nature communications

Article

https://doi.org/10.1038/s41467-026-69716-2

# A spatially resolved human glioblastoma atlas reveals distinct cellular and molecular patterns of anatomical niches

Received: 4 June 2024

Accepted: 6 February 2026

Published online: 20 February 2026

Check for updates

Pranali Sonpatki[1,21], Hyun Jung Park[2,3,21], Yao Lulu Xing[4], Kyung Yeon Han[2], Brett A. Schroeder[5], Hyeon Jong Yu[6,7], Hye Jin Kim[6,7], Tamrin Chowdhury[6,7], Jong Ha Hwang[6,7], Sun Mo Nam[6,7], Yoon Hwan Byun[6,7], Ho Kang[6,7], Joo Ho Lee[7,8], Soon-Tae Lee[7,9], Jae-Kyung Won[7,10], Tae Min Kim[7,11], Seung Hong Choi[7,12], Ja-Lok Ku[7,13], Sungyoung Lee[14], Hongseok Yun[14], Sung-Hye Park[7,10], Claudia K. Petritsch[15,16,17,18], Chul-Kee Park[6,7,19] ✉, Woong-Yang Park[2,20] ✉ & Nameeta Shah[1,2,20] ✉

Glioblastoma is an aggressive brain cancer with limited treatment options and poor patient survival, driven in part by cellular diversity within tumors. While individual cell types have been catalogued, how malignant, vascular, and immune cells are spatially organized inside human tumors remains incompletely understood. Here we show a spatially resolved, multi-modal atlas of human glioblastoma that integrates gene expression profiling across tissue sections with matched single-cell and protein measurements at subcellular resolution. Using a targeted 348 gene panel enriched for vascular and stromal markers, we identify less well-characterized endothelial, perivascular, and fibroblast-like cell states and define their spatial associations with malignant and immune compartments. We further identify a distinct oligodendrocyte population restricted to tumor core and perivascular regions that exhibits gene expression patterns associated with tumor recurrence and poor clinical outcome. This publicly accessible atlas provides a high-resolution framework for studying the spatial organization of glioblastoma and highlights region-specific cellular interactions that may represent therapeutically actionable vulnerabilities.

Glioblastoma (GBM) is the most lethal form of primary brain tumor in adults, with a median survival of only 15 months despite intensive treatment regimens, including surgical resection, radiation, and chemotherapy[1]. The inevitable recurrence of GBM and its resistance to therapies are largely driven by profound cellular and molecular heterogeneity[2]. Understanding how the tumor heterogeneity arises and contributes to GBM progression is crucial for improving patient outcomes and developing more effective treatments.

Early efforts in comprehensive genomic profiling suggested three molecular subtypes of GBM: classical, mesenchymal, and proneural[3]. Subsequently, single-cell RNA sequencing (scRNA-seq) has further elucidated the molecular diversity of GBM, revealing the coexistence of multiple malignant cell states such as mesenchymal-like (MES-like), neural progenitor cell-like (NPC-like), astrocyte-like (AC-like), and oligodendrocyte precursor cell-like (OPC-like) within individual tumors[4–7]. Transcriptional heterogeneity in tumor tissue is complex and shaped by cell-intrinsic as well as cell-extrinsic factors. Although scRNA-seq has advanced our understanding of GBM cell-intrinsic heterogeneity, it remains unclear whether tumor cell states are randomly dispersed throughout the tissue or instead shaped by regional microenvironmental pressures.

A full list of affiliations appears at the end of the paper. ✉e-mail: nsckpark@snu.ac.kr; woongyang@g.skku.edu; nameeta.shah@gmail.com

It has long been recognized that GBM exhibits significant histological heterogeneity, with distinct regions of the tumor displaying different patterns of cellular organization[8]. Hans-Joachim Scherer's pioneering work on the spatial architecture of gliomas highlighted the importance of these spatially organized patterns in the progression and invasiveness of gliomas[9]. In our previous work, we reported the Ivy Glioblastoma Atlas (Ivy GAP) study, a comprehensive molecular-histology map of GBM designed to address the tumor's spatial and molecular heterogeneity[10]. Building on the Ivy Glioblastoma Atlas (Ivy GAP), which linked molecular programs to GBM histological features, recent advances such as 10x Visium, Xenium, and GeoMx now provide finer resolution, enabling spatial mapping of tumor–microenvironment interactions at the cellular scale[11–15]. These studies showed that the tumor microenvironment, in particular hypoxia and immune cell infiltrates, reshapes the transcriptional and genetic heterogeneity of glioblastoma cells.

A major finding of recent scRNA-seq analyses was that unmutated oligodendrocytes are a part of the glioma microenvironment, infiltrating gliomas at substantial and varying frequencies across glioma subtypes and patient ages. Although the most frequent type of non-tumor cells in some adult high-grade glioma, surprisingly little is known about their spatial distribution. Ravi et al. have documented spatial segregation of neurodevelopmental neighborhoods from hypoxia-associated mesenchymal and inflammation-associated neighborhoods[13]. Greenwald et al. documented 14 metaprograms, including 6 non-malignant programs, which feature Oligo (oligodendrocytes) alongside Mac (macrophages/microglia), Reactive-AST (reactive astrocytes)[15]. The Oligo metaprogram was localized in Layer 5, corresponding to the infiltrating area of the brain parenchyma. Whether OL/OPCs exhibit temporally distinct and functional states in glioma, similar to normal OPCs which display spatiotemporal heterogeneity, depending on age and their anatomic location, remains unknown. While past studies have significantly enhanced our understanding of the microenvironmental factors driving glioma heterogeneity, their limited resolution reduces the capacity to fully capture the cellular composition and complex patterns of cell-cell interactions within GBM anatomical subdomains.

In this study, we present a comprehensive, high-resolution mapping of the cellular architecture within the anatomical niches of GBM. We generated whole spatial transcriptome data using Visium CytAssist, integrated with patient-matched single cell multi-omic data, including from matched tumor core and periphery. In addition, for better analysis of tissue architecture, we acquired high-resolution H&E images with spatial transcriptome at 40x in contrast to the more commonly used 5x or 10x magnifications. Additionally, we generated high-depth Spatial Transcriptome (ST) data, leading to a median of 4694 genes detected per section, compared to 2159 in a recent study and identifying an additional extra 482 genes, providing a more comprehensive view of the molecular landscape[13]. Furthermore, previous studies have primarily used fresh frozen (FF) sections for ST data generation[13,15] however, we generated ST data from not only FF ($n = 5$), but also FFPE sections ($n = 27$), which are more commonly used in clinical samples, offering broader applicability while preserving tissue architecture for comprehensive analysis. As a result, we obtained 115,914 spatial transcriptomes across 32 genotyped glioma sections, mapping 56 distinct cellular components, including oligodendrocyte subtypes not emphasized in prior studies, some localized to distinctive spatial domains in the tumor core. We identified a neutrophil population from the freshly processed samples, a cell type missing from most human GBM scRNA analyses[16]. In addition, we employed Xenium in situ technology with subcellular resolution to refine cell type annotations and validate the existence of the rare oligodendrocyte subtype. This study underscores the extensive spatial and cellular heterogeneity of GBM and emphasizes the intricate interplay among diverse cellular components. To facilitate broader access and utilization, our analytical results are presented as a public resource in a user-friendly format, available for all at https://github.com/nameetas/TSKGA.

## Results

### Spatially resolved GBM atlas

While considerable progress has been made in exploring molecular GBM tumor heterogeneity with spatial transcriptomics (ST) technologies, high-resolution cellular organization and cell–cell interaction analysis is still lacking. To generate a high-resolution spatial map of the cellular structure across the anatomical regions of GBM, we employed single cell, whole transcriptome ST (Visium CytAssist) as well as subcellular resolution in situ ST (Xenium) technologies on both fresh frozen (FF) and formalin-fixed paraffin embedded (FFPE) preserved blocks (Fig. 1A–E). To comprehensively capture GBM heterogeneity, we analyzed patients across different age groups with diverse genomic alterations (Supplementary Data 1 and Fig. 1F) and profiled ST on at least two tissue sections prepared from different locations with distinct anatomical features within the same tumor.

We investigated the cellular architecture of GBM in 28 patients (detailed patient and specimen information is provided in Supplementary Data 1 and Fig. 1F), including 18 with GBM, 5 with astrocytoma (IDH-mutant), 1 with oligodendroglioma (IDH-mutant, 1p19q co-deleted), and 4 with miscellaneous CNS pathologies (Fig. 1A). Complementary single-cell RNA sequencing (scRNA-seq) and cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) data were acquired from 25 patients, resulting in a collective analysis encompassing 223,113 cells (Fig. 1B). In addition, we cataloged the GBM single-cell reference atlas by compiling the newly generated single-cell data with publicly available sc/snRNA-seq datasets, containing 10 major cell types, which were further classified into 58 transcriptional states (Fig. 1B). Furthermore, we collected ST data for 17 cases, comprising 16 brain tumor patients (13 GBM, 3 Astrocytoma) and one healthy brain tissue control using the 10x Genomics Visium ST platform (Fig. 1C). As a result, we successfully generated a comprehensive dataset comprising 115,914 individual ST spots from 32 tissue slices. Among these, 14 cases have both matched ST and single-cell data. By integrating the reference single cell atlas with ST data, we identified recurrent molecular features associated with anatomical features (AF)-based niches including cell-type composition, and co-occurrence (Fig. 1D). Additionally, we generated subcellular resolution ST data using recently developed Xenium in situ technology (Fig. 1C) for 4 GBM patients, including a custom panel of 348 genes. From this, we annotated 729,307 cells within an area of 155 $mm^2$. Collectively, we generated an integrated spatially resolved GBM atlas which is publicly available through an interactive website at https://github.com/nameetas/TSKGA to enhance user accessibility (Fig. 1E).

### GBM cellular complexity revealed through matched core and periphery single-cell analysis

To investigate cell states in both the core and periphery (peri) of GBM, we generated single-cell transcriptomes from six matched core–peri fresh samples, processed on the same day (Supplementary Data 1). After quality filtering, 223,113 cells were pooled for single-cell transcriptomics, batch-corrected, and clustered using Recursive Consensus Clustering (RCC)[16]. The level 1 clusters of RCC were used to assign major cell types (microglia, oligodendrocyte, neoplastic, myeloid, lymphocyte) by canonical marker gene expression (Supplementary Data 2). The major cell types were then recursively clustered up to three levels and markers were identified for each cell subtype or cellular state (Fig. S1A–S1F). The cluster name indicates the major cell type followed by the cluster numbers separated by dashes for each level or an abbreviation based on previously described cellular states (see Methods, Fig. 2A, Supplementary Data 3). Notably, we were able to detect a granulocyte population from the freshly processed samples (4.6% vs 1% in freshly frozen), a cell type missing from most human

https://doi.org/10.1038/s41467-026-69716-2

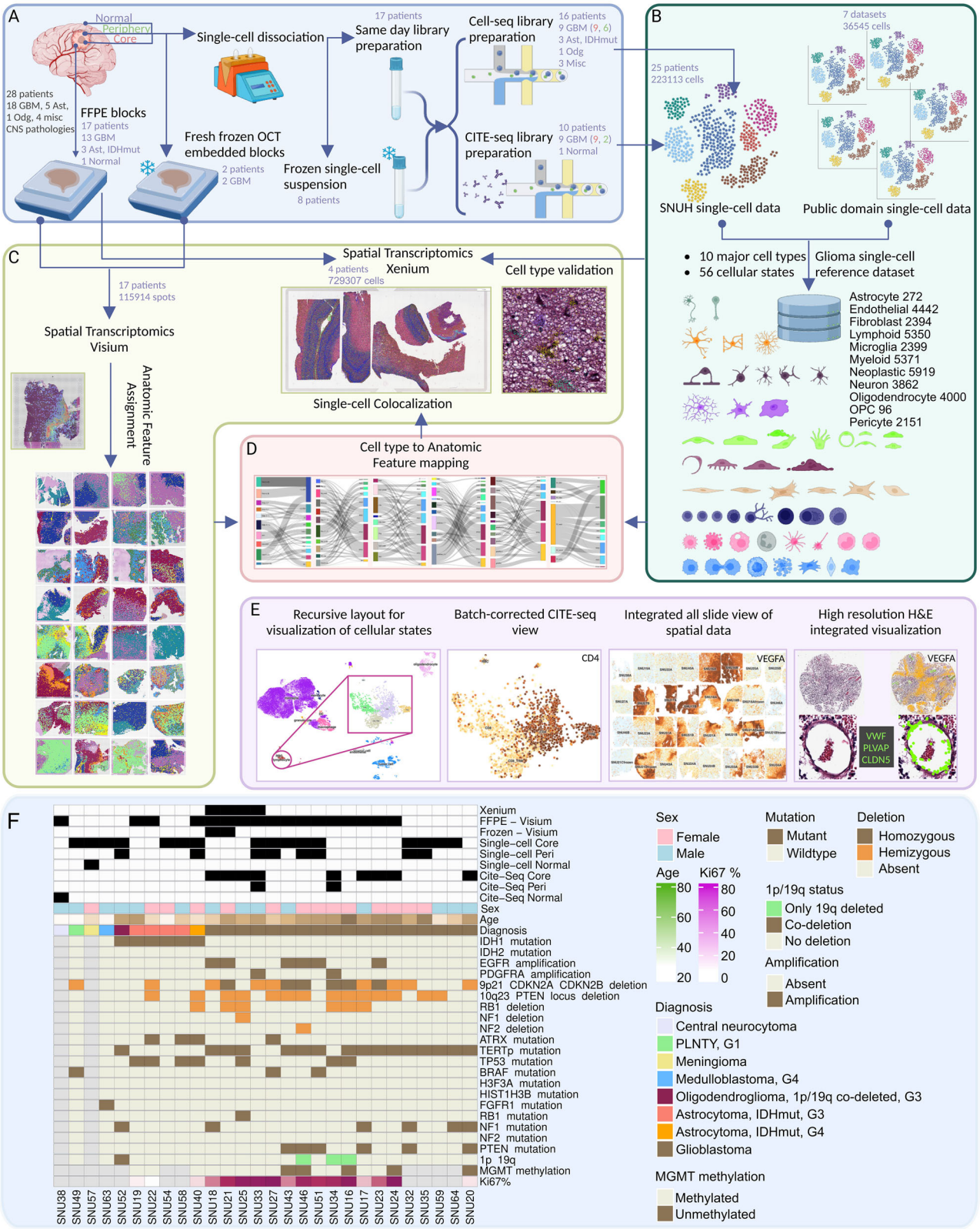
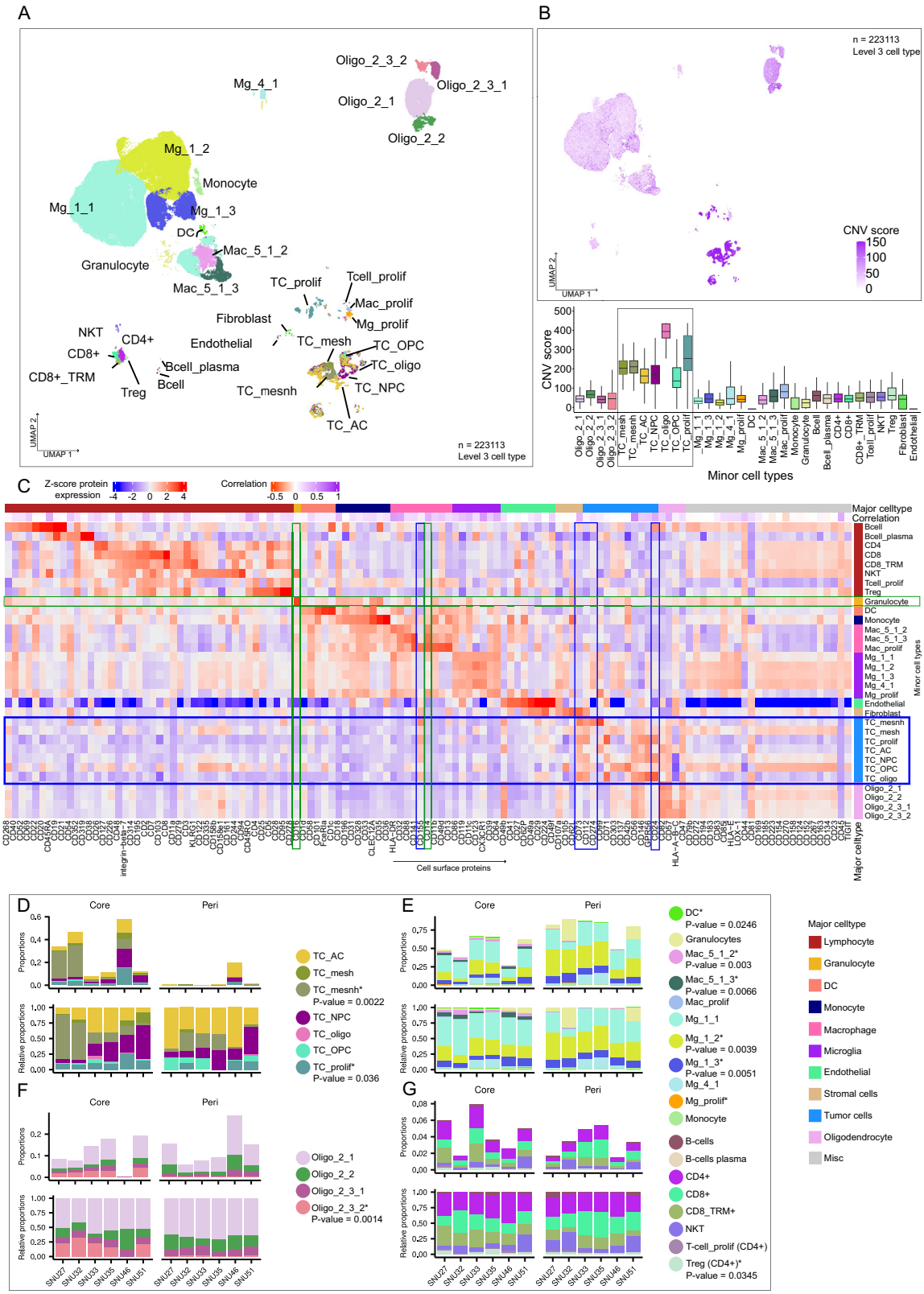



