

## ARTICLE



# Spatial multi-omics profiling of breast cancer oligo-recurrent lung metastasis

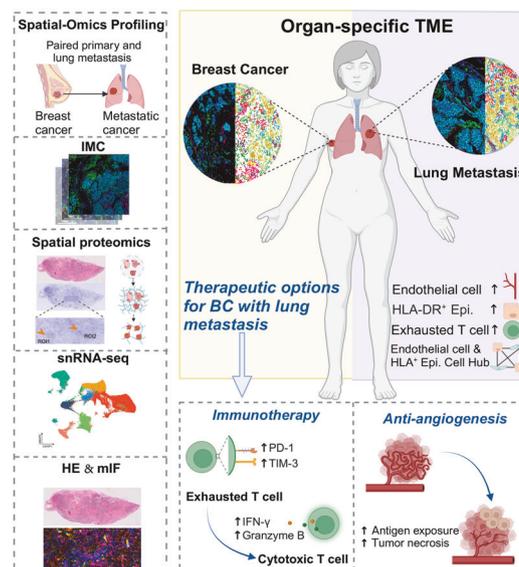
Yang Gao<sup>1,12</sup>, Bin Li<sup>1,12</sup>, Yuzhi Jin<sup>1,12</sup>, Jinlin Cheng<sup>2,12</sup>, Weihong Tian<sup>3</sup>, Lixiong Ying<sup>4</sup>, Libing Hong<sup>1</sup>, Shan Xin<sup>5</sup>, Bo Lin<sup>6</sup>, Chuan Liu<sup>1</sup>, Xuqi Sun<sup>1</sup>, Jun Zhang<sup>7</sup>, Haibo Zhang<sup>8</sup>, Jindong Xie<sup>9</sup>, Xinpei Deng<sup>9</sup>, Xiaomeng Dai<sup>1,10</sup>, Lulu Liu<sup>1</sup>, Yi Zheng<sup>1</sup>, Peng Zhao<sup>1,10</sup>, Guangchuan Yu<sup>11</sup>, Weijia Fang<sup>1,10</sup> and Xuanwen Bao<sup>1,10</sup>

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Primary breast cancer (BC) and metastatic tumors exhibit distinct tumor microenvironment (TME) ecosystems, and the heterogeneity of the TME of BC poses challenges to effective therapies. Evaluating the TME at the single-cell and spatial profiles offers potential for more precise treatments. However, due to the challenge of obtaining surgical specimens of both primary BC and oligo-recurrent lung metastasis simultaneously for high-resolution spatial analysis, the TME of lung-specific metastases using paired samples remains largely unexplored. In this study, we developed a comprehensive strategy using imaging mass cytometry (IMC), spatial proteomics, single-nucleus RNA-seq (snRNA-seq) and multiplex immunofluorescence to explore the spatial topology of lung-specific metastasis and the underlying biological mechanisms based on formalin-fixed paraffin-embedded (FFPE) samples from BC and paired lung metastasis. A total of 250,600 high-quality cells with spatial information revealed by IMC depicted the spatial differences in the TME between BC and lung metastasis. A significant increase in HLA-DR<sup>+</sup> epithelial cells, endothelial cells and exhausted T cells was detected in lung metastases compared to primary sites, with this difference accentuated in the triple-negative subtype. Moreover, a distinct cellular hub comprising endothelial cells and HLA-DR<sup>+</sup> epithelial cells implies the potential promising effect of anti-angiogenic therapy and immunotherapy in BC with lung metastasis, which was further validated by multiplex immunofluorescence analysis. Spatial proteomics further explored the underlying mechanism of TME components identified by IMC analysis. snRNA-seq validated the enrichment of endothelial cells in lung metastasis than that in BC at a whole FFPE slide level. In conclusion, this study determines the spatial multi-omics profiling of TME components at a single-cell resolution using paired samples of primary BC and lung oligo-metastasis. The comprehensive analysis may contribute to the development of therapeutic options.

*Oncogene*; <https://doi.org/10.1038/s41388-025-03388-y>

## Graphical Abstract



A full list of author affiliations appears at the end of the paper.

Received: 26 October 2024 Revised: 23 March 2025 Accepted: 31 March 2025

Published online: 15 April 2025

## INTRODUCTION

Breast cancer (BC) stands as one of the most prevalent malignancies globally, ranking second in incidence and fourth in cancer-related mortality [1]. Notably, 20–30% of metastatic breast cancer (mBC) cases are diagnosed with metastasis at their initial presentation, with over 90% of BC-related deaths attributed to complications from metastasis [2, 3]. mBC presents as a heterogeneous disease with varied prognoses, influenced by diverse clinicopathological characteristics, including pathological subtypes [4]. Clinically, BC is stratified into three subtypes: luminal (ER+, PR+/-), HER2 overexpression (HER2+, ER+/-, PR+/-), and triple-negative breast cancer (TNBC; ER-, PR-, HER2-), each exhibiting distinct relapse times and organ tropisms post-diagnosis [2]. Treatment strategies for mBC share commonalities yet diverge based on subtype and metastatic site. Recent advancements in understanding BC's molecular, genetic, and immune landscapes have led to more targeted and promising immunotherapies and targeted therapies [5]. However, the heterogeneity of the tumor microenvironment (TME) poses challenges to effective BC treatment. Interactions between tumor cells and the other cellular components within the TME significantly influence cancer therapeutic responses [6–9]. Precise techniques have significantly enhanced our understanding of TME, offering great potential for more effective treatments for metastatic BC. Nowadays, single-cell RNA sequencing (scRNA-seq) has been utilized to explore the heterogeneity of cellular composition and molecular characteristics in primary BC and partial metastatic lesions [10–14]. Spatially-resolved technologies have further provided high-precision mapping, enabling detailed assessments of tumor cell localization, immune cell infiltration, and stromal interactions within the TME [10, 12–17]. For instance, novel phenotypes of cancer-associated fibroblasts (CAFs) and myeloid cells have been explored as predominant immune modulators in antitumor immunity [14, 17]. Integrated cellular models for BC subgroup classification, as well as therapy-related prediction models, were also developed using these high-throughput techniques [10, 16]. Despite the lungs being a common BC metastatic site, solitary lung metastasis is rare [3, 4, 18]. Surgical treatment of primary BC and paired lung metastases is infrequent, hindering the acquisition of matched samples for molecular characterization. Moreover, most techniques require fresh tissue samples, while surgical specimens are typically formalin-fixed paraffin-embedded (FFPE), further complicating analysis. Recent advancements have enabled the application of precise techniques to analyze the complexity of the TME based on FFPE samples. For example, Patho-DBiT technology has substantially improved RNA abundance detection in FFPE samples, enabling high-yield, spatially genome-wide profiling of various RNA species [19]. This advancement enables more detailed analyses of RNA processing, microRNA-mRNA regulation, and splicing events. Additionally, the PLATO platform integrating microfluidics and transfer learning technologies, which achieved high-resolution mapping of thousands of proteins across entire tissue sections using FFPE tissue [20]. These emerging technologies will provide invaluable insights into the complex functional features of the TME in BC lung metastasis, as well as the cellular interactions within this environment. Considering the enriched immune cell populations in BC lung metastasis, the application of these spatial techniques—either through comprehensive RNA profiling enabled by Patho-DBiT or high-resolution protein mapping via PLATO—could significantly enhance our understanding of cellular interactions and spatial topology within the metastatic TME. Nevertheless, there are limited studies on single-cell and spatially resolved profiling of BC and paired lung metastases till now.

Here, we evaluate the spatial topology of TME in BC and paired lung metastases at a single-cell resolution. We employed FFPE samples for imaging mass cytometry (IMC), spatial proteomics,

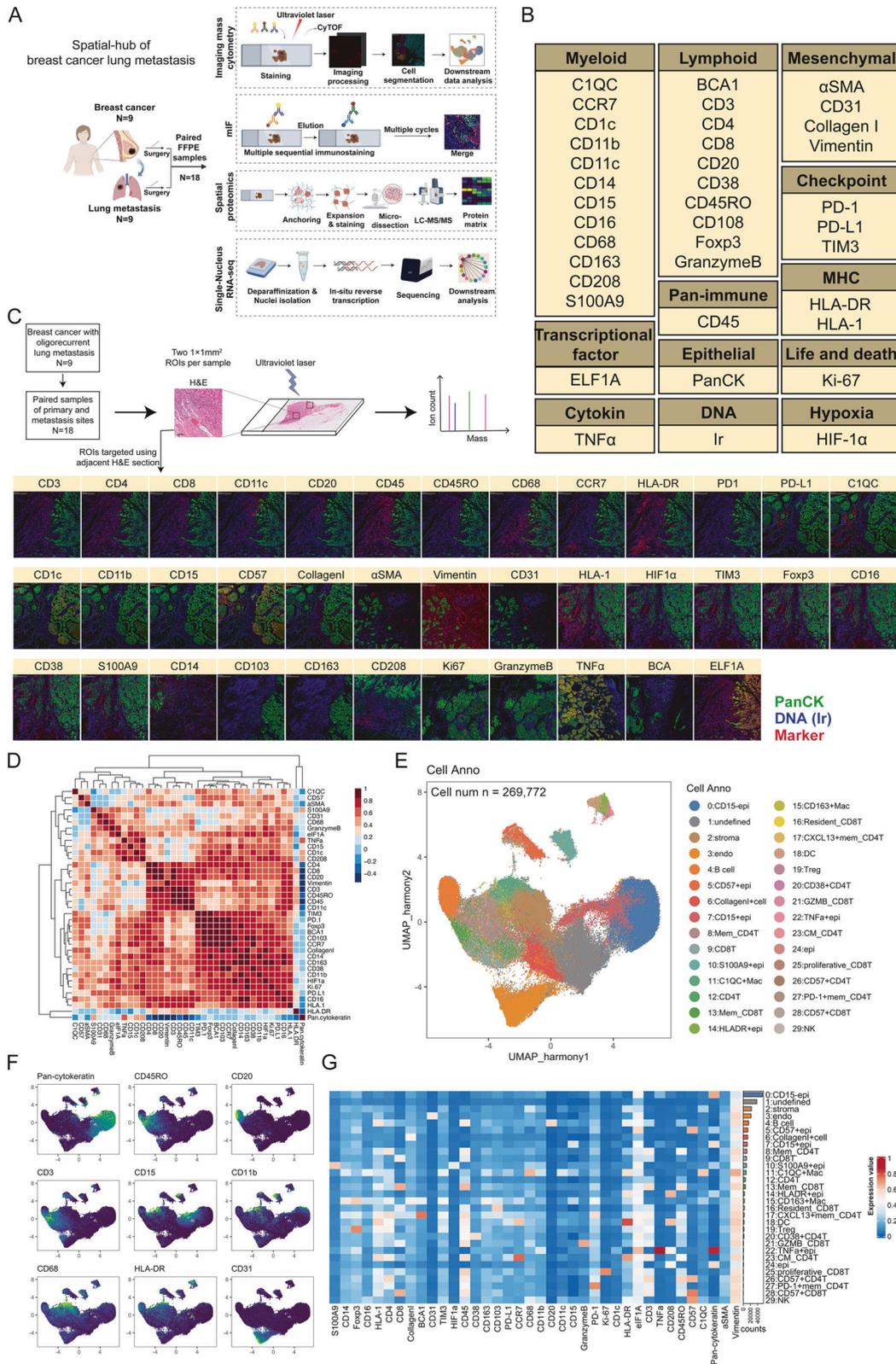
single-nucleus RNA-seq (snRNA-seq) and multiplex immunofluorescence (mIF) to elucidate BC and paired oligo-recurrent lung metastasis characteristics. A lung-specific endothelial cell and HLA-DR<sup>+</sup> epithelial cell hub was identified, offering potential targets for anti-angiogenic combined immunotherapy, thus presenting a novel therapeutic avenue for BC patients with lung metastasis.

