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

Nameeta Shah

Nameeta@skku.edu

Samsung Genome Institute

Hyun Jung Park

Seoul National University

Pranali Sonpatki

Amaranth Medical Analytics

Brett Schroeder

National Institutes of Health

Kyung Yeon Han

Samsung medical center

Hye Jin Kim

Seoul National University Hospital

Tamrin Chowdhury

Seoul National University College of Medicine https://orcid.org/0000-0002-6747-6316

Jong Ha Hwang

Seoul National University Hospital

Sun Mo Nam

Seoul National University Hospital

Yoon Hwan Byun

Seoul National University Hospital

Ho Kang

Seoul National University Hosptial https://orcid.org/0000-0003-2143-410X

Joo Ho Lee

Seoul National University Hospital

Soon-Tae Lee

Seoul National University Hospital https://orcid.org/0000-0003-4767-7564

Jae-Kyung Won

Seoul National University Hospital

Tae Min Kim

Seoul National University Hospital https://orcid.org/0000-0001-6145-4426
Seung Hong Choi
Seoul National University Hospital
Ja-Lok Ku
Seoul National University College of Medicine https://orcid.org/0000-0002-7090-537X
Sungyoung Lee
Seoul National University Hospital
Hongseok Yun
Seoul National University Hospital
Sung-Hye Park
Seoul National University College of Medicine https://orcid.org/0000-0002-8681-1597
Chul-Kee Park
Seoul National University Hosptial https://orcid.org/0000-0002-2350-9876
Woong-Yang Park
Samsung Genome Institute

Resource

Keywords:

Posted Date: May 30th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4468724/v1

Additional Declarations: There is **NO** Competing Interest.

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

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

[1]Department of Biochemistry, The Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul 08826, Korea

[2]Samsung Genome Institute, Samsung Medical Center, Seoul 06351, Korea

[3]Amaranth Medical Analytics, Bangalore 560103, India

[4]National Cancer Institute, National Institutes of Health, Bethesda, 20892, United States

[5]Department of Neurosurgery, Seoul National University Hospital, Seoul 03080, Korea

[6]Seoul National University College of Medicine, Seoul 03080, Korea

[7]Department of Radiation Oncology, Seoul National University Hospital, Seoul 03080, Korea

[8]Department of Neurology, Seoul National University Hospital, Seoul 03080, Korea

[9]Department of Pathology, Seoul National University Hospital, Seoul 03080, Korea

[10]Department of Internal Medicine, Seoul National University Hospital, Seoul 03080, Korea

[11]Department of Radiology, Seoul National University Hospital, Seoul 03080, Korea

[12]Korean Cell Line Bank, Laboratory of Cell Biology, Cancer Research Institute, Seoul National University College of Medicine, Seoul 03080, Korea

[13]Department of Genomic Medicine, Seoul National University Hospital, Seoul 03080, Korea

[14]Genomic Medicine Institute, Medical Research Center, Seoul National University, Seoul 03080, Korea

[15]Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Seoul, Republic of Korea.

[#]Contributed equally

*Corresponding authors

ABSTRACT

Glioblastoma (GBM) is a dynamic ecosystem with high plasticity where the complex interplay between different cellular components contributes to disease progression. Although single-cell RNA (scRNA)-seq has revealed remarkable cellular heterogeneity of GBM, the spatial organization of its diverse cell types is currently lacking. Here we present a comprehensive dataset of 115,914 spatial transcriptomes across 32 tissue sections of genotyped glioma samples. We present spatial maps of 56 fine-grained cellular components, including previously unrecognized oligodendrocyte subtypes and their interactive networks within each anatomical niche. Additionally, we utilized Xenium in situ technology to generate subcellular resolution spatial transcriptomic data, enhancing our understanding of GBM at a granular level. Our data provides novel insights into the cellular architecture of GBM. This valuable and openly shared resource will help develop new and effective combinatorial therapies to target multiple niches simultaneously to improve patient outcomes.

INTRODUCTION

Glioblastoma (GBM) is the most lethal form of primary brain tumor in adults. Despite surgical resection and intense systemic treatment, tumors inevitably recur and are subsequently treatment resistant. A major driver of this treatment resistance is the cellular and molecular heterogeneity of GBM in terms of its microenvironment. Therefore, understanding the spatial dynamics by which the heterogeneous tumor states emerge is critical to improving patient outcomes.

Multiple past classification efforts have attempted to capture the complexity of GBM. Early efforts in comprehensive genomic profiling suggested three molecular subtypes of GBM: classical, mesenchymal, and proneural[1]. Subsequent efforts classified GBM into states, including mesenchymal-like (MES-like), neural progenitor cell-like (NPC-like), astrocyte-like (AC-like), and oligodendrocytic precursor cell-like (OPC-like)[2]. GBM cells have also been classified based on the biological pathway enrichment, yielding two axes of variation, neuronal–proliferative/progenitor (NEU-PRR) and glycolytic/plurimetabolic–mitochondrial (GPM–MTC)[3].

Subsequent scRNA-seq technology showed that each tumor harbors multiple malignant cell states with varied inter- and intra-patient frequency[2-7]. The mechanism by which tumor cell states are determined remains elusive: whether they are stochastically and randomly dispersed across tumor tissue or confined regionally due to varying microenvironmental pressures influenced by genotype, or perhaps a combination of these factors, remains to be understood.

Transcriptional heterogeneity in tumor tissue is complex and shaped by cell-intrinsic as well as cell-extrinsic factors. Although scRNA-seq has greatly improved our understanding of cell-intrinsic molecular heterogeneity of tumor cells, the spatial information to discover cell-extrinsic factors, such as cell-to-cell interactions and tumor tissue organization, is inevitably lost during the single-cell dissociation process. Previously, we reported an anatomical feature (AF)-based transcriptional atlas, the Ivy Glioblastoma Atlas (Ivy GAP) study[8]. This work uncovered molecular programs specific to AFs, consistently observed across tumors. While this research elucidated the anatomical underpinnings of GBM molecular heterogeneity, comprehensive insights into the cellular composition and patterns of cell-cell interactions within these anatomical subdomains remain lacking.

Recent advances in spatial transcriptomic (ST) technologies allow us to understand the tumor microenvironment (TME) at higher resolution[9-13]. We performed ST with matched scRNA-seq and cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) to investigate the tissue architecture of the GBM anatomical niches. To fully capture the cellular heterogeneity of GBM TME,

we comprehensively curated reference data by compiling our patient-matched single-cell data with publicly available sc/snRNA-seq data[14-19].

We discovered novel roles for key constituents within the GBM TME, including GBM-associated oligodendrocytes and endothelial cells. Furthermore, we found that each anatomical niche harbored a unique cell type, and every tumor cell state coexisted with a distinct set of stromal cells within these niches. Our study underscores the extensive spatial and cellular heterogeneity of GBM and emphasizes the intricate interplay among diverse cellular components, which necessitates consideration in devising effective combination therapies for GBM. Further, we present our analytical results as a public resource made easily accessible to all through an innovative and intuitive format, available at https://github.com/nameetas/TSKGA.

RESULTS

Spatially resolved GBM atlas

Recent advances in spatial transcriptomic (ST) technology allow us to decipher tumor heterogeneity in spatial context and to interrogate relationships between malignant and non-malignant cells[9-13]. We investigated the cellular architecture of glioblastoma (GBM) by studying 28 patients, including 18 with GBM, 5 with astrocytoma (IDH-mutant), 1 with oligodendroglioma (IDH-mutant, 1p19q co-deleted), and 4 with miscellaneous CNS pathologies. We collected spatial transcriptomics (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, we generated a comprehensive dataset comprising 115,914 individual ST spots. Furthermore, 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 (Figure 1A). There are 14 cases with matched ST and single-cell data. In addition, we cataloged the GBM single-cell reference atlas by compiling the newly generated single-cell data with publicly available sc/sRNA-seq datasets[14-19], containing 10 major cell types, which were further classified into 58 transcriptional states (Figure 1B).

Additionally, we generated subcellular resolution ST data using recently developed Xenium in situ technology (Figure 1C) for 2 GBM patients, including a custom panel of 348 genes. From this, we annotated 351,823 cells within an area of 155 mm2. By integrating the reference single -cell atlas with ST data, we identified recurrent molecular features associated with AF-based niches including cell-type composition, cell-type co-occurrence as well as spatial LR interaction patterns (Figure 1D). Together, we generated a GBM spatial atlas with patient-matched single-cell RNA-seq and CITE-seq data (detailed patient and specimen information is provided in Table S1). To enhance user accessibility, we offer the ST analysis outcomes and the whole datasets through an interactive website at https://github.com/nameetas/TSKGA (Figure 1E).

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

In order to capture cellular states of major cell types in both the core and periphery (peri) of GBM, we profiled six matched core and peri samples through same day library preparation of fresh samples (Table S1). After quality filtering, 223,113 cells were pooled for single-cell transcriptomics, batch-corrected, and clustered using Recursive Consensus Clustering (RCC). The level 1 clusters of RCC were then used to assign major cell types by canonical marker gene expression (Figure 2A; Table S2). Notably, we were able to detect a granulocyte population from the freshly processed samples, a cell type missing from most human GBM scRNA-seq public datasets (Figure 2A). The major cell types were recursively

clustered up to three levels and 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 (Figure 2B-J; Table S3).

We further subclustered the major cell types, identified markers for each cell subtype (Figures S1A-S1F) and compared the cellular subtype proportions in core vs. peri samples. Core samples contained a higher proportion of tumor cells compared to peri samples (Figure 2C-D).

Our analysis uncovered an interesting discovery regarding oligodendrocyte subtypes. Specifically, we identified the Oligo_2_3_2 subtype exclusively within the core region across all six samples (see Figure 2E-F). Notably, this subtype was absent in both IDH-mutant gliomas and normal brain samples (see Figure S7A). 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 compared to Mg_1_2 and Mg_1_3 (resting microglia), which are more prevalent in the peri (see Figure 2G-H). Lymphocyte subtypes were not significantly different in core vs. peri except regulatory T-cells (Treg) which were mostly present in core (Figure 2I-J).

For 11 patients, we profiled 130 cell surface antigens on a total of 62,973 cells by CITE-seq (Table S1). CITE-seq analysis revealed canonical immunophenotypic markers for immune cells and other major cell types (Figures S2). We observed that tumor cells highly expressed immunomodulatory molecules, including CD24, CD44, CD112 (PVRL2), CD155 (PVR), CD73 (NT5E), and CD274 (PD-L1) (Figure S2C, highlighted in blue). This analysis provided a comprehensive list of cell surface markers for the major cellular components of the TME in GBM.

Mapping the Anatomic Features (AFs) of GBM through spatial transcriptomics

We generated ST data comprising 32 tumors from 17 patients, including 13 patients diagnosed with IDH-wildtype glioblastoma multiforme (GBM), 3 patients with IDH-mutant high-grade astrocytoma, and a normal brain region serving as a comparative control. We analyzed patients across different age groups with diverse genomic alterations (Figure 3A; Table S1). Moreover, we profiled at least two tissue sections from different locations within the same tumor (Figure S3). The ST specimens were chosen to encompass diverse AFs of GBM visible by H&E staining; leading edge (LE), cellular tumor (CT), microvascular proliferation (MVP) and pseudopalisading cells around necrosis (PAN)[8]. The median number of genes detected per section was 4,694 compared to 2,159 in a recent study[20]. Besides a higher read depth for most genes, we detected an additional 482 genes. Importantly, our ST data was generated from both frozen (n=5) and FFPE sections (n=27) (Figure S3). For better analysis of tissue

architecture, we scanned the H&E images with spatial transcriptome at 40x compared to the commonly used 5x or 10x magnifications (Figure 3B).

To investigate the spatial organization of GBM anatomical features (AFs), we first defined AFs of each ST spot. We utilized previously established marker gene sets for all AFs from Ivy GAP[8,21] (Figure 3C; Table S4). We further divided the LE area into LE_WM (white matter) and LE_GM (gray matter). Next, we defined potential GSC niches—vascular (PVN - perivascular niche) and hypoxic (PNZ - perinecrotic zone)— by integrating histological traits and transcriptional profiles of laser-captured micro-dissected regions expressing GSC markers[8,21]. Leveraging RNA-seq derived copy number analysis of ST data, we identified zones harboring tumor cells (IT, CT, PAN1/2, PNZ1/2, PVN1/2) and those devoid of tumor cells (LE_GM, LE_WM, PAN, PNZ, BV) (Figure 3D). Notably, normal sample spots (SNU38) were classified into LE_GM and LE_WM categories, while spots from patients with astrocytoma (Ast), IDH mutant (IDHm) (SNU19, SNU22) were classified into LE_GM, LE_WM, IT, and CT. In stark contrast, GBM patient spots exhibited classification across all AFs (Figure 3E-F; Figure S3). Furthermore, cell counts per spot were determined by identifying nuclei on H&E images using the DETR object detection algorithm (Figure 3G). As expected, spots without tumor cells have low counts and spots with tumor cells have higher counts.

Spatial mapping of GBM cellular components reveals AF-dependent heterogeneity

To better understand tumorigenesis and the intricate GBM ecosystem, it is important to dissect the TME into cellular components while maintaining spatial relationships. 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[14], human brain vascular cell scRNA-seq data[15,18,19], glioma associated T cell scRNA-seq data[16] and GBM myeloid cell enriched scRNA-seq data[17] (Figure 4A; Table S5). In total, we annotated 58 cellular subtypes across 10 major cell types; astrocytes, neurons, oligodendrocytes/OPC, neoplastic, microglia, lymphoid, myeloid, endothelial, fibroblast, and pericytes (Figure 4A). Using this reference dataset, we then identified highly specific markers for each cell subtype (Figure S2; Table S6). We excluded the cell subtypes without specific markers for spatial distribution, leaving 56 cell types for further analysis.

For every ST spot, we calculated the likely presence or absence of each of the 56 cell subtypes. Using the marker gene sets, we calculated an average expression score for each cell subtype. We also chose a cutoff to assign binary value of presence/absence of that cell subtype per ST spot (Methods). 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 (Figure 4B). 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 in Figure S4 and the interactive website, https://github.com/nameetas/TSKGA.

Spatial analysis of cellular subtypes revealed varying degrees of spatial and cellular heterogeneity within the GBM TME (Figures 4C; S5). We categorized the cellular landscape into four distinct niches. Firstly, the Proliferative Niche, which progresses from LE to IT and then CT, hosting neurons predominantly in LE_GM, while oligodendrocytes and astrocytes reside in LE and IT. The CT exhibits a notable abundance of non-mesenchymal tumor cells. Secondly, the Hypoxic Niche, encompassing PAN, PAN1, and PAN2, and harbors TC_mesh, TC_NPC, Mac_5_1_3, and gbmEndo_9_4. Thirdly, the Immune Niche, comprising PNZ, PNZ1, and PNZ2, predominantly houses immune cells and gbmEndo_9_4, along with mesenchymal-like tumor cells TC_mesh and TC_mesnh. Lastly, the Vascular Niche, encompassing BV, PVN1, and PVN2, predominantly comprises vascular and immune cell subtypes, with TC_mesnh tumor cell subtype exclusively observed in this niche (Figure S5A). Notably, the newly identified Oligo_2_3_2 is found within the Vascular Niche (Figure S5B).

Distinct spatial localizations of mesenchymal vs. non-mesenchymal tumor cells were evident. Non-mesenchymal subtypes, including TC_NPC, TC_OPC, TC_oligo, and TC_prolif, primarily occupy the proliferative niche, contrasting with mesenchymal counterparts that predominantly reside within hypoxic, immune, and vascular niches. TC_mesnh cells exhibit enrichment in immune and vascular environments, while TC_mesh cells are predominantly situated within immune and hypoxic niches (Figure S5A). This suggests that despite sharing mesenchymal gene signatures, these two tumor cell states may have arisen under different environmental pressures.

Analysis of CITE-seq data revealed heightened expression of immunomodulatory molecules including CD95, CD54, CD26, CD13, CD99, and CD274 on the surface of TC_mesnh cells compared to other tumor cell subtypes. TC_mesnh, resembling Mes1-like cells in Verhaak classification, demonstrate an association with immune infiltration, whereas TC_mesh (Mes2-like in Verhaak classification) are linked with chronic hypoxia. Notably, TC_mesnh do not express CD24, unlike other tumor cell subtypes (Figure S1E). In summary, TC_mesnh cells inhabit immunosuppressive environments and exhibit high expression of inhibitory molecules such as CD73 and CD274 (Figure S1E). Conversely, TC_oligo, TC_OPC, and TC_NPC predominantly reside in immune-desert niches, potentially evading immune surveillance through cell-autonomous mechanisms characterized by low expression of MHC class I molecules and upregulation of "don't eat me" signal molecule, CD24 (Figure 4C; Figure S1F-E). This implies the employment of diverse immune evasion mechanisms among tumor cell subtypes.