**Fig. 1 | Schematics of the spatial multi-omics glioma atlas, analysis workflow, and the interactive webpage. A** Study overview. **B** Schematic of the workflow for generating the glioma single-cell reference dataset. **C** Spatial transcriptomics data analysis workflow for anatomical feature assignment using Visium and validation using Xenium, the targeted spatial profiling method at subcellular resolution. **D** Spatial anatomic features to cell type mapping. **E** Example snapshot of a user-friendly webpage for exploring the dataset. **F** Summary of patient demographics and clinical information (see Supplementary Data 1 for details). Created with BioRender.com.

https://doi.org/10.1038/s41467-026-69716-2




GBM scRNA-seq public datasets (GBmap[17]—0.45%) (Fig. 2A, Fig. S2A). To validate the assignment of malignant cell types, we performed CNV analysis, which confirmed that these populations exhibited significantly higher CNV scores than non-malignant cells (Fig. 2B). Clustering cells recursively for three levels resulted in identifying cell populations like stromal and endothelial cells which are typically rare in glioblastoma scRNA-seq data (0.37%). For 11 patients, we also profiled 130 cell surface antigens on a total of 62,973 cells by CITE-seq (Supplementary Data 1). CITE-seq analysis identified key immunophenotypic markers for immune cells and other major cell types found in both the tumor and its surrounding microenvironment (Fig. S2B–C). We observed that tumor cells highly expressed immunomodulatory molecules, including CD24, CD44, CD112 (PVRL2), CD155 (PVR), CD73 (NT5E), and CD274 (PD-L1) (Fig. 2C, highlighted in blue). Granulocytes

**Fig. 2 | Single-cell classification and spatial Segregation of GBM Cellular Heterogeneity Between Tumor Core and Periphery. A** UMAP representation of minor cell types at level 3 of RCC. **B** UMAP displaying CNV scores for individual cells. The accompanying bar plot shows the distribution of CNV scores across minor cell types in the single-cell dataset. Neoplastic cells are highlighted with a rectangular box, indicating elevated CNV scores relative to other cell types. In boxplots, the centre line represents the median, the box spans the interquartile range (25th–75th percentiles), and whiskers extend to the most extreme data points within 1.5 * *IQR*. **C** Heatmap visualizing normalized cell surface protein expression across cellular states in the CITE-seq dataset. Blue rectangular boxes mark proteins with elevated expression in the neoplastic cell population. Green rectangular boxes mark proteins with differential expression in granulocytes. Overall and relative proportions of Neoplastic (**D**), myeloid (**E**), oligodendrocyte (**F**), and lymphoid (**G**) cell subtypes across patients in tumor core and periphery tissue. Minor cell types with significant proportional differences are marked with an asterisk. $n = 6$ biologically independent patients, with matched core and peri samples obtained from the same individuals. Significance was assessed with a two-sided paired t-test across patients on per-sample proportions. *$p$-value < 0.05.

show high expression of CD16 and low CD14 (Fig. 2C, highlighted in green) suggesting these are neutrophils consistent with other studies using protein markers[18,19]. This analysis provided a comprehensive list of cell surface markers for the major cellular components of the tumor microenvironment (TME) in GBM.

Next, we compared the cellular subtype proportions in core vs. peri samples. Core samples contained a higher proportion of tumor cells compared to peri samples (Fig. 2D–G). Among the subpopulation of neoplastic cells, there are cases that the TC_mesnh and TC_prolif cells are exclusively found in the core (Fig. 2D). For myeloid/microglia cell types, dendritic cells (DC), Mac_5_1_2 (blood-derived macrophages), Mac_5_1_3 (blood-derived hypoxic macrophages), and Mg_prolif (proliferating microglia) were more abundant in the core. In contrast, Mg_1_2 and Mg_1_3 (resting microglia) are more prevalent in the peri (see Fig. 2E). Among oligodendrocyte populations, we discovered a previously less studied subtype (Oligo_2_3_2) exclusively within the core region across all six samples except for sample SNU46, sample with highest tumor cell and lowest immune cell proportions (Fig. 2F). This population shows immune-activated features, reduced myelin-associated transcripts, and strong association with adverse outcomes—features not described in prior glioma atlases. Lymphocyte subtypes were not significantly different in core vs. peri except regulatory T cells (Treg) which were mostly present in core (Fig. 2G). Together, these findings highlight that the differences in cell types and proportions between the tumor core and peri-tumoral regions reflect variations in the TME and immune response.

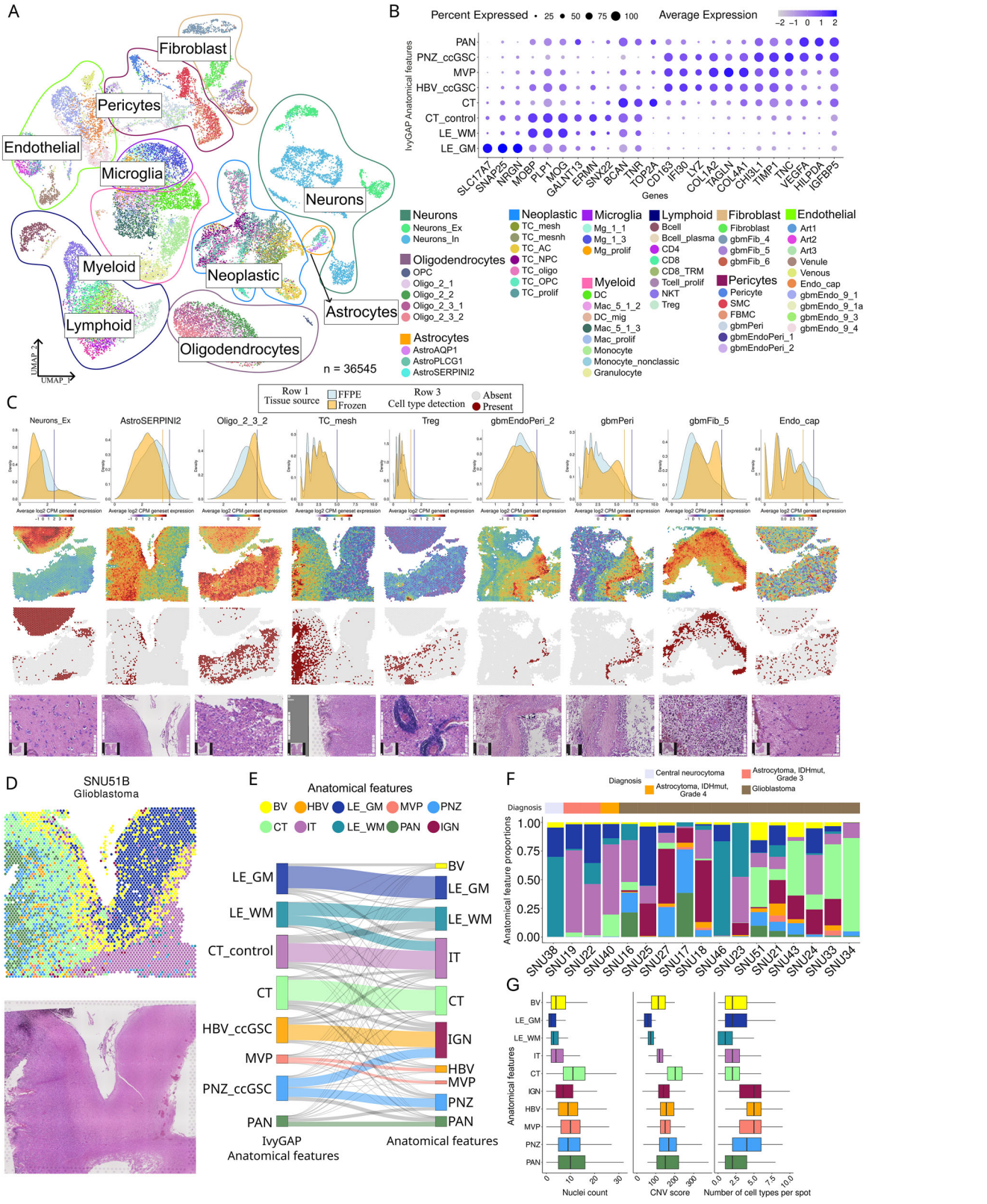
## Deconvolving GBM anatomical features with a curated single cell atlas

To construct a comprehensive GBM spatial map, we first created a reference single-cell atlas by combining newly generated scRNA-seq data with published single-cell transcriptomic data including human adult brain snRNA-seq data[20], human brain vascular cell scRNA-seq data[21–23], glioma associated T cell scRNA-seq data[24] and GBM myeloid cell enriched scRNA-seq data[25] (Fig. 3A and Supplementary Data 4). To optimize clustering while maintaining diversity, we constructed a reference dataset of 36,256 cells with balanced proportions of major cell types (2000–5000 cells per cell type, except for OPCs and astrocytes) (Fig. S2A). In total, we annotated 58 cellular subtypes across 10 major cell types; astrocytes, neurons, oligodendrocytes/OPCs, neoplastic, microglia, lymphoid, myeloid, endothelial, fibroblast, and pericytes. Using this reference dataset, we then identified highly specific markers for each cell subtype (Supplementary Data 5). We excluded the cell subtypes without specific markers for spatial distribution, leaving 56 cell types for further analysis. We were able to define more cellular states compared to 14 spatial metaprograms in Greenwald et al. study and 21 cell types at the finest level of annotations (level 3) in GBmap (Fig. S2A, Supplementary Data 6). Most identified states belonged to vascular cells, followed by immune cells and oligodendrocytes. There are three Greenwald spatial only malignant metaprograms; Chromatin.reg (malig.), Prolif.Metab (malig.), and MES.Ast (malig.) for which they were not able to assign a corresponding single cell type. Based on our analysis these spatial programs are dominated by vasculature associated cell types typically lost in glioblastoma scRNA-seq datasets unless captured using alternate protocols[21,26]. Since, we integrated these cell types in our reference dataset we were able to see that Chromatin.reg (malig.) program correlates to AstroSERPINI2, Prolif.Metab (malig. and shown to be strongly associated with vascular program in Greenwald et al.) to gbmEndoPeri_2, and MES.Ast (malig. and shown to be associated with immune activity in Greenwald et al.) to gbmFib_5 (Supplementary Data 6, Supplementary Data 13).

To investigate the spatial organization of GBM anatomical features (AFs), we generated ST data comprising 32 tumors from 17 patients, including 13 patients diagnosed with IDH-wildtype GBM, 3 patients with IDH-mutant high-grade astrocytoma, and a normal brain region serving as a comparative control (Supplementary Data 1, Fig. 1F). The ST specimens were carefully chosen to encompass diverse AFs of GBM identifiable by H&E staining; leading edge (LE), cellular tumor (CT), microvascular proliferation (MVP) and pseudopalisading cells around necrosis (PAN)[10]. We first defined AFs of each ST spot. We utilized previously established marker gene sets for all AFs from Ivy GAP[10,13] (Fig. 3B; Supplementary Data 7). Besides the histologically visible AFs, we also used glioma stem-like cell (GSC) niches-vascular (HBV_ccGSC - hyperplastic blood vessels with community curated glioma stem-like cell markers (ccGSC)) and hypoxic (PNZ_ccGSC - perinecrotic zone with ccGSC expression)-defined by integrating histological traits and transcriptional profiles of laser-captured microdissected regions expressing ccGSC markers[10,13]. Tumor regions lacking expression of any ccGSC markers were categorized as CT_control. We further divided the LE area into LE_WM (white matter) and LE_GM (gray matter) as these areas have distinct cellular profiles.

Each spot was scored for the 56 reference cell states using averaged marker expression, followed by binary presence/absence assignment (Fig. 3C, Supplementary Data 8 see Methods, Supplementary File 14). Visualization of Ivy GAP AFs over high-resolution H&E images indicated that further refinement was needed for AF assignment as HBV, MVP did not always overlap with hyperplastic/glomerulus blood vessels and PAN did not show any visible PAN on H&E. This can largely be attributed to overlapping immune and vascular signature across HBV, MVP, PNZ and overlapping hypoxic signature between PNZ and PAN. We were able to use our cell type assignment and H&E features to further stratify AFs in to 10 AFs with additional introduction of BV (Blood Vessel), and IGN (Immune-Glial Niche) (Fig. 3D–E, Fig. S3). Notably, normal sample spots (SNU38) were classified into LE_GM, LE_WM, and BV AFs, while spots from patients with astrocytoma (Ast), IDH mutant (IDHm) (SNU19, SNU22) were classified into LE_GM, LE_WM, BV, IT, and CT (Fig. 3F, Fig. S3). In stark contrast, GBM patient spots exhibited classification across all AFs (Fig. S3). We performed copy number analysis as an alternate method to estimate the presence of tumor cells for each spot. We also calculated the number of major cell types in each spot using our binary assignment. Furthermore, cell counts per spot were determined by identifying nuclei on H&E images using the DINO object detection algorithm. As expected, CT spots showed maximum CNV scores, highest cell density but low cell type diversity; IGN, HBV, and MVP followed by PNZ show lower CNV scores due to presence of nonmalignant cell types but highest cell type diversity (Fig. 3G, Fig. 4A).

https://doi.org/10.1038/s41467-026-69716-2




## Mapping GBM tumor architecture by anatomical feature specific cell states

In order to assess spatial AF preference for a given cell subtype, we compared the distribution of normalized expression values for that cell subtype in real data vs. AF randomized data. If the area under the curve for actual data is higher than randomized data then the cell subtype is preferentially present in that AF (Fig. 4B, Fig. S4). The distribution of the 56 cellular subtypes in 115,914 ST spots from 32 specimens, as well as detailed specimen analysis, including matched single cell UMAP, AF annotation, and spatial CNV analysis, are provided on the interactive website, https://github.com/nameetas/TSKGA (Supplementary Data 15).