## RESULTS

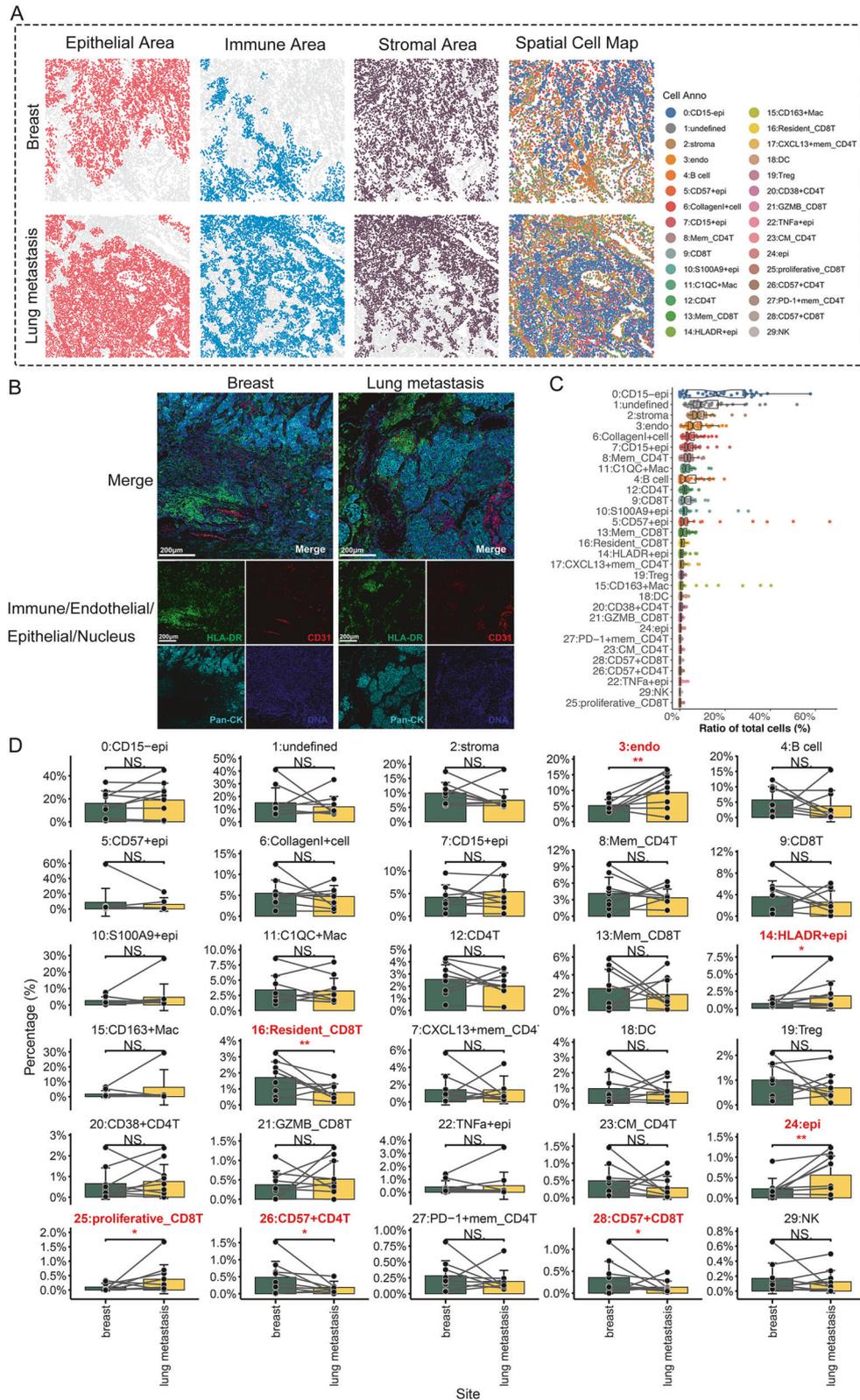
### Single-cell spatial analysis of paired BC lung metastasis

To unravel the spatial topology of BC lung metastasis, we conducted a comprehensive analysis strategy by integrating IMC, mIF, and spatial proteomics on paired BC and lung metastasis FFPE tissues (Fig. 1A). Initially, we employed a 40-marker panel of IMC to elucidate the spatial differences in the TME between BC and paired lung metastasis. The specimens were stained with the 40-plex antibody panel, targeting various cell types and immune checkpoints. Subsequently, we obtained high-dimensional histopathological images for each specimen, highlighting structural markers such as pan-cytokeratin (Pan-CK) for epithelial cells, collagen I for stromal cells, and immune cell markers including CD20 and CD3 for B cells and T cells, respectively (Fig. 1B, C). During the laser ablation process, we recorded the expression levels of markers and the relative spatial locations of cells. Using the IMC analysis pipeline, we generated cell masks for each cell, enabling us to calculate the spatial distribution of cells within the TME (Fig. S1), as well as the expression matrix of markers. The correlation heatmap of the staining markers revealed strong correlations between CD20 and CD45 or CD45RO (Fig. 1D), consistent with previous studies [21, 22]. The IMC images were segmented into 269,772 high-quality cells with spatial information, which were clustered into 29 distinct immune cell clusters, along with endothelial cells, epithelial cells, smooth muscle cells, lineage negative cells and collagen I (Fig. 1E). These cell clusters were subsequently clustered and annotated based on the expression of canonical markers (Fig. 1E, F), and the average expression of lineage markers for each cluster is illustrated in Fig. 1G. The lineage negative cluster contained the cells which could not be defined as any cell type. The distribution of cells from different subtypes (HER2 overexpression, luminal, triple negative) or originating from different sites (breast and lung metastasis) was depicted by UMAP in Fig. S2A and Fig. S2B. The proportions of each cluster relative to the total cells were also calculated and illustrated (Fig. S2C), demonstrating the distribution of TME cell abundance in each patient. CD57, typically found on active immune cells, was observed to be expressed in a cluster of epithelial cells (referred to as CD57<sup>+</sup> epithelial cell). A previous study also identified CD57 expression on epithelial cells in prostate cancer [23]. To confirm the expression of CD57 on epithelial cells, we conducted IHC staining. Our findings revealed a proportion of epithelial cells expressing CD57, particularly on ductal cells, consistent with the IMC analysis. Therefore, we confirmed the accuracy of IMC analysis in our results. Taken together, we applied IMC to depict the cell types in TME from BC and lung metastasis.

Using IMC data, we identified the epithelial, immune, and stromal areas of the TME, which were then merged to create a spatial cell map in BC lung metastasis (Fig. 2A, Figs. S3 and 4). Staining for HLA-DR, CD31, Pan-CK, and DNA revealed the immune, endothelial, and epithelial areas, as well as cell nuclei within the TME (Fig. 2B). The percentage of each cell type was visualized using boxplots, showing CD15<sup>+</sup> epithelial cells, stromal cells, and endothelial cells as the most abundant in the overall spatial cell atlas (Fig. 2C). We then conducted paired analysis to quantify cell type abundance in BC and lung metastasis (Fig. 2D and Fig. S5A). Endothelial cells, HLA-DR<sup>+</sup> epithelial cells, and proliferative CD8<sup>+</sup> T cells exhibited an increased tendency in lung metastasis compared to BC. Conversely, CD57<sup>+</sup> CD4<sup>+</sup> T cells and



**Fig. 1 Imaging mass cytometry (IMC) revealing the spatial ecosystem of breast cancer lung metastasis at a single-cell protein level.** **A** Overview of the study design and workflow. Formalin-fixed paraffin-embedded samples from paired primary sites and lung metastases were collected for IMC, multiplex immunofluorescence (mIF), spatial proteomics, single-nucleus RNA-sequencing. **B** Panel of all the markers used in IMC. **C** Schematic illustration of region of interest (ROI) selection for targeted multiplexed imaging by IMC. Pan-CK was labeled with green, DNA was labeled with blue and target marker was labeled with red. **D** The correlation heatmap of the staining markers. **E** Uniform Manifold Approximation and Projection (UMAP) showing the identified 29 distinct immune cell clusters. **F** Major clusters annotated with canonical markers. **G** Average expression of markers for each cluster.



**Fig. 2** IMC unraveling the cell component frequency differences in TME between breast cancer and paired lung metastasis. **A** Spatial cell map in breast cancer lung metastasis merged based on IMC data, including epithelial, immune, and stromal areas. **B** Representative IMC images of breast and lung metastasis with HLA-DR, CD31, Pan-CK, and DNA staining. Scale bar = 200  $\mu$ m. **C** The percentage of each cell type visualized using boxplots. **D** The abundance of different cell types between breast cancer and paired lung metastases quantified by paired analysis. \* $P < 0.05$ , \*\* $P < 0.01$ .

CD57<sup>+</sup> CD8<sup>+</sup> T cells showed a decreased tendency. Also, we examined the subtype-specific cell proportion in both BC and lung metastasis. The cell proportion in each subtype showed a similar tendency in both BC and lung metastasis. For instance, CD15<sup>+</sup> epithelial cells were enriched in HER2 overexpression subtype in both breast and lung site. Conversely, CD57<sup>+</sup> epithelial cells showed increased proportion in triple negative subtype in both breast and lung site (Fig. S5B, C).

We also performed the Ro/e analysis to assess organ-specific and subtype-specific enrichment of cell clusters (Fig. S6A, B). Endothelial cells and HLA-DR<sup>+</sup> epithelial cells were found to be increased in lung metastasis compared to BC, while CD57<sup>+</sup> epithelial cells were specifically enriched in BC. In terms of molecular subtype heterogeneity, CD57<sup>+</sup> epithelial cells, S100A9<sup>+</sup> epithelial cells, and CD163<sup>+</sup> macrophages were significantly enriched in TNBC, whereas CD15<sup>+</sup> epithelial cells were predominantly abundant in HER2 overexpression BC. Notably, endothelial cells and HLA-DR<sup>+</sup> epithelial cells showed significant enrichment in lung metastasis compared to BC, particularly in the triple-negative subtype (Fig. S6C).

