Article

Graph Fourier transform for spatial omics representation and analyses of complex organs

Yuzhou Chang [1,2,11], Jixin Liu[3,11], Yi Jiang[1], Anjun Ma [1,2], Yao Yu Yeo [4,5], Qi Guo[1], Megan McNutt[1], Jordan E. Krull [1,2], Scott J. Rodig[6,7], Dan H. Barouch [4,8], Garry P. Nolan [9], Dong Xu [10], Sizun Jiang [4,5,6], Zihai Li [2], Bingqiang Liu [3] ✉ & Qin Ma [1,2] ✉

Spatial omics technologies decipher functional components of complex organs at cellular and subcellular resolutions. We introduce Spatial Graph Fourier Transform (SpaGFT) and apply graph signal processing to a wide range of spatial omics profiling platforms to generate their interpretable representations. This representation supports spatially variable gene identification and improves gene expression imputation, outperforming existing tools in analyzing human and mouse spatial transcriptomics data. SpaGFT can identify immunological regions for B cell maturation in human lymph nodes Visium data and characterize variations in secondary follicles using in-house human tonsil CODEX data. Furthermore, it can be integrated seamlessly into other machine learning frameworks, enhancing accuracy in spatial domain identification, cell type annotation, and subcellular feature inference by up to 40%. Notably, SpaGFT detects rare subcellular organelles, such as Cajal bodies and Set1/COMPASS complexes, in high-resolution spatial proteomics data. This approach provides an explainable graph representation method for exploring tissue biology and function.

Advancements in spatial omics offer a comprehensive view of the molecular landscape within the native tissue microenvironment, including genome, transcriptome, microbiome, T cell receptor (TCR)[1], epigenome, proteome, transcriptome-protein markers co-profiling, and epigenome–transcriptome co-profiling[2] (Fig. 1a and Supplementary Fig. 1). These approaches enable the investigation and elucidation of functional tissue units (FTUs)[3], which are defined as over-represented multicellular functional regions with a unique physiologic function, with both cell-centric and gene-centric approaches. Specifically, cell-centric approaches involve the identification of spatial domains with coherent gene expression and histology[4], studying cell composition and neighborhoods within specific domains[5–7], and understanding inter-cellular mechanisms. In parallel, gene-centric approaches characterized FTUs by imputing gene expression[8] and identifying spatially variable genes (SVG)[9–11] in a highly complementary manner to cell-centric approaches.

[1]Department of Biomedical Informatics, College of Medicine, Ohio State University, Columbus, OH 43210, USA. [2]Pelotonia Institute for Immuno-Oncology, The James Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA. [3]School of Mathematics, Shandong University, 250100 Jinan, China. [4]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA. [5]Program in Virology, Division of Medical Sciences, Harvard Medical School, Boston, MA 20115, USA. [6]Department of Pathology, Dana Farber Cancer Institute, Boston, MA 02115, USA. [7]Department of Pathology, Brigham & Women's Hospital, Boston, MA 02115, USA. [8]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA. [9]Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA. [10]Department of Electrical Engineering and Computer Science and Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA. [11]These authors contributed equally: Yuzhou Chang, Jixin Liu. ✉e-mail: bingqiang@sdu.edu.cn; qin.ma@osumc.edu
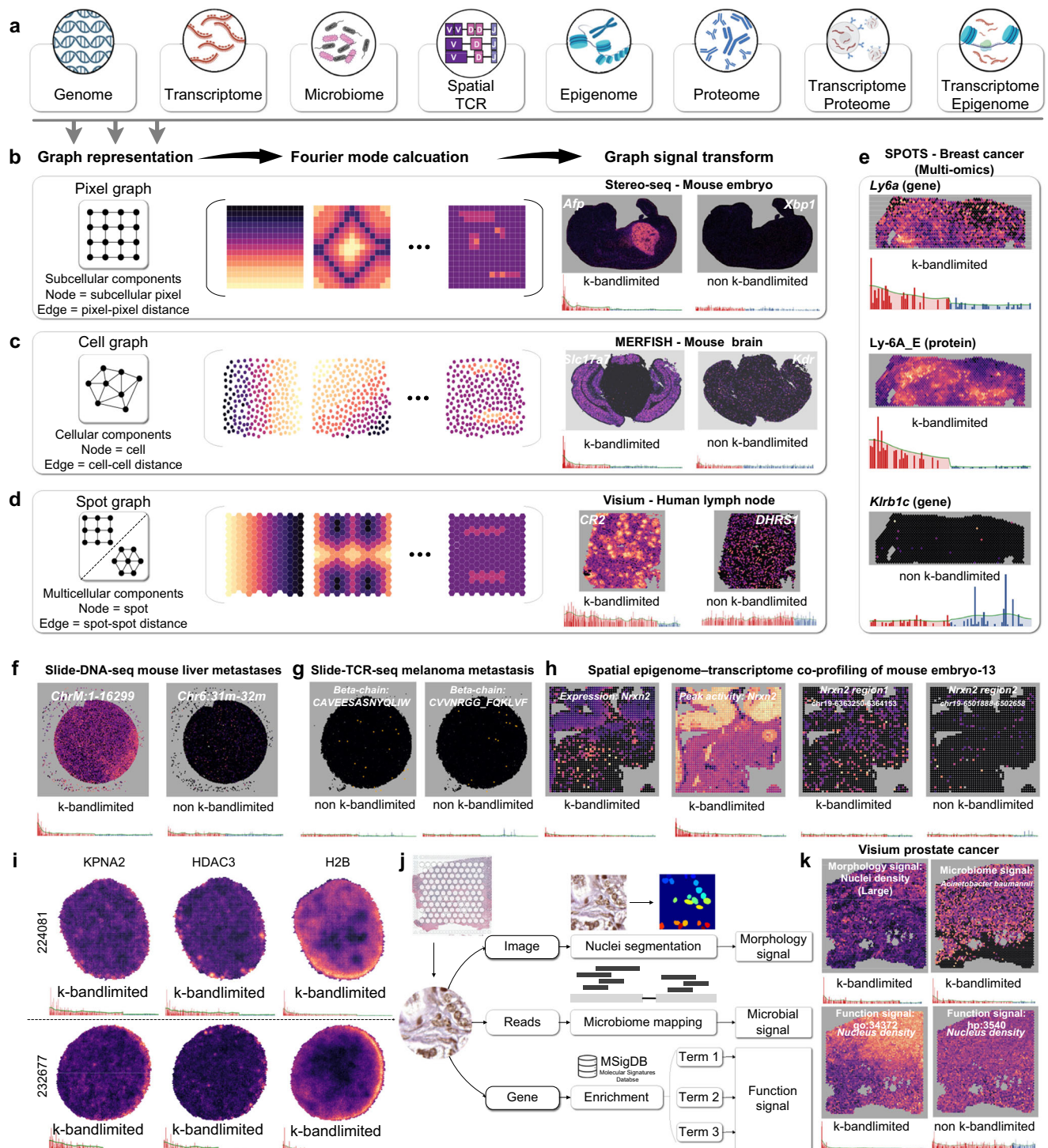
Classic statistical methods, such as SPARK[9], SPARK-X[12], and SpatialDE[11], have effectively modeled molecular variations and spatial relationships within a tissue. However, they did not fully explore the capacity to translate these relationships into understandable and analyzable features. In contrast, graph-based methods present a powerful alternative method that efficiently encodes and leverages spatial relationships within tissue in spatial omics data representation[13]. We postulate that an FTU can be intuitively considered a graph; its nodes represent spots or cells, and edges connect spatially adjacent or functionally related nodes. Within this representation of FTUs, a binary graph signal (e.g., 0,1), representing discrete two-state information at each node, and cellular or subcellular composition or omics features (e.g., genes) constitute continuous graph signals, encoding a range of values across the graph's nodes. These graph signals define the FTU's characterization, connect cell-centric and gene-centric analyses, and offer mutual interpretatibility[14], through the generation of a graph embedding that harmonizes the graph structure and signal magnitude. Furthermore, while graph-based machine learning methods are available to learn graph embeddings and carry out downstream tasks (e.g., graph classification), their learning progress is usually a "black box" and relies on an inductive bias (i.e., a hypothesis for a particular question) to train the model[15]. The characteristics of the produced graph embeddings are specifically tailored to perform optimally in certain targeted downstream tasks. Therefore, there is a need for a generic graph signal representation framework with a solid mathematic foundation to reveal intricate

Fig. 1 | Schema of SpaGFT for spatial omics representation. a The panel show-cases spatial omics technologies, including single and multi-modality methods. **b–d** the panels display the calculation of Fourier modes (FM), and the transformation of original graph signals into Fourier coefficients (FC) with different resolutions of technologies. **b** The figure presents pixel graphs with nodes at the subcellular level and edges denoting short Euclidean distances between connected pixels. This graph represents technologies like stere-seq and most spatial proteomics data, e.g., 4i. The two figures following panel **b** illustrate a *k*-bandlimited signal (e.g., *Afp*) and a non-*k*-bandlimited signal (e.g., *Xbp1*). **c** and **d** Cell graphs and spot graphs are composed of nodes at the cellular level resolution and multicellular level resolution, respectively, with edges representing short Euclidean distances between nodes in two panels. **e** The figure exhibits multi-modal data from a technology called SPOTS, which can measure both proteins and genes simultaneously. The *k*-bandlimited signal shown is for the *Ly6a* gene and its corresponding protein, while the non-*k*-bandlimited signal is for the *Klrb1c* gene. **f–h** The panels show examples of signals from Slide-DNA-seq, Slide-TCR-seq, and spatial epigenome–transcriptome co-profiling of mouse embryo-13. **i** The panel shows subcellular spatial proteome (i.e., 4i) are *k*-bandlimited signals. **j** This panel demonstrates data augmentation for sequencing-based spatial transcriptomics (e.g., Visium). The first step of augmentation involves using H&E images and Cell-pose for cell segmentation and counting the number of nuclei in each spot. The next step involves mapping reads to the microbiome genome, which then allows for the determination of microbiome abundance. Finally, gene lists (e.g., MSigDB) can be used to calculate the pathway activity score for each spot. **k** This panel displays the signals mentioned in panel **j**, including cell density, microbiome abundance, and pathway activity. Panel **a** is created with BioRender.com, created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

relations between molecular signatures and FTUs across multiple resolutions of spatial omics data.

To achieve this, we present the Spatial Graph Fourier Transform (SpaGFT), an analytical feature representation approach to encode smooth graph signals for representing biological processes within tissues and cells. It bridges graph signal processing techniques and spatial omics data, enabling various downstream analyses and facilitating insightful biological findings. Computationally, SpaGFT out-performed other tools in identifying SVGs with hundred-fold efficiency and gene expression imputation across human/mouse Visium data. Biologically, SpaGFT identified key immunological areas for B cell maturation processes from human lymph nodes Visium data and further illustrated secondary follicle cellular, morphological, and molecular diversity from exclusively in-house human tonsils CODEX data. Moreover, SpaGFT can be seamlessly integrated into other machine learning frameworks regarding domain identification (e.g., SpaGCN[4]), annotation transfer from cell types to spots (e.g., TACCO[6]), the cell-to-spot alignments (e.g., Tangram[7]), and subcellular hallmark inference (e.g., CAMPA[16]). Notably, enhanced CAMPA has enabled the discovery of rare subcellular structures like the Cajal body and Set1/COMPASS complex based on iterative indirect immunofluorescence image (4i) data[17], enhancing our understanding of cellular function using spatial omics technologies.

Results

### SpaGFT reliably represents the smooth signal of spatial omics data

We summarize current spatially resolved omics as three types of spatial graphs related to the granularity of nodes, ranging from subcellular level (i.e., pixel-level) to broader cellular (i.e., cell-level) and multicellular scales (i.e., spot-level) based on the spatial resolutions (Fig. 1b–k). This granularity can range from subcellular levels to broader cellular and multicellular scales. For example, based on the spatial graph of a spatially resolved transcriptomics (SRT) dataset, the transcriptomic profile of a specific gene is a graph signal and can be represented by the linear combination of its Fourier modes (FMs, Terminology Box). To elaborate, a low-frequency FM contributes to a low and smooth graph signal variation, representing a spatially organized pattern, while a high-frequency FM contributes to rapid graph signal variation and usually refers to noises in spatial omics data[18]. For example, if a gene exhibits a spatially organized pattern in SRT data, the Fourier coefficients (FCs) of corresponding low-frequency FMs are more dominant than FCs of high-frequency FMs in the graph Fourier representation. Notably, FMs are associated with graph structure and do not assume any predefined patterns[18], ensuring flexibility in representing both well-defined and irregular spatial signal patterns. Thus, regardless of single- (Fig. 1b–d, f, g, and i), multi-modalities (Fig. 1e and h), or augmented features (Fig. 1j and k), the spatial omics can be analytically trans-formed into FCs to quantify the contribution of FMs in the frequency domain[19], a feature space for enhancing the interpretability and generalizability in downstream analyses.

SpaGFT identifies spatially variable genes and enhances gene and protein signals

Using the representation framework of SpaGFT (Fig. 2a), the mathematical formulation of SVG identification can be derived as a *k*-bandlimited signal recognition problem, which determines the first *k* low-frequency FMs to best approximate the original graph signal (Fig. 2b, Supplementary Fig. 2, and S1 of Supplementary Note 1). This formulation can overcome three main limitations of SVG identification methods: (i) no pre-assumption of regular patterns in model design (e.g., radial hotspot, curve belt, or gradient streak)[9]; (ii) interpretable representation of SVG patterns[20] with spatial context; and (iii) high computational efficiency[12] when processing large-scale datasets. Essentially, we defined and implemented a GFTscore for each gene to quantify the contribution of low-frequency FMs by determining the first *k* low-frequency FMs, weighting, and summing corresponding FCs (S2 of Supplementary Note 1). Based on the definition, a gene is identified as an SVG if (i) its GFTscore is greater than the inflexion point based on the distribution of all genes' GFTscore and (ii) its FCs of the first *k* low-frequency FMs are significantly higher than FCs of high-frequency FMs (S3 of Supplementary Note 1). Consequently, we evaluated the performance of SVG identification using 31 public SRT datasets from human and mouse brains (Supplementary Data 1)[21–24]. As no golden-standard SVG database was available, we collected 849 SVG candidates from five existing studies[24–28], and 458 of them were used as curated benchmarking SVGs based on cross-validation with the in situ hybridization (ISH) database of Allen Brain Atlas[29] (Supplementary Data 2 and 3, see the "Methods" section). The SVG prediction performance of SpaGFT was compared with SPARK[9], SPARK-X[12], MERINGUE[30], SpatialDE[11], SpaGCN[4], and scGCO[31] in terms of six reference-based and two reference-free metrics (Supplementary Note 2). The grid search of parameter combinations was conducted on three high-quality brain datasets to evaluate each tool's performance, in which SpaGFT showed the highest median and peak scores (Fig. 2c and Supplementary Data 4). In addition, the computational speed of SpaGFT was two-fold faster than that of SPARK-X and hundreds-fold faster than those of the other four tools on the two Visium datasets (Supplementary Data 5). Although SpaGFT was slower than SPARK-X on the Slide-seqV2 dataset, it showed a remarkably enhanced SVG prediction performance compared to SPARK-X. We then performed an independent test on 28 independent datasets using the parameter combination with the highest median Jaccard Index among three datasets from the above grid-search test. The results revealed that SpaGFT promised supreme performance among the investigated tools based on the evaluation metrics (Fig. 2d, Supplementary Fig. 3a–d, and Supplementary Data 6). Within the top 500 SVGs from each of the above six tools, SpaGFT identified SVGs shared with other tools and also unique SVGs that were validated as the ground truth
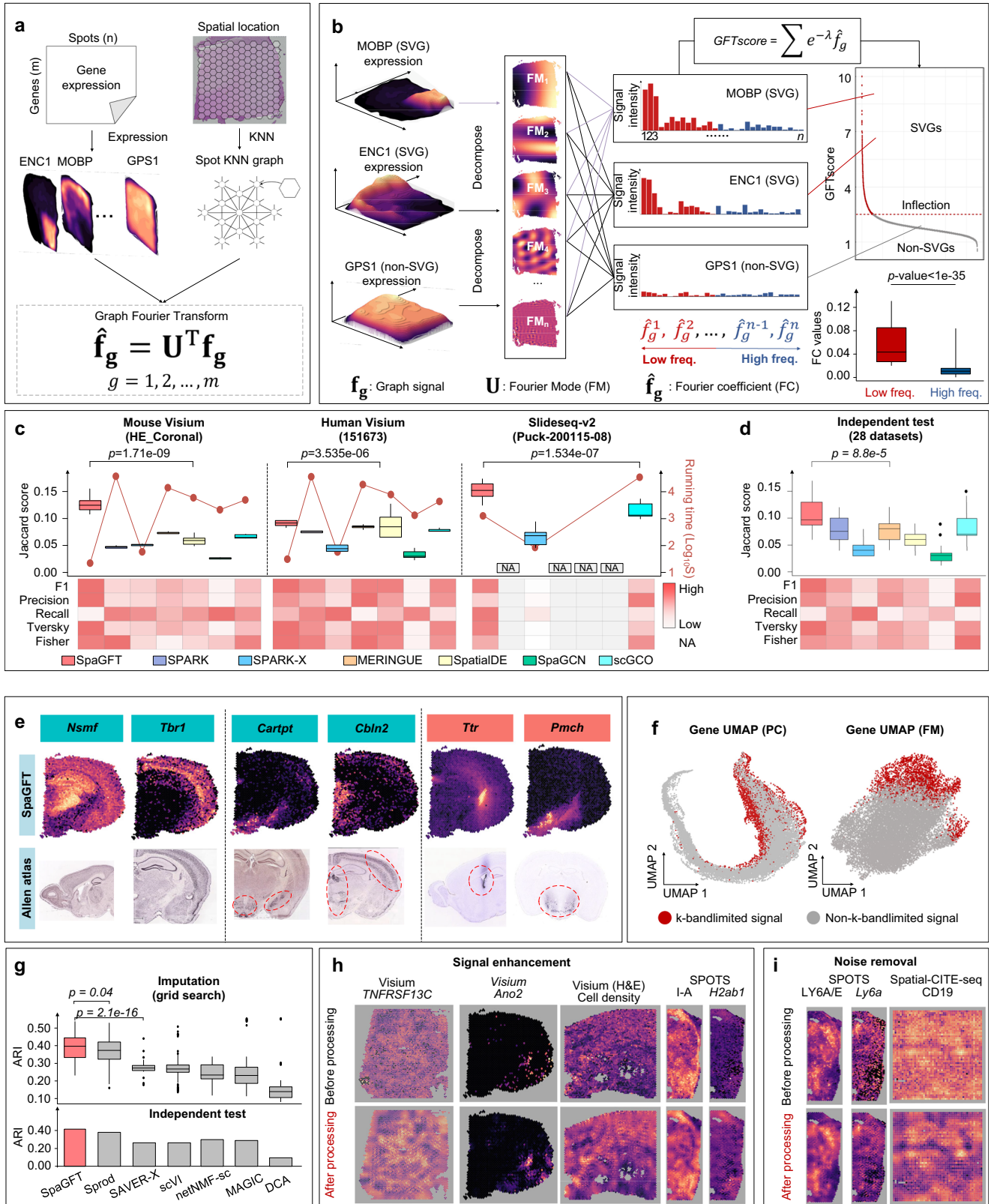
(Supplementary Fig. 3e and Supplementary Data 7). For example, *Nsmf* and *Tbr1* were identified by all six tools and showed clear structures of the hippocampus, cortical region, and cerebral cortex. On the other hand, *Cartpt, Cbln2, Ttr*, and *Pmch* were uniquely identified by SpaGFT and showed key functions in the brain, such as *Cartpt* participating in dopamine metabolism[29] (Fig. 2e, Supplementary Fig. 4, and Annotation 1 of Supplementary Note 3). These benchmarking results suggested

that SpaGFT is capable of leveraging upon the FM representation of gene expression for robust and accurate identification of SVGs from SRT data. SpaGFT takes advantage of FM representation of gene expression patterns in SVG identification, and the SVGs identified by SpaGFT were distinguishably separated from non-SVGs on the FM-based UMAP with a clear boundary, whereas SVGs were irregularly distributed on the principal component-based gene UMAP (Fig. 2f).

