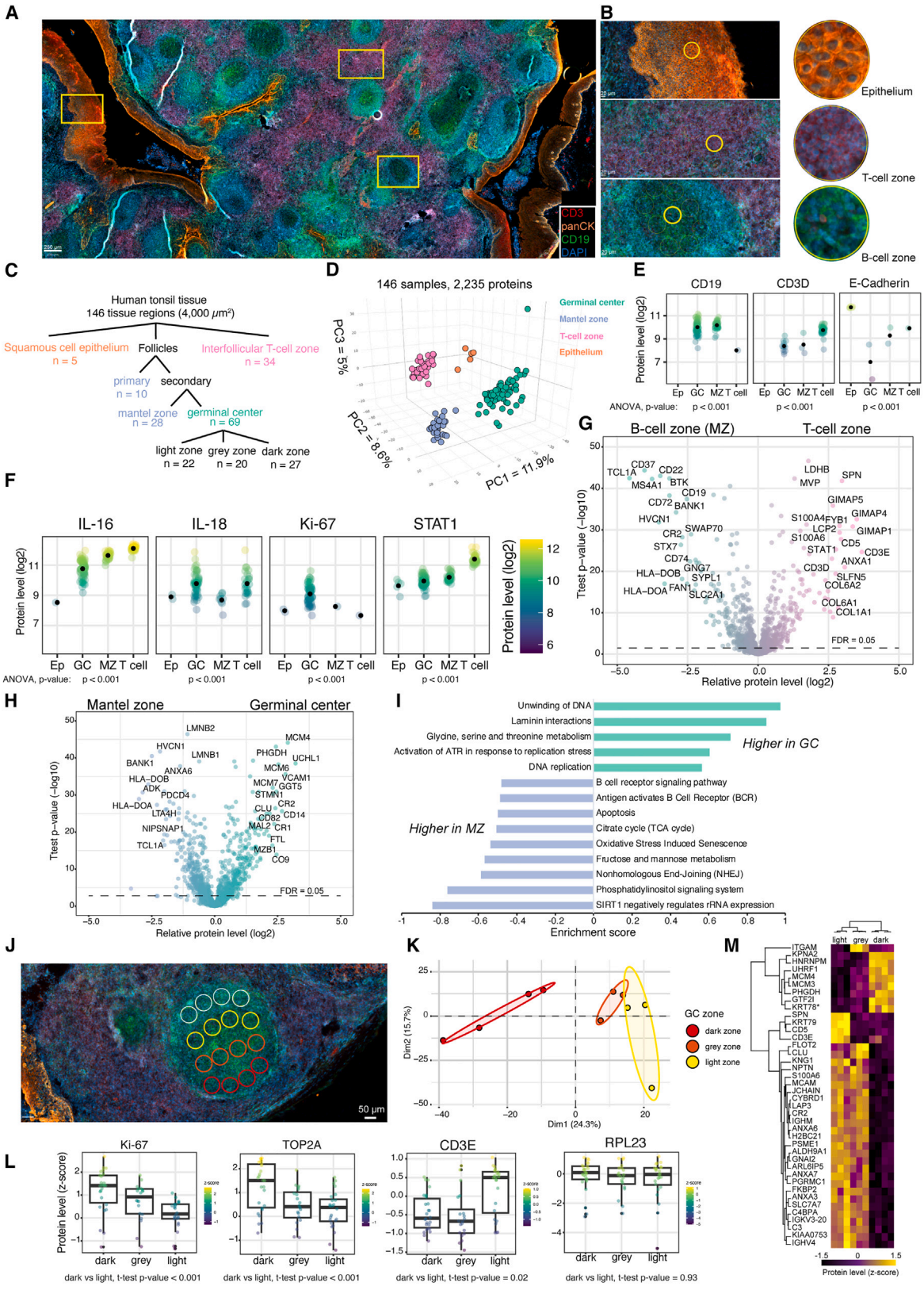
(E) Cumulative protein intensities of liver and tonsil tissue proteomes ranked from the highest to the lowest abundant protein. Density plot shows protein intensity distributions for both tissues. Top 20 proteins of each tissue are shown on the right.

(F) Dynamic range of protein abundance from different amounts of tonsil tissue.

(G) Histogram of protein intensities of the of 8,750 μm^2 tonsil tissue sample (43,750 μm^3 in volume). Proteins highlighted in pink belong to an RNA-seq-based signature of tonsil germinal centers (GSE12845), including known immune and cell-type-specific markers.

(H) Unsupervised hierarchical clustering of small (6,000 μm^2 × 10 μm) B cell, T cell, and mixed B/T cell regions, related to (B). *Z* scored protein levels indicate upregulated (red) or downregulated (blue) proteins.

CellPress
OPEN ACCESS



(legend on next page)

Spatially and cell-type-resolved proteomics of human tonsil tissue

Encouraged by our tonsil titration data, which revealed that known immune cell regulators, cytokines, and transcription factors were quantified from tissue regions of as little as 8,750 μm^2 (43,750 μm^3 in volume), we systematically investigated the impact of the spatial location on the B and T cell specific proteomes. Human tonsil represents a prime example of tissues that are organized into distinct microanatomical compartments to fulfill diverse biological functions critical for adaptive immunity. Following antigen encounter of naive B cells within the follicle, secondary (activated) follicles are rapidly formed as production centers of antigen-specific B cells. Following activation, B cells undergo cycles of maturation and selection within newly formed germinal centers (GCs), ultimately giving rise to highly antigen-specific antibody-secreting plasma cells and memory B cells, the key players of humoral immune response.[38]