GBM landscape visualized at sub-cellular resolution

Through the integration of single-cell and spatial transcriptome data, we identified novel cellular states (cell subtypes) within GBM niches. Given that Visium data offers a spatial resolution of 55 μm, it has limitations in pinpointing single-cells. To validate our findings at single-cell spatial resolution, we leveraged the Xenium platform, which provides transcriptome data at sub-micron resolution. A custom panel of 348 genes (Figure 5A; Table S7; methods) was designed to validate novel endothelial, pericyte, and oligodendrocyte cell subtypes and their spatial distribution. This panel was applied to data from two GBM patients from the Visium cohort, SNU18 and SNU25, utilizing one Xenium slide. A total of 351,823 cells were annotated within an area of 155 mm^2. There was an excellent correlation between the gene expression profiles of spatial single-cell data and our reference single-cell dataset using the 348 panel genes (Figure 5B). The Xenium slide was annotated at two levels, AF and single-cell (Figure 5 D-E), employing a semi-automated annotation process. GBM-associated pericytes (gbmPeri) and endothelial cells (gbmEndo*) expressing gene COL4A1 were identified, predominantly in the immune and vascular niche, a finding supported by our Xenium data (Figure 5A; Figure S6A). Furthermore, we sought to validate our findings for the Oligo_2_3_2 subtype, which was located in the tumor core in single-cell data and exhibited preferential localization in the vascular niche in spatial transcriptomic data. A similar trend was observed when comparing matched Visium and Xenium data (Figure S6B).

Understanding oligodendrocyte heterogeneity through our GBM atlas

Emerging evidence indicates the pivotal involvement of oligodendroglia in various neurodegenerative conditions[22,23]. Nonetheless, the precise extent of oligodendrocyte diversity and their role in GBM pathology remains elusive. We discovered four distinct oligodendrocyte cell states within GBM (Figure 2E-F) and delineated novel markers characterizing these states (Figure S1A). Notably, we observed a higher prevalence of HLA-A^{high} Oligo_2_3_2 within the tumor core compared to the periphery (Figure 2E-F). Moreover, Oligo_2_3_2 was uniquely detected in GBM, but not in IDH-mutant high-grade astrocytoma (Figure S7A).

To explore the potential roles of oligodendrocytes in GBM, we scrutinized the differentially expressed genes (DEGs) between oligodendrocytes from the tumor core and periphery (Figure 6A; Table S8). We found a significant upregulation of various immune-related genes including HLA-A, HLA-E, CLU, IFITM3, and MIF in Oligo_2_3_2 (Figure 6A; Table S8). Conversely, myelin-related genes such as OPALIN and MOBP were downregulated, particularly in Oligo_2_3_2 from the tumor core (Figure 6A). Consistently, gene ontology (GO) analysis highlighted enrichment in immune response pathways such as antigen processing and presentation, alongside negative regulation of natural killer cell cytokine production (Figure S7B). This pattern of expression was also corroborated in Xenium data, where Oligo_2_3_2 cells exhibited distinct transcript counts of CDH1, OPALIN, MBP, GSN, and NGFR, notably with low OPALIN and high GSN counts compared to Oligo_2_1 (Figure S7C). Furthermore, these Oligo_2_3_2 cells in Xenium data demonstrated high expression of SERPINA3, HLA-A,

TUBB2B, TIMP1, CLU, HLA-DRA, and S100A1 transcripts, consistent with the single-cell data (Figure 6A-B). Pseudotime analysis unveiled a trajectory transitioning from OPALIN Oligo_2_2 to HLA-Ahigh Oligo_2_3_2 (Figures 6C; Figure S8).

Additionally, our cell type co-occurrence analysis utilizing Visium ST data revealed a spatial correlation between Oligo_2_3_2 and various infiltrating immune cells and endothelial cells (Figure 6D). Similarly, spatial co-localization using Xenium data illustrated the proximity of Oligo_2_3_2 to TC_mesnh, Mg_1_1 (activated microglia), Mac_5_1_2 (Tumor associated macrophage), and gbmEndo (GBM-associated endothelial cells), in contrast to Oligo_2_1, which was found in close proximity to AstroPLCG1 (gray matter-associated astrocytes), Neurons_Ex (excitatory neurons), and normal endothelial cells (Figure 6E). The intermediate state represented by Oligo_2_3_1 exhibited preferential co-occurrence with reactive astrocytes. Thus, Oligo_2_3_2 emerges as a novel oligodendrocyte subtype likely shaped by environmental pressures of inflammation and tissue damage around leaky vessels, potentially contributing to the pathology of GBM. Next we investigated cell-cell interactions between Oligo_2_3_2 and its neighbors. We uncovered the most significant interactions between Oligo_2_3_2 and gbmEndoPeri, another newly identified cell type in this study. The analysis revealed two key interactions between Oligo_2_3_2 and gbmEndoPeri, gbmEndoPeri_2 in particular: WNT -SFRP1 and various subtypes of collagen-intergrins (Figure S9). Both WNT and integrin signaling play pivotal roles in modulating oligodendrocyte proliferation, differentiation and myelination, potentially influencing the GBM pathogenesis associated with oligodendrocyte dysfunctions.

DISCUSSION

Our study is one of the most comprehensive transcriptomic and spatially delineated analysis of GBM to date. Drawing upon data from 16 GBM patients and one healthy brain specimen, our approach combines cutting-edge technologies, including the 10x Genomics Visium Spatial Transcriptomics (ST) platform, alongside supplementary single-cell RNA sequencing (scRNA-seq) and cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) datasets from 26 patients. This meticulous integration of molecular profiles has yielded an expansive dataset, comprising 115,914 individual ST spots and 245,736 cells, providing novel insights into the cellular and molecular diversity within the GBM tumor microenvironment (TME). Importantly, this study marks the pioneering use of Xenium in situ technology in GBM analysis, achieving subcellular spatial resolution and enhancing our understanding of GBM heterogeneity to previously unattainable depths.

Our analysis reveals a distinct oligodendrocyte state, Oligo_2_3_2, marked by the expression of HLA-A and pronounced enrichment within the perivascular space[24]. Notably, Oligo_2_3_2 also displays high levels of the non-classical MHC class I molecule, HLA-E, known to interact with the inhibitory receptor CD94/NKG2A, modulating NK and CD8 T cell-mediated cytotoxicity[25,26]. Additionally, Oligo_2_3_2 upregulates LGALS1, a pleiotropic immunosuppressive molecule commonly found in immune-privileged organs[24]. This suggests a potential mechanism by which Oligo_2_3_2 promotes immune tolerance within the brain parenchyma, possibly exploited by GBM tumor cells to evade immune surveillance. Moreover, Oligo_2_3_2 downregulates myelin-associated genes OPALIN and MOBP while highly expressing FGFR2, implicating its involvement in inhibiting remyelination and contributing to the pathogenesis of multiple sclerosis[27,28]. Intriguingly, the upregulated genes in Oligo_2_3_2 overlap with the signature genes of neurodegenerative disease-associated oligodendrocytes, suggesting a potential link between neuroinflammation and the transformation of homeostatic oligodendrocytes into pathogenic states associated with impaired remyelination and neurodegenerative processes [22]. This underscores the complex relationship among neuroinflammation, neurovascular dysfunction, and oligodendrocyte pathology within the realm of neurodegenerative disease. Understanding the role of oligodendrocytes in GBM pathogenesis presents a promising avenue for identifying therapeutic targets.

Our investigation also unveiled recurrent molecular signatures and elucidated spatial organization of cell types within the GBM TME, underscoring their distinct contributions within AF-based niches [29]. Notable among our findings is the discernment of partially coinciding spatial segregation patterns among tumor cell states, delineating discrete AF-associated microenvironmental regions. Furthermore, each tumor cell state exhibited a preferential co-occurrence with specific non-malignant cell types. For

instance, TC_mesnh, a mesenchymal tumor cell state, demonstrated confinement within GBM-associated vascular niches, while TC_mesh was predominantly localized in hypoxic regions. Our consistent identification of macrophages within TC_mesnh suggests a potential macrophage-tumor cell crosstalk within the GBM perivascular niche, potentially driving the transition of tumor cells into mesenchymal states.

This spatial atlas provides a detailed molecular and cellular landscape of GBM and gives new insights into the complex heterogeneity of disease. Our findings underscore the potential of spatially resolved analysis in uncovering novel therapeutic targets and improving our understanding of GBM tumorigenesis and progression. Furthermore, making these datasets available through an interactive website, we aim to facilitate further research and discovery for a disease with minimally improved outcomes over the last three decades. Importantly, unlike other large ST datasets, ours were generated using FFPE specimens—the most widely used method of tissue preservation, and therefore the most applicable. The FFPE method preserves morphological details and permits high-quality histological images. Recently, several studies demonstrated that spatial gene expression of tissue sections could be predicted by integrating ST data with histological images[30,31]. We anticipate that our ST data with the high-resolution images will facilitate the development of such algorithmic tools and potentially help to find image-based molecular biomarkers that will serve as the foundation for novel targeted therapies.

METHODS

Patient samples

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).

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)

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)

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 -80ºC 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.

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.

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 hours 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 (600g, 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 5,000 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 25K reads per cell for gene expression libraries and 5K 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 cell-seq/cite-seq data was aligned using Cell Ranger software. 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 cellseq:

```
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
```

Raw_feature_bc_matrix files were used for further analysis. The counts data was then log2 (cpm + 1) normalized (cpm is counts per million). For single-cell data, the cells were filtered out if nCount_RNA < 1000 AND Mean of IFITM2, S100A12 and S100A8 > 2 (this was done to retain granulocytes). For 10x Visium data, the spots with nCount_RNA < 2000 were filtered out.

CITE-seq analysis

The mRNA data was normalized using the same process as cell seq data. For protein data, further normalization was done for comparison of protein expression across samples. Protein expression counts vary significantly across samples compared to mRNA expression. So, for each sample, the log2cpm data was z-scored using the mean and standard deviation of only oligodendrocyte cells. We chose oligodendrocyte cells as control cells for sample-wise normalization because they showed the lowest level of cell surface marker protein expression.

Xenium panel development

Gene selection for glioma xenium panel was done using the single-cell reference dataset. Significantly enriched genes in major and minor cell types were calculated using findMarkers function from scran package. 2689 genes that were enriched in individual cell types were selected for further analysis. Recursive feature elimination (rfe) from the caret [Kuhn, Max (2008). "Building Predictive Models in R Using the caret Package." Journal of Statistical Software, 28(5), 1–26. doi:10.18637/jss.v028.i05, https://www.jstatsoft.org/index.php/jss/article/view/v028i05.] package in R was used for filtering functionally redundant genes. Post rfe, 383 genes came up as significant markers. Using these 383 genes cell type assignment was again independently performed to check the accuracy of the gene profiles. Based on the assignment results and inter-gene correlations, 348 genes were finalized for the custom glioma xenium panel.

Xenium data analysis

Data normalization

The cell_feature_matrix.h5 files provided by 10x Xenium were used for further analysis. The data was normalized using the stlearn module's square root transformation method[32].

Cell type assignment

Major cell type

Based on the single-cell reference dataset and previously calculated markers, genes with exclusive enrichment in major cell types were selected (STable 6). From the transcripts file provided by 10x Xenium, only the cell_ids where nucleus_overlap is one are considered for further analysis. The sum of all the markers (geneset) per major cell type was calculated for each cell. Then a major cell type was assigned based on the max of the geneset sum. A xenium reference gene x major cell type matrix was generated where rows were 348 genes and columns were the average value of respective genes in the individual major cell types. All the cells that were left unassigned in the first run, were then correlated with the xenium reference major cell type matrix and assigned the cell type with maximum correlation value.

Minor cell type

Major cell types were further assessed for minor cell type assignment. Each major cell type was subsetted and clustered using Seurat using the following parameters:

PCA, ncps: 100

clustering algorithm: 2 (multi step louvain clustering) and

resolution: 0.2 - 1 (Based on the major cell type). All the minor cell types were assigned based on the markers calculated using the single-cell public reference dataset (Table S6).

Clustering of single-cell and spatial data

We used a modified version of Recursive Consensus Clustering (RCC) for clustering and marker calculations[33]. In brief, RCC first identifies major clusters at level 1. It then continues to take each cluster and then recursively clusters it to identify further sub-clusters at multiple levels. The major change involves using Seurat v4.2.0[34] as the underlying algorithm instead of ConsensusClusterPlus[35] used in the original publication. For integrating multiple samples, we used batch correction package Harmony v0.1.0 prior to clustering[36]. It also transforms the initial UMAP co-ordinates based on recursive clustering such that it allows better visualization of cluster hierarchy.

CNV calculation

CNV analysis for ST data was performed using the SPATA2 package in R[20]. For single-cell data inferCNV [inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV] was used for analysis. The single-cell data was split by major cell types for CNV analysis. Cells that were assigned as ‘Oligo_2_1', 'Mg_1_2', 'Mg_1_3’, and ‘Mg_1_1’ from the normal brain samples were used as a control for CNV analysis.

Reference single-cell dataset

The following datasets (Table S3) were used to build the reference single-cell dataset:

1. Allen institute: single-nucleus transcriptomes from 76,533 total nuclei derived from 2 post-mortem human brain specimens to survey cell type diversity in the primary motor cortex (M1C or M1)[14]. We used the original publication cell type annotation for the dataset.
2. Xie et al.: single-cell data generated from tumor core and paired surrounding peripheral tissue in 4 glioblastoma patients[18]. Specifically, for endothelial cells. RCC was performed to find clusters and cell types were identified using scMAGIC[37].

3. Winkler et al: single-cell data generated using dissociated vascular cells isolated from the adult human brain and arteriovenous malformations[15]. We used the original publication cell type annotation for the dataset.
4. Abdelfattah et al.: single-cell data generated from 44 GBM (new diagnosed + recurrent) and LGG samples[19]. RCC was performed to find clusters and cell types were identified using scMAGIC.
5. TAM-citeseq: CITE-seq data of myeloid cells from newly diagnosed and recurrent GBM samples[17]. We used the original publication celltype-annotation for the dataset.
6. Matthewson et al.: single-cell data of glioma-infiltrating T cells isolated from fresh tumor samples in 31 adult patients with either IDH-wildtype glioblastoma (GBM) or IDH-mutant glioma (IDH-G)[16]. We used the original publication cell type annotation for the dataset.

Marker calculation

We used scran for marker calculation[38]. To identify highly specific markers, we used FDR < 0.01 and minLog2FC > 0. Differential expression as log2 fold change (Log2FC) was calculated for each gene for the cluster (AF/cell type) for which markers are being calculated against all the clusters. minLog2FC is defined as the minimum log2 fold change value among all the comparisons. This identifies markers that have the highest average expression in that cluster.

Anatomic feature (AF) assignment for ST data

Log2cpm data from all the samples was merged. ssGSEA[39] was performed on the log2cpm data using gene sets for each AF (Table S4)[21]. The ssGSEA data was z-scored feature-wise. Z-scoring was done separately for FFPE and frozen samples. Each spot was assigned an AF based on max z-score for that spot. We also used the output from clustering of these spots to further refine the AF assignment. One of the clusters was found to be typically bordering the gray and white matter showing expression of both oligodendrocyte and neuron markers, which we annotated as LE_WGM.

Cell state assignment for single-cell data

We used four different ways to assign cell states.

1. Original publication - Annotations from the original publication were used. These include Art1, Art2, Art3, AstroAQP1, AstroPLCG1, AstroSERPINI2, DC, DC_mig, FBMC, Fibroblast, Monocyte_Nc, Neurons_Ex, Neurons_In, OPC, Pericyte, SMC, Tcell_prolif, Venous, and Venule.
2. Recursive clustering analysis (RCC) + well-established markers - Markers were calculated for the clusters at different levels. Review of these markers allowed us to then assign the cellular

subtypes. These include Bcell, Bcell_plasma, CD4, CD8, CD8_TRM, Endo_cap, gbmPeri, Granulocyte, Mac_prolif, Mg_prolif, Monocyte, NKT, RBC, TC_prolif, Treg.

