communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-026-09915-1

Single cell and spatial sequencing analysis of cancer associated fibroblasts in the brain metastasis tumor microenvironment

Thomas Simon[1], David N. Buckley[1], Zeyi Yang[1], Chikako Matsuba[1], Ben Y. Tew[1], Gerald C. Gooden[1], Kyle Hurth[2], Steven A. Toms[3], David D. Tran[4,5], Gabriel Zada[4,5], Matthew P. Salomon[1] & Bodour Salhia[1,5]

Brain metastasis (BM) remains largely incurable. Cancer-associated fibroblasts (CAFs) can either support or inhibit tumor growth in the tumor microenvironment (TME), yet their role in BM is not well described. In this study we define four transcriptionally distinct CAF subpopulations using single-cell and spatial sequencing of human BM tissues. The four CAF subpopulations we describe are termed extracellular matrix (*ECM*), *immune*, *contractile*, or *neural* CAFs, and each subpopulation shows distinct spatial distributions within the BM TME. Further analyses reveal that BM CAFs engage extensively in cell-cell communication and adopt distinct cell states, including an *ECM* CAF cell state marked by high levels of immunoglobulin superfamily containing leucine rich repeat expression (ISLR-CAFs). Functionally, ISLR-CAFs reduce BM tumor cell viability in vitro, consistent with a tumor-inhibitory role. These findings highlight the heterogeneity of CAFs in BM, emphasizing the importance of understanding stromal contributions in the underlying biology of BM.

Each year, nearly 300,000 people in the United States are diagnosed with brain metastasis (BM)—a largely incurable disease with a mean two-year overall survival rate below 2%[1]. Current National Comprehensive Cancer Network guidelines recommend combinations of surgery, radiosurgery, whole-brain radiotherapy and/or systemic therapy[2]. Nevertheless, the effectiveness of these therapeutic options remains suboptimal, and most patients ultimately succumb to disease[3].

BM tumors contain a complex tumor microenvironment (TME) comprised of a meshwork of extracellular matrix (ECM) proteins and a myriad of cell types, including endothelial cells, immune cells, and cancer-associated fibroblasts (CAFs)[4,5]. Tight crosstalk between cancer cells and the TME is essential to cancer properties such as invasion, metastasis, progression, and therapeutic resistance[5]. A deeper understanding of these interactions could enable the development of more effective strategies against BM.

CAFs are among the most prominent cell types within the TME, where they have been widely recognized as tumor-promoting[5]. However, emerging evidence from our group and others has pointed to the existence of tumor-inhibitory CAFs, along with many studies unable to prove the benefit of widely targeting CAFs as an anti-tumor strategy[4,6,7]. Our previous work showed that implanting a BM patient-derived CAF cell line with a BM patient-derived tumor cell (PDC) line reduced tumor growth and prolonged survival in mice with CAF admixed tumors[4]. These data indicate that BM CAFs are molecularly and functionally heterogeneous[4,5]. Yet, BM CAFs still need to be fully characterized.

Previously, we analyzed single cell (sc) transcriptomic data of CAFs in the TME of various solid tumors (breast, pancreatic ductal adenocarcinoma (PDAC), melanoma) and defined three subpopulations based on their gene expression patterns: (1) *desmoplastic/ECM* CAFs (properties of ECM remodeling and desmoplasia); (2) *immune* CAFs (gene expression patterns related to regulation of immune response); and (3) *contractile* CAFs (gene expression indicating properties related to cell contraction)[5]. To what extent these CAF subpopulations or others exist in the TME of BM and how they impact disease progression is yet to be evaluated.

In this study, we performed single-cell multiome sequencing (scMultiome-seq; RNA and ATAC) on nine human BM tissue samples. Our analysis revealed that the BM TME is a heterogeneous milieu comprising tumor cells, endothelial cells, immune cells, neurons, and CAFs. We highlighted the diverse CAF subpopulations within BM and their potential roles in shaping the TME. We confirmed the presence of *ECM*, *contractile*, and

[1]Department of Cancer Biology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [2]Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [3]Department of Neurosurgery and Medicine, Brown University Health, The Warren Alpert Medical School of Brown University, Providence, RI, USA. [4]Department of Neurosurgery, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [5]Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. e-mail: salhia@usc.edu

immune CAFs and defined a fourth subgroup, which we named *neural* CAFs. Spatial transcriptomics on four sc-cell–matched human BM tissue samples revealed the distribution and organization of CAF subpopulations, providing insight into how tissue architecture shapes interactions within the BM TME. Subclustering refined CAF functional states and revealed a subset of *ECM* CAFs marked by ISLR expression (ISLR-CAFs). Functional assays showed that ISLR-CAFs inhibited BM tumor cell growth in vitro. Together, this work provides a cellular, molecular, and functional characterization of BM CAFs, uncovering insights into tumor-inhibitory CAF subsets.

Results

Cell type distribution in the BM TME

ScMultiome-seq was performed on nine human BM tissue samples to evaluate cellular heterogeneity within the TME. Seven samples were from lung-to-brain metastases (LBM) and two from breast-to-brain metastases (BBM) (Fig. 1A, B, Supplementary Fig. 1A–E, and Table 1). Single-cell gene-expression (scRNA-seq) data were visualized using uniform manifold approximation and projection (UMAP) (Fig. 1A–D and Supplementary Fig. 1A–C, E)[8]. Cell-type prediction with scPred identified 14 distinct cell types (Fig. 1C)[9].

Copy number analysis with CopyKAT confirmed that breast- and lung-derived tumor cells were highly aneuploid (84% and 63%, respectively; Supplementary Fig. 1D and Supplementary Data 1). Interestingly, non-cancerous cells clustered according to cell type irrespective of BM cell of origin (breast or lung) (Fig. 1B, C). Conversely, cancer cells clustered according to their tissue sample of origin (Fig. 1B, C and Supplementary Fig. 1E).

Marker gene expression confirmed cell-type assignments for the six most abundant populations: *SFTPB* (lung cancer cells), *ERBB2* (breast cancer cells), *PTPRG* (macrophages), *CD38* (B cells), *PECAM1* (endothelial cells) and *FAP* (CAFs) (Fig. 1C, D)[10–15].

Cell-type distribution varied highly across patient samples (Fig. 1E and Supplementary Data 2). On average, tumor cells were the most abundant population, constituting 59% of cells in breast-to-brain and 76% of cells in lung-to-brain metastases. Non-cancer cell counts were the highest in BBM-USC-5-1 (52%) and LBM-USC-2-2 (51%), while LBM-RIH-1 (12%) and LBM-RIH-2 (9%) had the fewest. Among non-tumor cells, CAFs (8.5%), macrophages (6.5%) and endothelial cells (4.2%) were the most abundant overall. LBM-USC-2-2 had the highest levels of CAFs (22.5%) and LBM-RIH-5 the lowest (0.8%). Macrophages varied from 0.4% in LBM-RIH-1 to 19.7% in LBM-RIH-3, and endothelial cells from none in LBM-RIH-3 to in 8.7% in BBM-USC-5-2. Lymphocytes also varied by sample, with BBM-USC-5-1 containing high levels of B cells (16.9%) and T cells (6.6%), whereas others were below 5%. Neurons were present mostly in LBM-RIH-5 (4.8%) and LBM-USC-2-2 (6%). Oligodendrocytes were present in LBM-RIH-1 (5.8%) and LBM-RIH-5 (2.9%). Smooth muscle cells (SMCs) were only detected in BBM-USC-5-1 (3%) and BBM-USC-5-2 (1.6%).

Differences in cell type distribution in the BM TME were also observed between CNS regions, with non-tumor cells, including CAFs, tending to be more abundant in non-frontal-lobe BMs (Fig. 1E).

Altogether, the BM TME was highly heterogeneous in terms of cell type composition between patient samples. CAFs were among the most abundant non-tumor populations, indicative of a central role in BM pathology.

Identification of BM CAF subpopulations

To examine CAF heterogeneity in human BM, we analyzed CAF scRNA-seq profiles separately from other TME cell types. CAFs expressed higher levels of established markers, including *FAP, ACTA2, COL1A1, PDGFRB*, and *POSTN*, compared with other BM cell types (Supplementary Fig. 2A, B). Clustering analysis identified four transcriptionally distinct CAF subpopulations: *ECM*, *immune*, *contractile*, and a previously unrecognized *neural* CAF subpopulation (Fig. 2A, B). Notably, CAFs did not segregate according to tissue sample of origin (Fig. 2A and Supplementary Fig. 2C).

Differentially expressed genes (DEGs), identified between the four CAF clusters, and Ingenuity Pathway Analysis (IPA) were used to define the phenotype and putative function of each cluster (Fig. 2C and Supplementary Fig. 3)[5]. Cluster 1 (*ECM* CAFs) expressed genes involved in ECM composition, including *POSTN*, *LAMA2* and collagen *COL11A1* (Fig. 2B, C). Other ECM genes, such as collagens *COL8A1*, *COL1A1*, in addition to *DCN*, *LUM*, *FN1*, and *VCAN* were also highly expressed. CAF cluster 1 also expressed high levels of *PDGFRA* (Supplementary Data 3)[5]. IPA revealed activation of pathways such as "Wound Healing" and "GP6 Signaling" pathways (Supplementary Fig. 3). Altogether, *ECM* CAFs may play a role in producing and organizing the ECM within the BM TME.

Custer 2 (*Immune* CAFs) was highly heterogeneous with many subclusters and elevated expression of inflammation mediators and immune response genes such as *VWF* or *ST6GALNAC3* (Fig. 2B, C). Other highly expressed genes included components of the Human Leukocyte Antigen system (*HLA-A*, *HLA-C*, and *HLA-E*), immunoglobulin superfamily gene *ALCAM*, regulators of the complement system (*CD46* and *CD59*), and the monocyte/macrophage marker *CD163* (Supplementary Data 3)[5]. IPA analysis demonstrated enrichment of "Thrombin Signaling", "Neutrophil Extracellular Trap Signaling" or "CXCR4 Signaling" pathways (Supplementary Fig. 3).

Cluster 3 (*contractile* CAFs) showed strong expression of *PRKG1, ACTA2, and MYH11* (cell contraction) and *TRPC6* (regulator of actin stress fiber formation) (Fig. 2C and Supplementary Data 3)[5,16,17]. *Contractile* CAFs also expressed several pericyte markers, such as *RGS5*, *MCAM*, and *CSPG4* (Supplementary Data 3). Additionally, IPA revealed activation of pathways linked to mechanotransduction, tissue contraction or cytoskeleton changes, such as "Integrin", "Cardiac β-adrenergic" and "Paxillin" signaling pathways (Supplementary Fig. 3). Taken together, *contractile* CAFs may be involved in contractile force generation within the ECM.

Since genes related to neural functions, such as *LSAMP* and *NRCAM* were elevated in cluster 4, we named them *neural* CAFs (Fig. 2C). Other genes highly expressed in this cluster included neuronal marker *NCAM2* and glial marker *GFAP* (Supplementary Data. 3)[18,19]. IPA identified activation of "Synaptogenesis", "Glutamate Receptor", and "CREB Signaling in Neurons" pathways (Supplementary Fig. 3). These data point to the presence of a brain-specific subpopulation of CAFs that may be involved in regulating the neural compartment of the BM TME.

To further define CAF subpopulation regulatory programs, we examined chromatin accessibility profiles using scATAC-seq data (Supplementary Fig. 4 and Supplementary Data 4). Accessible transcription factor (TF) binding motifs associated with each CAF subpopulation were determined. TF motifs *ZBTB7B* and *ZBTB7C*, regulators of type I collagen and matrix metalloproteinase gene expression, were identified as most enriched in the *ECM* subpopulation (Supplementary Fig. 4, Supplementary Data 4)[20,21]. Motifs for TFs of the E-protein family (*TCF3* and *TCF4)*, linked to lymphocyte development and neuroinflammation, were observed in *immune* CAFs (Supplementary Fig. 4 and Supplementary Data 4)[22]. TF motifs contributing to muscle cell or SMC differentiation and function, like *MEF2A*, *MEF2B*, *MEF2C*, and *MEF2D*, were identified in *contractile* CAFs (Supplementary Fig. 4 and Supplementary Data 4)[23]. Members of the AP-1 TF family, recently linked to myofibroblastic CAFs and pro-tumorigenic properties of CAFs, were also represented in *contractile* CAFs, including *FOS::JUN(var.2)* and *FOSL2::JUN(var.2)* (Supplementary Data 4)[24]. TFs associated with neurodevelopment were observed in *neural* CAFs, including *EMX2* and members of the SOX family of TFs (associated with metastasis and immunosuppressive TME) like *SOX8* or *SOX15* (Supplementary Fig. 4 and Supplementary Data 4)[25]. Taken together, our scATAC-seq data further support the transcriptional programs for each CAF subpopulation identified by scRNA-seq (Fig. 2).

