Multimodal AI-enabled mass spectrometry-based expansion proteomics for whole-slide at single-cell resolution

Shuaiyao Wang[1,2,3]†, Zhen Dong[1,2,3]†*, Chunlong Wu[1,2,3], Jiayi Chen[1,2,3], Changao Li[1,2,3],
Jianpeng Sheng[4], Xiang Li[5], Yi Chen[1,2,3]*, Tiannan Guo[1,2,3]*

Affiliations:

[1]Affiliated Hangzhou First People's Hospital, State Key Laboratory of Medical Proteomics, School of Medicine, School of Future Biomedicine, 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]Zhejiang Provincial Key Laboratory of Pancreatic Disease, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang Province, China.

[5]School of Science, China Pharmaceutical University, Nanjing, Jiangsu Province, China.

*Corresponding author. Email: dongzhen@westlake.edu.cn (Z.D.); chenyi@westlake.edu.cn (Y.C.); guotiannan@westlake.edu.cn (T.G.)

†These authors contributed equally to this work

Abstract: Deep, quantitative proteome coverage at single-cell resolution across entire tissue sections remains a major challenge for mass spectrometry-based spatial proteomics. Here, we introduce an AI-empowered filter-aided expansion proteomics (FAXP) framework that combines FAXP with convolutional neural network (CNN)-based spatial inference to achieve whole-slide, single-cell-resolved proteomics. By combining tissue expansion with orthogonal laser capture microdissection to obtain whole-slide linear strip-resolved proteome measurements, we develop HetuNet, a CNN-based model that integrates these sparse orthogonal data with high-dimensional imaging-derived contexts to reconstruct comprehensive two-dimensional spatial protein expression landscapes. In mouse liver, this approach captures zone-specific protein patterns and continuous pathway-level gradients, validated by spatial transcriptomics. In colorectal cancer tissues, it resolves proteome-defined epithelial states, revealing functional divergence via differential epithelial-mesenchymal transition and necroptosis activation. Together, this framework enables deep, scalable spatial proteomic mapping across whole tissues at single-cell resolution, unlocking previously inaccessible insights into tissue organization and function.

Main

Tissue sections are central to pathological diagnosis and offer unique basis for studying disease mechanisms and therapeutic response (*1*). Accurate interpretation of tissue organization requires preservation of the spatial relationships among cells and their surrounding microenvironment (*2*, *3*). Spatial characterization of tissues has evolved from Hematoxylin and Eosin (H&E) staining to immunohistochemistry, and more recently to multiplexed immunofluorescence (IF) imaging platforms such as co-detection by indexing (CODEX) (*4*) and cyclic IF (*5*). These whole-slide approaches capture the entire slide's morphological features and selected protein expression. However, their protein detection capacity is inherently limited, as they rely on antibody availability and are restricted to predefined panels comprising at most dozens of known protein markers (*6*, *7*). Recently, spatial transcriptomics has enabled spatial profiling of mRNA expression across whole tissue sections (*8*, *9*); however, mRNA abundance provides only an indirect proxy of cellular state (*10*). Transcriptomic measurements are affected by limited capture efficiency and data sparsity, and transcript levels often show weak correspondence with protein abundance (*11*, *12*). In addition, spatial transcriptomics is limited in its ability to comprehensively capture extracellular components, such as matrix-associated and secreted proteins. These challenges are particularly pronounced in archived clinical specimens such as formalin-fixed paraffin-embedded (FFPE) tissues, where RNA-based assays are compromised due to unpredictable RNA fragmentation during paraffin embedding and further degradation associated with long-term or suboptimal storage (*13*). Together, these limitations motivate the development of approaches that directly and comprehensively measure the spatial proteome within whole-slide tissues.

Mass spectrometry (MS)-based spatial proteomics is rapidly progressing toward spatially resolved protein profiling within tissue contexts (*14*). Deep Visual Proteomics (DVP) is a representative single-cell, spatially resolved LC-MS workflow that integrates artificial intelligence (AI)-guided image analysis for region-of-interest (ROI) selection with laser capture microdissection (LCM) and high-resolution MS analysis (*15–17*). LCM enables spatially targeted proteomics, yet the requirement for physically isolatable regions limits access to small cells and subcellular features (*18*). Modest throughput and incomplete material recovery further constrain its scalability for whole-slide profiling (*19*). Inspired by expansion microscopy (*20*, *21*), tissue expansion has recently been integrated with mass spectrometry imaging (MSI) via hydrogel embedding to reduce molecular crowding and limit peptide diffusion, enabling improved dynamic range across whole-slide tissue sections (*22*). However, most MSI workflows primarily rely on MS1-level information and do not routinely incorporate MS/MS-based peptide sequencing, resulting in limited proteome granunarity and coverage (typically on the order of ~1,000 proteins) and thereby constraining biological interpretability (*22*, *23*). Other whole-slide proteomic approaches, including AI-driven frameworks such as PLATO and S4P, are constrained by limited effective spatial resolution and thus primarily capture multicellular signals rather than bona fide single-cell states (*24*, *25*). Microfluidics-based strategies such as PLATO typically sample strips with ~50% spatial coverage, inherently leaving at least 50% of the tissue unmeasured.

Here, we introduce a unified framework for reconstructing whole-slide, spatially resolved proteomic landscapes at single-cell resolution. This strategy generates strip-resolved MS measurements at single-cell scale, yielding a total of 3,883 protein groups in mouse liver and 6,501 protein groups in human colorectal cancer (CRC). Next, we develop HetuNet, a convolutional neural network (CNN)-based AI model that integrates these proteomic measurements with high-dimensional spatial molecular context derived from multiplexed IF imaging. Validation in mouse liver show that our method accurately reconstructs continuous metabolic zonation and proteomic gradients across tissue sections. *In silico* benchmarking with Xenium spatial transcriptomics data confirm the quantitative accuracy and structural robustness of the reconstructed spatial proteomes. Application to human CRC tissues uncovers proteome-defined cellular states beyond targeted imaging and delineated spatial functional divergence among epithelial populations near necrotic regions.

Results

Methodology establishment for whole-slide single-cell spatial proteomics

To establish a robust pipeline for whole-slide, single-cell spatial proteomics with high resolution and quantitative accuracy, we developed an integrated framework combining physical tissue expansion, multiplexed imaging, orthogonal LCM-based spatial proteomics, high-resolution MS, and deep learning-based inference (**Fig. 1a**). In this framework, two consecutive FFPE tissue sections (4~10 µm thick) were subjected to isotropic tissue expansion to enhance effective spatial resolution. One expanded section was used for high-resolution IF imaging, while both sections were subsequently stained with Coomassie Brilliant Blue (CBB) and processed by orthogonal LCM, yielding row- and column-oriented strip samples with single-cell-scale widths. All dissected samples were then processed using the filter-aided sample preparation (*26*, *27*) and analyzed by LC-MS/MS, enabling the identification and quantification of up to 5,448 proteins per sample.

On the basis of these measurements, spatial reconstruction was performed using the HetuNet framework, which infers high-resolution protein expression maps from sparse proteomic data and morphological features (**Fig. 1b**). HetuNet employs a shared CNN to extract multiscale spatial features from multi-channel IF images and tissue masks within local regions. These features are processed through an adaptive gating mechanism to generate location-specific representations, which are then decoded by protein-specific prediction heads to infer expression levels (**Fig. S1**). This architecture allows the spatial transfer of proteomic information from strip-level MS data to patch-level imaging data across the whole tissue section.

Next, we validated the physical integrity of the tissue following the expansion process. Bright-field imaging of expanded mouse liver sections showed that the global tissue morphology was well preserved (**Fig. 1c; Fig. S2a**). We assessed root-mean-square (RMS) measurement errors in feature lengths after non-rigid registration of pre- and post-expansion mouse liver tissue using 10-µm-thick FFPE slides. RMS errors were ~2% of the measurement distance and remained consistent across length scales up to 1,500 µm, indicating negligible spatial distortion introduced by expansion (**Fig. 1d**). To evaluate the depth of spatial proteomic profiling, we quantified the number of proteins identified across all spatial strips. Up to 3,600 proteins were identified in individual strips in both sampling directions, with lower coverage mainly observed at tissue edges (**Fig. 1e**). Each profiled spatial strip corresponded to a 12-µm width. Zoom-in views of DAPI staining revealed clearly resolved nuclear features, consistent with spatial resolution at the single-cell scale.

To ensure quantitative accuracy, we implemented quality control and normalization strategies. MS performance remained stable throughout data acquisition, as indicated by a consistent number of protein identifications across runs (**Fig. S2b**). Quantitative reproducibility was further supported by low coefficients of variation (CVs), with mean CVs below 0.2 at both precursor and protein levels, and by strong concordance among quality control samples, which showed Pearson correlation coefficients exceeding 0.99 (**Fig. S2c**, **d**). Given the variable tissue content in LCM strips, raw protein intensities were normalized by the proportion of IF image pixels within each sample area. This normalization reduced sampling-related variability, as reflected by stabilized expression levels of representative housekeeping proteins, including HSP90AB1 and RPL8, across individual row-wise samples (**Fig. 1f**), with consistent trends observed for additional housekeeping proteins (**Fig. S2e**). Together, these results support the consistency of the data generation and processing pipeline.

Reconstruction of continuous metabolic zonation gradients in the mouse liver

To validate the spatial fidelity, we utilized mouse liver tissue characterized by distinct metabolic zonation. IF staining for the central vein (CV) marker glutamine synthetase (GLUL) and the portal vein (PV) marker E-cadherin (E-CAD) was used to define ground-truth morphological landmarks. These reference images were visualized at their native high resolution (0.27 µm per pixel) and after downsampling to 12 µm per pixel to match the resolution of model outputs (**Fig. 2a**, Ground truth). We then evaluated HetuNet reconstructions for 20 zonated proteins (ten CV-enriched and ten PV-enriched) selected based on prior single-cell deep visual proteomics studies (*17*). The reconstructed spatial distribution of the CV-enriched protein CYP2C29 showed strong concordant with GLUL-positive CV regions, whereas the PV-enriched protein CYP2F2 aligned closely with E-CAD-positive PV regions (**Fig. 2a**, HetuNet). The remaining zonated proteins also displayed consistent spatial patterns (**Fig. S3**, **S4**).

We next quantified these spatial patterns by manually annotating representative regions within the tissue section as CV-like (Zone 3; pink) or PV-like (Zone 1; yellow) based on multiplexed IF signals of GLUL, asialoglycoprotein receptor 1 (ASGR1) and E-CAD (**Fig. 2b**; **Fig. S5a**, **b**). Using these annotations as spatial references, gene set enrichment analysis (GSEA) (*28*) of the predicted proteomes revealed clear zone-specific metabolic programs (**Fig. 2c**). Proteins associated with triglyceride metabolism, linoleic acid metabolism, steroid hormone biosynthesis, and retinol metabolism were enriched toward the CV region (Zone 3), whereas oxidative phosphorylation and urea cycle were preferentially enriched toward the PV region (Zone 1). These spatially resolved enrichment patterns were consistent with previously reported functional zonation of the liver (*17*, *29*, *30*).

We then compared the predictions of Flow2Spatial (*24*) with those generated by HetuNet to assess its compatibility with the orthogonal strip-based input format used in this study. While HetuNet produced coherent spatial distributions for representative CV- and PV-enriched proteins such as GULO and SDS, Flow2Spatial yielded discontinuous and spatially inconsistent patterns when applied to the same protein expression input data (**Fig. 2d**). These observations indicate a structural mismatch between the Flow2Spatial framework and the orthogonal strip-based IF input data format.

