Glioblastoma is spatially organized by neurodevelopmental programs and a glial-like wound healing response

Alejandro Mossi Albiach[1†], Jokubas Janusauskas[1†], Ivana Kapustová[1], Egle Kvedaraite[1], Simone Codeluppi[1,*], Johannes B. Munting[1], Lars E. Borm[1,3,4], Jesper Kjaer Jacobsen[1,5], Alia Shamikh[2,6], Oscar Persson[7,8], Sten Linnarsson[1#]

[1]Department of Medical Biochemistry and Biophysics, Karolinska Institute, 17177 Stockholm, Sweden. [2]Department of Clinical Pathology and Cancer Diagnostics, Karolinska University Hospital, 17164 Stockholm, Sweden. [3]Center for Brain & Disease Research, VIB, 3000 Leuven, Belgium. [4]Department of Human Genetics, KU Leuven, 3000 Leuven, Belgium. [5]Department of Neurology, Karolinska University Hospital, 17164 Stockholm, Sweden. [6]Department of Oncology-Pathology, Karolinska Institute, 17177 Stockholm, Sweden. [7]Department of Clinical Neuroscience, Karolinska Institute, 17177 Stockholm, Sweden. [8]Department of Neurosurgery, Karolinska University Hospital, 17164 Stockholm, Sweden

*Current affiliation: Moleculent, 17165 Stockholm, Sweden
†These authors contributed equally

Glioblastoma is the deadliest brain cancer, characterized by great cellular diversity and unique histology. To understand the spatial organization of transcriptional cell states, we mapped the expression of 888 genes in centimeter-scale tissue sections from a large patient cohort. We found a hierarchy of cellular states akin to normal brain development, including proliferating and differentiated cells interacting with the stroma. We discovered that mesenchymal-like glioblastoma cells comprised a major glial-like wound-response component and a distinct gliosarcoma-specific malignant fibroblast type. Our analysis highlighted hypoxia, tissue damage, and wound healing as major factors in glioblastoma spatial organization. Tumor microenvironment varied along the hypoxia gradient, inducing the recruitment of monocytes and pro-tumorigenic macrophages, and propagating wound response program activation in malignant glial cells. Our study reveals the dynamic progression of glioblastoma organization independent of its mutational profile, in response to tumor-induced injury.

Correspondence to:
A.M.A (alejandro.mossi.albiach@ki.se *@MossiAlejandro) S.L.* (sten.linnarsson@ki.se / *@slinnarsson* / linnarssonlab.org)

Introduction

Glioblastoma (GBM) is the most malignant form of brain cancer with a median survival time of 10 to 15 months and limited treatment options based on radio- and chemotherapy following surgical resection[1]. This rapidly progressing type of glioma commonly bears a combination of genetic aberrations affecting the receptor tyrosine kinase (RTK, e.g., *EGFR*, *PDGFRA*, *PTEN*, *PIK3CA*, and *NF1*), p53 (*MDM2*, *MDM4*, and *TP53*), retinoblastoma protein (*CDK4*, *CDK6*, *CDKN2A/B*, and RB), and telomerase (*ATRX*, *TERT*) pathways[2]. Altogether, the combined presence of such mutations and wild type isocitrate dehydrogenase (IDH) 1/2 confers a particularly aggressive phenotype compared to other brain tumors like high-grade oligodendroglioma and astrocytoma[2,3].

In their classic 1925 monograph examining the histology of a systematic series of gliomas, Percival Bailey and Harvey Cushing showed that distinct types of glioma could be aligned with developmental cell states. In particular, they noted the morphological similarities between GBM cells and spongioblasts (now named radial glia, RG, or neural progenitor cells, NPCs)[4]. GBM cells can in turn be classified into diverse transcriptional subtypes, the variety of which exceeds what could be expected from the mutational profile alone. These subtypes vary across and even within tumor specimens, and several subtypes generally co-exist[5]. GBM cells arise in the adult brain parenchyma, proliferating rapidly to generate and traverse along this variety of cell states, sometimes resembling neurogenesis and gliogenesis[6,7]. Indeed, more recently, single-cell RNA sequencing (scRNA-seq) studies have revealed marked transcriptomic similarities between GBM cell types and those present in the healthy adult brain or foetal development. An influential study proposed four canonical and interconverting GBM cell states resembling astrocytes (AC-like), oligodendrocyte progenitor cells (OPC-like), neural progenitor cells (NPC-like), and an additional mesenchymal state (MES-like)[8]. Others have proposed a developmental hierarchy in GBM with radial glia-like cells at the apex[9].

Notably, while OPC-like, AC-like, and NPC-like cell states represent related waypoints along a branching neurodevelopmental trajectory, mesenchymal-like cell states are not observed in normal brain development. The mesenchymal designation in glioblastoma has been variously used to suggest either a cell state similar to mesenchymal cells (e.g., mesenchymal stem cells, fibroblasts), or an ongoing epithelial-to-mesenchymal transition (EMT). Mesenchymal cells are small spindle-shaped cells of the connective tissues, including fibroblasts. In most tissues, fibroblasts are responsible for wound healing, acutely by secreting tissue factor (F3) that clots plasma components leaking into extravascular space (i.e., haemostasis), and subsequently by secreting collagens. In the brain, where fibroblasts are only present in the meninges and along arterioles, acute wound healing instead depends on astrocytes assisted by microglia; only the late phase of collagen secretion is performed by fibroblasts.

Contrary to human development, where cell types arise in a controlled spatiotemporal order in response to morphogen gradients, and where progenitor cells (RG, glioblasts, and OPCs) give rise unidirectionally to neurons, astrocytes, and oligodendrocytes, the GBM dynamics seem far more plastic. In fact, the ability to transition across these cell states is thought to be key in glioblastoma's resistance to therapeutics[8]. However, there is only limited mechanistic insight into how GBM cells navigate the cell state space, or how transcriptional cell states are related to tissue organization and the tumor microenvironment.

The advent of high-throughput spatial transcriptomics methods has led to some important developments in this regard. Early studies using array-based spatial transcriptomics have revealed a changing spatial organization of tumor and immune cells driven by hypoxia[5,10], and hybridization-based in situ sequencing (HybISS) has helped characterize intercellular interactions of tumor and stromal cells[11,12]. The recurring patterned co-existence of several transcriptional subtypes suggests that tumor cells form a complex system retaining certain functional and organizational aspects of the healthy brain to survive and evolve together with the tumor microenvironment (TME). The dynamic processes by which GBM cells organize into such impervious structures remain to be understood.

Here, we sought to understand the relationship between cell states and spatial GBM organization by examining transcriptionally defined cell types and states in situ. In this context, the term heterogeneity has been used excessively as a broad descriptor of cellular diversity and architectural complexity. Effectively, heterogeneity encompasses variability across multiple levels of description, including genomic instability, tumor progression, TME composition, cellular plasticity, and therapeutic response. Because of these complex cross-scale drivers and the indefinite definitions of tumor cell states, a large patient cohort, large gene panel, and large tissue areas are all required to resolve systemic organizational trends. Therefore, we used Enhanced ELectric single molecule Fluorescence in situ Hybridization (EEL-FISH)[13] to characterize the expression of 888 RNA species in situ in 57 sections from 41 unique surgical samples of 27 patients: 25 GBM (*IDH*-wildtype) cases and two oligodendrogliomas (*IDH*-mutant and 1p19q co-deleted).

Results

For each tumor specimen, we collected one or more sections for EEL-FISH, and three adjacent sections on each side for 12 multiplexed histological assays, including hematoxylin and eosin (H&E), Ki67 (proliferation), TUNEL (apoptosis/necrosis), lectins (blood vessels), and CD45 (immune cells). We used cell nuclei to register histological images with those from EEL-FISH, enabling the transfer of histological features (e.g., pseudopalisading formation, necrosis, microvascular proliferation, abnormal vasculature) onto local transcriptional neighborhoods.

EEL-FISH provided high-level multiplexing over large tissue areas (~1cm^2/sample) at the expense of sensitivity, as compared to gold-standard smFISH[14]. High cellular density precluded accurate direct cell segmentation of the EEL-FISH data. Therefore, we instead used an unsupervised segmentation-free approach to learn cell-free transcriptional states. We trained a graph attention neural network (GAT)[15–17] to identify common molecular profiles across all samples by gene expression and spatial proximity (**Figure 1A/B**). The GAT output was clustered into 42 distinct molecular states (m-states) that distinguished microscale regions throughout the tissue, roughly corresponding to single-cell transcriptional states. We then segmented each m-state separately and removed m-cells with less than a total of 10 molecules, resulting in a total of 6,463,910 segmented single cells (m-cells) with a median of 32 detected mRNA molecules per m-cell, from a total of 11,081,393 raw imaged nuclei.

Inferred tumor cell states resemble human brain development

As an initial annotation, we systematically classified the m-states by using Tangram[18] to transfer cluster labels from a GBM scRNA-seq meta-analysis[19] (**Figure S1A**). Most m-states matched either the vascular system, the immune system, or the tumor compartment; the latter included canonical astrocyte-like (AC-like), neural progenitor cell like (NPC-like), oligodendrocyte progenitor cell like (OPC-like), and mesenchymal-like (MES-like) states, as originally defined by Neftel *et al*,[8]. However, label transfer from published GBM datasets precludes the discovery of any previously unnoticed cell states. Indeed, the neoplastic m-states displayed a larger complexity compared to the four canonical GBM cell states. For instance, several m-states corresponded to the AC-like and MES-like subtypes each (**Figure S1A-I**). Therefore, we sought to refine our classification of m-states by drawing parallels to human brain development. We took advantage of a recently published comprehensive scRNA-seq dataset covering the first trimester of human brain development[20] and used Tangram to align m-states to these reference cell states.

A total of 3,860,755 tumor m-cells were re-classified by similarity to developmental cell types (**Figure 1C**). Some of the m-states were accordingly re-classified as either glioblast (GBL-like), preOPC-like, or RG-like (**Figure S1A-II**). The GBL-like and preOPC-like m-states were characterized by the expression of *EDNRB* and *EGFR*[21,22] (**Figure 1D**). In fact, pre-OPCs expressed the highest levels of *EGFR* of all m-states and had molecular features similar to those of outer radial glia (*PTPRZ1*, *ITM2C*, *HOPX*, *FABP7*, and *EDNRB*) or late glioblasts[20], which have been reported to be highly expressed in GBM[16].

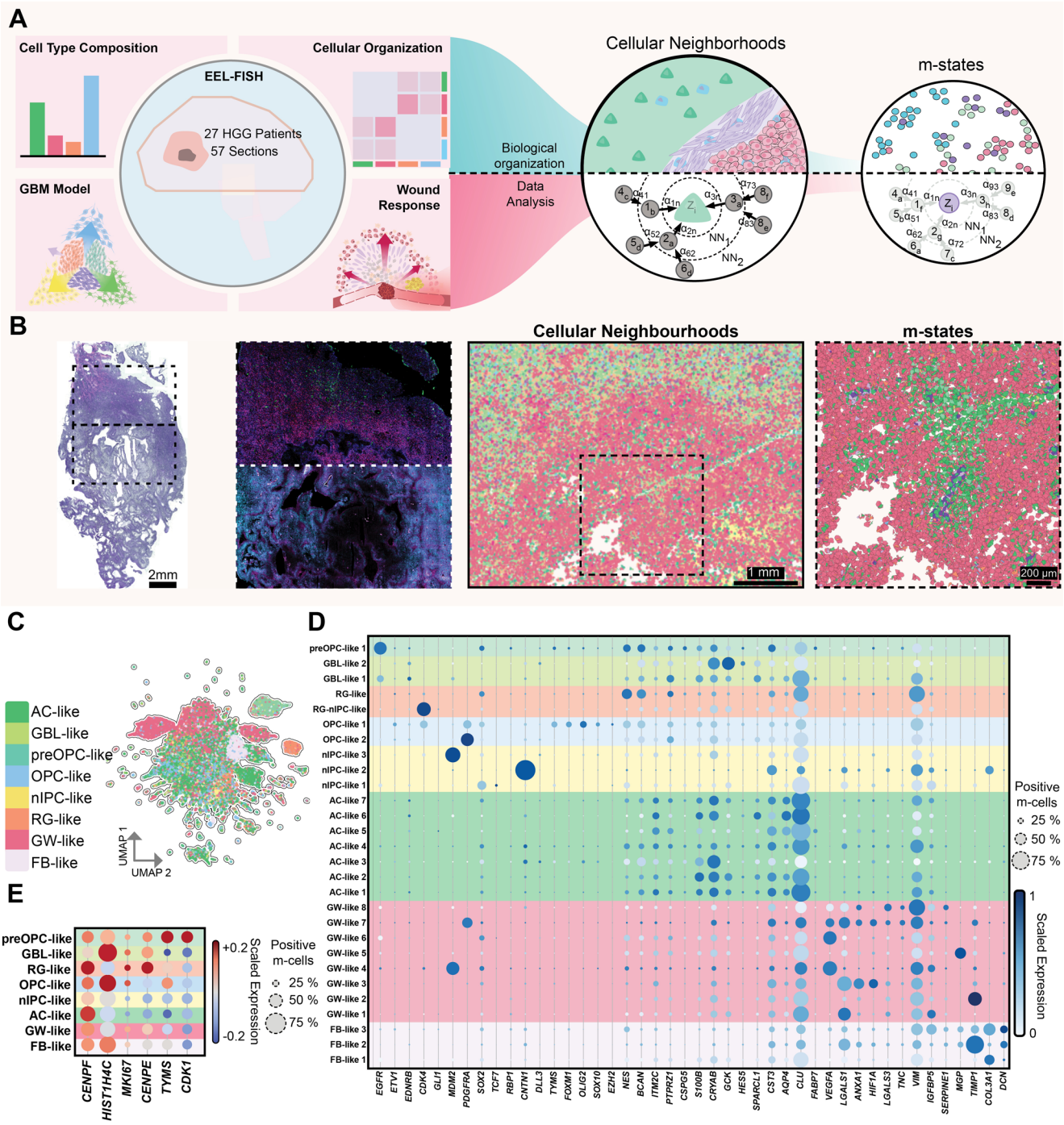


Figure 1. Data analysis schematic and results summary. (A) EEL-FISH on 57 tumor sections from 27 different high-grade glioma patients. Cell segmentation-free graph attention neural network applied on RNA-molecules for the inference of m-states (right-most panel). Graph attention network applied on the segmented m-states (m-cells) for the inference of macroscale cellular niches (middle panel). Schematic of results, cellular composition, cellular organization, wound response transition model and generalization of GBM model (left-most panel). **(B)** H&E staining on tumor sample SL035B (EEL-FISH adjacent section), proliferation (red) - endothelium (green) - nuclei (blue) staining, and cell membranes (magenta) - endothelium (yellow) -nuclei (blue) staining on two regions of SL005B sample (left-most panel). Cellular Neighborhoods spatial distribution in a region of tumor sample SL035B (middle panel). M-states spatial distribution around vasculature in a region of sample SL035B (right-most panel). **(C)** UMAP embedding of tumor m-states combined from all samples. **(D)** Dot plot showing the scaled marker gene expression for each tumor m-state. Dot size indicates the percentage of m-cells expressing each gene. **(E)** Dot plot showing normalized, log and scaled gene expression of cell cycle genes for each main class (preOPC/GBL/RG/OPC/nIPC/AC/GW/FB-like). Dot size indicates the percentage of m-cells expressing each gene.

States originally termed as NPC-like were defined by the expression of *SOX2*, *MDM2*, *CNTN1*, and/or *TCF7* (**Figure 1D**); these states mapped to developmental neural intermediate progenitor cells (nIPCs) (**Figure S1A-II**), and we accordingly termed them nIPC-like. In addition, we identified two subtypes by the expression of *SOX2*, *NES*, and *VIM* that mapped to radial glia (**Figure S1A-II**), and we termed these as RG and RG-nIPC, the latter subtype co-expressing *CDK4* and *GLI1* as well (**Figure 1D**). Expression of *CDK4* has been previously reported in NPC-like cells in GBM, and *GLI1* is expressed preferentially in RG cells in the neocortex[23]. RG-like, GBL-like, and pre-OPC-like cells all represented less differentiated states than the initially assigned canonical AC-like classification, in better agreement with Bailey and Cushing's spongioblast cell of origin.

More differentiated tumor cell states matched better to the previously described canonical GBM states. A higher expression of *CLU*, *AQP4*, *ITM2C*, *S100B*, and *CRYAB* defined a more mature AC-like state (**Figure 1D**). AC-like 3 subtype, originally classified as OPC-like/oligodendrocyte by Tangram, was re-classified as an AC-like m-state as it expressed high levels of *S100B* and *CRYAB* but not *PDGFRA*. Across the OPC-like populations, we identified two molecularly distinct subtypes. One highly proliferative state (OPC-like 1) expressing *TYMS* and *FOXM1*, two cell cycle related genes, higher levels of *NES* and *VIM*, indicating stem-like properties, and the *EZH2* histone remodelling enzyme; and a second more differentiated state (OPC-like 2) expressing higher levels of *PDGFRA* (**Figure 1D**). Thus, the canonical AC-like and OPC-like states were both represented in the tumors, but were distinct from the less differentiated nIPC-like, RG-like, GBL-like, and pre-OPC-like states.

Reevaluating the mesenchymal-like cell state in glioblastoma

We next applied a similar strategy to mesenchymal-like cells and aimed to elucidate their molecular complexity by searching for non-GBM analogues. The mesenchymal cell state in glioblastoma was originally described by Phillips *et al.* in 2006 based on bulk microarray gene expression profiles[24]. However, we found that a gene ontology analysis of that data revealed no terms related to mesenchymal cells or processes. Instead, many terms were related to neural development and to wound healing, blood coagulation, and apoptosis (**Table S1**). Phillips *et al*. also noted *CHI3L1* and *VIM* as mesenchymal cell markers. However, *CHI3L1* is expressed broadly beyond mesenchymal cells (highest in mammary glands)[25], and plays important roles in tissue remodelling, injury response, inflammation, and repair[25]; *VIM* is similarly broadly expressed, including in glia, and thus not a specific mesenchymal marker.

In 2010, Verhaak *et al*. refined the classification of glioblastoma tumors using a larger cohort of microarray profiles[6]. They defined one subtype as mesenchymal based on prior work and suggested an epithelial-to-mesenchymal transition linked to dedifferentiated and transdifferentiated tumors. Like Phillips, they pointed to the markers *CHI3L1* and *VIM*. Verhaak *et al*. performed gene ontology analysis (their Table S3) which revealed 123 statistically significant terms. However, none of the 191 known gene ontology terms related to mesenchymal states or processes were among the significant terms. Instead, among the most significant ontology terms were "immune response", "coagulation", "wound healing", "inflammatory response", and "haemostasis", among others. We repeated the analysis using updated ontology terms, with similar results (**Table S1**). Our reanalysis did uncover a significant overrepresentation of the term "epithelial-to-mesenchymal transition", with six out of 56 genes in the gene set. However, none of the transcription factors mechanistically linked to EMT[26] were in the set.

Neftel *et al.* subsequently used single-cell transcriptomics to define two mesenchymal-like glioblastoma cell states (MES1 and MES2)[8]. To justify naming the cells mesenchymal-like, they referred to previous work and performed a gene set enrichment analysis using MSigDb (their Table S3). For MES1, a single significant mesenchymal term was found, which however was the glioblastoma mesenchymal cell state defined by Verhaak *et al.*, a circular reference. The remaining reported MES1 terms were broad and unspecific, and they named this state "MES-like 1, hypoxia-independent". For MES2, nearly all the most significant terms were related to hypoxia, and none to mesenchymal cells or processes; they named this state "MES-like 2, hypoxia-dependent". To confirm and update these observations, we repeated the analysis using gene ontology instead of MsigDb (**Table S1**). For MES1, we found no terms related to mesenchymal states or processes; instead, common significant terms were related to complement activation, plasma membrane repair, wound healing, and coagulation. For MES2 we found terms related mainly to hypoxia, and none related to mesenchymal states or processes.

To further examine the possibility that glioblastoma contains mesenchymal-like cell states, we examined the expression of mesenchymal and fibroblast markers (*DCN*, *LUM*, *COL1A1*, *COL1A2*, *COL3A1*, and *COL4A1*, but not muscle-specific *COL2A1*), EMT transcription factors (*SNAI1*, *SNAI2*, *ZEB1*, *ZEB2*, and *TWIST1*), and the glial transcription factor *SOX2* in glioblastoma cell states and in normal human mesenchymal stem cells and fibroblasts[25] (**Figure S1B**). Whereas all tumor cell states maintained glial *SOX2* expression, neither fibroblasts nor mesenchymal stem cells expressed this gene. EMT transcription factors were inconsistently expressed and not specific to MES-like cell states. Mesenchymal and fibroblast marker genes were detected at low levels in AC-like and MES-like states.

However, as discussed in detail further below, we discovered true fibroblast-like cells in several of our tumors, particularly in gliosarcomas, which are glioblastomas that display a histological fibroblast component with spindle-shaped cells. Those fibroblast-like cells were molecularly distinct from previously defined mesenchymal-like glioblastoma cell states and expressed bona fide markers of mesenchymal cells such as collagens type I (*COL1A1*, *COL1A2*) and type III (*COL3A1*), *DCN*, and *LUM*. Re-examination of the Neftel *et al.* dataset showed that a single patient tumor, paediatric glioblastoma BT830, appeared to contain abundant fibroblast-like cells expressing both *COL1A1*, *COL1A2*, *COL3A1*, *DCN*, and *LUM*, but not *COL2A1*, similar to the fibroblasts and mesenchymal stem cells

from Tabula sapiens[25], while these genes were absent, very low, or inconsistent in the remaining 27 tumor specimens (BT920 expressed a subset of the genes; Figure S1B)8. About half of the cells in this tumor were fibroblast-like, and the remainder glial, consistent with the typical histology of gliosarcoma. One tumor out of 27 is consistent with the known ~2% frequency of gliosarcomas among glioblastomas[27]. Thus, both in our dataset (below) and in Neftel *et al.*, there are true mesenchymal cells in the form of fibroblast-like cells, but they are mostly present in the relatively rare gliosarcomas and are distinct from previously defined mesenchymal-like tumor cells. Interestingly, as shown in **Figure S1B**, these cells further expressed markers of neural crest stem cells (*SOX10* and *NGFR*) and developing meningeal fibroblasts (*FOXC1*), suggesting a differentiation process analogous to the emergence of the meninges from radial glia-derived neural crest in the early embryo.

If 'mesenchymal-like' tumor cells are not mesenchymal, what are they? Important clues, including the gene ontology analyses reported above, point to hypoxia, inflammation, injury, wound healing, blood clotting, and similar processes; that is, to a reactive state rather than a cellular differentiation process. To resolve the question, we again reanalysed Neftel *et al.*, while removing tumors with suspected gliosarcoma fibroblast involvement (**Figure S1C** and **Table S1**). MES2 was associated with hypoxia as before, and indeed a large fraction of MES2-specific genes were known direct targets of HIF1α, the master regulator of the hypoxic response (23 of the top 200 genes). In contrast, MES1 was associated with complement activation, plasma membrane repair, plasminogen activation, and coagulation, all of which are related to haemostasis, the extravascular clotting of extravasated blood or plasma. We further examined a recent integrated analysis of 12 published glioblastoma datasets, which had identified 27 core MES-like genes[28], most of which have known functions related to wound healing (**Table S1**). Notably, every single MES-core gene was robustly and significantly (FDR < 0.05, fold-change > 1.2) induced in a well-powered timecourse study of skin wound healing[29] (**Figure S1D**). With the exception of type I collagens (*COL1A1*, *COL1A2*), all MES-core genes were induced in the acute phase (within 24 hours); collagens were induced after 72 hours. We conclude that previously named mesenchymal-like cell states in glioblastoma more likely represent fundamentally glial (*SOX2*$^+$) tumor cells but with strong activation of hypoxia and wound healing responses analogous to those found in normal astrocytes during brain wound repair; we will henceforth refer to them as 'glial-like wound response' (GW-like) cell states. Based on the enriched gene sets and our gene ontology analyses, it is likely that additional related processes such as inflammation, apoptosis, and metabolic rewiring also characterize these cellular states.

