

RESEARCH

Open Access



Single-cell transcriptome sequencing reveals new epithelial-stromal associated mesenchymal-like subsets in recurrent gliomas

Jinwei Li^{1,2,11†}, Shengrong Long^{3,4,7†}, Yang Zhang^{5†}, Shuangqi Yu^{4,7†}, Hongyu Xu^{4,7}, Rui Liang⁶, Quan Liu⁸, Jinnan Zhang⁹, Xiang Li^{4,7}, Yixin Fu¹², Tao Xin^{10*} and Yinyan Wang^{1,11*}

Abstract

Gliomas, particularly glioblastomas, are highly malignant brain tumors with high recurrence rates and poor prognosis. Despite advances in treatment, recurrence remains a major challenge. Epithelial-mesenchymal transition (EMT) plays a key role in tumor invasion and recurrence. This study explores the transcriptional and regulatory mechanisms driving glioma recurrence, focusing on mesenchymal-like (MES-like) subpopulations. Single-nucleus RNA sequencing was performed on 52 IDH wild-type GBM specimens, including 26 primary and 26 recurrent tumors. Spatial transcriptomics data were also incorporated. Tumor subpopulations were identified through gene regulatory network analysis, copy number variation detection, and nonnegative matrix factorization. Functional validation was conducted using gene knockdown experiments, followed by xenograft studies. We discovered novel MES-like subpopulations in recurrent GBM enriched with EMT-related genes like *EGR1* and *SERPINE1*. These subpopulations exhibited increased transcriptional activity and were associated with poor prognosis and invasiveness. Knockdown of *SERPINE1* significantly reduced cell proliferation and migration. Spatial transcriptomics showed MES-like cells concentrated at the tumor margins, highlighting their role in invasion and recurrence. MES-like subpopulations, driven by *EGR1* and *SERPINE1*, are critical in GBM. Targeting these regulators could offer new therapeutic strategies to reduce glioma recurrence and improve outcomes.

Keywords Glioblastoma, Epithelial-mesenchymal transition, Single-cell RNA sequencing, Recurrence, MES-like subpopulations

[†]Jinwei Li, Shengrong Long, Yang Zhang and Shuangqi Yu authors contributed equally.

*Correspondence:
Tao Xin
drxintao@yeah.net
Yinyan Wang
tiantanyinyan@126.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Gliomas are the most common malignant tumors of the central nervous system. Despite receiving standard treatments, including maximal safe resection combined with radiotherapy and temozolomide chemotherapy, recurrence is inevitable [1]. High-grade gliomas are particularly prone to recurrence, with a 5-year survival rate of less than 10% [2, 3]. Additionally, the clinical significance of reoperation in recurrent glioblastoma patients remains unclear. Therefore, understanding the temporal evolution of tumors during treatment is crucial for addressing glioma recurrence. By analyzing gene expression data from primary and recurrent tumors, studying tumor evolution at the transcriptional level offers a novel perspective on the progression of glioblastoma.

Recurrent gliomas are often associated with increased malignancy, making subsequent treatments more challenging and leading to poorer prognosis. The epithelial-mesenchymal transition (EMT) is regulated by inducers, transcription factor families, and a series of signaling pathway genes, and is implicated in the invasion and progression of gliomas [4, 5, 6, 7, 8, 9]. The EMT process is frequently reactivated, enhancing the invasive and metastatic capabilities of cancer cells [10, 11]. This suggests that recurrent gliomas, potentially driven by residual invasive cells, may undergo a multi-step reactivation process. This process facilitates adhesion to the extracellular matrix, secretion of enzymes that degrade surrounding matrices, and migration within the glioma microenvironment prior to tumor formation [12, 13]. Therefore, investigating the mechanisms by which EMT influences GBM invasion and migration is crucial for developing new and effective therapeutic strategies.

In recent years, the rapid development of single-cell sequencing technology has significantly advanced the study of glioma tumor plasticity, the characteristics of the tumor microenvironment, and the role of the microenvironment in tumor plasticity. This progress has also brought potential new opportunities for glioma treatment. Currently, most studies categorize glioma tumor cells into neural progenitor cell-like (NPC-like), mesenchymal-like (MES-like), oligodendrocyte precursor cell-like (OPC-like), and astrocyte-like (AC-like) states [14, 15, 16]. Intratumoral heterogeneity and the dynamic plasticity of cell states are hallmarks of malignant brain tumors [14, 16, 17, 18]. Previous studies, including those from our group and others, have demonstrated that the MES-like cell state is associated with glioma invasion [16, 19, 20, 21]. However, research on tumor cell states remains limited, which may be a critical factor contributing to tumor recurrence and drug resistance.

A comprehensive understanding of the transcriptional regulatory mechanisms in primary and recurrent gliomas may aid in developing new therapeutic strategies and

improving patient outcomes. In this study, we utilized extensive single-cell sequencing data from matched primary and recurrent gliomas to identify genes with high transcriptional regulatory activity in different tumor cell states. Spatial transcriptomics analysis was conducted to determine the spatial localization of tumor cell subpopulations and the transcription factor *EGR1*. We validated the expression and differentiation of *EGR1* and *SERPINE* in tumor microenvironment subpopulations using single-cell sequencing data from wild-type mice at different stages. Additionally, we performed cellular functional assays and animal experiments to verify our findings. Our research suggests that in recurrent gliomas, *EGR1* may mediate EMT and that the *SERPINE*-MES subpopulation could be a key factor in glioma recurrence and drug resistance.

Materials and methods

Data collection

This study performed single-nucleus RNA sequencing analysis on 52 human IDH wild-type GBM specimens, comprising 26 primary tumors and 26 matched recurrent tumors from the same patients. All patients underwent standard surgical treatment, radiotherapy, and chemotherapy. The collection and processing of tumor specimens followed protocols described in previous literature based on original data [22]. Transcriptomic and clinical data was downloaded from the Chinese Glioma Genome Atlas (CGGA) database (<http://www.cgga.org.cn/>), including mRNA-693 cohort ($n=693$), mRNA-325 cohort ($n=325$), Rembrandt cohort ($n=475$), and glioma cohort from The Cancer Genome Atlas (TCGA) ($n=702$) (<https://www.cancer.gov/ccg/research/genome-sequencing/tcga>). Spatial transcriptomics sequencing data was obtained from previous studies and is now publicly available (GSE194329) [19]. Patient sample collection and study design for spatial transcriptome data have received ethical approval from West China Hospital, Sichuan University (2020.837). The spatial transcriptomics data included three human IDH wild-type primary GBM specimens, two recurrent glioblastoma specimens, and one peritumoral tissue sample. Early (day 7) and late (day 28) IDH wild-type GBM mouse single-cell transcriptomic data was obtained from the publicly available GEO database (GSE166218) [23].

Single-cell quality control processing workflow

The quality control parameters were set to allow a minimum of 500 genes and a maximum of 4000 genes per cell, with mitochondrial gene content limited to 20%. The FindVariableFeatures function was used to identify 2000 highly variable genes. Batch effects across different tumor samples were corrected using the R package harmony. Cell type identification and annotation were performed

by referencing the CellMarker database and previous literature [24, 25]. The R package SCP was used to create bar plots for cell proportion visualization and heatmaps for marker gene expression analysis.

Single-cell copy number variation (CNV) analysis

Single-cell CNV analysis was conducted based on single-cell transcriptome data to identify cells with copy number abnormalities, aiding in the inference of malignant tumor cells. To distinguish between tumor cells and non-malignant cells, we used the R package inferCNV [26]. T cells and macrophages were set as the reference cell types. After quantifying the degree of copy number variation using the CNV score, cells were classified into tumor cells and normal cells. Subsequently, we annotated subpopulations of tumor cells using marker genes identified from previous literature [25, 27, 28, 29, 30].

Gene regulatory network analysis

The transcriptional states of microenvironmental cells in single-cell sequencing were derived from underlying gene regulatory networks [31]. These networks included transcription factors, cofactors, and their downstream target genes. Using single-cell regulatory network inference and clustering (SCENIC) analysis, we effectively identified co-expression modules (regulons) between transcription factors (TFs) and potential target genes [18, 31]. We analyzed the differences in transcriptional activity between IDH wild-type primary and recurrent glioma tumor cell subpopulations.

Identifying MES subpopulation cells using non-negative matrix factorization (NMF)

We constructed an expression matrix using genes from the EMT signaling pathway and excluded cells without EMT gene expression. EMT-related genes were obtained from The Molecular Signatures Database (MSigDB) (Supplement Table 1). The R package NMF was used to decompose the MES cell subpopulation [32]. The function FindAllMarkers was employed to identify characteristic genes for each cluster. Subsequently, we redefined new MES cell subpopulations.

Single-cell subpopulation stemness analysis

We assessed the differentiation potential of glioma tumor cell subpopulations and MES subpopulations at single-cell level. The R package CytoTRACE was used to predict the differentiation capabilities of different subpopulation cells. The scores were visualized using bar plots and UMAP plots [33].

Constructing cellular trajectories

Cell subpopulation trajectory inference is crucial for studying the differentiation structure and sequence of

cell lineages. We employed the R packages “monocle2” and “Slingshot” to infer the differentiation trajectories of MES cell subpopulations and displayed the enriched signaling pathways within these MES subpopulations.

Pathway enrichment analysis

Metabolic changes in cells can significantly influence tumorigenesis and development. We further analyzed pathway enrichment in MES stromal subpopulations using the R package scMetabolism for metabolic pathway enrichment analysis. Single-cell flux estimation analysis (scFEA) was used to evaluate metabolic changes in the epithelial-mesenchymal MES subpopulation [34]. Heatmaps were generated to display the expression of immune-related signaling genes in the novel MES stromal subpopulation.

Spatial transcriptomics analysis

Spatial transcriptomics data were generated using the 10x Genomics Visium platform. Spatial cell type mapping was performed using single-cell transcriptome data as a reference [35]. The reference data were preprocessed and modeled using the Cell2location Bayesian framework. We set the expected number of cells per location to 30 and applied empirical Bayes priors to avoid overfitting. The model was trained for 30,000 epochs with default hyperparameters, as the authors of Cell2location recommended. After training, we extracted the posterior cell abundance estimates and applied log-transformation for downstream visualization. To determine spatial enrichment, we computed Z-scores across tissue sections and considered a Z-score > 1.5 as indicative of significant enrichment for a given cell type in a spatial location.

Computational gene perturbation

Geneformer has been applied in various biological fields, including identifying potential disease drug targets [36]. Geneformer effectively captures the relationships and importance of genes within single-cell transcriptomes. This focus mechanism allows Geneformer to prioritize the most relevant genes for specific learning objectives, optimizing prediction accuracy. In this study, we utilized Geneformer to investigate key genes associated with the transition from high EMT score MES cells to low EMT score MES cells in IDH wild-type recurrent patients.

Survival analysis of transcriptomic cohorts

We performed a prognostic analysis of MES stromal cell subpopulations. The R package GSVA was used to perform ssGSEA normalization scoring based on the gene sets of single-cell MES subpopulations within the expression matrix of tumor patients [37]. Using the CGGA-693 cohort, survival analysis was conducted by grouping patients based on the median value of the ssGSEA scores.

Immunostaining and analysis of human tissue microarrays

A high-density tissue microarray containing 180 cases of gliomas of varying grades and control was purchased from Superchip Biotech (catalog number HBraG180Su01). An anti-EGR1 antibody (catalog number 55117-1-AP, Proteintech) was used for signal detection (1:500 dilution). Tissue spots with >30% loss were excluded. The immunofluorescence (IF) score was calculated by multiplying the staining intensity by the percentage of stained cells Supplement Table 2.

Cell culture, transfection and viral transduction

The human glioma cell lines, U251 and T98G, were sourced from Procell (Wuhan, China) and cultured in Dulbecco's Modified Eagle Medium (DMEM; Servicebio), supplemented with 10% fetal bovine serum (FBS, Excel) and 10 μ L/mL penicillin-streptomycin solution (100 \times , Biosharp). The cells were incubated at 37 °C in a humidified atmosphere of 5% CO₂ (Hair).

siRNA transfection

For gene knockdown using siRNA, cells were seeded at 5×10^4 cells per well in six-well plates and allowed to adhere for 24 h. siRNAs targeting EGR1 (siEGR1) and a non-targeting control (siControl), both obtained from Ribobio, were used. Transfections were performed using the Ribo FECT™ CP Transfection Kit following the manufacturer's instructions when cells reached approximately 40% confluency. After six hours of incubation, the medium was replaced with fresh DMEM containing 10% FBS. Cells were collected 36 h post-transfection for subsequent experiments.

