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

Figure S1

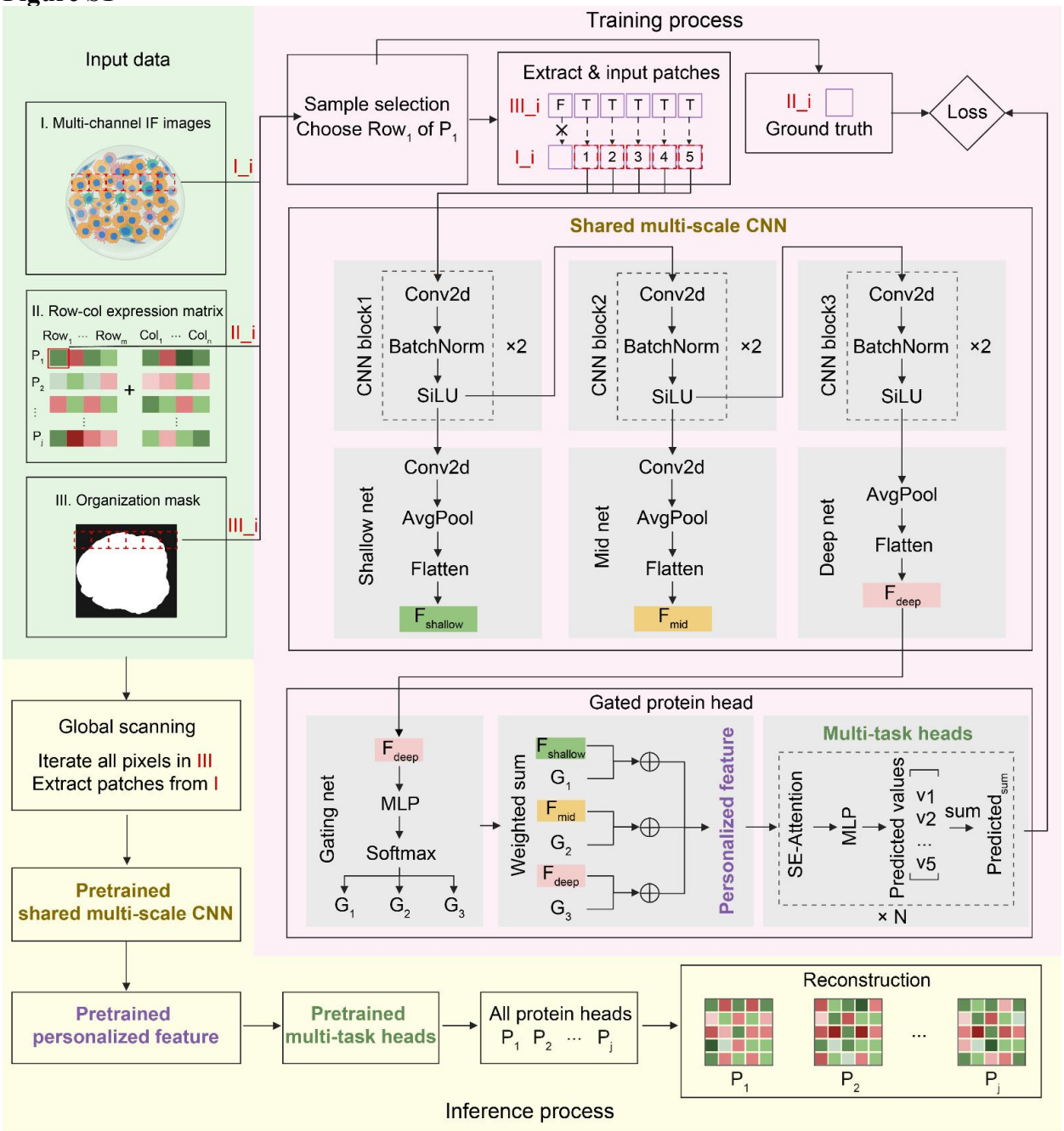

Fig. S1. Schematic overview of HetuNet framework for spatial protein expression reconstruction. Multi-channel immunofluorescence (IF) images, sparse protein expression measurements, and a tissue mask are provided as model inputs. During training, local image regions within the tissue mask are sampled and processed by a shared convolutional neural network to extract multiscale spatial features. These features are adaptively integrated via a gating mechanism to generate location-specific representations, which are then passed to protein-specific prediction heads to infer expression levels for individual proteins. During inference, the trained model is applied across the entire tissue area to reconstruct spatially resolved protein expression maps for all measured proteins.