Quantitative assessment of HetuNet reconstruction fidelity via *in silico* spatial transcriptomics

To assess the spatial reconstruction fidelity of HetuNet, we performed an *in silico* simulation using a high-resolution spatial transcriptomics dataset derived from human renal cell carcinoma

(*31*). We mathematically partitioned the tissue into orthogonal pseudostrips to mimic the approach's sampling strategy, and reconstructed spatial gene expression patterns from these aggregated inputs (**Fig. 3a**). Reconstruction fidelity was quantitatively evaluated using the structural similarity index measure (SSIM). SSIM values consistently exceeded 0.5 across all gene groups, indicating preservation of tissue-scale spatial organization (**Fig. 3b**). Notably, SSIM values increased significantly when computed using the top 100 or top 50 highly variable genes (HVGs) compared with using all remaining genes after quality filtering. While both the top 100 and top 50 HVGs achieved similarly elevated SSIM values, no systematic improvement was observed when the gene set was reduced from 100 to 50. Considering the balance between spatial fidelity and gene coverage, the top 100 HVGs were therefore selected for subsequent analyses.

Next, quantitative concordance between HetuNet-predicted and ground-truth expression profiles was assessed. Pearson correlation analysis of row- and column-aggregated expression for the top 100 HVGs showed uniformly high agreement, with mean correlation coefficients exceeding 0.98 **(Fig. 3c**), suggesting accurate recovery of large-scale spatial expression trends from aggregated inputs. We further examined model robustness by systematically increasing the number of reference markers used during training (**Fig. S6a**). SSIM values for the top 100 HVGs increased monotonically with the number of reference features, remaining above 0.6 across all conditions and reaching a maximum mean SSIM of 0.664 when 25 reference markers were used (**Fig. 3d**). A similar scaling behavior was observed in Pearson correlation analyses, where spatial Pearson correlation coefficients increased with marker number and reached their highest mean values under the 25-marker condition (mean PCC = 0.317; **Fig. S6b**, **c**).

At the level of whole spatial maps, increasing the number of reference markers was associated with improved reconstruction performance for representative genes. When trained with four reference markers, HetuNet achieved spatial Pearson correlation coefficients of 0.364 (KRT8), 0.460 (CXCR4), 0.441 (VIM), and 0.493 (MS4A1) relative to the ground truth (**Fig. 3e**). Expanding the reference set to 25 markers increased these correlations to 0.588, 0.601, 0.664, and 0.829, representing relative increases of ~62% (KRT8), ~31% (CXCR4), ~51% (VIM), and ~68% (MS4A1), respectively. Structural similarity analysis showed a parallel improvement, with SSIM values increasing from 0.54 to 0.72 (KRT8), 0.35 to 0.45 (CXCR4), 0.48 to 0.65 (VIM), and 0.76 to 0.88 (MS4A1) when increasing from four to 25 reference markers (**Fig. S6d**).

Together, these results indicate that HetuNet achieves improved spatial reconstruction fidelity with increasing numbers of reference markers when reconstructing spatial expression patterns from orthogonally aggregated inputs, as demonstrated using a heterogeneous cancer tissue. This relationship motivated subsequent analyses using high multiplexed reference marker sets.

Single-cell whole slide spatial proteomics reveals cellular architecture of human CRC tissue

We next applied this workflow to human CRC FFPE tissues for spatial proteomic profiling. Adjacent tissue sections were used for CODEX-based multiplexed IF imaging and orthogonal LCM-based proteomics (**Fig. 4a**). Bright-field images of CRC sections before and after expansion showed preserved global tissue morphology, with RMS measurement errors of ~2% (**Fig. S7a-c**). MS analyses along the horizontal (n = 134) and vertical (m = 123) sampling axes yielded comparable proteome coverage, with over 4,000 proteins identified in the majority of samples in both directions, while reduced coverage was observed at tissue edges (**Fig. 4b**, **c**). Totally 6,501 proteins were identified. Quality control analyses showed stable, normalized abundances of the housekeeping proteins (RPL13, RPS18, RPL8, and EEF1D) across column-wise samples (**Fig. S7d**). The CODEX images revealed spatial heterogeneity within the tumor microenvironment (**Fig. 4d**). Specifically, four representative regions of interest were selected for downstream analyses, including regions with high cellular heterogeneity (green and red boxes) and regions with more organized tissue structures (yellow and pink boxes).

Leveraging the high-dimensional spatial information provided by the CODEX platform, we trained HetuNet using 22 spatially resolved protein markers together with DAPI as reference inputs, while designating HLA-DRB1, CD44, VIM, and CDH1 as target proteins for prediction. These four markers were excluded from model training and reserved exclusively for evaluation (27 total features, **Fig. S8**). The predicted spatial distributions for these four markers were quantitatively compared with the ground-truth CODEX measurements, yielding global Pearson correlation coefficients of 0.355 (HLA-DRB1), 0.429 (CD44), 0.329 (VIM), and 0.332 (CDH1) across multiple regions of interest (**Fig. 4e**). For reference, in spatial transcriptomics datasets generated from 10-µm-thick sections at an 8-µm pixel resolution, the mean global Pearson correlations between modalities are typically below 0.2 (*32*), highlighting that the concordance achieved in our study is within a reliable and biologically meaningful range. The zoomed-in views correspond to the color-coded boxes are shown in **Fig. 4d**. When HetuNet was trained with only four reference markers (DAPI, CD31, FoxP3, and a-SMA), the global Pearson correlation coefficients between predicted and CODEX-measured expression were 0.312 (HLA-DRB1), 0.182 (CD44), 0.420 (VIM), and 0.314 (CDH1). Compared with the 23-marker training condition, the correlations for HLA-DRB1, CD44, and CDH1 decreased by approximately 12.1%, 57.6%, and 5.4%, respectively, indicating reduced reconstruction fidelity under the four-marker condition (**Fig. S7e**). In contrast, VIM exhibited a higher correlation under the four-marker condition, potentially attributable to its spatial similarity with α-SMA, one of the reference markers included in this reduced feature set (**Fig. S8**).

Deep proteome profiling resolves CRC cell states undefined by antibody-based imaging

To establish a reference cell type annotation for the tissue section, we first utilized the CODEX panel comprising 31 protein markers. After excluding five low-quality markers, HetuNet was trained to predict protein expression using the remaining 26 protein markers together with DAPI. Unsupervised clustering based on these markers resolved nine distinct clusters across the tissue section (**Fig. 5a**). By examining canonical marker expression profiles and uniform manifold approximation and projection (UMAP), we assigned these clusters to three major cell types, including epithelial cells, fibroblasts, and immune cells (**Fig. 5b** and **Fig. 5c**, left). A subset of cells lacked detectable expression of the targeted panel markers and could not be confidently assigned to these types; these cells were therefore annotated as “Unknown” in the spatial map (**Fig. 5c**, right).

Next, we examined whether HetuNet could provide additional insight into these unknown cells. Clustering was performed on the unbiased single-cell proteomes predicted by HetuNet, independent of the targeted CODEX panel. This analysis resolved 16 distinct cell clusters within the same tissue environment (**Fig. 5d**). Based on representative marker expression profiles (**Fig. 5e**), UMAP analysis (**Fig. 5f**, upper panel), and clinical annotations, these clusters were grouped into four major cell categories: epithelial cells, fibroblasts, immune cells, and necrotic cells (**Fig. 5f**, lower panel).

Using spatial regions of necrosis identified by an experienced pathologist (**Fig. 5g**, red box), comparison with CODEX annotations showed that cell classification based on HetuNet-predicted proteomes more completely delineated these regions. We next visualized the spatial distributions of the four cell-category annotations derived independently from CODEX and from HetuNet-predicted proteomes (**Fig. 5h**). Spatial overlap analysis showed substantial concordance between the CODEX-defined "Unknown" population and the proteome-defined necrotic cell category, indicating that proteome-based analysis resolves cell states not captured by targeted antibody-based imaging.

Spatially resolved functional divergence and signaling pathways in the epithelial compartment

To further characterize the functional heterogeneity within the epithelial compartment, we applied single-sample gene set enrichment analysis (ssGSEA) to the HetuNet-predicted proteomes. This analysis revealed spatially distinct activation patterns for multiple critical biological pathways, including epithelial-mesenchymal transition (EMT), extracellular matrix (ECM) receptor interaction, necroptosis, and the tricarboxylic acid (TCA) cycle (**Fig. 6a**). Spatial analysis showed that the epithelial cells surrounding the necrotic regions identified in the previous analysis exhibited distinct spatial patterns (**Fig. 6b**).

Then, we stratified epithelial cells into two distinct functional states based on the differential enrichment scores of the EMT and necroptosis pathways. The first state (E1) was characterized by elevated EMT activity and reduced necroptosis scores, whereas the second state (E2) exhibited the opposite pattern, with reduced EMT enrichment and elevated necroptosis scores **(Fig. 6c, d)**. We then examined representative proteins to characterize these epithelial states. Violin plots and spatial mapping showed that MACROH2A1 and CTNNB1 were more highly expressed in the E2 state, with enriched expression localized to epithelial cells adjacent to necrotic regions (**Fig. 6e, f**). Together, the spatial distributions of these proteins distinguished the E2 epithelial state from E1, linking this proteomic profile to epithelial cells in the vicinity of necrotic tissue.

Discussion

Single-cell, whole-slide spatial proteomics represents a central goal for decoding tissue organization and cellular heterogeneity, yet the field remains constrained by a fundamental trade-off among spatial resolution, proteome depth, and tissue-scale coverage (*14*, *33*, *34*). Antibody-based imaging offers high spatial resolution and tissue coverage but limited proteome depth (*6*). MSI-based approaches, such as iPEX, provides label-free molecular maps across tissue sections with restricted depth and quantitative robustness (*23*). LCM-based LC-MS strategies, including DVP, achieve deep and quantitative proteome profiling but are dependent on predefined regions of interest and therefore lack unbiased whole-slide scalability (*15–17*). AI-assisted inference frameworks such as PLATO and S4P have partially expanded spatial coverage from sparsely sampled measurements (*24*, *25*); however, their reliance on the underlying physical sampling scale constrains effective spatial resolution to ~25-500 μm, remaining above the single-cell level. Using reported section thicknesses (8 μm for PLATO and 10 μm for S4P) to translate these in-plane resolutions into voxel volumes, this corresponds to ~25 × 25 × 8 μm and ~500 × 500 × 10 μm, respectively, approximately 2.5-1,250 cell equivalents assuming an average cell volume of

~2,000 µm³ (*16*). Notably, the current state-of-the-art method PLATO uses an 8 µm section thickness combined with a 25 µm width resolution, within which a total of 1,986 protein groups can be identified. Together, these approaches remain unable to achieve single-cell, whole-slide proteomics with deep and quantitative coverage.

Building on our expansion proteomics workflow (*26*, *27*), we integrate isotropic tissue expansion with HetuNet-based AI inference to enable single-cell, whole-slide proteomics from only two consecutive 4-10-µm-thick FFPE sections. Using 4-µm-thick CRC tissue sections with a ~12-µm width resolution, we identified a total of 6,501 protein groups. The reconstructed voxel sizes, 12 × 12 × 10 µm for mouse liver and 12 × 12 × 4 µm for CRC, approximate 0.3-0.7 cell equivalents. This performance establishes a practical benchmark for spatial proteomics that jointly attains high resolution, deep proteome coverage, and extensive tissue continuity.