Overall, the differences in TME components between primary BC and lung metastasis were illustrated through IMC analysis.

#### Cellular neighborhoods and cell interaction/avoidance analysis revealed the organ-specific immune spatial topology

The spatial configuration of the TME plays a crucial role in tumor metastasis and response to therapy. Consequently, we conducted an in-depth analysis of regional CN to unveil the multicellular structures within BC and lung metastasis. We defined the CN as the nearest 20 cells to the center cell (Fig. S7A). To visualize organ-specific cell interactions and functional units within the TME of BC, we generated network, Voronoi, and CN patch plots for each IMC image (Figs. 3A and S8). A total of 15 CNs were identified and annotated based on major cell types (Fig. 3B). The distribution of each CN across BC and lung metastasis in different patients is illustrated in pie plots and histograms (Figs. 3C and S7B). The cellular compositions of these CNs fully mirrored the architectural features of the TME, including spots enriched with HLA-DR<sup>+</sup> epithelial cells alongside endothelial cells (CN6), regions rich in stromal cells & C1QC<sup>+</sup> macrophages (CN2), those abundant in CD57<sup>+</sup> epithelial cells and CD163<sup>+</sup> macrophages (CN15), and areas dominated by CD15<sup>+</sup> epithelial cells (Fig. 3D).

Subsequently, we proceeded to compare the enrichment levels of each CN between BC and lung metastasis (Fig. 3E, F). The Ro/e analysis indicated a higher abundance of CN6 in lung metastasis. Furthermore, CN6 was found to be enriched in the triple negative subtype compared to other molecular subtypes. In contrast, CN1 and CN5 exhibited enrichment in the HER2 overexpression subtype (Fig. S9A). Notably, CN6 displayed minimal enrichment in the breast site of the triple negative subtype, whereas the paired lung metastasis demonstrated a significant abundance of CN6 in this subtype (Fig. S9B). Furthermore, paired analysis demonstrated that the frequencies of CN6, enriched with HLA-DR<sup>+</sup> epithelial cells and endothelial cells, were significantly higher in primary breast tumors than in lung metastases. Finally, we conducted regional correlation analysis to delineate spatial interaction and avoidance pairs in the topology of BC lung metastasis (Figs. 3G, H and S7C). Endothelial cells were found to interact frequently with HLA-DR<sup>+</sup> epithelial cells, thereby confirming the cell module comprising HLA-DR<sup>+</sup> epithelial cells and endothelial cells (CN6).

In conclusion, utilizing advanced spatial analysis based on IMC, a distinct cell hub was identified, revealing organ-specific properties between the primary site and lung metastasis.

#### BC lung metastasis demonstrated a more exhausted TME compared to primary BC

As immunotherapy emerges as a pivotal therapeutic option in BC treatment, we conducted an investigation into the expression

patterns of key T cell regulators across primary breast tumors and lung metastases. Initially, the expression profiles of all IMC markers were visualized via a heatmap (Fig. S10A). At the patient level, markers such as  $\alpha$ SMA, CD68, GZMB, S100A9, CD31, CD15, and CD208 exhibited differential expression between BC and lung metastases. Notably, TIM3 and PD-1 showed significant upregulation on T cells derived from lung metastases compared to those from BC, particularly evident in the triple-negative subtype (Fig. S10B). Subsequently, we conducted mIF staining for validation (Fig. 4A). Antibodies of CD3, CD8, CD31, CK, PD-1, HLA-DR, DAPI were used to depict the TME ecosystem of BC lung metastasis. We firstly investigated the distance of HLA-DR<sup>+</sup> epithelial cells and exhausted T cells to endothelial cells. The distance of HLA-DR<sup>+</sup> epithelial cells was plotted (Fig. 4B) and the spatial plot showed significant cell network between endothelial cells and HLA-DR<sup>+</sup> epithelial cells in BC lung metastasis (Fig. 4C). Additionally, the pie plot revealed a shorter distance between HLA-DR<sup>+</sup> epithelial cells and exhausted T cells to endothelial cells in each subtype, with a notable emphasis on the triple negative subtype (Fig. 4D).

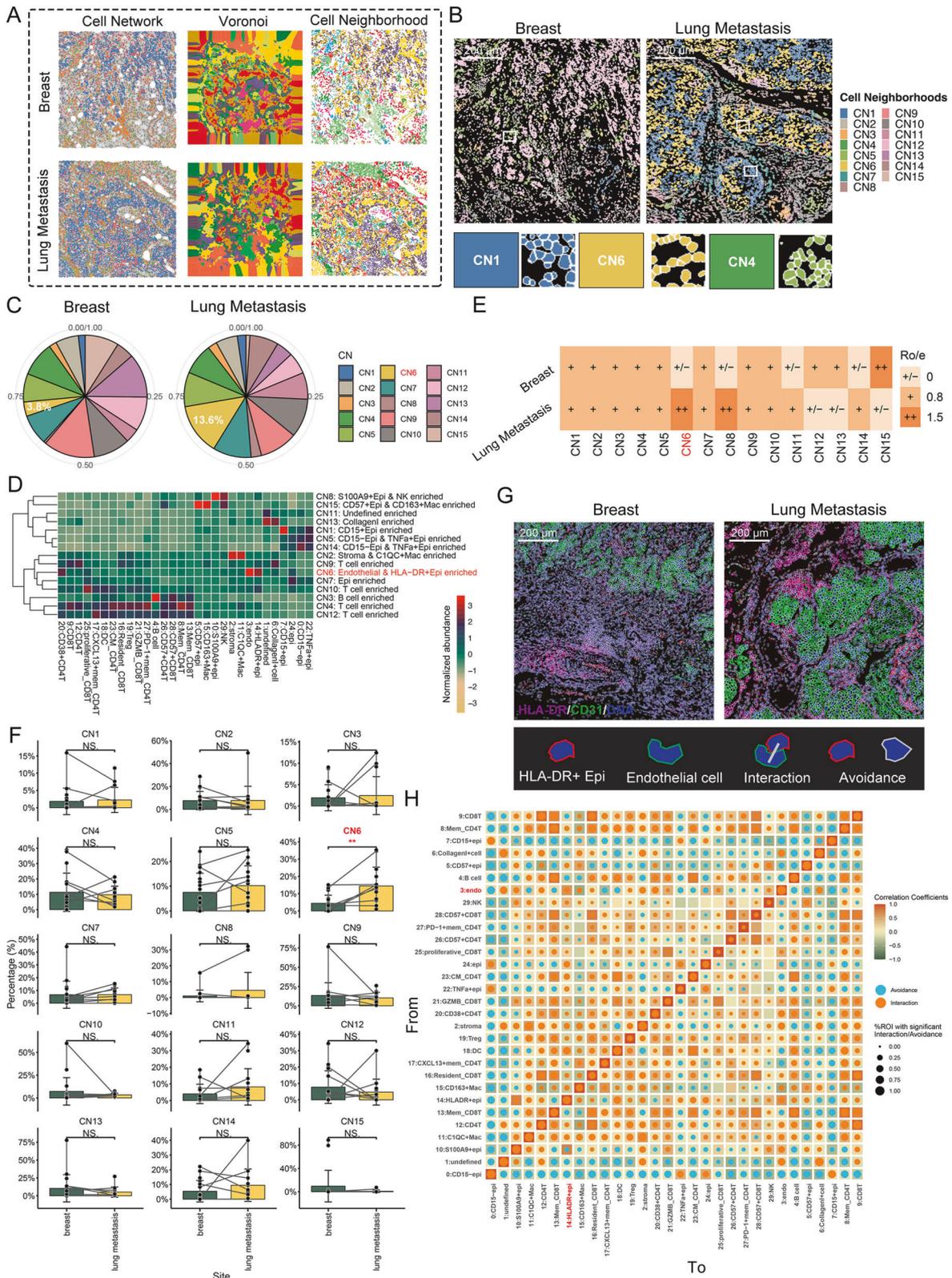
A recent study defined cells within 25  $\mu$ m as having “direct interaction” in IMC and mIF analyses [24]. To validate the cellular hub identified by IMC in human breast tumor lung metastasis, we performed similar analysis (Fig. 4E). Consistent with our previous definition, we considered cells within a 20  $\mu$ m window as having direct interaction. Our results demonstrated that all three BC subtypes showed an increased proportion of cell hubs in lung metastases compared to primary breast tissue, particularly in the triple-negative subtype (Fig. 4F).

Taken together, we revealed an exhausted TME in lung metastasis compared with that in BC, especially in triple negative subtype. The mIF staining and spatial analysis further validated the significant cell network between HLA-DR<sup>+</sup> epithelial cells and endothelial cells.

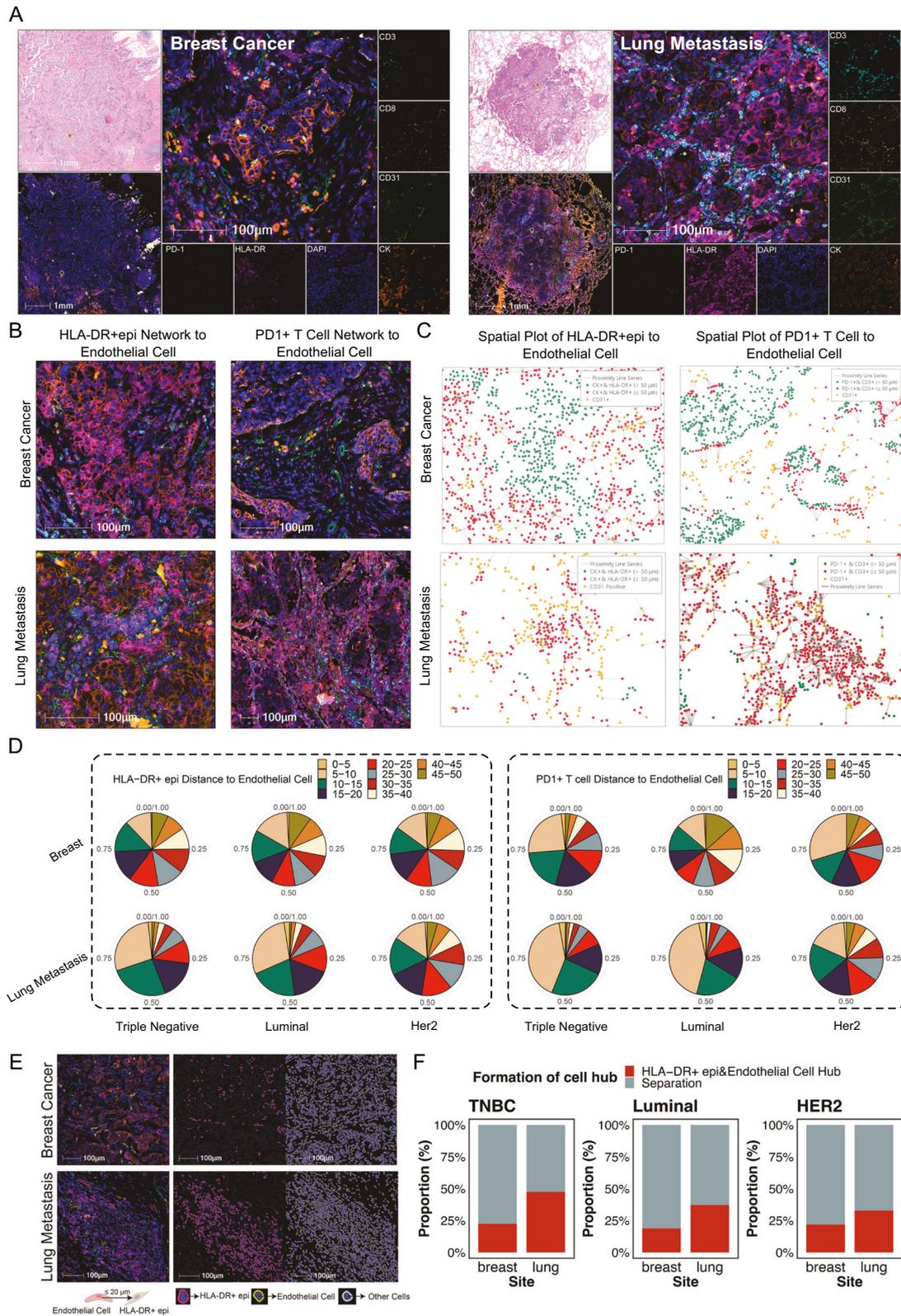
#### Paired spatial resolved proteomics analysis unveiled the underlying biological mechanism in BC lung metastasis