Figure S2

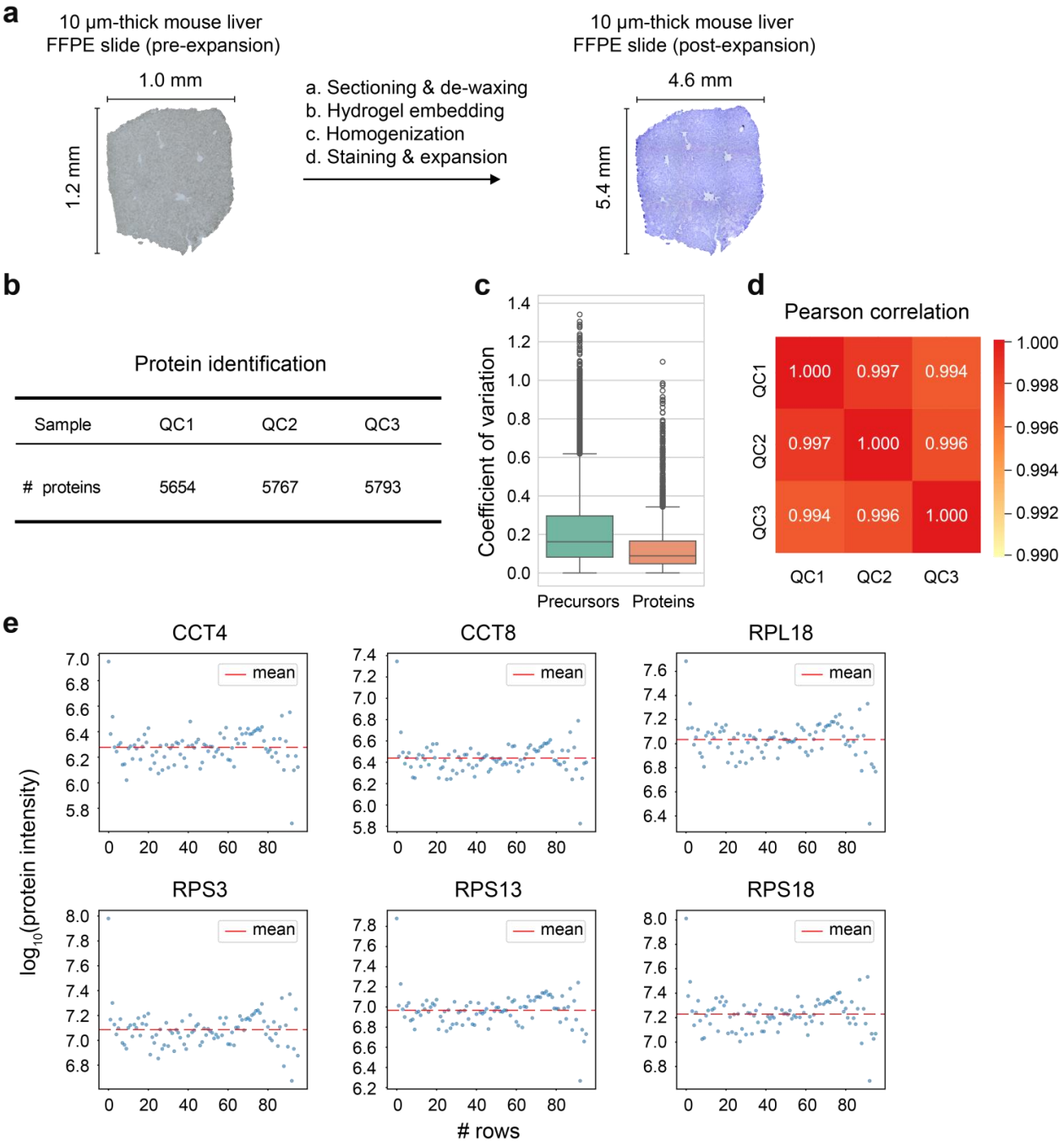

Fig. S2. Mass spectrometry performance and workflow reproducibility. a, Bright-field images of a mouse liver FFPE section before expansion (left) and after expansion (right), stained with Coomassie Brilliant Blue, shown for morphological comparison. **b,** Number of proteins identified in quality control (QC) samples, demonstrating stable mass spectrometry performance throughout data acquisition. **c,** Distribution of coefficients of variation for quantified proteins across QC samples. **d,** Pearson correlation of protein quantification across QC samples. **e,** Quality assessment of row-wise samples based on six additional housekeeping proteins. Scatter plots show the expression stability of these housekeeping proteins 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 S3