A key factor enabling this capability lies in the physical sampling architecture. Microfluidics-based strategies typically samples strips of ~25 µm width with ~50% spatial coverage, of which only ~25% is directly measured while the remaining ~75% is computationally inferred, making the final maps heavily reliant on imputation (*24*, *35*). While such designs enable parallelization, their discontinuous geometry introduces structured missing regions that pose challenges for AI-based spatial reconstruction, in contrast to continuously sampled, orthogonal strategies that more readily support learning of spatial continuity and cross-scale relationships. Conventional high-power LCM (Leica) approaches provide deep proteome measurements but are typically constrained to strip widths of ~100 µm and often require elevated laser energy, potentially compromising tissue integrity and necessitating multiple consecutive sections (*25*). By contrast, our strategy employs low-power LCM (MMI) combined with a 4-5× linear expansion factor (*26*, *27*, *36*), enabling near-lossless and near-complete spatial overlap and high-fidelity tissue cutting at ~12 µm effective resolution using only two sections. This design substantially mitigates the physical constraints inherent to microfluidic and high-power LCM sampling while preserving proteomic integrity. Importantly, these advances build upon these pioneering spatial proteomics technologies while extending the field toward scalable, continuous, single-cell proteome mapping.

To fully exploit the resulting deep but spatially sparse proteomics data, we developed HetuNet, a weakly supervised deep learning framework that formulates protein mapping as a tomographic reconstruction problem constrained by morphological guidance. Algorithmically, HetuNet distinguishes itself by employing a multi-scale gated convolutional architecture that extracts hierarchical spatial features from multiplexed IF imaging, coupled with a projection-consistency objective function. This design allows the model to learn complex, non-linear pixel-level distributions directly from aggregated strip-level measurements by maximizing the correlation between predicted spatial marginals and ground truth projections. By solving this inverse problem without requiring dense ground truth, HetuNet overcomes significant limitations in current methodologies: unlike PLATO (*24*), it eliminates the need for prior imputation of unmeasured regions or predefined tissue architectural references; and in contrast to S4P (*25*), which relies on unguided MLP inference across multiple consecutive sections, our approach explicitly leverages imaging-informed spatial priors. This enables the scalable reconstruction of continuous, single-cell-resolution protein maps across entire tissue sections, effectively bridging the trade-off between resolution, depth, and spatial coverage within a unified framework.

The biological fidelity of the reconstructed spatial proteome was examined in the structurally organized mouse liver, where well-defined metabolic gradients provide a stringent benchmark for spatial inference. HetuNet recapitulated continuous protein abundance gradients along the central-portal vein axis, closely matching IF patterns. Notably, the reconstructed proteomes preserved functional metabolic organization beyond canonical marker distributions; for instance, GSEA of the reconstructed data accurately delineated the spatial segregation of metabolic programs, with oxidative phosphorylation and the urea cycle enriched in PV regions, whereas steroid biosynthesis, triglyceride metabolism, linoleic acid metabolism, and retinol metabolism were preferentially enriched in CV regions, consistent with established liver physiological and LCM-based LC-MS measurements (*17*, *29*, *30*). Consistent *in silico* benchmarking on Xenium data showed that reconstruction fidelity (SSIM) scales monotonically with the number of reference markers, increasing from 4 to 25, thereby providing both a quantitative scaling relationship and an orthogonal validation of the HetuNet reconstruction framework. This behavior suggests that HetuNet captures intrinsic relationships between tissue architecture and protein expression, rather than relying on discrete pattern matching (*37*, *38*). Together, these observations support a generalizable principle: anchoring AI-based spatial inference in quantitative, experimentally measured proteomes, while constraining predictions with orthogonal structural references, enables biologically coherent and interpretable reconstruction of spatial proteomic landscapes at single-cell scale.

A notable strength of our method lies in its ability to resolve proteome-defined cellular states that remain inaccessible to targeted imaging-based annotation. In human CRC tissues, conventional CODEX imaging assigned a substantial fraction of cells to an “Unknown” category (**Fig. 5b**, **c**), reflecting intrinsic limitations of nucleus-centered segmentation and marker-restricted annotation, particularly in necrotic regions where nuclear integrity is compromised. By reconstructing spatial protein expression in an untargeted manner, our method enabled functional annotation within necrotic regions and, in turn, revealed distinct proteomic states of neighboring epithelial populations. These epithelial populations proximal to necrotic areas could be resolved into distinct proteome-defined states characterized by coordinated yet differential activation of EMT- and necroptosis-associated pathways. Notably, proteins show considerably greater cell-type specificity than their mRNA counterparts, highlighting the value of proteomic-level analysis (*39*). More broadly, this analysis highlights how deep, quantitative proteome inference can complement and extend antibody-based spatial imaging by uncovering spatial-functional heterogeneity beyond predefined marker panels. By coupling tissue expansion with deep learning-based spatial inference from sparse proteomic measurements, our method overcomes critical constraints of current spatial proteomics and enables deep, scalable and interpretable spatial proteome mapping across complex tissues. This framework establishes a foundation for systematic interrogation of tissue organization and disease-associated spatial programs, with potential applications spanning basic biology, disease mechanisms, and translational research.

While this framework advances spatial proteomics, several limitations should be noted. Spatial reconstruction relies on high-quality IF images as morphological references to provide spatial priors for computational inference. Although ablation analyses (**Fig. S6**) show that coherent predictions can be achieved with a minimal reference marker set, this dependence may limit performance in highly heterogeneous tissues; incorporating alternative spatial reference modalities, together with AI-based translation from histology to molecular features (*40*), may provide more accessible and complementary reference information. In addition, this framework requires LCM and LC-MS acquisition, introducing throughput constraints, particularly when scaling to larger tissue areas. While HetuNet-based orthogonal sampling reduces sampling complexity from a quadratic to a linear scale, increasing spatial resolution toward true single-cell and ultimately subcellular levels still entails trade-offs among resolution, throughput, and proteome depth. These trade-offs could be alleviated through higher tissue expansion factors (*41*, *42*), improved automation of LCM workflows, and continued advances in high-throughput, high-resolution MS (*43*, *44*). With continued technical refinement, our framework could further extend to enable subcellular spatial proteome analysis in intact tissues. Moreover, emerging AI-assisted peptide and protein identification strategies (*45*, *46*) may further increase proteome coverage, and the framework could be extended in the future to additional molecular layers, such as protein glycosylation (*27*). Finally, proteomic profiles are computationally inferred from strip-level measurements rather than obtained from physically isolated single cells, and the use of adjacent tissue sections introduces section-to-section variability. Increased instrumental performance enabling finer-grained, spatially resolved measurements beyond strip-level sampling, together with AI-guided inference and orthogonal sampling strategies applied to the same expanded tissue section, may help reduce inference uncertainty and improve spatial accuracy.

Materials and Methods

Animal care and tissue preparation

All animal experiments complied with the Animal Care Guidelines of Westlake University and were approved by the Institutional Animal Care and Use Committee (IACUC; protocol #24-039-GTN-2). Female C57BL/6J mice (Mus musculus, two months old) were used, and no sex-based comparisons were performed.

As described previously (*47*), mice were housed in a controlled barrier facility at 20-26°C and 40-70% relative humidity under a 12-hour light/dark cycle, with ad libitum access to food and water. Animals were deeply anesthetized with 1% sodium pentobarbital and perfused transcardially with 1× PBS followed by 4% paraformaldehyde (PFA; Sangon, China). Liver tissues were dissected and post-fixed in 4% PFA at 4 °C for 6 h.

Fixed tissues were dehydrated through a graded ethanol series (75%, 95%, and 100%; 30 min each) and infiltrated with paraffin at 60°C. Paraffin blocks were cored using a biopsy punch to obtain samples approximately 1.5 mm in diameter, and consecutive sections (10 µm thick) were prepared using a rotary microtome (RM2255, Leica, Germany) and mounted onto poly-L-lysine-coated glass slides.

Tissue microarray samples from colorectal cancer patients

Archival FFPE CRC tissue microarray (TMA) slides (product no. HCol-Ade075Pre-01) were obtained from Shanghai Outdo Biotech (Shanghai, China). Each slide contained 75 tissue cores representing samples from 25 patients, including adjacent normal tissue, carcinoma *in situ*, and invasive carcinoma. All specimens were histologically validated by the provider. In this study, one invasive carcinoma core was randomly selected, and supplier-provided consecutive FFPE sections (4 µm thick; ~1.5 mm in diameter) were used.

Written informed consent was obtained from all patients. Human tissue procurement was approved by the ethics committee of Shanghai Outdo Biotech (No. SHYJS-CP-1401001). The use of human tissue samples was approved by the ethics committee of Westlake University (Permission number: 20220913GTN001).

Histological staining and multiplexed imaging of TMA sections

Paraffin-embedded sections were deparaffinized by two 10-min incubations in heptane, followed by rehydration through a graded ethanol series (100%, 90%, and 75%; 5 min each) and a final rinse in double-distilled water for 5 min. H&E staining of CRC tissue sections were performed according to the manufacturer's instructions (Shanghai Yuanye Bio-technology, China). Brightfield images were acquired using an Zeiss Axioscan 7 slide scanner equipped with a 20× objective (NA = 0.8). Image acquisition was controlled using ZEISS ZEN (v3.11).

Multiplexed spatial imaging was performed using the PhenoCycler-Fusion platform (Akoya Biosciences, USA) according to the manufacturer's instructions. A CRC-TMA section consecutive to that used for H&E staining was deparaffinized, rehydrated, and subjected to heat-induced antigen retrieval (pH 9.0). The section was incubated with DNA-barcoded antibodies (**Supplementary Table 1**) and subsequently imaged on the PhenoCycler-Fusion system controlled by Fusion software (v2.2.6; Akoya Biosciences).

Tissue expansion

Paraffin-embedded mouse liver sections were dewaxed and rehydrated through a graded ethanol series, followed by rinsing in ddH_2O. H&E-stained human CRC slides were decoverslipped and rehydrated through a graded ethanol series, followed by rinsing in ddH_2O. Subsequent hydrogel-making procedures were performed as previously described (*27*). In brief, slides were incubated in NSA anchoring solution for 1 h at room temperature in the dark and washed three times with 100 mM MOPS anchoring termination buffer. Anchored tissues were infiltrated with an activated monomer solution containing N,N-dimethylacrylamide, sodium methacrylate, and pentaerythritol allyl ether, initiated with ammonium persulfate and N,N,N′,N′-tetramethylethylenediamine. Following overnight incubation at 4 °C, polymerization was carried out at 37 °C for 2 h in a nitrogen-filled vacuum oven. Tissue denaturation was performed using tissue-specific conditions. Mouse liver sections were incubated in 20% SDS denaturation buffer at 95 °C for 8 h. CRC sections were first incubated in 20% SDS denaturation buffer containing 20 mM TCEP at 105 °C in an autoclave for 1 h, followed by an additional 1 h incubation in 20% SDS denaturation buffer without TCEP at 105 °C in an autoclave. Samples were then washed in 1× PBS and subjected to reduction and alkylation.

For immunofluorescence imaging of expanded mouse liver samples (**Supplementary Table 1**), samples were blocked using blocking buffer (HUABIO, China) and incubated overnight at 4 °C with a primary antibody against asialoglycoprotein receptor 1 (ASGR1; 1:200; Proteintech, China). After washing with 1× PBS, samples were incubated with Alexa Fluor 488-conjugated goat anti-rabbit IgG (1:200; Abcam, UK) for 2 h at room temperature. Samples were then incubated with directly conjugated antibodies against glutamine synthetase (CoraLite Plus 647, 1:100; Proteintech) and E-cadherin (Alexa Fluor 555; BD Pharmingen, USA), the latter used at four times the manufacturer's recommended volume. Nuclear counterstaining was performed with DAPI (Abcam) for 30 min. Imaging was conducted using a Nikon CSU-W1 SoRa confocal microscope equipped with a 10× objective (NA 0.45) and controlled by NIS-Elements AR (v5.42.06).