3. Clustering analysis + literature-based marker analysis - We used this approach to assign cellular state to tumor cells from SNUH single-cell data. These include tumor cell (TC) subtypes TC_AC, TC_mesh, TC_mesnh, TC_MTC, TC_NPC, TC_oligo, and TC_OPC based on studies by Neftel et al., and Garofano et al[2,3]. We did not include glycolytic/plurimetabolic (GPM), neuronal (NEU), and proliferative/progenitor (PRR) in further analysis as they highly overlap with Neftel classification. The gene marker set used for this is given in Table S4. For all neoplastic cells we assigned cellular state using:
   - Mmax = max (TC_* l2cpmavg). [* is TC_AC, TC_NPC, TC_oligo, TC_OPC, TC_MTC]
   - If Mmax < 0 then label = 'misc' else label = TC_* with Mmax
4. Recursive clustering analysis (RCC) - Using RCC we were able to identify robust sub-clusters for GBM stromal cells, microglia, macrophages, and oligodendrocytes. We labeled them based on their cluster numbers as further studies would be required to establish their functional relevance. These include gbmEndo_9_1, gbmEndo_9_1a, gbmEndo_9_3, gbmEndo_9_4, gbmEndoPeri_1, gbmEndoPeri_2, gbmFib_4, gbmFib_5, gbmFib_6, Mac_5_1_2, Mac_5_1_3, Mg_1_1, Mg_1_3, Oligo_2_1, Oligo_2_2, Oligo_2_3_1, Oligo_2_3_2.

Integration of single-cell and ST data

Based on the marker set (Table S6), we calculated cell state scores for each spot. The score for cell state C is the average of log2cpm values of all the makers for cell state C. For binary assignment, we visualized the histogram for score of each cell type which typically had multimodal distribution and manually chose a score cut-off value. Supplementary Data file 1 has the histogram and cutoffs used for cell state scores, spatial plots for cell state score, and binary cell state assignment.

Co-occurrence score calculation

IOU calculation

Binary assignment of cell states to spots was used for calculating IOU (intersection over union) between two cell states. For a set of spots labeled as a specific anatomic feature (AF), we calculated the spots that have both cell states present and divided it by the total number of spots that have either of the cell states present.

IOmin calculation

Binary assignment of cell states to spots was used for calculating IOmin (intersection over minimum) between two cell states. For a set of spots labeled as a specific anatomic feature (AF), we calculated the spots that have both cell states present and divided it by the total number of spots for the less frequent

cell state. This score is useful for understanding co-occurrence for rarer cell states. Chord diagrams were used to visualize the IOmin between receptor-ligand pairs enriched across multiple cell types/AF. The chord diagrams were generated using the circlize package in R[40].

Differential gene expression calculation for single-cell data

Differential gene expression identification was done using t.test (R package) and log2FC was calculated by subtracting the average expression of control cells from the average expression of test cells. In order to avoid statistical artifacts due to an unbalanced number of cells in control and test groups, we made both groups have the same number of cells by randomly choosing cells from the group with higher number of cells. The volcano plots were drawn using the EnhancedVolcano package in R (Blighe, K, S Rana, and M Lewis. 2018. "EnhancedVolcano: Publication-ready volcano plots with enhanced coloring and labeling." https://github.com/kevinblighe/EnhancedVolcano.)

Survival analysis

The survival analysis was done using the survival package [Therneau T (2022). *A Package for Survival Analysis in R*. R package version 3.4-0, https://CRAN.R-project.org/package=survival] in R. Ecotype and MGMT fields were used as differentiators for survival analysis. The Kaplan-Meier plots were generated using the survmineR package in R.

DETR model for nuclei detection

We used DINO-DETR [arXiv:2203.03605] which is a flavor of the DETR [arXiv:2005.12872] model. The training data consists of 27 ROIs from 12 WSIs from the Visium dataset. For these ROIs the nuclei were marked, and it includes regions with a high level of noise. Small patches of 256 x 256 were generated from these ROIs for the purpose of training the DINO-DETR model. In total the number of training images was 352 and the number of validation images were 89.

Ligand receptor interactions

The ligand receptor interactions were analyzed using cellphoneDB v4[41]. The single-cell public reference dataset was used for analyzing the ligand-receptor interactions with default parameters.

DATA AND CODE AVAILABILITY

Raw data reported in this paper will be publicly available upon publication. The processed data and Interactive browsers for this study are available at https://github.com/nameetas/TSKGA. All code has been deposited at GitHub (https://github.com/nameetas/TSKGA) and will be publicly available as of the date of publication. Further information or resources will be available from the corresponding authors upon request.

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), and 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: Model Animal Cluster Project through the National Research Foundation (NRF) funded the Korea government (MSIT) (NRF-2021M3H9A1030155), New Faculty Startup Fund from Seoul National University (grant number: 550-20230104).

DECLARATIONS OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. 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).
2. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849 e21 (2019).
3. Garofano, L. *et al.* Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nat Cancer* **2**, 141-156 (2021).
4. Castellan, M. *et al.* Single-cell analyses reveal YAP/TAZ as regulators of stemness and cell plasticity in Glioblastoma. *Nat Cancer* **2**, 174-188 (2021).
5. Couturier, C.P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* **11**, 3406 (2020).
6. Patel, A.P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-401 (2014).
7. 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).
8. Puchalski, R.B. *et al.* An anatomic transcriptional atlas of human glioblastoma. *Science* **360**, 660-663 (2018).
9. Stahl, P.L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78-82 (2016).
10. Rodriques, S.G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463-1467 (2019).
11. Eng, C.L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235-239 (2019).
12. Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
13. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**(2018).
14. 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).
15. Winkler, E.A. *et al.* A single-cell atlas of the normal and malformed human brain vasculature. *Science* **375**, eabi7377 (2022).
16. Mathewson, N.D. *et al.* Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* **184**, 1281-1298 e26 (2021).
17. 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).
18. Xie, Y. *et al.* Key molecular alterations in endothelial cells in human glioblastoma uncovered through single-cell RNA sequencing. *JCI Insight* **6**(2021).
19. Abdelfattah, N. *et al.* Single-cell analysis of human glioma and immune cells identifies S100A4 as an immunotherapy target. *Nat Commun* **13**, 767 (2022).
20. Ravi, V.M. *et al.* Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639-655 e13 (2022).
21. Shah, N. *et al.* Gelatin methacrylate hydrogels culture model for glioblastoma cells enriches for mesenchymal-like state and models interactions with immune cells. *Sci Rep* **11**, 17727 (2021).
22. Kenigsbuch, M. *et al.* A shared disease-associated oligodendrocyte signature among multiple CNS pathologies. *Nat Neurosci* **25**, 876-886 (2022).
23. Pandey, S. *et al.* Disease-associated oligodendrocyte responses across neurodegenerative diseases. *Cell Rep* **40**, 111189 (2022).
24. Girotti, M.R., Salatino, M., Dalotto-Moreno, T. & Rabinovich, G.A. Sweetening the hallmarks of cancer: Galectins as multifunctional mediators of tumor progression. *J Exp Med* **217**(2020).

25. Brooks, A.G., Posch, P.E., Scorzelli, C.J., Borrego, F. & Coligan, J.E. NKG2A complexed with CD94 defines a novel inhibitory natural killer cell receptor. *J Exp Med* **185**, 795-800 (1997).
26. Rapaport, A.S. *et al.* The Inhibitory Receptor NKG2A Sustains Virus-Specific CD8(+) T Cells in Response to a Lethal Poxvirus Infection. *Immunity* **43**, 1112-24 (2015).
27. Kamali, S. *et al.* Oligodendrocyte-specific deletion of FGFR2 ameliorates MOG35-55 -induced EAE through ERK and Akt signalling. *Brain Pathol* **31**, 297-311 (2021).
28. Klimaschewski, L. & Claus, P. Fibroblast Growth Factor Signalling in the Diseased Nervous System. *Mol Neurobiol* **58**, 3884-3902 (2021).
29. Hara, T. *et al.* Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779-792 e11 (2021).
30. He, B. *et al.* Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat Biomed Eng* **4**, 827-834 (2020).
31. Bergenstrahle, L. *et al.* Super-resolved spatial transcriptomics by deep data fusion. *Nat Biotechnol* **40**, 476-479 (2022).
32. Pham, D. *et al.* Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nat Commun* **14**, 7739 (2023).
33. Sonpatki, P. & Shah, N. Recursive Consensus Clustering for novel subtype discovery from transcriptome data. *Sci Rep* **10**, 11005 (2020).
34. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e29 (2021).
35. Wilkerson, M.D. & Hayes, D.N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572-3 (2010).
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. Foroutan, M. *et al.* Single sample scoring of molecular phenotypes. *BMC Bioinformatics* **19**, 404 (2018).
40. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2 (2014).
41. Efremova, M., Vento-Tormo, M., Teichmann, S.A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506 (2020).

FIGURE LEGENDS

Figure 1 Spatial multi-omic glioma atlas schematic. (A) Sample preparation. (B) Glioma single-cell reference dataset construction. (C) Spatial transcriptomic data generated with Visium at 55 um resolution, validated with Xenium, and the targeted spatial profiling method at subcellular resolution. (D) Integrated analysis of single-cell and spatial data for mapping of cellular states to anatomic features. (E) User-friendly website for data visualization and exploration.

Figure 2 Spatial segregation of glioblastoma cellular heterogeneity between tumor core and periphery in SNUH dataset. (A) UMAP representation of major cell types. (B) UMAP representation using recursive layout L2 of cellular states. (C-J) UMAP and bar plots showing sample wise matched core and periphery proportions of (C-D) Neoplastic cells, highlighted TC_mesnh and TC_prolif cells are limited to the core (E-F) Oligodendrocytes, highlighted Oligo_2_3_2 cells are found in the core (G-H) Myeloid/Microglia, DC, Mac_5_1_2, Mac_5_1_3, Mg_prolif are in higher proportions in core compared to Mg_1_3 cells present in higher proportion in the periphery (I-J) Lymphocytes, Tregs are found in the core. *Significance is calculated using paired t-test, p-value < 0.05. See also Figure S1

Figure 3 Spatial transcriptomic data and anatomic feature (AF) annotation. (A) Patient information with genotype profile. (B) Representative high-resolution H&E images of the specimen from patient SNU51. (C) Dot plot of anatomical feature marker genes and the distribution of inferCNV scores across the anatomical features. Brown vertical lines indicate the threshold for LE, IT and CT segments segregation. All the spots with inferCNV core < 100 were assigned LE, between 100-150 were assigned IT and > 150 were assigned as CT. (D) Final anatomic feature assignment based on IvyGAP segments and CNV scores. (E) Spatial plot for AF annotation. (E) Relative proportions of AFs in tissue sections from each patient. (F) Nuclei were identified on H&E images using DINO-DETR object detection algorithm. The box plot shows nuclei count for each AF. See also Figure S3, S4

Figure 4 Spatial mapping of cell types over glioblastoma AFs. (A) UMAP representation of reference glioma single-cell atlas colored by major cell types and 58 cell subtypes. (B) Density plot for z-scored average cell subtype marker gene expression per AF. Solid lines show the results for actual data and dotted lines are for shuffled AF data. The ratio > 1 of Actual vs. Randomized AUC gives the strength of preferential presence of a given cell subtype in a given AF and ratio < 1 indicated preferential absence. Neurons_Ex are shown to be preferentially present in LE_GM. (C) Preferential presence of cell subtypes across AFs. The presence of connection shows ratio > 2 and the thickness indicates the strength. Leading edge to cellular tumor AFs (LE_GM, LE_WM, IT, CT) have neurons, oligodendrocytes, astrocytes, and non-mesenchymal neoplastic cells. Hypoxic AFs (PAN, PAN1,

PAN2) have astrocytes, neoplastic cells, macrophages, and glioblastoma associated endothelial cells. Immune AFs (PNZ, PNZ1, PNZ2) have mesenchymal neoplastic cells, microglia, macrophages, lymphoid cells, some fibroblasts, and glioblastoma associated endothelial cells. Vascular AFs (BV, PVN1, PVN2) have oligodendrocytes, astrocytes, mesenchymal neoplastic cells, all types of immune cells, fibroblasts, pericytes, and endothelial cells. See also Figure S5.

Figure 5 Xenium data for two glioblastoma patients SNU25 and SNU18 (A) Custom gene marker panel used for Xenium analysis. Genes used for major cell type assignment are highlighted. (B) Correlation between gene expression profiles of major cell types in single-cell and xenium data. Xenium slide (C) H&E (D) AF assignment for 110 um grid (E) minor cell types. See also Figure S6.

Figure 6 GBM-associated oligodendrocytes exhibit immunomodulatory features. (A) The single-cell trajectory reconstructed by Monocle 2 for GBM-associated oligodendrocytes, cells colored by oligodendrocyte cell subtypes. (B) DEGs between Oligo_2_3_2 vs Oligo_2_1, highlighted as red dots. (C) Common DEGs in single-cell vs. Xenium dataset. (D) Co-occurrence matrix for oligodendrocyte subtypes showing increased co-localization of Oligo_2_3_2 cells with immune and stromal cells relative to the other oligodendrocyte subtypes. (E) Spatial co-localization analysis in Xenium data showing Oligo_2_1 are found in close proximity to AstroPLCG1, Neurons_Ex, and normal Endothelial cells compared to Oligo_2_3_2 which are in close proximity to TC_mesnh, Mg_1_1, Mac_5_1_2, and glioblastoma associated endothelial cells. AstroAQP1 are closest to Oligo_2_3_1. See also Figure S6, S7, S8, and S9.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1 (A-F) Cell subtype marker genes for astrocytes, neurons and oligodendrocytes (A), microglia (B), lymphoid (C), myeloid (D), neoplastic (E) and endothelial/stromal cells (F).

Figure S2 (A) UMAP representation single-cells profiled using CITE-seq. (B) Canonical marker genes (left) and cell surface protein markers (right) for major cell types. (C) Normalized cell surface protein expression across cellular states.

Figure S3 AF annotation on each tissue section. H&E images and AF annotation of 32 spatial transcriptomic data

Figure S4 Single-cell and spatial transcriptomic data analysis of SNU51. (A) Spatial plot of SNU51 with AFs overlaid. Barplot of distribution of cell types across AFs in SNU51 spatial data. (B) UMAP of single-cell RNA-Seq data split by tissue type. Barplot of cell types across tissue in SNU51 single-cell data. (C) CNV analysis of SNU51 bulk tissue RNA-seq data. (D) Mutational profile of SNU51. (E) CNV analysis of SNU51 using spatial data across AFs. (F) CNV analysis of SNU51 using single-cell RNA-seq data across major cell types.

Figure S5 Cellular and spatial heterogeneity of GMB anatomical niches. (A-F) River plot depicting neoplastic (A), brain specific (B) microglia/myeloid (C), lymphoid (D), endothelial (E) and stromal (F) cell subtype contribution across AFs.

Figure S6 Validation of cell subtypes in Xenium data. (A) Blue dots representing endothelial, and pericyte subtypes. The top panel shows the normal endothelial (Endo_cap) and pericytes and the bottom one shows the GBM associated endothelial cells and pericytes (gbmPeri, gbmEndo*). The blood vessel in LE_GM shows the relative absence of COL4A1 (green dots) compared to the one in PVN area. The blue boxes highlight all the endothelial cells and pericytes. (B) The left side shows the oligodendrocyte subtypes map on the Xenium slide and the right side shows the corresponding Visium slides as blue dots. The solid yellow boxes highlight the LE_GM area of patient SNU25 enriched for Oligo_2_1, the dashed green boxes highlight the IT/PVN area of patient SNU25 enriched for Oligo_2_3_1, and the dotted red boxes highlight the PNZ/PVN area enriched for Oligo_2_3_2 of SNU18. The boxes are color coded to show the serial sections for both Xenium and Visium.

Figure S7 Oligodendrocyte heterogeneity across brain tumor pathologies. (A) Bar plots illustrating the correlation between oligodendrocyte subtypes and brain tumor types. The relative proportions of

oligodendrocyte cell subtypes across histologies reveal an increased presence of Oligo_2_3_2 subtype cells in tissue from the GBM core compared to periphery, control, and IDH mutant tissue samples (depicted in purple). (B) Gene Ontology enrichment analysis of Differentially Expressed Genes (DEGs) between Oligo_2_3_2 and Oligo_2_1 subtype. (C) Xenium data displaying the co-localization pattern of cell types for Oligo_2_1 and Oligo_2_3_2.

Figure S8 Genes related to oligodendrocyte trajectory.

Figure S9 Spatial cell-cell communication. Overview of ligand-receptor interactions between Oligo_2_3_2 and neighboring cell types using CellPhoneDB.