The distribution of CAF subpopulations varied across patient samples but followed a consistent pattern. On average, *ECM* CAFs were the most abundant among all BM tissue samples (42%), followed by *immune* CAFs (30%), *contractile* CAFs (22%) and *neural* CAFs (6%) (Fig. 2D and Supplementary Data 5). *ECM* and *immune* CAFs tended to be more abundant in frontal-lobe BMs, *contractile* CAFs were more prevalent in non-frontal regions, and *neural* CAFs tended to be more abundant in the cerebellum (Fig. 2D and Supplementary Data 5).

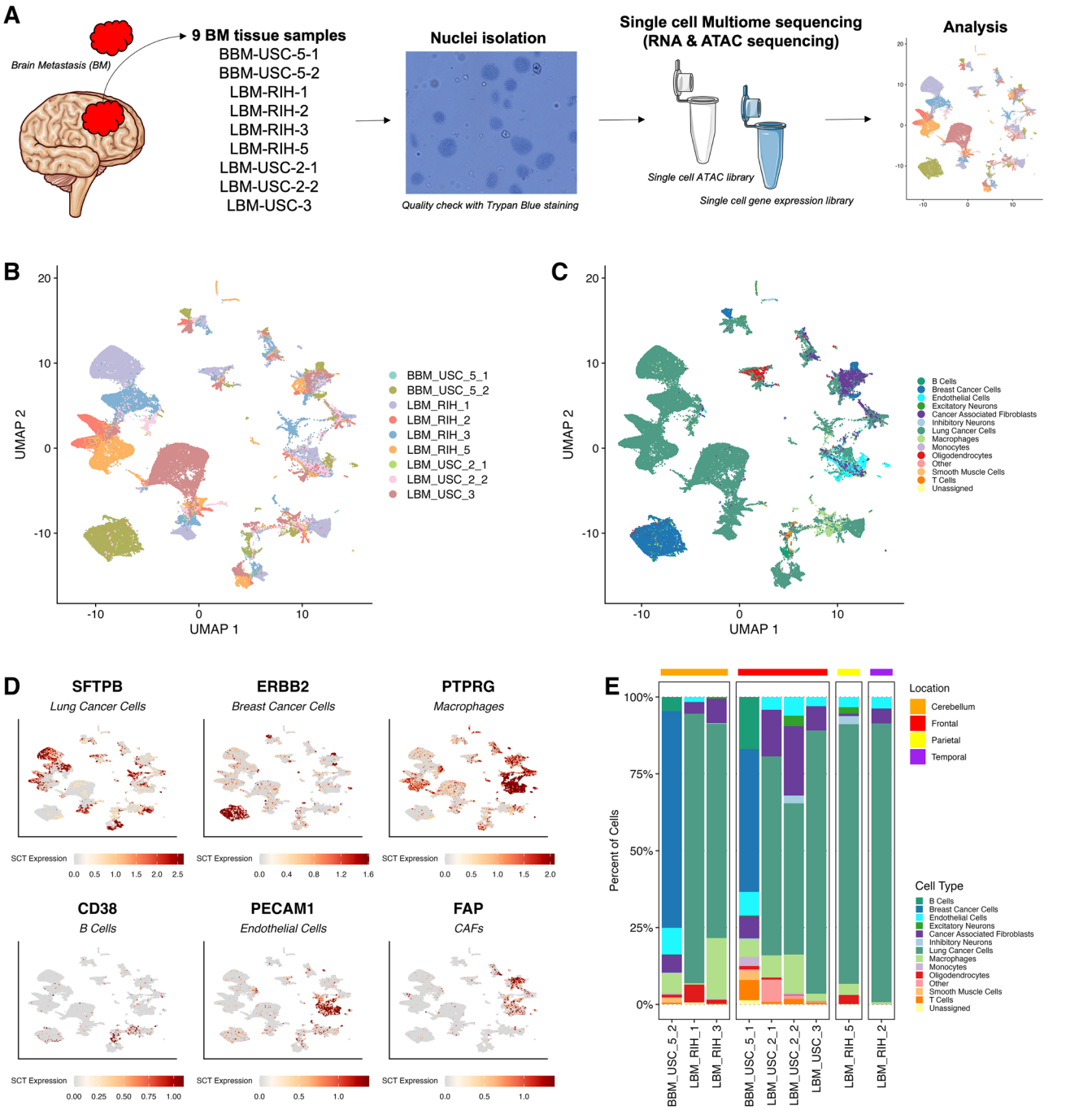


Fig. 1 | Description of the BM TME by scMultiome-seq. A General pipeline of the scMultiome-seq analysis of nine BM patient-derived tissue samples (Brain and microtube images provided by Servier Medical Art (https://smart.servier.com), licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).); **B** UMAP visualization of single cells from the scRNA-seq analysis of BM patient-derived tissues with different colors representing tissue samples of origin (BBM-USC-5-1: blue green, BBM-USC-5-2: green, LBM-RIH-1: purple, LBM-RIH-2: red, LBM-RIH-3: blue, LBM-RIH-5: orange, LBM-USC-2-1: light green, LBM-USC-2-2: pink, LBM-USC-3: hearth red); **C** UMAP visualization of the cell type prediction (scPred) from the scRNA-seq analysis of BM patient-derived tissues with different colors representing predicted cell types (B cells: teal, breast cancer cells: blue, endothelial cells: bright blue, excitatory neurons: green, cancer associated fibroblasts: purple, inhibitory neurons: blue gray, lung cancer cells: dark blue green, macrophages: light green, monocytes: light purple, oligodendrocytes: red, other: pink, smooth muscle cells: light orange, T cells: orange, unassigned: yellow); **D** UMAP visualizations of the expression of markers specific for shown cell types (Expression scale: gray to red; CAF Cancer Associated Fibroblast, CD38 Cluster of Differentiation 38, ERBB2 Erb-B2 Receptor Tyrosine Kinase 2, FAP Fibroblast Activation Protein, PECAM1 Platelet Endothelial Cell Adhesion Molecule-1, PTPRG Protein Tyrosine Phosphatase Receptor Type Gamma, SFTPB Surfactant Protein B); **E** Distribution of cell types in the TME of each BM tissue. (Cell type = B cells: teal, breast cancer cells: blue, endothelial cells: bright blue, excitatory neurons: green, cancer associated fibroblasts: purple, inhibitory neurons: blue gray, lung cancer cells: dark blue green, macrophages: light green, monocytes: light purple, oligodendrocytes: red, other: pink, smooth muscle cells: light orange, T cells: orange, unassigned: yellow; location = cerebellum: orange, frontal: red, parietal: yellow, temporal: purple). Image(s) provided by Servier Medical Art (https://smart.servier.com), licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/) (BBM Breast-to-Brain Metastasis, LBM Lung-to-Brain Metastasis, RIH Rhode Island Hospital, USC University of Southern California).

Table. 1 | Clinical information for BM patient-derived tissues analyzed by scMultiome-seq

Case ID	Sample ID	Source	Institute	Location	Sex	Age	Origin	Pre-op radiation? Y/N	Pre-op chemo? Y/N	Molecular details
LBM-USC-2	LBM-USC-2-1	Tumor	USC	Bi-frontal - Left	M	76	Lung - large cell neuroendocrine carcinomas	N	N	Synaptophysin+ Low TTF1 Ki67 70–90% tumor cells
	LBM-USC-2-2	Tumor	USC	Bi-frontal - Right	M	76	Lung - large cell neuroendocrine carcinomas	N	N	Synaptophysin+ Low TTF1 Ki67 70–90% tumor cells
LBM-USC-3	LBM-USC-3	Tumor	USC	Right frontal	F	58	Lung - adenocarcinoma	N	N	AE1/3+ (strong diffuse) TTF1+ (subset of cells) Napsin-A+ (subset of cells) P40- Mucicarmine highlights mucin
BBM-USC-5	BBM-USC-5-1	Tumor	USC	Right frontal	F	44	Breast	Y	Y (6 cycles)	HER2 + AE1/3- ER- PR-
	BBM-USC-5-2	Tumor	USC	Cerebellar	F	44	Breast	Y	Y (6 cycles)	HER2 + AE1/3- ER- PR-
LBM-RIH-1	LBM-RIH-1	Tumor	RIH	Left cerebellum	M	55	Non-small cell lung adenocarcinoma, poorly differentiated	N	N	CK20- CK7 + TTF1+
LBM-RIH-2	LBM-RIH-2	Tumor	RIH	Right temporal lobe	M	64	Non-small cell lung adenocarcinoma, poorly differentiated	N	N	CK20- CK7 + TTF1+
LBM-RIH-3	LBM-RIH3	Tumor	RIH	Left cerebellum	F	68	Non-small cell lung adenocarcinoma	N	N	PDL1- CK7 + TIF-1+ NapsinA+ CK20- GATA-3-, ATRX, DNMT3A, STK11, TP53 mutations identified
LBM-RIH-5	LBM-RIH-5	Tumor	RIH	Right mesial parietal lobe	F	65	Non-small cell lung adenocarcinoma, poorly differentiated	N	N	CK7 + CK20- TIF-1 + PDL1 + P40- (in tumor cells)

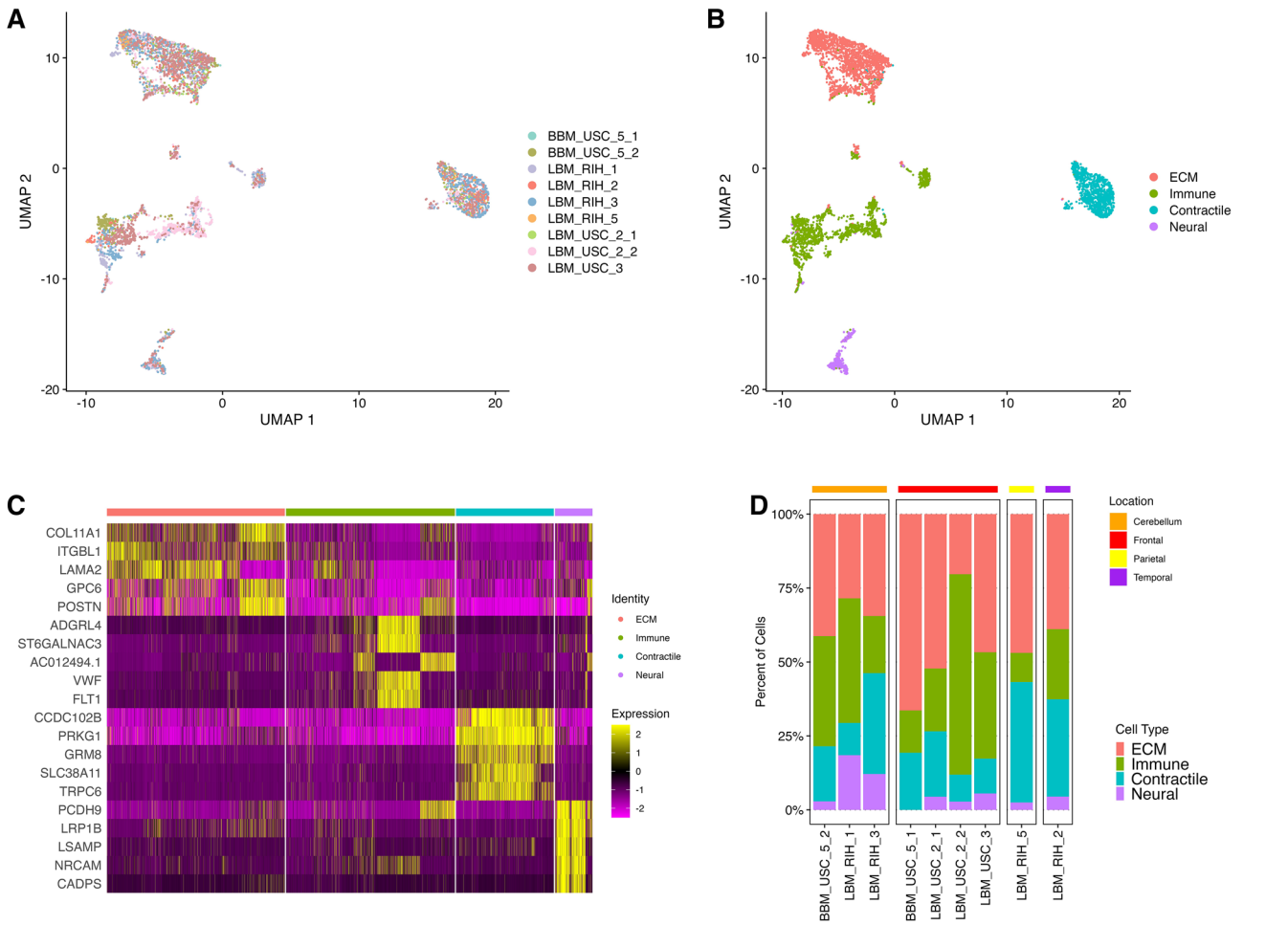
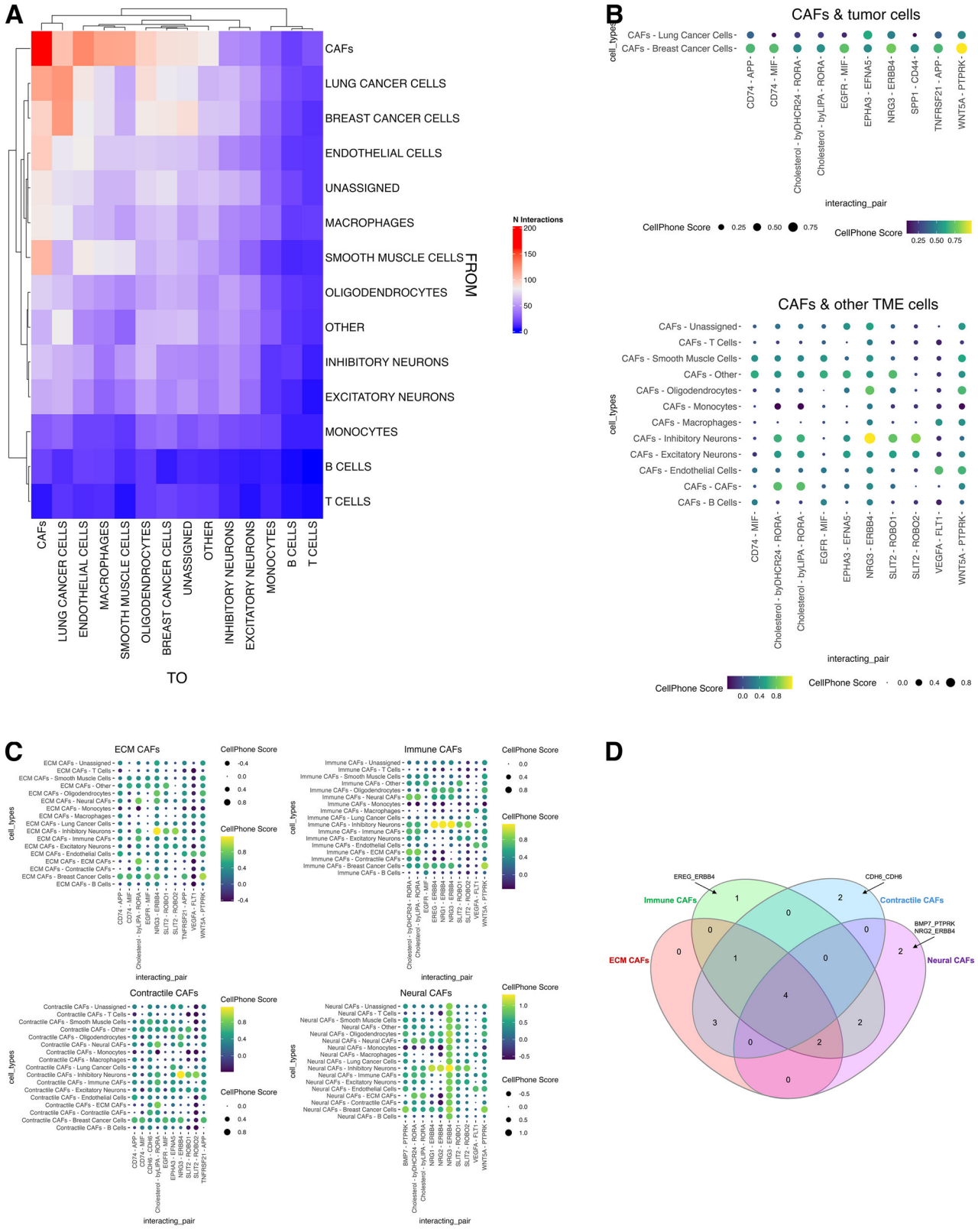


