nature communications

Article

https://doi.org/10.1038/s41467-024-53683-7

Spatial proteomics of single cells and organelles on tissue slides using filter-aided expansion proteomics

Received: 9 May 2024

Accepted: 21 October 2024

Published online: 30 October 2024

Check for updates

Zhen Dong[1,2,3,6], Wenhao Jiang[1,2,3,6], Chunlong Wu[1,2,3], Ting Chen[4], Jiayi Chen[1,2,3], Xuan Ding[1,2,3], Shu Zheng[4], Kiryl D. Piatkevich[2,3,5] ✉, Yi Zhu[1,2,3] ✉ & Tiannan Guo[1,2,3] ✉

Hydrogel-based tissue expansion combined with mass spectrometry (MS) offers an emerging spatial proteomics approach. Here, we present a filter-aided expansion proteomics (FAXP) strategy for spatial proteomics analysis of archived formalin-fixed paraffin-embedded (FFPE) specimens. Compared to our previous ProteomEx method, FAXP employed a customized tip device to enhance both the stability and throughput of sample preparation, thus guaranteeing the reproducibility and robustness of the workflow. FAXP achieved a 14.5-fold increase in volumetric resolution. It generated over 8 times higher peptide yield and a 255% rise in protein identifications while reducing sample preparation time by 50%. We also demonstrated the applicability of FAXP using human colorectal FFPE tissue samples. Furthermore, for the first time, we achieved bona fide single-subcellular proteomics under image guidance by integrating FAXP with laser capture microdissection.

Spatial proteomics can be achieved by affinity-based methods or mass spectrometry (MS)-based proteomics[1,2]. The latter enables comprehensive quantitative detection of proteins without the need for affinity reagents, the applicability and reliability of which largely rely on their specificity and stability. In addition, the antibody-based method requires intact reservation of the sample epitopes, posing a great challenge for clinical specimens. Two primary MS-based spatial proteomics methodologies have emerged. One concerns MS imaging (MSI)[3], utilizing matrix-assisted laser desorption-ionization (MALDI), which converts proteins or peptides from regions of interest (ROI) into gas-phase ions for subsequent MS analysis[4,5]. However, compared to the liquid chromatography (LC)-based fractionation coupled with electrospray ionization (ESI) MS method, MALDI-based MSI characterizes a limited number of peptides and proteins. Recently, with rapid improvement of the LC-MS instruments, even minute tissue samples dissected from laser capture microdissection (LCM) could be effectively analyzed[6–11]. The most prominent approach, deep visual proteomics (DVP), combines AI-driven image analysis, automated LCM, and high-sensitivity mass spectrometry[6]. Its extension to single-cell analysis (scDVP)[8] addresses the complexity of analyzing heterogeneous tissues at single-cell resolution. Coupled with other techniques, such as microfluidics[12,13], 3D-printed devices[14], or tissue clearing[15], LCM has achieved various applications in high-throughput and 3D proteome analysis.

Inspired by expansion microscopy[16–19], tissue expansion technology has recently been integrated with MS-based proteomics[20,21]. Coupled with antibodies, physical tissue expansion has recently been integrated with multiplexed ion beam imaging and imaging mass

[1]School of Medicine, Westlake University, Hangzhou, Zhejiang Province, China. [2]Westlake Center for Intelligent Proteomics, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang Province, China. [3]Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou, Zhejiang Province, China. [4]Cancer Institute (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education), The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, China. [5]School of Life Sciences, Westlake University, Hangzhou, Zhejiang Province, China. [6]These authors contributed equally: Zhen Dong, Wenhao Jiang. ✉e-mail: kiryl.piatkevich@westlake.edu.cn; zhuyi@westlake.edu.cn; guotiannan@westlake.edu.cn

cytometry platforms to enable the visualization of over 40 protein markers in archived tissue sections[22]. More recently, tissue expansion has been coupled to MALDI-based MS imaging[23].

In the expansion proteomics (ProteomEx) workflow, paraformaldehyde (PFA)-fixed tissue sections are treated with N-succinimidyl acrylate (NSA) to modify protein primary amino groups with an acryl group, enabling co-polymerization with hydrogel monomers[21]. Subsequently, the treated tissue undergoes infusion with hydrogel monomers, initiating an in situ polymerization reaction to create a tissue-hydrogel composite. To achieve isotropic expansion, the resulting tissue-hydrogel composite undergoes mechanical homogenization in a sodium dodecyl sulfate (SDS)-containing buffer at elevated temperatures (95 °C). Following the tissue expansion, the expanded sample is manually dissected to isolate a small ROI (below ~200 μm in size). These excised samples then undergo in-gel reduction/alkylation and tryptic digestion to extract peptides for subsequent LC-MS/MS analysis. Crucially, the chemical treatment of biological tissue does not introduce any chemical modifications and yields proteomics data identical to conventional sample preparation methods[21]. Moreover, ProteomEx operates without the need for sophisticated or custom equipment, relying solely on readily available reagents and accessories for sample preparation. However, limitations persist in the original ProteomEx protocol, with lateral and volumetric resolution capped at ~160 μm and ~0.61 nL, respectively[21]. Besides, the sample treatment usually takes ~54 h (excluding MS analysis). There are multiple manual steps that hinder the parallel processing of multiple samples, impeding large-scale analyses. Additionally, the applicability of ProteomEx to other types of archived tissue samples and their relevant tissue expandability has not been verified.

In this work, we present a filter-aided expansion proteomics (FAXP) approach, inspired by sample preparation methods such as filter-aided sample preparation[24], One-Tip[25], and simple and integrated spintip-based proteomics technology (SISPROT)[26], which utilize filter membranes[24] or tips[25,26] to achieve clean and comprehensive proteome profiling. Unlike One-Tip[25] and SISPROT[26], FAXP integrates optimized in-gel digestion[27] into self-assembled C18 filter tips, enabling high-throughput processing of tissue-hydrogel composites without sample size limitations. Compared to ProteomEx, FAXP offers enhanced lateral and volumetric resolution by 2.2- and 14.5-fold, respectively, and reduces processing time by 50%. For the first time, FAXP integrated with LCM enables precise isolation and analysis of a single mouse hepatocytic nucleus, identifying an average of 2368 proteins and showcasing its capability in subcellular proteomics. Moreover, FAXP is compatible with archived clinical FFPE samples and exhibits robust reproducibility. These advancements arise from improvements in tissue expansion, filter-aided in-gel digestion, and robotic integration. Using FAXP, we successfully characterize proteins associated with colorectal tumor progression from FFPE samples.

Results

Optimization of the tissue expansion procedures

Our initial focus was on expediting several time-consuming steps in the ProteomEx protocol, namely protein anchoring, hydrogel embedding, homogenization, and Coomassie staining (Fig. 1a). These optimizations collectively reduced the duration of these steps from 40.5 to 11.2 h (Fig. 1b) without compromising peptide and protein identifications (Supplementary Fig. 1). We found that autoclaving the tissue-hydrogel composite at 105–121 °C, which was intended for rapid homogenization, resulted in softer hydrogels post-expansion compared to the original treatment conditions. Since softer hydrogels were more difficult to handle during the dissection step and could potentially introduce larger distortions to the expanded tissue, we increased the hydrogel cross-linker ratio by 17.5 times (Supplementary Table 1) to harden it. This adjustment robustly enhanced the sturdiness of the hydrogel, consistent with a previous report[28].

To assess the impact of the modified hydrogel composition on tissue expansion, we conducted a quantitative analysis of the linear expansion factor and isotropy of expansion at both the cellular and whole tissue slide levels using a mouse liver FFPE slide. Protein anchoring time was selected as a variable in our experiments (Fig. 2a–e). At the cellular level, employing a 1-h anchoring condition yielded linear expansion factors (LEF) of the cell nucleus, as visualized via DAPI staining, comparable to those measured at the whole expanded tissue level (LEF = 4.42 ± 0.31 vs. 4.69 ± 0.17 for individual nucleus vs. whole expanded tissue, respectively; Fig. 2e), suggesting a consistent and uniform expansion pattern extending from the cell nucleus to the entire tissue structure. However, prolonged anchoring time (2 h) resulted in reduced expansion factors (LEF = 3.10 ± 0.58, p = 4.50e-6) at the cellular level compared to those measured at the whole expanded tissue level (Fig. 2e). Subsequently, we performed the anchoring for 1 h for all subsequent experiments. We then assessed the root-mean-square (RMS) measurement error in feature-length measurements after non-rigid registration of pre- and post-expansion tissue across length scales of up to 1500 μm, using 10-μm-thick mouse liver FFPE slides (n = 4). Our analysis revealed that the RMS errors were ~5.5% of the measurement distance, slightly improved compared to that of the original protocol (Fig. 2f, g). Comparable RMS errors were also observed for tissues with rich extracellular matrix, such as the kidney (n = 3; Supplementary Fig. 2).

Subsequently, we expedited the ProteomEx timeline and optimized peptide recovery from microdissected tissue-hydrogel samples by targeting the most laborious and unreliable steps associated with multiple manual procedures. We started with testing whether the reduction/alkylation steps, standard treatment to increase the yield of cysteine-containing peptides[29], could be performed on the whole tissue slice either before hydrogel embedding or after tissue-hydrogel composite homogenization using mouse brain slides (n = 3; Supplementary Fig. 3 and Supplementary Table 2). Compared to the original in-gel digestion protocol, reduction and alkylation performed after homogenization step on whole tissue section resulted in slightly increased peptide and protein identifications (by 22.8 and 13.0%, respectively) and improved distribution of peptide lengths while yielding similar carbamidomethyl modification rates and missed cleavage rates (Supplementary Fig. 3). These results show that performing reduction and alkylation on whole tissue-hydrogel composites after homogenization simplifies the protocol by reducing manual steps during in-gel digestion.

Then, we tested four reduction/alkylation conditions—10 mM dithiothreitol (DTT)/55 mM iodoacetamide (IAA), 20 mM Tris (2-carboxyethyl) phosphine (TCEP)/55 mM IAA, 45 mM TCEP/100 mM IAA, and 20 mM DTT/55 mM IAA—on entire tissue-hydrogel composites after homogenization using 10-μm-thick mouse liver FFPE slides (n = 3), comparing them to the in-gel digestion used in ProteomEx with one 10-μm-thick mouse liver FFPE slide (Fig. 2h). For the test groups, each condition was applied to a quarter of the hydrogel (n = 3). We found that 20 mM DTT/55 mM IAA resulted in more stable proteomic identification and better alkylation performance (Fig. 2i).