SUPPLEMENTAL TABLES

Supplementary Table 1: Patient characteristics and sequencing information with genomic alterations
Supplementary Table 2: Major cell-type markers based on RCC Level_1 cluster
Supplementary Table 3: Cell type annotations
Supplementary Table 4: Ivy GAP anatomic feature gene sets for spatial annotations
Supplementary Table 5: Public domain data summary
Supplementary Table 6: Cell type markers
Supplementary Table 7: Xenium 10x custom glioma gene panel
Supplementary Table 8: Oligodendrocyte DGE analysis between core and peri
Supplementary Table 9: CellPhoneDB ligand-receptor interactions

Figure 1

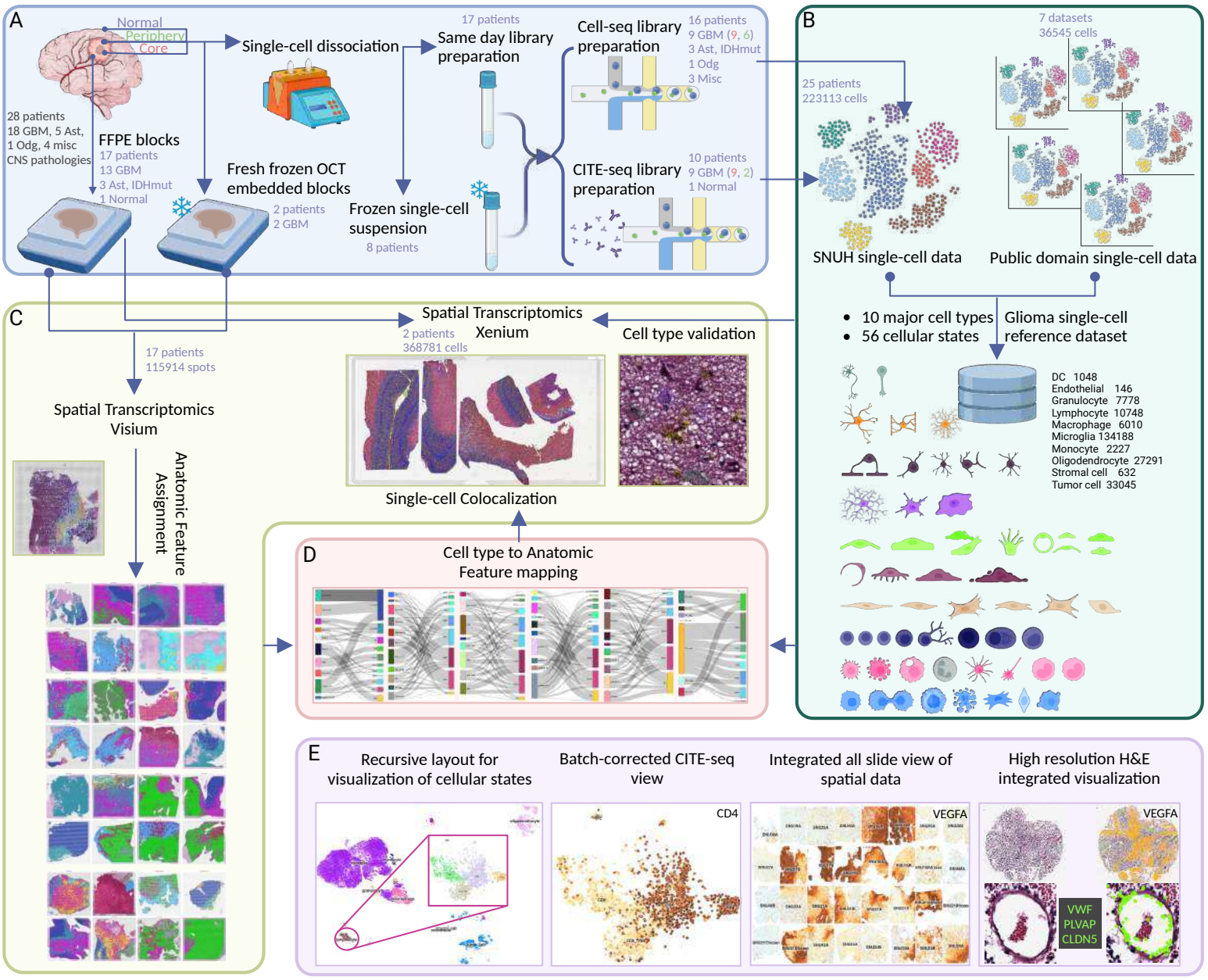

A
Normal
Periphery
Core
28 patients
18 GBM, 5 Ast,
1 Odg, 4 misc
CNS pathologies
FFPE blocks
17 patients
13 GBM
3 Ast, IDHmut
1 Normal
Single-cell dissociation
Fresh frozen OCT
embedded blocks
2 patients
2 GBM
Frozen single-cell
suspension
8 patients
17 patients
Same day library
preparation
Cell-seq library
preparation
16 patients
9 GBM (9, 6)
3 Ast, IDHmut
1 Odg
3 Misc
CITE-seq library
preparation
10 patients
9 GBM (9, 2)
1 Normal
B
25 patients
223113 cells
7 datasets
36545 cells
SNUH single-cell data
Public domain single-cell data
10 major cell types
56 cellular states
Glioma single-cell
reference dataset
DC 1048
Endothelial 146
Granulocyte 7778
Lymphocyte 10748
Macrophage 6010
Microglia 134188
Monocyte 2227
Oligodendrocyte 27291
Stromal cell 632
Tumor cell 33045
C
17 patients
115914 spots
Spatial Transcriptomics
Visium
Anatomic Feature
Assignment
Spatial Transcriptomics
Xenium
2 patients
368781 cells
Cell type validation
Single-cell Colocalization
D
Cell type to Anatomic
Feature mapping
E
Recursive layout for
visualization of cellular states
Batch-corrected CITE-seq
view
CD4
Integrated all slide view of
spatial data
VEGFA
High resolution H&E
integrated visualization
VEGFA
VWF
PLVAP
CLDN5

Figure 2

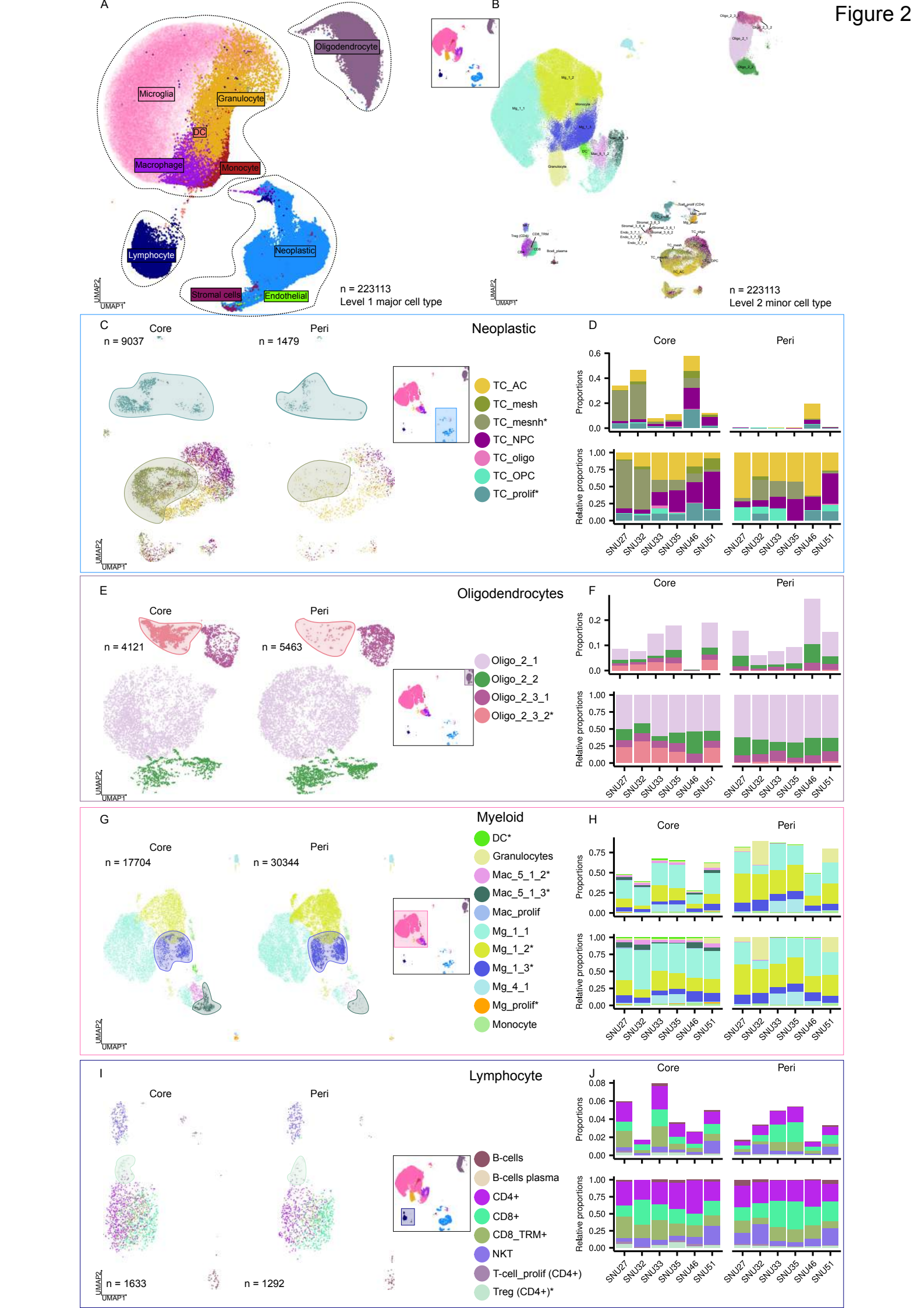

A
Oligodendrocyte
Microglia
Granulocyte
DC
Macrophage
Monocyte
Lymphocyte
Neoplastic
Stromal cells
Endothelial
UMAP2
UMAP1
n = 223113
Level 1 major cell type
B
n = 223113
Level 2 minor cell type
C
Core
n = 9037
Peri
n = 1479
Neoplastic
TC_AC
TC_mesh
TC_mesnh*
TC_NPC
TC_oligo
TC_OPC
TC_prolif*
D
Core
Peri
Proportions
Relative proportions
SNU27
SNU32
SNU33
SNU35
SNU46
SNU51
E
Core
n = 4121
Peri
n = 5463
Oligodendrocytes
Oligo_2_1
Oligo_2_2
Oligo_2_3_1
Oligo_2_3_2*
F
G
Core
n = 17704
Peri
n = 30344
Myeloid
DC*
Granulocytes
Mac_5_1_2*
Mac_5_1_3*
Mac_prolif
Mg_1_1
Mg_1_2*
Mg_1_3*
Mg_4_1
Mg_prolif*
Monocyte
H
I
Core
Peri
n = 1633
n = 1292
Lymphocyte
B-cells
B-cells plasma
CD4+
CD8+
CD8_TRM+
NKT
T-cell_prolif (CD4+)
Treg (CD4+)*
J

Figure 3

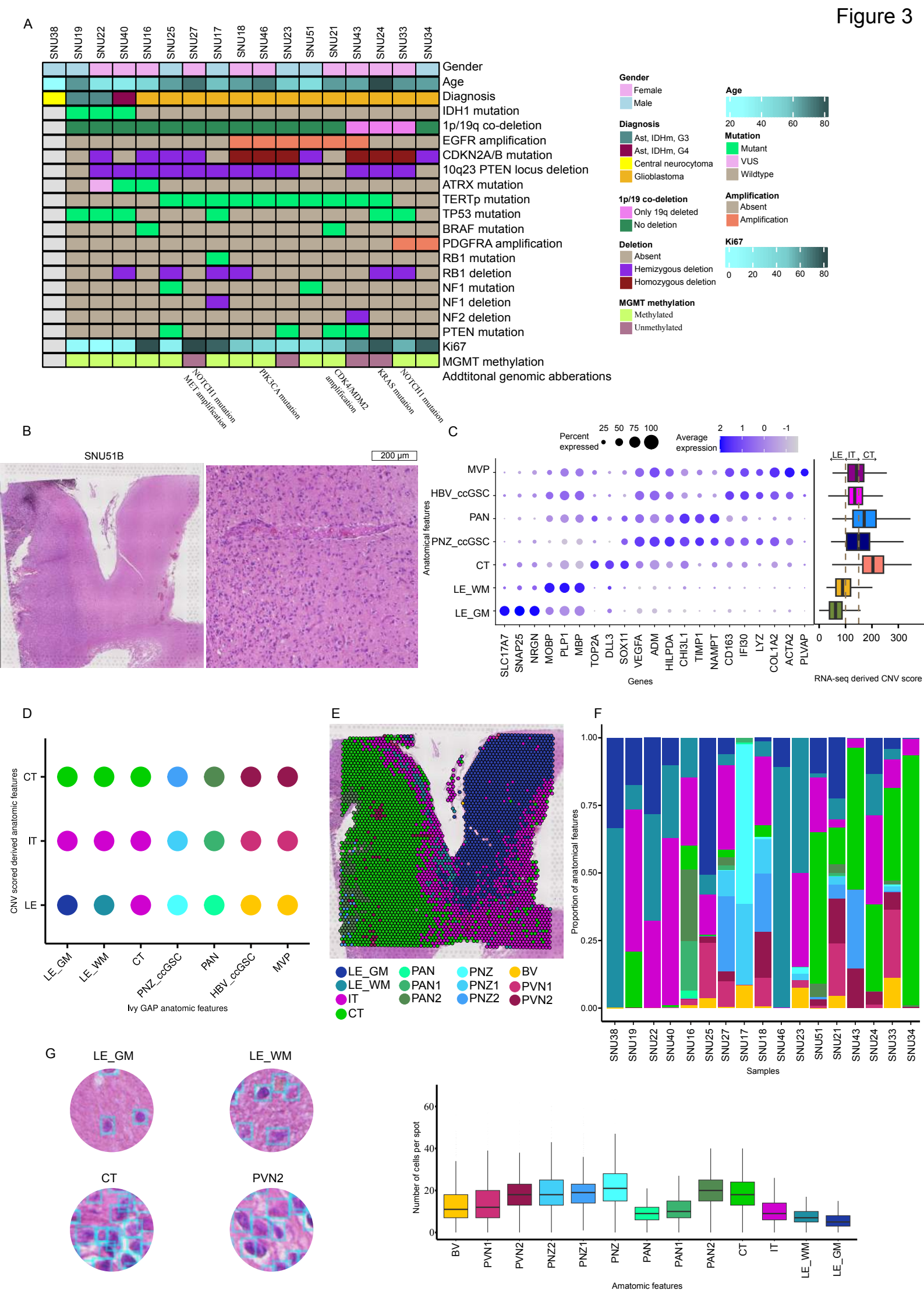
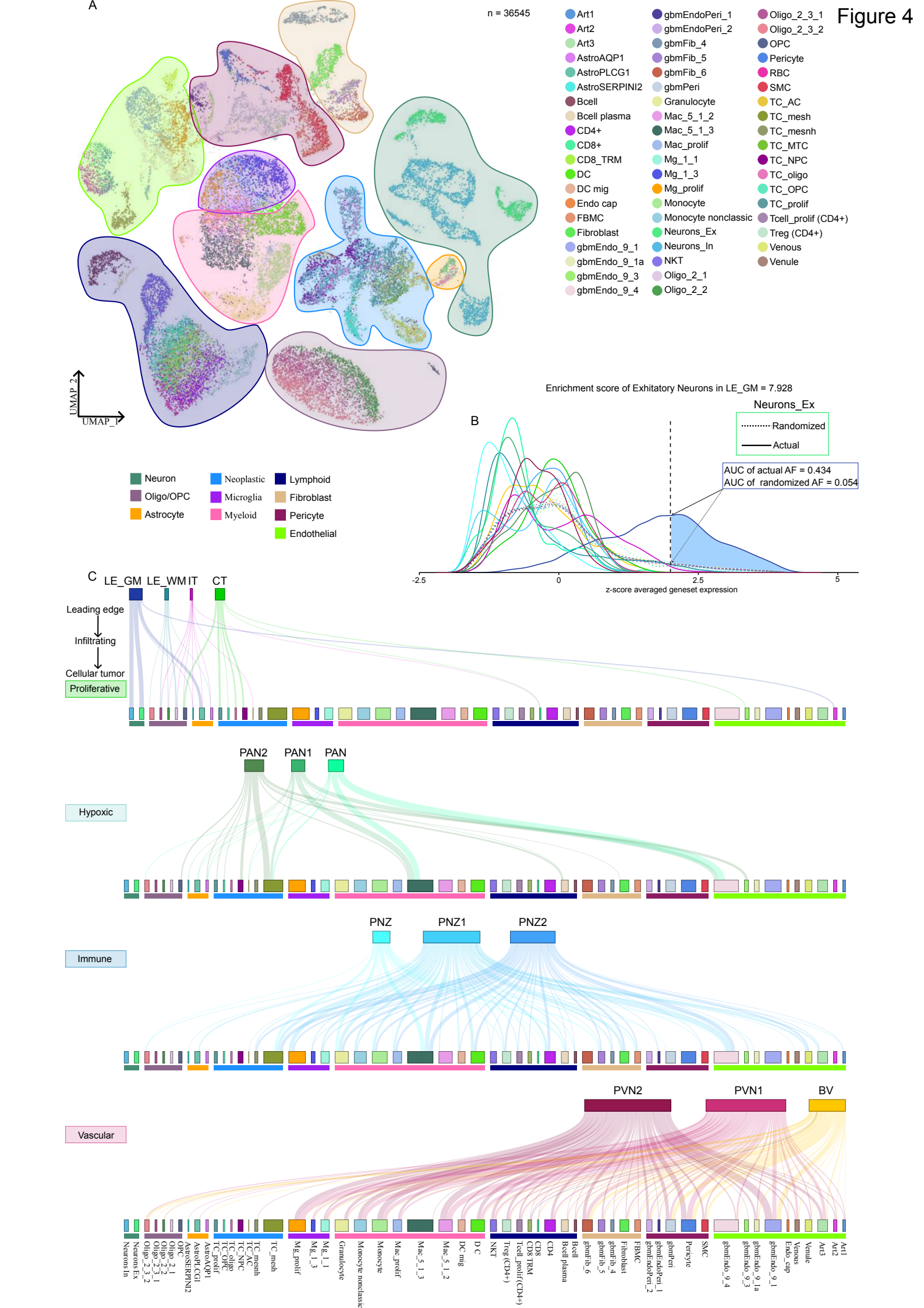