Fig. 2 | Description of CAF subpopulations in the BM TME by scMultiome-seq. **A** UMAP visualization of CAFs from the scRNA-seq analysis of BM patient-derived tissues with different colors representing tissue samples of origin (BBM-USC-5-1: blue green, BBM-USC-5-2: green, LBM-RIH-1: purple, LBM-RIH-2: red, LBM-RIH-3: blue, LBM-RIH-5: orange, LBM-USC-2-1: light green, LBM-USC-2-2: pink, LBM-USC-3: hearth red); **B** UMAP visualization of four CAF subpopulations defined by scRNA-seq analysis of BM patient-derived tissues with different colors representing each subpopulation (*ECM* CAFs: red, *immune* CAFs: green, *contractile* CAFs: blue, *neural* CAFs: purple); **C** Gene expression heatmap of top markers for each CAF subpopulation (Expression scale: purple to yellow; CAF subpopulation = *ECM* CAFs: red, *immune* CAFs: green, *contractile* CAFs: blue, *neural* CAFs: purple; ADGRL4 Adhesion G Protein-Coupled Receptor L4, CADPS Calcium-Dependent Secretion Activator, CCDC102B Coiled-Coil Domain Containing 102B, COL11A1 Collagen type XI Alpha 1 chain, FLT1 Fms Related Receptor Tyrosine Kinase 1, GPC6 Glypican-6, GRM8 Glutamate Metabotropic Receptor 8, ITGBL1 Integrin Beta-Like 1, LAMA2 Laminin subunit alpha 2, LRP1B Low-Density Lipoprotein Receptor-Related Protein 1B, LSAMP Limbic System Associated Membrane Protein, NRCAM Neuronal Cell Adhesion Molecule, PCDH9 Protocadherin 9, POSTN Periostin, PRKG1 Protein Kinase cGMP-Dependent 1, SLC38A11 Solute Carrier Family 38 Member 11, ST6GALNAC3 ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 3, TRPC6 Transient Receptor Potential Canonical 6, VWF Von Willebrand Factor); **D** Distribution of CAF subpopulations in the TME of each analyzed BM tissue (CAF subpopulation = *ECM* CAFs: red, *immune* CAFs: green, *contractile* CAFs: blue, *neural* CAFs: purple; location = cerebellum: orange, frontal: red, parietal: yellow, temporal: purple) (BBM Breast-to-Brain Metastasis, CAF Cancer Associated Fibroblast, ECM Extracellular Matrix, LBM Lung-to-Brain Metastasis, RIH Rhode Island Hospital, USC University of Southern California).

We validated these CAF subpopulation signatures identified in our dataset in three independent external BM scRNA-seq datasets (GSE131907, GSE186344, and GSE234832) (Supplementary Fig. 5)[26–28]. *Contractile* and *ECM* CAFs formed distinct clusters in the GSE131907 and GSE184344 datasets, *immune* CAFs clustered together in the GSE131907 and GSE234832 datasets, and *neural* CAFs were specifically observed in the GSE234832 dataset (Supplementary Fig. 5A). When combining the external datasets, we observed that *contractile* and *ECM* CAFs formed distinct clusters, while *immune* CAFs appeared scattered, consistent with their intrinsic heterogeneity we observed in our own dataset (Fig. 2B and Supplementary Fig. 5B). Expression of genes associated with each CAF subpopulation signature was also confirmed across the datasets, including ITGBL1 (ECM), ADGRL4 (Immune), TRPC6 (Contractile), and NRCAM (neural; Fig. 2C and Supplementary Fig. 5C)[26–28]. Together, these analyses confirm that CAFs in the BM TME are highly heterogeneous and reproducibly segregate into four distinct subpopulations: *ECM*, *immune*, *contractile*, and *neural*[5].

CellPhoneDB analysis places CAFs at the center of the cell-cell communication in the BM TME

We next examined the spectrum of cell–cell interactions involving CAFs and other cell types in the BM TME using CellPhoneDB. This analysis revealed extensive intercellular signaling, identifying 161 CAF–CAF interactions, and hundreds of additional interactions with other stromal and tumor cell types, including oligodendrocytes (166 interactions), breast cancer cells (184), macrophages (200), lung cancer cells (218), SMCs (221) and endothelial cells (225) (Fig. 3A and Supplementary Data 6). The *WNT5A–PTPRK* interaction represented the strongest CAF–cancer cell interaction in BBM, whereas *EPHA3–EFNA5* was the dominant CAF–cancer cell interaction in LBM (Fig. 3B). The *NRG3–ERBB4* and



SLIT2–ROBO1/2 pairs were central to CAF interactions with inhibitory neurons, with the *NRG3–ERBB4* signaling also being enriched between CAFs and oligodendrocytes (Fig. 3B, C). Prominent *VEGFA–FLT1* interactions between CAFs and endothelial cells or macrophages implicated CAFs in promoting neo-angiogenesis and driving the pro-metastatic phenotype of macrophages (Fig. 3B)[29].

We next analyzed communication patterns of each CAF subpopulation with other cell types in the BM TME. CAF subpopulations interacted most frequently with one another, with *immune* and *neural* CAFs showing the highest number of interactions (477 total; Supplementary Fig. 6 and Supplementary Data 7). Among non-tumor populations, *ECM* and *contractile* CAFs interacted most with *SMCs* (195 and 149 interactions,

Fig. 3 | CellPhoneDB analysis of the cell-cell communication in the BM TME. **A** Heatmap representation of the strength of interactions between the shown cell types identified by scRNA-seq analysis of the BM TME (N Interactions scale: blue to red); **B** Dotplot representation of the strength of shown molecular interacting pairs (*x*-axis) between shown communicating cell types (*y*-axis); **C** Dotplot representation of the strength of shown molecular interacting pairs (*x*-axis) between shown communicating cell types (*y*-axis) for each CAF subpopulation in the BM TME (B&C: CellPhone Score scale: blue to yellow); **D** Venn diagram of the top 10 cell–cell communication molecular interacting pairs identified for each CAF subpopulation. Venn diagram was created using InteractiVenn[70] (*ECM* CAFs: red, *immune* CAFs: green, *contractile* CAFs: blue, *neural* CAFs: purple; APP Amyeloid-beta Precursor Protein, BMP7 Bone Morphogenic Protein 7, CAF Cancer Associated Fibroblast, CD44 Cluster of Differentiation 44, CD74 Cluster of Differentiation 74, DHCR24 24-Dehydrocholesterol Reductase, ECM Extracellular Matrix, EFNA5 Ephrin-A5, EGFR Epidermal Growth Factor Receptor, EPHA3 Eph Receptor A3, ERBB4 Erb-b2 Receptor Tyrosine Kinase 4, EREG Epiregulin, LIPA Lipase A, MIF Macrophage Migration Inhibitory Factor, NRG1/2/3 Neuregulin 1/2/3, PTPRK Protein Tyrosine Phosphatase Receptor Type K, ROBO1/2 Roundabout Guidance Receptor 1/2, RORA RAR-related Orphan Receptor Alpha, SLIT2 Slit Guidance Ligand 2, SPP1 Secreted Phosphoprotein 1, TNFRSF21 TNF Receptor Superfamily Member 21, VEGFA Vascular Endothelial Growth Factor A, WNT5A Wnt Family Member 5A).

respectively) and with endothelial cells (194 and 153 interactions, respectively) (Supplementary Fig. 6 and Supplementary Data 7). Notably, one of the top 10 interactions involving *contractile* CAFs was *CDH6–CDH6*, a ligand–receptor pair associated with contraction, motility, and cancer progression, previously reported in SMCs and myofibroblast CAFs. This interaction scored highest between *contractile* CAFs themselves and between *contractile* CAFs and SMCs (Fig. 3C, D)[30].

Immune CAFs exhibited extensive predicted interactions with lung cancer cells (277), breast cancer cells (239), endothelial cells (270), and macrophages (247) (Supplementary Fig. 6, Supplementary Data 7). Among the top ligand–receptor pairs, *EREG–ERBB4*, an interaction associated with immune regulation and tumor-supportive functions, was particularly enriched between *immune* CAFs and *neural* CAFs, oligodendrocytes, and inhibitory neurons (Fig. 3C, D)[31].

Neural CAFs exhibited the highest number of predicted interactions with other cell types in the BM TME, including *immune* CAFs (477), lung (341) and breast cancer cells (275), SMCs (280), endothelial cells (296), and macrophages (273) (Supplementary Fig. 6). They also showed extensive interactions with CNS cells, including oligodendrocytes (256), inhibitory neurons (185), and excitatory neurons (163) (Supplementary Fig. 6 and Supplementary Data 7). Among the top ligand–receptor pairs, *NRG1–ERBB4, NRG2–ERBB4*, and *NRG3–ERBB4* were especially prominent between *neural* CAFs and inhibitory neurons, oligodendrocytes, and breast cancer cells (Fig. 3C, D)[32,33]. Notably, *NRG1–ERBB4* signaling with inhibitory neurons was also enriched in *immune* CAFs (Fig. 3C). Taken together, these analyses indicate that each CAF subpopulation communicates with a specific set of cell types through molecular interactions consistent with their distinct functions within the BM TME.