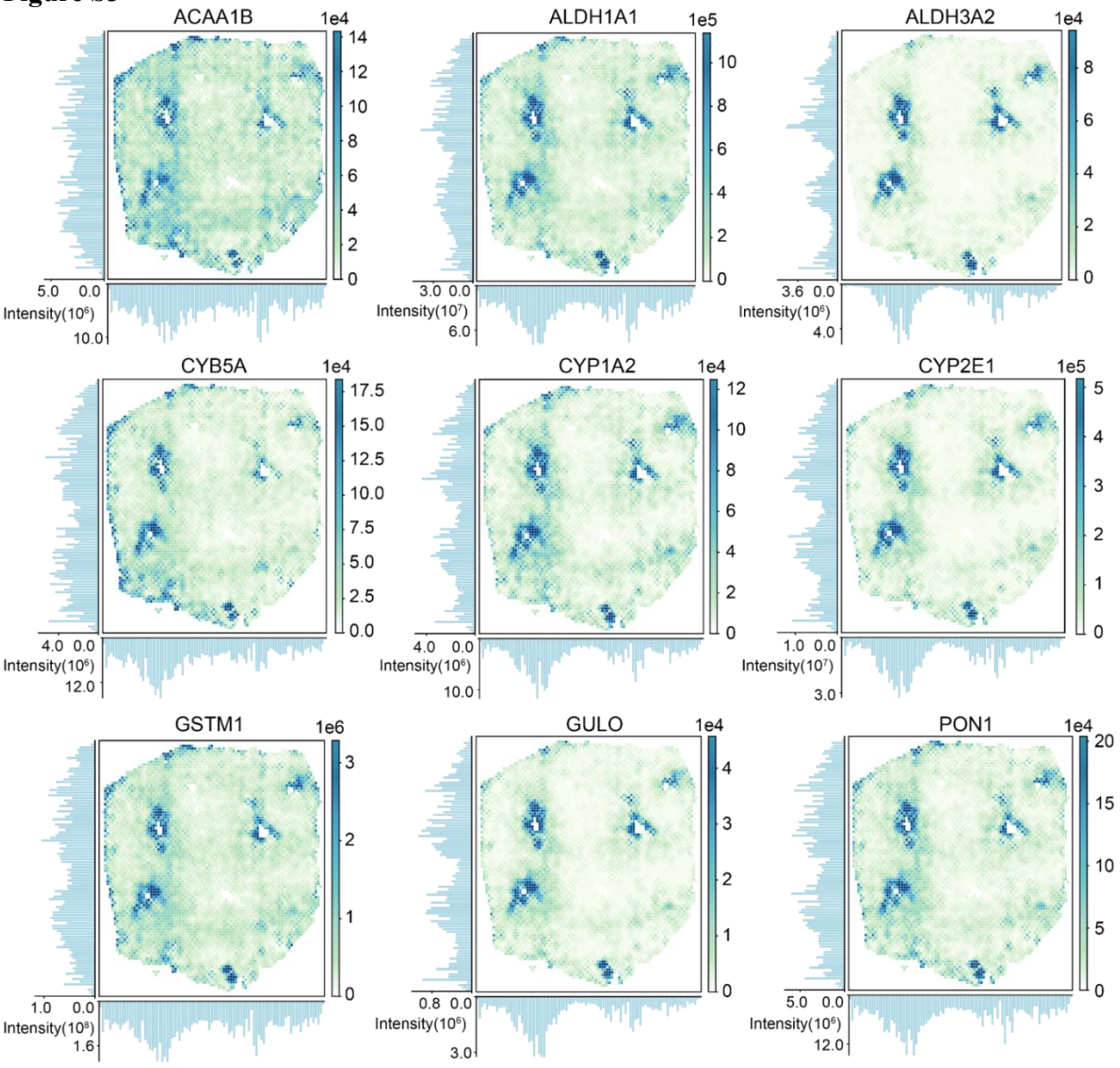

Fig. S3. Predicted spatial expression of CV marker proteins. The nine CV marker proteins were selected based on a previously published single-cell deep visual proteomics study (*17*). HetuNet-predicted protein expression profiles were converted into AnnData (adata) format and visualized spatially for each individual protein. For each spatial expression map, the accompanying marginal plots along the left and bottom axes show the corresponding mass spectrometry-measured protein intensities aggregated across the same row-wise and column-wise samples, respectively, providing a reference for the spatial expression patterns.