This conceptual reinterpretation is corroborated by our observation that GW-like m-states (corresponding to Neftel *et al.* MES-like states) showed a very distinctive molecular profile, preferentially expressing several of the following genes: *VEGFA*, *LGALS1*, *SERPINE1*, *IGFBP5*, *ANXA1*, *MGP*, and *TIMP1* (**Figure 1D**). We did not observe expression of true mesenchymal-specific genes such as type I or III collagens, indicating that these cells were not of true mesenchymal identity. Instead, *VEGFA*, *SERPINE1*, and *IGFBP5* are hallmark genes induced by hypoxia[30]; *ANXA1*, *VEGFA*, *SERPINE1*, *LGALS1*, and *VIM* are direct targets of the oxygen-sensing transcription factor HIF1α; *MGP* controls angiogenesis in mice[32] and in bovine cells[33]; and *TIMP1* is strongly induced by hypoxia in infant brain[34]. In addition, GW-like-specific *SERPINE1* is the plasminogen activator inhibitor (PAI-1), and *PLAU* is the urokinase plasminogen activator, both involved in the haemostatic response during wound healing. Furthermore, the diversity of GW-like states reflected the transcriptional diversity of other GBM cells. For instance, GW-like 4 expressed *MDM2*, suggesting a transition from nIPC-like, while GW-like 7 expressed *PDGFRA*, suggesting a transition from OPC-like state. These observations suggest that GW-like cells are tumor cells that react to hypoxia and plasma leakage from disrupted blood vessels, while retaining their underlying glial cell identity.

Finally, in addition to the conventional neoplastic developmental-like, and GW-like states, we also found three fibroblast-like (FB-like) m-states expressing *COL3A1* (**Figure 1D**), of which FB-like 2 and 3 co-expressed *COL3A1* and *DCN* likely representing true metaplasia (and not EMT), similar to BT830 from Neftel *et al.*[8] Such cells were found primarily, but not exclusively, in glioblastoma of subtype gliosarcoma, and will be discussed further below.

Tumor composition by developmental states and wound response

Next, we assessed GBM cellular composition in respect to the major m-states. Given that several cellular states may function as progenitors in GBM, proliferating and differentiating into distinct lineages[36], varied tumor compositions could be expected. Based on the expression of cell cycle genes, preOPC-like, RG, GBL-like, and OPC-like m-states were the most proliferative, whilst GW-like, nIPC-like, AC-like, and FB-like cells appeared to be less proliferative (**Figure 1E**). These results suggest a distinct tumor hierarchy where the most differentiated states are the least proliferative and the most proliferative are the least differentiated. However, regardless of the dominant cellular state, most tumors were characterized by a combination of both glial and wound response m-states, and only few samples (notably including both oligodendrogliomas) showed a very low percentage (<5%) of wound response m-states (**Figure 2A** and **Figure S1**).

Oligodendrocytes, vascular, and myeloid cells constituted most of the non-neoplastic compartment. The myeloid component was composed mainly of tumor associated macrophages/microglia, whereas lymphoid cells (NK cells, B/T cells, and plasma B cells) only amounted to a small proportion (<5%). While EEL-FISH sensitivity did not allow us to confidently map individual NK/B/T cells, we could nevertheless estimate their relative proportions across the different samples (**Figure 2A**).

Spatial relationships are conserved across GBM patients independently of genetic drivers

Each tumor specimen was characterized according to a series of clinical and molecular parameters (**Table SI2**). Genetic aberrations, in particular, can be one of the most critical tumor descriptors. Indeed, GBM

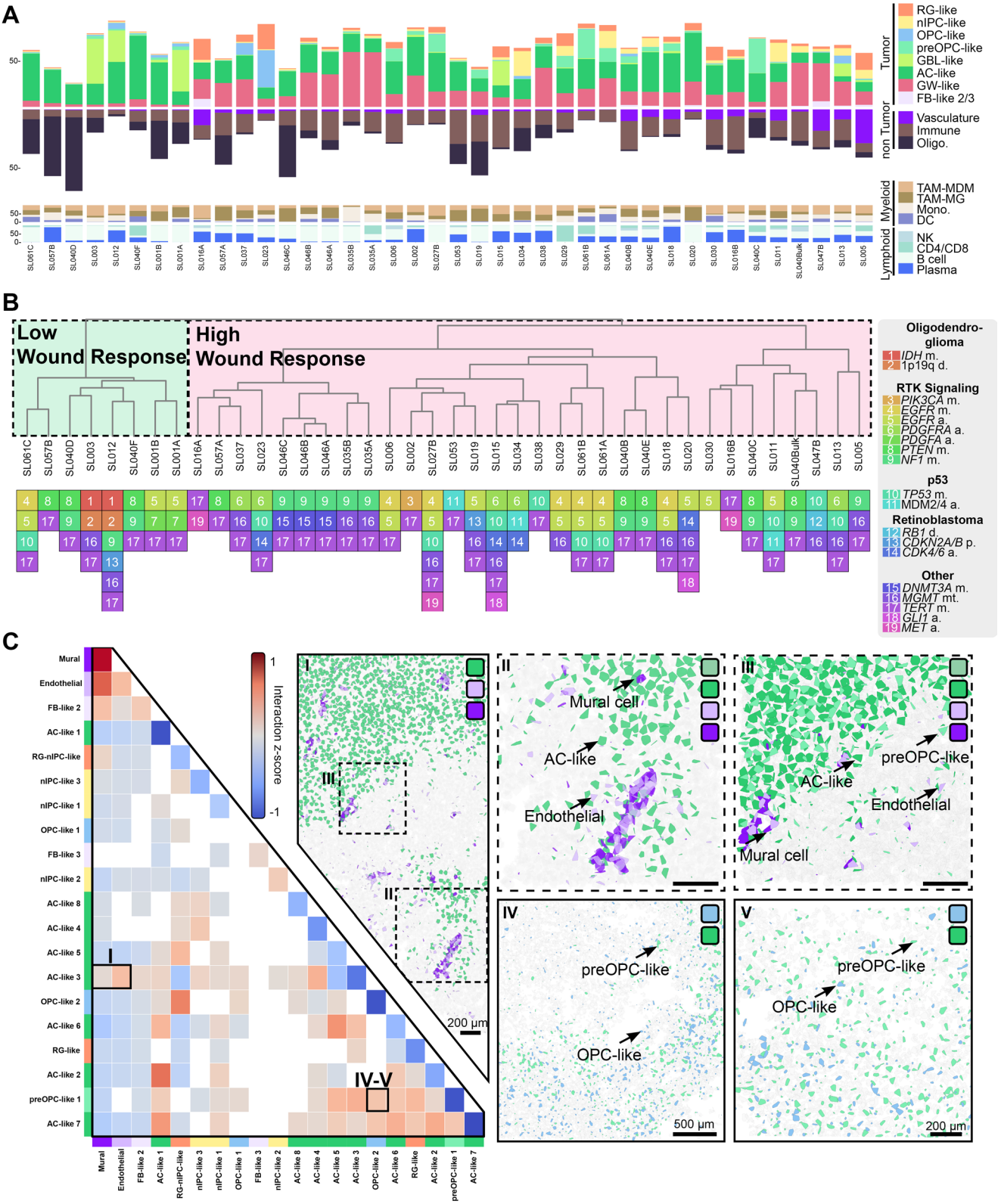


Figure 2. Molecular characterization of m-states and spatial architecture of HGG m-cells. (A) Percentage of tumor m-states (RG-like, nIPC-like, OPC-like, AC-like, FB-like 2/3, and GW-like) vs non-tumor states (myeloid and lymphoid cells, vasculature, and oligodendrocyte) per sample. Bar plots of the Tangram probabilities (percentage) for each scRNA-seq reference myeloid and lymphoid cell type as the sum of Tangram probabilities for all sample m-states. Percentage of myeloid cells per sample (TAM-MDM, TAM-MG, DCs, and monocytes). Percentage of

lymphoid cells per sample (NK, T, B, NK, and plasma B cells). **(B)** Dendrogram depicting the m-state spatial organization similarity across samples together with the mutational profile of each patient. Each square indicates mutations identified in the pathology report (**Table SI2**). Mutated (m.), amplified (a.), deleted (d.), positive signal (p.), methylated (mt.). **(C)** Neighborhood enrichment scores for the most frequent non-GW m-states and vasculature for all samples. I/II: Spatial distribution of AC-like 3 as well as endothelial and mural cells in situ in sample SL035B. III: Spatial distribution of AC-like and preOPC-like m-cells in situ in sample SL035B. IV/V: Spatial distribution of preOPC-like and OPC-like cells in situ in samples SL001A and SL001B. Gray polygons represent all other m-states.

m-state composition varied by the mutational spectrum (**Figure S1A**). Samples with *MDM2*, *CDK4*, and *PDGFRA* amplification showed higher proportions of nIPC-like and OPC/RG-like states, respectively, whilst samples with EGFR mutation and amplification were enriched in preOPC-like state. We sought to explore whether GBM exhibits a mutation-independent structural architecture. For this purpose, we quantified the overrepresentation of nearest neighbors for each of the major m-states present in most samples and used these m-state spatial proximity scores to build a sample-to-sample similarity matrix (**Figure 2B**). Generally, samples with matching mutational features displayed similar spatial organization. Oligodendroglioma samples and most of the GBM samples from the same patient were found closely together in the dendrogram: SL003 and SL012 (*IDH* mutated and 1p19q co-deleted), SL001A/B (*EGFR* and *PDGFA* amplified), SL046A/B/C (*NF1* and *DNMT3A* mutated), SL035A/B (*MGMT* methylated, *NF1* mutated). We observed a surprisingly large degree of spatial conservation given the expected GBM cell state variability for such large areas (>1cm^2) and samples from the same tumor separated by >1 cm.

Conversely, the most pronounced source of sample variability was the presence of the GW-like state, separating the glioma samples into two clades: with high and low GW-like proportions respectively (**Figure 2B**). In fact, such differences in GW-like state were observed even between samples originating from the same patient, placing the samples under different clades: SL040C/D and B/E, and SL061A/B and C. Thus, the abundance of GW-like cells, and the overall classification of a tumor, depended critically on the sampling location.

Notably, we observed consistent patterns of cell type spatial neighborhood organization regardless of the presence of GW-like m-state (**Figure 2C**). Specifically, AC-like 3 (*CRYAB*$^+$) was recurrently found near endothelial and mural cells (**Figure 2C-I/II/III**). We observed this functional association across tumors independently of the GW-like content (**Figure S2A**) or mutational status: oligodendroglioma (**Figure S2B**), *CDK4*/*PDGFRA*-amplified (**Figure S2C**), and *MDM2*-amplified (Figure S2D). These results suggest a conserved mechanism by which tumor cells maintain the association with vasculature, normally seen in healthy mature astrocytes[37], and suggest that the presence of GW-like hypoxic and wound healing responses does not fundamentally alter the organization of tumor cellular states.

Similarly, we observed preOPC-like 1 m-cells consistently neighboring OPC-like 2 and AC-like m-states (**Figure 2C-IV/V**), strengthening our classification of the *EGFR*$^+$ m-state as developmental preOPC-like, not only by the molecular profile, but also by the potential to generate neighboring OPC-like cells. The spatial proximity to AC-like state might also be indicative of an analogous potential of preOPC-like 1 cells to generate AC-like cells in GBM.

Cellular neighborhoods define glioblastoma niches

To better understand and represent conserved spatial relationships, we implemented a second GAT network connecting m-cells to their neighbors and thus learning neighborhood patterns which we termed 'cellular neighborhoods' (CNs). When training such a GAT, the number of CNs must be fixed in advance, which provides a tradeoff between fine resolution (many CNs) and interpretability (fewer, coarser CNs). We settled on ten CNs reflective of broad, common patterns of cellular niches. Generally, we observed high variability in the relative proportions of such CNs across the different tumor samples (**Figure S3A** and **Figure S4**). However, the constituent cell types showed a consistent and recurrent organization as revealed by their stable relative proportions within each CN. Tumors were composed mostly of varying proportions of CN2, CN3, CN4, CN5, and CN10. CN2 and CN3 revealed AC/preOPC-like m-states with and without endothelial association, respectively (CN2 - AC/preOPC/Endo, CN3 - AC/preOPC) (**Figure S3B**). CN4 captured niches with the highest amount of tumor-associated microglia (CN4 – TAM-MG), whereas CN5 captured niches with mostly monocyte-derived TAMs (CN5-TAM-MDM) (Figure S3B). CN10 captured the association between GW-like m-cells, vasculature (endothelial and mural cells), and infiltrating monocytes (CN10 - Wound) (**Figure S3B**). The remaining CNs revealed other less frequent cellular niches: mature oligodendrocytes and mature AC-like m-cells in the periphery (CN6) (**Figure S3B/C**), erythrocytes and *CNTN1*$^+$ nIPC-like 1 signature (CN1), OPC and RG-like cells (CN7), nIPC-like and GW-like cells (CN8), and GBL and FB-like cells (CN9).

Given the prior observation that samples from the same patient can differ substantially in m-state composition, we explored the differences in CN composition across multiple tumor regions of the same patient annotated according to their relative 5-aminolevulinic acid (5-ALA)-based fluorescence intensity at surgery: necrotic core, inner zone (high fluorescence), transitional zone (low fluorescence), and periphery (no fluorescence) (**Figure 3A**). The changes in CN composition revealed the variable cell niches from the tumor periphery to the tumor core. The necrotic core, captured in sample

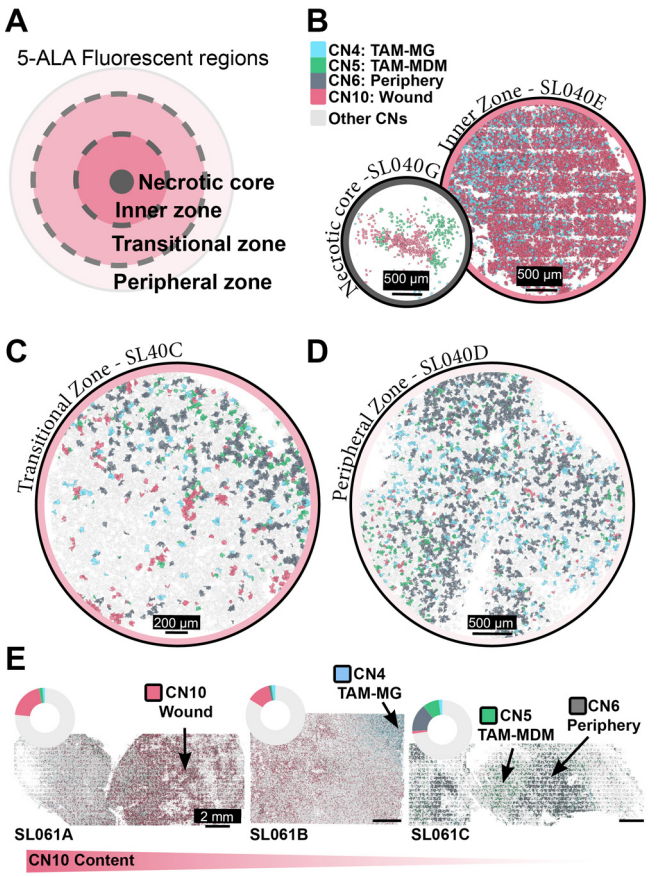


Figure 3. Core-to-periphery changes in tumor microenvironment defined by Cellular Neighborhood analysis. (A) Schematic of surgical 5-ALA contrast enhancing tumor region annotation: necrotic core and inner (high fluorescence), transitional (low fluorescence), and peripheral (no fluorescence) zones. Spatial distribution of CN4 (TAM-MG), CN5 (TAM-MDM), CN6 (Periphery) and CN10 (Wound) in the necrotic and inner zone **(B)**, in the transitional zone **(C)**, and in the periphery **(D)**. Color palette specific for CN. **(E)** Tumor architecture reconstruction in a multi-sample tumor, SL061, guided by CN composition of the three constituent samples. Pie charts show the CN composition in each sample: CN4 (TAM-MG), CN5 (TAM-MDM), CN6 (Periphery), CN10 (Wound), and others.

SL040G, was largely composed of CN10 (Wound) and CN5 (TAM-MDM), the former consisting of GW-like cells, vasculature and monocytes, and the latter of monocyte-derived TAMs (TAM-MDM; **Figure 3B**). Inner zone tumor regions, captured in sample SL040E, were primarily marked by CN10 (**Figure 3B**). The presence of CN4 (TAM-MG), consisting of microglia, and CN6 (periphery), consisting of oligodendrocytes and mature AC-like cells, was indicative of the transition from the inner tumor zone towards peripheral areas (**Figure 3C-D**, **Figure S3C**). These spatial axes captured the variability of cell type composition and interactions in another core-to-periphery spanning sample (**Figure 3E**), but a large sample size was required to resolve such transitions occurring from the tumor core to the peripheral areas.

Monocytes infiltrate hypoxic regions and associate with GW-like cells

CN analysis uncovered a varying tumor microenvironment along the core-periphery axis, with the interplay between TAMs, monocytes, and GBM states being a key characteristic. To further explore and validate the molecular diversity of our captured myeloid m-states (**Figure 4A**), we performed 28-color intracellular flow cytometry on 11 high-grade gliomas to match our RNA-based immune m-states (**Figure 4B**) with protein-level data (**Figure 4C** and **Figure S5**). We had identified two transcriptional monocyte m-states by low or high levels of *TIMP1*, termed Mono 1 and 2, respectively (**Figure 4B**). Of these, Mono 1 matched the protein expression of classical monocytes as identified by flow cytometry: CCR2, S100A9, and CD11b (*ITGAM*). Microglia (TAM-MG), mostly found in the tumor periphery, were identified by co-expression of *P2RY12* and *CX3CR1*, matching protein phenotype of microglia ($CD11b^{+}CD68^{+}CCR2^{-}CD49^{-}P2RY12^{+}CX3CR1^{+}$)[38,39] (**Figure 4C** and **Figure S5**). We identified four m-states of monocyte-derived macrophages (TAM-MDM 1-4). TAM-MDM 1 and 2 expressed higher levels of interferon response gene *TNFSF10* and *IDO2*, whereas TAM-MDM3 expressed higher levels of *CD14*, suggesting highest similarity to classical monocytes (Mono 1). TAM-MDM 4 expressed higher levels of hypoxia-induced genes *SPP1*[40] and *MT1H*[41]. Monocyte populations, TAM-MDM 3-4, and TAM-MG expressed *OSM*, which has been reported to mediate MES switch through its receptor *IL6ST*[42], which is expressed by reactive astrocytes[43]. TAM-MG and TAM-MDM 3 also expressed higher levels of *TREM2* and *CD274* (encoding the immune checkpoint ligand PD-L1), suggesting a pro-tumor role of these monocyte-derived macrophages[44]. We also validated this finding on protein level: the highest CD14-expressing macrophages showed significantly higher levels of immune checkpoint inhibitor PD-L1 (**Figure 4C** and **Figure S5**).

Next, we examined the spatial organization of these myeloid populations in relation to tumor cell states. TAM-MG m-cells were closer to peripheral and proliferative cellular states (preOPC-like 1, OPC-like 1) (**Figure 4D**) found in regions with high vascular proliferation. TAM-MDM 3 and TAM-MDM 4 were both spatially associated with Mono 1, suggesting a possibility of direct differentiation from Mono 1, in turn found near endothelial and mural cells (**Figure 4D**). TAM-MDM 4 were specifically associated with GW-like cellular states. Hence, we hypothesized that TAM-MDM 3 and TAM-MDM 4 differentiate from Mono 1, whereupon TAM-MDM 3 would acquire a pro-tumor phenotype (*TREM2* and *CD274*), whilst TAM-MDM 4 would acquire hypoxic characteristics as it was primarily associated with GW-like subtypes. We then aimed to contextualize these findings in respect to the spatiotemporal progression of interactions between immune and tumor cells. To this end, we explored a selection of tumors with varying proportions of GW-like cells. Firstly, we found areas with high density of

Mono 1 to be associated with areas of low GW-like m-cell density, and *HMOX1* hypoxia signature was highly associated with the presence of monocytes (**Figure 4E**). The same trend was observed in tumors with a higher GW-like content: monocyte-rich areas were surrounded by GW-like signature (**Figure 4F-I.a/c**) and GW-like m-cells were associated with TAM-MDM 3 and TAM-MDM 4 (**Figure 4G**). Among monocyte-rich hypoxic areas, we located some cells that highly expressed interferon-related genes *IFIT2*, *IFIT3*, *ISG15*, and *IL1B* (**Figure 4F-II/III**), similar to previously reported gene expression signature of CD16- monocytes[45]. Our data suggest a temporal progression wherein monocytes are recruited to hypoxic areas (*HMOX1*$^{+}$), in turn inducing a switch towards wound response in tumor cells. Monocytes are sequestered among GW-like cells and differentiate into immunosuppressive macrophages (*CD163*$^{+}$*TREM2*$^{+}$*CD274*$^{+}$) with hypoxia signature genes (*MT1H*$^{+}$*MT1X*$^{+}$*SPP1*$^{+}$) (**Figure 4G**). Our observations are in line with recent reports showing that hypoxic niches attract and sequester TAMs, ultimately leading to pseudopalisade formation[46].

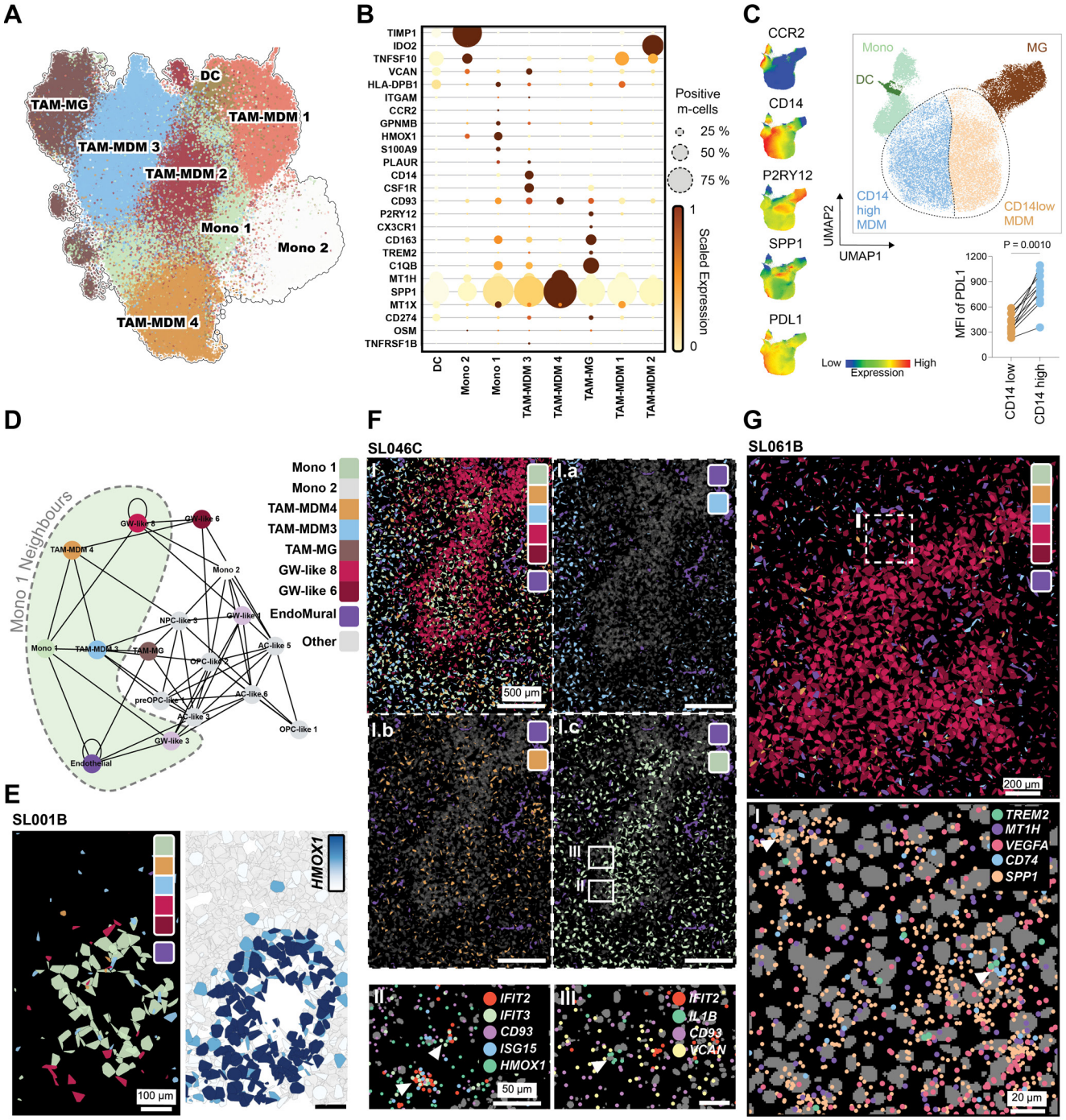


Figure 4. Molecular characterization and spatial distribution of myeloid m-cells in GBM. **(A)** UMAP embedding of myeloid m-states. Specific color palette for immune cells is maintained throughout panels B, D and E. **(B)** Dot plot showing the scaled marker RNA expression for each myeloid m-state. Dot size indicates the percentage of m-cells expressing each gene. **(C)** UMAP embedding of myeloid cells by protein expression from flow cytometry (top-right panel). Protein expression of CCR2, CD14, P2RY12, SPP1 and PDL1 across myeloid cells (left panel). Comparison of indicated marker mean fluorescence intensity (MFI) between CD14high and CD14low MDM in individual tumors; The Wilcoxon matched-pairs test was used to assess statistical significance; p value indicated. **(D)** Graph representation of neighborhood enrichment of most frequent m-states in the periphery (AC-like 3/5/6, OPC-like 1/2, preOPC-like 1, nIPC-like 3, GW-like 6/8, monocytes, TAM-MDM 3/4, and TAM-MG). Edges represent m-state pairs with z-score above 8.5. Non-GW m-states are colored gray. **(E)**

Spatial distribution of Mono 1, TAM-MDM 3/4, GW-like 6/8, and endothelial m-cells in a region of sample SL001B (low GW-like frequency). Spatial distribution of Mono 1 and TAM-MDM 3/4 with GW-like m-state. *HMOX1* expression in the same region. Gray polygons represent m-cells with zero expression. **(F)** I: Spatial distribution of Mono 1, TAM-MDM 3/4, GW-like 6/8, and endothelial m-cells in a region of a sample SL046C (high GW-like frequency). Spatial distribution of Mono 1, TAM-MDM 3/4 with GW-like m-state. I.a: Spatial distribution of endothelial m-cells and TAM-MDM 3. I.b: Spatial distribution of endothelial m-cells and TAM-MDM 4. I.c: Spatial distribution of Mono 1 and endothelial m-cells. II: EEL-FISH signal of *IFIT2*, *IFIT3*, *CD93*, *ISG15*, and *HMOX1* in region II. Gray polygons represent segmented nuclei masks. III: EEL-FISH signal of *IFIT2*, *IL1B*, *CD93*, and *VCAN* in region III. Gray polygons represent segmented nuclei masks. **(G)** Spatial distribution of Mono 1, TAM-MDM 3/4, GW-like 6/8, and endothelial m-cells in a region of a sample SL061B (high GW-like frequency). Spatial distribution of Mono 1, TAM-MDM 3/4 with GW-like m-state. I: EEL-FISH signal of *TREM2*, *M1H*, *VEGFA*, *CD74*, and *SPP1* in a region with high monocyte density. Gray polygons represent segmented nuclei masks.