Fig. 2 | Performance for SVG idenfication and spatial omics feature imputation.
a SpaGFT considers a gene-spot expression count matrix ($m \times n$) and spatial locations as input data, with *ENC1*, *MOBP*, and *GPS1* listed as examples. **b** Two known SVGs (*MOBP* and *ENC1*) and one non-SVG (*GPS1*) are shown as examples. The FMs can be separated into low-frequency (red) and high-frequency (blue) domains.
c The SVG prediction evaluation was compared to five benchmarking tools. The running time is represented as red lines. In addition, the other evaluation scores of all parameter combinations for each tool are shown as heatmaps. The two-sided Wilcox rank-sum test was used to calculate the *p*-value for the highest two tools (i.e., $N = 16$ and $N = 53$ for SpaGFT and MERINGUE in HE_coronal data; $N = 16$ and $N = 54$ for SpaGFT and MERINGUE in 151673 data; $N = 16$ and $N = 18$ for SpaGFT and SPARK-X in Puck-200115-08 data). The methods are not able to identify SVG in a reasonable time, showing NA in this panel. **d** The box plot shows independent test results. The two-sided Wilcox rank-sum test is used to calculate the *p*-values for the highest two tools ($N = 28$). Each box showcases the minimum, first quartile, median, third quartile, and maximum Jaccard scores in panels **c** and **d**. **e** SVG examples that all tools can identify (left panel) are uniquely identified by SpaGFT (middle and right panel). Green genes are reported in the literature, while orange is not. Expression of *Nsmf, Tbr1, Cartpt, Cbln2, Ttr,* and *Pmch* in adult mouse brain. Allen Mouse Brain Atlas, mouse.brain-map.org/experiment/show/74821712, mouse.brain-map.org/experiment/show/79591351, mouse.brain-map.org/experiment/show/72077479, mouse.brain-map.org/experiment/show/68632172, and mouse.brain-map.org/experiment/show/55. **f** Comparison of the UMAPs of the HE-coronal data in Principle Component features space and Fourier space. **g** Boxplot showcases the performance of SVG signal enhancement for grid search (top) and independent test using 151509 (bottom), where the y-axis is the ARI value. The two-sided Wilcox rank-sum test is used to calculate the *p*-values for the highest two tools ($N = 27$). Each box showcases the minimum, first quartile, median, third quartile, and maximum ARI scores. **h** and **i** The spatial map shows the signals before and after enhancement and noise removal for spatial omics features. Source data are provided as a Source Data file.

In addition, the distorted graph signal correction can be used as the mathematical formulation to impute a low-expressed gene or denoise a high-intensity but noisy protein in SpaGFT. Essentially, FCs are shifted towards a specific bandwidth by implementing a low-pass filter and are inversely transformed to an enhanced graph signal using an inverse graph Fourier transform (iGFT)[32]. To enhance the main signal and mitigate noise, a low-passing filter is employed to weigh and shift all FCs toward the low-frequency bandwidth (see the "Methods" section). In the end, these weighted FCs are transformed back to a corrected graph signal via iGFT (Supplementary Fig. 5a). In assessing the performance of gene expression correction, we used 16 human brains SRT datasets with well-annotated spatial domains[23,24] and utilized adjusted rand index (ARI) to measure the accuracy of predicting spatial domains using corrected gene expression. As a result, SpaGFT outperformed other gene enhancement tools in terms of ARI, including Sprod[8], SAVER-X, scVI, netNMF-sc, MAGIC, and DCA[33,34] (Fig. 2g, Supplementary Fig. 5b, and Supplementary Data 8). For example, SpaGFT enhanced the low-intensity spatial omics signal broadly across different technologies and species, such as gene *TNFRSF13C* for human lymph node, gene *Ano2* for mouse brain[29] (Supplementary Fig. 5c), cell density for human prostate tumor (from data of Fig. 1k), protein I-A, and corresponding gene *H2ab1* for mouse breast tumor. Similarly, *the* noisy background can also be removed, such as protein LY6A/E and corresponding gene *Ly6a* and protein CD19 (Fig. 2h, i and Annotation 2 of Supplementary Note 3).

SpaGFT identifies the germinal center, T cell zone, B cell zone, and crosstalking regions in the human lymph node

As low-frequency FC can represent smooth spatially variable patterns, they can be used for SVG clustering, and gene clusters can correlate with distinct FTUs from a gene perspective (Supplementary Fig. 6a). To demonstrate the application, we implemented SpaGFT in the publicly available Visium data of human lymph nodes, which, as secondary lymphoid organs contain well-known recurrent functional regions, such as T cell zones, B cell zones, and germinal center (GC)[20]. First, SpaGFT identified 1,346 SVGs and characterized nine SVG clusters (Fig. 3a and Supplementary Data 9). To recognize the FTUs of the T cell zone, B cell zone, and GC, we first used cell2location[35,36] to determine the cell proportion (Supplementary Fig. 6b and Supplementary Data 10) for the nine SVG clusters and investigate function enrichment (Supplementary Fig. 6c–e) for three selected FTUs. Based on the molecular, cellular, and functional signatures of three regions[35], we found that SVG clusters 3, 5, and 7 (Fig. 3b) were associated with the T cell zone, GC, and B cell zone, respectively (Annotation 3 of Supplementary Note 3).

In contrast to spatial domain detection tools, SpaGFT is not restricted to a rigid boundary for tissue-level identification of microenvironments[5]. Instead, SpaGFT allows overlapping regions to infer the functional coherence and collaboration among different FTUs. We therefore projected three FTUs represented by SVG clusters 3, 5, and 7 on the spatial map for visual inspection, and identified their close spatial proximity (Fig. 3c). These results are highly indicative of tissue regions of polyfunctionality amongst these three TFUs (four representative subregions are shown in Fig. 3c). To further investigate the crosstalk amongst these three TFUs, we projected spots (assigned to all three regions) to the Barycentric coordinates (the equilateral triangle in Fig. 3d), which displayed relations and abundance of the unique and overlapped regions regarding cell type components[37]. We identified 614 spots overlapped with B cell zone and GC, 158 spots overlapped with GC and T zone, 93 spots overlapped with T zone and B cell zone, and 26 spots overlapped across three FTUs (Supplementary Data 11), in support of the complex interactions within these three TFUs. We next hypothesized that the spots from the overlapped region would vary in functions and cell components to support the polyfunctionality of these regions. We thus investigated the changes in enriched functions (Supplementary Data 12) and cell types (Supplementary Fig. 7) across seven regions (i.e., GC, GC–B, B, B–T, T, T–GC, and T–B–GC). Our results identified lymph node-relevant pathways and cell types, such as B and T cell activity and functions, as significantly varied across those regions (Fig. 3e, f, Annotation 4 of Supplementary Note 3), in support of our hypothesis.

SpaGFT reveals secondary follicle variability based on CODEX data

The results of Visium in Fig. 3 showcased the ability of SpaGFT to identify FTUs via SVG clustering. Given that the current resolution of Visium (~50 µm per pixel) limited our ability to interpret the variability of finer follicle structures and their corresponding functions at the cellular level, we next performed single-cell level spatial proteomics on a human tonsil using a 49-plex CODEX panel at a ~0.37 µm per pixel resolution (Fig. 4a) to better characterize and interpret the follicle variability we observed and inferred using SpaGFT on the Visium data. Based on the anatomical patterns highlighted by B (e.g., CD20) and T cell (e.g., CD4) lineage markers, we selected fields of view (FOV) that would allow for a good representation of the complex tissue structures present in the tonsil (i.e., GC and interfollicular space[38]) while still highlighting the variability in follicle structure[39]. We first performed cell segmentation with DeepCell[40], followed by clustering with FlowSOM[41] and Marker Enrichment modeling[42] to identify the diverse cell phenotypes present in the data (Fig. 4b). Interestingly, we observed that the clear arrangement of T and B cell patterns (e.g., A3, A5, and A6) informed identifiable GC regions within the follicular structure, compared to others (e.g., A4) without clear T and B cell spatial organization (Fig. 4b). We, therefore, postulated that A4 is comprised of multiple follicles, unlike A5 and A6, to represent a more spatially complex FOV.
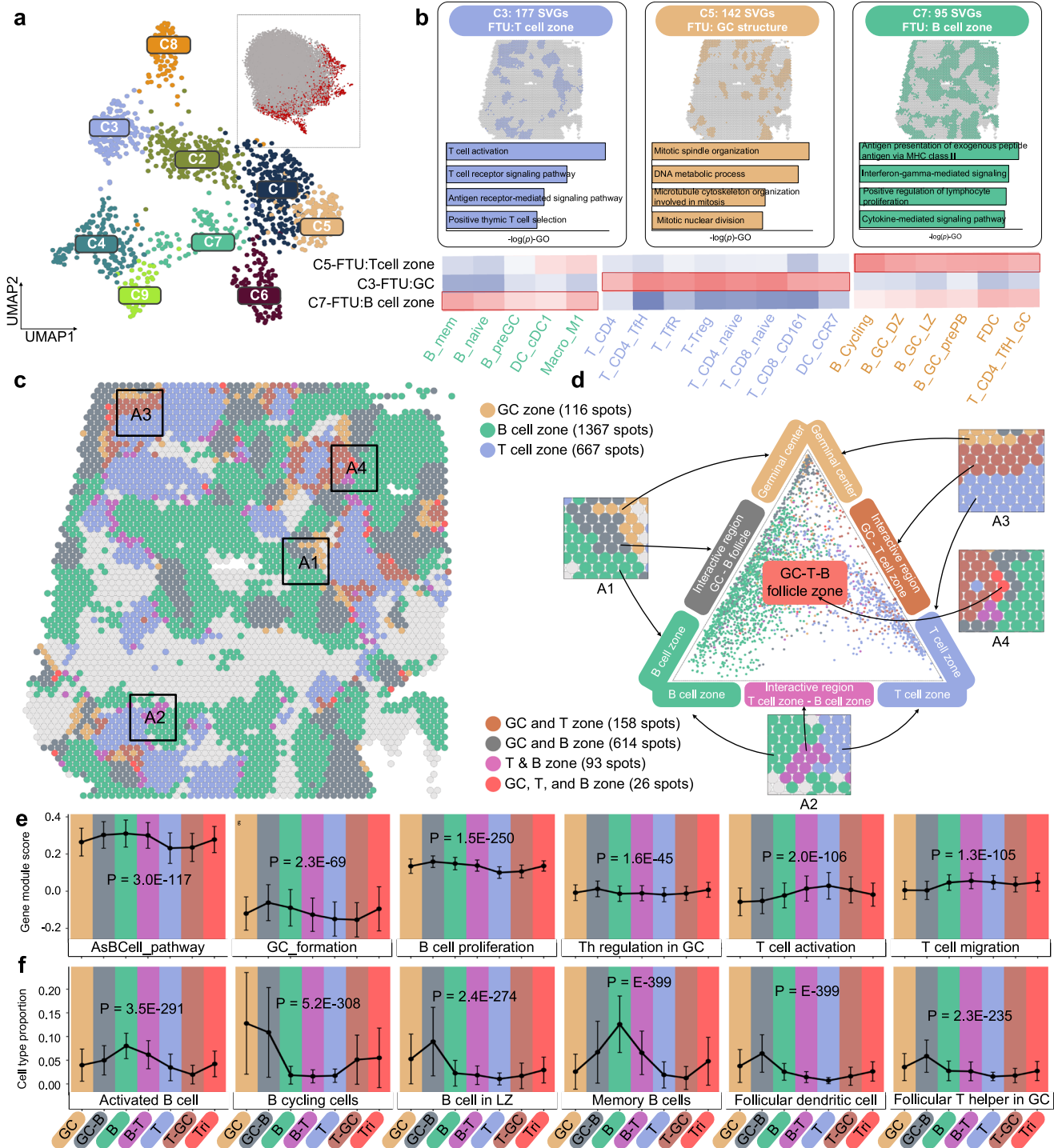
Fig. 3 | SpaGFT reveals the region-region interaction on the lymph node.
a UMAP visualization of nine SVG clusters from the human lymph node. Each dot represented SVGs. Upright UMAP showed SVGs in red and non-SVG in gray.
b Clusters 3, 5, and 7 were highly associated with the T cell zone, GC, and B follicle cell components based on molecular and functional signatures. The heatmap visualized the FTU-cell type correlation matrix. **c** The spatial map overlaid three FTUs and displayed the overlapped spots and unique spots. As different colors corresponded to spots, we selected four areas to showcase the region-to-region interaction. A1 showcased GC, GC-B interaction region, and B follicle. A2 showcased the B follicle, B−T interaction region, and T cell zone. A3 showcased the GC, GC-T interaction zone, and T cell zone. A4 displayed a B-GC-T interaction zone. **d** The barycentric coordinate plot shows cell-type components and the abundance of

spots in interactive and functional regions. If the spot is closer to the vertical of the equilateral triangle, the cell type composition of the spot tends to be signature cell types of the functional region. The spots were colored by functional region and interactive region categories. **e** and **f** The three plots displayed changes in enriched functions and cell type components across seven regions (GC, GC-B, B, B−T, T, T−GC, T−GC−B). The P-value was calculated using one-way ANOVA to test the differences among the means of seven regions. The number of sample sizes (i.e., spots) in the GC zone, B cell zone, T cell zone, GC and T zone, GC and B zone, T&B zone, and GC, T, and B zone are 116, 1367, 667, 158, 614, 93, and 26. The error bars show the standard deviation of enrichment scores. Source data are provided as a Source Data file.
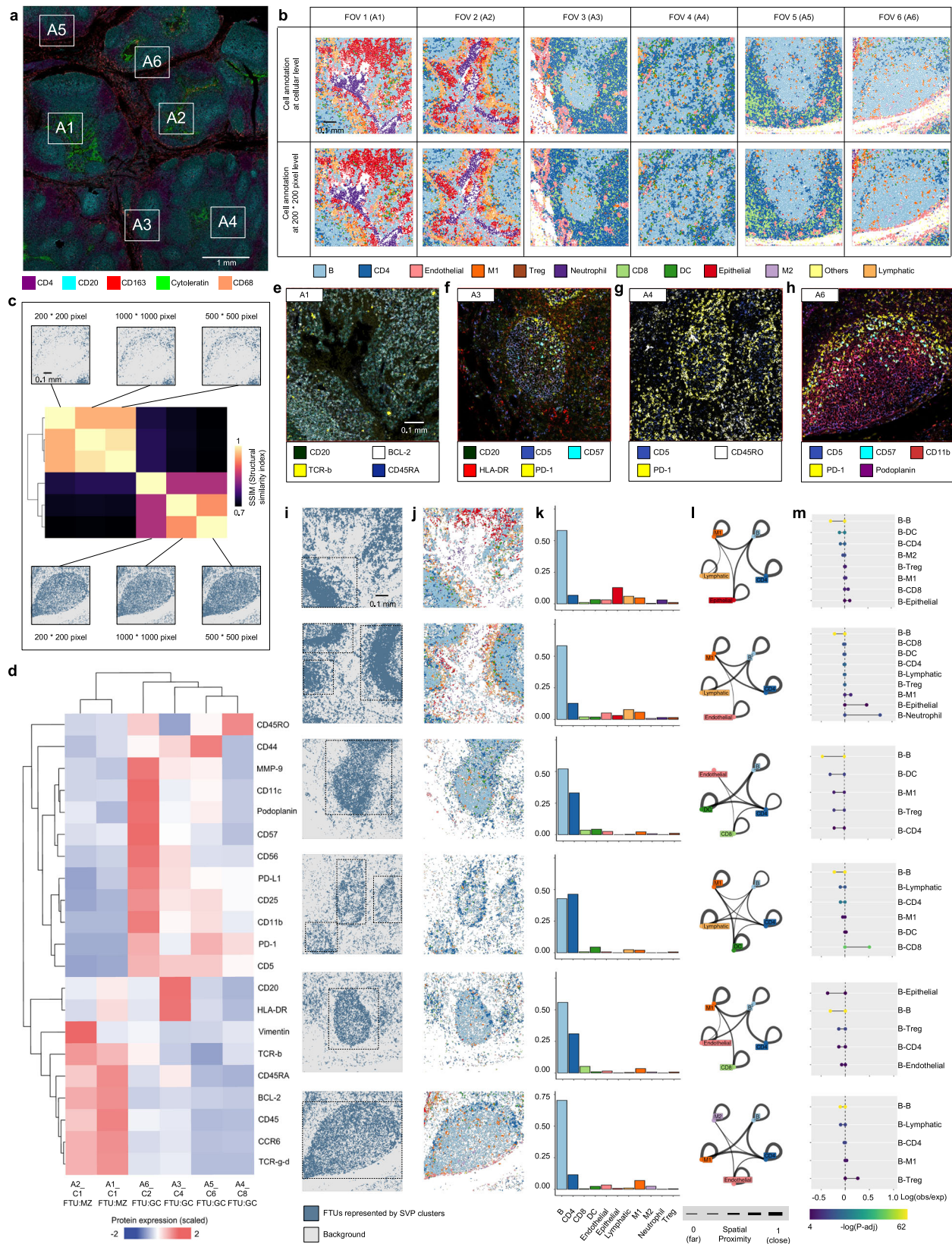
We investigated this further by directly using the raw CODEX images as inputs to identify FTUs formed from spatially variable protein (SVP) clusters within the tissue environment[43]. To verify whether downsampling the CODEX image (Supplementary Fig. 8a) would result in a loss in the power of characterizing FTUs, we first used FOV 6 to generate three images across different resolutions (with downsampling),

resulting in a (1) 1000-by-1000 pixel image (~0.8 μm per pixel size), (2) 500-by-500 pixel image (~1.6 μm per pixel size), and 3) 200-by-200 pixel image. Our results show that despite the generation of diverse low- and high-frequency FMs from three pixel-level images (as illustrated in Supplementary Fig. 8b), SpaGFT was stable to resolution changes, characterizing FTUs across different resolutions with consistent patterns