Next, we optimized in-gel proteolytic digestion conditions using mouse brain slides (n = 3; Supplementary Table 2) utilizing (1) one round of trypsinization (3 h); (2) two rounds of trypsinization (4 h + 12 h); (3) combination of LysC and trypsinization (4 h + 12 h). All conditions yielded a comparable number of identified peptides and proteins (Supplementary Fig. 4). However, the single-round trypsin digestion exhibited slightly higher missed cleavage rates compared to "LysC + Trypsin" digestion (Supplementary Fig. 4c). In conclusion, in-gel proteolytic digestion can be carried out in a wide range of conditions while extended time decreased missed cleavage rates.

https://doi.org/10.1038/s41467-024-53683-7

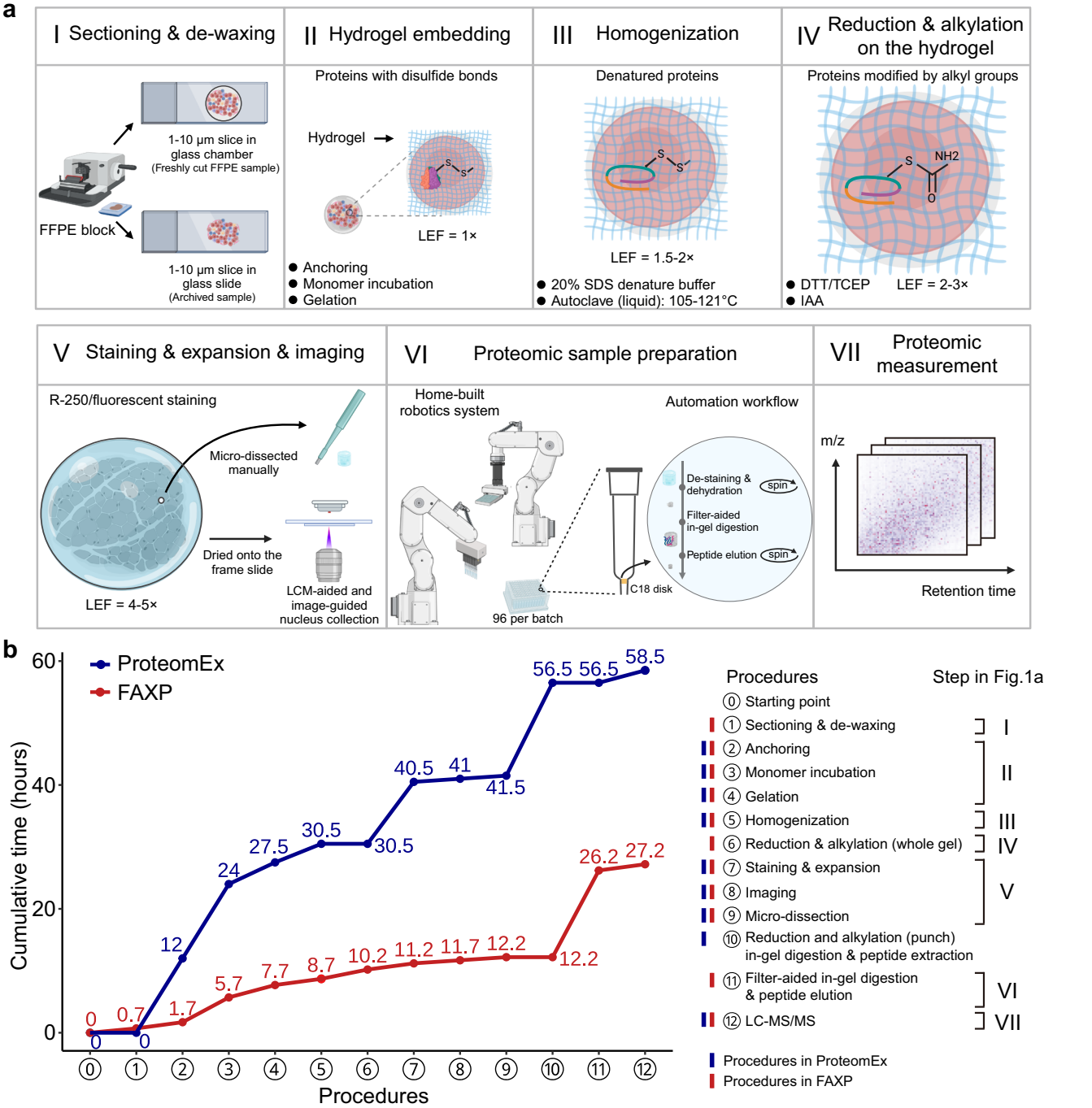
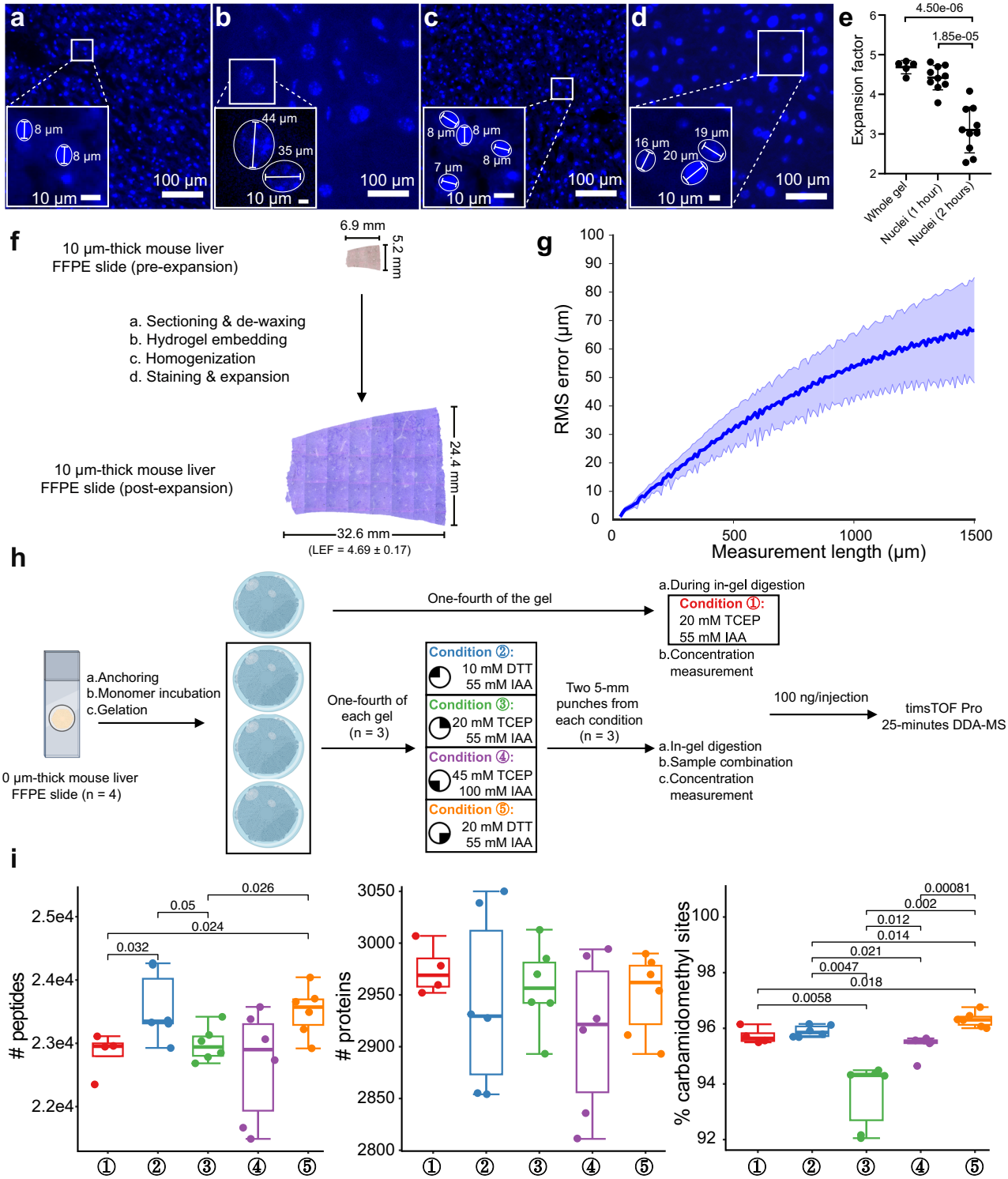


Fig. 1 | Comparative analysis of filter-aided expansion proteomics (FAXP) and ProteomEx. a Workflow of FAXP: The FAXP process initiates with formalin-fixed paraffin-embedded (FFPE) tissue block sectioning and dewaxing, followed by chemical treatment and embedding into a hydrogel. Subsequent steps involve autoclave-based homogenization, staining, expansion, imaging, and manual or laser capture microdissection. Filter-aided in-gel digestion occurs under robotic automation for DIA-MS, followed by proteomic measurement. (LEF linear expansion factor, which is calculated by comparing the expanded sample size to the original unexpanded sample size.) **b** Comparison of processing timelines: Timelines for each step in FAXP (red) and ProteomEx (blue) workflows are represented. Hands-on time is not delineated due to similarities between both workflows for small sample amounts. ProteomEx involves reduction and alkylation on individual punches during conventional in-gel digestion, while FAXP executes these steps on the entire gel. FAXP showcases faster processing times owing to shorter durations for various steps compared to ProteomEx. The FAXP data presented are derived from 5-μm-thick mouse liver FFPE samples. Source data are provided as a Source Data file. Figure 1a, created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

Adoption of filter-aided in-gel digestion

The ProteomEx method originally posed challenges in achieving high lateral and volumetric resolutions, limited to around 160 μm and 0.61 nL, respectively. This was mainly due to the difficulty in handling hydrogel samples under 1 mm in diameter. Specifically, these smaller hydrogel pieces tend to shrink through successive buffer exchanges, leading to sample loss during pipetting. To address this, we proposed immobilizing the tiny hydrogel pieces onto a membrane placed within a pipette tip. We tested the feasibility of this filter-aided in-gel digestion method, also known as "in-tip digestion", compared to the original

https://doi.org/10.1038/s41467-024-53683-7



in-gel digestion using 10-μm-thick mouse liver FFPE slides ($n = 6$ punches for each condition; Fig. 3a). Utilizing timsTOF Pro data-dependent acquisition (DDA) MS (50-min effective LC gradient) analysis for peptides generated from a 1.55 nL tissue-hydrogel composite, in-tip digestion yielded a remarkable increase in peptide and protein detections (by 96.6 and 31.1%, respectively; Fig. 3b). Additionally, this method enhanced protein digestion efficiency, with a missed cleavage rate of approximately 17.0%, lower than the 23.8% observed in in-gel digestion (Fig. 3c and Supplementary Fig. 5a). We also observed an unbiased identification across a series of physicochemical properties, like peptide hydrophobicity (Fig. 3d and Supplementary Fig. 5b) and increased detection of longer peptides (Fig. 3e), suggesting improved peptide coverage. Moreover, 74.9% of peptides and 97.5% of proteins identified through in-gel digestion were also detected via in-tip digestion. The in-tip method also identified 56.4% more peptides and 20.5% more proteins (Fig. 3f).