Changes in TME reveal the sequence of events for GW-like transition

To understand the events leading to a GW-like state induction, we took a closer look at the cellular organization of four GBM samples that contained variable percentages of such cells: SL001A, SL001B, SL006, and SL035B; from lowest to highest (**Figure 5**). Samples with a low GW-like percentage instead contained alternating percentages of AC-like, GBL-like, preOPC-like, nIPC-like, and/or OPC-like cell states. GBL-like and preOPC-like m-states were the predominant states in the GW-like low samples (e.g. SL001A/B), whereas GW-like areas were few and sparsely distributed (**Figure 5A**). We could locate a few regions within SL001B marked by CN10 (**Figure 5A-I**), which we had previously annotated as wound-like: GW-like, vascular, and monocyte recruiting. These regions exhibited higher expression of *VEGFA* (GW-like signature) and were surrounded by a higher density of TAM-MDM 3/4 and monocytes. We therefore postulated that these regions rich in classical or proneural states (AC, preOPC, OPC, nIPC, and RG) with few GW-like m-cells are undergoing a transition towards wound response phenotype, and that monocytes and TAM-MDMs, specifically highly expressing *SPP1* and *IL1B*, have an active role in this process (**Figure 5A-II.a**).

To substantiate this claim, we next examined the samples with higher GW-like content: SL006 and SL035B, the latter a recurrent GBM case with particularly extensive GW-like regions (**Figure 5B-C**). Samples with high GW-like content were also characterized by extensive necrosis and/or pseudopalisading histology (**Figure S6A/B/C**). Most notably, we observed a marked split between two compositional TME regimes in these samples. On the one hand, peripheral areas characterized by CN4 (peripheral TAMs) were generally accompanied by a high degree of vascular proliferation and cellular states with stem-like properties and high proliferative potential (preOPC-like, OPC-like 1, and RG-like; **Figure 5B-I**, **Figure 5B-I.a/c**), we identified as the leading edge. On the other hand, high GW-like regions seemed to displace other cellular states and were found closer to each other than to other tumor states (**Figure 4D**, **Figure 5B-II** and **Figure 5C-II.a**). Among the GW-like enriched areas in SL006 and SL035B, we observed a higher representation of TAM-MDM with higher expression of *SPP1* and *IL1B* compared to the leading edge (**Figure 5B-I.a/b**), where we instead found TAM-MG (CN4) and proliferating tumor cells.

In sample SL035B, we found perivascular niches with a wide diversity of non-GW-like states immediately surrounding vasculature: preOPC/RG/AC/OPC/RG/nIPC-like (**Figure 5C-II.a**), embedded into a greater wound response niche. Again, AC-like (*CRYAB*$^+$) m-states were found closest to endothelial m-cells (**Figure 5C-II.b**), whilst GW-like m-states were tightly aggregated and situated farthest from vasculature (**Figure 5-II.c**). It should be noted that the spatial organization of isolated perivascular niches nestled within GW-like regions (**Figure 5C-II**) highly resembled the organization of our previously described non-GW-like areas (**Figure 5C-I**) in the periphery. We further hypothesized that these niches (leading edge, wound, and perivascular) reflected the temporal sequence of GBM growth events (**Figure 5D**). The peripheral areas contained the most proliferative and undifferentiated m-states, serving as the leading edge of tumor expansion. Proliferating tumor cells would transition towards a quiescent and wound response phenotype associated with TAMs[42,47] and necrosis, whilst tumor cells in the perivascular intact niche maintained their glial identity.

The architecture of GW-like states is conserved across tumors

Next, we set out to characterize the spatial and molecular properties of GW-like subtypes to elucidate the intrinsic or extrinsic mechanisms that drive the wound response. We defined eight GW-like subtypes, all expressing the hypoxia-inducible factor *HIF1A* (master regulator of the hypoxia response), as well as several direct targets of *HIF1A* including *VEGFA*, *VIM*, and *LGALS1*. The diversity of GW-like states was characterized primarily by increased expression of *VEGFA*, and secondarily by varied expression of other marker genes: *EPAS1*, *MDM2*, *PDGFR*A, *MGP*, *TNC*, and *TIMP1* (**Figure 6A**). GW-like 4 (*MDM2*$^+$) and GW-like 7 (*PDGFRA*$^+$) likely corresponded to GW transitions from nIPC-like and OPC-like initial states, as these GW-like subtypes were found in samples with high abundance of the proposed initial states (**Figure S7A**). Thus, while all GW-like states showed response to hypoxia, GW-like 6/4/8/7 were the strongest responders with the highest expression of *VEGFA*. These states might correspond to Neftel *et al.* MES2 "hypoxia-dependent" states, while our GW-like 1/2/3/5 might correspond to Neftel *et al.* MES1 "hypoxia-independent" states[8]. Notably, the latter were defined by expression of the *HIF1A* target genes *LGALS1* and *VIM*, and thus likely still responding to hypoxia but at lower levels.

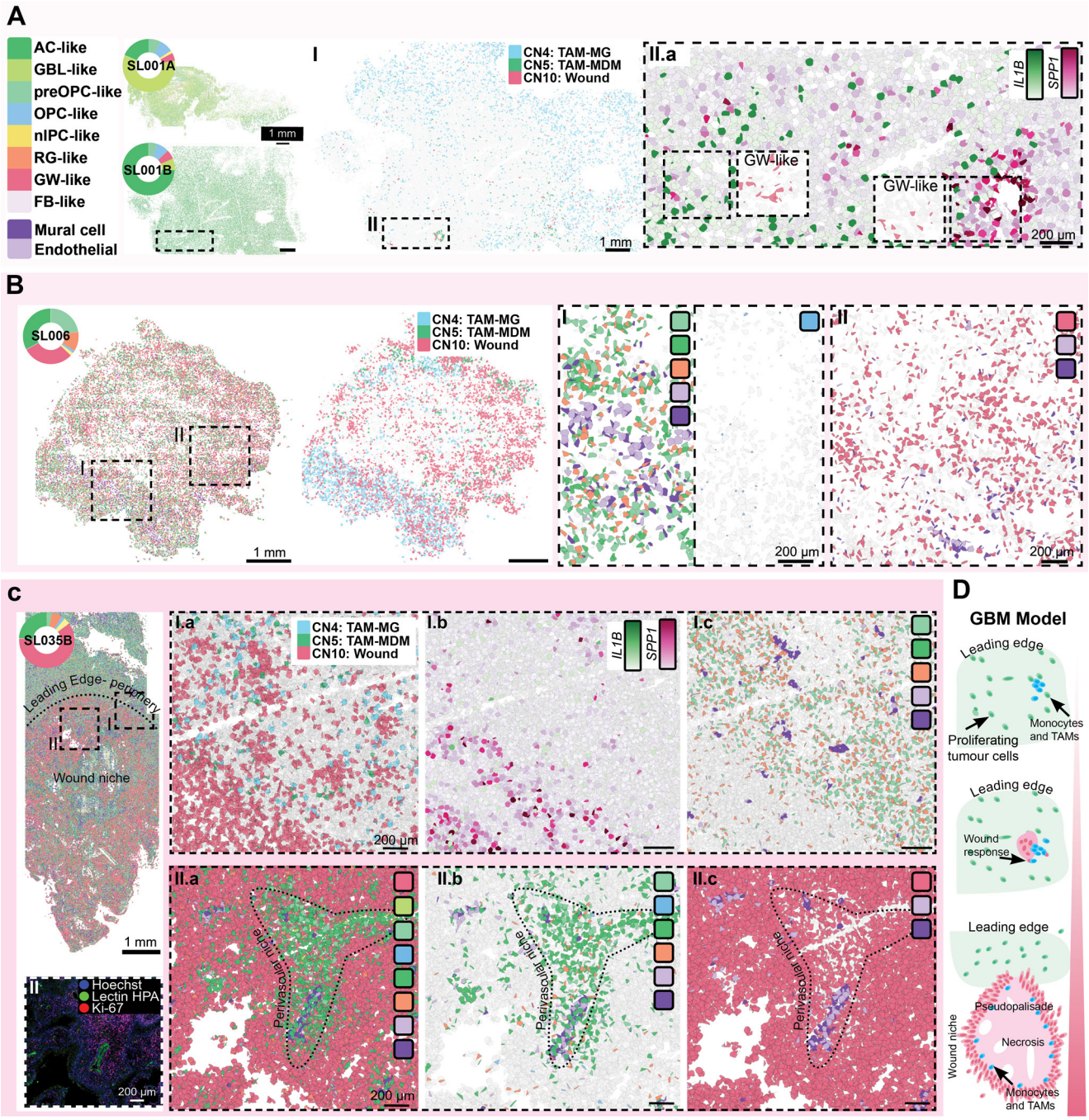


Figure 5. Spatial distribution of tumor m-states across tumors with variable percentage of GW-like cells. (A) Spatial distribution of tumor m-states in samples SL001A and SL001B; pie charts show the proportions of constituent m-states. II: Spatial distribution of CN4 (TAM-MG), CN5 (TAM-MDM), and CN10 (Wound) in sample SL001B. IIIa: Expression of *SPP1* and *IL1B* together with GW-like m-cells in region III of SL001B. Gray polygons represent m-cells with zero expression. **(B)** Spatial distribution of tumor m-states in sample SL006; pie chart shows the proportions of constituent m-states. I: Spatial distribution of CN4 (TAM-MG), CN5 (TAM-MDM), and CN10 (Wound) in sample SL006. II: Spatial distribution of preOPC/RG/AC-like, endothelial, and mural m-cells (left panel). Spatial distribution of OPC-like subtype (right panel). III: Spatial distribution of GW-like, endothelial, and mural m-cells. Gray polygons represent all other m-states. **(C)** Spatial distribution of tumor m-states in sample SL035B; pie chart shows the proportions of constituent m-states. I.a: Spatial distribution of CN4 (TAM-MG), CN5 (TAM-MDM), and CN10 (Wound) in region I of sample SL035B. Gray polygons represent all other CNs. I.b: Expression of *SPP1* and *IL1B* in region I of sample SL035B. Gray polygons represent m-cells with zero expression. I.c: Spatial distribution of preOPC/RG/AC-like, endothelial, and mural m-cells in region I of sample SL035B. Gray polygons represent all other m-states. II: Ki67 (proliferation), lectin (blood vessel), and Hoechst (nuclei) staining in region II of sample SL035B. II.a: Spatial distribution of GW-like and non-GW m-cells (preOPC/GBL/OPC/RG/AC-like) around endothelial and mural m-cells in region II of SL035B. II.b: Spatial distribution of non-GW m-cells (preOPC/OPC/RG/AC-like) around endothelial and mural m-cells. II.c: Spatial distribution of GW-like m-states around endothelial and mural m-cells. Gray polygons represent all other m-states.

Examining the spatial organization of our GW-like subtypes, we found robustly detectable patterns across different tumors (**Figure 6B** and **Figure 6C**). GW-like 5 (*MGP*$^{+}$) was found closest to endothelial and mural cells (**Figure 6B** and **Figure 6D-I.a**). Other non-hypoxic GW-like states displayed a loose spatial organization and were sparsely dispersed throughout the tumors (**Figure 6D-II**). In contrast, hypoxic states were organized in a layered serpentine pattern with GW-like 6 (highest *VEGFA*) residing farthest away from vasculature (**Figure 6D-I.b**) and other hypoxic GW-like states constituting parallel layers spanning until the vasculature where non-hypoxic GW-like and non-GW states could be found (**Figure 6B**, **Figure 6D-I.b/III.a**, and **Figure 5C-**

II). We thus observed a detailed architecture of GW-like GBM cell states, wherein hypoxia was the defining axis of organization.

Thereafter, we aimed to elucidate the dynamics leading to this spatial patterning of GW-like subtypes. The level of *VEGFA* expression varied across GW subtypes, with the highest expression in the most hypoxic subtypes (+HYPR). Given that GW-like 6 was found farthest away from vasculature at presumably the most hypoxic points and showed highest *VEGFA* expression, we chose *VEGFA* as a proxy for pseudo-spatiotemporal (latent time) hypoxic progression. As in many cancers, *VEGFA* expression promotes angiogenesis in response to hypoxia and is thus indeed a critical factor in tumor development. *VEGFA* is also a strong inducer of vascular leakage, resulting in extravasation of plasma, which coagulates on contact with extravascular tissue factor (F3). We sorted all GW-like m-states from all samples by *VEGFA* expression and found genes that significantly correlated with six latent time windows (**Figure 6E**). *HMOX1* was found at an earlier latent time, consistent with previous work reporting this gene as a hypoxia marker in GBM[5]. *HMOX1* was followed by the expression of monocyte recruiting chemokine *CCL2* and other hypoxia genes:

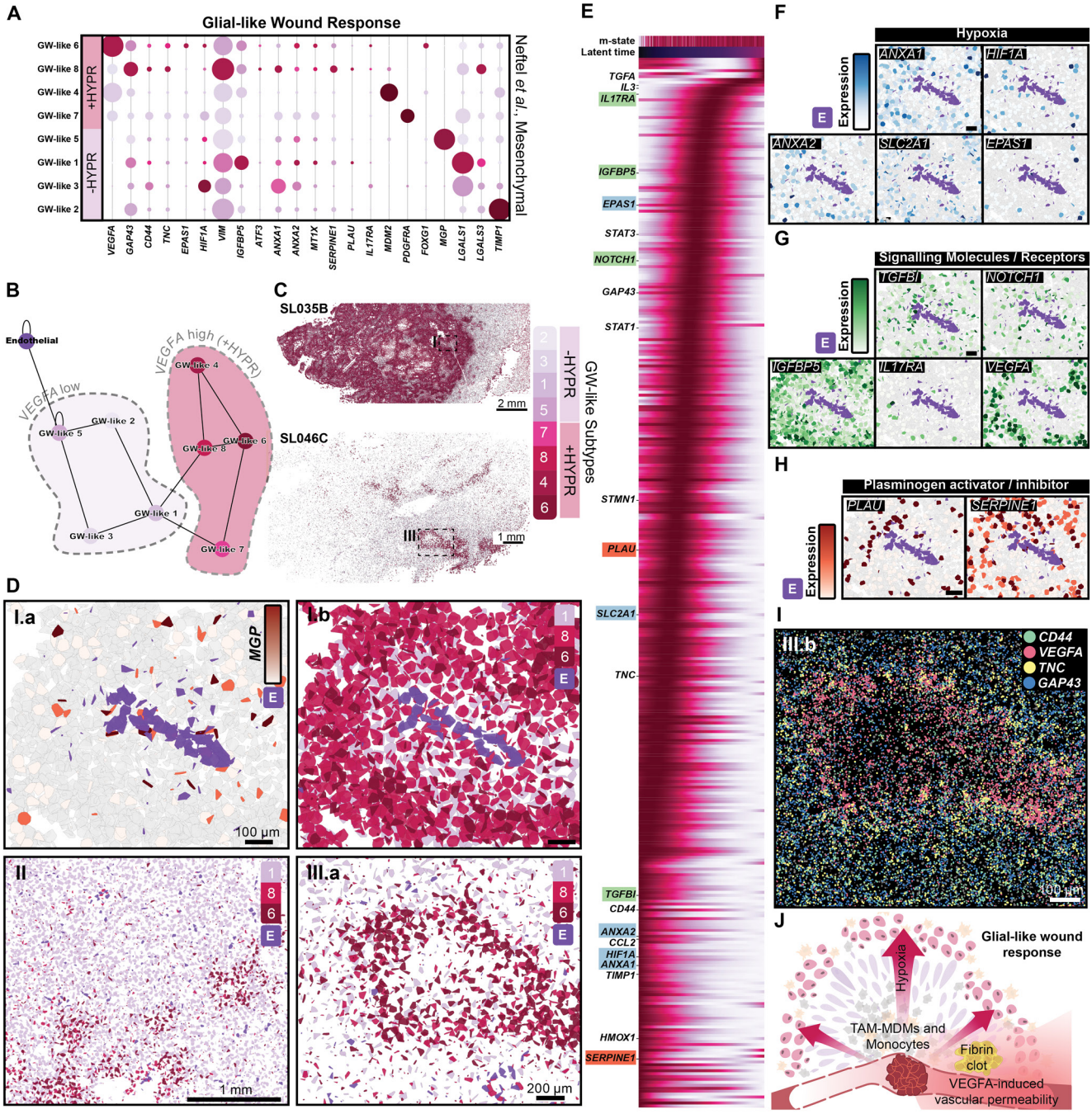


Figure 6. Molecular characterization and spatial architecture of GW-like m-states across tumors. (A) Dot plot showing the scaled marker gene expression for GW-like m-states. Dot size indicates the percentage of m-cells expressing each gene. **(B)** Graph representation of neighborhood enrichment of GW-like, endothelial and mural m-cells. Edges represent m-state state pairs with enrichment score above 8.5. **(C)** Spatial distribution of GW-like m-states in samples SL035B and SL046C. **(D)** I.a: Expression of *MGP* around endothelial and mural m-cells (denoted by "E") in region I of sample SL035B. Gray polygons represent m-cells with zero expression. I.b: Spatial distribution of GW-like 1/6/8 subtypes, endothelial and mural m-cells. II: Spatial distribution of GW-like 1/6/8, endothelial, and mural m-cells in region II of sample SL002. III.a: Spatial arrangement of GW-like 1/6/8 subtypes and endothelial/mural m-cells. **(E)** Genes correlated with VEGFA expression (latent time).

(F) Expression of hypoxia-related genes *ANXA1*, *HIF1A*, *ANXA2*, *SLC2A1*, and *EPAS1* in region I of sample SL035BC; ordered according to latent time and proximity to blood vessels. Gray polygons represent m-cells with zero expression. **(G)** Expression of signaling genes: *TGFBI*, *NOTCH1*, *IGFBP5*, *IL17RA* and *VEGFA* in region I of sample SL035B; ordered according to latent time and proximity to blood vessels. Gray polygons represent m-cells with zero expression. **(H)** Expression of signaling genes: *PLAU* and *SERPINE1* in region I of sample SL035B; ordered according to latent time and proximity to blood vessels. Gray polygons represent m-cells with zero expression. **(I)** EEL-FISH signal and nuclei showing the layered expression of *CD44*, *GAP43*, *TNC*, and *VEGFA* in region III of sample SL046C. **(J)** Proposed model for hypoxia and monocyte induced transcriptional changes in GBM cells and the transition to GW-like m-states (wound repair program).

HIF1A, *EPAS1*, *ANXA1*, *ANXA2*, and *SLC2A1* (**Figure 6F**). *ANXA1*, *HIF1A*, *ANXA2*, *EPAS1*, and *SLC2A1* were consecutively expressed at increasing distances from vasculature, indicating increased hypoxic stress and a simultaneous switch towards glycolysis. *ANXA2* has been linked to *HIF1A* mediated fibrin clearing[48] as well as astrogliosis after spinal cord injury[49], suggesting that this hypoxia-dependent progressions of gene expression might have even broader functional roles. Since hypoxia is known to stabilize HIF1/2α (*HIF1A* and *EPAS1*), thus activating *NOTCH1* which promotes cell quiescence[50], we hypothesized that similar events might be of mechanistic importance in GBM. Indeed, we observed changes in signalling genes across our latent time that suggest an order of events leading proliferating cells to a quiescent hypoxia-dependent tumor state. Namely, *TGFBI*, *NOTCH1*, *IGFBP5*, and *IL17RA* were consecutively expressed at increasing distances from blood vessels (**Figure 6G**). IGFBP5 has been reported to correlate with GBM cell invasiveness and decreased proliferation[51], and to be increased as a response to brain injury[52], whereas IL-17A has been reported to participate in neuroinflammation via its receptor IL17-RA, activating glial cells[53]. The latent time analysis also revealed a progression in the expression of urokinase plasminogen activator (*PLAU*) and its inhibitor *SERPINE1* (**Figure 6H**) with higher expression levels of both towards early in pseudo-time/space (**Figure 6E**). This is in concordance with our observation of monocytes associated with GW-like cells, as both plasmin (generated from plasminogen by PLAU) and SERPINE1 have been reported to induce monocytes (*PLAUR*$^+$) to infiltrate and polarize towards CD163$^+$ macrophages[54,55]. The expression of these genes suggests an activation of haemostasis gene regulatory programs (+HEMR) by GW-like cells to mitigate the effects of vascular leakage promoted by VEGFA[56]: plasma infiltration, fibrin clotting, and fibrinolysis.

With respect to cellular proliferation, cell cycle scores were higher in regions with lower GW-like presence (**Figure S7B**). We identified an association between *TNC* and cellular proliferation, as *TNC*-high regions were found near areas of elevated cellular proliferation, but they did not fully overlap (**Figure S7B**). As this gene was significantly correlated with GW-like transition in our latent time analysis, we hypothesized that it either marks the cells transitioning toward GW-like hypoxic states or plays an active supportive role in this process. This association was observed across different tumors, where *TNC* expression correlated with areas of GW-like states and decreased proliferation (**Figure S7B**). We observed the expression of other notable genes such as *CD44* and *GAP43* along our latent time axis (**Figure 6E**); both of these genes, together with *TNC*, have been previously associated with MES1 Neftel subtype, but are also expressed by astrocytes and reactive astrocytes[57–59]. We observed a layered structured defined by *GAP43*, *CD44*, *TNC*, and *VEGFA*, where *TNC* region corresponded to the layer found between *CD44*/*GAP43* and *VEGFA* (highest hypoxia) (**Figure 6I-III.b**). We thereby present evidence for potential organizational principles underlying GW transition and hypoxia-induced quiescent cellular states organization in GBM (**Figure 6J** and **Figure S7C/D**).

Malignant fibroblasts represent a truly mesenchymal cellular state present in gliosarcoma

Finally, we investigated the differentiation of GBM cells into FB-like m-states. Based on histopathological appearance and on the presence of FB m-states (**Figure 7A**), we identified two samples as gliosarcoma, a rare GBM subtype characterized by the simultaneous presence of glial and sarcomatous features[60]. Interestingly, these samples were not classified as gliosarcoma in the initial pathology reports, whereas samples initially classified as gliosarcoma did not show FB-like cells or gliosarcoma features on either EEL-FISH or H&E analysis (**Figure S1** and **Table S2**). Pathologist re-examination of the slides used for EEL confirmed that these tumors indeed had separate gliosarcoma and GBM regions. These findings show that features of gliosarcoma vary spatially across a single brain tumor and may be dominant in some regions while being entirely absent in others.

We firstly focused on tumor cell state composition in gliosarcoma (**Figure 7B**). Both samples were characterized by the presence of RG/AC m-cells (glial element), FB-like (sarcomatous element) m-cells, as well as GW-like m-cells. Similarly to our proposed transition through GW m-states, FB-like subtypes likely represented stages in an alternative, FB-like, differentiation trajectory, wherein *DCN*$^+$*TIMP1*$^+$*COL3A1*$^+$ FB-like 3, found in both gliosarcoma samples, was the most FB-differentiated m-state. The observed RG, AC, and GW subtypes, on the other hand, were not gliosarcoma-specific.