Spatial annotation of cellular subtypes revealed varying degrees of spatial and cellular heterogeneity within the GBM tumor

**Fig. 3 | Cell type identification, anatomical feature annotation, and spatial neighborhood interactions in Visium spatial transcriptomic data. A** UMAP visualization of minor cell types derived from the single-cell reference dataset (**B**) Dot plot illustrating the expression patterns of selected marker genes for anatomical features, curated based on IvyGAP annotations. Each dot represents the scaled average expression and proportion of cells expressing the gene across anatomical regions. **C** Distribution and spatial detection of cell types per spot in the Visium FFPE and frozen tissue sections. The first row depicts the density distribution of average cell type-specific geneset scores across all the spots, with vertical lines denoting detection thresholds orange for FFPE and blue for frozen samples. The second row presents the spatial maps of the continuous cell type score across the tissue section. The third row shows the binarized presence/absence classification of each cell type at the spot level, where red dots indicate cell type detection above the defined threshold. The fourth row provides magnified spatial views highlighting regions with detected cell types. **D** Spatial map of anatomical feature assignment for the SNU51 sample, based on integration of marker gene profiles and spatial proximity. **E** River plot visualizing the mapping of IvyGAP-defined anatomical features to the final inferred spatial anatomical annotations, reflecting both lineage and spatial continuity. **F** Relative proportions of anatomical features (AFs) across individual samples, highlighting the heterogeneity in spatial organization. The anatomical feature assignment was restricted to BV, LE_GM, LE_WM, IT and CT in non-GBM samples. **G** Distribution of nuclei counts, inferred copy number variation (CNV) scores, and number of detected cell types within each AF across 17 patients, 32 slides and 115914 Visium spots. Nuclei were identified from H&E-stained tissue sections using the DINO-DETR[47] deep learning-based object detection framework. In boxplots, the centre line represents the median, the box spans the interquartile range (25th–75th percentiles), and whiskers extend to the most extreme data points within 1.5 * *IQR*.

microenvironment (TME) (Fig. 4C and Fig. S4). BV are enriched with Endo_cap cells and consist mainly of normal vasculature, predominantly in LE_GM, followed by LE_WM, IT, and CT regions. LE_GM is enriched in neurons, AstroPLCG1 astrocytes, and Art2 cells, while LE_WM contains more oligodendrocytes, particularly Oligo_2_1. As the region transitions from LE_WM to IT, there is increased presence of other oligodendrocytes (Oligo_2_3_1), reactive astrocytes (AstroAQP1), microglia with an emergence of non-mesenchymal tumor cells. CT shows a clear enrichment of non-mesenchymal tumor cells. The Immune-Glial Niche (IGN) contains elevated levels of GBM-associated vascular cells, reactive astrocytes, fibroblasts, lymphoid and myeloid cells, as well as mesenchymal tumor cells (TC_mesnh). The identified Oligo_2_3_2 subtype is also found in IGN. HBV and MVP regions show the highest enrichment of vascular cells, with MVP notably containing more GBM-associated pericytes (gbmPeri, gbmEndoPeri). The PNZ is a hypoxic zone rich in mesenchymal tumor cells (TC_mesnh, TC_mesh), immune cells, and vascular cells. In contrast, the PAN - also hypoxic - features TC_mesh, TC_NPC, and hypoxic myeloid cells, and appears as pseudopalisades on H&E staining. The immune cell profile changes from IGN to HBV to MVP: microglia decrease while macrophages and lymphoid cells increase. In the hypoxic niches (PNZ and PAN), immune diversity drops further, with only macrophages remaining (Fig. 4C).

We compared our AFs with the spatial modules from Ravi et al. and the layers defined by Greenwald et al. (Fig. 4D, Supplementary Data 6 and 13). The Reactive Hypoxia module and Layer L1 align with PAN/PNZ; the Reactive Immune module and Layer L2 with IGN; the Radial Glia module and Layer L3 with HBV/MVP; the Regional OPC module and Layer L4 with LE_WM, IT, and CT; and the Regional NPC module and Layer L5 with LE_GM and BV (Fig. 4D). Greenwald et al. describe their layers as hypoxia-driven and are relevant to understanding GBM invasion from hypoxic core to brain parenchyma. However, accounting for tumor expansion likely requires additional insight into spatial structure. We analyzed the spatial relationships between AFs by calculating the proportion of neighboring AFs around each one (Fig. 4E, Fig. S4, see methods). AF pairs with proportions three times lower than expected by chance (based on shuffled data) are excluded from the graph, pairs with proportions at least three times higher than random are connected with green edges. Rest of the pairs are connected with blue edges. The thickness of self-loops indicates relative AF size: LE_WM, LE_GM, IT, CT, and IGN are the largest, followed by PAN and PNZ. Vascular AFs—BV, HBV, and MVP—are interspersed with these larger regions (Fig. 4E). Our neighborhood analysis reveals a bidirectional linear arrangement: LE (L5) ↔ IT/CT (L4) ↔ PNZ/PAN (L1). This organization reflects the known dual behavior of GBM—both expansive (mass forming tumor with hypoxic core) and invasive (infiltrative in brain parenchyma). In contrast, vascular and immune-related AFs (IGN, HBV, MVP) and layers (L2, L3) are more diffusely distributed between IT and PNZ and do not form clearly ordered layers (Fig. 4E). Unlike other Greenwald layers, L2 did not correspond to the Ivy GAP histology based anatomical features but it corresponds to IGN which resembles gliosis and fibrosis. GO enrichment analysis of L2 gene set shows the enrichment of wound response, vasculature development, ECM and myelin sheath supporting the presence of glial, vascular, fibroblasts and immune cells in IGN (Supplementary Data 9).

## GBM landscape visualized at sub-cellular resolution

The spatial resolution of Visium spatial transcriptomics (55 μm) enables the detection of gene expression patterns across tissue sections but has inherent limitations in resolving individual cells. To complement this, we employed the Xenium platform, which provides subcellular-resolution transcriptomic data, enabling spatial profiling at the level of individual cells. We designed a custom panel of 348 genes representing all major cell types (Methods, Supplementary Data 10, Fig. S5A) and used it to generate Xenium data from four GBM patients (SNU21, SNU33, SNU18, and SNU25) within our Visium cohort (Fig. 5A). This panel enabled spatial validation of 10 major cell types and their associated states, including previously uncharacterized subtypes of endothelial cells, pericytes, and oligodendrocytes.

We first annotated all cells using a semi-automated process, resulting in 729,307 annotated cells across a total area of 155 $mm^2$ (Fig. 5B–D). Gene expression profiles from the spatial single-cell data showed excellent correlation with our reference single-cell dataset using the 348 panel genes (Fig. 5E).

The Xenium data enabled analysis of anatomic features (AFs) across continuous regions spanning up to 1.2 cm. We selected 11 representative trajectories to capture transitions observed in the data (Fig. S5B). Grids were defined to match the area of a Visium spot (2376 $μm^2$). A subset of 40 genes (Supplementary Data 11) was chosen to assign AF labels to these grids based on known AF-specific cellular compositions (Fig. S5C). Spatial gene expression profiles were used to assign AFs to trajectory grids, with validation by H&E image visualization (Fig. 6 and Fig. S6). Representative areas for each AF with spatial transcripts overlaid on H&E image are shown in Fig. 6B.

Among these, IGN showed two distinct gene expression profiles. A subset of IGN grids displayed high expression of fibroblast and immune cell markers but lacked glial cell signatures. These were labeled as IFN (Immune Fibroblast Niche), corresponding to fibrotic scar. Fibrotic scarring was found to contribute to an immune-suppressive and pro-tumorigenic microenvironment that facilitates glioblastoma regrowth after therapy[27]. Remaining grids were assigned AF labels based on their shortest distance from the annotated trajectory grids (Fig. 7A).

We then assessed the cell type composition of AF-assigned grids (Fig. 7B–C). As expected, the cellular composition of AFs was consistent with findings from the Visium data (Fig. 7B–C). AF neighborhood analysis also revealed similar spatial relationships to those observed in the Visium dataset (Fig. 7D).

https://doi.org/10.1038/s41467-026-69716-2

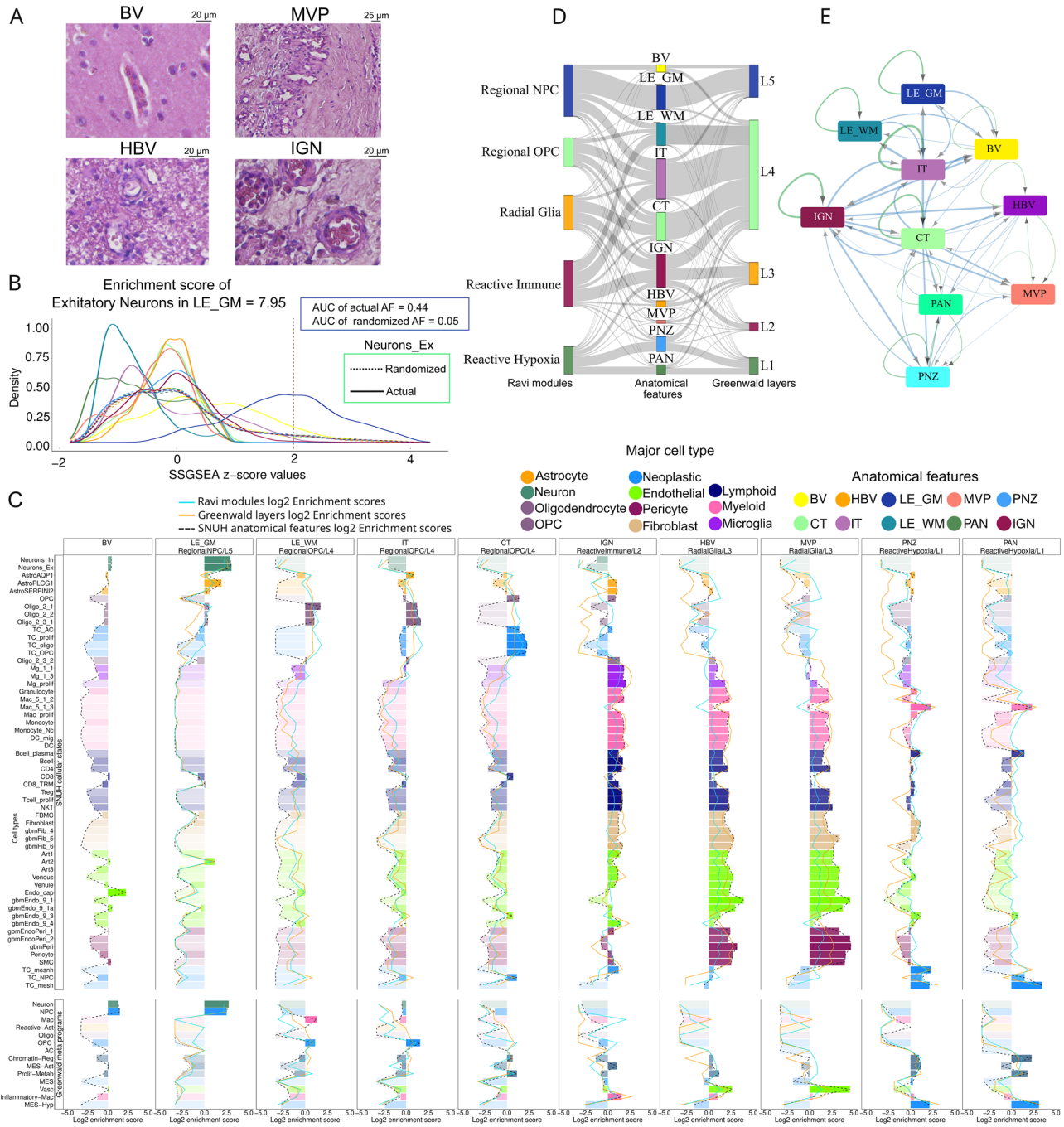



**Fig. 4 | Spatial mapping of cell types over glioblastoma AFs. A** Zoomed-in H&E representative snapshots highlighting distinct vascular morphologies observed in glioma samples, including BV, HBV, MVP, and IGN. Representative regions of interest are shown for illustrative purposes and reflect patterns observed consistently across the dataset. No independent experimental repetitions apply to these representative images. **B** Density plots showing z-scored ssGSEA values of cell subtype-specific marker genes across AFs. Solid lines represent actual AF assignments, while dotted lines represent results from shuffled randomized AF labels. The area under the curve (AUC) ratio (actual/randomized) quantifies preferential presence (ratio > 1) or absence (ratio <1) of specific subtypes within AFs. For example, Neurons_Ex exhibits a strong preferential presence in the LE_GM region. **C** Enrichment of our cell types and Greenwald et al metaprograms across AFs. Bar plots display the enrichment scores - defined as AUC ratios of actual versus randomized gene expression distribution, for cell types identified in the current study and meta-programs reported by Greenwald et al. AF panels are annotated with corresponding layers (Greenwald et al) and modules (Ravi et al). Overlayed lines indicate comparative enrichment: orange lines represent enrichment in Greenwald-defined layers, while cyan lines denote enrichment in Ravi-defined modules. **D** River plot depicting the correspondence between Greenwald layers and Ravi modules across anatomical features. The width of each stream represents the relative abundance and transition of features between layers and modules within distinct anatomical regions. **E** Network diagram from spatial neighborhood analysis, showing the degree of interaction and proximity between anatomical features. Edge thickness corresponds to the relative abundance of the anatomical features identified in the Visium dataset. Pairs with proportions at least three times higher than random are connected with green edges. Rest of the pairs are connected with blue edges. Scale bars: 20/25 μm.

https://doi.org/10.1038/s41467-026-69716-2

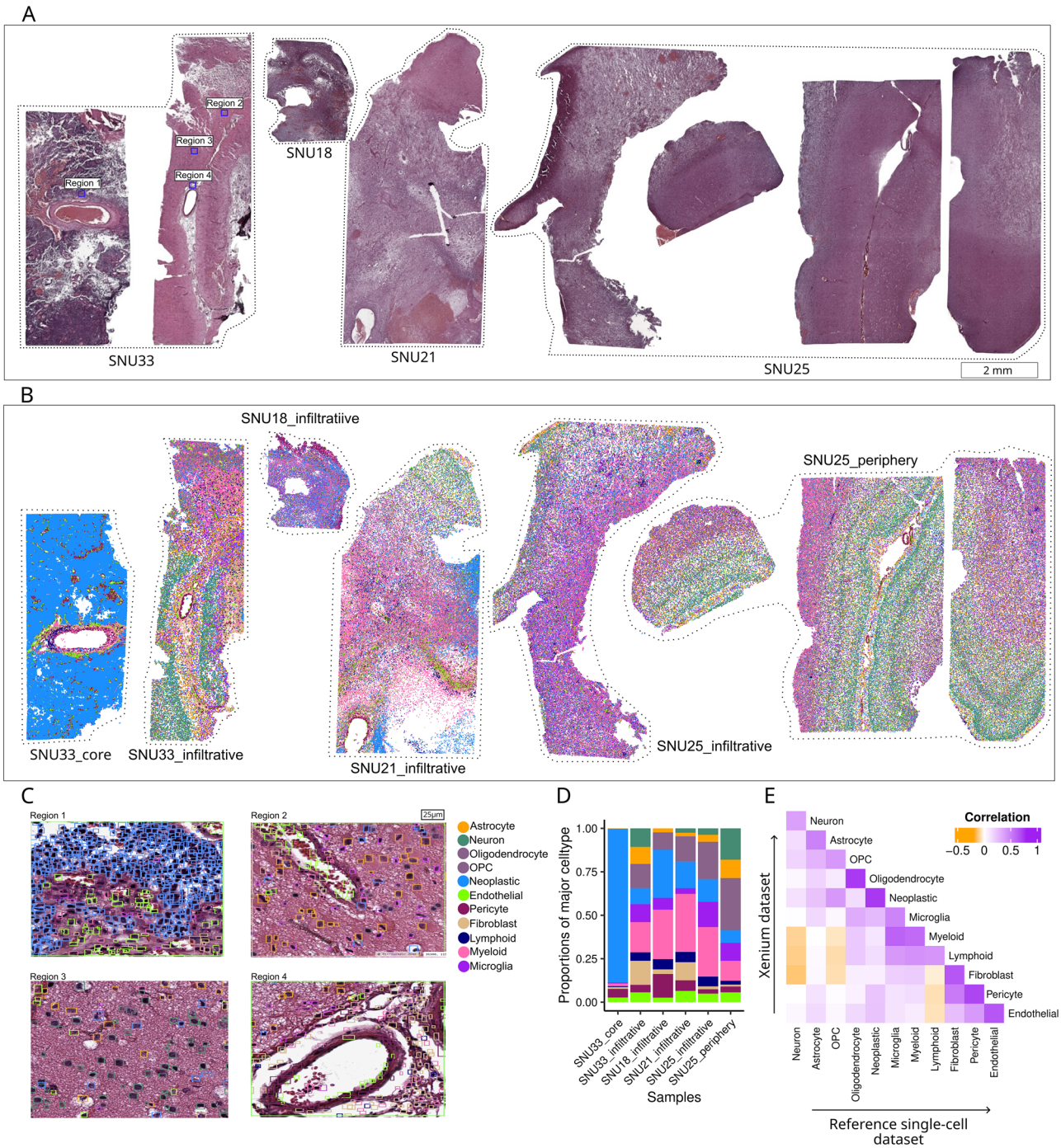



**Fig. 5 | Spatial mapping and validation of cell types in Xenium transcriptomic data. A** Representative H&E-stained tissue sections from SNU18, SNU21, SNU25, and SNU33 used for Xenium spatial transcriptomic analysis. These images are presented to document tissue coverage and sample inclusion and do not represent independently repeated experimental procedures. The marked region in the SNU33 section is shown at higher resolution in panel C, highlighting cellular heterogeneity. **B** Spatial distribution of major cell types across the four samples, derived from transcriptomic profiles and reference-based annotations. Each tissue core was further classified as cellular core, infiltrative, or periphery based on cellular composition and corresponding H&E morphology. **C** High-resolution zoomed-in views highlighting regions enriched for specific cell types, demonstrating spatial heterogeneity and localization patterns. The regions selected are highlighted in panel **A**. **D** Bar plots depicting the relative proportions of major cell types across all four tissue samples and their respective cores. **E** Pearson correlation analysis comparing the average expression profiles of major cell types identified in the Xenium dataset with those from a reference single-cell dataset, validating the accuracy of cell type assignments. Scale bar: 2 mm.