For proteomic visualization, mouse liver and CRC samples were stained with Coomassie Brilliant Blue R-250 (CBB; Sangon) overnight at room temperature and washed in 0.001× PBS to achieve complete expansion. Unexpanded mouse liver sections, as well as CBB-stained tissues after expansion, were imaged using a Zeiss Axio Zoom.V16 fluorescence stereo microscope and processed with ZEN software (v3.1). Isotropic distortion was quantified by registering corresponding pre- and post-expansion images using landmark-based nonrigid image registration implemented in MATLAB (MathWorks, R2024b), as previously described (*47*). Images were prepared and tissue areas were measured in ImageJ (v1.54p) (*48*) prior to registration. The linear expansion factor was calculated as the square root of the ratio of post-expansion to pre-expansion tissue areas.

LCM-based orthogonal sampling

The expanded hydrogel was placed onto an LCM steel frame slide (Leica, Germany) and allowed to dry. LCM was performed using an MMI CellCut system (Molecular Machines & Industries, Germany) controlled by mmiCellTools software (v6.1#578), with samples collected using diffuser caps. Transverse sampling was performed on one section and longitudinal sampling on the adjacent section to generate orthogonal sampling orientations.

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

Collected tissue-hydrogel strips were transferred into 10 µL FAXP tips (*26*, *27*). The C18 membrane was activated with 80% and 100% acetonitrile (ACN) to ensure device consistency and remove potential contaminants, followed by equilibration with double-distilled water and 50 mM ammonium bicarbonate (ABB). Samples were subsequently washed with 50% ACN/50% ABB and dehydrated with 100% ACN prior to digestion. Samples were digested using a trypsin (Hualishi Tech. Ltd, China) working solution (5 ng/µL) prepared in 50 mM ABB containing 10% (v/v) ACN and 0.005% (w/v) n-dodecyl β-D-maltoside (DDM). The digestion volume was adjusted to fully cover the tissue-hydrogel strip, typically ranging from 1-2 µL depending on strip length. Digestion was performed at 37 °C for 12 h. Peptides were sequentially eluted using stepwise increases in ACN concentration (2% and 70% ACN containing 0.1% TFA, followed by 100% ACN) to maximize peptide recovery. Eluates were pooled and vacuum-concentrated. The dried peptide pellet was reconstituted in 7 µL MS buffer (2% ACN, 0.1% FA) and subjected directly to LC-MS/MS analysis.

Proteomics data acquisition

Peptide samples were analyzed on an Orbitrap Astral mass spectrometer equipped with a FAIMS Pro interface and coupled to a Vanquish Neo UHPLC system (Thermo Fisher Scientific, USA). Peptides were loaded onto a trap cartridge and separated on a 15-cm in-house packed analytical column (75 µm inner diameter) operated in backward-wash mode. Solvent A consisted of 0.1% formic acid (FA) in 98% MS-grade water with 2% MS-grade ACN, and solvent B consisted of 0.1% FA in 98% MS-grade ACN with 2% MS-grade water. The chromatographic separation was performed using a 28-min gradient: 3-14% B (0-2 min, 550 nL/min), 14-21% B (2.1-6.1 min, 200 nL/min), and 21-40% B (6.1-22.1 min, 200 nL/min), followed by column washing and re-equilibration to 80% B at 550 nL/min.

The Orbitrap Astral was operated in data-independent acquisition (DIA) mode. Full MS scans were acquired at a resolution of 240,000 over an m/z range of 400-800 with a FAIMS CV of -48 V, RF lens of 45%, normalized AGC target of 500%, maximum injection time of 100 ms, and a carrier gas flow of 3.8 L/min. DIA MS/MS scans covered a precursor mass range of 400-800 m/z and a scan range of 150-2000 m/z using an isolation window of 8 m/z, maximum injection time of 14 ms, normalized collision energy of 26%, RF lens of 45%, normalized AGC target of 800%, and a loop control time of 0.6 s. All measurements were acquired in positive ion mode with a total runtime of 28 min. Data were acquired using Thermo Scientific Xcalibur (v4.7.69.37) and

Orbitrap Astral Tune Application (v1.0.100.40). System performance was monitored using 5 ng mouse liver QC peptides to ensure instrument stability and reproducibility across runs.

Proteomics data analysis

A total of 187 mouse liver spatial samples (96 rows and 91 columns) and 257 human CRC spatial samples (134 rows and 123 columns) were analyzed for protein identification and quantification using DIA-NN (v2.2.0). An in-house library was used for mouse liver studies, whereas DIA data from CRC samples were searched against the previously established pan-human DPHL v.2 spectral library (*49*). The DIA-NN (*50*) analysis was performed with the following parameters: --qvalue 0.01 --cut K*,R* --min-fr-mz 150 --max-fr-mz 2000 --missed-cleavages 2 --min-pep-len 7 --max-pep-len 50 --min-pr-mz 400 --max-pr-mz 800 --min-pr-charge 2 --max-pr-charge 4 --unimod4 --var-mods 1 --var-mod UniMod:35,15.994915,M --relaxed-prot-inf --rt-profiling --no-norm --matrices --verbose 1 --reanalyse. Proteins consistently detected in at least 70% of the samples, considering both row and column datasets, were selected for HetuNet training.

Overview of the HetuNet algorithm

To resolve single-cell-scale spatial protein expression maps from strip-resolved MS measurements and high-resolution morphological guidance, we developed HetuNet. This weakly supervised deep learning framework establishes a mapping function $\mathcal{F}_{\text{HetuNet}} : \mathcal{P} \rightarrow \mathbb{R}$, where $\mathcal{P}$ represents the space of high-resolution morphological image patches and $\mathbb{R}$ denotes the space of scalar protein abundance values. The model translates local morphological features into quantitative protein levels, constrained strictly by global spatial marginals.

Let $\mathbf{I} \in \mathbb{R}^{H' \times W' \times C}$ denote the input high-resolution multi-channel IF image, where H' and W' are the spatial dimensions and C represents the number of IF channels. The target protein expression is modeled as a latent matrix $\mathbf{X} \in \mathbb{R}^{H \times W}$, defined on a reconstruction grid where H and W are the height and width, respectively.

HetuNet operates under conditions where pixel-wise ground truth for $\mathbf{X}$ is inaccessible. Instead, supervision is provided by the observed row marginals $\mathbf{y}^R \in \mathbb{R}^H$ and column marginals $\mathbf{y}^R \in \mathbb{R}^W$. These marginals are defined as:

$$y_i^R = \sum_{j=1}^{W} \mathbf{X}_{i,j}, \quad y_j^C = \sum_{i=1}^{H} \mathbf{X}_{i,j}$$

where i and j are indices for rows and columns, and $\mathbf{X}_{i,j}$ represents the protein abundance at grid location (i, j). HetuNet aims to estimate the optimal reconstruction $\widehat{\mathbf{X}}$ such that its projections minimize the discrepancy with $\mathbf{y}^R$ and $\mathbf{y}^C$.

Pathway enrichment and functional analysis in mouse liver

Differential expression analysis was first performed between spatial regions (e.g., CV, PV, and Others) to identify region-specific protein markers. For each comparison, proteins were ranked according to their $\log_2$ fold-change values. Gene set enrichment analysis (GSEA) was then conducted using the gseapy (*51*) package (Python implementation of GSEA), with the KEGG 2021 Mouse and GO Biological Process 2021 libraries as reference databases. The normalized enrichment scores (NES) were calculated to assess the statistical significance of pathway enrichment. Terms with a false discovery rate (FDR q-value) < 0.05 were considered significantly enriched.

To visualize spatial trends of enriched pathways, we integrated the region annotation information into the enrichment results. For representative pathways, the normalized enrichment scores across regions were plotted as trend distributions or boxplots to reveal spatially ordered activation patterns (e.g., portal-to-central gradient). Functional modules with consistent spatial gradients were further examined to interpret region-specific biological processes, such as metabolic reprogramming and cell-type-specific signaling pathways.

In silico validation and model training

To assess the spatial reconstruction fidelity of HetuNet, an *in silico* simulation was performed utilizing a publicly available high-resolution spatial transcriptomics dataset derived from human renal cell carcinoma (*31*). A ROI measuring 2.4 × 1.8 mm was selected for validation purposes. IF features utilized for training underwent rigorous preprocessing to mitigate outliers and standardize input data. Specifically, pixel intensity values exceeding the 99.9th percentile within each channel were replaced with the median value of that respective channel. Subsequently, all channel values were normalized to a range of 0-255.

Prior to network training, the transcriptomic data were filtered to exclude genes with low expression levels based on a total count threshold (sum of rows and columns < 1,000). This filtration step reduced the initial dataset from 405 genes to 289 genes (excluding 116 genes). All 289 retained genes were utilized as input for the training of HetuNet. For specific downstream analyses, highly variable genes (HVGs) were identified using the *highly_variable_genes* function within the Scanpy suite (v1.9.8), with the *flavor* parameter set to 'seurat_v3' and *subset* set to *True*.

HetuNet was trained using consistent hyperparameters across all features. The model configuration was established as follows: patch size = 13, batch size = 128, base learning rate = 0.001, and training duration = 100 epochs. Additional model-specific parameters included a CNN learning rate fold (*cnn_lr_fold*) of 16/30, a total variation regularization weight (*tv_lambda*) of 0.0001, and a correlation weight (*correlation_lambda*) of 5. Training was performed on an NVIDIA A100 GPU. To quantify the reconstruction accuracy, the Structural Similarity Index Measure (SSIM) was calculated utilizing the *structural_similarity* function from the scikit-image library (v0.25.2) in Python (v3.12.12).

CODEX image segmentation and cell annotation

Nuclei were segmented based on DAPI staining intensity using the StarDist algorithm (v0.9.2). Subsequent morphological extraction and coordinate identification were performed using the *regionprops* function within the scikit-image library (v0.25.2). For each segmented cell, the mean fluorescence intensity was calculated across 31 protein marker channels and the DAPI channel. These quantitative data were converted into the AnnData format and subjected to analysis using Scanpy (v1.11.5). Preprocessing steps included total count normalization (*target_sum* = 1e4) and data scaling. Dimensionality reduction was performed via Principal Component Analysis (PCA) with *n_comps* = 15. A nearest-neighbor graph was constructed (*n_neighbors* = 10, *n_pcs* = 10), followed by Leiden clustering (*resolution* = 0.3, *random_state* = 888) and Uniform Manifold Approximation and Projection (UMAP) for visualization. Cell types were annotated based on the differential expression of lineage-specific markers within the resulting clusters(*52*, *53*). Specifically, clusters C1 and C4 were identified as epithelial cells; clusters C2, C3, C8 and C9 as fibroblasts; and clusters C6 and C7 as immune cells. Cluster C5 remained unclassified ("Unknown") (**Fig. 5b**).

Analysis of HetuNet-predicted protein profiles in CRC samples

HetuNet-predicted protein profiles were converted to AnnData format for downstream analysis. Prior to downstream analysis, proteins assigned to the “Human skin and hair” category were removed as potential contaminants based on the contaminant libraries (*54*). To ensure data quality, cell patches exhibiting extreme total counts (values ≤ 1st percentile or ≥ 99th percentile) were removed. Furthermore, to mitigate the impact of outliers, expression values for each protein exceeding the 99th percentile were clipped to the 99th percentile value.

The filtered dataset was analyzed using Scanpy (v1.11.5). Data were normalized (*target_sum* = 1e4) and log-transformed (*log1p*). Highly variable features were selected using the *highly_variable_genes* function (*flavor* = "seurat", *n_top_genes* = 2000). Following PCA, a neighborhood graph was constructed (*n_neighbors* = 10, *n_pcs* = 30), and cells were clustered using the Leiden algorithm (*resolution* = 1.0, *random_state* = 666). UMAP was employed for dimensionality reduction and visualization. Cell type annotation was performed based on marker expression profiles (*52*, *53*) and validated by clinical experts. Clusters were assigned as follows: C1, C2, C6, and C16 as immune cells; C3, C4, C5, C11, C12, and C13 as epithelial cells; C7, C8, C14, and C15 as fibroblasts; and C9 and C10 as necrotic cells (**Fig. 5e**).