The spatial m-state composition varied greatly across millimetre-scale distances in gliosarcoma, and the macroscale architecture of both samples mirrored the UMAP plot (**Figure 7B/C/D**). RG (*HES1*$^+$*PAX6*$^+$) and RG-nIPC (*GLI1*$^+$) m-cells co-localized with *AQP4*$^+$ AC-like m-cells, *VEGFA*+ GW-like m-cells, *TIMP1*+ GW-like m-cells, or *COL3A1*$^+$ FB-like 3 m-cells (**Figure 7C/D**). Both gliosarcoma samples were also characterized by

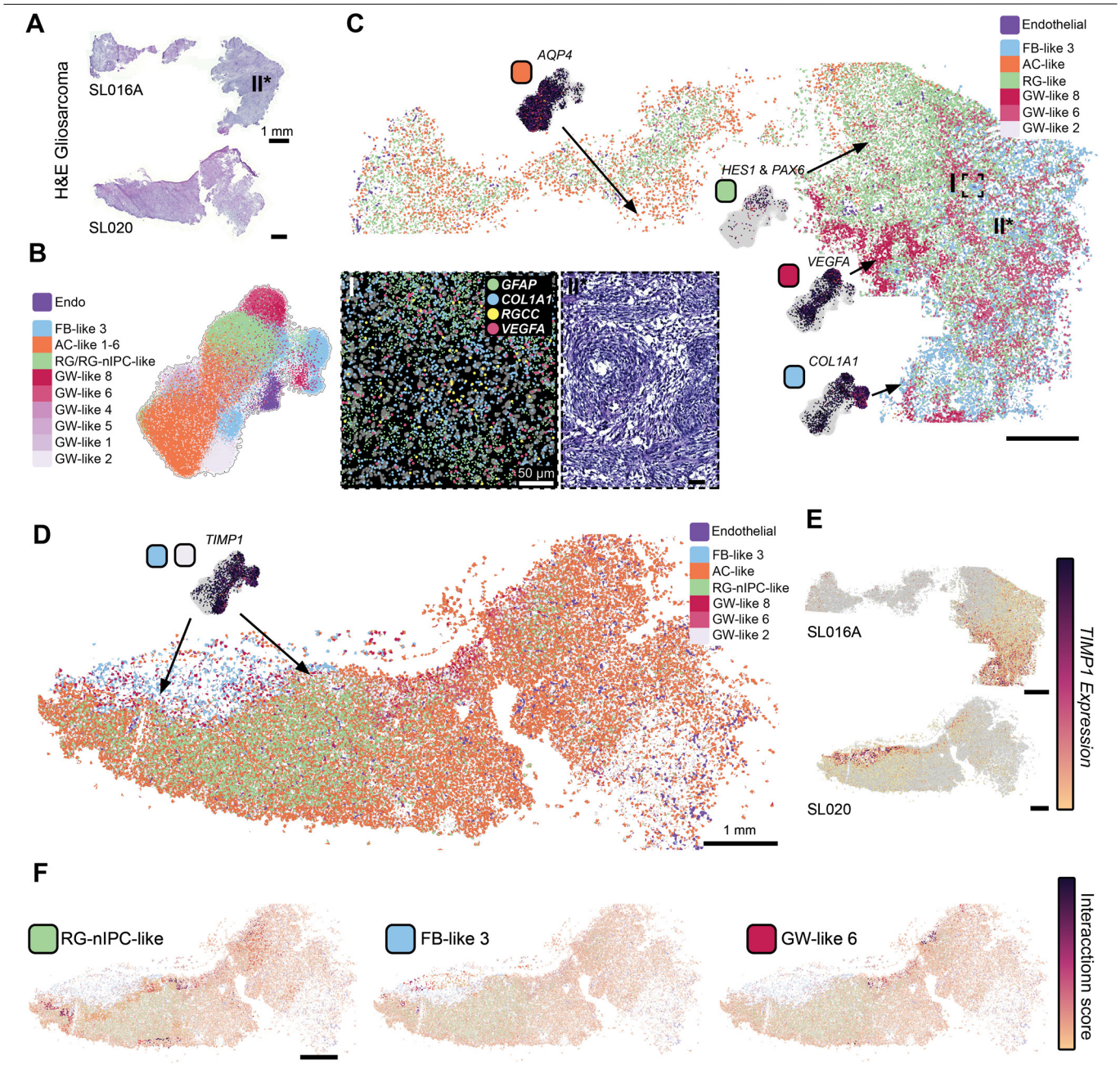


Figure 7. Molecular characterization and spatial distribution of cell types in gliosarcoma. (A) H&E staining of gliosarcoma samples SL016A and SL020. **(B)** UMAP embedding of tumor and endothelial cell m-states in samples SL016A and SL020. **(C)** Spatial distribution of m-states across gliosarcoma sample SL020, highlighting: *AQP4*$^+$ AC-like m-cells, *HES1*$^+$*PAX6*$^+$ RG-like m-cells, *VEGFA*$^+$ GW-like subtypes (6 and 8), and *COL1A1*$^+$ FB-like 3 m-cells. I: EEL-FISH signal and nuclei cytoarchitecture of gliosarcoma: *VEGFA*, *COL1A1*, *RGCC*, and *GFAP*. Gray polygons represent segmented nuclei masks. II*: Aligned cytoarchitecture in H&E staining of an adjacent section, showing biphasic glial and sarcomatous component. **(D)** Spatial distribution of m-states across gliosarcoma sample SL020, highlighting *TIMP1*$^+$ FB-like 3 and GW-like 2 m-cells. **(E)** Expression of *TIMP1* in gliosarcoma samples SL016A and SL020. **(F)** Neighborhood in situ enrichment score distribution for RG-nIPC-like (left), FB-like 3 (middle), and GW-like 6 (right) m-states.

high expression of metalloprotease inhibitor *TIMP1* (**Figure 7E**). Whilst the general macroscale architecture was conserved between the two samples, the microscale structure differed substantially. On the one hand, in sample SL016A, FB-like m-cells surrounded the endothelium and were in turn surrounded by the glial RG-like m-state (**Figure 7D-I/II**), forming a typical gliosarcoma cytoarchitecture; the RG-like cells were frequently further surrounded by *VEGFA*$^+$ GW-like m-cells (**Figure 7C-I/III**). This is consistent with our previous observation in GBM tumors, where *VEGFA*$^+$ hypoxic m-states were situated further away from vasculature. On the other hand, in sample SL020, FB-like 3 m-cells were highly aggregated and were not clustered around endothelial cells (**Figure 7D**). Lastly, we explored the spatial relationships between RG-like, AC-like, FB-like, and GW-like m-states inoma. We calculated the neighborhood enrichment scores for all permutations of m-state pairs to identify the most robust patterns of co-localization. RG-like areas surrounding either FB/AC/GW-like m-states had the highest interaction scores (**Figure 7F**), suggesting that the three latter m-states may indeed have differentiated from the RG-like m-state, similarly to how meningeal fibroblasts arise from radial glia-derived neural crest during brain development[61]. As noted above, this idea is supported by

the presence of tumor cells similar to neural crest stem cells in Neftel *et al.*[8] sample BT830 (**Figure S1B**).

Discussion

To date, glioblastoma cell state classification has largely rested upon the early bulk RNA-seq and scRNA-seq analyses. The former categorized glioblastomas into proneural, classical, and mesenchymal, while the latter refined this classification to four canonical states (AC, OPC, NPC, and MES). Subsequent spatial transcriptomics studies have followed this framework. Here, we revisited this foundational classification of GBM states from a spatial molecular organization standpoint. We acquired comprehensive spatial transcriptomic data from a large cohort of patient samples, typically imaging more than one square centimetre per specimen and in many cases including multiple samples per tumor. The large number of tumor specimens, large imaging area, and a high number of genes assayed together allowed systematic discovery and analysis of detailed transcriptional cell states and their spatial relationships. Furthermore, we exploited the availability of comprehensive transcriptional reference atlases of the normal developing human brain to refine the GBM molecular cell state classification.

Nevertheless, our study comes with several important limitations. The number of tumors studied was limited by time and resources. The size and location of tumor samples could only be partially controlled, guided by 5-ALA fluorescence during surgery and neuronavigation on MRI images. EEL-FISH is relatively insensitive, and cell segmentation in spatial analyses is challenging, leading to noisy expression data. The interpretation of cell types and states was limited by the 888-gene panel used, which had to be preselected based on prior knowledge.

We identified three least differentiated and most proliferative GBM m-states: preOPC, GBL, and RG; and four more differentiated and less proliferative states: OPC, AC, nIPC, and FB. It seems plausible that tumor growth is driven by the highly proliferative and immature cell states, which continuously self-renew and differentiate along approximately normal developmental trajectories. In support of this idea, we found, e.g., OPC-like cell states located in proximity of pre-OPCs and RG-like states in proximity of FB-like.

Most importantly, we propose that the previously described MES-like cell state has been underexplored and insufficiently resolved. We found that this cell state corresponds to a major hypoxia-reactive and wound-healing component (GW-like), and a minor fibroblast-like component (FB-like). GW-like states were broadly distributed across tumors, while the FB-like state was found primarily in gliosarcoma. Interestingly, as noted above, FB-like cells can be found under the MES category in the original Neftel *et al.* scRNA-seq dataset[8], particularly in a single sample (BT830) that may have been a gliosarcoma (definitive diagnosis would require reexamination of suitable tissue sections from the tumor). We believe that this conflation of FB-like and GW-like states has so far prevented a more nuanced understanding of these GBM components.

Given the expression of type I and III, but not muscle-specific type II, collagens, as well as the collagen crosslinking proteins Decorin (*DCN*) and Lumican (*LUM*), we consider FB-like cells to be truly mesenchymal. The emergence of such mesenchymal cells in rare gliosarcomas, as well as their occasional detection in tumors not diagnosed as gliosarcomas, and the fact that we observed distinct regions with and without presence of FB-like cells, all suggest that this cell state may arise sporadically from glial tumor cells. We noted that FB-like cells were spatially located near RG-like cells, suggesting that such immature progenitors may be differentiating towards fibroblast-like fates, similar to the differentiation of radial glia to neural crest in the early neural tube, and the subsequent maturation of neural crest stem cells into the fibroblasts that form the meninges. In agreement with this hypothesis, FB-like cells in Neftel *et al.*[8] sample BT830 expressed neural crest stem cell markers (*SOX10*, *NGFR*) and the mesenchymal transcription factor *FOXC1* which is required for normal development of the meninges[62].

In contrast, we propose the major GW-like component to be non-mesenchymal, instead representing hypoxic and wound-healing responses overlaid on essentially gliomatous cells maintaining expression of the glial transcription factor *SOX2*. Mesenchymal cells are involved in the process of wound healing in virtually all tissues except for the central nervous system[63]. There, wound healing instead entails the concurrent activity of reactive astrocytes, microglia, and other glial cells, in a process termed reactive gliosis[64,65]. Accordingly, in respect to GBM cellular state classification, we suggest that most previously described MES states might be functional modules rather than distinct cell types. That is, the GW-like ('MES') state is a modular and inducible functional state used by other GBM tumor cells. These functional GW-like modules, promoted by injurious tumor growth, were further characterized by expression of either hypoxia (+HYPR) or haemostasis (+HEMR) programs, as well as metabolic rewiring and inflammatory responses. Indeed, the process of injury-induced brain regeneration has been previously compared to certain GBM features[66].

Most GBM tumors have a high proportion of GW-like cells despite differences in mutational profiles, suggesting that this identity cannot be reduced to MES-associated mutations in, e.g., *NF1* or *PTEN*[67]. Instead, our results indicate that the wound response modules in GBM glial cells are induced by hypoxia. Moreover, GW-like cells

are found near monocytes and macrophages, both of which may thus also have active roles in maintaining, supporting, or otherwise enabling the GW-like state, similar to microglia in reactive gliosis[68]. Taken together, genetic and TME factors may have synergistic GW-like promoting effects, as *NF1* and *PTEN* deletions have been linked to increased recruitment of TAMs by GBM cells[69,70], and *NF1* deletion is implicated in pathological reactive gliosis[71]. Therefore, the maintenance of GW-like states likely depends on cross-scale systemic processes, which may serve as an important lead in understanding key causal relationships in GBM.

Hypoxia not only induces wound response cellular states, but also orchestrates a wound response tissue architecture in GBM. Our results reveal that local hypoxia is associated with induction of *HIF1A* expression and its downstream targets, leading to high *VEGFA* expression and transition to the most hypoxic cell states (+HYPR). We propose that these cells roughly correspond to MES2 as defined by Neftel *et al.*[8] In addition to promoting angiogenesis, VEGFA potently induces vascular permeability which in turn leads to plasma extravasation[56]. In healthy tissues, this extravasation initiates the coagulation cascade promoted by tissue factor (F3), causing fibrin clot deposition[72]. In the brain, tissue factor is expressed by astrocytes providing protection against haemorrhage[73]. Fibrin deposition has also been reported in GBM[74] and is known to cause monocyte recruitment[75]. This prior knowledge helps to contextualize our findings of spatial association between GW-like cells, monocytes, and macrophages and to synthesize an organizational model of the wound healing responses in GBM. We propose that monocytes associated with GW-like cells differentiate into pro-tumorigenic macrophages and acquire a hypoxic phenotype (+HYPR). The vascular leakage induced by VEGFA in these highly hypoxic zones in turn leads to a chronic elevated haemostasis response (+HEMR), which can be roughly likened to MES1 state as defined by Neftel *et al.*[8] Thereby, we demonstrate not only the presence of a haemostatic response in GBM, but also relate the chronic induction of haemostasis genes to the hypoxic organization. This functional complementarity illustrates that chronic wound response might be as fundamental in GBM organization as hypoxia.

Taken together, our results support the notion that GBM, despite its apparent complexity, follows reproducible compositional and organizational trends. In our model, GBM tumor progression can be broadly understood as a spatiotemporal transition from proneural/classical cell states into a wound response phenotype[5,7]. Cellular proliferation takes place predominantly in the periphery, where preOPC, GBL, and RG-like tumor cells divide and differentiate into their respectively mature AC/OPC/nIPC counterparts, forming structures that sustain tumor growth (**Figure S7C/D**). AC-like *CRYAB*$^{+}$ cells play a particularly significant structural role surrounding endothelial cells. Regardless of the underlying dominant tumor cellular state, rapid tumor growth leads to local hypoxia and *VEGFA* expression. VEGFA in turn causes plasma extravasation from local vasculature, which is coagulated by tumor cell-expressed tissue factor. A continuous cycle of hypoxia and haemostasis leads to chronic induction of hypoxic and haemostatic response genes — GW-like gene modules — in spatial domains organized by the distance to functional vessels. We therefore re-interpret the elusive MES-state as a point of functional convergence centred around wound healing and hypoxic responses, accessible from all GBM cell states and maintaining the underlying glial-like cell identity.

The results presented here help to unify and relate GBM cellular states to the underlying material processes of tumor expansion, and to reinterpret GBM as a wound that does not heal[76]. The distinction between glial identity and GW-like wound healing function could have important therapeutic implications, as cells of glial identity could be a promising target for cell-state specific therapies. In contrast, the GW-like activity is likely a secondary consequence of rapid tumor growth, and thus dependent on the proliferative capacity of immature glial-like progenitors. On the other hand, the FB-like cell states likely represent true metaplasia to a mesenchymal cell identity. This presents yet another novel target for cell-state specific therapies, particularly noteworthy in gliosarcomas and in gliomas with regional sarcomatous features.

Author contributions

S.L., A.M.A., and J.J. conceived the study. A.M.A., J.J., I.K., and S.L., interpreted the results. J.J. and A.M.A. performed EEL-FISH experiments. I.K. performed cryosectioning and additional stainings. L.E.B., A.M.A., and J.J. optimized the EEL-FISH protocol. A.M.A. performed data analysis. S.C. developed and performed image analysis. E.K. and J.B.M. performed Flow cytometry experiments and analysis. A.M.A. and S.L. designed the probes and the barcoding scheme. S.L performed scRNA-seq analysis. O.P. and J.K.J. provided tumor specimens and clinical annotation. A.S. performed pathology analysis. A.M.A., J.J., and S.L. wrote the manuscript. All authors provided critical feedback and helped shape the project, analysis, and the manuscript.

Acknowledgements

We thank Anna Johnsson, Karolinska Institute for grant management, Lijuan Hu, Karolinska Institute for lab management, and all members of the Linnarsson lab for helpful discussions. The authors also thank all neurosurgeons and nurses at the Department of Neurosurgery at Karolinska University Hospital, and

patients for donating tissue. This work was supported by Region Stockholm (support to J.K.J) and the Erling-Persson Family Foundation (Atlas of childhood disease), Hjärnfonden (FO2023-0309), the Swedish Research Council (2022-01248), and the Torsten Söderberg Foundation grants to S.L.

Declaration of interests

SL is a paid scientific advisor to Moleculent, Combigene, and the Oslo University Center of Excellence in Immunotherapy. S.L., A.M.A., J.J., S.C., and L.E.B. are shareholders in EEL Transcriptomics AB, which owns the intellectual property to the EEL method ("Spatial RNA localization method", provisional US patent application 63/139,701). All other authors declare that they have no competing interests.

Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sten Linnarsson (sten.linnarsson@ki.se).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw image data will be shared by the lead contact upon request. Gene expression data and RNA molecules spatial coordinates have been deposited at google storage and can be accessed at: https://github.com/linnarsson-lab/glioblastoma_spatial/. The m-cell segmented data is available in AnnData format compatible with scanpy, squidpy, and FISHspace. The EEL-FISH RNA xy coordinates are available as parquet files compatible with FISHscale. All original code has been deposited at https://github.com/linnarsson-lab/FISHscale for the analysis of and visualization of RNA molecules and https://github.com/linnarsson-lab/FISHspace for single cell analysis and visualization.

Experimental model and study participant details

Patient samples

Human glioblastoma samples were collected with informed consent from the patients and with ethical approval from the Swedish Ethical Review Authority (2020-03505). The use of samples was approved by the Swedish Ethical Review Authority (2020-02096). Information on patients' age and sex can be found in Table S2.

Method details

Tissue sample collection

Glioblastoma tumor samples were kept in Hibernate-E medium on ice during transportation from the surgery. Upon arrival, samples were promptly washed with and embedded in Tissue-Tek Optimal Cutting Temperature compound, and snap-frozen in 2-methylbutane and dry ice slurry. Samples were stored at −80°C until further use.

EEL-FISH coverslip functionalization

EEL-FISH coverslips were functionalized as outlined in a previously published protocol13, with some minor adjustments. Indium tin oxide (ITO)-coated coverslips (24 × 60 mm #1.5 thickness) with a surface resistivity of 30–60 Ω/Sq were placed in a beaker with acetone and cleaned by sonication for 20 min in a Ney ULTRAsonik 28X sonicator set to maximum power. Acetone was subsequently replaced with isopropanol, and the beaker with coverslips was sonicated again for 20 min. Isopropanol was replaced with nuclease-free dH2O for a final 20-min sonication cycle. The sonicator bath was refilled with cold water after each cycle to avoid excessive heating. Cleaned coverslips were stored in fresh nuclease-free dH2O for at least one week before further use (prolonged storage of over a month,

however, is not recommended as the ITO coating will deteriorate).

Immediately before surface functionalization, the cleaned coverslips were placed in an Epridia E103 rack and dried with nitrogen gas. Once dry, the coverslips were submerged in 2% (vol/vol) (3-glycidyloxypropyl)trimethoxysilane in acetone for 2 hours under a nitrogen atmosphere. Coverslips were rinsed once with acetone and allowed to dry under a nitrogen atmosphere. The ITO-coated side of the coverslip was identified with a resistance measurement using a multimeter; only this side was subjected to further functionalization. Firstly, 40 µl of 10 µM solution of 5' amine-modified oligo(dT)60 (/5AmMC6/UUUUGACTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT/iSuper-dT/TTTTTTTTTTTTTTTTTTTTTTTTTTTT/iSuper-dT/TT) in 1× Schott Spotting Solution was placed on the center of each coverslip and covered with a 22 × 22 mm HybriSlip. This reaction was carried out for 1 hour at room temperature before washing five times with 200 µl of 2× SSC buffer and two times with nuclease-free dH2O. Next, to improve RNA capture and to block residual epoxy groups, 60 µl of 0.1% (wt/vol) poly-D-lysine (molecular weight 70,000–150,000) solution in nuclease-free dH2O was placed on the center of each coverslip. The solution was covered with a 22 × 22 mm HybriSlip and kept on the coverslips for 30 min at room temperature before washing three times with 200 µl of nuclease-free dH2O. Europium-chelate beads with a diameter of 0.2 µm and 1 µm were respectively diluted 1:250 and 1:333 in nuclease-free dH2O and deposited onto the center of each coverslip by placing a 100 µl drop for 3 min. The europium bead mixture was subsequently removed, and the coverslips were allowed to dry in air. Coverslips were used within two days; coverslips which were not used immediately were stored under a nitrogen atmosphere.

EEL-FISH and histology section arrangement

Cryosections of 10 µm thickness were captured on the functionalized area of EEL-FISH coverslips or SuperFrost Plus slides (for histological analysis) and allowed to dry for a few minutes before transferring to a −80°C freezer for storage. The cryosections were collected in series of adjacent slices: three SuperFrost Plus slides with two tissue sections per slide were collected for histological analysis, followed by the EEL-FISH slide with one or more sections (depending on tissue area), followed by three more slides with two sections per slide for histological analysis. Five slides were stained with histological staining protocols and one slide was used as a negative control for tissue autofluorescence.

Probe design

The RNA-binding encoding probe design was based on previously reported procedures13 and was implemented in the software package Oligopy[76]. Each probe consisted of a 26- to 32-nucleotide-long sequence, reverse-complementary to the target RNA sequence, and two overhanging tails containing six detection sequences of 20 nucleotides each, to which the readout probes can bind. These six sequences were appended to the RNA-binding sequences in a random order to avoid detection bias arising from potential steric hindrance. Furthermore, readout tails were separated by 2 nucleotides (alternating TT, AT, TA, and AA).

Gene identities were encoded in 16-bit binary barcodes, with each barcode containing six positive bits and being read out in a single fluorescence channel, thus allowing up to 448-gene multiplexing per fluorescence channel over 16 rounds of imaging. In total, we used three different encoding probe panels, each designed to be read out in different fluorescence channels for orthogonal detection. Two of these panels constituted the 888 genes used in further analysis, whereas the third panel was used in some experiments for multiplexing proof-of-principle and for back-up detection of some of the 888 main genes.

RNA-binding regions were selected to have a Gibbs free energy (ΔG) at 37 °C of −28 kcal mol−1or less, G+C content of 40–60% and a minimum ΔG at 37 °C for hairpin or homodimer formation of −9.0 kcal mol−1. Then, probe specificity was analyzed using BLAST, and probes with more than 60% identity to five or more off-target RNA species were dropped. Finally, sets of a maximum of 28 probes that would tile with a minimum gap of 2 bp on the RNA were retrieved, and genes that did not reach a minimum of ten probes were dropped.

To prevent optical crowding, we optimized the distribution of genes over the available barcodes using prior knowledge from single-cell RNA-seq data. Whenever possible, the genes expressed in the same cell type were not labeled in the same decoding cycle.

Readout probes, ordered from IDT, consisted of 20-nucleotide-long sequences complementary to the detection sequences on the encoding probes and a disulfide-conjugated fluorophore on either 5' or both 5' and 3' ends. Readout probes corresponding to each of the three encoding probe panels were conjugated to two Cy3, Alexa Fluor 594, or Alexa Fluor 647 fluorophores, respectively.

Encoding probe amplification

Encoding probes were ordered from Twist Bioscience as Oligo Pools. Probe amplification was carried out separately for each of the three Oligo Pools but following the same previously published procedure[77].

Briefly, each Oligo Pool was resuspended in TE buffer to a final concentration of 1 ng/µl. Resuspended pools were PCR-amplified to generate an intermediate stock, which was subsequently purified using Zymo Oligo Clean and Concentrator kit and subjected to another PCR amplification round with purification. When working with new probe pools, it is recommended to perform a pilot qPCR reaction before proceeding to each PCR amplification round to establish how many cycles are needed (reaction should be stopped when the curve plateaus; the corresponding number of cycles ranged from 8 to 14 in our experience). Next, the probes were transcribed in vitro for at least 16 hours at 37°C, after which the RNA was purified using Zymo Oligo Clean and Concentrator kit. A reverse transcription reaction was carried out for 1 hour at 55°C using a 5'-amine-modified forward primer so that the encoding probes could be formaldehyde-fixed during EEL-FISH procedure. It should be noted here that we observed improved quality of the final probe product when the initial RNA concentration for the reverse transcription step was limited to approximately 150 ng/µl. To remove the RNA, alkaline hydrolysis was run for 15 min at 95°C. The final single-stranded DNA (ssDNA) probe product was purified using Zymo Oligo Clean and Concentrator kit and its concentration was measured using Qubit ssDNA assay kit. Probe and primer sequences are available in Table S3.