Spatial organization of CAFs in the BM TME

The spatial organization of BM TME cell types was investigated through spatial transcriptomic profiling of four BM tissue samples (BBM-USC-5-1, LBM-RIH-2, LBM-RIH-3, and LBM-USC-2-1; Fig. 4A). Consistent with our scRNA-seq data, cancer cell and CAF gene-expression signatures were among the most prominent features across samples (Fig. 4A). Cancer cell and CAF signatures were spatially segregated, showing mutually exclusive distributions across all spatial maps (Fig. 4A). Additional signatures corresponding to B cells, macrophages, endothelial cells, and SMCs were also observed. Cancer cell signatures were spatially distinct from SMC and endothelial signatures but co-localized with immune cell signatures (macrophages and B cells; Fig. 4A). In parallel, the mutually exclusive distributions of CAFs and cancer cells were also observed by a board-certified pathologist in an H&E-stained section of sufficient quality from one sample, providing histological evidence for the tissue architecture defined in the corresponding spatial transcriptomics map (Supplementary Fig. 7). Together, these findings support a model in which CAFs occupy stromal regions adjacent to cancer cells, while immune cells infiltrate into cancer-rich areas.

Spatial transcriptomics analysis supported the presence in the BM spatial architecture of the four CAF subpopulations defined in the single-cell dataset (Fig. 4B). *ECM* and *immune* CAF signatures were detected in all four tissues, with the *ECM* CAF signature strongest in BBM-USC-5-1, the *immune* CAF signature dominant in LBM-USC-2-1, and both signatures detected in LBM-RIH-2 (Figs. 4B and 2D). LBM-RIH-3 displayed signatures for all four CAF subpopulations, with *ECM* and *contractile* CAF signatures showing the strongest levels of detection, consistent with the scRNA-seq data (Figs. 4B, 2B–D and Supplementary Data 5). Notably, the *contractile* CAF signature was only detected by spatial transcriptomics in LBM-RIH-3, the same sample that exhibited the highest abundance of *contractile* CAFs in the single-cell analysis.

Spatial maps also revealed subpopulation-specific organization and interactions consistent with CellPhoneDB predictions. In BBM-USC-5-1, *immune* CAF signatures localized at the edges of CAF-rich regions, adjacent to signatures for tumor and infiltrating immune cells such as B cells and macrophages. By contrast, *ECM* CAF signatures localized more centrally within CAF compartments, often near endothelial or SMC signatures (Fig. 4B, C). In LBM-RIH-2 and LBM-RIH-3, tissue spots with macrophage signatures frequently overlapped with *immune* CAFs (Fig. 4A, B). Taken together, these spatial analyses highlight the architectural role of CAF subpopulations within the BM TME and provide spatial validation for the cell–cell interactions inferred from CellPhoneDB analysis.

Four major CAF subpopulations in BMs contain transcriptionally distinct sub-clusters

scRNA-seq analysis revealed that each BM CAF subpopulation could be further subdivided into transcriptionally distinct subclusters (Fig. 5). DEGs were identified for each subcluster, and key markers were used to assign descriptive labels. Gene expression profiles, combined with IPA, indicated distinct functional roles, including subclusters with putative tumor-supportive or tumor-inhibitory properties.

The *immune* CAF subpopulation was the most heterogeneous with 5 distinct subclusters: PCDH9-CAFs, CXCL12-CAFs, RAB3C-CAFs, PLXDC2-CAFs, and MIF-CAFs (Fig. 5 and Supplementary Data 8). PCDH9-CAFs were enriched for immunoregulatory pathways, such as "PD-1, PD-L1 Cancer Immunotherapy", "Neuregulin Signaling", "ERBB Signaling", and "IL-10 Signaling"[34]. In contrast, both CXCL12-CAFs and RAB3C-CAFs showed a pro-inflammatory profile, with CXCL12-CAFs showing elevated expression of *PDGFD* and activation of "IL-8" and "CXCR4" signaling, while RAB3C-CAFs expressed high *VWF* and *ALCAM*, linked to "FAT10 Cancer" and "IL-6" signaling pathways (Supplementary Fig. 8 and Supplementary Data 8). PLXDC2-CAFs were characterized by high expression of *CD163* and activation of "Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes", linking them to regulation of inflammation, EMT and M2 polarization of macrophages, consistent with a tumor-supportive phenotype[35]. MIF-CAFs showed high expression levels of the innate immunity regulator macrophage migration inhibitory factor (MIF) and MHC class I molecules (*HLA-A, B, C*), suggesting possible antigen-presenting roles, along with activation of the "Neutrophil Extracellular Trap Signaling Pathway" (Fig. 5, Supplementary Fig. 8, and Supplementary Data 8).

The *contractile* CAF subpopulation had four subclusters: RGS5-CAFs, MYH11-CAFs, DIAPH3-CAFs, and KCNMA1-CAFs (Fig. 5). RGS5-CAFs had strong expression of pericyte markers *RGS5* and *PDGFRB*, and activation of pathways such as "Inhibition of Angiogenesis by TSP1" and "PTEN Signaling", consistent with a normal pericyte-like phenotype[36]. In contrast, MYH11-CAFs demonstrated a tumor-supportive, myofibroblast-like phenotype, with high *PALLD* and *ACTA2* expression and enrichment

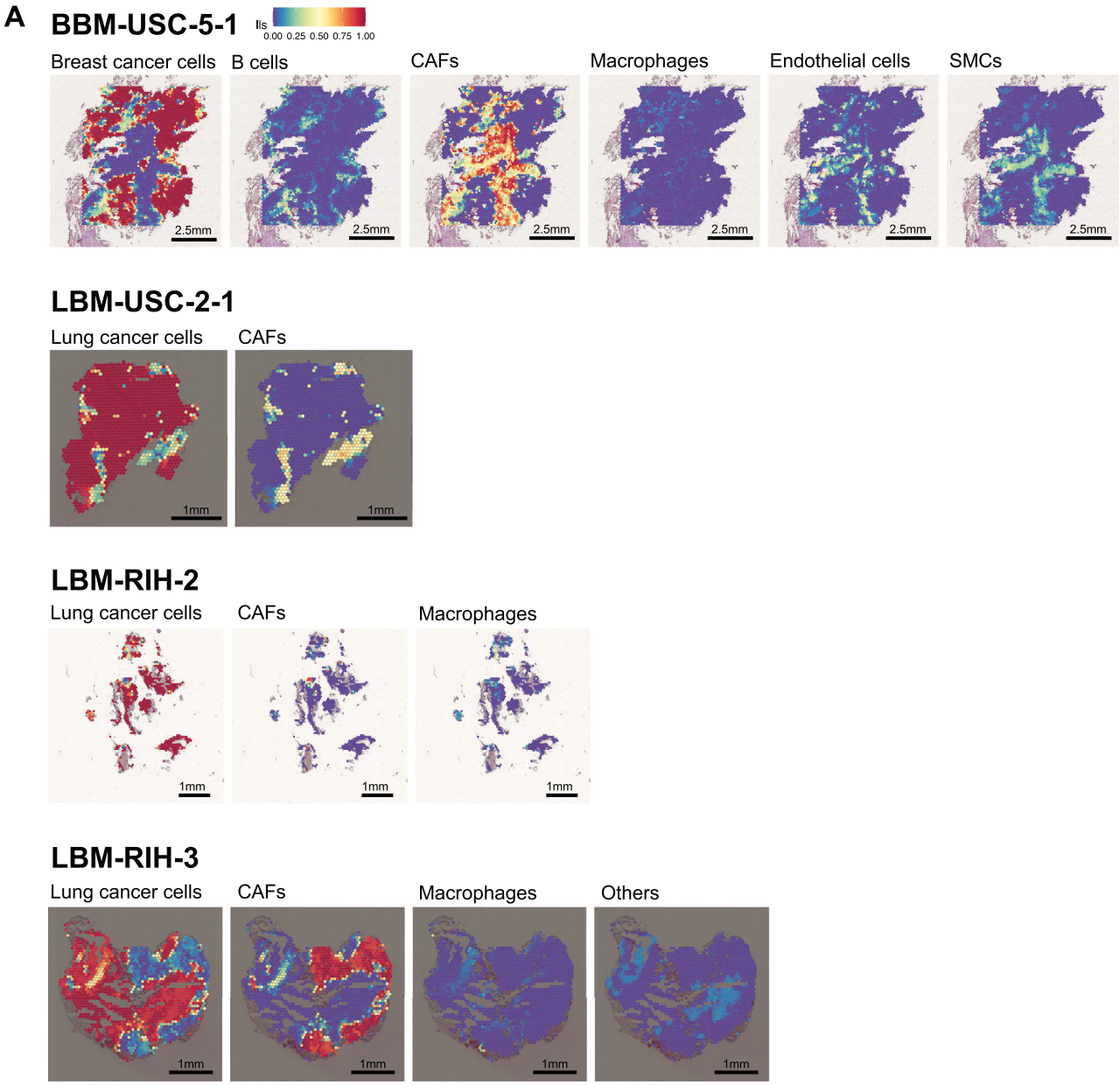


Fig. 4 | Description of the BM TME architecture by spatial transcriptomics. The spatial organization of the BBM-USC-5-1, LBM-USC-2-1, LBM-RIH-2, and LBM-RIH-3 BM patient-derived tissue samples was investigated by Visium spatial gene expression analysis. Visium gene expression spots in each tissue were annotated using gene expression signatures identified through scRNA-seq. Scale bars are shown. **A** Up to six cell types with highest levels of gene expression signature detection are shown in each tissue sample; **B** CAF subpopulations in the spatial architecture of the BM TME (purple arrows indicate signal from the "*neural* CAF" gene expression signature); **C** Description of spatial structures in the BBM-USC-5-1 tissue. Red arrows indicate a structure with co-localization of gene expression signatures for cancer cells, macrophages and *immune* CAFs, while black arrows indicate a structure with co-localization of endothelial cells, SMCs and *ECM*CAFs (Gene expression scale: blue to red, SMC Smooth Muscle Cells) (BBM Breast-to-Brain Metastasis, CAF Cancer Associated Fibroblast, ECM Extracellular Matrix, LBM Lung-to-Brain Metastasis, RIH Rhode Island Hospital, USC University of Southern California).

for "Wound Healing" and "Actin Cytoskeleton" pathways[5]. DIAPH3-CAFs expressed high levels of *DIAPH3* (an actin nucleation and elongation factor during cytokinesis), and DNA topoisomerase IIα *TOP2A*, with enrichment for the "Kinetochore Metaphase Signaling Pathway," suggesting a proliferative subcluster (Supplementary Fig. 9A and Supplementary Data 8)[37]. Cell-cycle score analysis performed on CAF subpopulations confirmed increased G2/M and S-phase marker expression in the *contractile* subpopulation relative to the other three CAF subpopulations (Supplementary Fig. 9B). KCNMA1-CAFs expressed high levels of *KCNMA1* (a regulator of smooth muscle contraction) and presented a DEG pattern that included immune markers *CD163* and *ALCAM*. KCNMA1-CAFs also showed activation of pathways such as "Natural Killer Cell Signaling", consistent with reports on CAFs with myofibroblast characteristics having roles in immunoregulation (Fig. 5, Supplementary Fig. 9, and Supplementary Data 8)[38].

We observed two subclusters in the *neural* CAF subpopulation: KCNIP4-CAFs and CHI3L1-CAFs (Fig. 5). KCNIP4-CAFs had strong expression of ECM-related genes such as *POSTN* and *COL1A1*, along with gene enrichment for signaling pathways like "Wound Healing", supporting an activated fibroblast-like state. CHI3L1-CAFs showed a neural-like profile, with high expression of Chitinase-3-like protein 1, expressed by astrocytes, and gene enrichment for pathways involved in normal neural function, such as "CREB Signaling in Neurons" (Fig. 5, Supplementary Fig. 10, and Supplementary Data 8)[39].