We first mounted a 10-μm-thick tonsil section obtained from a patient who underwent bilateral tonsillectomy on a metal frame LMD slide and performed four-color IF whole-slide imaging to detect B cells (CD19), T cells (CD3), epithelium (EP) (pan-CK), and DNA (DAPI) (Figures 4A and 4B). Using the open-source image analysis software QuPATH,[39] we selected a total of 146 microregions for automated LMD (STAR Methods) and quantitative proteomics (Table S4). Based on the well-defined tissue architecture of the human tonsil (Figure 4A), we included small circular regions of 4,000 μm^2 isolated from primary and secondary B cell follicles, including subregions of dark, light, and gray GC B cell niches, naive mantel-zone-derived B cells, various interfollicular T cell zones, and squamous cell EP (Figures 4A–4C). On average, we quantified 1,952 proteins per sample (Figure S4C) and 3,334 proteins in total. After data filtering and missing value imputation (STAR Methods), proteomes clearly separated by microanatomical regions dominated by the distinct cell types (Figure 4D). Known cell type markers such as CD3D (T cell marker), CD19 (B cell marker), and CDH1 (E-Cadherin, epithelial marker) were highest in the expected sample groups (Figure 4E), confirming the high specificity of our proteome data. GCs showed clear signs of increased proliferation (e.g., Ki-67, PCNA) and DNA repair (e.g., TOP2A), indicative of active sites of B cell expansion and somatic hypermutation.[40] T cell zones instead featured high levels of the STAT1 transcription factor as well as T cell modulating cytokines such as IL-16 (Figure 4F). Moreover, the global comparison of all T cells (n = 34) vs. B cell-specific proteomes (mantle zone [MZ], n = 36) revealed many bona fide B and T cell markers among the top regulated proteins (e.g., CD22, CD3E, CD72, and CD5) and potentially many other less-characterized ones (Figure 4G). We next focused on different B cell niches to assess if our spatially resolved data captured known functional differences of spatially defined B cell zones. Naive (MZ) and activated (GC) B cells showed strong proteome differences (Figure 4H), indicative of their unique biological functions. Although MZ B cells were characterized by higher metabolic activity (e.g., TCA cycle and fructose and mannose metabolism), senescence signatures, and phosphatidylinositol signaling, activated GC B cells showed strong replication and DNA repair signatures (Figures 4H and 4I; Table S4), in line with their known biological function. Encouraged by this, we next assessed whether our quantitative data even separated spatially defined GC sub-compartments of dark (sites undergoing active B cell proliferation and somatic hypermutation) vs. light (site of B cell selection) zones (Figures 4J and 4K). Clearly, dark-zone-derived B cells featured higher replication and DNA damage response profiles compared with the light zone (Figures 4L and 4M). The higher expression of the T cell receptor subunit CD3D in light zones on the contrary illuminated the presence of T helper cells, important for B cell selection.[40] We confirmed this finding by assessing our imaging data, which indeed revealed significantly higher CD3 signals in the light GC regions, which was not the case for the B cell marker CD19 (Figure S4D).

In summary, our data delineated how the robust microscopy-guided ultra-low-input tissue proteomics workflow introduced here can be applied to study cell type and spatially resolved proteomes in health and disease based on readily accessible archival FFPE specimens.

Figure 4. Spatially and cell-type-resolved proteomics of human tonsil tissue

(A) Immunofluorescence whole-slide image of a 10-μm thick tonsil tissue section stained for CD3 (T cells), CD19 (B cells), pan-CK (epithelium), and DNA (DAPI). Scale bars, 250 μm.

(B) Magnifications of the exemplary epithelium (EP), germinal center (GC), mantle zone (MZ), and T cell enriched regions used for laser microdissection and proteomic profiling. Scale bars, 20 μm.

(C) Sample collection strategy and the total number of samples for each tissue region.

(D) 3D principal-component analysis (PCA) of 146 samples based on 2,235 protein groups after data filtering and imputation.

(E and F) Log_2-protein levels of cell-type-specific and functional markers quantified in different tissue regions. Black dots indicate average values for each group. Asterisks indicate ANOVA p values of <0.001 after data imputation.

(G) Volcano plot of the pairwise proteomic comparison between the B cell zone (mantle zone) and T cell zone. Cell type-specific markers are highlighted in green and turquoise (two-sided t test, false discovery rate [FDR] < 0.05).

(H) Volcano plot of the pairwise proteomic comparison between mantle zone and germinal center samples. Cell type-specific markers are highlighted in green and blue (two-sided t test, FDR < 0.05).

(I) Pathway enrichment analysis (Reactome and KEGG) based on t test difference between mantel zone and germinal center samples. Selected pathways with a Benjamin-Hochberg FDR < 0.05 are shown.

(J) ROIs used for proteomic profiling of a secondary follicle region. Dark (red), gray (orange), and light (yellow) germinal center zones were selected for proteomic profiling. Mantel zone regions (light gray) are shown additionally. Scale bars, 50 μm.

(K) PCA of dark, gray, and light zone proteomes. Point concentration ellipses are shown for each group with a 95% confidence.

(L) Boxplots of relative protein levels (*Z* score) for selected markers. The boxplots define the range of the data (whiskers), 25th and 75th percentiles (box), and medians (solid line). Asterisks indicate two-sided t test p values (dark vs. light zone) of p < 0.001.

(M) Unsupervised hierarchical clustering of ANOVA significant proteins (permutation-based FDR < 0.05) from dark, gray, and light zone samples. Related to (J) and (K). Heatmap shows relative protein levels (*Z* score) of upregulated (yellow) and downregulated proteins (black). *KRT78 is marked as potential contaminant.

DISCUSSION

Spatial tissue proteomics connecting microscopy-based cell phenotyping with LMD guided MS-based proteomics is an emerging discovery concept for the study of cell function and heterogeneity in health and disease. Our group recently co-developed deep visual proteomics, an approach that combines high-parametric imaging and machine-learning-based single-cell phenotyping to guide precise tissue sampling for ultrasensitive LC-MS analysis. This enabled the profiling of as little as 100 tissue cells per sample to a depth of 3,000–5,000 proteins, dependent on the tissue and cell type of interest. However, the flexible and highly modular design of the DVP pipeline, enabling the profiling of single or few cells on one hand, or hundreds of phenotype-matched cells for deeper proteome interrogation (i.e., 5,000 proteins or more) on the other hand, also necessitates carefully designed tissue benchmarking experiments and detailed guidelines to extract most information for diverse biomedical applications. With this in mind, we here introduce, benchmark, and apply an optimized end-to-end workflow, starting from AR comparisons for immunostaining and microscopy, over guidelines for best-practice tissue collection by LMD, toward an optimal nanoflow dia-PASEF LC-MS scheme for highest-sensitivity MS-based proteomics. For the first time, we provide evidence that robust ultra-low-input proteomics (few- or even single-excised cells) from FFPE tissue slices is achievable and fully compatible with conventional AR and four-marker IF imaging protocols, paving the way for higher-plex IF combinations in the near future. At the same time, our flexible 384-well design makes the protocol easily adaptable to robotic workflows for further protocol automation. We find that low microliter volumes ($\sim$2 μL) for sample preparation are sufficient to achieve true single-cell sensitivity, which makes this workflow principally accessible to any laboratory. On the Bruker timsTOF SCP, this allowed us to quantify up to 2,000 proteins from single hepatocyte contours and around 5,000 proteins from 50-cell regions, demonstrating excellent workflow scalability. On the timsTOF Pro2 instrument, this translated to nearly 1,000 high-confidence proteins from 1 to 2 hepatocytes ($\sim$1,500 μm^2 regions) and a somewhat similar proteome depth for the 50-cell samples, further emphasizing the broad applicability of our protocol. The data from single-excised hepatocytes also revealed that a $\sim$25-μm spatial resolution is principally achievable for tissue types such as liver, on par with the spatial resolution of state-of-the-art spatial transcriptomics.[41,42] These data also show that highly abundant housekeeping proteins, such as ribosomal proteins, are quantifiable at a depth and precision similar to much higher sampling amounts, for example, from large-scale bulk measurements. Our data also show that many metabolic pathways (e.g., TCA cycle or fatty acid metabolism) are likewise amenable to single-cell-based proteomic analysis, as also shown recently.[18] Notably, housekeeping related proteins are often poorly correlated to mRNA abundances,[10,43] making these pathways particularly attractive for early applications of single-cell-based tissue proteomics.