The in-tip digestion approach also led to more reliable quantification outcomes, evidenced by considerable improvements in the coefficient of variance (below 0.2 for both peptides and proteins; Fig. 3g) and Pearson's correlation (close to 1.0 for both peptide and protein levels; Fig. 3h and Supplementary Fig. 5c). Furthermore, the in-tip digestion method identified 26.7% additional liver-specific proteins

Fig. 2 | Optimization of gel-making procedures and comparative proteomic analysis. DAPI staining of **a** a mouse liver formalin-fixed paraffin-embedded (FFPE) slide after a 1-h anchoring before expansion ($n = 1$), **b** post-expansion tissue-hydrogel composite following 1-h anchoring ($n = 1$), **c** a mouse liver FFPE slide after a 2-h anchoring before expansion ($n = 1$), and **d** post-expansion tissue-hydrogel composite following 2-h anchoring ($n = 1$). Magnified views highlighting nuclei with their measured diameters are included (**a**–**d**). Scale bars are provided within each image. **e** Comparison of macro-level colorimetric staining ($n = 5$ biological replicates, representing different FFPE slides) and micro-level DAPI nuclear staining under 1- and 2-h anchoring conditions ($n = 10$ biological replicates, representing different nuclear diameters after expansion compared to before expansion). *P* values are estimated by Welch's *t*-test (two-sided). Data were presented as mean values ± standard deviation. **f** Bright-field images of a mouse liver FFPE slide before expansion (top) and after expansion (bottom), stained with Coomassie Brilliant Blue. The linear expansion factor (LEF) was calculated by comparing the dimensions of the expanded sample to the unexpanded sample ($n = 4$). **g** Comparing root-mean-square (RMS) measurement length errors pre- and post-expansion in liver slide images ($n = 4$); an example shown in (**f**). The blue line represents the mean, and the shaded area indicates the SD. **h** Study design for comparative proteomic analysis involving four reduction and alkylation conditions compared to in-gel digestion ($n = 4$ biological replicates), used under ProteomEx in red[21], as the control group. In this study, mouse liver FFPE slides from three independent biological samples undergo conditions: 10 mM dithiothreitol (DTT)/55 mM iodoacetamide (IAA) in blue, 20 mM Tris (2-carboxyethyl) phosphine (TCEP)/55 mM IAA in green, 45 mM TCEP/100 mM IAA in purple, and 20 mM DTT/55 mM IAA in orange. One-quarter of each gel is used per condition ($n = 3$ biological replicates). After in-gel digestion, two 5-mm punches from each condition were pooled, and MS data were collected using 100 ng per injection with two injections per condition. **i** Comparative proteomic analysis for the reduction and alkylation conditions outlined in Fig. 2h, including identified peptides and proteins, and carbamidomethyl modification site percentages. *P* values are estimated by Welch's *t*-test (two-sided). For all boxplots, center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend to the most extreme data points within 1.5 times the interquartile range from the 25th and 75th percentiles (IQR). Data were presented as dots and points beyond whiskers are outliers. Source data are provided as a Source Data file. Fig. 2h, partially created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

than that by the in-gel digestion method (Fig. 3i), as curated by the UniProt database. Pathway enrichment analysis of the proteins identified by both methods showed a similar distribution of subcellular localizations and protein types (Fig. 3j). Collectively, these findings underscore that the filter-aided in-gel digestion not only facilitates sample handling but also substantially enhances proteome coverage and reproducibility.

FAXP enhances the throughput, resolution, and sensitivity of spatial proteomics

Combining filter-aided in-tip digestion with an accelerated tissue expansion workflow not only enhanced proteome coverage but also increased sample preparation throughput, spatial resolution, and sensitivity. To assess the impact of in-tip digestion on the lateral and volumetric resolution of ProteomEx, we investigated the volume-dependent limit of tissue micro-sampling using the optimized workflow, with peptide recovery performed using either in-gel or in-tip protocols using 10-μm-thick mouse liver FFPE consecutive slides ($n = 2$; Fig. 4a). Manual dissection of tissue microsamples with volumes ranging from 0.042 to 3.054 nL (equivalent to lateral resolutions from 73 to 624 μm; $n = 7$ punches each size per protocol) revealed that in-tip digestion, under both DDA-MS mode and data-independent acquisition (DIA) MS mode, identified a range of 5196 to 28,411 peptides and 19,499 to 48,330 peptides, respectively (Fig. 4b).

In-tip digestion consistently outperformed in-gel digestion in terms of peptide and protein identification in both DDA and DIA modes, especially for tissue samples with a volume less than 3.054 nL. Notably, in-tip digestion reached a peak for tissue volumes ranging between 0.763 and 1.357 nL, beyond which no further increase was observed. On the other hand, in-gel digestion encountered technical constraints, limiting its ability to process tissue volumes below 0.339 nL. For a tissue volume of 0.339 nL, in-tip digestion outperformed in-gel digestion by over 800 and 250% in peptide and protein identification, resulting in the identification of 2204 and 4844 proteins in DDA and DIA mode, respectively.

Remarkably, in-tip digestion demonstrated compatibility with tissue-hydrogel composites as small as 0.042 nL in volume, achieving a lateral resolution of 73 μm on 10 μm-thick sections, equivalent to eight murine hepatocytes[30]. These findings illustrate that in-tip digestion improves both lateral and volumetric resolution by 2.85- and 8.07-fold, respectively, while providing 8.07/2.55-fold higher peptide/protein identification compared to the ProteomEx protocol with in-gel digestion under its lower limit tissue volume using DIA mode.

Coupling FAXP with LCM to study single nucleus proteome

We then tried to integrate FAXP and LCM to analyze the proteome of a single nucleus, which can hardly be isolated and collected for downstream proteomics analysis using LCM directly from FFPE tissues (Supplementary Fig. 6a). The diameter of a typical mouse liver hepatocyte nucleus is about 8 μm (Fig. 2a). After expansion with FAXP, the diameter increased to about 35 μm, which could be effectively isolated for proteomic analysis (Supplementary Fig. 6b). Without expansion, LCM can produce thick laser-burning curves around the nucleus, potentially causing damage or contamination from the surrounding cytoplasm. This also makes it challenging to isolate the nucleus intact for downstream proteomic analysis. In contrast, with FAXP, the nucleus expands linearly by about four times, resulting in a volume increase of approximately 64 times. This expansion makes isolation by LCM significantly easier. Additionally, when surrounded by hydrogel, the weight increases, simplifying further processing for proteomic analysis.

We benchmarked this strategy to study mouse hepatic nuclei and cells using one 10-μm-thick mouse liver FFPE slide (Fig. 4c, d). After attaching the tissue-hydrogel composite onto the LCM frame slide (Fig. 4e), nuclei were successfully isolated, and hepatocytes were identified using asialoglycoprotein receptor 1 (ASGR1), a hepatocytic plasma membrane marker[31]. The proteome of these isolated nuclei and hepatocytes was then analyzed using an Orbitrap Astral mass spectrometer. Following quality control, which involved removing one low-correlation single nucleus sample and two outliers from the mononuclear cell group (Supplementary Fig. 7), an average of 2368 and 3312 proteins were identified from single nuclei and single mononuclear cells, respectively (Fig. 4f). The proteome correlations were strong, exceeding 0.86 among nuclei and over 0.92 among mononuclear hepatocytes (Fig. 4g). Protein expression analysis showed that nuclear-associated terms like nucleoplasm, chromosome, and nucleolus were predominantly expressed in nuclei, whereas terms related to mitochondrion, endoplasmic reticulum, Golgi apparatus and peroxisome were more prominent in cells (Fig. 4h). Pathway analysis revealed that the citric acid (TCA) cycle and metabolism were highly activated in cells, while pathways like gene expression and mRNA splicing were more active in nuclei (Fig. 4h). We used ASGR1 in conjunction with nuclear staining to differentiate between mononuclear and binuclear hepatocytes. The correlation between these cell types exceeded 0.9 (Supplementary Fig. 8a). Histone expression levels were consistent with the previous reports[8] (Supplementary Fig. 8b). Additionally, we identified 129 proteins exclusive to binucleated cells and 37 proteins unique to mononucleated cells (Supplementary Fig. 8c). These findings demonstrate FAXP's effectiveness in accurately analyzing single cells and nuclei.

https://doi.org/10.1038/s41467-024-53683-7

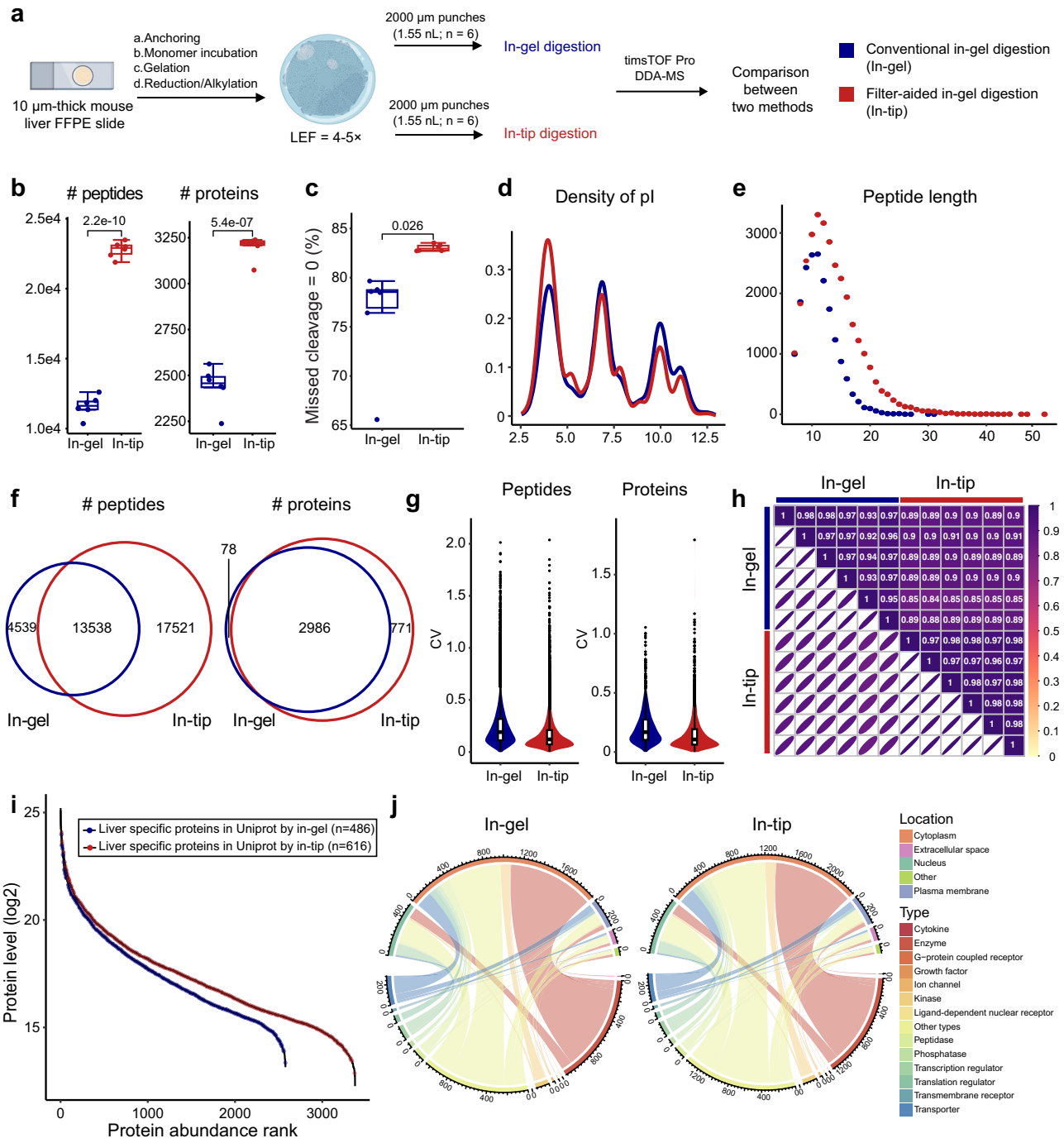


Fig. 3 | Comparative proteomic analysis of sample preparation methods applied to mouse liver FFPE slides. (**a**) Comparative proteomic study between conventional in-gel digestion (in-gel; blue) and filter-aided in-gel digestion (in-tip; red) using 2000-μm punches ($n = 6$ biological replicates) from mouse liver formalin-fixed paraffin-embedded (FFPE) slides (LEF linear expansion factor, which is calculated by comparing the expanded sample size to the original unexpanded sample size). **b** Comparison of identified peptides and proteins between in-gel and in-tip digestions ($n = 6$ biological replicates outlined in Fig. 3a). *P* values are estimated by Welch's *t*-test (two-sided). **c** Percentage of identified peptides with zero missed cleavage sites using in-gel and in-tip digestions ($n = 6$ biological replicates outlined in Fig. 3a). *P* values are estimated by Welch's *t*-test (two-sided). For both Fig. 3b, c, all boxplots show center lines representing the medians; box limits indicate the 25th and 75th percentiles, and whiskers extend to the most extreme data points within 1.5 times the IQR. Data were presented as dots, with points beyond whiskers representing outliers. **d** Density distribution of isoelectric points for identified peptides in both digestion methods. **e** Distribution of peptide lengths as shown in (b) in in-gel and in-tip digestions. **f** Venn diagram illustrating the overlap of identified peptides and proteins between in-gel and in-tip digestions. **g** Coefficient of variation (CV) showing quantified peptide and protein abundance variations between in-gel and in-tip digestions. **h** Pearson's correlation analysis for protein quantification between in-gel and in-tip digestions. **i** Log2-transformed mean protein intensities ranked by abundance, highlighting liver-specific proteins identified in both digestion methods. **j** Enriched protein subcellular locations and types presented in chord diagrams. Source data are provided as a Source Data file. Figure 3a, partially created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