Fig. 4 | SpaGFT reveals the follicle heterogeneity on Tonsil CODEX data. a A 49-plex CODEX data was generated from human tonsil tissue at a 0.37 μm/pixel resolution. Six FOVs were selected based on their varying tissue microenvironment and cellular organization. **b** Cell phenotype maps for each of the six FOVs, depicting the cellular composition and organization. **c** The results showed the characterization FTUs based on the gradient pixel-level images for A6. The heatmap depicts the SSIM score, where a higher score corresponds to a lighter color and greater structural similarity. **d** A heatmap showcasing the protein expression of each FTUs represented by the six SVP clusters, which were identified as FTUs resembling secondary follicles. The values in the heatmap are scaled by *z*-scores of protein expression. **e–h** Overlays of CODEX images for SVPs for FOVs 1, 3, 4, and 6, respectively. **i.** Spatial maps depicting the patterns of secondary follicle FTUs from six FOVs. Dash rectangles indicate the identified follicle regions. Note that panels d to h are ordered by FOV 1, 2, 3, 4, 5, and 6. **j** Cell phenotype maps of the FTUs identified in (**i**). **k** Barplots depicting the cell components of the identified FTU in (**i**). The cell type colors were depicted in (**b**). **l** The graph network depicting the spatial proximity of the top 5 abundant cell types in the FTU identified in **i**, as calculated by $\frac{1}{1+d}$, where *d* represents the average distance between any two cell types. **m** Dumbbell plots indicated significant cell–cell interaction among B cells and others. If the observed distance is significantly smaller than the expected distance, the two cell types tend to be contacted and interact. Line length represents relative distances, subtracting the expected distance from the observed distance. An empirical permutation test was used to calculate the *p*-value, and the point size was scaled using an adjusted *p*-value.

(Supplementary Fig. 8c). We subsequently calculated the structural similarity score (SSIM) to quantitatively evaluate pattern similarity among identified FTUs. Each gradient pixel size image identified six FTUs, and those patterns of FTUs showed pairwise consistency (Fig. 4c and Supplementary Fig. 8d), suggesting that 200-by-200 pixel downsampled images (an approximate factor of 105-fold from the original pixel size) were sufficient in characterizing FTUs to balance between computational efficiency and biological insights.

We next implemented SpaGFT to characterize FTUs for the six FOVs with 200-by-200 pixel images and annotated follicles for each FOV based on cell components (Supplementary Fig. 9) and protein signatures (Supplementary Data 13; Supplementary Figs. 10a, b). Specifically, FTUs represented by SVP cluster 1 of A1 and SVP cluster 1 of A2 displayed morphological features akin to that of a mantle zone (MZ). Molecularly, we uncovered that the B cell-specific marker[44] (CD20) and anti-apoptotic factor (BCL-2)[45] were SVPs for these two FTUs of A1 and A2 (Fig. 4d, e and Supplementary Fig. 10c). Our results confirmed the presence of CD20 in delineating the MZ structure, and additionally suggest that the presence of BCL-2 as an additional feature of MZ structures[46]. In another case, FTUs represented SVP cluster 4 of A3, SVP cluster 9 of A4, and SVP cluster 4 of A5 displayed GC-specific T cell signatures (Fig. 4f, g and Supplementary Fig. 10d) and corresponding molecular features, including PD-1[47] and CD57[48], indicating the presence of well-characterized GC-specific T follicular helper cells[49]. For FTUs represented by SVP cluster 2 of A6, we observed a complex molecular environment, where Podoplanin, CD11c, and CD11b were SVPs, thus showcasing the existence of follicular dendritic cell (FDC)[50] and GC-centric macrophages[51] networks (Fig. 4h). In addition to molecular heterogeneity, we further captured their variability in terms of length-scale and morphology (Fig. 4i), cell type (Fig. 4j and k), cell–cell distance (Fig. 4l), and cell-cell interactions (Fig. 4m). For example, from the tissue morphology perspective, A3–A6 captured clear oval shape patterns with different length-scales, but A1 and A2 captured multiple partial MZ patterns (Fig. 4i). Although visual inspection was unable to distinguish between the morphological patterns of GCs in A4 (Fig. 4b), SpaGFT was able to determine three small length-scale GC patterns at the molecular level (Fig. 4i).

Regarding cellular characteristics, six FTUs (i.e., two MZ from A1 and A2; four GCs from A3 to A6) were dominated by B and CD4 T cells with varying proportions (Fig. 4j-k; Supplementary Data 14). Specifically, MZs from A1 and A2 showed an average composition of 58% B and 10% CD4 T cells. GC from A3 and A5 with similar length-scale showed an average of 54% B and 32% CD4 T cells. A4 captured three length-scale GCs and showcased 43% B and 46% CD4 T cells, while the large-scale GC from A6 contained 70% B and 12% T cells, indicating B and T cell proportions varying in different length-scale GC. We could also infer cell–cell interaction based on distance (Fig. 4l, m). In general, MZ from A1 and A2 show that the observed B–B distance was smaller than the expected distance, which suggests the homogeneous biological process of the significant B–B interaction in the GC region. In

addition, cell-cell interaction also shows heterogeneity for two MZ. The interactions between CD4 T cells and B cells were observed in two MZ from A1 and A2, showcasing the infiltration of CD4 T cells into the B cell right mantle zone[52]. DC-B and CD4 T-B cell interactions in A3 and A4 suggest light zone functions for B cell selection[53,54]. Macrophage-B cell interactions in GC in A6 potentially indicated macrophage regulation on B cells (e.g., B cells that failed to trigger the cell proliferation signals during the B cell selection process underwent apoptosis and were subsequently engulfed by macrophages[55]). Our results demonstrate the applicability of SpaGFT at an initial subsampled lower resolution from high-plex spatial proteomics, thus efficiently identifying and characterizing high-attention tissue regions, including secondary follicles, to uncover cellular and molecular variability that can be further confirmed at the original single-cell resolution. We also affirmed that FTUs identified by SpaGFT were not simply regions of cell aggregation but reflected both the cellular and regional activity and cell–cell interactions based on spatially orchestrated molecular signatures.

SpaGFT can generate new features and be implemented as an explainable regularizer for machine-learning algorithms

SpaGFT can also be beneficial to enhance the performance of existing methods as an explainable regularizer through feature or objective engineering. To elucidate its applicative power, we exemplified three representative analyses of SRT as follows (Supplementary Fig. 11 and see the "Methods" section).

First, we showcase how spot clustering can identify spatial domains spatially coherently in both gene expression and histology. Here, we selected SpaGCN[4] as the demonstration to showcase the implementation of FC from the feature engineering aspect (Fig. 5a). To illustrate FCs being a feature, we extended the spatial expression matrix by concatenating a spot-by-FC matrix derived from the spot-spot similarity. Subsequently, the new feature matrix was input into the original SpaGCN model and predicted spatial domains. Same as the SpaGCN study, we utilized 12 datasets[24] of human brain SRT data for training (two datasets from the same tissue section) the number of new features and testing for improving SpaGCN on 10 datasets. The results indicated improvements in eight out of ten datasets (Supplementary Data 15) in identifying the spatial domains of the dorsolateral prefrontal cortex. Notably, the top five datasets exhibited enhancements between 7.8% and 42.6%.

Second, annotation transfer will solve the challenge of insufficient data labeling for the increasing emergence of SRT. We used TACCO[6] as an annotation transfer example tool to showcase the application of FC as a regularizer for the optimal transport (OT) method, which is a machine learning method that aimed to find the most efficient way (i.e., minimizing the overall cost associated with the movement) to move a probability distribution from one configuration to another. Specifically, TACCO allowed the transfer of phenotype-level annotation labels (e.g., cell type) from scRNA-seq to SRT using such an OT framework. Although TACCO has demonstrated algorithm
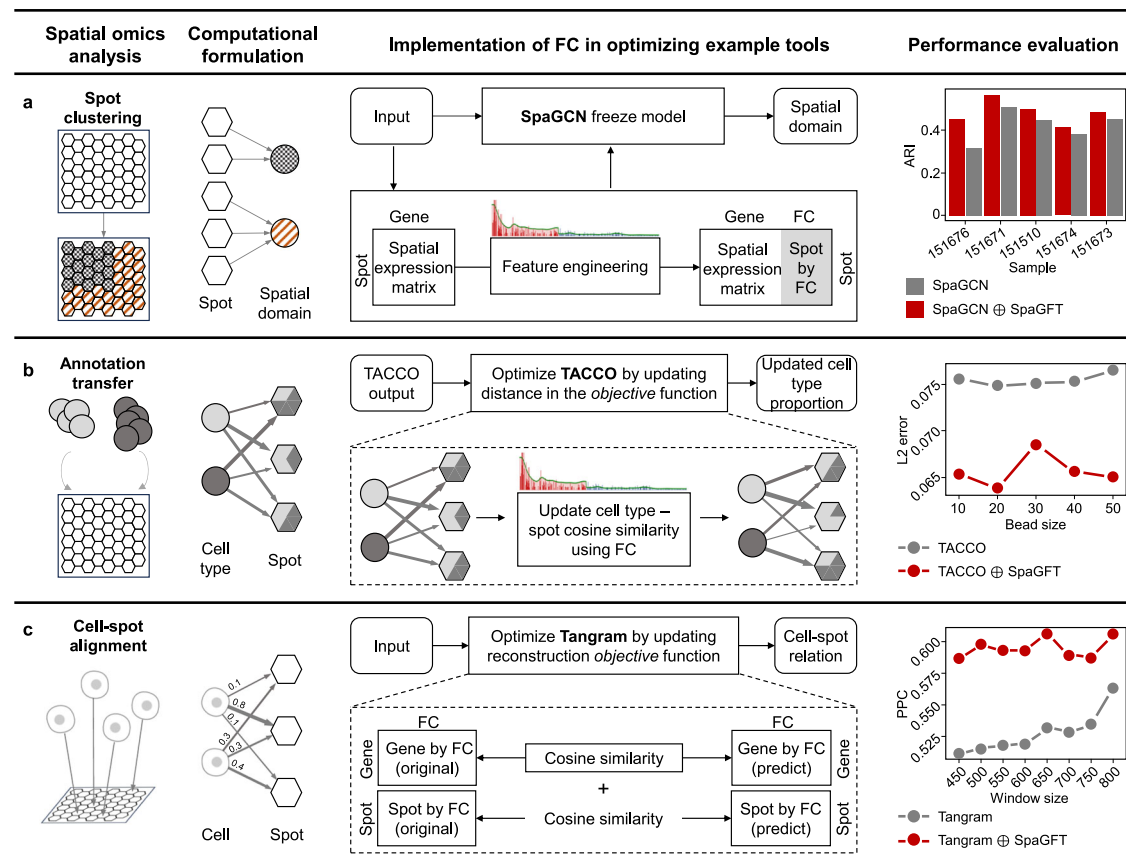
Fig. 5 | SpaGFT implementation for three cell-centric tools and the figure consists of four columns, each corresponding to spatial omics analysis, computational formulation, implementation of FC in optimizing examples tools, and performance evaluation. a Spot clustering can be formulated as a many-to-one mapping problem. Regarding the modified workflow of SpaGCN, we changed the original input of SpaGCN. A newly formed matrix was then placed into the frozen SpaGCN model for computation. The top 5 performance-increased samples are distinctly showcased, where the *y*-axis is the ARI value, and the *x*-axis is the sample number. **b** Annotation transfer is formulated as a many-to-many mapping problem. Regarding the modified workflow of TACCO, we modified the cost matrix for optimal transport. In the new cost matrix calculation method, we use weighted FCs as the feature to calculate the distance between CT and spots and then optimize the baseline mapping matrix (e.g., TACCO output). In the evaluation, we refer to TACCO methods to simulate spots with different bead sizes using scRNA-seq data and use L2 error to measure differences between predicted and known cell composition in each simulated spot. The y-axis is the bead size for a simulation data value, and the x-axis is the L2 error. Lower L2 error scores indicate better performance. **c** The cell-spot alignment can be formulated as a many-to-many mapping problem. Regarding the modified workflow of Tangram, we have added two additional constraint terms to its original objective function. The first constraint is designed from a gene-centric perspective, calculating the cosine similarity of the gene by FC matrix between the reconstructed and the original matrix. The second constraint is designed from a cell-centric perspective, calculating the cosine similarity on the spot by the FC matrix between the reconstructed and the original matrix. In the evaluation, we first simulate spatial gene expression data using different window sizes based on STARmap data. Subsequently, we measure the similarity between predicted and known cell proportions in each simulated spot using the Pearson correlation coefficient. A higher PPC indicates better performance (Source data are provided as a Source Data file).

effectiveness in consideration of cell similarity over all genes, we hypothesized that projecting cell similarity to the frequency domain and strengthening a topological regularization in OT's objective function will be a potential avenue for performance enhancement. In our modification, we integrated a topological regularization term into the original cost matrix to derive a new cost matrix (Fig. 5b and see the "Methods" section). Leveraging the evaluation metrics of the original TACCO study, our tests underscored an 8.7–14.9% L2-error decrease across five simulated bead sizes in terms of transferring annotated labels from scRNA-seq to unannotated SRT mouse brain data (Supplementary Data 16).

Third, aligning single-cell data (e.g., scRNA-seq) to low-/high-resolution SRT data was important to mutually benefit each other regarding spatial resolution and molecular diversity. We selected Tangram[7] as an alignment tool to demonstrate the topological regulation of genes and spots in the frequency domain. Tangram optimized the cell-to-spot mapping matrix through the gradient-based method, aiming to ensure the similarity between the reconstructed SRT based on scRNA-seq and the original SRT. The objective function

of Tangram is to measure cell density, gene-level similarity, and cell-level similarity in the vertex domain, respectively. In alignment with the hypothesis proposed in Fig. 5b, we constrained the similarity at both the gene- and cell-level in the frequency domain (Fig. 5c). As a result, our tests illustrated 7.4–15.9% Pearson correlation coefficient increase improvement regarding aligning scRNA-seq on simulated STARmap[56] mouse brain SRT data (Supplementary Data 17).