Figure S4

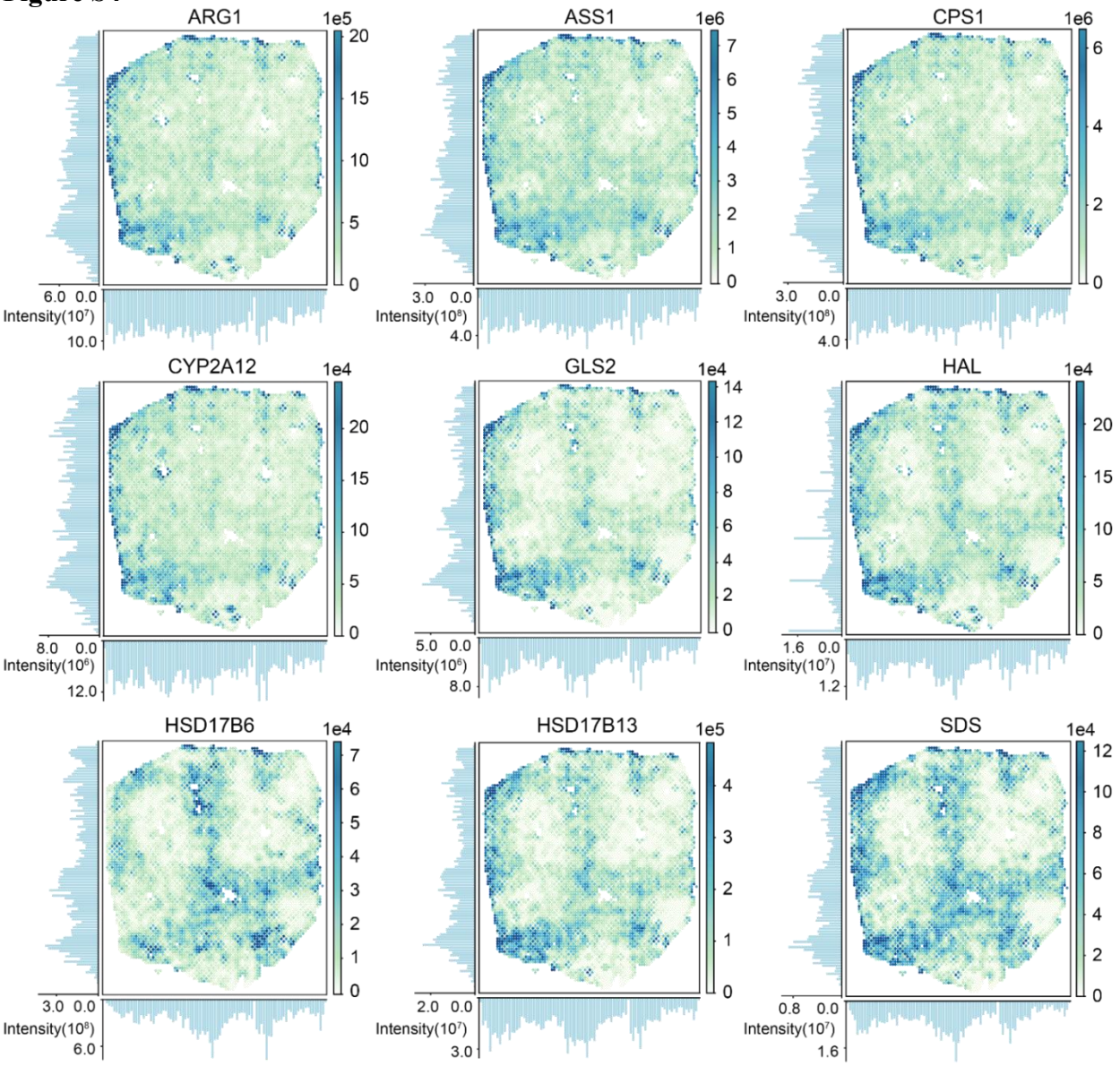

Fig. S4. Predicted spatial expression of PV marker proteins. The nine PV marker proteins were selected based on a previously published single-cell deep visual proteomics study (*17*). HetuNet-predicted protein expression profiles were converted into AnnData (adata) format and visualized spatially for each individual protein. For each spatial expression map, the accompanying marginal plots along the left and bottom axes show the corresponding mass spectrometry-measured protein intensities aggregated across the same row-wise and column-wise samples, respectively, providing a reference for the spatial expression patterns.