https://doi.org/10.1038/s41467-024-53683-7

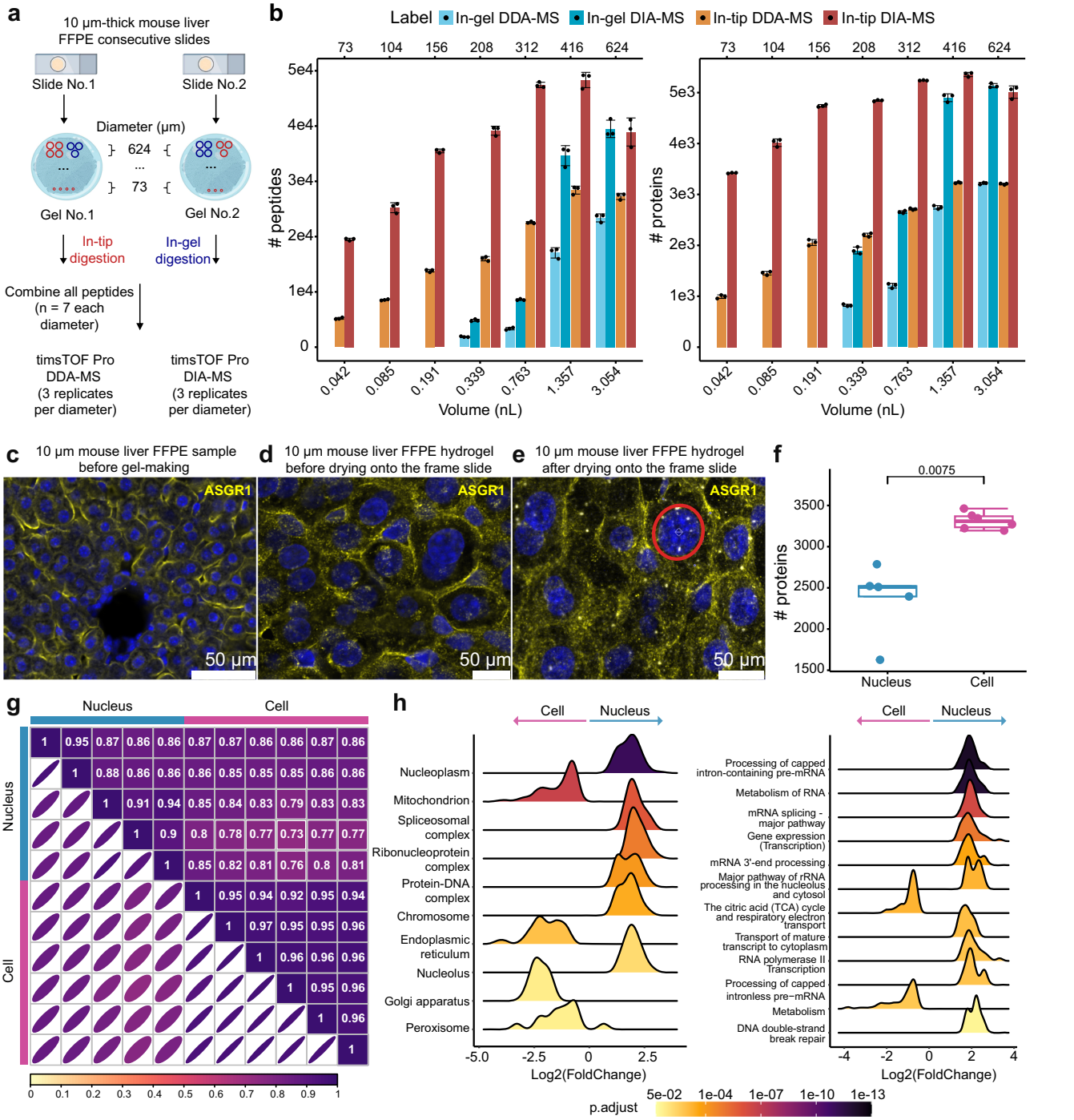


Fig. 4 | Detection range and limitation. a Study design for comparative proteomic analysis across various tissue volumes using in-gel (blue) and in-tip (red) digestions. Peptides from different diameters are analyzed using data-dependent acquisition mass spectrometry (DDA-MS; $n = 3$ technical replicates) and data-independent acquisition mass spectrometry (DIA-MS; $n = 3$ technical replicates). **b** Number of identified peptides and proteins for varied tissue volumes using in-gel and in-tip digestions as shown in (**a**), including in-gel DDA-MS (light blue), in-gel DIA-MS (blue), in-tip DDA-MS (orange), and in-tip DIA-MS (red). $n = 3$ technical replicates for each method and tissue volume. Data were presented as mean values ± standard deviation. **c** Asialoglycoprotein receptor 1 (ASGR1; yellow) and nuclear (blue) staining image of a mouse liver formalin-fixed paraffin-embedded (FFPE) slice before gel-making ($n = 1$). **d** ASGR1 and nuclear staining image of a mouse liver FFPE slice after gel-making ($n = 1$). **e** ASGR1 and nuclear staining image of a mouse liver FFPE slice after drying onto laser capture microdissection frame slide ($n = 1$). Scale bars are provided. **f** Comparison of proteins between single nucleus ($n = 5$ biological replicates; blue) and single mononuclear hepatocyte ($n = 6$ biological replicates; magenta). *P* value is estimated by Welch's *t*-test (two-sided). For all boxplots, center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend to the most extreme data points within 1.5 times the IQR. Data were presented as dots and points beyond whiskers are outliers. **g** Heatmap showing Pearson's correlations for protein quantification between each sample pair from a single mouse nucleus and mononuclear hepatocyte shown in (**f**). The color bar represents Pearson's correlation coefficient. **h** Ridge plot displaying protein expression distribution in enrichment analysis across major cellular compartments (left) and Reactome pathways (right) between single mononuclear hepatocyte and nucleus from (**f**). Selected items with FDR <0.05. Enrichment analysis was performed using fast gene set enrichment analysis with a Benjamini–Hochberg adjustment for multiple comparisons. Source data are provided as a Source Data file. Figure 4a, partially created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

Application of FAXP to understand intra-tissue variability of colorectal tumors

Next, we tried to apply the FAXP method to understand the intra-tissue variability of colorectal cancer (CRC). There were three patients, with the first having three consecutive FFPE slides, while each of the other two patients had one slide (Fig. 5a). Four regions, including normal (N), low-grade dysplasia (L), high-grade dysplasia (H), and carcinoma (C), were delineated from these samples based on hematoxylin and eosin (H&E) staining by an expert pathologist. Firstly, we validated the efficacy of the FAXP approach in achieving isotropic expansion of archived clinical samples (Fig. 5b). Of note, the RMS error for CRC tissue was found to be ~6% of the measurement distance (Fig. 5c), falling within the expected range based on mouse tissue expansion. Next, we manually performed a visual-guided microdissection of the CRC tissue-hydrogel composites from three distinct patients (referred to as P1, P2, and P3) and collected samples as outlined in Fig. 5a and Supplementary Table 3). A total of 131 microdissected tissue samples were processed using FAXP, leading to the identification of 6751 proteins. The systematic data analysis revealed that the number of identified proteins progressively increased with the severity of the cancer stage. Specifically, the numbers of protein identifications in L, H, and C were 6.67, 17.36, and 19.74% higher than that of the N samples based on median values, respectively (Fig. 5d), consistent with the previous reports[32]. To assess variations, we classified the identified proteins by disease region, MS batch, patient, and slide (Fig. 5e). Our data showed consistent protein identification across batches, patients, and slides (Fig. 5e and Supplementary Fig. 9), demonstrating the reproducibility and reliability of the FAXP workflow. Principal variance components analysis (PVCA)[33] indicated that the primary source of variability was due to different regions (48.94%), compared to patients (8.89%) and slides (0.7%), underscoring that regional differences are the main contributors to variability (Supplementary Fig. 9f).

We then compared the proteomes between different pairs of tissue regions (Fig. 6a). The UMAP plot effectively distinguished all CRC samples based on malignancy, particularly differentiating between normal and diseased regions (Fig. 6b). The number of differentially expressed proteins (DEPs) increased with disease progression, reflecting greater divergence between disease stages (Fig. 6c). Advanced-stage tumors exhibited a higher number of DEPs. We further employed Ingenuity Pathway Analysis (IPA) to explore DEPs across different malignancy stages compared to normal regions. This analysis revealed potential clinical biomarkers (Fig. 6d). Additionally, using the ADCdb[34], we identified potential antibody-drug conjugate (ADC) targets across different malignancy stages (Fig. 6e). Targets were generally present in all malignant stages compared to normal, with CEACAM5, PTGFRN, and ITGA3 consistently identified in every comparison. Besides, pronounced alterations in DEP expression were observed in advanced CRC stages compared to less severe stages (Supplementary Fig. 10a). We identified 19 DEPs associated with disease progression by comparing different malignant stages (Supplementary Fig. 10b). Notably, CEACAM5, a potential ADC target (Fig. 6e) and an oncogene associated with tumor progression and resistance to anoikis in colorectal carcinoma cells[35], exhibited increasing expression levels across disease stages (Supplementary Fig. 10b). This increase is consistent with CEACAM5's role in adhesion and invasion during CRC progression[36]. These findings underscore the reliability of the FAXP workflow and its capability to analyze variability in malignant changes within archived FFPE clinical samples.

Discussion

Compared to our previously published expansion proteomics method called ProteomEx[21], this FAXP approach has the following advantages (Fig. 1a). Firstly, we enhanced sample handling by using customized or commercial glass slides for FFPE samples, allowing greater flexibility in placement. FAXP accommodates any tissue size that fits on a standard glass slide (25 mm × 75 mm) and can process thin FFPE sections below 10 microns as well as thicker sections up to 120 microns via vibratome sectioning. It is also compatible with cryopreserved tissues, simplifying the process by removing the need for dewaxing and rehydration. We then introduced a 20% SDS denature buffer and autoclave-based tissue homogenization to expedite processing. Subsequently, we enhanced sample preparation reproducibility by conducting reduction/alkylation on the entire tissue sample embedded in the hydrogel. We also streamlined staining techniques after tissue expansion, ensuring faster and simpler procedures. Compared to ProteomEx, FAXP achieves a 2.2-fold improvement in lateral resolution (160 μm for ProteomEx vs. 73 μm for FAXP; Fig. 4b). Because FAXP can process 10-μm sections, while ProteomEx processes 30-μm sections, there is a 14.5-fold increase in volumetric resolution (0.61 nL for ProteomEx vs. 0.042 nL for FAXP; Fig. 4b). This enhancement in volumetric resolution demonstrates that FAXP can achieve high spatial resolution manually, without requiring high-end instruments like LCM, making it more accessible for laboratories. We further enabled LCM to analyze subcellular organelles. To surpass the 20 μm working limit of LCM[37], our study, for the first time, achieved bona fide single-nucleus proteomics under image guidance by integrating FAXP with LCM. We boosted sample throughput and robustness by employing filter-aided in-gel digestion. Moreover, we ensured comprehensive protein identification using DIA-MS with timsTOF Pro and Orbitrap Astral. Finally, we integrated robotics into our workflow. FAXP allows preparation of microdissected samples in a batch of up to 96, cutting total processing time by as much as 53.5% compared to ProteomEx (Fig. 1b). While 6% of gel samples are lost during peptide extraction using ProteomEx[21], FAXP, like other filter-aided sample preparation[24–26], not only eliminates the loss of gel samples and simplifies sample handling but also notably improves proteome coverage and reproducibility. To enhance reproducibility, Supplementary Note 1 provides detailed troubleshooting and critical steps for the FAXP workflow.