Given the limited number of markers available with the IMC technique, it becomes challenging to thoroughly explore potential pathway alterations and molecular activities within tumor tissues. To deepen our understanding of the spatial diversity and potential mechanisms underlying BC lung metastasis, we have developed a MS-based technique which extended compatibility to archived FFPE sections. This filter-aided expansion proteomics (FAXP) method combines manual tissue microdissection and significantly increase in volumetric resolution, peptide yield, as well as protein identifications, which enhances spatial resolution through tissue expansion, with bottom-up MS-based proteomic analysis [25, 26]. This methodology involves slicing tissue samples sequentially and conducting both IMC and spatial proteomics analysis to investigate the molecular characteristics within the ROIs identified by IMC (Fig. 5A). Sequential sectioning of tumor tissues was performed, with adjacent pairs selected for processing. One section was stained with H&E for ROI selection, while the other was used for IMC analysis. The H&E-stained sections were then employed for spatial proteomics studies. After HE staining, tissue expansion was carried out, and the selected ROIs were expanded for dissection. These dissected ROIs were subsequently analyzed by LC-MS/MS. By integrating IMC and spatial proteomics, we were able to uncover molecular features.

As demonstrated in Fig. 5A, we selected the exactly same 36 ROIs consistently across 18 samples. Tissue microdissection is facilitated by tissue expansion and staining via embedding it into a swellable hydrogel. Subsequently, peptide extraction and analysis are performed using a trapped ion mobility spectrometry (tims)TOF Pro-mass spectrometer in parallel accumulation-serial fragmentation combined with data-independent acquisition



**Fig. 3 Organ-specific immune spatial topology revealed by cellular neighborhoods (CNs) and cell-cell interaction and avoidance analysis.** **A** Representative Network, Voronoi and CN patch plots of TME in breast cancer and paired lung metastasis to visualize organ-specific cell interactions and functional units. **B** Identification of 15 distinct CNs from the 29 cell clusters in breast cancer and lung metastasis. **C** The distribution of each CN across breast cancer and lung metastasis illustrated in pie plots. **D** Cellular compositions of distinct CN and their corresponding abundances in each CN. **E** Tissue prevalence of each cluster on IMC images estimated by the Ro/e analysis. **F** Enrichment of each CN between breast cancer and lung metastasis. **G** Regional correlation analysis delineating spatial interaction and avoidance pairs in the topology of breast cancer lung metastasis. **H** Cell-cell interaction and avoidance analysis revealing the spatial cell interactions on IMC images.



(diaPASEF) mode. Quality control of the MS data showed robust reproducibility for both global precursors and proteins (Fig. S11A, B). Consequently, we have confirmed that paired spatial proteomics maintains high quality and provides detailed protein information regarding the components of the TME.

Therefore, we investigated the differences in signaling pathways underlying BC lung metastasis using ssGSEA analysis on spatial proteomics data. Unsupervised analysis of immunotherapy-related and hallmark pathways revealed significant intratumoral heterogeneity across breast and lung samples, indicating both

**Fig. 4 Differences in exhausted T cell phenotype between primary breast cancer and lung metastasis. A** The mIF staining for CD3 (cyan), CD8 (yellow), CD31 (green), PDCD1 (red), HLA-DR (purple), Pan-CK (orange) and DAPI (blue) in breast cancer and paired lung metastasis. Scale bar = 1 mm and 100  $\mu$ m, respectively. **B** The spatial networks of HLA-DR<sup>+</sup> epithelial cells and PD1<sup>+</sup> T cells to endothelial cells in breast cancer and lung metastasis. The white line indicates the connection between cells. The CD3 (cyan), CD8 (yellow), CD31 (green), PDCD1 (red), HLA-DR (purple), Pan-CK (orange) and DAPI (blue) was shown with the same color in Fig. 4A. Scale bar = 100  $\mu$ m. **C** The spatial plot of HLA-DR<sup>+</sup> epithelial cells and PD1<sup>+</sup> T cells to endothelial cells in breast cancer and lung metastasis. **D** The pie chart displaying the proportion of distance from HLA-DR<sup>+</sup> epithelial cells and PD1<sup>+</sup> T cells to endothelial cells. **E** HLA-DR<sup>+</sup> epithelial and endothelial cell hub in breast and lung metastasis sites validated by mIF spatial analysis. Cell hub was defined as the region within a 20  $\mu$ m proximity between the two cell types. HLA-DR<sup>+</sup> epithelial cells were labeled with rosein, endothelial cells were labeled with yellow, other cells were labeled with gray, and DNA was labeled with blue and target marker was labeled with red. **F** Difference in formation of cell hubs between primary breast tissue and lung metastases in different subtypes.

patient-related and site-related variability (Fig. 5B). Subsequently, we conducted an analysis of DEPs between BC tumors and their paired lung metastases. Notably, SEMA3B, MMP28, and CXCL17, among others, were up-regulated, while KRT20, COL2A1, and CD70, among others, were down-regulated (Fig. 5C). Leveraging the DEPs, we performed GO analysis (Fig. 5D), revealing enriched biological processes such as positive regulation of cell adhesion and regulation of chemotaxis, consistent with the functions of up-regulated proteins in lung metastasis. Conversely, processes like collagen fibril organization, cytoplasmic translation, and endodermal cell differentiation were enriched in BC tissues. Furthermore, ssGSEA analysis highlighted the upregulation of apical surface and KRAS signaling, as well as cholesterol homeostasis, in lung metastasis, while interferon alpha and gamma responses were positively regulated in BC tissues compared to lung metastasis. At the paired sample level, GSEA results unveiled a down-regulated interferon gamma response in lung metastasis and an upregulated estrogen response in breast tissue samples, consistent with ssGSEA and GO analyses. Hence, we demonstrated molecular differences between BC tissues and their paired lung metastases (Fig. 5E, F).

#### WGCNA indicated the enriched molecular features associated with spatial TME components

To explore deeper into the molecular mechanisms underlying the cellular components within the TME of breast lung metastasis, we conducted WGCNA. The frequency of TME component cells was determined using IMC in each ROI. This information was then correlated with the protein module identified by WGCNA. Subsequently, the potential biological processes associated with each TME component were investigated (Fig. 6A). Through analysis of mean connectivity and scale independence, a soft threshold of 8 was chosen (Fig. S12A, B). Subsequently, all proteins were assigned to 11 protein modules using an unsupervised clustering method (Fig. 6B), with 11 colors representing the distinct protein modules, each comprising proteins with similar expression patterns (Fig. 6C). We then correlated the cell frequencies of identified cell components via IMC with each module (Fig. 6C). For example, endothelial cell frequency exhibited a negative correlation with the yellow module, whereas CD57<sup>+</sup> epithelial cell frequency showed a positive correlation with the yellow module (Fig. 6D, E). Similarly, S100A9<sup>+</sup> and HLA-DR<sup>+</sup> epithelial cells were both positively correlated with the red module (Fig. 6F, G). Next, we identified proteins with the highest protein significance for each cell type, indicating those proteins most closely associated with the selected cell type (Fig. 6D–G, abs (protein significance) >0.6). Using these selected hub proteins for each cell type (Fig. 6H–K), we conducted GO analysis for the respective cell types (Fig. 6L–O). Endothelial cells exhibited terms related to cellular energy processes (Fig. 6L), CD57<sup>+</sup> epithelial cells were enriched with oxidative phosphorylation processes (Fig. 6M), S100A9<sup>+</sup> epithelial cells displayed metabolic process enrichments (Fig. 6N), while HLA-DR<sup>+</sup> epithelial cells were associated with lipid catabolic processes (Fig. 6O). In summary, WGCNA analysis revealed the molecular characteristics associated with the TME