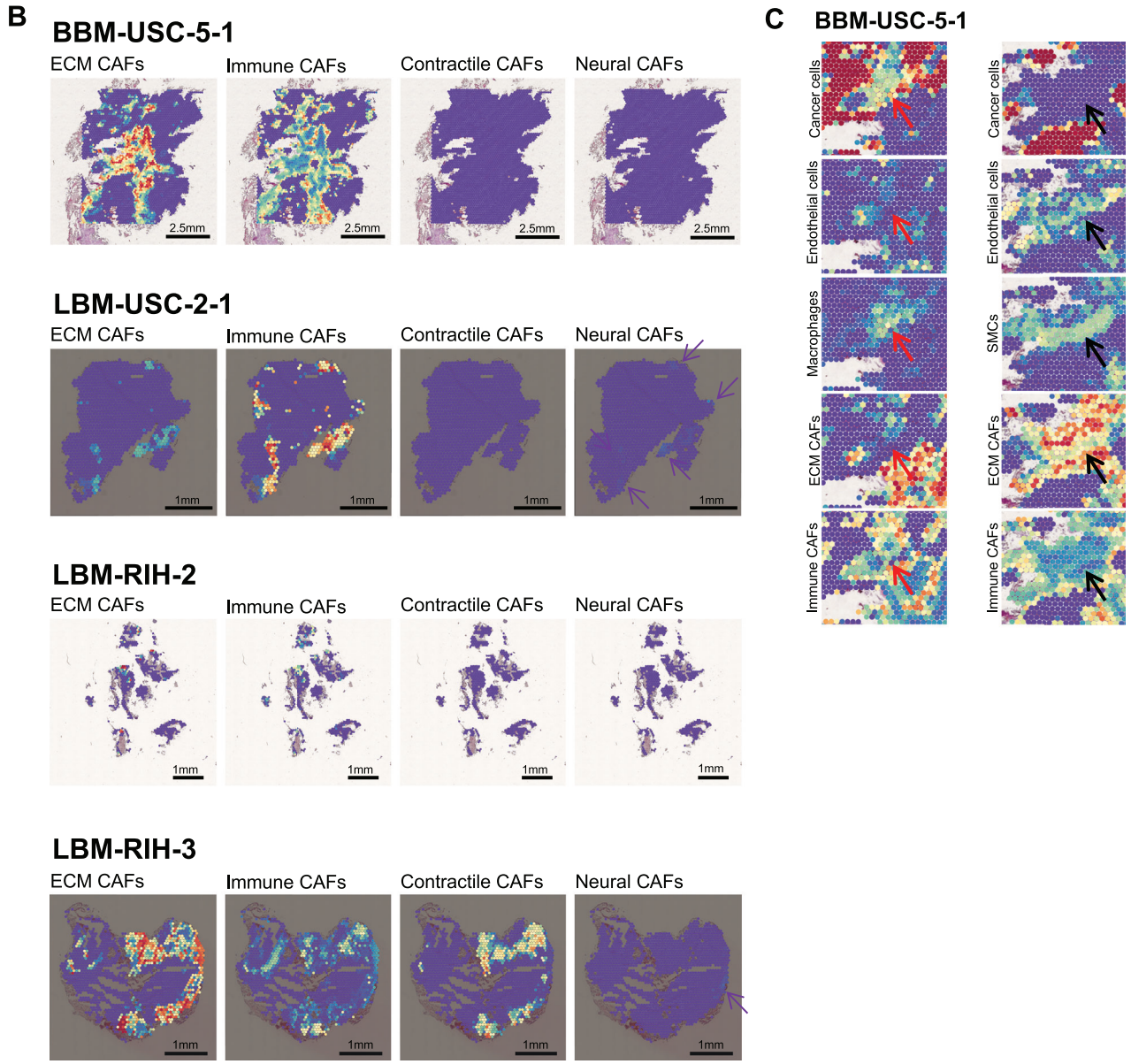


Fig. 4 | Continued.

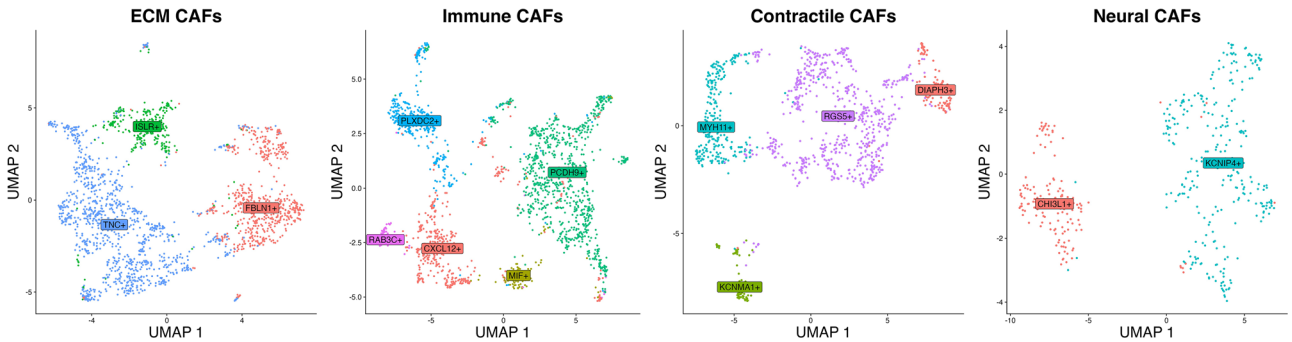


Fig. 5 | Analysis of subclusters in CAF subpopulations of the BM TME. UMAP visualization of each CAF subpopulation defined by scRNA-seq analysis of BM patient-derived tissues with different colors representing subclusters (*ECM* CAFs = ISLR+: green, TNC+: blue, FBLN1+: red; *Immune* CAFs = PLXDC2+: blue, RAB3C+: purple, CXCL12+: red, MIF+: green, PCDH9+: emerald green; *Contractile* CAFs = MYH11+: blue, RGS5+: purple, DIAPH3+: red, KCNMA1+: green; *Neural* CAFs = CHI3L1+: red, KCNIP4+: blue; CAF Cancer Associated Fibroblast, CHI3L1 Chitinase 3 Like 1, CXCL12 C-X-C motif chemokine ligand 12, DIAPH3 Diaphanous Related Formin 3, ECM Extracellular Matrix, FBLN1 Fibulin-1, ISLR Immunoglobulin Superfamily Containing Leucine-Rich Repeat, KCNIP4 Potassium Voltage-Gated Channel Interacting Protein 4, KCNMA1 Potassium Calcium-Activated Channel Subfamily M Alpha 1, MIF Macrophage Migration Inhibitory Factor, MYH11 Myosin Heavy Chain 11, PCDH9 Protocadherin 9, PLXDC2 Plexin Domain Containing 2, RAB3C RAB3C, member RAS oncogene family, RGS5 Regulator of G-protein signaling 5, TNC Tenascin-C).

The *ECM* CAF subpopulation was made up of three subclusters: FBLN1-CAFs, TNC-CAFs and ISLR-CAFs (Fig. 5). TNC-CAFs showed a reactive fibroblast phenotype with higher expression of *POSTN*, *PALLD* and reactive fibroblast marker *FAP* and activation of the "Tumor Microenvironment" pathway. FBLN1-CAFs showed a "normal" fibroblast profile, with expression of genes associated with normal CNS perivascular fibroblasts such as *ABCA9*, *NAV2*, or *FBLN1*, as well as activation of pathways like "RHOGDI Signaling"[40]. ISLR-CAFs showed strong expression of ECM genes such as *COL1A1* or *FN1*, as well as *VIM*. The *ISLR* gene has been shown to be highly expressed in CAFs with tumor-inhibitory capabilities in colorectal (CRC) and PDAC, suggesting that the *ECM* ISLR-CAFs are tumor-inhibitory in the BM TME[7,41,42]. Signaling pathways identified by IPA in ISLR-CAFs showed pathways linked to fibrosis regulation, as well as "Integrin Signaling" or "Wound Healing". Myofibroblast-related pathways, including "Regulation of Actin-based motility by RHO" were also activated in ISLR-CAFs (Fig. 5, Supplementary Fig. 11, and Supplementary Data 8)[41–43].

Together, our data revealed that each CAF subpopulation contains discrete subclusters of cells with either tumor-supportive or tumor-inhibitory capabilities, including a putative tumor-inhibitory ISLR-CAFs in the *ECM* CAF subpopulation.

Conditioned media from patient-derived CAF cell lines with high ISLR expression inhibit BM cell growth in vitro

BM CAF subpopulations consisted of subclusters whose gene expression profiles suggested specific roles in the TME, some appearing either tumor-supportive or -inhibitory. We previously established patient-derived CAF cell lines (CM01, CM02, CM03, and CM08) that we demonstrated to have either tumor-inhibitory or -supportive capabilities. Here, we further characterized their molecular and functional properties (Fig. 6A, B and Supplementary Fig. 12)[4].

First, scRNA-seq was performed on the CAF cell lines and revealed that CM01 and CM08 were more heterogeneous than CM02 and CM03 (Fig. 6A and Supplementary Fig. 12A). Next, we tested how closely the gene expression profiles of the CAF cell lines aligned with the gene expression profile of each CAF subpopulation. All four lines aligned most closely with the *ECM* CAFs (Fig. 6A). Accordingly, expression of the cell surface marker *ITGBL1*, which we have defined as specific for the *ECM* CAF subpopulation, was observed in CAF cell lines (Fig. 2C, Supplementary Fig. 12B, and Supplementary Data 3). Immune (*ADGRL4)* and contractile (*GRM8*) markers were low, and the neural marker (*PTPRZ1)* was undetectable (Supplementary Fig. 12B and Supplementary Data 3).

Previously, we demonstrated that CM08 inhibited tumor growth in a tumor cell/CAF mixing experiment in vivo[4]. One of the subclusters of *ECM* CAFs identified in the present analysis showed high expression of *ISLR*, a marker associated with tumor-inhibitory capabilities in other cancers[41,44]. Therefore, we examined the expression of *ISLR* in the scRNA-seq data for each CAF cell line. CM03 and CM08 CAFs expressed abundant levels of *ISLR*, but minimal to none in CM01 and CM02 CAFs (Fig. 6B). Western blot confirmed these findings, with ISLR protein detected in both the cell lysates and the conditioned media from CM03 and CM08 CAFs but not from CM01 or CM02 CAF cultures (Fig. 6C and Supplementary Fig. 13).

To assess if ISLR expression levels correlated with the tumor-inhibitory function, we treated a BM-derived tumor cell line (CM04) with conditioned media from high ISLR CAFs (CM03, CM08) or low ISLR CAFs (CM01, CM02) (Fig. 6D and Supplementary Data 9)[4]. CM08 CAF-derived conditioned media reduced CM04 viability by 21% at 72 h ($p = 0.0038$), 29% at 96 h ($p = 0.0201$) and 34% at 120 h ($p = 0.0053$), relative to controls. Treatment with conditioned media from CM03 decreased CM04 tumor cell viability by 16% at 120 h compared to cell viability at 72 h ($p = 0.0391$). In contrast, treatment with conditioned media from CM02 CAFs significantly increased the viability of CM04 cells by 11% at 24 h ($p = 0.0094$). We did not observe any significant difference in CM04 tumor cell viability upon treatment with conditioned media from CM01 (Fig. 6D). These data show that ISLR expression is correlated with the tumor-inhibitory functions of BM *ECM* CAFs.

Spatial transcriptomics confirmed the presence of the *ECM* ISLR-CAFs signature in all samples analyzed, suggesting that ISLR-CAFs participate in the BM tissue architecture (Supplementary Fig. 14A). High ISLR expression was also associated with *ECM* CAFs in an external BM scRNA-seq dataset (Supplementary Fig. 14B)[26–28]. Together, these data further support the relevance of *ECM* ISLR-CAFs within the BM TME.

To explore the relationships between *ECM* CAF subclusters, including the putative tumor-inhibitory ISLR-CAFs, we applied Monocle trajectory analysis to our scRNA-seq dataset. Monocle ordered single cells along a pseudotime continuum based on transcriptional similarity, predicting transitions between *ECM* CAF states (represented by gray circles along the trajectory) (Fig. 6E). Trajectory analysis revealed that FBLN1-CAFs, exhibiting the most "normal fibroblast-like" transcriptional profile, were positioned at the root of the pseudotime trajectory (white circle), suggesting that they represented an early or ground state within the *ECM* CAF lineage. In contrast, ISLR-CAFs, which may possess tumor-inhibitory features, occupied one terminus of the trajectory, and TNC-CAFs localized to intermediate states. These results suggest that ISLR-CAFs may represent a distinct cell state within the *ECM* CAF population, potentially arising from more fibroblast-like precursors (Figs. 5 and 6E).

To investigate the potential cellular origins of *ECM* CAFs in BM, we compared their transcriptional profiles with normal fibroblast and mural cell subtypes identified by Garcia et al. through scRNA-seq of the normal human brain vasculature[40]. Compared to 3 subtypes of perivascular fibroblasts (fibroblast type 1, 2, and 3), 2 subtypes of pericytes (pericyte type 1 and 2), and 2 subtypes of SMCs (arteriolar and venular SMCs, (aSMCs and vSMCs), *ECM* BM CAFs most closely resembled fibroblast type 1 (Supplementary Fig. 15A and Supplementary Data 10). Accordingly, both *ECM* CAFs and normal fibroblasts type 1 expressed *ABCA9*, top marker gene for normal fibroblasts type 1, and *FBLN1*, top marker gene of the "normal fibroblast-like" subcluster of *ECM* CAFs (Supplementary Fig. 15B)[40].

Discussion

CAFs have historically been underexplored in studies of BM, largely due to the assumption that fibroblasts are absent or play only a negligible role within the CNS parenchyma. Here, we show that CAFs are a structured, heterogeneous, and functionally relevant component of the BM TME, directly contributing to the tissue architecture and intercellular communication networks. Clustering analysis identified four transcriptionally distinct CAF subpopulations (*ECM*, *immune*, *contractile*, and *neural*), each defined by unique gene expression programs suggestive of diverse functional roles. This is consistent with literature highlighting that CAFs across diverse tumor types (PDAC, melanoma, CRC, and breast cancer) can be broadly categorized into *desmoplastic/ECM*, *immune*, and *contractile* populations, as we recently reported in a review[5,37,45–47]. Beyond this functional diversity, our data further indicate that BM CAF subpopulations differ in their spatial distribution, cellular origins, and degree of plasticity. These findings underscore that the CAF compartment in BM is more complex than the widely accepted but oversimplified "myofibroblast CAF versus inflammatory CAF" dichotomy[5,26,37].