Figure S5

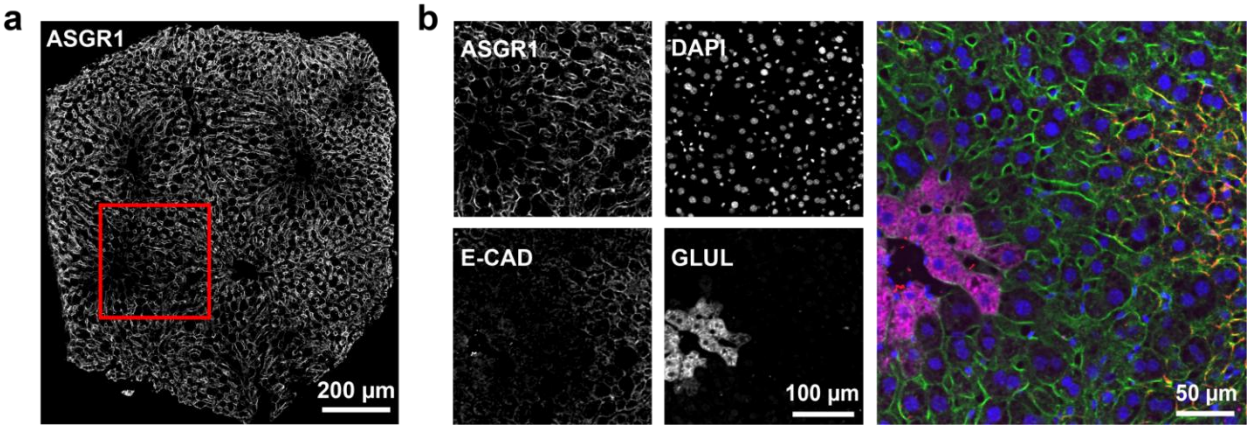

Fig. S5. Imaging references for expanded mouse liver tissue. a, Whole-slide immunofluorescence (IF) staining image of ASGR1 acquired from the expanded tissue section, providing a global imaging reference. **b,** Multiplexed IF staining results in the expanded tissue section. E-cadherin (E-CAD) marks PV regions, glutamate-ammonia ligase (GLUL) labels CV regions, ASGR1 labels the hepatocyte plasma membrane, and DAPI labels cell nuclei. Scale bars are provided.

Figure S6

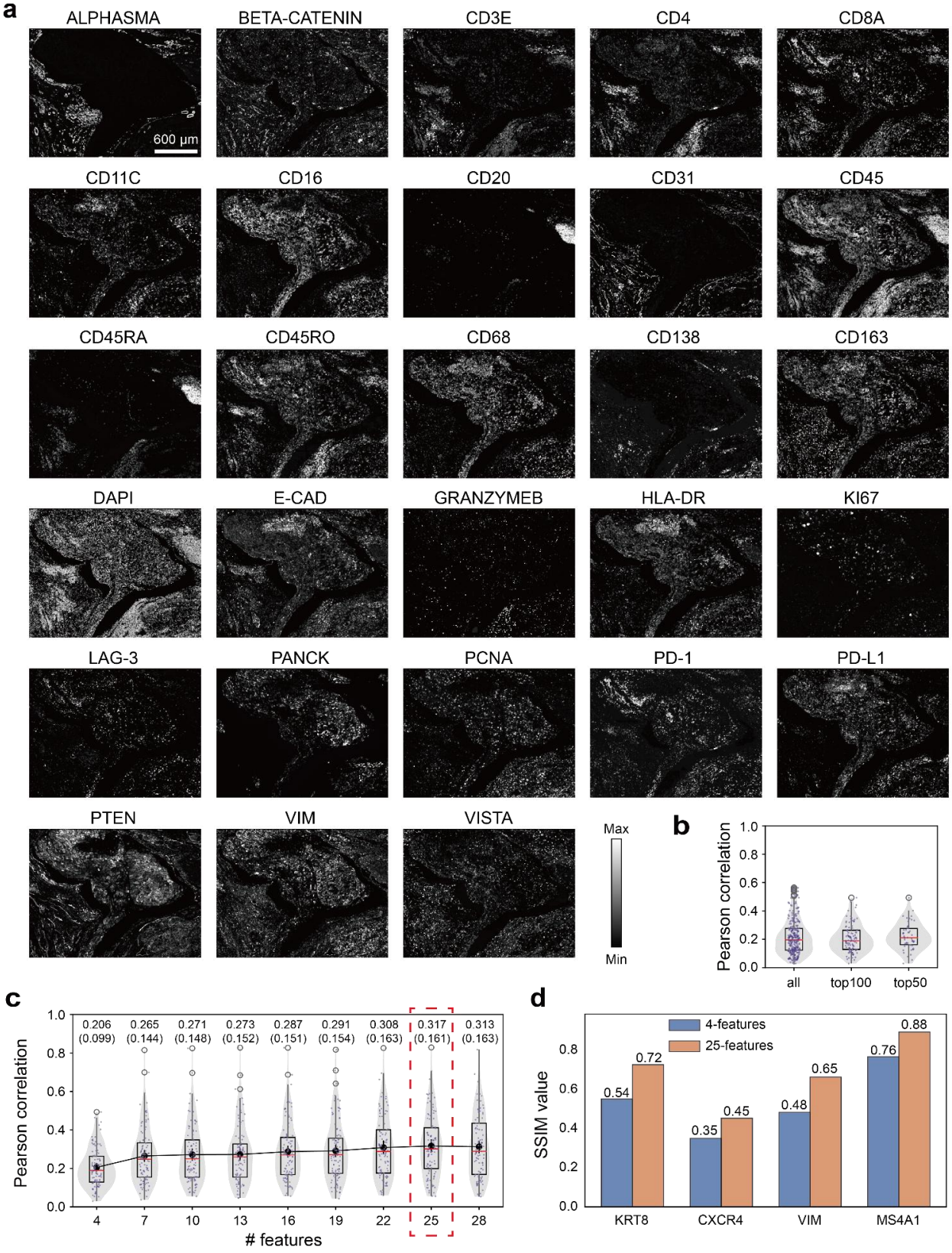
a
ALPHASMA
BETA-CATENIN
CD3E
CD4
CD8A
600 μm
CD11C
CD16
CD20
CD31
CD45
CD45RA
CD45RO
CD68
CD138
CD163
DAPI
E-CAD
GRANZYMEB
HLA-DR
KI67
LAG-3
PANCK
PCNA
PD-1
PD-L1
PTEN
VIM
VISTA
Max
Min
b
Pearson correlation
1.0
0.8
0.6
0.4
0.2
0.0
all
top100
top50
c
Pearson correlation
0.206 (0.099)
0.265 (0.144)
0.271 (0.148)
0.273 (0.152)
0.287 (0.151)
0.291 (0.154)
0.308 (0.163)
0.317 (0.161)
0.313 (0.163)
4
7
10
13
16
19
22
25
28
features
d
SSIM value
4-features
25-features
0.54
0.72
0.35
0.45
0.48
0.65
0.76
0.88
KRT8
CXCR4
VIM
MS4A1