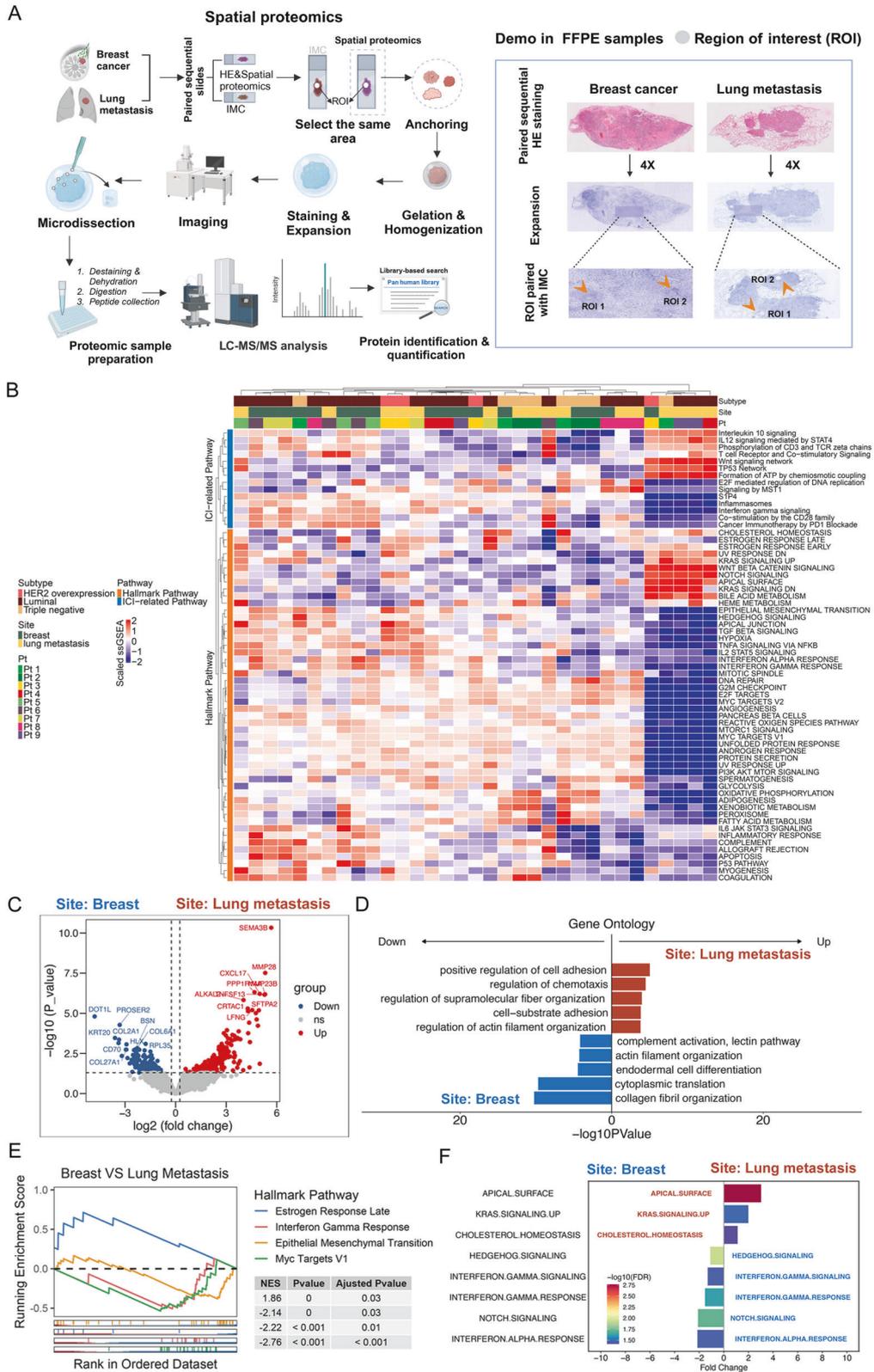
components during BC lung metastasis, as highlighted by spatial proteomics analysis.

#### snRandom-seq analysis of BC lung metastasis

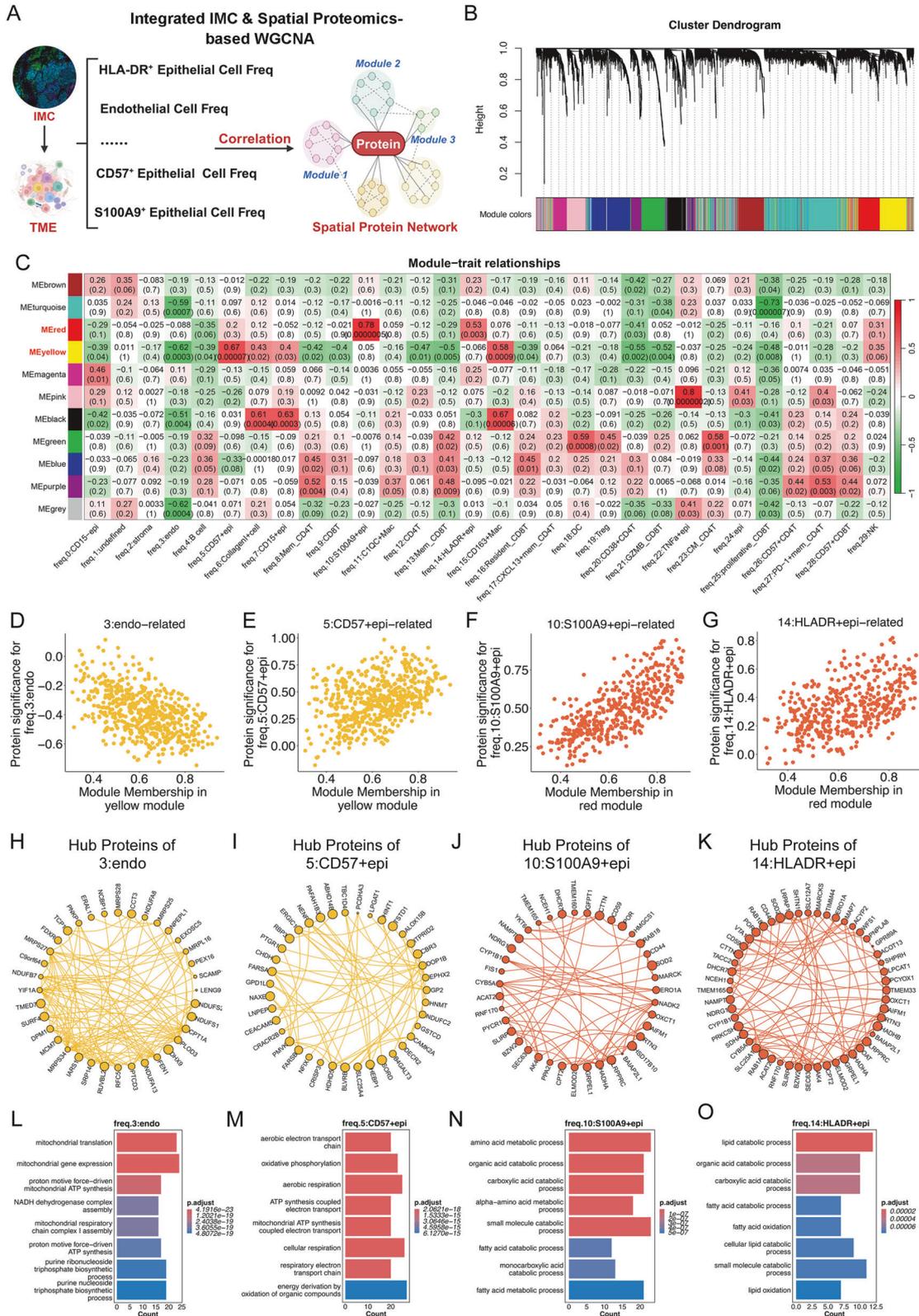
We revealed greater number of endothelial cells in BC lung metastasis at a spatial level, especially in TNBC. As IMC represents a small area of TME component in tumor tissues. To further validate the organ-specific TME of BC lung metastasis at single-cell resolution on the whole slide level, we performed snRandom-seq on 3 treatment-naïve FFPE tissue samples from primary BC ( $n = 3$ ) and lung metastases ( $n = 3$ ) using previously reported technology (Fig. 7A) [27]. The unsupervised clustering analysis classified the 100,235 cells into eight clusters (Fig. 7B). Each cluster was identified as a broad cell population with its canonical markers (Fig. 7C). We further calculated the inter-patient and inter-organ proportions and Ro/e analysis in each cell population to characterize the organ-specific TME of BC lung metastasis (Fig. 7D, E). The tissue enrichment of each cell population was evaluated by Ro/e analysis. Endothelial cells were relatively enriched in lung metastasis at a FFPE whole slide image level (Fig. 7E), which is consistent with the previous results from IMC and mIF. Due to the limitations of snRNA-seq, we were unable to identify exhausted T cells and HLA-DR<sup>+</sup> epithelial cells, as only RNA expression from nuclei can be extracted from FFPE samples. Consequently, we re-clustered the endothelial cells to delineate subpopulation specificity in TNBC lung metastasis, resulting in the identification of 14 distinct clusters (Fig. 7F). Utilizing markers for lymphatic and vascular endothelial cells (Fig. 7G), we identified 591 lymphatic endothelial cells and 5431 vascular endothelial cells (Fig. 7H). Further characterization revealed that the vascular endothelial cells could be classified into three subtypes: arterial endothelial cells, general capillary endothelial cells, and venous endothelial cells (Fig. 7I, J). Notably, general capillary endothelial cells exhibited relative enrichment in lung metastasis, as observed at the level of FFPE whole slide images (Fig. 7K). In summary, snRNA-seq analysis provided insights into the detailed subpopulations of endothelial cells enriched in lung metastasis compared to those present in primary BC.

#### DISCUSSION

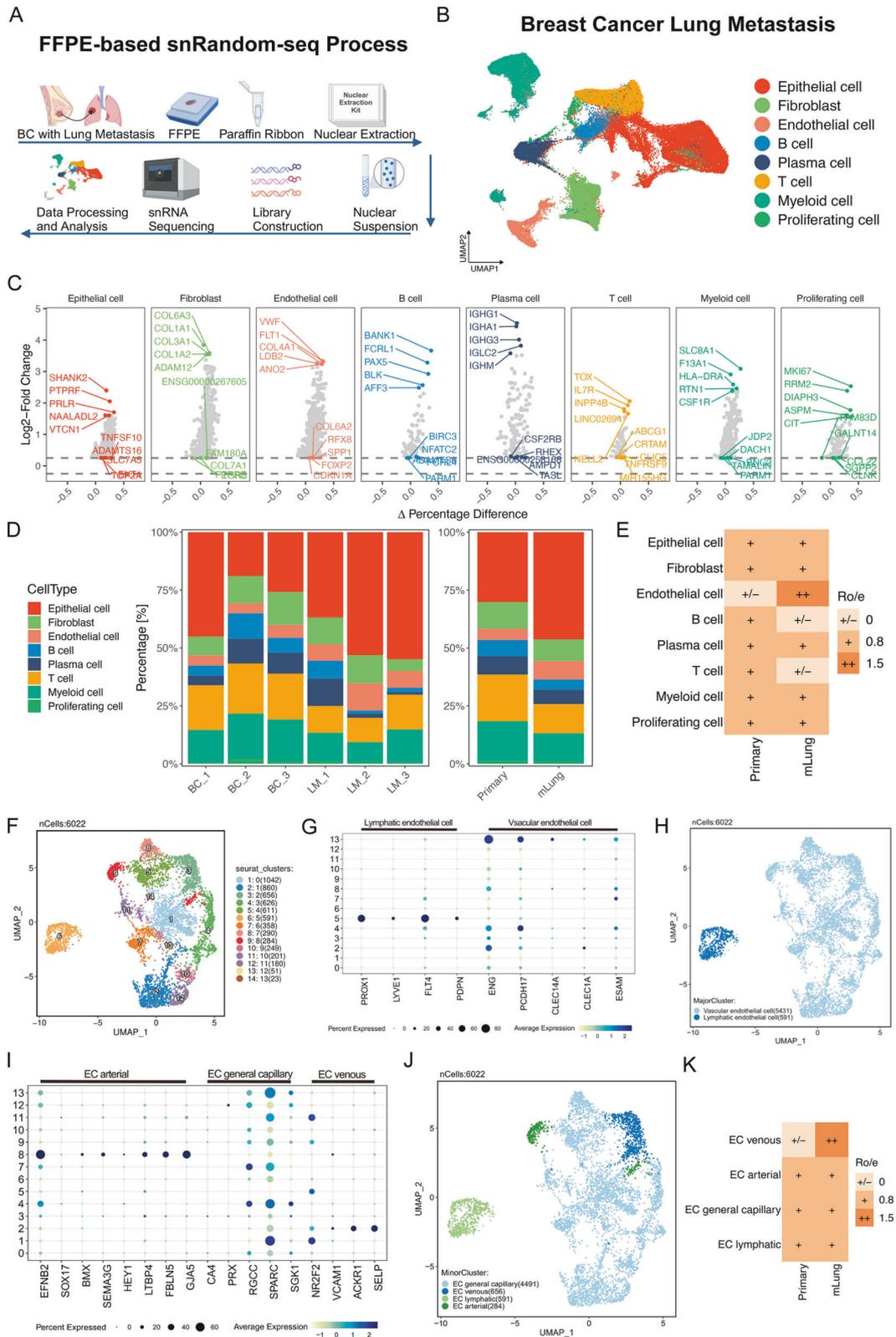
BC lung metastasis presents a key challenge in clinical management, indicating the importance of unraveling the complexities of the TME and its impact on treatment response [28, 29]. Recent studies have turned towards spatially organized cell populations within metastatic lesions as potential therapeutic targets across various tumor types [30, 31]. Notably, primary and metastatic tumors exhibit distinct spatial cellular hubs, influencing their responses to different therapies [32]. Exploring organ-specific metastasis mechanisms and their unique ecosystems holds promise for refining treatment approaches [33]. mBC pose a challenge in obtaining sufficient fresh tissue for conducting scRNA-seq or spatial transcriptome analysis, as patients with advanced TNM stage IV mBC often miss the window for surgical intervention. As a result, the TME of lung-specific metastases using