EEL-FISH

EEL-FISH was performed as previously described[13], with some minor modifications to adapt the procedure for glioblastoma tissue. For a more detailed overview of the method please refer to the original article. Unless stated otherwise, the volume per wash in every step of this part of the protocol was 200 µl. After the procedure has been initiated, the samples must not be allowed to dry out at any point, as excessive drying could cause steep and irreversible increase in background fluorescence.

Before beginning the experiment, two crosses flanking the tissue sample were drawn on the bottom side of the coverslip with a solvent-resistant Moist Mark Plus pen. These crosses would later be used to locate the sample area after tissue digestion. Nuclei were then stained with a 1 µg/ml propidium iodide solution in 2× SSC and washed five times with 2× SSC. To prevent the tissue from drying out during reference imaging, a parafilm gasket was placed around the tissue, and a drop of 2× SSC was added and sandwiched with a 24 × 32 mm coverslip. An overview image of the nuclei was taken with a ×10 objective, and the area of interest was traced to generate field of view (FOV) positions for ×40 and ×60 objectives. The FOVs for the ×60 objective and the corresponding coordinates for the reference crosses were saved for later use. Reference images of nuclei and europium beads were taken with the ×40 objective.

After the reference images were acquired, the sample was permeabilized with 0.1% (vol/vol) Triton X-100 and 10 mM dithiothreitol (DTT) in 1× TBE buffer for 5 minutes. The sample was then washed five times with 1× TBE, and any spillover or excess solution was removed with tissue paper to facilitate the adhesion of electrophoresis accessories. An electrical wire was then attached to the ITO-coated surface with conductive copper foil tape. Another piece of wire was attached with copper foil tape to a non-functionalized ITO-coated coverslip to be used as a cathode. Two 1.5-mm-thick polydimethylsiloxane spacers were placed on opposite sides of the tissue sample and covered with the cathode coverslip. This assembly was placed in a custom-made holder to prevent movement. An electrophoresis buffer containing 10 mM DTT and 1 M urea in 1× TBE was added between the two coverslips with a gel-loading pipette tip. The wires were then connected to a Keithley 2450 SourceMeter, and a potential of 1.5 V (or 10 V/cm) was applied for 20 min. Afterwards, the electrophoresis setup was carefully disassembled, and excess electrophoresis buffer was removed. To enhance the RNA capture, the sample was incubated with 6× SSC buffer for 5 minutes. The sample was then washed twice with 2× SSC and kept in 2× SSC while a digestion buffer containing 1% (wt/vol) sodium dodecyl sulfate (SDS), 20 mM tris-HCl pH 7.4, 100 U/ml Superase RNase inhibitor, and 0.8 U/ml proteinase K was prepared. The sample was then digested by incubating with the digestion buffer four times for five minutes at 30°C. We noticed high variability in glioblastoma tissue sensitivity to these digestion conditions, and therefore we strongly recommend adjusting the proteinase K concentration and the number of digestion buffer washes for each sample individually. For example, doubling the proteinase K concentration or adding one more five-minute wash with the digestion buffer helped to clear the more persistent samples. However, we do not recommend implementing the harsher conditions unless the sample is visibly not removed by the conventional digestion conditions, as excessive digestion might lead to poorer RNA capture. The digestion was followed by four washes with 5% (wt/vol) SDS in 2× SSC (pre-warmed to 30°C) for five minutes each at 30°C. The sample was subsequently washed five times with 2× SSC at room temperature, fixed with 4% formaldehyde in 1× PBS for 10 minutes, and washed five more times with 2× SSC.

Encoding probes, in amounts calculated to correspond to 1 nM concentration per unique probe once resuspended, were dried using a SpeedVac and resuspended in 25 µl of hybridization mix containing 30% (vol/vol) formamide, 0.1 g/ml dextran sulfate (molecular weight > 500,000), 1 mg/ml Escherichia coli tRNA, 2 mM ribonucleoside vanadyl complexes (RVC), and 200 µg/ml BSA in 2× SSC. The probe solution was then added onto the sample, carefully covered with a 22 × 22 mm plastic HybriSlip avoiding trapping any air bubbles. The encoding probes were allowed to hybridize in a Petri dish humidified with 2× SSC for at least 40 hours at 38.5°C.

After the hybridization, the HybriSlip was carefully removed by adding some 2× SSC to help detach the HybriSlip from the coverslip surface. Excess hybridization mix was removed from around the tissue area to allow tight mounting of the sample into a custom flow cell (Rebus Biosystems). The flow cell was then connected to the ROBOFISH fluidics system[78,79], which performed all subsequent steps automatically, and a TC-720 temperature controller (TE Technology). The volume of the flow cell chamber measured around 400 µl; all subsequent washing steps used at least double this volume to ensure proper replacement of chamber contents. When using an alternative setup, all volumes should be adjusted accordingly to be at least double of the chamber volume. To remove unbound encoding probes, the sample was first flushed with 2× SSC and then washed with 30% formamide in 2× SSC four times for 15 min at 47 °C, followed by four washes with 2× SSC. To prevent signal deterioration over multiple rounds, the encoding probes were fixed with 4% formaldehyde in 1× PBS for 10 min and then washed with 2× SSC. Fluorescent readout probes were dispensed to the sample at a concentration of 50 nM per unique probe in hybridization mix containing 10% formamide and hybridized for 10 min at 37°C. Unbound readout probes were flushed away with three washes with 20% formamide in 2× SSC for 3 minutes each, followed by four washes with 2× SSC. The temperature was then set to 20°C, and a degassed imaging buffer containing 2 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), 5 mM 3,4-dihydroxybenzoic acid (DBA, neutralized with NaOH), and 20 nM protocatechuate 3,4-dioxygenase from Pseudomonas sp. (PCD, initially resuspended at 2 µM in 50% glycerol, 50 mM KCl, 1 mM EDTA, and 100 mM Tris-HCl pH 8.0) in 2× SSC was injected into the flow cell. Imaging was initiated by the ROBOFISH system and performed on a Nikon Ti2 microscope. Once an imaging round had been completed, the sample was washed four times with 2× SSC, and the fluorophores were subsequently removed through reduction of the disulfide bond by washing twice with degassed 50 mM tris(2-carboxyethyl)phosphine (TCEP) solution in 2× SSC for 10 min each at 22 °C. TCEP solution was flushed out with ten washes with 2× SSC. Then, 15 more cycles of iterative readout hybridization, washing, imaging, and fluorophore stripping were performed as described here, thus covering the 16-bit barcode scheme. All solutions except 2× SSC were kept at 4°C, and all degassed solutions were covered with mineral oil.

EEL-FISH image acquisition

EEL-FISH image acquisition was performed on a Nikon Ti2 epifluorescence microscope equipped with a Nikon CFI Plan Apo Lambda ×60 oil immersion objective with a numerical aperture (NA) of 1.4, CFI Plan Apo Lambda ×40 objective with NA 0.95, CFI Plan Apo Lambda ×10 with NA 0.45, Nikon motorized stage, Nikon Perfect Focus system, Sona 4.2B-11 back-illuminated sCMOS camera with pixel size of 11 µm (Andor). The ×10 and ×40 objectives were only used for reference imaging, whereas the final EEL-FISH data acquisition was done using the ×60 objective. We used Lumencor Spectra light engine and matching filter cubes from Chroma. Namely, the following component configurations were used: europium channel was imaged using 360/28 bandpass filter (400 mW light source; power set to 30%), ET667/30m emission filter, and T647lpxr dichroic mirror; Cy3 channel was imaged using 534/20 bandpass filter (400 mW light source, power set to 100%), ET575/40m emission filter, and T550lpxr dichroic mirror; Alexa Fluor 594 channel was imaged using 586/20 bandpass filter (400 mW light source, power set to 100%), 89403X excitation filter, 89403m multiband emission filter, and 89403bs dichroic mirror; Alexa Fluor 647 channel was imaged using 631/28 bandpass filter (500 mW light source, power set to 100%), ET667/30m emission filter, and T647lpxr dichroic mirror. All exposure times were set to 100 ms. EEL-FISH image acquisition was performed using a custom Nikon NIS Elements imaging job automatically initiated by the ROBOFISH system (see https://github.com/linnarsson-lab/ROBOFISH for more details). Images of 12-bit depth were acquired for each FOV (8% overlap for the ×60 objective) with a z-stack of 17 steps of 0.3 µm.

Sample preparation for flow cytometry, intracellular flow cytometry, acquisition, and analyses

Tumor tissue was collected within 1h after surgical resection. Samples were first mechanically dissociated with a scalpel into approximately 1-3 mm pieces and enzymatically digested in RPMI-1640 medium (Cytiva) containing 0.25 mg/ml DNase I (Roche) and 0.25 mg/ml Collagenase (Sigma-Aldrich) at 37 °C for 25 min while stirring at 900 RPM (speed 4/6) on a hotplate stirrer (VWR). The digested suspension was strained through a 70-um strainer, neutralized by adding RPMI-1640 with 10% heat inactivated FBS (HI FBS, Gibco), and spun down at 200 x g for 10 minutes to pellet the cells, which were cryopreserved in HI FBS with 10% DMSO (Sigma). Cells were directly transferred to a -80 °C freezer in Coolcell freezing containers (Corning) to ensure a cooling rate of -1°C/minute and stored in N2.

For flow cytometry analyses, frozen cell suspensions were washed in 10 ml RPMI-1640 containing 10% HI FBS and intracellular staining was performed as previously described[80]. Briefly, cells were first stained for 20 min at RT with extracellular antibodies in the staining mix containing FACS buffer (dPBS (Sigma) with 2% HI FBS (Gibco) and 0.4% 0.5M EDTA (Invitrogen) added), 1:5 BD Horizon Brilliant Stain buffer (BD Biosciences), and 1:25 FcR blocking reagent (Miltenyi), followed by fixation using BD FixPerm buffer (BD Biosciences) for 30 min at RT, and stained with intracellular antibodies in BD Perm/Wash buffer (BD Biosciences) for 20 min at RT. The antibody panel (Key

resource table) was designed based on key markers from Fig. 4B and backbone from Kvederaite *et al*[81]. Flow cytometry was performed on a FACSymphony A5 (BD Biosciences) and fcs files were generated using DIVA (BD Biosciences). Analysis was done in FlowJo v10.5.3 (BD Biosciences) and statistics was performed using Prism 9 (GraphPad, San Diego, CA, USA).

Hematoxylin and eosin staining

A hydrophobic barrier was drawn on a slide around tissue sections using the ImmEdge Hydrophobic Barrier PAP Pen. The sections were fixed with 4% formaldehyde in 1× PBS for 10 min. The slide was subsequently washed three times with 1× PBS and three times with dH2O for 5 min each. Hematoxylin solution (Hematoxylin and eosin stain kit) was applied to each section for 1 min and rinsed in two changes of dH2O. Bluing agent (Hematoxylin and eosin stain kit) was applied to each section for 15 seconds and the slide was rinsed in two changes of dH2O. The slide was dipped in ethanol and excess ethanol was blotted off. Eosin Y solution (Hematoxylin and eosin stain kit) was applied to each section for 30 seconds and rinsed in ethanol. The sections were subsequently dehydrated in three changes of ethanol for 2 min each and rinsed in xylenes, before mounting with Entellan™ new. The slide was left to solidify for 24 hours in a ventilated hood.

Neuronal/RNA/cytoskeletal/vascular staining

A hydrophobic barrier was drawn on a slide around tissue sections using the ImmEdge Hydrophobic Barrier PAP Pen. The sections were fixed with 4% formaldehyde in 1× PBS for 10 min. The slide was subsequently washed three times with 1× PBS for 5 min each. The sections were permeabilized with 0.1% Triton X-100 in 1× PBS for 10 min and subsequently washed twice with 1× PBS for 5 min each. A stain mixture was prepared with following dilutions: 1:100 NeuroTrace™ 435/455 Blue Fluorescent Nissl Stain, 1:1000 SYTO™ RNASelect™ Green Fluorescent cell Stain, 1:200 Lycopersicon esculentum (Tomato) Lectin, and 1:400 Alexa Fluor™ Plus 647 Phalloidin with 1% BSA in 1× PBS solution. The sections were incubated with 200 µl of stain mixture per section for 30 min at room temperature in a dark staining chamber. The sections were permeabilized with 0.1% Triton X-100 in 1 × PBS for 10 minutes and subsequently washed three times with 1× PBS for 5 min each in a dark staining chamber. The slide was washed in 1 × PBS for 30 min in a dark staining chamber, before mounting with ProLong™ Gold Antifade Mountant and being coverslipped. The slide was left to solidify overnight in a dark staining chamber.

Apoptosis staining

A hydrophobic barrier was drawn on a slide around tissue sections using the ImmEdge Hydrophobic Barrier PAP Pen. The sections were fixed with 4% formaldehyde in 1× PBS for 10 min. The slide was subsequently washed three times with 1× PBS for 5 min each. The sections were permeabilized with 0.1% Triton X-100 in 1× PBS for 10 min and subsequently washed in two changes of 1× PBS for 5 min each. The slide was rinsed with deionized water. The sections were incubated with 100 µl of TdT reaction buffer (Click-iT™ Plus TUNEL Assay Kits for In Situ Apoptosis Detection) per section for 10 min at 38.5°C in a humidifying chamber. The TdT reaction buffer was removed, and sections were incubated with 100µl of TdT reaction mixture (made according to manufacturer's instructions, Click-iT™ Plus TUNEL Assay Kits for In Situ Apoptosis Detection) for 1 hour at 38.5°C in a humidifying chamber. The slide was rinsed in deionized water and subsequently washed with 3% BSA and 0.1% Triton X-100 in 1× PBS for 5 min. The slide was then rinsed in 1× PBS. The Click-iT™ Plus TUNEL reaction cocktail was made according to manufacturer's instructions (Click-iT™ Plus TUNEL Assay Kits for In Situ Apoptosis Detection) and mixed well by vortexing. Sections was incubated with 100 µl of The Click-iT™ Plus TUNEL reaction cocktail per section for 30 min at 38.5°C in a dark humidifying chamber. The reaction cocktail was then removed, and slide was washed with 3% BSA in 1× PBS in a dark staining chamber. The slide was rinsed in 1× PBS and subsequently washed in two changes of 1× PBS for 5 min each in a dark staining chamber. A nuclear stain mixture of 2 µg/ml Hoechst 33258 in 1× PBS was prepared. The sections were incubated with 200 µl of the nuclear stain mixture per section for 15 min at room temperature in a dark staining chamber. The slide was washed three times with 1× PBS for 5 min each in a dark staining chamber before mounting with ProLong™ Gold Antifade Mountant and being coverslipped. The slide was left to solidify overnight in a dark staining chamber.

Nuclear/endosomal membrane/cell membrane/vascular staining

A hydrophobic barrier was drawn on a slide around tissue sections using the ImmEdge Hydrophobic Barrier PAP Pen. The sections were fixed with 4% formaldehyde in 1× PBS for 10 min. The slide was subsequently washed three times with 1× PBS for 5 min each. The sections were permeabilized with 0.1% Triton X-100 in 1× PBS for 10 min and subsequently washed twice with 1× PBS for 5 min each. A stain mixture was prepared containing: 2 µg/ml Hoechst 33258, 1:200 Lectin HPA From Helix pomatia (edible snail) Alexa Fluor™ 647 Conjugate, 1:200 Wheat Germ Agglutinin, and 1:200 Concanavalin A Conjugates with 1% BSA in 1× PBS solution. The sections were incubated with

200 µl of stain mixture per section for 30 min at room temperature in a dark staining chamber. The sections were then permeabilized with 0.1% Triton X-100 in 1 × PBS for 10 min and subsequently washed three times with 1× PBS for 5 min each in a dark staining chamber. The slide was washed in 1 × PBS for 30 min in a dark staining chamber, before mounting with ProLong™ Gold Antifade Mountant and being coverslipped. The slide was left to solidify overnight in a dark staining chamber.

Vascular/Immune/Proliferation staining

A hydrophobic barrier was drawn on a slide around tissue sections using the ImmEdge Hydrophobic Barrier PAP Pen. The sections were fixed with 4% formaldehyde in 1× PBS for 10 min. The slide was subsequently washed with 1× TBS for 5 min. The slide was washed with three changes of 0.025% Triton X-100 in 1× TBS for 5 min each with gentle agitation. The sections were incubated with 200 µl of 10% normal goat serum and 1% BSA in 1× TBS blocking solution per section for 2 hours at room temperature. A mixture of primary antibodies was prepared with following dilutions: 1:1000 Recombinant Anti-Ki67 antibody [B56] and 1:250 Recombinant Anti-CD45 antibody [EP322Y] with 1% BSA in 1× TBS. The sections were incubated with 200 µl of primary antibody mixture per section in a humidifying chamber at 4°C overnight. The slides were washed with two changes of 0.025% Triton X-100 in 1× TBS for 5 min each with gentle agitation. A mixture of secondary antibodies was prepared with following dilutions: 1:500 Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody Alexa Fluor™ Plus 555 and 1:500 Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody Alexa Fluor™ 647 with 1% BSA in 1× TBS. The sections were incubated with 200 µl of secondary antibody mixture per section for 1 hour at room temperature in a dark staining chamber. Subsequently, the slide was washed with three changes of 1× TBS for 5 min each in a dark humidifying chamber. A stain mixture was prepared containing 2 µg/ml Hoechst 33258 and 1:200 Lectin HPA From Helix pomatia (edible snail) Alexa Fluor™ 488 Conjugate in 1× TBS. The sections were incubated with 200 µl of stain mixture per section for 30 min at room temperature in a dark staining chamber. The slide was washed with 0.025% Triton X-100 in 1 × TBS for 30 min and subsequently washed in three changes of 1× TBS for 5 min each at room temperature in a dark staining chamber. The slide was washed in 1 × TBS for 30 min in a dark staining chamber, before mounting with ProLong™ Gold Antifade Mountant and being coverslipped. The slide was left to solidify overnight in a dark staining chamber.

Histology image acquisition

Histology image acquisition was performed on a Zeiss Axio Scan.Z1 slide scanner with Zen Blue v3.1 acquisition software, using the Plan-Apochromat 20x/0.8 M27 objective with NA 0.8. For acquiring brightfield images (Hematoxylin and eosin staining) Hitachi HV-F203SCL camera (pixel size of 4.4 µm) with VIS-LED illumination source were used. For acquiring fluorescence images Hamamatsu ORCA-Flash4.0 V3 Digital CMOS camera (pixel size of 6.5 µm) with a Colibri 7 LED light source (100% intensity), PBS 405 + 493 + 575 + 654 + 761 beamsplitter, and PBP 425/30 + 514/31 + 592/25 + 681/45 + 785/38 emission filters. Namely, the following configurations were used: H3258 channel (Hoechst stain) was imaged using LED-Module 385 nm and BP 385/30 excitation filter; AF430 channel (NeuroTrace Nissl stain) was imaged using LED-Module 475 nm and BP 469/38 excitation filter; AF488 channel (SYTO RNASelect stain, Wheat Germ Agglutinin stain, Tissue autofluorescence, Lectin HPA Alexa Fluor 488 conjugate) was imaged using LED-Module 475nm and BP 469/38 excitation filter; AF555 channel (Goat anti-Mouse IgG Alexa Fluor 555) was imaged using LED-Module 567 nm and BP 555/30 excitation filter; AF594 channel (Tomato Lectin, Click-iT Plus TUNEL Assay for In Situ Apoptosis Detection, Concanavalin A) was imaged using LED-Module 567 and BP 555/30 excitation filter; AF647 channel (Phalloidin, Lectin HPA Alexa Fluor 647 conjugate, Goat anti-Rabbit IgG Alexa Fluor 647) was imaged using LED-Module 630 nm and BP 631/33. Histology image acquisition was performed at Karolinska Institutet's Biomedicum Imaging Core facility (https://ki.se/en/bic).

Quantification and statistical analysis

Image processing

We used pysmFISH to automatically detect and decode the EEL-FISH signal[82]. Shortly, after parsing, images were filtered, and large objects with a size over the 95th quantile were removed. Then, spots were counted and registered between hybridizations using the europium fiducial beads. The barcodes are then identified using the nearest-neighbor approach. To map the EEL-FISH signal (Å~60 objective) to the nuclei images (Å~40 objective), we first registered the europium beads imaged with both objectives by using a point set alignment algorithm based on a nearest-neighbor search, and the resulting transformation (scaling, rotation and shift) was applied to the EEL-FISH signal. Nuclei were segmented using Cellpose[83]. To process the large amount of data generated by each EEL-FISH experiment, the analysis is parallelized using Dask[84].

Image registration of histology images to EEL-FISH

For the alignment to EEL-FISH data, image registration was performed using the SimpleElastix python wrapper (v0.10.0) for Elastix software (v5.1.0). Each Hematoxylin and eosin (H&E) stained tissue section was aligned to the segmented nuclei image of the EEL-FISH tissue section (fixed image). RGB images of H&E tissue sections were split into individual red, green and blue value images. Binary masks of the H&E were created from the blue channel, based on their histogram to represent the nuclei. This binary nuclear H&E mask (moving image) was then aligned to EEL-FISH nuclei image (fixed image), using affine transformation with 15 resolutions and 2500 iterations. The original red, green and blue value images were aligned using the extracted transformation matrices.

M-states – Graph Attention Network

Transcriptional States, m-states, were inferred using a cell-free segmentation approach using a graph attention neural network (GAT) as previously proposed by Partel and Whalby17. Only the 888 RNA species common to all EEL-FISH experiments were used for further analysis. GFAP and MOG molecules were removed due to their high expression. First, molecules were connected into an undirected graph if the distance between two molecules was below the 95th quantile at which all molecules were connected to at least 3 neighboring molecules (~5-10 µm). Nodes in the graph represent RNA molecules and every node has a feature vector which is a one hot encoding the identity of the RNA transcript, edges in the graph connect neighboring molecules. We generated a graph for each EEL-FISH sample and pruned all graphs by removing molecules connected to less than 5 neighbors. The GAT was designed using pytorch and Deep Graph Library (https://www.dgl.ai/) with the following architecture: 2-hop neighbor aggregation using GAT2vCONV[85,86] module from DGL and a final linear layer; and the following hyperparameters: a random sampling strategy with 20 and 10 neighbors for 1st and 2nd hop, respectively, 64 units for each hidden layer and 48 units for generated lower dimensional embedding. The GAT was trained using Adam as optimizer[87] with a learning rate of 0.0001 and mini-batch sampling of 512 nodes per iteration. Model parameters are learned using an unsupervised method under the premise that neighboring RNA molecules in the should have similar embeddings, whilst distant molecules should have dissimilar embeddings. During training, embeddings are generated for positive nodes, true neighbors, and negative nodes, randomly sampled nodes, from each central node and the model's objective function consists of the binary cross-entropy between positive and negative embedding for each sampled node.

We trained our generalized GBM-GAT for 10 epochs on 2-hop subgraphs generated from 500,000 random nodes, from 30% of our EEL-FISH sections, constituting a graph for training with a total of 33.6 million nodes and 359.7 million edges. The generated embedding for all central nodes was clustered into 80 different m-states using sklearn function MiniBatchKMeans and a sklearn SGDClassifier was trained using the generated embedding to predict cluster id from GAT embedding vectors. Next, we applied our generalized GBM-GAT to each individual EEL-FISH section graph and classified each RNA into the 80 spatial domains using the trained classifier. RNA molecules with a prediction probability below 50% to a spatial domain were discarded. Finally, we reduced m-states to 42 after merging m-states with gene expression correlation over 85%.

M-cells – Segmentation

To facilitate further analysis, m-states were segmented using a combination of DBscan and QTclustering. First, sklearn function DBscan was applied to each individual m-state using 20 and 10 for the parameters_eps and min_samples, respectively. Next, we re-segmented DBscan outputs, if they had a maximum distance over 50 µm, using QTclustering algorithm with parameters max_radius and min_cluster_size, 20 and 10, respectively. We removed m-cells with less than 10 molecules. Finally, m-cells were kept if the distance to closest cell nuclei was below 50 µm.