We also provide guidance on how to empirically determine the optimal tissue amount to achieve the highest proteome coverage while avoiding instrument overloading, which is particularly pertinent for single-cell sensitivity MS setups. We introduce a simple normalization strategy using the extracted MS2 signal directly from the search results to find the optimal tissue amount for literally any low-input spatial tissue proteomics experiment. This strategy can also be applied to normalize qualitative sample differences, which are often observed for archival specimens due to, for example, varying archival times.[31] The importance of this can be illustrated through the eyes of a pathologist. For the histomolecular analysis of cancer progression states, samples are typically distributed over several FFPE tissue blocks (for example pre-cancer, primary tumor, and metastasis), potentially collected over many years. Survey experiments, as outlined here using the example of human tonsil tissue, could guide the best sampling strategy for deep and reproducible LMD-assisted proteomics.

However, although the latest generation MS instruments feature excellent sensitivity when combined with optimized sample preparation workflows, such as the one introduced here, MS throughput is still a major bottleneck for tissue sections that are often larger than 1 cm^2. Isobaric or non-isobaric multiplexing strategies generally offer good alternatives to label-free based methods further increasing sample throughput,[44–46] but they are still not sufficient to deal with this tremendous analytical bottleneck. One can estimate that for the profiling of one 1 × 1 cm tissue section, sampled by non-overlapping 50 × 50 μm squares, amounts, which we here show in liver and tonsil tissues to be amenable to the reproducible quantification of 2,000–3,000 proteins, 40,000 measurements would be required for gridded whole-slide sampling schemes. This is far beyond the reach of current LC-MS setups, which typically analyze 20–50 tissue proteomes per day. Instead, the integration of whole-slide IF imaging for detailed cell and cellular neighborhood phenotyping allows us to prioritize cells and ROIs subjected to global proteome analysis, thereby offering a powerful, cost-effective, and accessible spatial profiling strategy. We illustrate this in tonsil tissue, where we use four-marker whole-slide IF imaging to guide the sampling of over 140 microregions per single batch, covering naive and activated B cell niches, interfollicular T cell zones and squamous cell EP. From only $\sim$63 × 63 μm regions (4,000 μm^2 regions), we quantified spatially resolved proteomes of activated GC niches, illuminating sites of clonal B cell expansion and somatic hypermutation. Intriguingly, based on these 1,500–3,000 protein measurements per microregion, we quantified key players of immune cell function including cytokines and transcriptional regulators, emphasizing the power of our "biological fractionation" strategy to dig deep into the cell type and spatially resolved tissue proteome.

In conclusion, we here provide a scalable and optimized framework for MS-based spatial tissue proteomics of ultra-low-input archival specimens combining high-content imaging, LMD, and ultrasensitive MS.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2023.10.003.

ACKNOWLEDGMENTS

We would like to thank our colleagues at the Max Delbrück Center (MDC) for their support and fruitful discussions. In particular, we thank Ulrike Stein for her support to perform the mouse liver experiments. We thank Philipp Mertins and Simon Haas for their critical feedback on the manuscript and Christian Sommer for MS support. We also thank Andreas Mund (Center for Protein Research, University of Copenhagen) and Lisa Schweizer (MPI of Biochemistry, Munich) for their input and feedback on the LMD membrane comparisons. We thank Simon Schallenberg (Charité Pathology, Berlin) for his help with the tonsil tissue experiments. Furthermore, we acknowledge the MDC technology platforms "Proteomics" and "Advanced Light Microscopy" for their great support. All authors acknowledge support by the Federal Ministry of Education and Research (BMBF), as part of the National Research Initiatives for Mass Spectrometry in Systems Medicine, under grant agreement No. 161L0222.

AUTHOR CONTRIBUTIONS

Conceptualization, A.M. and F.C.; methodology, A.M., D.Q., J.K., J.N., and F.C.; experiments, A.M., D.Q., S.F., J.K., and F.C.; data analysis, A.M. and F.C.; figures, A.M. and F.C.; supervision, F.C.; funding acquisition, F.C.; writing the original draft, F.C. All the authors reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 1, 2023
Revised: August 3, 2023
Accepted: October 6, 2023
Published: October 30, 2023

REFERENCES

1. Palla, G., Fischer, D.S., Regev, A., and Theis, F.J. (2022). Spatial components of molecular tissue biology. Nat. Biotechnol. *40*, 308–318. https://doi.org/10.1038/s41587-021-01182-1.
2. Casasent, A.K., Schalck, A., Gao, R., Sei, E., Long, A., Pangburn, W., Casasent, T., Meric-Bernstam, F., Edgerton, M.E., and Navin, N.E. (2018). Multiclonal invasion in breast tumors identified by topographic single cell sequencing. Cell *172*, 205–217.e12. https://doi.org/10.1016/j.cell.2017.12.007.
3. Deng, Y., Bartosovic, M., Kukanja, P., Zhang, D., Liu, Y., Su, G., Enninful, A., Bai, Z., Castelo-Branco, G., and Fan, R. (2022). Spatial-CUT&Tag: spatially resolved chromatin modification profiling at the cellular level. Science *375*, 681–686. https://doi.org/10.1126/science.abg7216.
4. Marx, V. (2021). Method of the Year: spatially resolved transcriptomics. Nat. Methods *18*, 9–14. https://doi.org/10.1038/s41592-020-01033-y.
5. Longo, S.K., Guo, M.G., Ji, A.L., and Khavari, P.A. (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat. Rev. Genet. *22*, 627–644. https://doi.org/10.1038/s41576-021-00370-8.
6. He, S., Bhatt, R., Brown, C., Brown, E.A., Buhr, D.L., Chantranuvatana, K., Danaher, P., Dunaway, D., Garrison, R.G., Geiss, G., et al. (2022). High-plex multiomic analysis in FFPE at subcellular level by spatial. Mol. Imaging *40*, 1794–1806. https://doi.org/10.1101/2021.11.03.467020.
7. Baharlou, H., Canete, N.P., Cunningham, A.L., Harman, A.N., and Patrick, E. (2019). Mass cytometry imaging for the study of human diseases—applications and data analysis strategies. Front. Immunol. *10*, 2657. https://doi.org/10.3389/fimmu.2019.02657.
8. Schürch, C.M., Bhate, S.S., Barlow, G.L., Phillips, D.J., Noti, L., Zlobec, I., Chu, P., Black, S., Demeter, J., McIlwain, D.R., et al. (2020). Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. Cell *182*, 1341–1359.e19. https://doi.org/10.1016/j.cell.2020.07.005.
9. Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. Nature *537*, 347–355. https://doi.org/10.1038/nature19949.
10. Buccitelli, C., and Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene expression control. Nat. Rev. Genet. *21*, 630–644. https://doi.org/10.1038/s41576-020-0258-4.
11. Jackson, H.W., Fischer, J.R., Zanotelli, V.R.T., Ali, H.R., Mechera, R., Soysal, S.D., Moch, H., Muenst, S., Varga, Z., Weber, W.P., et al. (2020). The single-cell pathology landscape of breast cancer. Nature *578*, 615–620. https://doi.org/10.1038/s41586-019-1876-x.
12. Lin, J.R., Izar, B., Wang, S., Yapp, C., Mei, S., Shah, P.M., Santagata, S., and Sorger, P.K. (2018). Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. eLife *7*, 1–46. https://doi.org/10.7554/eLife.31657.
13. Sharma, K., D'Souza, R.C.J., Tyanova, S., Schaab, C., Wiśniewski, J.R., Cox, J., and Mann, M. (2014). Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and ser/Thr-based signaling. Cell Rep. *8*, 1583–1594. https://doi.org/10.1016/j.celrep.2014.07.036.
14. Sinitcyn, P., Richards, A.L., Weatheritt, R.J., Brademan, D.R., Marx, H., Shishkova, E., Meyer, J.G., Hebert, A.S., Westphall, M.S., Blencowe, B.J., et al. (2023). Global detection of human variants and isoforms by deep proteome sequencing. Nat. Biotechnol. https://doi.org/10.1038/s41587-023-01714-x.
15. Mund, A., Coscia, F., Kriston, A., Hollandi, R., Kovács, F., Brunner, A.-D., Migh, E., Schweizer, L., Santos, A., Bzorek, M., et al. (2022). Deep Visual proteomics defines single-cell identity and heterogeneity. Nat. Biotechnol. *40*, 1231–1240. https://doi.org/10.1038/s41587-022-01302-5.
16. Hollandi, R., Szkalisity, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., Grexa, I., Molnar, J., Balind, A., Gorbe, M., et al. (2020). nucleAIzer: A parameter-free deep learning framework for nucleus segmentation using image style transfer. Cell Syst. *10*, 453–458.e6. https://doi.org/10.1016/j.cels.2020.04.003.
17. Brunner, A.D., Thielert, M., Vasilopoulou, C., Ammar, C., Coscia, F., Mund, A., Hoerning, O.B., Bache, N., Apalategui, A., Lubeck, M., et al. (2022). Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. Mol. Syst. Biol. *18*, e10798. https://doi.org/10.15252/msb.202110798.
18. Rosenberger, F.A., Thielert, M., Strauss, M.T., Ammar, C., Mädler, C., Schweizer, L., Metousis, A., Skowronek, P., Wahle, M., Semenova, A., et al. (2023). Spatial single-cell mass spectrometry defines zonation of the