**Fig. 5 Paired spatial proteomics unveiled underlying biological mechanism of breast cancer lung metastasis. A** Overview of the design and workflow of mass spectrometry-based spatial proteomics applied to formalin-fixed paraffin-embedded tissue samples from paired primary and lung metastatic lesions. **B** Differences in immunotherapy-related and hallmark pathways in each patient and each site revealed by unsupervised analysis. **C** Differentially expressed proteins between breast cancer and paired lung metastasis. **D–F** The potential biological functions and relevant signaling pathways evaluated by gene ontology analysis, gene set enrichment analysis (GSEA) and single-sample GSEA analyses.



**Fig. 6 Weighted Gene Co-expression Network Analysis (WGCNA) showing enriched molecular features associated with each TME component. A** Overview of the design for integrated IMC and spatial-proteomics based WGCNA. **B** Proteins were assigned to 11 cell modules using an unsupervised clustering method. **C** Each cell module correlated with cell frequencies of identified cell components via IMC. **D–G** The association of protein significance and module membership for selected cell types and protein module showing by Pearson’s coefficient. **H–K** The hub proteins identified for the respective cell types. **L–O** Gene ontology analysis for the respective cell types.



**Fig. 7 TME atlas of primary triple negative breast cancer and lung metastases by scRandom-seq.** **A** Overview of the design for snRandom-seq. FFPE samples from primary TNBC and lung metastasis were collected for single-cell nuclei transcriptomic sequencing. **B** UMAP plot of major cell types from all the samples. **C** The volcano plot of highly variable genes for the major clusters. **D** The frequency of each cell cluster presented as a proportion of total cells in each sample. **E** Tissue prevalence of each cell cluster estimated by the Ro/e analysis. **F** UMAP plots of endothelial cells (ECs) showing distinct subclusters. **G** Dot plot showing the markers for lymphatic endothelial cell and vascular endothelial cell. **H** UMAP plots of lymphatic endothelial cell and vascular endothelial cell. **I** Dot plot showing the markers for EC arterial cell, EC general capillary and EC venous. **J** UMAP plots of all four cell clusters. **K** Tissue prevalence of each EC cluster estimated by the Ro/e analysis.

paired samples remains largely unexplored, especially at a single-cell spatial resolution. In our study, we developed a comprehensive strategy leveraging cutting-edge techniques such as IMC, spatial proteomics and mIF to explore spatial proteomic and RNA profiling within BC and paired lung metastatic lesions and identify potential therapeutic targets. Through the analysis of paired samples, we delineated the spatial landscape and characterized the heterogeneous atlas of the TME at a single-cell resolution. Importantly, all samples utilized in our study were FFPE samples, allowing for the application of these advanced techniques without reliance on fresh tissue.

BC reprograms the lung microenvironment to generate pre-metastatic niches, and pro-angiogenic factors (such as vascular endothelial growth factor) is one of the initial molecules involved [3]. Angiogenesis plays a key role in tumor growth and metastasis. After decades of endocrine and cytotoxic chemotherapy, targeted therapies have brought new treatment options such as angiogenesis inhibitors, HER2-targeted therapies, cyclin-dependent kinase 4 and 6 (CDK4/6) inhibitors [5]. Combination chemotherapy with bevacizumab and some small-molecule tyrosine kinase inhibitors (TKIs) targeting angiogenesis may improve the prognosis of BC patients [5]. Few studies have also explored the use of Endostar in BC [34, 35]. These clinical treatment options are consistent with our findings of increased endothelial cells in lung metastases. Regarding immunotherapy, the introduction of immune checkpoint inhibitors in chemotherapy resulted in a higher treatment response and longer survival, but is currently mainly limited to TNBC [5]. In the meanwhile, we found that C1QC<sup>+</sup> Mac, DC, GZMB<sup>+</sup> CD8<sup>+</sup> T cells, proliferative CD8<sup>+</sup> T cells were significantly increased in lung metastases, indicating an immune-exhausted TME. Hence, both immunotherapy and anti-angiogenic therapy may act as promising strategies for BCs with lung metastasis.

We further identified a cell module enriched with HLA-DR<sup>+</sup> epithelial cells and endothelial cells. This cellular composition suggests an intricate spatial interaction between HLA-DR<sup>+</sup> epithelial and endothelial compartments within the TME. The differential abundance of HLA-DR<sup>+</sup> epithelial and endothelial cell hub between primary breast tumors and lung metastases highlights its dynamic role in the metastatic process. Furthermore, the observed spatial interactions between endothelial cells and HLA-DR<sup>+</sup> epithelial cells underscore the complex interplay between angiogenesis and immune activation within the TME by IMC. This also suggests that interventions targeting both angiogenesis and immune evasion pathways, such as anti-angiogenic combined immunotherapy, may disrupt the spatial organization of HLA-DR<sup>+</sup> epithelial and endothelial cell hub to enhance therapeutic efficacy. Li et al. have demonstrated a dose-dependent synergy between anti-angiogenic therapy and programmed cell death protein-1 (PD-1) blockade, showcasing efficacy in 12 cases of TNBC [36]. However, the scope of this combination strategy remains limited, necessitating further research to address this limitation. In our study, the combined use of IMC and mIF enabled us to uncover the presence of HLA-DR<sup>+</sup> epithelial cells and exhausted T cells clustering around endothelial cells, a phenomenon that was particularly pronounced in the triple-negative subtype. Leveraging anti-angiogenic therapy has the potential to disrupt this specialized cellular hub, leading to the exposure of HLA-DR<sup>+</sup> epithelial cells and exhausted T cells. This, in turn, creates an opportunity for immune checkpoint inhibitor (ICI)-based immunotherapy to reverse the exhausted T cells into a cytotoxic phenotype, thereby making the HLA-DR<sup>+</sup> epithelial cells susceptible to targeted intervention.

Moreover, snRNA-seq revealed the TME exhibited a less Treg phenotype in lung metastasis than that in BC, which indicated the “reversed” un-exhausted T cells after ICI treatment may play more crucial roles. Therefore, based on our findings, we proposed that anti-angiogenic therapy may sensitize BC to PD-1 blockade in patients with BC lung metastasis, especially in TNBC.

Owing to the limited availability of samples, the comparison between different molecular subtypes was difficult in this study. Further collection of surgical specimens is necessary to investigate the immune heterogeneity between different pathological subtypes. In addition, more samples are needed to investigate the impact of differentiation status on the spatial topology of BC lung metastasis. Due to the limitations of snRNA-seq, the identification of HLA-DR<sup>+</sup> epithelial cells is not feasible. Consequently, this constraint precludes the analysis of ligand-receptor interactions between HLA-DR<sup>+</sup> epithelial cells and endothelial cells. The other criticism of this work is the lack of functional aspect of the features that have been identified from this data analysis.

In summary, we have utilized a comprehensive approach to depict the spatial multi-omics profiling of paired BC and lung metastasis samples. Our study identified a spatial endothelial-HLA-DR<sup>+</sup> epithelial cell hub within BC lung metastasis, which may play a pivotal role in determining response to anti-angiogenic therapy and immunotherapy. Therefore, our results may help develop personalized treatment strategies and improve patient outcomes in BC lung metastasis.

## METHODS

### Patients and samples

A total of 9 BC patients with oligo-recurrent lung metastasis after surgery in the First Affiliated Hospital, Zhejiang University School of Medicine (FAHZU) were retrospectively selected. Matched surgical specimens of primary breast tumor and lung metastasis were collected. The clinical features of these patients are shown in Supplementary Table 1. Samples for IMC, spatial proteomics and mIF were taken from tumor area. This study was carried out in strict accordance with the standards of the Declaration of Helsinki and was approved by the Ethics Committee of FAHZU (Ethical number: IIT20240530A). Informed consent was obtained from all the patients or their valid proxies.

### IMC analysis

Nine samples of primary BC and nine matched lung metastases were selected for IMC and downstream analysis. The FFPE samples were successively sliced, and one of the slices was taken from each patient for hematoxylin-eosin (H&E) staining. Experienced pathologists then identified one or two regions of interest (ROIs) of 500 × 500 μm<sup>2</sup> where immune cell infiltration was most abundant according to the morphological structure shown on H&E slides. Immune cell infiltration was characterized by a large number of small cell clusters that are predominantly blue stained, ranging from 1000 to 7000 square microns. Immune cells can be observed both peripherally and internally in tumor tissue. Adjacent sections were labeled with pre-designed IMC antibodies (Supplementary Table 2). Based on our previously published study [37], the ROIs were captured as square regions with a laser intensity of 400 Hz. Then the collected raw data were preprocessed by overflow signal compensation, image denoising, image contrast enhancement and cell segmentation. The individual cell or component in each channel of IMC image was segmented by the connection-sensing segmentation method [21]. The region props function in MATLAB was used to identify connected components within the image for cell segmentation. Artifacts were eliminated if the centroid of the nearest core exceeded 15 pixels in the case of other membrane channels. Each marker expression was normalized to the 99th percentile of each channel. The “Harmony” package (version 0.1.0) and “Rphenograph” (version 0.99.1) with in 100 nearest neighbors were used separately to correct batch processing effects and cell clustering. The cluster means were presented in the form of heatmaps and were utilized for further annotation. The “imcRtools” package (version 1.0.2) was used for IMC downstream data analysis. The 20 nearest neighboring cells of each cell were determined as cellular neighborhood (CN) based on Euclidean distance. Then the neighboring cells were clustered through K-means clustering (k = 15) based on the 29 cell clusters, along with endothelial cells and epithelial cells [38]. To verified CNs, the Voronoi diagrams were superimposed over the corresponding original IMC images. A permutation test method of the “imcRtools” package (version 1.0.2) was used to evaluate the interactions/avoidance between different cell clusters within each CN to explore spatial cell-cell interactions [39].