Our study shows that CAFs engage in extensive crosstalk with tumor cells, including through the *WNT5A_PTPRK* signaling axis, which has been linked to resistance to immune checkpoint inhibitors, suggesting a role for CAFs in the adaptation of tumor cells to the CNS microenvironment[48]. In addition, CAFs communicate with other non-tumor cell types of the BM TME through signaling pathways that regulate key processes of tumor growth, as well as neoangiogenesis and immunosurveillance (*VEGFA_FLT1*), and interactions with neural cell residents (*NRG3_ERBB4*)[49]. Interactions between CAFs and the neural cells of the BM TME, including with neurons and oligodendrocytes, suggest that CAFs may play a role in tuning the specific CNS milieu to tumor cell presence[32,49]. CAFs also show subpopulation-specific interactions in the TME. For instance, a strong *CDH6_CDH6* interaction between *contractile* CAFs and SMCs suggests a synchronized involvement of cell types able to increase ECM stiffness to promote BM growth. Such subpopulation-specific interactions may be

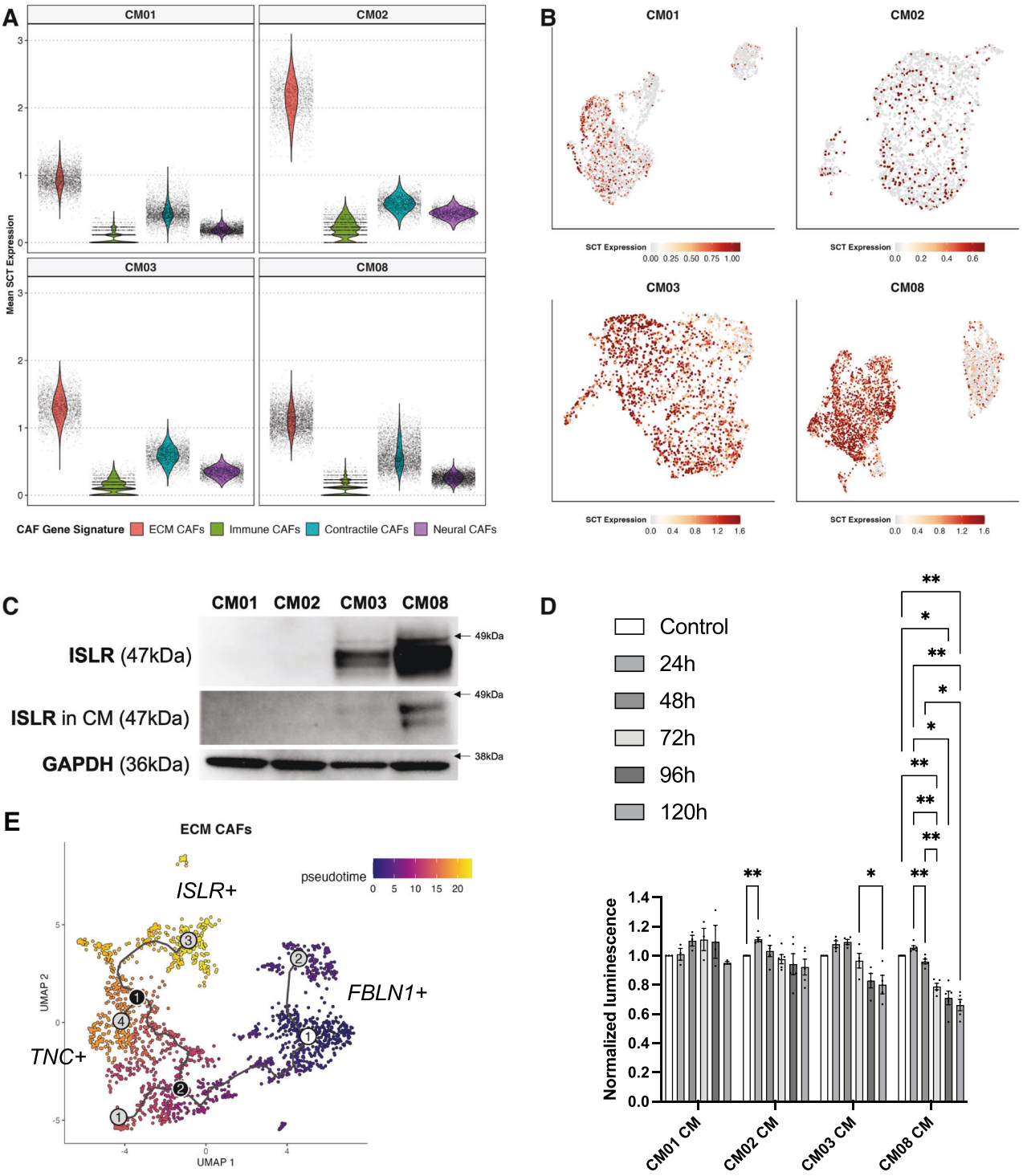


Fig. 6 | ISLR-CAFs derived from BM patients show tumor-inhibitory capabilities in vitro. A Expression of the BM CAF subpopulation gene signatures in each BM CAF cell line (*ECM* CAFs: red, *immune* CAFs: green, *contractile* CAFs: blue, *neural* CAFs: purple); **B** UMAP visualization of the expression of ISLR in scRNA-seq analysis of BM patient-derived CAF cell lines (Gene expression scale: gray to red); **C** Western blotting analysis of ISLR expression in BM patient-derived CAF cell lines and CAF cell line-derived conditioned media; Black arrows indicate size markers; **D** Cell viability analysis of BM patient-derived tumor cell line CM04 upon treatment with BM CAF cell line-derived conditioned media; results were reported as mean ± standard error of the mean (SEM); two-way repeated measures ANOVA was performed to determine the significance of the observed differences; $n > 3$ experiments (CM01 $n = 3$; CM02 $n = 5$; CM03 $n = 4$; CM08 $n = 5$), $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{***}p < 0.0001$; **E** UMAP visualization of pseudotime trajectory analysis performed in the *ECM* CAF subpopulation defined in scRNA-seq analysis of BM patient-derived tissues with number in a white circle being trajectory origin, numbers in black circles the trajectory nodes and numbers in gray circles the trajectory outcomes. CAF subpopulation subclusters defined in Fig. 5 are indicated (Pseudotime scale: blue to yellow; ISLR Immunoglobulin Superfamily Containing Leucine-Rich Repeat, TNC Tenascin-C, FBLN1 Fibulin-1) (CAF Cancer Associated Fibroblast, ECM Extracellular Matrix).

facilitated by the proximity of the CAF subpopulations to their partnering cell types, as revealed in our spatial transcriptomics analysis.

Our data also suggest that CAF subpopulations within the BM TME may originate from normal CNS fibroblast or SMC lineages in the CNS[40]. Nevertheless, it is possible that a subset of CAFs in the BM TME may have accompanied metastatic cells from the primary tumor site[5]. In addition, CAF subpopulations and their distribution are suspected to evolve over time or in response to therapy, along with the rest of the TME[5]. Such plasticity might, at least partially, explain the differences in cell type distribution we observed between samples, which were presumably collected at various time points during tumor progression.

Using patient-derived cell lines, we previously identified tumor-inhibitory CAFs, which expressed abundant amounts of matrix proteins and induced a desmoplastic reaction in vivo[4]. In the current study, we demonstrate that these CAFs might correspond to ISLR-CAFs within the *ECM* CAF subpopulation, which we showed have potential tumor-inhibitory functions in the BM TME[37,50]. Future investigations should assess how the presence or abundance of ISLR-CAFs correlates with survival in BM patients. Interestingly, data from patients with metastasis from prostate cancer or osteosarcoma showed that ISLR expression was linked to better survival[51,52]. Further work will also be needed to establish the molecular mechanisms underlying the tumor-inhibitory functions of ISLR-CAFs. ISLR has been reported to regulate fibrosis through reduction of collagen cross-linking and tissue stiffening, processes that may underlie its tumor-inhibitory functions[41,43,53–55].

While prior single-cell and spatial-omics studies have investigated CAFs in the BM TME, our report provides a more comprehensive characterization of the different BM CAF subpopulations, their interactions with other cell types, and their potential tumor-supportive vs. -inhibitory roles in the TME[26,27,56]. Although most of our scRNA-seq data were validated across independent external datasets, we observed some differences relative to previous studies, which could be explained by lower CAF numbers in the external datasets, as well as variabilities in tissue and data processing[26–28]. Further studies integrating larger datasets and comparisons to the TME of primary cancers would thus be needed to fully resolve these differences and help delineate BM TME specificities compared to their primary cancer counterparts. Investigating the TME of other cancers that metastasize to the brain, including melanoma, using a similar single-cell and spatial transcriptomics-based approach, would help determine whether the primary-tumor origin can impact the cell type and CAF distribution in the TME. Additional mechanistic and functional studies, both in vitro and in vivo, are needed to further dissect and understand the clinical and biological significance of the various CAF subpopulations. Overall, a better understanding of the role of CAFs in the BM TME will help develop therapeutic strategies that selectively target tumor-supportive CAFs while protecting tumor-inhibitory ones.

Materials and methods

Brain metastasis patient tumor sample collection

BM patient tissue samples were collected at Keck Medicine of the University of Southern California (USC, Los Angeles, CA) by Dr. Gabriel Zada and at Brown University Health (Providence, RI) by Dr. Steven A. Toms between October 2020 and February 2022. BM tissue samples were collected from 7 patients with primary breast (2 BMs) or lung (7 BMs) cancer. Tissue collection was approved by the USC Biomedical Institutional Review Board (IRB). Informed consent was obtained from all participants. All ethical regulations relevant to human research participants were followed. Detailed clinical information is summarized in Table 1. A total of 2 BMs were collected from both LBM-USC-2 and BBM-USC-5 patients. Fresh BM tissue samples were collected, transferred to a tube containing MACS© tissue storage solution (Miltenyi) and placed on ice before being processed for tissue snap freezing, OCT cryopreservation or cell dissociation within 24 h. The proportion of each tissue sample designated for scMultiome-seq was snap frozen in liquid nitrogen and stored at −80 °C until nuclei isolation. The proportion of each tissue sample designated for spatial transcriptomics analysis was embedded in OCT, immediately flash frozen in isopentane chilled by liquid nitrogen and stored at −80 °C until further processing.

Nuclei isolation and library preparation for scMultiome-seq

Nuclei isolation from frozen BM patient tissue samples was performed according to the 10x Genomics "Nuclei Isolation from Complex Tissues for Single Cell Multiome ATAC + Gene Expression Sequencing" protocol. Nuclei isolation from CAF cell lines was performed according to the 10x Genomics "Nuclei Isolation for Single Cell Multiome ATAC + Gene Expression Sequencing". Nuclei quality check and concentration measurement was performed at the end of the isolation process using a Countess II FL Automated Cell Counter (ThermoFisher Scientific) and ApoTome.2 microscope (ZEISS). Nuclei suspensions were then submitted to the Molecular Genomics Core (MGC) at the USC Norris Comprehensive Cancer Center (NCCC), where the rest of the scMultiome-seq procedures were performed. Nuclei suspensions were loaded onto a 10x Genomics Chromium Next gel beads in emulsion (GEM) Chip for transposed nuclei capture using 10x Genomics Chromium Controller. ScMultiome-seq libraries were performed using Chromium Next GEM Single Cell Multiome Reagent Kits according to the "Chromium Next GEM Single Cell Multiome ATAC + Gene Expression" user guide provided by 10x Genomics.

Sequencing was performed using a NovaSeq6000 instrument (Illumina) using NovaSeq6000 Reagent Kits v1.5 (100 cycles). For scMultiome-seq gene expression (RNA) libraries, dual-indexed sequencing runs were performed with the following read lengths: 28 bp for Read 1, 10 bp for i7 index, 10 bp for i5 index and 90 bp for Read 2. For scMultiome-seq ATAC libraries, dual-indexed sequencing runs were performed with the following read lengths: 50 bp for Read 1, 8 bp for i7 index, 8 bp for i5 index, 49 bp for Read 2.

Processing of scRNA-seq data from scMultiome-seq analysis

De-multiplexing, alignment, filtering and Unique Molecular Identifier (UMI) counting was performed using the Cell Ranger ARC Single Cell Software Suite version 2.0.2. Quality check and subsequent filtering were performed before analysis. For each sample, cells with fewer than 1000 UMIs or greater than 100,000 UMIs were removed. Cells with a mitochondrial content greater than 25% were also removed. A total of 89,364 individual cells from all 9 tissue samples were profiled, with a mean number per sample of 9929 cells (Supplementary Data 11). Number of cells per sample ranged from 745 to 19,875. Allosomal, mitochondrial, and ribosomal genes were removed for analysis. Samples had 7162 UMIs per cell on average (Supplementary Data 11). Patient-derived cell line samples had an average of 8230 UMIs per cell (Supplementary Data 11) with an average of 3276 cells per sample (min = 2132, max = 4783) (Supplementary Data 11).