A
SNU38
SNU19
SNU22
SNU40
SNU16
SNU25
SNU27
SNU17
SNU18
SNU46
SNU23
SNU51
SNU21
SNU43
SNU24
SNU33
SNU34
Gender
Age
Diagnosis
IDH1 mutation
1p/19q co-deletion
EGFR amplification
CDKN2A/B mutation
10q23 PTEN locus deletion
ATRX mutation
TERTp mutation
TP53 mutation
BRAF mutation
PDGFRA amplification
RB1 mutation
RB1 deletion
NF1 mutation
NF1 deletion
NF2 deletion
PTEN mutation
Ki67
MGMT methylation
Addtitonal genomic abberations
NOTCH1 mutation
MET amplification
PIK3CA mutation
CDK4/MDM2 amplification
KRAS mutation
NOTCH1 mutation
Gender
Female
Male
Age
20 40 60 80
Diagnosis
Ast, IDHm, G3
Ast, IDHm, G4
Central neurocytoma
Glioblastoma
Mutation
Mutant
VUS
Wildtype
1p/19 co-deletion
Only 19q deleted
No deletion
Amplification
Absent
Amplification
Deletion
Absent
Hemizygous deletion
Homozygous deletion
Ki67
0 20 40 60 80
MGMT methylation
Methylated
Unmethylated
B
SNU51B
200 μm
C
Percent expressed
25 50 75 100
Average expression
2 1 0 -1
MVP
HBV_ccGSC
PAN
PNZ_ccGSC
CT
LE_WM
LE_GM
Anatomical features
LE IT CT
0 100 200 300
RNA-seq derived CNV score
SLC17A7
SNAP25
NRGN
MOBP
PLP1
MBP
TOP2A
DLL3
SOX11
VEGFA
ADM
HILPDA
CHI3L1
TIMP1
NAMPT
CD163
IFI30
LYZ
COL1A2
ACTA2
PLVAP
Genes
D
CT
IT
LE
CNV scored derived anatomic features
LE_GM
LE_WM
CT
PNZ_ccGSC
PAN
HBV_ccGSC
MVP
Ivy GAP anatomic features
E
LE_GM
LE_WM
IT
CT
PAN
PAN1
PAN2
PNZ
PNZ1
PNZ2
BV
PVN1
PVN2
F
1.00
0.75
0.50
0.25
0.00
Proportion of anatomical features
SNU38
SNU19
SNU22
SNU40
SNU16
SNU25
SNU27
SNU17
SNU18
SNU46
SNU23
SNU51
SNU21
SNU43
SNU24
SNU33
SNU34
Samples
G
LE_GM
LE_WM
CT
PVN2
60
40
20
0
Number of cells per spot
BV
PVN1
PVN2
PNZ2
PNZ1
PNZ
PAN
PAN1
PAN2
CT
IT
LE_WM
LE_GM
Amatomic features

Figure 4


A
n = 36545
Art1
Art2
Art3
AstroAQP1
AstroPLCG1
AstroSERPINI2
Bcell
Bcell plasma
CD4+
CD8+
CD8_TRM
DC
DC mig
Endo cap
FBMC
Fibroblast
gbmEndo_9_1
gbmEndo_9_1a
gbmEndo_9_3
gbmEndo_9_4
gbmEndoPeri_1
gbmEndoPeri_2
gbmFib_4
gbmFib_5
gbmFib_6
gbmPeri
Granulocyte
Mac_5_1_2
Mac_5_1_3
Mac_prolif
Mg_1_1
Mg_1_3
Mg_prolif
Monocyte
Monocyte nonclassic
Neurons_Ex
Neurons_In
NKT
Oligo_2_1
Oligo_2_2
Oligo_2_3_1
Oligo_2_3_2
OPC
Pericyte
RBC
SMC
TC_AC
TC_mesh
TC_mesnh
TC_MTC
TC_NPC
TC_oligo
TC_OPC
TC_prolif
Tcell_prolif (CD4+)
Treg (CD4+)
Venous
Venule
UMAP_2
UMAP_1
Neuron
Oligo/OPC
Astrocyte
Neoplastic
Microglia
Myeloid
Lymphoid
Fibroblast
Pericyte
Endothelial
B
Enrichment score of Exhitatory Neurons in LE_GM = 7.928
Neurons_Ex
Randomized
Actual
AUC of actual AF = 0.434
AUC of randomized AF = 0.054
-2.5
0
2.5
5
z-score averaged geneset expression
C
LE_GM
LE_WM
IT
CT
Leading edge
Infiltrating
Cellular tumor
Proliferative
PAN2
PAN1
PAN
Hypoxic
PNZ
PNZ1
PNZ2
Immune
PVN2
PVN1
BV
Vascular

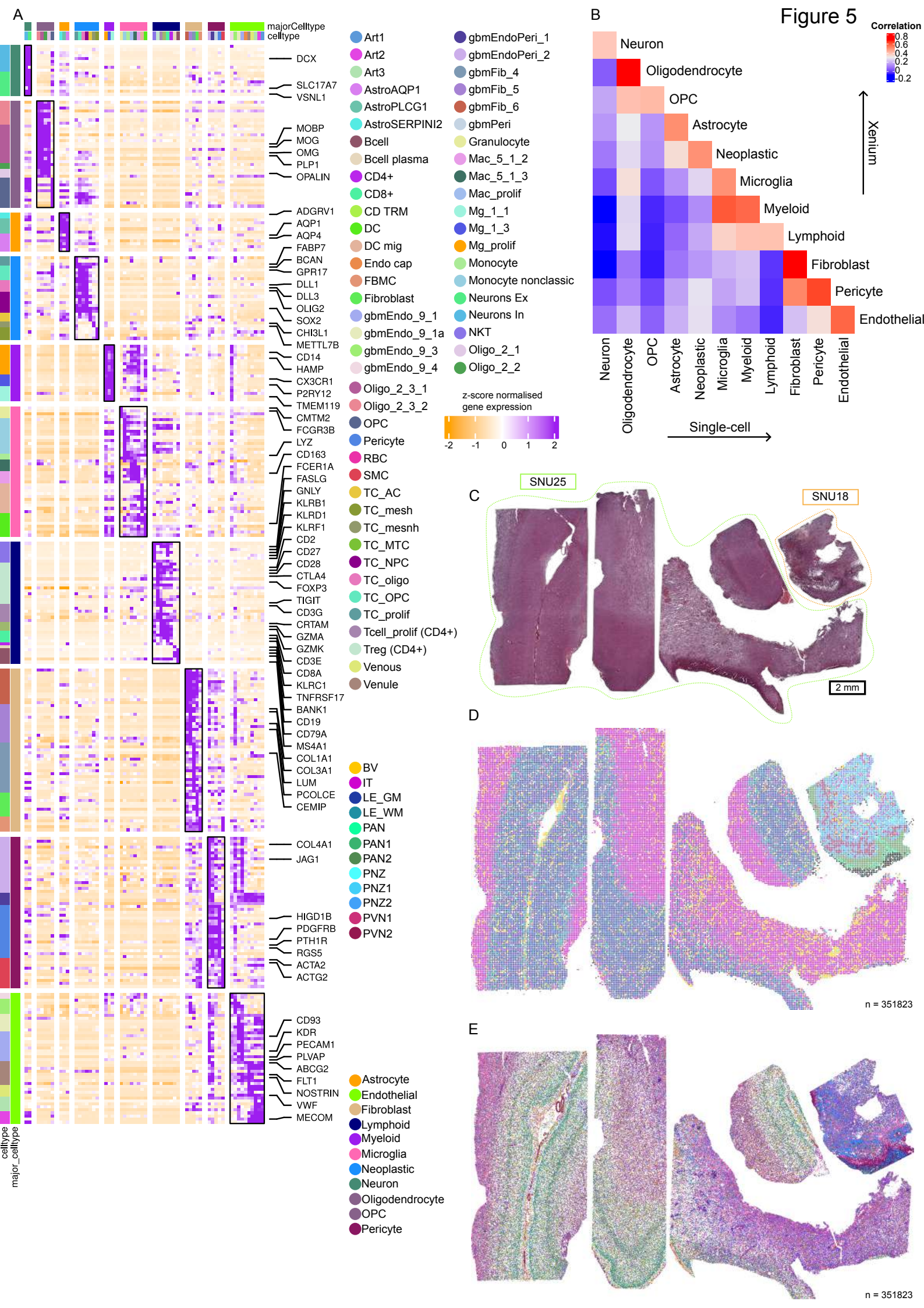
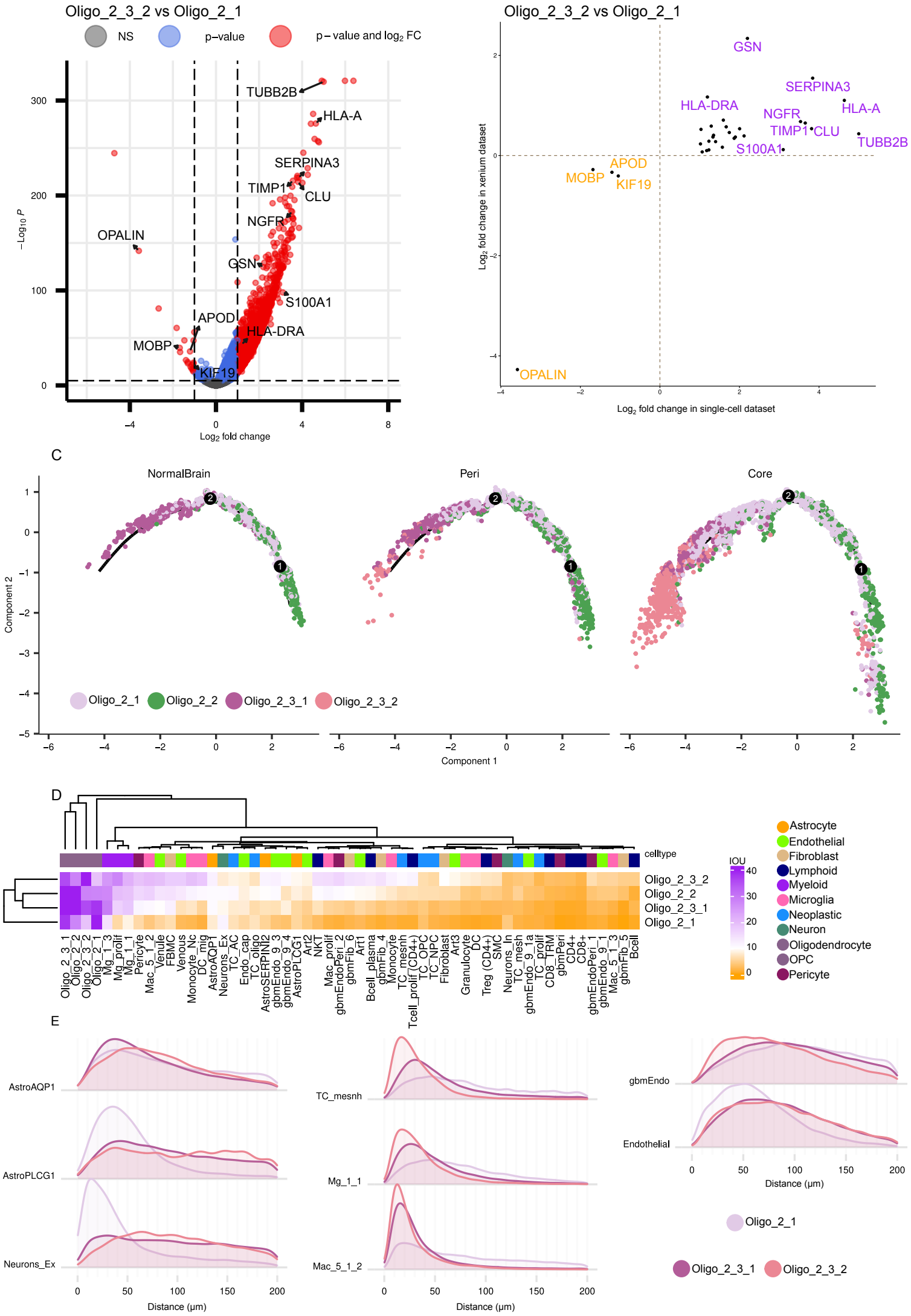


Figure 5
A
majorCelltype
celltype
DCX
SLC17A7
VSNL1
MOBP
MOG
OMG
PLP1
OPALIN
ADGRV1
AQP1
AQP4
FABP7
BCAN
GPR17
DLL1
DLL3
OLIG2
SOX2
CHI3L1
METTL7B
CD14
HAMP
CX3CR1
P2RY12
TMEM119
CMTM2
FCGR3B
LYZ
CD163
FCER1A
FASLG
GNLY
KLRB1
KLRD1
KLRF1
CD2
CD27
CD28
CTLA4
FOXP3
TIGIT
CD3G
CRTAM
GZMA
GZMK
CD3E
CD8A
KLRC1
TNFRSF17
BANK1
CD19
CD79A
MS4A1
COL1A1
COL3A1
LUM
PCOLCE
CEMIP
COL4A1
JAG1
HIGD1B
PDGFRB
PTH1R
RGS5
ACTA2
ACTG2
CD93
KDR
PECAM1
PLVAP
ABCG2
FLT1
NOSTRIN
VWF
MECOM
celltype
major_celltype
Art1
Art2
Art3
AstroAQP1
AstroPLCG1
AstroSERPINI2
Bcell
Bcell plasma
CD4+
CD8+
CD TRM
DC
DC mig
Endo cap
FBMC
Fibroblast
gbmEndo_9_1
gbmEndo_9_1a
gbmEndo_9_3
gbmEndo_9_4
Oligo_2_3_1
Oligo_2_3_2
OPC
Pericyte
RBC
SMC
TC_AC
TC_mesh
TC_mesnh
TC_MTC
TC_NPC
TC_oligo
TC_OPC
TC_prolif
Tcell_prolif (CD4+)
Treg (CD4+)
Venous
Venule
BV
IT
LE_GM
LE_WM
PAN
PAN1
PAN2
PNZ
PNZ1
PNZ2
PVN1
PVN2
Astrocyte
Endothelial
Fibroblast
Lymphoid
Myeloid
Microglia
Neoplastic
Neuron
Oligodendrocyte
OPC
Pericyte
gbmEndoPeri_1
gbmEndoPeri_2
gbmFib_4
gbmFib_5
gbmFib_6
gbmPeri
Granulocyte
Mac_5_1_2
Mac_5_1_3
Mac_prolif
Mg_1_1
Mg_1_3
Mg_prolif
Monocyte
Monocyte nonclassic
Neurons Ex
Neurons In
NKT
Oligo_2_1
Oligo_2_2
z-score normalised
gene expression
-2
-1
0
1
2
B
Correlation
0.8
0.6
0.4
0.2
0
-0.2
Neuron
Oligodendrocyte
OPC
Astrocyte
Neoplastic
Microglia
Myeloid
Lymphoid
Fibroblast
Pericyte
Endothelial
Xenium
Single-cell
C
SNU25
SNU18
2 mm
D
n = 351823
E
n = 351823