### Tissue distribution of clusters

We calculated the ratio of observed to expected cell numbers (Ro/e) for each cluster in different tissues to quantify the tissue preference of each cluster [40, 41]. The expected cell numbers for each combination of cell clusters and tissues were obtained from the chi-square test. One cluster was identified as being enriched in a specific tissue if  $Ro/e > 1$ .

### Spatial proteomics and downstream analysis

Paired FFPE samples of primary breast and lung metastases were prepared into 4  $\mu$ m slices. Hydrogel embedding, expansion, staining, imaging, microdissection, tissue polypeptide recovery, and mass spectrometry of recovered polypeptides were performed as previously reported [25]. After being treated with BT buffer and MES buffer, the slices were incubated with protein anchoring solution for 12 h. The samples were then washed with anchoring stop buffer, reacted with Activated Monomer Solution in gelation chamber at 4 °C for 12 h for gelation, and transferred to vacuum oven for polymerization reaction. The resulting tissue-hydrogel composite was immersed in a homogeneous buffer for expansion. After Coomassie staining and continuous washing, tissues were imaged using Zeiss Fluorescence Stereo Zoom Microscope. Microdissection was performed from ROIs in the expanded Coomassie-stained samples through destaining, dehydration, digestion and peptide collection. Finally, the samples were analyzed by liquid chromatography (LC)-mass spectrometry (MS)/MS. The peptide extraction and analysis are performed using a hybrid trapped ion mobility spectrometry (TIMS) quadrupole time-of-flight mass spectrometer (timsTOF-Pro) in parallel accumulation-serial fragmentation combined with data-independent acquisition Parallel Accumulation Serial Fragmentation (diaPASEF) mode as previous described [25, 42]. Bruker otofControl (version 6.2) and HyStar (version 5.1) were used for MS data acquisition. The data-dependent acquisition (DDA) data were analyzed using the FragPipe (version 15.0) platform and the MSFragger (version 3.1.1) [43, 44]. The self-built library is further used for analysis of PulseDIA data by DIA-NN (version 1.7.15) [45]. The differently expressed proteins (DEPs) analysis was performed using the “limma” package [46]. Gene Ontology (GO) analysis was performed using the “clusterProfiler” package (version 4.10.0) [47]. The single-sample gene set enrichment analysis (ssGSEA) score for each gene set was calculated using the “Gene Set Variation Analysis (GSVA)” package in R software (version 1.50.0) [48]. Weighted Gene Co-expression Network Analysis (WGCNA) was performed using the default parameter and standard pipeline [49].

### H&E staining

FFPE samples were continuously sliced into 4  $\mu$ m flakes. The slices were then successively dewaxed with xylene, rehydrated with graded ethanol, and rinsed with PBS. Next, the slices were dyed with hematoxylin at room temperature for half an hour and rinsed again with PBS. Soaking in ammonia turned the nuclei in the slices from red to bluish-purple. After cleaning with 75% alcohol for two minutes, the slides were dyed with eosin at room temperature for 1 h. Then we rinsed the slides directly with graded alcohol. Finally, anhydrous alcohol was replaced with xylene and then mounted on slides. The slides were examined using a light microscope (Leica), and the images were analyzed using Image-Pro Plus software (version 6.0).

### Immunohistochemistry (IHC) staining

BC tissue slides (4  $\mu$ m) were dewaxed with xylene and rehydrated with graded alcohol. Microwave heating was used to induce antigen epitope retrieval and blocking solution (Proteintech, B900780) was then used to block at room temperature for 1 h. The slides were incubated with primary antibody against CD57 (1:50, MA5-16948, Invitrogen) overnight at 4 °C. After washing with PBS for 3 times, the slides were then incubated with a peroxidase-conjugated (HRP) secondary antibody (31430, Invitrogen) for 10 minutes. The slides were examined by a light microscope (Leica) after dehydration, cleaning, and sealing. The images were analyzed using Image-Pro Plus software (version 6.0).

### mIF staining

FFPE samples of BC and lung metastases were sequentially stained using a TSA seven-color kit (H-D110071-50T, Yuanxibio). H&E and IHC staining were performed to test the antibody concentration for mIF staining followed by the Society for Immunotherapy of Cancer (SITC) mIF staining guidelines [50]. The slides were washed successively in tris buffered saline with tween 20 (TBST) buffer and then transferred to preheated ethylenediaminetetraacetic acid (EDTA) solution for 30 minutes with

microwave heating at 60 °C for dewaxing. After cooling to room temperature, the slides were incubated with anti-CD3 antibody (1:200, Abcam) for at least 30 minutes followed by washed with TBST for 3 times. Then a HRP secondary antibody (#DS9800, Leica) were added and incubated at room temperature for 10 minutes. TSA 520 working solution was added to the sample area for 10 minutes according to the manufacturer's instructions, followed by washing with TBST for 3 times. Repeat the same procedure for subsequent antibodies and fluorescent dyes in the following order: anti-CD3 (1:200, Abcam)/TSA 480, anti-CD31(1:200, CST)/TSA 520, anti-CD8 (1:100, Invitrogen)/TSA 570, anti-PD-1(1:100 CST)/TSA 620, anti-HLA-DR (1:1000, Abcam)/TSA 690, anti-Pan-CK (1:200, Abcam)/TSA 780. Wash each slide with distilled water and cover manually. The nuclei were stained with DAPI solution (Thermo Fisher Scientific, 62248) for 10 minutes. Panoramic 250 FLASH Tissue Imaging System (3D HISTECH) was used to scan the slides at 20 $\times$  magnification.

### Single-nucleus RNA-seq (snRNA-seq)

The nuclei of paired FFPE samples were extracted for snRNA-seq. The cross-sectional area of each sample embedded in the paraffin block was be greater than 0.3 cm<sup>2</sup>, and 2–3 paraffin rolls from each sample was cut in total, with the thickness of each roll about 20  $\mu$ m thick. Total RNA was extracted from a small amount of tissue for RNA quality control, and the integrity of RNA was evaluated based on the DV200 value. Follow-up experiments were only conducted for samples with DV200 greater than 40%. Droplet generation, single cell encapsulation, and cDNA capture were performed using the VITAcruizer Single-Cell Partitioning System (M20 Genomics). Using VITApilote High-Throughput Single-Cell Transcriptome Kit for FFPE samples (M20 Genomics), the procedures of dewaxing and blocking, RNA reverse transcription, single nucleus suspension preparation, single cell barcoding, cDNA amplification and library construction were performed successively according to the instructions. Qualified libraries were then selected for next-generation sequencing through the NovaSeq 6000 sequencing platform (Illumina). The FASTQ files were processed using VITAsier Bioinformatics Software (M20 Genomics). Specifically, sequencing reads in FASTQ format were aligned to the human reference genome (GRCh38) using the default parameters of the STAR software. A raw count matrix for each sample was generated through unique molecular identifiers (UMIs) counting and preliminary barcode screening. Following removal of low-quality or ambient barcodes based on UMI thresholds, the filtered gene expression matrix within retained high-confidence cell barcode were then used for downstream analysis.

### SnRNA-seq data analysis

The “Seurat” R software package (version 4.4.0) was used for quality control and integration. To filter out poor quality data, we first removed genes covered by fewer than three cells. Then, we filtered out cells expressing <500 or >5000 genes and containing <400 or >25,000 UMIs to exclude barcodes associated with empty partitions or double units. Doublets and multiplets were removed using scDBlFinder with default parameters. Cells with more than 15% mitochondria were also removed. The integrated analysis in the Seurat v4 function “IntegrateData” was used to integrate and embed individual cells from different individuals into a shared low-dimensional space. When the integral matrix was generated, an unsupervised graph-based clustering algorithm was used to cluster individual cells based on their expression, which was implemented in Seurat with default parameters. The “NormalizeData” function with default parameters was used to normalize the UMI count matrix. To transform the normalized gene expression matrix, the natural logarithm was carried out, and 2,000 highly variable genes were identified by “FindVariableFeatures” function with “vst” method. All the cell types were then clustered through these 2000 variable genes. After regressing the UMI-counts, 20 principal component analyses (PCA) were applied to the dataset to reduce dimensionality. The first round of clustering was performed using the “FindClusters” function on 20 PCs with a resolution of 1.2, each of which was annotated with known markers. The Uniform Manifold Approximation and Projection (UMAP) method was used for nonlinear dimensionality reduction. Each cluster was characterized using the “FindAllMarkers” procedure in Seurat, which identified the marker based on the average expressed log fold change (FC). The Wilcoxon Rank-Sum test by default was used. Characteristic genes and known lineage-specific markers were used for clustering annotation.

### Statistical analysis

R software (version 4.0.4) was used for statistical analyses. Statistical tests were selected based on data distribution and its variability. Student's t test,

Mann-Whitney U test and Kruskal-Wallis test were used to assess statistical significance. The *P*-value of DEPs was adjusted with Benjamini and Hochberg FDR method. The correlation was calculated with Pearson's coefficient method. Statistical significance was defined as  $P < 0.05$ .

## DATA AVAILABILITY

All data generated by this study have been deposited in the Chinese national genomics data center (<https://ngdc.cnbc.ac.cn>), under accession number NGDC: OMIX008849. All codes used for association studies are available on request.