SpaGFT introduces an inductive bias to regularize the deep learning method and identify rare subcellular organelles

We applied SpaGFT to obtain an interpretable spreading entropy regulation for a conditional variational autoencoder framework, CAMPA, to identify conserved subcellular organelles across multiple perturbed conditions on pixel-level 4i data (165 nm/pixel)[16,17]. To modify the model, we introduced an entropy term in the original reconstruction loss of CAMPA to regularize the spreading of graph signals[19]. Specifically, we constrained the entropy within the first *k* bandwidth and provided an inductive assumption for CAMPA to learn embeddings that represented *k*-bandlimited signals (Supplementary
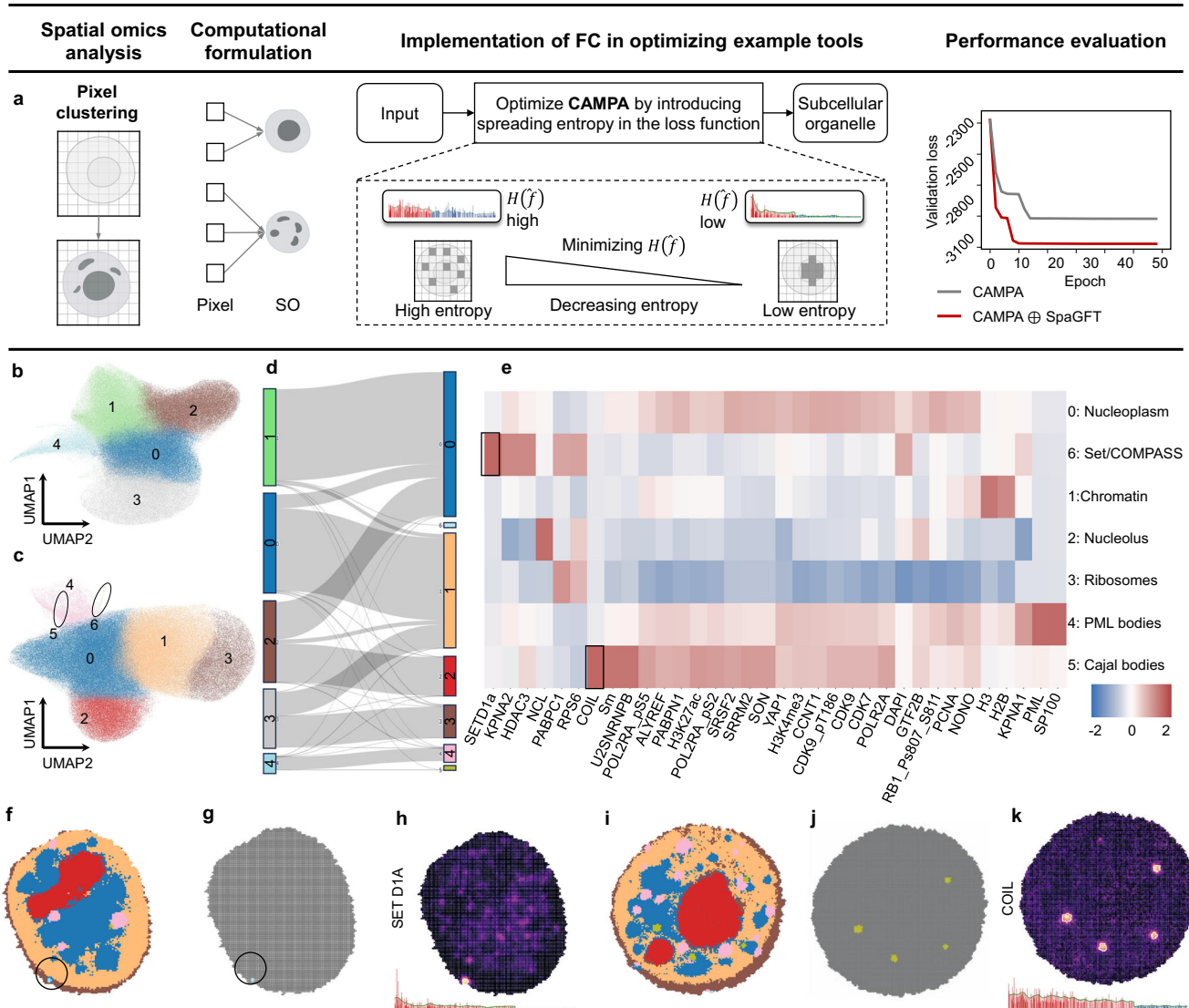
Fig. 6 | The case for implementing FC to CAMPA. a The first column shows pixel clustering concepts. In this bipartite graph (second column), pixel clustering can be formulated as a many-to-one mapping problem, where the source node represents the pixel, the target node represents the subcellular organelle, and the edges denote the corresponding mapping relationships. Regarding the modified workflow of CAMPA (third column), we have made a modification to the original loss function. The modified term aims to measure the spreading of graph signals in the reconstructed image. In the frequency domain, this spreading can be quantified using spreading entropy (see the "Methods" section). A spreading graph signal corresponds to high entropy, while a non-spread graph signal corresponds to low entropy. Therefore, the new regularizer term aims to minimize the spreading entropy. In the evaluation (fourth column), we used the validation loss, which was calculated using the same loss function and validation dataset to examine the contribution of the spreading entropy to the model training. The *y*-axis is the validation loss value, and the *x*-axis is the number of epochs for training the CAMPA model. **b**. UMAP shows five-pixel clusters predicted by the baseline model using the Leiden clustering algorithm at 0.2 resolution. **c** UMAP shows seven-pixel clusters predicted by the modified model using the Leiden clustering algorithm at 0.2 resolution. Two rare clusters were circled in this panel. **d** The Sanky plot shows the cluster changes from baseline model prediction to modified model prediction. **e** The heatmap shows the annotation of each cluster (modified model at resolution 0.2) using a human protein atlas. The column of the heatmap is the protein intensity in the cell nucleus, and the row corresponds to clusters. **f**–**h** The three figures showcase the overview of predicted pixel clusters, cluster 6, and marker protein for cell 224081. **i**–**k** The three figures showcase the overview of predicted pixel clusters, cluster 5, and marker protein for cell 367420.

Fig. 12a). Consequently, compared to the validation loss calculated from validation datasets (see the "Methods" section), the loss curve from the modified model showed a reduction and entered a stable state earlier (Fig. 6a). We observed that, by introducing the entropy term as a regularizer, the model enhanced the training efficacy in capturing and minimizing the reconstruction error and promoted faster convergence of the model.

Furthermore, we validated that the modified model significantly (*p*-value = 0.035) improved the baseline model regarding batch effect removal (Supplementary Fig. 12b–e) using kBET testing[57], indicating that the learned embeddings retained conserved structures of subcellular organelles across multiple perturbations. Next, compared with the baseline (Fig. 6b–d), the modified model additionally identified two rare clusters (Supplementary Data 18), including cluster 5 (with an average of 0.16% pixels per cell) and cluster 6 (with an average of 0.10% pixels per cell). Notably, the pixels assigned to these two clusters are very stable (not random signals computationally) regardless of different resolution parameters of the Leiden clustering algorithm (Supplementary Data 19 and Supplementary Fig. 12f). Subsequently, clusters 5 and 6 were annotated as Cajal bodies[58] and set1/COMPASS[59], respectively (Fig. 6e). Cluster 6 and its corresponding protein signature, SETD1A (Fig. 6f–h), displaying a highly concentrated pattern (with an average of 0.16% pixels per cell), were strongly shown as a *k*-bandlimited signal in the frequency domain. Furthermore, we also

observed similar characteristics of cluster 5 and the corresponding marker protein, COIL (Fig. 6i–k). Therefore, by integrating the regularization of low-frequency signals from SpaGFT, the CAMPA model's stability was enhanced in learning embeddings that represented subcellular organelles with k-bandlimited characteristics. This approach, which we term "explainable regularization," refines the detection and characterization of finer structures exhibiting spatially organized patterns.

Discussion

SpaGFT provides a reliable feature representation through graph Fourier transform that enhances our biological understanding of complex tissues. This method aligns with the advanced analytical capabilities required to dissect the intricate spatial components of tissue biology, from subcellular to multicellular scales. It eliminates the need for pre-defined expression patterns and significantly improves computational efficiency, as demonstrated in the benchmarking across 31 human/mouse Visium and Slide-seq V2 datasets. In addition, we also highlight our manually curated 458 mouse and human brain genes as close-to-optimization standard SVGs. This will bring an alternative evaluation metric based on realistic human/mouse data, which is complementary to simulation-based evaluation methods, such as BSP[60], SPARK-X, SpatialDE, SPARK, scGCO, and other benchmarking work[61]. Furthermore, implementing a low-pass filter and inverse GFT effectively impute low-expressed gene expression and denoise high-noisy protein intensity, leading to more precise spatial domain predictions, as showcased in the human dorsolateral prefrontal cortex. Notably, SpaGFT advances the interpretation of spatial omics data by enabling more accurate machine learning predictions. It has notably improved the performance of existing frameworks by 8−40% in terms of the accuracy of spatial domain identification, lower error of annotation transfer from cell types to spots, the correctness of the cell-to-spot alignments, and the validation loss of subcellular hallmark inference, respectively.

From a computational standpoint, SpaGFT and scGCO are two graph representation methods, among others, for spatial omics data analysis, with the former focusing on omic feature representation and the latter focusing on SVG detection. scGCO employs a graph-cut method to segment the tissue and compare the consistency between segmentations and gene expressions in support of SVG detection. SpaGFT uses the graph Fourier transform to find a novel latent space to represent gene expression and achieve various downstream tasks, including but not limited to, SVG identification, gene expression enhancement, and functional tissue unit inference.

In addition, there is a good potential for implementing SpaGFT into existing explainable spatial multi-modalities frameworks[2], such as UnitedNet[62], MUSE[63], and modalities-autoencoder[64]. Considering UnitedNet[62] as an example, it incorporates explainable machine learning techniques to dissect the trained network and quantify the relevance of features across different modalities, specifically looking at cell-type-specific relationships. To bring more spatial insight into UnitedNet, SpaGFT can provide (1) augmented features (e.g., modified SpaGCN in Fig. 5a) and (2) an explainable regularizer (e.g., modified CAMPA in Fig. 6). To generate the augmented spatial omics features, SpaGFT can first calculate cell-cell relations (e.g., calculation from H&E features, gene expression, or protein intensity) in the vertex domain and transform the relations to FCs, The FCs encode and quantify cell−cell variation patterns, which can be regarded as one of the inputs for UnitedNet. Regarding implementing SpaGFT as an explainable regularizer, the spreading entropy can be introduced into UnitedNet's reconstruction loss function, as UnitedNet has an encoder-decoder structure. By regularizing the entropy of encoded and decoded spatial omics features on the Fourier domain, the UnitedNet may be guided to learn spatially organized regions that

present a low-frequency signal (e.g., one functional tissue unit with a specific pattern and function). These enhancements are pivotal in characterizing complex biological structures using explainable regularization for deep learning framework, including identifying rare subcellular organelles, thus providing deeper insights into the cellular machinery.

Regarding the biological implications, SpaGFT offers alternative perspectives on spatial biology questions. Specifically, by grouping SVGs identified by SpaGFT, we can uncover distinct FTUs within organs. This has led to the identification of critical immunological regions in the human lymph node Visium data, enhancing our knowledge of B cell maturation and the polyfunctional areas it encompasses, such as the B cell zone, T cell zone, GC, B−T zone, GC−B zone, T−GC, and tri-zone Additionally, exclusive in-house CODEX data, SpaGFT has revealed secondary follicle differences in the morphology, molecular signatures, and cellular interactions in the human tonsil, offering a more nuanced understanding of B cell maturation. Additionally, SpaGFT introduces k-bandlimited signal entropy within the CAMPA framework. This has led to the groundbreaking identification of rare subcellular organelles, which are the Cajal body and the Set1/COMPASS complex. The former is integral to the regulation of gene expression, while the latter plays a critical role in epigenetic modifications. By enabling the investigation of these organelles with unprecedented detail, SpaGFT propels us closer to a comprehensive understanding of the spatial dynamics of gene expression and the epigenetic landscape within cells.

However, there is still room for improving prediction performance and understanding the FTU mechanism. First, SpaGFT discusses low-frequency signals in the frequency domain, but there is a lack of discussion on medium- and high-frequency signals. Although a previous study[65] described that most functionally related biological signals are presented in the low-frequency region, certain special signals are also found in the medium and high-frequency region. For instance, in the human brain fMRI (functional magnetic resonance imaging, a technique that measures brain activity by detecting changes associated with blood flow), low-frequency FMs capture the global variation signals (e.g., daydreaming and retrieving memories). Medium-frequency FMs capture brain networks with less global variation but more rapid processing (e.g., working memory or executive functions). High-frequency FMs capture responses to new or complex stimuli that involve local connections between close brain regions (e.g., acute, localized brain activities). Analogous to spatial omics data, we assume that medium and high-frequency signals may also have corresponding special biological signals with more local and less global variation (e.g., regions stimulation from the environment), complementing the current k-bandlimited signal approach of representing smooth global variation. Therefore, in future studies, we might focus more on multi-frequency signal interpretation. Second, although the SpaGFT computation speed is very competitive, it can be further enhanced by reducing the computational complexity from $O(n^2)$ to $O(n \times \log(n))$ using fast Fourier transform algorithms[66]. Third, the alteration of the spot graph and FTU topology represents a potential challenge in identifying FTUs across spatial samples from different tissues or experiments, which results in diverse FM spaces and renders the FCs incomparable. This is similar to the "batch effect" issue in multiple single-cell RNA sequencing (scRNA-seq) integration analyses. One possible solution to this challenge is to embed and align spatial data points to a fixed topological space using machine learning frameworks, such as optimal transport. Another possibility is to use H&E images as a common reference for all to make the embedding tissue-aware. Fourth, SpaGFT implementation on the CODEX image relies on experts' knowledge to pre-select functional regions. The future direction of analyzing multiplexed images is to develop a topological learning framework to automatically detect and segment functional objects based on SpaGFT feature representation. Overall, we believe

the value of our study is to bring an alternative view for explainable artificial intelligence in spatial omics modeling, including multi-resolution spatial omics data integration and pattern analysis across spatiotemporal data[13].

Methods

We introduce Spatial Graph Fourier Transform (SpaGFT) to represent spatial omics features. The core concept of SpaGFT is to transform spatial omics features into Fourier coefficients (FC) for downstream analyses, such as SVG identification, expression signal enhancement, and topological regularization for other machine algorithms. SpaGFT framework provides graph signal transform and seven downstream tasks: SVG identification, gene expression imputation, protein signal denoising, spatial domain characterization, cell type annotation, cell-spot alignment, and subcellular landmark inference. The detailed theoretical foundation of k-bandlimited signal recognition can be found in Supplementary Note 1.

Graph signal transform

K-nearest neighbor (KNN) Graph construction. Given a gene expression matrix containing n spots, including their spatial coordinates and m genes, SpaGFT calculates the Euclidean distances between each pair of spots based on spatial coordinates first. In the following, an undirected graph $G = (V, E)$ will be constructed, where $V = \{v_1, v_2, \ldots, v_n\}$ is the node set corresponding to n spots; E is the edge set while there exists an edge e_{ij} between v_i and v_j in E if and only if v_i is the KNN of v_j or v_j is the KNN of v_i based on Euclidean distance, where $i, j = 1, 2, \ldots, n$; and $i \neq j$. Based on the benchmarking results in Supplementary Data 4, the default K is defined as $1*\sqrt{n}$ among $0.5*\sqrt{n}$, $1*\sqrt{n}$, $1.5*\sqrt{n}$, and $2*\sqrt{n}$. Note that all the notations of matrices and vectors are bolded, and all the vectors are treated as column vectors in the following description. An adjacent binary matrix $\mathbf{A} = (a_{ij})$ with rows and columns as n spots is defined as:

$$a_{ij} = \begin{cases} 1, e_{ij} \in E \\ 0, \text{else.} \end{cases} \tag{1}$$

A diagonal matrix $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_n)$, where $d_i = \sum_{j=1}^{n} a_{ij}$ represents the degree of v_i.

Fourier mode calculation. Using matrices $\mathbf{A}$ and $\mathbf{D}$, a Laplacian matrix $\mathbf{L}$ can be obtained by

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \tag{2}$$

$\mathbf{L}$ can be decomposed using spectral decomposition:

$$\mathbf{L} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{T}}$$
$$\boldsymbol{\Lambda} = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{3}$$
$$\mathbf{U} = (\boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \ldots, \boldsymbol{\mu_n})$$

where the diagonal elements of $\boldsymbol{\Lambda}$ are the eigenvalues of $\mathbf{L}$ with $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, where λ_1 is always equal to 0 regardless of graph topology. Thus, λ_1 is excluded from the following analysis. The columns of $\mathbf{U}$ are the unit eigenvector of $\mathbf{L}$. $\boldsymbol{\mu}_k$ is the kth Fourier mode (FM), $\boldsymbol{\mu}_k \in \mathbb{R}^n$, $k = 1, 2, \ldots, n$, and the set $\{\boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \ldots, \boldsymbol{\mu_k}\}$ is an orthogonal basis for the linear space. For $\boldsymbol{\mu}_k = (\mu_k^1, \mu_k^2, \ldots, \mu_k^n)$, where μ_k^i indicates the value of the kth FM on node v_i, the smoothness of $\boldsymbol{\mu}_k$ reflects the total variation of the kth FM in all mutual adjacent spots, which can be formulated as

$$\frac{1}{2} \sum_{v_i \in V} \sum_{v_j \in V} a_{ij} \left(\mu_k^i - \mu_k^j\right)^2 \tag{4}$$

The form can be derived by matrix multiplication as

$$\frac{1}{2} \sum_{v_i \in V} \sum_{v_j \in V} a_{ij} \left(\mu_k^i - \mu_k^j\right)^2 = \frac{1}{2} \left[\sum_{v_i \in V} d_i \left(\mu_k^i\right)^2 \right.$$
$$\left. -2 \sum_{v_i \in V} \sum_{v_j \in V} a_{ij} \mu_k^i \mu_k^j + \sum_{v_j \in V} d_j \left(\mu_k^j\right)^2 \right]$$
$$= \sum_{v_i \in V} d_i \left(\mu_k^i\right)^2 - \sum_{v_i \in V} \sum_{v_j \in V} a_{ij} \mu_k^i \mu_k^j \tag{5}$$
$$= \boldsymbol{\mu_k}^{\mathsf{T}} \mathbf{D} \boldsymbol{\mu_k} - \boldsymbol{\mu_k}^{\mathsf{T}} \mathbf{A} \boldsymbol{\mu_k}$$
$$= \boldsymbol{\mu_k}^{\mathsf{T}} \mathbf{L} \boldsymbol{\mu_k}$$
$$= \lambda_k$$

where $\boldsymbol{\mu_k}^{\mathsf{T}}$ is the transpose of $\boldsymbol{\mu_k}$. According to the definition of smoothness, if an eigenvector corresponds to a small eigenvalue, it indicates the variation of FM values on adjacent nodes is low. The increasing trend of eigenvalues corresponds to an increasing trend of oscillations of eigenvectors; hence, the eigenvalues and eigenvectors of $\mathbf{L}$ are used as frequencies and FMs in our SpaGFT, respectively. Intuitively, a small eigenvalue corresponds to a low-frequency FM, while a large eigenvalue corresponds to a high-frequency FM.