The Seurat package was used to normalize gene expression, perform dimensionality reduction via PCA and UMAP, define clusters, and find DEGs between clusters using standard Seurat functions as described in the online manual[57]. Sample integration was performed using Harmony[58]. We ran dimensionality reduction using UMAP for cluster visualization. Cell type prediction was performed using scPred by first downloading single-cell reference data from the Human Protein Atlas[9,59]. Two models were then constructed: one trained using cell expression profiles from breast, brain, and lymphatic tissue to predict breast-to-brain metastases and the other trained using lung, brain, and lymphatic tissues to predict lung-to-brain metastases. UMAP embeddings were generated using the first 30 Harmony components (dims = 1:30) with Seurat's RunUMAP function (umap.-method = "uwot-learn", spread = 1.5, min.dist = 0.1), utilizing the Harmony-integrated dimensional reduction to correct for batch- and sample-specific trends. Cells predicted to be fibroblasts by the scPred models ($n = 5429$ cells) were isolated and re-processed in the same manner. Fibroblast clusters were determined using the FindClusters Seurat function using a resolution of 0.05; fibroblast subclusters were determined using the same method with a resolution of 0.25 to ensure low-abundance subpopulations were not overlooked. Differential expression between groups

was evaluated by Seurat's FindMarkers function using default parameters. Cell cycle phase and scores were determined using the Seurat CellCycleScoring function with default parameters.

Processing of scATAC-seq data from scMultiome-seq analysis

Cell by gene (scRNA-seq) and cell by peak (scATAC-seq) count matrices were generated using the Cell Ranger ARC pipeline as described above for the scRNA-seq analysis. Gene expression and ATAC peak data were processed using the R packages Seurat and Signac[57,60]. Low-quality cells were removed using cell filtering thresholds based on the distributions of quality metrics for each sample. Cells outside the 98-percentile for the number of fragments within peak regions, percent of reads in peaks, ratio of reads in ENCODE blacklist regions, nucleosome signal, and transcriptional start site enrichment were removed from the analysis. The remaining cells were then filtered to only include CAFs with subpopulation labels based on the scRNA-seq annotations as described above. A total of six samples had enough cells for downstream analysis. The number of cells per sample used for scATAC-seq analysis ranged from 71 to 1115. The average number of cells was 1062. MACS2 was used to call peaks and significantly differentially accessible regions among the four CAF subpopulations identified using the Seurat FindMarkers function with an adjusted $p < 0.1$ as the significance threshold[61]. Overrepresented transcription factor motifs were identified using the JASPAR 2020 database and the FindMotifs function from Signac with a significance threshold of adjusted $p < 0.1$[62].

Copy number variation analysis

Copy number variations (CNVs) were assessed in all main cell types. For CNV assessment, we used CopyKAT, which evaluates copy number profiles of each cell and assigns a "diploid" or "aneuploid" label to each cell[63].

Ingenuity pathway analysis

QIAGEN Ingenuity Pathway Analysis and DEG lists from each CAF subpopulations and associated subclusters (Figs. 2, 4) were used to perform gene enrichment for Canonical Pathways and their predicted activation (QIAGEN Inc., https://digitalinsights.qiagen.com/IPA)[64].

CellPhoneDB analysis

Cell-cell communication analysis was performed using the CellPhoneDB package. CellPhoneDB determines cell-cell communication networks based on single-cell expression of a curated list of ligands and their known receptors in the different cell types of the BM TME. It also allows measuring the degree of crosstalk between the different cell types identified in the BM TME, as represented by the "N Interactions" scale in Fig. 3A and "CellPhone Score" between receptors and ligands in Fig. 3B, C[65]. Because ligand/receptor interacting pairs involving FN1 were observed to widely dominate cell-cell interactions between all CAFs and other cell types of the TME, making up most of the top interactions in our analysis, they were excluded from Fig. 3B, C to allow the identification of more CAF subpopulation-specific ligand/receptor interactions between CAFs and other cell types.

Pseudotime trajectory analysis

Pseudotime trajectories were calculated from scRNA-seq data using the Monocle package[66–69]. Monocle cell data set (CDS) objects were created using scRNA-seq data extracted from previously created Seurat objects, so the same cells were used in both. Reductions were also extracted from the Seurat object so pseudotime could be directly overlaid on already generated UMAPs. The root node was determined for each CDS object by calculating which cells are nearest to each trajectory graph node and then picking the node with the highest fraction of nearby cells (https://cole-trapnell-lab.github.io/monocle3/docs/trajectories/).

Normal brain cell type and CAF subpopulation signature scoring

Normal brain cell type signature scoring. ScRNA-seq data were obtained from "Single-cell dissection of the human brain vasculature" by Garcia et al. (GSE173731) to establish gene signatures for "Fibroblast type 1", "Fibroblast type 2", "Fibroblast type 3", "Pericyte 1", "Pericyte 2", "Venular Smooth Muscle Cell", and "Arteriolar Smooth Muscle Cell" cell types described in the paper[40]. For each CAF subpopulation identified in the present analysis, expression of the marker genes for each of the normal brain cell types was combined using the MetaFeature Seurat function, which aggregates expression of multiple genes into one feature, representing how well they matched with a given normal cell subtype.

CAF subpopulation signature scoring. The top 100 upregulated gene markers for each CAF subpopulation were extracted. The mean expression for each of these gene sets was then calculated in each CAF cell line; we then compared the mean expression of each signature within each CAF cell line.

CAF subpopulation annotation of external datasets

We accessed scRNA-seq data from 3 external BM datasets generated by Gonzalez et al. (GSE186344), Kim et al. (GSE131907) and Song et al. (GSE234832)[26–28]. For each dataset, RNA counts were processed using the same Seurat-based method detailed above. Next, scPred was used to predict cell types, and fibroblast cells were isolated from both the GSE186344 dataset ($N = 5175$ CAFs), the GSE131907 dataset ($N = 443$ CAFs) and the GSE234832 dataset ($N = 134$ CAFs). To account for technical variation across datasets, fibroblast cells were merged and batch-corrected with Harmony, followed by dimensionality reduction (UMAP) and unsupervised clustering as described above. We then annotated CAF subpopulations in these external datasets by using our own dataset's CAF subpopulations as a reference. Specifically, anchors were identified between the integrated external fibroblasts and our reference CAF subpopulations using the Seurat function FindTransferAnchors, and CAF annotations were assigned via the Seurat function TransferData with default parameters. Finally, we generated UMAP plots to visualize how each fibroblast cluster in the external datasets corresponded to the CAF subpopulations identified in our study.

Sample and library preparation for Visium spatial gene expression analysis

BBM-USC-5-1, LBM-USC-2-1, LBM-RIH-2, and LBM-RIH-3 BM patient-derived tissue samples were used for spatial transcriptomics analysis using 10x Genomics Visium. OCT-embedded tissues were transferred to the Tissue Procurement Core (TPC) of the USC NCCC for further processing. Tissues were cut by TPC into 10 µm sections for the Visium protocol, as well as scrolls for RNA QC. Tissue sections and scrolls were then submitted to MGC at the USC NCCC, where the rest of the Visium procedures were performed. Suitability of tissue samples for Visium was carefully assessed by MGC. RNA QC was performed by measuring the RNA Integrity Number or DV200 of RNA to confirm tissues met the minimum requirements to proceed into spatial profiling. A tissue section per sample was stained with hematoxylin & eosin (H&E) to determine the region of interest (ROI) for spatial profiling. Ten-micrometer tissue sections were then mounted onto 10x Genomics Visium slides (6.5 × 6.5 mm spot) and further processed for fixation, H&E staining, whole transcriptome probe-mediated mRNA detection, reverse transcription, and library preparation, according to the "Visium CytAssist Spatial Gene Expression User Guide". Sequencing was performed on a NextSeq 2000 (Illumina).

Processing of Visium spatial gene expression data

Visium spatial transcriptomics data were processed using functions from the Seurat R package. Quality control of the raw count data was performed by filtering out low-quality spots where the number of genes detected per spot was <200 and the mitochondrial percentage per spot >15%. Counts were then normalized using the Seurat function LogNormalize. To annotate spots, we utilized our scRNA-seq dataset described above as a reference and performed an integration analysis between the scRNA-seq reference and Visium datasets. For the reference scRNA-seq dataset, counts per cell were normalized by LogNormalize and anchors between the reference and

Visium data were identified using the Seurat function FindTransferAnchors. The prediction scores for each cell type were transferred to the Visium data set using the Seurat function TransferData. We performed this analysis separately for the breast-to-brain BM samples and lung-to-brain BM samples. Scores for the four CAF subpopulations and the CAF subpopulation subcluster annotations were visualized using the SpatialFeaturePlot function in Seurat.

Cell culture

BM tumor cells (CM04) and BM CAFs (CM01, CM02, CM03, CM08) were established in our lab and maintained in culture as described before[4]. Briefly, cells were cultured in advanced DMEM/F12 media (Gibco), supplemented with 10% fetal bovine serum (FBS, Gibco), 1% Glutamax (Gibco), 10 ng/mL epidermal growth factor, 5 U/mL penicillin, 5 µg/mL streptomycin (Gibco), 2.5% Nu-Serum™ IV (Corning) and 10 ng/mL cholera toxin (Millipore Sigma). Cells were incubated at 37 °C in a humidified atmosphere at 5% CO_2. Media was changed twice a week. Cells were dissociated at confluence using trypsin/EDTA (Gibco). All cells were tested negative for mycoplasma at the beginning of the study.

Conditioned media were obtained from CAFs in culture by incubating cells at ~80% confluence with serum-free (Free from FBS and NuSerum IV) media for 24 h. At the time of conditioned media collection, cells were collected too, washed once with sterile PBS and stored at −20 °C until further processing. Cell concentration and viability were measured using a Countess II FL Automated Cell Counter. Only conditioned media from cells with >90% viability were stored at −20 °C until further use.

Western blot

Cell protein lysates were extracted using RIPA buffer (ThermoFisher Scientific), completed with Halt™ protease and phosphatase inhibitor cocktail (Sigma). 20–40 µg of protein extracts were loaded and run on a NuPAGE Bis-Tris protein gel (Thermofisher Scientific). Proteins were then transferred to a polyvinylidene fluoride (PVDF) membrane using an iBlot3 (Thermofisher Scientific). Blots were blocked in a 5% milk solution prepared in TBST, a Tris Base Solution (TBS) completed with 0.1% of Tween-20 (Millipore Sigma). Blots were incubated with primary antibodies (see below) at 4 °C overnight. Primary antibodies: ISLR (HPA050811, Millipore Sigma, 1/500 concentration); GAPDH (14C10, Cell Signaling Technology, 1/1000 concentration). Blots were then washed thrice with TBST for 5 min and incubated with secondary antibodies at room temperature (RT) for 1 h. Secondary antibody: Peroxidase AffiniPure Goat Anti-Rabbit IgG (H + L) (111-035-003, Jackson ImmunoResearch Laboratories Inc.). Blots were washed with TBST thrice before detection of the chemiluminescence signal. The chemiluminescence signal was visualized using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermofisher Scientific) and a ChemiDoc™ MP Imaging System (BioRad).

CellTiterGlow–cell viability assay

BM tumor cells (CM04) were plated in a 96-well plate (5000 cells per well) in triplicate in 100µL of complete media and incubated for 24 h. Cells were then washed once with sterile PBS and incubated with 100 µL of conditioned media obtained from cell culture of either CM01, CM02, CM03, or CM08 CAFs, and incubated for 24, 48, 72, 96 or 120 h. Controls were BM tumor cells incubated with serum-free media. At the time of cell viability assay, 100 µL of Cell-Titer-Glo® Reagent (Promega) was added to each well. The plate was agitated on a plate shaker for 2 min, then incubated for 10 min at RT before luminescence was recorded using a SpectraMax i3X plate reader (Molecular Devices). Graph shows an average of $n > 3$ experiments (CM01 $n = 3$; CM02 $n = 5$; CM03 $n = 4$; CM08 $n = 5$).

Statistics and reproducibility

Reproducibility was supported by independent biological replicates for single-cell ($n = 9$ patient-derived tissues) and spatial analyses ($n = 4$ patient-derived tissues) and by ≥3 independent experiments with biological and technical triplicates for in vitro assays. The sample size for in vitro cell viability experiments was determined based on the magnitude and consistency of differences between cells or conditions. Experimental findings were reliably reproduced across replicates. Results were reported as mean ± standard error of the mean (SEM). Two-way repeated measures ANOVA was performed to determine the significance of the observed differences. ANOVA: CAF cell line x Time: DF: 15, F (15, 65) = 3.652; CAF cell line: DF: 3, F (3, 13) = 18.05; Time: DF: 5; F (2.399, 31.19) = 19.02; Subject: DF: 13, F (13, 65) = 1.236; Residual: DF: 65. For Seurat CellCycleScoring, boxplots display S and G2M scores across CAF subpopulations; pairwise differences were assessed using two-sided Wilcoxon rank-sum tests. Differences were considered statistically significant at $p < 0.05$ (95% confidence interval, $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$).