## REFERENCES

- Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74:229–63.
- Harbeck N, Penault-Llorca F, Cortes J, Gnani M, Houssami N, Poortmans P, et al. Breast cancer. *Nat Rev Dis Prim*. 2019;5:66.
- Liang Y, Zhang H, Song X, Yang Q. Metastatic heterogeneity of breast cancer: Molecular mechanism and potential therapeutic targets. *Semin Cancer Biol*. 2020;60:14–27.
- Wang R, Zhu Y, Liu X, Liao X, He J, Niu L. The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC Cancer*. 2019;19:1091.
- Ye F, Dewanjee S, Li Y, Jha NK, Chen ZS, Kumar A, et al. Advancements in clinical aspects of targeted therapy and immunotherapy in breast cancer. *Mol Cancer*. 2023;22:105.
- Bao X, Li Q, Chen D, Dai X, Liu C, Tian W, et al. A multiomics analysis-assisted deep learning model identifies a macrophage-oriented module as a potential therapeutic target in colorectal cancer. *Cell Rep Med*. 2024;5:101399.
- Bao X, Wang D, Dai X, Liu C, Zhang H, Jin Y, et al. An immunometabolism subtyping system identifies S100A9+ macrophage as an immune therapeutic target in colorectal cancer based on multiomics analysis. *Cell Rep Med*. 2023;4:100987.
- Liu L, Zhang R, Deng J, Dai X, Zhu X, Fu Q, et al. Construction of TME and Identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. *Cancer Immunol, Immunother*. 2022;71:121–36.
- Bao X, Li Q, Chen J, Chen D, Ye C, Dai X, et al. Molecular subgroups of intrahepatic cholangiocarcinoma discovered by single-cell RNA sequencing-assisted multiomics analysis. *Cancer Immunol Res*. 2022;10:811–28.
- Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*. 2021;53:1334–47.
- Wang XQ, Danenberg E, Huang CS, Egle D, Callari M, Bermejo B, et al. Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature*. 2023;621:868–76.
- Liu YM, Ge JY, Chen YF, Liu T, Chen L, Liu CC, et al. Combined Single-Cell and Spatial Transcriptomics Reveal the Metabolic Evolvement of Breast Cancer during Early Dissemination. *Adv Sci*. 2023;10:e2205395.
- Zou Y, Ye F, Kong Y, Hu X, Deng X, Xie J, et al. The Single-Cell Landscape of Intratumoral Heterogeneity and The Immunosuppressive Microenvironment in Liver and Brain Metastases of Breast Cancer. *Adv Sci*. 2023;10:e2203699.
- Cords L, Tietscher S, Anzeneder T, Langwieder C, Rees M, de Souza N, et al. Cancer-associated fibroblast classification in single-cell and spatial proteomics data. *Nat Commun*. 2023;14:4294.
- Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, et al. The single-cell pathology landscape of breast cancer. *Nature*. 2020;578:615–20.
- Goodin DA, Chau E, Zheng J, O'Connell C, Tiwari A, Xu Y, et al. Characterization of the Breast Cancer Liver Metastasis Microenvironment via Machine Learning Analysis of the Primary Tumor Microenvironment. *Cancer Res Commun*. 2024;4:2846–57.
- Kuett L, Bollhagen A, Tietscher S, Sobottka B, Eling N, Varga Z, et al. Distant Metastases of Breast Cancer Resemble Primary Tumors in Cancer Cell Composition but Differ in Immune Cell Phenotypes. *Cancer Res*. 2025;85:15–31.
- Medeiros B, Allan AL. Molecular Mechanisms of Breast Cancer Metastasis to the Lung: Clinical and Experimental Perspectives. *Int J Mol Sci*. 2019;20:2272.
- Bai Z, Zhang D, Gao Y, Tao B, Zhang D, Bao S, et al. Spatially exploring RNA biology in archival formalin-fixed paraffin-embedded tissues. *Cell*. 2024;187:6760–79.e6724.
- Hu B, He R, Pang K, Wang G, Wang N, Zhu W, et al. High-resolution spatially resolved proteomics of complex tissues based on microfluidics and transfer learning. *Cell*. 2025;188:734–48.e722.
- Du J, Zhang J, Wang L, Wang X, Zhao Y, Lu J, et al. Selective oxidative protection leads to tissue topological changes orchestrated by macrophage during ulcerative colitis. *Nat Commun*. 2023;14:3675.
- Sun X, Teng X, Liu C, Tian W, Cheng J, Hao S, et al. A Pathologically Friendly Strategy for Determining the Organ-specific Spatial Tumor Microenvironment Topology in Lung Adenocarcinoma Through the Integration of snRandom-seq and Imaging Mass Cytometry. *Adv Sci*. 2024;2308892.
- Liu AY, Roudier MP, True LD. Heterogeneity in primary and metastatic prostate cancer as defined by cell surface CD profile. *Am J Pathol*. 2004;165:1543–56.
- Lemaitre L, Adeniji N, Suresh A, Reguram R, Zhang J, Park J, et al. Spatial analysis reveals targetable macrophage-mediated mechanisms of immune evasion in hepatocellular carcinoma minimal residual disease. *Nat Cancer*. 2024;5:1534–56.
- Li L, Sun C, Sun Y, Dong Z, Wu R, Sun X, et al. Spatially resolved proteomics via tissue expansion. *Nat Commun*. 2022;13:7242.
- Dong Z, Jiang W, Wu C, Chen T, Chen J, Ding X, et al. Spatial proteomics of single cells and organelles on tissue slides using filter-aided expansion proteomics. *Nat Commun*. 2024;15:9378.
- Xu Z, Zhang T, Chen H, Zhu Y, Lv Y, Zhang S, et al. High-throughput single nucleus total RNA sequencing of formalin-fixed paraffin-embedded tissues by snRandom-seq. *Nat Commun*. 2023;14:2734.
- Bao X, Shi R, Zhao T, Wang Y, Anastasov N, Rosemann M, et al. Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels tumour heterogeneity plus M2-like tumour-associated macrophage infiltration and aggressiveness in TNBC. *Cancer Immunol, Immunother*. 2021;70:189–202.
- Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate breast cancer metastasis to lung. *Nature*. 2005;436:518–24.
- Liu C, Xie J, Lin B, Tian W, Wu Y, Xin S, et al. Pan-Cancer Single-Cell and Spatial-Resolved Profiling Reveals the Immunosuppressive Role of APOE+ Macrophages in Immune Checkpoint Inhibitor Therapy. *Adv Sci*. 2024;2401061.
- Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discov*. 2022;12:134–53.
- Zou Y, Hu X, Zheng S, Yang A, Li X, Tang H, et al. Discordance of immunotherapy response predictive biomarkers between primary lesions and paired metastases in tumours: A systematic review and meta-analysis. *EBioMedicine*. 2021;63:103137.
- Tang Y, Tian W, Zheng S, Zou Y, Xie J, Zhang J, et al. Dissection of FOXO1-induced LYPLAL1-DT impeding triple-negative breast cancer progression via mediating hnRNPK/β-catenin complex. *Research*. 2023;6:0289.
- Chen J, Yao Q, Huang M, Wang B, Zhang J, Wang T, et al. A randomized Phase III trial of neoadjuvant recombinant human endostatin, docetaxel and epirubicin as first-line therapy for patients with breast cancer (CBCRT01). *Int J Cancer*. 2018;142:2130–8.
- Zhang X, Zhang Z, Cao M, Liu B, Mori M, Luoh SW, et al. A Randomized Parallel Controlled Phase II Trial of Recombinant Human Endostatin Added to Neoadjuvant Chemotherapy for Stage III Breast Cancer. *Clin Breast Cancer*. 2020;20:291–9.e293.
- Li Q, Wang Y, Jia W, Deng H, Li G, Deng W, et al. Low-Dose Anti-Angiogenic Therapy Sensitizes Breast Cancer to PD-1 Blockade. *Clin Cancer Res*. 2020;26:1712–24.
- Sun X, Teng X, Liu C, Tian W, Cheng J, Hao S, et al. A Pathologically Friendly Strategy for Determining the Organ-specific Spatial Tumor Microenvironment Topology in Lung Adenocarcinoma Through the Integration of snRandom-seq and Imaging Mass Cytometry. *Adv Sci*. 2024;11:e2308892.
- Cheung M, Campbell JJ, Whitby L, Thomas RJ, Braybrook J, Petzing J. Current trends in flow cytometry automated data analysis software. *Cytometry A*. 2021;99:1007–21.
- Windhager J, Zanotelli VRT, Schulz D, Meyer L, Daniel M, Bodenmiller B, et al. An end-to-end workflow for multiplexed image processing and analysis. *Nat Protoc*. 2023;18:3565–613.
- Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med*. 2018;24:978–85.
- Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*. 2018;564:268–72.
- Meier F, Brunner AD, Frank M, Ha A, Bludau I, Voytk E, et al. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods*. 2020;17:1229–36.
- Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods*. 2017;14:513–20.
- Yu F, Haynes SE, Teo GC, Avtonomov DM, Polasky DA, Nesvizhskii AI. Fast Quantitative Analysis of timsTOF PASEF Data with MSFragger and IonQuant. *Mol Cell Proteom*. 2020;19:1575–85.
- Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17:41–44.

46. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43:e47.
47. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS.* 2012;16:284–7.
48. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 2013;14:7.
49. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 2008;9:559.
50. Taube JM, Akturk G, Angelo M, Engle EL, Gnjatic S, Greenbaum S, et al. The Society for Immunotherapy in Cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. *J Immunother Cancer.* 2020;8.

## ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China 82101830 (to XB), 82202875 (to XS), 82373428 (to WF), 52233013 (to WF), and 82102817 (to X.Dai) and by Natural Science Foundation of Zhejiang Province LY23H160013 (to XB), LQ23H160029 (to XS), and the Zhejiang Province Public Welfare Technology Application Research Project No. LGF21H030009 (to JC) and by the Major Scientific Project of Zhejiang Province 2023C03061 (to WF). We would like to express our gratitude to professor Tiannan Guo and Dr. Zhen Dong for their help in performing the spatial proteomics. We also would like to acknowledge M20 Genomics for their help in VITA high-throughput single-cell transcriptome experiments for FFPE samples.

## AUTHOR CONTRIBUTIONS

YG, B.Li, YJ and JC conducted the experiments, analyzed multi-omics data, and drafted the original manuscript. WT, LY, LH, SX, B.Lin, CL, JZ and HZ provided support

for laboratory experiments and data analysis. XS, JX, X.Deng, X.Dai, LL, YZ and PZ contributed to sample collection and supervised the clinical data evaluation. XB, WF and GY conceived the project, designed the experiments, and developed the bioinformatics analysis workflow. All authors read and approved the final manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41388-025-03388-y>.

**Correspondence** and requests for materials should be addressed to Guangchuan Yu, Weijia Fang or Xuanwen Bao.

**Reprints and permission 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.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

<sup>1</sup>Department of Medical Oncology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310003, China. <sup>2</sup>State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310003, China. <sup>3</sup>Changzhou Third People's Hospital, Changzhou Medical Center, Nanjing Medical University, Changzhou, Jiangsu 213001, China. <sup>4</sup>Department of Medical Pathology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310003, China. <sup>5</sup>Department of Genetics, Yale School of Medicine, New Haven, USA. <sup>6</sup>Innovation Centre for Information, Binjiang Institute of Zhejiang University, Hangzhou, Zhejiang 310053, China. <sup>7</sup>School of Life Science and Technology, China Pharmaceutical University, Nanjing, Jiangsu 210009, China. <sup>8</sup>Cancer Center, Department of Radiation Oncology, Zhejiang Provincial People's Hospital (Affiliated People's Hospital), Hangzhou Medical College, Hangzhou, Zhejiang 310014, China. <sup>9</sup>State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China. <sup>10</sup>National Key Laboratory of Advanced Drug Delivery and Release Systems, Zhejiang University, Hangzhou 310058, China. <sup>11</sup>Department of Bioinformatics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong 510515, China. <sup>12</sup>These authors contributed equally: Yang Gao, Bin Li, Yuzhi Jin, Jinlin Cheng. ✉email: gcyu1@smu.edu.cn; weijiafang@zju.edu.cn; xuanwen.bao@zju.edu.cn