Graph Fourier transform. The graph signal of a gene g is defined as $\mathbf{f_g} = \left(f_g^1, f_g^2, \ldots, f_g^n\right) \in \mathbb{R}^n$, which is a n-dimensional vector and represents the gene expression values across n spots. The graph signal $\boldsymbol{f_g}$ is transformed into a Fourier coefficient $\hat{\mathbf{f}}_\mathbf{g}$ by

$$\hat{\mathbf{f}}_\mathbf{g} = \left(\hat{f}_g^1, \hat{f}_g^2, \ldots, \hat{f}_g^n\right) = \mathbf{U}^{\mathsf{T}} \mathbf{f_g} \tag{6}$$

In such a way, $\hat{f}_g^k$ is the projection of $\mathbf{f_g}$ on FM $\boldsymbol{\mu}_k$, representing the contribution of FM $\boldsymbol{\mu}_k$ to graph signal $\mathbf{f_g}$, k is the index of $\mathbf{f_g}$ (e.g., $k = 1, 2, \ldots, n$). This Fourier transform harmonizes gene expression and its spatial distribution to represent gene g in the frequency domain. The details of SVG identification using $\hat{\mathbf{f}}_\mathbf{g}$ can be found below.

SVG identification

GFTscore definition. We designed a GFTscore to quantitatively measure the randomness of gene expressions distributed in the spatial domain, defined as

$$\text{GFTscore}\left(f_g\right) = \sum_{k=1}^{n} e^{-\lambda_k} \tilde{f}_g^k \tag{7}$$

where λ_k is the pre-calculated eigenvalue of $\mathbf{L}$, and $e^{-\lambda_k}$ is used to weigh the $\tilde{f}_g^k$ to further enhance the smoothness of the spatial omics variation signal and reduce its noisy components (Supplementary Note 1S2.3)[18,67]. The normalized Fourier coefficient $\tilde{f}_g^k$ is defined as

$$\tilde{f}_g^k = \frac{|\hat{f}_g^k|}{\sum_{i=1}^{n} |\hat{f}_g^i|} \tag{8}$$

The gene with a high GFTscore tends to be SVG and vice versa. Therefore, all m genes are decreasingly ranked based on their GFTscore from high to low and denote these GFTscore values as $y_1 \geq y_2 \geq \ldots \geq y_m$. In order to determine the cutoff y_z to distinguish SVG and non-SVGs based on GFTscore, we applied the Kneedle algorithm[68] to search for the inflection point of a GFTscore curve described in Supplementary Note 1.

Wilcoxon rank-sum test implementation for determining SVGs. Although the above GFTscore is an indicator to rank and evaluate the

potential SVGs, a rigorous statistical test is needed to calculate the p-value for SVGs and control type I error. First, SpaGFT determines low-frequency FM and high-frequency FMs and corresponding FCs by applying the Kneedle algorithm to the eigenvalues of **L**. The inflection points are used for determining the low-frequency FMs and high-frequency FMs when the direction parameters are 'increasing' and 'decreasing', respectively. Second, the Wilcoxon rank-sum test is utilized to test the differences between low-frequency FCs and high-frequency FCs to obtain statistical significance. If a gene has a high GFTscore and significantly adjusted p-value, the gene can be regarded as an SVG. We use $f = (f_1, f_2, \ldots, f_n)$ to represent the expression of a random signal on n spots. If the gene corresponding to the graph signal is a non-SVG, the gene expressions on neighboring spots are independent. Otherwise, it will exhibit spatial dependence. Hence, we can assume that $(f_1, \ldots, f_n) \sim N(\mu_f, \sigma_f^2 I)$, similar in SpatialDE[11], where μ_f, σ_f^2 and I are the mean, variance, and identity matrix, respectively. In this case, each f_i follows a Gaussian distribution, which is independent and identically distributed. By implementing GFT on $(f_1, f_2, \ldots, f_n)$, we obtain Fourier coefficients $FC_1, FC_2, \cdots, FC_p$, where p is the number of low-frequency FCs and reflects the contributions from low-frequency FMs. We also obtain the $FC_{p+1}, FC_{p+2}, \cdots, FC_{p+q}$, where q is the number of high-frequency FCs and reflects the contributions from noise. Hence, we form the null hypothesis that no difference exists between low-frequency FCs and high-frequency FCs (Proof can be found in S3 of Supplementary Note 1). Accordingly, a non-parametrical test (i.e., Wilcoxon rank-sum test) is used for testing the difference between median values of low-frequency FCs and high-frequency FCs. Especially, the null hypothesis is that the median of low-frequency FCs of an SVG is equal to or lower than the median of high-frequency FCs. The alternative hypothesis is that the median of low-frequency FCs of an SVG is higher than the median of high-frequency FCs. The p-value of each gene is calculated based on the Wilcoxon one-sided rank-sum test and then adjusted using the false discovery rate (FDR) method. Eventually, a gene with GFTscore higher than y_z and adjusted p-value less than 0.05 is considered an SVG.

Benchmarking data setup
Dataset description. Thirty-two spatial transcriptome datasets were collected from the public domain, including 30 10X Visium datasets (18 human brain data, 11 mouse brain data, and one human lymph node data) and two Slide-seqV2 datasets (mouse brain). More details can be found in Supplementary Data 1. Those samples were sequenced by two different SRT technologies: 10X Visium measures ~55 μm diameter per spot, and Slide-seqV2 measures ~10 μm diameter per spot. Three datasets were selected as the training sets for grid-search parameter optimization in SpaGFT, including two highest read-depth datasets in Visium (HE-coronal) and Slide-seqV2 (Puck-200115-08), one signature dataset in Maynard's study[24]. The remaining 28 datasets (excluding lymph node data) were used as independent test datasets.

Data preprocessing. For all 32 datasets, we adopt the same preprocessing steps based on *squidpy* (version 1.2.1), including filtering genes that have expression values in <10 spots, normalizing the raw count matrix by counts per million reads method, and implementing log-transformation to the normalized count matrix. No specific pre-processing step was performed on the spatial location data.

Benchmarking SVG collection. We collected SVG candidates from five publications[24–28], with data from either human or mouse brain subregions. (i) A total of 130 layer signature genes were collected from Maynard's study[24]. These genes are potential multiple-layer markers validated in the human dorsolateral prefrontal cortex region. (ii) A total of 397 cell-type-specific (CTS) genes in the adult mouse cortex were collected from Tasic's study (2016 version)[28]. The authors performed scRNA-seq on the dissected target region, identified 49 cell

types, and constructed a cellular taxonomy of the primary visual cortex in the adult mouse. (iii) A total of 182 CTS genes in mouse neo-cortex were collected from Tasic's study[27]. Altogether, 133 cell types were identified from multiple cortical areas at single-cell resolution. (iv) A total of 260 signature genes across different major regions of the adult mouse brain were collected from Ortiz's study[25]. The authors utilized spatial transcriptomics data to systematically profile sub-regions and delivered the subregional genes using consecutive coronal tissue sections. (v) A total of 86 signature genes in the cortical region shared by humans and mice were collected from Hodge's study[26]. Collectively, a total of 849 genes were obtained, among which 153 genes were documented by multiple papers. More details, such as gene names, targeted regions, and sources, can be found in Supplementary Data 2.

Next, the above 849 genes were manually validated on the in-situ hybridization (ISH) database deployed on the Allen Brain Atlas (https://mouse.brain-map.org/). The ISH database provided ISH mouse brain data across 12 anatomical structures (i.e., Isocortex, Olfactory area, Hippocampal formation, Cortical subplate, Striatum, Pallidum, Thalamus, Hypothalamus, Midbrain, Pons, Medulla, and Cerebellum). We filtered the 849 genes as follows: (i) If a gene is showcased in multiple anatomical plane experiments (i.e., coronal plane and sagittal plane), it will be counted multiple times with different expressions in the corresponding experiments, such that 1327 genes were archived (Supplementary Data 3). (ii) All 1327 genes were first filtered by low gene expressions (cutoff is 1.0), and the FindVariableFeatures function ("vst" method) in the Seurat (v4.0.5) was used for identifying highly variable genes across twelve anatomical structures. Eventually, 458 genes were kept and considered as curated benchmarking SVGs. The evaluation criteria can be found in Supplementary Note 2.

Statistics and reproducibility
In our benchmarking experiment, we implemented a two-sided Wilcoxon-rank sum test to conduct a significance test. No data were excluded from the analyses. The experiments were not randomized. Randomization is not relevant to this study since each data was analyzed separately. We then computed the key evaluation metrics, including the Jaccard index, odds ratio, precision, recall, F1 score, Tversky index, Moran's Index, and Geary's C.

SpaGFT implementation and grid search of parameter optimization
A grid-search was set to test for six parameters, including ratio_neighbors (0.5, 1, 1.5, 2) for KNN selection and S (4, 5, 6, 8) for the inflection point coefficient, resulting in 16 parameter combinations. We set $K = \sqrt{n}$ as the default parameter for constructing the KNN graphs in SpaGFT. SVGs were determined by genes with high GFTscore via the KneeLocator function (curve='convex', direction='deceasing', and $S = 6$) in the kneed package (version 0.7.0) and FDR (cutoff is less than 0.05).

Parameter setting of other tools
(i) SpatialDE (version 1.1.3) is a method for identifying and describing SVGs based on Gaussian process regression used in geostatistics. SpatialDE consists of four steps, establishing the SpatialDE model, predicting statistical significance, selecting the model, and expressing histology automatically. We selected two key parameters, design_formula ('0' and '1') in the NaiveDE.regress_out function and kernel_space ("{'SE':[5.,25.,50.],'const':0}", "{'SE':[6.,16.,36.],'const':0}", "{'SE':[7.,47.,57.],'const':0}", "{'SE':[4.,34.,64.],'const':0}", "{'PER':[5.,25.,50.],'const':0}", "{'PER':[6.,16.,36.],'const':0}", "{'PER':[7.,47.,57.],'const':0}", "{'PER':[4.,34.,64.],'const':0}", and "{'linear':0,'const':0}") in the SpatialDE.run function for parameter tunning, resulting in 18 parameter combinations.

(ii) SPARK (version 1.1.1) is a statistical method for spatial count data analysis through generalized linear spatial models. Relying on

statistical hypothesis testing, SPARX identifies SVGs via predefined kernels. First, raw count and spatial coordinates of spots were used to create the SPARK object via filtering low-quality spots (controlled by min_total_counts) or genes (controlled by percentage). Then the object was followed by fitting the count-based spatial model to estimate the parameters via spark.vc function, which is affected by the number of iterations (fit.maxiter) and models (fit.model). Lastly, ran spark.test function to test multiple kernel matrices and obtain the results. We selected four key parameters, percentage (0.05, 0.1, 0.15), min_total_counts (10, 100, 500) in CreateSPARKObject function, fit.maxiter (300, 500, 700), and fit.model ("poisson", "gaussian") in spark.vc function for parameter tuning, resulting in 54 parameter combinations.

(iii) SPARK-X (version 1.1.1) is a non-parametric method that tests whether the expression level of the gene displays any spatial expression pattern via a general class of covariance tests. We selected three key parameters, percentage (0.05, 0.1, 0.15), min_total_counts (10, 100, 500) in the CreateSPARKObject function, and option ("single", "mixture") in the sparx function for parameter tuning, resulting in 18 parameter combinations.

(iv) SpaGCN (version 1.2.0) is a graph convolutional network approach that integrates gene expression, spatial location, and histology in spatial transcriptomics data analysis. SpaGCN consisted of four steps, integrating data into a chart, setting the graph convolutional layer, detecting spatial domains by clustering, and identifying SVGs in spatial domains. We selected two parameters, the value of the ratio (1/3, 1/2, 2/3, and 5/6) in the find_neighbor_cluster function and res (0.8, 0.9, 1.0, 1.1, and 1.2) in the SpaGCN.train function for parameter tuning, resulting in 20 parameter combinations.

(v) MERINGUE (version 1.0) is a computational framework based on spatial autocorrelation and cross-correlation analysis. It is composed of three major steps to identify SVGs. Firstly, Voronoi tessellation was utilized to partition the graph to reflect the length scale of cellular density. Secondly, the adjacency matrix is defined using geodesic distance and the partitioned graph. Finally, gene-wise autocorrelation (e.g., Moran's I) is conducted, and a permutation test is performed for significance calculation. We selected min.read (100, 500, 1000), min.lib.size (100, 500, 1000) in the cleanCounts function and filterDist (1.5, 2.5, 3.5, 7.5, 12.5, 15.5) in the getSpatialNeighbors function for parameter tuning, resulting in 54 parameter combinations.

(vi) scGCO (version 1.1.2) is a graph-cut approach that integrates gene expression and spatial location in spatial transcriptomics data analysis. scGCO consists of four steps: representing a gene's spatial expression with hidden Markov random field (HMRF), optimizing HMRF with graph cuts with varying hyperparameters, identifying best graph cuts, and calculating the significance of putative SVGs. We selected three parameters, the value of unary_scale_factor (50, 100, and 150) and smooth_factor (5, 10, and 15) in the identify_spatial_genes function for parameter tuning and fdr_cutoff (0.025, 0.05, and 0.075) in the final pipeline for identification of SVG, resulting in 27 parameter combinations.

Visualization of frequency signal of SVGs in PCA and UMAP

Mouse brain (i.e., HE coronal sample) with 2702 spots was used for demonstrating FCs on distinguishing SVG and non-SVG in the 2D UMAP space. SpaGFT determined 207 low-frequency FMs using the Kneedle Algorithm and computed corresponding FCs. PCA was also used for producing low-dimension representation. The transposed and normalized expression matrix was decomposed by using the sc.tl.pca function from the scanpy package (version 1.9.1). The first 207 principal components (PC) were selected for UMAP dimension reduction and visualization. The function sc.tl.umap was applied to conduct UMAP dimension reduction for FCs and PCs.

SVG signal enhancement

An SVG may suffer from low expression or dropout issues due to technical bias[8]. To solve this problem, SpaGFT implemented the low-pass filter to enhance the SVG expressions. For an SVG with an observed expression value $\mathbf{f_g} \in \mathbb{R}^n$, we define $\bar{\mathbf{f}}_\mathbf{g} \in \mathbb{R}^n$ as the expected gene expression value of this SVG, and $\mathbf{f_g} = \bar{\mathbf{f}}_\mathbf{g} + \boldsymbol{\epsilon_g}$, where $\boldsymbol{\epsilon_g} \in \mathbb{R}^n$ represents noises. SpaGFT estimates an approximated FCs $\mathbf{f_g}$ to expected gene expression $\bar{\mathbf{f}}_\mathbf{g}$ in the following way, resisting the noise $\boldsymbol{\epsilon_g}$. The approximation has two requirements (i) the expected gene expression after enhancement should be similar to the originally measured gene expression, and (ii) keep low variation within estimated gene expression to prevent inducing noises. Therefore, the following optimization problem is proposed to find an optimal solution $\mathbf{f_g}$ for $\bar{\mathbf{f}}_\mathbf{g}$

$$\mathbf{f_g^\star} = \operatorname{argmin}_\mathbf{f}[||\mathbf{f} - \mathbf{f_g}||^2 + c\frac{1}{2}\sum_{\boldsymbol{v}_i \in V}\sum_{\boldsymbol{v}_j \in V} a_{ij}(f^i - f^j)^2]$$
$$= \operatorname{argmin}_\mathbf{f}[||\mathbf{f} - \mathbf{f_g}||^2 + c\mathbf{f}^\mathrm{T}\mathbf{Lf}] \qquad (9)$$

where $|| \cdot ||$ is the $L2$-norm, $\mathbf{f} = (f^1, f^2, \ldots, f^n) \in \mathbb{R}^n$ is the variable in solution space, and $i, j = 1, 2, \ldots, n$. c is a coefficient to determine the importance of variation of the estimated signals, and $c > 0$. According to convex optimization, the optimal solution $\mathbf{f_g^\star}$ can be formulated as

$$2\left(\mathbf{f_g^\star} - \mathbf{f_g}\right) + 2c\mathbf{Lf_g^\star} = 0$$
$$\Longrightarrow (\mathbf{I} + c\mathbf{L})\mathbf{f_g^\star} = \mathbf{f_g}$$
$$\Longrightarrow \left(\mathbf{UU}^\mathrm{T} + c\mathbf{U\Lambda U}^\mathrm{T}\right)\mathbf{f_g^\star} = \mathbf{f_g}$$
$$\Longrightarrow \mathbf{U}(\mathbf{I} + c\mathbf{\Lambda})\mathbf{U}^\mathrm{T}\mathbf{f_g^\star} = \mathbf{f_g} \qquad (10)$$
$$\Longrightarrow \mathbf{f_g^\star} = \mathbf{U}(\mathbf{I} + c\mathbf{\Lambda})^{-1}\mathbf{U}^\mathrm{T}\mathbf{f_g} = \mathbf{U}(\mathbf{I} + c\mathbf{\Lambda})^{-1}\hat{\mathbf{f}}_\mathbf{g}$$

where $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, and $\mathbf{I}$ is an identity matrix. $(\mathbf{I} + c\mathbf{\Lambda})^{-1}$ is the low-pass filter and $(\mathbf{I} + c\mathbf{\Lambda})^{-1}\hat{\mathbf{f}}_\mathbf{g}$ is the enhanced FCs. $\mathbf{f_g^\star} = \mathbf{U}(\mathbf{I} + c\mathbf{\Lambda})^{-1}\hat{\mathbf{f}}_\mathbf{g}$ represents the enhanced SVG expression using inverse graph Fourier transform. Specifically, in HE-coronal mouse brain data analysis, we selected 1300 ($= 25\sqrt{n}, n = 2702$) low-frequency FCs for enhancing signal and recovering spatial pattern by using iGFT with $c = 0.0001$.

Data preprocessing on the human prostate cancer Visium data
Cell segmentation. The Visium image of human prostate cancer (adenocarcinoma with invasive carcinoma) from the 10X official website was cropped into patches according to spot center coordinates and diameter length. Each patch is processed by Cellpose for nuclei segmentation using the default parameter. Cell density in each patch is determined using the number of segmented cells.

Microbial alignment
Following the tutorial[69], the corresponding bam files were processed via Kraken packages by (1) removing host sequences and retaining microbial reads, (2) assigning microbial reads to a taxonomic category (e.g., species and genus), and (3) computing the relative abundance of different species in each spot.