## Distinct spatial and molecular characteristics of oligodendrocytes in GBM

Emerging evidence indicates the pivotal involvement of oligodendroglia in various neurodegenerative conditions[21,22]. However, the precise extent of oligodendrocyte diversity and their role in GBM pathology remains elusive. In this study, we identified four distinct oligodendrocyte cell states within GBM (Fig. 2F) and delineated markers characterizing these states (Fig. S1A). Notably, Oligo_2_3_2 was uniquely detected in GBM, but not in IDH-mutant high-grade astrocytoma, particularly within the tumor core rather than the periphery

https://doi.org/10.1038/s41467-026-69716-2

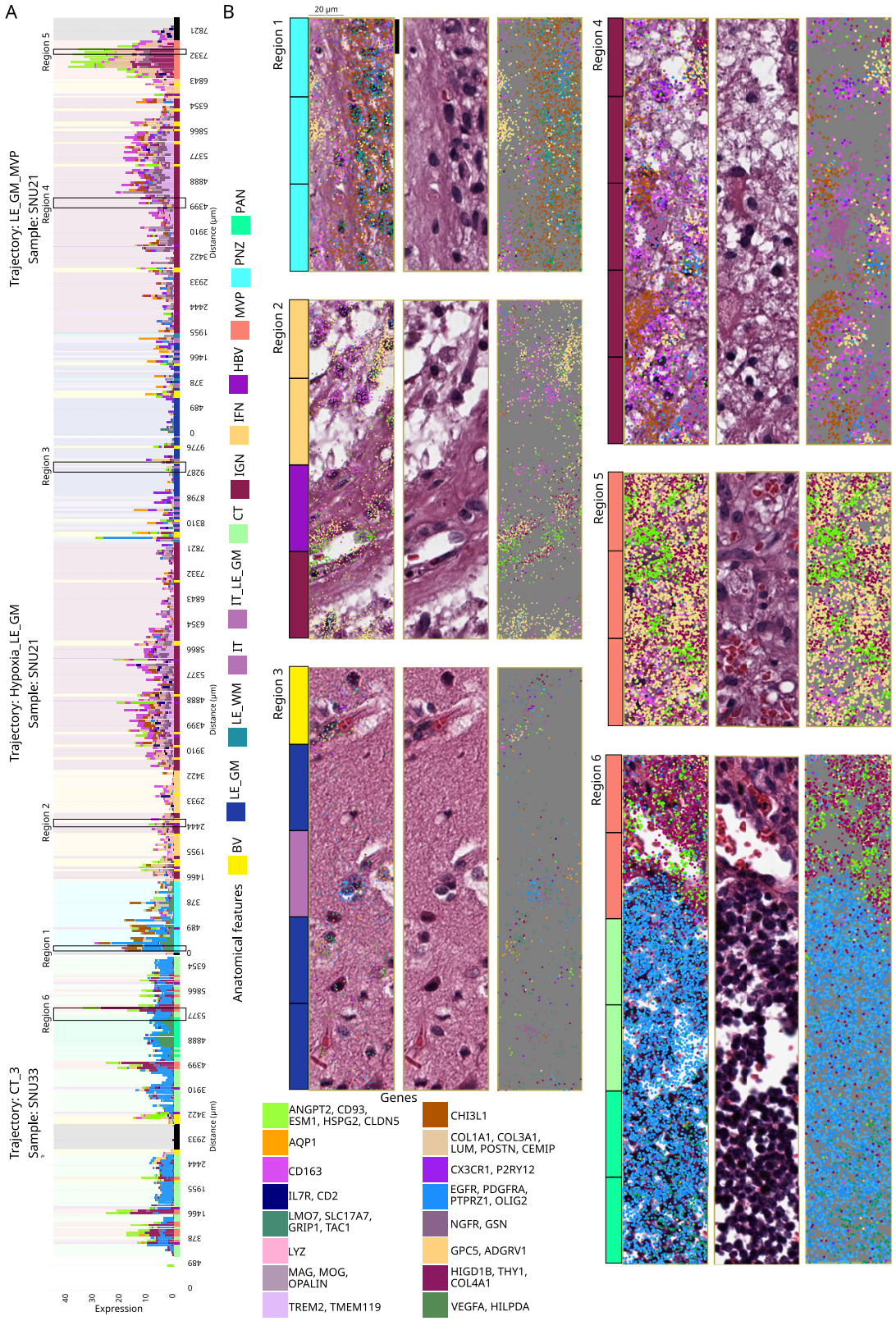



**Fig. 6 | Trajectory analysis on Xenium dataset. A** Distribution of transcript counts across AFs in three trajectories from SNU33 and SNU21, respectively. The bar plot shows the transcript count of canonical cell type marker genes in each grid. The area of the grid is set to mimic individual spots in Visium. The background highlights show the AF for each grid. The black boxes highlight the regions of interest, which are further zoomed in and shown. In total, eleven spatial trajectories were defined across four samples. Remaining trajectories are shown in Figure S6. **B** The spatial distribution of the transcripts on the H&E image. Six regions of interest from two trajectories are shown for visualization. The panel has three images per region of interest. The first image overlays the transcript expression on the H&E image. The second image is just an H&E image without transcript expression, and the third is just transcript expression without an H&E image. Scale bars: 20 μm.

https://doi.org/10.1038/s41467-026-69716-2

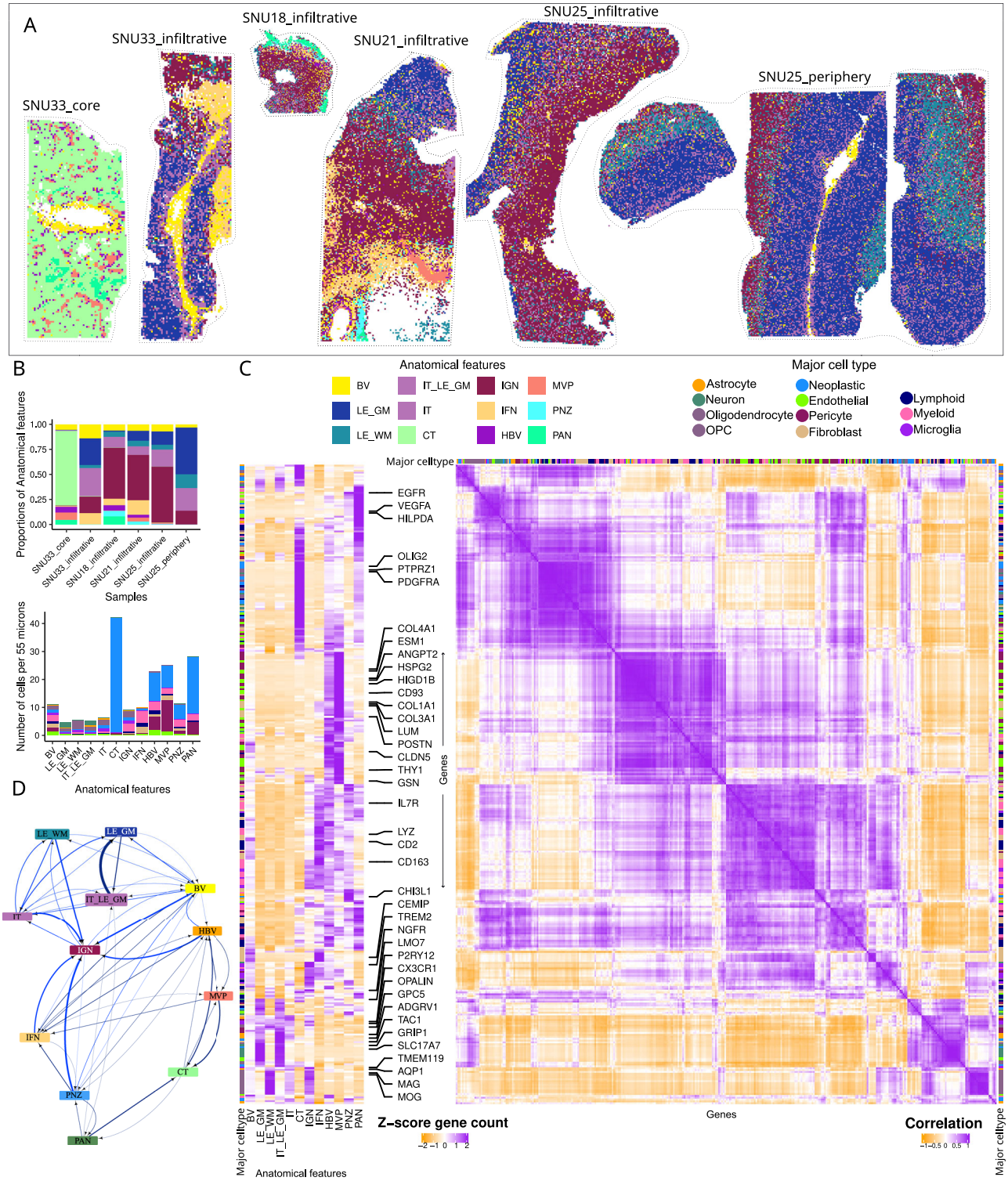



**Fig. 7 | Spatial characterization and interactions of anatomical features in the Xenium dataset. A** Spatial representation of anatomically defined features across Xenium-processed tissue sections, highlighting the spatial organization of glioma tissue architecture. **B** Bar plots showing the relative proportions of anatomical features across individual tissue samples. The accompanying plot shows the distribution of major cell types within each AF, illustrating the spatial composition and cellular heterogeneity of the tumor microenvironment. **C** Heatmap showing the average expression of genes from the glioma-specific Xenium gene panel across all AFs. Gene labels in the heatmap correspond to the 40 genes selected for anatomical feature assignment. A correlation matrix to the right illustrates co-expression patterns among genes, revealing modules of transcriptional co-regulation. **D** Network diagram derived from spatial neighborhood analysis, depicting the proximity and interaction strength between AFs in Xenium dataset. Edge thickness represents normalized neighborhood interaction scores, where thicker edges indicate stronger spatial co-occurrence or proximity between features.

(Fig. 8A). We observed the same in another dataset (Fig. S7A)[13]. To investigate the potential roles of Oligo_2_3_2 in GBM, we first examined the differentially expressed genes (DEGs) between Oligo_2_3_2 and the other oligodendrocyte subtypes using both single-cell and pseudo-bulk transcriptome datasets (Fig. 8B, Fig. S7B; Supplementary Data 12). Key upregulated genes in Oligo_2_3_2 were various immune-related genes including *GSN, TUBB2B, HLA-A, ALDOA, CLU, TIMP1, S100A1, SERPINA3*, and *NGFR* (collectively referred to as the Oligo_2_3_2 gene set). On the other hand, genes significantly downregulated in Oligo_2_3_2 were myelin-related genes such as *OPALIN, KIF19, ALDOC, PCDH9*, and *DOCK9* (referred to as the Oligodendrocytes gene set). Gene ontology (GO) analysis highlighted enrichment in antigen processing and presentation and terms related to metabolic shift towards glycolysis in Oligo_2_3_2 genes, while myelination and neural projection development in Oligodendrocyte genes (Fig. 8C). We further analyzed enrichment of these two gene signatures using ssGSEA[28] in bulk RNA-seq dataset from TCGA[29] and CGGA[30]. In both datasets, we observed the trend of increased enrichment of Oligo_2_3_2 gene set from IDHmut, 1p19q co-deleted gliomas to IDHmut, 1p19q non-co-deleted gliomas to IDHwt gliomas. Similarly, we observed the opposite trend for Oligodendrocytes gene set (Fig. S7C).

To investigate the spatial organization of Oligo_2_3_2 associated marker genes at single cell resolution, we analyzed the 10x Xenium dataset. Consistent with the single-cell data, we found oligodendrocyte populations with high expression of *SERPINA3, HLA-A, TUBB2B, TIMP1, CLU, HLA-DRA*, and *S100A1* and low *OPALIN* transcripts exclusively in the tumor core in Xenium data consistent with single-cell data (Fig. 8D–E). To identify cell types co-localized with Oligo_2_3_2 gene sets we analyzed both Visium and Xenium dataset. Correlation analyses of Visium dataset show that besides immune and stromal cell signatures, TC_mesnh signature is highly co-expressed with Oligo_2_3_2 genes in comparison to other subtypes of tumor cells (Fig. 8F). Using higher resolution Xenium data, we see that TC_mesnh and Mac_5_1_2 cells are found within a 20 μm radius of the 90% of Oligo_2_3_2 cells (Fig. S7D). In contrast Mg_1_3 and normal Endo_cap are in close proximity of Oligo_2_1 cells (Fig. S7D). In summary, Oligo_2_1 exhibited closer spatial association with normal brain cell types, including Mg_1_3 (inactive microglia), pericytes, and Endo_cap (normal endothelial cells). In contrast, Oligo_2_3_2 was more proximally located to TC_mesnh (neoplastic mesenchymal non-hypoxic cells), Mg_1_1 (active microglia), Mac_5_1_2, and glioblastoma-associated stromal and endothelial cell types (Fig. S7D). Interestingly, Oligo_2_3_2 does not show spatial proximity to TC_mesh (neoplastic mesenchymal hypoxic cells) cells. Our Visium ST data showed a mutually exclusive expression of TC_mesh and Oligo_2_3_2 signatures (Fig. S7E). Greenwald et al.[15] used a 40-marker antibody panel (CODEX) to characterize spatial organization of tumor cell states especially in the context of hypoxia at single-cell protein level. They annotated 428,395 cells from 12 samples. This analysis identified all major differentiated non-malignant cell types, including astrocytes, oligodendrocytes, neurons, vascular cells, T cells, B cells, and macrophages/microglia and key malignant cell states such as MES-like (TC_mesnh), MES-Hyp (TC_mesh), Chromatin-Reg, OPC-like (TC_OPC), NPC-like (TC_NPC), and AC-like (TC_AC). Their marker panel did not have markers to identify different oligodendrocyte subtypes but we were able to see that the GBM core samples with necrosis and high number of TC_mesh cells had almost none to few oligodendrocytes (Fig. S7F) validating the depletion of oligodendrocytes in hypoxic regions.

In order to validate Oligo_2_3_2 gene markers at protein level, we performed IHC for myelin marker Mag, actin cytoskeleton remodeling marker Gsn, and neuronal marker Ngfr on matching tissue from patients SNU33, and SNU21 used for Visium and Xenium samples. Based on Visium and Xenium spatial transcriptomic data Oligo_2_3_2 co-express these three markers. The gene and protein expression levels show high concordance for all three markers, as can be seen in the side-by-side Visium spatial gene profile and IHC protein expression where Visium equivalent tissue is highlighted in blue rectangle (Fig. S8). SNU33 showed highly heterogeneous and AF specific patterns. We analyzed three distinct regions (LE/IT, IGN, and CT) in SNU33 sample using the Oligo_2_3_2 markers, MAG, GSN, and NGFR. This was done for both Xenium and equivalent tissue areas in IHC dataset. Major cell type distributions within each region revealed a strong enrichment of neurons, and astrocytes in the LE/IT region, oligodendrocytes, macrophages, and endothelial cells in the IGN region, and neoplastic cells in the CT region (Fig. S9A). Notably, spatial overlay analysis confirmed the differential presence of *MAG, GSN*, and *NGFR* transcripts within these regions, aligning with expected cell type distributions, highest expression of all three markers in oligodendrocytes within the IGN region (Fig. S9B–C). The same pattern is observed in IHC data (Fig. S9D–E). We then performed IHC on an additional set of 11 patient samples (three grade 3 oligodendrogliomas, one grade 3 astrocytoma, two grade 4 astrocytomas and 5 glioblastomas). Four out of five glioblastoma cases showed same IGN pattern whereas none of the six IDH mutant cases exhibited such pattern (Fig. S10).