Cellular Neighborhoods – Graph Attention Network

Cellular Neighborhoods were obtained after training a second GAT using as input the output of our first GAT. First, m-cells were used to build a graph, similar to our molecule graph, where every m-cell was connected to its neighbors if they were at a distance below the 95th quantile at which molecules are connected to their 5th neighbor. This graph was pruned to remove m-cells connected to less than 2 neighbors (~15-30 µm). The CN-GAT was built using the following architecture: 2-hop neighbor aggregation using GAT2vCONV[85,86] module from DGL, deep residual learning[88] between aggregation layers, and a final linear layer; and the following hyperparameters: a random sampling strategy with 10 and 5 neighbors for 1st and 2nd hop, respectively, 64 units for each hidden layer and 24 units for generated lower dimensional embedding. We trained CN-GAT for 10 epochs using ADAM optimizer with 0.001 learning rate, 512 nodes per mini-batch and used 65% random nodes from all samples. Next, CN-GAT was applied to all m-cells and the generated embeddings were clustered into 10 CNs using MiniBatchKMeans.

M-states annotation

Annotation of m-states was done by combining Tangram predictions and gene enrichment scores. First, we used Tangram to map all m-states gene expression from all samples to GBMap averaged expression of tumor states, immune cells, and oligodendrocytes. Gene expression from each m-state was normalized to 10,000 molecules per m-cell and Tangram was run using the following parameters: mode ='clusters', num_epochs=1000, density_prior='uniform' and lambda_g2=1. The probability of each m-state for each reference cluster was averaged for all samples. The same procedure was repeated to map the m-states to human developmental scRNA-seq data

Gene enrichment was calculated for every m-state and sample using cytograph's enrichment score function Enrichment.

Cellular Neighborhood composition

The composition of each CN was calculated by aggregating the sum of the areas of each m-state class in each CN. The class abundance per CN was normalized and scaled to obtain the relative abundance of classes per CN.

Neighborhood enrichment and graphs

Squidpy functions were used to calculate the enrichment scores for all possible m-state pairs. First, spatial_neighbors function was used to build a neighborhood graph connecting each m-cell to the 10 nearest m-cells below 100 µm distance and used Delaunay triangulation. Then, we used neighborhood_enrichment to obtain the spatial proximity scores for each pair based on permutation test.

We calculated a spatial similarity dendrogram based on each tumor sample enrichment score. First, enrichment score matrices were vectorized and rearranged so that each vector entry corresponded to the enrichment score for the same m-state pair in each sample. Next, each vector was scaled and pairwise cosine distances calculated to build a dendrogram using ward method.

The TAM and GW neighborhood graphs were built using the previously calculated enrichment scores. To build the graph, we used each m-state in each pair as nodes and connected nodes whose enrichment score was above 8.5, to enrich for highly enriched interacting m-states. We used networkx spring_layout to facilitate graph visualization.

Neighborhood in situ enrichment

Neighborhood in situ interaction score was built based on the neighborhood enrichment scores. First, m-cells were binned in a hexagonal grid. Next, a vector field was computed by calculating the average enrichment interaction vector between the central hexagon and its neighboring hexagons. Each enrichment interaction vector was calculated as the original distance vector between each hexagon pair multiplied by the total number of m-state pairs found across hexagons multiplied by their enrichment score. The final enrichment in situ score was obtained by assigning m-cells to the value of their closest vector module.

Pseudotime analysis

Pseudo-time/space analysis was performed using VEGFA expression as latent time variable. First, _correlation_test function, from cellrank, was used to only keep genes that correlated with latent time. Then, m-cells were sorted according to their latent time and latent time was binned into 6-time windows using sextiles to define the start and end of each window. Then, the function rank_genes_groups from scanpy was used to find differentially expressed gene sets between time windows with p-value below 0.005 and log fold above 0.1.

Single m-cell Analysis

To analyze tumor m-states, m-cells were pooled from all samples. Cells with a total number of molecules below and above the 5th and 95th quantiles were removed. Top 500 highly variable genes were selected for downstream processing. We normalized the total number of molecules to l2-norm, performed log transformation and minmax scaling per each individual sample to remove batch effects. PCA was run and sample specific batch effects were corrected using harmony. UMAP dimensionality reduction was computed. Gliosarcoma and myeloid cell pools did reveal significant batch effects; in these cases, total counts were normalized, and log transformed, sample minmax scaling and harmony correction were not required. Clustering was not necessary as m-states were used for further analysis.

Data visualization

We used FISHscale for the spatial visualization of RNA molecules and FISHspace for the visualization of m-cells as polygons. FISHspace software was based on the python package stlearn, with some added functionalities for visualization and pseudotime analysis.

References

1. Stupp, R., Mason, W.P., van den Bent, M.J., Weller, M., Fisher, B., Taphoorn, M.J.B., Belanger, K., Brandes, A.A., Marosi, C., Bogdahn, U., *et al.* (2005). Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. New England Journal of Medicine 352, 987–996. 10.1056/NEJMOA043330/SUPPL_FILE/987SA1.PDF.

2. Brennan, C.W., Verhaak, R.G.W., McKenna, A., Campos, B., Noushmehr, H., Salama, S.R., Zheng, S., Chakravarty, D., Sanborn, J.Z., Berman, S.H., *et al.* (2013). The somatic genomic landscape of glioblastoma. Cell 155, 462. 10.1016/j.cell.2013.09.034.

3. Nobusawa, S., Watanabe, T., Kleihues, P., and Ohgaki, H. (2009). IDH1 mutations as molecular signature and predictive factor of secondary glioblastomas. Clin Cancer Res 15, 6002–6007. 10.1158/1078-0432.CCR-09-0715.

4. Bailey, P., and Cushing, H. (1925). Medulloblastoma cerebelli: a common type of midcerrebellar glioma of childhood. Arch Neurol Psychiatry 14, 192–224. 10.1001/archneurpsyc.1925.02200140055002.

5. Ravi, V.M., Will, P., Kueckelhaus, J., Sun, N., Joseph, K., Salié, H., Vollmer, L., Kuliesiute, U., von Ehr, J., Benotmane, J.K., *et al.* (2022). Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. Cancer Cell 40, 639-655.e13. 10.1016/j.ccell.2022.05.009.

6. Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., *et al.* (2010). Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17, 98–110. 10.1016/j.ccr.2009.12.020.

7. Phillips, H.S., Kharbanda, S., Chen, R., Forrest, W.F., Soriano, R.H., Wu, T.D., Misra, A., Nigro, J.M., Colman, H., Soroceanu, L., *et al.* (2006). Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 9, 157–173. 10.1016/j.ccr.2006.02.019.

8. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., *et al.* (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell 178, 835-849.e21. 10.1016/j.cell.2019.06.024.

9. Couturier, C.P., Ayyadhury, S., Le, P.U., Nadaf, J., Monlong, J., Riva, G., Allache, R., Baig, S., Yan, X., Bourgey, M., *et al.* (2020). Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. Nat Commun 11. 10.1038/S41467-020-17186-5.

10. Greenwald, A.C., Darnell, N.G., Hoefflin, R., Simkin, D., Gonzalez-Castro, L.N., Mount, C., Hirsch, D., Nomura, M., Talpir, T., Kedmi, M., *et al.* (2023). Integrative spatial analysis reveals a multi-layered organization of glioblastoma. bioRxiv, 2023.07.06.547924. 10.1101/2023.07.06.547924.

11. Ruiz-Moreno, C., Salas, S.M., Samuelsson, E., Brandner, S., Kranendonk, M.E.G., Nilsson, M., and Stunnenberg, H.G. (2022). Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma. bioRxiv, 2022.08.27.505439. 10.1101/2022.08.27.505439.

12. Liu, I., Jiang, L., Samuelsson, E.R., Marco Salas, S., Beck, A., Hack, O.A., Jeong, D., Shaw, M.K.L., Englinger, B., LaBelle, J., *et al.* (2022). The landscape of tumor cell states and spatial organization in H3-K27M mutant diffuse midline glioma across age and location. Nature Genetics 2022 54:12 54, 1881–1894. 10.1038/s41588-022-01236-3.

13. Borm, L.E., Mossi Albiach, A., Mannens, C.C.A., Janusauskas, J., Özgün, C., Fernández-García, D., Hodge, R., Castillo, F., Hedin, C.R.H., Villablanca, E.J., *et al.* (2022). Scalable in situ single-cell profiling by electrophoretic capture of mRNA using EEL FISH. Nature Biotechnology 2022 41:2 41, 222–231. 10.1038/s41587-022-01455-3.

14. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 5, 877–879. 10.1038/nmeth.1253.

15. Hamilton, W.L., Ying, R., and Leskovec, J. (2017). Inductive Representation Learning on Large Graphs.

16. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017). Attention Is All You Need.

17. Partel, G., and Wählby, C. (2021). Spage2vec: Unsupervised representation of localized spatial gene expression signatures. FEBS Journal 288, 1859–1870. 10.1111/febs.15572.

18. Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, Å., Zhang, M., *et al.* (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nature Methods 2021 18:11 18, 1352–1362. 10.1038/s41592-021-01264-7.

19. Ruiz-Moreno, C., Salas, S.M., Samuelsson, E., Brandner, S., Kranendonk, M.E.G., Nilsson, M., and Stunnenberg, H.G. (2022). Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma. bioRxiv, 2022.08.27.505439. 10.1101/2022.08.27.505439.

20. Braun, E., Danan-Gotthold, M., Borm, L.E., Vinsland, E., Lee, K.W., Lönnerberg, P., Hu, L., Li, X., He, X., Andrusivová, Ž., *et al.* (2022). Comprehensive cell atlas of the first-trimester developing human brain. bioRxiv, 2022.10.24.513487. 10.1101/2022.10.24.513487.

21. Huang, W., Bhaduri, A., Velmeshev, D., Wang, S., Wang, L., Rottkamp, C.A., Alvarez-Buylla, A., Rowitch, D.H., and Kriegstein, A.R. (2020). Origins and Proliferative States of Human Oligodendrocyte Precursor Cells. Cell 182, 594-608.e11. 10.1016/j.cell.2020.06.027.

22. van Bruggen, D., Pohl, F., Langseth, C.M., Kukanja, P., Lee, H., Albiach, A.M., Kabbe, M., Meijer, M., Linnarsson, S., Hilscher, M.M., *et al.* (2022). Developmental landscape of human forebrain at a single-cell level identifies early waves of oligodendrogenesis. Dev Cell 57, 1421-1436.e5. 10.1016/j.devcel.2022.04.016.

23. Wang, L., Hou, S., and Han, Y.G. (2016). Hedgehog signaling promotes basal progenitor expansion and the growth and folding of the neocortex. Nat Neurosci 19, 888. 10.1038/NN.4307.

24. Phillips, H.S., Kharbanda, S., Chen, R., Forrest, W.F., Soriano, R.H., Wu, T.D., Misra, A., Nigro, J.M., Colman, H., Soroceanu, L., *et al.* (2006). Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 9, 157–173. 10.1016/J.CCR.2006.02.019.

25. Jones, R.C., Karkanias, J., Krasnow, M.A., Pisco, A.O., Quake, S.R., Salzman, J., Yosef, N., Bulthaup, B., Brown, P., Harper, W., *et al.* (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. Science (1979) 376. 10.1126/SCIENCE.ABL4896/SUPPL_FILE/SCIENCE.ABL4896_MDAR_REPRODUCIBILITY_CHECKLIST.PDF.

26. Yang, J., Antin, P., Berx, G., Blanpain, C., Brabletz, T., Bronner, M., Campbell, K., Cano, A., Casanova, J., Christofori, G., *et al.* (2020). Guidelines and definitions for research on epithelial–mesenchymal transition. Nature Reviews Molecular Cell Biology 2020 21:6 21, 341–352. 10.1038/s41580-020-0237-9.

27. Kozak, K.R., Mahadevan, A., and Moody, J.S. (2009). Adult gliosarcoma: epidemiology, natural history, and factors associated with outcome. Neuro Oncol 11, 183. 10.1215/15228517-2008-076.

28. Chanoch-Myers, R., Wider, A., Suva, M.L., and Tirosh, I. (2022). Elucidating the diversity of malignant mesenchymal states in glioblastoma by integrative analysis. Genome Med 14, 1–12. 10.1186/S13073-022-01109-8/FIGURES/4.

29. Laurent, G.S., Seilheimer, B., Tackett, M., Zhou, J., Shtokalo, D., Vyatkin, Y., Ri, M., Toma, I., Jones, D., and McCaffrey, T.A. (2017). Deep sequencing transcriptome analysis of murine wound healing: Effects of a multicomponent, multitarget natural product Therapy-Tr14. Front Mol Biosci 4, 254339. 10.3389/FMOLB.2017.00057/BIBTEX.

30. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst 1, 417–425. 10.1016/j.cels.2015.12.004.

31. Percio, S., Coltella, N., Grisanti, S., Bernardi, R., and Pattini, L. (2014). A HIF-1 network reveals characteristics of epithelial-mesenchymal transition in acute promyelocytic leukemia. Genome Med 6, 1–12. 10.1186/S13073-014-0084-4/FIGURES/5.

32. Yao, Y., Jumabay, M., Wang, A., and Boström, K.I. (2011). Matrix Gla protein deficiency causes arteriovenous malformations in mice. J Clin Invest 121, 2993–3004. 10.1172/JCI57567.

33. Boström, K., Zebbondj, A.F., Yao, Y., Lin, T.S., and Torres, A. (2004). Matrix GLA Protein Stimulates VEGF Expression through Increased Transforming Growth Factor-β1 Activity in Endothelial Cells. Journal of Biological Chemistry 279, 52904–52913. 10.1074/JBC.M406868200.

34. Bednarek, N., Svedin, P., Garnotel, R., Favrais, G., Loron, G., Schwendiman, L., Hagberg, H., Morville, P., Mallard, C., and Gressens, P. (2011). Increased MMP-9 and TIMP-1 in mouse neonatal brain and plasma and in human neonatal plasma after hypoxia–ischemia: a potential marker of neonatal encephalopathy. Pediatric Research 2012 71:1 71, 63–70. 10.1038/pr.2011.3.

35. Suvà, M.L., and Tirosh, I. (2020). The Glioma Stem Cell Model in the Era of Single-Cell Genomics. 10.1016/j.ccell.2020.04.001.

36. Ramón y Cajal, S. (1895). Algunas conjeturas sobre el mecanismo anatómico de la ideación, asociación y atención. (No Title).

37. Friebel, E., Kapolou, K., Unger, S., Núñez, N.G., Utz, S., Rushing, E.J., Regli, L., Weller, M., Greter, M., Tugues, S., *et al.* (2020). Single-Cell Mapping of Human Brain Cancer Reveals Tumor-Specific Instruction of Tissue-Invading Leukocytes. Cell 181, 1626-1642.e20. 10.1016/J.CELL.2020.04.055.

38. Klemm, F., Maas, R.R., Bowman, R.L., Kornete, M., Soukup, K., Nassiri, S., Brouland, J.P., Iacobuzio-Donahue, C.A., Brennan, C., Tabar, V., *et al.* (2020). Interrogation of the Microenvironmental Landscape in Brain Tumors Reveals Disease-Specific Alterations of Immune Cells. Cell 181, 1643-1660.e17. 10.1016/J.CELL.2020.05.007.

39. Wei, J., Chen, Z., Hu, M., He, Z., Jiang, D., Long, J., and Du, H. (2021). Characterizing Intercellular Communication of Pan-Cancer Reveals SPP1+ Tumor-Associated Macrophage Expanded in Hypoxia and Promoting Cancer Malignancy Through Single-Cell RNA-Seq Data. Front Cell Dev Biol 9, 749210. 10.3389/FCELL.2021.749210/BIBTEX.

40. Murphy, B.J., Kimura, T., Sato, B.G., Shi, Y., and Andrews, G.K. (2008). Metallothionein induction by hypoxia involves cooperative interactions between metal-responsive transcription factor-1 and hypoxia-inducible transcription factor-1α. Molecular Cancer Research 6, 483–490. 10.1158/1541-7786.MCR-07-0341.

41. Hara, T., Chanoch-Myers, R., Mathewson, N.D., Regev, A., Suvà, M.L., Myskiw, C., Atta, L., Bussema, L., Eichhorn, S.W., Greenwald, A.C., *et al.* (2021). In brief Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. Cancer Cell 39, 779-792.e11. 10.1016/j.ccell.2021.05.002.

42. Zamanian, J.L., Xu, L., Foo, L.C., Nouri, N., Zhou, L., Giffard, R.G., and Barres, B.A. (2012). Genomic Analysis of Reactive Astrogliosis. The Journal of Neuroscience 32, 6391. 10.1523/JNEUROSCI.6221-11.2012.

43. Yu, M., Chang, Y., Zhai, Y., Pang, B., Wang, P., Li, G., Jiang, T., Zeng, F., Cheng, W., Meng, X., *et al.* (2023). TREM2 is associated with tumor immunity and implies poor prognosis in glioma OPEN ACCESS EDITED BY. 10.3389/fimmu.2022.1089266.

44. Mulder, K., Patel, A.A., Kong, W.T., Piot, C., Halitzki, E., Dunsmore, G., Khalilnezhad, S., Irac, S.E., Dubuisson, A., Chevrier, M., *et al.* (2021). Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. Immunity 54, 1883-1900.e5. 10.1016/j.immuni.2021.07.007.

45. Sattiraju, A., Kang, S., Giotti, B., Chen, Z., Marallano, V.J., Brusco, C., Ramakrishnan, A., Shen, L., Tsankov, A.M., Hambardzumyan, D., *et al.* (2023). Hypoxic niches attract and sequester tumor-associated macrophages and cytotoxic T cells and reprogram them for immunosuppression. Immunity 0. 10.1016/J.IMMUNI.2023.06.017.

46. Bhat, K.P.L., Balasubramaniyan, V., Vaillant, B., Ezhilarasan, R., Hummelink, K., Hollingsworth, F., Wani, K., Heathcock, L., James, J.D., Goodman, L.D., *et al.* (2013). Mesenchymal Differentiation Mediated by NF-κB Promotes Radiation Resistance in Glioblastoma. Cancer Cell 24, 331–346. 10.1016/j.ccr.2013.08.001.

47. Huang, B., Deora, A.B., He, K.L., Chen, K., Sui, G., Jacovina, A.T., Almeida, D., Hong, P., Burgman, P., and Hajjar, K.A. (2011). Hypoxia-inducible factor-1 drives annexin A2 system-mediated perivascular fibrin clearance in oxygen-induced retinopathy in mice. Blood 118, 2918. 10.1182/BLOOD-2011-03-341214.

48. Chen, J., Cui, Z., Yang, S., Wu, C., Li, W., Bao, G., Xu, G., Sun, Y., Wang, L., and Zhang, J. (2017). The upregulation of annexin A2 after spinal cord injury in rats may have implication for astrocyte proliferation. Neuropeptides 61, 67–76. 10.1016/J.NPEP.2016.10.007.

49. Man, J., Yu, X., Bebek, G., Bao, S., Correspondence, J.S.Y., Huang, H., Zhou, W., Xiang, C., Huang, H., Miele, L., *et al.* (2018). Hypoxic Induction of Vasorin Regulates Notch1 Turnover to Maintain Glioma Stem-like Cells In Brief Cell Stem Cell Hypoxic Induction of Vasorin Regulates Notch1 Turnover to Maintain Glioma Stem-like Cells. Stem Cell 22, 104-118.e6. 10.1016/j.stem.2017.10.005.

50. Dong, C., Zhang, J., Fang, S., and Liu, F. (2020). IGFBP5 increases cell invasion and inhibits cell proliferation by EMT and Akt signaling pathway in Glioblastoma multiforme cells. Cell Div 15, 1–9. 10.1186/S13008-020-00061-6/FIGURES/5.

51. Walter, H.J., Berry, M., Hill, D.J., and Logan, A. (1997). Spatial and Temporal Changes in the Insulin-Like Growth Factor (IGF) Axis Indicate Autocrine/Paracrine Actions of IGF-I within Wounds of the Rat Brain. Endocrinology 138, 3024–3034. 10.1210/ENDO.138.7.5284.

52. Zimmermann, J., Krauthausen, M., Hofer, M.J., Heneka, M.T., Campbell, I.L., and Müller, M. (2013). CNS-Targeted Production of IL-17A Induces Glial Activation, Microvascular Pathology and Enhances the Neuroinflammatory Response to Systemic Endotoxemia. PLoS One 8, e57307. 10.1371/JOURNAL.PONE.0057307.

53. Kubala, M.H., Punj, V., Placencio-Hickok, V.R., Fang, H., Fernandez, G.E., Sposto, R., and DeClerck, Y.A. (2018).

Plasminogen activator inhibitor-1 promotes the recruitment and polarization of macrophages in cancer. Cell Rep 25, 2177. 10.1016/J.CELREP.2018.10.082.

54. Saidak, Z., Soudet, S., Lottin, M., Salle, V., Sevestre, M.A., Clatot, F., and Galmiche, A. (2021). A pan-cancer analysis of the human tumor coagulome and its link to the tumor immune microenvironment. Cancer Immunology, Immunotherapy 70, 923–933. 10.1007/S00262-020-02739-W/FIGURES/6.

55. Eliceiri, B.P., Paul, R., Schwartzberg, P.L., Hood, J.D., Leng, J., and Cheresh, D.A. (1999). Selective requirement for Src kinases during VEGF-induced angiogenesis and vascular permeability. Mol Cell 4, 915–924. 10.1016/S1097-2765(00)80221-X.

56. Girgrah, N., Letarte, M., Becker, L.E., Cruz, T.F., Theriault, E., and Moscarello, M.A. (1991). Localization of the CD44 Glycoprotein to Fibrous Astrocytes in Normal White Matter and to Reactive Astrocytes in Active Lesions in Multiple Sclerosis. J Neuropathol Exp Neurol 50, 779–792. 10.1097/00005072-199111000-00009.

57. Dzyubenko, E., Manrique-Castano, D., Pillath-Eilers, M., Vasileiadou, P., Reinhard, J., Faissner, A., and Hermann, D.M. (2022). Tenascin-C restricts reactive astrogliosis in the ischemic brain. Matrix Biology 110, 1–15. 10.1016/J.MATBIO.2022.04.003.

58. Dalahmah, O. Al, Sosunov, A.A., Sun, Y., Liu, Y., Akpan, N., Madden, N., Connolly, E.S., Troy, C.M., II, G.M.M., and Goldman, J.E. (2023). The Matrix Receptor CD44 Is Present in Astrocytes Throughout the Human CNS and Accumulates in Hypoxia and Seizures. 10.20944/PREPRINTS202303.0444.V1.

59. Board, W.C. of T.E. (2022). Central Nervous System Tumours (WORLD HEALTH ORGANIZATION).

60. Etchevers, H.C., Vincent, C., Le Douarin, N.M., and Couly, G.F. (2001). The cephalic neural crest provides pericytes and smooth muscle cells to all blood vessels of the face and forebrain. Development 128, 1059–1068. 10.1242/DEV.128.7.1059.

61. Zarbalis, K., Siegenthaler, J.A., Choe, Y., May, S.R., Peterson, A.S., and Pleasure, S.J. (2007). Cortical dysplasia and skull defects in mice with a Foxc1 allele reveal the role of meningeal differentiation in regulating cortical development. Proc Natl Acad Sci U S A 104, 14002. 10.1073/PNAS.0702618104.

62. Gurtner, G.C., Werner, S., Barrandon, Y., and Longaker, M.T. (2008). Wound repair and regeneration. Nature 2008 453:7193 453, 314–321. 10.1038/nature07039.

63. Sofroniew, M. V. (2005). Reactive astrocytes in neural repair and protection. Neuroscientist 11, 400–407. 10.1177/1073858405278321.

64. Alonso, G. (2005). NG2 proteoglycan-expressing cells of the adult rat brain: possible involvement in the formation of glial scar astrocytes following stab wound. Glia 49, 318–338. 10.1002/GLIA.20121.

65. Demirci, Y., Heger, G., Katkat, E., Papatheodorou, I., Brazma, A., and Ozhan, G. (2022). Brain Regeneration Resembles Brain Cancer at Its Early Wound Healing Stage and Diverges From Cancer Later at Its Proliferation and Differentiation Stages. Front Cell Dev Biol 10. 10.3389/FCELL.2022.813314/FULL.

66. Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol 131, 803–820. 10.1007/S00401-016-1545-1/TABLES/5.

67. Gao, Z., Zhu, Q., Zhang, Y., Zhao, Y., Cai, L., Shields, C.B., and Cai, J. (2013). Reciprocal modulation between microglia and astrocyte in reactive gliosis following the CNS injury. Mol Neurobiol 48, 690. 10.1007/S12035-013-8460-4.

68. Chen, P., Zhao, D., Li, J., Liang, X., Li, J., Chang, A., Henry, V.K., Lan, Z., Spring, D.J., Rao, G., *et al.* (2019). Symbiotic Macrophage-Glioma Cell Interactions Reveal Synthetic Lethality in PTEN-Null Glioma. Cancer Cell 35, 868-884.e6. 10.1016/J.CCELL.2019.05.003.

69. Wang, Q., Hu, B., Hu, X., Kim, H., Squatrito, M., Scarpace, L., deCarvalho, A.C., Lyu, S., Li, P., Li, Y., *et al.* (2017). Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell 32, 42-56.e6. 10.1016/J.CCELL.2017.06.003.