SVG signal enhancement benchmarking
Sixteen human brain datasets with well-annotated labels were used for enhancement benchmarking[23,24]. Samples 151510, 151672, and 151673 were used for grid search. Other 13 datasets were used for independent tests. SpaGFT can transform graph signals to FCs, and apply correspondence preprocessing in the frequency domain to realize signal enhancement of genes. Briefly, it is composed of three major steps. Firstly, SpaGFT is implemented to obtain FCs. Secondly, a low-pass filter is applied to weigh and recalculate FCs. Lastly, SpaGFT

implements iGFT to recover the enhanced FCs to graph signals. We select c (0.003, 0.005, 0.007) and *ratio_fms* (13, 15, 17) in the low-pass_enhancement function, resulting in 9 parameter combinations. $c = 0.005$ and ratio_fms = 15 were selected for the independent test. For the parameters used for other computational tools, the details can be found as follows.

(i) SAVER-X (version 1.0.2) is designed to improve data quality, which extracts gene-gene relationships by adopting a deep auto-encoder and a Bayesian model simultaneously. SAVER-X is composed of three major steps roughly. Firstly, train the target data with an autoencoder without a chosen pretraining model. Secondly, filter unpredictable genes using cross-validation. Lastly, estimate the final denoised values with empirical Bayesian shrinkage. Two parameters were considered to explore the performance as well as the robustness of SAVER-X, including batch_size (32, 64, 128) in the saverx function and fold (4, 6, 8) in the autoFilterCV function, resulting in 9 parameter combinations.

(ii) Sprod (version 1.0) is a computational framework based on latent graph learning of matched location and imaging data by leveraging information from the physical locations of sequencing to impute accurate SRT gene expression. The framework of Sprod can be divided into two major steps roughly, which are building a graph and optimizing objective function for such a graph to obtain the de-noised gene expression matrix. To validate its robustness, two parameters were adjusted, including sprod_R (0.1, 0.5) and sprod_latent_dim (8, 10, 12), to generate nine parameter combinations.

(iii) DCA (version 0.3.1) is a deep count autoencoder network with specialized loss functions targeted to denoise scRNA-seq datasets. It uses the autoencoder framework to estimate three parameters (μ, θ, π) of zero-inflated negative binomial distribution conditioned on the input data for each gene. In particular, the autoencoder gives three output layers, representing for each gene the three parameters that make up the gene-specific loss function to compare to the original input of this gene. Finally, the mean (μ) of the negative binomial distribution represents denoised data as the main output. We set neurons of all hidden layers except for the bottleneck to (48, 64, 80) and neurons of bottleneck to (24, 32, 40) for parameter tuning, resulting in 9 parameter combinations.

(iv) MAGIC (version 3.0.0) is a method that shares information across similar cells via data diffusion to denoise the cell count matrix and fill in missing transcripts. It is composed of two major steps. Firstly, it builds its affinity matrix in four steps which include a data preprocessing step, converting distances to affinities using an adaptive Gaussian Kernel, converting the affinity matrix A into a Markov transition matrix M, and data diffusion through exponentiation of M. Once the affinity matrix is constructed, the imputation step of MAGIC involves sharing information between cells in the resulting neighborhoods through matrix multiplication. We applied the knn settings (3, 5, 7) and the level of diffusion (2, 3, 4) in the MAGIC initialization function for parameter tuning, resulting in 9 parameter combinations.

(v) scVI (version 0.17.3) is a hierarchical Bayesian model based on a deep neural network, which is used for probabilistic representation and analysis of single-cell gene expression. It consists of two major steps. Firstly, the gene expression is compressed into a low-dimensional hidden space by the encoder, and then the hidden space is mapped to the posterior estimation of the gene expression distribution parameters by the neural network of the decoder. It uses random optimization and deep neural networks to gather information on similar cells and genes, approximates the distribution of observed expression values, and considers the batch effect and limited sensitivity for batch correction, visualization, clustering, and differential expression. We selected n_hidden (64, 128, 256) and gene_likelihood ('zinb', 'nb', 'poisson') in the model.SCVI function for parameter tuning, resulting in 9 parameter combinations.

(vi) netNMF-sc (version 0.0.1) is a non-negative matrix decomposition method for network regularization, which is designed for the imputation and dimensionality reduction of scRNA-seq analysis. It uses a priori gene network to obtain a more meaningful low-dimensional representation of genes, and network regularization uses a priori knowledge of gene–gene interaction to encourage gene pairs with known interactions to approach each other in low-dimensional representation. We selected d (8, 10, 12) and alpha (80, 100, 120) in the netNMFGD function for parameter tuning, resulting in 9 parameter combinations.

SVG clustering and FTU identification

The pipeline is visualized in Supplementary Fig. 6a. As the pattern of one SVG cluster can demonstrate specific functions of one FTU, the FTU may not necessarily display a clear boundary to its neighbor FTUs. On the contrary, the existence of overlapped regions showing poly-functional regions is allowed. Computationally, the process of FTU identification is to optimize the resolution parameter of the Louvain algorithm for obtaining a certain number of biology-informed FTUs, which minimizes the overlapped area. Denote G' as the set of SVGs identified by SpaGFT. For each resolution parameter $res > 0$, G' can be partitioned to $\{G'_1, G'_2, \ldots, G'_{n_{res}}\}$ (i.e., $\bigcup_k G'_k = G'$ and $G'_k \cap G'_l = \varnothing, \forall k \neq l.$) by applying the Louvain algorithm on FCs, and the resolution will be optimized by the loss function below. Denote $X = (x_{s,g}) \in \mathbb{R}^{|S| \times |G'|}$ as the gene expression matrix, where S is the set of all spots. In the following, for each SVG group G'_k, $\text{pseudo}(s_{s,G'_k}) = \sum_{g \in G'_k} \log(x_{s,g})$ represents the pseudo expression value[4] for spot i. Apply k-means algorithms with $k = 2$ on $\{\text{pseudo}(s_{1,G'_k}), \text{pseudo}(s_{2,G'_k}), \ldots, \text{pseudo}(s_{|S|,G'_k})\}$ to pick out one spot cluster whose spots highly express genes in SVG group G'_k and such spot cluster is identified as a FTU, denoted as $S_i \in S$. Our objective function aims to find the best partition of G' such that the average overlap between any two S_i, S_j is minimized:

$$\text{argmin}_{res > 0} \frac{2 \times \sum_{k \neq l} |S_k \cap S_l|}{n_{res} \times (n_{res} - 1)}$$

SpaGFT implementation on the lymph node Visium data and interpretation

Lymph node SVG cluster identification and FTU interpretation. SVGs were identified on the human lymph node data (Visium) with the default setting of SpaGFT. To demonstrate the relations between cell composition and annotated FTUs, cell2location[35] was implemented to deconvolute spot and resolve fine-grained cell types in spatial transcriptomic data. Cell2location was first used to generate the spot-cell type proportion matrix as described above, resulting in a cell proportion of 34 cell types. Then, pseudo-expression values across all spots for one FTU were computed using the method from the FTU identification section. Then, an element of the FTU-cell type correlation matrix was calculated by computing the Pearson correlation coefficient between the proportion of a cell type and the pseudo-expression of an FTU across all the spots. Subsequently, the FTU-cell type correlation matrix was obtained by calculating all elements as described above, with rows representing FTUs and columns representing cell types. Lastly, the FTU-cell type matrix was generated and visualized on a heatmap, and three major FTUs in the lymph node were annotated, i.e., the T cell zone, GC, and B follicle.

Visualization of GC, T cell zone, and B follicles in the Barycentric coordinate system. Spot-cell proportion matrix was used to select and merge signature cell types of GC, T cell zone, and B follicles for generating a merged spot-cell type proportion matrix (an N-by-3 matrix, N

is equal to the number of spots). For GC, B_Cycling, B_GC_DZ, B_GC_LZ, B_GC_prePB, FDC, and T_CD4_TfH_GC were selected as signature cell types. For T cell zone, T_CD4, T_CD4_TfH, T_TfR, T_Treg, T_CD4_naive, and T_CD8_naive were selected as signature cell types. For B follicle, B_mem, B_naive, and B_preGC were regarded as signature cell types. The merged spot-cell type proportion matrix was calculated by summing up the proportion of signature cell types for GC, T cell zone, and B follicle, respectively. Finally, annotated spots (spot assignment in Supplementary Data 11) were selected from the merged spot-cell type proportion matrix for visualization. The subset spots from the merged matrix were projected on an equilateral triangle via Barycentric coordinate project methods[37]. The projected spots were colored by FTU assignment results. Unique and overlapped spots across seven regions (i.e., GC, GC–B, B, B–T, T, T–GC, and T–GC–B) from three FTUs were assigned and visualized on the spatial map. Gene module scores were calculated using the *AddModuleScore* function from the Seurat (v4.0.5) package. Calculated gene module score and cell type proportion were then grouped by seven regions and visualized on the line plot (Fig. 3e, f). One-way ANOVA using function aov in R environment was conducted to test the difference among the means of seven regions regarding gene module scores and cell type proportions, respectively.

CODEX tonsil tissue staining

An FFPE human tonsil tissue (provided by Dr. Scott Rodig, Brigham and Women's Hospital Department of Pathology) was sectioned onto a No. 1 glass coverslip (22x22mm) pre-treated with Vectabond (SP-1800-7, Vector Labs). The tissue was deparaffinized by heating at 70 °C for 1 h and soaking in xylene 2× for 15 min each. The tissue was then rehydrated by incubating in the following sequence for 3 min each with gentle rocking: 100% EtOH twice, 95% EtOH twice, 80% EtOH once, 70% EtOH once, ddH$_2$O thrice. To prepare for heat-induced antigen retrieval (HIER), a PT module (A80400012, Thermo Fisher) was filled with 1X PBS, with a coverslip jar containing 1X Dako pH 9 antigen retrieval buffer (S2375, Agilent) within. The PT module was then pre-warmed to 75 °C. After rehydration, the tissue was placed in the pre-warmed coverslip jar, then the PT module was heated to 97 °C for 20 min and cooled to 65 °C. The coverslip jar was then removed from the PT module and cooled for ~15–20 min at room temperature. The tissue was then washed in rehydration buffer (232105, Akoya Biosciences) twice for 2 min each then incubated in CODEX staining buffer (232106, Akoya Biosciences) for 20 min while gently rocking. A hydrophobic barrier was then drawn on the perimeter of the coverslip with an ImmEdge Hydrophobic Barrier pen (310018, Vector Labs). The tissue was then transferred to a humidity chamber. The humidity chamber was made by filling an empty pipette tip box with paper towels and ddH$_2$O, stacking the tip box on a cool box (432021, Corning) containing a −20 °C ice block, then replacing the tip box lid with a six-well plate lid. The tissue was then blocked with 200 μL of blocking buffer.

The blocking buffer was made with 180 μL BBDG block, 10 μL oligo block, and 10 μL sheared salmon sperm DNA; the BBDG block was prepared with 5% donkey serum, 0.1% Triton X-100, and 0.05% sodium azide prepared with 1X TBS IHC Wash buffer with Tween 20 (935B-09, Cell Marque); the oligo block was prepared by mixing 57 different custom oligos (IDT) in ddH$_2$O with a final concentration of 0.5 μM per oligo; the sheared salmon sperm DNA was added from its 10 mg/ml stock (AM9680, Thermo Fisher). The tissue was blocked while photobleaching with a custom LED array for 2 h. The LED array was set up by inclining two Happy Lights (6460231, Best Buy) against both sides of the cool box and positioning an LED Grow Light (B07C68N7PC, Amazon) above. The temperature was monitored to ensure that it remained under 35 °C. The staining antibodies were then prepared during the 2-h block.

DNA-conjugated antibodies at appropriate concentrations were added to 100 μL of CODEX staining buffer, loaded into a 50-kDa

centrifugal filter (UFC505096, Millipore) pre-wetted with CODEX staining buffer, and centrifuged at 12,500×g for 8 min. Concentrated antibodies were then transferred to a 0.1 μm centrifugal filter (UFC30VV00, Millipore) pre-wetted with CODEX staining buffer, filled with extra CODEX staining buffer to a total volume of 181 μL, added with 4.75 μL of each Akoya blockers N (232108, Akoya), G (232109, Akoya), J (232110, Akoya), and S (232111, Akoya) to a total volume of 200 μL, then centrifuged for 2 min at 12,500×g to remove antibody aggregates. The antibody flow through (99 μL) was used to stain the tissue overnight at 4 °C in a humidity chamber covered with a foil-wrapped lid.

After the overnight antibody stain, the tissue was washed in CODEX staining buffer twice for 2 min each before fixing in 1.6% paraformaldehyde (PFA) for 10 min while gently rocking. The 1.6% PFA was prepared by diluting 16% PFA in CODEX storage buffer (232107, Akoya). After 1.6% PFA fixation, the tissue was rinsed in 1X PBS twice and washed in 1X PBS for 2 min while gently rocking. The tissue was then incubated in the cold (−20 °C) with 100% methanol on ice for 5 min without rocking for further fixation and then washed thrice in 1X PBS as before while gently rocking. The final fixation solution was then prepared by mixing 20 μL of CODEX final fixative (232112, Akoya) in 1000 μL of 1x PBS. The tissue was then fixed with 200 μL of the final fixative solution at room temperature for 20 min in a humidity chamber. The tissue was then rinsed in 1X PBS and stored in 1X PBS at 4 °C prior to CODEX imaging.

A black flat bottom 96-well plate (07-200-762, Corning) was used to store the reporter oligonucleotides, with each well corresponding to an imaging cycle. Each well contained two fluorescent oligonucleotides (Cy3 and Cy5, 5 μL each) added to 240 μL of plate master mix containing DAPI nuclear stain (1:600) (7000003, Akoya) and CODEX assay reagent (0.5 mg/mL) (7000002, Akoya). For the first and last blank cycles, an additional plate buffer was used to substitute for each fluorescent oligonucleotide. The 96-well plate was securely sealed with aluminum film (14-222-342, Thermo Fisher) and kept at 4 °C prior to CODEX imaging.

CODEX antibody panel

The following antibodies, clones, and suppliers were used in this study:

BCL-2 (124, Novus Biologicals, 1:50), CCR6 (polyclonal, Novus Biologicals, 1:25), CD11b (EPR1344, Abcam, 1:50), CD11c (EP1347Y, Abcam, 1:50), CD15 (MMA, BD Biosciences, 1:200), CD16 (D1N9L, Cell Signaling Technology, 1:100), CD162 (HECA-452, Novus Biologicals, 1:200), CD163 (EDHu-1, Novus Biologicals, 1:200), CD2 (RPA-2.10, Biolegend, 1:25), CD20 (rIGEL/773, Novus Biologicals, 1:200), CD206 (polyclonal, R&D Systems, 1:100), CD25 (4C9, Cell Marque, 1:100), CD30 (BerH2, Cell Marque, 1:25), CD31 (C31.3 + C31.7 + C31.10, Novus Biologicals, 1:200), CD4 (EPR6855, Abcam, 1:100), CD44 (IM-7, Biolegend, 1:100), CD45 (B11 + PD7/26, Novus Biologicals, 1:400), CD45RA (HI100, Biolegend, 1:50), CD45RO (UCH-L1, Biolegend, 1:100), CD5 (UCHT2, Biolegend, 1:50), CD56 (MRQ-42, Cell Marque, 1:50), CD57 (HCD57, Biolegend, 1:200), CD68 (KP-1, Biolegend, 1:100), CD69 (polyclonal, R&D Systems, 1:200), CD7 (MRQ-56, Cell Marque, 1:100), CD8 (C8/144B, Novus Biologicals, 1:50), collagen IV (polyclonal, Abcam, 1:200), cytokeratin (C11, Biolegend, 1:200), EGFR (D38B1, Cell Signaling Technology, 1:25), FoxP3 (236A/E7, Abcam, 1:100), granzyme B (EPR20129-217, Abcam, 1:200), HLA-DR (EPR3692, Abcam, 1:200), IDO-1 (D5J4E, Cell Signaling Technology, 1:25), LAG-3 (D2G4O, Cell Signaling Technology, 1:25), mast cell tryptase (AA1, Abcam, 1:200), MMP-9 (L51/82, Biolegend, 1:200), MUC-1 (955, Novus Biologicals, 1:100), PD-1 (D4W2J, Cell Signaling Technology, 1:50), PD-L1 (E1L3N, Cell Signaling Technology, 1:50), podoplanin (D2-40, Biolegend, 1:200), T-bet (D6N8B, Cell Signaling Technology, 1:100), TCR β (G11, Santa Cruz Biotechnology, 1:100), TCR-γ/δ (H-41, Santa Cruz Biotechnology, 1:100), Tim-3 (polyclonal, Novus Biologicals, 1:50), Vimentin (RV202, BD Biosciences, 1:200), VISTA (D1L2G, Cell Signaling

Technology, 1:50), α-SMA (polyclonal, Abcam, 1:200), and β-catenin (14, BD Biosciences, 1:50). Readers of interest are referred to publication[70] for more details on the antibody clones, conjugated fluorophores, exposure, and titers.

CODEX tonsil tissue imaging

The tonsil tissue coverslip and reporter plate were equilibrated to room temperature and placed on the CODEX microfluidics instrument. All buffer bottles were refilled (ddH$_2$O, DMSO, 1X CODEX buffer (7000001, Akoya)), and the waste bottle was emptied before the run. To facilitate the setting up of imaging areas and z planes, the tissue was stained with 750 μL of nuclear stain solution (1 μL of DAPI nuclear stain in 1500 μL of 1X CODEX buffer) for 3 min, then washed with the CODEX fluidics device. For each imaging cycle, three images that corresponded to the DAPI, Cy3, and Cy5 channels were captured. The first and last blank imaging cycles did not contain any Cy3 or Cy5 oligos, and thus are used for background correction.

The CODEX imaging was operated using a ×20/0.75 objective (CFI Plan Apo λ, Nikon) mounted to an inverted fluorescence microscope (BZ-X810, Keyence) which was connected to a CODEX microfluidics instrument and CODEX driver software (Akoya Biosciences). The acquired multiplexed images were stitched, and background corrected using the SINGER CODEX Processing Software (Akoya Biosciences). For this study, six independent 2048 × 2048 field-of-views (FOV) were cropped from the original 20,744 × 20,592 image. The FOVs were selected to include key cell types and tissue structures in tonsils, such as tonsillar crypts or lymphoid nodules.