Fig. S6. Reference markers and additional *in silico* validation of HetuNet. a, Spatial expression profiles of Xenium protein panel markers used as reference inputs for *in silico* validation. Heatmaps show the spatial distribution and expression levels of Xenium protein panel markers used as reference inputs during HetuNet training. **b-c**, Spatial correlation analysis of reconstructed expression. **b**, Distribution of Pearson correlation coefficients between HetuNet-predicted and Xenium-derived ground-truth spatial expression profiles across different groups of highly variable genes (HVGs). **c**, Model robustness was evaluated by plotting Pearson correlation coefficients for the top 100 HVGs as a function of the number of reference markers used during training. In the boxplots, red lines indicate the median; black dots connected across boxplots represent the mean; bars denote the standard deviation (SD); and the numbers above each boxplot report the mean (SD). **d**, Structural similarity analysis for HetuNet-reconstructed expression of representative target genes. Bar plots show the structural similarity index measure (SSIM) for KRT8, CXCR4, VIM, and MS4A1 under two training conditions, using 4 or 25 Xenium protein panel markers as reference inputs. Colors indicate the corresponding training conditions.

Figure S7

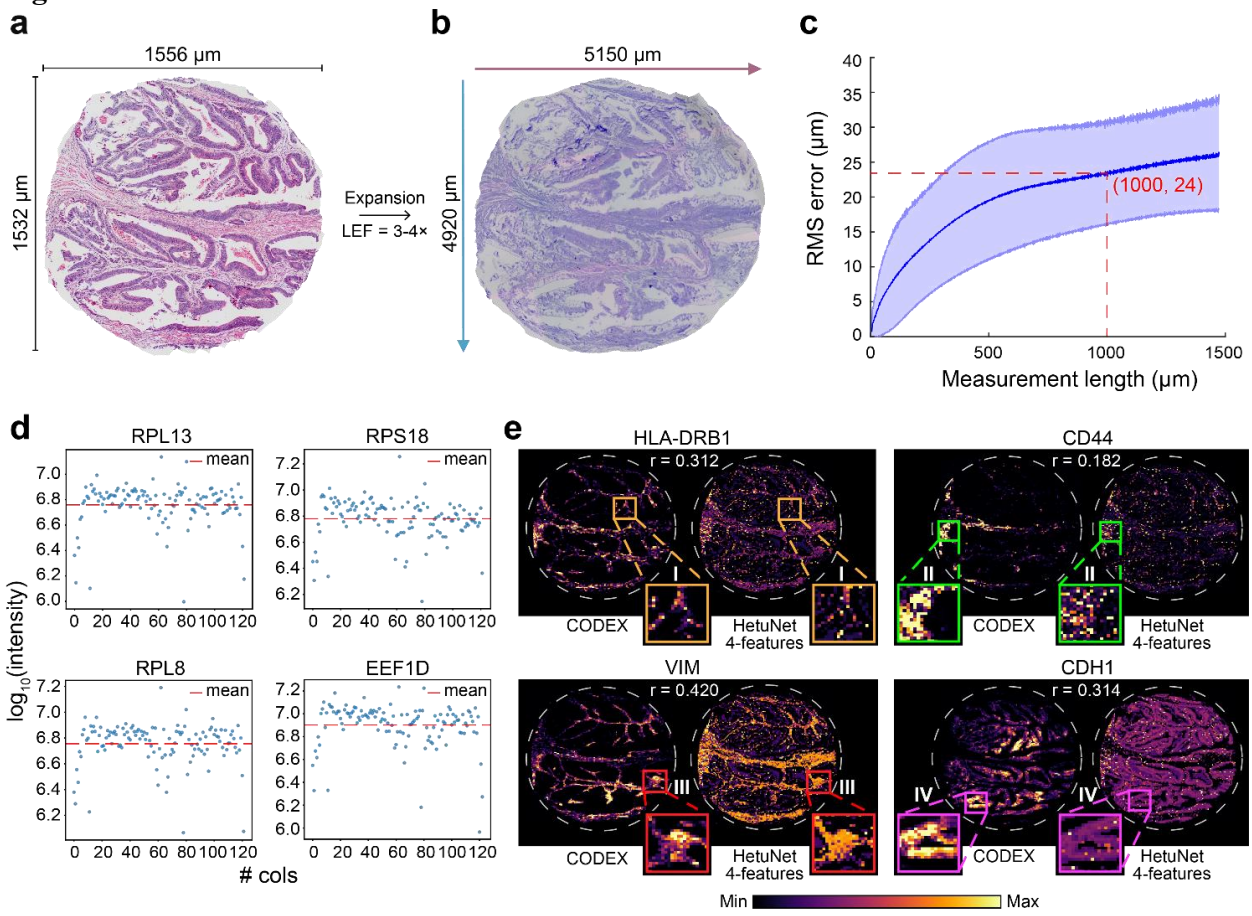

Fig. S7. Quantitative evaluation of human CRC data. a, Hematoxylin and eosin (H&E) staining of a human CRC FFPE tissue section before expansion. **b**, Coomassie Brilliant Blue staining of the corresponding tissue section after expansion. **c**, Comparison of root-mean-square (RMS) measurement length errors before and after tissue expansion in CRC samples. **d**, Quality control assessment of column-wise samples using housekeeping proteins. Scatter plots show the log10-transformed expression levels of four representative housekeeping proteins (RPL13, RPS18, RPL8, and EEF1D) across individual column samples. The y-axis represents the log10-transformed protein expression levels. To ensure comparability, raw expression values were normalized by the proportion of pixels captured in the CODEX data for each sample, correcting for variations in tissue area. The x-axis indicates column sample IDs. **e**, Spatial protein expression patterns of four representative markers (HLA-DRB1, CD44, VIM, and CDH1). For each marker, the left column shows ground-truth protein expression measured by CODEX, and the right column shows the corresponding HetuNet-predicted protein expression matrices inferred using four CODEX-derived spatial reference features (DAPI, CD31, FoxP3, and a-SMA). Colored bounding boxes indicate four regions of interest (ROIs), corresponding to the frames shown in **Fig. 4d**. Zoomed-in views compare ground-truth and HetuNet-predicted protein expression within each ROI.