Figure 6


A
Oligo_2_3_2 vs Oligo_2_1
NS
p–value
p – value and log2 FC
TUBB2B
HLA-A
SERPINA3
TIMP1
CLU
NGFR
OPALIN
GSN
S100A1
APOD
HLA-DRA
MOBP
KIF19
−Log10 P
Log2 fold change
B
Oligo_2_3_2 vs Oligo_2_1
GSN
SERPINA3
HLA-DRA
HLA-A
NGFR
TIMP1
CLU
TUBB2B
S100A1
APOD
MOBP
KIF19
OPALIN
Log2 fold change in xenium dataset
Log2 fold change in single-cell dataset
C
NormalBrain
Peri
Core
Component 2
Component 1
Oligo_2_1
Oligo_2_2
Oligo_2_3_1
Oligo_2_3_2
D
celltype
Oligo_2_3_2
Oligo_2_2
Oligo_2_3_1
Oligo_2_1
IOU
Astrocyte
Endothelial
Fibroblast
Lymphoid
Myeloid
Microglia
Neoplastic
Neuron
Oligodendrocyte
OPC
Pericyte
E
AstroAQP1
AstroPLCG1
Neurons_Ex
TC_mesnh
Mg_1_1
Mac_5_1_2
gbmEndo
Endothelial
Distance (μm)
Oligo_2_1
Oligo_2_3_1
Oligo_2_3_2

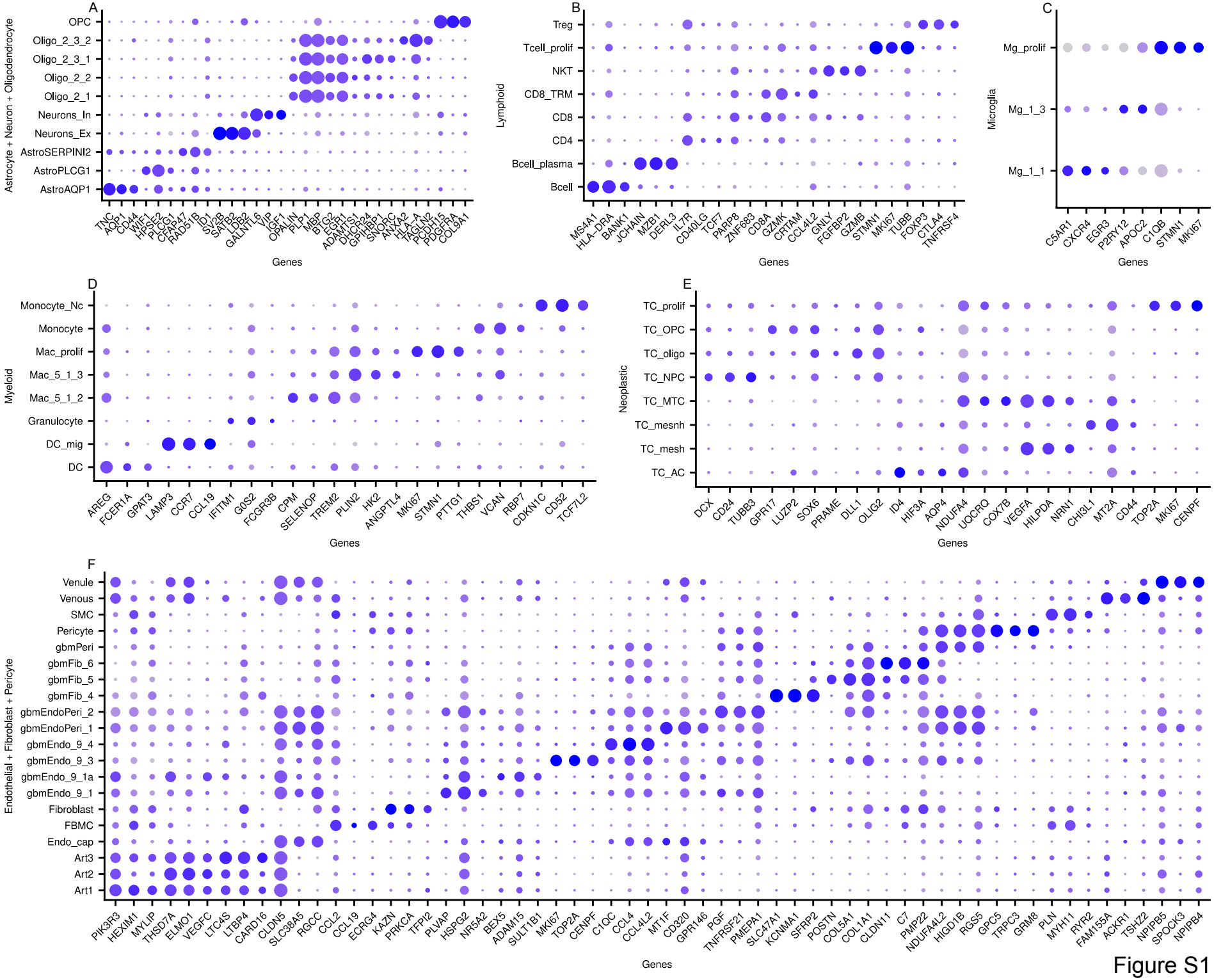

A
Astrocyte + Neuron + Oligodendrocyte
OPC
Oligo_2_3_2
Oligo_2_3_1
Oligo_2_2
Oligo_2_1
Neurons_In
Neurons_Ex
AstroSERPINI2
AstroPLCG1
AstroAQP1
TNC AQP1 CD44 WIF1 HPSE2 PLCG1 GFAP47 RAD51B ID1 SV2B SATB2 LDB2 GALNTL6 VIP IGF1 OPALIN PLP1 MBP BTG2 EGR1 ADAMTS1 DHCR24 GPHBP1 SNORC ANXA2 HLA-A TAGLN2 PODXL2 PDGFRA COL9A1
Genes
B
Lymphoid
Treg
Tcell_prolif
NKT
CD8_TRM
CD8
CD4
Bcell_plasma
Bcell
MS4A1 HLA-DRA BANK1 JCHAIN MZB1 DERL3 IL7R CD40LG TCF7 PARP8 ZNF683 CD8A GZMK CRTAM CCL4L2 GNLY FGFBP2 GZMB STMN1 MKI67 TUBB FOXP3 CTLA4 TNFRSF4
Genes
C
Microglia
Mg_prolif
Mg_1_3
Mg_1_1
C5AR1 CXCR4 EGR3 P2RY12 APOC2 C1QB STMN1 MKI67
Genes
D
Myeloid
Monocyte_Nc
Monocyte
Mac_prolif
Mac_5_1_3
Mac_5_1_2
Granulocyte
DC_mig
DC
AREG FCER1A GPAT3 LAMP3 CCR7 CCL19 IFITM1 G0S2 FCGR3B CPM SELENOP TREM2 PLIN2 HK2 ANGPTL4 MKI67 STMN1 PTTG1 THBS1 VCAN RBP7 CDKN1C CD52 TCF7L2
Genes
E
Neoplastic
TC_prolif
TC_OPC
TC_oligo
TC_NPC
TC_MTC
TC_mesnh
TC_mesh
TC_AC
DCX CD24 TUBB3 GPR17 LUZP2 SOX6 PRAME DLL1 OLIG2 ID4 HIF3A AQP4 NDUFA4 UQCRQ COX7B VEGFA HILPDA NRN1 CHI3L1 MT2A CD44 TOP2A MKI67 CENPF
Genes
F
Endothelial + Fibroblast + Pericyte
Venule
Venous
SMC
Pericyte
gbmPeri
gbmFib_6
gbmFib_5
gbmFib_4
gbmEndoPeri_2
gbmEndoPeri_1
gbmEndo_9_4
gbmEndo_9_3
gbmEndo_9_1a
gbmEndo_9_1
Fibroblast
FBMC
Endo_cap
Art3
Art2
Art1
PIK3R3 HEXIM1 MYLIP THSD7A ELMO1 VEGFC LTC4S LTBP4 CARD16 CLDN5 SLC38A5 RGCC CCL2 CCL19 ECRG4 KAZN PRKCA TFPI2 PLVAP HSPG2 NR5A2 BEX5 ADAM15 SULT1B1 MKI67 TOP2A CENPF C1QC CCL4 CCL4L2 MT1F CD320 GPR146 PGF TNFRSF21 PMEPA1 SLC47A1 KCNMA1 SFRP2 POSTN COL5A1 COL1A1 CLDN11 C7 PMP22 NDUFA4L2 HIGD1B RGS5 GPC5 TRPC3 GRM8 PLN MYH11 RYR2 FAM155A ACKR1 TSHZ2 NPIPB5 SPOCK3 NPIPB4
Genes

Figure S1

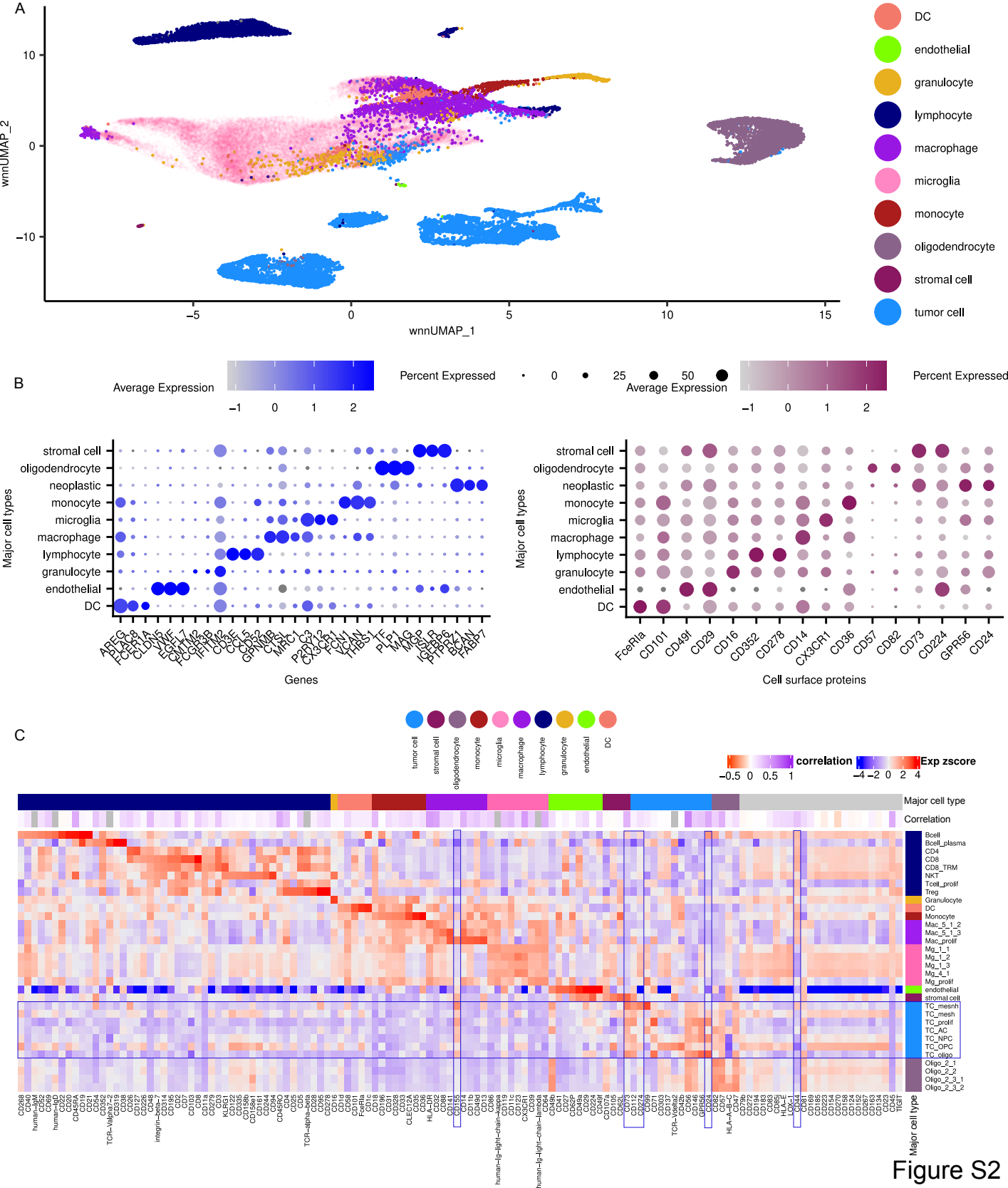

A
wnnUMAP_2
wnnUMAP_1
DC
endothelial
granulocyte
lymphocyte
macrophage
microglia
monocyte
oligodendrocyte
stromal cell
tumor cell
B
Average Expression
Percent Expressed
Major cell types
Genes
Cell surface proteins
C
correlation
Exp zscore
Major cell type
Correlation

Figure S2

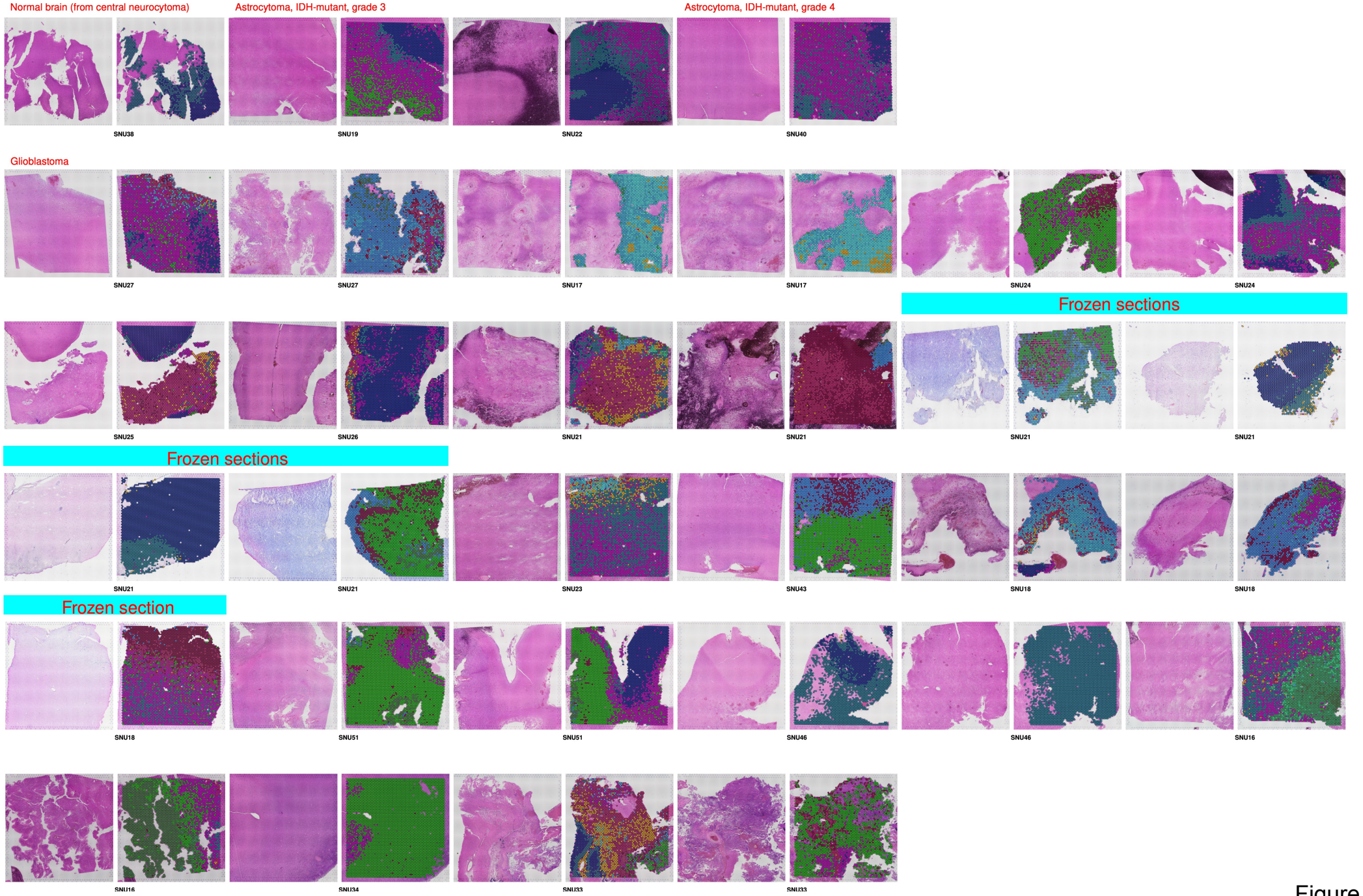

Normal brain (from central neurocytoma)
Astrocytoma, IDH-mutant, grade 3
Astrocytoma, IDH-mutant, grade 4
SNU38
SNU19
SNU22
SNU40
Glioblastoma
SNU27
SNU27
SNU17
SNU17
SNU24
SNU24
Frozen sections
SNU25
SNU26
SNU21
SNU21
SNU21
SNU21
Frozen sections
SNU21
SNU21
SNU23
SNU43
SNU18
SNU18
Frozen section
SNU18
SNU51
SNU51
SNU46
SNU46
SNU16
SNU16
SNU34
SNU33
SNU33