Recent advancements in spatial omics have significantly propelled spatial proteomics forward[1,2], with techniques like DVP and scDVP that integrate AI-assisted sample region selection, LCM, and advanced mass spectrometry for robust analysis of tissue samples down to single-cell resolution[6,8]. However, the spatial resolution of these approaches can be constrained by the limitations of LCM, particularly in maintaining the intactness of subcellular structures. In contrast, FAXP offers high spatial resolution through two distinct approaches. Firstly, it can achieve relatively high spatial resolution (down to 0.042 nL, equivalent to eight murine hepatocytes[30]) manually, without the need for high-end instruments, making it more accessible for a broader range of laboratories. As this method does not rely on LCM, the punch needle size can be further reduced, theoretically enabling even finer resolution. Secondly, when combined with LCM, FAXP further enhances spatial resolution beyond the capabilities of LCM alone, enabling precise analysis of individual subcellular organelles, such as a single nucleus, which can be challenging with conventional LCM techniques. FAXP also reduces sampling difficulties and improves accuracy compared to LCM alone, while minimizing laser-induced damage to tissue samples. Importantly, FAXP is not competing with DVP, which uses LCM to dissect cell shapes. Instead, FAXP is directly compatible with DVP. Cells or tissues could be expanded with FAXP and then followed by the DVP workflow, allowing for enhanced spatial resolution and sample integrity.

While the current FAXP workflow is not universally applicable to all tissue types due to its relatively complex preparation process, with operational challenges, potential contamination risks, and increased time costs, it does demonstrate significant adaptability across a diverse range of samples. For instance, tissues rich in extracellular matrix, such as CRC samples[38,39], require more rigorous homogenization, involving higher SDS concentrations, elevated temperatures, and extended processing times. These tailored optimizations, detailed in

https://doi.org/10.1038/s41467-024-53683-7

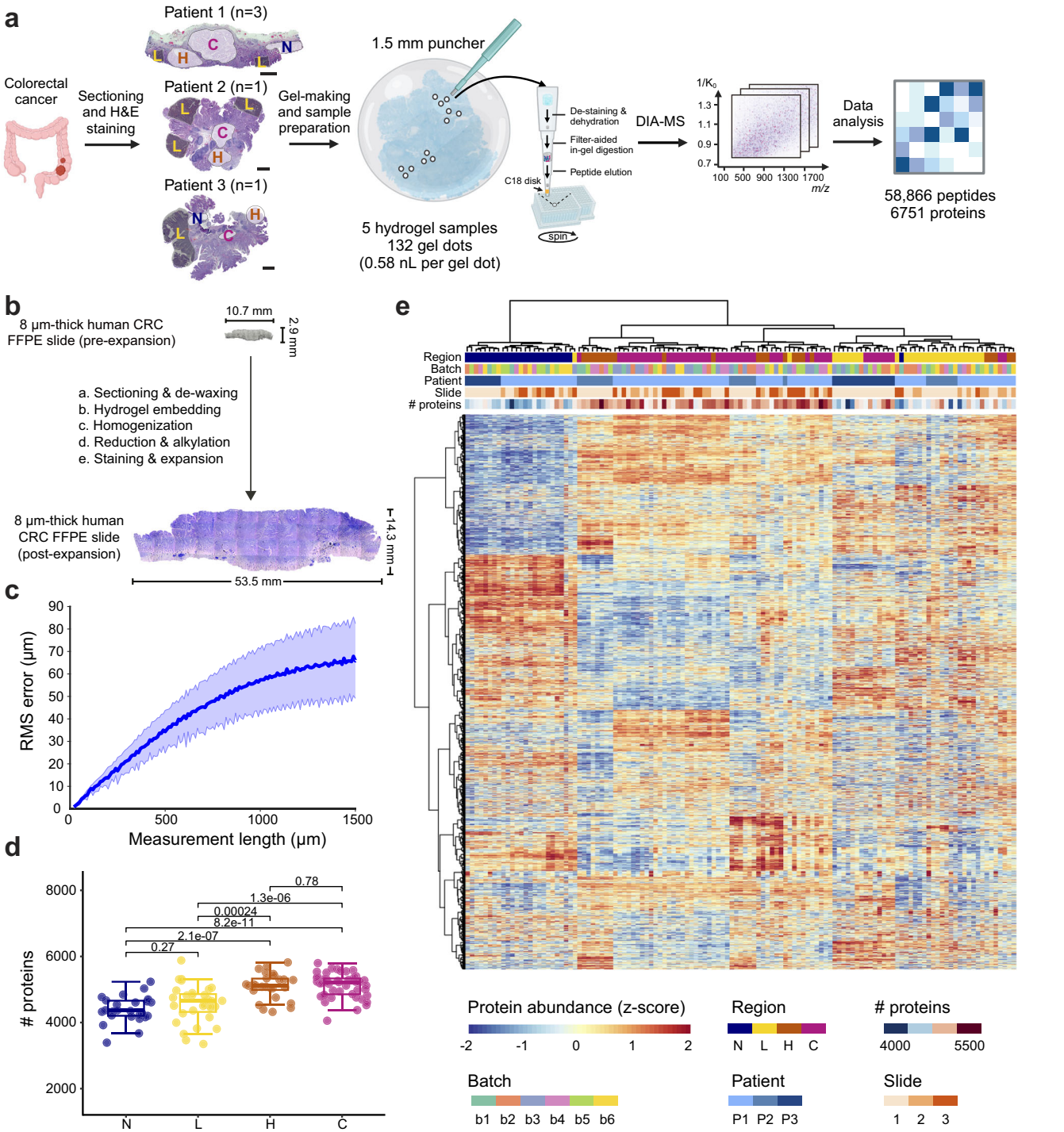
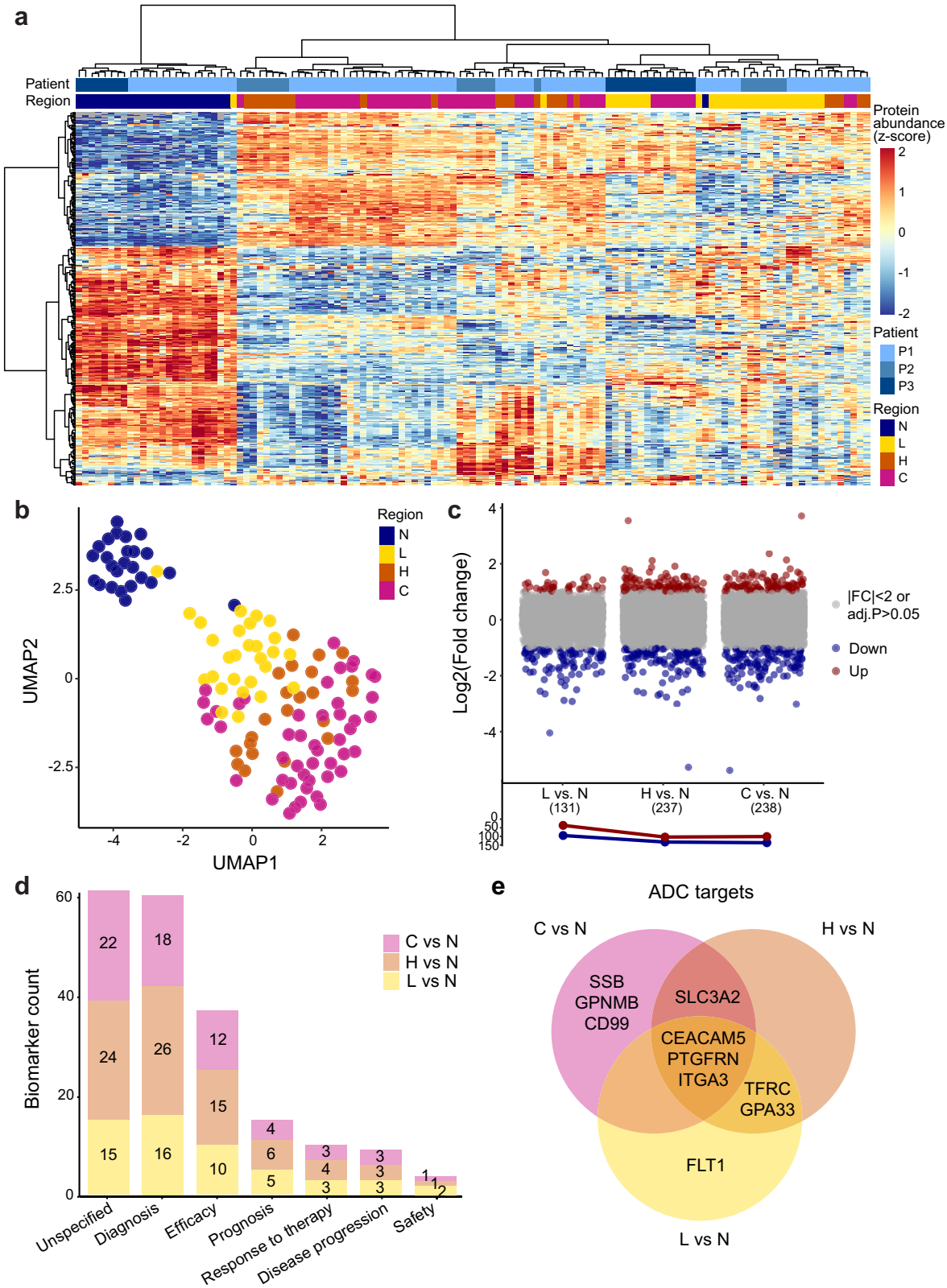


Fig. 5 | Application on colorectal cancer patient samples. a Study design overview: Analysis of colorectal cancer (CRC) samples from three patients (P1, P2, P3), including sequential slides from the first patient, to assess protein expression differences. Disease progression subtypes are labeled as follows: N normal (blue), L low-grade dysplasia (yellow), H high-grade dysplasia (orange), C carcinoma (magenta). Haematoxylin and Eosin (H&E) staining was performed on formalin-fixed paraffin-embedded (FFPE) slides. Mass spectrometry (MS) analysis was conducted on 1500-μm punches, followed by timsTOF Pro data-independent acquisition mass spectrometry (DIA-MS) and data analysis. Scale bars = 1000 μm. **b** Bright-field images of human CRC slides before (top) and after (bottom) expansion. Tissue-hydrogel composites were stained with Coomassie Brilliant Blue ($n = 3$). **c** Quantitative comparison of root-mean-square (RMS) measurement length error between original and expanded images ($n = 3$); an example is shown in (**b**). The blue line represents the mean, and the shaded area indicates the standard deviation. **d** Comparison of identified proteins across different subtypes. $n = 26$ biological replicates for N, $n = 31$ biological replicates for L, $n = 26$ biological replicates for H, and $n = 48$ biological replicates for C. These biological replicates represent different regions from multiple patients (as detailed in Supplementary Table 3). *P* values are estimated using Welch's *t*-test (two-sided) with a Benjamini–Hochberg adjustment. For all boxplots, center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend to the most extreme data points within 1.5 times the IQR. Data were presented as dots and points beyond whiskers are outliers. **e** Heatmap displaying protein abundance (z-score) with annotations for subtype regions, batch numbers (b1–b6), patient IDs (P1–P3), and slide numbers (1–3). Colors corresponding to each variable are defined in the legend. Source data are provided as a Source Data file. Figure 5a, partially created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

https://doi.org/10.1038/s41467-024-53683-7



Supplementary Fig. 11a and Supplementary Table 1, enhance FAXP's efficacy across various samples, underscoring its versatility for clinical applications. Although the FAXP method may be complex, its ability to be adapted for different tissue types suggests its potential for broad applicability in the future.