70. Zhu, Y., Romero, M.I., Ghosh, P., Ye, Z., Charnay, P., Rushing, E.J., Marth, J.D., and Parada, L.F. (2001). Ablation of NF1 function in neurons induces abnormal development of cerebral cortex and reactive gliosis in the brain. Genes Dev 15, 859–876. 10.1101/GAD.862101.

71. Aird, W.C. (2005). Coagulation. Crit Care Med 33. 10.1097/01.CCM.0000186267.66402.24.

72. Eddleston, M., De La Torre, J.C., Oldstone, M.B.A., Loskutoff, D.J., Edgington, T.S., and Mackman, N. (1993). Astrocytes are the primary source of tissue factor in the murine central nervous system. A role for astrocytes in cerebral hemostasis. J Clin Invest 92, 349–358. 10.1172/JCI116573.

73. Bárdos, H., Molnár, P., Csécsei, G., and Ádány, R. (1996). Fibrin deposition in primary and metastatic human brain tumours. Blood Coagul Fibrinolysis 7, 536–548. 10.1097/00001721-199607000-00005.

74. Clark, R.A.F. (2001). Fibrin and Wound Healing. Ann N Y Acad Sci 936, 355–367. 10.1111/J.1749-6632.2001.TB03522.X.

75. Dvorak, H.F. (2015). Tumors: wounds that do not heal-redux. Cancer Immunol Res 3, 1–11. 10.1158/2326-6066.CIR-14-0209.

76. Mossi Albiach, A. (2022). Oligopy: probe design for smFISH. https://github.com/linnarsson-lab/oligopy.

77. Moffitt, J.R., and Zhuang, X. (2016). RNA Imaging with Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH). Methods Enzymol 572, 1–49. 10.1016/BS.MIE.2016.03.020.

78. Borm, L.E. (2022). ROBOFISH: software to control automated microfluidics and imaging. https://github.com/linnarsson-lab/ROBOFISH.

79. Borm, L.E. (2022). ROBOFISH building instructions. https://www.protocols.io/view/robofish-construction-yxmvmxyj6l3p/v1.

80. Kvedaraite, E., Milne, P., Khalilnezhad, A., Chevrier, M., Sethi, R., Lee, H.K., Hagey, D.W., von Bahr Greenwood, T., Mouratidou, N., Jädersten, M., *et al.* (2022). Notch-dependent cooperativity between myeloid lineages promotes Langerhans cell histiocytosis pathology. Sci Immunol 7. 10.1126/SCIIMMUNOL.ADD3330.

81. Kvedaraite, E., Hertwig, L., Sinha, I., Ponzetta, A., Myrberg, I.H., Lourda, M., Dzidic, M., Akber, M., Klingström, J., Folkesson, E., *et al.* (2021). Major alterations in the mononuclear phagocyte landscape associated with COVID-19 severity. Proc Natl Acad Sci U S A 118. 10.1073/PNAS.2018587118/-/DCSUPPLEMENTAL.

82. Codeluppi, S., Borm, L.E., and Mossi Albiach, A. (2022). pysmFISH: image analysis software for multiplexed smFISH. https://github.com/linnarsson-lab/pysmFISH.

83. Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2020). Cellpose: a generalist algorithm for cellular segmentation. Nature Methods 2020 18:1 18, 100–106. 10.1038/s41592-020-01018-x.

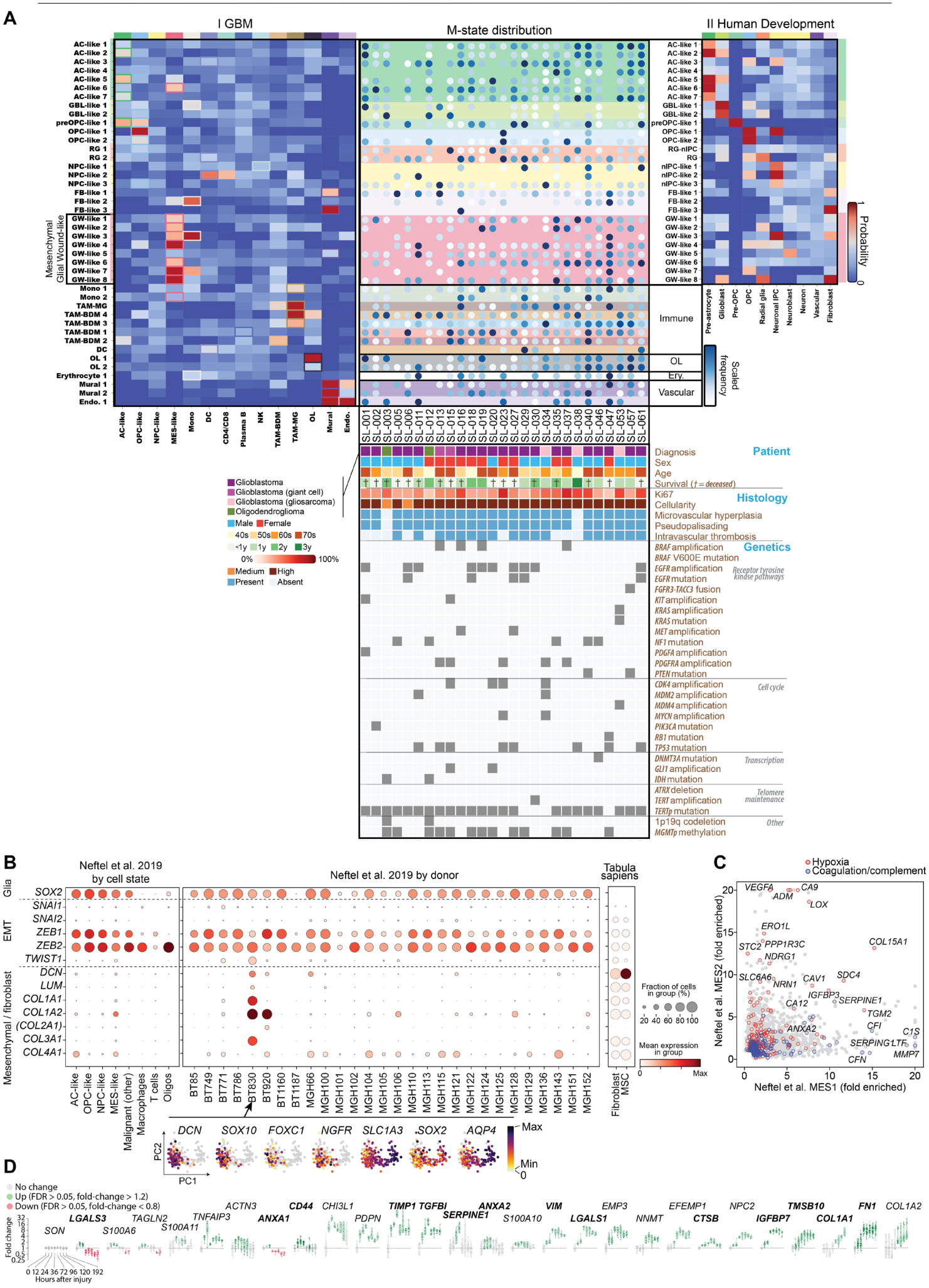
84. Rocklin, M. (2015). Dask: Parallel Computation with Blocked algorithms and Task Scheduling. PROC. OF THE 14th PYTHON IN SCIENCE CONF.

85. Hamilton, W.L., Ying, R., and Leskovec, J. Inductive Representation Learning on Large Graphs.

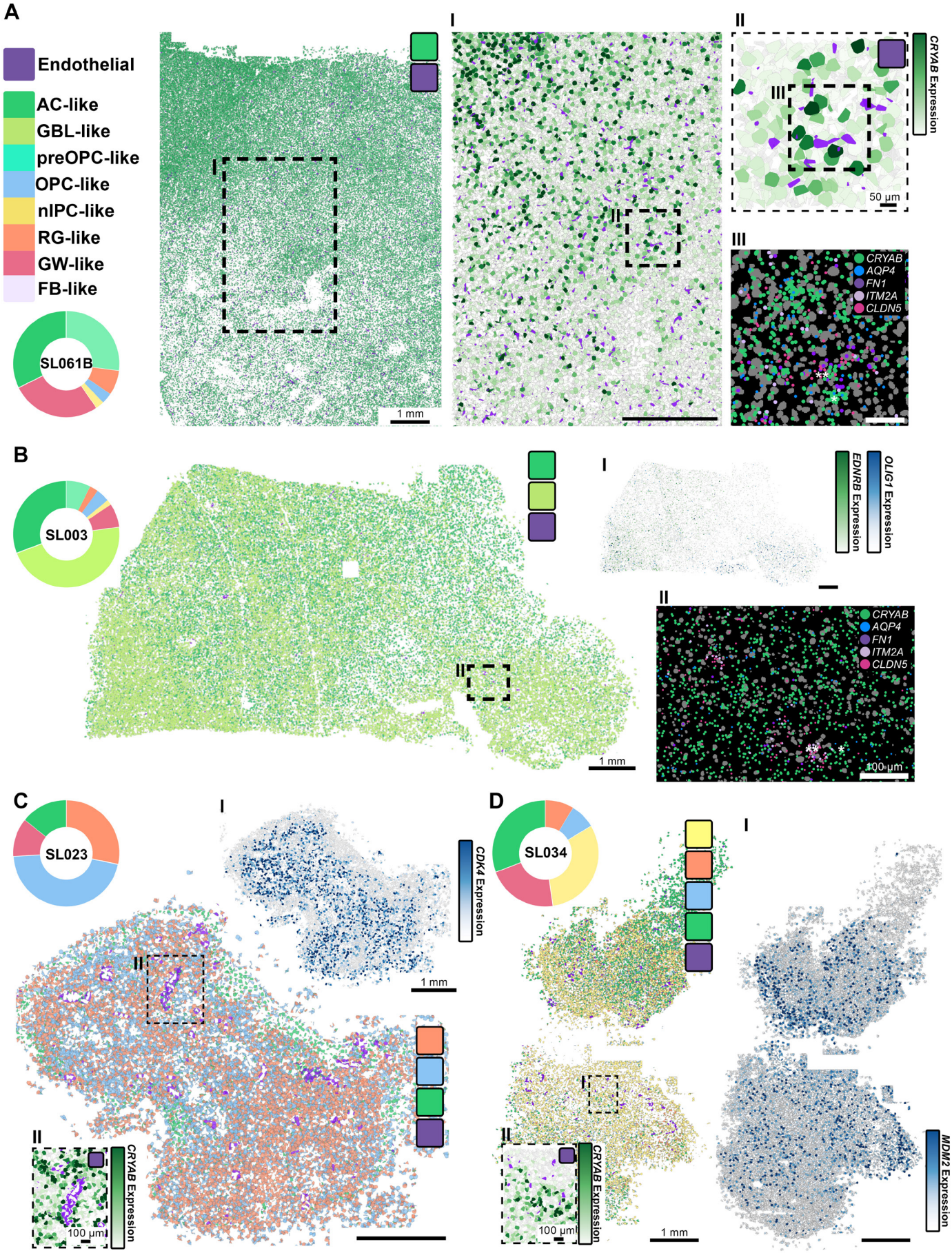
86. Brody, S., Alon, U., and Yahav, E. HOW ATTENTIVE ARE GRAPH ATTENTION NETWORKS?

87. Kingma, D.P., and Ba, J.L. (2014). Adam: A Method for Stochastic Optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

88. He, K., Zhang, X., Ren, S., and Sun, J. Deep Residual Learning for Image Recognition.

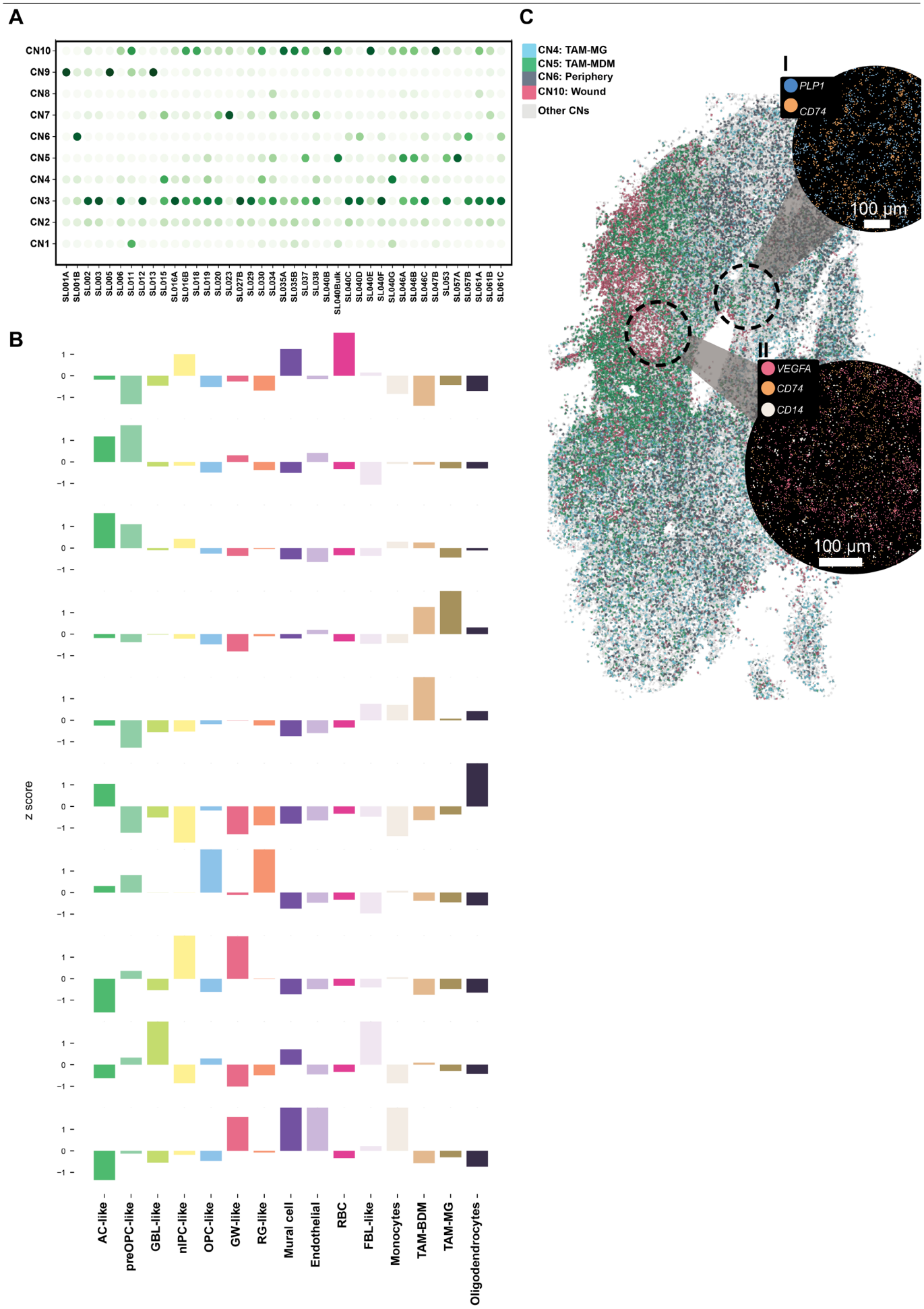

A
I GBM
M-state distribution
II Human Development
Mesenchymal
Glial Wound-like
Immune
OL
Ery.
Vascular
Probability
Scaled frequency
Pre-astrocyte
Glioblast
Pre-OPC
OPC
Radial glia
Neuronal IPC
Neuroblast
Neuron
Vascular
Fibroblast
Patient
Histology
Genetics
Diagnosis
Sex
Age
Survival († = deceased)
Ki67
Cellularity
Microvascular hyperplasia
Pseudopalisading
Intravascular thrombosis
BRAF amplification
BRAF V600E mutation
EGFR amplification
EGFR mutation
FGFR3-TACC3 fusion
KIT amplification
KRAS amplification
KRAS mutation
MET amplification
NF1 mutation
PDGFA amplification
PDGFRA amplification
PTEN mutation
Receptor tyrosine kinase pathways
CDK4 amplification
MDM2 amplification
MDM4 amplification
MYCN amplification
PIK3CA mutation
RB1 mutation
TP53 mutation
Cell cycle
DNMT3A mutation
GLI1 amplification
IDH mutation
Transcription
ATRX deletion
TERT amplification
TERTp mutation
Telomere maintenance
1p19q codeletion
MGMTp methylation
Other
Glioblastoma
Glioblastoma (giant cell)
Glioblastoma (gliosarcoma)
Oligodendroglioma
Male
Female
40s 50s 60s 70s
<1y 1y 2y 3y
0% 100%
Medium
High
Present
Absent
B
Neftel et al. 2019 by cell state
Neftel et al. 2019 by donor
Tabula sapiens
Glia
EMT
Mesenchymal / fibroblast
SOX2
SNAI1
SNAI2
ZEB1
ZEB2
TWIST1
DCN
LUM
COL1A1
COL1A2
(COL2A1)
COL3A1
COL4A1
Fraction of cells in group (%)
Mean expression in group
Max
C
Hypoxia
Coagulation/complement
Neftel et al. MES2 (fold enriched)
Neftel et al. MES1 (fold enriched)
VEGFA
CA9
ADM
LOX
ERO1L
STC2
PPP1R3C
NDRG1
COL15A1
SLC6A6
NRN1
CAV1
SDC4
IGFBP3
SERPINE1
CA12
TGM2
CFI
ANXA2
C1S
SERPING1
LTF
CFN
MMP7
D
No change
Up (FDR > 0.05, fold-change > 1.2)
Down (FDR > 0.05, fold-change < 0.8)
Fold change
Hours after injury
LGALS3
SON
S100A6
TAGLN2
TNFAIP3
S100A11
ACTN3
ANXA1
CD44
CHI3L1
PDPN
TIMP1
TGFBI
SERPINE1
ANXA2
S100A10
VIM
LGALS1
EMP3
NNMT
EFEMP1
CTSB
NPC2
IGFBP7
TMSB10
COL1A1
FN1
COL1A2

Supplementary Figure 1. Tangram mapping of m-states against GBM and human fetal development cell types, m-state distribution across samples, and pathology report summary. (A) Tangram mapping of m-states to reference scRNA-seq GBM cell types[19] (left-most panel). Tangram fine mapping of m-states to reference scRNA-seq cell types from human fetal development[20] (right-most panel). Dot plot of the distribution of m-states in samples from every patient (middle panel). Pathology report summary of tumor features, patient information, and tumor mutational profiles (bottom-panel). **(B)** Expression of relevant marker genes in previous datasets, Neftel et al[8]. and Tabula sapiens[25], as indicated. Mean expression is not necessarily comparable between datasets. Parentheses around *COL2A1* indicate that this gene is expected to be absent in mesenchymal cells. Arrow, putative gliosarcoma tumor. EMT, epithelial to mesenchymal transition. MSC, mesenchymal stem cell. At bottom, PCA projections of the cells from sample BT830, showing the expression of selected genes. Zeros are plotted in grey. **(C)** Genes enriched in Neftel et al. MES1 and MES2 cell states[8], after excluding putative gliosarcoma samples. Red, MSigDb hallmark genes related to hypoxia. Blue, MSigDb hallmark genes related to coagulation and complement activation. **(D)** Expression of MES-core genes[28] in a mouse skin wound healing timecourse[29] (RNA sequencing; n=6 or 7 animals per timepoint). Grey, no significant change from zero timepoint; green, significantly up-regulated (FDR < 0.05, fold-change > 1.2); red, significantly down-regulated (FDR < 0.05, fold-change < 0.8). The SON splicing co-factor is included for reference, being the gene with the lowest coefficient of variation in the dataset. Genes in bold were included in the EEL probe set.

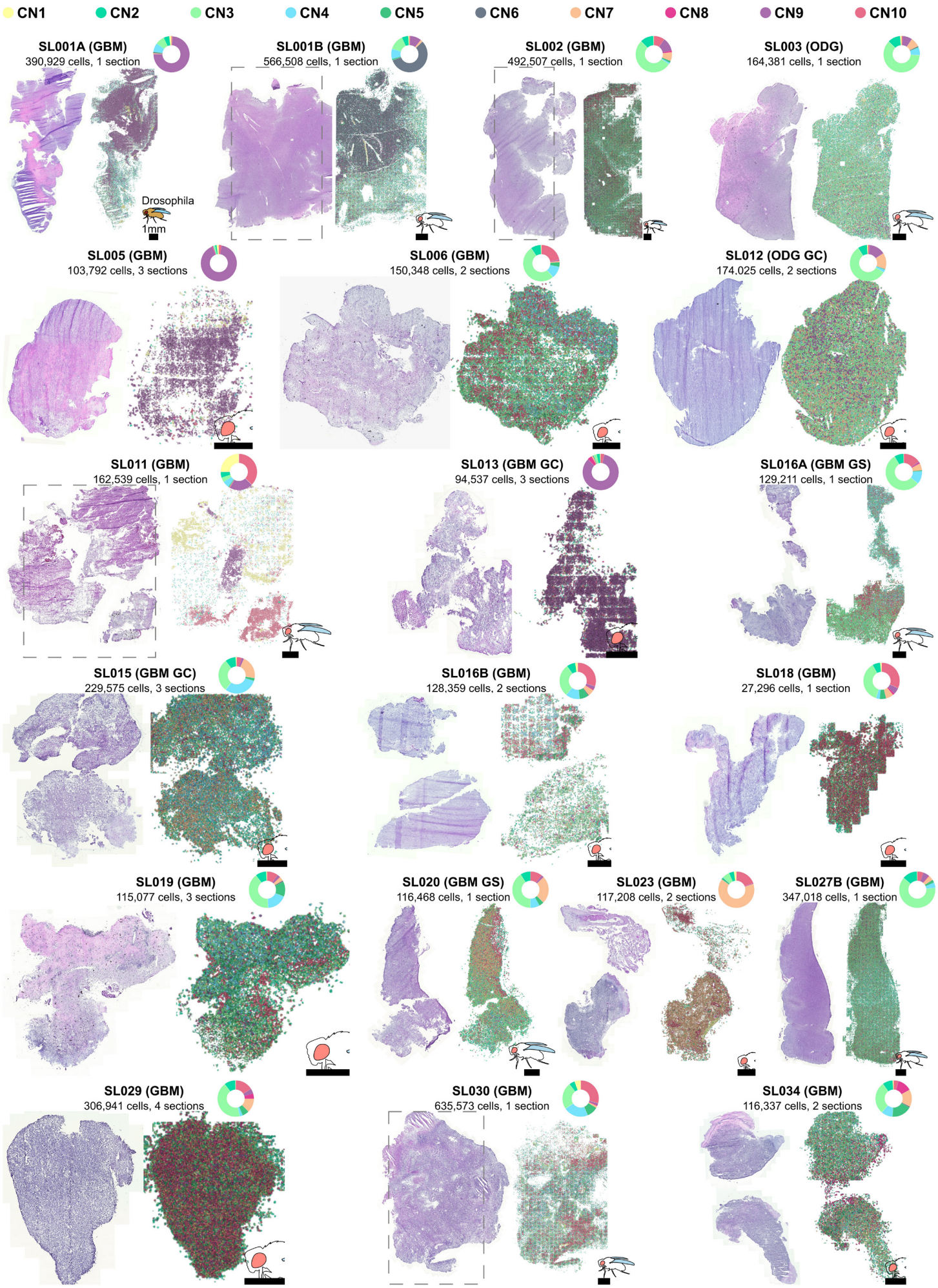


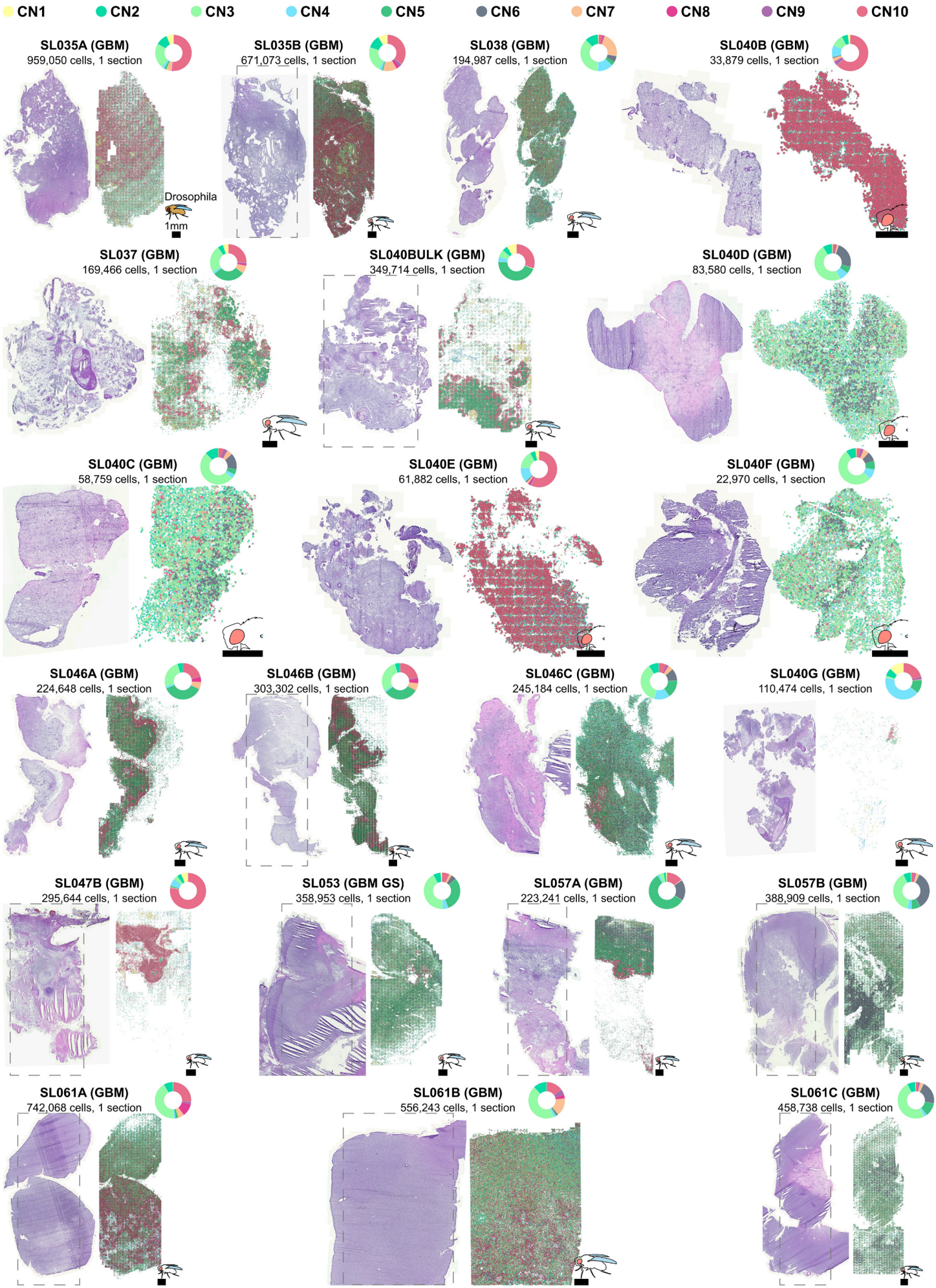
Supplementary Figure 2. AC-like and endothelial co-localization in GBM and oligodendroglioma. (A) Spatial distribution of AC-like 3 and endothelial m-cells in sample SL035B; pie chart shows the proportions of constituent m-states. I: Co-localization of *CRYAB* expression and endothelial m-cells. II: Close-up view of region II showing *CRYAB* expression and endothelial m-cells. III: EEL-FISH signal and nuclei showing the co-expression of AC-like marker genes CRYAB and *AQP4* (*) near endothelial cell marker genes *FN1*, *CLDN5*, and *ITM2A* (**). Gray polygons represent m-cells with zero expression. **(B)** Spatial distribution of AC-like 3 and endothelial m-cells in oligodendroglioma sample SL003; pie chart shows the proportions of constituent m-states. I: Expression of *OLIG1* and *EDNRB* in oligodendroglioma. Gray polygons represent m-cells with zero expression. II: EEL-FISH signal and nuclei showing the expression of AC-like marker genes *CRYAB* and *AQP4* (*) near endothe-