hepatocyte proteome author list. Nat Methods *20*, 1530–1536. https://doi.org/10.1038/s41592-023-02007.

19. Wiśniewski, J.R., Ostasiewicz, P., and Mann, M. (2011). High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. J. Proteome Res. *10*, 3040–3049. https://doi.org/10.1021/pr200019m.

20. Piehowski, P.D., Zhu, Y., Bramer, L.M., Stratton, K.G., Zhao, R., Orton, D.J., Moore, R.J., Yuan, J., Mitchell, H.D., Gao, Y., et al. (2020). Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100-μm spatial resolution. Nat. Commun. *11*, 8. https://doi.org/10.1038/s41467-019-13858-z.

21. Eckert, M.A., Coscia, F., Chryplewicz, A., Chang, J.W., Hernandez, K.M., Pan, S., Tienda, S.M., Nahotko, D.A., Li, G., Blaženović, I., et al. (2019). Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. Nature *569*, 723–728. https://doi.org/10.1038/s41586-019-1173-8.

22. Azimifar, S.B., Nagaraj, N., Cox, J., and Mann, M. (2014). Cell-type-resolved quantitative proteomics of murine liver. Cell Metab. *20*, 1076–1087. https://doi.org/10.1016/j.cmet.2014.11.002.

23. Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A., and Liotta, L.A. (1996). Laser capture microdissection. Science *274*, 998–1001. https://doi.org/10.1126/science.274.5289.998.

24. Ben-Moshe, S., and Itzkovitz, S. (2019). Spatial heterogeneity in the mammalian liver. Nat. Rev. Gastroenterol. Hepatol. *16*, 395–410. https://doi.org/10.1038/s41575-019-0134-x.

25. Nordmann, T.M., Schweizer, L., Metousis, A., and Thielert, M. (2023). Rahbek-gjerdrum, L.M., Stadler, P, and Bzorek, M. A standardized and reproducible workflow for membrane glass slides in routine histology and spatial proteomics. https://doi.org/10.1101/2023.02.20.529255.

26. Kawashima, Y., Kodera, Y., Singh, A., Matsumoto, M., and Matsumoto, H. (2014). Efficient extraction of proteins from formalin-fixed paraffin-embedded tissues requires higher concentration of tris (hydroxymethyl) aminomethane. Clin. Proteomics *11*, 4. https://doi.org/10.1186/1559-0275-11-4.

27. Meier, F., Brunner, A.D., Frank, M., Ha, A., Bludau, I., Voytik, E., Kaspar-Schoenefeld, S., Lubeck, M., Raether, O., Bache, N., et al. (2020). diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. Nat. Methods *17*, 1229–1236. https://doi.org/10.1038/s41592-020-00998-0.

28. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat. Methods *17*, 41–44. https://doi.org/10.1038/s41592-019-0638-x.

29. Hammad, S., Hoehme, S., Friebel, A., Von Recklinghausen, I., Othman, A., Begher-Tibbe, B., Reif, R., Godoy, P., Johann, T., Vartak, A., et al. (2014). Protocols for staining of bile canalicular and sinusoidal networks of human, mouse and pig livers, three-dimensional reconstruction and quantification of tissue microarchitecture by image processing and analysis. Arch. Toxicol. *88*, 1161–1183. https://doi.org/10.1007/S00204-014-1243-5/FIGURES/9.

30. Guo, T., Kouvonen, P., Koh, C.C., Gillet, L.C., Wolski, W.E., Röst, H.L., Rosenberger, G., Collins, B.C., Blum, L.C., Gillessen, S., et al. (2015). Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. Nat. Med. *21*, 407–413. https://doi.org/10.1038/nm.3807.

31. Coscia, F., Doll, S., Bech, J.M., Schweizer, L., Mund, A., Lengyel, E., Lindebjerg, J., Madsen, G.I., Moreira, J.M., and Mann, M. (2020). A streamlined mass spectrometry–based proteomics workflow for large-scale FFPE tissue analysis. J. Pathol. *251*, 100–112. https://doi.org/10.1002/path.5420.

32. Cunningham, R.P., and Porat-Shliom, N. (2021). Liver zonation – revisiting old questions with new technologies. Front. Physiol. *12*, 1433. https://doi.org/10.3389/FPHYS.2021.732929/BIBTEX.

33. Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Res. *38*, D750–D753. https://doi.org/10.1093/NAR/GKP889.