Lentiviral transduction

For stable SERPINE1 gene knockdown, cells were transduced with SERPINE1 shRNA lentivirus or control lentivirus (shControl) (Obio Technology). Polybrene (5 μ g/mL, Yeasen) was used to enhance viral infection efficiency. After 24 h of infection, the medium was replaced with fresh DMEM containing 2.5 μ g/mL puromycin (InvivoGen, USA) to select stably transduced cells for 10 days. siRNA and shRNA sequences used were listed in Supplement Table 3.

RNA extraction, cDNA synthesis, and quantitative real-time PCR

Total RNA was isolated from tissue samples stored at temperatures below -80 °C, utilizing Takara RNAiso Plus (Takara Bio, Otsu, Shiga, Japan) according to the manufacturer's instructions. For cDNA synthesis, 1 μ g of RNA was reverse transcribed using the HiScript IV QRT SuperMix for qPCR kit (Vazyme Medical Technology). Quantitative real-time PCR (qPCR) data were analyzed

using the $2^{-\Delta\Delta C_t}$ method. Primer sequences are provided in Supplement Table 4.

Cell counting Kit-8 proliferation assay

CCK8 assay was utilized to gauge cell growth. U251 cells and T98G cells were respectively seeded in 96-well-plates at a density of 5,000 cells per 100 μ L per well and treated for 24–72 h. Subsequently, 10 μ L of CCK8 reagent was added to each well, followed by a 30-minute incubation at 37 °C. Eventually, the Multiskan™ SkyHigh thermoscientific spectrophotometer was utilized to measure the absorbance at 450 nm.

Wound healing assay

The migration capacity of cells was gauged by performing wound healing assay. U251 cells and T98G cells were respectively seeded in 6-well-plates at a density of 500,000 cells per 1 mL per well. After 6 h, scratches were made using a 200- μ L pipette tip. The cells were then washed three times with PBS and continued to be cultured in a medium containing 5% FBS. The healing progress of the wounds was monitored at 0 and 24 h.

Flow cytometric cell cycle analysis

To assess DNA fragmentation, U251 and T98G cells were plated in six-well plates at a density of 500,000 cells per 1 mL per well for 24 h. After incubation, the cells were collected by trypsinization, rinsed with a complete medium, washed with PBS, and fixed in 70% cold ethanol for 24 h. The cells were then gently resuspended in 500 μ L of a hypotonic fluorochrome solution containing 450 μ L propidium iodide (PI) and 50 μ L RNase A to stain the nuclei. DNA content was measured using a CytoFLEX flow cytometer, and cell cycle distribution was analyzed with ModFit software, with 8,000 events recorded per sample.

Animal experiments

In a xenograft experiment, female athymic nude mice ($n=6$ /group) were acclimated for one week. U251 cells at 80% confluence were trypsinized after exposure to fresh medium for 24 h and inoculated 0.5×10^6 cells in 100 μ L volumes into mice of each group. On day 25, tumor volume was measured by calipers and calculated as $1/2 \times (\text{Length} \times \text{Width} \times \text{Height})$, and mice were euthanized for tumor fixation in formalin and paraffin embedding. Tissue sections were stained with H&E for proliferative analysis under a microscope. The animal experiments took place in the SPF Animal Laboratory at Wuhan University (ZN2023021).

Statistical analysis

Statistical analyses were performed using R (v4.3.1). For comparisons between two groups, unpaired two-tailed

Student's t-tests or Wilcoxon rank-sum tests were used as appropriate. For comparisons among more than two groups, one-way ANOVA with Tukey's post hoc test was used for normally distributed data, and Kruskal–Wallis test for non-parametric data. Survival curves were analyzed using the Kaplan–Meier method and log-rank test. Multiple hypothesis testing was adjusted using the Benjamini–Hochberg method where applicable. A p-value < 0.05 was considered statistically significant.

Results

Distribution of cell subpopulations in IDH wild-type primary and recurrent gliomas

The technical roadmap illustrates a multi-omics analysis of IDH wild-type primary and paired recurrent gliomas (Fig. 1A). We analyzed sequencing data from multiple cohorts to investigate the mechanisms of glioma recurrence. The findings were then validated through clinical tissue samples, cell-based experiments, and animal models. First, we analyzed transcriptional regulatory changes in cell subpopulations from single-cell data of primary and recurrent gliomas. Second, we identified key regulatory genes in specific MES-like subpopulations. Finally, we validated the roles of these key genes in gliomas through cellular and animal experiments. The UMAP plot shows the cell distribution of 52 IDH wild-type GBM samples, including 26 primary tumors and their 26 matched recurrent tumors (Supplement Fig. 1A). Cell subpopulations were annotated with markers for Oligodendrocytes, Astrocytes, Macrophages, Microglial cells, Neurons, Endothelial cells, and T cells (Fig. 1B). A total of 143,340 cells were included in the analysis for visualization (Fig. 1C). UMAP visualization depicts the distribution of cell subpopulations in primary and recurrent gliomas (Fig. 1D). The proportions of cell subpopulations in IDH wild-type primary gliomas were as follows: Astrocytes (60.5%), Oligodendrocytes (17.9%), Macrophages (7.3%), Microglial cells (5.1%), Neurons (7%), Endothelial cells (1.3%), and T cells (0.9%) (Fig. 1E). In recurrent gliomas, the proportions were higher for Oligodendrocytes (29.1%), Macrophages (7.9%), Microglial cells (6.4%), Neurons (7.6%), and Endothelial cells (4.1%) compared to primary gliomas. Enrichment analysis revealed that upregulated signaling pathways in recurrent gliomas include Salmonella infection, protein processing in the endoplasmic reticulum, endocytosis, Human T-cell leukemia virus 1 infection, and cell cycle pathways. Downregulated pathways include protein digestion and absorption, and cytoskeleton organization in muscle cells (Fig. 1F).

Distribution of tumor subpopulation expression and transcriptional activity

To dissect the transcriptional heterogeneity of GBM at single-cell resolution, particularly across primary and recurrent states, we applied transcriptional signature analyses to identify distinct tumor subpopulations. We further identified tumor cell subpopulations using infer-CNV. Violin plots illustrate the CNV scores of various cell clusters (Supplement Fig. 1B). Cell proportion plots revealed that primary gliomas comprised 86.4% tumor cells and 13.6% normal cells, while recurrent gliomas comprised 84.7% tumor cells and 15.3% normal cells (Supplement Fig. 1C). UMAP visualizations displayed the distribution of normal and tumor cells in both primary and recurrent gliomas (Supplement Fig. 1D). In the primary glioma, the number of malignant cells was 51,166, while the number of normal cells was 8,081. In recurrent glioma, the number of malignant cells increased to 71,221, and the number of normal cells was 12,872. Heatmaps showed the expression of marker genes in tumor cell subpopulations (Fig. 2A). UMAP visualizations revealed the spatial distribution of OC-like, OPC-like, AC-like, and MES-like tumor cell subpopulations (Fig. 2B). In the primary glioma, the numbers of the four types of malignant cells were as follows: OC-like (14,015), OPC-like (9,852), AC-like (26,709), and MES-like cells (590). In recurrent glioma, the numbers were: OC-like (28,222), OPC-like (11,632), AC-like (27,825), and MES-like cells (3,542). In primary gliomas, the proportions of OC-like, OPC-like, AC-like, and MES-like cells were 27.4%, 19.3%, 52.2%, and 1.2%, respectively (Supplement Fig. 1E). In recurrent gliomas, these proportions were 39.6%, 16.3%, 39.1%, and 5%, respectively. UMAP visualizations also showed gene expression and transcription factor activity in primary and recurrent tumor subpopulations (Fig. 2C). We conducted SCENIC analysis to reconstruct gene regulatory networks and identify cell states, generating nine modules in total (Fig. 2D). A total of 9 modules were formed through clustering. Modules 4 and 8 exhibited higher activity in recurrent glioma tissues (Fig. 2E). Our research primarily focused on mechanisms related to recurrent gliomas. Moreover, UMAP visualizations showed higher activity of MES subpopulations in Module 2 (Fig. 2F). CytoTRACE assessed the stemness and differentiation status of tumor cell subpopulations, with MES-like cells showing lower stemness scores (Fig. 2G, H).

Enrichment pathway analysis and Spatial localization reveal characteristics of tumor subpopulations

GSEA enrichment analysis of primary and recurrent glioma cell subpopulations indicated upregulated pathways in secondary MES-like subpopulation, including ECM-receptor interaction, WNT signaling, RAS signaling,

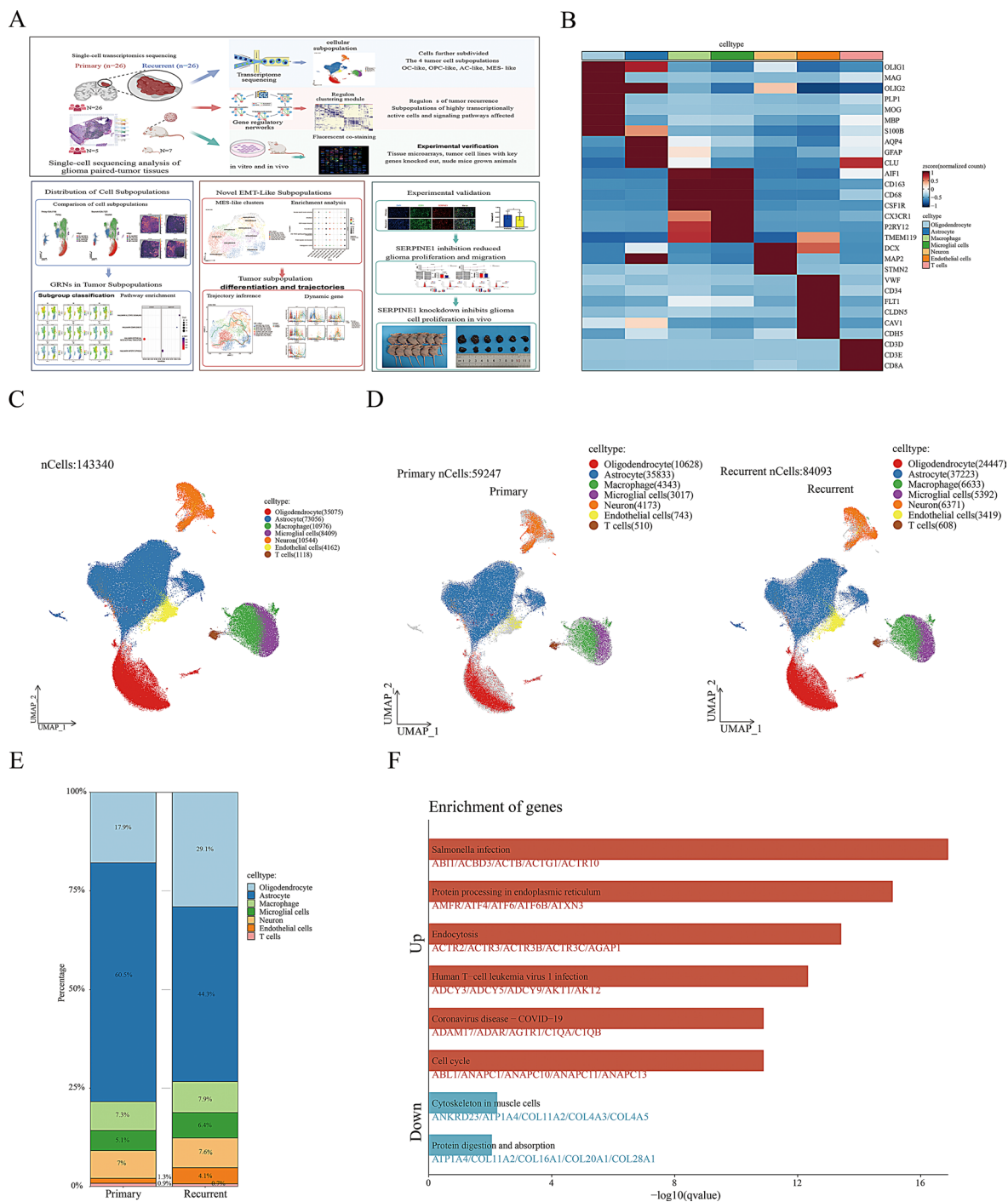


Fig. 1 Comprehensive analysis of cell subpopulations and gene expression in gliomas using single-cell RNA sequencing (scRNA-seq). **(A)** Schematic workflow outlining the steps of single-cell RNA sequencing, data analysis, subpopulation classification, and experimental validation. **(B)** Heatmap displaying the expression levels of key genes across different glioma cell types, with rows representing genes and columns representing different cell subpopulations. **(C)** UMAP plot visualizes cell type distribution in the entire dataset, with each color representing a distinct cell type as annotated in the legend. **(D)** UMAP plots comparing cell populations between primary and recurrent gliomas, revealing differences in cell type composition between the two conditions. **(E)** Bar chart showing the relative proportions of different cell types in primary and recurrent gliomas, highlighting the higher abundance of certain cell types in recurrent gliomas. **(F)** Bar plot representing gene enrichment analysis, showing significantly upregulated and downregulated genes associated with glioma progression