Figure S3

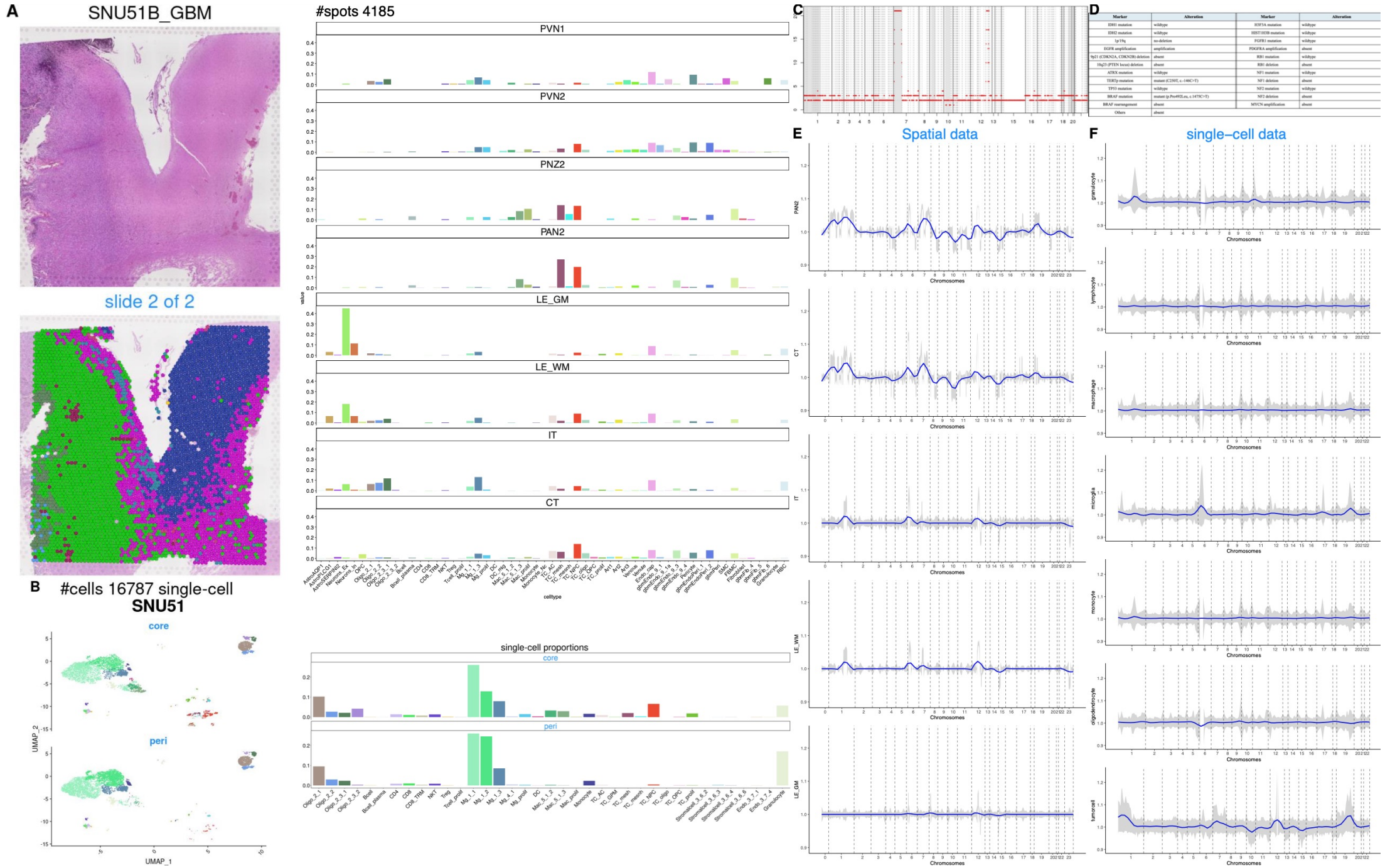

A
SNU51B_GBM
slide 2 of 2
B
#cells 16787 single-cell
SNU51
core
peri
UMAP_1
UMAP_2
#spots 4185
PVN1
PVN2
PNZ2
PAN2
LE_GM
LE_WM
IT
CT
value
celltype
single-cell proportions
core
peri
C
D
Marker
Alteration
IDH1 mutation
IDH2 mutation
1p/19q
EGFR amplification
9p21 (CDKN2A, CDKN2B) deletion
10q23 (PTEN locus) deletion
ATRX mutation
TERTp mutation
TP53 mutation
BRAF mutation
BRAF rearrangement
Others
H3F3A mutation
HIST1H3B mutation
FGFR1 mutation
PDGFRA amplification
RB1 mutation
RB1 deletion
NF1 mutation
NF1 deletion
NF2 mutation
NF2 deletion
MYCN amplification
E
Spatial data
F
single–cell data
PAN2
CT
IT
LE_WM
LE_GM
Chromosomes
granulocyte
lymphocyte
macrophage
microglia
monocyte
oligodendrocyte
tumorcell

Figure S4

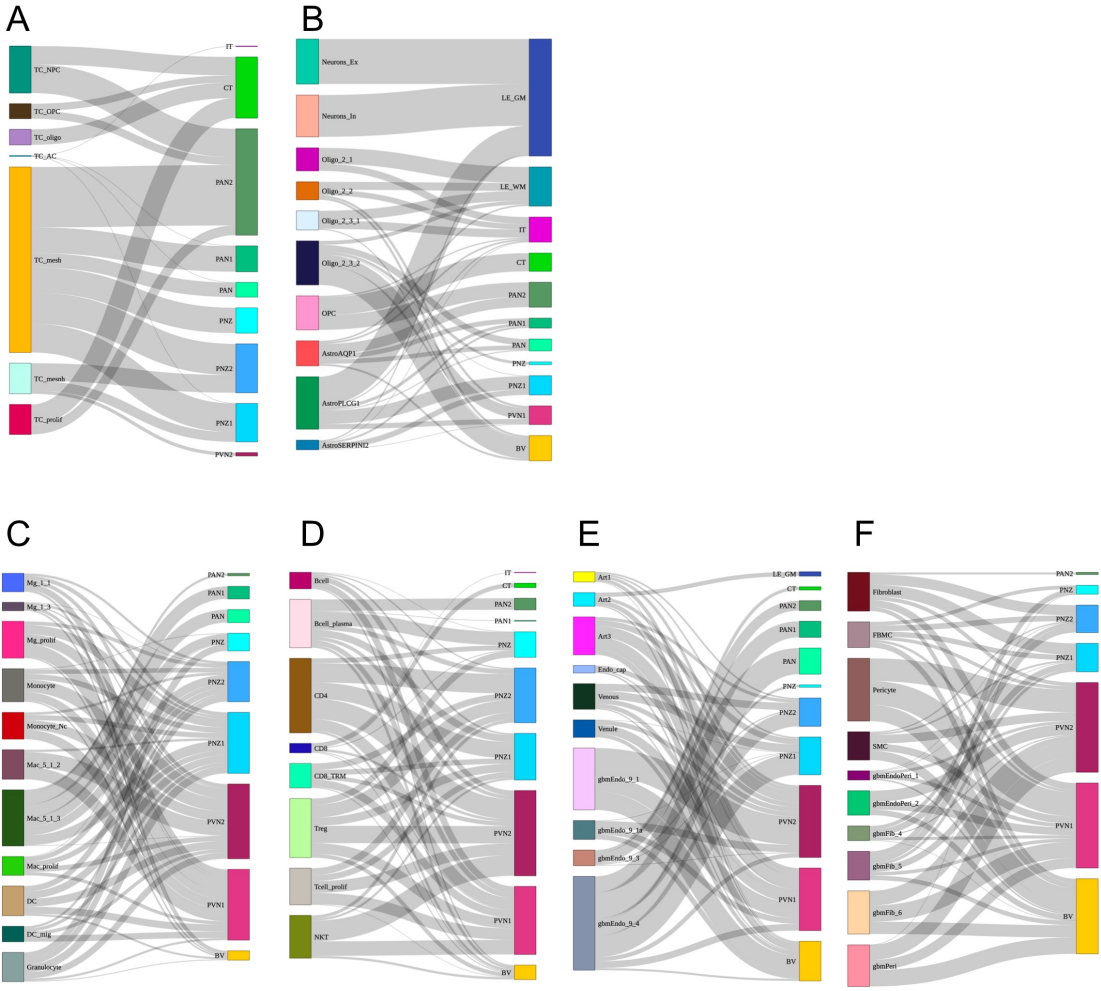

A
TC_NPC
TC_OPC
TC_oligo
TC_AC
TC_mesh
TC_meshh
TC_prolif
IT
CT
PAN2
PAN1
PAN
PNZ
PNZ2
PNZ1
PVN2
B
Neurons_Ex
Neurons_In
Oligo_2_1
Oligo_2_2
Oligo_2_3_1
Oligo_2_3_2
OPC
AstroAQP1
AstroPLCG1
AstroSERPINI2
LE_GM
LE_WM
IT
CT
PAN2
PAN1
PAN
PNZ
PNZ1
PVN1
BV
C
Mg_1_1
Mg_1_3
Mg_prolif
Monocyte
Monocyte_Nc
Mac_5_1_2
Mac_5_1_3
Mac_prolif
DC
DC_mig
Granulocyte
PAN2
PAN1
PAN
PNZ
PNZ2
PNZ1
PVN2
PVN1
BV
D
Bcell
Bcell_plasma
CD4
CD8
CD8_TRM
Treg
Tcell_prolif
NKT
IT
CT
PAN2
PAN1
PNZ
PNZ2
PNZ1
PVN2
PVN1
BV
E
Art1
Art2
Art3
Endo_cap
Venous
Venule
gbmEndo_9_1
gbmEndo_9_2
gbmEndo_9_3
gbmEndo_9_4
LE_GM
CT
PAN2
PAN1
PAN
PNZ
PNZ2
PNZ1
PVN2
PVN1
BV
F
Fibroblast
FBMC
Pericyte
SMC
gbmEndoPeri_1
gbmEndoPeri_2
gbmFib_4
gbmFib_5
gbmFib_6
gbmPeri
PAN2
PNZ
PNZ2
PNZ1
PVN2
PVN1
BV

Figure S5

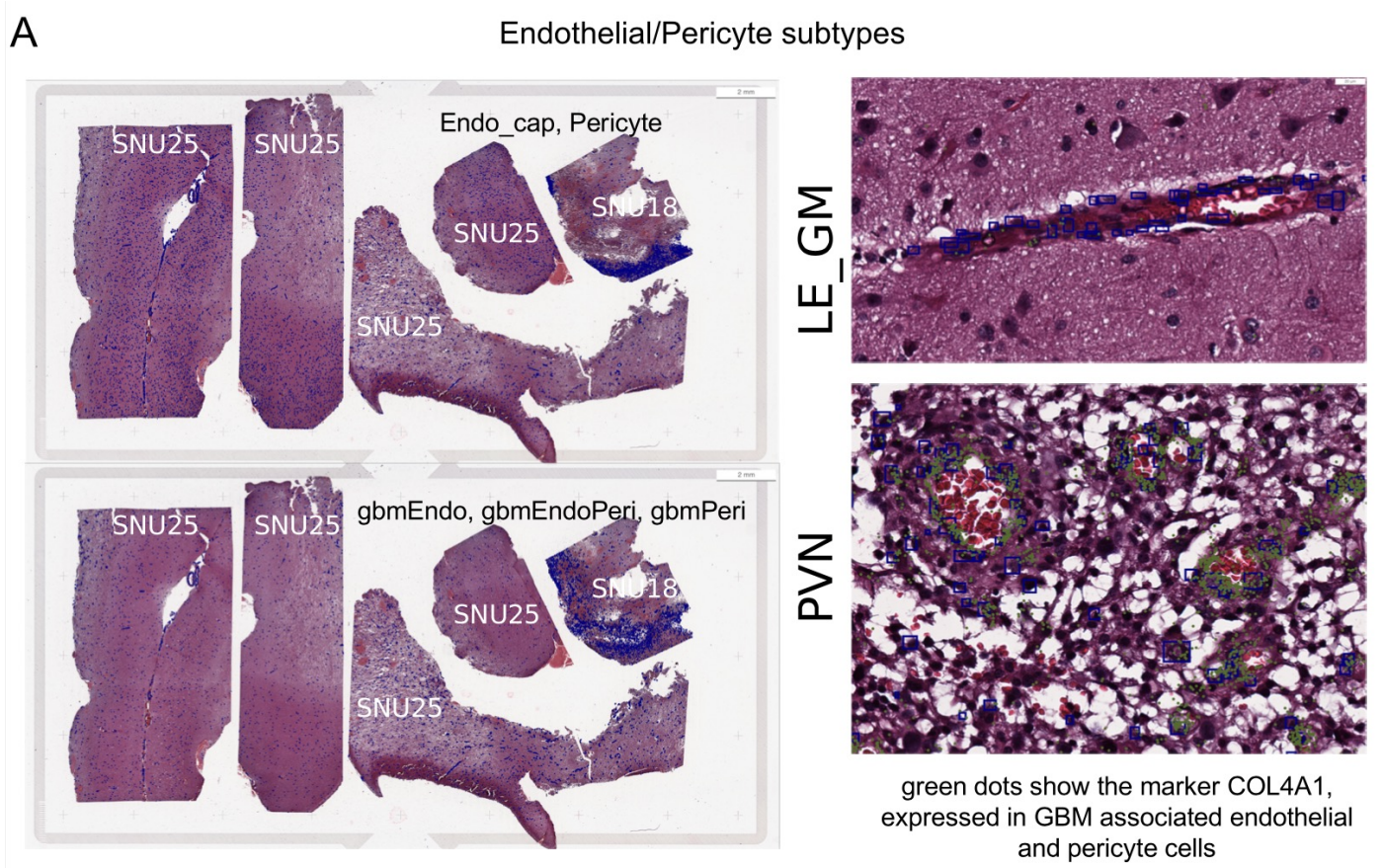
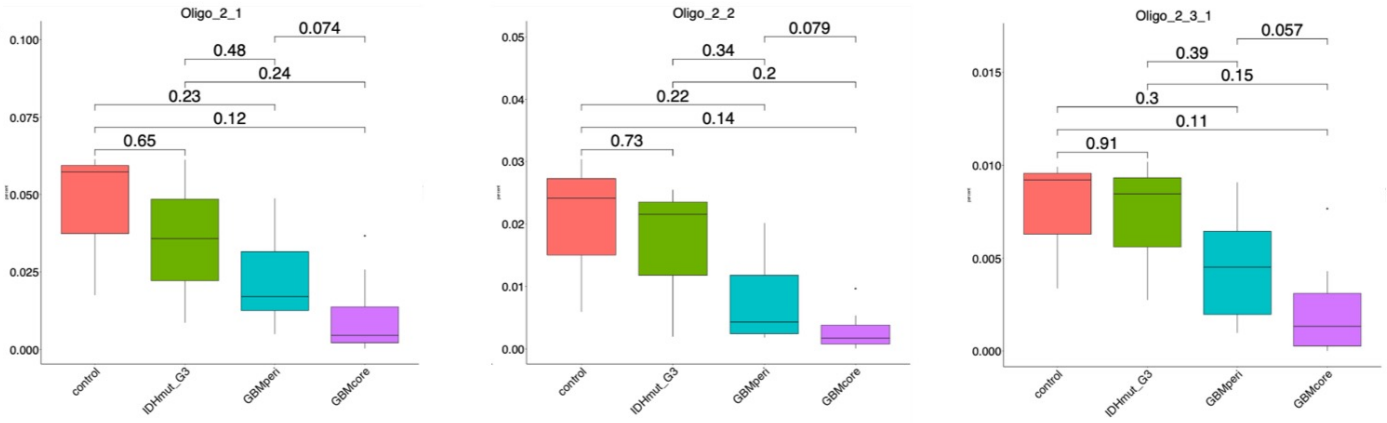
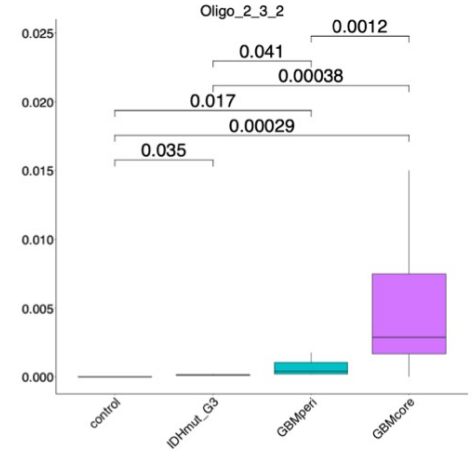