Epithelial state analysis in CRC samples

Pathway enrichment at the single-cell level was quantified using Single Sample Gene Set Enrichment Analysis (ssGSEA). Calculations were performed with the *gseapy* (version 1.1.11) Python library using the “custom” sample normalization strategy. Two curated pathway collections were employed as references: KEGG 2021 Human and MSigDB Hallmark 2020. The resulting Normalized Enrichment Scores (NES) were quality-controlled and integrated into the single-cell metadata matrix to enable downstream stratification and comparative analyses.

A spatially defined Region of Interest (ROI), corresponding to the area shown in **Fig. 5g**, was extracted for downstream analysis. Within this region, cells annotated as epithelial were selected for functional state classification. Two epithelial states were defined according to ssGSEA pathway activities: **State E1**: cells with positive enrichment of the Epithelial–Mesenchymal Transition (EMT) pathway (NES > 0) and negative enrichment of the Necroptosis pathway (NES < 0); **State E2**: cells with negative EMT enrichment (NES < 0) and positive Necroptosis enrichment (NES > 0).

Statistical significance was evaluated using a two-sided Mann–Whitney U test applied to non-zero expression values. Resulting p-values were corrected for multiple testing using the Benjamini–Hochberg False Discovery Rate (FDR) procedure.

References

1. S. Uderhardt, G. Neag, R. N. Germain, Dynamic Multiplex Tissue Imaging in Inflammation Research. *Annual Review of Pathology: Mechanisms of Disease* **19**, 43–67 (2024).
2. O. Elhanani, R. Ben-Uri, L. Keren, Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* **41**, 404–420 (2023).
3. R. dos Santos Peixoto, B. F. Miller, M. A. Brusko, *et al.*, Characterizing cell-type spatial relationships across length scales in spatially resolved omics data. *Nat Commun* **16**, 350 (2025).
4. Y. Goltsev, N. Samusik, J. Kennedy-Darling, *et al.*, Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
5. J.-R. Lin, M. Fallahi-Sichani, P. K. Sorger, Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat Commun* **6**, 8390 (2015).
6. J.-R. Lin, Y.-A. Chen, D. Campton, *et al.*, High-plex immunofluorescence imaging and traditional histology of the same tissue section for discovering image-based biomarkers. *Nat Cancer* **4**, 1036–1052 (2023).
7. Y. Liu, Y. Dai, L. Wang, Spatial omics at the forefront: emerging technologies, analytical innovations, and clinical applications. *Cancer Cell* **44**, 24–49 (2026).
8. D. Bressan, G. Battistoni, G. J. Hannon, The dawn of spatial omics. *Science* **381**, eabq4964 (2023).
9. Y. Cheng, S. Dang, Y. Zhang, *et al.*, Sequencing-free whole-genome spatial transcriptomics at single-molecule resolution. *Cell* **188**, 6953-6970.e12 (2025).
10. Y. Harnik, L. Buchauer, S. Ben-Moshe, *et al.*, Spatial discordances between mRNAs and proteins in the intestinal epithelium. *Nat Metab* **3**, 1680–1693 (2021).
11. M. Gry, R. Rimini, S. Strömberg, *et al.*, Correlations between RNA and protein expression profiles in 23 human cell lines. *BMC Genomics* **10**, 365 (2009).
12. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).
13. Z. Bai, D. Zhang, Y. Gao, *et al.*, Spatially exploring RNA biology in archival formalin-fixed paraffin-embedded tissues. *Cell* **187**, 6760-6779.e24 (2024).
14. T. Guo, J. A. Steen, M. Mann, Mass-spectrometry-based proteomics: from single cells to clinical applications. *Nature* **638**, 901–911 (2025).
15. F. A. Rosenberger, S. C. Mädler, K. H. Thorhauge, *et al.*, Deep Visual Proteomics maps proteotoxicity in a genetic liver disease. *Nature* **642**, 484–491 (2025).
16. A. Mund, F. Coscia, A. Kriston, *et al.*, Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat Biotechnol* **40**, 1231–1240 (2022).
17. F. A. Rosenberger, M. Thielert, M. T. Strauss, *et al.*, Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat Methods* **20**, 1530–1536 (2023).
18. A. G. Bury, A. Pyle, F. Marcuccio, *et al.*, A subcellular cookie cutter for spatial genomics in human tissue. *Anal Bioanal Chem* **414**, 5483–5492 (2022).

19. H. Chen, Y. Zhang, H. Zhou, *et al.*, Routine Workflow of Spatial Proteomics on Micro-formalin-Fixed Paraffin-Embedded Tissues. *Anal. Chem.* **95**, 16733–16743 (2023).

20. F. Chen, P. W. Tillberg, E. S. Boyden, Expansion microscopy. *Science*, doi: 10.1126/science.1260088 (2015).

21. G. Wen, V. Leen, T. Rohand, *et al.*, Current Progress in Expansion Microscopy: Chemical Strategies and Applications. *Chem. Rev.* **123**, 3299–3323 (2023).

22. F. Wang, C. Sun, T. W. Wu, *et al.*, iPEX enables micrometre-resolution deep spatial proteomics via tissue expansion. *Nature* **649**, 505–514 (2026).

23. J. Taylor, J. K. Lukowski, C. R. Anderton, Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics. *J Am Soc Mass Spectrom* **32**, 872–894 (2021).

24. B. Hu, R. He, K. Pang, *et al.*, High-resolution spatially resolved proteomics of complex tissues based on microfluidics and transfer learning. *Cell* **188**, 734-748.e22 (2025).

25. R. Qin, J. Ma, F. He, *et al.*, In-depth and high-throughput spatial proteomics for whole-tissue slice profiling by deep learning-facilitated sparse sampling strategy. *Cell Discov* **11**, 21 (2025).

26. Z. Dong, W. Jiang, C. Wu, *et al.*, Spatial proteomics of single cells and organelles on tissue slides using filter-aided expansion proteomics. *Nat Commun* **15**, 9378 (2024).

27. Z. Dong, C. Wu, J. Chen, *et al.*, Filter-aided expansion proteomics for the spatial analysis of single cells and organelles in FFPE tissue samples. *Nat Protoc*, 1–25 (2025).

28. A. Subramanian, P. Tamayo, V. K. Mootha, *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).

29. S. Ben-Moshe, S. Itzkovitz, Spatial heterogeneity in the mammalian liver. *Nat Rev Gastroenterol Hepatol* **16**, 395–410 (2019).

30. W. Birchmeier, Orchestrating Wnt signalling for metabolic liver zonation. *Nat Cell Biol* **18**, 463–465 (2016).

31. Xenium In Situ Gene and Protein Expression data for FFPE Human Renal Cell Carcinoma, *10x Genomics*. https://www.10xgenomics.com/datasets/xenium-protein-ffpe-human-renal-carcinoma.

32. A. Schroeder, M. L. Loth, C. Luo, *et al.*, Scaling up spatial transcriptomics for large-sized tissues: uncovering cellular-level tissue architecture beyond conventional platforms with iSCALE. *Nat Methods*, 1–12 (2025).

33. A. Mund, A.-D. Brunner, M. Mann, Unbiased spatial proteomics with single-cell resolution in tissues. *Molecular Cell* **82**, 2335–2349 (2022).

34. A. Goossen, L. Lagache, M. Salzet, *et al.*, MS-based spatial proteomics: deciphering tissue microenvironment by bridging MS imaging and large-scale proteomics. *Trends Chem* **7**, 659–677 (2025).

35. D. Zhang, L. A. Rubio Rodríguez-Kirby, Y. Lin, *et al.*, Spatial dynamics of brain development and neuroinflammation. *Nature* **647**, 213–227 (2025).

36. Z. Dong, W. Jiang, T. Guo, Tissue expansion enables proteomics at improved spatial resolution. *Trends in Biochemical Sciences* **50**, 945–946 (2025).

37. Z. Li, Y. Li, J. Xiang, *et al.*, AI-enabled virtual spatial proteomics from histopathology for interpretable biomarker discovery in lung cancer. *Nat Med* **32**, 231–244 (2026).

38. E. Wu, M. Bieniosek, Z. Wu, *et al.*, ROSIE: AI generation of multiplex immunofluorescence staining from histopathology images. *Nat Commun* **16**, 7633 (2025).

39. T. Wu, L. Jiang, T. Mukhtar, *et al.*, Single-cell proteomic landscape of the developing human brain. *Nat Biotechnol*, 1–14 (2026).

40. J. M. J. Valanarasu, H. Xu, N. Usuyama, *et al.*, Multimodal AI generates virtual population for tumor microenvironment modeling. *Cell*, doi: 10.1016/j.cell.2025.11.016 (2025).

41. J.-B. Chang, F. Chen, Y.-G. Yoon, *et al.*, Iterative expansion microscopy. *Nat Methods* **14**, 593–599 (2017).

42. S. Wang, T. W. Shin, H. B. Yoder, *et al.*, Single-shot 20-fold expansion microscopy. *Nat Methods* **21**, 2128–2134 (2024).

43. U. H. Guzman, M. Rykar, I. A. Hendriks, *et al.*, Higher-Throughput Proteome Profiling Enabled by Parallelized Pre-Accumulation and Optimized Ion Processing in the Orbitrap Astral Zoom Mass Spectrometer. *Molecular & Cellular Proteomics*, 101504 (2026).

44. Z. Dong, W. Xiang, W. Jiang, *et al.*, Expansion omics: from expansion microscopy to spatial omics. *Mol Syst Biol*, doi: 10.1038/s44320-025-00171-9 (2025).

45. Z. Liu, P. Liu, Y. Sun, *et al.*, DIA-BERT: pre-trained end-to-end transformer models for enhanced DIA proteomics data analysis. *Nat Commun* **16**, 3530 (2025).

46. G. Wallmann, P. Skowronek, V. Brennsteiner, *et al.*, AlphaDIA enables DIA transfer learning for feature-free proteomics. *Nat Biotechnol*, 1–10 (2025).

47. L. Li, C. Sun, Y. Sun, *et al.*, Spatially resolved proteomics via tissue expansion. *Nat Commun* **13**, 7242 (2022).

48. J. Schindelin, I. Arganda-Carreras, E. Frise, *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).

49. Z. Xue, T. Zhu, F. Zhang, *et al.*, DPHL v.2: An updated and comprehensive DIA pan-human assay library for quantifying more than 14,000 proteins. *Patterns* **4** (2023).

50. V. Demichev, C. B. Messner, S. I. Vernardis, *et al.*, DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, 41–44 (2020).

51. Z. Fang, X. Liu, G. Peltz, GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757 (2023).

52. D. Fawkner-Corbett, A. Antanaviciute, K. Parikh, *et al.*, Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* **184**, 810-826.e23 (2021).

53. C. Hu, T. Li, Y. Xu, *et al.*, CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* **51**, D870–D876 (2023).

54. A. M. Frankenfield, J. Ni, M. Ahmed, *et al.*, Protein Contaminants Matter: Building Universal Protein Contaminant Libraries for DDA and DIA Proteomics. *J. Proteome Res.* **21**, 2104–2113 (2022).