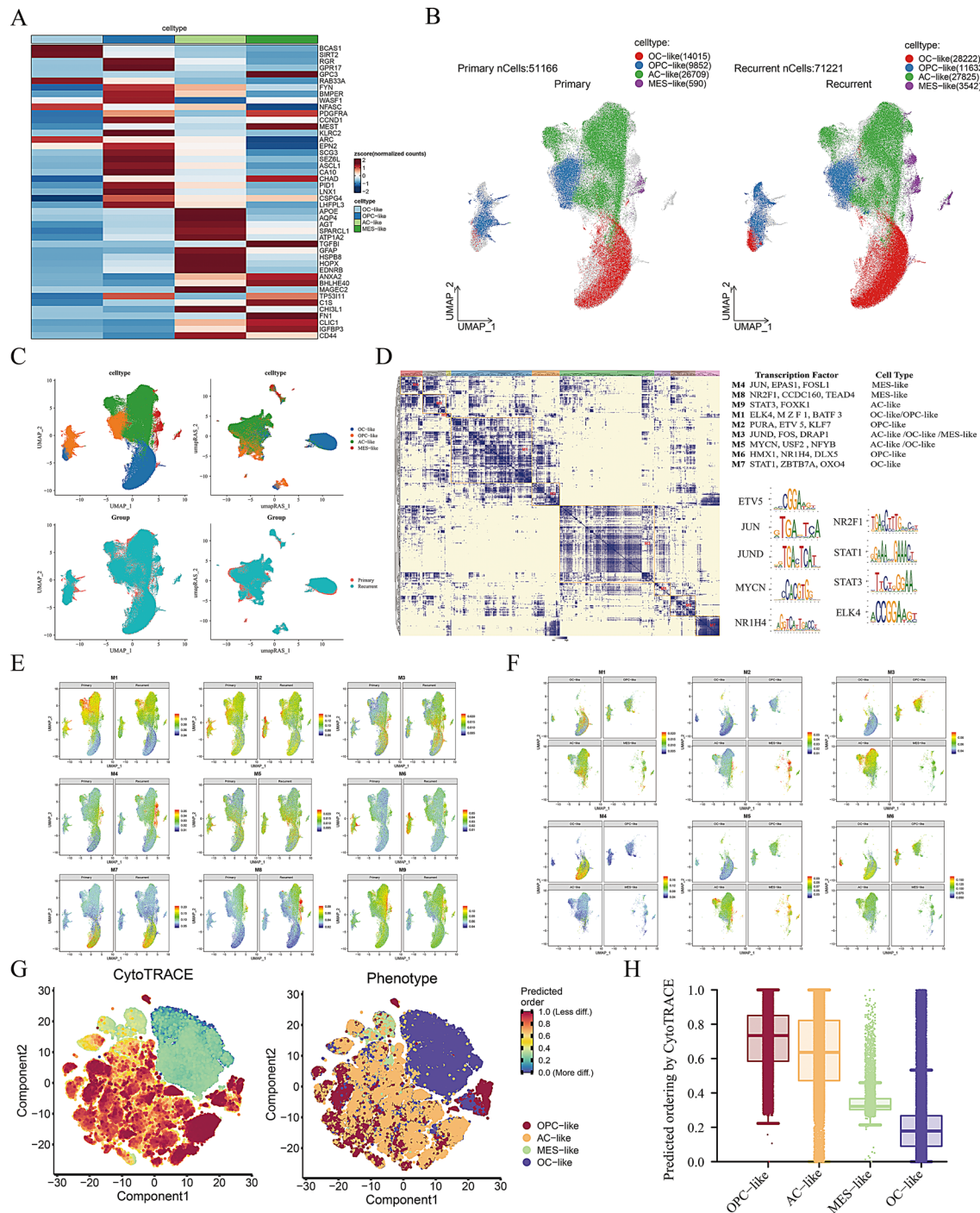


Fig. 2 Comparative analysis of primary and recurrent gliomas using single-cell RNA sequencing (scRNA-seq). **(A)** Heatmap showing the expression levels of key genes across different glioma subtypes in both primary and recurrent gliomas. Rows represent genes, and columns represent subtypes. **(B)** UMAP plots illustrate the distribution of 51,166 cells in primary gliomas and 71,221 cells in recurrent gliomas, with cells color-coded by subtype (OPC-like, AC-like, MES-like, OC-like). **(C)** UMAP plots comparing cell subtype distribution across different groups, showing changes in subtype composition between primary and recurrent gliomas. **(D)** Heatmap and transcription factor motif analysis highlighting key transcription factors associated with different glioma subtypes, including ETV5, JUN, and MYCN. **(E-F)** Expression patterns of various marker genes across glioma subtypes, visualized as UMAP feature plots, with intensity corresponding to gene expression levels. **(G)** CytoTRACE and phenotype plots showing differentiation states of cells across glioma subtypes, with predicted differentiation levels color-coded. **(H)** Boxplot representing the predicted ordering of cell differentiation levels by CytoTRACE across different glioma subtypes, indicating varying differentiation states between subtypes

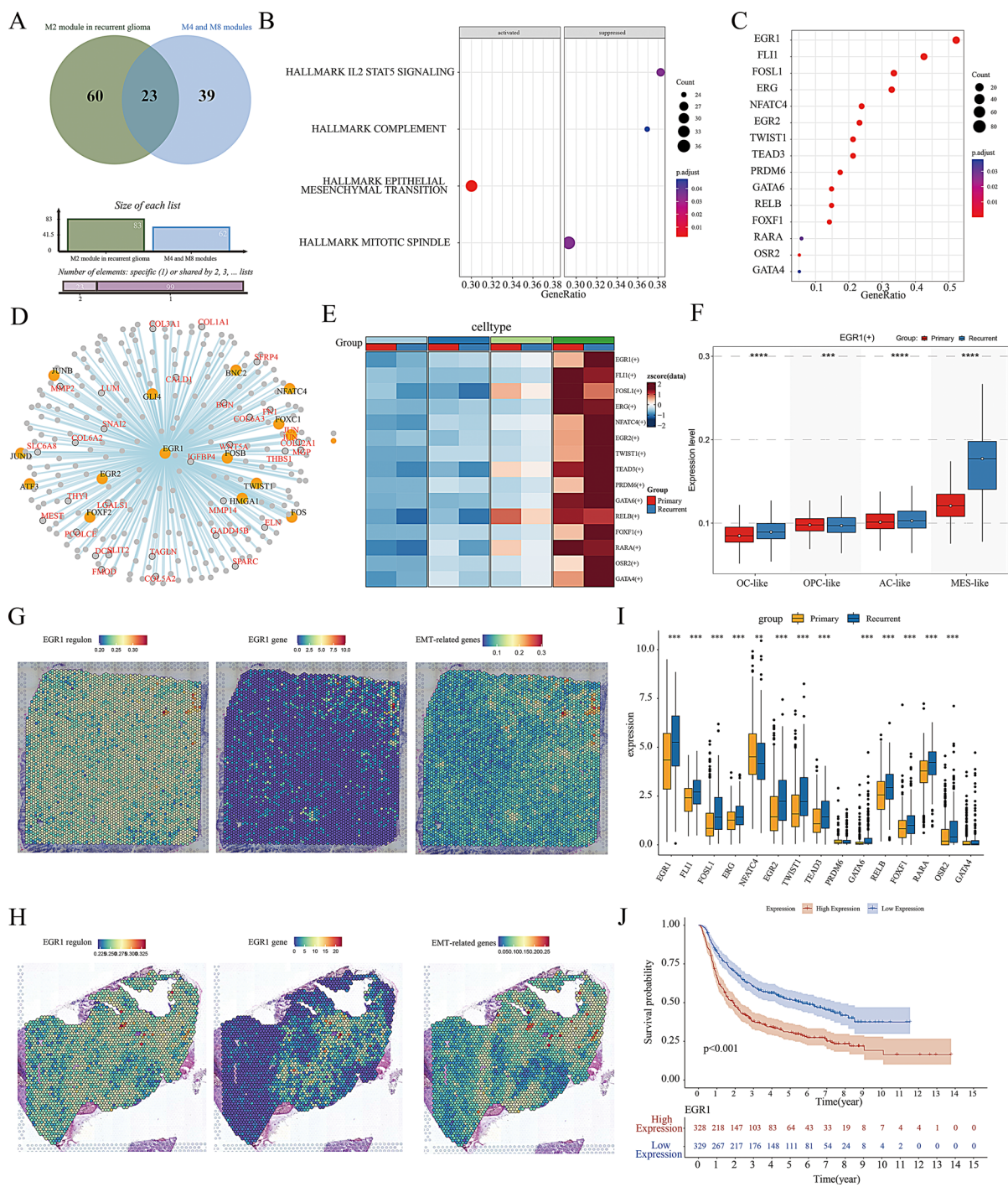


Fig. 3 (See legend on next page.)

and PI3K-AKT signaling pathways (Supplement Fig. 2A). This study supports previous findings on the significance of the interaction between the PI3K/AKT/mTOR and WNT/ β -catenin pathways, suggesting that concurrently targeting these pathways may enhance the treatment of glioblastoma multiforme [38, 39, 40]. In the AC-like subpopulation, upregulated pathways included inflammatory mediator regulation of TRP channels, oxytocin signaling,

and cAMP signaling pathways. The OPC-like subpopulation showed upregulation in glioma, GABAergic synapse, transcriptional misregulation in cancer, MAPK signaling, and cGMP-PKG signaling pathway. The recurrent OC-like subpopulation had upregulated dopaminergic synapses and rap1 signaling pathways. Spatial transcriptomics analysis revealed the spatial localization of tumor cell subpopulations. The AC-like subpopulation occupied

(See figure on previous page.)

Fig. 3 Gene expression analysis and survival prediction in glioblastoma subtypes

(A) Venn diagram showing the overlap of differentially expressed genes between primary and recurrent glioblastomas, with the bar plot below illustrating the number of genes unique to each group. **(B)** Pathway enrichment analysis highlighting the activation and suppression of key signaling pathways, including IL2/STAT5 signaling, complement pathway, epithelial-mesenchymal transition (EMT), and mitotic spindle pathway, with dot size representing gene ratio and color indicating p-values. **(C)** Dot plot depicting the differential expression of transcription factors (*EGR1*, *FLI1*, *FOXL1*, etc.) between primary and recurrent glioblastomas. The dot size corresponds to the gene ratio, and the color indicates p-values. **(D)** Interaction network of transcription factors associated with glioblastoma subtypes, centered around *EGR1* and related factors. **(E)** Heatmap showing the expression of key genes across glioblastoma subtypes (OC-like, OPC-like, AC-like, MES-like) in both primary and recurrent tumors. **(F)** Boxplot of *EGR1* expression levels across glioblastoma subtypes, demonstrating elevated expression in MES-like cells, particularly in recurrent glioblastomas. **(G)** Spatial transcriptomic heatmaps showing the spatial distribution of *EGR1* and EMT marker expression in secondary glioblastomas, with color intensity corresponding to expression levels. **(H)** Boxplot comparing the expression levels of various genes between primary and recurrent glioblastomas, showing significant differences in key regulators. **(I)** Kaplan-Meier survival curve showing the association between *EGR1* expression levels and patient survival, with high *EGR1* expression correlating with poorer prognosis ($p < 0.001$)

most of the spatial distribution in secondary tumors, consistent with our single-cell subpopulation analysis (Supplement Fig. 2B-F). Additionally, compared to primary gliomas, the spatial localization of MES-like cells was less prominent in secondary glioblastomas.