We further validated the expression of protein key markers for Oligo_2_3_2 such as S100a1, Serpina3n, B2M (MHC class I subunit), *Opalin, and Sox10* as a marker for oligodendrocyte lineage cells and CRE as a marker for tumor cells, using immunofluorescence of murine high-grade glioma in situ[31]. Glioma-associated oligodendrocytes - identified as Sox10+CRE- cells - were found within the tumor core and periphery, including infiltrating tumors. Quantifications of marker-positive cells revealed higher proportions of Sox10+ cells positive for S100a1 and Serpina3n in the tumor core, compared to normal mouse brain and the tumor periphery, including infiltrative regions. Sox10+ cells with Opalin expression show a decreasing trend in the tumor core without reaching significant differences (Fig. S11).

### Clinical implications of GBM associated oligodendrocytes

To explore the clinical significance of the GBM associated oligodendrocyte (Oligo_2_3_2) markers, we examined multiple publicly available datasets, including single-cell and bulk transcriptomic data with associated clinical outcomes. To assess the prognostic value of the Oligodendrocyte and Oligo_2_3_2 gene set scores in the context of established clinical and molecular parameters, a multivariable Cox proportional hazards regression analysis was performed using both the TCGA ($n = 603$) and CGGA ($n = 1018$) glioma cohorts. The models incorporated age, IDH mutation and 1p/19q co-deletion status, and MGMT promoter methylation status alongside the two transcriptional signatures (Fig. 9A). In both TCGA and CGGA cohorts, the Oligodendrocyte gene set score was significantly associated with favorable prognosis, whereas the Oligo_2_3_2 gene set score showed a strong association with poor overall survival. These associations remained consistent and independent of key clinical variables, including age, IDH mutation/1p19q co-deletion status, and MGMT promoter methylation, highlighting the distinct prognostic relevance of these transcriptional programs in glioma. While multivariable Cox regression strengthens the prognostic analysis, it remains challenging to completely disentangle IDH-associated transcriptional effects from subtype-specific gene programs in bulk RNA datasets, and residual confounding cannot be completely excluded.

We then evaluated the prognostic significance of the Oligodendrocytes and Oligo_2_3_2 gene signatures using TCGA and CGGA, by classifying patients into four groups based on the expression of the respective gene signatures. The analysis revealed that patients with high scores for the Oligo_2_3_2 gene set and low scores for Oligodendrocytes gene set exhibited significantly poorer survival outcomes, whereas those with high scores for the Oligodendrocytes gene set and low scores for the Oligo_2_3_2 gene set demonstrated the most favorable prognosis (Fig. 9B). Consistent trends were observed when patients were further stratified by IDH mutation status and 1p/19q co-

https://doi.org/10.1038/s41467-026-69716-2

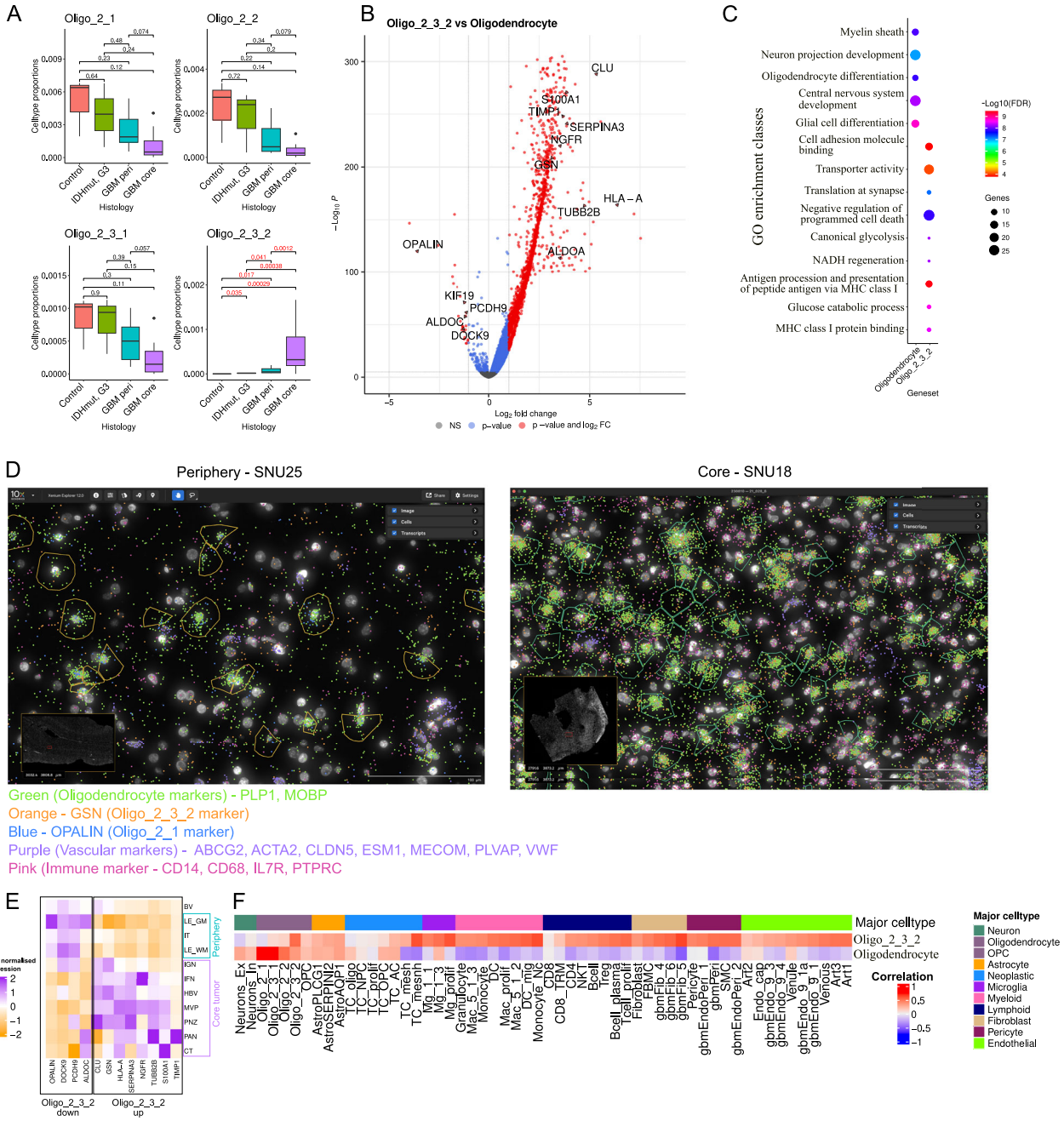



**Fig. 8 | Molecular, spatial, and clinical characterization of oligodendrocyte subtypes. A** Bar plots illustrating the correlation between oligodendrocyte subtypes and brain tumor types. The relative proportions of oligodendrocyte cell subtypes across histology reveal an increased presence of Oligo_2_3_2 subtype cells in tissue from the GBM core ($n = 16$) compared to periphery ($n = 7$), control ($n = 3$), and IDH mutant G3 ($n = 3$) tissue samples (depicted in purple). To identify significant proportional differences, unpaired two-sided t-test was used. Significant $p$-values (<0.05) are highlighted in red. Each data point represents cell type proportions aggregated per patient sample. In boxplots, the centre line represents the median, the box spans the interquartile range (25th–75th percentiles), and whiskers extend to the most extreme data points within 1.5 * *IQR*. with values beyond this limit plotted as individual outliers. **B** Volcano plot depicting genes significantly upregulated or downregulated in the Oligo_2_3_2 subtype compared to all other oligodendrocyte subtypes. Differential expression analysis was performed using the clusterDE package in R to identify genes specifically enriched in the Oligo_2_3_2 subpopulation. *P* values were adjusted for multiple comparisons using the Benjamini–Hochberg procedure, and genes with a false discovery rate (FDR) < 0.05 were considered significantly differentially expressed. **C** Gene Ontology enrichment analysis of differentially expressed genes (DEGs) between Oligo_2_3_2 and Oligodendrocyte subtypes. **D** The spatial distribution of specific oligodendrocyte subtype markers in LE (Periphery) and IGN/PNZ (Core) regions. Green dots represent oligodendrocyte markers (*PLP1* and *MOBP*), orange dots indicate expression of the Oligo_2_3_2 marker (GSN), blue dots denote expression of the Oligo_2_1 marker (*OPALIN*), purple dots highlight vascular markers (*ABCG2, ACTA2, CLDN5, ESM1, MECOM, PLVAP*, and *VWF*), and pink dots represent immune cell markers (*CD14, CD68, IL7R*, and *PTPRC*). **E** Average normalized gene expression of Oligo_2_3_2 and Oligodendrocyte genesets across anatomical features in Xenium dataset. **F** Correlation of Oligo_2_3_2 and Oligodendrocyte gene sets with other cell type gene sets in Visium dataset. Scale bars: 100 μm.

https://doi.org/10.1038/s41467-026-69716-2

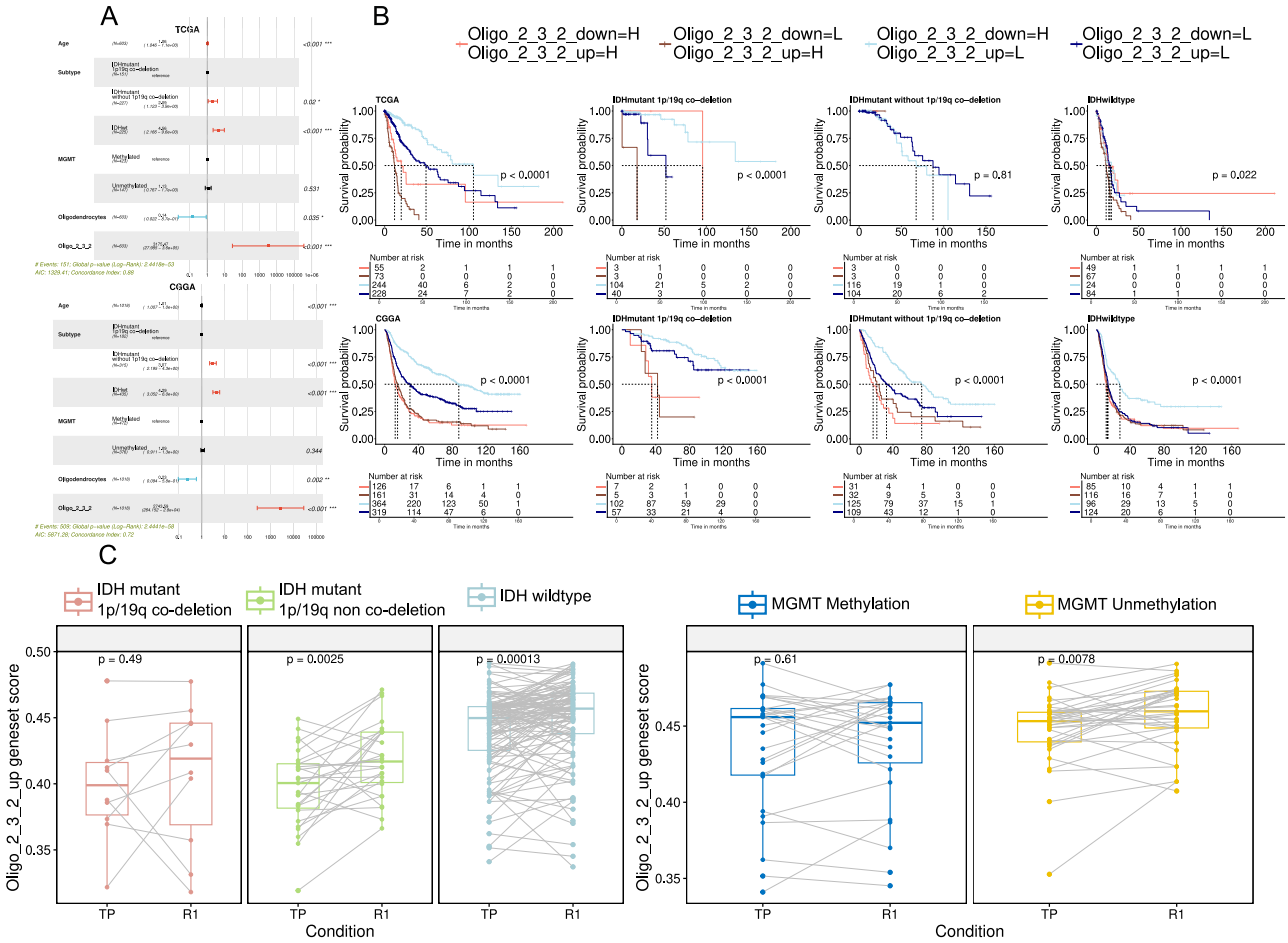



**Fig. 9 | Prognostic relevance of Oligodendrocyte subtypes. A** Multivariable Cox proportional hazards analysis in TCGA ($n = 1018$) and CGGA ($n = 603$) glioma cohorts. Forest plots show hazard ratios (HR) with 95% confidence intervals for key molecular and clinical variables. **B** Kaplan–Meier survival curves stratifying patients based on ssGSEA scores of the Oligodendrocytes and Oligo_2_3_2 gene sets in the CGGA and TCGA datasets. Survival profiles are shown across four clinically relevant subgroups: the entire cohort, IDH-mutant with 1p/19q co-deletion, IDH-mutant without 1p/19q co-deletion, and IDH-wildtype patients. Significant survival differences suggest prognostic relevance of the Oligo_2_3_2 gene program across molecular glioma subtypes. Statistical differences between survival curves were assessed using a two-sided log-rank test, with exact *P* values reported. **C** Boxplots showing the expression levels of the Oligo_2_3_2 gene set in primary versus recurrent glioblastoma (GBM) samples from the GLASS cohort[31]. Samples are stratified by molecular subtype: IDH-mutant/1p19q codeleted ($n = 10$), IDH-mutant/1p19q non-codeleted ($n = 27$), and IDH-wildtype ($n = 117$). Additional stratification by MGMT promoter methylation status is shown ($n = 27$ methylated, $n = 35$ unmethylated). Each data point represents one independent patient sample. A significant increase in Oligo_2_3_2 gene set expression is observed in recurrent tumors, particularly in IDH-mutant/1p19q non-codeleted and IDH-wildtype patients. A similar trend is observed in tumors lacking MGMT promoter methylation, suggesting a potential role of this gene program in therapy resistance and tumor recurrence. In boxplots, the centre line represents the median, the box spans the interquartile range (25th–75th percentiles), and whiskers extend to the most extreme data points within 1.5 * *IQR*. with values beyond this limit plotted as individual outliers. Significance was assessed using two-sided paired t-test.

deletion status, underscoring the prognostic relevance of these gene sets across molecularly defined glioma subgroups (Fig. 9B). The bulk transcriptomic data (GLASS dataset)[32] from matched primary (TP) and recurrent (R1) gliomas, was analyzed to determine the dynamic changes in Oligo_2_3_2 gene signatures at recurrence. ssGSEA scores demonstrated that genes upregulated in Oligo_2_3_2 exhibited increased expression in recurrent gliomas in both IDH-mutant, 1p/19q-codeleted and IDH-wildtype samples (Fig. 9C). Additionally, within IDH-wildtype tumors, recurrent samples with unmethylated MGMT promoters showed elevated expression of the Oligo_2_3_2 gene set, suggesting a potential association with treatment resistance and tumor recurrence. At single-cell level this was observed in Abdelfatteh et al. dataset where we see higher prevalence of Oligo_2_3_2 in an unmethylated recurrent tumor (Fig. S7A).

## Discussion

This study provides a high-resolution, spatially resolved atlas of the GBM tumor microenvironment by integrating single-cell and spatial transcriptomics across both fresh frozen and FFPE tissue samples. By using FFPE-compatible protocols and applying multi-platform spatial profiling technologies (Visium and Xenium), we expand the applicability of spatial biology into clinically relevant and diagnostically compatible formats. This approach enabled us to chart the spatial architecture of GBM at comparatively higher resolution and depth.

Unlike prior GBM atlases that emphasized malignant and immune components[13], our work highlights previously underappreciated but functionally critical glial and stromal compartments. Notably, we uncovered a distinct immune oligodendrocyte population, Oligo_2_3_2, enriched in the immune–glial niche (IGN), a transcriptionally defined, spatially restricted domain marked by immune, glial, vascular, and fibrotic elements. This population demonstrates an immune-activated, non-myelinating phenotype and is absent in IDH-mutant gliomas, showing spatial proximity to mesenchymal tumor cells (TC_mesnh) and immunosuppressive macrophages. These findings establish a functional context for Oligo_2_3_2 within the immunomodulatory landscape of the GBM core.