Acknowledgments: This work is supported by grants from the National Natural Science Foundation of China (Major Research Plan, 92259201), the National Natural Science Foundation of China (Key Joint Research Program) (grant no. U24A20476), the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0533300), the National Key R&D Program of China (2021YFA1301600), Zhejiang Provincial Natural Science Foundation of China (LMS26C050003) and the State Key Laboratory of Medical Proteomics (SKLP-Y202405). **Fig.1a**, **b**, **Fig.3a**, **Fig.4a** and **Fig.S1** were created with BioRender.com and icons in Fig.4a-c were made by Nano Banana Pro. We thank Q. Zhang for maintaining the Astral instrument. We thank the Microscopy Core Facility of Westlake University for the facility support and thank technicians F. Xiao, M. Liao and G. Fang for technical assistance. We thank Dr. J. Xiang from the Affiliated Hangzhou First People's Hospital of Westlake University for expert pathological annotation of the CRC H&E section. The language of the initial draft was refined using ChatGPT before undergoing proofreading by co-authors.

Author contributions: T.G. conceived and supervised the study. Z.D. co-conceived the study and led method development and experimental implementation. Y.C. and S.W. contributed to algorithm design and benchmarking. C.W., Z.D., and J.C. performed expansion proteomics analyses. S.W., Z.D., C.L., and X.L. conducted data analysis. J.S. performed the CODEX experiment. S.W. and Z.D. drafted the manuscript. T.G. provided critical revisions. All authors approved the final version.

Competing interests: T.G. is the shareholder of Westlake Omics (Hangzhou) Biotechnology Co., Ltd. The other author declares no competing interests.

Figure 1

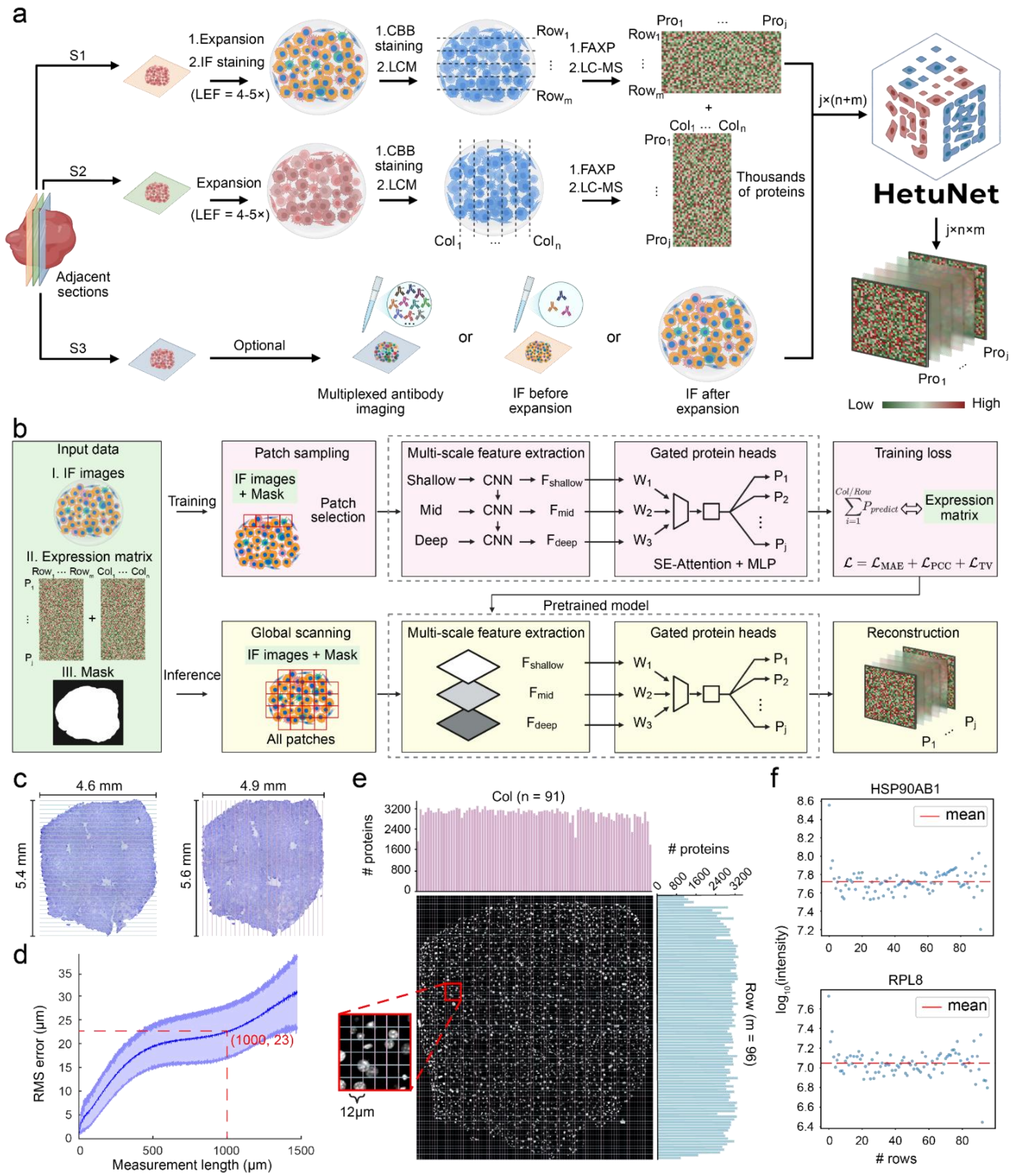


Fig. 1. Workflow and experimental validation. a, This workflow integrates tissue expansion, orthogonal laser capture microdissection (LCM) after Coomassie Brilliant Blue (CBB) staining, filter-aided sample preparation with LC-MS data acquisition, and HetuNet-based spatial inference. The linear expansion factor (LEF) was calculated by comparing sample dimensions before and after expansion. Tissue expansion enables near single-cell spatial resolution and supports multiple imaging references, including multiplexed antibody imaging,

immunofluorescence (IF) before expansion, and high-resolution IF after expansion, which together facilitate HetuNet reconstruction. Following HetuNet reconstruction, thousands of proteins are mapped back onto the whole tissue section. **b**, Schematic of the HetuNet algorithm for spatial inference from multi-modal input data. **c**, Bright-field images of two consecutive expanded mouse liver FFPE sections stained with CBB. Blue and red lines indicate the LCM cutting directions. **d**, Root-mean-square (RMS) measurement length errors quantified before and after expansion in liver section images. The blue line indicates the mean value, and the shaded area represents the standard deviation. **e**, Proteome coverage across all spatial samples, shown as the number of identified proteins. Nuclei are stained with DAPI, and a zoomed-in region of interest highlights fine structural details. Each spatial patch corresponds to a width of 12 μm. **f**, Quality control of row-wise samples based on housekeeping protein expression. Scatter plots show the log10-transformed expression levels of HSP90AB1 and RPL8 across individual row samples. Protein intensities were normalized by the proportion of IF image pixels for each sample to correct for variations in sampled tissue area. To ensure comparability, raw expression values were normalized by the proportion of pixels captured in the IF image data for each sample, correcting for variations in tissue area. The x-axis corresponds to row sample IDs.

Figure 2

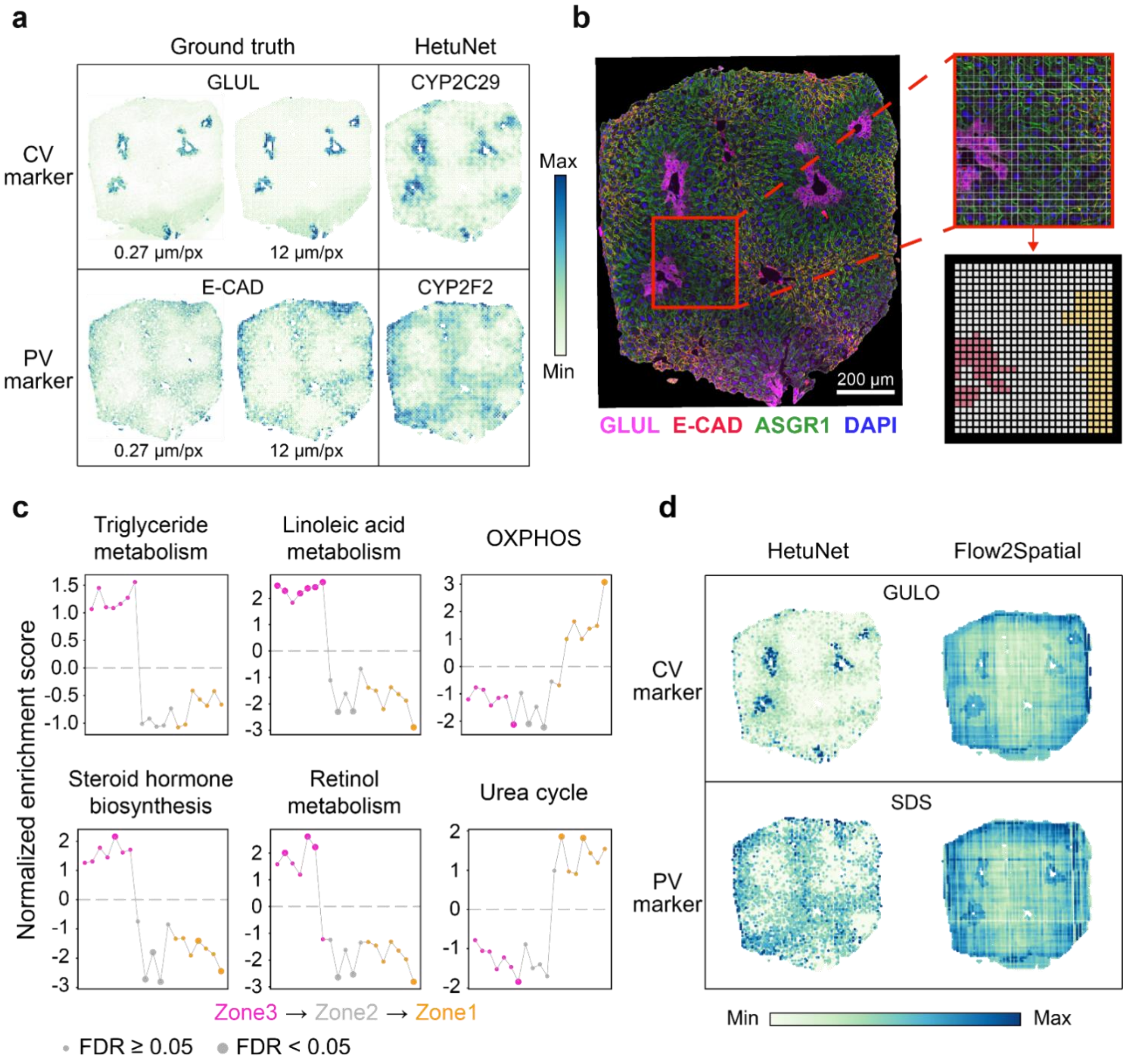
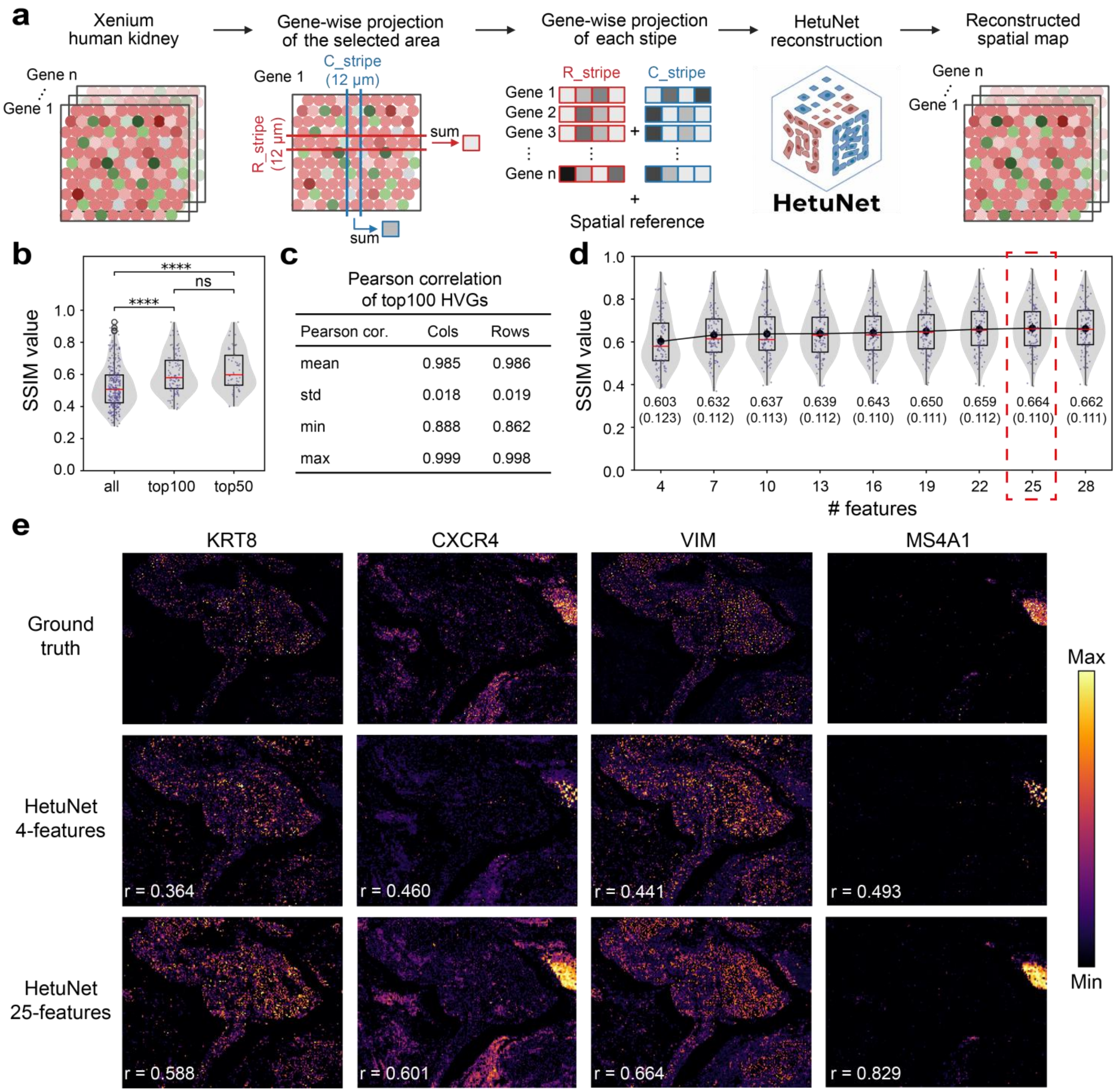