Figure S8

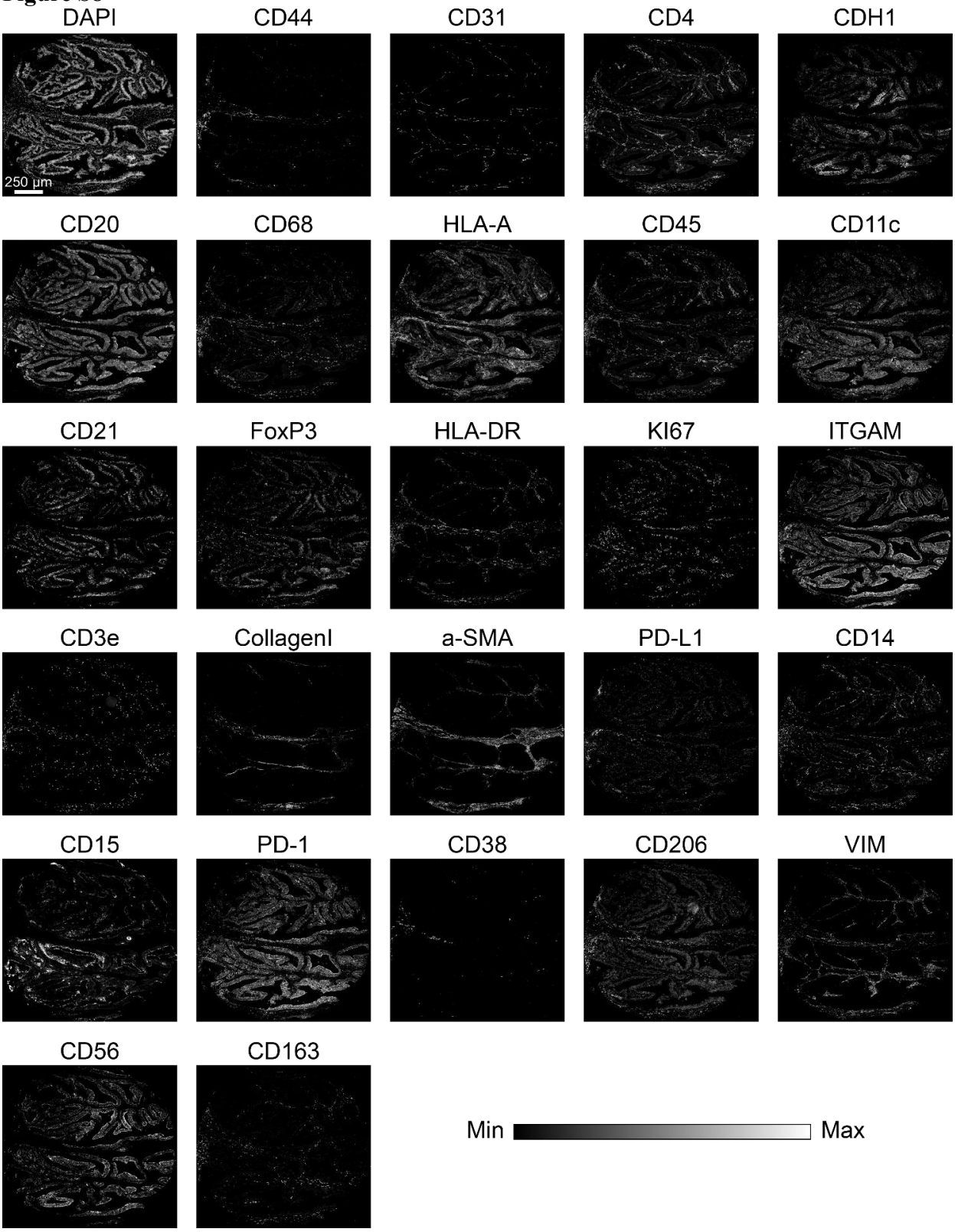

Fig. S8. Spatial expression profiles of the CODEX protein panel used for HetuNet reconstruction and validation. Heatmaps show the spatial distribution and expression levels of

27 protein markers selected from the 31-marker CODEX protein panel (with DAPI used for nuclear staining). A subset of these markers was used as spatial reference input features and supervision signals for HetuNet training, whereas four markers (HLA-DRB1, CD44, Vim, and CDH1) were withheld and used exclusively for validation.

Supplementary Table 1. Antibody Panel for CODEX and Conventional Immunofluorescence Imaging

Name	Supplier	Catalog number
*For CODEX**		
DAPI	Akoya	232121
CD44	Akoya	240059
CD31	Akoya	232172
CD4	Akoya	232174
E-cadherin	Akoya	232185
CD20	Akoya	240230
CD68	Akoya	232176
CD45RO	Akoya	232188
HLA-A	Akoya	200067
CD45	Akoya	240060
CD8	Akoya	232151
CD16	Abcam	ab256582
CD11c	Akoya	200074
CD21	Akoya	240003
FoxP3	Akoya	240170
HLA-DR	Akoya	240017
Ki67	Akoya	232179
CD11b	Abcam	ab209970
CD3e	Akoya	200066
CollagenI	Abcam	ab316223
a-SMA	Akoya	240068
PD-L1	Akoya	240228
CD14	Akoya	240066
CD15	Abcam	ab240092
PD-1	Akoya	240035
CD38	Akoya	200071
GranzymeB	Akoya	240074
CD206	Akoya	200098
Vimentin	Akoya	240069
CD56	Abcam	ab214436
CD57	Akoya	200094
CD163	Akoya	200100
For conventional immunofluorescence imaging		
ASGR1 Polyclonal antibody	Proteintech	11739-1-AP
Alexa Fluor 555 Mouse anti-E-Cadherin	BD Pharmingen	560064
CoraLite Plus 647-conjugated Glutamine Synthetase Monoclonal antibody	Proteintech	CL647-66323-2
Goat Anti-Rabbit IgG H&L (Alexa Fluor 488) preadsorbed	Abcam	ab150081
DAPI Staining Solution	Abcam	ab228549-2ml

**Antibodies were applied within the manufacturer-recommended dilution range (1:50-1:200). Imaging parameters, including exposure time, were guided by manufacturer recommendations and adjusted as needed to ensure detectable signals across markers.*