Strikingly, Oligo_2_3_2 mirrors the disease-associated oligodendrocytes (DAOs) described in multiple sclerosis and neurodegeneration models, cells that exhibit immune reprogramming and spatial restriction to diseased areas[33]. This suggests a broader neuroinflammatory mechanism in which oligodendrocytes undergo dysfunctional state transitions across CNS diseases. The presence of Oligo_2_3_2 in recurrent and MGMT-unmethylated tumors, as well as its correlation with adverse clinical outcomes, further implies a role in therapy resistance and recurrence.

Beyond glial contributions, we delineated a second spatial domain, an immune–fibrotic niche (IFN), which shares features with fibrotic scars described in GBM following anti-CSF1R therapy[27]. These regions, characterized by fibroblast accumulation and immune suppression, may act as sanctuary sites promoting tumor survival under therapeutic pressure. The IFN and IGN together reflect dynamic remodeling of the tumor microenvironment, involving wound healing, immune modulation, and stromal support.

The translational implications are substantial. Our approach, compatible with archival FFPE material, allows development of digital spatial profiling across retrospective clinical cohorts. Furthermore, the parallels between GBM and broader CNS pathologies, such as fibrotic remodeling and glial immunoregulation point to shared therapeutic vulnerabilities. Targeting programs like TGF-β signaling in IFN or immune-activated oligodendrocytes in IGN could open new avenues for combination therapy.

Several limitations must be acknowledged. While we map spatial associations of Oligo_2_3_2, IGN, and IFN, we do not demonstrate causal mechanisms underlying their emergence. It remains to be determined whether these niches drive tumor behavior or simply reflect reactive changes. Similarly, although Oligo_2_3_2 correlates with recurrence and poor survival, its direct contribution to tumor progression requires functional validation in experimental models.

Future studies should include perturbation experiments to dissect the role of these cell types and niches in GBM biology. For instance, whether modulating TGF-β–driven fibrosis or altering the immune state of oligodendrocytes affects tumor trajectory remains an open and important question. Our atlas provides a foundational platform to prioritize such targets and evaluate them within their spatial and cellular contexts.

In summary, our atlas redefines the spatial and cellular complexity of GBM, emphasizing glial and fibrotic programs as key components of tumor ecology. By mapping the context-specific interactions between tumor, stroma, and immune cells, we hypothesize that targeting these interacting cell populations in combination, rather than in isolation, may be more effective, a concept our atlas can help guide.

## Methods

### Patient samples

This study was conducted in accordance with all relevant ethical regulations. Tissue samples for this study were obtained from patients undergoing surgical resection with informed consent for their usage for research purposes in accordance with the guidelines of the Institutional Review Board of Seoul National University Hospital, which approved this study (IRB Nos. H-0507-509-153 and H-2010-123-1166). No individually identifiable information is reported in this manuscript. All clinical and demographic data are presented in anonymized and aggregated form, such that individual participants cannot be identified. Sex and gender information were collected via patient self-reports. Disaggregated data for all individual experiments are included in the supplementary data submitted with the study. Overall distribution by sex is reported in the Supplementary Data 1. Consent for reporting and sharing individual-level disaggregated data was obtained from all participants prior to study commencement. No sex-based stratified analyses were performed, as the study was not designed or powered to assess sex-specific effects and analyses focused on tumor-intrinsic transcriptional programs. No financial or material incentives were provided to patients who donated their samples. These samples were entirely consumed during the process of genetic sequencing and data acquisition; consequently, no biological material remains. All resulting genomic data, however, are available in public database or upon request.

### Animals

All animal procedures were performed in accordance with institutional and federal guidelines (Stanford University APLAC protocol # 33633). Mice were housed in groups of three to five and maintained on a 12-h light/dark schedule. Food and water were available *ad libitum*. Mice were fed a standard chow diet (Inotiv, #2018, Tekland Global 18% protein rodent diet) starting at weaning (p21). Compound transgenic mice were generated by crossing Cre-dependent BRAF CA with Ink4a/Arf flox/flox mice, as previously described (Xing et al 2025). *BRAF* $^{CA/WT}$*; Ink4a/Arf* $^{flox/flox}$ mice were bred on a pure C57Bl/6 genetic background in our animal facility to produce sufficient mice for experiments.

No formal statistical methods were used to predetermine sample sizes. Sample sizes were determined by the availability of high-quality human tumor specimens and experimental material generated in this study, consistent with standard practice in exploratory spatial transcriptomic and histopathological analyses. Sample sizes for spatial transcriptomics and histological experiments reflect the number of biologically independent patient samples meeting predefined quality-control criteria. For experimental validation and animal studies, sample sizes were chosen based on prior studies in the field and feasibility considerations and were sufficient to observe reproducible effects across biologically independent samples. No data generated in this study were excluded from analysis on the basis of sample size.

### Intracranial injections of adenovirus-Cre

*BRAF*$^{CA/WT}$*; Ink4a/Arf* $^{flox/flox}$ mice were injected with 2 μl of adenovirus expressing cre recombinase under the control of the ubiquitous CMV promoter (Ad-Cre; Titer = 1×10^9 PFU/ml; Vector Biolabs, Malvern, PA) to induce BRAF V600E expression and deletion of tumor suppressor genes p16 $^{Ink}$ and p19 $^{Arf}$. The stereotaxic coordinates selected to target the cortical layer, relative to bregma, are anterior-posterior−1.00 mm, medial-lateral−1.80 mm, and dorsal-ventral−1.5 mm. Experimental mice were monitored and humanely sacrificed in accordance with institutional animal ethics committee-approved humane endpoints, which defined the maximal allowable tumour burden as the loss of more than 15% of original body weight and/or the development of clinical signs consistent with intracranial tumor formation, such as hydrocephalus, lethargy, and hunched posture, and these limits were not exceeded in any animal. Brains of symptomatic mice were collected, and sectioned and tumor-bearing brains were subjected to immunofluorescence staining.

### Spatial transcriptomics (ST)

ST experiments were performed using the 10X Visium Spatial Gene Expression kit, following manufacturer's instructions. Here, we briefly describe the methods.

### Sample preparation (FFPE patient samples)

A total of 2 FFPE sections (5um each) per sample were used to determine RNA integrity. Total RNA was extracted using RNA Isolation Kit (Qiagen, 73504) according to the manufacturer's protocol. DV200 (the percentage of total RNA fragments >200) were determined using TapeStation (HS RNA Screen Tape, Agilent, 5067-5579) according to the manufacturer's protocol. Only samples with DV200 ≥ 50% were qualified and used.

### Deparaffinization, staining, imaging, and FFPE library construction

5 μm thick sections were immediately placed on Visium array slide (Visium Spatial Gene Expression slides, 10× Genomics). Slides were deparaffinized and stained with hematoxylin and eosin. Brightfield imaging was carried out at 100x magnification with Nikon Eclipse Ti2, and post-processing was performed using NIS software. After imaging, spatially barcoded, ligated probe products were amplified using KAPA SYBR FAST qPCR Master Mix (Roche, KK4600). Fragments in the size of interest were selected using SPRIselect reagent (Beckman Coulter, B23318). Library quality check was performed using TapeStation (HS D1000 ScreenTape, Agilent, 5067-5584).

### Sample preparation (Fresh frozen patient samples)

Fresh GBM tissue was collected immediately after resection and embedded in cryomolds using Tissue Tek OCT (Sakura, 4583) using a bath of isopentane and liquid nitrogen. Following freezing in OCT, blocks were stored at −80 °C until processing. A total of 2 sections (10um each) per sample were used to determine RNA integrity. Total RNA was extracted using RNA Isolation Kit (Qiagen, 74104) according to the manufacturer's protocol. RIN values were determined using TapeStation (HS RNA Screen Tape, Agilent, 5067-5579) according to the manufacturer's protocol. Only samples with an RNA integrity value > 7 were qualified and used.

### Sample preparation (Fixed frozen mouse samples)

Mice were intraperitoneally injected with ketamine (100 mg/kg) and xylazine (10 mg/kg) to induce deep anesthesia and subsequently transcardiac perfusion was performed with phosphate-buffered saline (PBS) followed by 4% paraformaldehyde (PFA)/PBS. Brains perfused with 4% PFA were then resected, post-fixed in 4% PFA/PBS for 2 h on ice, cryoprotected in 15% sucrose/PBS for 24 h and then 30% sucrose/PBS for at least 48 h, followed by embedding in Tissue-Tek® O.C.T. Compound (Sakura Finetek USA, Inc., Torrance, CA). The embedded tissues were stored at −80 °C until sectioned. 10μm-thick coronal sections of the brain were cut on a cryostat (Leica, Wetzlar, Germany), collected onto Superfrost Plus slides (Menzel Glaser, VWR) and air dried for 1 h before storing at −80 °C until stained.

### Fixation, staining, imaging, and construction of fresh frozen cDNA libraries

10 μm thick sections were immediately placed on Visium array slide (Visium Spatial Gene Expression slides, 10× Genomics). Slides were fixed in 100% methanol, stained with hematoxylin and eosin. Brightfield imaging was carried out at 100x magnification with Nikon Eclipse Ti2, and post-processing was performed using NIS software. After imaging, sections were permeabilized to release and capture mRNA to oligonucleotides on the capture areas. Following permeabilization, the on-slide reverse transcription (RT) reaction and second strand synthesis was performed. Spatially barcoded cDNA was amplified using KAPA SYBR FAST qPCR Master Mix (Roche, KK4600). Fragments in the size of interest were selected using SPRIselect reagent (Beckman Coulter, B23318). Library quality check was performed using TapeStation (HS D1000 ScreenTape, Agilent, 5067-5584). To optimize cDNA fragments for Illumina NextSeq, fragmentation and double-sided size selection was performed using SPRIselect reagent.

### Immunofluorescence staining on frozen mouse brain sections

Frozen 10 μm mouse brain sections on Fisherbrand™ Superfrost™ Plus Microscope Slides (ThermoFisher Scientific) slides were washed briefly in PBS and blocked with PBS containing 0.3% Triton X-100 and 10% normal donkey serum for 1 hr at room temperature. The sections were then incubated with primary antibodies at room temperature overnight in a humidified chamber, washed briefly in PBS three times next day, followed by 1 h incubation with secondary antibodies, the slides were then washed three times with PBS, and coverslipped with a DAPI-containing mounting solution (VectaShield H20002; Vector Laboratories, Newark, CA) and subjected to confocal microscopic analysis. Some primary antibodies were incubated simultaneously. The following primary antibodies were used: mouse anti-Cre recombinase (1:500; MAB3120, Millipore Sigma, Burlington, MA), rabbit anti-Serpina3n (1:200; MA5-41236, ThermoFisher Scientific), rabbit anti-Opalin (1:300; ab121425, Abcam), rabbit anti-S100a1 (1:300; PA1-932, ThermoFisher Scientific), rabbit anti-B2M (1:500; PI701250, ThermoFisher Scientific) and goat anti-Sox10 (1:100; AF2864, R&D Systems). Secondary antibodies raised in donkey and conjugated to Alexa Fluor 488, Alexa Fluor 594 or Alexa Fluor 647 were purchased from Jackson ImmunoResearch Labs (West Grove, PA) or Invitrogen (now ThermoFisher Scientific) and used at 1:200 dilution.

### Confocal imaging, image processing and statistical analysis

Stained 10 μm thick coronal sections were imaged by laser scanning confocal microscopy (Zeiss LSM700) which was used to detect up to four fluorophores by laser excitation at 405, 488, 561, and 639 wavelengths. Confocal images were acquired using a 20X objective. Images were imported into Fiji/ImageJ software for quantification of cellular density in the regions of interest. All analyses were performed in a blinded fashion. Statistical analysis of quantitative results was performed using GraphPad Prism software. Statistical significance was determined using two-way ANOVA with Sidak's multiple-comparison test. Quantitative data are reported as mean ± SEM, and any values of $P < 0.05$ were considered statistically significant.

### Immunohistochemical staining on tumor tissues from human patients

Formalin-fixed paraffin-embedded tumor tissues were sectioned at 4 μm thickness for IHC. IHC staining was performed using the BenchMark ULTRA system (Ventana-Roche, Manheim, Germany) with antibodies for GSN (PA5-29910, Invitrogen), MAG (D4G3, Cell Signaling), and NGFR (MA5-31968, ThermoFisher Scientific).

### ST sequencing

The average length of the final libraries was quantified using TapeStation (HS D1000 ScreenTape, Agilent, 5067-5584). The paired-end sequencing was performed on the Illumina NextSeq 550 using the following read protocol; 28 cycles for read 1; 10 cycles for i7 index read: 10 cycles for i5 index read; 90 cycles for read 2.

### Single-cell RNA-sequencing

**Tissue dissociation.** Tissue specimens were collected and immersed in RPMI media at room temperature immediately after surgical resection. The MACS brain tumor dissociation kit (Miltenyi Biotec, Auburn, CA, USA) and gentleMACSTM Dissociator (Miltenyi Biotec) were used to dissociate the tissue samples within 2 h after collection, and debris was removed using MACS Debris Removal Solution (Miltenyi Biotec) according to the manufacturer's protocol. Only the samples with cell viability > 80% after dissociation were used in this study. Isolated single cells were cryopreserved until further use.

**Construction of cDNA libraries and sequencing.** The library preparation was performed using Chromium Next Gem Single-cell 5′ kit v2 (10x Genomics, USA, 1000263) with a cell recovery target of 10,000, following the manufacturer's instruction. The libraries were processed according to manufacturer's recommendation (10x Genomics, USA). The libraries were sequenced on a HiSeq 2500 system using 100-bp paired-end sequencing, and the raw sequencing data was processed using Cell Ranger v5.0.1 with default setting. The reads were aligned to the human reference genome (hg38).

**CITE-seq.** Approximately 1 million cells were incubated with Fc receptor blocking solution (Human Trustain FcX™, BioLegend, 422301) at 4'C for 10 min and were stained with TotalSeq™-C Universal Cocktail, V1.0 (BioLegend, 399905) at 4'C for 30 min. After a centrifuge (600 g, 4'C, 5 min), cells were resuspended with 3 ml of Cell Staining Buffer (BioLegend, 420201). Cells were washed twice more and resuspended in 500 ul of Cell Staining Buffer. Antibody-stained cells were processed by Chromium controller (10x Genomics, USA) with a cell recovery target of 5000 for the generation of gene expression libraries and Feature Barcode libraries. Gene expression libraries and Feature Barcode libraries were prepared using Chromium Next Gem Single-cell 5' kit v2 (10x Genomics, USA, 1000263) and 5' Feature Barcode kit (10x Genomics, USA, 1000256) following the manufacturer's instruction. The libraries were sequenced on HiSeq 2500 system using 100-bp paired-end sequencing yielding 25 K reads per cell for gene expression libraries and 5 K reads per cell Feature Barcode libraries, respectively. The raw sequencing data was processed using Cell Ranger v5.0.1 with default setting. The reads were aligned to the human reference genome (hg38).

### Quantification and statistical analysis

**Preprocessing of 10x files.** The 10x Visium data was aligned using Space Ranger, and the Cell-seq/CITE-seq data was aligned using the Cell Ranger software. The following are the commands used,

Cell Ranger command for CITE-seq:

*cellranger count –id = [sampleid] citeseq --libraries=library.csv --transcriptome=refdata-gex-GRCh38-2020-A*

*--feature-ref=TotalSeqC_universal_GBM_feature_ref.csv*

Cellranger command for Cell-Seq:

*cellranger count --id = [sampleid] --fastqs = [fastq file location] --sample = [sample name] --transcriptome = refdata-gex-GRCh38-2020-A*

Spaceranger command for Visium:

*spaceranger count --id = [sampleid] --transcriptome = refdata-gex-GRCh38-2020-A/ --probe-set =Visium_Human_Transcriptome_Probe_Set_v1.0_GRCh38-2020-A.csv --fastqs = [fastq file location] --sample = [sample name] --image=image.tif --slide = [slide name] --area = A1 --localcores = 32 --localmem = 64*

**Data normalization of single-cell and spatial transcriptome datasets.** Raw count matrices (raw_feature_bc_matrix) were used for downstream analysis. The data was normalized using the counts per million (CPM) method, followed by log-transformation. To prevent undefined values during log transformation, a pseudocount of 1 was added. For single-cell RNA-seq data, cells with a total RNA count (nCount_RNA) below 1000 were excluded. Additionally, to enrich for granulocytes, cells were retained if the mean expression of *IFITM2*, *S100A12*, and *S100A8* exceeded a threshold of 2. For 10x Genomics Visium spatial transcriptomics data, spots with nCount_RNA less than 2000 were filtered out.