Fig. 2. Model validation using zonated hepatocytes. a, Visualization of spatial protein abundance derived from immunofluorescence (IF) staining (left, ground truth) and HetuNet reconstructions (right). For the ground-truth IF data (left), the panel is further displayed at two resolutions: the original IF image (left) and a downsampled version matching the HetuNet reconstruction resolution (12 μm per pixel; right). GLUL and E-CAD serve as central vein (CV) and portal vein (PV) markers, respectively. CYP2C29 represents a CV-enriched protein, and CYP2F2 represents a PV-enriched protein. **b,** Multiplexed IF images of GLUL, E-CAD, ASGR1, and DAPI acquired from the same tissue section. A representative zonated region is shown as a zoomed-in view. Based on the spatial expression patterns of these IF markers, all spatial patches (12 × 12 μm) within the zoomed region were manually annotated as CV-like (Zone 3; pink), PV-like (Zone 1; yellow), or remaining regions (gray). These annotations were used as reference spatial labels for downstream analyses. Scale bar is provided. **c,** Gene set enrichment analysis comparing proteins enriched in individual spatial patches with those in all other patches,

summarized by normalized enrichment scores. Enrichment patterns are shown along the CV-PV axis (Zone 3 → Zone 2 → Zone 1). Dot size indicates statistical significance after multiple testing correction. **d,** Compatibility of Flow2Spatial(*24*) with the orthogonal strip-based input format used in this study. The heatmap illustrates the structural mismatch between the Flow2Spatial framework and the orthogonal strip-based input format, indicating limited compatibility with the current data structure. GLUL and SDS are shown as representative CV- and PV-enriched proteins, respectively.

Figure 3



Pearson correlation of top100 HVGs

Pearson cor.	Cols	Rows
mean	0.985	0.986
std	0.018	0.019
min	0.888	0.862
max	0.999	0.998

Fig. 3. Validation of HetuNet performance via *in silico* simulation. a, Validation of the HetuNet algorithm using an image-based spatial transcriptomics dataset generated by 10x Genomics Xenium profiling of human renal cell carcinoma tissue. To simulate the sampling strategy of laser capture microdissection, the tissue was sliced *in silico* along two orthogonal axes. Along each axis, the tissue was partitioned into consecutive 12 μm pseudostrips spanning the selected region. Within each pseudostrip, gene expression was obtained by summing the expression values of all spatial spots covered by that strip. HetuNet was then applied to reconstruct spatial gene expression patterns from these aggregated inputs. **b,** Distribution of structural similarity index measure (SSIM) values across different groups of highly variable genes (HVGs). **c,** Pearson correlation analysis comparing HetuNet-predicted and ground-truth row- or column-aggregated expression for the top 100 HVGs. **d,** Evaluation of model robustness, showing SSIM values of the top 100 HVGs as a function of the number of reference markers

selected from the Xenium protein panel used during training. The red lines in the boxplots indicate the median; the black dots connected across boxplots represent the mean; the bars denote the standard deviation (SD); and the numbers below each boxplot show the mean (SD). **e,** Comparative heatmaps illustrating the spatial fidelity of HetuNet-reconstructed expression maps relative to the Xenium-derived ground truth for representative genes (KRT8, CXCR4, VIM, and MS4A1). The top row shows the Xenium ground truth, the middle row shows the HetuNet reconstructions trained with four Xenium protein panel markers, and the bottom row shows the HetuNet reconstructions trained with 25 Xenium protein panel markers. The value r represents the Pearson correlation coefficient calculated globally across all spatial patches.

Figure 4

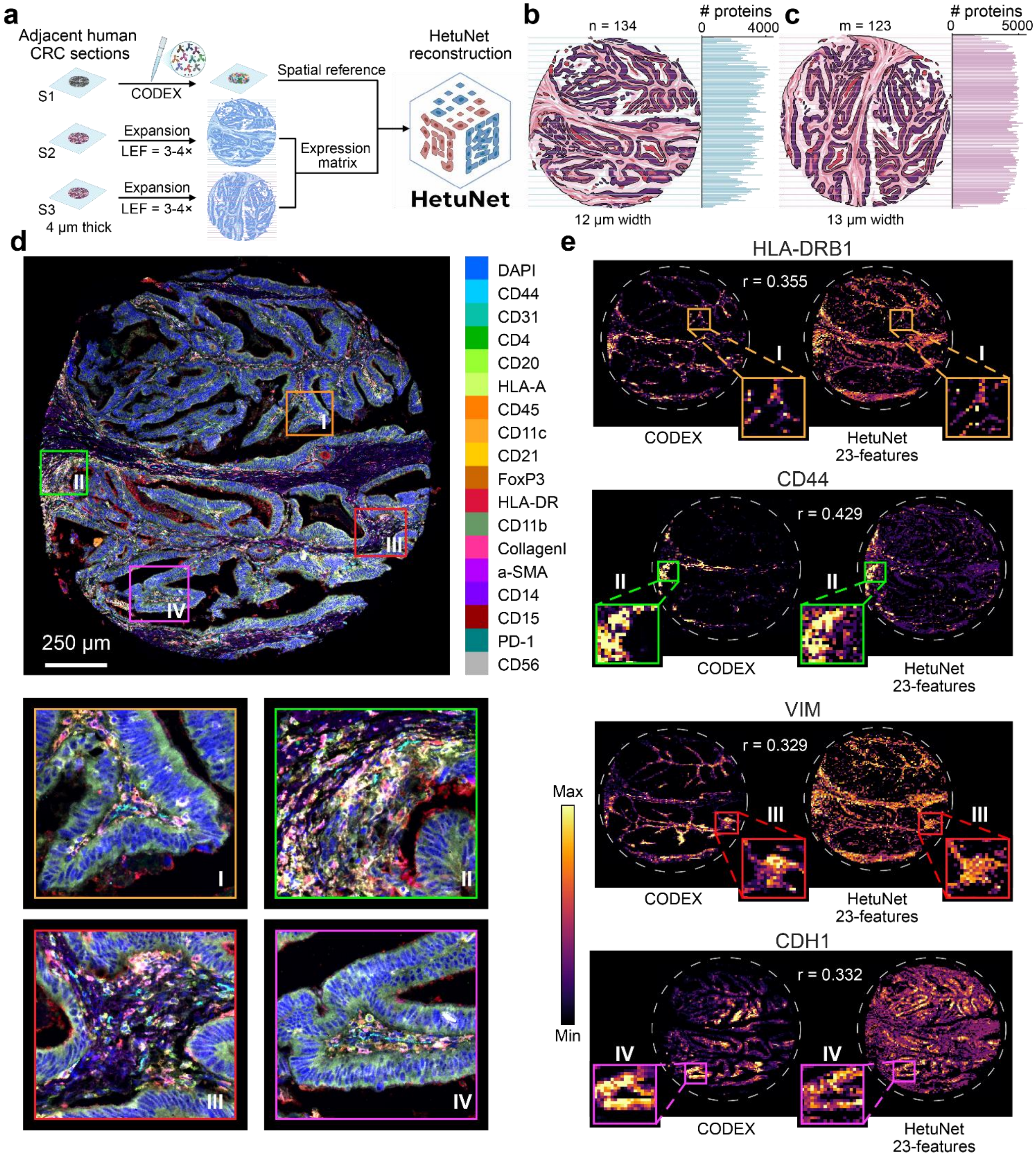


Fig. 4. Application of FFPE human colorectal cancer tissues. a, Schematic workflow illustrating the experimental design in human colorectal cancer (CRC) FFPE tissues using three consecutive tissue sections. One section was subjected to multiplexed immunofluorescence staining using CODEX with 31 protein markers and nuclear staining. The other two sections were processed by tissue expansion, followed by orthogonal laser capture microdissection-based spatial proteomic profiling, and subsequently used for spatial protein prediction by HetuNet, together with spatial references derived from the CODEX results. **b-c,** Overview of protein

identification across two adjacent CRC tissue sections. Proteins were measured along orthogonal stripe-based sampling directions, with one section sampled along the horizontal axis (**b**) and the other along the vertical axis (**c**; 90° rotated). Each stripe represents a narrow tissue band (**b**, n = 134 stripes with 12 μm width; **c**, m = 123 stripes with 13 μm width), demonstrating comparable protein coverage and spatial consistency between the two sections. **d,** Representative multiplexed immunofluorescence images of human CRC tissues acquired using CODEX. Zoomed-in insets indicate heterogeneous regions of interest (ROIs). Scale bar is included. **e,** Spatial protein expression patterns of four representative markers (HLA-DRB1, CD44, Vim, and CDH1). For each marker, the left column shows protein expression measured by CODEX (ground truth), and the right column shows the corresponding HetuNet-predicted protein expression matrices inferred from a curated set of 23 CODEX-derived spatial reference features (**Fig. S8**), with HLA-DRB1, CD44, Vim, and CDH1 withheld for validation. Colored bounding boxes indicate four ROIs, corresponding to the frames shown in panel **d**. Zoomed-in views compare ground truth and HetuNet-predicted protein expression within each ROI, illustrating the spatial patterns and local heterogeneity captured by the model.