34. Nwosu, A.J., Misal, S.A., Truong, T., Carson, R.H., Webber, K.G.I., Axtell, N.B., Liang, Y., Johnston, S.M., Virgin, K.L., Smith, E.G., et al. (2022). In-Depth mass spectrometry-Based proteomics of formalin-Fixed, Paraffin-Embedded Tissues with a Spatial Resolution of 50–200 μm. J. Proteome Res. *21*, 2237–2245. https://doi.org/10.1021/acs.jproteome.2c00409.

35. Balgley, B.M., Guo, T., Zhao, K., Fang, X., Tavassoli, F.A., and Lee, C.S. (2009). Evaluation of archival time on shotgun proteomics of formalin-fixed and paraffin-embedded tissues. J. Proteome Res. *8*, 917–925. https://doi.org/10.1021/pr800503u.

36. Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. *15*, e8503. https://doi.org/10.15252/MSB.20188503.

37. Massoni-Badosa, R., Soler-Vila, P., Aguilar-Fernández, S., Nieto, J.C., Elosua-Bayes, M., Marchese, D., Kulis, M., Vilas-Zornoza, A., Bühler, M.M., Rashmi, S., et al. (2022). An atlas of cells in the human tonsil. https://doi.org/10.1101/2022.06.24.497299.

38. De Silva, N.S., and Klein, U. (2015). Dynamics of B cells in germinal centres. Nat. Rev. Immunol. *15*, 137–148. https://doi.org/10.1038/nri3804.

39. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., et al. (2017). QuPath: open source software for digital pathology image analysis. Sci. Rep. *7*, 16878. https://doi.org/10.1038/s41598-017-17204-5.

40. Victora, G.D., and Nussenzweig, M.C. (2012). Germinal centers. Annu. Rev. Immunol. *30*, 429–457. https://doi.org/10.1146/annurev-immunol-020711-075032.

41. Lewis, S.M., Asselin-Labat, M.L., Nguyen, Q., Berthelet, J., Tan, X., Wimmer, V.C., Merino, D., Rogers, K.L., and Naik, S.H. (2021). Spatial omics and multiplexed imaging to explore cancer biology. Nat. Methods *18*, 997–1012. https://doi.org/10.1038/s41592-021-01203-6.

42. Merritt, C.R., Ong, G.T., Church, S.E., Barker, K., Danaher, P., Geiss, G., Hoang, M., Jung, J., Liang, Y., McKay-Fleisch, J., et al. (2020). Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nat. Biotechnol. *38*, 586–599. https://doi.org/10.1038/s41587-020-0472-9.

43. Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., et al. (2014). Proteogenomic characterization of human colon and rectal cancer. Nature *513*, 382–387. https://doi.org/10.1038/nature13438.

44. Derks, J., Leduc, A., Huffman, R.G., Specht, H., Ralser, M., Demichev, V., and Slavov, N. (2021). Increasing the throughput of sensitive proteomics by plexDIA. Nat. Biotechnol. *41*, 50–59.

45. Friedrich, C., Schallenberg, S., Kirchner, M., Ziehm, M., Niquet, S., Haji, M., Beier, C., Neudecker, J., Klauschen, F., and Mertins, P. (2021). Comprehensive micro-scaled proteome and phosphoproteome characterization of archived retrospective cancer repositories. Nat. Commun. *12*, 3576. https://doi.org/10.1038/s41467-021-23855-w.

46. Thielert, M., Itang, C., Ammar, C., Schober, F., Bludau, I., Skowronek, P., Wahle, M., Zeng, W.-F., Zhou, X.-X., Brunner, A.-D., et al. (2022). Robust dimethyl-based multiplex-DIA workflow doubles single-cell proteome depth via a reference channel Author list. https://doi.org/10.15252/msb.202211503.

47. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450. https://doi.org/10.1093/nar/gky1106.

48. Schmacke, N.A., Mädler, S.C., Wallmann, G., Metousis, A., Bérouti, M., Harz, H., Leonhardt, H., Mann, M., and Hornung, V. (2023). SPARCS, a platform for genome-scale CRISPR screening for spatial cellular phenotypes. https://doi.org/10.1101/2023.06.01.542416.

49. Frankenfield, A.M., Ni, J., Ahmed, M., and Hao, L. (2022). Protein contaminants matter: building universal protein contaminant libraries for DDA and DIA proteomics. J. Proteome Res. *21*, 2104–2113. https://doi.org/10.1021/acs.jproteome.2c00145.

50. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740. https://doi.org/10.1038/nmeth.3901.

51. Cox, J., and Mann, M. (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinformatics *13*, S12. https://doi.org/10.1186/1471-2105-13-S16-S12.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Recombinant Anti-Sodium Potassium ATPase antibody [EP1845Y]	Abcam	Cat#ab76020; RRID: AB_1310695
Recombinant Alexa Fluor® 647 Anti-CD3D antibody [EP4426]	Abcam	Cat#ab198937; RRID: AB_2889190
CD20 Monoclonal Antibody (L26), Alexa Fluor™ 488, eBioscience	Thermo Fisher Scientific	Cat# 53-0202-80, RRID: AB_10734357
Pan Cytokeratin Monoclonal Antibody (AE1/AE3), eFluor™ 570, eBioscience	Thermo Fisher Scientific	Cat# 41-9003-80, RRID: AB_11217482
Biological samples		
Murine liver FFPE tissue from C57BL/6 mice	The Jackson Laboratory	https://www.jax.org/strain/000664
Human tonsil FFPE tissue	Institute of Pathology at Charite University Hospital	https://pathologie-ccm.charite.de/
Chemicals, peptides, and recombinant proteins		
n-Dodecyl-beta-Maltoside (DDM)	Sigma-Aldrich	Cat# D4641-500MG
Endoproteinase Lys-C	Promega	Cat# VA1170
Proteomics grade modified trypsin	Promega	Cat# V5117
Tris(2-carboxyethyl)phosphine hydrochloride	Sigma-Aldrich	Cat# C4706-2G
Acetonitrile (ACN) HPLC-grade	VWR	Cat# 83640.290
Isopropanol (ISO)	Sigma-Aldrich	Cat# 1070222511
Triethylammonium bicarbonate pH 8.5 (TEAB)	Merck	T7408-100ML
Formic acid	Merck Millipore	Cat# 1.00264.1000
Trifluoroacetic acid	Sigma-Aldrich	Cat# 96924-250ML-F
2-chloroacetamide	Sigma-Aldrich	Cat# C0267-100G
EnVision FLEX Target Retrieval Solution High pH (50X)	Agilent Dako	Cat# K8004
Microscopy Neo-Clear	Sigma-Aldrich	Cat# 1.09843.5000
Odyssey Blocking Buffer	LI-COR Biosciences	Cat# 927-70001
Prolong Diamond antifade mounting medium	Invitrogen	Cat# P36961
Aqua Poly Mount	Polysciences Europe GmbH	Cat# 18606-20
Pepsin solution for antigen retrieval	Agilent Dako	Cat# S3002
Deposited data		
Mass spectrometric raw files	This paper	ProteomeXchange: PXD042367
List of hepatocyte specific markers (Azimifar et al.[22])	10.1016/j.cmet.2014.11.002	1-s2.0-S1550413114004999-mmc5.xlsx
Software and algorithms		
DIA-NN version 1.8.1	Demichev et al.[28]	https://github.com/vdemichev/DiaNN
R version 4.2.2	The R Project for Statistical Computing	https://www.r-project.org/
Perseus version 1.6.15.0	Tyanova et al.[50]	https://maxquant.net/perseus/
Leica Laser Microdissection software version 8.3.0.08259	Leica Microsystems	https://www.leica-microsystems.com/products/microscope-software/p/leica-lmd-software/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Zeiss ZEN version 3.7	Carl Zeiss AG	https://www.zeiss.com/microscopy/de/produkte/software/zeiss-zen.html
LAS X software version 3.7.524914	Leica Microsystems	https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls/
Qupath version 0.4.3	Bankhead et al.[39]	https://qupath.github.io/
Biological Image Analysis Software (BIAS)	Single Cell Technologies	https://single-cell-technologies.com/bias-2/BioStudies Archive accession number S-BSST820
Bruker Compass Data Analysis Software version 6.0	Bruker Daltonik GmbH	https://www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software.html
Qupath_to_LMD function: Contour export from Qupath to the Leica LMD7	This paper	https://doi.org/10.5281/zenodo.8414787
Other		
C18 Evotips (EV2013, Evotip Pure, Evosep)	Evosep Biosystems	https://www.evosep.com/evotip/
96-well plate	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/de/en/AB1300
384-well low-binding plate	Eppendorf	Cat# 0030129547
Super PAP-pen liquid blocker mini	Science Services	Cat# N71312-N
Cover glass	Corning	Cat# CLS2980223, #1.5
PPS frame slides	Leica	Cat# 11600294
PEN glass slides	Carl Zeiss	Cat# 15350731