Reporting summary

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

Data availability

Data will be made available upon reasonable request. Sequencing data, including scMultiome-seq and Visium spatial gene expression data, are available from the National Center for Biotechnology Information Gene Expression Omnibus database, with accession codes GSE322964 (scMultiome-seq) and GSE320505 (Visium spatial gene expression).

Code availability

Computer code will be made available upon reasonable request.

Received: 30 July 2024; Accepted: 12 March 2026;
Published online: 01 April 2026

References

1. Proescholdt, M. A. et al. The management of brain metastases—systematic review of neurosurgical aspects. *Cancers* **13**, 1616 (2021).
2. Mahmoodifar, S. et al. A quantitative characterization of the spatial distribution of brain metastases from breast cancer and respective molecular subtypes. *J. Neurooncol.* **160**, 241–251 (2022).
3. Tew, B. Y. et al. Patient-derived xenografts of central nervous system metastasis reveal expansion of aggressive minor clones. *Neuro Oncol.* **22**, 70–83 (2020).
4. Tew, B. Y. et al. Isolation and characterization of patient-derived CNS metastasis-associated stromal cell lines. *Oncogene* **38**, 4002–4014 (2019).
5. Simon, T. & Salhia, B. Cancer-associated fibroblast subpopulations with diverse and dynamic roles in the tumor microenvironment. *Mol. Cancer Res.* **20**, 183–192 (2022).
6. Ozdemir, B. C. et al. Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell* **25**, 719–734 (2014).
7. Miyai, Y., Esaki, N., Takahashi, M. & Enomoto, A. Cancer-associated fibroblasts that restrain cancer progression: Hypotheses and perspectives. *Cancer Sci.* **111**, 1047–1057 (2020).
8. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
9. Alquicira-Hernandez, J., Sathe, A., Ji, H. P., Nguyen, Q. & Powell, J. E. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol.* **20**, 264 (2019).
10. Sin, D. D. et al. Pro-surfactant protein B as a biomarker for lung cancer prediction. *J. Clin. Oncol.* **31**, 4536–4543 (2013).
11. Jiang, M., Liu, J., Li, Q. & Xu, B. The trichotomy of HER2 expression confers new insights into the understanding and managing for breast cancer stratified by HER2 status. *Int. J. Cancer* **153**, 1324–1336 (2023).
12. Lissandrini, D. et al. Receptor-type protein tyrosine phosphatase gamma (PTPgamma), a new identifier for myeloid dendritic cells and specialized macrophages. *Blood* **108**, 4223–4231 (2006).

13. Romero-Ramírez, H., Morales-Guadarrama, M. T., Pelayo, R., López-Santiago, R. & Santos-Argumedo, L. CD38 expression in early B-cell precursors contributes to extracellular signal-regulated kinase-mediated apoptosis. *Immunology* **144**, 271–281 (2015).
14. Crouch, E. E. & Doetsch, F. FACS isolation of endothelial cells and pericytes from mouse brain microregions. *Nat. Protoc.* **13**, 738–751 (2018).
15. Fan, C. et al. Single-cell transcriptome integration analysis reveals the correlation between mesenchymal stromal cells and fibroblasts. *Front. Genet.* **13**, 798331 (2022).
16. Karimi, A. & Milewicz, D. M. Structure of the elastin-contractile units in the thoracic aorta and how genes that cause thoracic aortic aneurysms and dissections disrupt this structure. *Can. J. Cardiol.* **32**, 26–34 (2016).
17. Hofmann, K. et al. Classical transient receptor potential 6 (TRPC6) channels support myofibroblast differentiation and development of experimental pulmonary fibrosis. *Biochim. Biophys. Acta Mol. Basis Dis.* **1863**, 560–568 (2017).
18. Bregin, A. et al. Expression and impact of Lsamp neural adhesion molecule in the serotonergic neurotransmission system. *Pharm. Biochem. Behav.* **198**, 173017 (2020).
19. Moore, A. et al. Loss of function of the neural cell adhesion molecule NrCAM regulates differentiation, proliferation and neurogenesis in early postnatal hypothalamic tanycytes. *Front. Neurosci.* **16**, 832961 (2022).
20. Jeon, B. N. et al. Zbtb7c is a molecular 'off' and 'on' switch of Mmp gene transcription. *Biochim. Biophys. Acta* **1859**, 1429–1439 (2016).
21. Beauchef, G. et al. The p65 subunit of NF-κB inhibits COL1A1 gene transcription in human dermal and scleroderma fibroblasts through its recruitment on promoter by protein interaction with transcriptional activators (c-Krox, Sp1, and Sp3). *J. Biol. Chem.* **287**, 3462–3478 (2012).
22. Robinson, K. F., Narasipura, S. D., Wallace, J., Ritz, E. M. & Al-Harthi, L. β-Catenin and TCFs/LEF signaling discordantly regulate IL-6 expression in astrocytes. *Cell Commun. Signal.* **18**, 93 (2020).
23. Pon, J. R. & Marra, M. A. MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget* **7**, 2297–2312 (2016).
24. Mazzeo, L. et al. ANKRD1 is a mesenchymal-specific driver of cancer-associated fibroblast activation bridging androgen receptor loss to AP-1 activation. *Nat. Commun.* **15**, 1038 (2024).
25. Jiang, J. et al. SOX on tumors, a comfort or a constraint? *Cell Death Discov.* **10**, 67 (2024).
26. Song, Q. et al. Single-cell sequencing reveals the landscape of the human brain metastatic microenvironment. *Commun. Biol.* **6**, 760 (2023).
27. Kim, N. et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* **11**, 2285 (2020).
28. Gonzalez, H. et al. Cellular architecture of human brain metastases. *Cell* **185**, 729–745.e720 (2022).
29. Qian, B. Z. et al. FLT1 signaling in metastasis-associated macrophages activates an inflammatory signature that promotes breast cancer metastasis. *J. Exp. Med.* **212**, 1433–1448 (2015).
30. Qin, Y. et al. A cancer-associated fibroblast subtypes-based signature enables the evaluation of immunotherapy response and prognosis in bladder cancer. *iScience* **26**, 107722 (2023).
31. Wang, Y. et al. Epiregulin reprograms cancer-associated fibroblasts and facilitates oral squamous cell carcinoma invasion via JAK2-STAT3 pathway. *J. Exp. Clin. Cancer Res.* **38**, 274 (2019).
32. Zhao, W. J., Yi, S. J., Ou, G. Y. & Qiao, X. Y. Neuregulin 2 (NRG2) is expressed in gliomas and promotes migration of human glioma cells. *Folia Neuropathol.* **59**, 189–197 (2021).
33. Ogier, C. et al. Targeting the NRG1/HER3 pathway in tumor cells and cancer-associated fibroblasts with an anti-neuregulin 1 antibody inhibits tumor growth in pre-clinical models of pancreatic cancer. *Cancer Lett.* **432**, 227–236 (2018).
34. Kumagai, S., Koyama, S. & Nishikawa, H. Antitumour immunity regulated by aberrant ERBB family signalling. *Nat. Rev. Cancer* **21**, 181–197 (2021).
35. Guan, Y. et al. Overexpression of PLXDC2 in stromal cell-associated M2 macrophages Is related to EMT and the progression of gastric cancer. *Front. Cell Dev. Biol.* **9**, 673295 (2021).
36. Figueiredo, A. M. et al. Phosphoinositide 3-kinase-regulated pericyte maturation governs vascular remodeling. *Circulation* **142**, 688–704 (2020).
37. Sebastian, A. et al. Single-cell transcriptomic analysis of tumor-derived fibroblasts and normal tissue-resident fibroblasts reveals fibroblast heterogeneity in breast cancer. *Cancers* **12**, 1307 (2020).
38. Hu, S. et al. TDO2+ myofibroblasts mediate immune suppression in malignant transformation of squamous cell carcinoma. *J. Clin. Investig.* **132**, e157649 (2022).
39. Li, F., Liu, A., Zhao, M. & Luo, L. Astrocytic Chitinase-3-like protein 1 in neurological diseases: potential roles and future perspectives. *J. Neurochem.* **165**, 772–790 (2023).
40. Garcia, F. J. et al. Single-cell dissection of the human brain vasculature. *Nature* **603**, 893–899 (2022).
41. Mizutani, Y. et al. Meflin-positive cancer-associated fibroblasts inhibit pancreatic carcinogenesis. *Cancer Res.* **79**, 5367 (2019).
42. Kobayashi, H. et al. The balance of stromal BMP signaling mediated by GREM1 and ISLR drives colorectal carcinogenesis. *Gastroenterology* **160**, 1224–1239.e1230 (2021).
43. Takahashi, M. et al. Roles of the mesenchymal stromal/stem cell marker Meflin/Islr in cancer fibrosis. *Front. Cell Dev. Biol.* **9**, 749924 (2021).
44. Iida, T. et al. Pharmacologic conversion of cancer-associated fibroblasts from a protumor phenotype to an antitumor phenotype improves the sensitivity of pancreatic cancer to chemotherapeutics. *Oncogene* **41**, 2764–2777 (2022).
45. Davidson, S. et al. Single-cell RNA sequencing reveals a dynamic stromal niche that supports tumor growth. *Cell Rep.* **31**, 107628 (2020).
46. Dominguez, C. X. et al. Single-cell RNA sequencing reveals stromal evolution into LRRC15(+) myofibroblasts as a determinant of patient response to cancer immunotherapy. *Cancer Discov.* **10**, 232–253 (2020).
47. Li, H. et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).
48. Jiang, Y. Q. et al. Investigating mechanisms of response or resistance to immune checkpoint inhibitors by analyzing cell-cell communications in tumors before and after programmed cell death-1 (PD-1) targeted therapy: an integrative analysis using single-cell RNA and bulk-RNA sequencing data. *Oncoimmunology* **10**, 1908010 (2021).
49. Zhao, W. J., Ou, G. Y. & Lin, W. W. Integrative analysis of neuregulin family members-related tumor microenvironment for predicting the prognosis in gliomas. *Front. Immunol.* **12**, 682415 (2021).
50. Toledo, B., Picon-Ruiz, M., Marchal, J. A. & Perán, M. Dual role of fibroblasts educated by tumour in cancer behavior and therapeutic perspectives. *Int. J. Mol. Sci.* **23**, 15576 (2022).
51. Hu, K. et al. A novel overall survival prediction signature based on comprehensive research in prostate cancer bone metastases. *Front. Med.* **9**, 815541 (2022).
52. Ma, T. et al. Immune-based prognostic biomarkers associated with metastasis of osteosarcoma. *Gen. Physiol. Biophys.* **42**, 1–12 (2023).
53. Hara, A. et al. Roles of the mesenchymal stromal/stem cell marker meflin in cardiac tissue repair and the development of diastolic dysfunction. *Circ. Res.* **125**, 414–430 (2019).

54. Ichihara, R. et al. Matrix remodeling-associated protein 8 is a marker of a subset of cancer-associated fibroblasts in pancreatic cancer. *Pathol. Int.* **72**, 161–175 (2022).
55. Nakahara, Y. et al. Fibroblasts positive for meflin have anti-fibrotic properties in pulmonary fibrosis. *Eur. Respir. J.* **58**, 2003397 (2021).
56. Christensson, G. et al. Spatial multiomics reveals intratumoral immune heterogeneity with distinct cytokine networks in lung cancer brain metastases. *Cancer Res. Commun.* **4**, 2888–2902 (2024).
57. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
58. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
59. Karlsson, M. et al. A single–cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, eabh2169 (2021).
60. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
61. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
62. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–d92 (2020).
63. Gao, R. et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* **39**, 599–608 (2021).
64. Krämer, A., Green, J., Pollard, J. Jr. & Tugendreich, S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30**, 523–530 (2014).
65. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
66. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
67. Van den Berge, K. et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).
68. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
69. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
70. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform.* **16**, 169 (2015).

Acknowledgements

The authors would like to thank the Keck Genomics Platform, the Molecular Genomics Core (NCI grant #P30CA014089) and the Translational Pathology Core of the USC Norris Comprehensive Cancer Center for their help with scMultiome-seq and Visium spatial gene expression processing.

Author contributions

Study conception and design: B. Salhia; *Data collection*: B. Salhia, T. Simon, D. N. Buckley, B. Y. Tew, M. Salomon, C. Matsuba, Z. Yang, G. C. Gooden, and K. Hurth; *Data analysis and interpretation*: B. Salhia, T. Simon, D. N. Buckley, M. Salomon, C. Matsuba, Z. Yang, and K. Hurth; *Manuscript preparation*: B. Salhia, T. Simon, D. N. Buckley, B. Y. Tew, M. Salomon, C. Matsuba, Z. Yang, G. C. Gooden, K. Hurth, D. Tran, S. A. Toms, and G. Zada.

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/s42003-026-09915-1.

Correspondence and requests for materials should be addressed to Bodour Salhia.

Peer review information *Communications Biology* thanks Qianqian Song and Spencer Watson for their contribution to the peer review of this work. Primary handling editor: Jasmine Pan and Johannes Stortz.