Figure 5

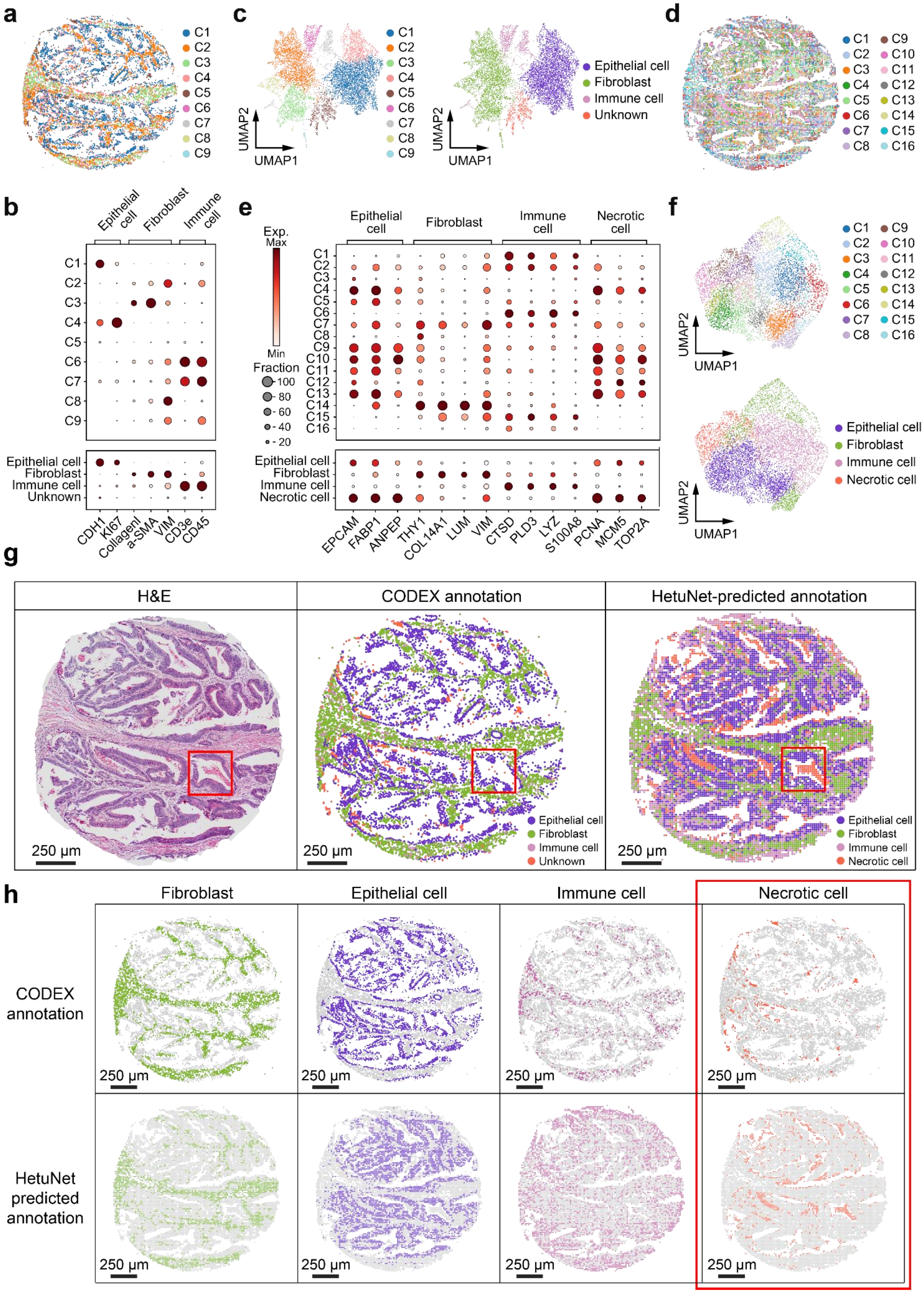

a
C1
C2
C3
C4
C5
C6
C7
C8
C9
c
UMAP2
UMAP1
Epithelial cell
Fibroblast
Immune cell
Unknown
d
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15 C16
b
Epithelial cell
Fibroblast
Immune cell
C1
C2
C3
C4
C5
C6
C7
C8
C9
Epithelial cell
Fibroblast
Immune cell
Unknown
CDH1
KI67
CollagenI
a-SMA
VIM
CD3e
CD45
e
Exp.
Max
Min
Fraction
100
80
60
40
20
Epithelial cell
Fibroblast
Immune cell
Necrotic cell
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15 C16
Epithelial cell
Fibroblast
Immune cell
Necrotic cell
EPCAM
FABP1
ANPEP
THY1
COL14A1
LUM
VIM
CTSD
PLD3
LYZ
S100A8
PCNA
MCM5
TOP2A
f
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15 C16
UMAP2
UMAP1
Epithelial cell
Fibroblast
Immune cell
Necrotic cell
g
H&E
CODEX annotation
HetuNet-predicted annotation
250 µm
Epithelial cell
Fibroblast
Immune cell
Unknown
Epithelial cell
Fibroblast
Immune cell
Necrotic cell
h
Fibroblast
Epithelial cell
Immune cell
Necrotic cell
CODEX annotation
HetuNet predicted annotation
250 µm

Fig. 5. Our workflow reveals additional proteome-defined cell states beyond CODEX imaging. a, Leiden clustering of single cells based on the expression of 31 CODEX protein markers and nuclear staining, revealing nine clusters across the tissue section. **b,** Dot plot showing canonical marker expression across CODEX-derived clusters, used for manual cell-type annotation and cluster merging. **c,** UMAP visualization of CODEX single-cell data before (left) and after (right) manual cell-type annotation. **d,** Leiden clustering of single cells based on HetuNet-predicted protein expression profiles, identifying sixteen clusters across the same tissue section. **e,** Dot plot showing marker expression patterns across HetuNet-predicted protein clusters, used for manual cell-type annotation and cluster merging. **f,** UMAP visualization of HetuNet-predicted single-cell proteomes before (top) and after (bottom) manual cell-type annotation. **g,** Spatial overview of a colorectal cancer tissue section. Left, H&E staining. Middle, cell-type annotation based on CODEX imaging, identifying epithelial cells, fibroblasts, immune cells, and an unclassified (Unknown) population. Right, cell-type annotation derived from HetuNet-predicted single-cell proteomes, identifying epithelial cells, fibroblasts, immune cells, and a necrotic cell state. **h,** Comparison of spatial distributions of cell types identified by CODEX imaging and HetuNet-predicted proteomes. The CODEX “Unknown” population shows a spatial pattern highly similar to the necrotic cell state identified from predicted proteomic data. Scale bars are provided.

Figure 6

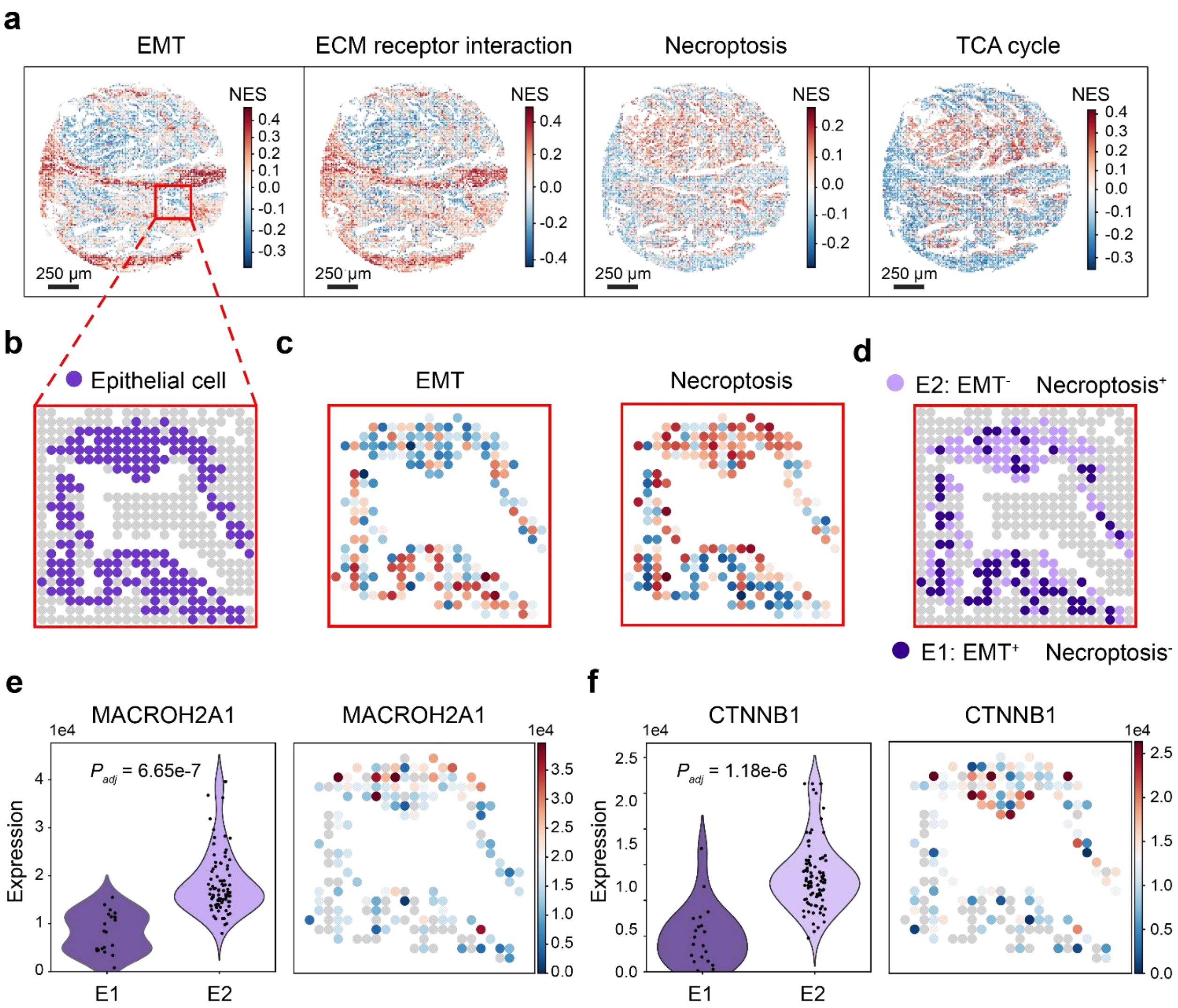


Fig. 6. Our workflow resolves epithelial cell state divergence around necrotic regions. a, Spatial single-sample gene set enrichment analysis (ssGSEA) performed on HetuNet-predicted single-cell proteomes. Spatial activation patterns of selected pathways, including epithelial-mesenchymal transition (EMT), extracellular matrix (ECM) receptor interaction, necroptosis, and tricarboxylic acid (TCA) cycle, are shown. **b,** Regional dissection of epithelial cell states in areas proximal to necrotic regions. **c-d,** Based on ssGSEA scores of EMT and necroptosis pathways (**c**), epithelial cells were stratified into two functional states: E1, characterized by elevated EMT and reduced necroptosis scores; and E2, characterized by reduced EMT and elevated necroptosis scores (**d**). **e-f,** Violin plot of non-zero MACROH2A1 (**e**, left) and CTNNB1 (**f**, left) expression across E1 and E2 cells, with violins colored by epithelial state and individual cells overlaid as black dots. The corresponding spatial distributions of MACROH2A1 (**e**, right) and CTNNB1 (**f**, right) expression are shown within the tissue.