In exploring CRC applications, substantial differences arise among different stages of malignant changes. This emphasizes the criticality of analyzing specific CRC subtypes to unravel deeper molecular insights into the disease[40,41]. Particularly noteworthy is FAXP's ability to discern variability in disease progression within the slides. This feature is crucial in clinical research, allowing for precise sample analysis and enhancing our knowledge of diseases and treatment options.

In summary, FAXP presents a notable advancement in expansion proteomics by offering substantially accelerated sample processing,

Fig. 6 | FAXP applied to archived colorectal carcinoma samples. a Heatmap showing the 335 significantly differentially expressed proteins (DEPs) across different stages of malignant changes and normal regions. Regions are color-coded as N (Normal, blue), L (Low-grade dysplasia, yellow), H (High-grade dysplasia, orange), and C (Carcinoma, magenta), with patient IDs annotated as P1, P2, and P3. Protein abundance is displayed as z-scores. **b** Uniform manifold approximation and projection (UMAP) visualization of all regions by the DEPs shown in (**a**). **c** Differential protein expression analysis showing up- and down-regulated proteins (blue: downregulated, red: upregulated) across different stages of malignant changes compared to normal regions, with the line chart showing the number of DEPs. Statistical analysis was performed using Welch's *t*-test (two-sided), with a Benjamini–Hochberg adjustment for multiple comparisons. **d** Biomarker distribution for each stage of malignant change, with color-coded bars: C vs N (magenta), H vs N (orange), and L vs N (yellow). **e** Venn diagram comparing antibody-drug conjugate (ADC) targets across different stages of malignant changes compared to normal regions. Source data are provided as a Source Data file.

adaptability to diverse samples, improved spatial resolution, and enhanced sensitivity and reproducibility.

The FAXP workflow presented here has several constraints. The complexity of the protocol introduces the risk of potential contamination, particularly with keratin. To successfully execute this protocol, it is essential to maintain a clean experimental environment and use clean tubes, tips, and pipettes. The workflow supports a linear expansion factor of about five times to prevent FFPE tissue cracking, which is generally sufficient. However, while FAXP combined with LCM shows promise for analyzing larger subcellular structures like the nucleus, further optimization of expansion factors is needed for smaller organelles such as vacuoles, peroxisomes, and mitochondria. Additionally, reliance on Coomassie Brilliant Blue for staining CRC FFPE tissue-hydrogel composites limits the robustness and versatility required for accurate clinical assessments. As a demonstration of this technology, we analyzed samples from several CRC patients. Clinical interpretation from the CRC data should be further validated in independent cohort samples in the future. Finally, the current robotics system requires partial manual assistance, underscoring the need for future efforts to achieve full automation of the FAXP protocol. Improved automation will not only reduce operational errors but also minimize potential contamination.

Methods

This study complies with all relevant regulations for human study participants and was conducted in accordance with the Declaration of Helsinki. The collection and use of human tissue were approved by the Institutional Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine and the Ethics Committee of Westlake University.

Preparation of FFPE mouse organ tissue samples

All animal procedures adhered to Westlake University Animal Care Guidelines and received approval from the Institutional Animal Care and Use Committee (IACUC) under protocol #22-083-GTN. Male and female mice were used in separate benchmarking experiments, without comparing sexes, so results are not disaggregated by sex. Sample size estimation for animal studies was not based on statistical methods. As in previous work[21], all mice were housed in controlled barrier facilities, where the macroenvironmental temperature ranged from 20 to 26 °C and humidity from 40 to 70%. Food and water were available ad libitum, and the mice were kept on a 12-hour light/dark cycle. The housing conditions were carefully monitored and maintained.

Four male C57BL/6J wild-type mice (Mus musculus, four months old) and one female C57BL/6J wild-type mouse (Mus musculus, 2 months old) were obtained from the Laboratory Animal Resources Center of Westlake University. In the benchmarking studies, all experiments used male mice except for the single-nucleus proteomics study, which used the female mouse. All mice were deeply anesthetized with 1% sodium pentobarbital. Subsequently, transcardial perfusion was conducted using 1× phosphate-buffered saline (PBS) and 4% paraformaldehyde (PFA) obtained from Electron Microscopy Sciences, USA. The liver and kidney were dissected and fixed in 4% PFA for 6 h at 4 °C.

Fixed tissues underwent dehydration through sequential ethanol concentrations (75, 95, and 100%) for 30 min each. Following dehydration, the tissues were infiltrated with paraffin wax at 60 °C. The embedded tissues were then sliced into either 5-μm or 10-μm thickness using a rotary microtome (Leica RM2255, Germany) according to standard protocols and mounted onto glass slides.

Collection of FFPE samples from colorectal cancer patients

Archival FFPE colorectal tissue samples were from three CRC patients diagnosed at the Second Affiliated Hospital of Zhejiang University School of Medicine (Hangzhou, China). These samples were sliced to an 8-μm thickness using a rotary microtome (Leica RM2255, Germany). Tumor grading was conducted following the tumor, node, and metastasis (TNM) classification system for CRC. Tissue regions from both male and female patients were randomly selected from FFPE slides, with no analysis of sex or gender differences.

Prior to surgeries, informed consent was obtained from all patients. The human tissue collection adhered to the approved institutional review board protocol at the hospital. The collection and usage of human tissue samples were approved by the Institutional Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine (No. 2020-322) and by the ethics committee of Westlake University (Permission number: 20220913GTN001).

Update of tissue expansion and staining workflow

Supplementary Table 4 outlines reagents' compositions and storage conditions for tissue expansion, staining, and sample preparation. Detailed supplier and catalog information for FAXP-related chemicals, reagents, and accessories are provided in Supplementary Table 5.

Deparaffinization of paraffin-embedded sections involved two 10-min heptane treatments. Subsequent rehydration steps happen in 100, 90, and 75% ethanol for 5 min each followed by rehydration in ddH_2O for 5 min. For CRC tissue sections, H&E staining were performed following the guidelines from the manufacturer (Shanghai Yuanye Biotechnology, China).

The tissue expansion procedure, adapted from ref. 21, included PBS rinsing, anchoring for 1 or 2 h at 25 °C, and subsequent washing of the samples three times with the anchoring termination buffer for 5 min each. The tissue anchored with NSA underwent infusion with an activated monomer solution in a gelation chamber at 4 °C. Subsequently, it underwent polymerization in a vacuum oven at 37 °C in a nitrogen gas atmosphere for 3 h. For homogenization, sample autoclavation, akin to the dExPath protocol[42], was employed. The tissue-hydrogel composite was transferred to a dish containing a protein denaturation buffer, derived from the dExPath protocol[42]. This buffer contained 50 mM tris(hydroxymethyl)aminomethane (Tris), 20% SDS, 25 mM ethylenediaminetetraacetic acid, disodium salt, dihydrate (EDTA-Na_2·$2H_2O$), and 200 mM NaCl. Homogenization occurred at a temperature range of 105–121 °C under 1.22–2.03 bars for 60–90 min in the autoclave. To enhance the hydrogel's mechanical stability post-autoclaving, the cross-linker concentration was increased by 17.5-fold from the original concentration (Supplementary Table 1). Following homogenization, the samples were transferred to 9 cm petri dishes and washed thrice with 1× PBS. Reduction and alkylation of proteins involved adding 20 mM DTT in 1× PBS for reduction and 55 mM IAA in 1× PBS for alkylation, each incubated for 30 min. Subsequent washes

with 1× PBS occurred thrice for 10 min each. The hydrogel sample was then incubated in ddH_2O, allowing it to expand to its maximum size.

By employing Coomassie Brilliant Blue R-250 staining (Sangon, China), the process was accelerated to less than 1 h compared to the original 10-hour duration. Imaging utilized a Zeiss Fluorescence Stereo Zoom Microscope (Axio Zoom.V16) by ZEN v3.1 software.

Improved reduction and alkylation conditions

To ensure the proper unfolding and stabilization of proteins for accurate mass spectrometry analysis, we tested four groups of reduction and alkylation conditions on three 10-μm mouse liver hydrogel samples. Among the four conditions (10 mM DTT and 55 mM IAA; 20 mM DTT and 55 mM IAA; 20 mM TCEP and 55 mM IAA; 45 mM TCEP and 100 mM IAA), the combination of 20 mM DTT/55 mM IAA exhibited the highest carbamidomethyl modification rate, over an average of 96% (Fig. 2i). This method not only showed comparable rates of identifying cysteine-containing peptides and proteins to the established ProteomEx method (marked as "Condition 1" in Fig. 2i) but also demonstrated superior peptide identification performance.

Tissue-hydrogel sample preparation via conventional in-gel digestion

Manual microdissection was performed on Coomassie-stained tissue-hydrogels following reduction and alkylation. The excised samples underwent de-staining in a solution of 50% acetonitrile (ACN) and 50% ddH_2O, followed by dehydration in a solution of 50% ACN and 100 mM ammonium bicarbonate (ABB). After drying in a SpeedVac, trypsin digestion was carried out using 12.5 ng/μL trypsin (Hualishi Tech. Ltd, China) in 10% ACN and 90% 50 mM ABB for 4 h at 37 °C, followed by overnight incubation with an additional 10 mM ABB.

The resulting peptide solutions were collected, and subsequent steps involved the addition of 25 mM ABB followed by 50% ACN and 2.5% formic acid (FA) with vortexing, repeated three times with supernatant collection. 100% ACN was added until the hydrogel samples were reduced to small white dots. The peptide solutions were concentrated using a SpeedVac to 20–30 μL, followed by de-salting using C18 micro spin columns (Thermo Fisher Scientific, USA). After purification, the samples were dried with a SpeedVac, stored at −80 °C, and prepared for subsequent LC-MS/MS analysis.

Tissue-hydrogel sample preparation via filter-aided in-gel digestion

The C18 disk (Empore, USA) was assembled into either a 10 or 250 μL tip to create a spintip device, termed as the FAXP tip. After rinsing and activating the spintip with 20 or 100 μL of 80% ACN, microdissected tissue-hydrogel samples were added, followed by incubation with 50% ACN and 100 mM ABB. Dehydration was achieved by adding 100% ACN. Protein digestion utilized 12.5 ng/μL trypsin (Hualishi Tech. Ltd, China) in 10% ACN and 90% 50 mM ABB. The digestion process happened at 37 °C for 4 h, followed by overnight incubation with an extra 10 mM ABB. Following digestion and overnight incubation, the resulting peptide solutions underwent sequential steps: addition of 6 or 12 μL of 2% ACN and 0.1% trifluoroacetic acid (TFA), elution with 6 or 20 μL of 70% ACN and 0.1% TFA (twice), and a final elution with 6 or 20 μL of 100% ACN. These peptide solutions were concentrated using a SpeedVac and stored at −80 °C for LC-MS/MS analysis.