Analysis of key signaling pathways, cell subpopulations, and EMT-related gene expression in recurrent gliomas

We investigated the key signaling pathways and cell subpopulations involved in recurrent gliomas. Modules 4 and 8 were primarily expressed in the MES cell subpopulations of recurrent gliomas (Supplement Table 5). Additionally, transcriptional regulatory Module 2 was mainly expressed in the MES cell subpopulations of recurrent gliomas (Supplement Table 6). A Venn diagram of transcription factors from Modules 4, 8, and 2 in recurrent gliomas was constructed (Fig. 3A). Enrichment analysis revealed activation of the EMT pathway in the MES subpopulations, while IL2-STAT5 signaling, complement, and mitotic spindle pathways were inhibited (Fig. 3B). Key transcription factors activating the EMT pathway included *EGR1*, *FLI1*, and *FOSL1* (Fig. 3C). An interaction map of *EGR1* with its target genes was constructed (Fig. 3D). To further characterize intratumoral heterogeneity related to epithelial-mesenchymal transition (EMT), we performed unsupervised non-negative matrix factorization (NMF) clustering on malignant cells from both primary and recurrent glioblastoma samples, using the expression of EMT pathway-related genes. This approach enabled us to identify previously unrecognized EMT-like subpopulations, which may play distinct roles in tumor progression and treatment resistance. The heatmap displayed the expression of EMT-related genes in recurrent MES cell subpopulations (Fig. 3E). Compared to primary gliomas, the *EGR1* regulon was highly expressed in the MES subpopulations of recurrent gliomas ($p < 0.001$) (Fig. 3F). *EGR1* gene, the *EGR1* regulon, and the EMT pathway were spatially localized in secondary glioblastomas (Fig. 3G, H). Transcriptome sequencing validation showed high expression of EMT-related genes, including *EGR1*, *FLI1*, *FOSL1*, *ERG*, and *EGR2*, in recurrent gliomas ($p < 0.001$) (Fig. 3I). Kaplan-Meier

survival analysis indicated that high *EGR1* expression was associated with poor prognosis in glioma patients ($p < 0.001$) (Fig. 3J). The results also reveal a difference in tumor grading and the decision to undergo chemotherapy between patients with recurrent gliomas and those with primary gliomas (Supplement Table 7).

Identification of novel EMT-like subpopulations and analysis of their signaling pathways and stemness differentiation

We identified novel EMT-like subpopulations through unsupervised NMF clustering based on EMT pathway-related genes [41]. The identified subpopulations included *MGP*+MES-like-C1 (15.7%), *SLIT2*+MES-like-C2 (20.4%), *SNTB1*+MES-like-C3 (7.9%), *DPYSL3*+MES-like-C4 (7.1%), *VEGFC*+MES-like-C5 (6.8%), *SERPINE1*+MES-like-C6 (10.4%), *COL12A1*+MES-like-C7 (5.3%), unclear-MES-like-C8 (17.9%), and nongene-MES-like-C9 (8.4%) (Fig. 4A, B). Moreover, glioma patients with high expression of EMT-related genes, including *MGP*, *SERPINE1*, *SNTB1*, *VEGFC*, and *COL12A1*, had poor prognosis (Supplement Fig. 3A-G). Specifically, glioma patients with high expression of *MGP*+MES-like-C1, *VEGFC*+MES-like-C5, *SERPINE1*+MES-like-C6, and *COL12A1*+MES-like-C7 subpopulation cells had poor prognosis (Supplement Fig. 4A-G). The signaling pathways enriched in different EMT-like subpopulations included vascular smooth muscle contraction, ECM-receptor interaction, and focal adhesion pathways. Specifically, the *SERPINE1*+MES-like-C6 and *COL12A1*+MES-like-C7 subpopulations were enriched in the ECM-receptor interaction pathway (Fig. 4C). We assessed the stemness differentiation status of these tumor subpopulations using CytoTRACE (Fig. 4D). Box plots indicated that the *MGP*+MES-like-C1 subpopulation had the highest stemness score (Fig. 4E). Spatial transcriptomics analysis of recurrent gliomas revealed that most subpopulations were located at the tumor margins, with a broader distribution observed for the *DPYSL3*+MES-like-C4 subpopulation (Fig. 4F, G). Using the Geneformer deep learning model, we inferred the key genes influencing the MES-like tumor subpopulations.

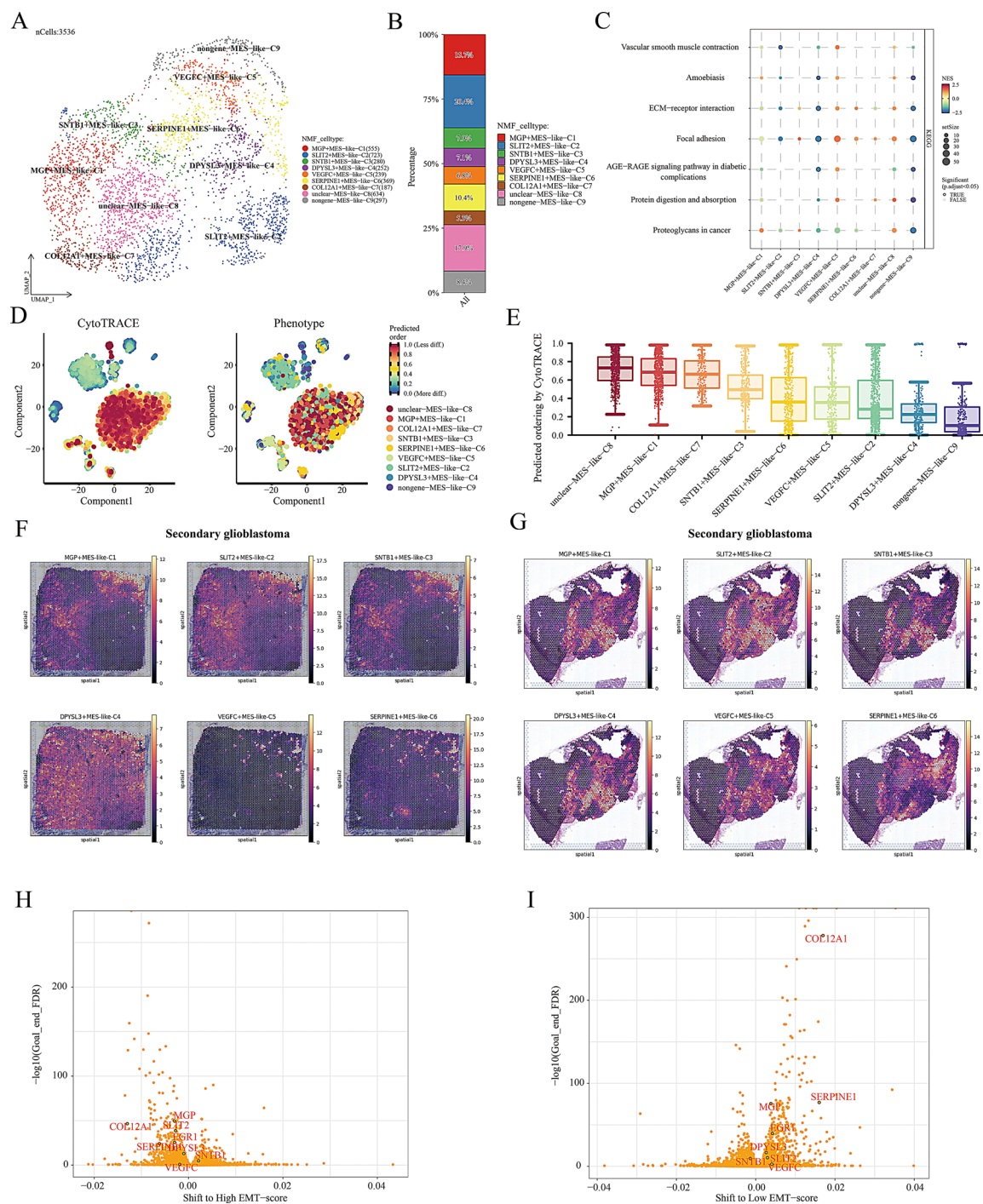


Fig. 4 Subtype analysis and pathway enrichment in secondary glioblastomas. **(A)** UMAP plot showing the distribution of secondary glioblastoma cells, color-coded by non-negative matrix factorization (NMF) subtypes. Each cluster represents a distinct subtype, annotated with marker genes. **(B)** Bar chart representing the relative proportions of different NMF-derived subtypes across secondary glioblastomas. Subtype MES-like C7 shows a higher prevalence compared to other subtypes. **(C)** Dot plot illustrating the pathway enrichment analysis for different subtypes. Pathways such as ECM-receptor interaction, focal adhesion, and protein digestion and absorption are highlighted, with dot size indicating the number of enriched genes and color representing significance (p-value). **(D)** CytoTRACE and phenotype UMAP plots showing differentiation states of the glioblastoma cells. Predicted differentiation states are color-coded, with less differentiated cells clustered in MES-like subtypes. **(E)** Boxplot comparing CytoTRACE-predicted ordering scores for glioblastoma subtypes, showing variation in predicted differentiation levels across subtypes. **(F-G)** Histological sections of secondary glioblastoma corresponding to each MES-like subtype. Color intensity reflects expression levels, highlighting the spatial heterogeneity of gene expression across tumor sections. **(H-I)** Volcano plots showing shifts in EMT (epithelial-mesenchymal transition) scores and differential gene expression in secondary glioblastomas. Key genes such as SERPINE1, MOAP1, and COLEC12 are highlighted based on their significance and EMT shift

The key genes identified were *MGP*, *SLIT2*, *SNTB1*, *DPYSL3*, *VEGFC*, *SERPINE1*, and *COL12A1*, which are associated with the transition from high EMT-score MES cells to low EMT-score MES cells in IDH wild-type recurrent gliomas (Fig. 4H, I).

The trajectory analysis and changes in cell communication of MES-like tumor subpopulations were examined

To explore the state changes of novel EMT-related MES-like tumor subpopulations, we conducted cell trajectory and cell communication analyses. The pseudotime analysis revealed that MES-like tumor cells could be classified into three states (Supplement Fig. 5A). Furthermore, the novel EMT-related MES-like tumor subpopulations formed the cell differentiation trajectory, with *MGP* and *SERPINE1* located at the terminal stage of cell differentiation (Supplement Fig. 5B, C). Subsequently, the network diagram illustrates the number and strength of interactions between cell subpopulations (Supplement Fig. 5D, E), with *MGP*+MES-like-C1 showing a relatively high number and strength of interactions. Additionally, the heatmap indicated that both *MGP*+MES-like-C1 and *COL12A1*+MES-like-C7 had prominent input and output signals (Supplement Fig. 5F). *MGP*+MES-like-C1 exhibited outgoing signaling pathways such as MK, BMP, IGF, and SEMA3, while incoming pathways included MK, BMP, PTN, EDN, and VISFATIN. *COL12A1*+MES-like-C7 showed outgoing signals in the MK, BMP, IGF, SEMA3, and HGF pathways, with incoming signals from MK, BMP, PTN, IGF, and EDN. A hierarchy plot further demonstrated that the *MGP*+MES-like-C1 subpopulation interacted more frequently with other subpopulations through the MK and IGF signaling pathways (Supplement Fig. 5G, H). Lastly, the bubble plot displayed ligand-receptor relationships between subpopulations, highlighting that MDK, IGF, and BMP6 signaling pathways were involved in the regulation of *MGP*+MES-like-C1 and *SERPINE1*+MES-like-C6 over the other subpopulations (Supplement Fig. 5I).