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Fabian Coscia (fabian.coscia@mdc-berlin.de).

Materials availability

This study did not generate new materials.

Data and code availability

- The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner[47] with the dataset identifier PXD042367. All data reported in this paper will be shared by the lead contact upon request.
- The source code for processing the shapes used for laser microdissection has been deposited and is freely available at github.com/CosciaLab/Qupath_to_LMD. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse experiments and organ harvesting

For the mouse liver proteome experiments, 6- 8 weeks old female C57BL/6 mice from Jackson Laboratory were used. C57BL/6 mice were housed in individually ventilated cages in a specific pathogen-free mouse facility at the Max-Delbrück Center for Molecular Medicine (Berlin, Germany).

For liver excision, anesthetized mice were sacrificed by cervical dislocation, and the livers were removed, rinsed twice in ice-cold PBS, and transferred to 4% formaldehyde solution for fixation (fixation for at least 24h to 48h). Thereafter, livers were paraffin-embedded for further histological analyses. The animal experiments were performed in accordance with the United Kingdom Coordinated Committee on Cancer Research (UKCCR) guidelines and were approved by local governmental authorities (Landesamt für Gesundheit und Soziales Berlin, Germany).

Human tissue samples

Tonsil tissues were obtained from two female patients aged 34 and 36. Both presented with recurrent chronic tonsillitis and underwent bilateral tonsillectomy. The pathological examination revealed hypertrophy and hyperplasia of the lymphoid follicles as well as enlarged germinal centers. Furthermore, an abundance of collagen fibers within the stroma could be observed. There were no signs of active inflammation or malignancy.

Resected tonsils were fixed in 10% buffered formalin before gross processing. After overnight fixation, the specimens were weighed, measured, and macroscopically evaluated. Afterward, the specimens were cut into 5-mm-thick slices. Two representative slices were embedded in paraffin. The embedded tissue blocks were cut into 2-μm-thick sections, and stained with hematoxylin and eosin for histological examination. The tissue blocks were then stored at room temperature at the archive of the Institute of Pathology at Charité University Hospital, Campus Mitte. The study was performed according to the ethical principles for medical research of the Declaration of Helsinki and approval was obtained from the Ethics Committee of the Charité University Medical Department in Berlin (EA1/222/21).

METHOD DETAILS

Hematoxylin-eosin staining

Briefly, PEN glass slides (Carl Zeiss, 15350731) were treated with UV light for 1 hour. PPS frame slides (Leica, 11600294) were used directly for the next steps. FFPE tissue sections were cut with a microtome (5 μm or 10 μm-thick), air dried at 37 °C overnight and heated at 60°C for 10 minutes to facilitate better tissue adhesion. Next, tissue sections were deparaffinized by washing 2x5 minutes in Neo Clear (Sigma Aldrich, 1.09483.5000), followed by a series of 99%, 80% and 70% ethanol for 2 minutes, respectively, and rehydrated by immersing in milliQ water three times. Then, slides were stained in Mayer's hematoxylin for three minutes and immersed in tap water for another ten minutes, rinsed in milliQ water and stained with eosin for 30 seconds. Subsequently, the slides were dehydrated by submerging in 70%, 80%, and 99% ethanol serially. Samples were finally air-dried and stored at RT until imaging. Before imaging, a cover glass (Corning, CLS2980223, #1.5) was mounted with Aqua Poly Mount medium (Polysciences Europe GmbH, 18606-20).

Immunofluorescence staining

Following tissue sectioning, mounting on LMD slides and de-paraffinization (described above), two epitope retrieval methods were compared: heat-induced (HIER) and protease (pepsin)-induced (PIER, 10 min) epitope retrieval. HIER was done by submerging in EnVision FLEX Target Retrieval Solution High pH solution (diluted to 1X) (Agilent Dako, cat.no. K8004) and heating in a steamer at 95°C for 20 min, and subsequently cooled down in a pre-heated PBS buffer at room temperature for 30 min. Odyssey Blocking Buffer (LI-COR BioScience, 927-70001) was used for blocking in a humidified chamber for 30 min at room temperature. PIER was done by applying pepsin (Agilent Dako, cat.no. S3002) on the tissue slide for 5 minutes at 37°C, and immersed in PBS solution to stop the reaction. The slides were washed two times with PBS and air dried before antibody staining.

For murine liver tissue stains, the primary antibody targeting Na/K-ATPase (stock concentration 0.563 mg/ml, dilution 1:100, Abcam, ab76020) was diluted in Odyssey Blocking Buffer and incubated overnight at 4 °C in a humidified chamber. Next, tissue specimens were washed 3x in PBS and secondary antibodies for the visualization of Na/K-ATPase (Alexa Fluor 488 donkey anti-rabbit, stock concentration 2 mg/ml, dilution 1:250, A32790, Invitrogen) were diluted in Odyssey Blocking Buffer and applied for 1 hour at room temperature in the dark. After staining, slides were washed 3x in PBS, counterstained by Hoechst (dilution 1:1000 in PBS, Thermo Fisher Scientific, #62249) for 10 minutes and followed by three washes in PBS and two washes in milliQ. Before imaging, a cover glass (#1.5) was mounted using ProLong™ Diamond anti-fade mounting medium (Thermo Fisher Scientific, P36961).