**Data normalization of CITE-seq.** mRNA expression data were normalized using the same approach applied to the single-cell RNA-seq data, involving CPM normalization followed by log-transformation with a pseudocount. Protein expression data were further normalized to enable comparison across samples, as protein counts exhibited greater inter-sample variability compared to mRNA expression. To address this, log-transformed CPM protein expression values were z-scored within each sample using the mean and standard deviation derived specifically from oligodendrocyte cells. Oligodendrocytes were selected as internal controls for normalization due to their consistently low levels of cell surface marker protein expression.

**Clustering of single-cell data.** A modified version of Recursive Consensus Clustering (RCC) was employed for cell clustering and marker gene identification[16]. Briefly, RCC identifies primary clusters at the first level and then recursively applies clustering within each cluster to reveal hierarchical substructures across multiple levels. In our implementation, the major modification involved replacing the original ConsensusClusterPlus algorithm[34] with Seurat v4.2.0[35] for improved scalability and integration with single-cell analysis workflows. Before clustering, batch correction was performed using *Harmony* v0.1.0[36] to integrate multiple samples. Additionally, UMAP coordinates were recalculated after recursive clustering to enhance the visualization of hierarchical cluster relationships.

**Cell-type assignment in single-cell.** Level 1 RCC clusters were annotated as major cell types based on the expression of canonical marker genes. These major cell types were then subjected to recursive clustering up to three levels to resolve finer cellular subtypes. Marker genes were identified for each sub-cluster. Cluster names reflect the hierarchical structure, with the major cell type followed by cluster numbers separated by dashes for each level, or by abbreviations corresponding to previously characterized cellular states. Cell type assignments were validated using CITE-seq protein expression data. Neoplastic cells were further confirmed based on their inferCNV-derived copy number variation (CNV) scores.

**CNV calculation.** Copy number variation (CNV) analysis of spatial transcriptomics (ST) data was performed using the *SPATA2* package in R[13]. CNV profiles were inferred by smoothing gene expression signals across spatially adjacent spots and comparing relative expression along chromosomal coordinates to reference non-malignant cell populations. Spatial spots exhibiting large-scale, concordant chromosomal gains and losses were classified as malignant, whereas spots lacking such patterns were classified as non-malignant. For single-cell RNA-seq data, CNV inference was conducted using inferCNV[5]. The single-cell data was split by major cell types for CNV analysis. Cells labeled as 'Oligo_2_1', 'Mg_1_2', 'Mg_1_3', and 'Mg_1_1' from normal brain samples were used as reference controls for CNV inference.

**Reference single-cell dataset.** The reference single-cell dataset was constructed using the datasets listed in Supplementary Data 3:

1. Allen Institute: Single-nucleus transcriptomes from 76,533 nuclei obtained from two post-mortem human brain specimens were used to survey cell type diversity in the primary motor cortex (M1C or M1)[20]. Cell type annotations from the original publication were retained.
2. Xie et al.: Single cell RNA-seq data from tumor cores and matched peripheral tissue of four glioblastoma patients[22] were used, particularly for endothelial cell identification. RCC was applied to identify clusters, and cell types were annotated using *scMAGIC*[37].
3. Winkler et al.: Single cell data from dissociated vascular cells of the adult human brain and arteriovenous malformations[21] were incorporated. The original cell type annotations were used.
4. Abdelfattah et al.: Single-cell RNA-seq data from 44 glioma samples (newly diagnosed and recurrent GBM and LGG)[23] were included. Clustering was conducted using the RCC algorithm, and cell types were subsequently annotated using scMAGIC based on canonical markers.
5. Pombo Antunes et al.: CITE-seq data from myeloid cells in newly diagnosed and recurrent GBM samples were incorporated[25], using the cell type annotations provided in the original study.
6. Matthewson et al.: Single-cell RNA-seq data of glioma-infiltrating T cells isolated from fresh tumor samples in 31 adult patients with either IDH-wildtype GBM or IDH-mutant glioma[24] were included. Cell type annotations from the original publication were retained.

To capture further biologically significant subtypes identified in our cohort, we also incorporated selected cells from our single-cell dataset into the reference dataset.

**Cell state assignment for reference single-cell data.** Cell states were assigned using four complementary approaches:

1. Original publication annotations: Cell state annotations from the original publications were adopted without modification. These included a range of cell types and subtypes such as Artery subtypes (Art1, Art2, Art3), Astrocyte subtypes based on single-gene enrichment (AstroAQP1, AstroPLCG1, AstroSERPINI2), DC, migrating DC, FBMC, Fibroblast, non-classical Monocyte, Excitatory Neurons, Inhibitory Neurons, OPC, Pericyte, SMC, proliferating Tcells, Venous, and Venule.
2. Recursive Consensus Clustering (RCC) combined with canonical markers: Marker genes were identified across multiple clustering levels using RCC. These markers were then reviewed and matched with well-established cell type signatures to assign cell identities. This approach allowed classification of immune and stromal subsets including B cell, B cell (plasma), CD4 T cell, CD8 T cell, CD8 tissue-resident memory T cell (TRM), capillary endothelial cell (Endo_cap), glioblastoma-perivascular cells (gbmPeri), granulocyte, proliferating macrophage (Mac_prolif), proliferating microglia (Mg_prolif), monocyte, natural killer T cell (NKT), red blood cell (RBC), proliferating T cell (TC_prolif), and regulatory T cell (Treg).
3. Literature-based marker analysis for tumor cell states: To assign states to tumor cells from the SNUH single-cell dataset, we integrated clustering with gene signature analysis based on studies by Neftel et al. (2019) and Garofano et al. (2021). Identified tumor cell subtypes included TC_AC, TC_mesh, TC_mesnh, TC_MTC, TC_NPC, TC_oligo, and TC_OPC. Other signatures, such as glycolytic/plurimetabolic (GPM), neuronal (NEU), and proliferative/progenitor (PRR), were excluded due to substantial overlap with the Neftel classification. For each neoplastic cell, we computed:
   - *Mmax* = *max(TC_* log2 CPM average)
   - If *Mmax* < 0, the cell was labeled as misc; otherwise, it was assigned to the corresponding TC subtype with the highest *Mmax*. The gene signatures used are listed in Supplementary Data 4.
4. RCC-based de novo stromal subtype identification: RCC enabled the identification of robust subclusters within glioblastoma stromal cells, microglia, macrophages, and oligodendrocytes. These subclusters were labeled using a hierarchical cluster naming scheme (e.g., gbmEndo_9_1, Mac_5_1_3, Oligo_2_3_1), reflecting their recursive clustering structure. As the functional roles of these subpopulations remain to be fully elucidated, we retained cluster-based labels to preserve analytical granularity for future studies.

**Marker calculation.** Marker gene identification was performed using the scran package[38]. Genes were considered highly specific markers if they met a false discovery rate (FDR) threshold of <0.01 and a minimum log2 fold change > 0. Differential expression was calculated as the log2 fold change of each gene in the target cluster (annotated feature/cell type) relative to all other clusters. The minimum log2 fold change represents the smallest log2 fold change across all pairwise comparisons, ensuring that selected markers consistently exhibit higher average expression within the target cluster.

**Differential gene expression calculation for single-cell data.** Differentially expressed genes were identified using the clusterDE[39] R package, which corrects for the *double-dipping* bias that occurs when clustering and differential expression testing are performed on the same dataset. clusterDE was applied to the normalized Seurat object to identify genes whose expression patterns robustly defined each cluster while controlling the false discovery rate (FDR). Genes with FDR < 0.05 were considered significant and retained as differentially expressed genes. Log2 fold change values were calculated by subtracting the average expression of control cells from that of test cells. Alternatively, differential gene expression analysis was also performed using pseudo-bulk data derived from the single-cell transcriptomic dataset. DESeq2[40] was used to normalize and identify differentially expressed genes. For the pseudobulk differential expression analysis, DESeq2 was run with a design formula that included both patient identity and cell type as covariates to account for inter-patient variability and differences in cellular composition. Genes with adjusted $p < 0.01$ and $|\log_2 FC| > 0.8$ were considered significant, with $\log_2 FC > 0.8$ classified as up-regulated and $\log_2 FC < -0.8$ as down-regulated. Volcano plots were generated using the *EnhancedVolcano* package in R[41], which enables the creation of publication-ready visualizations with enhanced coloring and annotation features.

**Cell type proportions calculation for core vs periphery tissue.** To evaluate whether specific cell states were enriched in the tumor core compared to the periphery, we calculated cell type frequencies for each sample in both regions. Statistical differences between core and peripheral tissues were assessed using a paired t-test on the cell proportions across all cellular states.

**Integration of single-cell and spatial transcriptome data.** Cell state scores for each spatial transcriptomics spot were calculated using the marker gene sets listed in Supplementary Data 6. For a given cell state *C*, the score was defined as the average of the log2 CPM values of all marker genes associated with *C*. To enable binary classification of cell states, we visualized histograms of cell state scores across all spots. These distributions were typically multimodal, allowing manual selection of score thresholds for each cell state. Supplementary Data 14 includes the histograms with corresponding cut-off values, spatial plots of cell state scores, and binary cell state assignments.

**Anatomic feature (AF) assignment for spatial transcriptome data.** Based on binary cell type assignments and expression profiles, spot identities were refined using a hierarchical rule-based approach:

1. LE_GM (Leading Edge - Gray Matter): Spots positive for either Neurons_Ex or Neurons_In were assigned as LE_GM.
2. MVP (Microvascular Proliferation): Spots co-expressing both gbmEndo_9_1 and gbmPeri markers were assigned as MVP.
3. HBV (Hyperplastic Blood Vessel): Spots expressing either gbmEndo_9_1 or gbmPeri, and not already assigned MVP, were assigned as HBV.
4. PAN (Pseudopalisading cells Around Necrosis): Spots expressing TC_mesh but lacking TC_mesnh expression, were assigned as PAN.
5. PNZ (Peri Necrotic Zone):
   Spots previously labeled as PAN and additionally positive for Mac_5_1_2 were reassigned as PNZ.
   Spots co-expressing TC_mesh and TC_mesnh, while still unassigned, were also assigned to PNZ.
6. IGN (Immune-Glial Niche):
   Spots positive for TC_mesnh and negative for TC_mesh, while still unassigned, were annotated as IGN.
   Additionally, spots positive for either Mac_5_1_2 or gbmFib_6, and originally labeled LE_GM or were yet unassigned, were also classified as IGN.
7. CT (Cellular Tumor): Spots with the presence of any of the non-mesenchymal neoplastic cell populations (TC_oligo, TC_prolif, TC_OPC, TC_NPC, or TC_AC), and yet unassigned, were designated as CT.

8. IT (Infiltrating Tumor): Spots expressing Oligo_2_1, Oligo_2_2, or Oligo_2_3_1, and previously classified as CT, were relabeled as IT.
9. LE_WM (Leading Edge - White Matter): Spots positive for Oligo_2_1, Oligo_2_2, Oligo_2_3_1, or OPC, and still unassigned, were assigned as LE_WM.
10. BV (Blood Vessels): Spots expressing Endo_cap or SMC and previously labeled LE_GM or yet unassigned were designated as BV.
11. PNZ (Refined): Among spots labeled as PAN, those with average geneset expression of TC_mesh <6 were reclassified as PNZ.

This rule-based spatial reannotation enabled consistent, biologically informed classification of microenvironmental regions across the tumor tissue.

**Xenium panel development.** Gene selection for the custom glioma Xenium panel was performed using a reference single-cell dataset. Marker genes significantly enriched in both major and minor cell types were identified using the findMarkers function from the scran package. A total of 2689 cell type-enriched genes were selected for initial analysis. To eliminate functionally redundant genes, recursive feature elimination (RFE) was applied using the caret package in R (Kuhn, 2008)[42]. This process yielded 383 genes as significant discriminative markers. To validate their cell type specificity and utility, cell type assignment was independently reassessed using only these 383 genes. Based on assignment accuracy and inter-gene correlation analyses, a final set of 348 genes was selected for inclusion in the glioma *Xenium* custom panel.

**Xenium data analysis.** Xenium data was generated for four patients, namely SNU18, SNU21, SNU25, and SNU33.

**Data normalization.** The cell_feature_matrix.h5 files generated by 10x Genomics *Xenium* were used for downstream analysis. Gene expression data were normalized using the square root transformation method implemented in the *stlearn* module[43], which has been shown to provide robust normalization for spatial transcriptomics data.

### Cell type assignment

**Major cell type.** Based on the single-cell reference dataset and previously identified marker genes, a subset of genes with exclusive enrichment in major cell types was selected. From the *Xenium* transcript file provided by 10x Genomics, only cells with nucleus_overlap were retained for downstream analysis to ensure high-confidence cell assignments. For each cell, the sum of expression values for marker genes corresponding to each major cell type was computed. An initial cell type was assigned based on the gene set with the highest cumulative expression. To support this, a *Xenium* reference matrix was generated, consisting of 348 selected genes (rows) and their average expression across individual major cell types (columns). Cells that remained unassigned in the initial round were subsequently annotated by calculating their Pearson correlation with the reference matrix and assigned to the cell type exhibiting the highest correlation coefficient.

**Minor cell type.** Major cell types were further resolved into minor subtypes by subsetting and reclustering each major population using *Seurat*. Dimensionality reduction was performed using PCA with 100 principal components. Clustering was carried out using the Louvain algorithm, with the resolution parameter adjusted between 0.2 and 1, depending on the heterogeneity observed within each major cell type. Minor cell type identities were assigned based on marker gene expression, using reference markers derived from the reference single-cell dataset (see Supplementary Data 5).

**Anatomical feature assignment.** To investigate anatomic features (AFs) within the high-resolution Xenium spatial transcriptomics data, we defined spatial grids (2376 μm²; area-matched to Visium spots) across tissue sections from four GBM patients (SNU21, SNU33, SNU18, and SNU25).

**Trajectory selection.** We selected 11 representative spatial trajectories across the four samples to capture key biological transitions and microenvironmental heterogeneity observed in our dataset. Trajectories were chosen based on anatomical landmarks, visible transitions in gene expression, and areas of interest identified from H&E morphology and Visium correspondence. Each trajectory was designed to traverse multiple relevant AFs, such as perivascular regions, tumor parenchyma, infiltrative margins, and fibrotic or immune-infiltrated zones. Each spatial trajectory comprised a variable number of spatial grids, reflecting differences in tissue architecture and trajectory length (range: 85–235 grids per trajectory; exact counts provided in the Source Data files).

**40-gene selection for AF assignment.** A curated panel of 40 genes was used to assign AF identities to each grid. Selection was based on established markers of canonical GBM anatomical structures—incorporating markers that distinguish normal versus GBM-associated vasculature[26], as well as our in-house oligodendrocyte markers and established microglia/macrophage/monocyte[25].

**Manual AF assignment.** Initially, AF labels were assigned to grids along selected trajectories using a combination of spatial gene expression profiles (from the 40 selected genes), focused cell type composition, and direct comparison to H&E images. Special attention was given to subregions such as the immune fibroblast niche (IFN), which was defined by the presence of fibroblast and immune markers and the absence of glial signatures within the Immune-glial Niche (IGN).

**Distance-based AF label propagation.** For the remaining grids outside the manually annotated regions, AF labels were assigned based on the shortest Euclidean distance to an AF-annotated grid within the same trajectory. This approach allowed continuous annotation of AFs across tissue while minimizing manual bias and ensuring spatial coherence.

### Gene ontology analysis

Functional enrichment was assessed using ToppFun[44], based on the top 100 markers from the upregulated and downregulated gene sets.

### Survival analysis

Survival analysis was performed using the *survival* package in R[45]. Patient stratification was based on Oligodendrocytes and Oligo_2_3_2 gene set scores computed by sssGSEA. Stratification thresholds were selected such that samples with the highest Oligodendrocytes scores were enriched for *IDH*-mutant tumors, whereas those with the highest Oligo_2_3_2 scores predominantly corresponded to *IDH*-wildtype cases. This approach was motivated by consistent observations across multiple datasets indicating that the Oligo_2_3_2 signature is largely, though not exclusively, associated with *IDH*-wildtype gliomas. Kaplan–Meier survival curves were generated using the *survminer*[46] package to visualize survival distributions across groups. In addition, multivariable Cox proportional hazards regression models were constructed to evaluate the independent prognostic value of molecular and clinical attributes. Hazard ratios (HR) with 95% confidence intervals were estimated, and forest plots were generated using the ggforest function to visualize the relative effect of each variable on survival. The Oligodendrocytes and Oligo_2_3_2 gene signatures identified in this study were incorporated as continuous covariates in the Cox models alongside patient age, IDH mutation status, 1p/19q co-deletion status and MGMT status to assess their prognostic relevance.