Immunostaining and gel-making for single-nucleus proteomics

A 10 μm FFPE liver section from a 2-month-old female mouse was used for this study. The tissue section was incubated for 30 min in citrate antigen retrieval solution (Leagene, China) and subsequently in blocking buffer (Sangon, China) at room temperature for 1 h. The section was then incubated with ASGR1 polyclonal antibody (Proteintech, China) at a 1:200 dilution. Following primary antibody incubation, the samples were treated with goat anti-rabbit IgG H&L secondary antibody (Alexa Fluor® 555) (Abcam, UK) at a 1:1000 dilution. The section was washed three times with 1× PBS and counterstained with DAPI (Abcam, UK). The section was imaged using a Nikon CSU-W1 SoRa confocal microscope with a 20× objective (0.8 NA), controlled by NIS-Elements AR v5.42.01 software.

About 50–100 μL of protein anchoring solution was added to the tissues, followed by a 1-h incubation at room temperature. The sample was then washed three times with 50–100 μL of anchoring termination buffer. After washing, the sample was dried in a fume hood, and an activated monomer solution was added to the surface. The assembly was placed in a humidified chamber and incubated overnight at 4 °C. The following day, the sample was covered with a glass coverslip and placed in a preheated vacuum drying oven at 37 °C for 3 h under anaerobic conditions. The tissue-hydrogel composite was transferred to a dish and submerged in protein denaturation buffer for homogenization, followed by reduction and alkylation.

For imaging, after blocking (Sangon, China), a tissue-hydrogel composite was incubated overnight at 4 °C with ASGR1 polyclonal antibody (Proteintech, China) diluted to 1:200, then washed in 1× PBS three times. Next, the hydrogel was incubated for 2 h at room temperature with goat anti-rabbit IgG H&L (Alexa Fluor® 555) (Abcam, UK) diluted to 1:200. The hydrogel was washed three times with 1× PBS, and counterstained with SYTOX Deep Red Nucleic Acid Stain (Thermo Fisher Scientific, USA) for 30 min at room temperature. The expanded sample was imaged using a Nikon CSU-W1 SoRa confocal microscope with a 20× objective (0.8 NA), controlled by NIS-Elements AR v5.42.01 software. The expanded hydrogel was then placed onto an LCM steel frame slide (Leica, Germany) and allowed to dry.

Proteomics data acquisition using Orbitrap Astral

For LCM using the Leica LMD7 system controlled by Leica LMD v8.4 software, the Y5 filter cube was used to visualize nuclei, with the following cutting parameters: power at 60, aperture set between 17 and 20, speed ranging from 8 to 12, bridge size set to 0, a final pulse at 6, head current at 100%, and pulse frequency between 120 and 150. The same cutting parameters were applied for hepatocytes, using the RHOD ET filter cube.

Each sample obtained using the Leica LMD7 was transferred into the 10 μL FAXP tip, and 20 μL of 100% ACN was added. The samples were incubated at room temperature for 10 min, followed by centrifugation at 150 × *g* for 5 min to remove the supernatant. Subsequently, following ref. 43, we added 0.015% *n*-dodecyl β-D-maltoside (DDM) to the trypsin digestion buffer, achieving a final trypsin concentration of 5 ng/μL. The samples were then incubated at 37 °C for 4 h. Following digestion, the resulting peptide solutions underwent sequential steps: addition of 6 μL of 2% ACN/0.1% TFA/0.015% DDM, elution with 6 μL of 70% ACN/0.1% TFA/0.015% DDM (twice), and a final elution with 6 μL of 100% ACN. These peptide solutions were then concentrated using a SpeedVac and reconstituted for LC-MS/MS analysis.

Single nuclei ($n = 6$), single mononuclear cells ($n = 8$), single binuclear cells ($n = 7$), and blank controls ($n = 5$) were analysed. Peptides were analyzed via an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific, USA) controlled by Thermo Scientific Xcalibur v4.7.69.37 and Orbitrap Astral Tune Application v1.0.100.40. Separation of peptide samples utilized an Aurora Elite TS 15 × 75 C18 UHPLC column (IonOpticks, Australia) in a Vanquish™ Neo UHPLC system (Thermo Fisher Scientific, USA). The liquid chromatography procedure involved a 19.5-min gradient running step and a 5.5-min column washing step. The mobile phase comprised buffer A (98% MS grade H_2O, 2% MS grade ACN, and 0.1% FA) and buffer B (98% MS grade ACN, 2% MS grade H_2O, and 0.1% FA). The injection workflow was direct injection.

The Orbitrap MS full scan settings included a total carrier gas flow of 3.8 L/min, an Orbitrap resolution of 240k, a scan range of 400 to 800 m/z, a FAIMS CV of −42 V, an RF Lens of 45%, a normalized automatic gain control (AGC) target of 500%, and a maximum IT of 100 ms.

The Astral DIA-MS^2 scan parameters covered a precursor mass range of 400 to 800 m/z and a scan range of 150 to 2000 m/z. The DIA isolation window was set at 20, the maximum injection time at 40 ms, with HCD normalized collision energy at 25%, RF lens at 45%, maintaining a normalized AGC target of 800%, and a loop control time of 0.6 s.

Proteomics data acquisition using timsTOF Pro

The tissue collection was performed using commercially available biopsy punches (Integra Miltex, USA) ranging from 0.35 to 3 mm in diameter. A total of four mouse liver FFPE slides (10-µm-thick) were analyzed under four different reduction/alkylation conditions ($n = 3$ per condition, each with two technical replicates), with in-gel digestion under ProteomEx serving as the control ($n = 4$), using the Parallel Accumulation Serial Fragmentation (PASEF) mode of data-dependent acquisition (ddaPASEF). For the comparative study between in-gel ($n = 6$) and in-tip digestion ($n = 6$) methods, 2000-µm punches from a single mouse liver FFPE slide (10-µm-thick) were analyzed using dda-PASEF. Peptides from different tissue volumes of two consecutive mouse liver FFPE slides (10-µm-thick) were processed by in-gel and in-tip digestion and analyzed with three replicates per digestion method per volume using both ddaPASEF and data-independent acquisition using PASEF (diaPASEF).

The peptide samples, separated on a custom-packed silica column in a nanoElute® system (Bruker Daltonics, Germany), utilized a mobile phase composed of MS grade H_2O + 0.1% FA (buffer A) and MS grade ACN + 0.1% FA (buffer B) at a flow rate of 300 nL/min. Employing both 25-min and 50-min effective linear LC gradients, buffer B (%) increased from 5 to 27% over 25 and 50 min, respectively. Subsequently, it rose to 40% in 5 and 10 min, followed by a boost to 80%. The 50-min LC gradient resulted in enhanced peptide and protein identification, and a higher mean peak full width at half maximum (FWHM) averaging from ~2.5 to 3.5 (Supplementary Fig. 11b).

For MS analysis, eluted peptides were examined using a trapped ion mobility spectrometry (TIMS) combined with a quadrupole time-of-flight mass spectrometer (timsTOF Pro, Bruker Daltonics, Germany) through a CaptiveSpray nano-electrospray ion source. MS data were acquired using the Bruker Otofcontrol v6.2 and HyStar v5.1 mass spectrometer control software.

In the ddaPASEF mode, each cycle included ten PASEF scans, operating with a ramp time of 100 ms for a total cycle duration of 1.17 s. Ion mobility scanning ranged from 0.7 to 1.3 Vs/cm^2, and MS1 and MS2 acquisition covered the m/z range of 100 to 1700 Th. To manage precursor intensities, those reaching a target intensity of 20,000 arbitrary units were dynamically excluded for 0.4 min. Additionally, singly charged precursors were excluded based on their position in the m/z-ion mobility plane.

In the diaPASEF mode, most settings were the same as those of the ddaPASEF mode, except for a ramp time adjusted to 166 ms. The isolation windows were defined as outlined in Supplementary Table 6.

Proteomic data analysis

DDA data analysis was performed utilizing FragPipe (version 18.0)[44,45], integrating the MSFragger search engine (version 3.5). The analysis involved querying a FASTA file from SwissProt, housing 16,985 mouse protein entries (downloaded in September 2018), alongside their respective decoy sequences. Furthermore, a separate FASTA file from SwissProt, comprising 20,386 human protein entries (downloaded in November 2022), along with corresponding decoy sequences, was utilized for human sample analysis. "IM-MS" was chosen as the MS data type, employing a fragment mass tolerance of 0.05 Da. Trypsin was specified as the digestion enzyme, cleaving after "KR" residues, except when followed by "P". Carbamidomethylation was assessed for cysteine modification rates and treated as a variable modification in this analysis, differing from other DDA files where it was set as a fixed modification.

In DIA data analysis, parameters were set in DIA-NN (version 1.8.1)[46,47]: a 1% false discovery rate for precursors, fixed modifications including "C carbamidomethylation" and variable modifications including "N-term M excision" and "Methionine oxidation". The criteria included a peptide length range of 7 to 50, a precursor charge range of 2 to 4, and m/z ranges for precursor and fragment ions between 100 and 1700. Our approach involved selecting "Unrelated runs", "Use isotopologues", and "MBR" options. For specific studies, heuristic protein inference was applied for mouse liver-based analysis. Protein inference was set as "off", and the quantification strategy was set as "Robust LC (high precision)". Other settings remained as default values. In-house libraries were utilized for mouse liver-based and CRC studies, respectively.

For subcellular studies, a library-free mode was employed, using a FASTA file from SwissProt containing 17,217 mouse protein entries (downloaded in August 2024). Modifications were made to the precursor (400–800 m/z) and fragment ion (150–2000 m/z) ranges, while the "Unrelated runs" option was not selected. The "--relaxed-prot-inf" command was included under Additional Options. Protein inference was set to "Genes," and the quantification strategy was configured as "Robust LC (high precision)." All other settings remained at their default values.

To enhance clarity, the terminology for peptides and proteins was made consistent. Peptides generally referred to amino acid sequences without accounting for PTMs (post-translational modifications) or charge states, while proteins referred to razor proteins (in FragPipe reports) or protein groups (in DIA-NN reports). Specifically, in Figs. 2i, 3b, f, g, razor proteins and indistinguishable proteins (in FragPipe reports) were considered. In contrast, in Supplementary Fig. 9, PTMs and charge states of precursor peptides were taken into consideration.

Protein matrix preprocessing

Unless otherwise specified, protein matrices from DIA-NN were converted to a logarithmic scale for all subsequent analyses, except for mean intensity calculations. Prior to the differential expression analysis, proteins were filtered to control the missing data ratio. For CRC studies, the missing ratio was less than 20% in at least three out of the four groups.

Proteomic quality control analysis

The CRC application study involved six batches (batch design available in the source data for Fig. 5e), resulting in a total of 131 files due to an MS bug in one occurrence. On average, 28,763 peptides and 5480 proteins were identified across the pooled samples using DIA-MS (Supplementary Fig. 9a). Besides, robust reproducibility was observed in global peptides and proteins, as indicated by the coefficient of variation and Pearson's correlation analysis in pooled samples (Supplementary Fig. 9b, c). Furthermore, throughout the analysis, a steady and low average missed cleavage rate of 18.27% was consistently maintained, as evaluated by DDA-MS (Supplementary Fig. 9d).