For tonsil tissue stains following tissue sectioning, mounting on LMD slides and de-paraffinization, the tissues were subjected to heat-induced epitope retrieval as described above. Odyssey Blocking Buffer was used for blocking in a humidified chamber for 30 min at room temperature. Next, conjugated primary antibodies targeting CD20 (stock concentration 0.5mg/ml, dilution 1:50, Thermofisher, 53-0202-80, Alexa Fluor 488), CD3 (stock concentration 0.5 mg/ml, dilution 1:100, Abcam, ab198937, Alexa Fluor 647), and pan-cytokeratin (stock concentration 0.2 mg/ml, dilution 1:100, Thermofisher, 41-9003-80, eFluor 570) were diluted in Odyssey Blocking Buffer (and incubated overnight at 4 °C in a humidified chamber. Tissue specimens were washed 4x in PBS, counterstained by Hoechst (dilution 1:1000 in PBS, Thermo Fisher Scientific, 62249) for 10 minutes, washed 4x in PBS and 2x in milliQ water. Subsequently, the slides were dehydrated by submerging them in 70%, 80%, and 99% ethanol serially. Before imaging, a cover glass was mounted with ProLong™ Diamond anti-fade mounting medium.

High-resolution microscopy

Images of immunofluorescence-labeled tonsil tissue sections were acquired using an Axioscan 7 system (Zeiss), equipped with wide-field optics, a Plan-A photochromat 10x/0.45 M27 objective and a quadruple-band filter set for Alexa fluorescent dyes. The wide-field acquisition was performed using the Colibri 7 LED light source and an AxioCam 712m camera. Images were obtained automatically with Zeiss ZEN 3.7 (blue edition) at non-saturating conditions (16-bit dynamic range).

Images of immunofluorescently-labeled and hematoxylin-eosin stained murine liver tissue sections were acquired on the Leica LMD7 system using the LAS X software (version 3.7.524914, Leica Microsystems), an HC PL FLUOTAR 10x/0.32 DRY objective and the DF7000T camera. The microscope was equipped with the following filter sets: LMD-Dapi, LMD-Cy3, LMD - Alexa594,

LMD-YFP, LMD-CY5, LMD-BGR, GFP. After imaging, the cover glass was finally removed through gentle agitation of the slide in PBS. After one dip in milliQ water for salt removal, the slide was air-dried and stored at 4°C until laser microdissection.

Image analysis and contour export for laser microdissection

Image analysis was performed in QuPath (version 0.4.3) and BIAS (BioStudies Archive accession number S-BSST820). Annotations of different regions of interest were manually created in QuPath after image analysis. The assignment of three reference points (x-y coordinates) is required for precise contour transfer between the screening and laser microdissection microscopes. Contours and reference points were exported in a geojson format and translated into the .XML format compatible with the Leica LMD7 software. The code for processing the shapes is available at github.com/CosciaLab/Qupath_to_LMD, it uses geopandas (Version 0.12.2) and py-lmd[48] (Version 1.0.0).

Laser microdissection

We used the Leica LMD 7 system and Leica Laser Microdissection V 8.3.0.08259 software for the collection of tissue contours. Depending on the contour size, tissue was cut with a 20x or 63x objective in fluorescence or brightfield mode. The following laser settings were used for the 20x objective (HC PL FL L 20x/0.40 CORR): power 56, aperture 1, speed 15, middle pulse count 1, final pulse -1, head current 37 – 45% (depending on tissue type and section thickness), pulse frequency 801 and offset 101. For the 63x objective (HC PL FLUOTAR L 63x/0.70 CORR XT), the settings were: power 59, aperture 1, speed 50, middle pulse count 1, final pulse -1, head current 45-50%, pulse frequency 3000, offset 101.

Contours were cut and sorted into a low-binding 384-well plate (Eppendorf 0030129547) configured over the 'universal holder' function with one empty well between samples.

Sample preparation for LC-MS analysis

A detailed sample preparation protocol is provided in the supplemental information **(**Methods S1**)**. Tissue samples were collected by manual cutting or by automated cutting after contour import into low-binding 384-well plates. For mouse liver tissue samples (5- or 10-μm-thick sections cut with a microtome), regions of 600 μm^2 - 100,000 μm^2 were collected. To concentrate tissue pieces at the bottom of each well after LMD collection, 15 μl of acetonitrile was added to each well, briefly vortexed and vacuum dried (15min at 60°C). Another well inspection is recommended before proteomics sample preparation to ensure high collection efficiency.

We tested three different protocols, DDM-based, ACN-based and a combination of DDM and ACN. The lysis buffer for the DDM-based protocol consisted of 0.1% DDM, 5mM TCEP, 20mM CAA and 0.1M TEAB in water. 2μl of lysis buffer was added to each sample well using the MANTIS Liquid Dispenser (Formulatrix, V3.3 ACC RFID, software version 4.7.5) and the high-volume diaphragm chips (Formulatrix, cat.no. 233128). The plate was closed with a PCR ComfortLid (Hamilton), and heated at 95°C for 60 minutes. Then, samples were shortly cooled down, and 1μl of LysC was added (prediluted in water to 2 ng/μl) and digested for minimum 2 hours at 37°C in the thermal cycler (50°C lid temperature). Subsequently, 1μl of trypsin was added (prediluted in water to 2 ng/μl) and incubated overnight at 37 °C in the thermal cycler. The next day, digestion was stopped by adding trifluoroacetic acid (TFA, final concentration 1% v/v), and samples were vacuum dried before peptide clean-up.

For the ACN-based protocol, the lysis buffer consisted of 5mM TCEP, 20mM CAA, 0.1M TEAB diluted in water. 2 μl of lysis buffer was added to each sample well, the plate was closed with PCR ComfortLid, and heated at 95°C for 60 minutes in a thermal cycler (Bio-Rad S1000 with 384-well reaction module) at a constant lid temperature of 110°C. Next, 1 ul of 100% ACN was added to each well, and the plate was incubated for an additional 60 min in a thermal cycler at 75°C at a constant lid temperature of 110°C. Then, samples were shortly cooled down, and 1 μl of LysC was added (prediluted with 0.1M TEAB, 30% ACN water to 2 ng/μl) and digested for 2 hours at 37°C in the thermal cycler (50°C lid temperature). Subsequently, 1 μl of trypsin was added (prediluted with 0.1M TEAB, 10% ACN water to 2 ng/μl) and incubated overnight at 37 °C in the thermal cycler. The next day, digestion was stopped by adding trifluoroacetic acid (TFA, final concentration 1% v/v), and samples were vacuum dried before peptide clean-up.