### Single sample gene set enrichment analysis (SSGSEA)

To assess specific cell type enrichment patterns in both single-cell and spatial transcriptomics datasets, the Single Sample Gene Set Enrichment Analysis (ssGSEA) method was applied. Gene sets were derived from the marker identification approaches described above. The resulting ssGSEA scores were subsequently z-scored and used for downstream analyses.

### Cell-type enrichment in anatomical features

We first calculated the average log2 CPM values of each cell type gene set for every spot. Using the density function in R, we computed the distribution of each gene set score across all anatomical features (AFs). To assess the significance of these distributions, AF assignments were randomized 1000 times, and density curves were recalculated for each iteration. For both the original and shuffled datasets, the area under the curve (AUC) was computed for gene set scores greater than 2. The enrichment score was defined as the ratio of the AUC from the original data to the mean AUC obtained across the 1000 shuffled iterations, thereby minimizing bias from random sampling. For each AF-celltype pair, empirical *p*-values were computed as the AUC of shuffled statistics greater than or equal to the observed AUC, following the standard permutation formula:

$$p = (\text{count of permuted values} \geq \text{observed}) + 1/\text{Nperm} + 1$$

False discovery rate (FDR) correction across all comparisons was then applied using the Benjamini–Hochberg method.

### Neighborhood analysis

Neighborhood analysis was performed to characterize spatial associations between anatomical features (AFs) in both the Visium and Xenium transcriptomic datasets. For each dataset, spatial proximity was defined based on immediate neighbors - spots in the Visium dataset and grids (each grid area equivalent to a Visium spot) in the Xenium dataset. The immediate neighbors for each spot/grid were first identified using their spatial coordinates, and neighboring units were then aggregated according to their assigned anatomical features. This process yielded an AF × AF adjacency matrix, where each entry represented the number of neighboring interactions between two anatomical features. To correct for differences in feature abundance, the raw interaction counts were normalized by the total number of neighboring pairs, resulting in a normalized neighborhood frequency matrix.

To assess the statistical significance of spatial associations, a permutation-based null model was generated. The anatomical feature assignments were randomly shuffled 1000 times, and the neighborhood analysis was recomputed for each iteration to establish a background distribution of random neighbor counts. For each pairwise interaction, the ratio of the observed normalized neighborhood count to the mean randomized count was computed, providing an enrichment score that quantified the strength of spatial association. Ratios greater than 1 indicated enrichment (features co-localizing more frequently than expected by chance), whereas ratios below 1 indicated spatial exclusion. For each spatial pair, empirical *p*-values were computed as the area under the curve (AUC) of shuffled statistics greater than or equal to the observed AUC, following the standard permutation formula:

$$p = (\text{count of permuted values} \geq \text{observed}) + 1/N\text{perm} + 1$$

where Nperm = 1000 corresponds to the number of randomizations performed. False discovery rate (FDR) correction across all pairwise comparisons was applied using the Benjamini–Hochberg method to control for multiple testing. These resulting ratios were used as edge weights, while the anatomical features served as nodes, to construct an interaction network, which was visualized using Cytoscape[47].

### DETR model for nuclei detection

We employed DINO-DETR[48] a variant of the DETR[49] model, for object detection tasks. The training dataset consisted of 27 regions of interest (ROIs) derived from 12 whole slide images (WSIs) from the 10x Genomics Visium dataset. Within these ROIs, nuclei were manually annotated, including areas with substantial background noise to ensure robustness. For model training, each ROI was subdivided into smaller image patches of size 256 × 256 pixels. In total, 352 patches were used for training and 89 for validation.

### Mutual exclusivity test

The Rediscover package[50] in R used to identify mutually exclusive patterns in ST data.

### Reporting summary

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

## Data availability

Raw single-cell RNA-sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession code PRJNA1337938. Raw Visium spatial transcriptomics data have been deposited under the same Bioproject ID. Processed spatial transcriptomics data, corresponding images, and associated metadata are available on Zenodo under the following DOIs: https://doi.org/10.5281/zenodo.17622242 (Xenium) [https://doi.org/10.5281/zenodo.17622242] and 10.5281/zenodo.17572905 (Visium) [https://doi.org/10.5281/zenodo.17572905]. Single-cell RNA-seq, CITE-seq, and spatial transcriptomics data generated in this study, together with associated cell type annotations, are available via the UCSC Cell Browser under the dataset identifier multiomic-gbm [https://cells.ucsc.edu/?ds=multiomic-gbm]. These data are publicly available and do not require restricted access. Bulk RNA-sequencing datasets analysed in this study were obtained from the following public repositories: • TCGA glioma dataset [https://portal.gdc.cancer.gov/] • CGGA glioma dataset [http://www.cgga.org.cn/] • GLASS consortium glioma dataset [https://glass-consortium.org] Previously published spatial transcriptomics datasets analysed in this study include: • GSE237183 (Greenwald et al., 10x Visium glioblastoma dataset) • Ravi et al. 10x Visium glioblastoma dataset [https://zenodo.org/records/16505469] Previously published single-cell and single-nucleus RNA-sequencing datasets analysed in this study include: • Allen Institute human motor cortex single-nucleus RNA-seq dataset [http://www.brain-map.org] • GSE162631 (Xie et al.) • Winkler et al. vascular single-cell RNA-seq dataset [https://adult-brain-vasc.cells.ucsc.edu/] • GSE131928 (Abdelfattah et al.) • GSE163120 (Pombo Antunes et al., CITE-seq) • GSE163108 (Mathewson et al.) The Source Data files are provided on Zenodo (https://doi.org/10.5281/zenodo.18093125).

## Code availability

All the required codes have been deposited on Zenodo (https://doi.org/10.5281/zenodo.17622242) as *main_figures_scripts.zip*. All annotations required to reproduce the analyses are available through the UCSC Cell Browser under the dataset ID *multiomic-gbm* (https://cells.ucsc.edu/?ds=multiomic-gbm). Tutorials and videos for accessing annotations and metadata are also provided in GitHub (https://github.com/nameetas/TSKGA).

## References

1. Vargas Lopez, A. J. Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO)

consensus review on current management and future directions. *Neuro Oncol.* **23**, 502–503 (2021).
2. Nomura, M. et al. The multilayered transcriptional architecture of glioblastoma ecosystems. *Nat. Genet* **57**, 1155–1167 (2025).
3. Verhaak, R. G. et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98–110 (2010).
4. Couturier, C. P. et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat. Commun.* **11**, 3406 (2020).
5. Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
6. Richards, L. M. et al. Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nat. Cancer* **2**, 157–173 (2021).
7. Neftel, C. et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* **178**, 835–849 e821 (2019).
8. Burger, P. C. & Green, S. B. Patient age, histologic features, and length of survival in patients with glioblastoma multiforme. *Cancer* **59**, 1617–1625 (1987).
9. Scherer, H. J. Structural development in gliomas. *Am. J. Cancer* **34**, 333–351 (1938).
10. Puchalski, R. B. et al. An anatomic transcriptional atlas of human glioblastoma. *Science* **360**, 660–663 (2018).
11. Jain, S. et al. Single-cell RNA sequencing and spatial transcriptomics reveal cancer-associated fibroblasts in glioblastoma with protumoral effects. *J. Clin. Invest.* **133** https://doi.org/10.1172/JCI147087 (2023).
12. Ren, Y. et al. Spatial transcriptomics reveals niche-specific enrichment and vulnerabilities of radial glial stem-like cells in malignant gliomas. *Nat. Commun.* **14**, 1028 (2023).
13. Ravi, V. M. et al. Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639–655 e613 (2022).
14. Wang, L. et al. A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nat. Cancer* **3**, 1534–1552 (2022).
15. Greenwald, A. C. et al. Integrative spatial analysis reveals a multilayered organization of glioblastoma. *Cell* **187**, 2485–2501 e2426 (2024).
16. Sonpatki, P. & Shah, N. Recursive Consensus Clustering for novel subtype discovery from transcriptome data. *Sci. Rep.* **10**, 11005 (2020).
17. Ruiz-Moreno, C. et al. Charting the single-cell and spatial landscape of IDH-wild-type glioblastoma with GBmap. *Neuro Oncol.* **27**, 2281–2295 (2025).
18. Maas, R. R. et al. The local microenvironment drives activation of neutrophils in human brain tumors. *Cell* **186**, 4546–4566 e4527 (2023).
19. Raghavan, J. V. et al. Immuno-phenotyping of IDH-mutant grade 3 astrocytoma and IDH-wildtype glioblastoma reveals specific differences in cells of myeloid origin. *Oncoimmunology* **10**, 1957215 (2021).
20. Sunkin, S. M. et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res.* **41**, D996–D1008 (2013).
21. Winkler, E. A. et al. A single-cell atlas of the normal and malformed human brain vasculature. *Science* **375**, eabi7377 (2022).
22. Xie, Y. et al. Key molecular alterations in endothelial cells in human glioblastoma uncovered through single-cell RNA sequencing. *JCI Insight* **6**, https://doi.org/10.1172/jci.insight.150861 (2021).
23. Abdelfattah, N. et al. Single-cell analysis of human glioma and immune cells identifies S100A4 as an immunotherapy target. *Nat. Commun.* **13**, 767 (2022).
24. Mathewson, N. D. et al. Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* **184**, 1281–1298 e1226 (2021).
25. Pombo Antunes, A. R. et al. Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. *Nat. Neurosci.* **24**, 595–610 (2021).
26. Bejarano, L. et al. Single-cell atlas of endothelial and mural cells across primary and metastatic brain tumors. *Immunity* **58**, 1015–1032 e1016 (2025).
27. Watson, S. S. et al. Fibrotic response to anti-CSF-1R therapy potentiates glioblastoma recurrence. *Cancer Cell* **42**, 1507–1527 e1511 (2024).
28. Barbie, D. A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
29. Cancer Genome Atlas Research, N. et al. The cancer genome atlas pan-cancer analysis project. *Nat. Genet* **45**, 1113–1120 (2013).
30. Zhao, Z. et al. Chinese Glioma Genome Atlas (CGGA): a comprehensive resource with functional genomic data from Chinese glioma patients. *Genom. Proteom. Bioinforma.* **19**, 1–12 (2021).
31. Xing, Y. L. et al. BRAF/MEK Inhibition Induces Cell State Transitions Boosting Immune Checkpoint Sensitivity in BRAFV600E -mutant Glioma. *bioRxiv* https://doi.org/10.1101/2023.02.03.526065 (2025).
32. Consortium, G. Glioma through the looking GLASS: molecular evolution of diffuse gliomas and the Glioma Longitudinal Analysis Consortium. *Neuro Oncol.* **20**, 873–884 (2018).
33. Pandey, S. et al. Disease-associated oligodendrocyte responses across neurodegenerative diseases. *Cell Rep.* **40**, 111189 (2022).
34. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
35. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
36. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
37. Zhang, Y., Zhang, F., Wang, Z., Wu, S. & Tian, W. scMAGIC: accurately annotating single cells using two rounds of reference-based classification. *Nucleic Acids Res.* **50**, e43 (2022).
38. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* **17**, 75 (2016).
39. Song, D., Li, K., Ge, X. & Li, J. J. ClusterDE: a post-clustering differential expression (DE) method robust to false-positive inflation caused by double dipping. *Res. Sq* (2023). https://doi.org/10.21203/rs.3.rs-3211191/v1
40. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
41. Blighe, K., Rana, S. & Lewis, M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. https://github.com/kevinblighe/EnhancedVolcano (2018).
42. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
43. Pham, D. et al. Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nat. Commun.* **14**, 7739 (2023).
44. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009).
45. Therneau, T. M. & Grambsch, P. M. Modeling Survival Data: Extending the Cox Model. Springer, New York. R package survival version 3.8-6. https://CRAN.R-project.org/package=survival (2000).
46. Kassambara, A., Kosinski, M. & Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.5.1. https://CRAN.R-project.org/package=survminer (2021).

47. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
48. Geng, H., Jiang, J., Shen, J. & Hou, M. Cascading alignment for unsupervised domain-adaptive DETR with improved DeNoising anchor boxes. *Sensors (Basel)* **22**, https://doi.org/10.3390/s22249629 (2022).
49. Carion, N. et al. End-to-End Object Detection with Transformers. In *Computer Vision – ECCV 2020, Lecture Notes in Computer Science*, (eds Vedaldi, A., Bischof, H., Brox, T. & Frahm, J. M.) Vol. 12346, 213–229 (Springer, 2020).
50. Ferrer-Bonsoms, J. A., Jareno, L. & Rubio, A. Rediscover: an R package to identify mutually exclusive mutations. *Bioinformatics* **38**, 844–845 (2022).

## Acknowledgements

The authors would like to acknowledge the patients who generously donated the samples that made this study possible. This work was supported by the following grants; N.S.: The Pershing Square Foundation, C.-K.P.: Seoul National University College of Medicine Research Foundation (grant number: 800-20210327), National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00335143), W.-Y.P.: Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science & ICT (NRF-2017M3A9A7050803), a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HR20C0025). H.-J.P.: The National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (grant number: RS-2024-00357204). New Faculty Startup Fund from Seoul National University (grant number: 550-20230104). C.K.P.: BRAF LGG consortium research fund, National Institute of Health [17×074] Bio-X Seed Grant, Stanford Cancer Institute, National Institute of Health U54 [CA261717].

## Author contributions

P.S. conducted data analysis, prepared figures, and contributed to manuscript writing. H.J.P. performed experiments, collected data, and contributed to manuscript writing. B.A.S. provided intellectual input and revised the manuscript. K.Y.H., H.J.Y., H.J.K., T.C., J.H.H., S.M.N., Y.H.B., and H.K. performed experiments and collected data. Y.L.X. data and analyses, manuscript revision. J.H.L., S.-T.L., T.M.K., and S.H.C. contributed to clinical analysis. J.-K.W. and S.-H.P. carried out histological and molecular analyses. J.-L.K. established cell lines. S.L. and H.Y. performed genetic analyses. C.K.P. provided intellectual input, resources, mouse model, data, manuscript revision. C.-K.P., W.-Y.P., and N.S. supervised the study, provided conceptual input and critical review; N.S. additionally performed data analysis and manuscript writing. C.-K.P. provided clinical supervision. C.-K.P., W.-Y.P., and N.S. are corresponding authors.

## 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/s41467-026-69716-2.

**Correspondence** and requests for materials should be addressed to Chul-Kee Park, Woong-Yang Park or Nameeta Shah.

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

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

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



[1]Amaranth Medical Analytics, Bangalore, India. [2]Samsung Genome Institute, Samsung Medical Center, Seoul, Republic of Korea. [3]Department of Biochemistry, The Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, Republic of Korea. [4]Department of Neurosurgery, Stanford School of Medicine, Stanford, CA, USA. [5]National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [6]Department of Neurosurgery, Seoul National University Hospital, Seoul, Republic of Korea. [7]Seoul National University College of Medicine, Seoul, Republic of Korea. [8]Department of Radiation Oncology, Seoul National University Hospital, Seoul, Republic of Korea. [9]Department of Neurology, Seoul National University Hospital, Seoul, Republic of Korea. [10]Department of Pathology, Seoul National University Hospital, Seoul, Republic of Korea. [11]Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea. [12]Department of Radiology, Seoul National University Hospital, Seoul, Republic of Korea. [13]Korean Cell Line Bank, Laboratory of Cell Biology, Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea. [14]Department of Genomic Medicine, Seoul National University Hospital, Seoul, Republic of Korea. [15]Department of Neurology and Neurological Sciences, Stanford University, Palo Alto, CA, USA. [16]Maternal & Child Health Research Institute, Stanford University School of Medicine, Stanford, CA, USA. [17]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. [18]Cancer Model Development Center, Stanford, Palo Alto, USA. [19]Genomic Medicine Institute, Medical Research Center, Seoul National University, Seoul, Republic of Korea. [20]Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea. [21]These authors contributed equally: Pranali Sonpatki, Hyun Jung Park. ✉e-mail: nsckpark@snu.ac.kr; woongyang@g.skku.edu; nameeta.shah@gmail.com