Cell segmentation

Custom ImageJ macros were used to normalize and cap nuclear and surface image signals at the 99.7th percentile to facilitate cell segmentation. Cell segmentation was performed using a local implementation of Mesmer from the DeepCell library (deepcell-tf 0.11.0)[40], where the multiplex_segmentation.py script was modified to adjust the segmentation resolution (microns per pixel, mpp). model_mpp = 0.5 generated satisfactory segmentation results for this study. Single-cell features based on the cell segmentation mask were then scaled to cell size and extracted as FCS files.

Cell clustering and annotation

Single-cell features were normalized to each FOV's median DAPI signal to account for FOV signal variation, arcsinh transformed with cofactor = 150, capped between 1st–99th percentile, and rescaled to 0–1. Sixteen markers (cytokeratin, podoplanin, CD31, αSMA, collagen IV, CD11b, CD11c, CD68, CD163, CD206, CD7, CD4, CD8, FoxP3, CD20, CD15) were used for unsupervised clustering using FlowSOM[41] (66 output clusters). The cell type for each cluster was annotated based on its relative feature expression, as determined via Marker Enrichment Modeling[42], and annotated clusters were visually compared to the original images to ensure accuracy and specificity. Cells belonging to indeterminable clusters were further clustered (20 output clusters) and annotated as above.

SpaGFT implementation on tonsil CODEX data and interpretation

Resize CODEX images and SpaGFT implementation. As each FOV consisted of 2048 by 2048 pixels (~0.4 μm per pixel size), the CODEX image needed to be scaled down to 200 by 200 pixels (~3.2 μm per pixel size) to reduce the high computational burden (Supplementary Fig. 8a). Therefore, original CODEX images (2048 by 2048 pixels) were resized to 200 by 200 images by implementing function "resize" and selecting cubic interpolation from the imager package (v.42) in R environments. SpaGFT was then applied to the resized images by following default parameters.

Structural similarity (SSIM) calculation. The Structural Similarity (SSIM) score was a measurement for locally evaluating the similarity between two images regardless of image size[71]. The SSIM score ranged from 0 to 1; a higher score means more similarity between two images. It was defined as follows:

$$\text{SSIM} = l(x,y)^{\alpha} \cdot c(x,y)^{\beta} \cdot s(x,y)^{\gamma}$$

x and y were windows with 8 by 8 pixels; $l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$ was the luminance comparison function for comparing the average brightness of the two images regarding pixels x and y. C_1 is constant, and α is the weight factor of luminance comparison. $c(x,y) = \frac{2\sigma_x\sigma_y + C_1}{\sigma_x^2 + \sigma_y^2 + C_2}$ was the contrast comparison function for measuring the standard deviation of two images. C_2 is constant, and β is the weight factor of contrast comparison. $s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$ was the structure comparison by calculating the covariance between the two images. C_3 is constant, and γ is the weight factor of structure comparison.

Cell–cell distance and interaction analysis. To compute cell–cell distance within one FTU, we first select cells assigned to each FTU. An undirected cell graph was then constructed, where the cell was a node and edge connected by every two cells defined by the Delaunay triangulation using *the deldir* function from the deldir package (v.1.0-6). Subsequently, the edge represented the observed distance between the connected two cells, and Euclidean distance was used for calculating the distance[72]. Lastly, the average distance among different cell types was computed by taking the average of the observed cell–cell distance to generate the network plot. Regarding the determination of the cell–cell interaction, the spatial location of cells assigned in each FTU was permutated and re-calculated cell–cell distance as expected distance. If the cell–cell distance is lower than 15 μm[73] (~5 pixels in the 200 by 200-pixel image), the cells will contact and interact with each other. Wilcoxon rank-sum test was used for the computed p-value for expected distance and observed distance. If the expected distance was significantly smaller than the observed distance, it suggested that cells would interact with each other.

SpaGFT implementation in SpaGCN

Let $\mathbf{X}_{\text{spa}}$ be the SRT gene expression matrix with the dimension $n_{\text{spot}} \times n_{\text{gene}}$, in which n_{spot} and n_{gene} represent the numbers of spots and genes, respectively. Upon normalization, the spot cosine similarity matrix $\mathbf{X}_s$ is computed by the formula $\mathbf{X}_s = \mathbf{X}_{\text{spa}}\mathbf{X}_{\text{spa}}^{\mathsf{T}}$, yielding a matrix with dimension $n_{\text{spot}} \times n_{\text{spot}}$. Denote $\mathbf{U} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_{n_{\text{FC}}})$, where each $\boldsymbol{\mu}_l$ is the lth eigenvector of the Laplacian matrix of the spatial graph and n_{FC} is the number of Fourier coefficients. Hence, graph Fourier transform is implemented to transform $\mathbf{X}_s$ into the frequency domain by:

$$\hat{\mathbf{X}}_s = \mathbf{U}^{\mathsf{T}}\mathbf{X}_s$$

Subsequently, the newly augmented spot-by-feature matrix is obtained by concatenating SRT gene expression matrix $\mathbf{X}_{\text{spa}}$ and transformed signal matrix $\hat{\mathbf{X}}_s$:

$$\mathbf{X}_{\text{new}} = \text{concat}(\mathbf{X}_{\text{spa}}, \hat{\mathbf{X}}_s)$$

Finally, the matrix $\mathbf{X}_{\text{new}}$ is inputted into SpaGCN as a replacement for the original gene expression matrix to predict the spatial domain cluster labels across all spots.

To evaluate the performance of such modification, 12 human dorsolateral prefrontal cortex of 10x Visium datasets were applied in benchmarking based on annotations from the initial study of SpaGCN[4]. The adjusted Rand index (ARI) was selected as the evaluation metric to measure the consistency between the predicted spot clusters and

manually annotated spatial domain. The parameter num_fcs, which controlled the count of FCs, was determined by utilizing a grid search methodology executed on datasets 151508 and 151670. The search spanned a range of values from 600 to 1400, sampled per 100 steps. Upon analysis, the optimal parameter value was established at 1000 (Supplementary Data 15), while the other parameters were set to the default in SpaGCN. Next, the performance was compared on the 10 remaining datasets for the independent test.

SpaGFT implementation in TACCO

SpaGFT was implemented to improve the performance of TACCO, which leveraged optimal transport (OT) to transfer annotation labels from scRNA-seq to spatial transcriptomics data. The core objective function of TACCO is denoted by a cost matrix $\mathbf{C} = (c_{tb})$ and a proportion matrix $\mathbf{\Gamma} = (\gamma_{tb})$:

$$\Phi(\mathbf{\Gamma}) = \sum_{tb} \gamma_{tb} c_{tb}$$

Specifically, c_{tb} quantifies the cost that transports an object b to an annotation t. In TACCO, principal component analysis (PCA) was used to reduce the dimension of scRNA-seq and spatial transcriptomics gene expression matrices to the PC matrices by keeping the first 100 PCs, respectively. Subsequently, $\mathbf{C}$ is computed by calculating the Bhattacharyya coefficients between cell type-averaged scRNA-seq and spatial transcriptomics PC matrices. Finally, the OT's optimization is solved by using the Sinkhorn–Knopp matrix scaling algorithm to yield a 'good' proportion matrix $\mathbf{\Gamma}$.

For finding $\mathbf{\Gamma}$, the cost matrix $\mathbf{C}$ plays the most important role in the OT's optimization process. Based on the originally calculated $\mathbf{C}$, an updated cost matrix $\mathbf{C}^{\mathbf{update}}$ considering spatial topology information is fused. To incorporate this topology information from the spatial data, the coordinates of spatial spots are used to construct a spatial graph, which is as the input with gene expression and initial TACCO-calculated mapping $\mathbf{\Gamma}$, which represent cell-type proportions into SpaGFT for calculating FCs of genes and cell types (CT). Subsequently, these gene FCs matrices were weighted and averaged by spot expression value to obtain the spots' FCs for obtaining spot level constraints. The cosine distance is calculated between the FCs of spatial spots and the FCs of cell types to create the updated CT-spot cost matrix $\mathbf{C}^{\mathbf{update}}$. The $\mathbf{C}'$ is a united cost matrix fused by $\mathbf{C}$ and $\mathbf{C}^{\mathbf{update}}$ with a balancing parameter β as

$$\mathbf{C}' = \beta \mathbf{C} + (1 - \beta) \mathbf{C}^{\mathbf{update}}$$

This updated $\mathbf{C}'$ is then fed back into TACCO's OT algorithm to predict revised cell type proportions for the spatial data. In addition, we used a simulated validation dataset with the setting of bead size = 5 to conduct a grid search on the input parameters S, the sensitivity in the Kneedle algorithm from SpaGFT, and β for determining these hyperparameters. While maintaining computational efficiency, we ascertained that the updated TACCO with $\beta = 0.8$ and $S = 24$ can achieve the best performance. Our experiments reveal that the updated TACCO, enriched with SpaGFT features, outperforms the baseline TACCO model in the simulated independent test dataset with the setting of bead size $\in [10, 20, 30, 40, 50]$.

SpaGFT implementation in Tangram

Denote $\mathbf{X_{sc}}$ as the gene expression matrix of scRNA-seq with the dimension $n_{cell} \times n_{gene}$, in which n_{cell} and n_{gene} represent the numbers of cells and genes, respectively. $\mathbf{X_{spa}}$ is the SRT gene expression matrix with dimension $n_{spot} \times n_{gene}$, and n_{spot} represents the number of spots. Tangram aims to find a mapping matrix $\mathbf{M} = \left(m_{ij} \right)_{n_{cell} \times n_{spot}}$, where

$0 \le m_{ij} \le 1$, $\sum_{i}^{n_{spot}} m_{ij} = 1$ and m_{ij} reflects the probability of cell i mapping to spot j. Hence, $\mathbf{M}^{\mathrm{T}} \mathbf{X_{sc}}$ can be treated as the reconstructed SRT gene expression matrix using scRNA-seq. Let $\mathbf{X_{re}} = \mathbf{M}^{\mathrm{T}} \mathbf{X_{sc}}$. The regularization part of the original objective function of Tangram is as follows:

$$\Phi(\mathbf{M}) = w_1 \sum_{k=1}^{n_{gene}} \mathrm{cosine}(\mathbf{X_{re}^{\cdot,k}}, \mathbf{X_{spa}^{\cdot,k}}) + w_2 \sum_{j=1}^{n_{spot}} \mathrm{cosine}(\mathbf{X_{re}^{j,\cdot}}, \mathbf{X_{spa}^{j,\cdot}})$$

where the first term describes the cosine similarity of gene k across all spots in reconstructed SRT gene expression matrix and real SRT gene expression matrix, weighted by w_1; and the second term describes the cosine similarity of spot j across all genes in reconstructed SRT gene expression matrix and real SRT gene expression matrix, weighted by w_2. By maximizing the objective function, the optimal mapping matrix $\mathbf{M}^*$ can be obtained.

Denote $\mathbf{U} = \left(\mathbf{\mu_1}, \mathbf{\mu_2}, \ldots, \mathbf{\mu_{n_{FC}}} \right)$, where each $\mathbf{\mu_l}$ is the lth eigenvector of the Laplacian matrix of the spatial graph and n_{FC} is the number of Fourier coefficients. Hence, we can implement graph Fourier transform for genes by

$$\hat{\mathbf{X}}_{\mathbf{spa}} = \mathbf{U}^{\mathrm{T}} \mathbf{X_{spa}}$$

$$\hat{\mathbf{X}}_{\mathbf{re}} = \mathbf{U}^{\mathrm{T}} \mathbf{X_{re}}$$

Therefore, both $\hat{\mathbf{X}}_{\mathbf{spa}}$ and $\hat{\mathbf{X}}_{\mathbf{re}}$ are the representations of genes in the frequency domain with the dimension $n_{FC} \times n_{gene}$. In addition, $\mathbf{X}'_{\mathbf{spa}} = \mathbf{X_{spa}} \mathbf{X_{spa}^{\mathrm{T}}}$ can be considered as the spot similarity matrix calculated by gene expression from real SRT data with dimension is $n_{spot} \times n_{spot}$. Similarly, $\mathbf{X}'_{\mathbf{re}} = (\mathbf{M}^{\mathrm{T}} \mathbf{X_{sc}})(\mathbf{M}^{\mathrm{T}} \mathbf{X_{sc}})^{\mathrm{T}}$ represents the spot similarity matrix calculated by gene expression in reconstructed SRT data. In this way, we can implement graph Fourier transform for spots by:

$$\widetilde{\mathbf{X}}_{\mathbf{spa}} = \mathbf{U}^{\mathrm{T}} \mathbf{X}'_{\mathbf{spa}}$$

$$\widetilde{\mathbf{X}}_{\mathbf{re}} = \mathbf{U}^{\mathrm{T}} \mathbf{X}'_{\mathbf{re}}$$

Therefore, both $\widetilde{\mathbf{X}}_{\mathbf{spa}}$ and $\widetilde{\mathbf{X}}_{\mathbf{re}}$ are the new representations of spots in the frequency domain with the dimension $n_{FC} \times n_{spot}$. Therefore, we improved the objective function of Tangram by adding the similarity measurements of genes and spots in the frequency domain. The new objective function is

$$\Phi(\mathbf{M}) = w_1 \sum_{k=1}^{n_{gene}} \mathrm{cosine}(\mathbf{X_{re}^{\cdot,k}}, \mathbf{X_{spa}^{\cdot,k}}) + w_2 \sum_{j=1}^{n_{spot}} \mathrm{cosine}(\mathbf{X_{re}^{j,\cdot}}, \mathbf{X_{spa}^{j,\cdot}})$$
$$+ w_3 \sum_{k=1}^{n_{gene}} \mathrm{cosine}(\hat{\mathbf{X}}_{\mathbf{re}}^{\cdot,k}, \hat{\mathbf{X}}_{\mathbf{spa}}^{\cdot,k}) + w_4 \sum_{j=1}^{n_{spot}} \mathrm{cosine}(\widetilde{\mathbf{X}}_{\mathbf{re}}^{\mathbf{j},\cdot}, \widetilde{\mathbf{X}}_{\mathbf{spa}}^{\mathbf{j},\cdot})$$

where w_1 weights similarities of genes in the vertex domain; w_2 weights similarities of spots in the vertex domain; w_3 weights the similarities of genes in the frequency domain and w_4 weights the similarities of spots in the frequency domain.

To evaluate the performance of such modification. We adopted the evaluation scheme from Bin Li et al. study. In addition, we simulated this SRT dataset by 'gridding' a dataset (STARmap) using various window sizes (400, 450, ..., 1200). In addition, simulated datasets of window sizes 400 and 1200 were used for grid search to determine the hyperparameters. In this way, w_3 and w_4 were set to 11 and 1, respectively, and other parameters (including w_1 and w_2) were the default parameters of Tangram. Our experiments reveal that the updated

Tangram, enriched with SpaGFT features, outperforms the baseline Tangram model.

SpaGFT implementation in CAMPA

Overall, the CAMPA framework, a conditional variational autoencoder for identifying conserved subcellular organelle on pixel-level iterative indirect immunofluorescence, was modified by adding an entropy term on its loss function to regularize graph signal (e.g., protein intensity) spreading or concentration. Specifically, compared with the baseline CAMPA loss function, which computed the mean squared error (MSE) loss for each pixel, the modified loss function additionally considered protein global spreading at the cell level.

Data preparation for model training, testing, and validation. Following the baseline CAMPA paper and guidelines[16], 292,548 (0.05% of full data) pixels datasets were down-sampled from processed cell nuclei of I09 (normal), I10 (Triptolide treatment), I11 (normal), and I16 (TSA treatment) wells based on 184A1 cell line. The training, testing, and validation data were set to 70%, 10%, and 20%, respectively.

Entropy regularization. For cell $i \in I$, where I was the complete set of all cells in the down-sampled data, the corresponding original protein signatures in each cell were denoted as $\mathbf{X}^i$ with the dimension $n_{pixel} \times n_{channel}$, where n_{pixel} and $n_{channel}$ represented the number of pixels in one cell and the number of proteins, respectively. Similarly, $\hat{\mathbf{X}}^i$ was denoted as reconstructed protein signatures for cell i. To measure the spreading of reconstructed protein signatures in the frequency domain, $\hat{\mathbf{X}}^i$ and the coordinates of pixels were input into SpaGFT for computing the FC $\hat{\mathbf{F}}^i$ with the dimension, in which n_{FC} was the number of FC. Denote $\mathbf{U} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_{n_{FC}})$, where each $\boldsymbol{\mu}_k$ was the kth eigenvector of the Laplacian matrix of the spatial neighboring graph for cell i. Hence, FCs of reconstructed protein signatures for cell i was calculated by

$$\hat{\mathbf{F}}^i = \mathbf{U}^T \hat{\mathbf{X}}^i$$

Subsequently, $\hat{\mathbf{F}}^i = (\hat{\mathbf{f}}_1^i, \hat{\mathbf{f}}_2^i, \hat{\mathbf{f}}_{n_{FC}}^i)$ was used to calculate entropy by the entropy function, which regularized a concentrated graph signal[19,74]

$$L_{\text{Entropy}} = -\sum_{i \in I} \sum_{k=1}^{n_{FC}} \frac{|\hat{\mathbf{f}}_k^i|^2}{\|\hat{\mathbf{F}}^i\|_2} \log \frac{|\hat{\mathbf{f}}_k^i|^2}{\|\hat{\mathbf{F}}^i\|_2}$$

where $\|\cdot\|_2$ presents L^2-norm.