Cell lineage differentiation and gene correlation analysis of MES-like tumor subpopulations

We inferred the cell lineage differentiation of MES-like tumor subpopulations using Slingshot analysis, which identified two lineages (Fig. 5A, B). The *MGP*+MES-like-C1 and *SERPINE1*+MES-like-C6 subpopulations were located at the end of the differentiation trajectories (Fig. 5C, D). Dynamic analysis revealed six cell clusters that were enriched to analyze biological processes, molecular functions, and cellular components (Fig. 5D). The correlation between *EGR1* and *MGP* was high in both recurrent and primary gliomas ($R_R=0.45$, $R_P=0.60$) (Fig. 5E). Similarly, the correlation between *EGR1* and *SERPINE1* was high in both recurrent and primary

gliomas ($R_R=0.54$, $R_P=0.60$) (Fig. 5F). We subsequently analyzed the immunofluorescence expression of *EGR1* and *SERPINE1* in primary and recurrent gliomas using tissue microarrays (Fig. 6A, B). Immunofluorescence experiments on tissue microarrays of primary and recurrent gliomas showed a high correlation between *EGR1* and *SERPINE1* ($R=0.47$) (Fig. 6C). *EGR1* exhibited high correlation in both primary and recurrent gliomas (Fig. 6D). Transcriptome sequencing analysis revealed a positive correlation between the expression of *SERPINE1* and *EGR1* ($p<0.001$) (Fig. 6E, F).

Expression and enrichment analysis of EMT-related genes, immune characteristics, and metabolic pathways in MES-like tumor subpopulations

We examined the expression of EMT-related genes in various EMT-associated MES-like subpopulations (Supplement Fig. 6A). Next, we analyzed the expression of chemokines, chemokine receptors, MHC-related molecules, immune characteristics, and interferons in these subpopulations (Supplement Fig. 6B). Chemokines, chemokine receptors, and interferon-related genes were primarily expressed in the *SERPINE1*+MES-like-C6 and *VEGFC*+MES-like-C5 subpopulations. The *MGP*+MES-like-C1 subpopulation was enriched in metabolic pathways such as sphingolipid metabolism, primary bile acid biosynthesis, and cysteine and methionine metabolism (Supplement Fig. 6C). The *SERPINE1*+MES-like-C6 subpopulation was mainly enriched in steroid hormone biosynthesis, nicotinate and nicotinamide metabolism, nitrogen metabolism, and N-glycan biosynthesis. The *DPYSL3*+MES-like-C4 and *VEGFC*+MES-like-C5 subpopulations enriched most metabolic pathways. Furthermore, we used scFEA to infer the metabolic flux of cell subpopulations through a graph neural network model (Supplement Fig. 6D). The fatty acid signaling pathway was upregulated in the *MGP*+MES-like-C1 and *COL12A1*+MES-like-C7 subpopulations. The *VEGFC*+MES-like-C5 subpopulation showed down-regulation of glucose and other metabolic signaling pathways.

Expression and differentiation of epithelial-mesenchymal transition (EMT) related genes in IDH wild-type gliomas at different stages

Given that our analysis of glioma single cells used nuclear sequencing, which contains fewer immune cells, we next examined the changes in the tumor immune microenvironment in IDH wild-type glioma mice at early and late stages. After annotating the cell subpopulation markers, the tumor microenvironment subpopulations were identified as microglial cells, monocytes, macrophages, T cells, NK cells, dendritic cells (DCs), B cells, and oligodendrocytes (Supplementary Fig. 7A). Compared to

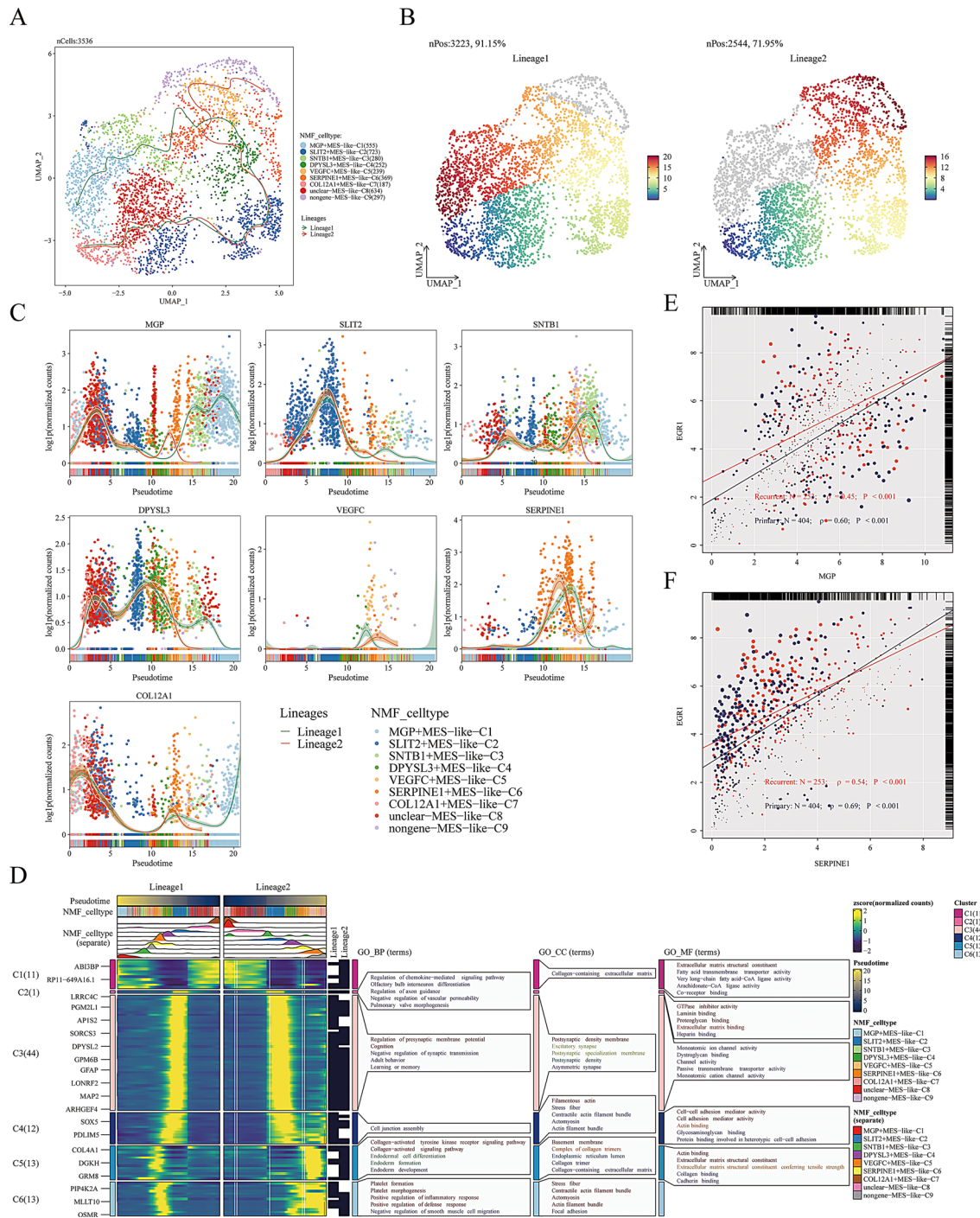


Fig. 5 Lineage tracing and molecular characterization of glioblastoma subtypes

(A) UMAP plot showing the distribution of secondary glioblastoma cells, color-coded by NMF-derived subtypes. Overlaid lines represent two distinct cell lineages identified within the dataset. **(B)** UMAP plots showing the two main lineages, Lineage 1 (left) and Lineage 2 (right), with pseudotime progression color-coded to indicate cellular differentiation along each lineage. **(C)** Expression trends of key genes (MGP, SLIT2, SNTB1, DPYSL3, VEGFC, SERPINE1, COL12A1) along the pseudotime axis in each lineage. The solid lines represent the smoothed gene expression curves for each gene, indicating dynamic expression changes during glioblastoma progression. **(D)** Heatmap displaying the gene expression profiles across subtypes in both lineages, with corresponding Gene Ontology (GO) enrichment analysis for biological processes (BP), cellular components (CC), and molecular functions (MF) shown on the right. The vertical lines divide the NMF subtypes into their respective lineages. **(E-F)** Scatter plots showing the correlation between SERPINE1 and MCP expression, highlighting significant correlations within the two main lineages. The color intensity reflects the strength of the correlation, with red lines representing the linear regression fits.

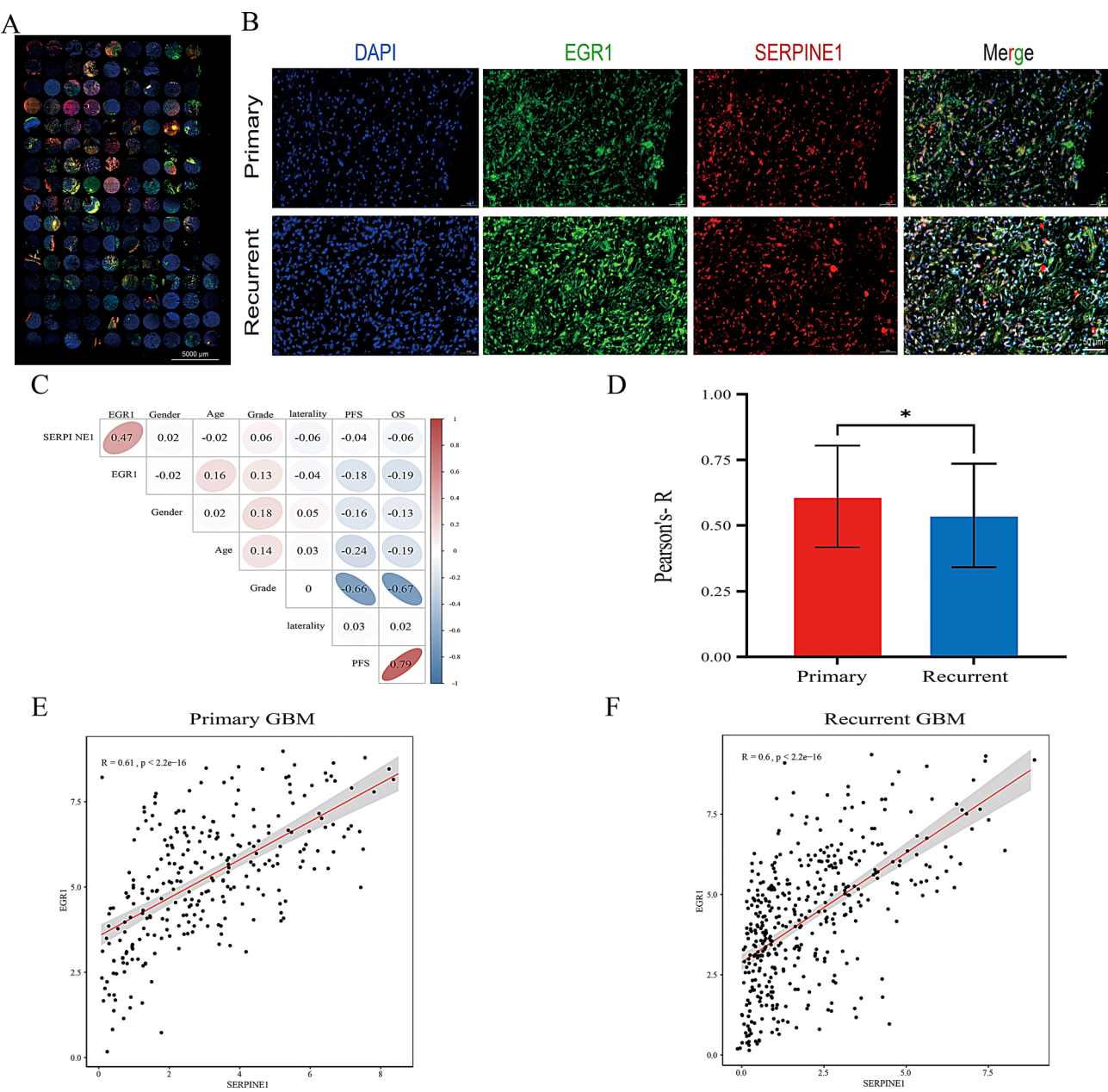


Fig. 6 Correlation between EGR1 and SERPINE1 expression in glioblastoma and functional validation. **(A)** Tissue microarray showing immunofluorescence staining for various markers across multiple glioblastoma samples. **(B)** Immunofluorescence images displaying EGR1 (green) and SERPINE1 (red) expression in primary and recurrent glioblastoma tissues. DAPI (blue) marks the nuclei, and the merged images show the co-localization of EGR1 and SERPINE1. **(C)** Correlation matrix showing relationships between EGR1, SERPINE1, and clinical factors, including gender, age, grade, laterality, progression-free survival (PFS), and overall survival (OS). A strong correlation between EGR1 and SERPINE1 ($R = 0.47$) is observed. **(D)** Bar plot showing Pearson's correlation coefficient (R) between EGR1 and SERPINE1 in primary and recurrent glioblastomas, with a significant difference ($*p < 0.05$) between the two groups. **(E)** The relationship between SERPINE1 and EGR1 in the transcriptome sequencing of primary gliomas. **(F)** The relationship between SERPINE1 and EGR1 in the transcriptome sequencing of recurrent gliomas. The red regression line in the figure represents linear fitting, and the gray shaded area indicates the 95% confidence interval