Statistical analysis

For all numerical comparisons, Welch's *t*-test (two-sided) was executed, while statistical visualizations were made using PRISM (version 10.0.3) and R (version 4.4.0). MATLAB (version R2021a) was utilized for isotropic analysis, following a previously established protocol[21].

The comparison of DEPs was conducted using the limma package in R, which applies linear models and incorporates the duplicateCorrelation function to account for within-patient correlations. "Region" was included as the primary fixed effect of interest without an intercept, and "Patient" was set as the blocking factor. We defined a set of planned contrasts to conduct hypothesis testing between specific pairs of regions. An empirical Bayes moderated *t*-test was applied with FDR adjustment to account for multiple testing. For any pairwise

comparison, proteins with an adjusted p value <0.05 and |logFC|> log2 were considered statistically significant, while ensuring that the overall linear model had an adjusted p value <0.05.

The Pearson's correlation coefficient was calculated after the exclusion of missing values. Hierarchical cluster analysis was performed using Ward's method and Euclidean distance. The uniform manifold approximation and projection (UMAP) dimension reduction technique was employed to identify the fuzzy topological representation of high-dimensional data.

The pvca R package was utilized to conduct principal variance components analysis (PVCA). The top-N principal components were selected based on a 90% threshold for explained variance. Patient, slide, and region were included as random effects to account for their inherent variability. The magnitude of each effect as a proportion of the total variance was estimated by weighted average.

Gene ID conversion was conducted using the clusterProfiler (version 4.11.0). The Gene Ontology (GO) term annotations were performed using org.Mm.eg.db (version 3.18.0), and redundant GO terms were simplified manually. The fast gene set enrichment analysis (GSEA) was performed using clusterProfiler to identify Cellular Component Ontology terms ranging in size from 10 to 300. The Reactome pathway-based analysis was conducted using the ReactomePA (version 1.47.0) package to identify gene sets ranging in size from 5 to 300. Gene sets with an adjusted p value of less than 0.05, as reported by GSEA, were considered statistically significant.

Reporting summary

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

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[48] partner repository with the dataset identifier PXD041412. The raw images for Figs. 2a–d, 4c–e, and 5a, along with all isotropic analysis datasets used in this study, are available on Zenodo at https://doi.org/10.5281/zenodo.13843661[49]. Source data are provided with this paper.

Code availability

Code relevant to data analysis in this study is available at GitHub [https://github.com/guomics-lab/FAXP] and Zenodo [https://doi.org/10.5281/zenodo.13843661[49]].

References

1. Mund, A., Brunner, A. D. & Mann, M. Unbiased spatial proteomics with single-cell resolution in tissues. *Mol. Cell* **82**, 2335–2349 (2022).
2. Baysoy, A., Bai, Z., Satija, R. & Fan, R. The technological landscape and applications of single-cell multi-omics. *Nat. Rev. Mol. Cell Biol.* **24**, 695–713 (2023).
3. Taylor, M. J., Lukowski, J. K. & Anderton, C. R. Spatially resolved mass spectrometry at the single cell: recent innovations in proteomics and metabolomics. *J. Am. Soc. Mass Spectrom.* **32**, 872–894 (2021).
4. Moore, J. L., Patterson, N. H., Norris, J. L. & Caprioli, R. M. Prospective on imaging mass spectrometry in clinical diagnostics. *Mol. Cell. Proteom.* **22**, 100576 (2023).
5. Claes, B. S. R. et al. MALDI-IHC-guided in-depth spatial proteomics: targeted and untargeted MSI combined. *Anal. Chem.* **95**, 2329–2338 (2023).
6. Mund, A. et al. Deep visual proteomics defines single-cell identity and heterogeneity. *Nat. Biotechnol.* **40**, 1231–1240 (2022).
7. Makhmut, A. et al. A framework for ultra-low-input spatial tissue proteomics. *Cell Syst.* **14**, 1002–1014.e1005 (2023).
8. Rosenberger, F. A. et al. Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat. Methods* **20**, 1530–1536 (2023).
9. Xu, R. et al. Spatial-resolution cell type proteome profiling of cancer tissue by fully integrated proteomics technology. *Anal. Chem.* **90**, 5879–5886 (2018).
10. Nordmann, T. M. et al. Spatial proteomics identifies JAKi as treatment for a lethal skin disease. *Nature*. https://doi.org/10.1038/s41586-024-08061-0 (2024).
11. Kabatnik, S. et al. Spatial characterization and stratification of colorectal adenomas by deep visual proteomics. *iScience* **27**, 110620 (2024).
12. Kwon, Y. et al. Hanging drop sample preparation improves sensitivity of spatial proteomics. *Lab Chip* **22**, 2869–2877 (2022).
13. Zhu, Y. et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* **9**, 882 (2018).
14. Ma, M. et al. In-depth mapping of protein localizations in whole tissue by micro-scaffold assisted spatial proteomics (MASP). *Nat. Commun.* **13**, 7736 (2022).
15. Bhatia, H. S. et al. Spatial proteomics in three-dimensional intact specimens. *Cell* **185**, 5040–5058 (2022).
16. M'Saad, O. & Bewersdorf, J. Light microscopy of proteins in their ultrastructural context. *Nat. Commun.* **11**, 3850 (2020).
17. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
18. Ku, T. et al. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* **34**, 973–981 (2016).
19. Tillberg, P. W. et al. Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016).
20. Drelich, L. et al. Toward high spatially resolved proteomics using expansion microscopy. *Anal. Chem.* **93**, 12195–12203 (2021).
21. Li, L. et al. Spatially resolved proteomics via tissue expansion. *Nat. Commun.* **13**, 7242 (2022).
22. Bai, Y. H. et al. Expanded vacuum-stable gels for multiplexed high-resolution spatial histopathology. *Nat. Commun.* **14**, 4013 (2023).
23. Chan, Y. H. et al. Gel-assisted mass spectrometry imaging enables sub-micrometer spatial lipidomics. *Nat. Commun.* **15**, 5036 (2024).
24. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
25. Ye, Z. et al. One-Tip enables comprehensive proteome coverage in minimal cells and single zygotes. *Nat. Commun.* **15**, 2474 (2024).
26. Chen, W. et al. Simple and integrated spintip-based technology applied for deep proteome profiling. *Anal. Chem.* **88**, 4864–4871 (2016).
27. Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).
28. Damstra, H. G. J. et al. Correction: Visualizing cellular and tissue ultrastructure using ten-fold robust expansion microscopy (TREx). *eLife* **11**, e85169 (2022).
29. Suttapitugsakul, S., Xiao, H. P., Smeekens, J. & Wu, R. H. Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS- based proteomics. *Mol. Biosyst.* **13**, 2574–2582 (2017).
30. Hammad, S. et al. 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 (2014).

31. Zhang, Z. et al. Deficiency of ASGR1 promotes liver injury by increasing GP73-mediated hepatic endoplasmic reticulum stress. *Nat. Commun.* **15**, 1908 (2024).
32. Shao, Y. K. et al. Proteomics profiling of colorectal cancer progression identifies PLOD2 as a potential therapeutic target. *Cancer Commun.* **42**, 164–169 (2022).
33. Boedigheimer, M. J. et al. Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genomics* **9**, 285 (2008).
34. Shen, L. et al. ADCdb: the database of antibody–drug conjugates. *Nucleic Acids Res.* **52**, D1097–D1109 (2024).
35. Ordoñez, C., Screaton, R. A., Ilantzis, C. & Stanners, C. P. Human carcinoembryonic antigen functions as a general inhibitor of anoikis1. *Cancer Res.* **60**, 3419–3424 (2000).
36. Blumenthal, R. D., Hansen, H. J. & Goldenberg, D. M. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (carcinoembryonic antigen). *Cancer Res.* **65**, 8809–8817 (2005).
37. Bury, A. G. et al. A subcellular cookie cutter for spatial genomics in human tissue. *Anal. Bioanal. Chem.* **414**, 5483–5492 (2022).
38. Yuan, Z. N. et al. Extracellular matrix remodeling in tumor progression and immune escape: from mechanisms to treatments. *Mol. Cancer*. **22**, 48 (2023).
39. Karlsson, S. & Nystrom, H. The extracellular matrix in colorectal cancer and its metastatic settling-alterations and biological implications. *Crit. Rev. Oncol. Hematol.* **175**, 103712 (2022).
40. Baker, A. M. et al. Evolutionary history of human colitis-associated colorectal cancer. *Gut* **68**, 985–995 (2019).
41. Wang, W. et al. Molecular subtyping of colorectal cancer: recent progress, new challenges and emerging opportunities. *Semin. Cancer Biol.* **55**, 37–52 (2019).
42. Valdes, P. A. et al. Improved immunostaining of nanostructures and cells in human brain specimens through expansion-mediated protein decrowding. *Sci. Transl. Med.***16**, eabo0049 (2024).
43. Tsai, C.-F. et al. Surfactant-assisted one-pot sample preparation for label-free single-cell proteomics. *Commun. Biol.* **4**, 265 (2021).
44. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520 (2017).
45. Yu, F. et al. Fast quantitative analysis of timsTOF PASEF data with MSFragger and IonQuant. *Mol. Cell. Proteomics* **19**, 1575–1585 (2020).
46. Demichev, V. et al. dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nat. Commun.* **13**, 3944 (2022).
47. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
48. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
49. Dong, Z. et al. Spatial proteomics of single cells and organelles on tissue slides using filter-aided expansion proteomics. *Zenodo*. https://doi.org/10.5281/zenodo.13843661 (2024).

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Major Research Plan, Grant No. 92259201 to T.G.), the National Key R&D Program of China (Grant No. 2022YFF0608403 to Y.Z. and Grant No. 2021YFA1301600 to T.G.), the Key Research and Development Program of Zhejiang Province (Grant No. 2022C03037 to T.G.), the National Natural Science Foundation of China (Grant No. 32171093 to K.D.P.), the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (Grant No. 2024SSYS0031 to K.D.P.), and the Westlake Education Foundation. Figure 1a was created with Biorender.com. Icons from BioRender.com were used in Figs. 2h, 3a, 4a, 5a. We would like to thank Cuiji Sun for teaching gel-making approaches based on ProteomEx, Dr. Liang Yue for sharing the home-built pan-human timsTOF Pro DIA library, Dr. Xiangqing Li and his team, as well as Qi Xiao for helping in robotics. The language of the first draft has been polished with ChatGPT before proof-reading by the co-authors.

Author contributions

T.G., Y.Z., and K.D.P. conceived the project. Z.D., Y.Z., K.D.P., and C.W. developed and benchmarked the method. T.C. and S.Z. provided the clinical samples. Z.D., C.W., and J.C. performed the proteomic experiments. X.D. optimized the Astral method. W.J. and Z.D. conducted proteomic data analysis and Z.D. performed the isotropic analysis. Z.D., K.D.P., W.J., and T.G. wrote the manuscript, and the others revised the manuscript. T.G., Y.Z., and K.D.P. supervised the project.

Competing interests

T.G. and Y.Z. are shareholders of Westlake Omics (Hangzhou) Biotechnology Co., Ltd., where ProteomEx and FAXP-related technologies are commercialized. T.G., Z.D., Y.Z., C.W., and K.D.P. are inventors of a granted patent (Chinese patent number: ZL 202310889444.4) submitted by Westlake Laboratory of Life Sciences and Biomedicine that cover the FAXP methods. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53683-7.

Correspondence and requests for materials should be addressed to Kiryl D. Piatkevich, Yi Zhu or Tiannan Guo.

Peer review information *Nature Communications* thanks Fabian Coscia, Ali Shaib and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.