In addition, the η parameter was used as a weighting term to balance the initial loss function and the entropy-decreasing loss function, assigned with 0.3 as default. The formula of the modified loss function L_{modified} was as follows:

$$L_{\text{modified}} = \eta \sum_i^I D \ln \sigma + \frac{D}{2\sigma^2} \text{MSE}(\hat{\mathbf{X}}^i, \mathbf{X}^i) + (1-\eta) L_{\text{Entropy}}$$

where D is a constant, which was used the same as the baseline mode ($D = 0.5$). The initial decoder loss function was a part of the objective function in CAMPA, which used an analytical solution from σ-VAE[75] to learn the variance of the decoder. The MSE and the logarithm of the variance were minimized through σ, which was a weighting parameter between the MSE reconstruction term and the KL-divergence term in the CAMPA objective function. There was an analytic solution to compute the value of σ:

$$\sigma^{*2} = \text{MSE}(\mathbf{X}^i, \mathbf{\nu}^i)$$

σ^{*2} was estimated value for σ^2 and $\mathbf{\nu}^i$ presented the estimated latent mean for $\mathbf{X}^i$.

Regarding the implementation, the training and testing datasets were selected to build the modified and baseline models, respectively. Subsequently, to fairly compare the two models' training efficiency, the same validation dataset and initial loss were implemented to evaluate the convergence of validation loss.

To interpret the modified CAMPA training efficiency improvement regarding biological perspective, batch effect removal and prediction accuracy were evaluated. Regarding batch effect removal, a proportion of 1% of pixels were subsampled from prepared data. First, UMAP embeddings calculated from the CAMPA latent representations were generated to visualize the mixture of three perturbation conditions. To quantitatively compare the batch effect removal between the baseline and modified model, the kBET[57] score was computed using the CAMPA latent representations across perturbation conditions. Following the kBET suggestion, 0.5% pixels (~1500 pixels) were iteratively selected for calculating the kBET score (a higher rejection rate suggested a better batch effect removal result) 10000 times using 1–100 neighbors.

Subsequently, the CAMPA latent representations were clustered utilizing the Leiden algorithm[16] at resolutions of 0.2, 0.4, 0.6, 0.8, 1.2, 1.6, and 2.0. To understand the identity of each cluster predicted by the modified CAMPA under the resolution of 0.2, the protein intensities in each pixel cluster were visualized in the heatmap. Each pixel's channel values were averaged at the cluster level and scaled by channel (column-level) z-score. Clusters were annotated based on the highest expressed markers and human protein atlas.

To evaluate the conserveness and homogeneity of the predicted cluster across different resolutions, we implemented high-label entropy to quantify the trend of diverging from one cluster into two clusters[76]. For example, at the resolution of 0.2, all pixels of cluster 6 predicted by the modified model were used to calculate entropy via a probability vector with two lengths. The first element of this vector was a percentage of pixels at the current resolution (i.e., 0.2), which tended to be the largest cluster at the next resolution (e.g., 0.4). The second element of this vector was the percentage of the rest of the pixels at the current resolution, which tended to be other clusters at the next resolution. The high-label entropy was repeatedly calculated on the same pixels of one cluster within/across baseline and modified model across gradient resolutions (i.e., 0.2, 0.4, 0.6, 0.8, 1.2, 1.6, and 2.0). To visualize intact cells and summarize the relation between pixel and cell in Supplementary Data 19, seven clusters predicted by the modified model based on resolution 0.2 were transferred to all pixels from full-size data via function *project_cluster_data* in the CAMPA package. The illustrated examples (id: 367420 and 224081) were extracted to calculate the FC of COIL and SETD1A and visualize.

Reporting summary

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

Data availability

All datasets from 10x Visium can be accessed from https://www.10xgenomics.com/products/spatial-gene-expression. Slide-DNA-seq data is available as accession code SCP1278 in the Single Cell Portal. Slide-TCR-seq data is available as accession code SCP1348 in the Single Cell Portal. The GSM5519054_Visium_MouseBrain data can be accessed via the GEO database with an accession code GSM5519054. Regarding the human brain dataset, twelve samples can be accessed via endpoint "jhpce#HumanPilot10x" on Globus data transfer platform at http://research.libd.org/globus/. The other six human brain datasets (2-3-AD_Visium_HumanBrain, 2-8-AD_Visium_HumanBrain, T4857-AD_Visium_HumanBrain, 2-5_Visium_HumanBrain, 18-64_Visium_HumanBrain, and 1-1_Visium_HumanBrain) can be accessed via the GEO database with an accession code GSE220442 and https://bmbls.bmi.osumc.edu/scread/stofad-2. The two Slide-seqV2 datasets are available

as accession code SCP815 in the Single Cell Portal. MERFISH data (Slice1_Replicate1-Vizgen_MouseBrainReceptor) can be accessed from https://console.cloud.google.com/marketplace/product/gcp-public-data-vizgen/vizgen-mouse-brain-map?pli=1&project=vizgen-gcp-share. Xenium data (Rep1-Cancer_Xenium_HumanBreast) is downloaded from https://www.10xgenomics.com/products/xenium-in-situ/human-breast-dataset-explorer. Spatial-CITE-seq data can be accessed via the GEO database with an accession number of GSE213264. Spatial epigenome−transcriptome co-profiling data (spatial_ATAC_RNA_MouseE13) can be accessed via the GEO database with an accession code GSE205055. The 184A1 datasets used to train modified CAMPA reported in this manuscript can be found at https://doi.org/10.5281/zenodo.7299516. SPOT data can be accessed via the GEO database with an accession number of GSE198353. The CODEX tonsil data generated in this study have been deposited in the Zenodo database under accession code 10433896. Source data are provided in this paper. Source data are provided with this paper.

Code availability
SpaGFT is a Python package for modeling and analyzing spatial transcriptomics data. The SpaGFT source code and the analysis scripts for generating results and figures in this paper are available at https://github.com/OSU-BMBL/SpaGFT. The source code is also available on Zenodo[77] with link https://doi.org/10.5281/zenodo.12595086.

References
1. Liu, S. et al. Spatial maps of T cell receptors and transcriptomes reveal distinct immune niches and interactions in the adaptive immune response. *Immunity* **55**, 1940–1952.e1945 (2022).
2. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. Methods and applications for single-cell and spatial multi-omics. *Nat. Rev. Genet.* **24**, 494–515 (2023).
3. Jain, S. et al. Advances and prospects for the Human BioMolecular Atlas Program (HuBMAP). *Nat. Cell Biol.* **25**, 1089–1100 (2023).
4. Hu, J. et al. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).
5. Schürch, C. M. et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* **182**, 1341–1359.e1319 (2020).
6. Mages, S. et al. TACCO unifies annotation transfer and decomposition of cell identities for single-cell and spatial omics. *Nat. Biotechnol.* **41**, 1465–1473 (2023).
7. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* **18**, 1352–1362 (2021).
8. Wang, Y. et al. Sprod for de-noising spatially resolved transcriptomics data based on position and image information. *Nat. Methods* **19**, 950–958 (2022).
9. Sun, S., Zhu, J. & Zhou, X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* **17**, 193–200 (2020).
10. Liu, Y. et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* **183**, 1665–1681.e1618 (2020).
11. Svensson, V., Teichmann, S. A. & Stegle, O. SpatialDE: identification of spatially variable genes. *Nat. Methods* **15**, 343–346 (2018).
12. Zhu, J., Sun, S. & Zhou, X. SPARK-X: non-parametric modeling enables scalable and robust detection of spatial expression patterns for large spatial transcriptomic studies. *Genome Biol.* **22**, 184 (2021).
13. Velten, B. & Stegle, O. Principles and challenges of modeling temporal and spatial omics data. *Nat. Methods* **20**, 1462–1474 (2023).
14. Lake, B. B. et al. An atlas of healthy and injured cell states and niches in the human kidney. *Nature* **619**, 585–594 (2023).
15. Chen, F., Wang, Y.-C., Wang, B. & Kuo, C.-C. J. Graph representation learning: a survey. *APSIPA Trans. Signal Inf. Process.* **9**, e15 (2020).
16. Spitzer, H., Berry, S., Donoghoe, M., Pelkmans, L. & Theis, F. J. Learning consistent subcellular landmarks to quantify changes in multiplexed protein maps. *Nat. Methods* **20**, 1058–1069 (2023).
17. Gut, G., Herrmann, M. D. & Pelkmans, L. Multiplexed protein maps link subcellular organization to cellular states. *Science* **361**, eaar7042 (2018).
18. Ricaud, B., Borgnat, P., Tremblay, N., Gonçalves, P. & Vandergheynst, P. Fourier could be a data scientist: from graph Fourier transform to signal processing on graphs. *C. R. Phys.* **20**, 474–488 (2019).
19. Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. & Vandergheynst, P. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **30**, 83–98 (2013).
20. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* **40**, 308–318 (2022).
21. Buzzi, R. M. et al. Spatial transcriptome analysis defines heme as a hemopexin-targetable inflammatoxin in the brain. *Free Radic. Biol. Med.* **179**, 277–287 (2022).
22. Stickels, R. R. et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).
23. Chen, S. et al. Spatially resolved transcriptomics reveals genes associated with the vulnerability of middle temporal gyrus in Alzheimer's disease. *Acta Neuropathol. Commun.* **10**, 188 (2022).
24. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).
25. Ortiz, C. et al. Molecular atlas of the adult mouse brain. *Sci. Adv.* **6**, eabb3446 (2020).
26. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
27. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
28. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
29. Science, A.I.f.B. Allen Institute for Brain Science (2004). *Allen Mouse Brain Atlas [dataset]* (Allen Institute for Brain Science, 2011).
30. Miller, B. F., Bambah-Mukku, D., Dulac, C., Zhuang, X. & Fan, J. Characterizing spatial gene expression heterogeneity in spatially resolved single-cell transcriptomics data with nonuniform cellular densities. *Genome Res.* gr. 271288.271120 (2021).
31. Zhang, K., Feng, W. & Wang, P. Identification of spatially variable genes with graph cuts. *Nat. Commun.* **13**, 5488 (2022).
32. Ortega, A., Frossard, P., Kovačević, J., Moura, J. M. F. & Vandergheynst, P. Graph signal processing: overview, challenges, and applications. *Proc. IEEE* **106**, 808–828 (2018).
33. Hou, W., Ji, Z., Ji, H. & Hicks, S. C. A systematic evaluation of single-cell RNA-sequencing imputation methods. *Genome Biol.* **21**, 218 (2020).
34. Elyanow, R., Dumitrascu, B., Engelhardt, B. E. & Raphael, B. J. netNMF-sc: leveraging gene-gene interactions for imputation and dimensionality reduction in single-cell expression analysis. *Genome Res.* **30**, 195–204 (2020).
35. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
36. Li, B. et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nat. Methods* **19**, 662–670 (2022).
37. Bhate, S. S., Barlow, G. L., Schürch, C. M. & Nolan, G. P. Tissue schematics map the specialization of immune tissue motifs and their appropriation by tumors. *Cell Syst.* **13**, 109–130.e106 (2022).

38. Kerfoot, S. M. et al. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity* **34**, 947–960 (2011).

39. Natkunam, Y. The biology of the germinal center. *Hematology* **2007**, 210–215 (2007).

40. Greenwald, N. F. et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* **40**, 555–565 (2022).

41. Van Gassen, S. et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry Part A* **87**, 636–645 (2015).

42. Diggins, K. E., Greenplate, A. R., Leelatian, N., Wogsland, C. E. & Irish, J. M. Characterizing cell subsets using marker enrichment modeling. *Nat. Methods* **14**, 275–278 (2017).

43. Liu, C.C. et al. Robust phenotyping of highly multiplexed tissue imaging data using pixel-level clustering. *Nat Commun.* **14**, 4618 (2023).

44. Pavlasova, G. & Mraz, M. The regulation and function of CD20: an "enigma" of B-cell biology and targeted therapy. *Haematologica* **105**, 1494–1506 (2020).

45. Meda, B. A. et al. BCL-2 Is consistently expressed in hyperplastic marginal zones of the spleen, abdominal lymph nodes, and ileal lymphoid tissue. *Am. J. Surg. Pathol.* **27**, 888–894 (2003).

46. Hockenbery, D. M., Zutter, M., Hickey, W., Nahm, M. & Korsmeyer, S. J. BCL2 protein is topographically restricted in tissues characterized by apoptotic cell death. *Proc. Natl Acad. Sci. USA* **88**, 6961–6965 (1991).

47. Heit, A. et al. Vaccination establishes clonal relatives of germinal center T cells in the blood of humans. *J. Exp. Med.* **214**, 2139–2152 (2017).

48. Chtanova, T. et al. T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells1. *J. Immunol.* **173**, 68–78 (2004).

49. Dorfman, D. M., Brown, J. A., Shahsafaei, A. & Freeman, G. J. Programmed death-1 (PD-1) is a marker of germinal center-associated T cells and angioimmunoblastic T-cell lymphoma. *Am. J. Surg. Pathol.* **30**, 802–810 (2006).

50. Marsee, D. K., Pinkus, G. S. & Hornick, J. L. Podoplanin (D2-40) is a highly effective marker of follicular dendritic cells. *Appl. Immunohistochem. Mol. Morphol.* **17**, 102–107 (2009).

51. Gray, E. E. & Cyster, J. G. Lymph node macrophages. *J. Innate Immun.* **4**, 424–436 (2012).

52. Johansson-Lindbom, B., Ingvarsson, S. & Borrebaeck, C. A. Germinal centers regulate human Th2 development. *J. Immunol.* **171**, 1657–1666 (2003).

53. Nakagawa, R. & Calado, D. P. Positive selection in the light zone of germinal centers. *Front. Immunol.* **12**, 661678 (2021).

54. Allen, C. D. C. et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat. Immunol.* **5**, 943–952 (2004).

55. Allen, C. D., Okada, T. & Cyster, J. G. Germinal-center organization and cellular dynamics. *Immunity* **27**, 190–202 (2007).

56. Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018).

57. Büttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A. & Theis, F. J. A test metric for assessing single-cell RNA-seq batch correction. *Nat. methods* **16**, 43–49 (2019).

58. Morris, G. E. The Cajal body. *Biochim. Biophys. Acta (BBA) - Mol. Cell Res.* **1783**, 2108–2115 (2008).

59. Tajima, K. et al. SETD1A protects from senescence through regulation of the mitotic gene expression program. *Nat. Commun.* **10**, 2854 (2019).

60. Wang, J. et al. Dimension-agnostic and granularity-based spatially variable gene identification using BSP. *Nat. Commun.* **14**, 7367 (2023).

61. Li, Z. et al. Benchmarking computational methods to identify spatially variable genes and peaks. Preprint at *bioRxiv* https://doi.org/10.1101/2023.12.02.569717 (2023).

62. Tang, X. et al. Explainable multi-task learning for multi-modality biological data analysis. *Nat. Commun.* **14**, 2546 (2023).

63. Bao, F. et al. Integrative spatial analysis of cell morphologies and transcriptional states with MUSE. *Nat. Biotechnol.* **40**, 1200–1209 (2022).

64. Chang, Y. et al. Define and visualize pathological architectures of human tissues from spatially resolved transcriptomics using deep learning. *Comput Struct Biotechnol J.* **20**, 4600–4617 (2022).

65. Huang, W. et al. Graph frequency analysis of brain signals. *IEEE J. Sel. Top. Signal Process.* **10**, 1189–1203 (2016).

66. Lu, K.-S. & Ortega, A. Fast graph Fourier transforms based on graph symmetry and bipartition. *IEEE Trans. Signal Process.* **67**, 4855–4869 (2019).

67. Magoarou, L. L., Gribonval, R. & Tremblay, N. Approximate fast graph Fourier transforms via multilayer sparse approximations. *IEEE Trans. Signal Inf. Process. Netw.* **4**, 407–420 (2018).

68. Satopaa, V., Albrecht, J., Irwin, D. & Raghavan, B. Finding a 'kneedle' in a haystack: detecting knee points in system behavior. In *Proc. 2011 31st International Conference on Distributed Computing Systems Workshops* (ed. Du) 166–171 (IEEE, 2011).

69. Lu, J. et al. Metagenome analysis using the Kraken software suite. *Nat. Protoc.* **17**, 2815–2839 (2022).

70. Phillips, D. et al. Highly multiplexed phenotyping of immunoregulatory proteins in the tumor microenvironment by CODEX tissue imaging. *Front. Immunol.* **12**, 687673 (2021).

71. Zhou, W., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

72. Jiang, S. et al. Combined protein and nucleic acid imaging reveals virus-dependent B cell and macrophage immunosuppression of tissue microenvironments. *Immunity* **55**, 1118–1134.e1118 (2022).

73. Fang, R. et al. Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. *Science* **377**, 56–62 (2022).

74. Ricaud, B. & Torrésani, B. A survey of uncertainty principles and some signal processing applications. *Adv. Comput. Math.* **40**, 629–650 (2014).

75. Rybkin, O., Daniilidis, K. & Levine, S. Simple and effective VAE training with calibrated decoders. In *Proc. International Conference on Machine Learning* (ed. Meila) 9179–9189 (PMLR, 2021).

76. Sikkema, L. et al. An integrated cell atlas of the lung in health and disease. *Nat. Med.* **29**, 1563–1577 (2023).

77. Liu, J. et al. *SpaGFT Github: OSU-BMBL/SpaGFT: 0.1.1.* Zenodo (2024).

Acknowledgements

Author contributions

Conceptualization: Q.M.; methodology: J.L., Y.C., B.L., Q.M.; software coding: J.L., Y.J., and Y.C.; data collection and investigation: Y.C., Q.G., and M. M.; experiment and interpretation: Z.L., D.X., Y.Y.Y., S.J., S.R., G.N., and D.B.; data analysis and visualization: Y.C., Y.J. and J.L.; case study design and interpretation: Y.C., J.L., S.J., J.E.K. and A.M.; software testing and tutorial: J.L., Y.J., and Y.C.; graphic demonstration: Y.C., Y.J., and A.M.; manuscript writing, review, and editing: all the authors.

Competing interests

S.J. is a co-founder of Elucidate Bio Inc., has received speaking honorariums from Cell Signaling Technology, and has received research support from Roche unrelated to this work. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at
https://doi.org/10.1038/s41467-024-51590-5.

Correspondence and requests for materials should be addressed to Bingqiang Liu or Qin Ma.

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

Reprints and permissions information is available at
http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.