IDH-mutant glioma subgroups, IDH-wildtype gliomas exhibit higher proportions of monocytes (57.8%), macrophages (66.4%), T cells (65.5%), NK cells (74.2%), and dendritic cells (DCs) (65.1%) (Supplementary Fig. 7B, C). UMAP visualizations showed the distribution of the tumor microenvironment in IDH wild-type glioma mice at early and late stages (Supplementary Fig. 7D). Compared to early-stage IDH wild-type mice, the expression of *Egr1* was significantly lower in immune cells within the tumor microenvironment of late-stage IDH wild-type mice ($p < 0.001$) (Supplementary Fig. 7E). Additionally, the CytoTRACE2 stemness analysis revealed that

late-stage IDH wild-type cells exhibited a higher degree of stemness differentiation (Supplementary Fig. 7F–J). The pseudo-temporal analysis revealed that advanced-stage gliomas predominantly exist in the terminal differentiation phase (Supplementary Fig. 8A). The differential expression of *Egr1* and *Serpine1* also varied across different stages of glioma progression (Supplementary Fig. 8B, C).

SERPINE1 Inhibition reduced glioma proliferation and migration

To further explore the oncogenic role of *SERPINE1* in gliomas, we designed siRNA sequences targeting *SERPINE1* and assessed their knockdown efficiency in glioma cells for subsequent experiments (Fig. 7A). Previous studies have demonstrated that silencing *EGR1* expression significantly inhibits glioma cell proliferation and induces G1 phase cell cycle arrest, indicating that *EGR1* plays a crucial role in glioma cell proliferation [42, 43]. The inhibition of *SERPINE1* was found to diminish the proliferation and migration of glioma cells (Fig. 7B, C). The siRNA targeting *SERPINE1* was designed and its efficacy in reducing *SERPINE1* levels in U251 and T98G cells was confirmed. The CCK8 assay indicated a significant reduction in OD450 values in U251 cells and T98G cells 72 h after transfection, suggesting a decrease in cell proliferation (Fig. 7B). The wound healing assay showed a reduced healing rate in *SERPINE1*-depleted cells at 24 h, indicating impaired migration of glioma cells in vitro (Fig. 7C).

Cell cycle analysis

The cell cycle distribution of U251 cells and T98G cells treated with the siRNA sequence against *SERPINE1* was measured by comparing the proportion of cells in the G0/G1, S, and G2/M phases. After treatment with the siRNA sequence against *SERPINE1*, the proportion of cells in G1 and S was significantly increased, while the proportion in the G2/M phase was decreased (Fig. 7D). These results indicate that the siRNA sequence against *SERPINE1* facilitated the arrest of cells in the G1 and S phases of the cell cycle, thereby inhibiting U251 cell and T98G cell proliferation.

SERPINE1 knockdown inhibits glioma cell proliferation in vivo

Additionally, we examined the effects of *SERPINE1* in vivo (Fig. 7E). The results from the tumor xenograft studies demonstrated that sh-*SERPINE1* significantly reduced the size of the transplanted tumors, as evidenced by decreases in both tumor volume and weight (Fig. 7F, G). These findings are consistent with our in vitro results. In the xenograft tumor model, a comparison between the si-NC and si-*SERPINE1* groups revealed that, while both groups formed tumor tissue based on HE staining,

the tumor volume in the si-*SERPINE1* group was significantly smaller (Fig. 7H). IHC staining for Ki67 indicated that the proportion of Ki67-positive cells, a marker of proliferation, was notably lower in the si-*SERPINE1* group compared to the si-NC group, suggesting a marked reduction in tumor cell proliferation in the si-*SERPINE1* group. Additionally, the expression level of *SERPINE1* was significantly decreased in the si-*SERPINE1* group, further supporting its potential role in inhibiting tumor growth. Overall, the results demonstrated that si-*SERPINE1* exhibited a pronounced tumor-suppressive effect in the xenograft model, primarily reflected in reduced tumor proliferative activity and decreased *SERPINE1* expression.

Discussion

In this study, we identified key regulatory genes and cellular mechanisms driving the recurrence of IDH wild-type gliomas. Through extensive single-cell RNA sequencing and spatial transcriptomics, we found that the transcription factor *EGR1* significantly influences the EMT in glioma cells. Our analysis revealed that MES-like cell subpopulations exhibit high transcriptional regulatory activity, contributing to glioma invasion and drug resistance. Furthermore, we discovered that the *SERPINE1*-MES subpopulation is a critical factor in glioma recurrence, as evidenced by its high expression in recurrent gliomas and its association with poor patient prognosis. These findings were validated through functional assays and animal models, demonstrating the pivotal role of *EGR1* and *SERPINE1* in glioma recurrence and highlighting potential therapeutic targets to improve patient outcomes.

The recurrence of gliomas, particularly IDH wild-type gliomas, is a significant clinical challenge. Our findings that *EGR1* mediates EMT in MES-like subpopulations align with previous studies that have implicated EMT in tumor progression and metastasis [4, 6, 44, 45, 46]. *EGR1*'s role in enhancing the invasive capabilities of glioma cells supports the notion that transcription factors are crucial in regulating the plasticity and adaptability of cancer cells under therapeutic pressure [5, 42]. The upregulation of EMT-related genes in gliomas suggests that these factors could serve as biomarkers for aggressive glioma phenotypes and potential targets for novel therapeutic interventions [7, 8, 42, 47, 48].

Our spatial transcriptomics analysis revealed that MES-like subpopulations are primarily located at the tumor margins in recurrent gliomas, indicating their role in tumor invasion and recurrence [20]. This spatial distribution pattern is consistent with the enhanced invasive properties of these cells, as they interact with and degrade the extracellular matrix to facilitate tumor spread [13, 25]. The identification of specific subpopulations, such

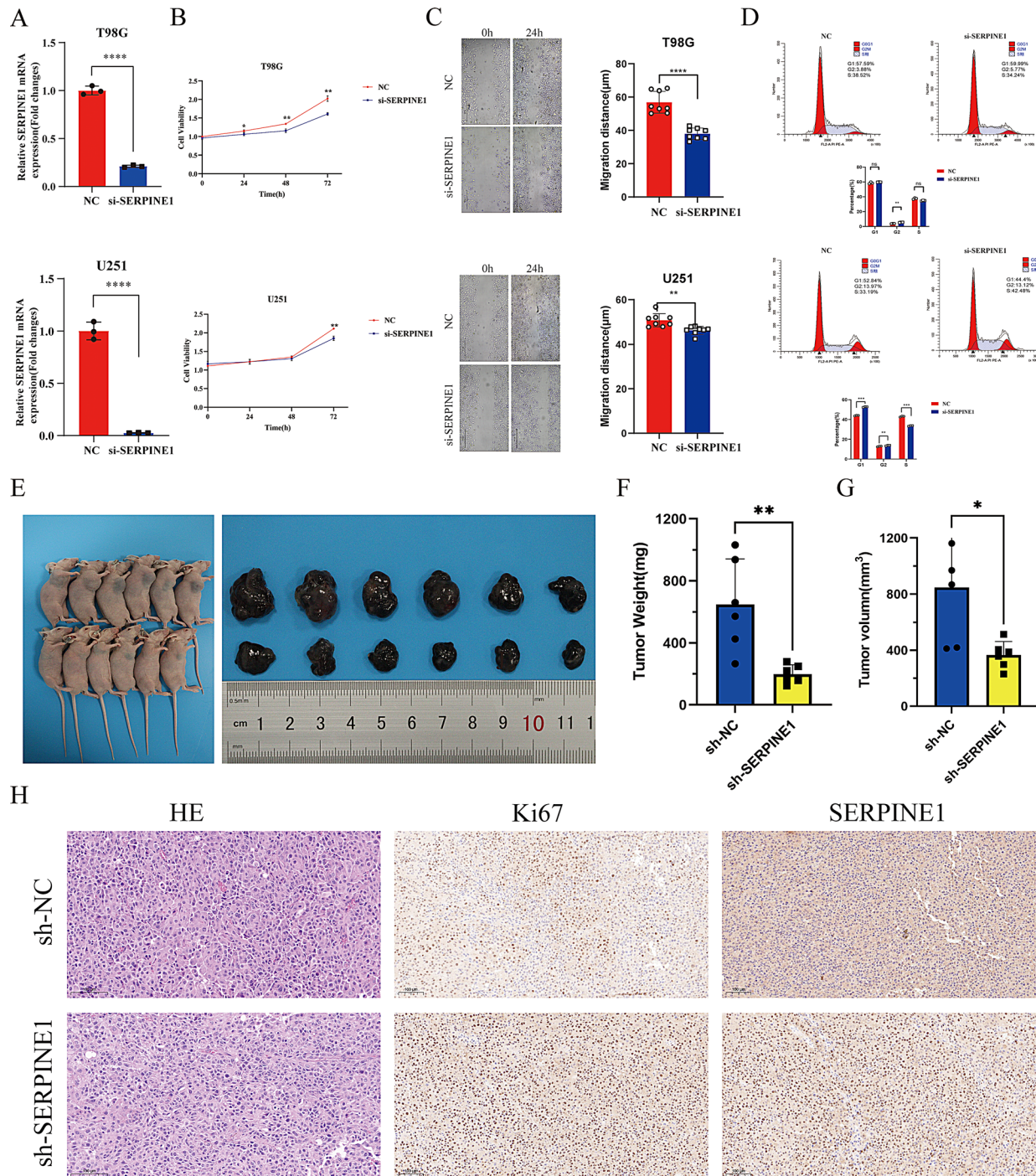


Fig. 7 Functional validation of SERPINE1 knockdown in glioblastoma cells and in vivo tumor models. **(A)** Quantitative PCR results showing a significant reduction in SERPINE1 mRNA expression after si-SERPINE1 treatment in both T98G and U251 glioblastoma cells (**** $p < 0.0001$). **(B)** Proliferation curves of T98G and U251 cells treated with si-SERPINE1, demonstrating reduced cell proliferation over 72 h (** $p < 0.01$). **(C)** Wound healing assay showing reduced migration distance in si-SERPINE1 treated T98G and U251 cells at 24 h. Quantification of migration distance shows a significant reduction compared to the negative control (NC) (**** $p < 0.0001$). **(D)** Flow cytometry analysis of cell cycle distribution in si-SERPINE1 and NC-treated T98G and U251 cells. si-SERPINE1 treated cells show an increased percentage of cells in the G1 phase and a reduction in S phase (** $p < 0.01$, *** $p < 0.001$). **(E)** In vivo tumor growth in nude mice. Representative images of mice and excised tumors are shown for the NC and sh-SERPINE1 treated groups. Tumor weight is significantly lower in the sh-SERPINE1 group (** $p < 0.01$). **(F)** Tumor weight of excised tumors from sh-SERPINE1 and NC-treated mice. Tumor weight is significantly lower in the sh-SERPINE1 group (** $p < 0.01$). **(G)** Tumor volume measurement showing reduced tumor size in sh-SERPINE1 treated mice compared to NC-treated mice (* $p < 0.05$). **(H)** Histological analysis of tumor sections stained with H&E, Ki67, and SERPINE1. The sh-SERPINE1 treated group shows reduced Ki67 and SERPINE1 staining, indicating decreased proliferation and target knockdown efficacy

as the *SERPINE1*-MES-like cells, which are enriched in ECM-receptor interaction pathways, further underscores the importance of the tumor microenvironment in glioma progression [12]. These findings are corroborated by previous research demonstrating the critical role of the extracellular matrix in providing structural support and biochemical signals that drive tumorigenesis [10].