For the combined DDM/ACN-based protocol, the lysis buffer for the DDM-based protocol consisted of 0.1% DDM, 5mM TCEP, 20mM CAA and 0.1M TEAB in water. 2μl of lysis buffer was added to each sample well using the MANTIS Liquid Dispenser and the high-volume diaphragm chips. The plate was closed with a PCR ComfortLid, and heated at 95°C for 60 minutes. Then, samples were shortly cooled down, and 1 μl of LysC was added (2 ng/μl in 0.1M TEAB [pH 8.5] and 30% ACN in milliQ water) and digested for minimum 2 hours at 37°C in the thermal cycler (50°C lid temperature). Subsequently, 1 μl of trypsin was added (2 ng/μl containing 10% ACN and 0.1M TEAB [pH 8.5] in milliQ water.) and incubated overnight at 37 °C in the thermal cycler. The next day, digestion was stopped by adding trifluoroacetic acid (TFA, final concentration 1% v/v), and samples were vacuum dried before peptide clean-up.

Peptide clean-up with C-18 tips

Evotip (Evosep, Odense, Denmark) based peptide clean-up was performed as recommended by the manufacturer. Briefly, 20 ul of buffer B (99.9% ACN, 0.1% FA) was added to each C-18 tip (EV2013, Evotip Pure, Evosep) and centrifuged at 700 rpm for 1 minute. Then, 20 ul of buffer A (99.9% water, 0.1% FA) was added from the top of each C-18 tip, activated in isopropanol for 20 seconds and centrifuged again at 700 rpm for 1 minute. Digested tissue samples were then loaded onto Evotips, washed once with 20 ul buffer A and finally eluted with 20 ul buffer B to a 96-well plate (Thermo Fisher Scientific, AB1300), and vacuum dried (15 min at 60°C).

Samples were stored at −20 °C until liquid chromatography–mass spectrometry (LC–MS) analysis. For LC-MS analysis, 4.2 μl of MS loading buffer (3% acetonitrile, 0.1% TFA in water) was added, the plate was vortexed for 10 seconds and centrifuged for 1 minute at 700g. 4 μl were finally injected into the mass spectrometer.

Liquid chromatography–mass spectrometry (LC – MS) analysis

LC-MS analysis was performed with an EASYnLC-1200 system (Thermo Fisher Scientific) connected to a trapped ion mobility spectrometry quadruple time-of-flight mass spectrometer (timsTOF SCP and timsTOF Pro2, Bruker Daltonik) with a nano-electrospray ion source (CaptiveSpray, Bruker Daltonik). The autosampler was configured to pick samples from 384- and 96-well plates.

Peptides were loaded on a 20-cm home-packed HPLC column (75-μm inner diameter packed with 1.9-μm ReproSil-Pur C18-AQ silica beads, Dr. Maisch).

Peptides were separated using a linear gradient from 7-30% buffer B (0.1% formic acid and 90% ACN in LC-MS grade water) in 14 minutes, followed by an increase to 60% for 1 minute and a 1.5-minute wash in 90% buffer B at 250 nl min^{-1}. Buffer A consisted of 0.1% formic acid in LC-MS grade water. The total gradient length was 21 minutes. A column oven was used to keep the column temperature constant at 40°C.

For dia-PASEF analysis, we used a dia-PASEF method with 8 dia-PASEF scans separated into 3 ion mobility windows per scan covering a 400-1000 m/z range by 25 Th windows and an ion mobility range from 0.64 to 1.37 Vs cm^{-2}. The mass spectrometer was operated in high sensitivity mode, with an accumulation and ramp time at 100 ms, capillary voltage set to 1750V and the collision energy as a linear ramp from 20 eV at 1/K0 = 0.6 Vs cm^{-2} to 59 eV at 1/K0 = 1.6 Vs cm^{-2}. The collision energy was ramped linearly as a function of ion mobility from 59 eV at 1/K0 = 1.6 V s cm^{-2} to 20 eV at 1/K0 = 0.6 V s cm^{-2}.

Proteomics raw data analysis

We used DIA-NN[28] (version 1.8.1) for dia-PASEF raw file analysis and spectral library generation.

Spectral library generation

For spectral library generation to analyze dia-PASEF data, human and mouse FASTA files were downloaded from Uniprot (2022 release, UP000000589_10090 Mus Musculus, UP000005640_9606, downloaded on April 10th and April 8th 2022 respectively). DIA-NN in silico predicted libraries were generated by providing the human or mouse FASTA file and frequently found contaminants[49] (mouse + mouse tissue contaminants, or human + universal contaminants). Deep learning-based spectra, RTs and IMs prediction were enabled for the appropriate mass range of 300-1200 m/z. N-terminal M excision was enabled and cysteine carbamidomethylation was enabled as a fixed modification. A maximum of 2 missed cleavages was allowed, and the precursor charge set to 2 - 4. For the generation of project-specific refined libraries, in-silico-generated mouse and human libraries were used to search 20-50 raw files of high-sample amounts, such that the optimal total ion current was reached (see Figures 2 and 3). The refined murine liver library consisted of 68,006 precursors, 61,554 elution groups and 8,225 protein groups. The refined human tonsil library consisted of 47,999 precursors, 44,675 elution groups and 8,137 protein groups.

Search of dia-PASEF raw files with refined libraries

DIA-NN was operated in the default mode with minor adjustments. Briefly, MS1 and MS2 accuracies were set to 15.0, scan windows to 0 (assignment by DIA-NN), isotopologues were enabled, no MBR for project-specific DIA-NN-refined libraries, heuristic protein inference and no shared spectra. Enabling MBR for refined libraries improved data completeness of single-cell samples (Figure S2G). No difference was observed for higher input samples. Proteins were inferred from genes, neural network classifier was set to single-pass mode, quantification strategy as 'Robust LC (high precision)'. Cross-run normalization was set to 'RT-dependent', library generation as 'smart profiling', speed and Ram usage as 'optimal results'.

QUANTIFICATION AND STATISTICAL ANALYSIS

Proteomics data analysis was performed with Perseus[50] (version 1.6.15.0) and within the R environment (https://www.r-project.org/, version 4.2.2) with the following packages: ggplot2 (v3.4.2), FactoMineR (v2.8), factoextra (v 1.0.7.999), plotly (v4.10.1), reshape2 (v1.4.4), viridis (v0.6.3), UpSetR (v1.4.0). For differential expression analysis (t-test or ANOVA, Figures 3 and 4), data were filtered to keep only proteins with 70% non-missing data in at least one group. Missing values were imputed based on a normal distribution (width = 0.3, downshift = 1.8) before statistical testing. For multi-sample (ANOVA) or pairwise proteomic comparisons (two-sided unpaired t-test), a permutation-based FDR of 5% was applied to correct for multiple hypothesis testing. Principal component analysis was performed in R (see packages above). 1D pathway enrichment analysis[51] (Figure 4I) was done in Perseus based on the KEGG (https://www.genome.jp/kegg/) and Reactome Pathway Database (reactome.org), using a Benjamini-Hochberg FDR cut-off of 0.05. The minimum category size was set to 5.