A
Endothelial/Pericyte subtypes
SNU25
SNU25
Endo_cap, Pericyte
SNU18
SNU25
SNU25
SNU25
SNU25
gbmEndo, gbmEndoPeri, gbmPeri
SNU18
SNU25
SNU25
LE_GM
PVN
green dots show the marker COL4A1, expressed in GBM associated endothelial and pericyte cells
B
Oligodendrocyte subtypes
SNU25
SNU25
SNU18
SNU25
Oligo_2_1
Oligo_2_3_1
Oligo_2_3_2

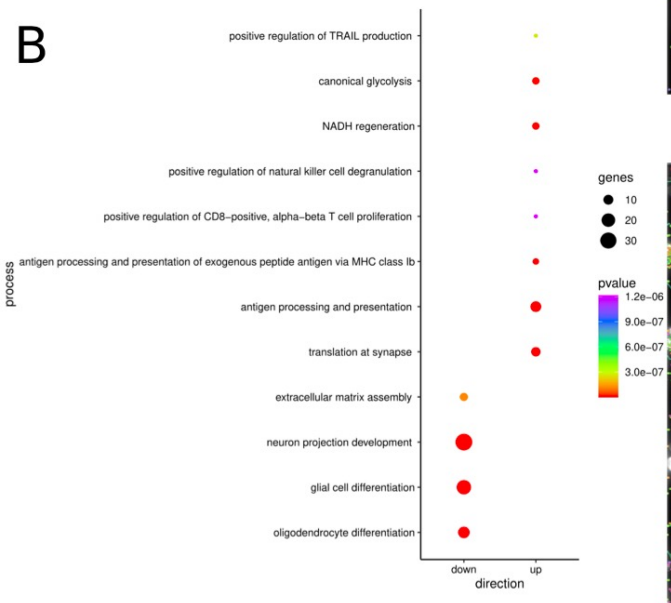
Figure S6

A



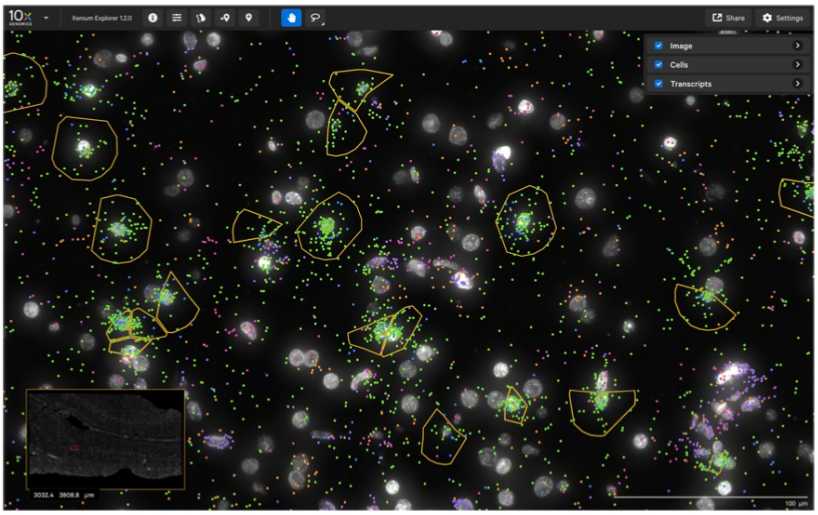


B

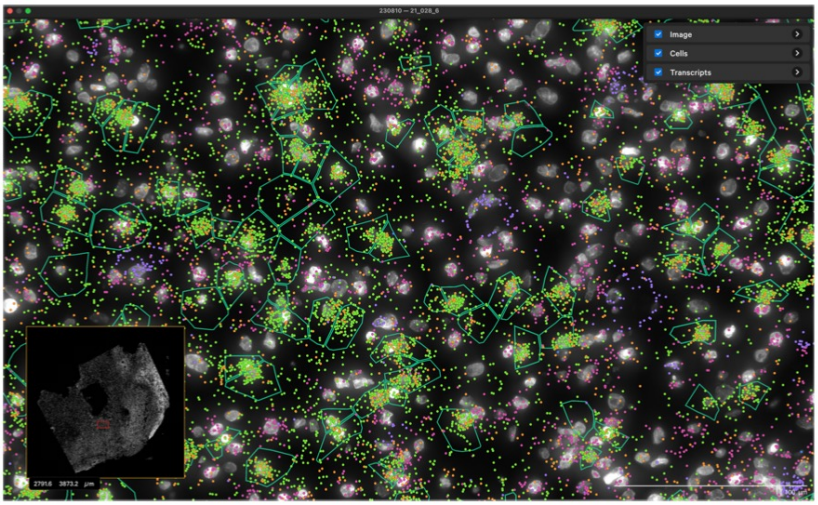


C

Oligo_2_1



Oligo_2_3_2



green (oligo markers) - plp1, mobp,
orange - GSN (2_3_2 marker)
blue - Opalin (2_1 marker)
purple (vascular markers) - abcg2, acta2, cldn5, esm1, mecom, plvap, vwf,
pink (immune markers) - cd14, cd68, IL7R, ptprc

Figure S7

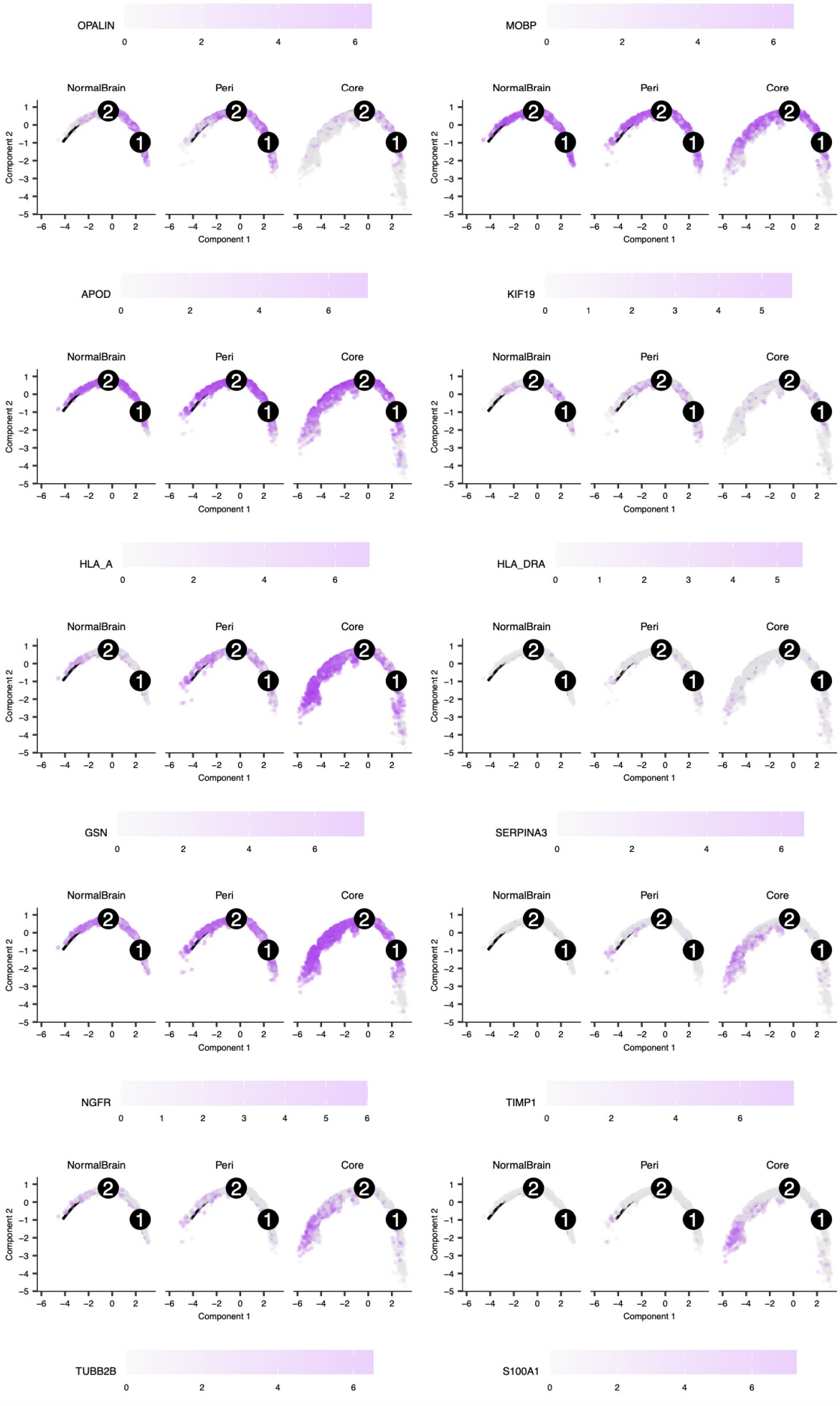

OPALIN
MOBP
APOD
KIF19
HLA_A
HLA_DRA
GSN
SERPINA3
NGFR
TIMP1
TUBB2B
S100A1
NormalBrain
Peri
Core
Component 1
Component 2

Figure S8

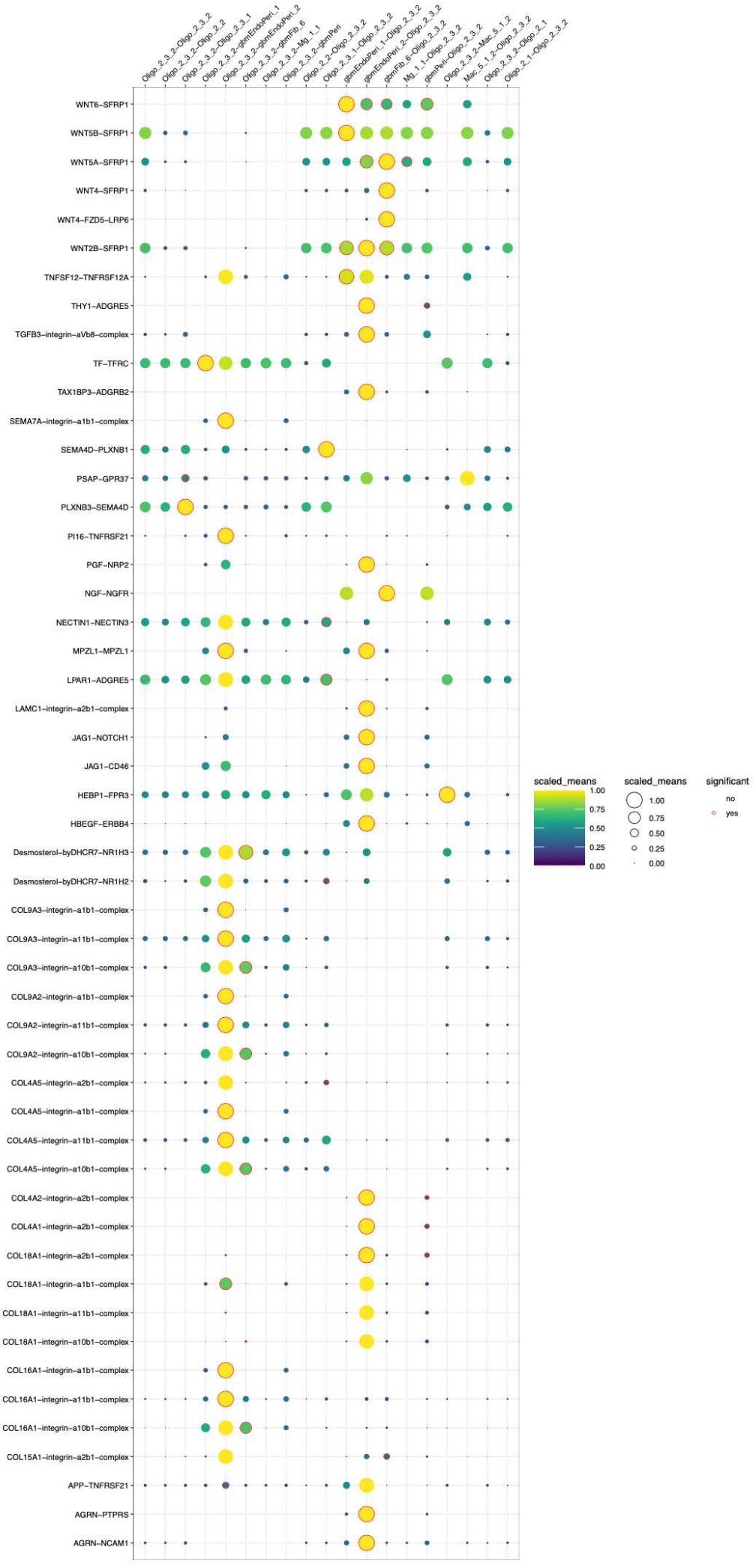

Oligo_2_3_2–Oligo_2_3_2
Oligo_2_3_2–Oligo_2_2
Oligo_2_3_2–Oligo_2_3_1
Oligo_2_3_2–gbmEndoPeri_1
Oligo_2_3_2–gbmEndoPeri_2
Oligo_2_3_2–gbmFib_6
Oligo_2_3_2–Mg_1_1
Oligo_2_3_2–gbmPeri
Oligo_2_2–Oligo_2_3_2
gbmEndoPeri_1–Oligo_2_3_2
gbmEndoPeri_2–Oligo_2_3_2
gbmFib_6–Oligo_2_3_2
Mg_1_1–Oligo_2_3_2
gbmPeri–Oligo_2_3_2
Oligo_2_3_2–Mac_5_1_2
Oligo_2_3_2–Oligo_2_2_1
Oligo_2_1–Oligo_2_3_2
WNT6–SFRP1
WNT5B–SFRP1
WNT5A–SFRP1
WNT4–SFRP1
WNT4–FZD5–LRP6
WNT2B–SFRP1
TNFSF12–TNFRSF12A
THY1–ADGRE5
TGFB3–integrin–aVb8–complex
TF–TFRC
TAX1BP3–ADGRB2
SEMA7A–integrin–a1b1–complex
SEMA4D–PLXNB1
PSAP–GPR37
PLXNB3–SEMA4D
PI16–TNFRSF21
PGF–NRP2
NGF–NGFR
NECTIN1–NECTIN3
MPZL1–MPZL1
LPAR1–ADGRE5
LAMC1–integrin–a2b1–complex
JAG1–NOTCH1
JAG1–CD46
HEBP1–FPR3
HBEGF–ERBB4
Desmosterol–byDHCR7–NR1H3
Desmosterol–byDHCR7–NR1H2
COL9A3–integrin–a1b1–complex
COL9A3–integrin–a11b1–complex
COL9A3–integrin–a10b1–complex
COL9A2–integrin–a1b1–complex
COL9A2–integrin–a11b1–complex
COL9A2–integrin–a10b1–complex
COL4A5–integrin–a2b1–complex
COL4A5–integrin–a1b1–complex
COL4A5–integrin–a11b1–complex
COL4A5–integrin–a10b1–complex
COL4A2–integrin–a2b1–complex
COL4A1–integrin–a2b1–complex
COL18A1–integrin–a2b1–complex
COL18A1–integrin–a1b1–complex
COL18A1–integrin–a11b1–complex
COL18A1–integrin–a10b1–complex
COL16A1–integrin–a1b1–complex
COL16A1–integrin–a11b1–complex
COL16A1–integrin–a10b1–complex
COL15A1–integrin–a2b1–complex
APP–TNFRSF21
AGRN–PTPRS
AGRN–NCAM1
scaled_means
1.00
0.75
0.50
0.25
0.00
scaled_means
1.00
0.75
0.50
0.25
0.00
significant
no
yes

Figure S9

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryTables13May.pdf