The use of single-cell RNA sequencing allowed us to uncover the heterogeneity within glioma cell populations and identify subpopulations with distinct transcriptional profiles. This approach provides a more nuanced understanding of the tumor's cellular landscape compared to bulk RNA sequencing [17]. Our study revealed that recurrent gliomas exhibit higher proportions of oligodendrocyte-like and MES-like cells compared to primary gliomas, suggesting a shift in cellular composition that may contribute to therapy resistance and tumor aggressiveness [16]. The differentiation potential of these subpopulations, as assessed by CytoTRACE, indicates that MES-like cells possess lower stemness scores, aligning with their specialized role in invasion rather than self-renewal [33].

Functional assays further validated the roles of *EGR1* and *SERPINE1* in glioma recurrence. Knockdown of *EGR1* in glioma cell lines led to a significant reduction in *SERPINE1* expression, decreased cell proliferation, and impaired migratory capabilities, highlighting the interdependence of these factors in promoting tumor progression [26]. In vivo experiments using xenograft models confirmed these findings, with *SERPINE1* knockdown resulting in smaller tumor sizes and reduced invasiveness. These results suggest that targeting the *EGR1*-*SERPINE1* axis could be a promising strategy for inhibiting glioma recurrence and improving patient outcomes [21].

Despite these significant findings, our study has several limitations. First, while single-cell RNA sequencing provides detailed insights into cellular heterogeneity, it does not capture the full complexity of tumor-microenvironment interactions. Additional studies incorporating other omics approaches, such as proteomics and metabolomics, could provide a more comprehensive understanding of the mechanisms driving glioma recurrence. Second, our findings are based on samples from a relatively small cohort of patients, which may limit the generalizability of our results. Larger studies with diverse patient populations are needed to validate these findings. Third, while our in vivo models provided valuable insights into the roles of *EGR1* and *SERPINE1*, they do not fully recapitulate the human tumor microenvironment. Future research should explore the use of patient-derived xenografts and organoid models to better mimic human glioma biology. While our functional validation experiments using T98 and U251 glioblastoma cell lines provided initial evidence supporting a

regulatory relationship between *EGR1* and *SERPINE1*, we acknowledge that these established cell lines do not fully recapitulate the molecular and phenotypic complexity of recurrent GBM. Specifically, they lack the therapy-induced evolutionary pressures and tumor microenvironmental context present in patient-derived recurrent tumors. As such, the functional implications of the *EGR1*-*SERPINE1* axis in recurrence remain to be fully elucidated. Future studies utilizing patient-derived glioma stem-like cells, tumorsphere cultures, or recurrent GBM xenograft models will be essential to validate and extend these findings in a clinically relevant context.

Conclusion

Our study provides a comprehensive analysis of the regulatory mechanisms underlying the recurrence of IDH wild-type gliomas, emphasizing the critical role of the transcription factor *EGR1* and the *SERPINE1*-MES subpopulation in mediating EMT and tumor invasion. This research highlights the complexity and heterogeneity of glioma cell subpopulations, revealing novel insights into their dynamic plasticity and interaction with the tumor microenvironment. These findings underscore the importance of targeting specific transcriptional regulators and signaling pathways to develop more effective therapeutic strategies for glioma treatment.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40478-025-02036-6>.

Supplementary Material 1: Supplement Figure 1. Copy number variation (CNV) analysis in primary and recurrent glioblastoma. **(A)** UMAP visualization of 143,340 cells from glioma samples, showing distinct cell populations, including oligodendrocytes, astrocytes, macrophages, microglia, endothelial cells, neurons, and T cells, color-coded based on cell type. **(B)** Violin plots showing CNV scores across different Seurat clusters. Each cluster represents a distinct cell population, and CNV scores are color-coded based on cluster identity. Higher CNV scores suggest a greater degree of malignancy. **(C)** Bar plot comparing the proportion of malignant and normal cells between primary and recurrent glioblastoma samples. Primary tumors consist of 86.4% malignant cells, while recurrent tumors contain 84.7% malignant cells. **(D)** UMAP plots visualizing the distribution of malignant (red) and normal (blue) cells in primary (left) and recurrent (right) glioblastomas. Malignant cells dominate both primary and recurrent samples, but recurrent tumors show a slightly higher proportion of normal cells. **(E)** Bar chart showing the relative proportions of cell subtypes in primary and recurrent gliomas. Recurrent gliomas display a higher percentage of MES-like and OC-like cells compared to primary gliomas.

Supplementary Material 2: Supplement Figure 2. Spatial and pathway analysis of primary and secondary glioblastomas using spatial transcriptomics. **(A)** Pathway enrichment analysis showing significantly altered pathways across glioblastoma subtypes, including inflammatory mediator regulation of TRP channels, Wnt signaling pathway, ECM-receptor interaction, and MAPK signaling pathway. The x-axis represents $-\log_{10}(q\text{-value})$ and the dot size indicates the number of altered genes in each pathway. **(B-C)** Spatial transcriptomic heatmaps showing gene expression patterns for AC-like, MES-like, OC-like, and OPC-like cells in secondary glioblastoma samples. Each plot represents the spatial distribution of gene expression across the tumor sections, with color intensity corresponding to expression levels. **(D-F)** Spatial transcriptomic heatmaps illustrating gene expression patterns

for AC-like, MES-like, OC-like, and OPC-like cells in primary glioblastoma samples. Similar to secondary glioblastoma, these heatmaps highlight the spatial heterogeneity of gene expression in the different subtypes.

Supplementary Material 3: Supplement Figure 3. Kaplan-Meier survival curves for glioblastoma patients based on the expression levels of key genes. **(A)** Kaplan-Meier survival curve for MGP expression. Patients with high MGP expression show significantly worse survival ($p < 0.001$). **(B)** Kaplan-Meier survival curve for SERPINE1 expression. Higher SERPINE1 expression is associated with poorer survival ($p < 0.001$). **(C)** Kaplan-Meier survival curve for SLIT2 expression, showing no significant difference between high and low expression groups ($p = 0.827$). **(D)** Kaplan-Meier survival curve for SNTB1 expression. High SNTB1 expression correlates with worse survival ($p = 0.010$). **(E)** Kaplan-Meier survival curve for VEGFC expression. Patients with high VEGFC expression have significantly worse survival outcomes ($p = 0.035$). **(F)** Kaplan-Meier survival curve for COL12A1 expression. Higher COL12A1 expression is associated with significantly reduced survival ($p < 0.001$). **(G)** Kaplan-Meier survival curve for DPYSL3 expression. There is no significant difference in survival between high and low DPYSL3 expression groups ($p = 0.155$).

Supplementary Material 4: Supplement Figure 4. Kaplan-Meier survival curves for glioblastoma patients based on gene expression in MES-like subtypes. **(A)** Kaplan-Meier survival curve for DPYSL3+MES-like C4 subtype. No significant difference in survival is observed between high and low expression groups ($p = 0.722$). **(B)** Kaplan-Meier survival curve for MGP+MES-like C1 subtype. High expression is significantly associated with worse survival ($p < 0.001$). **(C)** Kaplan-Meier survival curve for SERPINE1+MES-like C6 subtype. Patients with high expression have significantly poorer survival outcomes ($p < 0.001$). **(D)** Kaplan-Meier survival curve for SLIT2+MES-like C2 subtype. No significant difference in survival is observed between high and low expression groups ($p = 0.231$). **(E)** Kaplan-Meier survival curve for SNTB1+MES-like C3 subtype. No significant difference in survival is observed ($p = 0.149$). **(F)** Kaplan-Meier survival curve for VEGFC+MES-like C5 subtype. High expression is associated with significantly worse survival ($p = 0.001$). **(G)** Kaplan-Meier survival curve for COL12A1+MES-like C7 subtype. High expression is significantly associated with poorer survival ($p < 0.001$).

Supplementary Material 5: Supplement Figure 5. Pseudotime analysis and signaling pathway interactions in glioblastoma subtypes. **(A)** Monocle2 trajectory plots showing the pseudotime progression of 3,536 cells across various NMF-derived subtypes. Each point represents a single cell, color-coded by subtype and pseudotime state. **(B)** Expression dynamics of MGP along the pseudotime trajectory. Cells with higher MGP expression cluster at the end of the trajectory, indicating association with later differentiation stages. **(C)** Expression dynamics of SERPINE1 along the pseudotime trajectory, showing similar clustering at later pseudotime points. **(D)** Cell-cell communication network analysis showing signaling interactions between different MES-like subtypes. The size of the nodes reflects the strength of outgoing signals, and the thickness of edges indicates interaction strength between subtypes. **(E)** Detailed signaling interaction network of MES-like subtypes, focusing on specific pathways, including those involving MGP, VEGFC, and SERPINE1. **(F)** Heatmap showing outgoing and incoming signaling patterns between glioblastoma subtypes. The intensity of green color indicates the strength of signaling interactions between subtypes. **(G)** MK signaling pathway network showing specific source-target relationships between subtypes, with key signals originating from MGP and targeting various MES-like subtypes. **(H)** IGF signaling pathway network highlighting interactions between MES-like subtypes, with MGP as a central node. **(I)** Dot plot showing pathway enrichment analysis for key transcription factors involved in glioblastoma subtype differentiation. Dot size represents the number of enriched pathways, and color reflects p-value significance.

Supplementary Material 6: Supplement Figure 6. Metabolic pathway analysis and gene expression profiling across glioblastoma subtypes. **(A)** Heatmap showing the expression of key metabolic genes across NMF-derived glioblastoma subtypes. Rows represent genes, and columns represent subtypes, with color intensity reflecting normalized expression levels (z-scores). **(B)** Heatmap illustrating the expression of genes involved in the glycosylation pathway across glioblastoma subtypes, with similar color-coding to indicate gene expression differences between subtypes. **(C)** Dot plot depicting the enrichment of various metabolic pathways

across NMF-derived subtypes. Dot size corresponds to the number of genes involved in each pathway, and color represents the significance of enrichment (p-value). **(D)** Pathway-specific heatmap visualizing key metabolic reactions (e.g., glucose to glucose-6-phosphate, fatty acid metabolism, and methionine to cysteine conversion) across glioblastoma subtypes, with each subtype displaying a distinct metabolic profile. Rows represent pathways, and columns represent subtypes, with color intensity reflecting pathway activity levels.

Supplementary Material 7: Supplement Figure 7. Single-cell RNA sequencing analysis of immune cell populations and differentiation states in IDH wild-type (WT) and mutant (MUT) glioblastoma subtypes. **(A)** Heatmap showing differential gene expression across various immune cell subtypes in IDH WT and MUT glioblastomas. Rows represent genes, and columns represent cell subtypes. **(B)** Bar chart illustrating the relative proportions of immune cell types in WT and MUT glioblastomas. **(C)** Stacked bar chart showing the distribution of specific immune cell subtypes in WT and MUT glioblastomas, highlighting significant differences in monocyte and microglia populations. **(D)** UMAP plots visualizing the distribution of late-stage (left) and early-stage (right) glioblastoma cells across different immune cell types, including microglia, monocytes, macrophages, T cells, NK cells, dendritic cells (DCs), and oligodendrocytes. **(E)** Violin plots comparing Egr1 expression levels across immune cell subtypes in WT and RII glioblastomas, with significant differences marked (**** $p < 0.0001$). **(F)** UMAP plot showing cell phenotypes across WT and RII glioblastoma samples, color-coded based on phenotype classifications. **(G)** UMAP plot showing CytoTRACE-predicted differentiation potential, with less differentiated cells clustered on the left and more differentiated cells on the right. **(H)** UMAP plot showing CytoTRACE-derived differentiation categories, dividing cells into undifferentiated and differentiated groups. **(I)** Boxplot comparing developmental potential between phenotypic groups, with stemness scores significantly higher in the WT group. **(J)** UMAP plot showing relative differentiation states as predicted by CytoTRACE, with color intensity indicating levels of differentiation.