lial marker genes FN1, CLDN5, and ITM2A (**). **(C)** Spatial distribution of OPC/RG/AC-like and endothelial m-cells in sample SL023; pie chart shows the proportions of constituent m-states. I: Expression of *CDK4* in sample SL023. Gray polygons represent m-cells with zero expression. II: Region highlighting *CRYAB* expression is highest near endothelial m-cells. (D) Spatial distribution of nIPC/RG/OPC/AC-like and endothelial m-cells in sample SL034; pie chart shows the proportions of constituent m-states. I: Expression of *MDM2* in sample SL034. Gray polygons represent m-cells with zero expression. II: Region highlighting *CRYAB* expression is highest near endothelial m-cells.

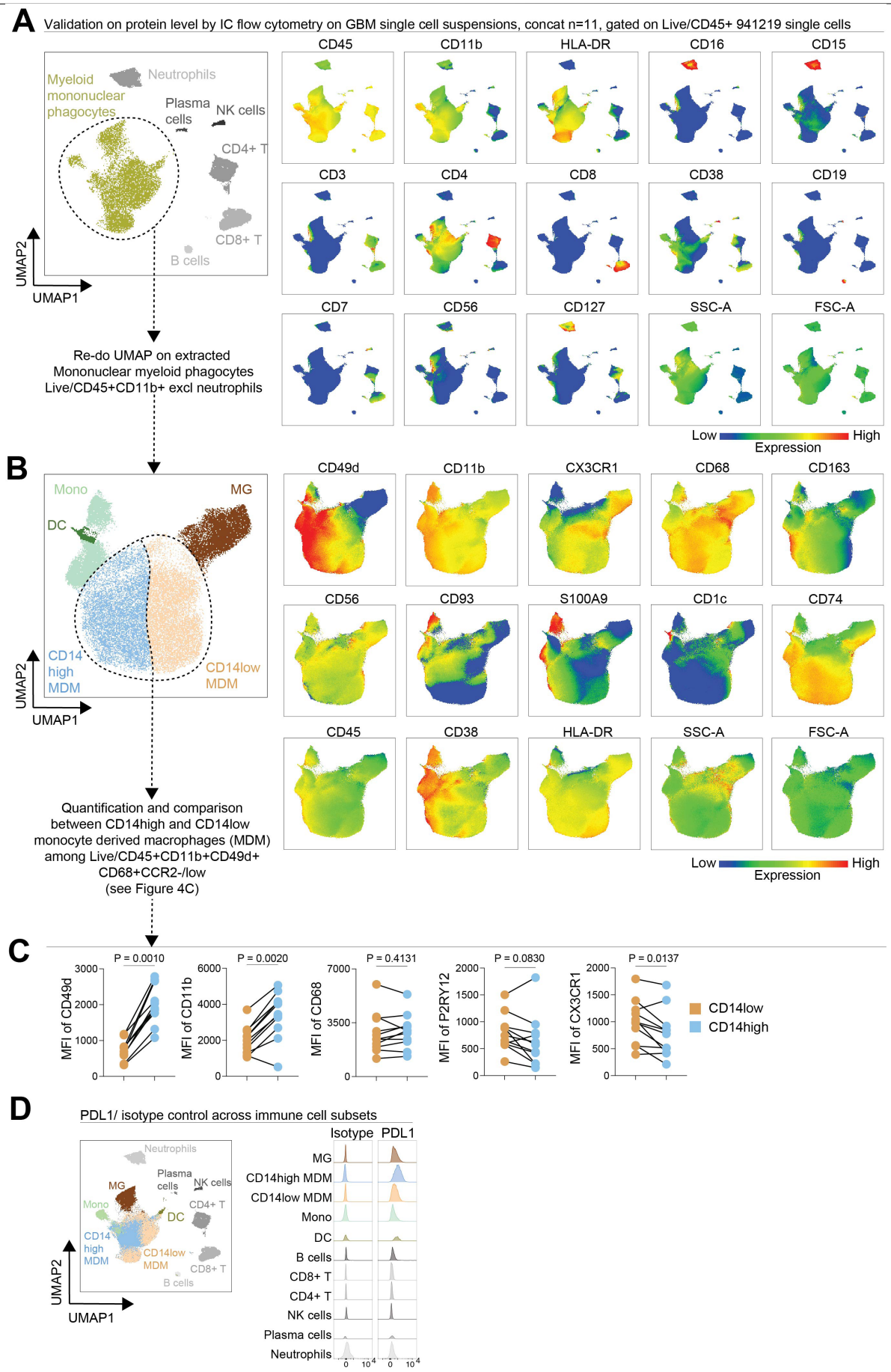


Supplementary Figure 3. Cellular Neighborhood composition and distribution across all samples. (A) Dot plot of scaled CN distribution across patient samples. (B) Bar plot of scaled m-state composition of each CN. (C) Spatial distribution of CN4 (TAM-MG), CN5 (TAM-MDM), CN6 (Periphery) and CN10 in tumor sample SL046C. I: EEL-FISH signal of *PLP1* and *CD74* in the peripheral areas dominated by CN4 (TAM-MG) and CN6 (Periphery). II: Expression of *CD14*, *CD74*, and *VEGFA* in a region dominated by CN10 (Wound).

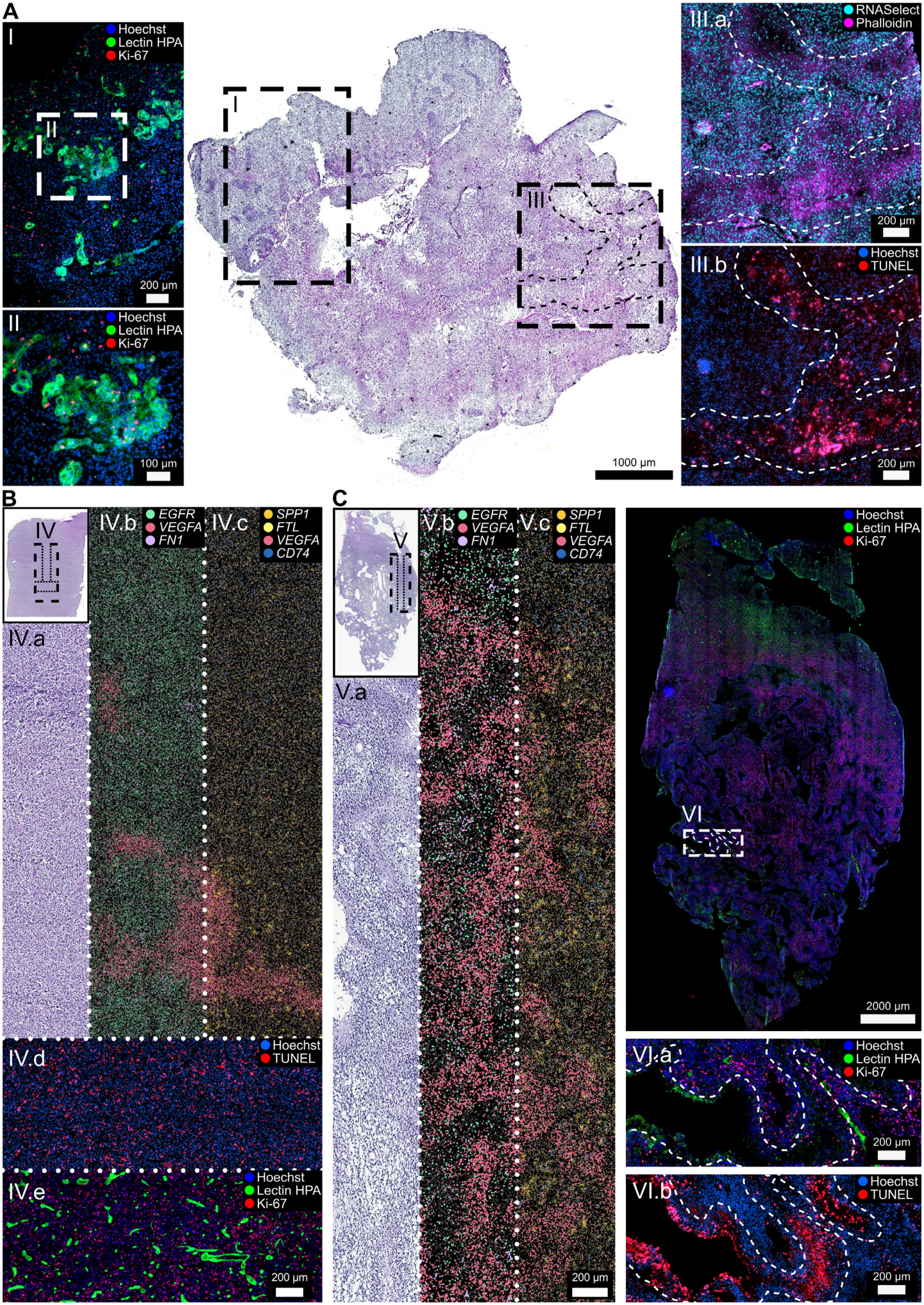

CN1
CN2
CN3
CN4
CN5
CN6
CN7
CN8
CN9
CN10
SL001A (GBM)
390,929 cells, 1 section
SL001B (GBM)
566,508 cells, 1 section
SL002 (GBM)
492,507 cells, 1 section
SL003 (ODG)
164,381 cells, 1 section
Drosophila
1mm
SL005 (GBM)
103,792 cells, 3 sections
SL006 (GBM)
150,348 cells, 2 sections
SL012 (ODG GC)
174,025 cells, 2 sections
SL011 (GBM)
162,539 cells, 1 section
SL013 (GBM GC)
94,537 cells, 3 sections
SL016A (GBM GS)
129,211 cells, 1 section
SL015 (GBM GC)
229,575 cells, 3 sections
SL016B (GBM)
128,359 cells, 2 sections
SL018 (GBM)
27,296 cells, 1 section
SL019 (GBM)
115,077 cells, 3 sections
SL020 (GBM GS)
116,468 cells, 1 section
SL023 (GBM)
117,208 cells, 2 sections
SL027B (GBM)
347,018 cells, 1 section
SL029 (GBM)
306,941 cells, 4 sections
SL030 (GBM)
635,573 cells, 1 section
SL034 (GBM)
116,337 cells, 2 sections


CN1
CN2
CN3
CN4
CN5
CN6
CN7
CN8
CN9
CN10
SL035A (GBM)
959,050 cells, 1 section
Drosophila
1mm
SL035B (GBM)
671,073 cells, 1 section
SL038 (GBM)
194,987 cells, 1 section
SL040B (GBM)
33,879 cells, 1 section
SL037 (GBM)
169,466 cells, 1 section
SL040BULK (GBM)
349,714 cells, 1 section
SL040D (GBM)
83,580 cells, 1 section
SL040C (GBM)
58,759 cells, 1 section
SL040E (GBM)
61,882 cells, 1 section
SL040F (GBM)
22,970 cells, 1 section
SL046A (GBM)
224,648 cells, 1 section
SL046B (GBM)
303,302 cells, 1 section
SL046C (GBM)
245,184 cells, 1 section
SL040G (GBM)
110,474 cells, 1 section
SL047B (GBM)
295,644 cells, 1 section
SL053 (GBM GS)
358,953 cells, 1 section
SL057A (GBM)
223,241 cells, 1 section
SL057B (GBM)
388,909 cells, 1 section
SL061A (GBM)
742,068 cells, 1 section
SL061B (GBM)
556,243 cells, 1 section
SL061C (GBM)
458,738 cells, 1 section

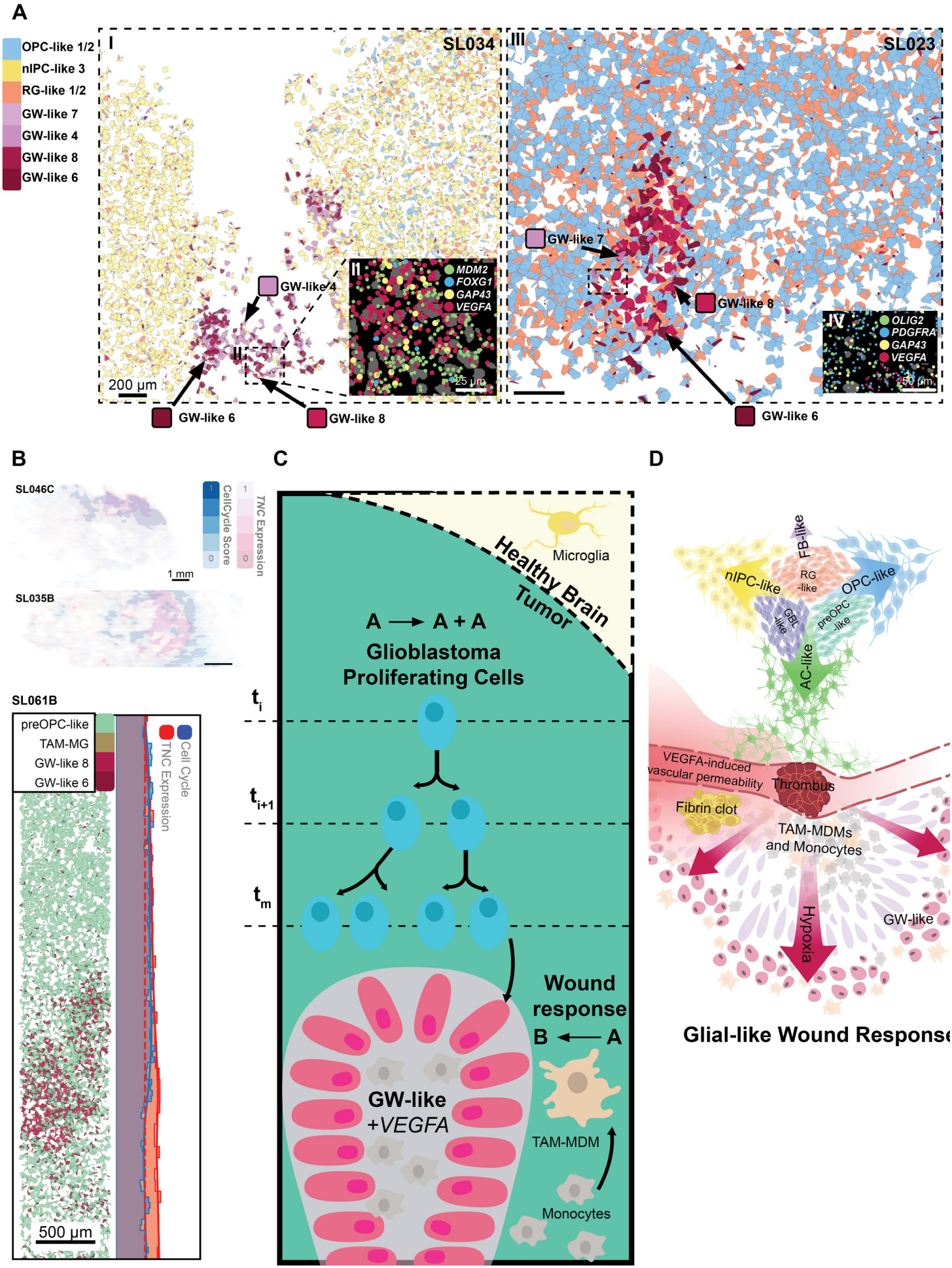
Supplementary Figure 4. Hematoxylin and eosin staining and Cellular Neighborhood composition of all samples. Each sample panel contains: H&E staining (left subpanel), CNs (right subpanel), and information on the total number of sections analyzed per sample (one shown) and the total number of nuclei scanned with EEL-FISH for all sections. Drosophila for size reference. Pie charts show the proportions of constituent CNs (GBM = glioblastoma, ODG = oligodendroglioma, GC = giant cell, GS = gliosarcoma).



Supplementary Figure 5. Validation of tumor associated macrophages on protein level using high-parameter flow cytometry. (A) Flow cytometry data from concatenated live immune cells (CD45+) from 11 tumors in a UMAP plot with gated myeloid mononuclear phagocytes shown in yellow (left); cell identities established based on marker expression (right). **(B)** Myeloid mononuclear phagocytes extracted from A in a UMAP plot (far right; from Figure 4B) showing monocytes (Mono), dendritic cells (DC), microglia (MG) and gated CD14^{high} (blue) and CD14^{low} (orange) monocyte-derived macrophages (MDM) (right); cell identities established based on marker expression (left). **(C)** Comparison of indicated marker mean fluorescence intensity (MFI) between CD14^{high} and CD14^{low} MDM in individual tumors; The Wilcoxon matched-pairs test was used to assess statistical significance; p values indicated. **(D)** Representative histograms of PDL1 and isotype control within indicated cell populations.


A
I
Hoechst
Lectin HPA
Ki-67
II
200 µm
II
Hoechst
Lectin HPA
Ki-67
100 µm
I
III
1000 µm
III.a
RNASelect
Phalloidin
200 µm
III.b
Hoechst
TUNEL
200 µm
B
IV
IV.a
IV.b
EGFR
VEGFA
FN1
IV.c
SPP1
FTL
VEGFA
CD74
IV.d
Hoechst
TUNEL
IV.e
Hoechst
Lectin HPA
Ki-67
200 µm
C
V
V.a
V.b
EGFR
VEGFA
FN1
V.c
SPP1
FTL
VEGFA
CD74
200 µm
Hoechst
Lectin HPA
Ki-67
VI
2000 µm
VI.a
Hoechst
Lectin HPA
Ki-67
200 µm
VI.b
Hoechst
TUNEL
200 µm

Supplementary Figure 6. Histological features of tumors with high GW-like composition. (A) H&E staining of sample SL006, showing numerous pseudopalisades. I: Region of microvascular proliferation showing glomeruloid vascular structures, high vascularization (Lectin HPA), and actively proliferating cells (Ki67); presented in higher magnification in II. IIIa: Pseudopalisade areas devoid of cell nuclei (RNASelect) together with cytoskeletal staining (Phalloidin). IIIb: Pseudopalisades showing apoptotic/necrotic cells (TUNEL staining) in the core. **(B)** H&E staining of sample SL061B. IV.a: Close-up view of H&E staining in a part region IV. IV.b: EEL-FISH signal showing the expression of *EGFR*, *VEGFA*, and *FN1* in a part region IV of sample SL061B. IV.c: EEL-FISH signal showing the expression of *SPP1*, *FTL*, *CD74*, and *VEGFA* in a part region IV. IV.d: High level of apoptosis (TUNEL) in a part region IV. Nuclei (Hoechst). IV.e: High vascularization (Lectin HPA) and cellular proliferation (Ki67)) in a part region IV. Nuclei (Hoechst). (C) H&E staining of sample SL035B showing a large region of pseudopalisading necrosis. V.a: Close-up view of H&E staining in a part region V showing pseudopalisading cellular formations. V.b: EEL-FISH signal showing the expression of *SPP1*, *FTL*, *CD74*, and *VEGFA* in a part region V. V.c: EEL-FISH signal showing the expression of *EGFR*, *VEGFA*, and *FN1* in a part region V. VI: Sample SL035B stained for vasculature (Lectin HPA), nuclei (Hoechst) and proliferation (Ki67). VI.a: Proliferating cells (Ki67) outside of the Pseudopalisade (white dotted line) and blood vessels (Lectin HPA). VIb: Apoptotic/necrotic cells (TUNEL staining) inside the pseudopalisade (white dotted line).



Supplementary Figure 7. GW-like subtypes show subtype specific transitions from OPC/nIPC-like and GBM model. (A) Spatial distribution of m-states in samples SL034 (*MDM2* amplified) and SL023 (*CDK4* and *PDGFRA* amplified). I: Spatial distribution of nIPC/RG/OPC-like and GW-like subtypes (4, 8 and 6) in region I of SL034. II: EEL-FISH signal and nuclei showing the expression of *MDM2*, *FOXG1*, *GAP43*, and *VEGFA* in

region II of SL034. III: Spatial distribution of nIPC/RG/OPC-like and GW-like subtypes (7, 8, and 6) in region III of SL023. IV: EEL-FISH signal and nuclei showing the expression of *PDGFRA*, *OLIG2*, *GAP43*, and *VEGFA* in region IV of SL023. **(B)** Spatial distribution of binned *TNC* expression and binned cell cycle score in samples SL046C and SL035B. Expression and cell cycle score were binned into 5 levels. Spatial distribution of preOPC-like, GW-like 6/8, and TAM-MG (left panel) in SL061B. Collapsed and binned total *TNC* expression and collapsed cell cycle score over the shown tissue length (right panel). **(C)** Generalization of GBM proliferation and transition into quiescent GW-like phenotype under TAM and monocyte interaction. **(D)** GBM hierarchical model. Undifferentiated cell types at the center (GBL/preOPC/RG-like) differentiate into AC/OPC/nIPC-like subtypes. Subtype independent transformation into GW-like state. AC/Endothelial cell interaction.

Supplementary Table 1. Reanalysis of mesenchymal cell states in glioblastoma. Excel workbook containing the results of gene ontology analyses of mesenchymal-like cell states from Phillips et al., Verhaak et al. and Neftel *et al.*[6,8,24]. Each tab contains the results from one such analysis and includes links to the original publication, the gene set used, and a list of all statistically siginificant gene ontology terms with statistics.

Supplementary Items Table 1. Pathology report.

Supplementary Items Table 2. Oligonucleotide probe pools for EEL-FISH.