Supplementary Material 8: Supplement Figure 8. Pseudotime trajectory and gene expression dynamics in glioblastoma. **(A)** Monocle2 trajectory plots depicting the pseudotime progression of 38,931 cells across various states. Each point represents a cell, with colors indicating distinct pseudotime states (left), timepoints (middle), and overall pseudotime progression (right). **(B)** Pseudotime trajectory highlighting the expression of Egr1. Cells with higher Egr1 expression are predominantly located at specific branches of the trajectory, suggesting its association with certain differentiation stages. **(C)** Pseudotime trajectory showing the expression of Serpine1. Similar to Egr1, higher Serpine1 expression clusters at particular branches, indicating its role in specific cell states along the differentiation trajectory.

Supplementary Material 9

Supplementary Material 10

Supplementary Material 11

Supplementary Material 12

Supplementary Material 13

Supplementary Material 14

Supplementary Material 15

Acknowledgements

This study was funded by the National Natural Science Foundation of China (82072786) and the Major Scientific and Technological Innovation Project (2021ZD0201605). We acknowledge the CGGA, TCGA, and the GEO database for providing access to the datasets used in this study. Thanks to the main data sources provided by the Chinese Glioma Genome Atlas Database and the participation of all patients.

Author contributions

Jinwei Li, Shengrong Long, and Yang Zhang designed and performed the experiments, collected and analyzed the data, and wrote the manuscript. Shuangqi Yu and Hongyu Xu assisted in experimental design, data analysis,

and provided critical revisions to the manuscript. Rui Liang, Jinnan Zhang, and Quan Liu provided essential reagents and materials, contributed to data interpretation, and assisted with manuscript preparation. Yixin Fu, Xiang Li, and Tao Xin supervised the research and provided expertise in glioma biology. Tao Xin, Yinyan Wang, as the corresponding author, conceived the project, oversaw all aspects of the study, and provided final approval of the version to be published.

Funding

This study was funded by the National Beijing Municipal Natural Science Foundation (No. JQ23040) and the Major Scientific and Technological Innovation Project (2021ZD0201605).

Data availability

The datasets generated and analyzed during the current study are available in the following repositories: Single-nucleus RNA sequencing data: The data from the 52 human IDH wild-type GBM specimens, including 26 primary tumors and 26 matched recurrent tumors, are available from the Gene Expression Omnibus (GEO) under accession number GSE174554. The Spatial transcriptomics data are publicly available under the accession number GSE194329. The IDH wild-type GBM mouse single-cell transcriptomic data is available from the GEO database under the accession number GSE166218.

Declarations

Ethics approval and consent to participate

Patient sample collection and study design for spatial transcriptome data have received ethical approval from West China Hospital, Sichuan University (2020.837).

Consent for publication

The authors declare no conflict of interest relevant to the present manuscript.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China

²Department of Neurosurgery, West China Hospital, Sichuan University, Chengdu, Sichuan 610000, China

³Department of Thoracic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

⁴Department of Neurosurgery, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China

⁵Department of Vascular Surgery, Fuwai Yunnan Cardiovascular Hospital, Affiliated Cardiovascular Hospital of Kunming Medical University, Kunming, Yunnan, China

⁶College of Bioengineering, Chongqing University, Chongqing 400030, China

⁷Brain Research Center, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China

⁸Department of Neurosurgery, The Fourth Affiliated Hospital of Guangxi Medical University, Liuzhou, Guangxi 545000, China

⁹Department of Neurosurgery, China-Japan Union Hospital of Jilin University, Changchun 130000, China

¹⁰Department of Neurosurgery, The First Affiliated Hospital of Shandong First Medical University, Jinan 250014, China

¹¹Beijing Neurosurgical Institute, Capital Medical University, Beijing 100070, China

¹²Department of Neurosurgery, First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

Received: 17 March 2025 / Accepted: 10 May 2025

Published online: 07 June 2025

References

1. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS (2020) CBTRUS statistical report: primary brain and other central nervous system

- tumors diagnosed in the united States in 2013–2017. *Neuro Oncol* 22:iv1–iv96
2. Jiang T, Mao Y, Ma W et al (2016) CGCG clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett* 375:263–273
3. Jiang T, Nam DH, Ram Z et al (2021) Clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett* 499:60–72
4. Iser IC, Pereira MB, Lenz G, Wink MR (2017) The Epithelial-to-Mesenchymal Transition-Like process in glioblastoma: an updated systematic review and in Silico investigation. *Med Res Rev* 37:271–313
5. Kauffhold S, Bonavida B (2014) Central role of Snail1 in the regulation of EMT and resistance in cancer: a target for therapeutic intervention. *J Experimental Clin cancer Research: CR* 33:62
6. Lamouille S, Xu J, Derynck R (2014) Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* 15:178–196
7. Talasila KM, Soentgerath A, Euskirchen P et al (2013) EGFR wild-type amplification and activation promote invasion and development of glioblastoma independent of angiogenesis. *Acta Neuropathol* 125:683–698
8. Wang Y, Shen N, Yang Y et al (2024) ZDHHC5-mediated S-palmitoylation of FAK promotes its membrane localization and epithelial-mesenchymal transition in glioma. *Cell Communication Signaling: CCS* 22:46
9. Yao Z, Zhang F, Qi C et al (2024) SECTM1 promotes the development of glioblastoma and mesenchymal transition by regulating the TGFβ1/Smad signaling pathway. *Int J Biol Sci* 20:78–93
10. Jiang J, Chen HN, Jin P et al (2023) Targeting PSAT1 to mitigate metastasis in tumors with p53-72Pro variant. *Signal Transduct Target Therapy* 8:65
11. Saitoh M (2018) Involvement of partial EMT in cancer progression. *J Biochem* 164:257–264
12. Osswald M, Jung E, Sahm F et al (2015) Brain tumour cells interconnect to a functional and resistant network. *Nature* 528:93–98
13. Galvao RP, Kasina A, McNeill RS et al (2014) Transformation of quiescent adult oligodendrocyte precursor cells into malignant glioma through a multistep reactivation process. *Proc Natl Acad Sci USA* 111:E4214–4223
14. Ravi VM, Will P, Kueckelhaus J et al (2022) Spatially resolved multi-omics Deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* 40:639–655e613
15. Couturier CP, Ayyadury S, Le PU et al (2020) Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* 11:3406
16. Neftel C, Laffy J, Filbin MG et al (2019) An integrative model of cellular States, plasticity, and genetics for glioblastoma. *Cell* 178:835–849e821
17. Tirosh I, Venteicher AS, Hebert C et al (2016) Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature* 539:309–313
18. Venteicher AS, Tirosh I, Hebert C et al (2017) Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science*, New York, NY, p 355
19. Ren Y, Huang Z, Zhou L et al (2023) Spatial transcriptomics reveals niche-specific enrichment and vulnerabilities of radial glial stem-like cells in malignant gliomas. *Nat Commun* 14:1028
20. Verhaak RG, Hoadley KA, Purdom E 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
21. Wang Q, Hu B, Hu X et al (2018) Tumor evolution of Glioma-Intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell* 33:152
22. Wang L, Jung J, Babikir H et al (2022) A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nat cancer* 3:1534–1552
23. Friedrich M, Hahn M, Michel J et al (2023) Dysfunctional dendritic cells limit antigen-specific T cell response in glioma. *Neurooncology* 25:263–276
24. Zhang X, Lan Y, Xu J et al (2019) CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res* 47:D721–d728
25. Greenwald AC, Darnell NG, Hoefflin R et al (2024) Integrative Spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* 187:2485–2501e2426
26. Patel AP, Tirosh I, Trombetta JJ et al (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Sci (New York NY)* 344:1396–1401
27. Xiong A, Zhang J, Chen Y, Zhang Y, Yang F (2022) Integrated single-cell transcriptomic analyses reveal that GPNMB-high macrophages promote PN-MES transition and impede T cell activation in GBM. *EBioMedicine*;83:104239
28. Zheng Y, Carrillo-Perez F, Pizurica M, Heiland DH, Gevaert O (2023) Spatial cellular architecture predicts prognosis in glioblastoma. *Nat Commun* 14:4122

29. Couturier CP, Nadaf J, Li Z et al (2022) Glioblastoma scRNA-seq shows treatment-induced, immune-dependent increase in mesenchymal cancer cells and structural variants in distal neural stem cells. *Neurooncology* 24:1494–1508
30. Dai X, Ye L, Li H et al (2023) Crosstalk between microglia and neural stem cells influences the relapse of glioblastoma in GBM immunological microenvironment. *Clin Immunol (Orlando Fla)* 251:109333
31. Aibar S, González-Blas CB, Moerman T et al (2017) SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 14:1083–1086
32. Townes FW, Engelhardt BE (2023) Nonnegative Spatial factorization applied to Spatial genomics. *Nat Methods* 20:229–238
33. Gulati GS, Sikandar SS, Wesche DJ et al (2020) Single-cell transcriptional diversity is a hallmark of developmental potential. *Sci (New York NY)* 367:405–411
34. Alghamdi N, Chang W, Dang P et al (2021) A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. *Genome Res* 31:1867–1884
35. Kleshchevnikov V, Shmatko A, Dann E et al (2022) Cell2location maps fine-grained cell types in Spatial transcriptomics. *Nat Biotechnol* 40:661–671
36. Theodoris CV, Xiao L, Chopra A et al (2023) Transfer learning enables predictions in network biology. *Nature* 618:616–624
37. Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14:7
38. Daisy Precilla S, Biswas I, Kuduvalli SS, Anitha TS (2022) Crosstalk between PI3K/AKT/mTOR and WNT/ β -Catenin signaling in GBM - Could combination therapy checkmate the collusion? *Cell Signal* 95:110350
39. Yu JS, Cui W (2016) Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. *Development* 143:3050–3060
40. Shorning BY, Dass MS, Smalley MJ, Pearson HB The PI3K-AKT-mTOR pathway and prostate cancer: at the crossroads of AR, MAPK, and WNT signaling. *Int J Mol Sci* 2020;21.
41. Yang J, Antin P, Berx G et al (2020) Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* 21:341–352
42. Knudsen AM, Eilertsen I, Kielland S et al (2020) Expression and prognostic value of the transcription factors EGR1 and EGR3 in gliomas. *Sci Rep* 10:9285
43. Chen DG, Zhu B, Lv SQ et al (2017) Inhibition of EGR1 inhibits glioma proliferation by targeting CCND1 promoter. *J Experimental Clin cancer Research: CR* 36:186
44. Wang Y, Qin C, Zhao B et al (2023) EGR1 induces EMT in pancreatic cancer via a P300/SNAI2 pathway. *J Translational Med* 21:201
45. Zhao Z, Zhang KN, Wang Q et al (2021) Chinese glioma genome atlas (CGGA): A comprehensive resource with functional genomic data from Chinese glioma patients. *Genom Proteom Bioinform* 19:1–12
46. Fontana R, Mestre-Farrera A, Yang J (2024) Update on Epithelial-Mesenchymal plasticity in Cancer progression. *Annu Rev Pathol* 19:133–156
47. He H, Liang L, Jiang S et al (2024) GINS2 regulates Temozolomide chemosensitivity via the EGR1/ECT2 axis in gliomas. *Cell Death Dis* 15:205
48. Fu X, Ren C, Dai K, Ren M, Yan C (2024) Epithelial-Mesenchymal Transition Related Score Functions as a Predictive Tool for Immunotherapy and Candidate Drugs in Glioma. *Journal of chemical information and modeling*

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.